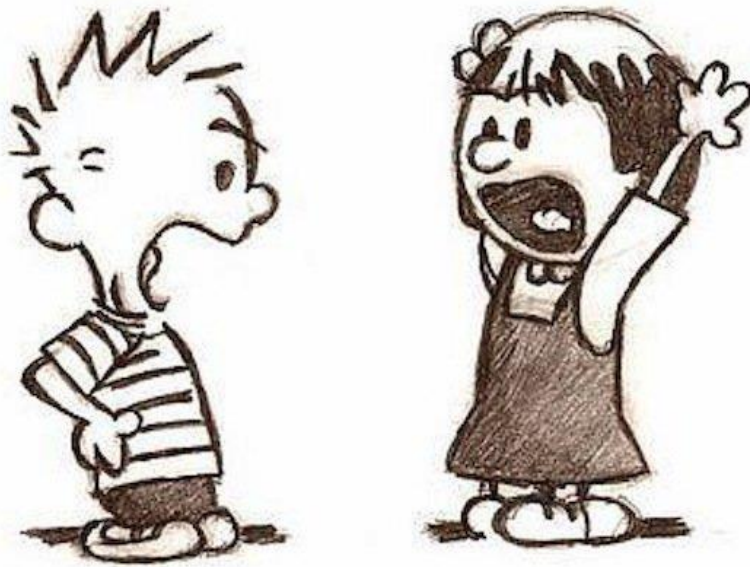




INTERNETSOCIALITET PÅ REDDIT

Et studie af et underforum på reddit: r/unpopularopinion

RESUME

Contemporary analysis of internet sociality on the social media platform, reddit

Eskil Andersen

Master's Thesis – Sociology,
AAU

INDHOLD

Abstract.....	
1 Indledning	4
1.1 De sociale medier.....	5
1.2 Hvad er reddit	7
1.3 Reddits struktur	8
1.3.1 Strukturernes påvirkning på det sociale liv.....	9
1.4 Hvorfor r/unpopularopinion	10
2 Problemformulering.....	10
2.1 Eksisterende forskning.....	12
2.1.1 Det sociale liv i reddits struktur	12
3 Teori	15
3.1 Om paradigmer	16
3.2 Neo-tribalism in an ordered world	17
3.3 Traditionernes magt i stjernetågens krystalliseringer	19
3.3.1 Hysterese og det sociale livs foranderlighed	20
3.4 Neo-tribal kultur på den sociale medieplatform reddit.....	22
4 Videnskabsteori	23
4.1 Den kritisk realistiske virkelighed	24
4.2 Ontologisk og epistemologisk videnskabsforståelse	24
4.3 Abstraktion og retroduktion	26
5 Metode.....	26
5.1 Dataindsamling	27
5.1.1 r/unpopularopinion som observationsobjekt.....	27
5.2 Brugerpopulationen på reddit	28
5.2.1 Aktørerne på reddit.....	29
5.2.2 Praktisk dataudvælgelse	30
5.2.3 Analysens datasæt	31
5.3 Analysestrategi.....	32
5.3.1 Tidsserieanalyse	33
5.3.2 Univariat analyse.....	34
5.3.3 Emneanalyse/Topic Modelling.....	35
5.3.4 LDA og topic modelling i praksis	36
6 Analyse.....	39

6.1	Tidsserier	40
6.1.1	2016	41
6.1.2	2017	42
6.1.3	2018	42
6.1.4	2019	43
6.2	Univariat analyse.....	43
6.2.1	2016	44
6.2.2	2017	46
6.2.3	2018	49
6.2.4	2019	51
6.3	Emneanalyse (Topic Modelling).....	54
6.3.1	2016	55
6.3.2	2017	58
6.3.3	2018	59
6.3.4	2019	61
7	Diskuterende konklusion	62
7.1	Refleksioner over den deskriptive analyse	62
7.2	Tematisering af den samlede indholdsdeling	64
7.3	Emneanalytisk diskussion	66
7.3.1	Mænd, Kvinder og kampen mellem kønnene.....	66
7.3.2	<i>Black/White</i>	68
7.4	Afrunding	70
8	Litteraturliste	73
9	Bilagsoversigt	80
10	Slutnoter	81

ABSTRACT

This thesis concerns itself with internet sociality, based on the interactions between users on the social media platform, reddit. Specifically, on the subreddit r/unpopularopinion. The empirical foundation is based on more than 3 years of interaction between the users of the textually based debate forum. Through the thesis, reddit will be framed, both as a social institution and a sociological phenomenon. The sociality of the actors on reddit is deemed a valuable target of observation, considering the influx of social media and its massive impact on contemporary society. Interaction on reddit is masked by user anonymity, which poses numerous opportunities and ethical problems for implicated actors and for the global and local societies alike, meanwhile complicating general sociological conceptualizations of the user population and its sociality. The theoretical scope of the thesis is based around the work of Michel Maffesoli in his book *Time of the Tribes – The Decline of Individualism in Mass Society* (1996), and will discuss the social and paradigmatic changes the contemporary society and its actors is undergoing, as the internet and social media is becoming an increasingly vital part of social life. Scientifically, the thesis takes a *critical realist* approach, and seeks to illuminate some of the underlying mechanisms that play out in the interaction between actors and help shape sociality on r/unpopularopinion as well as reddit in general. The analytical framework is based around a longitudinal collection of shared opinions and the related interactions among the user population on the forum. Through this, it will be sought to gain insight into the general sociality on the forum and, more importantly, to cast a light on any waves or fluctuations in the discursive and behavioural development in the observation period. By conceptualizing some of the workings of internet sociality, through the lens of an anonymous debate forum of unpopular opinions, it is hoped to better grasp some of the prospects and complications, an anonymous social media platform presents for both its users and for society. Methodologically, the thesis seeks to gain insight into ways of understanding the masses of vast amounts of data that flow through the internet daily, and how to subtract valuable insights from the cornucopia of information presented on reddit. In effect, the scope of the thesis works with *topic modelling* and will use the textual universe, presented in the written interactions between the users of r/unpopularopinion, to illuminate the discursive fluctuations through the empirical epochs, and try to thematize the overall main topics of discussions through the years. The Thesis aims to create a way of exploration within anonymous social media, and to better the understanding of social interaction in the internet era.

“Technology is neither good nor bad; nor is it neutral...technology’s interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves.”

(Kranzberg 1986:545)

1 INDLEDNING

Dette speciale søger at belyse internetsocialitet og det sociale liv på det tekstuelte baserede debatforum for upopulære holdninger r/unpopularopinion på den sociale medieplatform reddit. Reddit markerer sig blandt andre online sociale medier ved at være anonymt, brugermodereret og stort set ucensureret (Kilgo et al. 2018:2; Massanari 2016:4). Reddit fungerer som en paraplyplatform for et utal af brugerskabte debat- og delingsfora, på tværs af interesser og emner, fra det festlige til det direkte morbide (Aladağ et al. 2018; Literat and van den Berg 2019:233). Reddit som fænomen har fået større opmærksomhed over de senere år, både blandt den brede befolkning og medier, men også blandt forskere i forskellige videnskabsgrene. Selvom enkelte undersøgelser har arbejdet med reddit som helhed (Massanari 2016; Matias 2018; Panek et al. 2018), er langt størstedelen af den videnskabelige forskning rettet mod specifikke underfora kendt som subreddits, der fremstår som uafhængige myndigheder under reddit's paraply (Cole, Watkins, and Kleine 2016; Eschler and Menking 2018; Mills 2018; Sharma et al. 2016). Brugernes anonymitet og relative autonomi i kongruens med den emnespecifikke indholdsdeling, har skabt grobund for dybe videnskabelige indsigter i såvel interessefelter, som kontroversielle adfærdsmønstre blandt brugergrupperinger på platformen (Marwick 2017; Massanari 2017; Matias 2018; Mills 2018). Med den brede opdeling i forskellige fora, er der også stor variation i oplevelsen af platformen, og enhver undersøgelse af reddit som emnespecifikt fænomen, vil være begrænset i sin optik.

I dette speciale arbejdes der med interaktionen omkring holdningsudmeldinger og den relationelle udvikling i takt med reddit's, og r/unpopularopinions, stigende popularitet og anvendelse. Undersøgelsen fokuserer på internetsocialitet og udviklingen i interaktion, i hvad der kan opleves som et relativt uforstyrret direkte demokrati, under emnemæssig afgrænsning af upopulære holdningsdelinger (Mills and Fish 2015:236). Interessefeltet for undersøgelsen ligger i et tidspænd fra reddit's begyndende popularitet til en anerkendt og relativt bredt anvendt social medieside. Specialet er udarbejdet som et longitudinelt casestudie af den kollektive adfærd orienteret mod holdningsudmeldinger og på de diskursive bølgninger, der følger af den samlede brugersammenslutning på r/unpopularopinions udvikling og ekspansion. Specialet udarbejdes med

det formål at beskrive nogle af de samfundsmæssige tendenser der gør sig gældende i det 21. århundredes globale massekultur, hvor den teknologiske udvikling og internettets udbredelse har skabt nye muligheder for såvel individer som organisationer for at begå sig i samfundet (Curran, Fenton, and Freedman 2016:93; Maffesoli 1996). Blandt den brede middelklasse, på tværs af det globale landskab, har de sociale mediers komme været nok den af internettets virkemidler, der har haft størst indvirkning på det man kan betegne som massernes kommunikationskultur (Curran et al. 2016:92–95; Leavitt 2015; Massanari 2016:7; Van Mieghem 2011). Specialet søger sociologisk at afdække nogle af de essentielle træk, der opleves under massetilstrømningen af brugere til reddit, og herunder r/unpopularopinion (Baumgartner 2019; Subredditstats.com 2019). Indramningen af internet-socialiteten på r/unpopularopinion, specifikt i forhold til dets udvikling og popularitet, har til formål at understøtte forståelsen for det sociale liv der udspiller sig i internetalderen, set i lyset af de kommunikative og sociale virkemidler, der gør sig gældende på reddit og specifikt på r/unpopularopinion. Dette bearbejdes, blandt andet ved hjælp af emneanalyse over dominerende og tematiserende termer for indholdsdelingen.

1.1 DE SOCIALE MEDIER

Aktøren i samfundets måde at kommunikere med såvel nære relationer som perifere bekendtskaber, har undergået en forandring i takt med, at tilgængeligheden og anvendelsesmulighederne af kollektive medier har udviklet sig (Curran et al. 2016:85; Neal 2012:227–32). Internettets sociale medier er en bred viftebetegnelse for et utal af formidlings-, delings- og underholdningsorienterede platforme, der tager udgangspunkt i brugergenereret indhold. Indholdsdelingen er som oftest defineret i platformens typemæssige afgrænsning, nogle mere emnespecifikke end andre (Ibid.). Siden de globale sociale mediers fremkomst i slut 90'erne og i starten af det nye årtusind, er der opstået og bortgået en myriade af forskellige typer af sociale mediesider og platforme, der i en tilstand af usikker midlertidighed, er defineret af, og fuldstændig afhængige af, brugernes engagement og interesse (Neal 2012:227–30). Sideløbende med den teknologiske udvikling og tilgængelighed, har de sociale medier måtte tilpasse sig eller forgå til fordel for nye delings- og kommunikationsformer, der fanger aktørernes interesse og behov (Curran et al. 2016:86–87). Den adfærdsmæssige interaktion mellem internettets aktører, har siden de sociale mediers tidligste eksistens, dannet grobund for mange kontroverser og etiske problematikker, såvel i global som lokal kontekst (Curran et al. 2016:117; Marwick 2017; Massanari 2017; Mills 2018). Internetaktivisten John Barlow reagerede i et manifest, sin modvilje mod *the Communication Decency Act (CDA)*, den først indførte amerikanske lov mod internetadfærd, der for første gang implicerede regulering af forskellige typer af deling på internettet:

"Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather. We have no elected government, nor are we likely to have one, so I address you with no greater authority than that with which liberty itself always speaks. I declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us. You have no moral right to rule us nor do you possess any methods of enforcement we have true reasons to fear. (Barlow 1996; Curran et al. 2016:117)

Internettets altomspændende tilgængelighed har nedbrudt nogle geografiske begrænsninger, men har også i høj grad sat nogle ekstremer ved den menneskelige socialitet under belysning (ibid.). Selvom ovenstående citat ikke forholder sig direkte til reddit, kan det anskues at udmeldingen på mange tangerer den gennemgående holdning, delt blandt mange brugere på reddit (Kilgo et al. 2018:2). Internettets sociale kommunikationsmedier, og herunder også reddit, har dannet grobund for både radikale og normdannende kulturelle strømninger, der har haft såvel positive som negative følgevirkninger (Massanari 2017; Matias 2018; Park and Conway 2018). Nedbrydelsen af geografiske begrænsninger i onlineverdenen, har resulteret i, at sociale grupperinger der ellers har været inddæmmet i deres lokale afgrænsning, har været i stand til at finde samhørighed og fællesskab på tværs af landegrænsere og socioøkonomiske skel (Amer and Noujaim 2019; Curran et al. 2016:93; Mills 2018). Parallelt med forskellige kulturelle og subkulturelle adfærdstendenser, der har fået nyt liv på de sociale medier, har individets selvfremsstilling været problematiseret og underlagt forskellige former for kritik. Set i et metaperspektiv, har f.eks. Facebooks indsamling af brugerdata, sat privatlivets ret under skæmme, senest, blandt andet i relation til problematikken omkring *Cambridge Analytica* og deres anvendelse af brugerdata til at påvirke, f.eks. politiske valgkampe (Amer & Noujaim 2019). Samtidig er mange sociale medier blevet kritiseret for at fremtvinge en forvredet selviscenesættelse blandt aktørerne, hvor Twitter, Instagram og igen Facebook kan nævnes som eksempler (Marwick and Boyd 2011; Tiggemann and Zaccardo 2015; Twenge et al. 2018). Selvom reddit ingenlunde er gået fri af kontroverser og kraftig kritik gennem sin eksistens, adskiller platformen sig ved at være "totalt" anonym for aktøren på platformen (Kilgo et al. 2018:2–3). Der kræves ingen personlige informationer for at blive bruger, ikke engang en e-mail skal tilknyttes, hvorved aktøren er fritaget en personlig selvfremsstilling (ibid.). Derimod bliver der rettet fokus på indholdsdelingen og dens sammenhæng, opdelt i forskellige specifikke fora. Der opstår subkulturer og interessesammenslutninger, hvor den enkelte aktør kan tilpasse sin oplevelse og engagement ud fra egne interesser (Massanari 2016:4; Mills 2018). Det er derved interaktionen der kommer i fokus, frem for at være rettet mod de enkelte aktører i den store mængde af brugere på reddit's platform.

1.2 HVAD ER REDDIT

Reddit benytter sig af forskellige udtryk som ikke nødvendigvis findes andre steder. I dette speciale kan der henvises til slutnote, hvis uddybning af fagterminologi ønskes.¹

Reddit er et open-source online fællesskab, der har eksisteret siden 2005. Siden dets oprindelse har populariteten været støt stigende på nuværende tidspunkt rangerer det som den 5. mest besøgte hjemmeside i USA, og den 12. mest besøgte hjemmeside på verdensplan (Alexa.com 2019). Samtidig er det også den side blandt de mest populære, hvor folk opholder sig længst tid, gennemsnitligt set, og på nuværende tidspunkt er Reddit mere populær end både Twitter og Instagram i USA, med godt 330 millioner månedlige aktive brugere (ibid.). Selve siden Reddit er ikke meget i sig selv. Ganske simpelt designet, selvom den er blevet mere strømlinet gennem tiden, er der ikke meget flair eller bravur over sidens fremtoning. En enkelt hvid baggrund, der, hvis ønskes, kan ændres til mørk for bedre læsning i de sene nattetimer, og en utilitybar (UI) med søgefelt og forskellige justeringsmuligheder for brugeren udgør Reddits facade (Douglas 2014:329–30; Massanari 2016:4). Resten er brugerindhold der som standard fremvises for den nye bruger efter popularitet. Efterhånden som man bliver del af eksisterende subreddits og fællesskaber, eller skaber sine egne, ændres ens feed også, så det bliver tilpasset ens interesser og ønsker (Kilgo et al. 2018:2–3; Massanari 2016:4; Mills 2018).

Reddit blev skabt af Steve Huffman og Alexis Ohanian, mens de var bofæller på universitet i Virginia, den ene studerede computervidenskab og den anden studerede handel. Ifølge skaberne, repræsenterer Reddit alt det der er nyt og populært på internettet; et fællesskab af millioner samlet i underfællesskaber, hvor folk deler, diskuterer og stemmer på nogenlunde alt hvad der er interessant i verden (Roshan 2019). Navnet reddit er en sammentrækning af "read it" (har læst det), og blev valgt for at indramme udtrykket "i already read it on reddit", og for at understrege deres position som internettes forside (ibid.). Reddit bliver af samme grund i udgangspunktet skrevet med småt, som det også vil blive i dette speciale.

Reddit er svært at definere. Det er et onlinefællesskab, og et forum for forskellig- og ensartede mennesker kan mødes om interesser og holdninger. Men for nogle brugere er det også mere end det.

"(...) reddit is like life, it is what you make it. I think this is also true about Facebook, Twitter, etc. Anyone who complains about these sites sucking are basically mostly blaming their own decisions. If your Facebook experience is horrible, it's probably because your friends are boring. If your Twitter experience is lame, you're probably following the wrong people. Similarly, if you're having a bad experience on reddit, you might want to reconsider the subreddits you're subscribed to."

User PhoX jf. (Roshan 2019)

Reddit bliver af nogle brugere fremstillet som en slags metavirkelighed; et socialt opholdssted i internetalderen (ibid.). Brugernes delings- og skabelsesengagement, sammenholdt med reddit's mission om at være internettes forside, har gjort platformen til en form for central togstation for informationsflow i internetalderen, hvor indhold bliver spredt over den globale, webbaserede infrastruktur (ibid.). Platformen er baseret på en lang række individuelle diskussionsrum kendt som subreddit, hvor hvert enkelt emne har sin egen side, moderatorer og tilhængere. Rummene er skabt og bliver vedligeholdt af brugerne der ligeledes kan dele og kommentere på egne og andres indhold. Den overliggende platform har en enorm og alsidig brugerbase, der repræsenterer forskellige sociodemografiske og -økonomiske forskelligheder på tværs af både alder og lokalitet. Ligeså enorm og alsidig Reddit er, ligeså varierende og konstant udviklende er dets emneindhold på tværs af platformen (Roshan 2019). Folk kunne bruge, og bruger, ufattelige mængder af tid på at surfe gennem links, kommentarer og billeder, uden nogensinde at kradse i overfladen af informationsindhold på Reddit, der i en nærmest skræmmende eufemisme for universets tilstand, ekspanderer (eksponentielt og eksplosivt) hurtigere end man kan fatte. De sociale medier er i dag en fast og uadskillelig del af samfundets sammenhængskraft, og netop brugernes anonymitet på reddit præsenterer en socialitet, som på mange måder kan beskrives som tidligere uset, som minimum i volumen og tilgængelighed. Reddit fremstilles som internettets forside, og har sin eksistensberettigelse i brugernes interne interaktion og delingsiver. Reddit er dog stadig en økonomisk virksomhed, der som andre sociale medier søger profit (Page 2018; Roshan 2019). De følgende afsnit vil undersøge reddit's virksomhedsstruktur, og belyse hvilke forstyrrende elementer, der kan være i det sociale livs frihed på reddit.

1.3 REDDITS STRUKTUR

Reddit blev som benævnt stiftet i 2005, og opnåede bred anvendelse og popularitet i løbet af 2008. I sin tidlige opstart overtog Advance Publications ejerskab, og oprettede reddit som datterselskab. I 2011, opnåede platformen igen selvstændighed som uafhængig organisation, men stadig med Advance Publications i aktiemajoritet (Page 2018; Roshan 2019). Tilstanden som uafhængig organisation har dog gjort at de har kunnet arbejde med projektet reddit som onlinefællesskab, uden at skulle jage profitmaksimering. Hertil har det været påkrævet at ingen private investorer ejer over 1 procent af aktiemajoriteten (ELI5 2015; RedditOriginals 2013; Roshan 2019). Selv om reddit igennem sin eksistens har været en nonprofitorganisation, er der stadig et behov for at skabe omsætning, for at betale serverplads, medarbejdere, for at udvikle på platformen og forblive relevante. Denne indtjening er, som de fleste andre sociale mediesider, skabt primært gennem reklamer (Bhatt 2018; Page 2018; Roshan 2019).

Reklameelementet på reddit har siden lanceringen været et debatteret og betændt emne mellem brugere og administratorer af platformen, og det har derfor aldrig været noget, der har fået lov at fylde for meget i oplevelsen af siden (Page 2018; r/TheoryOfReddit 2018). Reddit begyndte at sælge reklameplads i 2009, efter siden begyndte at blive aktivt anvendt. Reklamepladsen er underlagt et regelsæt, der fastholder at reklamer skal være tydeligt markeret som reklamer, og at de ikke må være invasive for brugernes oplevelse (Page 2018; Reddithelp.com 2019; Roshan 2019). Som et kompromis til brugere der ønsker at være fri af reklameoplevelsen, har reddit blandt andet indført premium medlemskaber for interesserede brugere. medlemskabet giver, blandt andet, brugerne mulighed for at fravælge reklamerne i deres oplevelse af siden (Reddit.com/premium 2019). Premium medlemskabet, eller *reddit gold* som det hedder, fungerer dermed som en mellemvejsløsning for både reddit som administration, og for brugerskabets ønske om at forblive ureguleret og ucensureret af reklamer og virksomheder (Bhatt 2018; Matias 2018). For 6 dollars om måneden kan aktøren opnå relativ selvstændighed på platformen. Samtidig støtter de det sociale fællesskabs uafhængighed, og ved at kunne fravælge reklamer får brugerne direkte udbytte af virksomhedens indtjening på medlemskaber (Reddit.com/premium 2019; Roshan 2019)

1.3.1 Strukturernes påvirkning på det sociale liv

Strukturen og ejerskabet omkring reddit som internetfællesskab kan være meget afgørende for den relative frihed aktørerne oplever på siden. Mængden og graden af censur og kontekstforstyrrende elementer bliver i høj grad påvirket af hvilken mission og intention der ligger bag for virksomhedens ejerskab. I dette tilfælde er der meget der peger på at reddit, i hvert fald for nu, opretholder deres oprindelige intention om at skabe et *internetnets forside*; et rum for åben diskussion, nyheds- og indholdsdeling på tværs af sociale, økonomiske og geografiske skel (Kilgo et al. 2018:2–3; Massanari 2016:4; Roshan 2019). Selve reddit's fundament og eksistensgrundlag er dets brugere og deres interaktion med sidens platform, og både brugerne og virksomhedsledelsen er bevidste om dette. Forsøg på at moderere eller kontrollere brugernes frihed er i stor stil blevet mødt med kamp, og f.eks. reklamer er kun modvilligt blevet accepteret fordi der både har været adblockere og premium medlemskaber som alternativer (Kilgo et al. 2018; Matias 2018). For at kunne give den oplevelse brugerne efterspørger, og for at sikre at de bliver ved at med at benytte lige præcis reddit over mange andre internetfællesskaber og *news aggregators*, eller nyhedsdelere, er reddit nødsaget til at imødekomme brugernes ønsker. Efterhånden som reddit bliver støbt og formet af både brugere og virksomhed, opstår der strukturer som også brugerne må underlægge sig, for at kunne opretholde det fællesskab de selv har været medskabende i (Matias 2018; Mills 2018; Mills and Fish 2015; Page 2018).

1.4 HVORFOR R/UNPOPULAROPINION

Af de over én million emnespecifikke subreddits på den samlede platform, er det et relativt fåtal der opnår bred anvendelse og anerkendelse blandt den samlede brugerpopulation. Langt de fleste subreddits fremstår som nicher, henvendt til et mindre udsnit af reddit's brugere, alt efter interesse. Nogle subreddits, opstår som behov for specifik informationsudveksling, f.eks. med politiske bevæggrunde (Curran et al. 2016:186–87; Mills 2018). Mange kan ses som samtidshistoriske sociale fænomener, der opstår og forgår, efterhånden som folk mister interessen, og finder nye udtryksformer og emneinteresser. r/unpopularopinion kan nok bedst beskrives som et socialt fænomen, der udspringer af et behov blandt brugerne, for at udtrykke sig kritisk til den samfundsmæssige populærkultur. Underforummet er et af de ældre blandt reddit's mange fora, skabt i 2012, men har først vundet aktiv anvendelse omkring 2016. Selvom det er svært at vurdere præcist hvor populære og hvor mange der reelt anvender de forskellige subreddits, kan det antages, at r/unpopularopinion i dag er blandt de absolut mest populære og bredt anvendte (Baumgartner 2019; Subredditstats.com 2019). Engagementet på de forskellige fora fluktuerer og varierer fra dag til dag, og fra minut til minut, men undervejs i specialeforløbet har r/unpopularopinion konstant været omkring top-10 af mest aktive underfora på reddit's platform, med over 500.000 subscribers (ibid.). Den upopulære holdning som udtryksform, kan nævnes som et godt eksempel på reddit's udbredelse til andre sociale medier, hvor konceptet har spredt sig til såvel andre sociale medier som visuelle formidlingstyper (Adam 2018; Nembhard 2018). På reddit har r/unpopularopinion afgrænset sig ved at være tekstuel orienteret, og med et simpelt, men rammesættende regelsæt. Underforummet forholder sig til apolitisk og reelt indhold, og der bliver derfor modereret for blandt andet, spam, satire, had og polarisering (R/unpopularopinion 2019a). Forummets censur har fokus på en civil tone, mere end det holdningsmæssige indhold, og i udgangspunktet, vil posts der er originale og ordentligt formuleret ikke blive modereret, så længe de ikke er direkte hadefulde eller polariserende over for specifikke befolkningsgrupper, f.eks. racer, politiske eller seksuelle orienteringer (ibid.).

2 PROBLEMFORMULERING

Det udvalgte underforum r/unpopularopinion udmærker sig på mange måder til det intendede longitudinelle case-studie af kollektiv adfærd på det sociale medie, reddit. At forummet udelukkende er tekstuel orienteret, simplificerer en dataindsamling, da der ikke skal filtreres for andre indholdsformater, som billeder eller videoer. Der kan argumenteres for, at forummets alsidige anvendelse og stigende popularitet, kan siges at afspejle reddit's udvikling som helhed, og det er derfor en forhåbning, at r/unpopularopinion også tilnærmelsesvis kan afspejle den generelle og

gennemsnitlige brugeradfærd på reddit som samlet platform. Da det er en forholdsvis gammel subreddit, der i lyset i reddit's stigende popularitet har fundet nyt liv, er det en vurdering, at forummet har såvel veteraner som nybegyndere blandt den aktive brugerskare. Det populære emnemæssige indhold på forummet er defineret ud fra brugernes diskussion om, og tilslutning til, forskellige kontroversielle holdningsudmeldinger. Indholdsmæssigt, er der dermed mulighed for indblik i den generelle diskurs på forummet, hvordan brugerne forholder sig til forskellige diskussionspunkter, og hvordan de forskellige interaktionsmuligheder, der udgør reddit's platform, ændrer sig efterhånden som reddit's, og r/unpopularopinions, totale brugerpopulation bliver større og forandres. Dette speciales problemformulering kan dermed beskrives således:

Hvordan udvikler de dominerende termer sig på forummet r/unpopularopinion, i perioden 2016 til og med marts 2019?

Det empiriske materiale bearbejdes med udgangspunkt i årmæssige epoker, for den samlede observationsperiode. Igennem den analytiske proces, vil data derved behandles med udgangspunkt i 4 epoker, 2016, 2017, 2018 og 2019. På nær 2019, der forholder sig til årets første tre måneder, indgår det fulde år i de resterende observationsepoker. Specialets primære tilgang forholder sig til at belyse udvikling i indholdsdeling blandt brugerne på r/unpopularopinion. I forståelsen af internetsocialitet og interaktionen mellem brugerne på forummet, er tidshorisonten ligeledes vurderet et vigtigt element, specielt i bearbejdelsen af den observerede popularitetsudvikling for forummet og reddit som helhed. Denne analysedel bearbejdes ved hjælp af tidsserieanalyser, med udgangspunkt i arbejdsspørgsmålet: **Hvordan er brugerudviklingen på r/unpopularopinion i perioden 2016 til 2019.** Selvom det primære udgangspunkt for specialets empiriske indramning ligger i det tekstuelle univers, der bliver delt som holdningsudmeldinger på forummet, bygger en del af den analytiske forståelse af interaktionen mellem brugerne på tilgængelige, relationelle variable. De udvalgte indholdsrelaterede variable anvendes til at danne en dybereliggende indblik i hvordan det sociale liv udleveres på reddit og r/unpopularopinion. Dette element besvares ved hjælp af arbejdsspørgsmålet: **Hvordan udspiller interaktionen sig mellem brugerne på r/unpopularopinion i den indsamlede observationsperiode.**

Det tekstuelle univers der forefindes på forummet, bearbejdes i dette speciale med udgangspunkt i dominerende termer, som indikatorer for eventuelle forandringer og fluktuationer for debatten. Termer skal forstås som en metodologisk tilgang, der giver mulighed for at anvende en kvantitativ fundering over det tekstuelle indhold på r/unpopularopinions brugerflade. De opdelte epoker anvendes som grundlag for at bearbejde indholdet kvantitativt, og fremhæve bærende temaer for den store mængde af forskelligartet indhold, der udgør forummet r/unpopularopinion. Ved at anvende emneanalyse, der i videnskabeligt regi er benævnt *topic modelling*, er det intentionen at udtrække

værdifulde kerneinformationer fra den massive mængde af indhold der bliver delt hver dag (Blei, Ng, and Jordan 2003; Chang and Boyd-Graber 2009). Den empiriske afgrænsning kan på mange måder betegnes som en niche af den sociale virkelighed, både set i forhold til reddit's brugerpopulation, såvel som den generelle socialitet i internetalderen. Alligevel kan emneanalysen af de delte holdninger på r/unpopularopinion, være med til at sige noget om hvilke følgevirkninger der kan være i brugertilstrømningen til det sociale medie reddit, der ellers har været en notorisk underdog blandt internettets sociale opholdssteder. Analysen fungerer i praksis som en art *proof-of-concept* formulering, eller som en deskriptiv vurdering af det sociale liv på r/unpopularopinion. *Proof-of-concept*, fordi specialet formidler en metodetilgang, der bidrager til forståelsen af, hvad der foregår på forummet, blandt de mange brugere og interaktioner. Samtidig tilbyder specialet en metoderamme til at inddæmme internetsocialitet, specifikt rettet mod det sociale medie, reddit. Belysningen af udviklingen i emneorientering, kan viderebearbejdes, for eksempel gennem en diskursanalyse af emneanalysens fremhævede tematikker.

2.1 EKSISTERENDE FORSKNING

Følgende afsnit vil fremhæve undersøgelsens primære forskningsgrundlag, specifikt i relation til den eksisterende forskning på den sociale medieplatform, reddit. Som undersøgelsesobjekt, er reddit blevet anvendt til en myriade af forskellige formål, og med en lang række forskellige metodetilgange (Bilag 1). Dette afsnit vil søge at definere en grundlæggende forståelse for reddit's brugerpopulation, og hvilke udfordringer der kan være i forståelsen af det sociale liv der udspiller sig mellem mange forskelligartede aktører.

Reddit har, med udgangspunkt i det anvendte forskningsgrundlag, været et forskningsmæssigt interesseområde siden slutningen af 2010 og starten af 2011 (Bilag 1). Anvendelsen af reddit som forskningsobjekt tog fart fra omkring 2015, og er i stigende grad blevet fundet interessant på tværs af en bred række videnskabelige områder (ibid.). Dette speciale forholder sig til livet på internettet, og dets socialitet, som værende udgjort af en lang række forskelligartede elementer og konstruktioner, der hver især repræsenterer forskellige sociale virkeligheder for de engagerede aktører. Det sociale medie, reddit, markerer sig i denne kontekst ved både sin relative brugeranonymitet, såvel som sin åbne og løst censurerede platformstruktur (Kilgo et al. 2018:2–4; Leavitt 2015:318).

2.1.1 Det sociale liv i reddit's struktur

I forskningsmæssigt regi, er brugeranonymiteten blandt de mest primære interessante faktorer, der adskiller reddit fra andre sociale medier i samtidens internetliv. Adgangskravene for at blive bruger på platformen, og derved tilgå dets fora, er minimale, og stiller få eller ingen krav til personlige

oplysninger. Anonymiteten og de lempelige adgangskrav har, ifølge en del forskning, været medskabende til det sociale liv, der udspiller blandt brugerne på det sociale medie, reddit. Reddits strukturelle opsætning har været understøttende for, hvad der beskrives som *throwaway accounts*, eller ligegyldige og midlertidige brugerprofiler (Leavitt 2015; Pavalanathan and De Choudhury 2015:316). *throwaways*, er kendetegnet ved at være brugerprofiler, der er oprettet med det specifikke formål at skjule brugerens identitet, og give aktøren mulighed for at agere uden relation til sin opbyggede identitet, enten som reddit bruger, eller i den virkelige verden (Koerber 2014). Blandt de mange fora der udgør reddit samlede hele, er brugermængden generelt præget af en form for pseudoanonymitet, hvor nogle brugere ofte og offentligt refererer til deres personlige liv, og hvor andre er *throwaways*, eller aldrig afslører noget om deres liv uden for den sociale virkelighed på reddit (Pavalanathan and De Choudhury 2015:316–17). Reddits succes og popularitet som social medieplatform, bliver i mange tilfælde tilskrevet anonymiteten og den lave grad af censur der bliver opretholdt, på tværs af platformens fora. Brugernes identitet og opbygning af samme, bliver på reddit fritaget den personlige og relationelle fremstilling der, for eksempel gør sig gældende på medier som Facebook og Instagram (Marwick and Boyd 2011; Tiggemann and Zaccardo 2015; Twenge et al. 2018). I forskningen på området antages det, at anonymiteten appellerer til mange, da det tillader aktørerne at forholde sig til etisk problematiserede og socialt stigmatiserede debatområder, uden at person skal sættes under skue for deres udmeldinger eller holdninger (Leavitt 2015; Pavalanathan and De Choudhury 2015:317). F.eks. i arbejdet med brugeranonymitet og mental sundhed, er det blevet påvist, blandt andet, hvordan brugere, i mange tilfælde er mere ærlige i anonyme tilstande, og censurerer sig selv i relation til hvem de opfatter som modtagere af informationen (Bagroy, Kumaraguru, and De Choudhury 2017; De Choudhury et al. 2016:2108; Leavitt 2015:325; Pavalanathan and De Choudhury 2015:317). Af samme årsag, er der ofte fundet anvendelse af *throwaway accounts* i mental sundhed, og specielt i mere socialt stigmatiserede emneområder, hvor brugere ikke vil sætte deres personlige fremstilling i relation til deres udmelding (ibid.).

Reddit har i sin anonyme og relativt censurfrie tilstand været medskabende til en forholdsvis sammentømret brugerkultur der, blandt andet, bliver beskrevet som *participatory culture*, eller som en deltagende kultur, set i forståelsen af, at indholdet, og derigennem reddit's værdi, bliver skabt igennem brugernes interaktioner og relationer (Kilgo et al. 2018; Massanari 2017). Igennem forskningen af reddit's brugerplatform, er det blevet observeret hvordan anonymiteten i kombination med online kommunikation kan skabe en *online disinhibition effekt*, der fremavler, blandt andet, fjendtlig og *troll-like* adfærd. I denne forbindelse, er *trolls* en betegnelse for internetaktører, der bevist og med provokation som formål, på negativ vis deltager i online debatter og fora (Kilgo et al. 2018:2; Koerber 2014). Massanaris artikel fra 2017 fremhæver i et etnografisk studie af reddit, blandt andet

hvordan strukturen og den dominerende kultur fungerer som en neksus for toksiske teknokulturer og som en fødegrund for misogyne og antifeministiske tendenser (Massanari 2017:333). Dette bliver eksemplificeret gennem, blandt andet #GamerGate og #TheFappening, to hændelser, der blev populariseret og spredt gennem reddit's brugerflade (ibid.). Begge hændelser forholder sig til den ufrivillige deling af information og indhold om offentlige personligheder, specifikt i lyset af reddit's anonyme brugerflade. Massanari argumenterer for, at der i de to forløb fremstod en ekstremt misogyn hegemonisk kultur blandt den generelle brugerpopulation, hvor ethvert forsøg på at kritisere indholdsdelingen eller den nedladende og objektiviserende diskurs omkring den ufrivillige fremstilling af de implicerede, blev underkendt og nedstemt af den dominerende brugergruppering (Kilgo et al. 2018:2–3; Massanari 2017:335–36). Både #Gamergate og #TheFappening blev i forskningsmæssigt regi fundet interessant, fordi der var flere eksempler på, at brugere uden relation til problematikken indgik som "krigere" for antifeminisme og misogynitet (ibid.).

"Attempts to battle misogyny and racism in open systems like Reddit are undermined by the justification that the site "requires a significant tolerance for adversariality that may be alienating to some participants" (Reagle, 2013, para. 33). Using freedom of speech as a foundation, some users justify hostility toward others by adhering to an argument of choice equivalent with self-regulated exclusion (i.e., they don't like it, they don't have to be here)." (Kilgo et al. 2018:2)

I forlængelse af ovenstående, bliver der argumenteret for, at der opstår hvad der beskrives som et *odd paradox* i den deltagende kultur, hvor individer, med forskellige grupperinger som samlingspunkt, udfører adfærd de selv anser som nobelt, eller som minimum uproblematisk, selvom det på overfladen kan anskues som uetisk aktivitet (Kilgo et al. 2018:2; Massanari 2017:341–42). Den deltagende kultur på reddit har vist sig på mange måder og i mange former, og ikke kun med misogyne og hegemoniske tendenser, som anskuet af, blandt andet Massanari 2017 og Kilgo et al. 2018. I sin stigende popularitet over årene, har reddit's brugerpopulation opnået relativ stor indflydelse på den offentlige debat vedrørende forskellige emner og temaer, på tværs af reddit's fora. I eksemplet med #GamerGate og #TheFappening viste reddit sig som en kontrast til mere konventionelle og anerkendte sociale medier som Twitter og Youtube, specielt i forhold til den antifeministiske tone. I andre tilfælde, har dele af reddit's brugerpopulation engageret sig, og haft indflydelse på temadiskurser, i alt fra politik over populærkultur, til internetregulativer og online adfærd (Matias 2018; Mills 2018; Mills and Fish 2015)

I det ovenstående afsnit, er det forsøgt at indramme reddit's kontroversielle tilstand og udvikling, der igennem tiden har været med til at definere reddit som social medieplatform. Dette speciale arbejder mod at uddybe forståelsen for det sociale liv blandt brugerne på underforummet r/unpopularopinion,

med udgangspunkt i de dominerende termer for indholdsdelingen. Relationen mellem den eksisterende forskning omkring forskellige delelementer, såvel som reddit som helhed, bidrager til at forstå, hvad der kan betegnes som en generel adfærd blandt den samlede brugerpopulation, eller primære kutyper i den gængse interaktion. Her er det vigtigt at være opmærksom på, at reddit er et meget dybt og bredt netværk af brugere, interesser og emneorienterede nicher, der anfægter en hver form for holistisk afdækning af reddit som fænomen. Hvor den anvendte eksisterende forskning for det meste forholder sig til reddit eksistens frem til omkring 2016, tager dette speciale udgangspunkt i forløbet fra 2016, og frem til samtiden i 2019. Specialet tager ligeledes udgangspunkt i den figurative popularitetsekspllosion, reddit oplever som socialt medie i den givne periode. Selvom reddit er såvel berømt som berygtet for dets lave grad af censur, er der i dag indført visse kontrolsystemer for at overkomme de mest ekstreme udmeldinger og direkte illegal adfærd (Kilgo et al. 2018:3; Redditinc 2018). Herunder er det blevet forbudt at dele, blandt andet, børneporno og ufrivillig pornografi (ibid.).

3 TEORI

I det følgende afsnit, vil specialets teoretiske rammesætning blive gennemgået og sat i relation til de sociologiske og analytiske interessepunkter for specialets forskningsobjekt. Det teoretiske interessefelt i specialet, er fokuseret mod reddit's mange subkulturelle nicher og forskelligartede brugere og brugergrupperinger, og interaktionen imellem dem på reddit's platform. Forståelsen for reddit's platform som en naturlig eller selvfølgelig del af det 21. århundredes globale samfund, ses i dette speciale i lyset af den samfundsmæssige og samtidshistoriske udvikling. Det offentlige og brede samlingspunkt for nicher og specifikke interesseområder repræsenteret gennem reddit's platform, forstås i dette speciale, ud fra Michel Maffesolis neo-tribale subkulturteorier i *"The Time of the Tribes – The Decline of Individualism in Mass Society"* (1996). Maffesolis teoretiske arbejde fungerer som det gennemgående teoretiske analyseværktøj i specialets analyse, men der argumenteres i dette speciale for, at den subkulturelle forståelse, præsenteret af Maffesoli, må uddybes med andre teoretiske forståelser, for at indfange nogle af de bagvedliggende mekanismer i undersøgelsesobjektet. Dette ses ud fra, hvad Maffesoli selv beskriver som forskellige paradigmer i den samfundsvidenskabelige forskning. Hertil tages der udgangspunkt i Pierre Bourdieus teoretiske arbejde, som supplement og kontrast til Maffesoli. Derfor vil følgende teoriafsnit afdække såvel Maffesolis subkulturelle samfundsforståelse som specialets rammesættende afdækning af den intenderede forståelseshorisont.

3.1 OM PARADIGMER

Maffesoli tager selv afstand fra, hvad han beskriver som *Grand Theories*, eller forsøg på at definere det totalte samfund. Bogen er opstillet ud fra en forståelse for socialitet som forandrende, og ser variationer over tidligere sociologisk forskning, både som forskellige måder at observere det sociale, og som udtryk for forskellige paradigmer i den sociale virkelighed. Selvom Maffesoli bygger sine teoretiske bevæggrunde på en lang række sociologiske og filosofiske udgangspunkter, og trækker tråde tilbage til såvel Weber som Durkheim, bliver Maffesolis relation til Bourdieus samfundsforståelse, i dette speciale oplevet som den mest relevante at inddrage i den analytiske kontekst.

*“Although (la) **Distinction** is perhaps applicable to modernity, it is by contrast totally inadequate in explaining the varied forms of social groupings that are today at the forefront. Their outlines are ill-defined: sex, appearance, lifestyles – even ideology – are increasingly qualified in terms (‘trans’, ‘meta’) that go beyond the logic of identity and/or binary logic. Briefly, and taking the terms in their most accepted sense, we can say that we are witnessing the tendency for a rationalized ‘social’ to be replaced by an empathetic ‘sociality’, which is expressed by a succession of ambiances, feelings and emotions.” (Maffesoli 1996:11)*

I forlængelse af ovenstående citat, argumenterer Maffesoli i *Time of the Tribes* for, at hvor Bourdieus *la Distinction* kunne give stor forståelse for det moderne paradigme, var *Time of the Tribes* et udtryk for en del af den postmoderne samfundsforståelse (ibid.). Paradigmebegrebet skal ses ud fra en forståelse af samfundet og samfundets socialitet som varierende og forandrende sideløbende med den samtidshistoriske udvikling. Selvom der opstår nye realiteter og tilstande for samfundets medlemmer, og selvom der må opstå nye forståelses- og begrebsformer til at forklare tilstandene, bliver gamle forklaringsformer ikke derved forlagt (ibid.). Der argumenteres for, at samfundsmæssige tendenser må forstås i lyset af samtidens omstændigheder, men også ud fra hvad der har bragt samfundet til sin umiddelbare position. På grund af denne teoretiske forforståelse, og på grund af Maffesolis diskuterende tilgang, vil forholdet mellem de to teoretiske udgangspunkter blive anvendt såvel supplerende, som forståelseshorisont, men mest af alt som diskussionsgrundlag for de analytiske fund. På grund af analysens og reddit's datagrundlag, kan Bourdieus fulde begrebsapparat ikke anvendes i fuldt omfang. Både på grund af de specialets empiriske vilkår, men vigtigst på grund af brugernes anonymitet. Brugernes anonymitet, og reddit's struktur generelt opleves i dette som understøttende for Maffesolis teoretiske forklaringskraft i internetalderen.

3.2 NEO-TRIBALISM IN AN ORDERED WORLD

Reddits eksistensberettigelse og grundlag kan, med sikkerhed, tilskrives en lang række forskellige omstændigheder og vilkår, men præsenteres i dette speciale ud fra Maffesolis neo-tribale samfundsforståelse. Centralt for Maffesolis teoretiske univers, er forståelsen for det sociale, eller socialiteten, som dynamisk, fluktuerende og under konstant forandring. Derfor argumenterer han for, at ethvert forsøg på at gribe, indfange eller fastlåse et interesseobjekt til udmattelse, per definition er fejlagtigt. I stedet må forklaringen henholde sig til at beskrive objektets form, bevægelser, tøven, dets bedrifter og forskellige fluktuationer (Maffesoli 1996:5). Socialiteten som analyseobjekt, bliver i *Time of the Tribes* opstillet i interaktionen mellem aktører i samfundet. Det sociale liv er afgjort i forholdet imellem, og internt i, forskellige grupperinger og aktørinteraktioner, og hvordan samhandlingen udspiller sig. Fænomener der opstår og udspiller sig i det sociale liv, kan derved ikke tilskrives individuelle eller enkeltstående handlinger, men som tilstande der fremgår i forlængelse af, eller som kontrast til, tidligere og eksisterende tilstande.

" (...) just as there is no individual existence, there can be no singular thought. Our consciousness is but a meeting ground, the crystallization of various currents which, with precise moments of symmetry, intersect, attract or repel one another." (Maffesoli 1996:68)

Denne orientering mod den kollektive samhandling, og objektforklaringer som krystalliseringer af forskellige strømninger og bølger, der udspilles og udvikles i interaktionen mellem forskellige grupperinger og kulturelle tilstande, kan i Maffesolis forståelseshorisont ligeledes overføres til den overliggende forståelse for, hvordan civilisationer opstår og forgår, og hvordan konceptualiseringer af samfundet, er konstitueret af samtidige og refleksive forhold. I forlængelse af paradigmeforståelsen, skal Maffesolis teoretisk fundament, ses i lyset af såvel den postmoderne virkelighed, han præsenterer, som i det samfundsmæssige og sociale vilkår, der leder op til den oplevede virkelighed i det 21. århundredes internetalder.

"As we know, 'Rome is no longer in Rome' but at a certain point, it must measure itself against the other, that is, its heterogeneous empire." (Maffesoli 1996:109)

Hertil må det tilføjes, at Maffesoli anskuer enhver civilisation og ethvert samfunds opstandelse og eksistens som interaktionen mellem en heterogen masse. Massen skal, som kontrast til såvel den Marxistiske pøbel (versus bourgeoiserne) som den Bourdieuske klasseopdeling, ikke ses som udtryk for en kampplads eller som skue for hegemoni, men derimod som en abstrakt stjernetåge (*nebula*) af aktører der, uden afklaret centrum eller begrænsning, sammentrækker og forskubber forskellige og ensartede interesser, grupperinger og typer. Stjernetågens interaktioner mødes og eksploderer på tværs af den samlede masse; skaber nye former for socialitet og social adfærd, efterhånden som andre

forhærdes, forandres eller forgår (Maffesoli 1996:144–45). Maffesoli afgrænser sig i sin samfundsvidenskabelige forskning til nichesociologisk fokus på subkulturelle tendenser og de vilkår, der spiller ind i udviklingen af massernes socialitet i det postmoderne samfund. Han udformer sine teorier omkring adfærden blandt den brede befolkning i samtidens *megalopolises*; omkring det sociale liv blandt storbyens enorme netværk af aktører. Maffesoli beskriver hvordan de flygtige og midlertidige fællesskaber der opstår i storbyens liv, alt fra en joggingtur i den offentlige park eller en tur på den lokale café, til de sociale handlinger i menneskets færde på internettet, er repræsentative for en ny og gennemgående samfundsmæssig tilstand, han beskriver som en massestamme, eller som den neo-tribale kultur (Maffesoli 1996:76). Maffesoli argumenterer for, at der sket en forandring i samfundet fra, f.eks. 70'ernes bysociologiske afdækning af bander og gruppemiljøer i Chicagos undergrund, og tidligere samfunds tætte familieband, til en tilstand af mere åbne og midlertidige sammenslutninger. I det Maffesoli beskriver som samtidens single kultur, hvor en større og større del af samfundet er ugifte og uden tætte tilknytninger, er der opstået en tilknytning til mere objektorienterede interessefællesskaber (ibid.:139-40). Grundlaget for at anvende og modernisere stammekulturbegrebet som pendant til det samtidige, vestlige samfunds sociale og subkulturelle vilkår, sker ud fra en kontrast til tidligere samfundsforsknings positivistisk orienterede mekaniske eller individualistiske forklaringsformer. I stedet argumenterer Maffesoli for et mere organisk perspektiv, hvor det sociale liv kun kan determineres i relation til gruppen, eller stammen, enten direkte eller *a contrario* (Maffesoli 1996:79). I storbyens fysiske virvar og på internettets varierende fællesskaber og tilslutninger, argumenterer Maffesoli for, at der er udtryk for en massestammedialektik:

“(...) the mass being the all-encompassing pole, the tribe being the pole representing a particular crystallization. All social life is organized around these two poles in an endless movement; a more or less rapid movement, more or less intense, more or less stressful, according to the place and the people involved. In a certain manner, the ethic of the moment created by this endless movement allows us to reconcile the static (spaces, structures) and the dynamic (histories, discontinuities) that are generally seen as opposites.” (Maffesoli 1996:127).

Maffesoli argumenterer i forlængelse af ovenstående for en ny forståelse for *space-time* (rumtid), i forbindelse med samtidens *megalopoliser* og, måske mest relevant for dette speciale, internettets nye socialitetsformer. Den brede masses tilgængelighed til forskellige interessefællesskaber, og rum til at udføre og udforme disse, har skabt nye former for lokalitet, og samtidig fremhævet stedets betydning i den sociologiske forskning. Maffesoli argumenterer for, at der i samtiden er fremgået en antropologisk struktur, der forbinder enhver form for socialitet til en given lokalitet, eller *space*. Ifølge Maffesoli, bliver kerneforståelsen i hans type af socialforskning derved fokuseret på to centrale begreber – sted og socialitet (ibid.:131).

“This relates to my hypothesis of the underground centrality that characterizes sociality, hence the importance of the ‘genius of place’. This collective sentiment carves out a space, which has in turn an effect on the sentiment in question. This makes us aware of the fact that every social form is part of a pattern traced by the passage of the centuries, that it is an offshoot of this, and that the modes of living that constitute it can only be understood in terms of this substrate, in short, the entire thematic of the Thomist habitus or the Aristotelian exis.” (Maffesoli 1996:130)

Stedet skal derved, i Maffesolis optik, ses som et abstrakt samlingspunkt, mere end en specifik fysisk lokation. Stjernetågens krystalliseringer udtrykkes i, hvordan massens aktører samler sig i midlertidige og porøse stammer fokuseret på et fælles objekt. Samlingen om objektet skaber et rum, der tillige påvirker og udvikler den kollektive opfattelse af objektet. Vekselvirkningen mellem samfundets mange aktører, elementer og historiske udvikling, udgør derved, ifølge Maffesoli, de fundamentale vilkår for det sociale livs forfatning, skrøbelighed, og til tider, enorme kollektive påvirkningskraft:

“Just like chemistry, it is all a question of combination: through the differentiated association of elements, we obtain such and such a specific compound; but, with a minimal change or by moving an element, the entity may take on a new form. This is how, in the end, one social equilibrium passes into another.” (Maffesoli 1996:119–20)

3.3 TRADITIONERNES MAGT I STJERNETÅGENS KRYSTALLISERINGER

Selvom samfundets socialitet og sociale liv, i Maffesolis teoretiske univers, er determineret ud fra dets fluktuation og midlertidighed, afskriver han sig ikke fra, at der opstår og eksisterer mere rigide og længevarende kulturelle institutioner på tværs af massens stjernesky. Den historiske udviklingsproces, er essentiel for den neo-tribale samfundsforståelse, og socialitetens krystalliseringer, ses som udformninger af strømninger og samfundsmæssige bølgninger. Det sociale liv skal altså, ifølge Maffesoli, ses som et komplekst system af mange forskelligartede og -rettede aktører, interesser og objekter, hvis udtryk er afgjort af dets abstrakte lokalitet, krystalliseringernes interne og eksterne relation og interaktion, såvel som dets historiske udvikling og forankring. Netop den historiske disposition, i relation til dette speciales longitudinelle fokusobjekt, kræver indsigt i hvilke mere rigide samfundsstrukturer, spiller ind i de adfærdsmæssige normdannelse og diskurser der udspiller sig på forummet r/unpopularopinions. Hvor Maffesolis *Time of the Tribes* indrammer og konstituerer formen af socialiteten i internetalderens paradigme, tildeler den stadig værdi til tidligere paradigmers forklaringsformer i forståelsen af de mere dybereliggende mekanismer, som påvirker det sociale livs ekvilibrisme (Maffesoli 1996:11).

Trods store forskelligheder mellem Bourdieu og Maffesoli i deres teoretiske arbejde, er de begge orienterede mod interaktionen mellem det sociale livs aktører og strukturer, som det centrale for den sociologiske forskning. Bourdieu har i sin praktiske begrebsudvikling, orienteret sig mod en del af det sociologiske virke som en form for symbolsk kampsport, hvor interaktionen i både afdækningen og i udformningen af det sociale liv, udspiller sig som en konflikt; konflikt om plads, om dominans, om sandhed:

"Jeg siger gerne at sociologi kan være en form for symbolsk kampsport, der kan bruges som et forsvarsmiddel mod de forskellige former for symbolsk vold, som samfundets medborgere udsættes for specielt via mediefeltet, noget der er meget almindeligt nutildags." Pierre Bourdieu, Forelæsninger på Collège de France 2000-2001, jf. (Mathiesen 2005:127).

3.3.1 Hysterese og det sociale livs foranderlighed

Det sociale liv, og relationerne og interaktionerne mellem de forskellige aktører og institutioner i samfundet, bliver i Bourdieus arbejde forklaret ud fra en praktisk orienteret begrebsværktøjskasse, der fokuserer på at filtrere og segmentere flere elementer i det sociale livs konstellation. Selvom begrebsapparatet spreder sig udover en bred række udtryk, er de forskellige forhold, Bourdieu beskriver, internt relaterede. Samtidig påvirker de hinanden, og er gensidigt afhængige (Grenfell 2012:33, 1–2). Af samme grund kan det have problematiske og uheldige følgevirkninger at anvende enkelte begreber i uafhængig kontekst (ibid.). I dette speciale, vil Bourdieu nærmere anvendes i sin generelle og praktiske overvejelser om den sociale virkelighed, og derfor vil specifikke begreber eller teknikker fra Bourdieus begrebsapparat ikke anvendes. I socialforskningen har Bourdieus teoretiske univers haft stor indflydelse på samtidens forståelse og konceptualisering af menneskelig socialitet, og hans formuleringer af, for eksempel *habitus* og *felt*, er bredt anerkendt og anvendt på tværs af forskellige forskningsregi (Grenfell 2012:1; Mathiesen 2005:136–37). *Habitus* er defineret af aktørens sociodemografiske baggrund, og er fundamental for forståelsen af, hvordan individer positionerer sig i samfundet, ifølge Bourdieus perspektivering (Grenfell 2012:49–52). For dette speciales fokusområde; reddit's anonyme brugergruppering, er det selvsagt ikke muligt at danne forståelse for aktørernes sociale baggrunde, *habitus* og de resterende begreber, der udgør Bourdieus forståelseshorisont, men det vurderes undersøgelsesmæssigt, at Bourdieus samfundsforståelse til stadighed har en effekt på det sociale liv, der udspiller sig i internetalderen, i forlængelse af Maffesolis paradigmeforståelse. Selvom Maffesoli er den bærende teoretiske forståelse for dette speciale, argumenteres der for, at Bourdieus samfundsforståelse kan bidrage til indramningen af det sociale liv på forummet r/unpopularopinion på den anonyme sociale medieplatform, reddit. Dette ses i relation til Bourdieus overvejelser om det sociale liv, socialiteten og specielt forholdet mellem brugerkonstellationer og socialitetens *sted*, der udspringer af hans grundlæggende undren:

“all of my thinking started from this point: how can behaviour be regulated without being the product of obedience to rules?” (Grenfell 2014:49)

Bourdieu argumenterer for, at for at forstå det sociale liv og dets fænomener, er det ikke nok at anskue blot interaktionerne mellem de forskellige aktører og strukturer, men at *feltet* eller det sociale ståsted også må inddrages (Grenfell 2014:65). *Feltet* bliver i dette speciale anskuet som en pendant til Maffesolis perspektivering over *place* eller *stedet*, der, ifølge Bourdieu, ligeledes er determineret af dets historiske og relationelle kontekst (Ibid.). Derved argumenteres der, igennem dette speciale for, at Maffesolis kernebegreber om *socialitet* og *sted*, til dels, tangerer den overliggende forståelse i Bourdieus *habitus* og *felt* (Grenfell 2014:65-66, Maffesoli 1996:131). Maffesoli argumenterer for, at socioanalyser baseret på det moderne samfunds konstituering, ikke tilnærmelsesvis kan indfange det, han betegner som den neo-tribale kultur, på grund af den gennemtrængende tilstand af midlertidighed og foranderlighed, der gennemsyrrer samtidens storbykulturer (Maffesoli 1996:11). Selvom det kan siges at Bourdieus praktiske virke, primært har været orienteret mod samfundsvilkår i tiden før internettets sociale revolution, beskriver Bourdieu gennem begrebet *hysteres*, til dels, de samfundstendenser Maffesoli udpensler i *Time of the Tribes* (Grenfell 2014:126). *Hysteres* er et udtryk for den foranderlighed samfundet undergår i interaktionen mellem aktører og strukturer; mellem *habitus* og *felter*. Nærmere betegnet, er *hysteres*, ifølge Bourdieu, et udtryk for disruption mellem relativt stabile *habitus* og *felter*.

“(...) habitus changes constantly in response to new experiences. In other circumstances, and at times of crisis in particular, habitus must respond to abrupt, sometimes catastrophic, field changes, but that response always takes time. In such circumstances where new and stable field structures have yet to emerge, novel field opportunities, often transitory, come into being.” (Grenfell 2014:127)

Om reddit som platform er et transitivt eller mere permanent socialt fænomen, er stadig op til historiens udvikling at afgøre, men der kan uomtvisteligt argumenteres for, at internettet og de sociale mediers indtog har forandret det vestlige samfunds socialitet, og skubbet til hvad der kan observeres som *felter*, eller *steder*, i det sociale livs landskab (Maffesoli 1996:128-29). Selvom grundforståelsen for det neo-tribale paradigme er baseret på et opbrud i modernitetens stabilitet, er samtidens socialitet i internetsamfundet påvirket af det sociale livs *habitus*, eller kollektive hukommelse, som viderefører normer og strukturer på tværs af generationelle og paradigmemæssige skel (Maffesoli 1996:136-37). På trods af Maffesolis kritik af modernitetens sociologiske forståelse af *feltet* i samtidige socioanalyser, frasiger han sig altså ikke de stjernetågens krystalliseringer og de neo-tribale stammers relation til tidligere *felters*, eller *steders*, strukturelle interaktions- og adfærdsnormer.

3.4 NEO-TRIBAL KULTUR PÅ DEN SOCIALE MEDIEPLATFORM REDDIT

En signifikant determinant ved den neo-tribale kultur i det postmoderne samfunds storbyer og på internettets sociale medier, er *stedets* afgørende, men abstrakte konstruktion. *Feltet* er ikke længere afgjort af fysiske, lokale og historiske afgrænsninger på samme måde som under modernitetens fastlagte strukturer. Internettet og de sociale mediers brede tilgængelighed, har ændret opfattelsen af, hvad der kan konstitueres som et *sted*, eller et *felt*, og har haft stor påvirkning på det sociale liv, og den generelle opfattelse af socialitet (Maffesoli 1996:147–48). Det at være alene, er ikke længere ensbetydende med ensomhed, og selvom en større del i dag bor som singler end tidligere tider, har den enkelte aktør i den brede middelklasse, langt flere interaktions- og netværksskabende muligheder end nogensinde før (Maffesoli 1996:139). Reddit er en eksemplificering af såvel de nye tiders forandrede socialitetsformer, såvel som den forandrede *stedforståelse*. De mange subreddits repræsenterer hver for sig et specifikt og unikt *sted*, der dog på samme tid er underlagt den samlede platforms overliggende normer og adfærdsforventninger, der, i modsætning til storbyens kalas, ikke er fastlåst til en specifik lokation. Selvom det må påpeges at Maffesolis arbejde er fra 1996, et godt stykke tid før reddit's eksistens, og før de sociale medier for alvor bragte det private liv ud i den offentlige æter, muliggør Maffesoli i høj grad fremkomsten af en form teknologisk stjernetåge (*electronic nebula*); En virtuel virkelighed, i kombination af internettets virkemidler:

“(...) a communicational matrix in which groups with various goals will appear, gain strength and die; groups which recall somewhat the archaic structures of village clans or tribes. The only notable difference which characterizes the electronic nebula is of course the very temporality of these tribes. Indeed, as opposed to what is usually meant by this notion, the tribalism we are exploring here can be completely ephemeral, organized as the occasion arises” (Maffesoli 1996:139).

I forståelsen af Maffesolis udredning af det postmoderne samfunds subkulturelle tilstand, i relation til Bourdieus praktiske begrebsorientering mod det moderne paradigmes dybere mekaniske forklaringer, er det hastigheden i og flygtigheden af de sociale fællesskaber der eksisterer og opstår i internetalderen, der er afgørende for forskellen. Hvad Bourdieu oplever som *hysterese* eller social disruption, ser Maffesoli som etablering af en ny paradigmemæssig socialitet, orienteret mod det hastige liv i storbyen og på internettet (Maffesoli 1996:6-8). Hvor de *sociale klasser* og *kapitaler* i Bourdieus begrebsunivers, har stor indvirkning på aktørens evne til og mulighed for at begå sig i *felter*, som ofte er stærkt determineret og påvirket af dets historicitet og sociodemografiske position, er *stedet*, ifølge Maffesoli, langt mere orienteret med objektet, interessen eller den interaktion der finder sted (Grenfell 2014:116, Maffesoli 1996:6-8). Selvom der til stadighed er institutioner og *steder* der bærer det moderne paradigmes vilkår videre, er reddit's forskellige rum netop determineret af deres

relative flygtighed og deres tilstand af manglende rumspecifik historicitet (Ibid.). Reddits forskellige fora viderefører i mange tilfælde selvsagt emnerelaterede diskurser fra tidligere og lignende emnedebatter, men reddit's åbenhed og direkte demokrati gør samtidig, at adfærden har mindre at gøre med traditions- og lokalitetsbundne regelsæt, og mere at gøre med vekselvirkningen mellem de forskellige aktører i stjernetågen. Derved er mængden og graden af aktivitet på reddit's forskellige fora, afgørende for typen og formen af repræsenterede diskurser.

På grund af den store tilgængelighed og brede anvendelse, fremstår reddit som en unik strømning af krystalliseringer, der tilnærmelsesvis fremstår som et meta- eller parallelsamfund for et udsnit af af den samlede stjernesky's aktører. På samme tid, har reddit formået at opretholde en simpel men forankret struktur omkring adfærden på reddit, hvorved platformen, på trods af den store forskellighed i indhold på tværs af reddit's mange mikrostammer, fremstår som en samlet enhed. Reddit som platform, bliver i dette speciale, observeret som en organisk helhed, der er såvel påvirkende som påvirket af dets afgrænsede dele (Maffesoli 1996:6). Anvendelsen af underforummet r/unpopularopinion, sker med udgangspunkt i fokus på den diskursive interaktion og de gruppemæssige konstellationer, der former sig i vekselvirkningen mellem forskellige typemæssige repræsentanter af reddit's hele, og hvordan debatten udformer sig i et relativt åbent, emneafgrænset afsnit over tid.

4 VIDENSKABSTEORI

I det følgende afsnit vil den gennemgående videnskabsteoretiske fundering blive gennemgået og sat i relation til de sociologiske overvejelser og det empiriske grundlag. Den overliggende forståelseshorisont repræsenteret i det videnskabelige arbejde i dette speciale, tager udgangspunkt i den videnskabsteoretiske og sociologiske forståelse præsenteret af Maffesoli. I dette aspekt, forstås samfundet og socialiteten som samspillet og kombinationen af mange forskelligartede og varierende bagvedliggende mekanismer, der som kemiske konstruktioner, forandres og tager nye former i det sociale livs kontinuum (Maffesoli 1996:119). I forståelsen af det sociale liv der udspiller sig mellem brugerne på r/unpopularopinion, ses interaktionen som den bærende forståelseshorisont. Derved bliver de individuelle aktørers tilstand og ageren mindre vigtig, og der ligges større vægt på den observerede udvikling og samhandling i mødet mellem forskellige aktører og sammenslutninger (ibid.). Dette fokus på interaktionen, frem for de enkelte enheder, som præsenteret i Maffesolis teoretiske ramme, bidrager til forståelsen af internetsocialiteten på reddit, uden at møde konflikt i brugernes tilstand af anonymitet.

4.1 DEN KRITISK REALISTISKE VIRKELIGHED

Den oplevede virkelighed, der søges afdækket i dette speciale, tager i forlængelse af denne antagelse, udgangspunkt i den kritiske realisme. Den kritiske realisme er i sig selv bred betegnelse, der er blevet bearbejdet og anvendt på en lang række forskellige måder og udgangspunkter. For dette speciale, skal den videnskabsteoretiske retning forstås i sin kontrast til positivismen, og i sin objektivering af virkeligheden som dyb og lagdelt, der tangerer den organiske og helhedsorienterede forståelse, præsenteret af Maffesoli (Brante 2003:169; Maffesoli 1996:2–5). I sin kontrast til den positivistiske tankegang, tager den kritiske realisme afstand fra en empiristisk kausalitetsforståelse, hvor B altid følger som konsekvens af A. Derimod anser den kritiske realisme virkelighedens effekter og følgevirkninger som udtryk for en dyb ontologi, hvor verden eksisterer, uafhængigt af menneskelig påvirkning og afmåling (Buch-Hansen and Nielsen 2008:88). Hertil kan den dybereliggende mekanismeforståelse ses i relation til den kritiske realisme forhold mellem det ontologiske og det epistemologiske (Ibid.:22).

4.2 ONTOLOGISK OG EPISTEMOLOGISK VIDENSKABSFORSTÅELSE

I den kritiske realisme arbejdes der med en todelt forståelse for det videnskabelige objekt; den transitive og intransitive dimension (Buch-Hansen & Nielsen 2008:21). Den transitive dimension er determineret ved at være menneskets samlede viden om verden, og bliver eksemplificeret gennem den menneskelige, videnskabelige teoridannelsesproces, eller epistemologi:

”Teorierne er dermed videnskabens transitive objekt; de udgør den dimension som indirekte forbinder videnskaben med virkeligheden.” (Danermark et al. 1997:33; Lund 2017:19)

På den måde skal det samtidig forstås, at den objektiverede og observerede videnskab ikke er udtryk for det sociale livs dybe ontologiske virkelighed, men derimod en proces mod at forbinde den transitive og den intransitive; eller menneskets søgen mod dybere mening og forståelse, i det paradoksale forhold mellem epistemologi og ontologi (Buch-Hansen & Nielsen 2008:22). I denne forbindelse, er den intransitive dermed et udtryk for den dybe og komplekse sammenhæng mellem samfundets mange aktører, mekanismer og naturlige omstændigheder, og et udtryk for verdens *sande* konstituering; det videnskaben søger at afdække. Den kritiske realisme positionerer sig metodologisk ved at anskue forholdet mellem de to dimensioner, og observere og problematisere den paradoksale tilstand mellem menneskets videnskabelige *sandhed* og den dybere virkelighed, som udspiller sig uafhængigt af menneskelig handling (Ibid.:23). Den transitive dimension er, modsat den intransitive, historicitets-, sted- og kontekstafhængig (Danermark et al. 1997:34). I relation til det præsenterede teoretiske fundament, tangerer det forholdet til *stedet*, som en central del af den sociale interaktion

(Maffesoli 1996:129-139). I den forskningsmæssige bearbejdning i den *kritiske realisme*, er der derved en kontrast imellem den reelle virkelighed, og såvel den oplevede som objektiviserede virkelighed (Grenfell 2014:197). Forholdet mellem den transitive og den intransitive dimension, skal ses ud fra menneskets orienteringsmæssige position i verden, der per definition påvirker objektivisering af virkeligheden; den transitive dimension vil aldrig til fuldstændighed kunne afdække den intransitive dimension (Danermark et al. 1997:34). Med udgangspunkt i denne indramning af den kritiske realismes positionering, fremhæver Brante 3(4) hovedpointer i den samlede forståelseshorisont:

1. *There is a reality existing independently of our representation or awareness of it (ontological postulate)*
 - 1a. *There is a social reality existing independently of social 'scientists' representation or awareness of it (ontological postulate for social sciences)*
2. *It is possible to achieve knowledge about this reality (epistemological postulate)*
3. *All knowledge is fallible – and correctable (methodological postulate)*

(Brante 2001:172).

I forlængelse af denne indramning, bearbejder den kritiske realisme virkeligheden ud fra en forestilling om 3 eksisterende domæner: det *empiriske*, det *faktiske* og det *virkelige* (Buch-Hansen & Nielsen 2008:24). Det *empiriske domæne* forholder sig til den erfarede virkelighed. Det er det direkte beskrevne og observerbare, der udspiller sig i det empiriske. Det *faktiske domæne* forholder sig til den udlevede virkelighed, fænomener og begivenheder der, upåagtet af forskningens opmærksomhed, opstår og forgår i den sociale virkelighed. Det *virkelige domæne* er et udtryk for den dybe ontologiske virkelighed, som med udgangspunkt i ikke observerbare, underliggende strukturer og mekanismer, flyder over i, og påvirker det *faktiske* og *empiriske domæne*. Hvor f.eks. positivismen primært forholder sig til det observer- og målbare i de første to domæner, adskiller den kritiske realisme sig ved at forholde sig dybdegående til forholdet mellem samtlige tre domæner, og mere specifikt orientere sig mod det *virkelige domænes* påvirkning på de resterende to (Ibid.). Hermed kan det beskrives hvordan sociale fænomener udtrykkes som krystalliseringer i større strømninger og bølginger der spreder sig og forandres på tværs af massens stjernesky, og hvordan mødet mellem forskelligartede kombinationer af aktører, institutioner, strukturer og mekanismer udformer det sociale livs forandring og samtidige ekvilibrisme på tværs af tid og rum (Maffesoli 1997:118-119). Den kritisk realisme fastholder, i dette speciale i relation til stjerneskyens abstrakte kompleksitet, virkeligheden som bestående af åbne og interagerende systemer, hvorved fænomener aldrig opstår

arbitrært, men i effekt af de ikke observerbare strukturer og mekanismers samspil i det *virkelige domæne*, som flyder over i det *faktiske* og det *empiriske domæne* (Buch-Hansen & Nielsen 2008:26).

4.3 ABSTRAKTION OG RETRODUKTION

Den metodologiske anvendelse og bearbejdning af den kritiske realisme i videnskabeligt arbejde, forudsætter en åben virkelighedsforståelse, hvor der søges forbindelse mellem oplevelser i det *empiriske* eller *faktiske domæne* over i det *virkelige domæne*. Ved at søge at forklare nogle af de bagvedliggende mekanismer, strukturer og relationer ved det oplevede fænomen, bliver anvendelsen af såvel induktive som deduktive tilgange problematiseret i deres dogmatisk empiriske orientering, og i stedet anvendes der i den kritiske realisme ofte retroduktion som metodologisk forståelsværktøj til fænomener i de *faktiske* og *empiriske domæner* (Buch-Hansen & Nielsen 2008:60-61). Retroduktion og i denne forbindelse også abstraktion, forstås ud fra, den kritiske realisme forsøger at indfange nogle af de generative mekanismer for undersøgelsesobjektet. Med udgangspunkt i den oplevede virkelighed i det *empiriske* eller *faktiske domæne*, bliver forskningen orienteret mod at grave under de observerede, faktuelle fænomener. Undersøgelsesobjektet bliver altså ikke orienteret mod at forstå og forklare fænomenet til fuldstændighed, så meget som mod at formulere og indramme nogle af de dybereliggende strømninger og bølgninger der påvirker objektets umiddelbare tilstand. Anskuelsen af forbindelserne mellem de observerede og faktiske fænomener mod de dybereliggende generative mekanismer søges gennem abstraktion i den kritiske realisme, at bidrage til forståelsen af det *virkelige domæne* (Ibid.). Derved bliver den historiske, samtidige og relationelle afdækning af reddit's eksistensgrundlag, såvel som dets brugerpopulation, i relation til den teoretiske forforståelse, anset som abstraktionsgrundlag for at forbinde den observerede virkelighed i den indsamlede empiri til de generative mekanismer og samfundsmæssige strømninger og bølgninger der udformer sig i samtidens og internettets sociale liv. Herigennem er det hensigten, at der kan opnås dybere indsigt i udformningen af virkelighedens 3 domæner, i forhold til adfærden på den sociale medieplatform reddit's underforum r/unpopularopinion.

5 METODE

Den empiriske og analytiske bearbejdning er i dette speciale foretaget med udgangspunkt i programmeringssproget Python (MindfireSolutions 2017; Python 2019). igennem metodeafsnittet vil der blive forholde sig til forskellige, primære elementer af databearbejdningen, hvor udvalgte teknikker og metoder vil blive fremhævet, i relation til deres anvendelse i specialet. Dataindsamlingen er baseret på et udtræk af informationer fra platformen reddit, og afsnittet om dataindsamling vil

derfor også inddrage overvejelser over hvilke data, der antages som fordelagtige i besvarelsen af, hvad der kan være påvirkende på udviklingen i dominerende termer på r/unpopularopinion igennem den udvalgte observationsperiode.

5.1 DATAINDSAMLING

Empiriindsamlingen har i dette speciale været orienteret mod det longitudinelle aspekt, i afdækningen af udviklingsperspektivet, reddit undergår mod bred anvendelse og popularitet. Desuden har der fokus på *data mining*, eller indsamling af empiri, baseret på den aktive interaktion, direkte fra reddit's datagrundlag. Til denne proces er der blevet anvendt en såkaldt *API Wrapper*, igennem Pythons sprogflade. En *API* kan forstås på flere måder, men i dette speciale, anvendes det som en metode til at kommunikere med og indsamle empiri fra reddit's platform, uden at tilgå siden fysisk (Boe 2016; Boe et al. 2016; Gazarov 2016). Da reddit, ligesom Python, er open-source, er der stor tilgængelighed til deres datastruktur, og reddit's administration opfordrer ligefrem virksomheder og andre interesserede, til at tilgå platformen ved hjælp af deres, til formålet strukturerede *API* (Wardle 2015). Selvom dataindsamlingen i dette speciale forholder sig til et "uudforsket" område, i den forstand at det givne datamateriale ikke forefindes andre steder, er processen ikke ukendt, og andre har haft til formål at indsamle specifikke elementer af reddit's datafundament på samme måde som der ønskes i denne udarbejdning (ibid.). I dette speciale anvendes Pythonprojekterne *Pushshift* og *PRAW*. *Pushshift* fungerer som en database over den samlede indholdsdeling på reddit, der giver mulighed for at tilgå større dataudtræk uden at forbyrde sig mod reddit's regler om indsamlingsmængde, og fungerer i syntaks med reddit's *API*, og *PRAW* (Baumgartner 2019; Reddit.com/pushshift 2019; Wardle 2015). *PRAW* er et akronym for *Python Reddit API Wrapper*, tilegnet anvendelse i Python, og er designet som et interaktivt interface, hvor reddit's generelle regler og vilkår er inkorporeret (ibid.). I reddit's massive mængde af posts, interaktioner og metainformationer, giver denne indsamlingsproces mulighed for at indfange nogle af de mange strømninger, der ellers ville blive tabt, men tildeler også en vis kvalitativ udvælgelsesproces til undersøgelsesobjektet (Boe 2016; Boe et al. 2016).

5.1.1 r/unpopularopinion som observationsobjekt

Med udgangspunkt i reddit's totale platform, vil en kvantitativ dataindsamling, specielt i et longitudinelt perspektiv, hurtigt blive forpurret af den enorme masse af posts, kommentarer, upvotes, downvotes, og andre interaktionsformer der deles på tværs af den brede og dybe flade. På grund af de mange forskellige typer af indhold, f.eks. billeder, videoer, links, som bliver delt og fremkommer på platformens populære hovedforsider, ville anvendelse af disse, som udgangspunkt for analyse, være problematiseret af den meget forskelligartede indholdsdeling. Den relative ensartethed i den typemæssige indholdsdeling, som præsenteret på r/unpopularopinion giver mulighed for at anskue

den generelle brugermængde på forummets interaktion i et mere direkte lys, end hvis man benyttede f.eks. reddit's *hot* (populære) forside. I antagelsen af reddit som en form for *electronic nebula*, eller en art virtuel storby, fremstår de mange subreddits som neo-tribale stammer der, med sine egne lovmæssigheder og regler, stadig indgår som en del af det organiske og fluktuerende hele (Maffesoli 1996:139). I denne forståelse, ses reddit's underfora, som defineret af de bagvedliggende strømninger og bølger, der former og danner skyen. Selvom det igen må understreges, at den samlede masse ikke kan forstås med udgangspunkt i et mindre udsnit, er det forhåbningen at der, ved at anvende en såvel bred som dyb subreddit, der opretholder en specifik forventning til forummets brugere, kan opnås indsigt i en generel brugerinteraktion og -udvikling. I forlængelse af dette, opretholder forummet en apolitisk, og lighedsorienteret debat, hvormed det kan forventes at tendentielle og epokedefinerede tilstande i den overliggende brugeradfærd, til dels kan tilskrives forandringer eller udvikling i den samlede brugermængde (R/unpopularopinion 2019a). I den neo-tribale stammekultur, er der stor vandring mellem de forskellige stammer, og midlertidigheden præger mange sociale fællesskabers tilstand (Maffesoli 1996:76). Blandt brugerpopulationen på reddit, kan det antages at en stor del af brugerne, tilgår mange forskellige stammer, eller subreddits, og at foraenes popularitet kan være et udtryk for krystalliseringer i mødet mellem forskellige strømninger og bølger (ibid.). r/unpopularopinion har i analysens anvendte observationsperiode opnået relativ stor brugertilstrømning, uden at appellere til en specifik brugergruppering (R/unpopularopinion 2019b; Subredditstats.com/unpopularopinion 2019).

5.2 BRUGERPOPULATIONEN PÅ REDDIT

Selvom anonymiteten udfordrer afdækningen af reddit's samlede brugerpopulation, er der foretaget en række undersøgelser og surveys, der har skabt en grundlæggende forståelse for demografien for reddit's aktive brugergruppe, der dog skal tages med forbehold for reddit's popularitetsmæssige udvikling (Sattelberg 2019). Det antages, at omkring 50 procent af brugerpopulationen er fra Nordamerika, USA og Canada, og de resterende er spredt over resten af verden, primært i vestligt orienterede, engelskkyndige kulturer, med Australien og England som nærmeste følgere til Nordamerika (ibid.). Selvom engelsk er det mest populært anvendte sprog på reddit, findes der en lang række subreddits hvor andre sprog er primært anvendt. Historisk, har reddit kønsmæssigt været domineret af mænd, hvor det i nogle undersøgelser har været vurderet til over 70 procent (ibid.). Selvom der har været en udvikling i brugerbasen, kan det stadig antages, at der på tværs af reddit's platform er en overvægt af mænd. Udover få spørgeundersøgelser foretaget i USA, er der meget lidt information om den etnicitetsmæssige fordeling af reddit's brugerpopulation, men det er forventeligt, at der er et flertal med kaukasisk etnicitet, op mod 65 procent. Den næstmest repræsenterede

kategori er brugere med latinske rødder. Aldersmæssigt er reddit klart mest anvendt blandt yngre generationer, dog under lovmæssig forudsætning af en minimumsalder på 13 (Sattelberg 2019). En stor del af brugerne er under uddannelse eller tidligt i deres arbejdsliv. Selvom der er bred variation i den totale brugerpopulation på reddit, kan det antages de yngre generationers brede middelklasse, er stærkest repræsenteret (ibid.). Det må understreges at reddits population med udgangspunkt i denne undersøgelse, ikke kan afdækkes hverken fyldestgørende i sin udvikling eller som et stillbillede af den nuværende tilstand. I dette speciale er det antagelse, at reddits population forandrer sig i den udvalgte tidshorisont. Samtidig kan der være stor variation på tværs af de forskellige subreddits, hvormed de gennemgående tendenser ikke nødvendigvis gør sig gældende i specifikke fora. I forhold til at forstå udviklingen af dominerende termer, anvendelsen og populariteten af r/unpopularopinion, vil der blive taget udgangspunkt i en brugermæssig for forståelse, hvor det forventes at forummet i sin tidlige opstart er køns- og etnicitetsmæssigt domineret af yngre kaukasiske mænd. Denne vurdering er baseret på den aktive brugerskare blandt reddits samlede population, men reddits åbenhed gør, at der kan være mange forskellige aktører og brugertyper der oplever interaktionen på de forskellige fora.

5.2.1 Aktørerne på reddit

Som nævnt, er der meget få opstillede krav for at få adgang til reddits indhold. Gennemgående kan det siges, at reddits brugerskare er udgjort af to forskellige typer: Den ikke deltagende, og den deltagende. Den ikke-deltagende aktør på reddit kan i princippet indebære alle der benytter sig af internettet. Om man har oprettet sig som bruger på reddit eller ej, kan man relativt frit få adgang til reddits fora, links og nyhedsdelinger. Tilstanden som tilskuer er populært betegnet som "lurkers", og udgør generelt set størstedelen af brugermængden på enhver given social medieplatform (Garfield 2016; Koerber 2014; Nielsen 2006; R/dataisbeautiful 2019). Både tilmeldte brugere og anonyme udefrakommende kan være lurkers, hvilket gør det svært at få et overblik over præcis hvor mange unikke besøgende, der dagligt benytter reddits platform. Den ikke-deltagende bruger af reddit kan defineres, som én der læser og observerer interaktionen på platformen uden selv at indgå, hverken med votes eller posting (ibid.). De deltagende aktører på reddit falder alle under bruger kategorien. For at kunne interagere med siden med votes og posts skal man have oprettet en bruger. Mellem de deltagende brugere, er der bred variation i aktivitetsniveau og deltagelse. Ifølge en undersøgelse foretaget på 301 unikke brugere over en periode på et år, er der en gennemgående passivitet blandt reddits aktive brugerskare, og det store overtal forholder sig relativt tavst, og benytter primært votingsystemet til at lade deres mening blive hørt (Glenski, Pennycuff, and Weninger 2017; Grigonis 2018; R/dataisbeautiful 2019; Smith 2019).

Fordi platformen er i konstant forandring og udvikling, er det svært at danne et præcist overblik over hvor mange brugere og posts der månedligt gennemstrømmer siden, men ifølge reddit selv, drejer det sig om cirka 330 millioner unikke brugere og 11 millioner posts (R/dataisbeautiful 2019; RedditOriginals 2013; Smith 2019). Selvom forklaringerne på brugernes gennemgående passivitet kan være mange, peger flere forskere på, at mange brugere positionerer sig tilbageholdende på platformen, og først efter længere tid som passive, deltager i fora og debatter. En del af forklaringen herpå kan være, at man som passiv deltager kan opleve de sociale reglementer og adfærdsnormer uden selv at risikere konsekvenser af normbrud, og uden at sætte sine egne holdninger på højkant for omverdenens bedømmelse (Glenski et al. 2017; Mills and Fish 2015). reddit's brugere er både dømte, jury, dommer og bøddel for det delte indhold på reddit's platform, og hvad der fanger massernes eller subgruppernes opmærksomhed og opnår popularitet, kan variere på tværs af, f.eks. tidspunkt, emneindhold, formidlingsform og tilsyneladende humør blandt brugerne (Literat and van den Berg 2019:237–39). Alle brugere på reddit er princippet i stand til at oprette en ny subreddit. I 2017 blev der målt omtrent 1,2 millioner subreddits der hver forholder sig til et bestemt emne. Mængden af subreddits er eksploderet siden 2008, hvor der var cirka 10.000 (Richter 2017). Langt de fleste af disse er dog små subreddits med få aktive brugere. Ens for dem alle er, at de er modereret af brugerne selv. Brugeren der opretter en given subreddit, har magten til at sætte regler og normer for siden, og kan efter forgodtbefindende tildele magten til andre interesserede. Populære emner, hvor subreddits kan vokse sig til veritable kollektive bevægelser, kan resultere i, at moderatorer får en betydelig relativ magt for diskursen inden for et givent område (Curran et al. 2016:92–95; Matias 2018; Mills 2018). Dette skal dog ses i lyset af den frihed brugerne har på platformen som helhed. Der er intet der afholder brugere fra at oprette en ny subreddit om et emne, hvis moderatorerne misbruger deres magt. Brugerne i ligeledes fritstillet i deres valg af hvilke og hvor mange subreddits de vil subscribe eller tilknytte sig til, og den personlige oplevelse af hvad brugeren oplever på reddit som platform, varierer markant ud fra individuelle valg. Den enkelte brugers oplevelse vil dog på den umiddelbare forside være påvirket af, hvad den samlede brugerpopulation er mest optaget af, da meget populære posts får en mere fremtrædende position på forsiden, uagtet personlige præferencer (Mills 2018:40–41).

5.2.2 Praktisk dataudvælgelse

Selv med afgrænsningen omkring r/unpopularopinion som undersøgelsesobjekt, kan der tilknyttes en lang række meta- og kerneinformationer til det tekstuelle univers der udspiller sig. Kort opsummeret, er indholdsdelingen i r/unpopularopinion baseret på en tekstuel holdning, der hvis ønskes kan uddybes i en forklarende tekst (R/unpopularopinion 2019a, 2019b). Andre brugere kan kommentere eller stemme på holdningen, og på den måde kritisere eller sympatisere udmeldingen. For hver

holdning er der informationer tilknyttet omkring blandt andet brugeren, votes, en samlet score for holdningen, etc. På grund af den store mængde af såvel observationer som tilgængelig information, har der, i relation til specialets tids- og emnemæssige afgrænsning, være flere datafravalg og overvejelser. For nærmere overblik over de tilgængelige informationer for observationerne kan bilagene til dataindsamling anvendes (Bilag 2, Bilag 3). Det mest prominente fravalg i forhold til at opnå dybere forståelse for interaktionerne, er de specifikke holdningers kommentarspor. Dette ville forordre en anderledes analysetilgang end den anvendte i specialet, men ville samtidig være mere fyldestgørende i sin kontekstuelle forståelseshorisont. I stedet, har dette speciale taget sit udgangspunkt i udviklingen omkring hvad der på reddit konstitueres som en upopulær eller kontroversiel holdning, og vil forsøgsvis anvende variable, der kan indramme udviklingen, specielt i forhold til de relaterede brugerinteraktioner. Samtidig anvendes den indsamlede empiri til at give et indblik i adfærden omkring de forskellige diskursive strømninger gennem undersøgelsesobjektets tidslinje. Hertil er der udvalgt 7 variable, der hver især kan bidrage til indramningen af de enkelte holdningers relationelle position i massen af udmeldinger og interaktioner. Som den primære variabel, er den enkelte holdning blevet inddraget. Muligheden for at inddrage den uddybende tekst er fravalgt, da det er en opsat forventning på forummet, at holdningen skal være udtrykt klart og uden sarkasme (R/unpopularopinion 2019a). Det må derfor forventes at udmeldingen kan observeres uden behov for uddybning. De resterende 6 variable for hver observation forholder sig til; *bruger*, *delingstidspunkt*, *upvote*, en *upvote ratio* (der viser forholdet mellem up- og downvotes), *kommentarmængde* (antal kommentarer tilknyttet observationen på indsamlingstidspunktet), samt en samlet *score* (Bilag 2, Bilag 3). Reddit holder det præcise grundlag for den samlede score skjult, men det kan antages at være en kombination af blandt andet *upvote ratio*, antallet af kommentarer, og interaktionen i kommentarsporet (Coldewey 2016; Salihefendic 2015). I kombinationen af disse udvalgte variable forventes det, at eventuelle svingninger og forandringer på r/unpopularopinion kan indrammes med udgangspunkt i interaktionen mellem forummets brugere. Ved at centrere omkring f.eks. kontroversielle, populære, eller negativt ombruste holdninger undervejs forummets udvikling, kan det anskues om der er tendentielle diskurser eller dominerende grupperinger i forskellige epoker af undersøgelsesobjektet.

5.2.3 Analysens datasæt

Analysens anvendte datasæt tager udgangspunkt i ovenstående variable, og er observationsmæssigt den samlede interaktion på r/unpopularopinion, over en periode på omtrent tre et kvart år, fra januar 2016 frem til april 2019. I dataindsamlingsprocessen, blev det observeret, med udgangspunkt i den anvendte tilgang med PRAW, at r/unpopularopinion oplevede meget lidt aktivitet, i perioden før 2016. Det har ikke været muligt at teste om den manglende aktivitet tangerer virkeligheden for perioden,

eller om det er en konsekvens af manglende data i indsamlingsprocessen, men gennem den anvendte *PRAW API*, observeres der kun 10 holdninger i perioden 2012 frem til starten af 2016. For perioden 2016 og frem, er der i kontrast langt mere stabil og konstant aktivitet på forummet (Bilag 2, Bilag 3, [Subredditstats.com/unpopularopinion](https://subredditstats.com/unpopularopinion) 2019).

Selvom der foretaget fravalg i forhold til det variabelmæssige indhold, har det været en vigtig del af indsamlingen, at samtlige observationer i det givne forum inddrages. Derved indfanges både de posts der opnår popularitet og anerkendelse, såvel som dem der forsvinder i massen og bliver glemt. I praksis, har dataindsamlingsprocessen samtidig visualiseret, hvor stor en mængde brugere og aktivitet der reelt tilgår *r/unpopularopinion* i observationsperioden, og for at indfange denne variation, vil analysen som udgangspunkt behandle data i epoker. Samlet set består det empiriske grundlag for dette speciales analyse af: 7 variable og ca. 300.000 observationer, ulige fordelt over de repræsenterede år. Som kontekst er der omtrent 5000 observationer i 2016 mod næsten 115.000 på 3 måneder i 2019 (Bilag 4, Bilag 7).

5.3 ANALYSESTRATEGI

Afsnittet om analysestrategi vil formulere undersøgelsens tilgang til den indsamlede empiri, og vil indramme de forskellige metodemæssige overvejelser og valg i besvarelsen af specialets problemformulering. Afsnittet har til formål at diskutere og argumentere for anvendelsen af analysens modeller og metoder, og de anvendte analyseredskaber vil blive fremhævet hver for sig, og belyses med henblik på eventuelle styrker og svagheder i forståelsen af undersøgelsesobjektet. Specialet er analytisk opbygget efter to hoveddele: En deskriptiv analyse, samt en emneanalyse. Begge analyser forholder sig til det samlede observationsobjekt fordelt over epoker, med år som delingskriterie. Derved fremstår der 4 enkeltstående datagrundlag, som hver især vil blive gennemgået og belyst. Epokerne der indeholder årene fra 2016 til og med 2018 er samtlige 12 måneder, hvorimod den sidste epoke indeholdende 2019, kun har observationer for de første 3 måneder. Den deskriptive analysedel består af en række analytiske redskaber til at formidle både udviklingen i epokerne, samt indholdet i datagrundlaget. Indledningsvist anvendes tidsserieanalyser til give et overblik over den aktivitetsmæssige udvikling for hver epoke. Dette opfølges af en univariat analysedel, der belyser empiriens udvalgte variable, og sætter dem i relation til det tekstuelle univers i interaktionen på *r/unpopularopinion*. Den deskriptive analyse opfølges af en emneanalyse, bedre kendt som *topic modelling*, der anvender en type maskinlæring til at sortere indholdsdelingen i forskellige klynger, baseret på deres interne relation (Blei 2012:77). I det følgende afsnit, vil de anvendte analytiske redskaber blive gennemgået og frembragt i forhold til deres anvendelse i analysen.

5.3.1 Tidsserieanalyse

Tidsserieanalysen vil præsentere den udvikling r/unpopularopinions undergår i observationsperioden, og forsøge at indramme den generelle udvikling i brugerpopulationen. Til dette vil der anvendes to visualiseringsmodeller, et *scatterplot* og en *rolling average*.

I fordelingen mellem to variable, også kaldet en bivariat analyse, anvendes der i det deskriptive ofte *scatterplots*. Modellen fungerer ved at fordele observationerne i et datasæt på en x- og en y-akse. I dette speciale, anvendes modellen til at visualisere den tidsmæssige udvikling, og vil derfor repræsentere tidsdimensionen på x-aksen. På y-aksen visualiseres aktivitetsniveauet, der tager udgangspunkt i et gennemsnitligt dagligt niveau i mængden af indholdsdeling. Analysens *scatterplots* vil derfor gennemgående fremstå som en todimensionel fremstilling af den tidsmæssige udvikling, fordelt over 365 observationer (Agresti and & Finlay 2009:57–59).

Rolling, eller *moving average* (RA), er et udtryk for en tidsseriemodel der aggregerer observationerne, og giver et gennemsnitligt overblik over den tidsmæssige udvikling. Med samme princip som *scatterplottet*, inddeles observationerne i udvalgte indramninger, og udglattes med en linje. I stedet for at se de enkelte observationer i datamængden, fremstår der, i dette speciale, derved en lineær relation mellem tid og aktivitetsniveau (Datacamp.com 2018b; Vincent 2017). Her kan det desuden noteres, at den anvendte aggregering i analysen er baseret på et ugentligt gennemsnit. RA modellen giver mulighed for yderligere forståelse gennem dens 3 primære visualiseringslinjer. En generel linje, der viser den gennemsnitlige udvikling, samt to andre, én øvre og én nedre, som viser henholdsvis den højeste og den laveste værdi i observationsvinduet (Ibid.)

For at belyse den gennemgående udvikling af aktivitetsniveau for r/unpopularopinion, vurderes postmængden som relativt fyldestgørende for formidlingen af interaktionen på forummet. Selvom den sociale interaktion blandt brugerne, ikke udelukkende er udgjort af selve holdningsdelingen, men ligeledes af såvel kommentarer, votes som subscribers, antages det, at udviklingen af delte holdninger, i mange tilfælde vil tangere den overliggende interaktionsudvikling, blandt reddit's brugere på r/unpopularopinion. *Scatterplottet* giver mulighed for at observere hver dag som en enkeltstående enhed, der kan belyse såvel udsving som generel udvikling. Ved at inddrage en gennemsnitsudvikling ved hjælp af *rolling average*, bliver den visualiserede fremstilling udpenslet, der tydeliggør eventuelle fluktuationer, fald eller markante udsving i fordelingen af observationer. Ved at kombinere disse to modeller, bliver visualiseringen af aktivitetsudviklingen på forummet simplificeret og læsbart, uden at gå på kompromis med eventuelle udsving i den empiriske virkelighed (Vincent 2017).

5.3.2 Univariat analyse

Den univariate analyse, fungerer som en relativt slask gennemgang af den empiriske virkelighed, frembragt i specialet. Som nævnt tidligere, er der igennem dataindsamlingen, blevet udvalgt syv variable, som hver især forventes at kunne give et indblik i interaktionen mellem brugerne på r/unpopularopinion. I tidsserieanalysen bliver tidsvariablen *created* anvendt til at observere udvikling i indholdsdeling fra dag til dag, igennem de forskellige epoker. I den univariate analyse, anvendes de resterende relaterede variable, til at anskue om udviklingen har indflydelse på den kollektive interaktion blandt r/unpopularopinions brugermængde. Til at belyse de forskellige variable, anvendes to forskellige modeller: Et *distplot* og en *probability mass function*.

Et distplot fungerer som et histogram med et inkorporeret *kernel density estimat*. Histogrammet fungerer ved at visualisere den relative frekvens distributionen af observationer i en kvantitativ variabel. Modellen tager udgangspunkt i intervaller, der observeres på x-aksen, og på y-aksen observeres distributionen er hver intervalfrekvens (Agresti and & Finlay 2009:35). Samtidig giver modellen mulighed for at observere forskellige elementer, f.eks. gennemsnit, percentiler, etc. (Ibid.). *Kernel density* estimatet fungerer som supplement til det visualiserede i histogrammet, hvor fordelingen af observationerne nærmere kan observeres. Ved at udglatte histogrammets fund, er det nemmere at antyde om variablen er, f.eks. normalfordelt, eller om der er asymmetriske tendenser (Agresti and & Finlay 2009:37–39; Waskom 2012).

En anden metode til at repræsentere kvantitative variables distribution, er ved at anvende en såkaldt *PMF*, en *probability mass function*. Modellen tager samme udgangspunkt som histogrammet, og vil derved observere præcis det samme objekt, men tager et anderledes perspektiv. PMF'en visualiserer, observation for observation, hvordan den værdi positionerer sig i forhold til mængden af alle de andre observationer. Modellen visualiserer derved et todimensionelt rum at værdifordelingen i en given variabel. X-aksen viser hver enkelt observation, og y-aksen viser hvor den enkelte observation ligger i rummet, eller, hvordan den placerer sig i det mulige rum af observationer. Derved bliver koncentrationer i populationen i PMF'en en del af massen, og eventuelle outliers bliver nemmere at observere, som afvigere i sandsynligheden (Datacamp.com 2018a; Downey 2016). I praksis anvendes forholdet mellem de to modeller til at visualisere såvel koncentrationer i de værdimæssige fordelinger, samt at uddybe forståelsen for eventuelle outliers, og brede variationer i variabelenes udtryk.

Begrundelsen for at anvende disse specifikke univariate visualiseringsmetoder til at formidle data i dette speciale, sker med en ambition og at formidle data så simpelt, men samtidig så uddybende som muligt. Ved at anvende et distplot sammen med en PMF, kan der dannes et gennemgående overblik, der sikrer at man både kan anskue markante outliers og dominerende frekvenser. Selvom specialet

ikke søger at påvise nogen relation mellem de givne variable i det empiriske materiale, antages det, at der er flere af de udvalgte informationer der, i større eller mindre grad, er internt korreleret. Her fungerer PMF'en fordelagtigt, på grund af dens opstilling af observationerne hen ad x-aksen. Hermed bliver analysens visualiseringer ikke blot sammenlignelige mellem de 2 modeller, mens også mellem de udvalgte variable, i hvert fald fra et skønsmæssigt perspektiv (Downey 2016).

5.3.3 Emneanalyse/Topic Modelling

I dette afsnit, vil specialets anvendelse af *topic modelling* beskrives. *Topic modelling*, eller emneanalyse, er en relativt bredt anvendt metode til at danne overblik over store og spuriøse tekstbaserede data (Blei et al. 2003; Chang and Boyd-Graber 2009). Metoden fungerer som en kombination af kvantitativ dataanalyse og kvalitative vurderingsprocesser. Der anvendes en form for *unsupervised* maskinlæring, hvor algoritmer søger efter mønstre og sammenhænge i det givne datamateriale (Blei 2012:82). For dette speciale, er *Latent Dirichlet Allocation (LDA)* og *Latent Semantic Indexing (LSI)* anvendt som det algoritmiske udgangspunkt for at tematisere indholdsdelingen på r/unpopularopinion igennem observationsperioden. *LDA* er, hvad der kan beskrives som *state-of-the-art* inden for visualisering af indholdsfordeling og tematisering af tekstdata, hvor *LSI* er en ældre metode, der anvendes til at understøtte og belyse det visualiserede indhold i *LDA*'en analysetilgang (Blei 2012; Blei et al. 2003:993; Chang and Boyd-Graber 2009; Klein, Clutton, and Polito 2018). *LDA* og *LSI* minder på mange måder om hinanden, set fra et mekanisk perspektiv, men *LDA*'en er ofte anvendt da den simplificerer en kvalitativ vurdering af indholdsfordelingen. I kontrast er *LSI* anvendt, da den giver mulighed for at bearbejde forholdet mellem de dominerende termer, og gå mere i dybden med de observerede fund fra *LDA*'en. I den overliggende forståelse af *topic modelling*, herunder såvel i *LDA* som *LSI*, akkumuleres datagrundlaget i et samlet korpus, i dette speciale ét for hver epoke. I hvert korpus tælles og opdeles indholdet i ord, med udgangspunkt i en simplificering af det tekstuelle. Ved at forarbejde de delte holdninger, så ord bliver fremstillet i deres stammeform, og med de mest dominerende fyldord censureret væk, opstår der et latent rum af definerende ord, hvor der kan opstå forbindelser, og derved er det muligt at der fremkommer emnemæssige afgrænsninger for det givne datagrundlag (Blei et al. 2003:993–94; Chang and Boyd-Graber 2009)(Bilag 9). I fagterminologi betegnes de modererede ord som *tokens*, der i den analytiske bearbejdning ligeledes er problemformuleringens *termer*. Optællingen og opgørelsen af determinerende ord, beskrives som korpussets ordbog, eller *dictionary*. I *topic modelling* er korpus altid udgjort af en række enkeltstående observationer, kaldet dokumenter (Blei et al. 2003:994-95). I dette speciale refererer dokumenter til de enkelte posts der bliver delt, og én holdning udgør derved ét dokument.

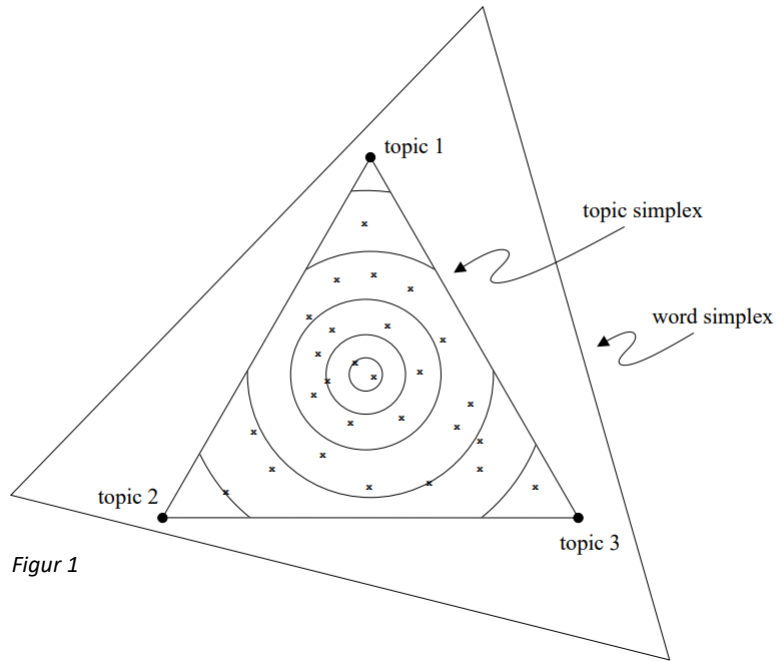
5.3.4 LDA og topic modelling i praksis

Latent Dirichlet Allocation eller *LDA*, er en metode inden for *topic modelling* der anvender en bayesisk, generativ og probabilitisk tilgang til at finde mønstre og gennemgående temaer i et givent korpus (Blei et al. 2003; Chang and Boyd-Graber 2009; Rehurek 2019). *LDA'en* adskiller sig fra tidligere anvendte modeller, f.eks. *LSI'en*, ved at forholde sig til hver enkelt ord i et givent dokument. Hvert dokument i et korpus, kan altså sandsynlighedsmæssigt tilgå flere forskellige af de konstruerede temaer, alt efter hvilke ord der indgår. I forlængelse af dette, bliver de givne temaer dannet med udgangspunkt i de repræsenterede ord, og hvordan de fremstår i kongruens. *LDA'en* bearbejder informationerne i de forskellige korpora med udgangspunkt i de Finettis teorem (1990) der fastslår at: "*Enhver samling af omskiftelige og tilfældige variable kan repræsenteres i et uendeligt antal af forskellige distributioner*" (Blei et al. 2003:994, oversat). I forlængelse skal denne formulering forstås som at ethvert dokument i en samling af dokumenter, er konditionelt uafhængig af den samlede enhed, men samtidig relationelle og omskiftelige så de bliver tilpasset et bagvedliggende latente rum (ibid.). Hertil skal det forstås at repræsentationen antager, at konstrueringen af sætningen ikke er lige så vigtig som hvilke ord, der indgår i dokumentet. Ordenes orden og position i en sætning, er altså ikke i udgangspunktet udslagsgivende for typificeringen i *LDA'en*. Det latente rum i *LDA'en* er defineret af opstillingen af overliggende emner, da *LDA* som metode, antager at antallet af emner er kendt på forhånd (Blei et al. 2003:82–83; Chang and Boyd-Graber 2009:3; Rehurek 2019)

Det latente rum i *topic modelling* og inddelingen af emner i *LDA'en* skal forstås ud fra de subjektive og kvalitative overvejelsesprocesser, der indgår i indramningen af emnefordeling for et givent korpus. Selvom *LDA'en* som metode er bredt anerkendt, begrænser den sig ved at være primært opført til at belyse datagrundlag, hvor indholdet til en vis grad er kendt i forvejen. I dette speciale arbejdes der eksplorativt, og formålet med tilgangen, er at højne forståelsen for det sociale liv i internetalderen, og hvordan den relationelle og tilgængelige information kan behandles til dette formål. Grundlaget for at anvende *LSI* som supplerende, ligger derved i *LDA'ens* begrænsninger. Hvor *LDA* generelt opfattes som en af de mest prominente til at danne et overblik til kvalitativ vurdering, inddrages den ældre og mere determinerende *LSI* til at skabe grundlaget for bearbejdning af de givne korpora. Kort opsummeret, er *LDA* og *LSI* tangerende metoder, der primært adskiller sig ved at *LDA* forholder sig til hver enkelt token, hvor *LSI* forholder sig til hvert dokument som helhed. *LSI* giver samtidig mulighed for at inkorporere en principal komponentanalyse (*PCA*), en metode til at uddybe den emnemæssige fordeling i det latente rum af tokens. I dette speciale bliver *PCA'en* analyseret gennem *K-means clustering*, en simpel *unsupervised* maskinlærings algoritme der fordeler det givne data med udgangspunkt i et specificeret antal klynger (*k*) (Ding and He 2004). Algoritmen er naiv, fordi *k* skal præciseres, men ved at inddrage en såkaldt *elbow method*, kan man kombinere bearbejdningen fra

PCA med *K-means clustering*, og give indblik i, hvor ideelle forskellige klyngekonstellationer vil være, for det givne datagrundlag (ibid.). *Elbow method* fungerer i praksis ved at opstille en lineær funktion, hvor k bliver fremstillet i et givent interval. Den visualiserede linje, fremstår som en arm, og hvor der fremkommer en "albue", vil umiddelbart være der, den bedste klyngekonstellation skal findes, med udgangspunkt i det givne interval (Blei et al. 2003:1005–6; Ding and He 2004; McIntyre, Bengfort, and Lacanlale 2016). I praksis, kan der derved opstilles et parameter for at kontrollere en optimal klyngefordeling, på baggrund af en ønsket indramning, f.eks. for at finde det ideelle antal klynger, mellem 1 og 50 for et givent korpus (Abdi and Williams 2010; Blei et al. 2003:1005–7; Ding and He 2004; McIntyre et al. 2016). Herfra overføres *PCA*'ens vurdering til *LDA*'ens konstruktion, hvorved der fremstilles en visualisering af det latente rum, med udgangspunkt i den fremførte klyngefordeling. Visualiseringen af de givne korpora igennem *LDA* fungerer som en interaktiv og modellerbar brugerflade, hvor de bærende termers interne interaktion og kongruens kan belyses, med udgangspunkt i den udvalgte klyngefordeling (Sievert and Shirley 2014).

De dominerende og primære termer, der ønskes belyst i dette speciale, fremkommer derved i bearbejdning af de forskellige epokers korpusindhold. I en uddybende forståelse af *LDA*'en i praksis, forholder metodens algoritme sig til flere elementer. Som nævnt behandler *LDA*'en korpusets ordindhold i et latent rum af interaktionerne mellem ordene. Antallet af emner afgør dimensionsforholdet i det latente rum, og definerer den visuelle fremstilling af ordenes interaktion. *LDA*'en fordeler indholdet i et givent korpus, med udgangspunkt i distributionen over de givne tokens for korpusets ordbog. Samtlige ord og deres distribution i det samlede korpus, fordeles i det latente rum som punkter, også kaldet et ordsimpleks (*word simplex*) (Blei et al. 2003:1000-2). Med en opsat klyngefordeling (k) fordeler *LDA*'en indholdet i en reduceret dimensionsfordeling over fordelingen og relationen mellem de forskellige ordsimpleks, kaldet emnesimpleks (*topic simplex*). I praksis udvælger modellen et tilfældigt punkt i det latente rum, og alle ord i et givent korpus, fordeles med udgangspunkt i afstanden og interaktionen mellem tokens, ud fra en bayesisk sandsynlighedsmodel (ibid.).



Figur 1

Ovenstående figur 1 er en visualisering af denne distribution i det latente rum, hvor der er anvendt tre emner (Blei et al. 2003:1002). Hvert x i figuren skal i *LDA'en* ses som et ord, og dets afstand til hvert emnepunkt afgør hvilken klynge det tilgår. Positionen af x^n er afgjort af den interne korrelation og interaktion mellem ordene i det samlede korpus. Hvis ord ofte fremstår i samme dokument, er der større sandsynlighed at de er nær hinanden i det latente rum. Den visualiserede figur skal samtidig ses som et udgangspunkt for den dimensionsreduktion der finder sted i *LDA* (Blei et al. 2003:1000–1002; Chang and Boyd-Graber 2009:3–4; Ma 2018; Sievert and Shirley 2014). Ordenes sandsynlighed for at tilgå et givent emne, er afgjort af de parametermæssige vilkår der bliver opsat, blandt andet i antallet af givne emner, for det latente rum.

“- Rocco: Fuckin'! What the fuckin'. Fuck. Who the fuck fucked this fucking...How did you two fucking fucks...Fuck!

- Connor: Well, that certainly illustrates the diversity of the word.”

(Rocco & Connor MacManus, *Boondock Saints* (1999), jf. (Moviequotes.com 2019))

De fremkomne tokens i behandlingen af det tekstuelle univers, kan ofte have tvetydige anvendelser hvilket maskinalgoritmen ikke nødvendigvis lærer automatisk. Ord er ofte afhængige af deres kontekst og relative anvendelse, som eksemplificeret i ovenstående citat fra kultfilmen *Boondock Saints* (1999). Alt efter hvor præcis klyngekonstellationen er, i relation til det anvendte datamateriale, har algoritmen mulighed for at lære sammenhænge mellem de forskellige tokens, hvorved denne udfordring kan overkommes i en grad (Blei 2012:78–79; Chang and Boyd-Graber 2009:3–4; Sievert and Shirley 2014:69–70). I dette speciale forventes det, at der er stor variation i ordvalg, uden at der nødvendigvis er tematisk sammenhæng, hvilket kan problematisere den semantiske forståelse af

analysens emneopdeling. Af samme grund vurderes det vigtigt at bearbejde emneanalysens fund med et kritisk, kvalitativt perspektiv, hvilket der vil blive sat fokus på i specialets diskuterende konklusion. Læsbarheden og visualiseringen af *LDA's* fund, pålægges variation i kraft af de subjektive valg, der ligger i definitionen af det latente rum i emneanalysen. Udover valget af dimensionsreduktion, gennem klyngemængde og behandling af det tekstuelle materiale, er der flere effekter, der kan forandres: Sådanne hyperparametre for algoritmens behandling af data. *LDA'en* trækker på bayesisk statistik, i den måde den forholder sig til variationen af ord. I praksis skal *LDA* "trænes", i den forstand, at den søger mønstre imellem ord, med udgangspunkt i den måde man behandler og fodrer den med data (Rehurek 2019; Sievert and Shirley 2014). Det bayesiske element ses, i forlængelse af Finettis teorem, ved at *LDA* antager at mønstre der findes i den primære tekstindsamling, også vil gøre sig gældende for videre bearbejdning. Dette skal desuden ses ud fra, at *LDA'en* kan anvendes som en prætrænet model, hvor observerede mønstre fra en dataindsamling kan overføres til en anden (Blei et al. 2003:994, 998; Rehurek 2019). Da den tidsmæssige udvikling, og forståelsen for eventuelle fluktuationer over den indholdsmæssige fordeling er et vigtigt element, er prætræning ikke anvendt i dette speciale. Derimod er hver epoke behandlet for sig, hvor der desuden er opsat et *passes* parameter, der kontrollerer at modellen går igennem det samlede datagrundlag flere gange, fra forskellige indgangspunkter, for at sikre mod eventuelle tidsmæssige forandringer for de enkelte epoker (Rehurek 2019). De anvendte modeller i specialets emneanalyser fremstår derved som et samlet overblik de givne epoker, og vil fungere til at uddybe den kollektive indholdsdeling på tværs af året. Derfor vil eventuelle forandringer i indholdet primært kunne observeres på tværs af epoker, og ikke så meget inden for den givne epoke. Selvom det må antages, at udviklingen i brugerinteraktionerne og diskursive fluktuationer sker flydende, er det såvel praktisk som behandlingsmæssig vurdering, at de epokemæssige samlinger er fyldestgørende i belysningen af den tekstuelle forandringsproces på r/unpopularopinion.

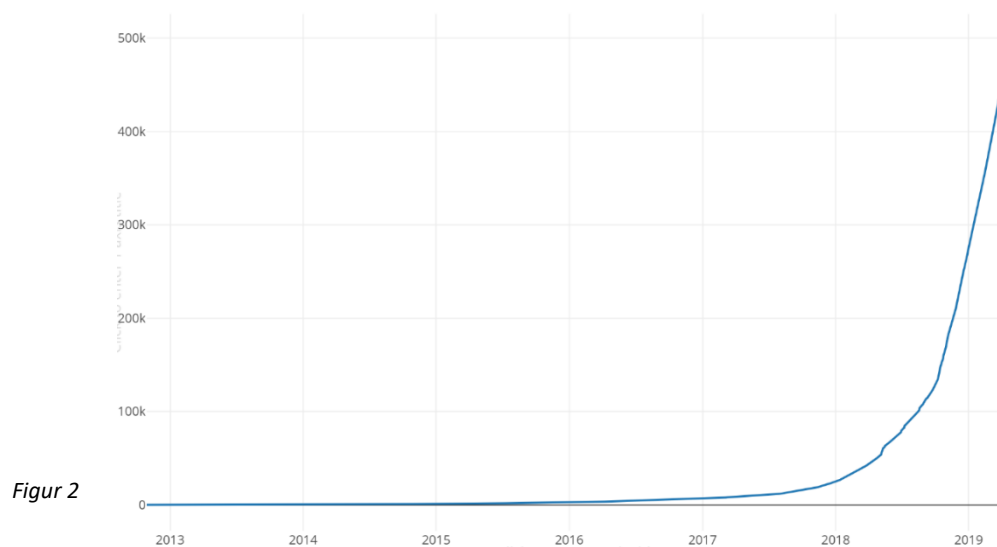
6 ANALYSE

Igennem afsnittet om analysestrategi, er de forskellige overvejelser omkring specialets analytiske proces gennemgået. Specialet har til formål at belyse reddit's anonyme internetsocialitet, og søge at bidrage til forståelsen af det sociale liv der udspiller sig i internetalderen. For hver del af analysen er der inddraget metodiske overvejelser og analytiske strategier, der antages at være behjælpelige til at forstå den brede og dybe variation af interaktioner der udfolder sig på reddit's underforum r/unpopularopinion. Gennemgående skal analysen forstås som en deskriptiv belysning af det sociale liv der udleveres på r/unpopularopinion i den samlede observationsperiode. Indledningsvist behandler

analysen, hvad der kan beskrives som metainformationer omkring det tekstuelle univers på den udvalgte subreddit, gennem tidsserier og univariat analyse. Brugerengagementet vil herigennem blive belyst, og sætte fokus på såvel udvikling som forandring i det sociale liv på r/unpopularopinion. Herefter vil der blive søgt efter dominerende termer og emnemæssige fordelinger over det tekstuelle indhold, for at belyse effekten af den observerede brugertilstrømning til forummet. Specialets analyse tager udgangspunkt i en epokemæssig undersøgelse af indholdsdelingen og de relationelle interaktioner for brugerne på r/unpopularopinion, i perioden fra januar 2016 frem til april 2019. Specialet er opdelt i tre hoveddele, bestående af en række tidsserier, en univariat analysedel samt en emneanalyse af det tekstuelle indhold i de anvendte epoker. Afslutningsvist vil analysens fund blive disputeret og relateret igennem specialets diskuterende konklusion.

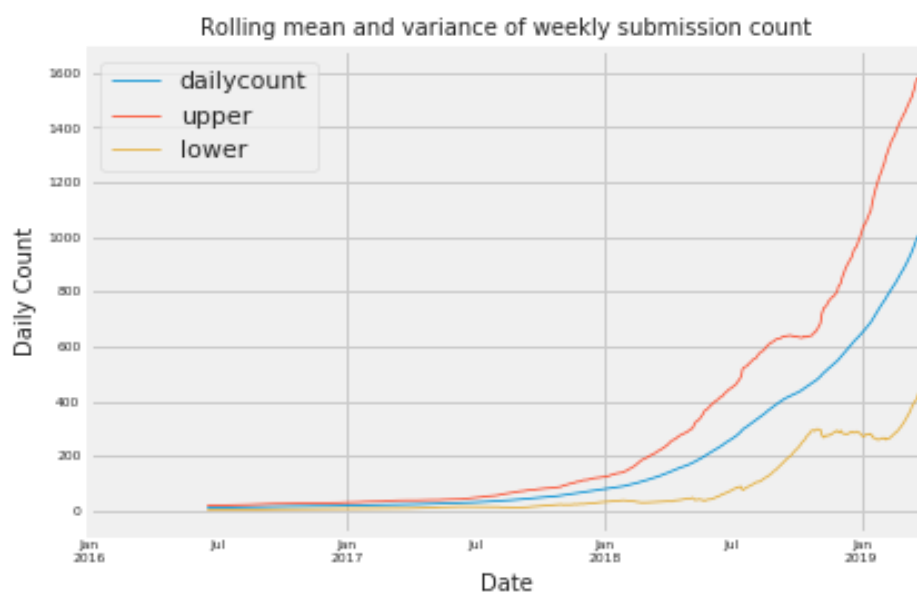
6.1 TIDSSERIER

I den følgende analysedel, vil den aktivitetsniveauet på r/unpopularopinion blive vurderet med udgangspunkt i antallet af daglige posts, over tid gennem observationsvinduet.



Figur 2

Figur 2 viser antallet af subscribers på forummet r/unpopularopinion siden dets opstart i slutningen af 2012. Herfra kan det vurderes, at forummet har oplevet meget lidt aktivitet frem til 2016, hvor der opleves en begyndende stigning i subscribers. Denne udvikling følger løbende gennem 2016 og 2017, hvor der i 2018 opleves en eksplosion i brugertilstrømningen der fortsætter ufortrødent ind i 2019. Antallet af subscribers går fra omtrent 3.000 i starten af analysens observationsperiode (januar 2016), og omkring 500.000 ved dens afslutning (april 2019) (Subredditstats.com 2019).

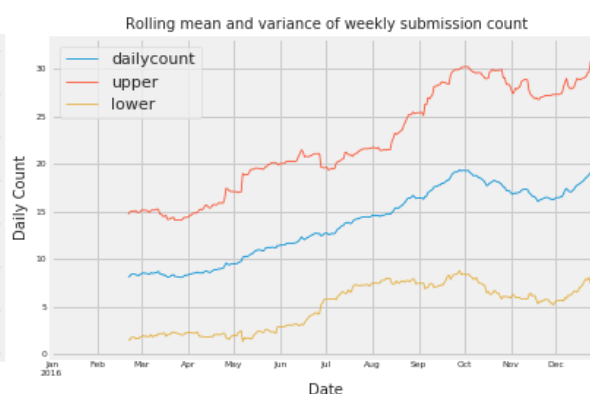
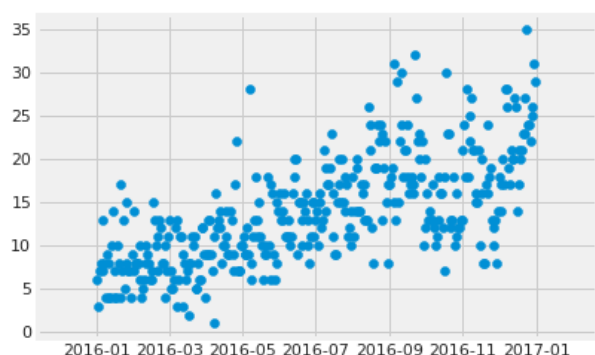


Figur 3

Figur 3 giver i forlængelse af figur 2, indblik i hvordan aktivitetsniveauet udvikler sig på r/unpopularopinion gennem observationsperioden. Her er der anvendt den gennemsnitlige postmængde til at visualisere engagementet blandt forummets subscribers, baseret på de ugemæssige opdelinger. Det kan vurderes, at aktivitetsniveauet tilnærmelsesvist følger brugertilstrømningen i observationsperioden. I 2016 opleves der en mindre stigning, der øges i løbet af 2017, men først for alvor tager fart i løbet af 2018 og i 2019. Det kan vurderes, at der er omkring 100 daglige posts i gennemsnit i januar 2018, mod omkring 1.000 daglige posts i slutningen af marts 2019.

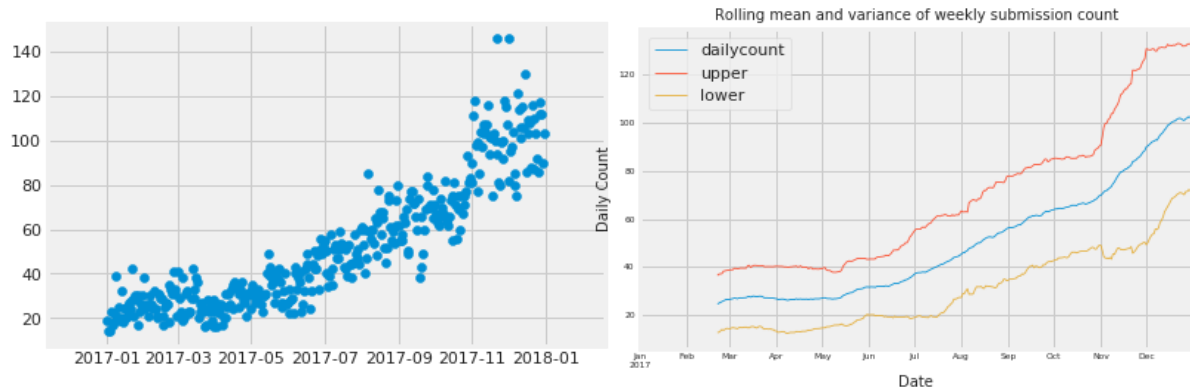
Det vurderes, at denne aktivitetsmæssige udvikling er generelt stigende på tværs af analysens observationshorisont, men for at udspecificere engagementet vil de epokemæssige udviklingstrin blive belyst, med et scatterplot og en RA model for hvert år.

6.1.1 2016



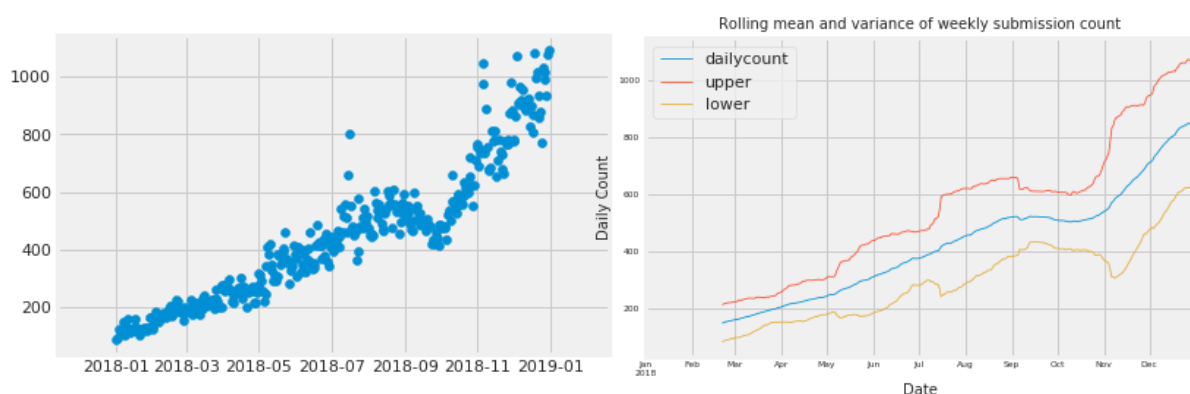
I starten af 2016 opleves der i udgangspunktet mellem 5 og 15 daglige posts, der langsomt men sikkert stiger mod cirka 25 gennemsnitlige daglige posts mod slutningen. Selvom scatterplottet viser en relativt stor spredning mellem observationerne, må dette ses i relation til den lave variation mellem plottets ekstremiteter, hvor der forment i denne epoke maksimalt opnås 35 posts på en enkelt dag (Bilag 4).

6.1.2 2017



Allerede i 2017, kan der fra scatterplottet observeres en mindre spredning, men større variation mellem observationspunkterne. Udviklingen følger tendensen fra 2016, og starter omkring 25 daglige posts i gennemsnit, og udvikler sig løbende gennem året, mod cirka 100 gennemsnitlige posts i slutningen af 2017, hvor der dog er dage der når op mod 150 individuelle submissions, hen mod slutningen af året (Bilag 5).

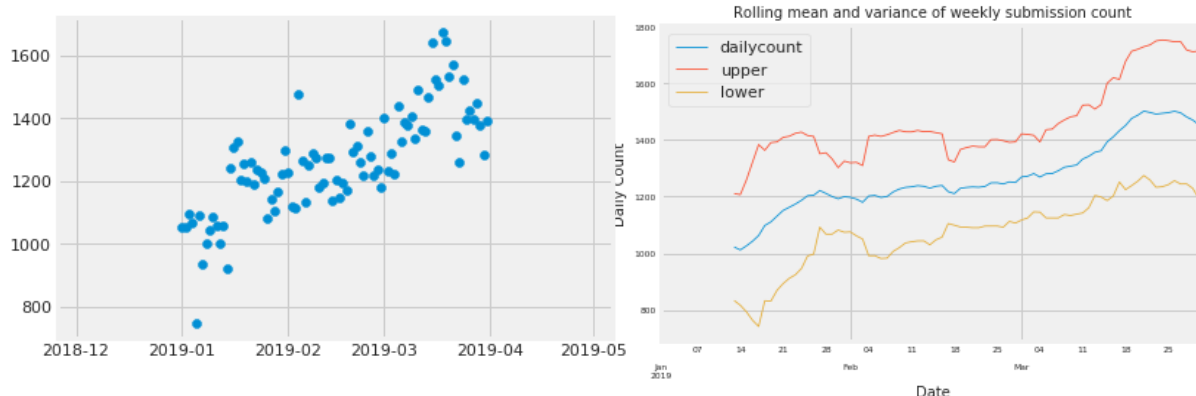
6.1.3 2018



2018 følger igen udviklingen fra tidligere år, hvor aktivitetsudviklingen formindsker spredningen, og tydeliggør korrelationen for postmængden over tid. Der kan observeres et mindre fald i aktivitetsniveauet omkring oktober 2018, hvor der modsat den tendentielle udvikling, maksimalt er omkring 500 daglige posts, i en kortere periode. Trods dette midlertidige fald i aktivitetsniveau, der ikke umiddelbart kan sidestilles med brugertilstrømningen i samme periode, opleves der

efterfølgende en, om muligt, endnu større stigning end tidligere, som fremkommer i løbet af oktober. I slutningen af 2018 er der omkring 1.000 daglige posts i gennemsnit, med de yderste op mod 1.100, i skarp kontrast til de omkring 100 i dagligt gennemsnit, i starten af året (Bilag 6).

6.1.4 2019



Observationernes fordeling og den generelle udvikling antydnet i scatterplottet og *RA* modellen, skal ses med udgangspunkt i, at observationsperioden for 2019 kun indeholder de første 3 måneder af året, og derfor kun fremstår med omkring 90 observationer frem for 365, som i de resterende tidsserieanalyser. På nær en enkelt udstikker med under 800 posts, bliver den tendentielle udvikling fulgt. I 2019 er der mellem 1.000 og 1.100 daglige posts i starten af året, og i slutningen af marts er der ca. 1.500 daglige posts, dog med relativ stor spredning. Ud fra *RA* modellen, kan der observeres et mindre fald mod slutningen af marts 2019. Det kan ikke fra denne undersøgelse vurderes om der er tale om et fald, en stagnering, eller blot en forvikling på grund af afslutningstidspunktet for empirien, men det er ikke forventeligt at det observerede fald er ensbetydende med et direkte opbrud i aktivitetsudviklingen (Bilag 7).

6.2 UNIVARIAT ANALYSE

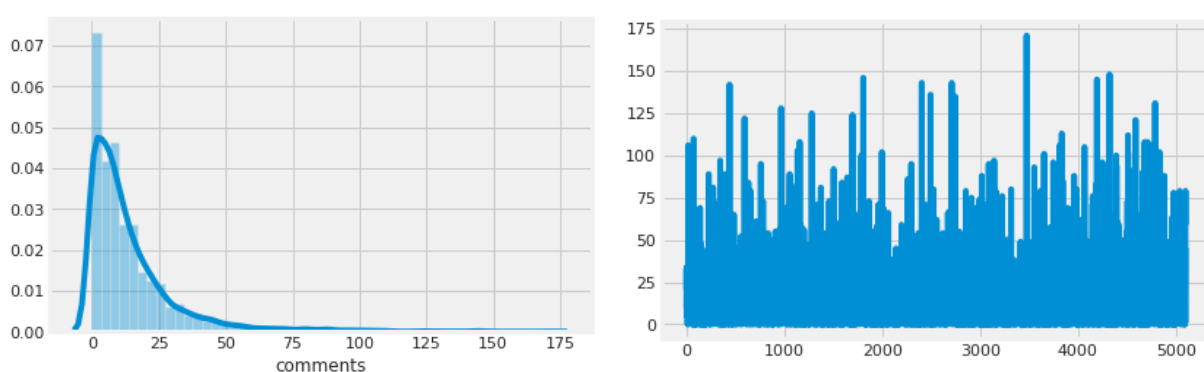
Brugernes interaktion på *r/unpopularopinion*, er defineret af mere end blot indholdsdeling, og bliver i dette speciale undersøgt ved hjælp af relationelle variable til at indramme brugernes individuelle udmeldinger, og forummets reaktion på dem. For hver holdningsudmeldelse er der seks variable tilknyttet, henholdsvis; forfatternavn, kommentarmængde, antal upvotes, en upvote ratio og en samlet score. I det følgende afsnit, vil variablene blive gennemgået epokemæssigt, for at udpensle den observerede aktivitetsudvikling.

6.2.1 2016

6.2.1.1 Forfatterskab

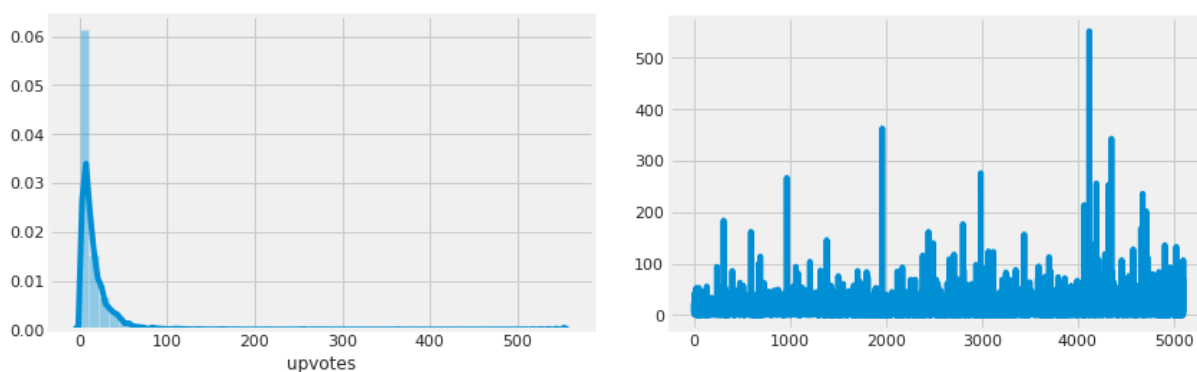
Det første der kan påpeges, med udgangspunkt i forfattervariablen, er at en relativt stor del af de delte holdninger, ikke har forfatternavn tilknyttet, hvilket er en gennemgående tendens gennem den samlede observationsperiode. I 2016 blev der delt 5.099 holdninger, hvoraf 2.855 har forfatternavn tilknyttet. Af disse, er 1.540 unikke brugere. Der opleves ikke store variationer i frekvensen; blandt de kendte forfattere, har langt de fleste én eller få posts, og den mest aktive har 28 posts fordelt over året (Bilag 4).

6.2.1.2 Kommentarer



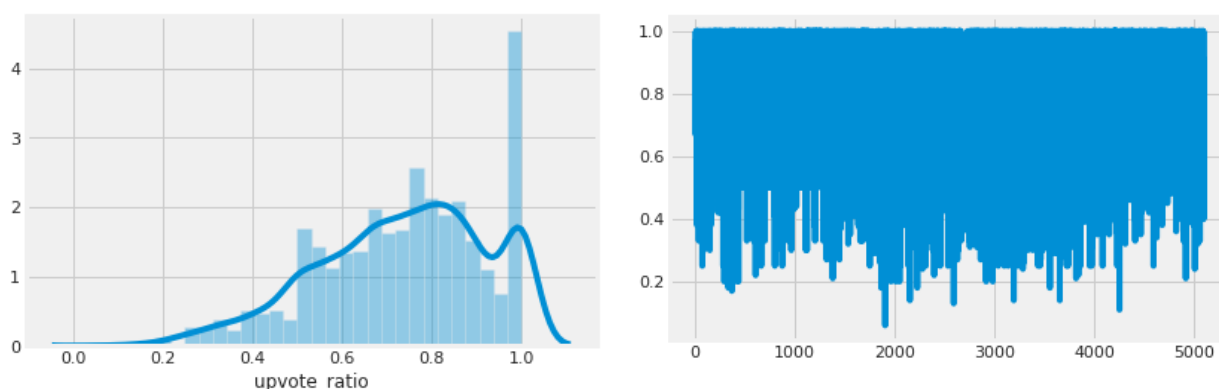
Ovenstående distplot, hvor mængden af kommentarer er i x-aksen, viser, at der er en klar koncentration omkring de laveste tal, men op til omkring 175 kommentarer ved den yderste observation. En overvejende andel af observationerne har ingen kommentarer tilknyttet, og den absolutte størstedel befinder sig mellem 0 og 50 kommentarer. PMF'en, med observations ID på x-aksen og kommentarmængden på y-aksen, viser den ledende observation foran et felt af holdninger med mellem 75 og 150 kommentarer. En ekstra feature ved PMF'en er, at på grund af datasættets struktur, fungerer x-aksen også som en tidslinje for upload dato. Derfor kan det også antages, at der ikke er en tydelig sammenhæng mellem tidspunkt og kommentarmængde for 2016 (Bilag 4). Den mest kommenterede holdning har opnået 171, efterfulgt af en række observationer med mellem 140 og 150, dog maksimalt 148 (ibid.).

6.2.1.3 Upvotes



Igen ses den primære koncentration af observationer i den nedre del, mellem 0 og 50 upvotes er normen. Ved upvotevariablen opleves der en større spredning, med over 500 upvotes på én holdning. Her er det også bemærkelsesværdigt, at den mest upvotede udmelding er markant foran andre populære holdninger, med over 100 ned til den næste observation. Med udgangspunkt i PMF'en kan det antages at der opleves en vag stigning i antallet af upvotes, som sandsynligvis kan forklares af den udvikling r/unpopularopinion oplever i tidsrummet. I 2016 opnår den mest upvotede holdning 552 upvotes, de nærmeste efterfølgere har 363 og 343 upvotes. For samme tidsrum gælder det, at 795 har fået 1, og 746 har opnået 0 upvotes. Da brugerne er i stand til at upvote deres egne posts, må det antages, at det vil være en gennemgående tendens gennem årene, at mange observationer vil have 0 eller 1 upvotes (ibid.).

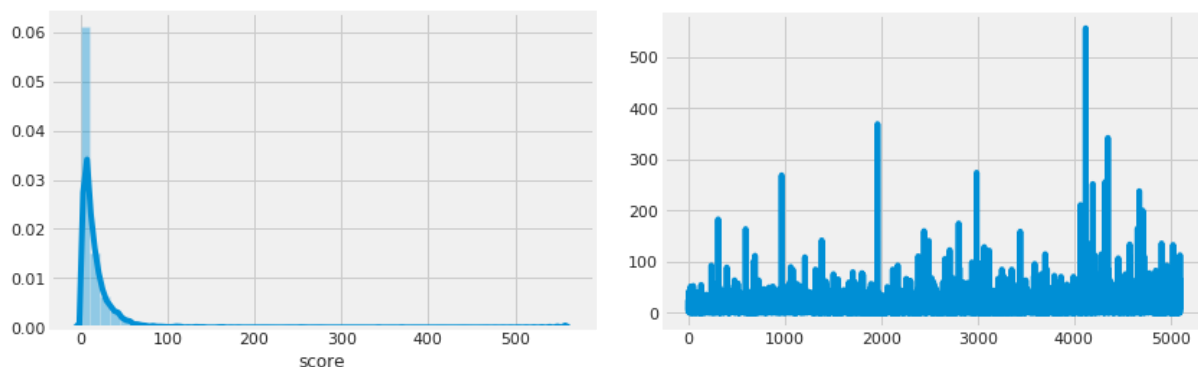
6.2.1.4 Upvote Ratio



Distplottet viser at der er en gennemgående positiv upvote tendens, med flere upvotes en downvotes gennemgående. Igen er der en koncentration med en ratio på 1.0 der sandsynligvis kan forklares med samme grundlag som upvote variablen. Der er 657, der udelukkende har fået upvotes, og 233 der har fået lige mange upvotes og downvotes. Selvom de største koncentrationer skal ses i 1.0 og 0.5, er det bredeste felt af observationer centreret mellem 0.67 til 0.87 (Bilag 4). Fra PMF'en kan de lavest

placerede outliers observeres. Der er en koncentration af upopulære holdninger med en ratio omkring 0.2, mens den mest ekstreme er helt nede på 0.06, fulgt af en række observationer med en ratio mellem 0.1 og 0.15. Tidsdimensionen i PMF'en viser ikke de store udsving, selvom der måske kan observeres et mindre fald omkring observation 2.000 og 4.000, hvor flere holdninger opnår meget lav ratio (Bilag 4).

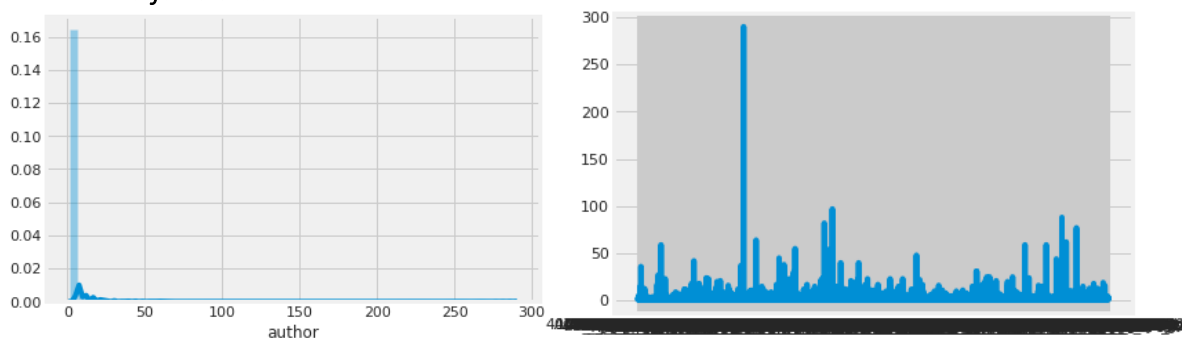
6.2.1.5 Score



Variablen for score viser nogenlunde de samme tendenser som blev oplevet under upvotes, hvormed sammenhængen mellem disse understøttes. Igen med en tydelig koncentration fra 0 til 50 og marginaler der tangerer upvotes. 75 procent af observationerne har fået 15 eller under i samlet score, hvilket også gør sig gældende for upvote variabelen. Den yderste observation har fået 556 point, og de næstkommende har henholdsvis 369 og 342 i samlet score. Selvom det kan vurderes, at upvotes er en god indikator for, om en post har opnået høj score, er det ikke ensbetydende med at der er direkte sammenhæng. For 2016 kan det siges, at posts der har fået mange upvotes, generelt har været positivt modtaget på r/unpopularopinion (ibid.).

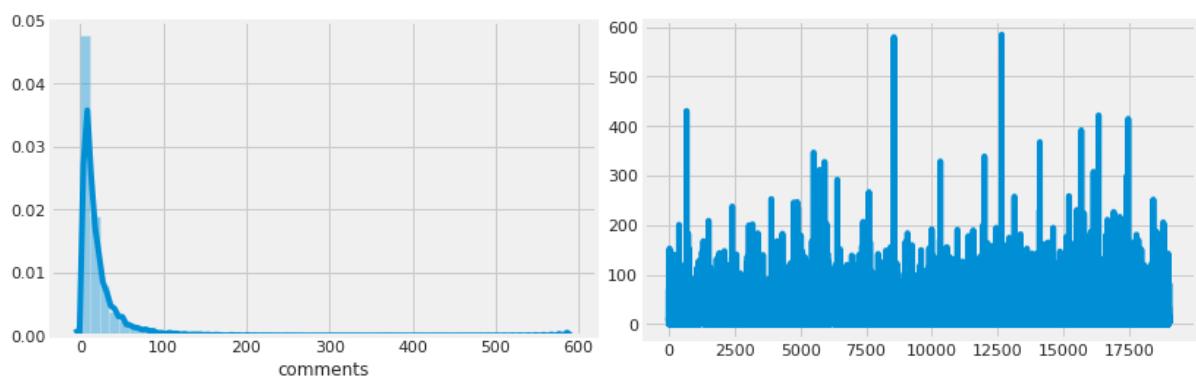
6.2.2 2017

6.2.2.1 Forfatterskab



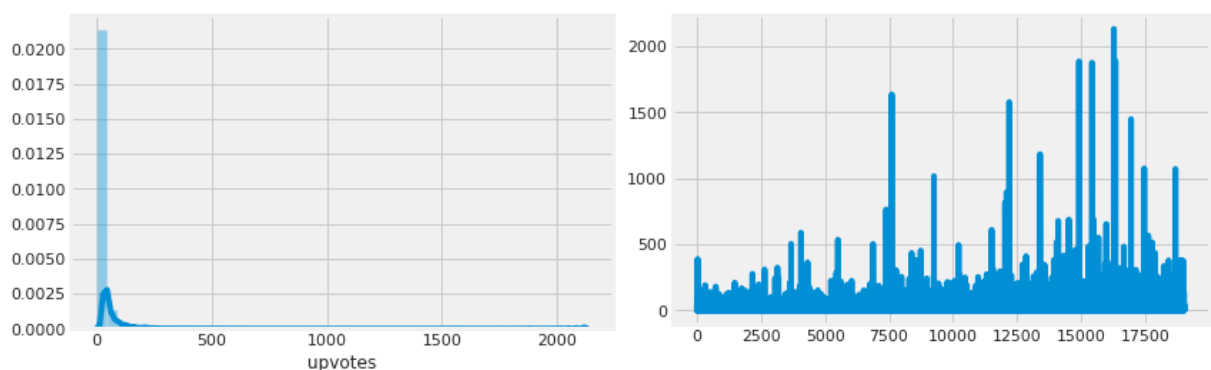
I 2017 bliver der delt 19.015 holdninger. 10.579 udmeldinger har forfatternavn tilknyttet, og 4.569 er unikke brugere. Der er stadig en klar koncentration omkring én eller få udmeldinger, men brugertilstrømningen til forummet har også bidraget til flere, mere aktive brugere. De mest aktive brugere på r/unpopularopinion i 2017, har generelt omkring 50 posts i løbet af året, men en enkelt tydelig outlier har 290 holdningsudmeldinger. Selvom der er en markant forskel i aktivitetsniveau for denne bruger i forhold til de resterende, er det ikke antaget, at den enkelte bruger har haft indvirkning på forummets generelle udtryk. De 290 posts udgør omkring 1,5 procent af den samlede indholdsdeling for året, og det vurderes at brugerens indholdsdeling ikke adskiller sig på andre parametre end aktivitet (Bilag 5).

6.2.2.2 Kommentarer



I 2017 har langt de fleste posts mellem 0 og 25 kommentarer tilknyttet; 75 procent af holdningerne har 21 kommentarer eller mindre, hvorefter det stagnerer hen mod 100 kommentarer. Blandt de næsten 20.000 delte holdninger, er der en mindre gruppering med mellem 100 og 400 kommentarer, og to markante outliers med henholdsvis 585 og 580 kommentarer tilknyttet (ibid.).

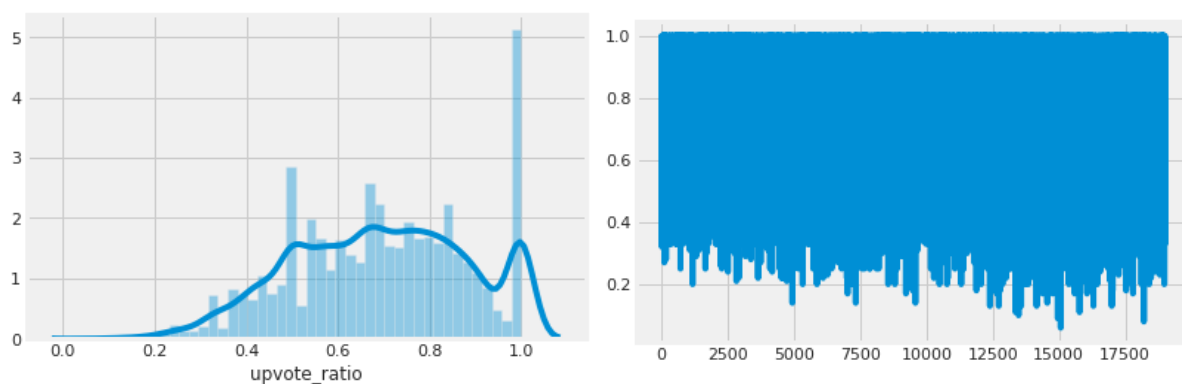
6.2.2.3 Upvotes



Ligesom med kommentarer, er der en klar koncentration i den lave ende, langt størstedelen har fået 0 eller 1 upvotes; henholdsvis 3.874 og 2.845 (Bilag 5). Der er en tydelig samling mellem 0 og 50, hvoraf den største gruppering har under 20 upvotes; 75 procent med 16 upvotes eller mindre. Der kan

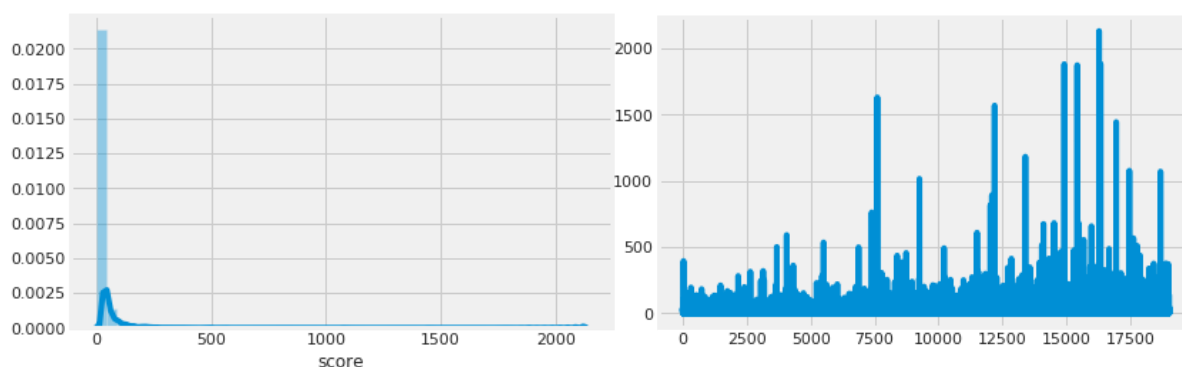
observeres en noget større spredning end i 2016, hvor en enkelt holdning har fået 2.130 upvotes, efterfulgt af et mindre felt af posts med mellem 1.000 og 2.000 upvotes (Bilag 5).

6.2.2.4 Upvote Ratio



Upvote ratioen viser sig i 2017 en smule mere normalfordelt omkring 0.5, men der er stadig et klart overtal af positive tilkendegivelser i interaktionen på r/unpopularopinion. De tydeligste koncentrationer ses i 1.0 og 0.5, med henholdsvis 1.999 og 1.130 observationer, hvilket kan anskues som understøttende for, at mange holdninger får lidt eller ingen opmærksomhed på platformen. Dette vurderes ud fra, at jo flere der har interageret med en holdning, jo mere varieret vil ratioen være i sit udtryk. Den bredere fordeling i distplottet, kan derfor være et udtryk for mere aktivitet, der muligvis følger den relationelle udvikling på forummet. I 2017 positionerer upvote ratioen sig i den højere ende, med 0.69 som gennemsnitlig værdi. De laveste outliers når fra 0.2 og ned, og de mest radikale har en ratio på 0.06 og 0.08 (Ibid.).

6.2.2.5 Score



Scorevariablen belyser, som flere af de andre variable, en koncentration i den lave ende. 75 procent af observationerne har en score på 16 eller under. Med PMF'en kan det desuden anskues, at der sideløbende med aktivitetsudviklingen i slutningen af 2017, sker en mindre udvikling i, hvor højt folk generelt scorer på deres indhold. Det vurderes ligeledes at være i den tidsramme de højeste outliers

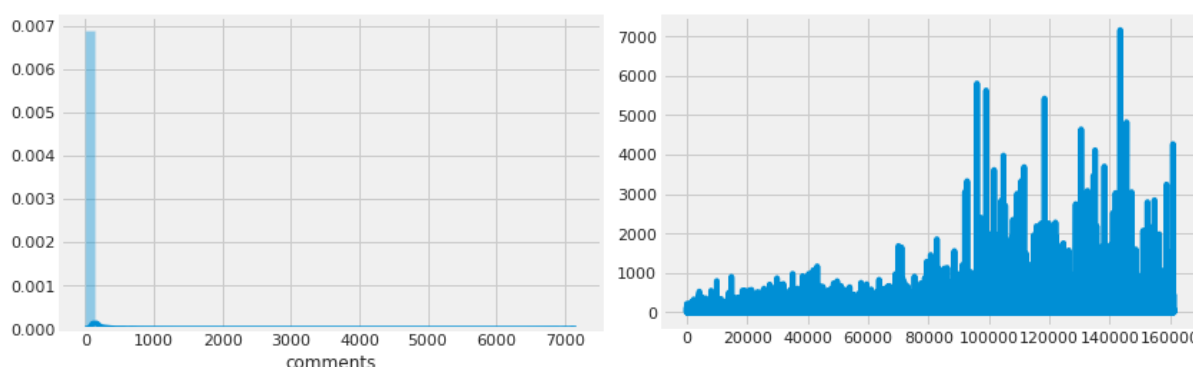
fremkommer, den højest scorende med 2.133, og de næstkommende tre med op mod 1.900 point (Bilag 5).

6.2.3 2018

6.2.3.1 *Forfatterskab*

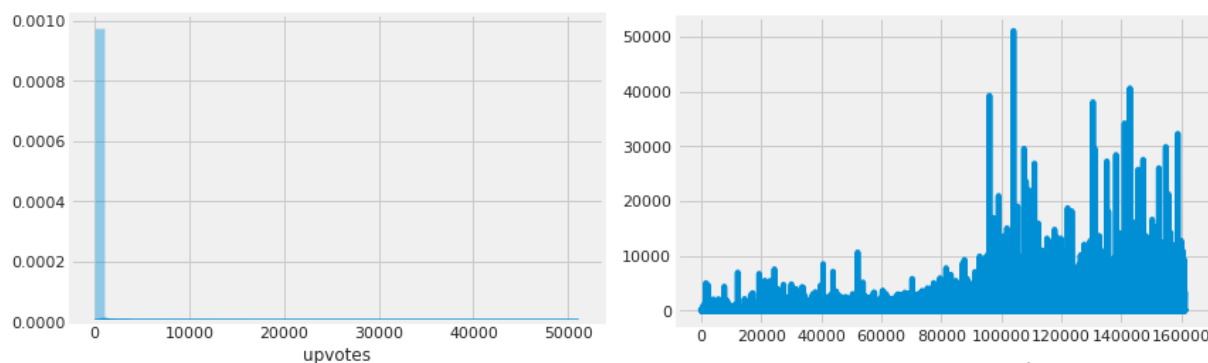
I 2018 bliver indholdet på r/unpopularopinion mangedoblet i forhold til året før, med 161.275 delte holdninger. Af den samlede indholdsdeling, har 93.867 tilknyttet forfatterinformationer, hvoraf 41.907 er unikke brugere. I 2018 er der en del flere aktive brugere, de mest prominente med nær ved én post per dag i gennemsnit: Den yderste observation med 373 posts, fordelt over året (Bilag 6). Selvom der er flere, meget aktive brugere på forummet, udgør deres engagement en forsvindende lille del af den samlede interaktion i 2018, og det vurderes igen, at de ikke afviger på andre parametre end deres virkelyst (ibid.)

6.2.3.2 *Kommentarer*



Med den store udvikling i holdningsdelinger, kan der ligeledes observeres et spring i de relaterede variable, og specielt spredningen mellem observationerne fremstår markant. Trods dette, er der stadig en klar koncentration af holdninger med få eller ingen kommentarer. Over 75 procent af de delte holdninger har opnået under 20 kommentarer, hvoraf 25 procent opnåede 0 eller 1. I ekstremerne, er der en mindre gruppering af observationer med over 5.000 kommentarer, og den mest omdiskuterede holdning har opnået 7.163. Udover den store spredning mellem holdningerne i 2018, sker der også en observerbar og markant udvikling inde i året. Fra omkring observation 90.000 og frem, er der et klart spring i antallet af kommentarer, i forhold til tidligere. Observation 90.000 positionerer sig i slutningen af september, og det kan derfor antages, at det observerede spring kan sidestilles med udviklingen i indholdsdeling, visualiseret i tidsserieanalysen (Bilag 6).

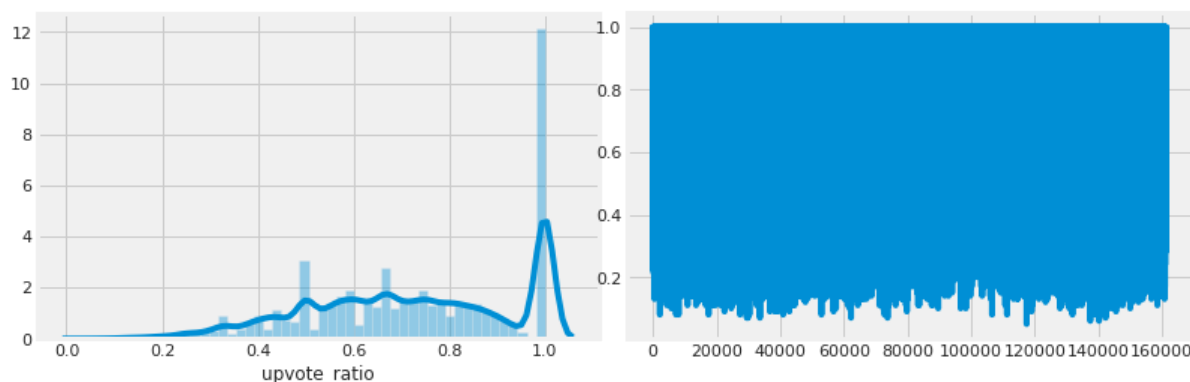
6.2.3.3 Upvotes



Den samme antagelse kan overføres til upvote

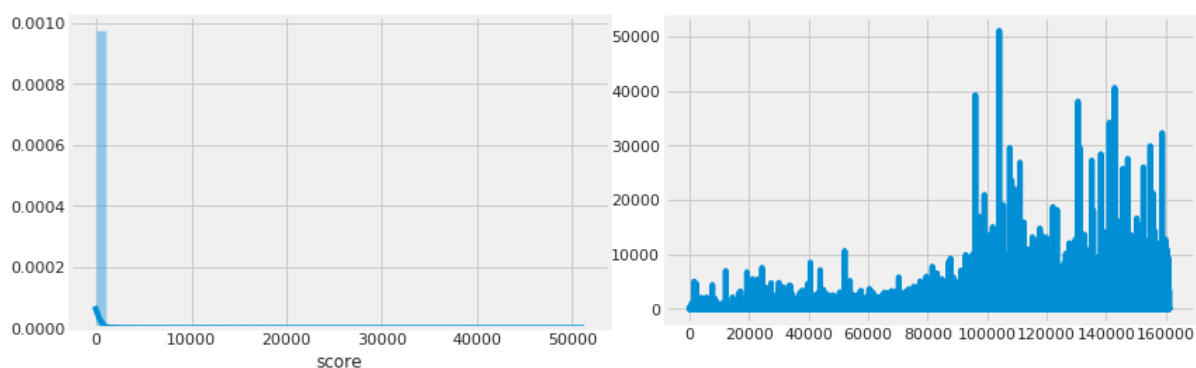
variablen, der, med samme samling omkring de lavere mængder, har en stor spredning, med én holdning der har opnået 51.125 upvotes. Selvom den yderste observation er relativt fremtrædende i forhold til de resterende, er der et felt af observationer mellem 20.- til 40.000 upvotes. Stort set samtlige af observationer der kommer over 10.000 upvotes, befinder sig omkring oktober og fremefter (ibid.).

6.2.3.4 Upvote Ratio



Selvom distplottet fremstår en smule mere sammenpresset, på grund af den store mængde af posts, er der ikke store gennemgående variationer fra tidligere. De største grupperinger ligger, som tidligere, på 1.0 og 0.5, med henholdsvis 35.733 og 9.243 observationer. Det kan derved antages, at der er en større relativ samling omkring holdninger med en ratio på 1.0 i forhold til 2017, hvilket også kan være medforklarende til en højere gennemsnitlig ratio på 0.72. Trods den store koncentration omkring ratio på 1.0, vurderes det stadig, at holdningerne generelt positionerer sig i den positive ende af skalaen, med størstedelen over 0.5 (ibid.). Det spring der observeres i de resterende relationelle variable, er ikke umiddelbart udslagsgivende for hvor kontroversielt holdningerne bliver modtaget, med en rimelig jævn fordeling gennem observationsperioden. Selvom tidshorisonten generelt ikke opleves som påvirkende for upvote ratioen, skal den laveste stadig findes hen mod slutningen af 2018, med en ratio på 0.05 (Bilag 6).

6.2.3.5 Score



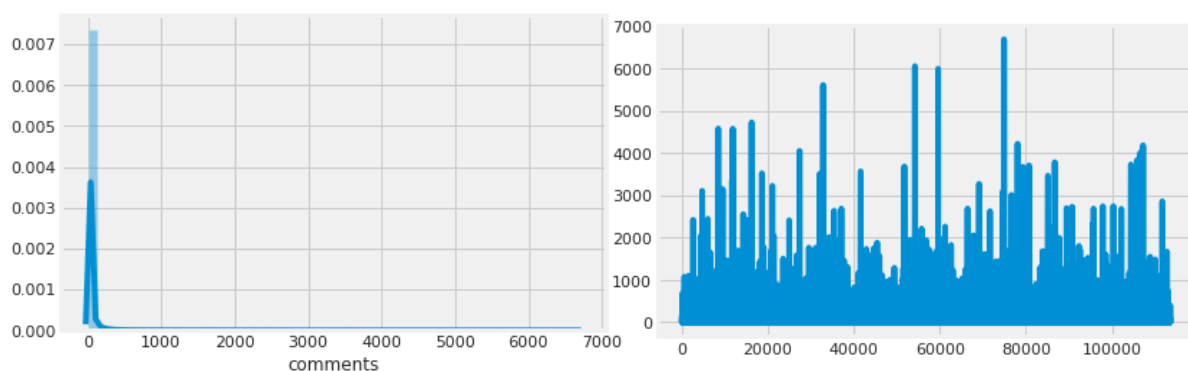
Som med kommentarmængde og upvotes, kan der observeres en mærkbar stigning i, hvor mange point, de delte holdninger scorer på r/unpopularopinion. Trods denne udvikling hen mod oktober, er langt størstedelen af indholdsdelingen samlet omkring lavere scores; 75 procent af holdningerne har 10 i score eller under. Blandt de højest scorende posts, er der, sideløbende med interaktionen i upvotevariablen, et felt mellem 20.000 og 40.000 i score. Den yderste med en samlet score på 51.129 (ibid.).

6.2.4 2019

6.2.4.1 Forfatterskab

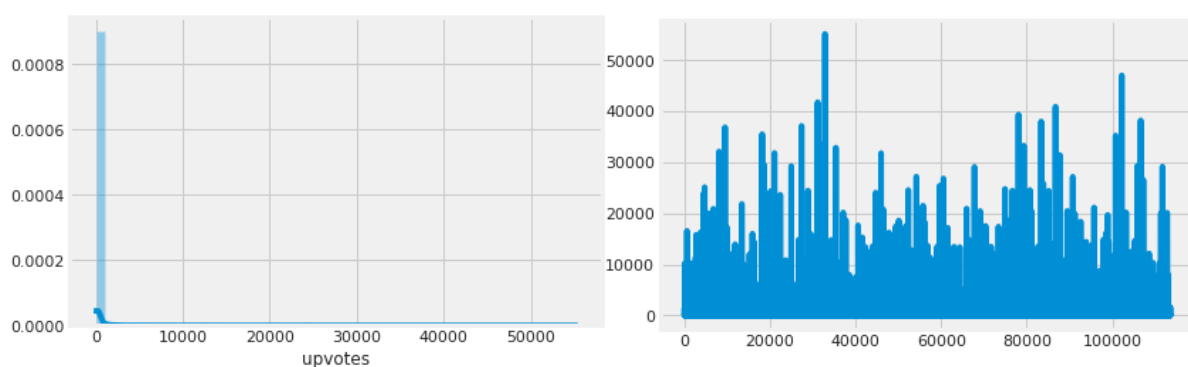
Denne analysedel behandler de første tre måneder 2019, fra 1. januar til og med 31. marts. I observationsperioden bliver der delt 113.576 holdninger, hvoraf 73.533 har tilknyttet forfatterinformation. Blandt disse, er 43.563 unikke brugere. De mest af aktive deler gennemsnitligt op til flere posts per dag: Den mest prominente med 247 posts, fordelt over 90 dage. Som i tidligere epoker, er det vurderet at de enkelte brugeres aktivitet, trods det relativt høje niveau, ikke er udslagsgivende for forummets samlede interaktionsmønster (Bilag 7).

6.2.4.2 Kommentarer



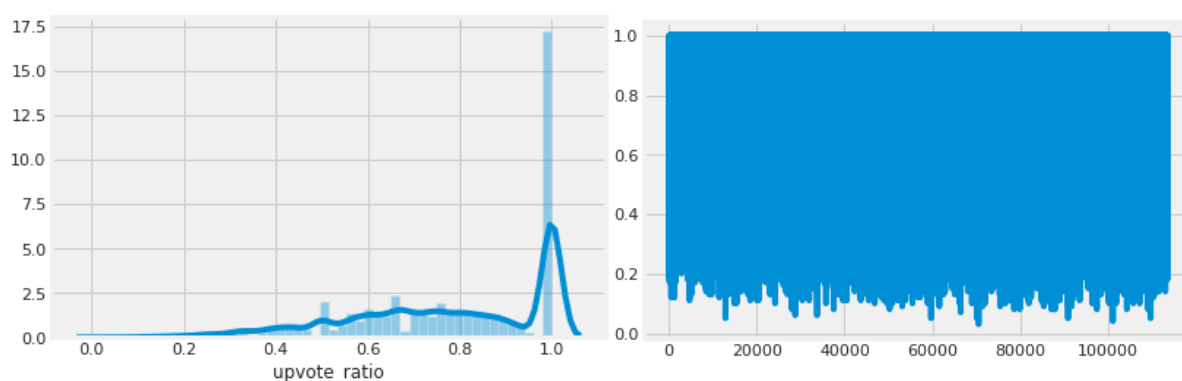
Variablen for kommentarmængde følger på mange måder tendenserne observeret i slutningen af 2018, hvor flere posts i stor grad er debatteret blandt forummets brugere. Gennem de tre måneders observationsperiode, kan der ikke ses nogen tydelig udvikling i kommentarmængden, hvorimod der muligvis kan antages et svagt fald, eller en stagnering, hen imod slutningen. Hoveddelen af holdningerne har en lav mængde af kommentarer tilknyttet: 75 procent med 15 eller mindre. En interessant observation ved 2019 er, at der for første gang i den samlede observationsperiode, er flere posts med 1 kommentar tilknyttet, end der er posts med 0 kommentarer. 29.763 holdninger har 1 enkelt kommentar, og kun 8.030 i relation har ingen kommentarer overhovedet (Bilag 7). Blandt de mest omdiskuterede, er der en relativt stor gruppering med mellem 2.- og 4.000 kommentarer tilknyttet, og nogle få med omkring 6.000. Den mest prominente med 6.687 (ibid.).

6.2.4.3 Upvotes



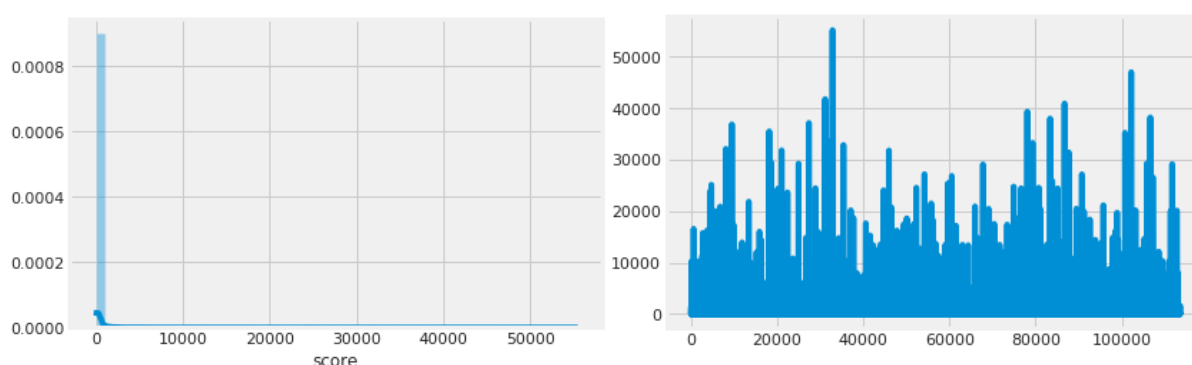
Ligesom med kommentarmængden, opleves der igen en klar koncentration i den lave ende, og endnu engang kan det observeres, at der er langt flest posts med blot én enkelt upvote tilknyttet i 2019. Dette gør sig gældende for 39.013 observationer, hvor den nærmeste følger er 0 upvotes, i 16.062 tilfælde. Udviklingsmæssigt, er der ikke tydelige tegn på forandring i upvotemængden, muligvis et mindre fald i den samlede upvote aktivitet cirka mellem observation 40.000 til 60.000. Fordelt over observationsperioden, men primært i starten og i slutningen, kan der observeres en gruppering af posts med op omkring 40.000 upvotes. Selvom én enkelt desuden kommer nær 50.000, placerer den mest upvotede post sig markant foran de andre med 55.171 individuelle positive tilkendegivelser fra brugerne på r/unpopularopinion (Bilag 7).

6.2.4.4 Upvote Ratio



I 2019 er der, om muligt, en tydeligere relativ samling i den højere ende af skalaen end tidligere. Selvom 1.0 og 0.5 stadig er de mest repræsenterede ratioer, med henholdsvis 36.965 og 4.463 observationer, kan der observeres betydelige grupperinger i den positive ende af skalaen. Det skal tilføjes, at distplottet i 2019 har en fordeling hvor hver *bin* repræsenterer to placeringsmæssige positioner, hvormed f.eks. ratio 0.66 og 0.67 samlet har flere observationer end omkring 0.50 og 0.51 (Bilag 7). Posts med upvote ratio 1.0 udgør en betydelig del af indholdsdelingen for 2019, med over 32 procent af den samlede observationsmængde. 2019 er samtidig den epoke hvor de mest yderliggående observationer kan findes, hvor én holdning har en ratio på 0.03. På tværs af de tre måneder i observationsperioden, er der en jævn fordeling af observationer der tilnærmelsesvis tangerer denne ratio, uden at være helt så radikalt modtaget blandt forummets brugere. Trods disse meget lave ratioer, er det gennemsnitlige niveau endnu engang steget, og er i 2019 på over 0.77 som gennemsnitlig ratio, der nok alligevel primært skal ses i lyset af den store koncentration i ratio 1.0 (ibid.).

6.2.4.5 Score



Variablen for score kan, som i de resterende epoker, ses specielt i relation til variabelen for upvotes. En betydelig mængde af observationerne er placeret i bunden af scorehierarkiet; 75 procent med score 8 eller under, hvoraf 39.008 har en score på 1, og 16.076 har en score på 0 (samlet over 48 procent)

(Bilag 7). Generelt, tangerer adfærden i scorevariablen, det observerede i upvotevariablen. Den højest scorende post for 2019 har opnået 55.199, og de nærmeste følgere positionerer sig mellem 40.- og 50.000 i samlet score (ibid.).

6.3 EMNEANALYSE (TOPIC MODELLING)

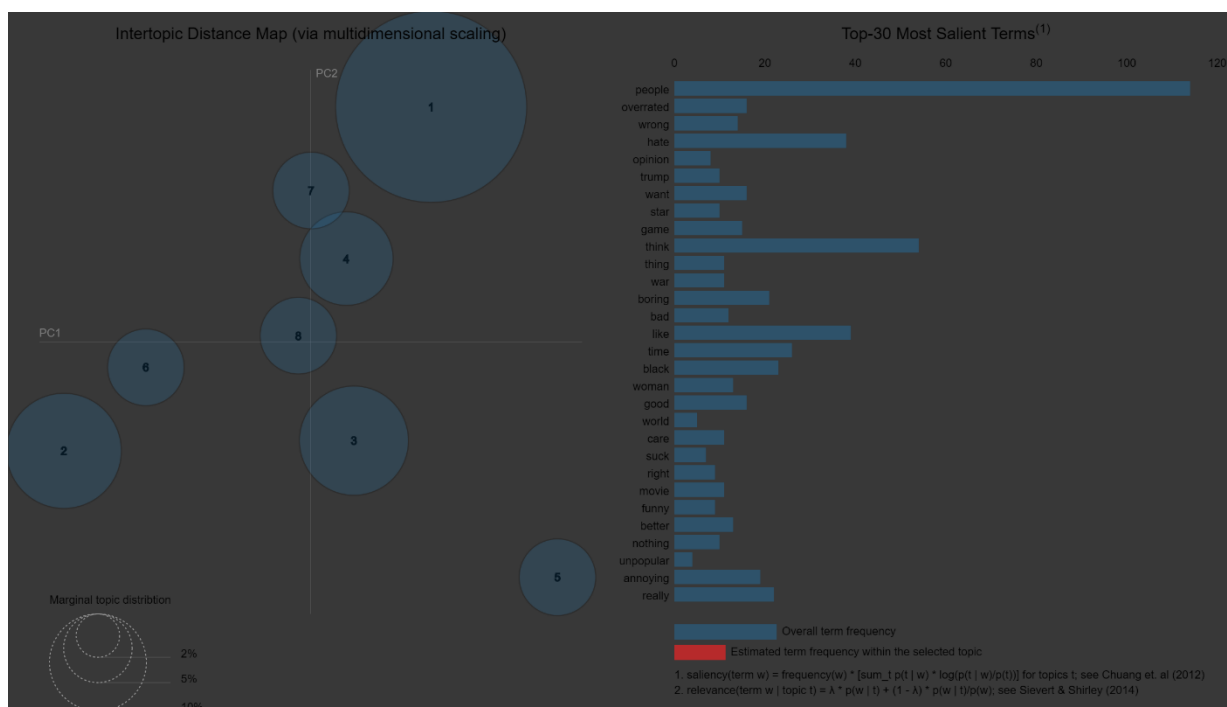
I denne analysedel, søges det at indramme forskellige gennemgående temaer, på tværs af undersøgelsen epoker. Gennemgående, er interaktionen på forummet r/unpopularopinion defineret af at være et åbent rum, hvor brugerne frit kan dele holdningsorienteret og tekstuel indhold, og hvor der ikke pålægges nogle forventninger til den emnemæssige kontekst, andet end en ordentlig tone. I forståelsen af forummet som et socialt rum, hvor det sociale liv udspringer af en vekselvirkning mellem forummets mange forskellige brugere, forventes det, at der vil fremstå forskellige overliggende temaer, der i højere grad end andre emner, fylder i den epokemæssige debat og diskurs på r/unpopularopinion. Der bliver taget udgangspunkt i en analysestrategi, der beror på anvendelsen af maskinlæring og klyngeanalyse til at indramme den udefinerede masse af posts, der bliver delt i de forskellige epoker. Selvom denne tilgang på mange måder kan beskrives som state-of-the-art inden for *Topic Modelling*, er analysens fund i mange tilfælde begrænset af de kvalitative udvælgelsesprocesser, der indgår i udformningen af den anvendte klynge- og emneanalyse (Blei 2012:83-84). Samtidig må analysens fund, ses i lyset af LDA'ens algoritmiske konstruktion, der ikke nødvendigvis er skabt til de formål som specialet søger at belyse (ibid., Song 2018:8). Indramningen af de forskellige epokers overliggende temaer, skal derved ses i relationelt lys, hvor de forskellige klynger og konstellationers udtryk, i høj grad er defineret af forskerens valg. Som nævnt i afsnittet om analysestrategi, er der foretaget visse greb for at overkomme de subjektive vurderinger i nærliggende grad, men analysens emnefund skal ses som en udformning af den empiriske virkelighed, og ikke som en absolut sandhed af det observerede.

Analysen foretages med en forventning om, at der, sideløbende med den observerede udvikling i den deskriptive analyse, vil ske en forandring i den overliggende emneorientering på forummet. Både i variationen af, og indholdet i forskellige tematiske hovedpunkter for r/unpopularopinion. Analysen tager samme udgangspunkt som den deskriptive del, og behandler den samlede observationsperiodes epoker hver for sig. For hver epoke, er der blevet foretaget en *LSI* baseret analyse af den optimale klyngemængde, hvorefter den endelige emnemæssige opdeling er foretaget ved brug af *LDA* (jf. afsnit om analysestrategi). Klyngeinddelingen er samtidig defineret af en intention om at gøre visualiseringen og emneforståelsen relativt simpel og læsbar, og selvom det er muligt at et højere klyngeantal vil øge forståelsen af kompleksiteten og variationen af interaktioner på r/unpopularopinion, er den samlede mængde klynger fastholdt under 20 i hver af analysens epoker.

For hver epoke, vil det samlede korpus af posts blive internt relateret og positioneret, hvor der kan opstå mønstre af sammenhænge mellem brugernes indholdsdeling. Disse sammenhænge kan tolkes som forskellige temaer, hvor primære ord eller sætninger bliver bærende for den kontekstuelle forståelse af det samlede indhold (Chang & Boyd-Graber 2009:1). Simpelt forklaret, betyder det, at hvis de delte posts forholder sig til et bestemt ord i mange tilfælde, bliver dette ord bærende for forståelsen af et emne. Hvis der, f.eks., bliver talt meget om en bestemt film eller person, vil denne fremstå markant i en klynge, men det er ikke nødvendigvis sådan, at man kan tolke på om debatten er primært negativt eller positivt ladet (ibid.). Emneanalysen søger altså ikke at forklare adfærden, eller give indblik i hvordan brugerne forholder sig til de overliggende temaer, men derimod at belyse hvilke emner der fylder mest i debatten mellem r/unpopularopinions brugerpopulation for de givne epoker. Hertil vurderes det interessant at beskue eventuelle forandringer og fluktuationer i den emnemæssige udvikling på r/unpopularopinion, i den samlede observationsperiode. De enkelte posts, der udgør den samlede vurdering af det emnemæssige indhold, kan i nogle tilfælde indgå i flere klynger, fordelt ud fra en relationel sandsynlighedsmodel (Song 2018:1, Blei 2012). For den tekstuelle forståelse, er der taget udgangspunkt i holdningernes primære emneklynge; den klynge hvor de har størst sandsynlighed for at placere sig, i forståelsen af det gennemgående indhold i de forskellige temaer. Indsnævringen af emner, kan have stor påvirkning på den kollektive positionering af de forskellige holdninger, og nogle posts fremstår uden at være i én bestemt temamæssig klynge, men som mere bredt orienteret. Selvom disse også udgør en vigtig del af den empiriske virkelighed, er det blevet valgt at filtrere de observerede klynger ud fra holdningernes specifikke tilhørsforhold. Hermed er posts, der ikke placerer sig tydeligt inden for en klynge, fravalgt for at fremhæve det mere temadefinerende indhold. Denne filtrering påvirker altså ikke den kollektive klyngeanalyse, men anvendes til at belyse det mest markante indhold i hvert hovedtema. I praksis, er denne filtrering foretaget ved at frasortere de holdninger der har under 0.65 i sandsynlighed for at positionere sig i en specifik klynge. Hele emneanalysen foretages med udgangspunkt i det, til formålet, opbyggede bilag om *topic modelling* (Bilag 9). Eftersom visualiseringen af de epokemæssige LDA-konstruktioner fungerer som interaktive modeller, kan bilaget, hvis ønskes, anvendes som værktøj til aktiv undersøgelse af analysens fund.

6.3.1 2016

I 2016 er der udvalgt 8 overliggende emneklynger, der fordeler sig relativt ligeligt på tværs af den samlede indholdsdeling for året.



Dette speciales emneanalyse bearbejdes med udgangspunkt i ovenstående model. Som bilag, kan modellen anvendes som interaktivt værktøj, til at forstå LDA'ens konstruktion, og hvilke vægtninger der ligger til grund for klyngernes fordeling. Denne figur, er en visualisering af den overliggende klyngefordeling for 2016, samt et overblik over hvilke termer der har været mest påvirkende for holdningernes fordeling mellem klyngerne. Figuren til venstre fremstiller klyngerne i et todimensionelt rum, baseret på en bagvedliggende faktoranalyse. Med udgangspunkt i to principalakser, er LDA'ens kompleksitet reduceret, for lettere læsbarhed (jf. analysestrategi) (Chuang et al. 2012). De forskellige klyngers position og størrelse er afgjort af, hvor prævalente termerne i de givne temaer er, eller, hvor ofte de forskellige termer er nævnt, i det samlede *dictionary* af ord (Sievert & Shirley 2014:63). Derved er klyngerne ikke nødvendigvis et udtryk for, hvor mange holdninger indgår, men derimod et udtryk for, hvilke termer der er mest dominerende, og oftest fremstår sammen (ibid.). Figuren til højre viser de 30 mest determinerende udtryk for LDA'ens samlede konstellation. Frekvensfordeling er en visualisering af, hvor ofte ordet fremstår iblandt den samlede postmængde. *People* er den mest anvendte term, blandt de mest prominente. Selvom de fleste fyldord er blevet modereret væk, er der stadig en del termer, som ikke siger meget om det emnemæssige indhold, i forhold til den overliggende tematisering af indholdsdelingen i 2016. Gennemgående, vil emneanalysen søge at undersøge klyngefordelingerne med udgangspunkt i hvad der vurderes som interessante, eller primære termer for en menneskelig kontekstualisering af LDA'en. Herfra foretages analysen af 2016 med udgangspunkt i termer som *trump*, *star*, *game*, *war*, *black*, *woman*, *world* og *movie*.

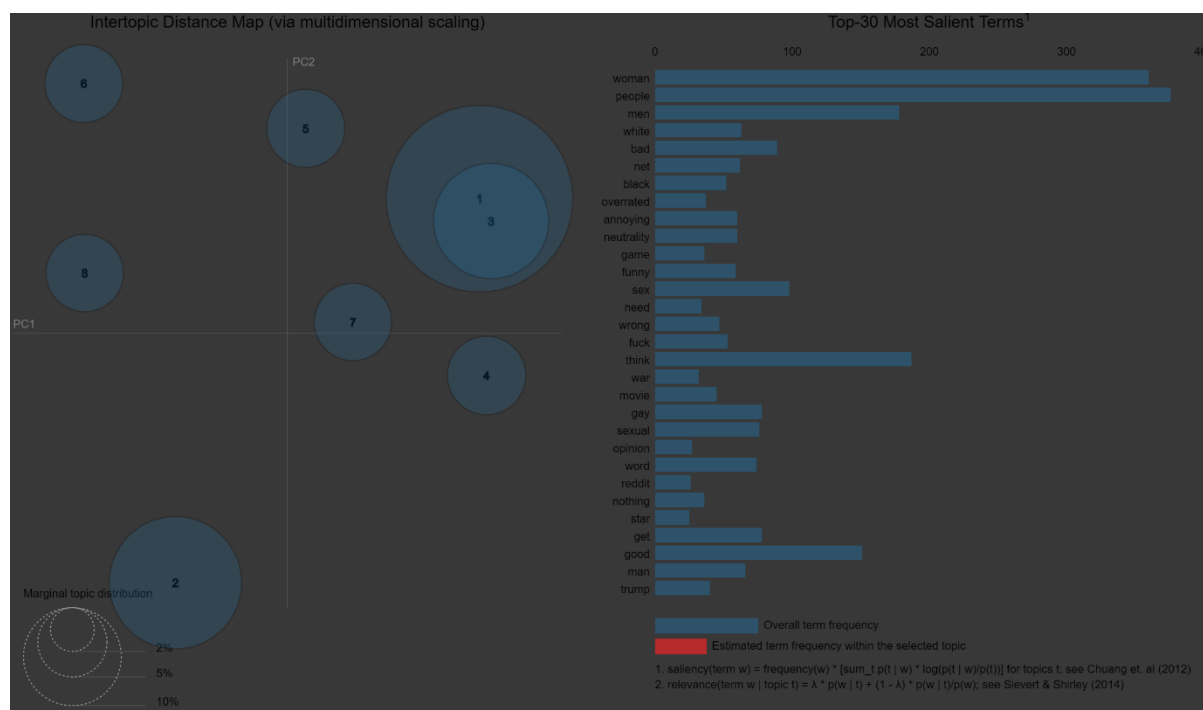
Klyngefordelingen danner et overblik over, hvilke termer der ofte fremstår i kongruens, og samler indholdsdelingen på r/unpopularopinion ud fra den ordmæssige opstilling i de forskellige holdninger. De hyppigst anvendte termer, som visualiseret i den højre del af ovenstående model, kan altså udover at belyse hvilke ting der bliver diskuteret mest på forummet, samtidig give en idé om, hvordan de mest anvendte termer fremstår i samspil med andre ord i klyngernes konstellation. Klynge 1, er den største i konstellationen for 2016, er blandt andet udgjort af udtryk omkring *black, white, dog, cat, culture, american* og *food*. Klynge 2 er determineret af termer som, for eksempel *game, war, star, video, movie, character, show, death* og *force*, såvel som *christmas, pizza, college, teacher, football* og *clinton*. Selvom der kan være forskellige forklaringer på, hvorfor disse termer indgår i samme klynge, er det sandsynligt at, f.eks. *star* og *war* ofte falder i samme kategori, på grund af spillefilmen *Star Wars: The Force Awakens*, fra december 2015. Det kan også være medforklarende til at en term som *force* har en determinerende rolle for klynge 2.

Termen *game*, indgår i flere klynger, men er den eneste blandt de primært fremhævede, der indgår i klynge 3. Denne er desuden determineret af, blandt andet *music, law, abortion, gun, joke, family* og *islam*. I klynge 4 er *trump* blandt de determinerende termer, sammen med, f.eks. *donald, hillary, clinton, internet, woman* og *gay*. *Trump* og *woman* er begge repræsenterede termer i flere klynger, hvor *trump* også bliver nævnt i kongruens med, f.eks. termer som *race, religion, liberal* og *gender*, i klynge 6, hvor der i samme klynge også indgår, blandt andet *music, movie* og *harambe*. De resterende klynger forholder sig i forskellig grad til de mest determinerende termer, hvor specielt *woman*, synes at indgå i flere. klynge 8 forholder sig, som eksempel, til termer som *woman, men, trans, sex* og *feminism*.



6.3.2 2017

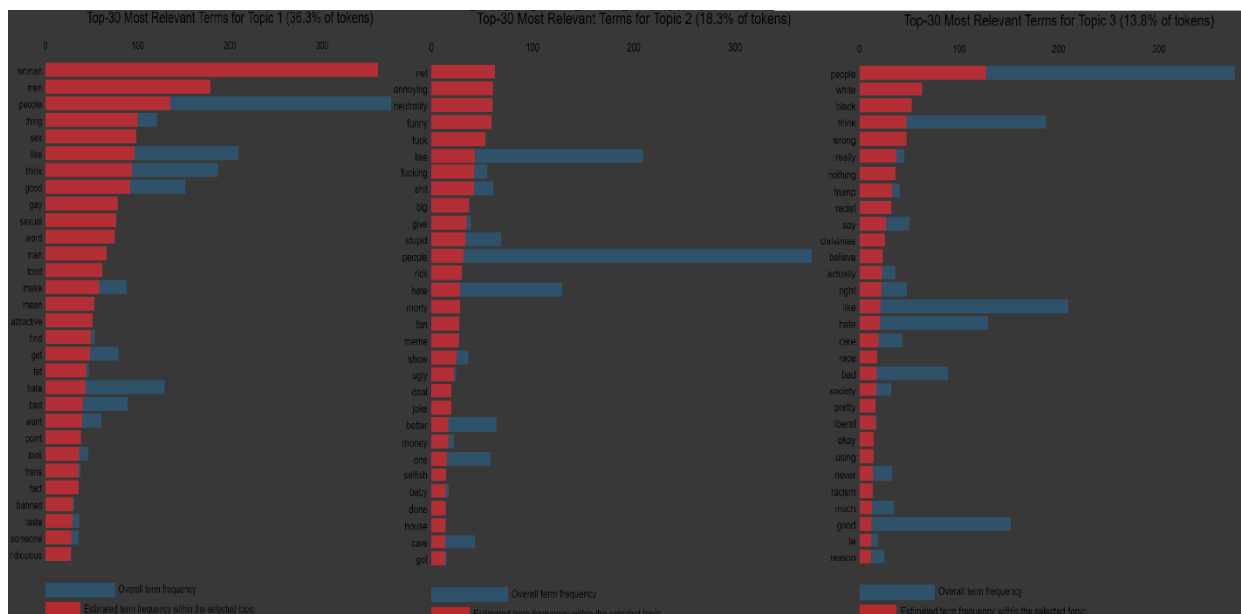
På overfladen fremstår epoken for 2017 tilnærmelsesvis nær det foregående år. Konstellationen af klynger er stadig primært determineret af bærende termer som *woman*, *men*, *movie*, *white*, *black*, *game*, *trump*, *star* og *war*, og er igen blevet fordelt med udgangspunkt i 8 emneinddelinger. Udover gengangerne fra tidligere, fremstår, blandt andet *net*, *neutrality*, *sex*, *sexual* og *gay* som de mest dominerende termer. *People* er stadig den mest prominente blandt de anvendte udtryk i 2017. Termen *woman* fremstår næsten lige så ofte; op mod 400 gange, og er samlet i den største af klyngerne klynge 1, der samtidig forholder sig til udtryk omkring *men*, *sex*, *sexual*, *gay*, *trans* og *food*. Klynge 3, fremstår i 2017's emnekonstellation, som en integreret del af klynge 1. Klynge 3 indeholder samtlige af de observerede termer om *black*, *white* og *racist*. *Trump* er samtidig en markant determinant i denne klynge, sammen med termer som *liberal*, *society*, *race* såvel som *christmas*.



Blandt de nye termer der kan observeres i de prominente temaord, er klynge 2 udgjort af *net* og *neutrality*, såvel som, for eksempel *rick*, *morty*, *meme*, *money*, *show* og *got*. Der kan selvsagt ikke drages nogle sikre antagelser udelukkende fra de forskellige termers kongruens i LDA'en, men det er forventeligt at *net* og *neutrality* er populært anvendte termer på grund af den samtidige udvikling inden for statslig kontrol af internetudbud i USA, betegnet som *net neutrality*, og at *rick* og *morty* sandsynligvis refererer til tv-serien *Rick & Morty*. I forlængelse af dette, er det muligt at *got* skal forstås som et akronym over serien *Game of Thrones*.

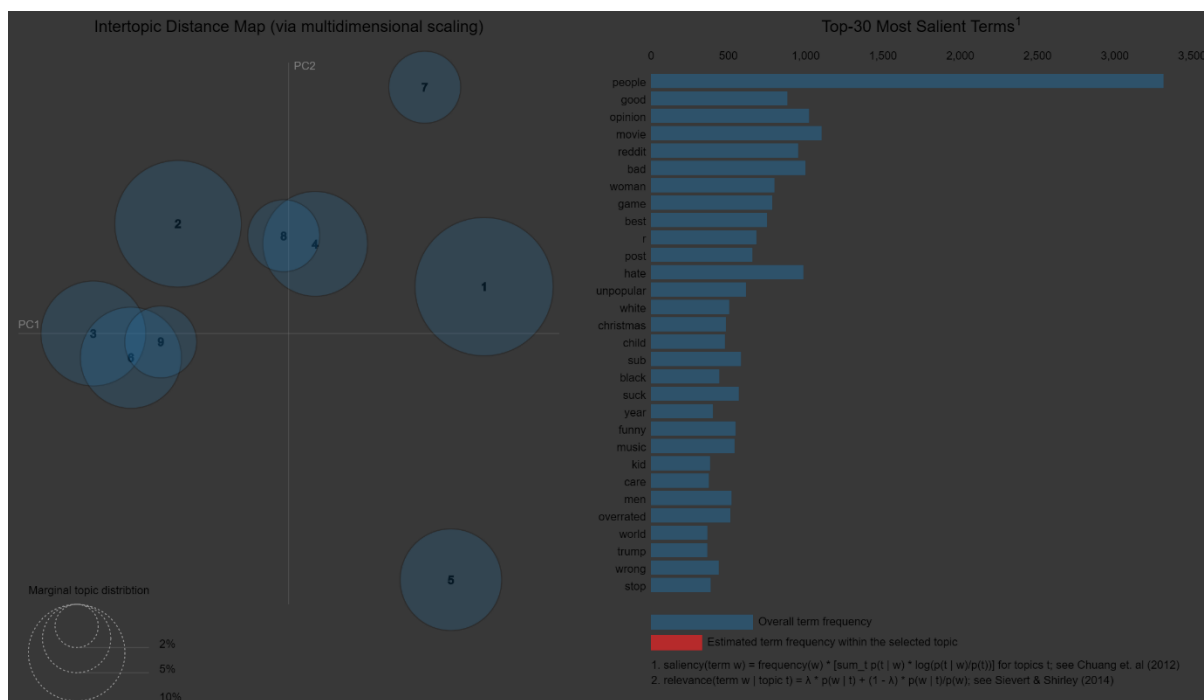
Selvom termen *trump* primært er præsenteret i klynge 3, indgår det også i klynge 5, der desuden er domineret af udtryk som *country*, *american*, *world*, *president*, *political*, *asian* såvel som *music* og *rap*. Termerne *star*, *war* og *movie* er igen samlet i én klynge, i denne epoke er det klynge 6. Her indgår også en stor del af frekvensfordelingen over *movie*, såvel som termerne *film*, *girl*, *guy*, *music*, *animal*, *pizza* og *porn*.

I klynge 8, fremstår *movie* ligeledes, sammen med blandt andet *game*, *show*, *character*, *story*, *youtube* og *business*. Klynge 4 og 7 er begge primært defineret af termer der ikke fremstår blandt de mest prominente. Klynge 4 af præget af termer som *internet*, *rape*, *victim*, *assault*, *retarded* og *political*, og klynge 7 er determineret af, blandt andet *child*, *dog*, *parent*, *school*, *death*, *life*, *government*, *gun*, *gender*, *kill*, *college* og *sport*.



6.3.3 2018

Emneanalysen for 2018 markerer sig ved at have flere forandringer i de 30. mest determinerende udtryk for året. Samtidig er der præsenteret 9 klynger i 2018, med udgangspunkt i LSI, hvor det kan antages at aktivitetsudviklingen også har haft indvirkning på klyngernes konstellation. *People* er stadig den mest benyttede term, anvendt op mod 3.500 gange på tværs af epokens samlede postmængde i r/unpopularopinion. Hverken *net* eller *neutrality* er længere fremtrædende, og *star* og *war* er heller ikke at finde blandt de mest prominente udtryk. *Woman*, *men*, *black*, *white*, *game*, *movie* og *trump* er stadig blandt de mest determinerende. Som noget nyt, er *reddit* med som term, sammen med *post*, *christmas*, *child*, *kid* og *world*.

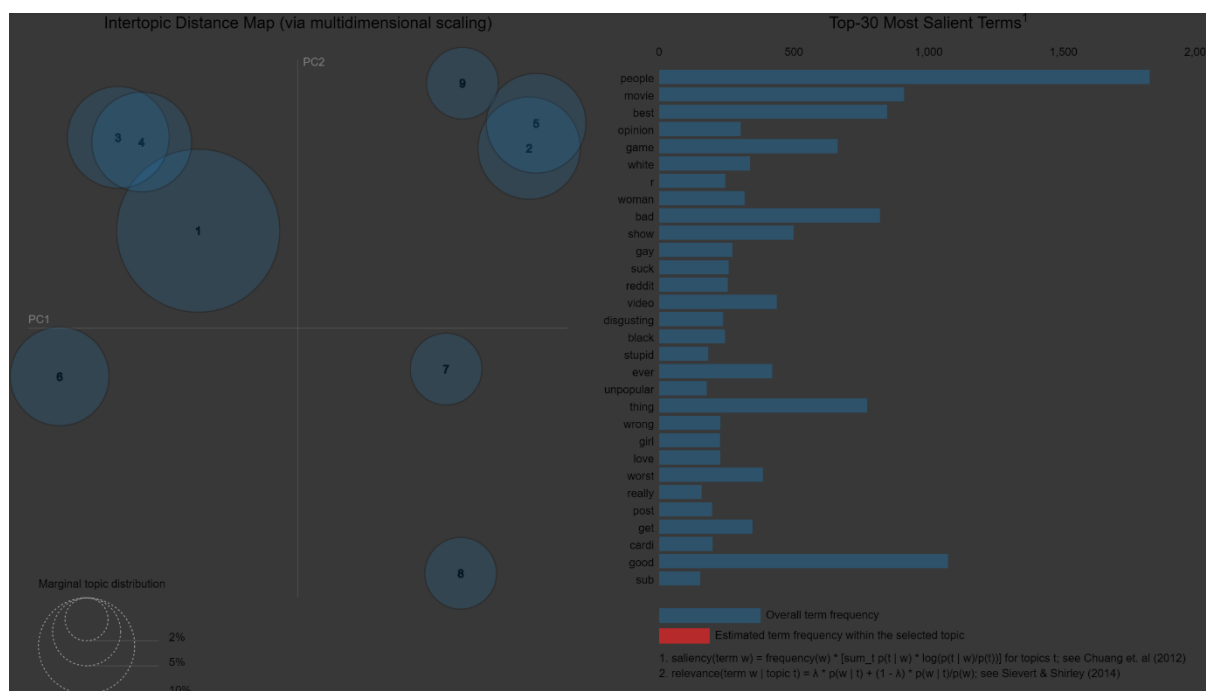


Den største af klyngerne, klynge 1, består af termer som *movie*, *game*, *show*, *video*, *meme*, *culture*, *character sport*, *car* og *porn*. I klynge 2 fremstår flere af de nye termer, *reddit* og *post*, sammen med termer som *subreddit*, *comment*, *human* og *system*. *Woman* udgør, selvom det stadig er et dominerende udtryk i debatten på r/unpopularopinion, en mindre del, i forhold til 2017, og er i denne epoke spredt over to klynger, klynge 3 og 6. Klynge 3 er desuden af såvel, for eksempel *male*, *female*, *gay*, *abortion* såvel som termer sandsynligvis relateret mere til racespørgsmål, *black*, *white*, *racist*, *race* og *nigger*. I samme klynge fremstår termer som *animal*, *pet*, *rewind*, *drug* og *eat*.

I klynge 4, er udtryk om *trump* samlet, sammen med *world*, *country america*, *president*, *liberal*, *conservative*, *government* og *state*. I samme emnegruppering findes termer som *high*, *school* og *gun*. Klynge 5 består af, blandt andet *christmas*, *family*, *money*, *trash*, *dead* og *suicide*. I klynge 6 er der, udover *woman*, en række termer som *child*, *kid*, *men*, *parent*, *adult*, *sex*, *relationship* og *food*.

Klynge 7, 8 og 9 er igen præget af, ikke at være determineret af nogen af de mest prominente termer for 2018. Klynge 7 bevæger sig over termer som *society*, *religion*, *film*, *wall* og *water*. Klynge 8 er determineret af, blandt andet *youtube*, *pewdiepie*, *politics*, *racism* og *rape*. I klynge 9 fremstår, blandt andre termer, de tidligere determinanter, *star* og *war*, som en mindre del af den kollektive debat. De er i denne epoke placeret sammen med termer som, for eksempel *pizza*, *holiday*, *guy* og *girl*.

6.3.4 2019



2019, hvis samlede korpus er en smule mindre end 2018, viser stadig mange af de samme tendenser fra tidligere. Konstellationen af klynger er i 2019 ligeledes inddelt i 9 overliggende temaer. *People* ligger som i de resterende epoker i den absolutte top, blandt de mest fremtrædende, fulgt af termer som *movie*, *game*, *woman*, *show*, *gay*, *reddit*, *video*, *black*, *girl*, *post*, *sub* og *cardi*. *Trump* er for første gang ikke fremtrædende, og *post*, *sub* og *cardi* er nye termer blandt de meste prominente.

Movie, *game*, *show* og *film* er i 2019 samlet i klynge 1, sammen med termer som *character*, *marvel*, *pizza*, *music* og *song*. Klynge 2 er udgjort primært af udtryk omkring *black*, *white*, *racism*, *culture*, *america*, *social*, *medium* og *internet*. Termen *woman* er i 2019 fordelt over 2 klynger, men i klynge 3 fremstår den sammen med, blandt andet *man*, *parent*, *baby*, *family*, *sexual*, *mental* og *illness*, der ligeledes figurerer sammen med termer som *military*, *war*, *star*, *attack*. Klynge 3 og 4 er i LDA'ens konstellation tæt emnemæssigt forbundet. Klynge 4 er udgjort af termer som *gay*, *straight*, *love*, *god*, *muslim*, *sport*, samlet med hvad der kan vurderes som mad- og drikkerelaterede termer, for eksempel *food*, *water*, *milk* og *apple*.

Det resterende klynger er mere varieret i deres udtryk, og forholder sig i mindre grad til de determinerende termer for 2019. Klynge 5 er udgjort af, for eksempel *school*, *kid*, *child*, *dog*, *animal*, *suicide* og *friend*. Den nyligt observerede term, *cardi*, refererer muligvis til den amerikanske performer Cardi B, og er i klynge 6 præsenteret sammen med *girl*, *guy*, *music* og *sound*, såvel som, blandt andet *cheese*, *chocolate*, *pewdiepie* og *pet*. I klynge 7 og 8 er begge forholdsvis små grupperinger i den samlede termfordeling. Klynge 7 har termer som *trump*, *joke*, *meme* og *college*, og klynge 8 er udgjort

af, blandt andet *reddit*, *subreddit*, *community*, *politics*, *porn* og *unpopularopinion*. I klynge 9, er *woman* sammen med, f.eks. *sex*, *men*, *gender*, *abortion* og *child*.

7 DISKUTERENDE KONKLUSION

Igennem dette speciale, er det empiriske materiale blevet belyst med det formål at uddybe forståelsen for socialiteten i internetaalderen, mere specifikt med fokus på det sociale liv på den anonyme foraplatform *reddit*. Anvendelsen af arbejdsspørgsmål har bidraget til en bred afdækning af interaktionerne blandt brugerne på forummet *r/unpopularopinion*, hvor såvel udviklingen af forummets form og tilstand, som det diskursive univers der udspiller sig mellem brugerne, har været søgt belyst. Problemformuleringen er besvaret ved hjælp af primært kvantitative metodetilgange, og søger mod at simplificere en forståelseshorisont over den massive mængde af data der gennemstrømmer forummet dagligt. Intentionen med specialet er at bearbejde indramningen af store datamængder i relation til socialiteten på internettets sociale medier, hvis indflydelse og anvendelse kun er blevet en større og mere vital del af samfundets sociale eksistens siden internettets spæde opstart. Indsamlingen af empiri fra *reddits* brugerflade har, udover at bidrage til specialets analytiske fundament, også bidraget til forståelsen af, hvordan anonyme online fællesskaber kan anvendes i sociologisk forskning, og hvordan programmeringssprog kan bruges interaktivt og på tværs af specialerammen. I forlængelse af dette, har anvendelsen af Python givet mulighed for at samle specialets forskellige elementer og strømlinet processen fra indsamling til analyse.

7.1 REFLEKSIONER OVER DEN DESKRIPTIVE ANALYSE

Specialets fokusområde forholder sig til det tekstuelle forum om upopulære holdninger på *reddit*, *r/unpopularopinion*. Analytisk, har der været fokus på interaktionen mellem de aktive brugere på platformen, og hvordan det udspiller sig i en kvantitativ forståelseshorisont. Selvom det tekstuelle element spiller en vigtig rolle i forståelsen af den interaktion der udspiller sig, beskæftiger en stor del af specialets analyse sig med det relationelle, omkring udvalgte indikatorer for social interaktion mellem forummets brugere. Dette er sket for at uddybe, præcist hvordan forummets udtryksform forandres gennem observationsperioden, og efterhånden som der kommer flere subscribers.

Tidsserieanalysen forholder sig specifikt til, hvordan brugertilstrømningen påvirker aktiviteten på *r/unpopularopinion*, med udgangspunkt i antallet af dagligt delte holdninger på forummet. Det er tydeligt, at brugertilstrømningen til *reddits* samlede platform også har haft stor indflydelse på aktiviteten på *r/unpopularopinion*, og dette kan specielt observeres i den udvalgte observationsperiode, hvor 2018 og 2019 for alvor ser stor stigning i delte holdninger. Sammenholdt

med den tilgængelige, eksisterende information vedrørende r/unpopularopinion, har tidsserien fungeret som en understøttende belysning af den observerede brugertilstrømning til forummet. Forholdet mellem hvor mange brugere der er subscribers, og hvor meget aktivitet der er på forummet, anskues som forholdsvis ensartet gennem den samlede observationsperiode. Den store grad af passivitet, der generelt opleves blandt brugerne på sociale medier, kan altså siges ligeledes at gøre sig gældende for brugerne på r/unpopularopinion, hvor der er stor forskel på antallet af brugere og antallet af holdningsdelinger.

Den univariate analyse har været anvendt til at belyse forskellige implikationer af brugerinteraktionen på r/unpopularopinion, i lyset af den eksisterende forskning og teoretiske overvejelser for rummet. Igennem forfattervariablen er der blevet observeret flere interessante ting. Den store grad af manglende tilknyttede forfatternavne, kan i praksis være resultatet af flere forskellige effekter; det er muligt at der er fejl i den præsenterede empiriske virkelighed, og at en del af manglen er et resultat heraf. Det vurderes dog, at det primært er resultatet af, at de givne brugere sidenhen er blevet slettet, og derved ikke fremstår i systemet. Derved kan det antages, at der er implikationer for *throwaway* kultur, også på forummet r/unpopularopinion. Samtidig understøttes teorien om de aktive brugeres gennemgående passivitet af, at langt størstedelen af den samlede aktive brugergruppe på forummet kun har én eller få delte holdninger, på tværs af observationsperioden.

Kommentarmængden kan, ligesom forfattervariablen, give indblik i flere interessante tendenser. Som en del af reddit's platform, positionerer r/unpopularopinion sig konstant højt, i forhold til mængden af kommentarer, og er udslagsgivende for forummets høje aktivitetsniveau i eksisterende statistiske overblik (Subredditstats.com 2019). Trods den relativt høje kommentaraktivitet på forummet, er der stadig en lang række observationer med få eller ingen kommentarer tilknyttet, og gennemsnittet ligger konstant i den lave ende, også i de sidste epoker, hvor der er et højt antal subscribers, og større variation mellem observationerne. Uden at kunne gisne om indholdet og adspredelsen af brugere i kommentarsporene for de forskellige observationer, er det tydeligt at brugertilstrømningen har stor indvirkning på mængden af kommentarer i r/unpopularopinion. Specielt hen mod slutningen af 2018 kan der observeres et spring i kommentaraktivitet, der virker voldsomt i forhold til den generelle brugertilstrømning. I 2018, blev der observeret en popularisering af den upopulære holdning som udtryksform på et andet populært socialt medie, Twitter, der kan være medforklarende både til tilstrømningen til r/unpopularopinion, og til den store aktivitetsudvikling i kommentarmængde (Adam 2018, Nembhard 2018). Her må det selvsagt understreges, at denne vurdering ikke er kontrolleret, og at det må ses som en overvejelse over vilkårene i den empiriske virkelighed for forummet. Denne argumentation skal ses i lyset af Maffesolis indramning af det sociale liv i det postmoderne samfund. Der argumenteres for, at reddit's brugerpopulation har stor bevægelsesfrihed, ikke blot mellem de

forskellige fora på reddit's platform, men også på internettets andre, og meget forskelligartede sociale medier, herunder også, f.eks. Twitter. På denne måde opleves der interessesammenslutninger der, som med r/unpopularopinion kan vokse sig til veritable bevægelser med massiv brugertilslutning. Den relativt markante stigning i aktivitet i perioden vurderes med afsæt i argumentet for de sociale mediers gennemgående brugerpassivitet. Mange nye brugere føler de skal lære de *sted*specifikke spilleregler og normer, før de engagerer sig i debatten. I dette tilfælde er det muligt at denne aktivitetsmæssige udvikling i slutningen af 2018, kan være et resultat af en brugertilstrømning, hvor en stor del af de aktive brugere allerede er bekendt, som minimum, med dele af de sociale spilleregler for *stedet*. Ovenstående argumentation, sker med udgangspunkt i Maffesolis udtryk om massernes bevægelsesfrihed, på tværs af forskellige interessesammenslutninger og konstellationer på internettets *nebula* af interaktionsmuligheder. Populariseringen af udtryksformen omkring upopulære holdninger, kan derved givetvis have trukket brugere fra Twitter mod reddit's underforum r/unpopularopinion, eller i hvert fald sat fokus på dets eksistens og form. Da udtryksformen allerede er oplevet gennem Twitter, er det muligt at nye brugere på r/unpopularopinion, føler sig trygge omkring den bekendte udtryksform.

Tesen om den gennemgående passivitet blandt brugerpopulationer på sociale medier, bliver i specialets empiriske materiale understøttet af ideen om, at selv aktive brugere i udgangspunktet oftest engagerer sig primært med votes. *Upvotevariablen* belyser, at selv om der er rimelig høj kommentaraktivitet på r/unpopularopinion, er aktivitetsniveauet for votes, konstant og relationelt langt højere gennem den samlede observationsperiode. I relation til mængden af upvotes, er der igennem den univariate analyse, gennemgående blevet konstateret et overtal af upvotes, i forhold til downvotes.

7.2 TEMATISERING AF DEN SAMLEDE INDHOLDSDELING

I bearbejdningen af det empiriske materiale for dette speciale, har én af de metodologiske udfordringer været at indramme det tekstuelle indhold, uden kvalitativt at dykke ned og belyse indholdet med en specifik hypotese, eller teoretisk fundering. I stedet, har formålet været at anvende kvantitative tilgange, der for så vidt muligt, kan give indblik i brugernes delte indhold på tværs af det enorme datagrundlag. Reddit's stigende popularitet og brede platform gør, at der er konstant og alsidig indholdsdeling, hvor temaer og mening kan være vidt forskellige. Selv inden for et mindre afgrænset område, som r/unpopularopinion, kan der være stor adspredelse i, hvad brugerne deler, og hvordan de går til debatten. *Topic Modelling*, eller *Latent Dirichlet Allocation*, som den specifikke metode anvendt i dette speciale, har mange praktiske anvendeligheder, der gør den fordelagtig, i en formidling af varieret og massivt dataindhold (Blei et al. 2003, Chang et al. 2009). Den interaktive form, der er

visualiseret i emneanalysen, giver mulighed for at forholde sig dybdegående til relationen mellem de forskellige bærende termer, for *LDA's* emner (ibid.). I specialet har det været problematiseret, hvordan *LDA* arbejder med udgangspunkt i, at forskeren er bekendt med det reelle antal af specifikke emner for den givne *LDA* Model, modsat den praktiske og empiriske virkelighed for specialets analyse (Chuang, Manning, and Heer 2012; Sievert and Shirley 2014:67–69). Igennem analysen, har der da også været en tydelig tendens af "rod" i emnerne, eller situationer, hvor det har været svært at udrede et bestemt emne fra den givne klynge. Trods dette, bidrager *LDA* modellen med en lang række objektive sandheder for specialets empiri, der bidrager til en vis grad af tolkning. F.eks., er det uagtet subjektive overvejelser og behandling af *LDA*, altid de mest prominente termer for indholdet, der bliver fremhævet i modellens visualisering (Sievert and Shirley 2014, Blei et al. 2003). Selvom det i sig selv ikke er nok til at understøtte de misogyne tendenser fremhævet af, bl.a. Massanari 2017, og Kilgo et al. 2018, er termen *woman*, én af de få der er repræsenteret som dominerende på tværs af samtlige epoker.

Summen af brugernes indholdsdeling, er udgjort af en betydelig mængde forskelligartet indhold, der trods forsøg på at objektivisere indholdet, til en vis grad må analyseres kvalitativt for at uddybe forståelsen. Gennemgående er der mange interessante kvalitative vurderinger, der kan udtrækkes fra specialets emneanalyse. De primære dominerende termer på tværs af observationsperioden er udgjort af nogle tydelige dialektiske modsætninger mellem *woman* og *men* (selvom *men* ikke indgår som primær i 2019), og *black* og *white*. I *LDA'en*, fremstår antonymerne gemenligt i samme klynger, gennem observationsperiodens epoker. *Woman* er et mere dominerende term gennem epokerne på *r/unpopularopinion*, ift. *men*, men fremstår begge ofte sammen med termer som *sexuality*, *sex*, og andre køns- og seksualitetsrelaterede termer. *Black* og *white*, er generelt set sat i relation til hvad der kan vurderes som racemæssige termer, på nær 2016, hvor det er sværere at tyde, hvilken forbindelse termene har til de resterende. På *r/unpopularopinion* er det en regel at indholdsdeling ikke, bl.a. ikke må være politiserende. Trods dette, er det en gennemgående tendens, at politiske termer er en markant del af det samlede indhold. Dette er specielt gældende omkring 2016 og 2017, hvor *trump* fylder en stor del af debatten. Termen *movie*, er konsistent i sin tilstedeværelse blandt de mest determinerende, og der er en interessant forbindelse mellem *movie* og termene *star* og *war*. Det har i analysen været argumenteret for en forbindelse med film franchisen *Star Wars*, og de nyere film der er blevet udgivet.

Emneanalysens fund, fungerer i forståelsen af, hvilke interessepunkter der er dominerende blandt de aktive brugere på *r/unpopularopinion* igennem observationsperioden. Selvom det er nødvendigt at tolke på indholdet, og dykke dybere i forholdene mellem *LDA's* klyngekonstellationer, er der vigtige tematiske grundinformationer der kan udtrækkes fra den visuelle fremstilling i analysen. Det vurderes

at, der på tværs af epokerne, bliver fastholdt mange lignende hovedtemaer, men at der samtidig bliver introduceret nye debattemner for indholdsdelingen løbende. Det anskues, at en del af indholdet på r/unpopularopinion forholder sig til populærkultur, og i denne forbindelse også hvad der generelt fylder meget i den offentlige debat – på internettet som i andre offentlige medier. Dette eksemplificeres gennem, blandt andet, *net* og *neutrality* for 2017, *youtube* og *pewdiepie* i 2018 og *cardi* i 2019.

Selvom popkultur kan vurderes til at være en markant del af emneindholdet på r/unpopularopinion, virker forummet domineret af hvad der kan anskues som kontroversielle diskurser i den offentlige debat, fra racespørgsmål over køns- og seksualitetsdebat til politik (primært amerikansk). Det amerikanske fokus skal sandsynligvis ses som produkt af flere effekter. I overblikket over reddit, er der blevet argumenteret for, at op mod 50 procent af den samlede brugerpopulation er amerikanske. Samtidig har præsidentvalget i 2016 været anset som ekstremt kontroversielt af mange årsager, og er noget der for alvor har fået folk til tasterne på online sociale medier, og det virker ikke til at r/unpopularopinion er gået fri af dette (Mills 2018).

7.3 EMNEANALYTISK DISKUSSION

Dette speciale arbejder med at indramme de dominerende termer for r/unpopularopinion i sin udvikling gennem den samlede observationsperiode. Rent diskursivt, har analyserne ikke bearbejdet eventuelle fluktuationer i, hvordan brugerne forholder sig til forskellige emner, og om der sker en forandring gennem løbet. I den univariate analyse, er det blevet fremhævet, at nogle delte holdninger opnår mere opmærksomhed, og bliver mere anerkendt og anset blandt massen af interaktioner på r/unpopularopinion. I bilag 8 kan der dannes et overblik over de mest opmærksomhedsskabende holdningsdelinger på tværs af analysens epoker. I det følgende afsnit, vil der blive inddraget interessante og profilerede holdninger, der vil sættes i relation til de mest dominerende termer for de givne undersøgelsesepoker. Dette afsnit skal ikke ses som en diskursanalyse over indholdet, men nærmere som en fundering over det dikotomiske forhold mellem de dominerende termer *women* og *men*, og *black* og *white*. I forlængelse af dette, vil der blive vurderet hvilke ting der opnår anerkendelse og give et kortere indblik i, hvordan brugerne generelt udtrykker sig omkring de bærende termer, igennem den samlede observationsperiode.

7.3.1 Mænd, Kvinder og kampen mellem kønnene

Der bliver gennemgående observeret en stor interesse omkring kvinder, på tværs af de empiriske epoker. Termerne der forholder sig til kvinder varierer, men *woman* er meget konsistent i sin tilstedeværelse blandt de mest dominerende. For 2016, har flere de mest profilerede holdninger også

kvinder som omdrejningspunkt. Den mest omdiskuterede *'The only Privilege in this day & age is Attractive Female Privilege.'* (u/DarkBloodSoulsBorne, 4. oktober 2016), og den mest upvotede *"The best thing about Hillary Clinton not winning is that the first female president won't be Hillary Clinton."* (u/bodan94, 11. november 2016), forholder sig til termen *female*, og det kan vurderes, at holdningerne er en smule polariserende i forhold til ligestillingsdebatten. Blandt de mest primære holdninger der kan relateres til termen *woman*, kan det antydes, at debatten bærer præg af at være om kvinder, fremfor af kvinder. Dette er en selvsagt en vurderingssag, men udmeldinger som *'Women should return to subservience'*, *'Women turn a corner at 30'*, *'Women over 35 should not breed'* og *'Housewives are infinitely more valuable than career women'*, fremstår alle som primære i forhold til termen *woman*, og der observeres generelt hvad der kan betegnes som skelsættende holdninger blandt brugerne. Selvom *men* ikke fremtræder blandt de mest dominerende termer i 2016, fremstår det som i de resterende epoker oftest i kongruens med *woman*. Forholdet mellem debatten om mænd og kvinder, viser en tydelig forskel i, hvordan brugerne tilgår emnerne. Selvom der er eksempler på undertrykkende kommentarer mod hankønnet (f.eks. *'MEN AR THE WORST!!!!!!!!!!!!!!1'*, *'Men are disgusting pigs'*, og *'Men have to change'*), opleves det, at flere holdninger hvor *men* indgår, også forholder sig til *woman* som term, og i flere tilfælde med en fordelagtig fremstilling af mænd, i det dikotomiske forhold mellem kønnene i debatten på r/unpopularopinion. Dette kan ses med holdninger som: *'Generally speaking, women are objectively more shallow than men'*, *'Men suffer enormously from the patriarchal/traditional gender roles'* og *'Women really aren't as funny as men'* (Bilag 9).

I 2017 er *woman* som term, ekstremt dominerende, og er i høj grad determinerende for konstellationen af emneklynger. Blandt de mest profilerede holdninger, fremstår kønsdebatten ikke direkte, og der er ikke på samme måde tydelige konnotationer mod misogyne tendenser, selvom det er tydeligt fra emneanalysen, at kønsdebatten stadig er en markant del af indholdsdelingen for 2017 (Bilag 8, Bilag 9). Ud fra det analogiske overblik over *woman* som term, kan det vurderes at der til stadighed primært bliver diskuteret omkring kvinder, frem for af kvinder på forummet r/unpopularopinion i 2017. Trods ét enkelt observeret eksempel på en selverkendt kvindelig bruger: *'I am a woman, and I am prolife'*, er en markant del af indholdet omkring termen *woman*, ligesom i 2016, hvad der kan vurderes som objektiviserende omkring kvinders rolle og position i samfundet (f.eks. *'Pregnant women disgust me'*, *'Women arrested with meth or crack cocaine should be forcibly sterilized'* *'Women are overrated'* og *'Women over 30 are expired raisins'* (Bilag 9)). *Men* er i 2017 også blevet en del af de mest dominerende termer, og der opleves da også større variation i indholdet omkring mænd på r/unpopularopinion. Det kan vurderes, at flere holdningsudmeldinger stiller sig negativt ladet overfor hankønnet, f.eks. *'Men are disposable'*, *'Men should be sterilized when they're*

young’, og *‘Men are too easily offended’*. Selvom der opleves flere eksempler på mandekritiske udmeldinger end i 2016, opleves der stadig en mere positiv stemning omkring termen *men* ift. *women*, og forskellen mellem mænd og kvinder fylder mere under termen *men* (*‘Men are held to a higher beauty standard’*, *‘Men and women are on average equally intelligent, but men make up more of the genius population’*, *‘Women are empirically more sexist than men’* og *‘Men are more objectified than women’*) (Bilag 9).

I 2018 opleves der et interessant skift i indholdet i holdningerne relateret til *men* såvel som *women*. Hvor der tidligere har været langt tydeligere forskelle i udmeldingerne omkring mænd og kvinder, virker det i 2018 som om der er mere balance i forholdet mellem kønnene. Selvom *woman* stadig er et mere dominerende term end *men*, er der mere kritisk indhold om mænd end tidligere, og holdningerne opleves generelt noget mere varieret end tidligere. Det må understreges, at der ikke kan drages konklusioner om kønsfordelingen, eller fluktuationer i brugerpopulationen, men det er en anskuelse for specialet, at der opleves en mere ligelig fordeling mellem forskellige subgrupperinger på forummet *r/unpopularopinion* i 2018, i forhold til tidligere epoker. Selvom kvinder til stadighed bliver objektiviseret i debatten på forummet (f.eks. *‘women shouldn’t be over 120lb’*, *‘All women should be 90-110lbs’* og *‘Women should be housewives’*), går mænd ikke, på samme måde som tidligere, fri af den kritiske objektivisering af hankønnet (f.eks. *‘All men are paedos’*, *‘Men are whack’*, *‘Men are for reproduction and catching bullets’*, og *‘Men. Are. Such. PIGS’*) (ibid.).

2019 viser en tydelig tilstedeværelse af den kvindelige stemme i debatten på *r/unpopularopinion*. Langt flere holdninger end tidligere tager udgangspunkt i variationer af holdninger fra et selverkendt, kvindeligt synspunkt: *‘As a woman...’/‘Because I am a woman...’/ ‘AS A WOMAN...’*. På tværs af kønsspektrummet opleves der i 2019 såvel kritik og objektivisering, som forherligelse og helliggørelse, og debatten opleves generelt langt mere ligelig end tidligere epoker (Bilag 9).

7.3.2 *Black/White*

Hvor dikotomien mellem *men* og *woman* på *r/unpopularopinion* opleves som en relativt tydelig kønsmæssig debat, er forholdet mellem *black* og *white* som termer en smule mere tvetydigt. I 2016 er der relationer mellem såvel mediekultur, f.eks. *‘Black Mirror is overrated’*, *‘Orange is the New Black is not a good show’*, *‘Black Widow is a boring character’*, over design *‘The Matte Black iPhone 7 is boring’*, til *Black Lives Matter*, en omdiskuteret politisk bevægelse, der var et populært diskussionsemne på sociale medier omkring 2016, f.eks. *‘Black lives matter.’* Og *‘Black Lives Don’t Matter’*. Selvom der af variationer i brugen af termen *black*, er der, specielt i forholdet mellem *black* og *white*, tydelige racemæssige konnotationer for debatten på *r/unpopularopinion*. Det er i specialet antaget, at brugerpopulationen er domineret af yngre kaukasiske mænd, og der er flere observationer,

der understøtter denne antagelse. Der er holdninger der antyder tilstedeværelsen af andre etniciteter, f.eks. *'I don't feel sorry for white people'* og *'I'm black and can't stand black lives matter'*, men en stor del af holdningerne omkring såvel *black* som *white*, opleves, hvis ikke direkte racistisk, som favoriserende for *white* etniciteter. Dette opleves generelt, ikke som direkte nedladende mod *blacks* specifikt, men mere som en grundlæggende retfærdiggørelse af at være *white*: *'Whites are a misunderstood ethnic group'*, *'It is possible to be racist towards white people'*, *'Minorities are just as racist as whites'* og *'I think the fight against "White Male Privilege" has gone too far'*. Trods dette, vurderes det at der, med *black* som mere dominerende end andre etnicitetsrelaterede termer i 2016, kan observeres en del indhold der kan betegnes som reelt racistisk, f.eks. *'Asians are worse than blacks'*, *'Black People suck'*, *'Black people need their own country'* og *'Black people really should go back to Africa'*.

På samme måde som tidligere, er der variationer over brugen af termerne *black* og *white* i 2017 (f.eks. *'I like black licorice'*, *'I like black jellybeans'* *'I don't like Black Sabbath'* og *'white grains are cucking you nutritionally'*). Igen er der en relativt tydelig dominans af racerelaterede debatemner. Det er mindre udtalt end i 2016, men der kan stadig observeres et overtal af, hvad der kan beskrives som racistiske udmeldinger blandt brugerne, f.eks. *'Blacks are god's biggest mistake of creation'* og *'I think whites are genetically superior'*. Der er i epoken for 2017, relativt stor interesse omkring emner om *white supremacy*, *white genocide*, og *white privilege*, hvor der opleves mange holdninger der taler til fordel for de *white* etniciteters position i det globale samfund. Trods den stadige dominans af en undertrykkende retorik omkring termer som *black* og *white*, opleves der er en udvikling, og flere eksempler på, at debatten varierer og ikke nødvendigvis fastholder hegemoniske tendenser, eller retfærdiggør en specifik retorik (*'it's okay to be black'*, *'it's okay to be white'*, *'is this just a place for angry white people to vent?'*) (Bilag 9). I forlængelse af den univariate analyse blev det observeret, at én af de mest omdiskuterede holdninger for 2017 på r/unpopularopinion, forholder sig til racedebat: *'The amount of white people celebrating they're very quickly becoming a minority in their own country makes me sick'* (bilag 8).

I det dikotomiske forhold mellem *black* og *white*, virker det på mange måder oplagt, at racespørgsmål fylder en stor del af indholdet. Det skal understreges, at inddragelsen af disse termer, såvel som *men* og *woman* ikke sker for at understrege hverken misogyn, racistisk eller hegemonisk kultur blandt brugerpopulationen, men for at tage udgangspunkt i nogle af de termer der fylder meget, og derved observere en eventuel forandring. Med *black* og *white*, anses der er en betydelig forandring fra tidligere epoker på r/unpopularopinion i 2018, hvor det vurderes at direkte racistiske udmeldinger fylder langt mindre end tidligere. Der er en interessant forbindelse mellem termerne *black* og *white* i 2018. Hvor holdninger forbundet til *black* fremstår mindre direkte raceladet eller undertrykkende end

tidligere, er der mange holdninger omkring *white*, der fastholder en markant forsvarsposition (*'it's okay to be white'*, *'whites are discriminated against'*). Indholdsdelingen omkring termene *black* og *white*, minder på mange måder om hinanden på tværs af 2018 og 2019. Selvom der oplevedes mere tydelige racistiske undertoner i 2016 og 2017, anskues der stadig at være en skævvridning i debatten omkring racetematikker (*'we have to accept that black people have lower avg IQ'*, *'Blacks are rapists'*, *'There will soon be a white uprising!'*, *'White supremacy is widely misunderstood'*. *Black* og *white*, forholder sig til mange forskellige ting, fra populær- og mediekultur til mad, f.eks. (*'Jack Black is obnoxious'*, *'Black Widow is a useful avenger (MCU)'*, *Black Jellybeans are good'* og *'Black Panther was just OK'*) (Bilag 9). Gennemgående opleves det, at indholdet på r/unpopularopinion i 2018 og 2019, er bredt varieret, og repræsenterer en bred brugerpopulation og et bredt holdningsspektrum. Trods dette, opleves der til stadighed tendenser for en dominerende kultur, der kan opleves undertrykkende for nogle grupperinger i samfundet, som én bruger beskriver det; *'Reddit can be frustrating when you're black'* (Bilag 9).

7.4 AFRUNDING

Anvendelsen af termene *black* og *white*, og *woman* og *men*, fremhæver muligvis nogle ekstremer ved indholdsdelingen, der ikke nødvendigvis er udtryk for brugerpopulationens generelle holdninger til emner, men de ovenstående vurderinger, må samtidig ses i lyset af, at netop disse termer gennemgående fylder en stor del af den samlede debat. Det antages at der sker markante forandringer i holdningsdelingerne, både på grund af den blotte mængde af indhold, men det vurderes også, at der indtræder større variation i den sociodemografiske population på forummet r/unpopularopinion i observationsperioden. Det antages, at indholdet bærer præg af at være mødt af en lang række forskellige brugertyper, der forholder sig forskelligartet til varierende tematikker, der dominerer forummet. Massanari 2017 og Kilgo et al. 2018 funderer begge over den meget specifikke retorik, der er fremherskende på reddit som helhed, og det vurderes, at specielt de tidlige observationsepoker understøtter forholdet om *online inhibition*, eller en tilstand der retfærdiggør en undertrykkende eller hegemonisk adfærd blandt brugerne på r/unpopularopinion. Det direkte demokrati, der i udgangspunktet gør sig gældende på hele reddits platform, viser samtidig, at brugertilstrømningen til reddits forskellige fora, kan have stor indvirkning på indholdet, mængden af deling og engagement på de specifikke fora, ved hjælp af den deltagende kultur der præger det sociale liv på reddits forskellige fora.

Selvom reddit som platform har været såvel berømt som berygtet for dets brugerskabte indhold og lave grad af censur, argumenterer dette speciale for, at udviklingen af de dominerende termer, og hvordan brugerne gennemgående forholder sig til dem, ikke nødvendigvis er skadeligt for internettets

sociale liv. Aktivitsudviklingen på r/unpopularopinion er understøttende for, at brugerpopulationen oplever en forandring der til lige påvirker formen og udtrykket for forummet. Med udgangspunkt i Maffesolis koncept om samfundet, argumenteres der for, at online sociale fora, i dette tilfælde reddit's r/unpopularopinion, er en del af en større helhed i den sociale virkelighed, der udspiller sig på internettet. Selvom der opleves understøttende tendenser for *online inhibition*, misogyne og hegemoniske tendenser i de første epoker, anskues det at forummet i sin popularisering oplever en mere åben interessesammenslutning omkring den upopulære holdning som udtryksform. Forummet som *sted*, er stadig fastlåst i grundvilkåret om denne udtryksform, men på grund af den åbne struktur, anses det som tydeligt at forummet bliver et mødested for mange forskelligartede og foranderlige brugersammenslutninger igennem specialets observationsperiode.

Igennem den analyse-mæssige behandling af specialets empiriske grundlag, er der oplevet såvel muligheder som udfordringer i anvendelsen af, hvad der kan beskrives som *state-of-the-art* metoder til *topic modelling*, eller emneinddeling af store tekstmængder. Gennemgående er anvendelsen af LDA'en oplevet som en succes, der har bidraget til stor indsigt i fordelingen af emner på tværs af indholdsdelinger fra mange forskellige aktører. Den store masse af indhold, viser samtidig nogle af de udfordringer, de ypperste metoder på området på nuværende tidspunkt møder i praktisk anvendelse. På populære sociale mediesider, som eksemplificeret her gennem r/unpopularopinion, bliver der dagligt delt enorme mængder af data og information. Set fra et sociologisk perspektiv, udgør denne indholdsdeling en del af det sociale liv, og der kan argumenteres for, at der ligger stor værdi i at kunne behandle og belyse dette samspil af aktører. Hvordan man kvantificerer, eller funderer over denne informationsudveksling mellem mange forskelligartede aktører, er en udfordring, og mulighederne for at tilgå materialet analytisk, forandres efterhånden som samfundets teknologiske grundlag udvikles. I dette speciale, er der blevet formidlet en metodetilgang, der giver mulighed for at dykke ned i et komplekst og varieret indholdsgrundlag, uden at der ofres vigtig information i processen. I bearbejdningen af indholdsdelingen på r/unpopularopinion, er det oplevet, at *state-of-the-art* metoder til *topic modelling*, har problemer i tematiseringen af datagrundlag med store variationer. Oprindeligt, blev specialets problemstilling muligvis tilgået en smule naivt, i forventningen om at der med algoritmisk indgriben kunne skabes et fast og cementeret overblik over det givne indhold. Selvom det ikke lykkedes til perfektion, var anvendelsen af LDA givtig til at danne overblik, uden at kunne skabe ideelle emneinddelinger. Et af de tydeligste problemer med anvendelsen af *topic modelling* i dette speciale, har været at udrede de underliggende temaer, med udgangspunkt i bærende termer. Igennem det komplekse og dybt varierede indhold på forummet, er det blevet tydeligt, at emneinddeling er en metodologisk såvel som semantisk udfordring. Dette skal ses ud fra, at ords brug og anvendelse altid skal ses i deres kontekst, og deres betydning og interne associationer er afgjort af

relationelt indhold. Derfor er der ikke nødvendigvis tydelige konnotationer mellem primære termer, og bestemte emnemæssige afgrænsninger. Emneinddeling er yderligere besværliggjort af forholdet mellem semantisk og maskinel indholdsforståelse. Hvordan mennesker i praksis vil inddele forskelligartet indhold, er ikke nødvendigvis det samme som en maskinel vurdering. Som i eksemplet med *Black/White*, er det tydeligt at termerne spreder sig over en lang række forskellige temaer, selv inden for mere afgrænsede områder, f.eks. racedebat. Selv med et alvidende perspektiv over observationsperiodens empiri, ville en fuldstændig kvalitativ emneinddeling være problematiseret af semantiske vekselvirkninger og flertydigt indhold. Ud fra denne perspektivering, anskuer dette speciale derved, at forståelsen af det sociale liv, der udspiller sig på sociale mediesider som reddit, og herunder r/unpopularopinion, må ses i relation til denne fluktuation af indhold og semantisk konceptualisering. I det åbne internetlandskab der udgør reddit's brugerflade, mødes mange forskelligartede interesser og aktører, der i dette møde, påvirkes og forandres i en konstant udviklingsproces. Der argumenteres i dette speciale for, at enhver metodologisk og sociologisk behandling af reddit's sociale liv, må ses i denne indramning. I forlængelse af Maffesoli's antagelser, at den sociologiske forskning må forholde sig til at: *"surfing over the waves of sociality"*, skal de analytiske fund ses i lyset af den udvikling der sker gennem observationsperiodens epoker (Maffesoli 1996:5). I denne forstand har såvel LDA'en som de resterende anvendte metoder, bidraget til en gennemgående konceptualisering af hvad der sker i samspillet mellem de mange aktører, der udgør r/unpopularopinions aktive brugerpopulation. Specialets ligger, intentionen tro, op til en viderebearbejdning af de analytiske fund for den givne empiri. Specielt med henblik på diskursive fluktuationer, opleves de analytiske fund givtige for en eventuel viderebearbejdning. De primære termer, siger meget om indholdet, men trods diskussionsafsnittets behandling af *Woman/Men* og *Black/White*, er det stadig muligt at udarbejde en dybere konceptualisering af den sociale virkelighed. Som *proof-of-concept* anses det, at specialet giver mulighed for at forstå den gennemgående udvikling og det umiddelbare indhold, men at der kan opnås dybere indsigt ved hjælp af, f.eks. diskursanalyser eller andre variationer af indholdsanalyser.

8 LITTERATURLISTE

- Abdi, Hervé and Lynne J. Williams. 2010. "Principal Component Analysis." *Wire's Computational Statistics* 2(4):433–59.
- Adam. 2018. "Unpopular Opinions." *Know Your Meme*. Retrieved (<https://knowyourmeme.com/memes/unpopular-opinions>).
- Agresti, Alan and Barbara & Finlay. 2009. *Statistical Methods for the Social Sciences*. 4th ed. Pearson, Prentice Hall.
- Aladağ, Ahmet Emre, Serra Muderrisoglu, Naz Berfu Akbas, Oguzhan Zahmacioglu, and Haluk O. Bingol. 2018. "Detecting Suicidal Ideation on Forums: Proof-of-Concept Study." *JMIR Publications* 20(6).
- Alexa.com. 2019. "The Top 500 Sites on the Web." *@Alexa - an Amazon Company*. Retrieved (<https://www.alexa.com/topsites>).
- Amer, Karim and Jehane Noujaim. 2019. *The Great Hack*.
- Bagroy, Shrey, Ponnurangam Kumaraguru, and Munmun De Choudhury. 2017. "A Social Media Based Index of Mental Well-Being in College Campuses." Pp. 1634–46 in *Proc SIGCHI Conf Hum Factor Comput Syst*.
- Barlow, John Perry. 1996. "A Declaration of the Independence of Cyberspace." *Electronic Frontier Foundation*, February 8.
- Baumgartner, Jason. 2019. "Pushshift.io - Learn about Big Data and Social Media Ingest and Analysis." *Pushshift.io*. Retrieved (<https://pushshift.io/>).
- Bhatt, Nachiket. 2018. "How Does Reddit Work? Insights into Reddit Business Model and Revenue Model of Reddit." *Ncrypted Websites - Blog*.
- Blei, David M. 2012. "Probabilistic Topic Models - Surveying a Suite of Algorithms That Offer a Solution to Managing Document Archives." *Communications of the ACM* 55(4):77–84.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* (3):993–1022.
- Boe, Bryce. 2016. "PRAW: The Python Reddit API Wrapper." *Praw.Readthedocs.io*. Retrieved (<https://praw.readthedocs.io/en/latest/#>).

- Boe, Bryce, Joe RH, Levi Roth, and Randy Goodman. 2016. "PRAW - PRAW-Dev." *Github*. Retrieved (<https://github.com/praw-dev>).
- Brante, Thomas. 2003. "Consequences of Realism for Sociological Theory-Building." *Journal for the Theory of Social Behaviour* 31(2):167–95.
- Buch-Hansen, Hubert and Peter Nielsen. 2008. *Kritisk Realisme*. 3rd ed. Roskilde Universitetsforlag.
- Chang, Jonathan and Jordan Boyd-Graber. 2009. "Reading Tea Leaves: How Humans Interpret Topic Models." *Neural Information Processing Systems (NIPS)* 22.
- De Choudhury, Munmun, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. "Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media." Pp. 2098–2110 in *Proc SIGCHI Conf Hum Factor Comput Syst*.
- Chuang, Jason, Christopher D. Manning, and Jeffrey Heer. 2012. "Termite: Visualization Techniques for Assessing Textual Topic Models." *Stanford University Computer Science Department*.
- Coldewey, Devin. 2016. "Reddit Overhauls Upvote Algorithm to Thwart Cheaters and Show the Site's True Scale." *Techcrunch.Com*, December.
- Cole, Jennifer, Chris Watkins, and Dorothy Kleine. 2016. "Health Advice from Internet Discussion Forums: How Bad Is Dangerous?" *JMIR Publications* January 6t.
- Curran, James, Natalie Fenton, and Des Freedman. 2016. *Misunderstanding the Internet*. Routledge, Tailor & Francis Group.
- Danermark, Berth, Mats Ekström, Liselotte Jakobsen, and Jan C. Karlsson. 1997. *Att Förklara Samhället*. 1st ed. Lund Studentlitteratur.
- Datacamp.com. 2018a. "Exploratory Data Analysis in Python." *Datacamp.Com*. Retrieved (<https://www.datacamp.com/courses/exploratory-data-analysis-in-python>).
- Datacamp.com. 2018b. "Visualizing Time Series Data in Python." *Time Series Python Course*. Retrieved (<https://www.datacamp.com/courses/visualizing-time-series-data-in-python>).
- Ding, Chris and Xiaofeng He. 2004. "K-Means Clustering via Principal Component Analysis." *ICML '04 Proceedings of the Twenty-First International Conference on Machine Learning* 29-.
- Douglas, Nick. 2014. "It's Supposed to Look Like Shit: The Internet Ugly Aesthetic." *Journal of Visual Culture* 13(3):314–39.
- Downey, Allen. 2016. "What Is a Distribution." *Allen Downey Blog*.

- ELI5. 2015. "Who Owns Reddit / Who Pays for It / Are There Shareholders Who Control Its Posts?" in *r/Explainlikeimfive*.
- Eschler, Jordan and Amanda Menking. 2018. "'No Prejudice Here': Examining Social Identity Work in Starter Pack Memes." *Social Media + Society*.
- Garfield, Stan. 2016. "90-9-1 Rule of Thumb: Fact or Fiction?" *Medium*.
- Gazarov, Petr. 2016. "What Is an API? In English, Please." *Freecodecamp*.
- Glenski, Maria, Corey Pennycuff, and Tim Weninger. 2017. "Consumers and Curators: Browsing and Voting Patterns on Reddit." *IEEE Transactions on Computational Social Systems* 40(4).
- Grenfell, Michael. 2012. *Pierre Bourdieu: Key Concepts*. 2nd ed. edited by M. Grenfell. Acumen.
- Grigonis, Hillary. 2018. "Reddit Scatters In-Feed Sponsored Posts inside Its Mobile Apps." *Digital Trends* March 15.
- Kilgo, Danielle K., Yee Man Margaret Ng, Martin J. Riedl, and Ivan Lacasa-Mas. 2018. "Reddit's Veil of Anonymity: Predictors of Engagement and Participation in Media Environments with Hostile Reputations." *Social Media + Society* 4(4):1–9.
- Klein, Colin, Peter Clutton, and Vince Polito. 2018. "Topic Modeling Reveals Distinct Interests within an Online Conspiracy Forum." *Frontiers in Psychology* 9(FEB):1–12.
- Koerber, Brian. 2014. "The TL;DR Guide to Reddit Lingo." *Mashable*.
- Kranzberg, Melvin. 1986. "Technology and History: 'Kranzbergs Laws.'" *Technology and Culture* 27(3):544–60.
- Leavitt, Alex. 2015. "'This Is a Throwaway Account': Temporary Technical Identities and Perceptions of Anonymity in a Massive Online Community." Pp. 317–27 in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Vancouver: ACM, New York.
- Literat, Ioana and Sarah van den Berg. 2019. "Buy Memes Low, Sell Memes High: Vernacular Criticism and Collective Negotiations of Value on Reddit's Meme-Economy." *Information, Communication & Society* 22(2):232–49.
- Lund, Peter C. 2017. "Alternative Behandlinger – Hvorfor Gør Raske Individer Brug Af Det?" Aalborg University.
- Ma, Edward. 2018. "2 Latent Methods for Dimension Reduction and Topic Modeling." *Towards Data*

Science.

Maffesoli, Michel. 1996. *The Time of the Tribes – The Decline of Individualism in Mass Society*. Sage Publications, London.

Marwick, Alice E. 2017. "Scandal or Sex Crime? Gendered Privacy and the Celebrity Nude Photo Leaks." *Ethics and Informations Technology* 19(3):177–91.

Marwick, Alice E. and Danah Boyd. 2011. "I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience." *New Media & Society* 13(1):114–33.

Massanari, Adrienne. 2017. ""#Gamergate and the Fappening: How Reddits Algorithm, Governance and Culture Support Toxic Technocultures." *New Media & Society* 19(3):329–46.

Massanari, Adrienne L. 2016. "Participatory Culture, Community, and Play." *Participatory Culture, Community, and Play* (November).

Mathiesen, Anders. 2005. "Bourdieu's Videnskabsteori - i Anledning Af: Viden Om Viden Og Refleksivitet – Forelæsninger På Collège de France 2000-2001. Hans Reitzels Forlag 2005 (190 Sider)." *Dansk Sociologi* 16(4):127–38.

Matias, Nathan J. 2018. "Going Dark: Social Factors in Collective Action Against Platform Operators in the Reddit Blackout." *Designing Quality in Social Media* chi4good,:1138–51.

McIntyre, Kristen, Benjamin Bengfort, and Jonathan Lacanlale. 2016. "Elbow Method." *Scikit Learn - Yellowbrick*. Retrieved (<https://www.scikit-yb.org/en/latest/api/cluster/elbow.html>).

Van Mieghem, Piet. 2011. "Human Psychology of Common Appraisal: The Reddit Score." *IEEE Transactions on Multimedia* 13(6):1404–6.

Mills, Richard A. 2018. "Pop-up Political Advocacy Communities on Reddit.Com: SandersForPresident and The Donald." *AI and Society* 33(1):39–54.

Mills, Richard and Adam Fish. 2015. "A Computational Study of How and Why Reddit.Com Was an Effective Platform in the Campaign Against SOPA." Pp. 229–41 in *Social Computing and Social Media*.

MindfireSolutions. 2017. "Python: 7 Important Reasons Why You Should Use Python." *Medium.Com*.

Moviequotes.com. 2019. "'The Boondock Saints' Quotes (1999)." *Moviequotes*. Retrieved August 19, 1BC (<https://www.moviequotes.com/s-movie/the-boondock-saints/>).

Neal, Diane R. 2012. "Social Media for Academics - A Practical Guide." *Chandos Publishing Social*

Media Series.

Nembhard, Candice. 2018. "'Unpopular Opinion' Threads Are Taking Over Twitter." *High Snobiety*.

Nielsen, Jakob. 2006. "The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities." *NN/G*.

Page, Vanessa. 2018. "How Reddit Makes Money." *Investopedia*. Retrieved (<https://www.investopedia.com/articles/investing/093015/how-reddit-makes-money.asp>).

Panek, Elliot, Connor Hollenbach, Jinjie Yang, and Tyler Rhodes. 2018. "The Effects of Group Size and Time on the Formation of Online Communities: Evidence From Reddit." *Social Media + Society*.

Park, Albert and Mike Conway. 2018. "Harnessing Reddit to Understand the Written – Communication Challenges Experienced by Individuals with Mental Health Disorders: Analysis of Texts From Mental Health Communities". *Journal of Medical Internet Research* 20(4).

Pavalanathan, Umashanthi and Munmun De Choudhury. 2015. "Identity Management and Mental Health Discourse in Social Media." Pp. 315–21 in *roc Int World Wide Web Conf*.

Python. 2019. "What Is Python? Executive Summary." *Python Software Foundation*. Retrieved (<https://www.python.org/doc/essays/blurb/>).

R/dataisbeautiful. 2019. "Let's Hear It for the Lurkers! The Vast Majority of Reddit Users Don't Post or Comment. [OC]." *Reddit.Com/r/Dataisbeautiful*. Retrieved July 30, 2019 (https://www.reddit.com/r/dataisbeautiful/comments/b5f9wi/lets_hear_it_for_the_lurkers_the_vast_majority_of/).

r/TheoryOfReddit. 2018. "How Do/Will Reddit Make Money." in *r/TheoryOfReddit*.

R/unpopularopinion. 2019a. "R/Unpopularopinion/Rules." *Reddit.Com*. Retrieved (<https://www.reddit.com/r/unpopularopinion/>).

R/unpopularopinion. 2019b. "R/Unpopularopinion." *Reddit.Com*. Retrieved (<https://www.reddit.com/r/unpopularopinion/>).

Reddit.com/premium. 2019. "Reddit Premium." *Reddit.Com/Premium*. Retrieved (<https://www.reddit.com/premium>).

Reddit.com/pushshift. 2019. "R/Pushshift." *Reddit.Com/r/Pushshift*. Retrieved (https://www.reddit.com/r/pushshift/comments/bcxguf/new_to_pushshift_read_this_faq/).

Reddithelp.com. 2019. "Reddit Advertising Policy." *Reddithelp.Com*. Retrieved

- (<https://www.reddithelp.com/en/categories/advertising/policy-guidelines/reddit-advertising-policy>).
- Redditinc. 2018. "Reddit User Agreement." *Redditinc.Com/Policies*. Retrieved (<https://www.redditinc.com/policies>).
- RedditOriginals. 2013. "Reddit Myth Busters." *Redditblog.Com*. Retrieved (<https://redditblog.com/2013/08/06/reddit-myth-busters/>).
- Rehurek, Radim. 2019. "Model.Ldamodel - Latent Dirichlet Allocation." *Gensim*. Retrieved (<https://radimrehurek.com/gensim/models/ldamodel.html>).
- Richter, Felix. 2017. "The Explosive Growth of Reddit's Community." *Statista*, November 17.
- Roshan, Vishnu. 2019. "How Does Reddit Make Money? Reddit Business Model." *Feedough*. Retrieved (<https://www.feedough.com/reddit-make-money-reddit-business-model/>).
- Salihefendic, Amir. 2015. "How Reddit Ranking Algorithms Work." *Medium.Com*.
- Sattelberg, William. 2019. "The Demographics of Reddit: Who Uses the Site?" *Techjunkie - Your Personal Geek*.
- Sharma, Ratika, Britta Wigginton, Carla Meurk, Pauline Ford, and Coral E. Gartner. 2016. "Motivations and Limitations Associated with Vaping among People with Mental Illness: A Qualitative Analysis of Reddit Discussions." *Environmental Research and Public Health*.
- Sievert, Carson and Kenneth E. Shirley. 2014. "LDavis: A Method for Visualizing and Interpreting Topics." *Proceedings of the Workshop on Interactive Language Learning , Visualizations and Interfaces* 63–70.
- Smith, Craig. 2019. "80 Amazing Reddit Statistics and Facts (2019) | By the Numbers." *DMR Business Statistics*.
- Subredditstats.com/unpopularopinion. 2019. "Subredditstats.Com/r/Unpopularopinion." Retrieved (<https://subredditstats.com/r/unpopularopinion>).
- Subredditstats.com. 2019. "Subredditstats/Unpopularopinion." Retrieved (<https://subredditstats.com/r/unpopularopinion>).
- Tiggemann, Marika and Mia Zaccardo. 2015. "'Exercise to Be Fit, Not Skinny': The Effect of Fitspiration Imagery on Women's Body Image." *BODY IMAGE* 15:61–67.
- Twenge, Jean M., Thomas E. Joiner, Megan L. Rogers, and Gabrielle N. Martin. 2018. "Increases in

Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time.” *Clinical Psychological Science* 6(1):3–17.

Vincent, Thomas. 2017. “Visualizing Time Series Data in Python.” *Datacamp.Org* - via Thomas Vincent, Head of Data Science at Getty Images. Retrieved (<https://www.datacamp.com/instructors/tlfrvincent>).

Wardle, Josh. 2015. “API.” *Reddit Archive - Reddit Wiki / API*. Retrieved (<https://github.com/reddit-archive/reddit/wiki/API>).

Waskom, Michael. 2012. “Seaborn: Statistical Data Visualization.” *Seaborn*. Retrieved (<https://seaborn.pydata.org/>).

9 BILAGSOVERSIGT

Bilag 1: VosViewer – Oversigt over anvendt forskningsgrundlag

Bilag 2: Python Notebook: *PRAW* metodologi, og indledende dataindsamling.

Bilag 3: Python Notebook: Inddragelse af *Pushshift* og endeligt datagrundlag.

Bilag 4-7: Python Notebook: Univariat analyse og Tidsserie

- **Bilag 4:** 2016
- **Bilag 5:** 2017
- **Bilag 6:** 2018
- **Bilag 7:** 2019

Bilag 8: Overblik over de mest prominente holdninger gennem observationsperiodens epoker.

Bilag 9: Python Notebook: *Topic Modelling*/Emneanalyse

10 SLUTNOTER

1.1 ⁱ REDDITSKE BEGREBER

Subreddit: Subreddit er reddit's betegnelse for de forskellige fora og diskussionsrum, som udgør platformen. For at holde den enorme mængde af indholdsdeling emnerelevant og organiseret, kan brugerne dele deres interesser og synspunkter i emnespecifikke, mere eller mindre åbne, undergrupper. Subreddits styres og modereres af den eller de aktive brugere der har oprettet kanalen, hvor der er mulighed for at lave bestemte regler og retningslinjer for den givne subreddit (Koerber 2014).

Posting: Posting er en paraplybetegnelse for den aktive brugers interaktionsmuligheder med onlinefællesskabet. Her er det muligt at dele links, billeder, diskussioner, emnetråde etc. Brugeren poster sit indhold i en given subreddit, som andre brugere kan interagere med, enten med deres egne posts og kommentarer eller ved at vote (ibid.).

Votes: Votes er en interaktionsmulighed for reddit's brugere. Ved enten at upvote eller downvote, kan brugere tilkendegive sin mening om andre brugeres posts og kommentarer. Votesystemet har indflydelse både på de enkelte posts position på platformen, og på de enkelte brugeres såkaldte karmapoint. En bruger der har fået mange upvotes på tværs af sine posts, vil have mange karmapoints og dermed fremstå mere positivt, hvorimod en bruger der har fået mange downvotes, vil fremstå mindre heldigt (ibid.).

Subscribers: Subreddits vokser sig større og når ud til flere, ved at brugere kan tilskrive, eller subscribe sig til det specifikke forum. Når en bruger subscriber sig til en bestemt subreddit, påvirker det dennes holistiske oplevelse af platformen, da det emnespecifikke indhold får en større plads (ibid.).