AALBORG UNIVERSITY
DENMARK

Master Thesis

THE VALUE OF METACRITIC AND ITS RELATIONSHIP WITH VIDEO GAME SALES

FLAVIO TONA SHNEIDER

MIKE-E

January, 2020

Contents

ABSTRACT

The Electronic Entertainment market has been widely discussed in the media generating many discussions and missinformations. One of the controversies comes in the form of using MetaScore, provided by MetaCritic, as a comparison tool between games. Within this study we are going to dwelve into the value of Metacritic together with the correlation between software sales and MetaScore, just like the year of launch and the original platform of the games. Using descriptive and statistical analysis, the results concluded that MetaScore has relationships with game sales, showing its importance for the industry and value for the consumer and industry.

KeyWords: Metacritic, Games, Sales

1 INTRODUCTION

Video games have increasingly become part of the daily lives of the world's population, whether through home consoles, laptops or apps aimed at ios, android, and other mobile mediators. In addition, there has been increasing attention within the global marketplace, increasing media news coverage, new laws and taxes especially tailored for this industry. Recently prestigious championships games like eSport (Electronic Sport), with prizes of up to six millions of dollars for the team placed first.

For a degree of comparison between industries, cited by Sieberg (2011) the US music industry generated revenue of $ 6.9b, the movie industry generated $ 10.6b, while the video game industry generated revenue of $ 23b in 2011, even above the sum of its competitors in the entertainment market, this shows the weight of this industry.

It is a common practice for consumers before making the decision to buy a product to first get the opinion of others, it can be from close relatives and friends to complete strangers online. It is a practice worth of noting especially on websites like Amazon where consumers from around the globe post their experiences and grade (on a 5 stars system) the product and delivery service.

Inside the Video Game industry there is also this practice but not only done by the general public but especially by professional reviewers that has a history of doing it for years. Those reviewers are known by the community and the industry and commonly receive the products before the launch for the general public making it possible for them to produce the grade before the games reach the shelves.

Since there are many different reviweres utilizing different grading systems, ranging from a 4-Star grade system to 0-100, make it harder to directly compare their opinions. Recently one website called MetaCritics decided to compile those different grades and translated those grading systems into a single one, easy to understand, called MetaScore.

This system has been used by not only the general public but even on financial and Investor Relation reports of big companies like Nintendo as a marketing strategy to show the quality of

the games using MetaScore as a parameter. With this in mind we reach the scope of this study: What is the value of MetaCritic? To understand and reach an answer we first need to see how the MetaScore is aggregated, how is it viewed by the general public and how is it viewed by the industry and companies, and by accomplishing this it can become another step in understanding consumer behavior within this market.

The focus of this study is the value of Metcritic, in especial its understanding by the general public as a parameters of quality, meaning that the higher the number the higher the quality of the product. We are going to see how the industry and media have been using this number and the importance they give to it.

One important aspect of the relationship for the companies is to see if this quality instrument correlates with sales, if a higher grade is followed by higher sales thus increasing the value of this metric. To test this we will focus mainly on checking for correlations between MetaScore and sales in this market. Sales can be separated into Software and Hardware. Within this study we will not delve into more specific and recent sectors, such as *Downloable Contents* (DLCs), which today are a constant part of sales, and peripheral utensils such as games in commemorative editions, premium controllers and so on due to the lack of information provided by companies and resellers.

A database of sales of 15,289 games over the last 20 years has been gathered, as well as its release year, the console on which it was distributed and its genre. The data was collected in the VGchartz system, a website that compiles sales numbers publicly provided by different companies, in deep explanation can be found later in the Database section. Some modifications and censorships were necessary to be made in the database so that it could be studied in a more appropriate way.

Sales data were cross-referenced with the variable MetaScore, our possible instrument for quality, which is a specifically weighted average of ratings given by professional critics. Within the specialized media there is also a discussion about the functionality, damages and benefits that this average can bring.

Total hardware sales data also were collected in a website called VGChartz. The data was organized monthly from January 2004 until December 2014. Subject to the trends of the market, especially the behavior near the end of the year. On average, within the 10 years studied, November along with December accounted for 42.8% of annual sales, showing strong seasonality at the end of the year, often outweighing other market effects. In addition to seasonality, a growing trend can be observed, peaking in 2008 and then declining.

In the software segment, after collection, the data were organized in a way to make it easier to read and modify. The study was done in two parts, the first with construction and descriptive analysis of the database and graphs, crossing software sales with release year and MetaScore, thus being able to observe movement trends, outliers, observation density and boundaries. For example, only games with MetaScore over 70 achieved sales of over 1.6m (Global) on Playstation3, which includes 110 games.

As for software, the following platforms were initially studied: Playstation1 (PS1), Playstation2 (PS2), Playstation3 (PS3), PlayStationPortable (PSP), Gamecube (GC), Wii , Xbox, Xbox 360 (X360),
Nintendo DS (NDS), GameBoy Advance (GBA), WiiU , XboxOne (XOne), PlayStation4 (PS4) , Nintendo 3DS (3DS), Playstation Vita (PSV). Later the reasons will be explained in the Methodology section, but the most recent consoles were removed from the sample (PS4, Xone, 3DS, PSV, WiiU).

The statistical study was performed by calculating Pearson's correlation and equations calculated by Least Squares. OLS system was used to verify whether there is a correlation between global software sales and their designated MetaScore. The reading was made observing the result of t-student, within a margin of 10%, 5% and 1%, in addition to the coefficient, its sign and intuitions that link the result with market.

By understanding the value of Metacritic different segments can benefit. If it can be considered as a reliable source of a quality parameter so it becomes easier for the general public to make a rational choice before buying certain products, making it a reliable metric to look for before

choosing. For the industry it can be of great importance to know not only that the metric can represents quality but also its possible correlation with final sales, making it an important step to the possibility of predicting sales.

2 LITERATURE REVIEW

The main issue with this definition is that value changes from individual to individual, and also changes over time and culture making this a very fluid concept. Inside business the focus usually revolves around how to test it, how to transfer the value of your company to your product avoiding misalignment and especially inside consumer behavior studies, trying to understand what the consumer desires at given time and presenting him with a product that fits his needs.

When studying the value of Metacritic we need to keep in mind that value fluctuates, so the value for the industry could be of different nature from the value for the consumer, but in some cases the value can overlap.

For instance, if the consumer believes that this number represents quality then the companies will be willing to use this number as a parameter not only because they believe it's a good measurement but also because the consumers believe, and the more the large companies publicly uses MetaScore as a parameter the more credibility it gives to the measurement, increasing the reliability for the consumers, creating a sort of feedback loop.

For the companies one way that some are using this metric is to evaluate developers before hiring, seeing the games they already produced and the Score they achieved on each one, thus being another way to quantify and simplify the qualities of a producer. When used poorly this can create problems since he final value for a game, or its quality, do not solely depends on the developer alone, it is obviously the sum of different counterparts so evaluating a developer based on the quality of his final products could create a number of bias, over/undershooting the real participation of that developer in the product.

On another hand this metric could be of great value and pinpoint the quality of a developer in the case of very small and new companies. In those cases the developers rarely depend on a vast number of employees, and in many cases those small companies are compose of a single person.

They are usually called Independent Companies or Indie and the main characteristic is the lack of financial support from any large company.

One important aspect of this industry is that the price range is considerably small between products from large companies, even though the cost can vastly differ. In especial the games that represents the mascots of the companies, the flagship products, the companies expects them to sell a lot and the consumers expects them to be of high quality, those products are called Triple A or simply AAA.

Since the discussion is going to also revolves around sales we need to be cautious with some aspect, including price changes over time, over countries and even the hardware necessities for said product to work. The price of the product influence the decision of the buyers, and the average price changed over time due to inflation and other factors, so before we delve into the relationship between MetaScore and Sales we first will be addressing some of those aspects and adding disclaimers when necessary.

Within the electronic entertainment market, there are few previous specific studies on the relationship between quality, MetaScore and sales. Most of the information and research in the area is from writers in journalistic media and general news, such as Forbes and CBS, or websites specializing in this market such as Kotaku and IGN.

Looking at the studies and research, they do not appear to have consensus on the existence of a significant and observable relationship between MetaScore and Sales. Each research used a different database with different methodologies, we are going to delve in them later. Aside from researches it is also possible to find statements from certain individuals from the industry, usually during presentations or conferences, that concerns the use of MetaCritic by companies.

One of the most prestigious events of this industry is called DICE, (an acronym for *Design, Innovate, Communicate and Entertain*) and is held by a non-profit organization from California, the Academy of Interactive Arts & Science (AIAS). The aim of this summit is to present and award companies, products and individuals based on their career and impact they had in the industry.

During a presentation in 2008 the Activision Vice-President of Marketing Robin Kaminsky started with the following sentence: "*For every additional five points over an 80 percent average review score, sales may as much double.* ". This statement became famous within the market and made some researchers in the field motivated to check if this relationship really exists, or if it would just be another form of advertising.

Within motivated studies, we have Rob Savillo (2010). In his first study he used games released in May 2009, and checks their sales during the first 10 weeks after launch. The consoles included were NDS, Wii, PS2, PS3, PSP, X360. By looking at the database, setting the percentage of games based on their sales, and matching their MetaScore, no significant relationship was found between sales and MetaScore. The author later retraces the study by increasing his observations for games released between October 1 and December 31 (2009), including 218 games. As a conclusion he wrote "*My analysis demonstrates that in the link between Metacritic scores and sales exists"*.

Greenwood (2013) tests the relationship using 196 games, only from the PS3 and Xbox360 consoles, also dividing by genre (Action , Rpg and Fps). Through a descriptive analysis cited appearing a relationship, and testing the correlation between them found a positive and significant relationship between MetaScore and sales. Other authors, cited by Greenwood, found relationships indicating that value in Metacritic impacts sales (Murdoch 2010; Wingfield 2007; Everiss 2008). Greenwood also cites 3 authors who discredit the relationship between MetaScore and quality (Dodson 2006; Periera 2012; MacDonald 2012), showing that the market is still divided within these results.

Wingfield, co-creator of MetaCritic, believes that the system he helped to devise greatly facilitates the search for information. Everiss (2008) mentions that even large producers like Eletronic Arts use MetaCritic as a quality metric. Periera (2012) already mentions that MetaCritic may end up leading developers and publishers to rely on the weighted note to make their decisions, which may not necessarily bring good results to the quality of the final product.

Other authors have a more pessimistic view of the subject. Macdonald (2012) mentions that the existence of a relationship, and especially the perception and use of this relationship by publishers, can be bad within this market. Some producers may begin to direct their efforts to create games that seek a high score, but not necessarily to appeal to their target audience and produce innovations. The author quotes in his article: "*it's bad for developers, it's bad for critics, and ultimately that means it's bad for gamers too*".

The low amount of studies of this type of comparison, Quality x Sales, is understandable at some degree. Sales is a common number to have, most companies presents them, if not by obligation on Investor Relation, as a marketing strategy to show that a product is selling well. It is not an easy number to get with extreme precision since normally the sales of a product changes over time making it too dynamic, but depending on the market and how long it has been on shelves it is stable enough to round up to hundreds of millions copies.

Quality, on the other hand, is a very complicated measure to get since it revolves personal choices for variables and what to measure. Depending on the market and the price of the product you can find different approaches and different levels of depth.

When comparing different authors within the same study subject it is possible to see that there is no obviously consensus on the relationship between MetaScore and Sales, with some finding it positive while others finding it negative. The database utilized where not consistent over the studies, and that may be the reason why they had different results, not only the number of games but also how many consoles they covered.

Savillo based his research on the first 10 weeks after the release of certain games, 218 observations made the cut to the final database. He crossed MetaScore and Sales of those games and presented a visual analisis. Later it was concluded based on the top 5 games of each console of the time that no connection or trend was perceived between the variables, simply by observing the sales and score of the top products.

Greenwood took a different approach. The data wasn't selected based on time of launch, but instead based on the genre of the game. In his case he selected 196 games classified as First Person Shooter (FPS), Role Playing Game (RPG) and Action. Aside from this he also focused only on two consoles, XBOX360 and PS3. The choice was made based on the sense that those two were competitors at the time and considered by the author to have similarity of marketing.

To test the data Greenwood used two methods, fist was the visual identification of patterns based on graphics plotted from the data, and the second was a statistical measure known as Pearson Product-Moment Correlation Coefficient, or "Pearon's R". For both methods he found the correlation to be significant and positive, presented a correlation of .55 for the non-modified curve and .72 for a log transformed data set.

Comparing exclusively the number of observations this study presents a considerably higher database collection in comparison to others, within software it covers 4461 products, more than 20 times when compared with the previous author with the biggest database (Savillo). On consoles/hardware the other studies covered 5 at most, while this study covered 9 different ones. The other studies, mainly Greenwood and Savillo constrained the data set based on time, console or genre.

The difference in database size and how the data was selected could present a more robust result. The selection and censor in this study was firstly based on all possible games that do not sell anymore, making this less time-dynamic and dependent, the censors applied are more discussed on a different chapter.

One of the scopes of this paper would be to shed more light on the relationship between MetaScore and global software sales, another way would be to make room for further studies. To begin the study itself, we first need to understand how the game market works.

3 MARKET ANALYSIS

The analysis will provide a brief explanation of how MetaCritic works, as it is a fundamental part of the analysis. We'll look at how it is viewed by specialized media, how it has been used by the industry, and whether it is appropriate to use the unprofessional user-based average known as UserScore. As it will be an important part, and not necessarily publicly known and intuitive, follows a brief explanation of the functionality, reading, construction, criticism and defenses of the system created by MetaCritic. In addition, a brief explanation will be given about features of the Hardware and Software Market.

As for hardware, it will be explained which systems were studied, how they behave, when they are concurrent and when they are complementary as products, what happens when an update of an earlier version of the system is created (transition from PS2 to PS3, for example) and other situations.

It is quite unique for this segment to have a professional metric for quality, since such thing does not exist in this scale for other products like coffee machines or clothes. In most cases, especially when you observe the review system of Amazon, it is done by the general public. The biggest problem with this is consistency, some public reviews can be tainted by emotions and are willing to give very high or very low passionate grades, and no one is actually checking the quality of the reviews to refrain the final compiled value to demonstrate a fully trusted number.

It is hard to quantify quality, and even harder to translate different values into one singular number. Each reviewer possesses its own pattern, its own value and unique style. This is not different from what schools face when they need to grade students and create a grading system, the quantification of quality values.

As for Software, it will be explained which games were used, which consoles, the possibilities of different editions of the same game, existence of combo sales and other situations.

3.1 METASCORE

MetaCritic is a website that aims to make weighted averages for games, movies, tv series and music. In this study the focus will be exclusively on games. To understand the use of MetaScore, we first need to understand some aspects of the system itself.

There are currently numerous discussions about the benefits and harms of using MetaScore as a tool for quality. First we would like to explain how the value is compiled. Numerous websites produce game *reviews*, some with videos, some with text, and some with both. Critics receive the games, or demos, a few days or weeks before the official sale, and produce a review on the subject, weighing qualities and defects, to a note based on numbers or letters.

To better organize the system, and satisfy the consumer with a faster search speed, a website gathered the numerous websites that produce *reviews* and created a weighted average, generating just one number for the consumer, easy to understand. The exact calculation and its weighting is not disclosed to the public, apparently to protect developers and *reviewers* , but the scope of this study does not take into account the calculation to arrive at MetaScore , but rather the importance that this number reflects on sales. The site also specifies all the notes of each *reviewer*, with direct links to reading and video of them, making it clear who are the notes that feed the final average.

In addition to professional *reviewers*, MetaCritic has a space for UserScore, based solely on user ratings and reviews, feeding another average, ultimately achieving two results: MetaScore and UserScore.

Not all sites have the same rating metric, some range from 0 to 100, some use 0 to 5 stars, so for MetaCritic to produce only one result, it needs to normalize the ratings. An example of how it translates notes would be Table 1, shown later in the Note Conversion section.

MetaCritic is not only used by buyers, to demonstrate the importance and how it has been used, let's look at some cases. Recently, in his IR presented in March / 2015 Nintendo used MetaScore for comparisons with other companies, talking about the importance of having

a Metacore above 85: " *For Nintendo's current platforms , Nintendo 3DS and Wii U, 19 titles met these criteria . In contrast, the current platforms of other companies , PS4, Xbox One and PS Vita , totaled eight titles* . Nintendo's attitude was later criticized by sites like Kotaku and Forbes. Other companies, such as Irrational Games, demanded that developer *applications* must show MetaScore of the games they produced.

Obsidian , a game developer, made a *royalty deal* with his employees in the development of Fallout New Vegas (FNV), where they would get a bonus if their grade on the site was over 85, but they only got 84. The case became famous and prompted There is much discussion as to whether MetaScore should be used as a quality tool , and the harm it could bring to developer-producer agreements.

After the incident with FNV, MetaCritic received heavy criticism from sites like IGN: " *What is wrong is the way that Metacritic averages are used by the gaming industry to determine how games are made and sold, and the negative effect that they are having on. criticism* "written by MacDonnald (2012) . Some companies, such as Funcom, blame MetaCritic for falling *share prices* , and reaffirm the importance of a consistent *review* : " *A game like The Secret World, which is not based on a well-known brand, is normally dependent on positive. press reviews to achieve successful initial sales, in addition - but not limited - to other factors like word of mouth* . "

Other authors advocate the use of MetaCritic, and attack some of the sites that went against the system: " *Game reviewers don't like Metacritic because it takes their work and aggregates it on a separate site, thus greatly reducing unique visitor numbers to their site* ". This reduction in the number of views on the annotation sites, and the consequent increase in views in MetaCritic, has not yet been officially tested.

Take-Two CEO spoke at a conference: "In fact, if your ratings are below a certain level, they could seriously impair your ability to sell a security, and above a certain level you could make a big difference in your success.".

Regardless of how you view it, positively or negatively, this value is being used as a parameter by both the supply pool and the demand itself. The weighting and method used to calculate the Score is not relevant to the scope of this study, since the final result has been the only point observed by both parties, and even without knowing the weights, the goal would be to understand the relationship of this number and the final sales.

The system has some problems. UserScore is entirely based on *reviews* of users., Even though a few users try to be objective and meticulous, most shown passion with notes 10, maximum, and 0, minimal, deliberate games that simply " not loved" in contrast to professional MetaScore , where sites bring together a wealth of experience, argument, and consideration to come up with a result. As such, UserScore can be very polarized. To illustrate one of UserScore problems, we will use a famous case in which two developers Telltale created fake accounts and made *reviews* of own games, adulterating the UserScore positively.

Another problem is that both Scores are dynamic. If all grades that MetaCritic considers important for the final average have not yet been released , MetaScore may fluctuate strongly with each entrant. This will be better discussed later.

A more recent problem would be that with the age of the console-tight internet, some games also get updates, have DLCs, fix bugs, better balance and harmonize the game even after release. Some *reviewers* adapt their rating, and MetaCritic generally updates the change, but they are rarely radical changes. Regarding expansions and DLCs, MetaCritic encourages *reviewers to* give different grades for the game and DLCs (if they are expansions, not just details like Skins, customizations, etc.), thus differentiating the notes from those who buy only the game. , and who expands using DLC.

Another point would be that the aggregate note uses websites of different nationalities. Some games are released at different times in different countries, other times have censorship in countries like Germany and Australia, so as much as the game has been released in countless countries, it may have slight variations. Another factor would be that different cultures have different tastes, and that Oriental *reviewers* would rate based on their target audience. The same

would happen in Europe and America. In the words of the creators of the MetaCritic system themselves: "*We believe that multiple opinions are better than one, user voices can be as important as critics, and opinions must be scored to be easy to use. "*

When a user opens the website and tries to search for a product it is possible to immediately see two number, one is the MetaScore and the other is the UserScore. They both represent similar ideas, a compilation of reviews condensed into one number, but the origin of the review is what differs both values.

As explained before, MetaScore is based on professional and seasoned reviewers, usually well known by the public and already with a high degree of respect and consistency on their opinions. This is one of the strongest points of the MetaScore since it is not tainted by a vast number of anonymous sources that populates the UserScore. Usually Amazon and other websites they use UserScore, and for many reasons it is not suited to use as a strong metric, we are going to cover this on another segment.

For instance, inside the automobilist industry it is also possible to find professional reviewers who constantly grade and comment on products. One major example is the British magazine and tv show called Top Gear. They grade new products in a very professional manner constantly utilizing a system that makes it easy to compare cars from different ages and functions while respecting each category.

The problem consists in only having a few reviewers inside the automobilist industry so the consumer would end up having only a few professional opinions. The particularity of Metacritic is that it reaches a vast number of reviewers, but in order to compare and compile those it first needs to convert different systems into one that makes sense, is easy to read and doesn't lose much information from the original source.

3.1.1 GRADE CONVERSION

Not all sites use a number system. Some sites use a star system, ranging from zero to 4 stars, intermediating a half star. No stars converting directly to a zero rating, and 4 stars

converting directly to a maximum rating (100). One, two and three stars reflecting ¼, ½ and ¾ in whole numbers and the intermediates, by generating broken numbers, are rounded up, as shown in Table 1 below.

Table 1 - Example of Metascore adaptation

Star grade system	4	3.5	3	2.5	2	1.5	1	0.5	0
MetaScore	100	88	75	63	50	38	25	12	0

Source : MetaCritic System

Other sites use letters, similar to the American education system. In MetaCritic, they define A or A + as 100 and F as 0, the intermediate symbols (B, C-, D +, etc.) are their respective fractions. This can lead to problems, because a one letter within the school system represents a space notes, not a note in itself, and F is a note below the average 5, not necessarily a zero. So the sensitivity for site readers who have become accustomed to giving a score based on this alphabetic system may lose some of the sensitivity when seeing a B- being submitted as 67. The same is true for stars, as rating as 3 stars, for example, means that the game was in a spectrum between 2.5 stars and 3.5 stars, not necessarily a fixed integer, but a representative area.

Most sites use 0 to 100 or 0 to 10 with decimal numbers, which facilitates transformation. Sites that use 0 to 10 with integers, their representative are multiplied by 10.

Finally, when a weighted average is taken, in addition to the number known as MetaScore, it is given a color and meaning defined by the site depending on the field it is in, for better visual perception by the target audience. Between 75 and 100 receives a green background (90-100 universally acclaimed, 75-89 General Favorable *Reviews*), 50 to 74 Yellow Background (Mixed *Reviews*), 0 to 49 Red Background (20-49 generally unfavorable , 0-19 universal disgust). This system works for both MetaScore and UserScore . The system may vary when facing Movies, Series and Music.

3.2 HARDWARE MARKET

Hardwares are the systems that run the software, the games themselves. The study included Playstation1 (PS1), Playstation2 (PS2), Playstation3 (PS3), PlayStationPortable (PSP), Gamecube (GC), Wii, Xbox, Xbox 360 (X360), Nintendo DS (NDS), GameBoy Advance (GBA) , WiiU, XboxOne (XOne), PlayStation4 (PS4) and Nintendo 3DS (3DS). Part of these are Video Game Consoles, generally home appliances, connected to a television, electricity provided by external means such as sockets such as PS1, PS2, PS3, PS4, GC, Wii, WiiU, Xbox, X360, Xone. Others are HandHeld Game Consoles (portable, own screen, with internal battery or battery) like PSP, GBA, NDS, 3DS.

In general the market competes separately between table and handheld consoles, ideally the handheld is complementary to the table consoles, even games often try to integrate between the two such as Pokemon Stadium on N64 and Pokemon Blue / Red on Gameboy, GBA as secondary screen for GC Splinter Cell, etc. Among the notebooks studied, PsP and PSVita are from Sony and NDS, GBA and 3DS are from Nintendo.

Within the features of the consumer in this market, there are rare situations when buying the same console numerous times, there is no regular pattern of repeat purchase of the same system by the same consumer, for example, it is unusual for a residence to have more than one PS3.

When a newer version of a console is released, its previous version is not immediately withdrawn from the market, the previous version remains a while until software ceases to be released to it. For example, Metal Gear Solid V was released for both PS3 and PS4. It has become commonplace to release games for both systems (ps3 and ps4, for example), thus smoothing the adaptation to a new console, not losing consumers who only have the previous system.

Console life varies widely, there is still no standard number, some have been supported and games for over 8 years, and others have been out of production in less than 2 years. Also, there is not necessarily an obligation for backwards compatibility, for example, Snes games do not work on Nintendo64, however PS1 games work on PS2, Wii games work on WiiU. The presence of a

virtual store in the consoles gives the possibility of not having backward physical compatibility (cd, dvd, cartridges, etc.) but can keep old games can be bought and played on new consoles.
There may be variations of the same hardware. PS2 and PS2 Slim, Wii and Wii Mini, all variations play the same games, only exception would be 3DS and New 3DS, but it is debatable whether it is new hardware or just a variation of the old one.

In addition to running the games themselves, some consoles play movies (dvd and / or blueray), applications such as Netflix, and have internet access, allowing use of search sites, youtube, facebook and allowing you to navigate the global network.

3.3 SOFTWARE MARKET

Software is the games themselves. They do not work alone, needing a mediator to function. Games can have both physical and virtual sales. Physical selling games can be based on CD, DVD, Blueray, Cartridges and Tapes. Virtual Sales are Downloadable releases of the game directly to the console. Virtual games do not have a resale system after use, but some systems allow the purchase of a single use code that can be resold.

Some of the consoles have backwards compatibility, allowing anyone who owns only PS3 to buy PS2 games as well. In addition to the release date and source console, the games have genres such as: Adventure, Platform, Sport, etc., many games use features of various genres. Platform games with role playing features for example. Games can also cross platforms, example of this would be Fifa14 (X360, XOne, PS2, PS3, PS4, PSP, PSV, Wii, 3DS, etc.).

When the new Hardwares are released, they usually already have games to accelerate its initial sale. Usually companies launch consoles along with launching games of some brand already rooted in the market, and these games are known as *System Sellers*. These games are recognized for being almost mandatory for those who have the system, usually get very high penetration on that console, grades and sales.

There are also games that come with consoles, such as Wii Sports or Kinect Adventures. May be with the value of the software embedded or not in the price. Sometimes

these promotions are temporary, sometimes permanent. Apparently, when a company wants to increase its console sales, and she sees the opportunity to use a game released after the console's release , she creates these combos, and takes advantage of a potential System Seller. An example would be the WiiU, launched in 2012, being sold alongside Xenoblade Chronicles X at Christmas 2015 as a promotion.

Within this market there are awards given to the games, the most famous award being called Game Of The Year (GOTY), a prize given by some specialized websites for the best game released that year. Different sites may choose different games.
Some games have special editions, whether it's a company birthday celebration, or a collection already containing all of their DLCs and Season Passes. Sometimes when the game is named for GOTY gets a commemorative edition, examples we have Red Dead Redemption and Gears of War 2.

Recently we have differentiated systems with media purely virtual, as is the example of Steam, one of the largest of them. However we will not go into detail because it is not the current scope of this work. To get into the central point of the research, let's dig deeper into the methodology and studies that have been applied. First by understanding the value of this system to the general public and the industry, followed by a deeper study of the relationship between final sales and MetaCritic, in order to validate part of the value considered by the media and some companies.

4 - VALUE

4.1 CONSUMER

For the general consumer it is always interesting to see the quality of a product before buying, especially if the quality is not only demonstrated by the producer but also from other consumers. This practice can be seen in other platforms like Amazon or Ebay and it is of great value to acquire this previous knowledge. MetaCritics takes another step and presents not only the public’s opinion but also from professional unbiased reviewers, so one of the biggest values of this system comes from being a condensed quantification of qualities.

Another aspect is how easy it is to understand the final value presented, the metascore is a singular number that varies from 0 to 100. When the consumer looks for reviews by himself he can encounter a myriad of different metrics applied, from star system to letter, one particular quality of metacritic is its easiness to read.

4.2 - VALUE FOR THE INDUSTRY AND USAGE

The concept of public reviewers and even professional reviewrs isn't something new in this industry, two decades ago it was possible to find certain specialized magazines that used to cover new products, launches, and normally giving them grades based on their own stardards and systems, this doesn't differ from what we have today, but in the past those number were more escarse and harder to come by, ever worse to organize and compare, so it was harder to find companies actually using this as an argument or even internal feedback.

Nowadays not only it is easier to find those numbers and reviews, but also MetaCritic provides an easier compilation of those numbers.

The MetaCritic system in its nature is a compilation of reviews, and those reviews are nothing less than the quantification of qualities as it is the case of most grading systems. For the industry to have a quality measurement in this scale is something quite singular, especially when done by the general public containing the final buyer and user; and a third party (reviewers).

It's a great way to have feedback if properly used, some companies tend to have a hard time communicating with the user after the purchase, hence the whole segment of after-sales. To understand the problems and especially the good things of your product is something that some companies need a whole department to do, acquiring this information is already a hard task, but when it is naturally given to you the whole process becomes easier.

Since the MetaScore is a measurement of quality some developers use this system to present their portfolio, their curriculum, and even companies asking this as a requirement for a job. One famous case, and possibly the first one, came from 2012 when a company called Irrational

Games posted a job offer in which they asked for a few requirements: Six-plus years as a game designer; Four-plus years of experience managing direct reports; Shipped a minimum of three games from pre-production through ship; Credit on at least one game with an 85+ average Metacritic review score. Considering the scores observed in this study it is fair to say that 85 is a considerably high grade to achieve.

Another example of companies using metacritic as a metric for developers comes in the form of a payment bonus. In this case a developer, Chris Avellone, from a company called Obsidian publicly explained that in one job that was ordered from Bethesda to develop a game: Fallout New Vegas (FNV) they received a straight payment instead of the usual royalty applied for developers. In this case Bethesda promised that if the game reached a score of at least 85 on MetaCritic they will give a bonus payment to the developers, this shows another way that this system is being used by the industry. In this particular case the game didn't reach the target score, missing from just one point, reaching a max of 84 and the developers couldn't receive their bonus payment.

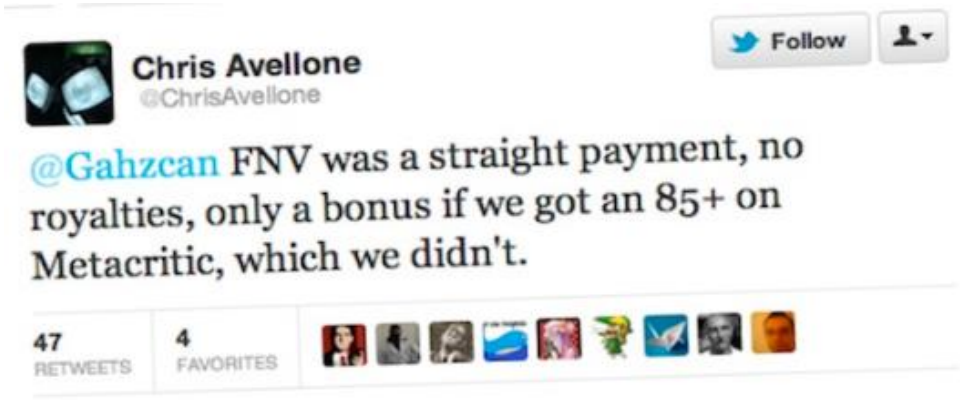


Figure 1 - Developer talking about a deal with a company involving a Metacritic value as quality target

One properly expected use of this system is to show the MetaScore reached by your products, which happened for small developers looking for a job, but also for large companies like Nintendo in their Financial Demonstration. Nintendo's Investor Relationship (IR) is normally

presented by the president of the company and on 2015 they used the Metascore of different games as an argument to show how good that year was.

Inside the IR it is possible to find a comparison between Nintendo and the other two big companies (Sony and Microsoft) how many games in that year they had with a metascore above 85, this shows the importance of the system to big companies.

"It shows a list of packaged game titles that received a Metascore of 85 or more and a User Score of 8.5+ or more on Metacritic.com, which we can admit received high scores both from the professional reviewers and consumers. For Nintendo's current platforms, Nintendo 3DS and Wii U, 19 titles met these criteria. In contrast, the current platforms of other companies, PS4,Xbox One and PS Vita, totaled eight titles."
Nintendo IR, 2015

Figure 2 - Usage of MetaScore as a comparison tool by Nintendo (IR 2015)

With all put together we have that Metacritics can be used as quality measument and a natural feedback provided not only by the public but also by third parties that do not necesseraly belong on the demand or the target audience. Aside from this, companies have already been using the system as a metric for developers, as a requirement for jobs and even for bonus payment. And more recently used by one of the biggest companies in this industry as a standard comparison between companies inside its own IR.

There is a possible feedback loop when you put together the industry and the consumers. This metric is important for the industry and is being used because it is important for the consumers, and the consumers can also take it seriously since the large companies are already taking it seriously too.

4 METHODOLOGY

In this topic we will go into more detail about how data were collected, sources used, legitimacy of sources, how they were organized, and how data was censored.

4.1 DATA COLLECTION

The MetaScore and UserScore notes were taken directly from the website that produces them, MetaCritic. The value of sales were taken from VGChartz, VGSales, WikiSales, GiantBomb, Corporate IR , Geimin.net (sales in Japan). The year the game was released, publisher, genre and platform were also taken directly from the VGChartz system.

Savillo and Greenwood made a choice to pick games based on certain criterias. Savillo used only games that were launching at the time and he had access to the sales figures of only the first 10 weeks. This is interesting from the idea that we are looking at the initial sales impact of a new product.

In the case of this study we are not going to follow this choice to observe only the first initial impact, since in doing so we would be ignoring the explosive seasonal sales that happen during the end of the year. We are going to look further into this on the next segment, but this is a choice that can drastically impact the final result.

Greewood took the data selection approach of picking only 3 genres (RPG, Action, FPS) and only on two consoles (XBOX360 and PS3). He didn't limit the data based on time, but based on marketing similarities since those two consoles were considered competitors at the time. In the case of this study we are not going to limit the dataset based on console or genre, since this would limit the number of observations by a harshly amount.

The database was built utilizing every genre and every possible console. The consideration here will be based on games that are not being sold anymore or games that the sales number has stagnated. In this sense we are not dealing with a dynamic notion of sales that are still possibly going to increase, we are dealing with the final sale of a given product and his final grade, solid and static values.

VGchartz numbers are dynamically corrected, each time companies post results on their IR, the site checks and adjusts the values. The older the numbers, the more likely they are to be corrected. For this reason, recently released games were disregarded, as they can still be corrected with each release of new IRs.

Nothing prevents old games sales from being corrected, but within this market the sales trend of a game is decreasing over time, generally having strong impacts only at the end of the year. Eventually sales shocks may occur, such as specific promotions, collections, collector editions, or DLC releases.

Within this study, the price of games was not taken into account because it is a considerably deeper analysis in another explanatory area, with another data collection system, inflation correction and relationship with different markets.

4.2 **HARDWARE - DATA**

The Hardware study was used to gain greater insight into how the market behaves if there is any trend. As we can see in Table 2 below, hardware sales were separated on a monthly basis, which is the sum of all major hardware offered on the market. Data collected begins in January 2004

and ends in December 2014. All hardware was used in the sum, including that produced in the current generation.

Table 2 - Total Global Hardware Sales (millions)

Ano	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Vendas	9.65	25.51	43.28	75.38	89.72	82.06	74.57	68.99	54.22	48.42	45.11

Source : Author

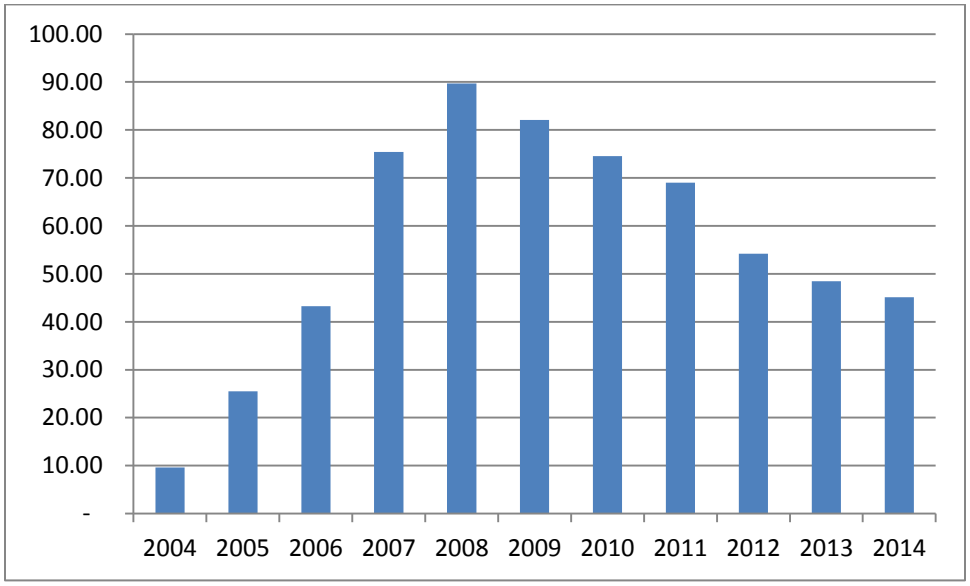
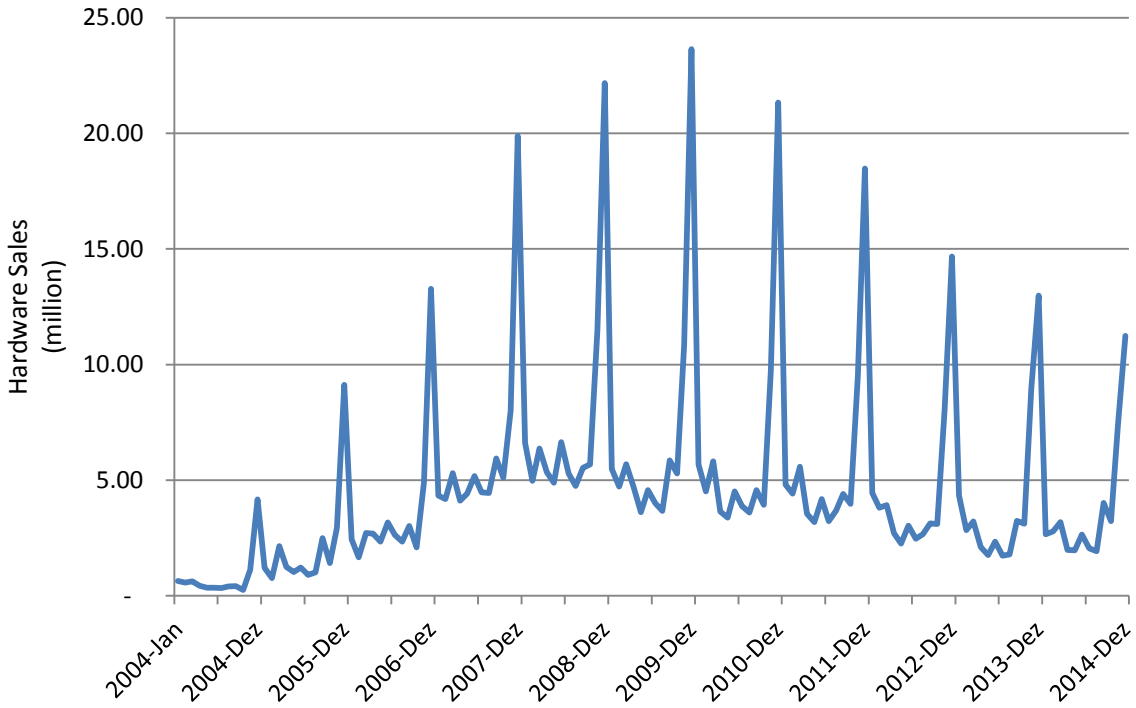


Figure 3 Distribution of yearly hardware sales (millions)

Sales were organized in such a way that a graph could be created (Figure 1) crossing the month / year with the total value of sales, thus observing the existence of any trend or seasonality. Visually you can already see a strong seasonality at the end of each year, probably due to the presence of religious holidays. In addition, there is an increase in sales, mainly in 2008 and 2009.

Figure 4 - Total Hardware Sales (monthly)



Source: Author

After 2009 begins a fall in annual sales, reaching in 2014 with approximately 45 million, representing an amount below 2009 , but above 2006, showing that in the long run there was a growth, but compared to the peak in 2008, we are in a fall. The study of the data was made based on the number of consoles sold, but in relation to the revenue, world industry does not show such a strong decay.

Table 3 - Annual revenue of the Video Game Industry (worldwide), corrected by dollar inflation

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Revenue	54.55	44.25	48.41	67.92	72.22	64.50	65.90	66.20	63.00	76.00

US dolar (million)

Source: VgSales

In addition to the revenue shown in Table 3, as a measure of comparison, in the United States in 2011, the Video Games industry surpassed sales of the movie ($ 10.6b) and music ($ 6.9b) industry with sales, a total of $ 23b,

An interesting feature to note would be that even with the crisis in 2008, it remained one of the highest revenue years in the period.

Another interesting feature to note is the peaks present at the end of each year. A distribution of the fraction in which each month represents each year showed that November and December have a relatively higher percentage of annual sales. If it were an equal distribution, we would expect 8.33% each month, but we see small peaks during the year, and a large percentage change at the end. As shown in Table 4 below:

Table 4 - Percentage Occupation that November and December represent annually

Year	**Nov**	**Ten**	**Nov + Dec**	**Rest of the Year**
2004	11.59%	43.17%	54.76%	45.24%
2005	11.50%	34.74%	47.25%	52.75%
2006	11.34%	30.67%	42.01%	57.99%
2007	10.62%	26.37%	37.00%	63.00%
2008	12.80%	24.71%	37.51%	62.49%
2009	13.17%	28.81%	41.97%	58.03%
2010	13.05%	28.60%	41.65%	58.35%
2011	13.80%	26.77%	40.57%	59.43%
2012	14.81%	27.06%	41.86%	58.14%
2013	18.53%	26.82%	45.35%	54.65%
2014	16.48%	24.91%	41.39%	58.61%

Source: prepared by the author

Combining November and December sales, we have averaged 42.85% of concentrated sales over the past 10 years. A considerable amount, and that could get a degree of attention, especially for those who want to choose the best time of year to launch a game or console, or even for inventory control.

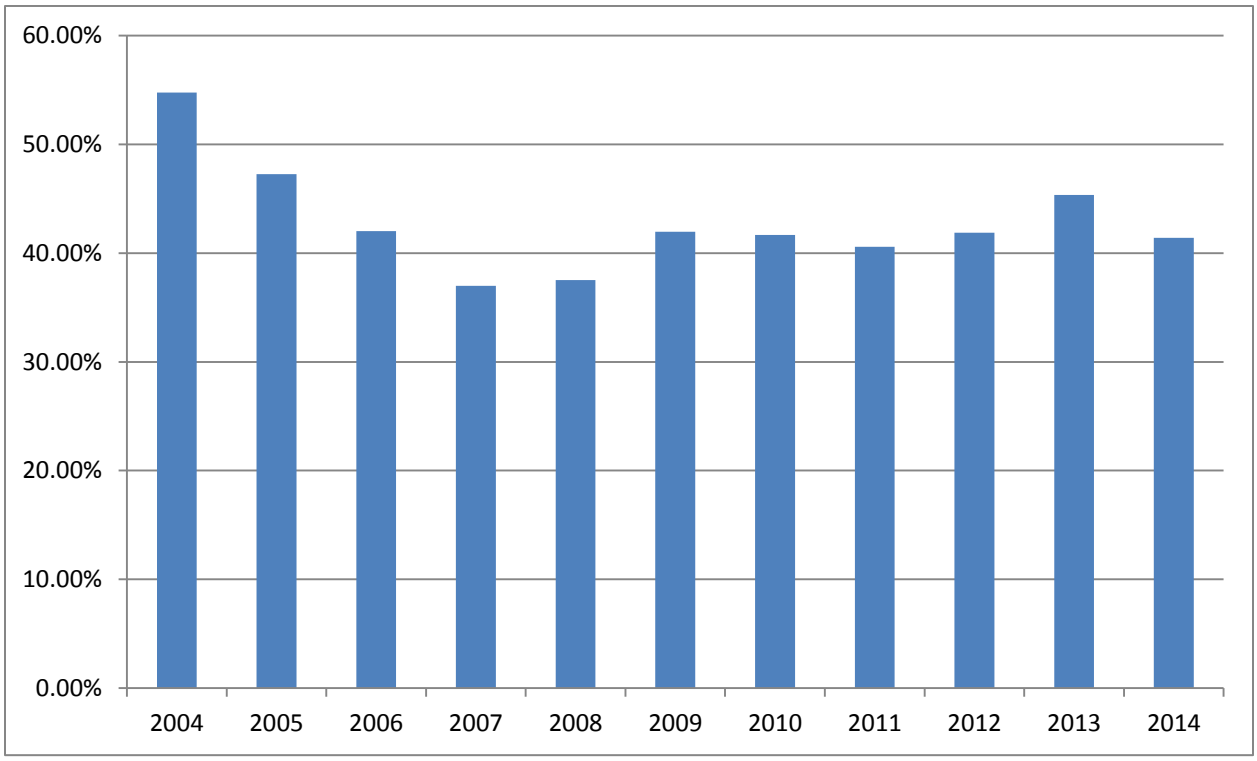


Figure 5 Percentage of the yearly sales represented by November plus December

If a product launches during the course of the year, its sales can be measured soon after launch, but as noted in the year-end impact, it would be best to wait at least one year after launch to ensure its presence in the market and your global sales.

4.3 SOFTWARE - DATA

Software sales were captured from different sources. Initially most were collected from the VGChartz system, later numerous data were compared with other sources, mainly corporate IR, WikiSales, Giantbomb, etc. If the numbers were very different between sources, the observed game would be removed from the database, if they were close values, the priority would be to

correct to resemble those shown in IR companies or their publications, but the difference occurred with low frequency, not significantly affecting the database.

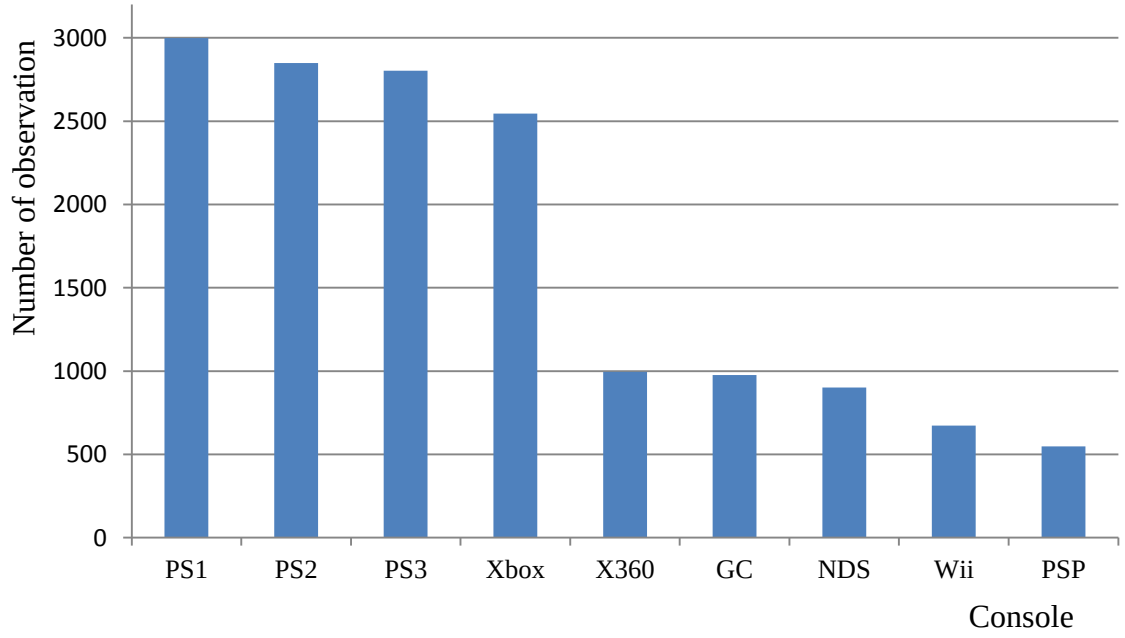
One of the initial approaches to this study would be to see if there is a correlation between MetaScore and global sales, mainly because it is a hotly debated subject in the specialized media (as seen earlier in the Literature Review section).

In addition to MetaScore, the site offers UserScore. As explained earlier, it comes from user opinions, and has no sample limit and can vary dynamically all the time. There are two problems with using this number. First, users are passionate, a quick note shows that high and low grades are extremely common, and often accompanied by innocuous explanations. Another problem happens when UserScore moves too far from MetaScore, which could mean that public opinion differs too much from professional opinion, regardless of the reason. When this difference proves to be significant, further investigation is appropriate.

The numbers are organized by year of release, the company that released the game, Console that was published, sales in North America, sales in Europe, sales in Japan, sales in the rest of the world (calculated using Global sales - (NA + EUR + JP)), Global Sales, MetaScore, UserScore, Meta-User Difference and game penetration on that console. To facilitate the bottom line, we will use only global sales value, MetaScore and the year of release.

As noted in Figure 2 below, in total data from 15,289 games were acquired (996 PS1, 2,848 PS2, 2,545 PS3, 549 PSP, 977 Xbox, 2,999 X360, 2,802 NDS, 901 Wii, 672 Gamecube). Within the observations, not necessarily all were used; the reasons will be clarified in the Data Censorship segment.

Figure 2 - DISSEMBUTION OF OBSERVATIONS



Source: prepared by the author

Within the database, we would have the possibility of errors, and the value of the sale is not absolutely accurate, the buyer has acquired the same game more than once, thus reaching twice the same console. But these would be rare cases within this market. There is also a strong resale market for used games in some countries, which could increase the penetration effect of consoles without directly changing the value of sales for companies, but this type of market

4.4 **CENSORS APPLIED**

All games with sales below 0.1m were removed, as a small nominal variation between sources would cause a large percentage variation in game sales, and the margin of error would be large due to the lack of precision for fine calculations. A game that was considered with sales of 0.4m by one source, and 0.6m by another, would have a variation of 1.5 times from first to second, but nominally would be only 0.2m. On the other hand, a 5.0m game on one source and 5.2m on another has a nominal difference of 0.2m, but the percentage variation is considerably smaller.

Games that have no official notes in Metacritic, Year of Release, or Release Console have been removed. As much as they passed the first 0.1m sales censorship, without MetaScore they could not be used. There are different reasons why the game does not have a rating on the site, it may be that few critics have *reviewed* , the site only computes the rating if it has at least four *reviwers* , or if an error has been found, such as some form of system abuse.

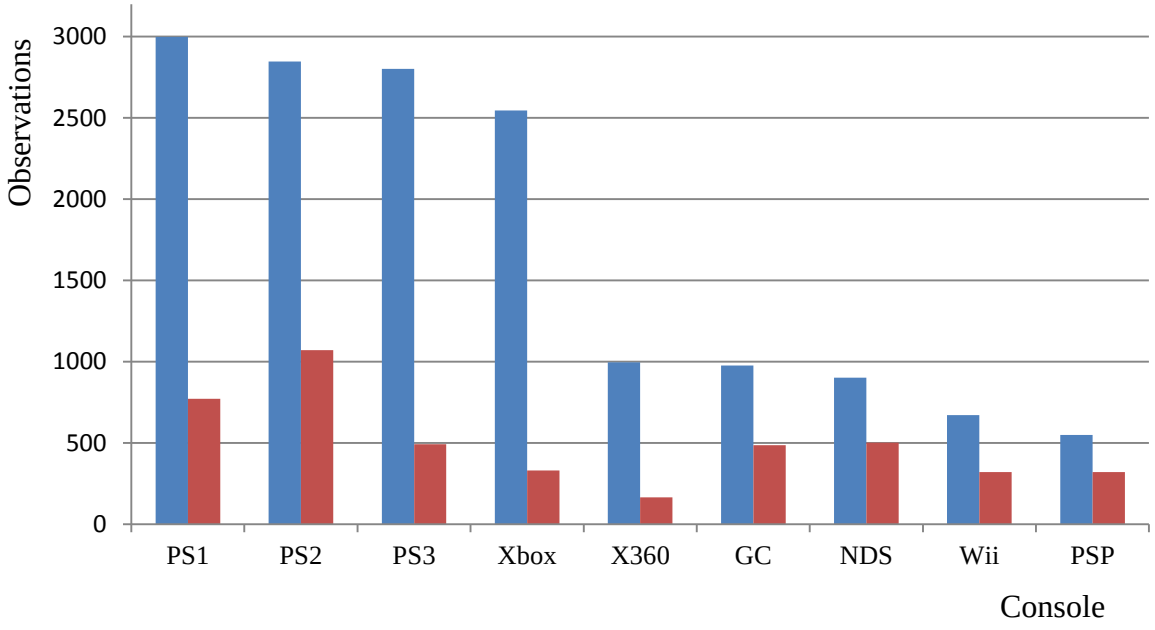
There have been instances where developers used UserScore's anonymity to positively tamper with the notes, but once identified, were removed from the sample. Other times there have been cases where the game was not previously distributed to certain *review* sites , leading to a temporarily biased score until the game was released and rated.

Recent console games (PS4, XOne, WiiU, 3DS, PS Vita) have not been used as their sales are still being made and their Metascore can still be changed. GBA was not used as many game sales figures varied widely across sources, an unusual situation.

In addition to sales figures being dynamic, as we can see in the Hardware-Data section, there are strong shocks at the end of the year. Even though a game was released in August, its sales can be strongly modified in the last two months of the year, making it impossible to use games released in 2015 from being used within the database.

The games that were considered as sales-married were removed. These are games that were initially sold on consoles (Wii Sports coming along with most Wiis sold globally) or along with devices for such a console, as an example of the Wii Party game sold along with Wii Mote, Wii controller, Kinect Adventures coming along with Kinect, handset sold for the X360. Generally observed as outliers, mainly for the very high penetration that they obtained in their consoles.
In Figure 3 below we can see how the database was modified by the censorship applied. Prior to censorship the database had 15,289 games (996 PS1, 2,848 PS2, 2,545 PS3, 549 PSP, 977 Xbox, 2,999 X360, 2,802 NDS, 901 Wii, 672 Gamecube), after the data correction, more than one third of the observations were dropped, leaving 4461 games (165 PS1, 1072 PS2, 331 PS3, 321 PSP, 486 Xbox, 772 X360, 492 NDS, 502 Wii, 320 Gamecube).

Figure 3 - BEFORE AND AFTER DATA CORRECTION (Blue Before, Red After)



From the moment in which the data was corrected, with all applied censure, one can begin to study both graphics and econometric, for verification of the corr elation between Software Sales and metascore, thus approaching the goal of the study.

5 STUDIES

This part has been separated into Descriptive Analysis and Econometric Analysis. Descriptive Analysis is based on reading charts and tables, which, along with market knowledge, can bring to light important results. Econometric Analysis would mainly use statistical and econometric programs and techniques to calculate the existence of correlations and their impacts on Software sales.

5.1 DESCRIPTIVE ANALYSIS

After organizing the data, the objective would be to create graphs to analyze the relationship between MetaScore and Sales. To get a better view of the relationship, the observations were separated into different source consoles, so it resulted in different graphs for each console, thus obtaining an individualized study. A study was also made with all the crowded games.

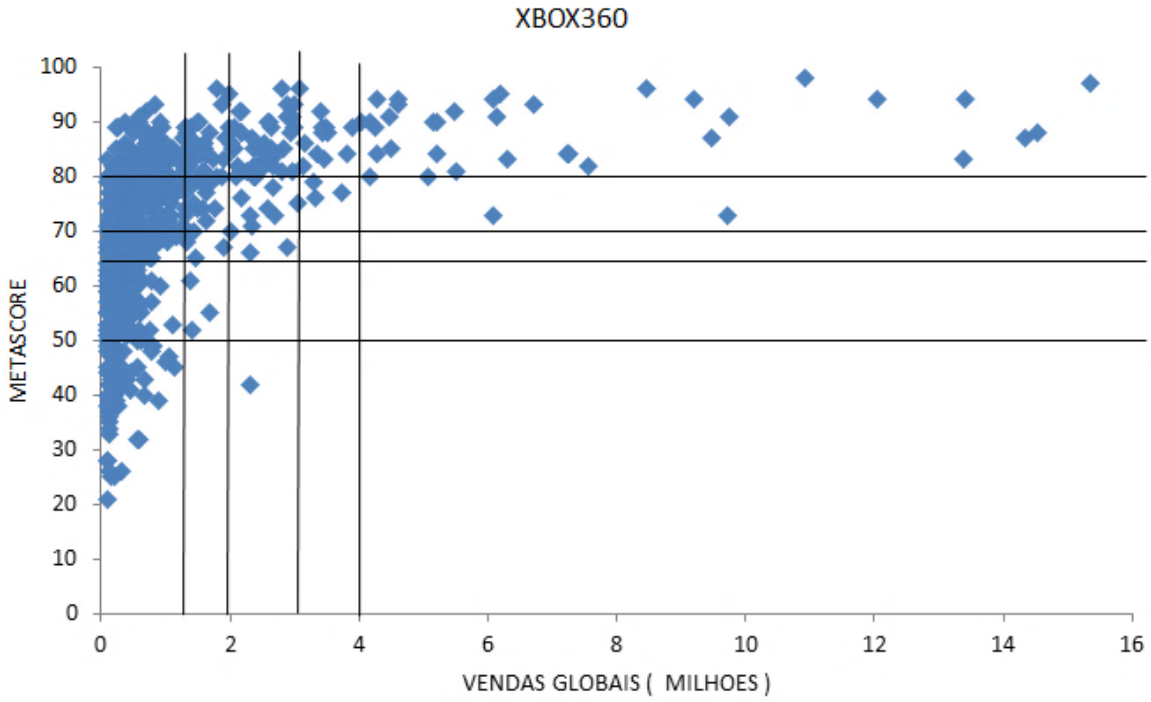
Separately, different game numbers were used for each console, so we have: 170 PS1, 1091 PS2, 337 PS3, 326 PSP, 501 Xbox, 792 X360, 498 NDS, 519 Wii, 326 GC. Charts have been assembled so that the X axis is sales, starting at 0 and in millions of units, and the Y axis is the notes in MetaCritic. The individual graphs and their respective analyzes can be found in Appendix B.

Visually the 9 graphs have repeated patterns between them with similar characteristics. In all there is a high observation density in the left corner, corresponding to games with low sales, regardless of the score they obtained and the year they were released. Showing that all games, regardless of the console that apply, have the possibility of low sale, which is true. Even something counterintuitive like a 90 note game can have sales close to 0.5m.

Another peculiarity that is observed was that the distribution tends to get around to the right (higher sales) as increases representative note MetaCritic, which is a first visual indicator of a relationship between them, showing that the higher the note received, the greater the chance of a bigger sale.

Looking at the sales only, one can see that there are many games that got low sales, and few games that got a high sale. The terms “low” and “high” vary from console to console, when viewed all together we can quantify, but individually we have very large variations at the top.
Another situation that can be observed analyzing the database is the proportions that sales grow, compared to the grades. Since it does not have a linear growth, the study needs a mathematical deepening, which will be seen later. What you can get from chart information is slight censorship of the note, which limits the value of the sale. In order not to overload reading, we will only give two examples, Xbox360 and the union of the entire database.

Figure 4 - Relationship between sales and metascore (XBOX360)



Source: prepared by the author

Within the sample, we have a game called Zumba Fitness, with score 42 and sales 2.31m, it will be considered an outlier for being too far from the standard of other observations.
Within Figure 4, the lines show the observed censorship intersections. From what we can see, within this sample, only games with grades above 50 managed to overcome the 1.5m sales barrier (above these sales, the base has 135 games); only games with a score above 65 managed to overcome the 2m sales barrier (with 107 games above this line); above 3m only notes above 70 (55 games); and above 4m only games rated above 80 (35 games), excluding two observations.

These limits that the notes appear to be can be understood as a form of growth barrier. High scoring games did not necessarily sell much, but only high scoring games could sell in high quantity. The other side occurs with low grade games, when less than 50, there was none that

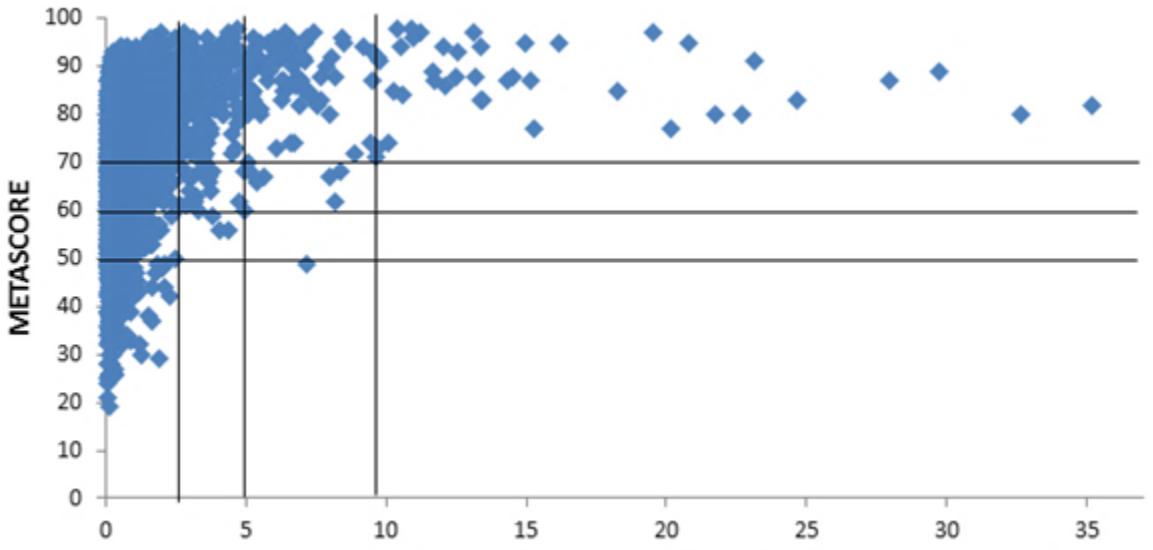
could exceed 1.5m of copies sold. One million unit sales may not be considered low, but in perspective, the higher the sale, the better.

Similar analyzes are repeated for all other consoles, modifying the note to be capped and the sale it delimits, as well as other outliers. Individual studies are found in Appendix B. Another factor that may be taken into account would be that some consoles have more observations than others; in addition different consoles had different sales in the market. Software released for a high-end console may have a better chance of reaching a larger, already wider audience, but it also has to be borne in mind that a console with more games has greater internal competitiveness. More about the relationship between consoles and games will be seen in the Analysis section.

The graphs of the crossover between MetaScore and Sales for the different consoles, and their respective observed results, are found in Appendix. They all have similarities in appearance, even though they are from different times and have sales over different consoles.
This can start an idea about the repetition of this pattern.

As we can see, all consoles actually have visible points limiting sales by grade, so now let's look at how they behave when all games on all consoles are together in Figure 5.

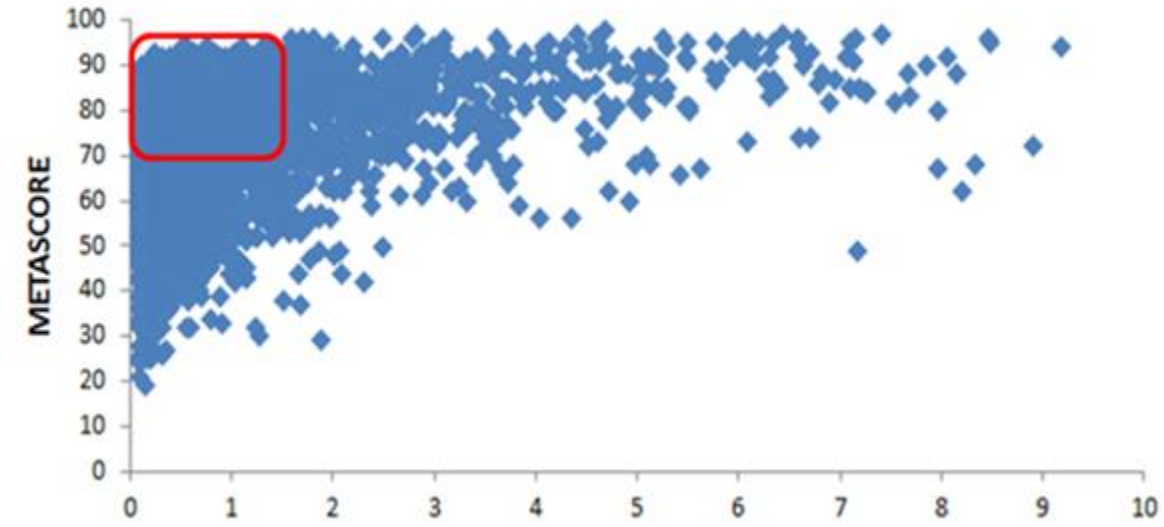
Figure 5 - RELATIONSHIP BETWEEN SALES AND METASCORE (ALL)



Source: prepared by the author

Initially we can observe that only games with grades above 50 achieved sales above 2.5m, only games with grades above 60 achieved sale above 5m. With the exception of only 6 observations, it jumps to 70 the grade limit to exceed 5m of sales.
To better observe, in Figure 6 were removed the top 1% of games (45 games excluded) with higher sales, so the graph is more accurate to be analyzed visually.

Figure 6 - PROBLEM AREA STATEMENT

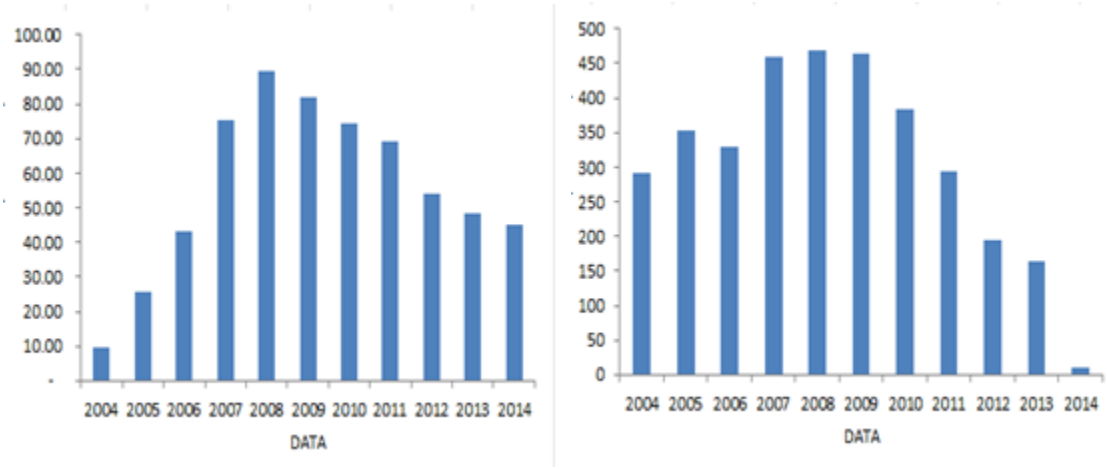


Source: prepared by the author

Having a better view of the upper left region in Figure 6, it has a very high observation density. This zone identifies the region with high-grade but low-selling games (under 1.5m), thinking of grade as an instrument for quality, we have a problematic region. Intuitively one would think that high quality games should sell more than low quality games, but the region delimited in the graph gives us counter intuitive information that this does not necessarily occur in this market.

In addition to sales correlating with the Metacritic scores, we can see from Figure 7 below a relationship between the annual console sales of the last 10 years, and the sum of database software sales. There is a peak in 2008, with a drop to 2014. The biggest problem directly comparing would be that we are not using all available hardware for 2012, 13 and 14. So these numbers are biased downwards in the Software part, obviously not. The use of PS4, WiiU, 3DS, PSVita and Xone prevents the entry of numerous games that could bring the two curves even closer, but the peak in 2008 would remain unchanged.

Figure 7 - HARDWARE AND SOFTWARE SALES COMPARISON



Not only the console that the game was released on is important, but possibly also the year it was released seems to have relevance, along with MetaScore. To better define the relationships between these variables, a more specific statistical study is needed.

5.2 ECONOMETRIC ANALYSIS AND RESULTS

In addition to the Descriptive Analysis, it is interesting to note also a Statistical Analysis, to quantify the correlations. An econometric analysis would firstly have a correlation between sale, note and year in which it was launched. From the Descriptive Analysis we can see that the relationship between note and sales is not linear, so we would have to consider the use of a differentiated function, whether quadratic, exponential, log-log, log-lin, and others.

Within the econometric study we will use two strands, the Score crossing Sales and using the years as dummies, and later also using Consoles as dummies. Observing the t-student for degree of acceptance and the coefficient to quantify the effect applied by the variable.

First, to test correlation between Sales and Score, we used Pearson's correlation coefficient as a parameter. Two forms were tested in Score (linear and quadratic), and two forms in sales (linear and logarithmic). We found the following result, as shown in Table 5 below:

Table 5 - Pearson Correlation Test Results

Pearson	**Score**	**Score** ^{2}
Sales	0.307897	0.334338
Ln (Sales)	0.460464	0.482919

Source: prepared by the author

All tests had a significance level of 1%, showing that there is a positive and considerable relationship between them, especially when Score is in logarithmic format and Score in quadratic format, which initially directs us to a format to be tested in subsequent equations. in the Least Squares model. A high value in the correlation coefficient would be desirable, but consideration should be given to sample size and the use of actual values. This result, together with the Descriptive Analysis, reinforces the idea that there is indeed a relationship between the variables. For greater accuracy and quantification of the ratio, we will next test econometrically.
Using the database of all games, every year, all platforms, running M [inios Squares in different ways. Within the equation formatting, one of the quality criteria would be the student's t or the p-value of the variable. Student's T uses as a null hypothesis the resolution that the estimated coefficient would be zero, nonexistent. Objective to have a significant variable would be that the null hypothesis was rejected, but it can be rejected at different levels of significance: 1% when student's t is outside the range between -2,576 and 2,576, or p value below from 0.01; 5% when t of student is outside the range between -1.96 and 1.96 or p value below 0.05; and 10% when t of student is outside the range between -1,645 and 1,645 or p value below 0.1. Any value within] -1,645, 1,645 [and p value over 0.1 was considered insignificant.

Four formats were used to calculate the equation: Score, Constant Score, Square Score and Score, and Constant Square Score and Score. Total sales maintained the logarithmic format. As an initial objective, observe the existence of correlation between MetaScore and Sales. As a result, in all 4 systems, significance was found at 1% level (p-value <0.01) for the Score variable, as seen in Table 6.

Not using current-generation consoles could cause some problems with the individual study of the effect of the years. Examples are the 3DS that was released in 2011, WiiU was released in 2012, and since sales of these consoles are not being considered, there could be a problem identifying the correct effects of time on sales. With the release of a higher-generation console, sales of the previous one can be expected to fall, so there would be a bias not easily circumvented without the use of recent data. Another factor would be that within the database itself there are consoles that started and ended, for example GameCube, which was produced both hardware and software from 2001 to 2007, and we have observations dating from 1996 to 2014, so there would be years that some consoles would not own games. At the same time, there would be years that would have more consoles in life, again biasing the result. For these reasons, we move the study to the side of the impact of the hardware on which the game is located, rather than the year itself, but the analysis in question has not been ruled out, its results are in Appendix A.

The current generation has not been used as they include games released during 2015 and, as seen in the Hardware - Data section, the end of the year has a huge impact on sales. No matter how much software was released before the end of the year, its sales could be underestimated without the effects of Christmas.

For the regression assembly, the Least Squares system was used. Equations were assembled with 4 variations of Score and constant. Linear score, Linear score with constant, Score with squared score, and Score with squared score with presence of constant. In addition, the presence of all hardware (Table 7) and each year (Table 9, Appendix A) was tested. Table 6 below shows how the results behave using only Score variations and a constant.

Table 6 - Equations using only Score and Constant

Dependent Variable: log (Global Sales)				
	(1)	(2)	(3)	(4)
Score	-0.008474 *** (-35.75864)	0.037055 *** (34,63846)	-0.068393 *** (-53.03104)	-0.0827 *** (-10.54761)
SCORE ²	-	-	0.000796 *** (46.99441)	0.000899 *** (15,41049)
Const	-	-3.334096 *** (-43,31412)	-	0.478193 * (1.849839)
R-squared	-0.119511	0.212027	0.251306	0.25188

Number of observations: 4461

Student's T is in parentheses

*** 1% significance, ** 5% significance, * 10% significance

Source: prepared by the author

As we can see from Table 6, the Score variable always proves to be significant at the 1% level, regardless of format or constant usage, again identifying a potential correlation between MetaScore and overall game sales.

In Table 7 below, we selected the model that obtained the best results, both regarding its combination of p-values, as well as its R^2 and explanation within the market. In addition, for degree of comparison the decay that has been made from the equation with all dummies, until it can have all its variables within a significance level of 1% (P value <0.01). Decay occurred by the individual elimination of the variable with the highest p value. In this case, in order, we eliminated PsP, NDS, Ps1, Xbox360 and Ps2, so as to arrive at a result where we had no more variables to remove, all at a significance level of 1%.

Table 7 - Explanatory Equations between Score and Consoles

Dependent Variable: log (Global Sales)						
	(1)	(2)	(3)	(4)	(5)	(6)
Score	-0,07537 *** (-10,0551)	-0.071702 *** (-39,00314)	-0.0696 *** (-48,433)	-0.069026 *** (-48,73284)	-0.068022 *** (-51,1756)	-0.06725 *** (-52.9468)
SCORE ²	0.000842 *** (15,04541)	0.000815 *** (42.64508)	0.000799 *** (46,8489)	0.000797 *** (46.76908)	0.000791 *** (47,17618)	0.000785 *** (47.52339)
XBOX	-0.32066 (-1.28346)	-0.442689 *** (-7.096556)	-0.5093 *** (-10,044)	-0.538994 *** (-10,97673)	-0.577335 *** (-12,7155)	-0,60299 *** (-13.8547)
PS3	0.930016 *** (3,718357)	0.808699 *** (11,778,146)	0.74296 *** (12.6923)	0.711924 *** (12.48405)	0.673104 *** (12,51169)	0.647631 *** (12,39539)
GC	-0.16587 (-0.65978)	-0.287866 *** (-4.197377)	-0.35461 *** (-6,1006)	-0.383825 *** (-6,759,326)	-0.421938 *** (-7.86284)	-0.4476 *** (-8.59367)
WII	0.494855 ** (1,992,329)	0.373496 *** (6,053,326)	0.306565 *** (6,16313)	0.279191 *** (5,777,326)	0.242063 *** (5,398,451)	0.216502 *** (5.041143)
PS2	0.32375 (1,302122)	0.201422 *** (3,6674039)	0.13432 *** (3,229,127)	0.105591 *** (2.713013)	0.067639 ** (1.976049)	-
X360	0.3031 (1,226,369)	0.181797 *** (3,187526)	0.115368 *** (2,61946)	0.086531 ** (2,046314)	-	-
PS1	0.358848 (1,435171)	0.239891 *** (2,887,443)	0.175082 ** (2,32831)	-	-	-
NDS	0.235768 (0.944123)	0.113759 * (1,832471)	-	-	-	-
PSP	0.128375 (0.504418)	-	-	-	-	-
R-squared	0.331323	0.331284	0.33078	0.329965	0.329335	0.328747

Number of observations: 4461

Student's T is in parentheses

*** 1% significance, ** 5% significance, * 10% significance

Source: prepared by the author

As we can see in table7, in the chosen equation, as in all other equations, all the variations of Score are quite significant, reinforcing again the relationship between both variables. An interesting situation to note would be that the linear part of the Score is represented with a negative coefficient, while the quadratic part has a positive and much lower

coefficient. Recalling that, as a logarithmic system was used in sales, it prevents a negative sales result, which ensures a correct economic answer, since negative sales would not make sense.

Because it is a quadratic equation, it starts at a negative point near zero, finds its minimum, and then rises. This initial drop would show that very low-rated games could have higher sales than slightly higher-rated games, which would not necessarily make sense. However, it reaches its minimum quickly, close to 40 and 45 of Score. Looking at the database, we have the lowest score would be 19, and between 19 and 30 we have only 26 games. Below 40, we have 109 games. Then the possible problem would be minimized by the proportion of the sample, as it would explain a very low percentage of observations.

Observing the assembly of the chosen equation, the score p-value always remained significant at 1%, which reinforces again the idea that the score affects sales. Regarding the coefficient sign and its magnitude, we can observe certain movement patterns between variations of the equation. Gamecube and Xbox always have negative coefficients, Ps3 always has the highest coefficient, second being always the Wii. When present, Ps1 always has the third highest coefficient.

To understand if these movement patterns make sense within the market, let's look at some database details through Table 8:

Table 8 - Demonstration of Number of Games on Each Console, Total Sales, and Average

Console	No. of Games	Total Sales (millions)	Average
PS3	331	653.69	1.92
Ps1	165	231.44	1.40
Wii	502	596.71	1.19
NDS	492	497.61	1.08
X360	772	807.89	1.05
PS2	1072	975.79	0.91
PsP	321	231.28	0.63
GameCube	320	292.5	0.56
Xbox	486	230.57	0.48

Source: prepared by the author

Using the average game sales of each console, we can see an interesting relationship with the coefficients found. Within the selected sample, Gamecube and Xbox have the two lowest averages, which corresponds to the fact that they have a negative coefficient. Ps3 has the largest sample mean, which also corresponds to the fact that its coefficient is the largest. Ps1 and Wii also obtained high coefficients and are in high positions within the mean analysis. This shows a possible explanatory relationship within the market for the coefficients result, giving more consistency to the analysis.

6 CONCLUSION

The scope of this work initially would be to verify the existence of a relationship between quality and sales, within the electronic entertainment market. As a quality instrument was used MetaScore, calculated by MetaCritic, and sales values were obtained through the basis of VgChartz.

Two methods were used to reach the desired conclusion, descriptive analysis and econometric analysis. Through Descriptive Analysis we were able to observe a positive relationship between sales and Score, the higher the Score, the higher the sales. It was also observed that they appear to have a nonlinear relationship. In addition, we may observe certain censorship points, such as

Score barriers to be overcome so that the sale has the possibility of exceeding certain values. For example, using all observations, only games with a score above 50 achieved sales above 2.5m.
For the Econometric Analysis , Pearson's correlation coefficient was initially used , which resulted in a positive and significant relationship . To improve the identification and quantify the relationship, the Least Squares method was used . For the choice of the equation the natural sales logarithm combined with variations in the Score formulation was used , besides a series of dummies were used to represent the effects of consoles and effects of years. Regardless of format, Score has always been significant at 1%, which again proves the relationship between them.

Within all options, the best equation was shown using quadratic score and certain console dummies . The coefficients fit within the reality of the database, since the lowest coefficients are the consoles that have the lowest sales averages, and the highest coefficients are the ones that have the highest sales averages.

Combining all the analyzes and studies, we can reach the conclusion that there is a strong possibility of existence of the relationship between Score and Sales within the study database. The database does not have a high number of games with high sales and low grades, but there are games with high grade but low sale. From this observation we can conclude that a high score does not necessarily guarantee a high sale, but it reduces the risk of having a low sale, possibly enhancing the effects of other variables such as advertising and marketing, location, pricing and others.

For further study of the subject, it is appropriate to study the effects that advertising and marketing have on high score and low score games. Mostly answer the question: Are high score games better affected by optimized advertising or is there no relationship?

Another factor that can be understood and improved in the study would be that not using the current 5 consoles causes problems in the study between years, for example, ps4 was born with ps3 still active, so recent years generate negative bias because they do not contain existing sales. .This can be corrected after the end-of-year effects are released, so we would have more

accuracy in sales data. Another study appreciated would be a possibility of calculating the demand expectation for the hardware market, being able to study separately the effect of the end of the year (November along with December).

Regarding the econometric technique, the use of panel could also be used as another argument for the final result, using fixed effect for time and consoles, but generating a balanced database proved to be a problem. If eventually corrected and tested, the results can be compared with the current study done here and see if they maintain the same type of result.

Within this study, the price of games was not taken into account because it is a considerably deeper analysis in another explanatory area, with another data collection system, inflation correction and relationship with different markets. But it would be an extremely valid study to be done in the future, with an adequate database.

Bibliography

Adams Greenwood-Ericksen, Scott R. Poorman, Roy Papp Eludamos. *On the Validity of Metacritic in Assessing Game.* Value Journal for Computer Game Culture. 2013; 7 (1), pp. 101-127

Alves, L. (2011) *Tale-Two: Metacritic key to games success.* Available at:< http://www.eurogamer.pt/articles/2011-03-09

Ashcraft, B. (]2015) *Nintendo really likes Metacritic.* Available at:< http://kotaku.com/nintendo-really-likes-metacritic-1686254849

Carter, G. (2012). Funcom Blames MetaCritic for Share Price Drop. Available at: < http://www.escapistmagazine.com/news/view/119015-Funcom-Blames-MetaCritic-For-Share-Price-Drop

Everiss, B. (2008) *Metacritic has changed the games industry*. Avaliable at:< http://www.bruceongames.com/2008/06/04/metacritic-has-changed-the-games-industry/

KASTENSMIDT, C., Apontamentos para o Estudo de Adaptações de Jogos Digitais (in Portuguese). Revista Translatio, No. 6, 2013

Kohler, C .*Power up: How Japanese Video Games gave the world an extra life*. Set, 2004

Marchand, André, and Thorsten Hennig-Thurau. "Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities." *Journal of Interactive Marketing* 27.3 (2013): 141-157.

McDonald, K. (2012) *Is metacritic ruining the games industry?* IGN, 6/16/2012. Available at:< http://www.ign.com/articles/2012/07/16/is-metacritic-ruining-the-games-industry

Murdoch, J. (2010) *Metacritic: Gaming the score*

Nintendo, Investor Relation (2015). Available at:< www.nintendo.co.jp/ir/en/library/events/150217/index.html

Periera, C. (2012): *Metacritic presents real problems for the industry*. Available at:< http://www.1up.com/news/metacritic-presentsproblems-industry

Savillo, R. *Metacritic Scores vs. Sales: 2009 Holiday Season Analysis*. http://venturebeat.com/community/2010/04/28/metacritic-scores-vs-sales-2009-holiday-season-analysis/

Sheff, D .*Game over: press start to continue.* (1999)

Sieberg, D. (2011) *Video Game ratings made by anonymous panel.* Available at:< http://www.cbsnews.com/news/video-game-ratings-made-by-anonymous-panel/

Tassi, P. (2015) . *Why Nintendo is wrong in using Metacritic*. Available at:< http://www.forbes.com/sites/insertcoin/2015/02/18/why-is-nintendo-wrong-to-be-proud-of-their-metacritic-scores/

UnSubject. (2013) . *In Defence of Metacritic for Video Games.* Available at:< http://evilasahobby.com/2013/07/22/in-defence-of-metacritic-for-video-games/

White, M. *The Senescence of Creativity: How Market Forces are Killing Digital. Games* LDG Vol 3, No 4 (2009)

Wingfield, N. (2007) *High scores matter to game makers, too*. The Wall Street Journal, September 20, 2007. Available at :< http://wsj.com/articles/SB119024844874433247

Weimar, J. (2011) *Telltale Developers Caught 'Cheating' on Metacritic.* Available at:< http://mashthosebuttons.com/2011/11/telltale-developers-caught-cheating-on-metacritic/

Domain references

MetaCritic. Available at :< www.metacritic.com > Available at: 2019

VGChartz . Avaiable at :< www.vgchartz.com >

VGSales. Revenue of the Video Game Industry. Available at:< www.vgsales.wikia.com/wiki/Video_game_industry#cite_ref-30

Dota2. Prize in dolar of Esport Championship. Available at:< www.dota2.prizetrac.kr/

TheESA. The Electronic Software Association. Available at:< www.theesa.com/about-esa/industry-facts >

Appendix

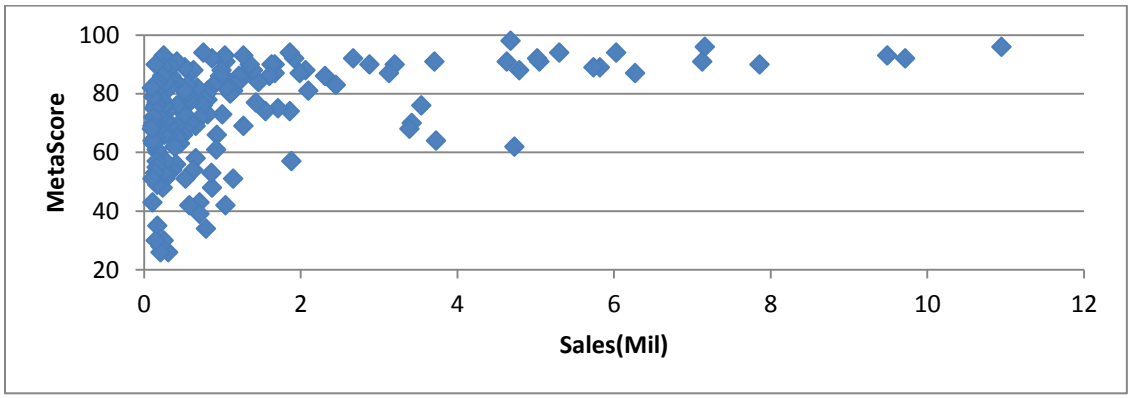
Table 4 -

Dep. Var. : log (Global Sales)					
	(1)	(2)	(3)	(4)	(5)
SCORE	0.036936***	0.036936***	0.037067***	-0.077972***	-0.076176***
	(34.67261)	(34.67261)	(34.93767)	(-10.09526)	(-51.78139)
$SCORE^2$	-	-	-	0.000865***	0.000852***
				(15.01402)	(47.98249)
C	-	-2.951505***	-2.982905***	-	-
		(-7.50996)	(-33.10202)		
1996	-2.951505***	-	-	0.384908	-
	(-7.50996)			(0.86851)	
1997	-2.762025***	0.18948	-	0.705414**	-
	(-10.06602)	(0.411149)		(1.995171)	
1998	-2.864338***	0.087167	-	0.626884**	-
	(-12.91018)	(0.201638)		(1.97325)	
1999	-2.915052***	0.036453	-	0.623807**	0.644067**
	(-15.13645)	(0.087015)		(2.069481)	(2.531829)
2000	-3.282125***	-0.33062	-0.308294***	0.340629	0.565548***
	(-28.74585)	(-0.844021)	(-3.178767)	(1.281692)	(2.81654)
2001	-3.349659***	-0.398154	-0.376209***	0.262604	0.562269***
	(-33.55913)	(-1.028778)	(-4.923455)	(1.01175)	(3.283157)
2002	-3.570519***	-0.619014	-0.597018***	0.071312	0.27915***
	(-39.56398)	(-1.608823)	(-9.339382)	(0.276351)	(3.132074)
2003	-3.593589***	-0.642084*	-0.620251***	0.061034	0.200901***
	(-39.19952)	(-1.668733)	(-9.644841)	(0.235355)	(2.96944)
2004	-3.459688***	-0.508183	-0.486197***	0.192486	0.130244**
	(-37.94617)	(-1.320092)	(-7.465083)	(0.743172)	(2.351456)
2005	-3.621571***	-0.670066*	-0.648058***	0.045663	-
	(-41.27844)	(-1.74403)	(-10.71847)	(0.176427)	
2006	-3.426135***	-0.47463	-0.452445***	0.252685	0.190255***
	(-38.62133)	(-1.233491)	(-7.102435)	(0.972377)	(3.539295)
2007	-3.174694***	-0.223189	-0.200737***	0.493878*	0.43186***
	(-37.52831)	(-0.580624)	(-3.304)	(1.914951)	(8.60742)
2008	-3.177204***	-0.225699	-0.203215***	0.483114*	0.421248***
	(-37.85108)	(-0.587274)	(-3.376364)	(1.878497)	(8.518406)
2009	-3.260168***	-0.308663	-0.286337***	0.385026	0.323223***
	(-37.95279)	(-0.802908)	(-4.675536)	(1.499064)	(6.373178)
2010	-3.144067***	-0.192562	-0.170369***	0.487798*	0.426056***
	(-34.8089)	(-0.499932)	(-2.581233)	(1.894727)	(7.586946)
2011	-3.04466***	-0.093155	-	0.574792**	0.512901***
	(-30.70934)	(-0.240911)		(2.212831)	(7.771415)
2012	-3.005479***	-0.053974	-	0.62496**	0.562349***
	(-25.97425)	(-0.138349)		(2.341912)	(6.680137)
2013	-2.81386***	0.137645	-	0.816449***	0.753921***
	(-21.68911)	(0.348754)		(2.99151)	(7.372754)
2014	-3.429298***	-0.477793	-0.455771**	0.265459	-
	(-14.73146)	(-1.084417)	(-2.035115)	(0.792707)	
R-squared	0.250676	0.250676	0.249833	0.286882	0.286575

As we can see, we have equations (1), (2) and (4) using all the studied years as dummies and varying in relation to the score and constant assembly. Equations (3) and (5) are the end result of equations (2) and (4), respectively. Dummies were taken individually, using the highest p-value as the parameter of choice for the year taken. Equations were considered acceptable when p-value below 0.05 was obtained in all their variables. Further study could shed light on the economic motives behind the coefficients found. Because 2008 was a peak year for hardware sales, one of the highest coefficients was to be expected, but did not yield a striking result by comparison. On the other hand we have 2013 with high coefficient. As explained during the study, the most recent years have a bias in the number of observations, which makes an accurate result for the analysis of the final result unfeasible.

Appendix - By console

Fig 6 - PS1



On PS1, only games rated above 50 managed to overcome the 1m sales barrier, above 1.3m only above 70 and above 2m only above 80 (except for five games).

Fig 7 - PS2

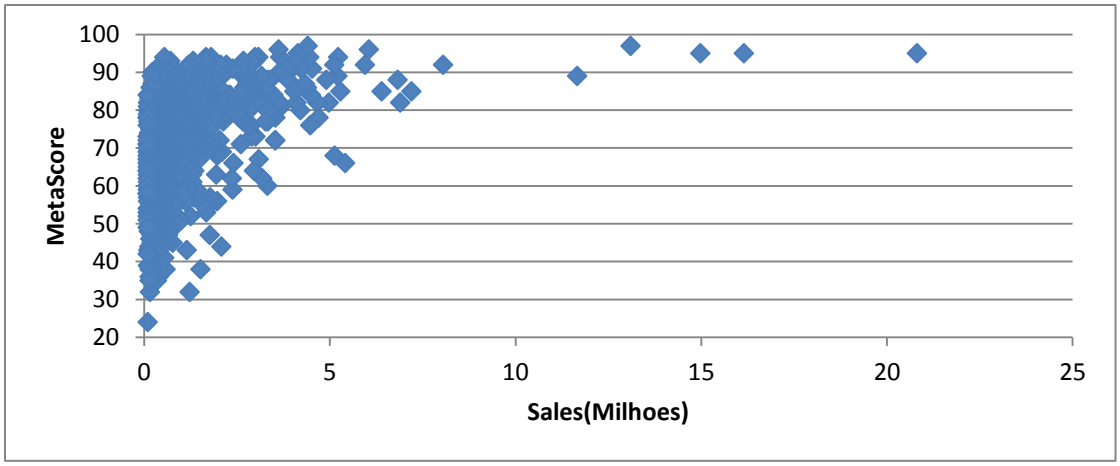


Fig 8 - PS3

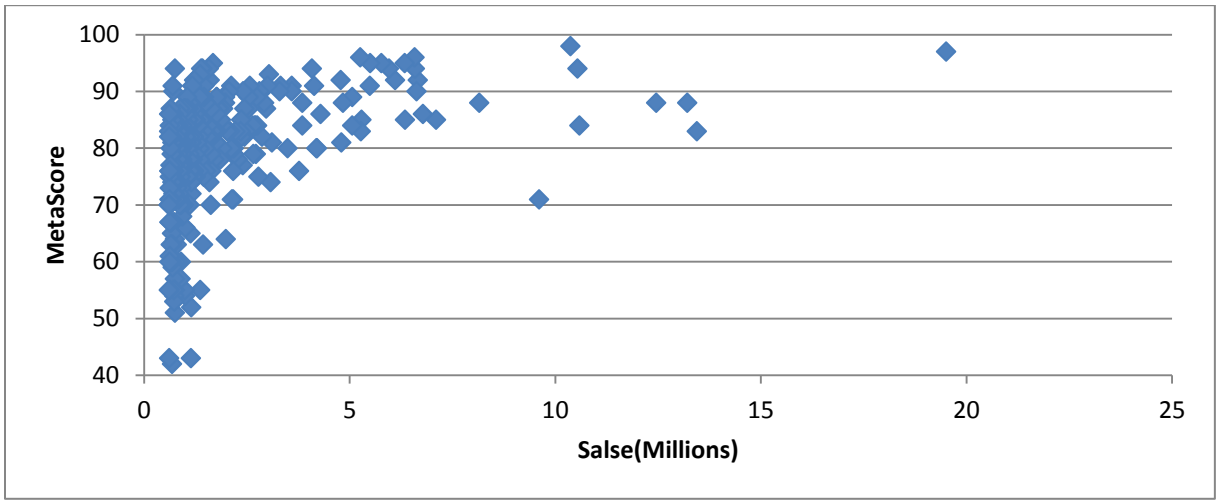


Fig 9 - PSP

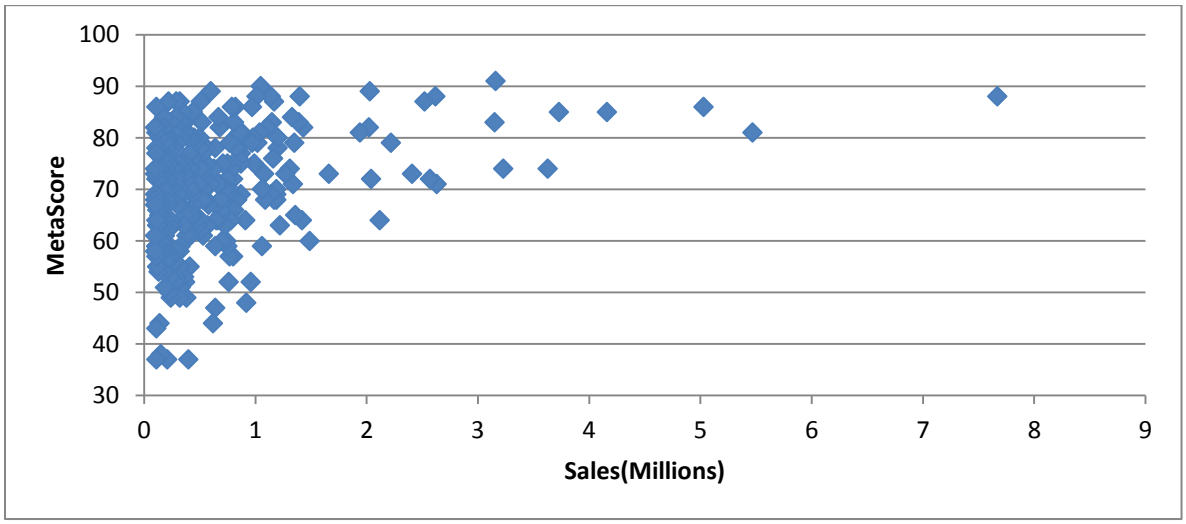


Fig 10 - Xbox

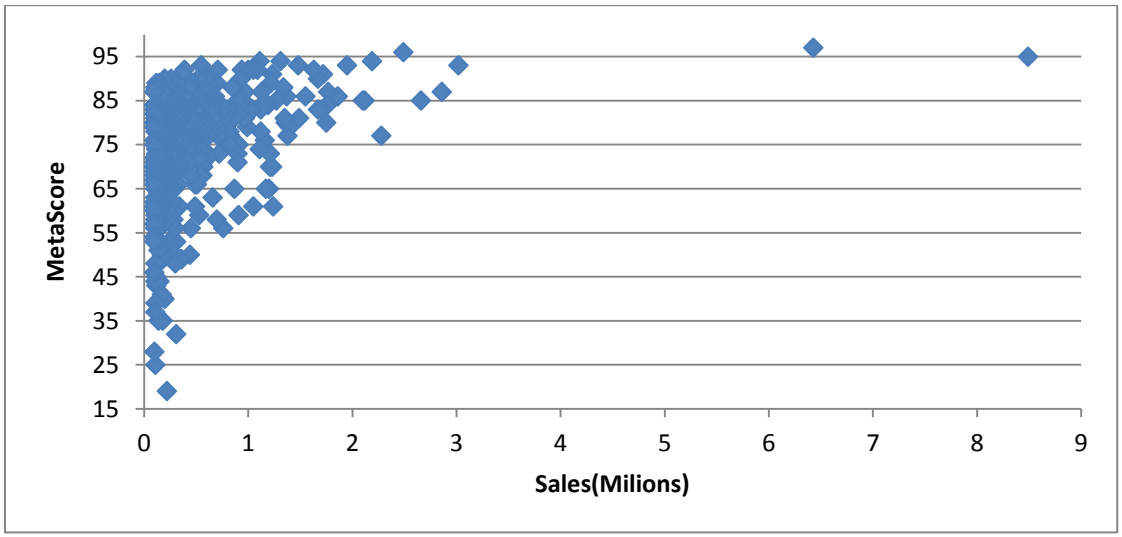
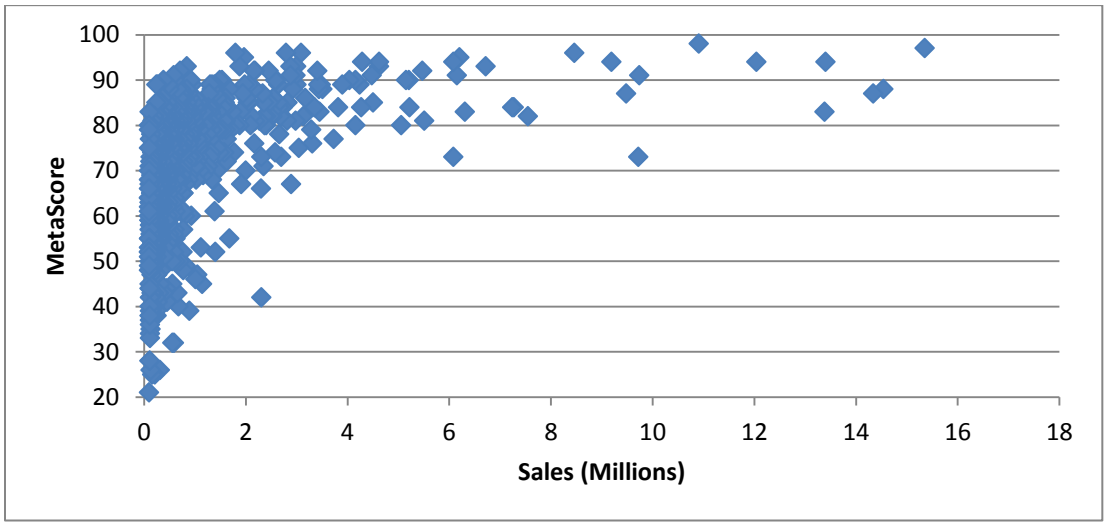


Fig 11 - X360



In the X360, only games with grades above 50 managed to overcome the 1.5m sales barrier; over 65 to overcome the 2m sales barrier; above 70 to pass the 3m barrier; over 80 to overcome the 4m sales barrier.

Fig 12 - NDS

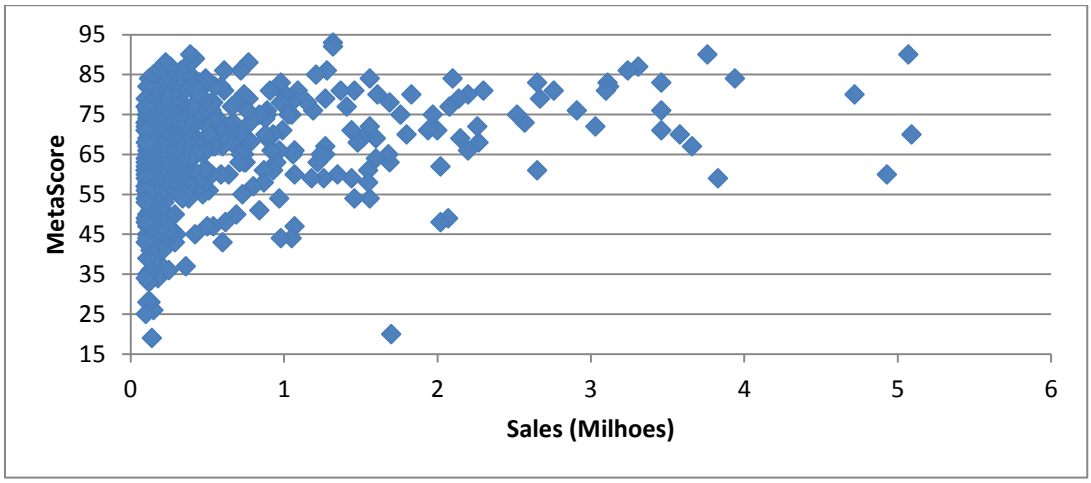
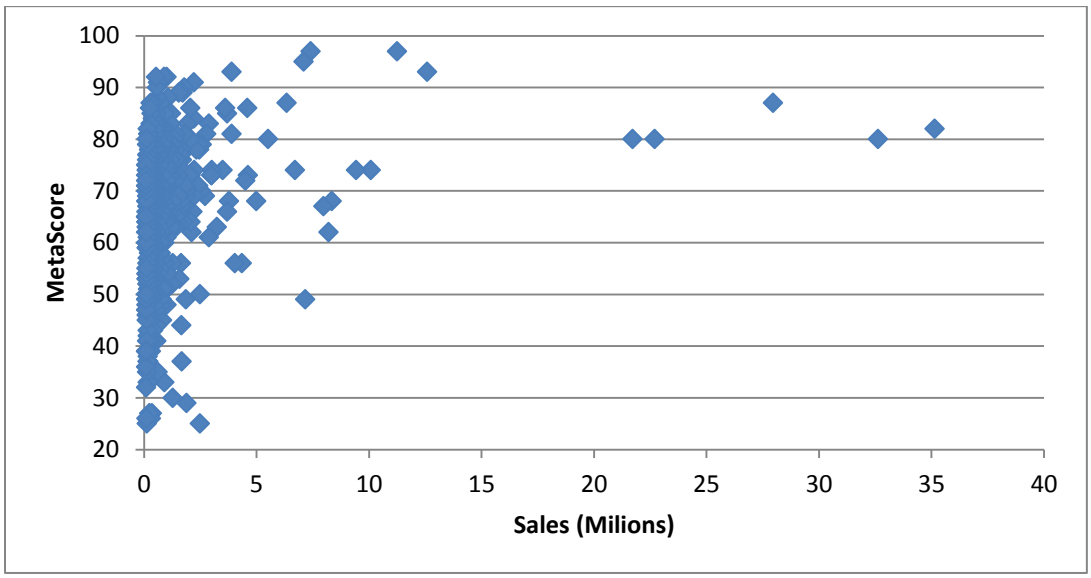


Fig 13 - Wii



The Wii has slightly different features than other consoles, the only visually significant significance being that the observation density above 1 million sales increases considerably from 60 note.

Fig 14 - GameCube