Reonne Benoy

A00268832 |

Warner Brother’s DC Projects Solution Proposal

Descriptive Analysis of DC and Marvel Projects for WB’s improved DC project performance

Contents

[Introduction 2](#_Toc27363650)

[Analysis results 2](#_Toc27363651)

[Movie gross 2](#_Toc27363652)

[Descriptive Statistics: DC Vs Marvel 3](#_Toc27363653)

[Comic books – DC vs Marvel 8](#_Toc27363654)

[Comic books from 2000s 12](#_Toc27363655)

[Record Setting gross data 13](#_Toc27363656)

[Directors influence 15](#_Toc27363657)

[Conclusion 16](#_Toc27363658)

### List of Figures

[Figure 1: Comic book movie dataset 3](#_Toc27392707)

[Figure 2: Descriptive Statistics: DC vs Marvel 4](#_Toc27392708)

[Figure 3: Boxplot of worldwide gross for MCU and DC movies 4](#_Toc27392709)

[Figure 4: Boxplot of DC and Marvel movies gross 5](#_Toc27392710)

[Figure 5: Probability plot of MCU and DC movies: Normality check 6](#_Toc27392711)

[Figure 6: Summary Report of DC films 7](#_Toc27392712)

[Figure 7: Summary Report of Marvel films 8](#_Toc27392713)

[Figure 8: Descriptive Statistics on comic books issued 9](#_Toc27392714)

[Figure 9: Bar-graph of sales for Comic books 9](#_Toc27392715)

[Figure 10: Bar-graph of sales for comic books 10](#_Toc27392716)

[Figure 11: Boxplot of comic book issued (All Time) 11](#_Toc27392717)

[Figure 12: Boxplot of Sales of comic books (All Time) 11](#_Toc27392718)

[Figure 13: Summary report of DC comics (All Time) 12](#_Toc27392719)

[Figure 14: Summary Report of Marvel Comics (All Time) 13](#_Toc27392720)

[Figure 15: Normality test for comics from 2000s 14](#_Toc27392721)

[Figure 16: Bar-graph record-setting gross 14](#_Toc27392722)

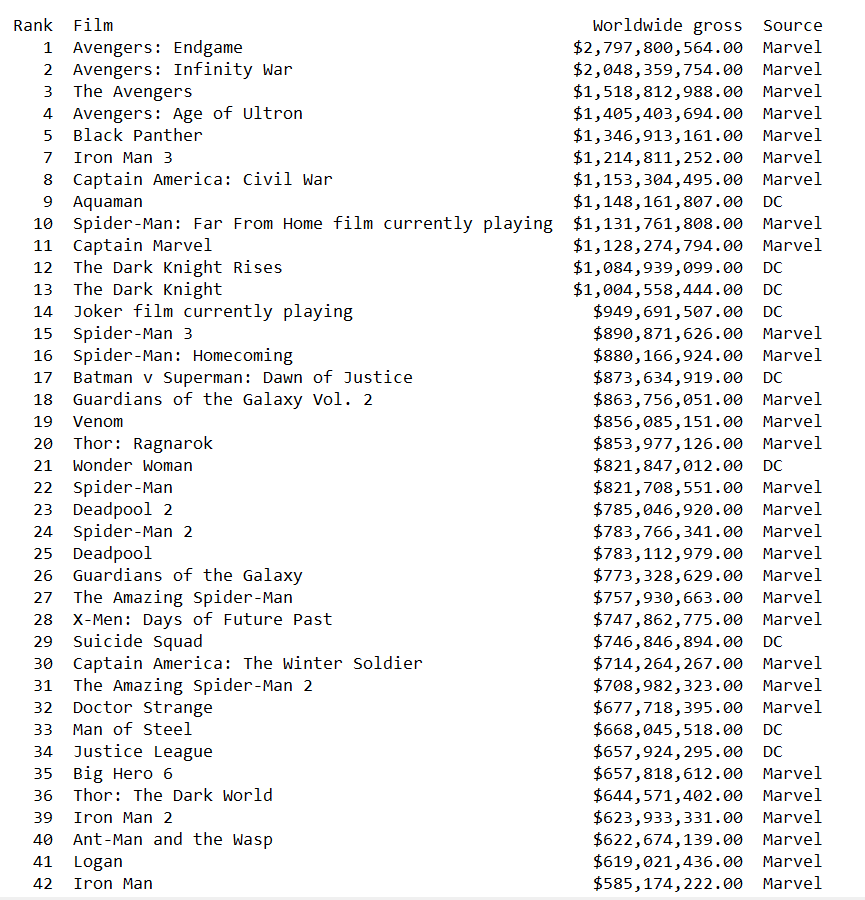
[Figure 17: Bar-graph of record setting (Years) 15](#_Toc27392723)

# Introduction

The final report focuses on the analysis of data we considered. The analysis we use will be simple statistical approaches such as mean, median, standard deviation and normality testing. The report only consists of a descriptive analysis and would not be dealing with predictive analysis and hence, any opportunity to include any of the Z- test and t-test would be avoided as of now since, we wouldn’t be testing on different confidence intervals to check for its significance and normality. We also look at the bar-graphs and the boxplots for a visual representation of the description deduced.

# Analysis results

## Movie gross



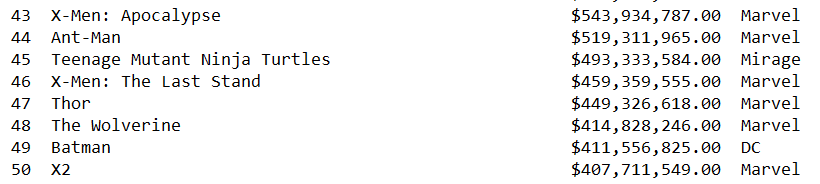


Figure 1: Comic book movie dataset

The MCU and DCEU movie gross and their ranks show us that the MCU has out-grossed the DCEU worldwide and also by a large margin. While only three of the DC projects, shown above, have crossed the 1 billion-mark, multiple Marvel projects, as revealed, has crossed the 1 billion mark.

### Descriptive Statistics: DC Vs Marvel

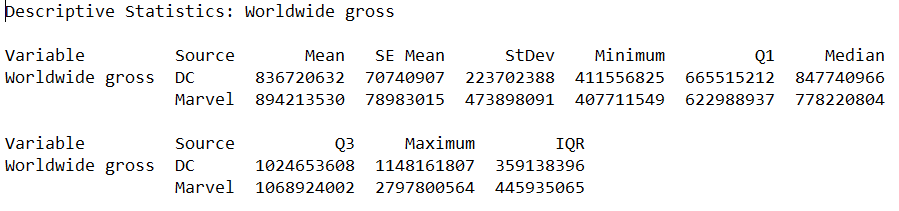


Figure 2: Descriptive Statistics: DC vs Marvel

From the descriptive statistics of the worldwide gross of DC projects and Marvel projects, if we look at the means, we can see that, the Marvel projects tops DC at 894 million. The standard deviation shows that Marvel project gross are influenced by more outliers relative to DC project gross. Further analysis on the gross will show how and why the outliers exists. The outliers influence other factors such as comic book sales as well.

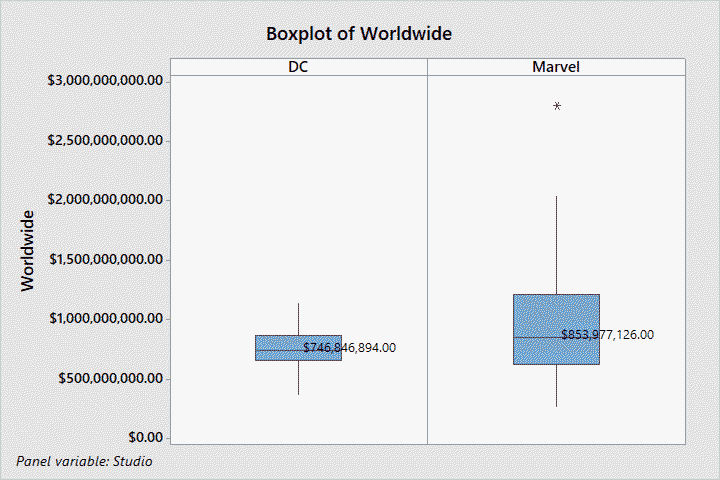


Figure 3: Boxplot of worldwide gross for MCU and DC movies

The boxplot of the worldwide gross is taken to show that the DC projects gross falls between the minimum and the maximum of the worldwide gross of the Marvel projects, which means that the worst performed DC project has still made more profit than the worst Marvel project. However, in terms of who made the most out of the profits, it can be easily said to be Marvel.

The above boxplot only adds one of the outliers and hence, the sole data value, indicated using a ‘\*’ at close to the 3 billion mark. We now look at the boxplot which adds all the outliers into the calculation, thereby increasing the median and the mean altogether.

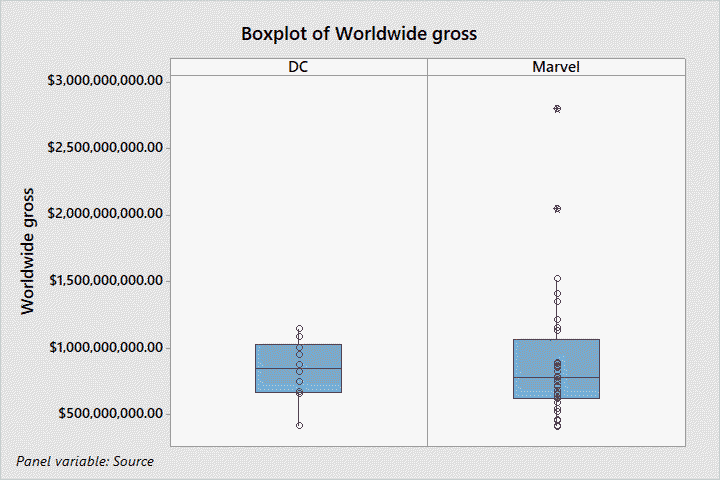


Figure 4: Boxplot of DC and Marvel movies gross

The above boxplot shows the outliers and how they affect the overall statistics. In the marvel section of the Boxplot, we see two outliers, one close to the 3 billion mark and the other just crossing the 2 billion mark. We shall check for the normality for the datasets of DC and Marvel.

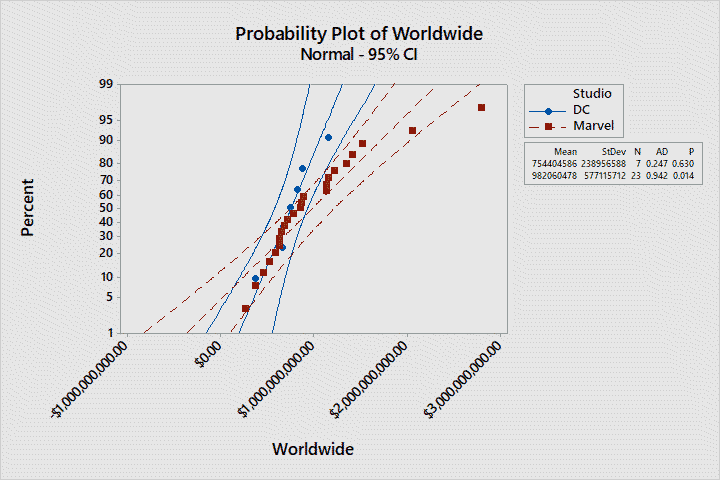


Figure 5: Probability plot of MCU and DC movies: Normality check

We see, from the above probability plot of the worldwide gross that the gross graph of the DC projects come close to the normality curve. When we look at the curve made by the Marvel projects, we see that the outliers affect the curve. The p-value of for the Marvel gross data is less than 0.05 (P-value < 0.05) and hence, we have sufficient evidence that the Marvel gross data don’t follow the normal curve. Similarly, if we look at the p-value for the DC gross data, the value is 0.630, which is greater than 0.05 and hence, we can say that the DC gross data follows the normal curve. This also means that there are no outliers to affect the curve for the DC data.

From the probability plot of the worldwide gross, we see that 90% of the individual data from the DC gross data lies above the regression line and also, the data starts to deviate from the normality curve on the last data value. Similarly, looking at the Marvel data values, we see that there are multiple outliers and it affects the curve. The two outliers lie below the normality curve in the farther end, which made more profit, whereas, there are some outliers lying below the normality line in the nearer end, with not much profit or borderline failures as far as relative box office results are concerned.

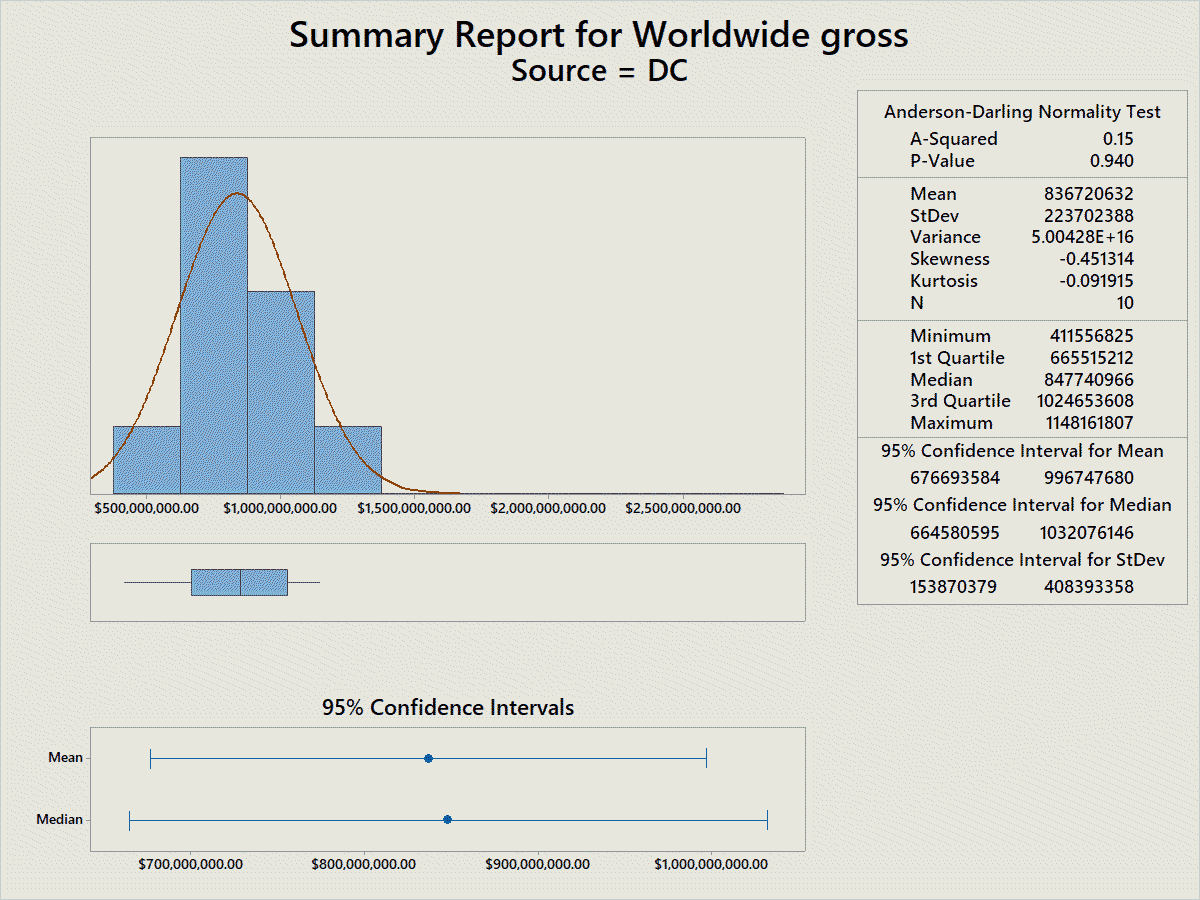


Figure 6: Summary Report of DC films

From the summary report, we see a summary of what was previous analysed. In the above graph, we see that the worldwide gross of the DC projects follows the normal distribution graph. The mean and the median don’t differ by much. Since, it is approximately similar, we can verify that there is enough evidence to claim that the dataset is not affected by outliers.

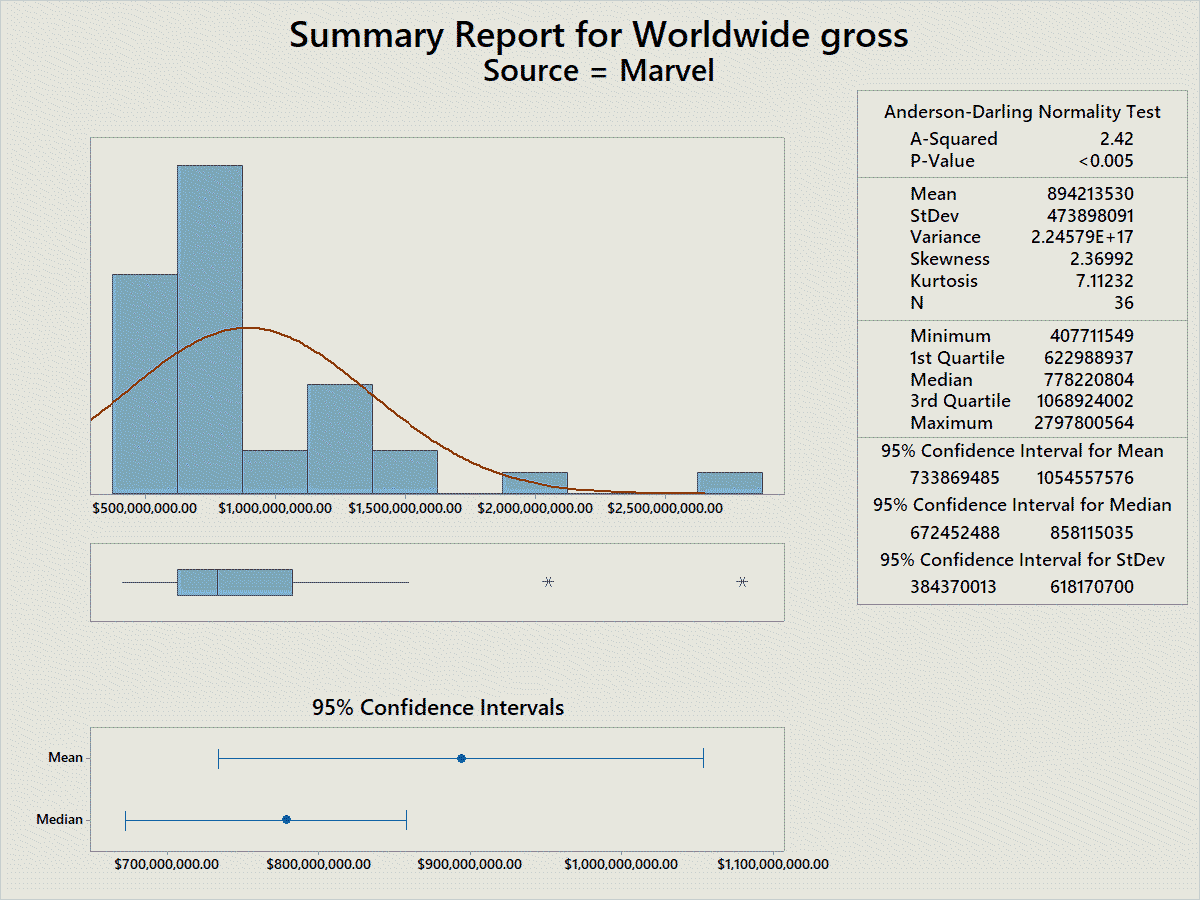


Figure 7: Summary Report of Marvel films

The summary report for the Marvel projects worldwide gross summarises the previously analysed data. The normality, the mean, standard deviation and other statistics are shown along with their respective confidence intervals for 95% confidence level. We can also verify that, the data for Marvel worldwide gross is affected by outliers. The difference between the mean and the median for the data is significant. This can only be possible due to outliers. Moreover, it doesn’t follow the normal distribution curve as the p-value is less than 0.05.

## Comic books – DC vs Marvel

In this section we see how the comic sales for both the studios were, before and after the advent of the popularity of having live-action comic book movies on the big screen. For this, we look at the number of issues and the approximated sales for the comic books

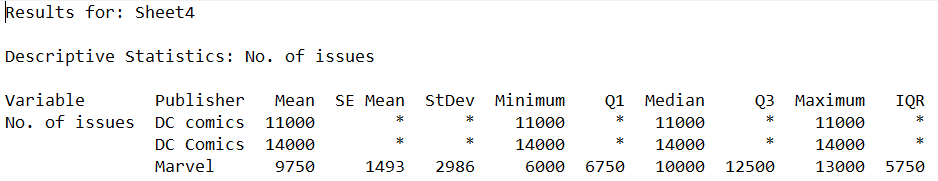


Figure 8: Descriptive Statistics on comic books issued

From the descriptive statistics, we see that the mean of the number of comic books issued for DC comics is greater than that of the Marvel comic books.

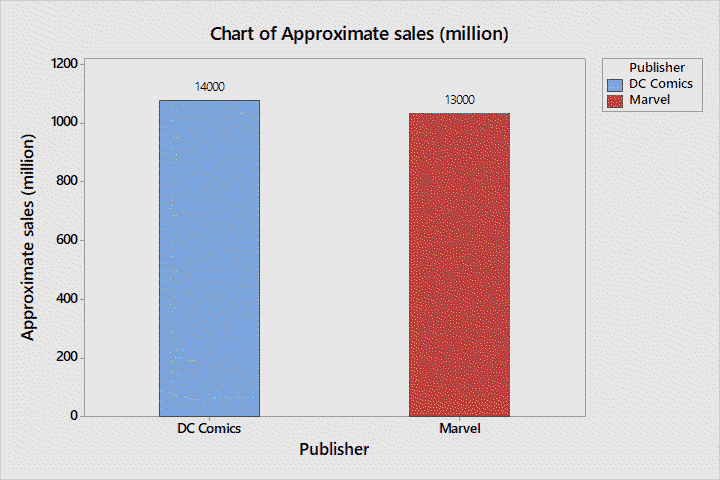


Figure 9: Bar-graph of sales for Comic books

From the above bar-graph, we see that for DC comics, the approximated sale is about 1.1 billion dollars for 14000 units sold whereas Marvel sold 13000 units for about 1 billion dollars.

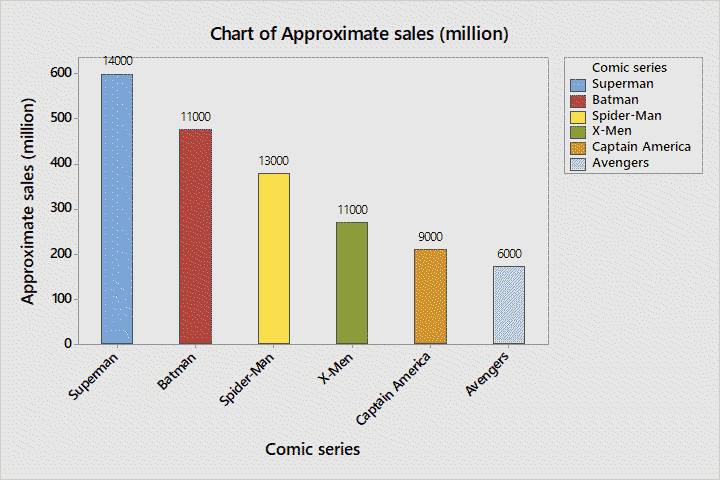


Figure 10: Bar-graph of sales for comic books

The above bar-graph just shows the same results as the previous bar-graph with the complete details of the other comic books. Here we see Superman and Batman comics issues from DC comics at the peak, signifying that they have the upper hand in terms of comic books sold. This is directly proportional to how many audiences know the character. Since, the two characters have the highest approximate sales than its competitors, we can deduce that people know more about the DC comic characters than its competitors and also, they would be more interested to watch anything about that character when it comes to live action movies on big screen. To observe the same, we look at the boxplot for the number of issues and the approximate sales.

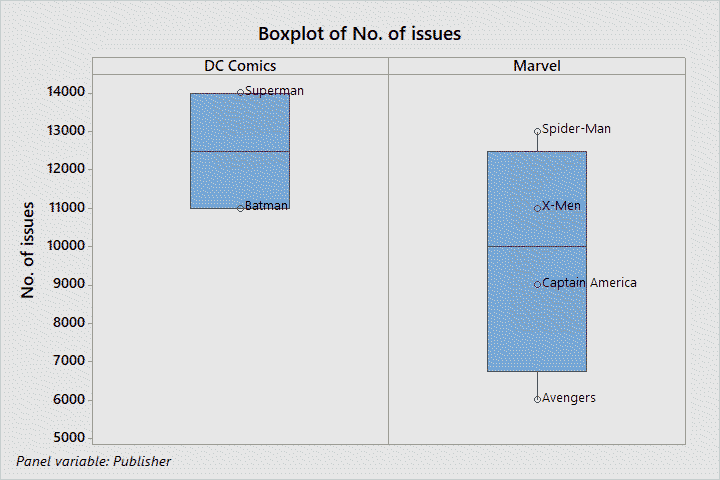


Figure 11: Boxplot of comic book issued (All Time)

From the above box-plot, we can see that the number of comic book issued by both Marvel and DC. We see Superman comics at the highest, from the DC comics, followed by spider-man comics, which is a Marvel comic. Now, we look at the approximated sales for the same. We do this because, it is possible that even though a greater number of issues were sold, the unit prices might differ according to the popularity of the character and their likability.

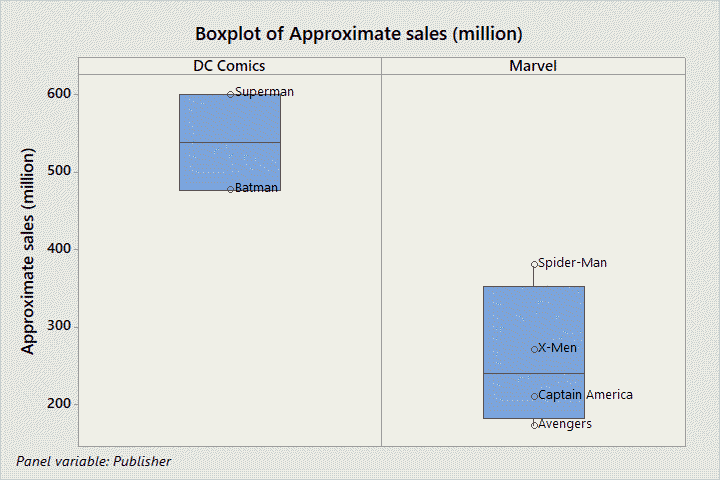


Figure 12: Boxplot of Sales of comic books (All Time)

When we look at the approximated sales of the comic books, DC comics stands unparalleled with both Batman and Superman having at least 450 million to Marvel’s best at about 350 million. This shows DC comics has popular characters which has a lot of potential.

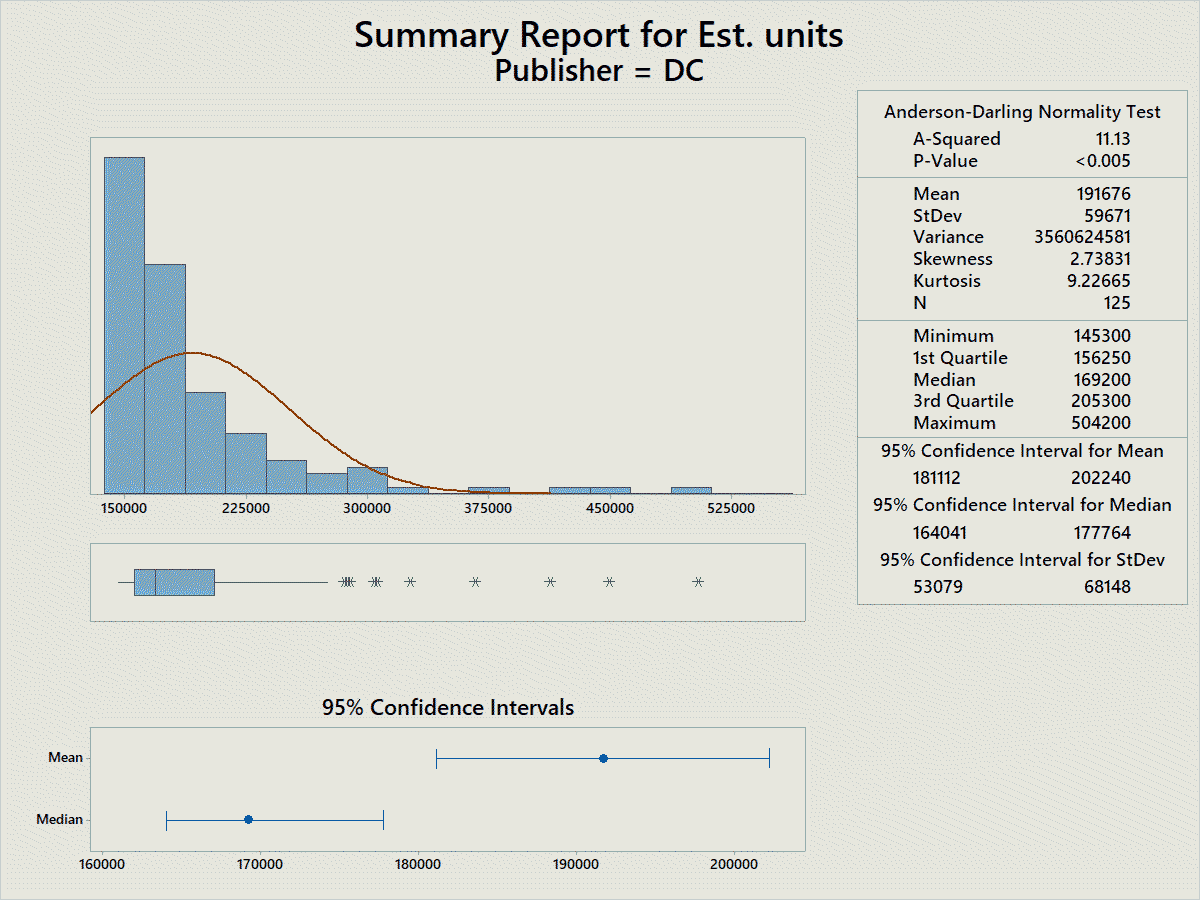


Figure 13: Summary report of DC comics (All Time)

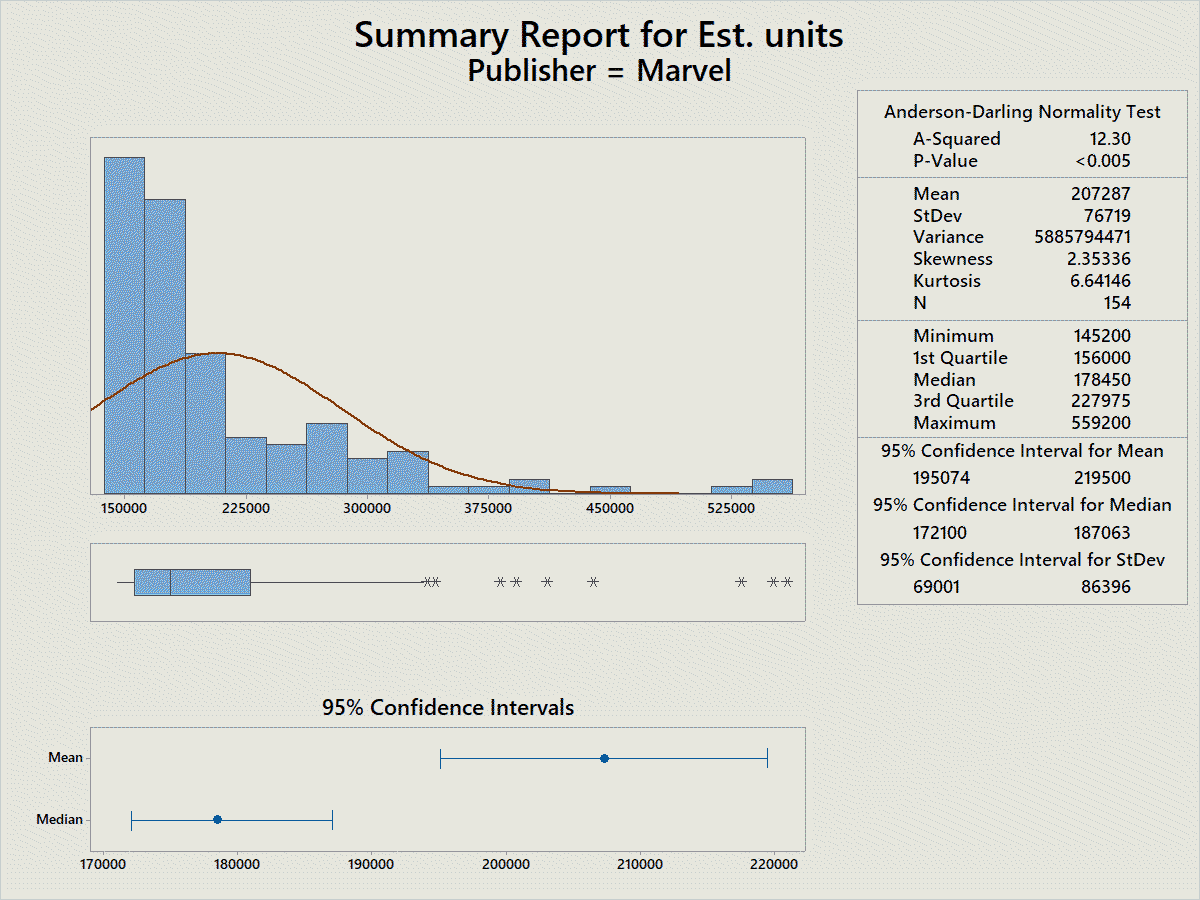


Figure 14: Summary Report of Marvel Comics (All Time)

The summary report for both DC comics and Marvel shows that the data is highly influenced by outliers. The difference between the mean and the median is significantly high. Moreover, the p-value is less than 0.05 for both. Therefore, they don’t follow the normality curve.

## Comic books from 2000s

From this data we see how the films have influenced the comic book sales as well. We could deduce that before the reception of the films, these comic books would not have worked its way through in the similar way, instead, it would have lagged behind its competitors.

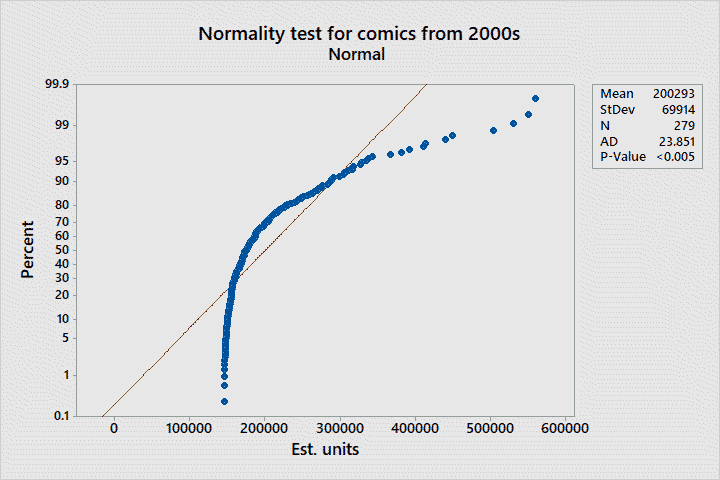


Figure 15: Normality test for comics from 2000s

We see a shift in the normality curve from the line. This shows the presence of multiple outliers. From the data, we can observe that the comic book movies have affected the comic book sales.

## Record Setting gross data

Record setting gross data shows the data about the record setting comic book films at their times.

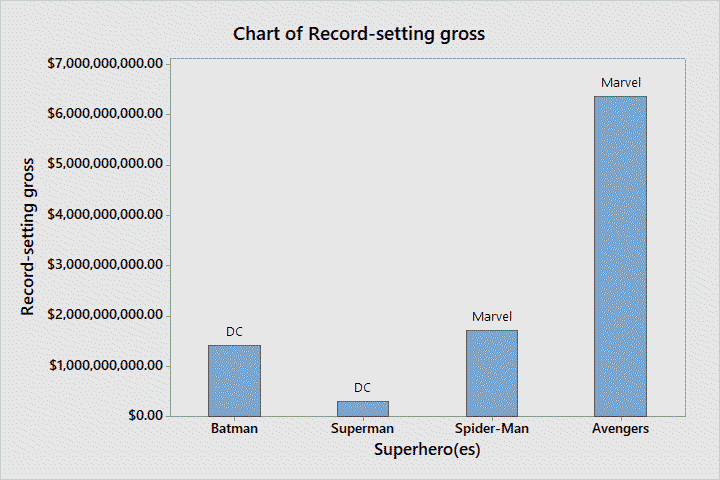


Figure 16: Bar-graph record-setting gross

Here we see the record setters in history. Two films from DC and the other from Marvel. We will need to look at the year in which the film set the record.

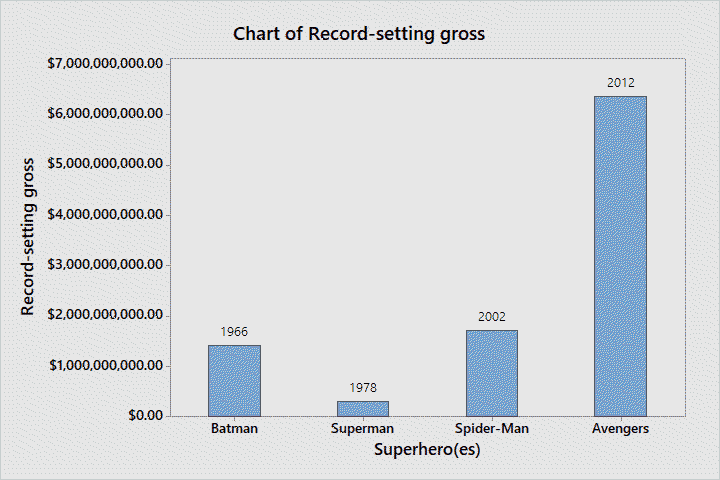


Figure 17: Bar-graph of record setting (Years)

The above bar-graph is the same as the previous bar-graph except this one shows the record-set year. The importance of showing the year is to correlate the data abstract to the ongoing situation as to how building up franchises, irrespective of their corresponding comic book familiarity and sales, works well in the film industry and studios could use it to earn much profit. “Batman” and “Superman” were record setters and to this day, those movies remain a precedent of greatness. As we can see, with time, the record setting movies started to be part of trilogies and franchises. And also, the time interval between two record setters reduced.

## Directors influence

|  |  |  |  |
| --- | --- | --- | --- |
| Sl. no | Movie | Director | Director Credits |
|  | Birds of Prey | Cathy Yan | 5 |
|  | Wonder Woman 1984 | Patty Jenkins | 13 |
|  | The Batman | Matt Reeves | 13 |
|  | The Suicide Squad | James Gunn | 15 |
|  | Black Adam | Jaume Collet Serra | 2 |
|  | The Flash | Andy Muschietti | 8 |
|  | Shazam 2 | David F. Sandberg | 20 |
|  | Aquaman 2 | James Wan | 17 |

The more the director credits, the more credibility the film will have. The above table shows the list of upcoming DC films with their directors and their director credits. A relatively higher director credits is observed for the directors of the upcoming DC projects. For a direct comparison, we would need the similar dataset for the directors of upcoming Marvel projects. However, due to the validity and reliability of the data about the directors, we omit the factual data for the time being and focus on the DC project and their potential. Since the project is towards providing a solution proposal for Warner Brothers on their DC projects, the focus on the project directors are a massive boost in building the solution.

# Conclusion

Therefore, we can say that the franchise creation of movies is a money-making scheme if we put it simply. Warner Brothers faced much competition and criticism for following Disney’s franchise scheme and had failed. The descriptive analysis provided shows that and more. Since, a predictive analysis is not carried out, we hadn’t dwelled into predicting when and how they could make a profit. However, a different approach to film making, and data analytics on a large scale to analyse what people would want from the films. Some of the ways to use data analytics and film direction together are:

* **Measurable target outcomes**: Target audience and profit prediction results.
* **Audience Behavioural proxies**: Production of films based on audience’s behaviours to the previous project of similar approach.
* **Predictive models and demand scoring matrices**: Quality of the predictive model and how well it could improvise according to situations.
* **Business decision prediction**: Quality of business decision making: long-lasting or short lived.
* **Size of Movies** A 3- hour film or below would test the model.
* **Genre**: Different genres attract different audiences.
* **Studio**: Projects of branded studios like Marvel, DC, Blumhouse and A24
* **Seasonality**: Christmas season, spring and summer, holiday seasons
* **Political Climate**: SJWs, political power and movements.
* **Time-Quality movies**: Divided into in-time bound and out-time bound movies.
  + **In-time bound movies**: Movies which are prevalent only for limited years (1-4)
  + **Out-time bound movies**: Movies which remain intact for a long time (10-40)
* **Geography**: Certain Genres of movies would not release at certain geographical locations. It may be due to societal status, beliefs or other factors.
* **Representation**: Gender, racial and sexuality representation is welcomed well.
* **Competition**: Other competitions apart from studious and companies like Netflix, Amazon Prime, Torrent downloads and video piracies.
* **Population and population estimate**: Making movies for large population locations and including into the business model to profit from the large population.
* **Online presence**: Data from not only reviews on YouTube but also the YouTube comment section, Instagram fan-pages, Facebook data, blogs, reviews, press, Twitter volume and sentiment and more.