# Q1

There is no strong correlation between. Between the budget of a movie and its profitability. Interestingly, investing more money into a movie didn’t seem to reduce risk. It did however, seem to be that movies with higher budget tended to be able to turn some more extreme profits. The linear regression line plotted on the scatter plot has a slope of just over 0.73. Our Pearson correlation coefficient was just over 0.05. Spearman Rank’s correlation came out to under -.08 with a P-value of . This is a line with a negative slope. This all suggests that there is likely no correlation between the values. We concluded that spending much over 250 million dollars to produce a movie was a bad idea. Based on the few movies around this area range and above, we don’t tend to see any wildly successful movies and the amount that stands to be lost is high. More importantly, movies with smaller budgets in the 150 million range can often see the same if not bigger profit margins without the high risk investment.

# Q2

On average movies tend to run at about 109.80 minutes long with a standard deviation of 22.75 minutes. Interestingly, the most profitable of movies are between 75 and 200 minutes but also the most failures. As movies become longer the extremes of profit taper out with most long movies losing money. We applied ridge regression using a 3rd degree polynomial. There are very few movies on the long side of the graph and found that it was hard to fit any trend line the data that we were happy with. Based on the plot using this polynomial we did glean some information upon closer examination of the data. Runtimes between 150 and 200 minutes tended to be a bit more profitable (or at least less unprofitable) than movies in the much shorter ranges. This could be due to lack of data for higher ranges, or that these longer movies can hold people’s attention longer. We took the same scatter plot and colored by genre. Looking at these ranges we learned some genres can be a little bit longer and be successful. In particular, we noticed several instances of long movies that were dramas, or adventure movies. (For convenience, we provided a zoomed version of the graph on the 150 to 200 minute range.

# Q4

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Mean Gross Profit | Standard Deviation Gross Profit | Successful | Unsuccessful | Total | Success Rate |
| G | $42,944,275.94 | $79,194,200.35 | 46 | 25 | 71 | 64.79% |
| PG | $27,960,527.49 | $69,304,377.37 | 300 | 160 | 460 | 65.22% |
| PG-13 | $18,588,209.40 | $64,524,998.19 | 579 | 480 | 1059 | 54.67% |
| R | $11,173,006.05 | $36,607,924.80 | 728 | 578 | 1306 | 55.74% |
| NC-17 | ($288,010.75) | $685,550.31 | 2 | 2 | 4 | 50.00% |
|  |  |  |  |  |  |  |

For this problem we focused on the profitability of movies with respect to their genre. We used a set of box and whisker plots to see what the distribution of this data was like. We learned that in general, the industry hovers just over breaking even on average. What is more interesting is that G and PG movies were more likely to be successful than their higher rated counterparts. We determined a movie was successful if it broke even or better. PG-13 movies could occasionally turn somewhat large profits and contained the most profitable of all movies. Rated R movies often make relatively small profits. Rated G movies have the most spread out profits. In addition, NC-17 movies are scarce making them more or less impossible to glean data from. There was only 4 of them in the entire dataset (with profit information). Due to the nature of content ratings, we believe G rated movies are the most accessible and have the widest possible audience while R and NC-17 are the least accessible. This explains the patterns we see in the data. Mean profit consistently goes down as content rating goes up.

# Q5

Graphs: UK line and US Line:

US: Animation, Western

Unfortunately the data for the second country with the second most movies, the UK is far more sparse. The US produces 78.9% of all the movies in our database with the required financial data. The UK produced 8.2% of the data.

In the US:

From 2000 through 2002 Westerns lost large sums of money. In 2003, they became the 4th most profitable. From 2004 to 2007 they continued their decline losing money each year without fault. Finally 2008 and 2009 saw no movies of this genre. When the genre came back in 2010 it was revived making an average of $48,285,381 making it the second most profitable genre in 2010 right after Animations.

This pattern of consistent losses with large one-year resurgences suggests that Western movie are not a very safe genre since they jump back in forth from least successful to most successful spots.

Animations have consistently performed very well and seem to be among the safest of options. In 2000 they lost $18,104,135 on average. Every year since, they always made money. In 2003 all through 2007 they were the most profitable of genres.

In the UK:

In general, the UK movies tend to have less genres appear each year than the US does and most of them lose money each year. In 2000, Animation was the most popular genre but no additional animations were produced in the UK until 2005 where they made a modest on average profit. During over the next 10 years only 2008, 2011, 2012 and 2014 would make Animation movies. All would have significant losses on average except for in 2014 where the genre became the 3rd most profitable on average. Looking at the data, there is only one data point for every year in the UK chart except for in 2005. This makes it hard to draw any conclusions about the market from the production country. Only one western movie was ever made in the UK.