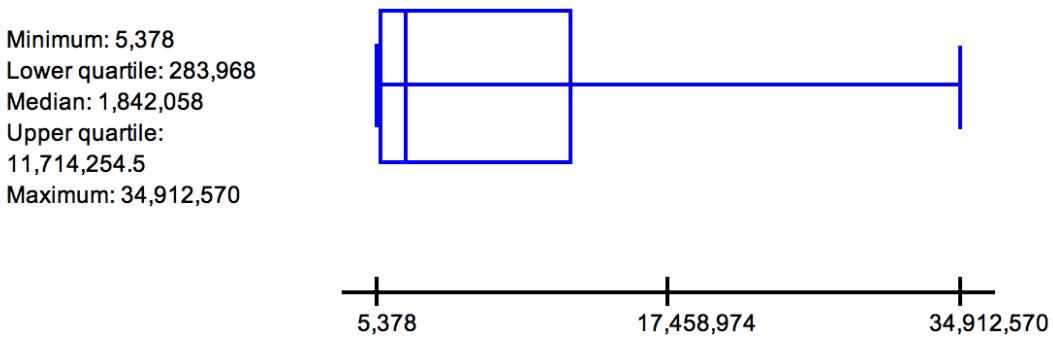
Introduction:

For this assignment I created 2 datasets, one is market analysis of 2011 and 2012 which includes distributors' names, movie numbers of 2011 and 2012, inflation adjusted gross of 2011 and 2012, Tickets number of 2011 and 2012, and ticket share of 2011 and 2012.
Another dataset is about the number of Academy Award nominees of different categories in 2010 and 2011, which include categories of Awards, number of nominees in 83 th and 84th Academy Award.
For dataset creation, I delete the data that only appears in 2011 or 2012. And for graph creation I basically used Excel.
In this report there are 5 sections, which answer the 4 questions in the assignment.
Section1 is for question1, section2 is for question 2, section3 is for the first question in question3, section4 is for the second question in question3 and finally section5 is for question4.

Section1

Question1:
A box and whisker graphic comparing the inflation adjusted gross with respect to the distributors (combine both years of movies for this exercise)

Graph and analysis:
For this question I basically used 2 columns data in the dataset, which are inflation adjusted gross of 2011 and 2012. I added the 2 columns so to get a new column of total inflation adjusted gross. Then I use this column of data to generate the Box and Whisker graph.
I used a online Box and Whisker graph tool to generate the graph.

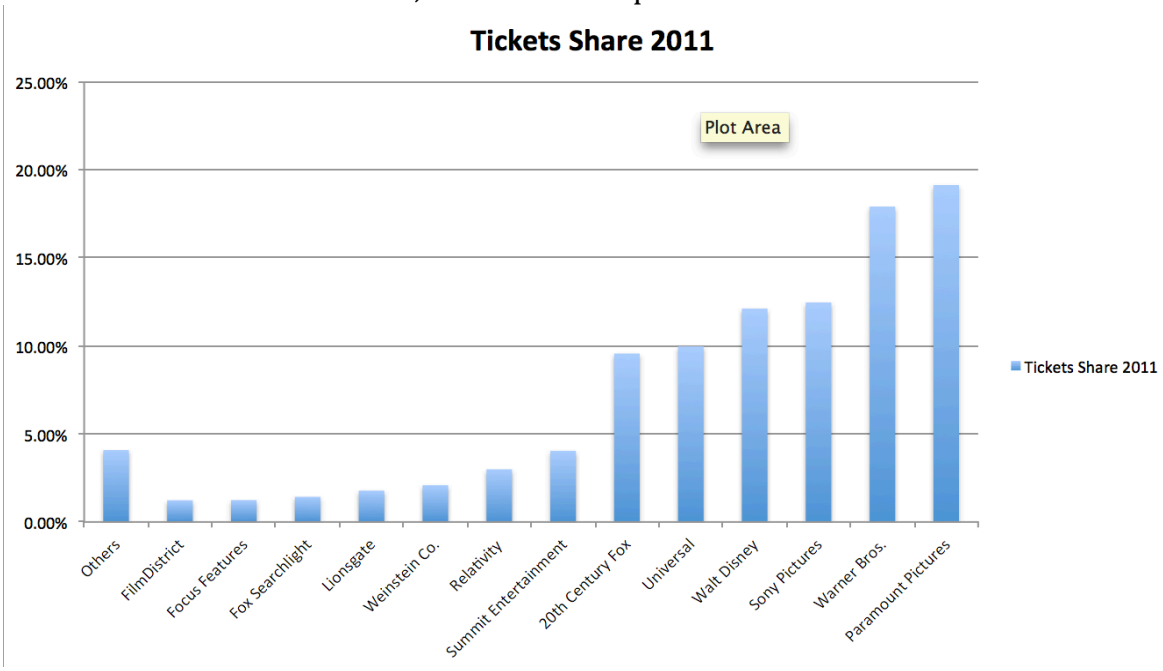


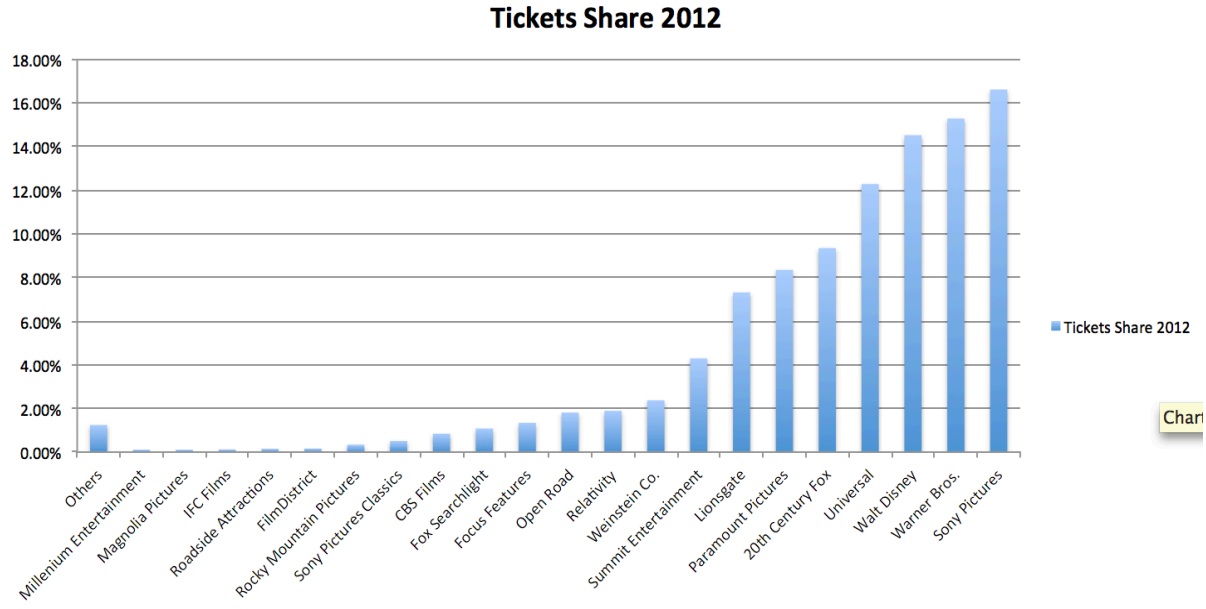
As the graph above shows, most of distributors do not have too much gross, basically around 283968 and 117142545 USD. And the median gross is nearer to the lower quartile. The upper extreme shows that the gross of a distributor can be really high as it gets to 34912570.

Section2

Question2:
Create two histograms showing the distribution of ticket sales in the year 2011 and 2012 respectively

Graphs and Analysis:
I used the data of tickets share to generate the two graphs. For those percentage of tickets share lower than 0.1, I added them up to be others.



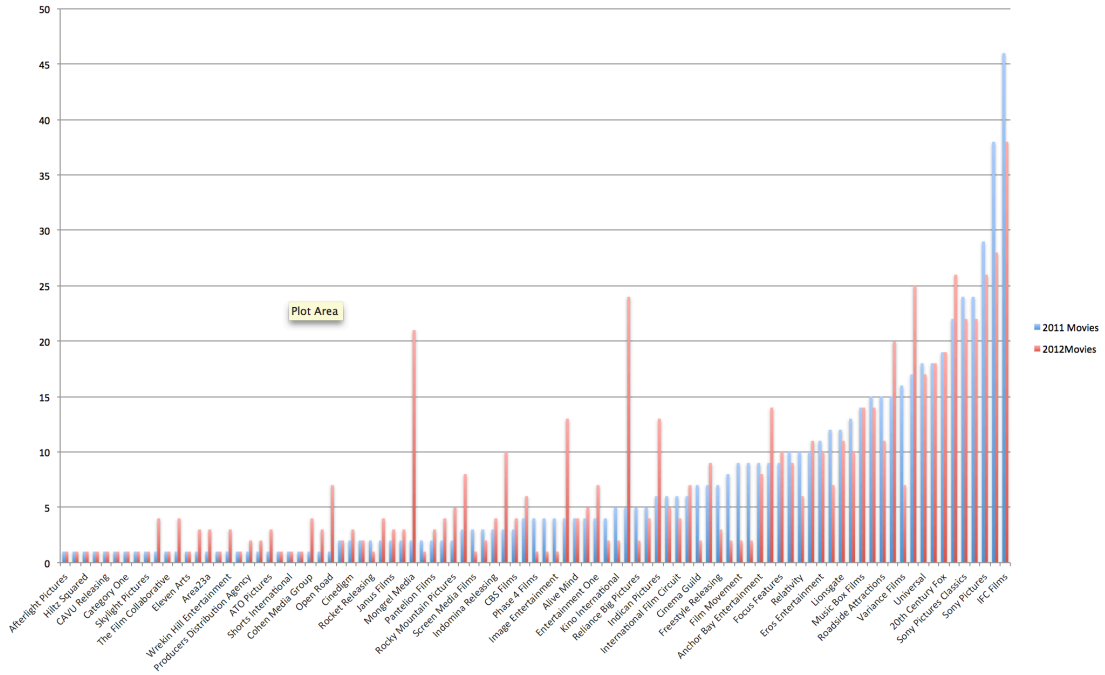


From the 2 histograms above we can see clearly that the distributors that have high percentage of tickets share do not change a lot from 2011 to 2012. Also the maximum percentage of tickets share of both years are lower than 20%. In 2012 there are more distributors which have more than 0.1% of tickets share.

Section3

Question3.1:
How does movie numbers of different distributors changes from 2011 to 2012?

Graph and analysis:
The distributors of this graph are ordered by ascending order of movie numbers in 2011.

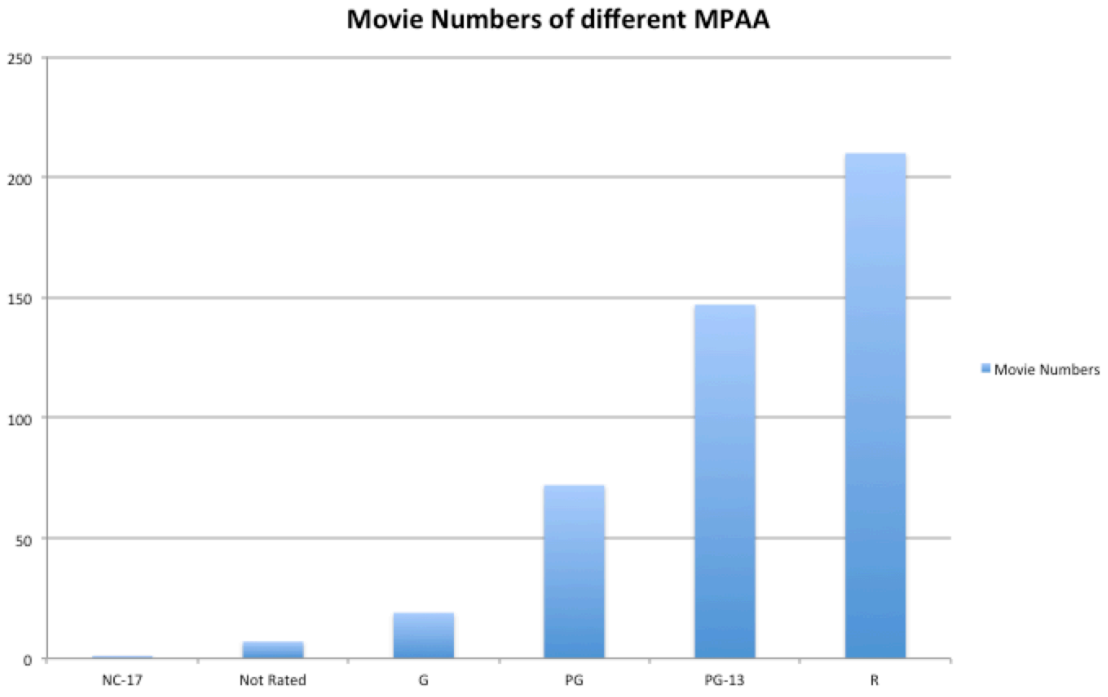


From the graph above we can see that the movie numbers produced by big distributors (produce more movies than other distributors) does not change o lot. But for small distributors movie numbers may change greatly, for example Mongrel Media produced 2 movies in 2011 and 21 movies in 2012.

Section4

Question3.2:
Show movies numbers of different MPAA ratings.

Graph and analysis:
The NC-17, not rated, G type movies do not have too much market share. And people are more intended to watch movie of PG, PG-13 and R MPAA ratings.



Section5

Question4:

Create a comparison of the Best of XXX category for the 83rd and 84th Academy Award Nominees (which correspond to the 2011 and 2012 season respectively). For example, compare the gross distribution of Best Picture Nominees between years. Information on Academy Award nominees can be found on Wikipedia (http://en.wikipedia.org/wiki/84th_Academy_Awards)

Graph and Analysis:

From the graph it's clear that numbers of nominees of most of the the categories didn't change a lot from 83th to 84th.

The 3 changed categories are Best picture, Best Animated Feature and Best Original Song.

Most of the categories have 5 nominees and the Best Picture has the most nominees.