

# The motion picture industry

[Entertainment](#), [Movie](#)



The motion picture industry is a competitive, multi-billion dollar industry involving production crews, marketing crews, actors, directors, distribution companies, movie theaters, and movie rental companies. Four variables are commonly used to measure the success of a motion picture; these variables include the opening weekend gross sales, total gross sales, number of theaters showing the movie, and weeks in the top sixty of gross sales. Using a sample of 100 motion pictures from 2005 and numerical methods of descriptive statistics, one can learn how these variables play a part in the success of a motion picture. Opening Weekend Gross Sales After calculating several central tendency measures (mean, median, and mode), one can create a histogram and see the opening weekend gross sales are heavily skewed to the right. The median opening gross sales of 0.39, indicating 50% of the opening gross sales values were less than 0.39 and 50% were greater than 0.39, would be a more appropriate central tendency measure than the mean due to this skewness (3.43). The range of opening gross sales was 108.43, from 0.01 (Caterina in the Big City) to 108.44 (Star Wars: Episode III), and the standard deviation was 18.87 ( $\hat{\sigma}^2 = 356.2544$ ). However, the interquartile range of 12.37 would be a more appropriate measure of variability because the distribution is skewed and has extreme values. Once a box-plot is created, one can see the extreme values are 108.44 (Star Wars: Episode III), 102.69 (Harry Potter and the Goblet of Fire), 77.06 (War of the Worlds), 50.34 (Mr. and Mrs. Smith), 48.75 (Batman Begins), and 33.90 (Wedding Crashers). Values that are less than z-score of -3 or larger than z-score of +3 should also be considered outliers, so, if using only this criteria, War of the Worlds (z-score: 3.586), Harry Potter and the Goblet of Fire (z-

score: 4.944), and Star Wars: Episode III (z-score: 5.248) would have still been considered outliers. Total Gross Sales As with the opening weekend gross sales, after creating a histogram and calculating the mean, median, and mode, one can see the total gross sales are also heavily skewed to the right. The median total gross of 5.85, indicating 50% of the total gross sales values were less than 5.85 and 50% were greater than 5.85, would be a more appropriate central tendency measure than the mean due to this skewness (3.28). The range of total gross sales was 380.15, from 0.03 (Sons of Provo and Jiminy Glick in La La Wood) to 380.18 (Star Wars: Episode III), and the standard deviation was 63.16 ( $\hat{\sigma}^2 = 3989.7784$ ). However, the interquartile range of 47.03 would be a more appropriate measure of variability because the distribution is skewed and has extreme values. Once a box-plot is created, one can see the extreme values, once again, are 380.13 (Star Wars: Episode III), 287.18 (Harry Potter and the Goblet of Fire), 234.21 (War of the Worlds), 209.22 (Wedding Crashers), 205.28 (Batman Begins), and 186.22 (Mr. and Mrs. Smith). Values that are less than z-score of -3 or larger than z-score of +3 should also be considered outliers, so, if using only this criteria, Star Wars: Episode III (z-score: 5.496), Harry Potter and the Goblet of Fire (z-score: 4.023), and War of the Worlds (z-score: 3.185) would have still been considered outliers. Number of Theaters Similar to the previous two variables, one can see the number of theaters is also heavily skewed to the right after drafting a histogram and calculating the mean, median, and mode. The median number of theaters of 410, indicating 50% of the number of theaters values were less than 410 and 50% were greater than 410, would be a more appropriate central tendency measure

than the mean due to this skewness (0.56). The range of the number of theaters was 3,905, from 5 (Lost Embrace and Head On) to 3,910 (War of the Worlds), and the standard deviation was 1378.69 ( $\hat{\sigma} = 1900784.582$ ). However, the interquartile range of 2,580.75 would be a more appropriate measure of variability because the distribution is skewed. There was not any number of theaters outliers observed using the box-plot method or z-score criteria.

Weeks in Top 60 The weeks in the top sixty variable does have a right tail after creating a histogram and calculating the mean, median, and mode, however, the data is not as heavily skewed to the right as the data from the opening weekend gross sales, total gross sales, and number of theaters variables. The median weeks in the top sixty of 7, indicating 50% of the weeks in the top sixty values were less than 7 and 50% were greater than 7, would be a more appropriate central tendency measure than the mean due to this skewness (0.67). The range of the weeks in the top sixty was 26, from 1 (fourteen different films) to 27 (The Wild Parrots of Telegraph Hill), and the standard deviation was 6.39 ( $\hat{\sigma} = 40.82586$ ). However, the interquartile range of 10 would be a more appropriate measure of variability because the distribution is skewed. There was not any weeks in the top sixty outliers observed using the box-plot method or z-score criteria.

Relationships Using a sample of 15 motion pictures, the Pearson correlation between total gross sales and opening weekend sales is 0.969<sup>1</sup>, indicating a high positive correlation. This implies that motion pictures with high total gross sales tend to have high opening weekend gross sales, or movies with low total gross sales tend to have low opening weekend gross sales. The Pearson correlation between total gross sales and the number of theaters is 0.761<sup>2</sup>, indicating,

again, a high positive correlation. Thus, motion pictures with high total gross sales tend to have been shown in more theaters, or movies with low total gross sales tend to have been shown in less theaters. The Pearson correlation between total gross sales and weeks in the top sixty is 0.508<sup>3</sup>, indicating a medium positive correlation. This shows us that motion pictures with high total gross sales tend to have been in the top sixty longer than movies with low total gross sales. The competitiveness of the movie industry continues to grow as numerous studios produce more and more motion pictures each year. The financial success of each movie varies considerably and can be measured using the opening weekend gross sales, the total gross sales, the number of theaters, and the number of weeks in the top sixty variables; also, one can use descriptive statistics to discover correlations between each of these variables, which can ultimately help the studios when deciding to produce and release movies. Calculations:  $r = \frac{\sum XY - (\sum X \sum Y)/N}{\sqrt{(\sum X^2 - (\sum X)^2/N)(\sum Y^2 - (\sum Y)^2/N)}}$   $\sum X = 1,104.070$   $\sum Y = 349.950$   $\sum X^2 = 187,369.754$   $\sum Y^2 = 18,851.324$   $\sum XY = 58,402.566$   $r = \frac{58,402.566 - (1,104.070 \times 349.950)/15}{\sqrt{(187,369.754 - (1,104.070^2/15))(18,851.324 - (349.950^2/15))}}$   $r = \frac{32,644.613}{\sqrt{1,133,943,624.961}}$   $r = 0.969$   $\sum X = 1,104.070$   $\sum Y = 31,524$   $\sum X^2 = 187,369.754$   $\sum Y^2 = 101,879,282$   $\sum XY = 3,799,746.170$   $r = \frac{3,799,746.170 - (1,104.070 \times 31,524)/15}{\sqrt{(187,369.754 - (1,104.070^2/15))(101,879,282 - (31,524^2/15))}}$   $r = \frac{1,479,432.658}{\sqrt{3,780,357,774,398.713}}$   $r = 0.761$   $\sum X = 1,104.070$   $\sum Y = 176$   $\sum X^2 = 187,369.754$   $\sum Y^2 = 2,786$   $\sum XY = 17,394.930$   $r = \frac{17,394.930 - (1,104.070 \times 176)/15}{\sqrt{(187,369.754 - (1,$

$104.070^2/15) (2,786 - (176^2/15))]$   $r = 4,440.509 \hat{\sigma}^2 76,494,667.072$   $r = 0.508$