Mini-Project 3-3

QNT/275

Mini-Project 3-3

In this assignment, a movie data set was provided. This paper will provide a calculation including mean, SD, 5 number summary and interquartile range. In addition, the paper will analyze the genre that has a greater variability in the terms of income. Finally, the paper will include plot of movie’s length of time.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Action | Adventure | Animation | Biography | Comedy | Crime | Documentary | Drama | Horror |
| Mean |  93,602,863  |  122,371,299  |  94,178,308  |  31,174,421  |  28,653,208  |  26,341,667  |  6,417,583  |  24,463,030  |  31,619,028  |
| Standard Deviation |  93,305,573  |  104,675,749  |  82,412,674  |  32,675,664  |  33,263,412  |  16,671,950  |  6,294,492  |  36,532,450  |  26,121,782  |
| Minimum |  1,558,836  |  2,589,811  |  2,553,002  |  2,107,925  |  1,359,497  |  3,322,127  |  1,514,782  |  1,359,910  |  2,756,333  |
| Quartile 1 |  19,736,060  |  45,058,978  |  50,151,543  |  4,721,703  |  3,461,509  |  13,757,804  |  1,539,680  |  2,641,241  |  13,993,563  |
| Quartile 2 |  58,607,007  |  113,729,151  |  64,251,541  |  22,592,566  |  10,616,104  |  26,307,600  |  2,799,318  |  12,846,116  |  29,136,184  |
| Quartile 3 |  130,800,900  |  169,483,956  |  131,538,435  |  44,755,349  |  45,216,161  |  33,680,992  |  10,422,129  |  26,412,730  |  37,060,782  |
| Maximum |  350,126,372  |  337,135,885  |  257,760,692  |  115,637,895  |  150,157,400  |  52,543,632  |  17,780,194  |  167,767,189  |  84,273,813  |
| Inter Quartile Range |  111,064,841  |  124,424,979  |  81,386,892  |  40,033,646  |  41,754,652  |  19,923,188  |  8,882,449  |  23,771,490  |  23,067,219  |

The average total gross earning of adventure genre that has a value of 122,371,199 has the highest variability among the styles that include action, comedy, drama, horror, biography, and documentary. The average gross income of animation and action follow afterward. They have respective values of 94,178,308 and 93,602,863. The genre that has the lowest average gross income is the documentary that recorded a value of 6,417,583. In the analysis of the median gross income of the genres, adventure still provided the highest value of 113,729,151. The genre that follows adventure regarding the median total gross earnings is animation with a value of 64,251,541. The third placed genre is the action that had median gross earnings of 58,607,007. Interestingly, the documentary genre is still last with a value of 2,799,318.

In the above case, the information regarding the total gross income of all the genres skew rightwards. As such, it is an indication that each of the genres has few movies that have higher levels of total gross income. Regarding the maximum and the minimum, the top-ranked genre of adventure has 337,135,885 and 2,589,811 respectively. On the other side, the documentary genre has a maximum value of 17,780,194 and a minimum total gross income of 1,514,782 show that is ranked lowest in almost every component analyzed.

Therefore, the adventure genre is observed to have the highest level of variability regarding the total gross income. Notably, the standard deviation and interquartile range often measure the dispersion with standard deviation measuring the dispersion in right units, (Gravetter and Wallnau, 2009). The first quartile is usually subtracted from the third quartile (Q3-Q1) to get the interquartile range. As such, these values of interquartile range and standard deviation which are high because they show that the adventure genre has the highest level of variability. The earlier explained values also show that the adventure genre has the highest degree of variability.





In this case, a film of each genre will be analyzed to determine its length, which would help in finding out possible differences the lengths across the given genres as well as determine the presence of any outliers.

The use of Boxplot usually indicates symmetry. Therefore, it implies that when the line is nearer to the center of the box, and the whisker length being same, then the sample was generated from a symmetric population, (Welkowitz et al., 2006). In case the top whisker has a longer length compared to the bottom with the line moving towards the bottom, then it implies that sample is got from a rightwards skewed population, (Triola, 2005). On the other side, a sample generated from a population skewed towards the left is characterized with a longer bottom whisker than the top and line moving towards the top of the box.

In this situation, the lengths of movies of horror, crime, comedy, action, drama, and documentary are skewed to the right showing that there are fewer movies of these genres that run for several minutes. On the other hand, the adventure, animation, and biography movies are skewed to the left shows that few movies have large lengths. Interestingly, all the movies are skewed and do not follow normal distribution showing that the best measure of central tendency would be the use of median.

References

Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth.

Triola, M. F. (2005). *Elementary statistics*. Boston: Pearson/Addison-Wesley.

Welkowitz, J., Cohen, B. H., & Ewen, R. B. (2006). *Introductory statistics for the behavioral sciences*. Hoboken, NJ: Wiley.