**List Of Appendices**

This Dataset displays Top 1000 Highest Grossing Movies of All Time IMBD and contains following columns

**Movie Title**: The name of the movie.

**Year of Release**: The year the movie was released.

**Genre**: Categories where the movie belongs.

**Movie Rating**: Ratings given by IMDb registered users (on a scale of 1 to 10)

**Duration**: Movie running time in minutes.

**Gross:** Gross earnings in U.S. dollars.

**Worldwide LT Gross**: Worldwide Lifetime Gross (International + Domestic totals.

**Metascore**: Weighted average of many reviews coming from reputed critics (on a scale of 0 to 100)

**Votes**: Number of votes cast by IMDb registered users.

**Logline**: A one or two sentence summary of the film.

**Appendix 1**

µ = 6.8

σ = 0.9

X = Random variable

**Appendix 2**

= Sample Mean

E = Margin of Error

where is the critical value for the level of confidence, C=1−α, such that the area under the standard normal distribution to the right of is equal to , σ is the population standard deviation, and n is the sample size.

= 117.67, = 22, n = 1000

= 0.025

= 1.96

**Appendix 3**

t =

T-test statistic is appropriate to use in this case because the claim is about a population mean, the population is normally distributed, the population standard deviation is unknown, and the sample is a simple random sample

**Appendix 4 & 5**

Table 4.1

|  |  |
| --- | --- |
| *Regression Statistics* | |
| Multiple R | 0.663601159 |
| R Square | 0.440366498 |
| Adjusted R Square | 0.439243863 |
| Standard Error | 0.675038629 |
| Observations | 1000 |

Table 4.2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 2 | 357.4892412 | 178.7446206 | 392.2615399 | 2.1379E-126 |
| Residual | 997 | 454.3101188 | 0.45567715 |  |  |
| Total | 999 | 811.79936 |  |  |  |

p-value for the regression equation is found under the "Significance F" column,

Since P-value , 2.1379E-126, is less than the alpha given, 0.05, we can say that there is sufficient evidence at the 0.01 level of significance to conclude that the linear relationship between the independent variables and the dependent variable is statistically significant.