Technology Project 1

Exploring and Understanding Data

Group Members: 1. _____

2. _____

3. _____

We are going to describe the movies currently playing at the box office. Follow the Minitab instructions under the technology projects link on the website to collect the information and for help finishing the report.

1. Fully describe the data collected using the Who, What, Where, When, Why, and How (if appropriate) as covered in chapter 2 of the text.
2. Display the categorical data for the "critic", "yahoo", and "rating" variables. Use a frequency table, bar chart, and pie chart (you pick which one to use with which variable).
3. Create a contingency table between the "fresh" and "rating" variables. Make it a marginal distribution where the percent is based on the rating (that is, what percent of each rating is fresh and what percent is rotten)?
4. Display the quantitative data for the weekend revenues, number of theaters, length of movie, number of weeks in release, and number of tomatoes. Use a dot plot, histogram, and stem and leaf plot at least once each. You decide which graph is appropriate for which variable.
5. Pick one of the quantitative variables just described that is skewed to the right and apply a logarithmic transformation to the data. Pick a graph and show the data after the transformation.
6. Create a box plot of the number of theaters showing a movie, broken down by the MPAA rating of the movie.
7. Numerically describe the weekend revenues, weeks in release, number of theaters showing the movie, and number of tomatoes.
8. Create a normal probability plot for the weekend revenues. Tell about whether or not the data appear normal? If not run a Box-Cox Transformation to transform your data and then generate a normal probability plot for the transformed data. Comment on the normality of the transformed data.
9. Standardize the number of theaters and then describe that variable. In particular comment on the mean and the standard deviation of the standardized variable.