Project Preparation

1. Data Cleaning---kunal
2. Nidhi and foram:

* Plots – Histogram for numerical, barplots for categorical data
* Distribution of hit/not hit movies for relevant genre.
* Count for different genre hit/not hit.
* Scatter plots/box plots
* Distribution of revenue to budget issue
* Heatmap for correlation

1. Modelling---anjali

* Logistic Regression
* Decision Tree
* KNN
* Random Forest

1. Comparing and Selecting Best Algorithm Heatmap for correlation---kunal and anjali

* Performance of various models

(F1, accuracy, senisitivity, aoc)