Movie Analysis Project

By: Sibi Tiruchi

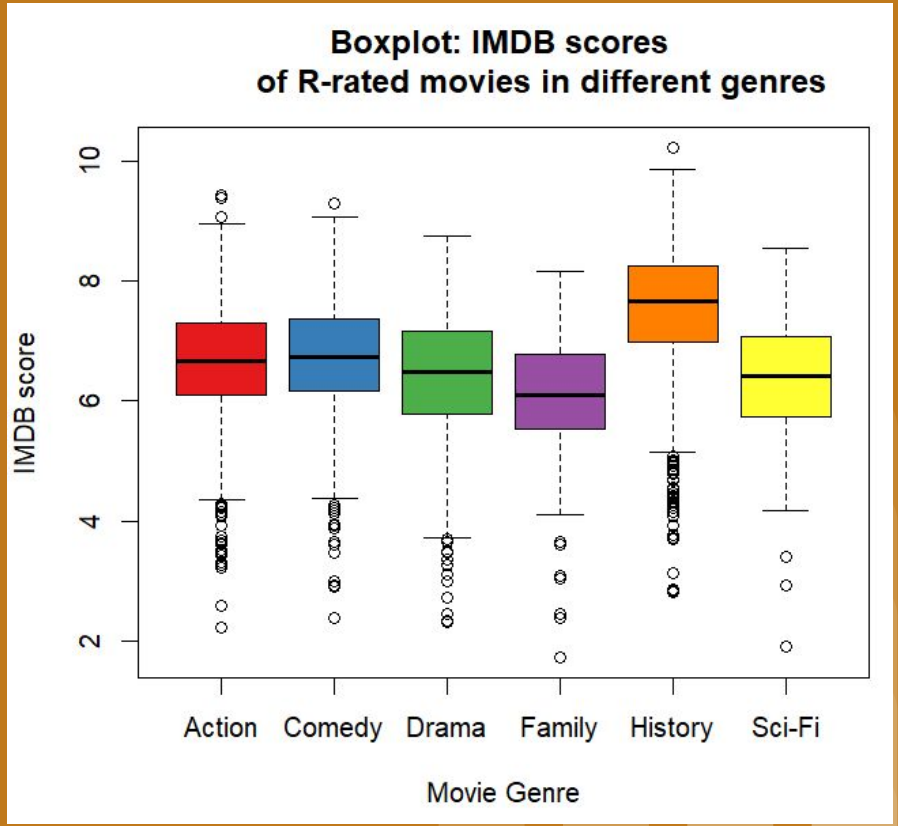
Part A

Hypothesis #1

Hypothesis: R-rated History movies have higher imdb scores than other genres.

The boxplot indicates that the median line is much higher for History R-rated movies than for other genres.

This indicates that R-rated movies in the History typically do better.



Hypothesis #1

Assessing the actual means displays that R-rated History is the only genre with an average IMDB score above a 7

```
> tapply(moviesR$imdb_score, moviesR$genre, mean)
  Action   Comedy    Drama   Family  History   Sci-Fi
6.638439 6.725157 6.411240 6.145942 7.512010 6.366116
```

Hypothesis #2

Hypothesis: For Sci-Fi movies, a higher budget translates to a higher imdb score (directly proportional)

After looking at the mean IMDB scores for High budget, Medium budget, and Low budget Sci-Fi movies, we can see that the budget gives a slight edge to the mean score although it is not significant.

```
> mean(sciFiHi$imdb_score)
[1] 6.495592
> mean(sciFiMed$imdb_score)
[1] 6.255022
> mean(sciFiLo$imdb_score)
[1] 6.111266
```

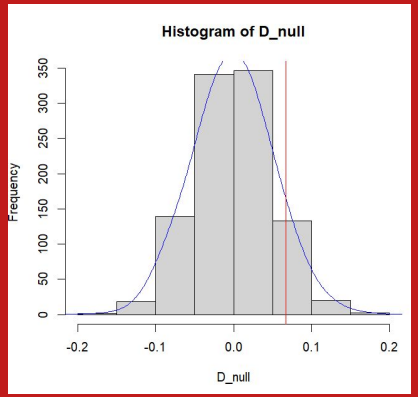
Hypothesis #3

Hypothesis: High grossing Family movies have higher IMDB reviews than high grossing Comedy movies on average.

The subframe ‘familyHighGross’ has the all family movies that are high grossing. The subframe ‘comedyHighGross’ has the all family movies that are high grossing. We see that mean IMDB scores are higher for Family movies, verifying that the hypothesis is true.

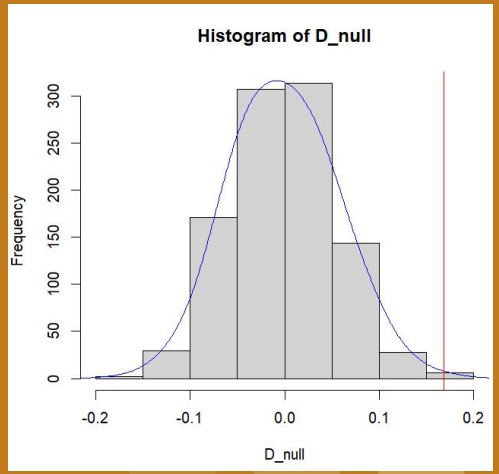
```
> mean(familyHighGross$imdb_score)
[1] 7.648296
> mean(comedyHighGross$imdb_score)
[1] 6.771587
```

Part B


Histogram of D_null
Frequency
0
50
100
150
200
250
300
350
-0.2
-0.1
0.0
0.1
0.2
D_null

Hypothesis 1

Alternative Hypothesis (H_a): Sci-Fi movies have an average imdb score higher than Action movies.
Null Hypothesis (H_o): Sci-Fi movies have the same average imdb score as Action movies.

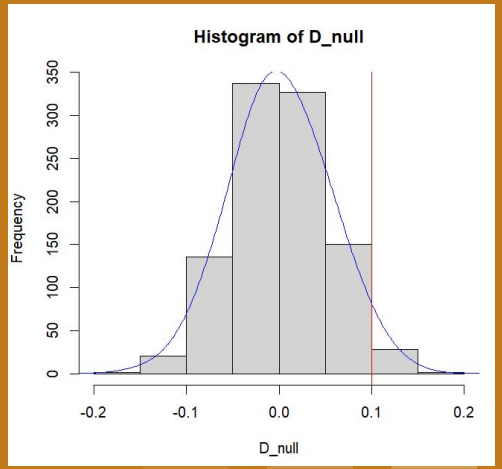


P-Value: 0.001
0.001 < 0.05 → Reject H_o

```
> PermutationTestSecond::Permutation(d, "Cat", "Val", 1000, "Sci-Fi", "Action")
[1] 0.001
```

Hypothesis 2

Alternative Hypothesis (H_a): Sci-Fi movies have an average imdb score higher than Comedy movies.
Null Hypothesis (H_o): Sci-Fi movies have the same average imdb score as Comedy movies.

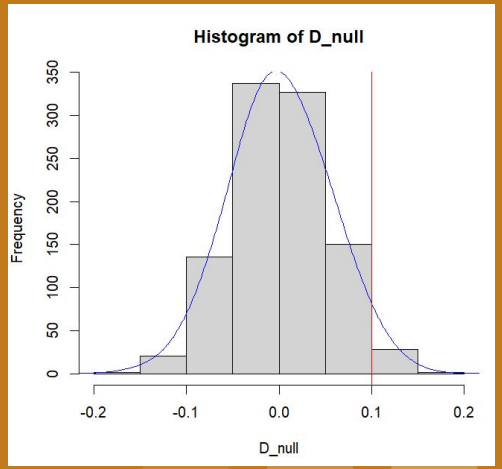


P-Value: 0.0
0.029 < 0.05 → Reject H_o

```
> PermutationTestSecond::Permutation(d, "Cat", "Val", 1000, "Sci-Fi", "Comedy")
[1] 0.029
```

Hypothesis 3

Alternative Hypothesis (H_a): Comedy movies have an average imdb score higher than Action movies.
Null Hypothesis (H_o): Comedy movies have the same average imdb score as Action movies.



P-Value: 0.084
0.084 > 0.05 → Do Not Reject H_o

```
> PermutationTestSecond::Permutation(d, "Cat", "Val", 1000, "Comedy", "Action")
[1] 0.084
```