
title: "MovieRatingAnalysis"

author: "17F-8075"

date: "11/19/2021"

output: pdf_document

```
``{r setup, include=FALSE}
```

```
knitr::opts_chunk$set(echo = TRUE)
```

```
``
```

```
``{r}
```

```
library(ggplot2)
```

```
library(ggthemes)
```

```
library(scales)
```

```
library(dplyr)
```

```
library(lessR)
```

```
``
```

-extracting data from .dat file where separator is ':'

-making extracted data a dataframe

-displaying its top 6 rows

```
``{r}
```

```
MOVIE_DATASET = do.call(rbind, strsplit(readLines('C:/Users/Talha/Downloads/movies.dat'), '::', fixed=T))
```

```
MOVIE_DATASET <- as.data.frame(MOVIE_DATASET)
```

```
head(MOVIE_DATASET)
```

```
``
```

-assigning column names to dataframe

-making ID column numeric

-displaying its top 6 rows

```
``{r}
```

```
colnames(MOVIE_DATASET) <- c("ID", "Title", "Genre")
```

```
MOVIE_DATASET$ID <- as.numeric(MOVIE_DATASET$ID)
```

```
head(MOVIE_DATASET)
```

```
...
```

-extracting data from .dat file where separator is ':'

-displaying its top 6 rows

```
``{r}
```

```
Rating_of_movies = read.delim("ratings.dat", header= FALSE ,sep = ':', colClasses = c(NA, "NULL"))
```

```
head(Rating_of_movies)
```

```
...
```

-assigning column names to dataframe

-making ID column numeric

-displaying its top 6 rows

```
``{r}
```

```
colnames(Rating_of_movies) <- c("User", "ID", "Rating_of_movies", "Timestamp")
```

```
head(Rating_of_movies)
```

```
...
```

-merging dataframes based on ID

-displaying its top 6 rows

```
``{r}
```

```
data = merge(MOVIE_DATASET, Rating_of_movies, by = "ID")
```

```
data = data[with(data, order(ID)),]
```

```
head(data)
```

```
```
```

-ploting pie chart

```
```{r}
```

```
Rating_of_movies<- dplyr::count(data, Rating_of_movies, sort = TRUE) %>%
```

```
  mutate(perc = `n` / sum(`n`)) %>%
```

```
  arrange(perc) %>%
```

```
  mutate(labels = scales::percent(perc))
```

```
ggplot(Rating_of_movies, aes(x = "", y = perc, fill = factor(Rating_of_movies)),alpha = 0.8) +
```

```
  geom_col() +
```

```
  geom_text(aes(label = labels),
```

```
    position = position_stack(vjust = 0.5)) +
```

```
  coord_polar(theta = "y")
```

```
```
```