```r
#Netflix Original data
# Import Libraraies
library(ggplot2)
  library(dplyr)
  library(choroplethr)
  library(choroplethrMaps)
  library(openintro)
  library(tidyverse)
  library(scales)
#Data Cleaning
getwd()
setwd('C:/Users/admin/Downloads')
getwd()
Dataset=read.csv('NetflixOriginals.csv')
Dataset
View(Dataset)
print(Dataset)
print(head(Dataset))
print(tail(Dataset))
print(summary(Dataset))
print(str(Dataset))
run=max(Dataset$Runtime)
run
print(glimpse(Dataset))
fil=filter(Dataset, Language=='English')
fill
#col-wise sum
apply_sum_col=apply(Dataset[,4:5],2,sum)
apply_sum_col
#row-wise sum
apply_sum_row=apply(Dataset[,4:5],1,sum)
apply_sum_row
#unique value
uni=unique(Dataset$Title)
uni
#Count of unique value
uni1=length(unique(Dataset$Runtime))
uni1

#Sort
sorr=arrange(Dataset,desc(Runtime))
sorr
sor=head(sorr)
sor
#summary
View(Dataset)
summary(group_by(Dataset,Genre,Avg=mean(Runtime)))

arrangeWindows()
#filterwise groupby
filt=Dataset%>% group_by(Genre)%>%filter(Genre=='Thriller' &
Language=='Hindi')
print(filt)
```

```
filt_id=Dataset%>% group_by(Genre)%>%filter(Genre=='Horror' & IMDB.Score>=5.0)
print(filt_id)


#groupwise Language wise sum of runtime
Name=Dataset %>% group_by(Language) %>% summarise(Run=sum(Runtime)) %>%
arrange((desc(Run)))
View(Name)

#groupwise Language wise sum of IMDB.Score
idb=Dataset %>% group_by(Language) %>% summarise(Run=max(IMDB.Score)) %>%
arrange((desc(Run)))
print(idb)


#Title wise groupby
Name1=Dataset %>% group_by(Title) %>% summarise(Run=max(Runtime)) %>%
arrange((desc(Run)))
na=print(head(Name1))
View(head(Name1))
#Title wise groupby of IMDB.Score
ID=Dataset%>%group_by(Genre)%>%summarise(idb=max(IMDB.Score))%>%
arrange((desc(idb)))
ID


# Conclusions
# Min. runtime of movie is: 4.00and min score(rating) is 2.50 and length of
languages i 584 and all are character datatype
# Sum of runtime is 54649 and rating sum is 3662.7
# There is 124 unique title in the dataset.
# The highest runtime is 209 of the movie title The Irishman, genre: Crime
drama relaise on the premiere Nov 27, 2019 with the score 7.8 language
English.
# The highest hindi Thriller movie runtime of  Raat Akeli Hae is 149 with the
score rating 7.3 realised on jul3=y 31, 2020
# The horrow hindi movie bulbul got the highest rating in netflix i.e. 6.6
with runtime 94
# The highest sum of runtime in netflix is of english language with the sum of
runtime 36819 and max rating is 9.
# Genre "Documentary" has the highest score i.e. 9

#Data Visualizations

# ggplot of Language wise Runtime analysis

d=ggplot(Dataset,aes(x=Language,y=Runtime,fill=Language))+geom_col()
+ggtitle("Language wise analysis")+coord_flip()+theme(legend.position =
"None",axis.text.y =element_text(size=6))
print(d)
# ggplot of Language wise IMDB.Score analysis
d=ggplot(idb,aes(x=Run,y=Language,fill=Language))+geom_boxplot(alpha=0)
+geom_jitter(alpha = 0.3, color = "tomato") + xlab("Rating") +
ylab("Language")+ ggtitle("Language wise analysis of IMDB.Score")
```

```
print(d)



#ggplot of Title wise ploting

d1=ggplot(na,aes(x=Title,y=Run,fill=Title))+geom_col()+ggtitle("Title wise
analysis")+coord_flip()+theme(legend.position = "None",axis.text.y
=element_text(size=6))
print(d1)

d1=ggplot(na,aes(x=Title,y=Run,group=1))+geom_violin(fill="yellow", alpha=0.2)
+ggtitle("Title wise analysis")+
  xlab("Year")+ylab("Lowest Pressure")
print(d1)

#bar chart

barplot(ID$idb,names.arg = ID$Genre,col='green',xlab = 'Genre',density=50,ylab
= 'idb Score',main='Bar Plot for score wise Genre')



#histogram
hist(Dataset$Runtime, breaks=15,col = 'green',density = 15,xlab='Runtime')
#Scatter plot
plot(sor$Runtime,sor$IMDB.Score,main='scatter
plot',xlab='Runtime',ylab='IMDB.Score')
#pie chart
d1=head(Dataset)
d1
pie(d1$Runtime,main = "Pie Chart for Runtime",col =
rainbow(length(d1$Runtime)))

# Conclusion
# Top 5 highest runtime movie in netflix is:- The Irishman: 209, The forest of
love: 150, Springsteen on Broadway: 155, Da5 Bloods: 155, Citation 150 and
ludo is 148 runtime movie in netflix.
# Top score of netflix movie is English:9,Portuguese: 8.7, English/Ukranian/
Russia: 8.4, Spanish 8.3
#The English lanuguage movie has the highest runtime movie in netflix
# genre documentary has the highest runtime
```