#Adding Libraries and Importing Data

library(tidyverse)

library(lubridate)

library(dplyr)

library(ggplot2)

library(janitor)

library(stringr)

naruto<- read\_csv("E:/DA\_Course/Naruto/Dataset/naruto\_ratings\_data\_v04.csv")

#Cleaning Data

naruto<-clean\_names(naruto)

naruto <- subset(naruto, select=-c(2,3,4,7,8,11))

n\_distinct(naruto$episode\_number\_overall)

str(naruto)

naruto$episode\_number\_overall<- as.character(naruto$episode\_number\_overall)

write.csv(naruto,file="E:/DA\_Course/Naruto/Dataset/naruto1.csv") #Exporting to split directors and writers columns

naruto<-read\_csv("E:/DA\_Course/Naruto/Dataset/naruto.csv")

#Analyzing Data

naruto$weight<- ((naruto$votes)/sum(naruto$votes))

naruto$fr<- (naruto$rating\*naruto$weight)

naruto$fr<- naruto$fr\*100

naruto$fr<-round(naruto$fr,2)

summary(naruto$fr)

naruto <- subset(naruto, select=-c(7))

write.csv(naruto,file = "E:/DA\_Course/Naruto/Report/for\_viz.csv")

n\_distinct(naruto$writer)

top\_n(group\_by(naruto$director),10)

test <-top\_n(naruto %>%

group\_by(naruto$director) %>%

summarise(avgFR = mean(fr)),20) #Taking Top 20 directors cause its hard to visualize 108

write.csv(test, file= "E:/DA\_Course/Naruto/Report/topdir.csv")