library(tidyverse)

library(scales)

setwd("C:/Users/di11m/Downloads/Coursera/Data Analysis/RStudio")

#Election Data .csv files: Washoe and NV

washoe <- read.csv("Election\_Results\_Washoe\_CLEANED.csv"); View(washoe)

nevada <- read.csv("Election\_Results\_NV\_CLEANED.csv"); View(nevada)

#Sum of votes by candidate: Washoe and NV

bidenElecW <- filter(washoe, Candidate == "BIDEN, JOSEPH R.") #Washoe

bidenTotalW <- sum(bidenElecW$Votes)

bidenElecN <- filter(nevada, Candidate == "BIDEN, JOSEPH R.") #State

bidenTotalN <- sum(bidenElecN$Votes)

blankElecW <- filter(washoe, Candidate == "BLANKENSHIP, DON")

blankTotalW <- sum(blankElecW$Votes)

blankElecN <- filter(nevada, Candidate == "BLANKENSHIP, DON")

blankTotalN <- sum(blankElecN$Votes)

jorgElecW <- filter(washoe, Candidate == "JORGENSEN, JO")

jorgTotalW <- sum(jorgElecW$Votes)

jorgElecN <- filter(nevada, Candidate == "JORGENSEN, JO")

jorgTotalN <- sum(jorgElecN$Votes)

noneElecW <- filter(washoe, Candidate == "None Of These Candidates")

noneTotalW <- sum(noneElecW$Votes)

noneElecN <- filter(nevada, Candidate == "None Of These Candidates")

noneTotalN <- sum(noneElecN$Votes)

trumpElecW <- filter(washoe, Candidate == "TRUMP, DONALD J.")

trumpTotalW <- sum(trumpElecW$Votes)

trumpElecN <- filter(nevada, Candidate == "TRUMP, DONALD J.")

trumpTotalN <- sum(trumpElecN$Votes)

#Percentage of votes by candidate: Washoe and NV

candidate <- c("BIDEN", "BLANKENSHIP", "JORGENSEN", "NONE", "TRUMP")

votesTotalW <- bidenTotalW+blankTotalW+jorgTotalW+noneTotalW+trumpTotalW

votesTotalN <- bidenTotalN+blankTotalN+jorgTotalN+noneTotalN+trumpTotalN

percentTotalW <- data.frame("Candidate" = candidate, "Percent" = c(bidenTotalW/votesTotalW, blankTotalW/votesTotalW, jorgTotalW/votesTotalW, noneTotalW/votesTotalW, trumpTotalW/votesTotalW))

percentTotalN <- data.frame("Candidate" = candidate, "Percent" = c(bidenTotalN/votesTotalN, blankTotalN/votesTotalN, jorgTotalN/votesTotalN, noneTotalN/votesTotalN, trumpTotalN/votesTotalN))

percentTotalW %>% gt #data table Washoe

percentTotalN %>% gt #data table NV

#Graph of Election Results: Washoe

electionTableW <- data.frame(candidate,TotalVotesW = c(bidenTotalW, blankTotalW, jorgTotalW, noneTotalW, trumpTotalW))

electionGraphW <-

ggplot(electionW, aes(x=candidate, y=TotalVotesW, fill= candidate))+

geom\_col()+

theme(legend.position = "none")+

scale\_y\_continuous(breaks = scales::pretty\_breaks(n = 6))+

scale\_x\_discrete(guide = guide\_axis(n.dodge = 2))+

scale\_fill\_manual(values = c("BIDEN" = "blue", "BLANKENSHIP" = "purple", "JORGENSEN" = "yellow", "NONE" = "black", "TRUMP" = "RED"))+

labs(title="2020 Presidential Election Results: Washoe County", x="Candidate", y="Votes")

#Graph of Election Results: NV

electionTableN <- data.frame(candidate,TotalVotesN = c(bidenTotalN, blankTotalN, jorgTotalN, noneTotalN, trumpTotalN))

electionGraphN <-

ggplot(electionN, aes(x=candidate, y=TotalVotesN, fill= candidate))+

geom\_col()+

theme(legend.position = "none")+

scale\_y\_continuous(breaks = scales::pretty\_breaks(n = 6), labels = label\_number())+

scale\_x\_discrete(guide = guide\_axis(n.dodge = 2))+

scale\_fill\_manual(values = c("BIDEN" = "blue", "BLANKENSHIP" = "purple", "JORGENSEN" = "yellow", "NONE" = "black", "TRUMP" = "RED"))+

labs(title="2020 Presidential Election Results: Nevada", x="Candidate", y="Votes")

#Graph of Benford's Law distribution

benlaw <- function(d) log10(1 + 1 / d)

digits <- 1:9

bensData <- data.frame(Digits = digits, Percentage = benlaw(digits))

bensGraph <- ggplot(bensData, aes(x= Digits, y=Percentage)) +

geom\_line(color="red")+

geom\_point(color="red", size=3)+

scale\_x\_continuous(n.breaks = 9) +

scale\_y\_continuous(n.breaks = 8) +

labs(title = "Benford's Law", x = "Leading Digit", y= "Proportion of Occurrence")

#Leading Digits: Washoe

ldOccurW <- c(sum(washoe$LeadDigit == 1), sum(washoe$LeadDigit == 2), sum(washoe$LeadDigit == 3), sum(washoe$LeadDigit == 4), sum(washoe$LeadDigit == 5), sum(washoe$LeadDigit == 6), sum(washoe$LeadDigit == 7), sum(washoe$LeadDigit == 8), sum(washoe$LeadDigit == 9));

ldTotalW <- sum(ldOccurW)

ldPercentW <- c(ldOccurW[1]/ldTotalW,

ldOccurW[2]/ldTotalW,

ldOccurW[3]/ldTotalW,

ldOccurW[4]/ldTotalW,

ldOccurW[5]/ldTotalW,

ldOccurW[6]/ldTotalW,

ldOccurW[7]/ldTotalW,

ldOccurW[8]/ldTotalW,

ldOccurW[9]/ldTotalW)

ldDataW <- data.frame(Digits = digits, Percentage = ldPercentW)

ldGraphW <- ggplot(ldDataW, aes(x=Digits, y=Percentage))+

geom\_line(color="blue")+

geom\_point(color="blue", size=3)+

scale\_x\_continuous(n.breaks = 9) +

scale\_y\_continuous(n.breaks = 8) +

labs(title = "Leading Digit of Total Votes\nin Washoe Presidential Election", x= "Leading Digit", y= "Proportion of Occurrence")

#Leading Digits: NV

ldOccur <- c(sum(nevada$LeadDigit == 1), sum(nevada$LeadDigit == 2), sum(nevada$LeadDigit == 3), sum(nevada$LeadDigit == 4), sum(nevada$LeadDigit == 5), sum(nevada$LeadDigit == 6), sum(nevada$LeadDigit == 7), sum(nevada$LeadDigit == 8), sum(nevada$LeadDigit == 9));

ldTotal <- sum(ldOccur)

ldPercent <- c(ldOccur[1]/ldTotal,

ldOccur[2]/ldTotal,

ldOccur[3]/ldTotal,

ldOccur[4]/ldTotal,

ldOccur[5]/ldTotal,

ldOccur[6]/ldTotal,

ldOccur[7]/ldTotal,

ldOccur[8]/ldTotal,

ldOccur[9]/ldTotal)

ldTable <- data.frame(Digits = digits, Percentage = ldPercent)

ldGraph <- ggplot(ldTable, aes(x=Digits, y=Percentage)) +

geom\_line(color="orange")+

geom\_point(color="orange", size=3)+

scale\_x\_continuous(n.breaks = 9) +

scale\_y\_continuous(n.breaks = 8) +

labs(title = "Leading Digit of Total Votes\nin Nevada Presidential Election", x= "Leading Digit", y= "Proportion of Occurrence")

#Benford's Law comparison: Washoe

tableWB <- data.frame(Digits=digits,

Percentage=c(benlaw(digits), ldPercentW),

Focus = c(rep("Benford",9), rep("Washoe",9)))

graphWB <- ggplot(tableWB, aes(x=Digits, y=Percentage, group=Focus)) +

geom\_line(aes(color=Focus))+

geom\_point(aes(color=Focus))+

scale\_x\_continuous(n.breaks = 9)+

scale\_y\_continuous(n.breaks = 8)+

scale\_color\_manual(values=c("red", "blue"))+

labs(title = "Leading Digits of Votes in Washoe v. Benford", x= "Leading Digit", y= "Proportion of Occurrence")

#Benford's Law comparison: NV

tableNB <- data.frame(Digits=digits,

Percentage=c(benlaw(digits), ldPercent),

Focus = c(rep("Benford",9), rep("Nevada",9)))

graphNB <- ggplot(tableNB, aes(x=Digits, y=Percentage, group=Focus)) +

geom\_line(aes(color=Focus))+

geom\_point(aes(color=Focus))+

scale\_x\_continuous(n.breaks = 9)+

scale\_y\_continuous(n.breaks = 8)+

scale\_color\_manual(values=c("red", "orange"))+

labs(title = "Leading Digits of Votes in Nevada v. Benford", x= "Leading Digit", y= "Proportion of Occurrence")

#Benford's Law comparison: NV, Washoe, Benford

tableAll <- data.frame(Digits = digits,

Percentage = c(benlaw(digits), ldPercentW, ldPercent),

Focus = c(rep("Benford", 9), rep("Washoe", 9), rep("Nevada", 9)))

graphAll <- ggplot(tableAll, aes(x=Digits, y=Percentage, group=Focus)) +

geom\_line(aes(color=Focus))+

geom\_point(aes(color=Focus))+

scale\_x\_continuous(n.breaks = 9)+

scale\_y\_continuous(n.breaks = 8)+

scale\_color\_manual(values=c("red", "orange", "blue"))+

labs(title = "Washoe and Nevada v. Benford", x= "Leading Digit", y= "Proportion of Occurrence")