Tian Qiu

00265-35063

1. 26 variables.

Categorical or qualitative variables: Year, Month, DayOfMonth, DayOfWeek, DepTime, ArrTime, UniqueCarrier, FlightNum, Origin, Dest.

Quantitative or numeric variables: rest of the variable.

2. What is the standard deviation of Distance in this dataset?

What is the range of distance in this dataset?

1. airline2008NovS <- read.delim("~/Desktop/STAT350/STAT350/Labs/Lab1/airline2008NovS.txt")
2. There are 3 missing values (NA) in the dataset.

(a)

cleaned <-airline2008NovS[complete.cases(airline2008NovS),]

dim(cleaned)

View(cleaned)

write.table(helicon\_cleaned,file="airCleaned.txt",quote=F,

row.names=F, sep="\t")

(b) 9997 observations.

(c) ~/Desktop/STAT350/STAT350/Labs/Lab1

5. (a)

airline2008NovS\_new <- airline2008NovS\_cleaned

airline2008NovS\_new$Dest <- as.character(airline2008NovS\_new$Dest)

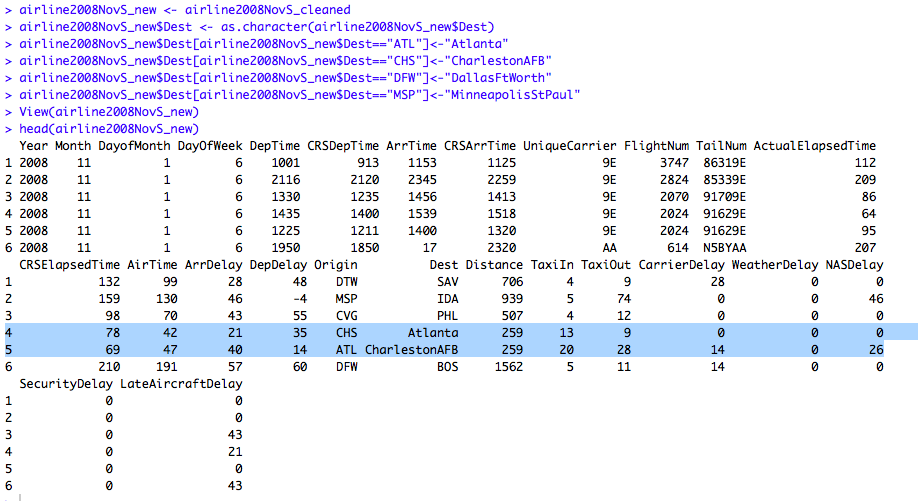
airline2008NovS\_new$Dest[airline2008NovS\_new$Dest=="ATL"]<-"Atlanta"

airline2008NovS\_new$Dest[airline2008NovS\_new$Dest=="CHS"]<-"CharlestonAFB"

airline2008NovS\_new$Dest[airline2008NovS\_new$Dest=="DFW"]<-"DallasFtWorth"

airline2008NovS\_new$Dest[airline2008NovS\_new$Dest=="MSP"]<-"MinneapolisStPaul"

View(airline2008NovS\_new)

(b) 

6. (9 points) We are going to see if the variable "ActualElapsedTime" can be calculated from other variables in the data set.

1. (3 pts.) Write down a mathematical equation to calculate "ActualElapsedTime" from “AirTime”, “TaxiIn” and “Taxiout”.

ActualElapsedTime = AirTime + TaxiIn +Taxiout

1. (3 pts.)

airline2008NovS\_new$ActualElapsedTime = airline2008NovS\_new$AirTime + airline2008NovS\_new$TaxiIn +airline2008NovS\_new$TaxiOut

head(airline2008NovS\_new)

1. (3 pts.) 