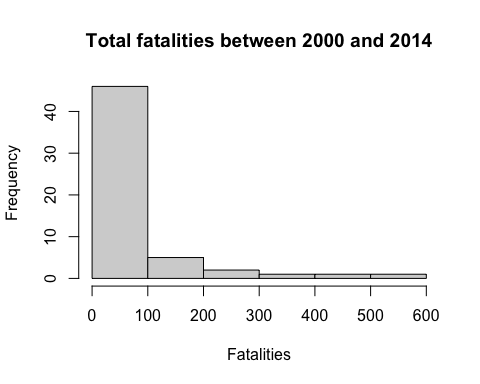
Week 5 Visualizations Activity

Iryna Barvi

4/17/2022

# Section 2: Reading the data into R. (5 points)  
airlines <- "https://raw.githubusercontent.com/fivethirtyeight/data/master/airline-safety/airline-safety.csv"  
airline\_safety <- read.csv(airlines)

hist(airline\_safety$fatalities\_00\_14, main= "Total fatalities between 2000 and 2014", xlab = "Fatalities", breaks=4)



airline\_safety$total\_incidents = airline\_safety$incidents\_85\_99 + airline\_safety$incidents\_00\_14  
airline\_safety$total\_seat\_km\_per\_incident = airline\_safety$avail\_seat\_km\_per\_week / airline\_safety$total\_incidents  
segdata = data.table(airline\_safety)  
  
working\_data = segdata[,.(incidents\_00\_14,avail\_seat\_km\_per\_week)]  
  
kmean\_output = kmeans(working\_data, center=3)  
cluster = kmean\_output$cluster  
centers = kmean\_output$centers  
qplot(x=avail\_seat\_km\_per\_week, y=incidents\_00\_14, data=working\_data,color=factor(cluster))

