Week 5 Assignment

Dan Brooks

February 28, 2016

Extract the CSV file from my Github Page and get all of the libraries that are needed for this assignment.

```
library(RCurl)
```

```
## Loading required package: bitops
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
```

```
##
## Attaching package: 'tidyr'

## The following object is masked from 'package:RCurl':
##
##     complete
```

```
URL <- getURL("https://raw.githubusercontent.com/DanielBrooks39/IS607/master/Week_5/Flight_Information.
FlightData <- read.csv(text = URL, header = TRUE)
```

Gives names to the vectors that are in the data frame and add information to the table where there is a blank spot.

Create a Tidy dataset with the columns Airline, Info, Destinations and the total flights that were delayed or ontime for waech airline and destination

```
names(FlightData) <- c("Airline", "Info", "Los Angeles","Phoenix", "San Diego", "San Francisco", "Seatt
FlightData$Airline[2] <- "Alaska"
FlightData$Airline[4] <- "AM West"
Tidy <- FlightData %>% gather("Destination", "Timing", 3:7)
```

Separate the full tidy dataset into delayed and ontime flights

```
Delay <- Tidy %>% filter(Info == "Delay")
OnTime <- Tidy %>% filter(Info == "OnTime"|Info == "Ontime")
```

Find the mean number of lfights per airline that is ontime and delayed

```
AvgDelay <- Delay %>% group_by(Airline) %>% summarise(mean = mean(Timing))
AvgDelay
```

```
## Source: local data frame [2 x 2]
##
##   Airline  mean
##    (fctr) (dbl)
## 1  Alaska 100.2
## 2 AM West 157.4
```

```
AvgOntime <- OnTime %>%  group_by(Airline) %>% summarise(mean = mean(Timing))
AvgOntime
```

```
## Source: local data frame [2 x 2]
##
##   Airline   mean
##    (fctr)  (dbl)
## 1  Alaska  654.8
## 2 AM West 1287.6
```

Join the delayed and ontime datasets together and figure out the ration between the number of ontime flights to the number of delayed flights per airline

```
Joined <- inner_join(AvgDelay, AvgOntime, by = "Airline")
names(Joined) <- c("Airline", "AvgDelay", "AvgOnTime")
FlightInfo <- Joined %>% mutate("Ratio(OnTime/Delay)" = AvgOnTime/AvgDelay)
FlightInfo
```

```
## Source: local data frame [2 x 4]
##
##   Airline AvgDelay AvgOnTime Ratio(OnTime/Delay)
##    (fctr)    (dbl)     (dbl)               (dbl)
## 1  Alaska    100.2     654.8            6.534930
## 2 AM West    157.4    1287.6            8.180432
```

Find the mean number of flights that were ontime or delayed according to their destination

```
AvgDelay <- Delay %>% group_by(Destination) %>%  summarise(mean = mean(Timing))
AvgDelay
```

```
## Source: local data frame [5 x 2]
##
##     Destination  mean
##           (chr) (dbl)
## 1   Los Angeles  89.5
## 2       Phoenix 213.5
## 3     San Diego  42.5
## 4 San Francisco 115.5
## 5       Seattle 183.0
```

```
AvgOntime <- OnTime %>%  group_by(Destination) %>% summarise(mean = mean(Timing))
AvgOntime
```

```
## Source: local data frame [5 x 2]
##
##     Destination   mean
##           (chr)  (dbl)
## 1   Los Angeles  595.5
## 2       Phoenix 2530.5
## 3     San Diego  297.5
## 4 San Francisco  411.5
## 5       Seattle 1021.0
```

Find the ratio between the avg number of flights that were ontime and the average number of flights that were delayed based on their destination

```
Joined <- inner_join(AvgDelay, AvgOntime, by = "Destination")
names(Joined) <- c("Destination", "AvgDelay", "AvgOnTime")
DestInfo <- Joined %>% mutate("Ratio" = AvgOnTime/AvgDelay) %>% arrange(desc(Ratio))
DestInfo
```

```
## Source: local data frame [5 x 4]
##
##     Destination AvgDelay AvgOnTime     Ratio
##           (chr)    (dbl)     (dbl)     (dbl)
## 1       Phoenix    213.5    2530.5 11.852459
## 2     San Diego     42.5     297.5  7.000000
## 3   Los Angeles     89.5     595.5  6.653631
## 4       Seattle    183.0    1021.0  5.579235
## 5 San Francisco    115.5     411.5  3.562771
```

If I had to pick an airline to fly on, I would pick AM West, because their ratio of ontime flights to delayed flights is higher thatn Alaska. Also, I would pick Phoenix as my destination because their ratio is also the best compared to the other destinations.