DATA 607 - Assignment 5

Chitrarth Kaushik

2/27/2020

Importing the CSV file into R using readr function - read_csv. The function automatically lables missing column names and also guesses the variable type.

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------------------------- tidyverse 1.3

## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.4
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts -------------------------------------------------------------------- tidyverse_conflicts
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
flight_det<-read_csv("https://raw.githubusercontent.com/chitrarth2018/607-Assignment_5/master/Flight_det
```

```
## Warning: Missing column names filled in: 'X1' [1], 'X2' [2]

## Parsed with column specification:
## cols(
##   X1 = col_character(),
##   X2 = col_character(),
##   `Los Angeles` = col_double(),
##   Phoenix = col_double(),
##   `San Diego` = col_double(),
##   `San Francisco` = col_double(),
##   Seattle = col_double()
## )
```

Transforming the wide data to generate the comparison between the two airlines

```
final_flight<-flight_det%>%gather(City, Num_flights, 'Los Angeles', Phoenix, 'San Diego', 'San Francisc

final_flight_ana<-final_flight

analysis<-final_flight_ana%>%group_by(Flight_name)%>%mutate(Per_tot=(Num_flights/sum(Num_flights))*100)%
view(analysis)
```