Analytics Foundations: Lab 2

1. Revisit the *NormTemp* dataset from Lab 1, where we examined the observed mean body temperature (**temperature**) in comparison to the well-known "average" of 98.6.

   a. Perform a statistical test (alpha = 0.05) to determine whether this well-known number is actually the mean body temperature. What is your p-value? Explain in words what this p-value means. What is your conclusion?

   b. Give the 95% Confidence Interval for **temperature**. Explain in words what a 95% confidence interval represents.

   c. If we restrict our analysis to only the females in this dataset, would our conclusion change?

   d. Is there any difference (alpha=0.05) in temperature between the two genders recorded in this dataset?

2. The *Airline* dataset contains information regarding the number of international airline travelers (variable **air**) across different months of the year from 1949-1960. To obtain this data set, you will need to:

   data(AirPassengers).

   We are interested in knowing if during this time period there was a significant difference between air travel in the Summer months of June, July, and August vs. the remainder of the year? Use a statistical hypothesis test (alpha=0.05) to support your answer. In order to get month information, you will need to:

   install.packages('tseries')
   install.packages('forecast')

   library(tseries)
   library(forecast)
   cycle(AirPassengers)

Discuss the assumptions in this hypothesis test and if you feel comfortable with your conclusion?

You might find the following code useful:

```
air1 =  data.frame(AirPassengers)
air2 = air1 %>% mutate(summer=ifelse(cycle(AirPassengers) %in%
6:8,1,0))
```