Week 3 Tutorial Worksheet

AY22/23 Semester 2

Question 1 Base `R` syntax

1. Create a vector of all integers from 13 to 21, and save it as `a1`.
2. Create a vector of *odd* integers from 13 to 21, and save it as `a2`.
3. What is the sum of `a1` and `a2`?
4. The formula for the first `n` integers is $n(n+1)/2$. Compute the sum of all integers from 1 to 100 and verify that this formula holds for `n=100`.

Question 2 Base `R` data manipulation and plotting

The following code loads a built-in `R` data set `airquality`. The data contain daily air quality measurements in New York in 1973.

```
data(airquality)
```

To learn more about the data set, type `?airquality` in your `R` console and read the documentation file.

1. What does the command `summary(airquality)` do? What does the command `class(airquality)` do?
2. Using the `str()` command to view the structure of this data set. How many months are there in the data set?
3. Is there any missing entries in the `airquality` data frame? If there is, eliminate row(s) that contain missing values and save it in a new data frame called `df`.
4. For the remaining questions, we will work on the data frame `df`. Create two new data frames, one for summer months (May and June) and the other for fall months (July, August, and September) and name them as `df_summer` and `df_fall`, respectively.

5. Do summer and fall months have similar temperature? Create a factor variable `Season` in `df` that takes value of `Summer` for May and June, and `Fall` for July, August, and September. Compute the summary statistics of temperature in summer and fall.

6. Recode the `Month` variable from 5 - 9 to May through September. Recreate the boxplot below comparing temperature across all months.