# Introduction to Data Scientist 11372 (UG)

# Assignment 1 – Data Wrangling and Exploration

Tuan Anh (Vincent) Nguyen – u3196825

**Part D – Insight:** *As a data scientist, you need to practice extracting insights and valuable information from the analysis you conduct on the data. This can be done by raising some questions that can be answered by doing this analysis. Question such as, "Based on the weather analysis, what is the best time of the year, in which you would recommend people living outside of ACT to come to visit it?"*

*Can you list at least 2 questions that can be answered by running analysis on this data set?*

1. What is the time in the year that ACT would be most vulnerable to bushfire according to 2019 data?
2. What is the best time of the year for Hot Air Balloon ride in Canberra according to 2019 data?

**Answer for question 1:**

- During December and January was the time Canberra most vulnerable to Bushfire due to temperature can get peak to 41.1 and 41.6 with comparable high average temperature. Especially humidity during December is exceptionally low, only 33.3 average compare January and Feb, resulted in fire control challenging and easily to spread wildly.

```
temp_month
A tibble: 12 x 4
 Month min_temp_month max_temp_month avg_temp_month
  <ord>          <dbl>          <dbl>          <dbl>
1 1               13.4           41.6           26.1
2 2                5.5           36             21.0
3 3                3.1           34.1           19.0
4 4                0.2           26.1           15.0
5 5               -2.7           23.8            9.96
6 6               -4.9           17.4            6.99
7 7               -5.1           17.3            7.17
8 8               -5.2           18.9            7.15
9 9               -3.1           24.7           10.4
0 10              -0.2           31.7           14.9
1 11               1.6           39             18.2
2 12               4.6           41.1           22.4
temp_quarter
A tibble: 4 x 4
Quarter min_temp_quarter max_temp_quarter avg_temp_quarter
<ord>              <dbl>            <dbl>            <dbl>
1                    3.1             41.6             22.1
2                   -4.9             26.1             10.7
3                   -5.2             24.7              8.23
4                   -0.2             41.1             18.5
```

-

- humid_month

A tibble: 12 x 4

| Month | min_humid_month | max_humid_month | average_humid_month |
|---|---|---|---|
| <ord> | <dbl> | <dbl> | <dbl> |
| 1 | 23.5 | 69 | 47.9 |
| 2 | 26 | 65 | 47.5 |
| 3 | 31.5 | 78.5 | 56.5 |
| 4 | 40.5 | 76 | 58.5 |
| 5 | 53 | 89 | 68.6 |
| 6 | 56.5 | 92 | 72.5 |
| 7 | 49.5 | 88.5 | 67.7 |
| 8 | 34.5 | 79.5 | 61.1 |
| 9 | 27 | 76 | 48.2 |
| 10 | 22 | 68.5 | 42.2 |
| 11 | 16 | 75 | 35.2 |
| 12 | 11.5 | 67 | 33.3 |

humid_quarter

A tibble: 4 x 4

| Quarter | min_humid_quarter | max_humid_quarter | average_humid_quarter |
|---|---|---|---|
| <ord> | <dbl> | <dbl> | <dbl> |
| 1 | 23.5 | 78.5 | 50.8 |
| 2 | 40.5 | 92 | 66.5 |
| 3 | 27 | 88.5 | 59.1 |
| 4 | 11.5 | 75 | 36.9 |

**Answer for question 2:**

- In ACT first 2 quarter (from January to April), the average wind speed tends to lower compared to the second half, where its ranging between 12.9 to 15.3 km/h which is a safe and good for riding balloon with the temperature is mild not too hot or extreme (average 22.1)

- ws_month

A tibble: 12 x 4

| | Month | min_ws_month | max_ws_month | avg_ws_month |
|---|---|---|---|---|
| | <ord> | <dbl> | <dbl> | <dbl> |
| 1 | 1 | 8.5 | 32.5 | 14.8 |
| 2 | 2 | 5.5 | 33 | 15.6 |
| 3 | 3 | 3 | 37 | 15.4 |
| 4 | 4 | 3 | 28 | 11.5 |
| 5 | 5 | 3.5 | 38.5 | 15.1 |
| 6 | 6 | 1 | 38 | 11.9 |
| 7 | 7 | 3.5 | 39.5 | 17.5 |
| 8 | 8 | 4.5 | 35 | 17.0 |
| 9 | 9 | 6.5 | 35 | 18.0 |
| 0 | 10 | 7.5 | 41.5 | 18.4 |
| 1 | 11 | 4.5 | 38 | 21.6 |
| 2 | 12 | 8 | 36 | 18.0 |

- > ws_quarter

# A tibble: 4 x 4

| | Quarter | min_ws_quarter | max_ws_quarter | average_ws_quarter |
|---|---|---|---|---|
| | <ord> | <dbl> | <dbl> | <dbl> |
| 1 | 1 | 3 | 37 | 15.3 |
| 2 | 2 | 1 | 38.5 | 12.9 |
| 3 | 3 | 3.5 | 39.5 | 17.5 |
| 4 | 4 | 4.5 | 41.5 | 19.3 |