

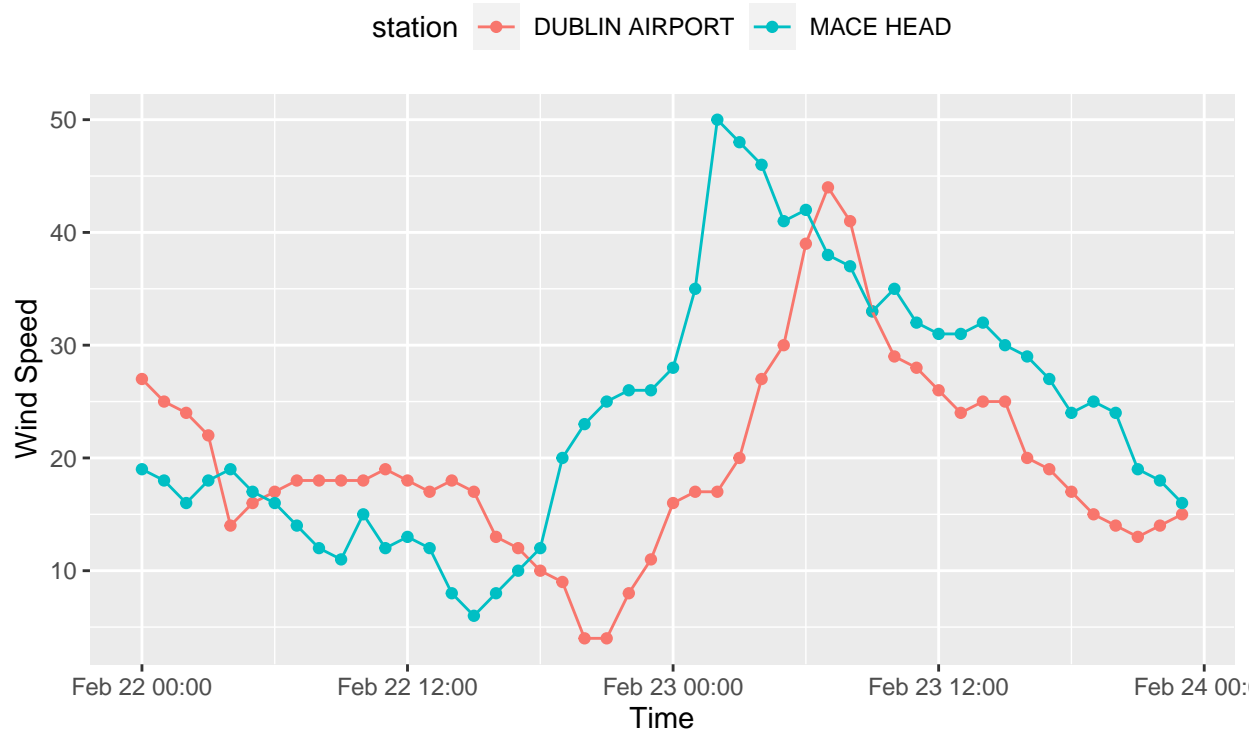
Lab Test 2 - dplyr and ggplot2

Allocated Time - 1 hour 45 mins

1. Using a combination of `filter()` and the library `ggplot2`, generate the following graph for the weather stations “MACE HEAD” and “DUBLIN AIRPORT” for the 22nd and 23rd of February.

Wind Speed Plot

February 2017



2. Using a combination of **filter()**, **arrange()** and **slice()**, display the top five wettest hours of the year, in the order in which the observations appeared

```
top5
```

```
## # A tibble: 5 x 12
##   station      year month   day  hour date                rain  temp  rhum  msl
##   <chr>      <dbl> <dbl> <int> <int> <dtm>                <dbl> <dbl> <dbl> <dbl>
## 1 CASEMENT    2017     5    27    10 2017-05-27 10:00:00  13.5  14.4   98 1007.
## 2 BALLYHAISE  2017     7    19    16 2017-07-19 16:00:00  14.3  15.9   96 1000.
## 3 MALIN HEAD  2017     8    17    18 2017-08-17 18:00:00  15.4  13.9   89 1002.
## 4 MALIN HEAD  2017     8    22    19 2017-08-22 19:00:00  16.4  15.8   94 1010.
## 5 MALIN HEAD  2017     8    22    20 2017-08-22 20:00:00  16.6  16.8   95 1009.
## # i 2 more variables: wdsp <dbl>, wddir <dbl>
```

3. Using a combination of **filter()**, **select()** and **arrange()**, create the following tibble for temperatures on June 21st for “PHOENIX PARK” and “VALENTIA OBSERVATORY”.

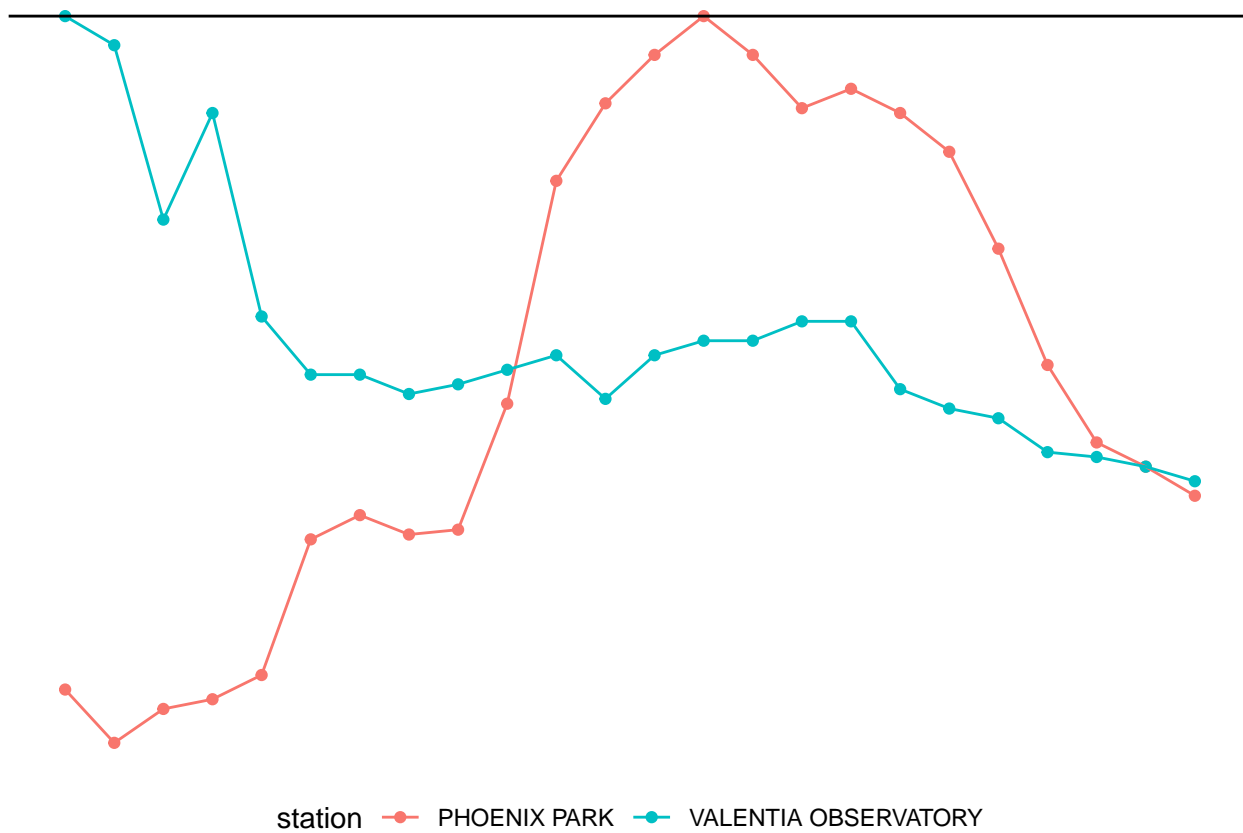
```
temp21
```

```
## # A tibble: 48 x 3
##   hour station          temp
##   <int> <chr>          <dbl>
## 1     0 PHOENIX PARK      14.4
## 2     0 VALENTIA OBSERVATORY 23.5
## 3     1 PHOENIX PARK      13.3
## 4     1 VALENTIA OBSERVATORY 22.9
## 5     2 PHOENIX PARK      14
## 6     2 VALENTIA OBSERVATORY 19.3
## 7     3 PHOENIX PARK      14.2
## 8     3 VALENTIA OBSERVATORY 21.5
## 9     4 PHOENIX PARK      14.7
## 10    4 VALENTIA OBSERVATORY 17.3
## # i 38 more rows
```

4. Using `mutate()` and any other relevant `dplyr` function to generate the following new columns for `temp21`, and plot the difference between the actual and the max for each station in each hour. In the plot make use of the appropriate theme, and draw a line to indicate zero.

```
temp21
```

```
## # A tibble: 48 x 5
## # Groups:   station [2]
##   hour station          temp DailyMaxTempS DiffDailyMax
##   <int> <chr>          <dbl>         <dbl>         <dbl>
## 1     0 PHOENIX PARK      14.4          28.3         -13.9
## 2     0 VALENTIA OBSERVATORY 23.5          23.5           0
## 3     1 PHOENIX PARK      13.3          28.3         -15
## 4     1 VALENTIA OBSERVATORY 22.9          23.5        -0.600
## 5     2 PHOENIX PARK      14           28.3         -14.3
## 6     2 VALENTIA OBSERVATORY 19.3          23.5         -4.2
## 7     3 PHOENIX PARK      14.2          28.3         -14.1
## 8     3 VALENTIA OBSERVATORY 21.5          23.5          -2
## 9     4 PHOENIX PARK      14.7          28.3         -13.6
## 10    4 VALENTIA OBSERVATORY 17.3          23.5         -6.2
## # i 38 more rows
```



5. Using **summarise()** and any other relevant **dplyr** function to generate the graph of total annual rainfall by hour of the day, for “DUBLIN AIRPORT” and “NEWPORT”. Note the function **geom_col()** can be used, and has a similar use of the **position** argument as **geom_bar()**.

