Lab Test 3 - purrr

Allocated Time - 1 hour 45 mins

Load the following libraries.

```
library(aimsir17)
library(purrr)
library(dplyr)
library(ggplot2)
library(tidyr)
```

1. Filter three stations from the tibble **observation**, and stored in the variable **a17**.

```
a17
```

```
## # A tibble: 26,280 x 12
##    station    year month   day  hour date                rain  temp  rhum   msl
##    <chr>     <dbl> <dbl> <int> <int> <dttm>              <dbl> <dbl> <dbl> <dbl>
##  1 DUBLIN A~  2017     1     1     0 2017-01-01 00:00:00   0.9   5.3    91 1020.
##  2 DUBLIN A~  2017     1     1     1 2017-01-01 01:00:00   0.2   4.9    95 1020.
##  3 DUBLIN A~  2017     1     1     2 2017-01-01 02:00:00   0.1   5      92 1020.
##  4 DUBLIN A~  2017     1     1     3 2017-01-01 03:00:00   0     4.2    90 1020.
##  5 DUBLIN A~  2017     1     1     4 2017-01-01 04:00:00   0     3.6    88 1020.
##  6 DUBLIN A~  2017     1     1     5 2017-01-01 05:00:00   0     2.8    89 1020.
##  7 DUBLIN A~  2017     1     1     6 2017-01-01 06:00:00   0     1.7    91 1020.
##  8 DUBLIN A~  2017     1     1     7 2017-01-01 07:00:00   0     1.6    91 1021
##  9 DUBLIN A~  2017     1     1     8 2017-01-01 08:00:00   0     2      89 1022.
## 10 DUBLIN A~  2017     1     1     9 2017-01-01 09:00:00   0     2.6    84 1023.
## # i 26,270 more rows
```

```
## # i 2 more variables: wdsp <dbl>, wddir <dbl>
```

```
a17 %>%
  select(station) %>%
  pull() %>%
  unique()
```

```
## [1] "DUBLIN AIRPORT"       "MACE HEAD"            "VALENTIA OBSERVATORY"
```

2. Add a new column (using **case_when()**) that adds the season to the tibble. Winter includes months 11, 12 and 1; Spring includes months 2, 3 and 4; summer contains months 5, 6 and 7; and autumn includes months 8, 9 and 10. Note that we reduce the number of columns as part of this process.

```
a17
```

```
## # A tibble: 26,280 x 5
##    station        season date                 wdsp   msl
##    <chr>          <chr>  <dttm>              <dbl> <dbl>
##  1 DUBLIN AIRPORT Winter 2017-01-01 00:00:00    12 1020.
##  2 DUBLIN AIRPORT Winter 2017-01-01 01:00:00     8 1020.
##  3 DUBLIN AIRPORT Winter 2017-01-01 02:00:00     8 1020.
##  4 DUBLIN AIRPORT Winter 2017-01-01 03:00:00    12 1020.
##  5 DUBLIN AIRPORT Winter 2017-01-01 04:00:00    11 1020.
##  6 DUBLIN AIRPORT Winter 2017-01-01 05:00:00    12 1020.
##  7 DUBLIN AIRPORT Winter 2017-01-01 06:00:00    13 1020.
##  8 DUBLIN AIRPORT Winter 2017-01-01 07:00:00    13 1021
##  9 DUBLIN AIRPORT Winter 2017-01-01 08:00:00    13 1022.
## 10 DUBLIN AIRPORT Winter 2017-01-01 09:00:00    13 1023.
## # i 26,270 more rows
```

3. Use the appropriate function from **tidyr** to generate the following tibble.

```
a17_n
```

```
## # A tibble: 12 x 3
## # Groups:   station, season [12]
##    station              season data
##    <chr>                <chr>  <list>
##  1 DUBLIN AIRPORT       Winter <tibble [2,208 x 3]>
##  2 DUBLIN AIRPORT       Spring <tibble [2,136 x 3]>
##  3 DUBLIN AIRPORT       Summer <tibble [2,208 x 3]>
##  4 DUBLIN AIRPORT       Autumn <tibble [2,208 x 3]>
##  5 MACE HEAD            Winter <tibble [2,208 x 3]>
##  6 MACE HEAD            Spring <tibble [2,136 x 3]>
##  7 MACE HEAD            Summer <tibble [2,208 x 3]>
##  8 MACE HEAD            Autumn <tibble [2,208 x 3]>
##  9 VALENTIA OBSERVATORY Winter <tibble [2,208 x 3]>
## 10 VALENTIA OBSERVATORY Spring <tibble [2,136 x 3]>
## 11 VALENTIA OBSERVATORY Summer <tibble [2,208 x 3]>
## 12 VALENTIA OBSERVATORY Autumn <tibble [2,208 x 3]>
```

4. Add a column which shows the correlation coefficent for each observation, for the variables **wdsp** and **msl**. In the call to **cor**, the argument use="complete.obs" can be used to filter out any missing values.

```
a17_n
```

```
## # A tibble: 12 x 4
## # Groups:   station, season [12]
##    station              season data                 Correlation
##    <chr>                <chr>  <list>                     <dbl>
##  1 VALENTIA OBSERVATORY Autumn <tibble [2,208 x 3]>      -0.553
##  2 MACE HEAD            Autumn <tibble [2,208 x 3]>      -0.516
##  3 VALENTIA OBSERVATORY Spring <tibble [2,136 x 3]>      -0.486
##  4 DUBLIN AIRPORT       Autumn <tibble [2,208 x 3]>      -0.475
##  5 VALENTIA OBSERVATORY Winter <tibble [2,208 x 3]>      -0.462
##  6 MACE HEAD            Spring <tibble [2,136 x 3]>      -0.453
##  7 MACE HEAD            Winter <tibble [2,208 x 3]>      -0.426
##  8 DUBLIN AIRPORT       Winter <tibble [2,208 x 3]>      -0.404
##  9 VALENTIA OBSERVATORY Summer <tibble [2,208 x 3]>      -0.327
## 10 DUBLIN AIRPORT       Spring <tibble [2,136 x 3]>      -0.309
## 11 MACE HEAD            Summer <tibble [2,208 x 3]>      -0.290
## 12 DUBLIN AIRPORT       Summer <tibble [2,208 x 3]>      -0.144
```

5. Display the results (absolute values of the correlation) in the following plot.

```
ggplot(a17_n,aes(x=station,y=abs(Correlation),fill=season))+
  geom_col(position="dodge")
```