Examining the influence of wind speed on TMACC patrons(2014-2017)

Akira Govender (3540386) & Yameen Badrodien (2340061)

15 March 2018

Introduction

Table Mountain is one of Soth Africa's most renowned attractions, drawing in thousands of tourists annually. Tourists have the choice of either hiking up the mountain or as many prefer, using the cableway. The Table Mountain Aerial Cableway Company has been providing these tourists with a world-class experience since October 4, 1929. The TMACC operates in a National Park and World Heritage Site.

The data collected gave insight into the number of patrons making use of the cableway. The peak season takes place during the summer months, and commences 14 December and concludes 8 January. This period is characterized by unpredictable weather, especially with regards to wind. It is a well known fact that wind and the relative speed of winds experienced at the TMACC site have the capacity to influence periods of operation. This study examines the wind trends and patron trends for the peak season period of 2014- 2017, and attempts to understand the relationship between these two variables.

Methods and Meterials

It is a well-known fact that wind conditions are highly influential in its ability to affect the operating hours of the cableway- which subsequently has the capacity to influence the number of patrons making use of the service.

Data collection:

Data was obtained from a correspondent at the TMACC. The dataset included the number of patrons making use of the cableway during the peak season periods from 2014 to 2017. The patron data, in its raw state itself was separated in terms of the direction being travelled. For the purpose of this study, this dataset needed to be corroborated with additional data, namely, the average wind speeds (km/h), experienced during the peak season periods. The wind speed data was obtained from an open source weather data website www.timeanddate.com. The wind speed data relative to the peak season periods for 2014-2017 were extracted and appended to the patron dataset obtained from TMACC. The raw data was sorted using Microsoft Excel, ensuring that the data was separated by travel direction, and collectively expressed as the total number of patrons. The data was then converted to a comma separated value file format (.csv). The conversion of the spreadsheet to this file format would ensure that the data could be imported into the graphical and statistical manipulation software- R studio. The manner in which the original data set was manipulated can be seen in the appended R-script.

Setup

```
library(tidyverse)
library(dplyr)
library(lubridate)
library(ggpubr)
library(readr)
```

```r
library(ggmap)
library(ggsn)
library(rmarkdown)
library(knitr)
```

```r
knitr::opts_chunk$set(echo = TRUE)
```

Load data and assign to object

```r
cable <-  read.csv2("Intro_R_Workshop-master/data/cable_data_new.csv")
```

Creating a satellite map indicating the loacation of the TMACC: Upper and Lower Stations

Loading the data

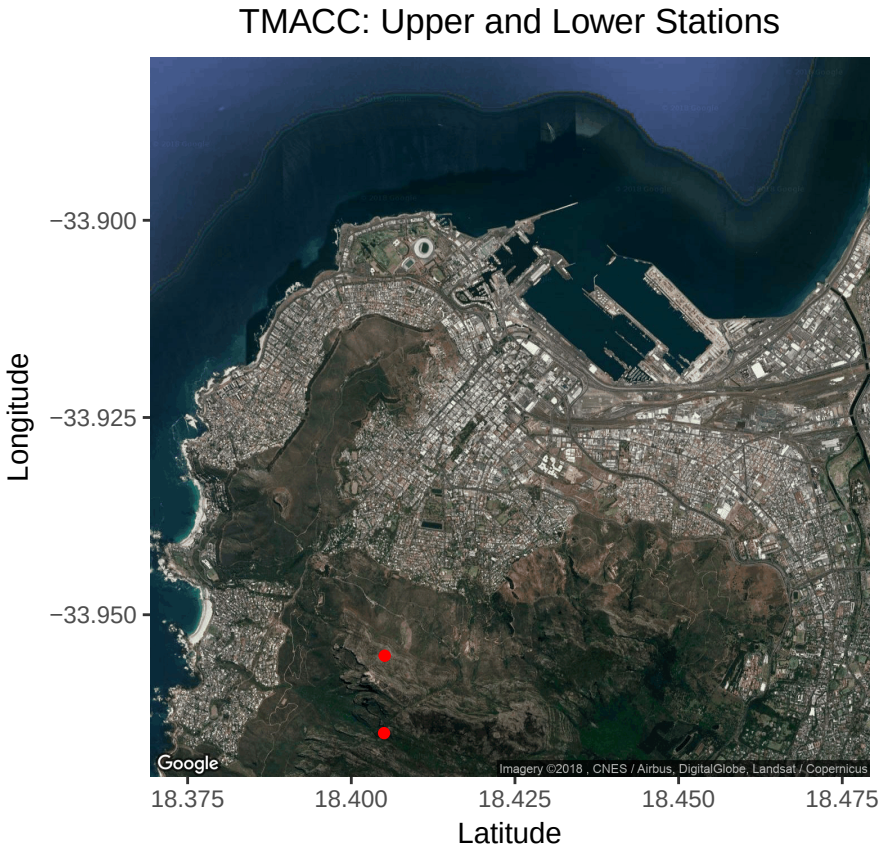
```r
load("Intro_R_Workshop-master/data/cape_point.RData")
```

Loading the map data and assigning to an object

```r
capepoint <- get_map(location = c(lon = 18.4241, lat = -33.9249),
                     zoom = 13, maptype = 'satellite')
```

Plotting the map and coordinates of the Upper and Lower Stations

```r
ggmap(capepoint) +
  geom_point(data = cable, aes(x = lon+0.002, y = lat-0.007),
             colour = "red", size = 1.5) +
  labs(x = "Latitude", y = "Longitude", title = "TMACC: Upper and Lower Stations") +
  theme(plot.title = element_text(hjust = 0.5))
```



Assigning the plot to an object

```
cable1 <- ggmap(capepoint) +
  geom_point(data = cable, aes(x = lon+0.002, y = lat-0.007),
             colour = "red", size = 1.5) +
  labs(x = "Latitude", y = "Longitude", title = "TMACC: Upper and Lower Stations",
       caption = "Figure 1: Satellite image of TMACC: Upper and Lower Cablestations ") +
  theme(plot.title = element_text(hjust = 0.5), plot.caption = element_text(hjust = 0.5))
```

Adding labels and assigning the plot to the final object

```
cable2 <- cable1 +
  geom_text(data = cable, aes(lon+0.0135, lat-0.0045, label = site),
            hjust = 1.05, vjust = 0.7,
            size = 3, colour = "white")

cable2
```

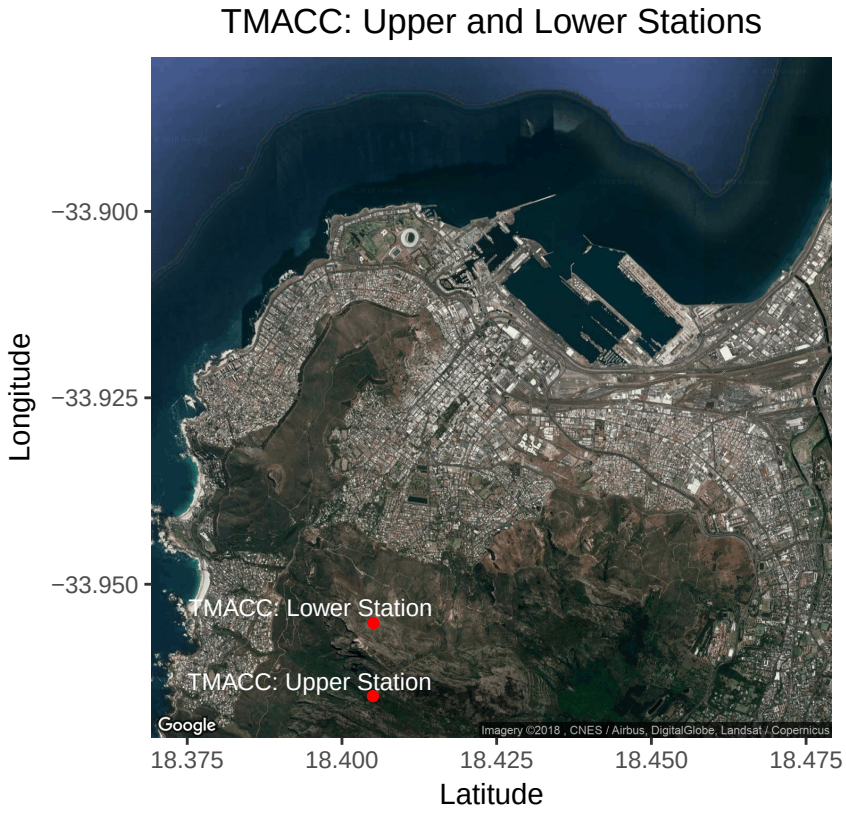


Figure 1: Satellite image of TMACC: Upper and Lower Cablestations

Creating a graph indicating the TMACC: Peak season patrons 2014-2017

Group and mutate

```
direction_data <- cable %>%
  mutate(year = as.numeric(year)) %>%
  group_by(Travel.Direction, year_raw) %>%
  summarise(mn_amount = mean(amount_raw, na.rm = TRUE),
            sd_amount = sd(amount_raw, na.rm = TRUE))
```

This created an average for the 4 time periods

Creating the graph

```
direction_graph <- ggplot(direction_data, aes(x = year_raw, y = mn_amount,
                                              fill = Travel.Direction)) +
  geom_col(aes(fill = Travel.Direction), position = "dodge",width = 0.5) +
  labs(x = "Year",y = "Average number of patrons",
       title = "TMACC Peak Season Patrons 2014-2017",
       caption = "Figure 2: The average number of patrons recorded (2014-2017)" ) +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5)) +
  geom_errorbar(aes(ymin = mn_amount - sd_amount,
```

```
                    ymax = mn_amount + sd_amount),
                position = "dodge",width = 0.5)
direction_graph
```

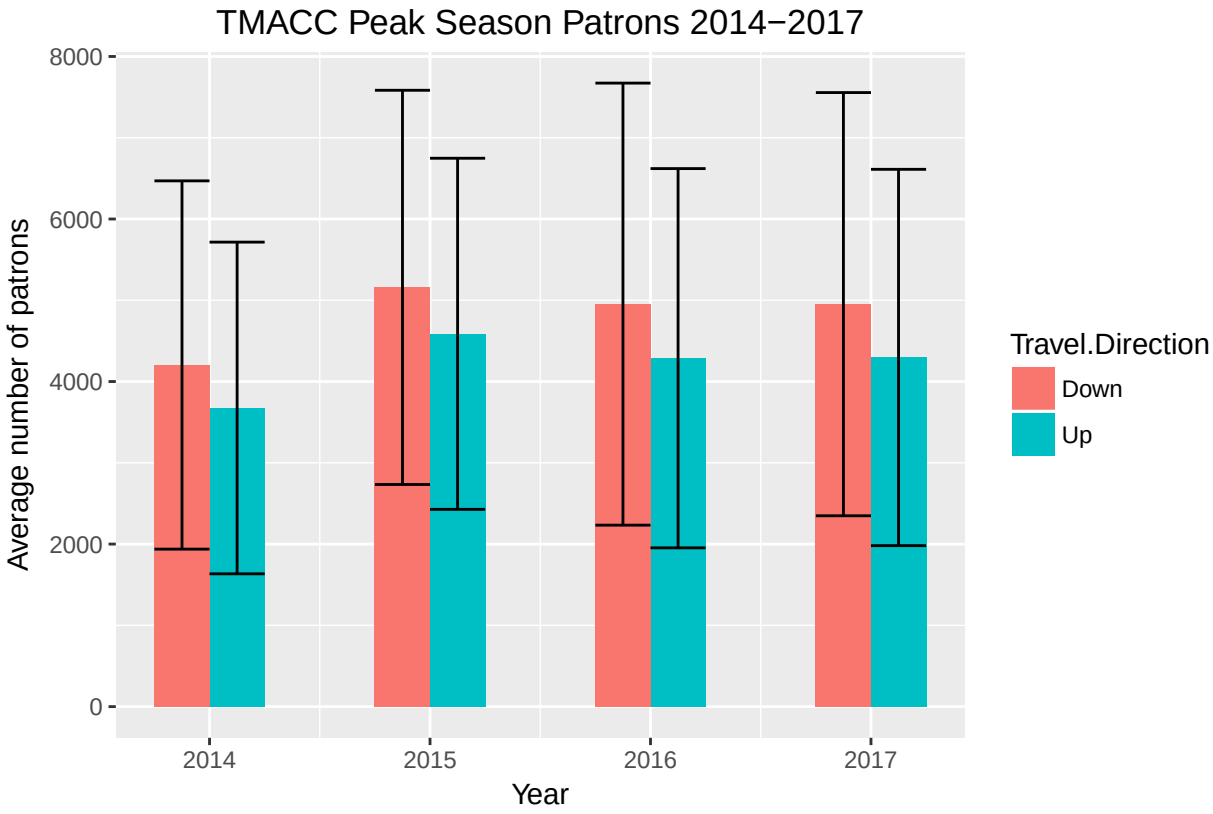


Figure 2: The average number of patrons recorded (2014−2017)

Creating a graph showing the comparison between weekends and weekdays

Group and mutate

```
period_new <- cable %>%
  mutate(year = as.numeric(year)) %>%
  group_by(year, Period) %>%
  summarise(mn_visitors = mean(visitors, na.rm = TRUE),
            sd_visitors = sd(visitors, na.rm = TRUE))
```

Removing the unwanted 9th row with NA values

```
period_new <- period_new[-c(9), ]
```

Creating the graph

```
period_graph <- ggplot(period_new, aes(x = year, y = mn_visitors,
                                       fill = Period)) +
  geom_col(aes(fill = Period), position = "dodge",width = 0.5) +
  labs(x = "Year",y = "Average number of patrons",
       title = "TMACC Patrons: Weekends vs Weekdays (2014-2017)",
       caption = "Figure 3: The average number of patrons recorded
       for both weekdays and weekends during peak season (2014-2017)" ) +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5)) +
  geom_errorbar(aes(ymin = mn_visitors - sd_visitors,
                    ymax = mn_visitors + sd_visitors),
                position = "dodge",width = 0.5)
period_graph
```

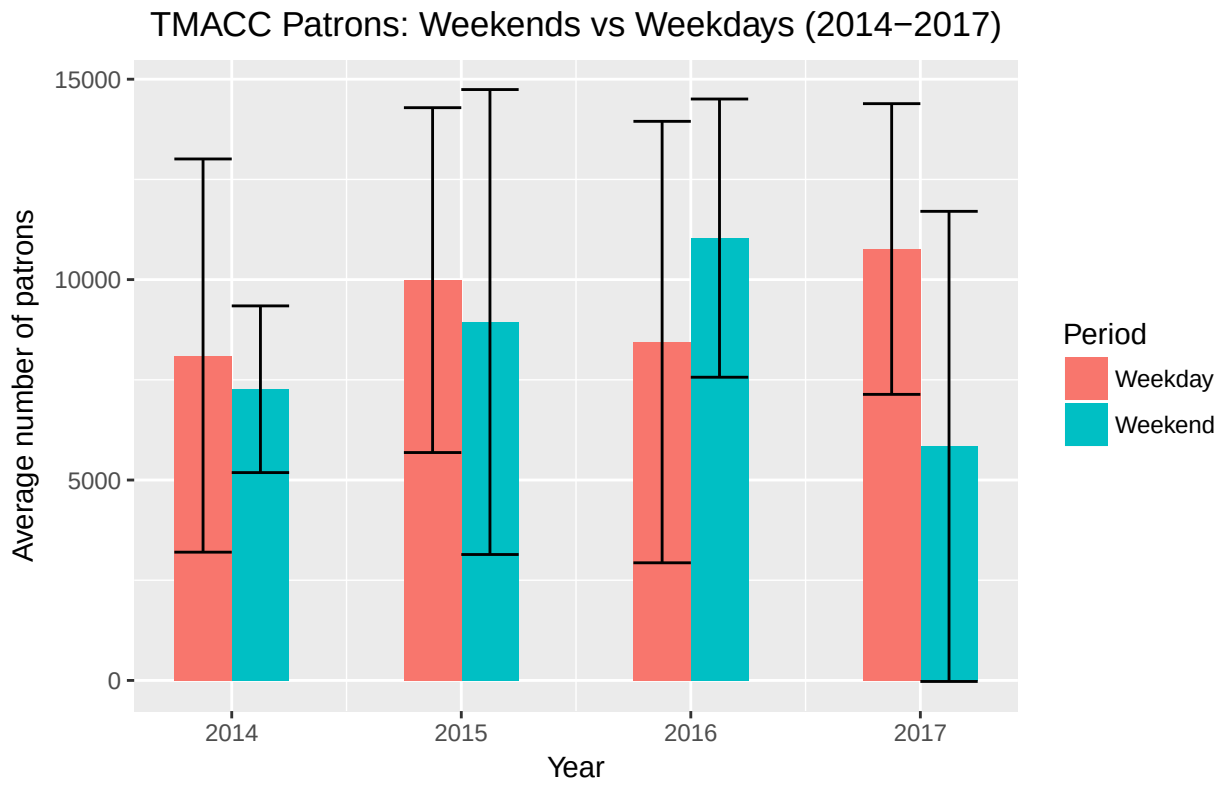


Figure 3: The average number of patrons recorded
for both weekdays and weekends during peak season (2014−2017)

Conducting a t-test on the number of Patrons vs. the Period

Group and mutate

```
tdata <- cable %>%
  group_by(visitors, Period) %>%
  summarise()
```

Remove the unwanted 92nd row with NA values

```
tdata_new <- tdata[-c(92), ]
```

Conducting the T-test

```
test1 <- tdata_new %>%
  filter(Period == "Weekend" | Period == "Weekday") %>%
  select(Period, visitors)

t.test(visitors ~ Period, data = test1)
```

```
##
##  Welch Two Sample t-test
##
## data:  visitors by Period
## t = 2.0346, df = 37.694, p-value = 0.04898
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##     8.869692 3743.009937
## sample estimates:
## mean in group Weekday mean in group Weekend
##             10746.125              8870.185
```

Analysing annual wind trends for Cape Town (2014-2017)

```
annual_wind_speed <- ggplot(data = cable, aes(x = day, y = wind_raw,
                                              colour = as.factor(year_raw))) +
  geom_line(aes(group = year_raw), size = 1) +
  geom_smooth(method = lm, colour = "Orange", se = FALSE) +
  geom_point(size = 2) +
  facet_wrap(~year_raw) +
  labs(x = "Days (14 Dec-08 Jan)", y = "Wind Speed (km/h)", colour = "Year",
       caption = "Figure 4: The annual wind trends observed from 14 Decemeber
       till 8 January for 2014-2017") +
  ggtitle("TMACC: Annual Wind Speed (km/h) Trends (2014-2017)") +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5))

annual_wind_speed
```

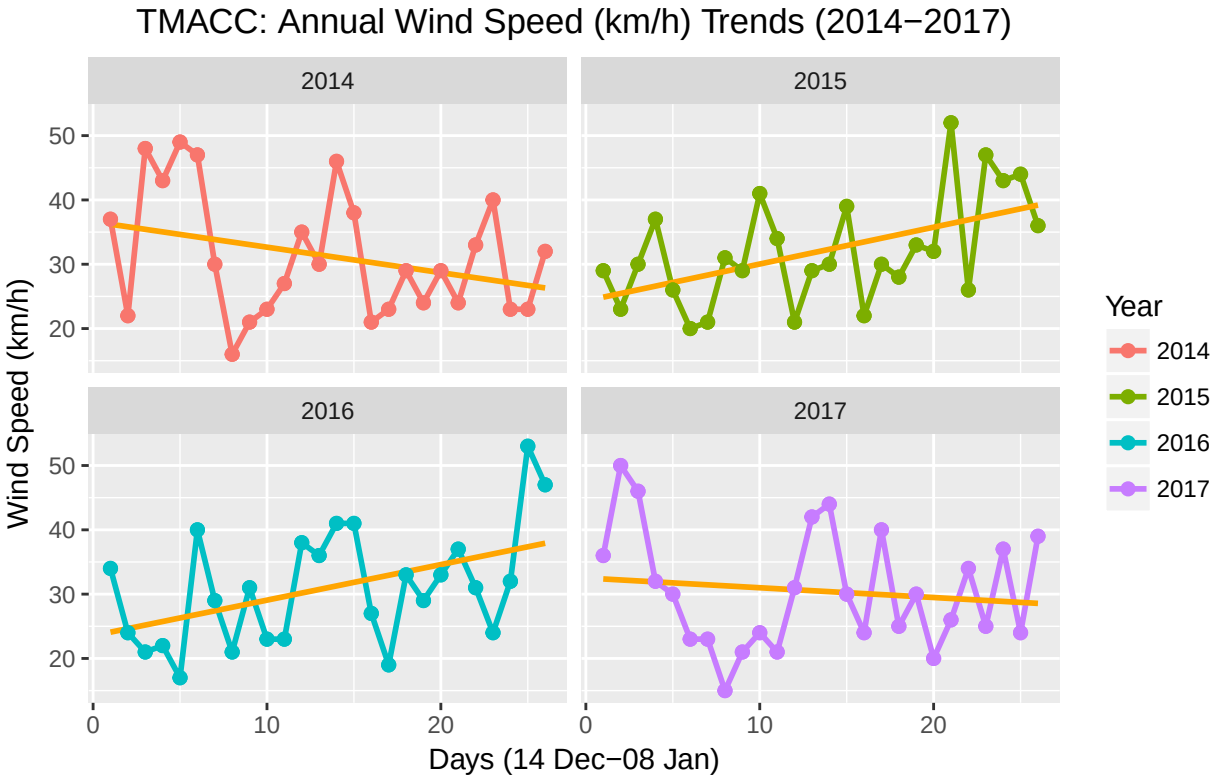


Figure 4: The annual wind trends observed from 14 Decemeber till 8 January for 2014−2017

Analysing the annual patrons per day (2014-2017)

```
annual_patrons <- ggplot(data = cable, aes(x = day, y = amount_raw,
                                           colour = as.factor(year_raw))) +
  geom_line(aes(group = year_raw), size = 1) +
  scale_y_continuous(breaks = c(2500, 5000, 7500, 10000, 12500, 15000),
                     minor_breaks = NULL) +
  scale_x_continuous(breaks = c(2, 10, 15, 20, 25, 30),
                     minor_breaks = NULL) +
  geom_smooth(method = lm, colour = "Orange", se = FALSE) +
  geom_point(size = 2) +
  facet_wrap(~year_raw) +
  labs(x = "Days (14 Dec - 08 Jan)", y = "Patrons", colour = "Year",
       caption = "Figure 5: The annual number of patrons recorded from
       14 Decemeber till 8 January for 2014-2017") +
  ggtitle("TMACC: Annual Patrons Per Day (2014-2017)") +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5))


annual_patrons
```

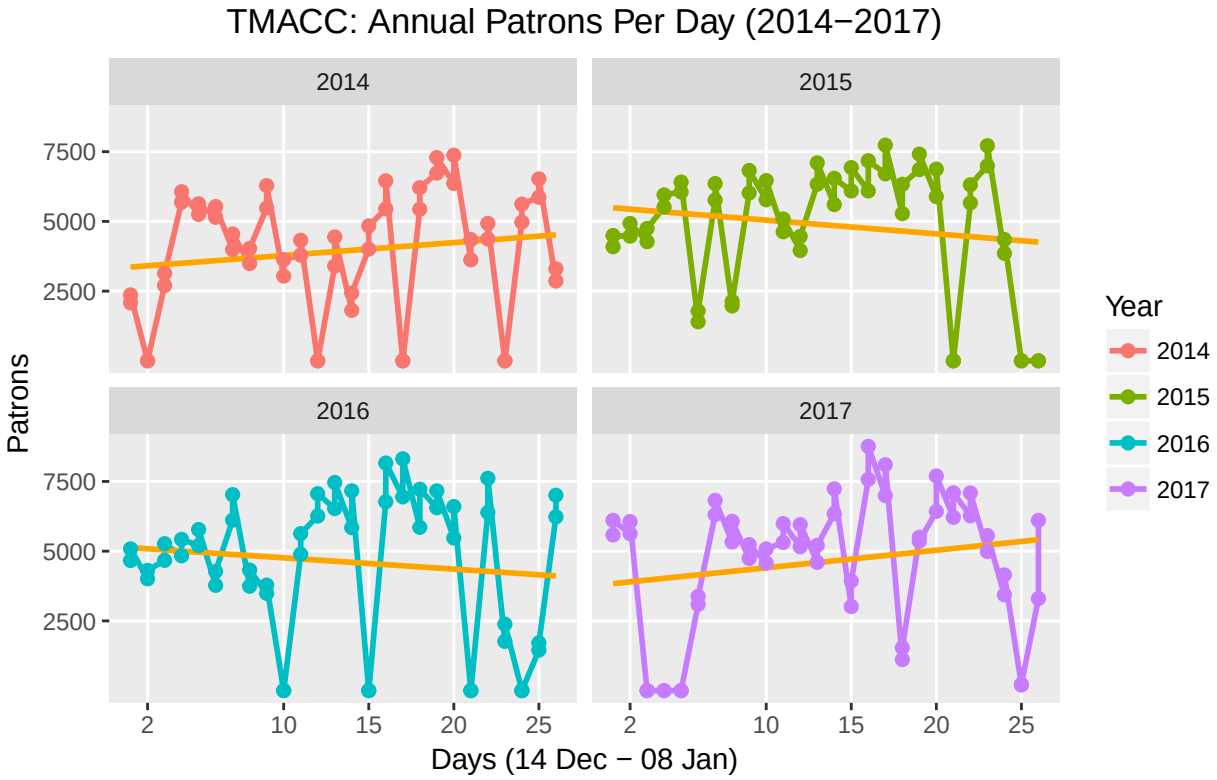


Figure 5: The annual number of patrons recorded from 14 Decemeber till 8 January for 2014−2017

Analysing the correlation between wind speed (km/h) vs. Patrons (2014-2017)

```
correlation <- ggplot(data = cable, aes(x = wind_raw, y = amount_raw)) +
  geom_point(pch = 16, cex = 1.5, colour = "Navy") +
  scale_y_continuous(breaks = c(2500, 5000, 7500, 10000, 12500, 15000),
                     minor_breaks = NULL) +
  geom_smooth(method = lm, colour = "Red", se = FALSE) +
  labs(x = "Average Wind Speed (km/h)", y = "Patrons",
       caption = "Figure 6: Correlation between wind speed (km/h) and
       average number of patrons for 2014-2017") +
  ggtitle("TMACC: Wind Speed vs Patrons (2014-2017)") +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5))

correlation
```

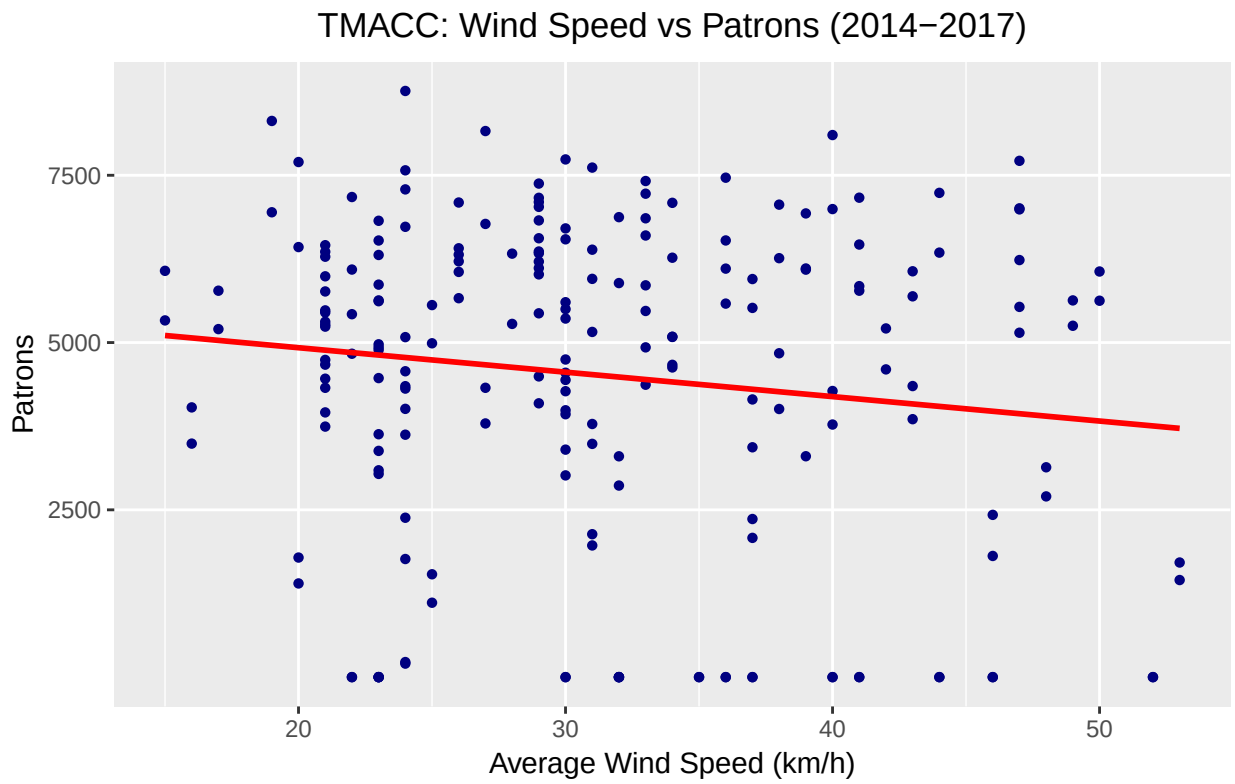


Figure 6: Correlation between wind speed (km/h) and average number of patrons for 2014−2017

Running the correlation test: Wind vs. Patrons (2014-2017)

```
cor.test(cable$wind_raw, cable$amount_raw, method = "pearson",
         conf.level = 0.95)
```

```
##
##  Pearson's product-moment correlation
##
## data:  cable$wind_raw and cable$amount_raw
## t = -1.9809, df = 206, p-value = 0.04893
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2677815811 -0.0006928596
## sample estimates:
##        cor
## -0.1367211
```

Conducting the ANOVA test for the number of patrons over the four years

Summarizing the data to get the two needed columns

```
stat_data <- cable %>%
  mutate(year = as.numeric(year)) %>%
```

```
  group_by(year_raw, amount_raw) %>%
  summarise()
```

Conducting the ANOVA analysis

```
stat_data.aov <- aov(year_raw ~ amount_raw, data = stat_data)
summary(stat_data.aov)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## amount_raw    1   2.19   2.190   1.749  0.188
## Residuals   182 227.79   1.252
```

Discussion

Analyses of the wind and patron data which was obtained for the TMACC revealed the trends for either of the aforementioned variables. The average number of individuals travelling up the cableway, for the periods of 2014-2017 ranged between 3675-4588. In contrast the average number of individuals travelling down the cableway for the peak season periods of 2014-2017, ranged between 4204-5159 (Figure: 2). This revealed that during the peak season period, there are more individuals travelling down the cableway in contrast to those travelling up. The highest patronage over the past four seasons was exhibited during 2015, while the lowest patronage was exhibited during 2014–travelling both up and down (Figure: 2).

Peak season patronage was grouped into two categories and represented graphically (Figure: 3). The graphical representation of this comparison, revealed a higher number of patrons during the week, in contrast to patronage over weekends. This trend was exhibited for all peak seasons from 2014-2017 with the exception of 2016. A students t-Test (Welch Two Sample t-Test), revealed that a significant difference was exhibited between weekday patronage and weekend patronage of the cableway ($P < 0.05$).

Examination of annual winds speeds (2014-2017), indicated that average wind speeds decreased over the peak seasons of 2014 and 2017 (Figure: 4). In contrast we find that the average wind speed increased over the peak seasons of 2015 and 2016. The trends in the annual number of patrons revealed an inverse relationship between the annual number of patrons (Figure: 5) and the average wind speeds for each individual season (Figure: 4). This relationship was further examined with a correlation test (Pearson's Product-Moment Correlation) between the average wind speed (2014-2017) and the total number of patrons (2014-2017) (Figure: 6). The correlation test revealed a negative, albeit weak, correlation between average wind speed(km/h) and the number of patrons per day for the peak seasons 2014-2017.

Differences in the total number of patrons per year (2014-2017) were examined using a one way analysis of variance (ANOVA). The test revealed that there was no significant difference in the number of patrons for the peak seasons 2014-2017.