**PROGRAMMING FOR DATA ANALYSIS**

**U.S Weather Report Data Management**

APU2F2102CS(IS)

Abiyyu Taj Mahasin Bagindo

TP058652

Date Assigned: March 17th 2021

Date Completed:

**Introduction**

The objective of this project is to analyze the annual weather data report from the United States using the R programming language to help decision making. This expansive weather data contains multiple nation-wide variables such as temperature, rainfall, air pressure, sunshine, and more. The data records the aforementioned variables of every day in United States during a span of a whole year. Therefore, a wide analysis can be made, and multiple conclusions can be reached.

**First Question – When is the best time to dry your clothes?**

**Analysis 1-1: Which days have zero rainfall?**

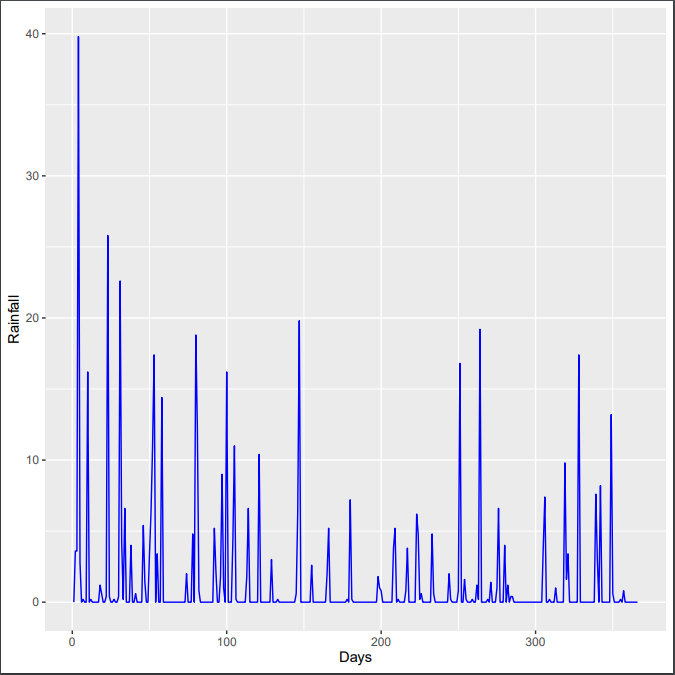


Figure Rainfall graph

The first analysis is to filter out days that do not rain, because the last thing you want when drying your clothes is for them to be wet again. The code below starts with the initialization of a NULL vector with the name *rainyday* which will be used to store the days that rain. A variable called *rainydayindex* is then assigned a value of 1. A for loop that loops 366 times for each day of the year that contains an if statement then starts. The if statement checks whether the contents of the second column of the *rainfallframe* data frame, which is used to store the rainfall data is below 0.2 or not because 0.2 is the lowest value of rainfall. If it is lower than 0.2, that day does not rain and the value of the counter variable *i* is then assigned to the index *rainydayindex* of the vector *rainyday*. The *rainydayindex* variable is then incremented by 1 and the loop repeats.

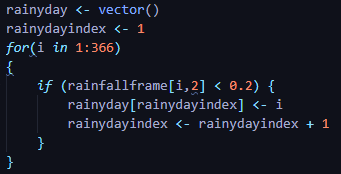


Figure Rainy days filter code

**Analysis 1-2: Which days have high evaporation rate?**

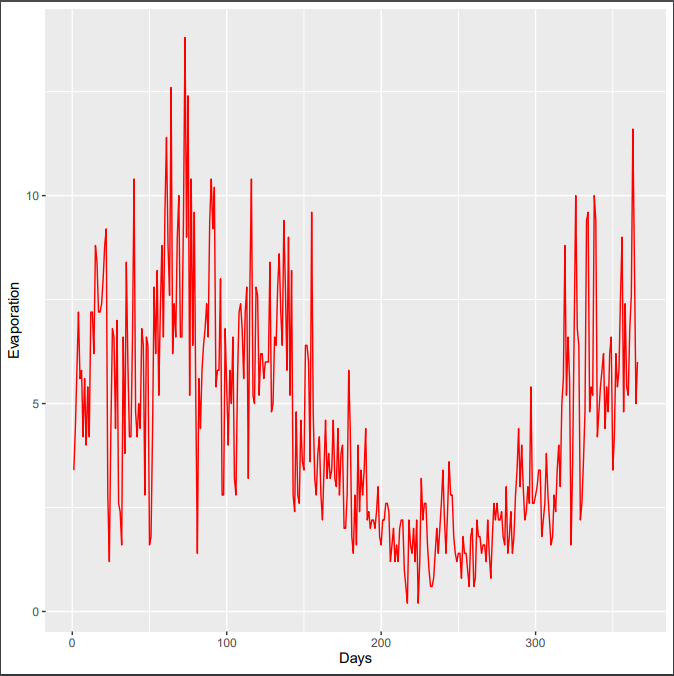


Figure Evaporation graph

The second analysis is to filter out the days that have above average evaporation rate. The higher the evaporation rate, the faster the water on your clothes will evaporate, therefore the faster it will dry. The algorithm used in the code below is the same as the one used in the previous analysis only with the variables’ names changed with the only addition being the way the filter value is obtained. As opposed to it being 0.2 due to 0.2 being the lowest value of rainfall. The value that is used to filter out the days based on evaporation rate is the average rate of evaporation which is 4.5; which has been rounded down to a 4.

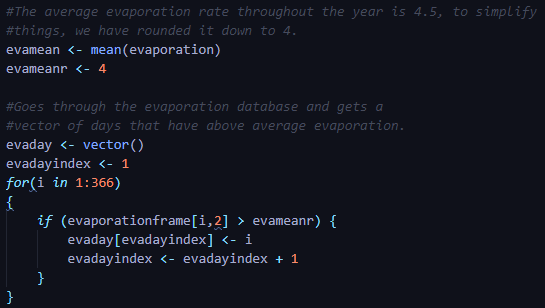


Figure Evaporation mean and filter code

**Analysis 1-3: Which days have above average sunshine?**

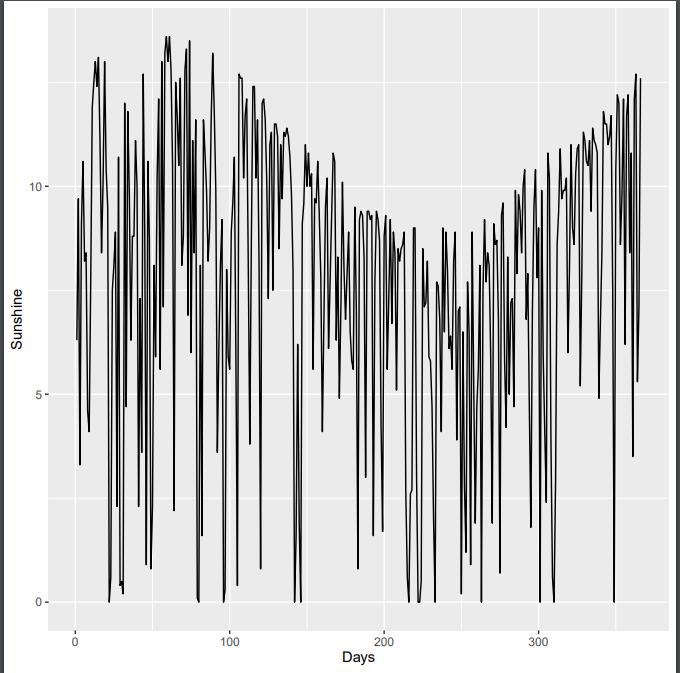


Figure Sunshine graph

The whole activity of drying clothes naturally, is only possible due to the energy carried by sunlight. Therefore, the third analysis is to filter out days with above average sunshine. The algorithm used is the same as the evaporation algorithm in Analysis 1-2. With only variables’ names being changed.

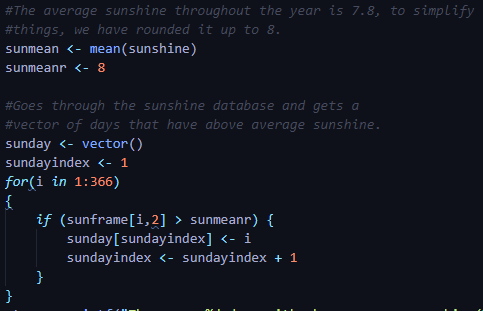


Figure Sunshine mean and filter algorithm

**Question 1: Results**

By using the three aforementioned analysis methods, it has been found that there are 264 days that do not rain, 185 days with above average evaporation rate (4), and 214 days with above average sunshine (8). However, out of all these hundreds of days, only one day is a subset of all three categories. Day 366 is the best day to dry your clothes with 0 rainfall, 6 evaporation rate, and 12.6 sunshine.

**Second Question – When is the best time to jog in the evening?**

**Analysis 2-1: Which days have zero rainfall?**

**Analysis 2-2: Which days have comfortable levels of humidity in the evening?**

**30-50**

**Analysis 2-3: Which days do not have windy evenings?**

**Third Question – when is the best time to fly from the east to the west?**

**Analysis 3-1: when does the wind blows to the west?**

**Analysis 3-2: which days do not have cloudy mornings?**

**Analysis 3-3: which days do not have heavy rain?**

**Fourth Question – ski resort travel**

**Ana 4-1: when winter**

**Ana 4-2: when cloud**

**Ana 4-3:**

**Fifth Question – when storm?**

**Ana 5-1: when low pressure**

**Ana -52: when heavy rain?**

**Ana 5-3: when fast gust?**

**Conclusion**