**Project 2 Summary Report**

**Tamani Kaid**

**Marcus Garnham**

**Abdulahi Ghelle**

Tamani Kaid

The purpose of the project was to see if the temperature and the rainfall has increased in England in the period of 1999 to 2019.

At first we were planning to use OpenWeather and Meteomatics, however, they were not free and were not usable. So Marcus has found free data from Kaggle in csv format, which I used and was able to extract required data.

I read the data through python which was simple, it then gave me 8 columns which were the year, months, tmax, tmin, af, rain, sun and stations.

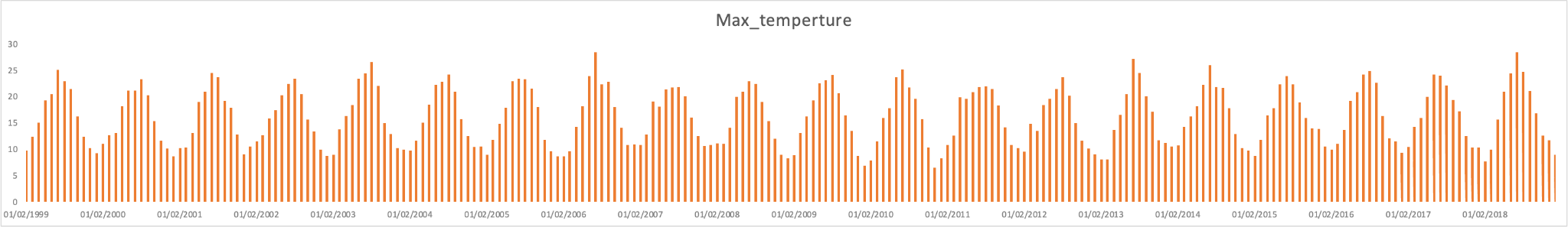
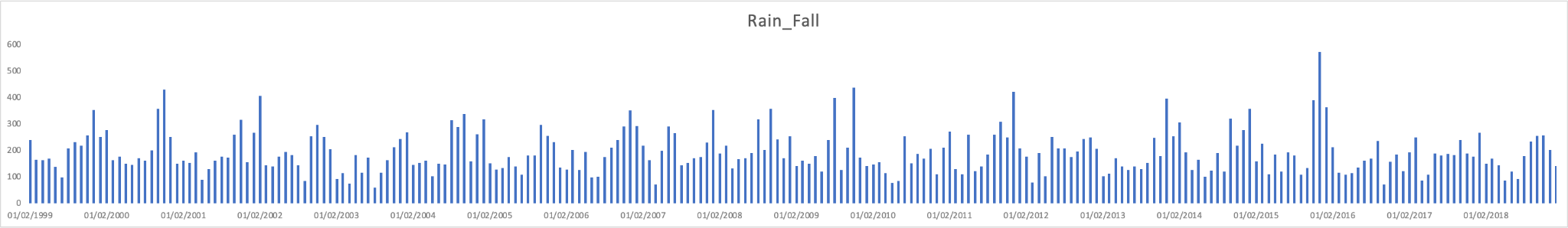
It also gave dates from 1941 to 2020.

I then filtered out the data as i didnt need it, I only kept the stations, max temperature and the rainfall column. I also created a new datetime column from the year and the months. I then set the daternage from 1999/01/01 to 01/01/2019 - “which was wrong, it was meant to be 01/12/2019 but it was already sent to the group.”

I then exported the python dataframe into csv and sent it to abdulahi to complete his section.

From this data I have found out that on average the max temperature has increased and the rainfall has decreased over the years in england.

Below are the graphs that I produced in excel.



Marcus Garnham Work Done Project 2

During this project the first thing I did was looking up data we could use for our weather task/database. I managed to find some free data on Kaggle that contained Met Office data for their weather stations for several decades. The reason we used the website Kaggle version and not the Met Office official website is we couldn’t get data very far back without paying for it and it was all neatly packed in one file. This was the dataset that Tamani worked on and was an excel csv file.

I had to search much longer for some complementary data for myself to work on. Which I eventually found on the Met Office website at the following address;

<https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-and-regional-series>

I selected the region as the uk as our stations from the previous dataset are located all over the UK and not limited to just one. So we couldn’t use data from just England, Wales, Scotland or Northern Island but had to use UK data as a whole. Maybe if we had more time we could have done a regional breakdown but an overall view had suffice. So I chose to set the region to the UK and selected three parameters which gave me three separate tables. The parameters I selected were Rainfall, Max Temp (Temperature) and the Min Temp (Temperature). This way we can compare the overall picture from the country as a whole with the individual weather stations. See if they follow the rough trend of the data.

Note the Met Office records this data in mm for the rainfall and Celius for the temperature.

So for my three tables I got from the Met Office I had to save the output pages as a txt file. I then had to save these files as a csv file as they were pretty much impossible to use as they were since they had different lengths of separation between the data columns. So I could read the txt but there was no set delimiter so everything was a complete mess. By saving the file as an csv/excel file I was able to use the data tab and then the text to columns to break everything out of a single column. I had to select the fixed width as there was no set delimiter from the txt file.

Now with the new csv files and tables I was able to read them into my python. Once I had my new data frames all I had to do was cut down what data to what I needed. I needed to keep the year and the observations for each month. So I cut from each table the columns win (winter), spr(spring), sum(summer), aut(autumn) and ann (annual). I also had to cut the data so we dropped the unwanted years. Then I renamed each column besides the ‘year’ column as that would be shared by all of them and how we would join the three tables later in Abdulahi’s part.

The reason the month columns were renamed is so it would be easily recognisable what they stood for. I couldn’t have my data with column name month as that would be confusing and cause problems joining them later on. So I added rainfall, min\_temp, and max\_temp to the month names. So you would know what month it was but also if the column represented temperature or rainfall in final database.