

7COM1079-0901-2024 - Team Research and Development Project

Final report title: Analyzing Weather Data in R

Group ID: A82

Dataset number: DS031

Prepared by: Asif Ali

University of Hertfordshire  
Hatfield, 2024

## Table of Contents

1. Introduction	...	[3]
1.1. Problem statement and research motivation	...	[3]
1.2. The data set	...	[3]
1.3. Research question	...	[3]
1.4. Null hypothesis and alternative hypothesis ( $H_0/H_1$ )	...	[3]
2. Background research	...	[4]
2.1. Research papers (at least 3 relevant to your topic / DS)	...	[4]
2.2. Why RQ is of interest (research gap and future directions according to the literature)	...	[4]
3. Visualisation	...	[4]
3.1. Appropriate plot for the RQ <i>output of an R script (NOT a screenshot)</i>	...	[4]
3.2. Additional information relating to understanding the data (optional)	...	[5]
3.3. Useful information for the data understanding	...	[6]
4. Analysis	...	[7]
4.1. Statistical test used to test the hypotheses and output	...	[7]
4.2. The null hypothesis is rejected /not rejected based on the p-value	...	[8]
5. Evaluation – group’s experience at 7COM1079	...	[8]
5.1. What went well	...	[8]
5.2. Points for improvement	...	[8]
5.3. Group’s time management	...	[8]
5.4. Project’s overall judgement	...	[9]
6. Conclusions	...	[9]
6.1. Results explained.	...	[9]
6.2. Interpretation of the results	...	[9]
6.3. Reasons and/or implications for future work, limitations of your stud	...	[9]
7. Reference list	...	[9]
8. Appendices	...	[10]
R code used for analysis and visualisation.	...	[10]

1. Evaluation – group’s experience at 7COM1079

1.1. What went well

Cooperation and proper coordination were the keys to the performance of the project. That means one had to have regular meetings where goal statements were discussed; decentralization of tasks helped to use the members’ strengths. Statistical tests and models were employed against the obtained dataset and offered precise and robust outcomes. Furthermore, the use of graphs and charts was helpful only in explaining trends that were arrived at. These coding tools such as R helped to make the analysis easier in terms of reliability in addition to minimizing possible mistakes.

1.2. Points for improvement

Some issues were identified as follows, nevertheless, there are some difficulties. The process of data preprocessing had some time consumption issues in the beginning, which pointed to the problem of severe time management. Some of the issues found were that the dataset had some data cleaning problems which could have been minimized by experience in data management. Moreover, the methods used in the statistical data analysis were rather complicated and hence took time for the group members to master.

1.3. Group’s time management

Time management both for teaching and learning activities was satisfactory though needed enhancement. Even though specific deliverables were produced within targeted deadlines, some of the project stages, for instance, data cleaning and hypothesis testing, were given more time. These concerns could be managed better in future projects if more specific timelines were set, and assessments of the project’s progress were more thorough.

1.4. Project’s overall judgement

In any case, the execution of the project was able to provide answers to the research questions and affirm the hypotheses. The use of statistical techniques complemented with data visualization was effective in providing insights into the given data set. Thus, the findings are of practical use to help support the detailed understanding of variations in climate and effects on specific weather.

---