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

Final report title: An analysis into the seasonality of the GEEQ coin

Group ID: A226

Dataset number: DS101

Prepared by: [Arshad Sultan - 23074483],

[Hariharan Baskar – 23074498,

Lewis Bailey – 22062632,

Vishvha Natarajan - 23065612,

Abdul Mateen Syed – 19064164]

University of Hertfordshire

Hatfield, 2024

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Serial Number** | **Index** | **Page Number** |
| **1** | **Introduction** | **4** |
| 1.1 | Problem statement and research motivation | 4 |
| 1.2 | The data set | 4 |
| 1.3 | Research question | 4 |
| 1.4 | Null hypothesis and alternative hypothesis (H0/H1) | 4 |
| **2** | **Background research** | **5** |
| 2.1 | Research papers (at least 3 relevant to your topic / DS) | 5 |
| 2.2 | Why RQ is of interest (research gap and future directions according to the literature) | 5 |
| **3** | **Visualisation** | **6** |
| 3.1 | Appropriate plot for the RQ output of an R script (NOT a screenshot) | 6 |
| 3.2 | Additional information relating to understanding the data (optional) | 6 |
| 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 | 7 |
| **5** | **Evaluation – group’s experience at 7COM1079** | **7** |
| 5.1 | What went well | 7 |
| 5.2 | Points for improvement | 7 |
| 5.3 | Group’s time management | 7 |
| 5.4 | Project’s overall judgement | 7 |
| 5.5 | Comment on GitHub log output | 8 |
| **6** | **Conclusions** | **8** |
| 6.1 | Results explained. | 8 |
| 6.2 | Interpretation of the results | 8 |
| 6.3 | Reasons and/or implications for future work, limitations of your study | 8 |
| **7** | **Reference list** | **9** |
| **8** | **Appendices** | **9** |
| A | R code used for analysis and visualisation | 9 |
| B | GitHub log output | 12 |

**1. Introduction**

* 1. **Problem statement and research motivation**

Investor behaviour in the financial markets is usually shaped by several factors. One of the biggest factors is seasonality or, more precisely, the monthly variations. Additional examination of these monthly variations in trading volume can highlight patterns that demonstrate how investors are likely to act in various months throughout the year. In this report, we explore a seasonality of high/low average trading volume in a specific month to understand the market behaviour and investor psychology.

There are several implications of this, though, such as the liquidity of the market, price movements, investment strategies, and market efficiency in general. Understanding this seasonal character of the trading volume, as highlighted by Bryman (2008), is pivotal. Campbell and Shiller (1988) assert that differences in trading volume produce significant effects on price and return.

This research focuses on identifying patterns that can improve trading, manage risk, and explain the mechanisms that determine the operation of the financial markets.

* 1. **The data set**

The dataset contains trading volume data collected over a certain period, including:

• Date (Month): The independent variable is the month of the year, an interval data.

• Volume: The dependent variable, trading volume for each corresponding month, is also interval data.

Such data would be good for understanding investor behaviour through fluctuations in trading volumes across months and how they could indicate patterns in the market at specific parts of the year. It will also show the consequence it is drawing on trading strategy and market trends.

* 1. **Research question**

Is there a correlation between average trading volume in GEEQ coins and months from August 2020 to May 2022?

The research question aims to explore the seasonal patterns in trading activity and their implications on investor behaviour and market trends.

To answer this question, we will conduct a statistical analysis of the dataset, looking into the relationship between months and trading volumes. Using techniques of correlation, we hope to find significant trends or patterns in trading activity throughout the year. We will conduct statistical tests using techniques of correlation to find patterns in trading activity.

* 1. **Null hypothesis and alternative hypothesis (H0/H1)**

Null Hypothesis (H₀): There is no correlation between average trading volume and the months of the year.

This hypothesis assumes that any variations in average trading volume across the months are purely random and not attributable to consistent seasonal trends.

Alternative Hypothesis (H₁): There is a correlation between average trading volume and the months of the year.

It is assumed that the following hypothesis will then be tested for a significant statistical relationship between the months and the average volume of trade to accept whether the trend is seasonal or not.

Consequently, it becomes meaningful to check through appropriate statistical analyses whether enough evidence can be drawn to reject H0 in support of the alternative hypothesis indicating that the month and the average trading volumes vary. **(I’m thinking something could replace this but I do not know at the moment)**

1. **Background research**
   1. **Research papers (at least 3 relevant to your topic / DS)**

Geeq coin is the native cryptocurrency that is used on the GEEQ platform. It serves as a transactional cryptocurrency with a main use case being able to pay validation networks on the GEEQ platform for their services (Conley, 2019). There are three main roles that the coin plays in the platform:

* Compensation for validators for providing provably accurate data services
* Micropayments
* Fees for ledger rental

(Geeq, 2024)

The 2018 whitepaper for the GEEQ project (Conley) highlights additional use cases. These include enabling payments for smart city for services such as parking or paying a toll to auctions on sites such as eBay or Craig’s List. These use cases are suitable due to the low transactional cost making it highly scalable as well.

Previous studies on cryptocurrencies have examined calendar effects, including time-of-day, day-of-week, and month impacts on volatility, returns, and trading volume. Kaiser (2019) looked at the Monday and weekend effect, the January effect, and the Halloween effect, finding no consistent effects in the trading activity of cryptocurrencies, but there are some reoccurring trading behaviours. In another paper (Baur, 2019) it was determined that there were no patterns in trading activity over time but there was some consistent trading occurring on certain days of the week.

* 1. **Why RQ is of interest (research gap and future directions according to the literature)**

The research question is of interest because whilst similar papers have concluded that there are not many regular patterns in cryptocurrencies in terms of trading volume and returns it might be different for the GEEQ coin. There are however research gaps when it comes to coins that have low transactional costs and can be used regularly for micropayments. Researching the monthly volume of the GEEQ coin might show regular trading behaviours across certain months or it might be like other papers where there was no regular pattern when it comes to trading.

1. **Visualisation**
   1. **Appropriate plot for the RQ**

Scatterplot is used to visualise the correlation between the two variables. Here, the X-axis represents the independent variable (Month and Year) and the Y-axis represents the dependent variable (Average Trading Volume of GEEQ coin). A regression line is also added to identify the correlation of the variables, which turns out to be a negative correlation.

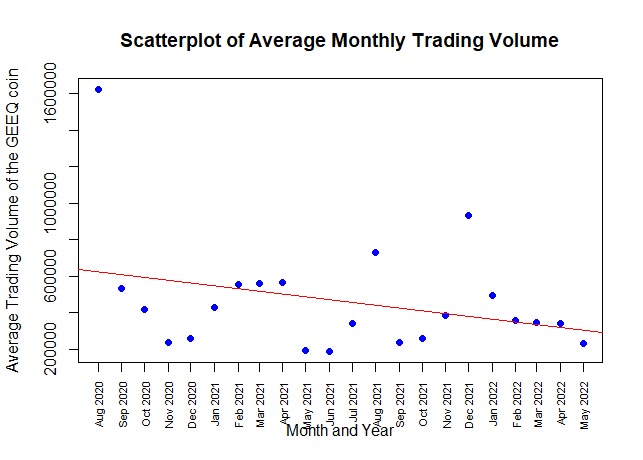


Figure 1 Output of R code for scatterplot

* 1. **Additional information relating to understanding the data**

The regression line helps to identify the density and trend of the correlation between the two variables which confirms a negative correlation. Additionally, histogram is also used to visualise the distribution of frequency of the dependent variable. By adding a normal bell curve, a right skewed distribution was identified.

* 1. **Useful information for the data understanding**

The scatterplot represents that there is a downward trend from August 2020 to May 2022 in the average monthly trading volume of GEEQ coin, which implies a negative correlation. Unlike most datapoints, few months shows high trading volume which might be due to some other external factors like decrease in price or market hype etc.

1. **Analysis**
   1. **Statistical test used to test the hypotheses and output**

The Anderson-Darling and Shapiro-Wilk tests were used to check the normality of the Average monthly trading volume data. where the result of the tests which is, the p-value was less than 0.05 (p < 0.05) meaning the data does not follow normal distribution. Therefore, the Spearman's rank correlation test was used to evaluate the relationship between time and trading volume, which produced a value of rho = -0.2343 suggesting that it is a weak negative correlation.

* 1. **The null hypothesis is rejected /not rejected based on the p-value**

The Spearman correlation test resulted in ρ value of -0.2343(rho), suggesting a weak negative correlation between time and average trading volume. However, the p-value of 0.2925 is greater than the significance threshold of 0.05, meaning the result is not statistically significant.

Thus, we fail to reject the null hypothesis, which assumes no correlation. This indicates that the weak negative trend observed in the data could be due to some random variation or noise. Therefore, there is no strong evidence to support a meaningful relationship between time (month and year) and the average trading volume in the data.

1. **Evaluation – group’s experience at 7COM1079**
   1. **What went well**

The process of implementing the hypothesis testing methods in R has been done successfully by the group, also includes Shapiro-Wilk normality test and Spearman’s rank correlation test. These tests proved the need for non-parametric tests and analysis for variable relationships. As the code made best practices, extraction, cleaning and analysis were made clear. Additionally, the group’s overall proficiency for providing clear results and manage with the complex statistical tasks were guaranteeing the project’s success.

* 1. **Points for improvement**

Although the project had strengths, the tutor’s feedback highlighted areas for improvement. To ensure the clarity and alignment in variable order the research question and hypotheses were revised. There was improvement made in dataset like outliers were addressed properly and adding the contextual information like time periods and currency. Additionally, outliers might have affected the corelation results while re-examining the scatterplot analysis. These modifications were made improvement for the quality of work.

* 1. **Group’s time management** The team did a fair job of handling the deadline, but they faced some delays due to troubleshooting code issues and missing data. To make sure the smoother progress more time was allotted to testing and debugging process. Frequent progress checks have helped to mitigate problems and ensure timely completion.
  2. **Project’s overall judgement**

The group’s continuous efforts reflect in the GitHub log, where regular commits add the projects overall progress. Further reviews addressed the corrections need to be fixed such as inconsistencies in variable and missing initialization. The codebase’s accuracy was guaranteed and collaboration was enhanced.

* 1. **Comment on the GitHub log output**
* "Consolidated all R scripts into 1 file"

- Combined the individual Histogram and Scatterplot R-Script into a single one.

* "Adding Normality and Correlation tests to the code"

- Incorporated key statistical tests.

* "Formatted and finalized the document"

- Corrected errors from the feedbacks, and structured the report.

For complete commit log, refer to Appendix B.

1. **Conclusions**
   1. **Results explained**

In this statistical analysis of the average monthly trading volume of the GEEQ coin, the relationship between time (month and year) and trading volume was a weak negative correlation, ρ = -0.2343. However, the associated p-value, 0.2925, was greater than the 0.05 significance level. The result was not statistically significant. We were thus unable to refute the null hypothesis of no correlation. These findings agree with previous works like Kaiser (2019) and Baur (2019), who report inconsistent or weak seasonal trends in trading volumes of some cryptocurrencies.

* 1. **Interpretation of the results**

The absence of statistical significance suggests that variables other than simple seasonality may influence monthly trading patterns in GEEQ, even though the data indicated a reduction. This is consistent with other studies that show calendar impacts on financial assets are occasionally detectable but are frequently erratic or eclipsed by certain market occurrences. In the case of GEEQ, variables like volatility in prices or market hype may have a greater direct impact on trading volumes.

* 1. **Reasons and/or implications for future work, limitations of your study**

To gain a better understanding of seasonal trading behaviour, future research could investigate the role of key market events, regulatory updates, or advancements in technology. One limitation of this study is the short time span of the dataset, which might not be sufficient to capture longer-term trends. Adding price data and studying how trading volume and price are connected could offer a more complete picture of GEEQ coin's market activity.

1. **Reference list**

Bryman, A. (2008) *Social research methods. 3rd edition*. Oxford: Oxford University Press.

Campbell, J.Y. and Shiller, R.J. (1988a) *'The dividend-price ratio and expectations of future dividends and discount factors'*, Review of Financial Studies, 1(3), pp. 195-228.

Campbell, J.Y. and Shiller, R.J. (1988b) *'Stock prices, earnings, and expected dividends'*, The Journal of Finance, 43(3), pp. 661-76.

Conley, J. P., 2018. *The Geeq Project White paper Version 2.0*. [Online] Available at: <https://geeq.io/wp-content/uploads/2018/08/White-paper.pdf> [Accessed 31 12 2024].

Conley, J. P., 2019. *The Geeq™ White Paper*. [Online] Available at: <https://geeq.io/geeq-white-paper-2/> [Accessed 31 12 2024].

Drik G. Baur, D. C. K. G. Z. (. L., 2019. *Bitcoin time-of-day, day-of-week and month-of-year effects in returns and trading volume*. Finance Research Letters, Volume 31, pp. 78-92.

Geeq, 2024. *Powered by Geeq*. [Online] Available at: <https://geeq.io/tokenomics-update-as-of-testnet-v1/> [Accessed 31 12 2024].

Kaiser, L., 2019. *Seasonality in cryptocurrencies*. Finance Research Letters, Volume 31, pp. 232-238.

Baur, D.G., Cahill, D., Godfrey, K. and Liu, Z.F. (2019) *'Bitcoin time-of-day, day-of-week and month-of-year effects in returns and trading volume'*, Finance Research Letters, 31, pp. 78-92.

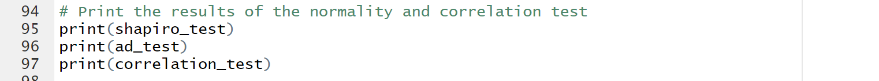
1. **Appendices**
2. **R code used for analysis and visualisation**

**Histogram, Scatterplot and test script.R**

A screenshot of a computer program

Description automatically generated

Figure 2

A screenshot of a computer program

Description automatically generated

Figure 3

Figure 4

**Rscript.log**

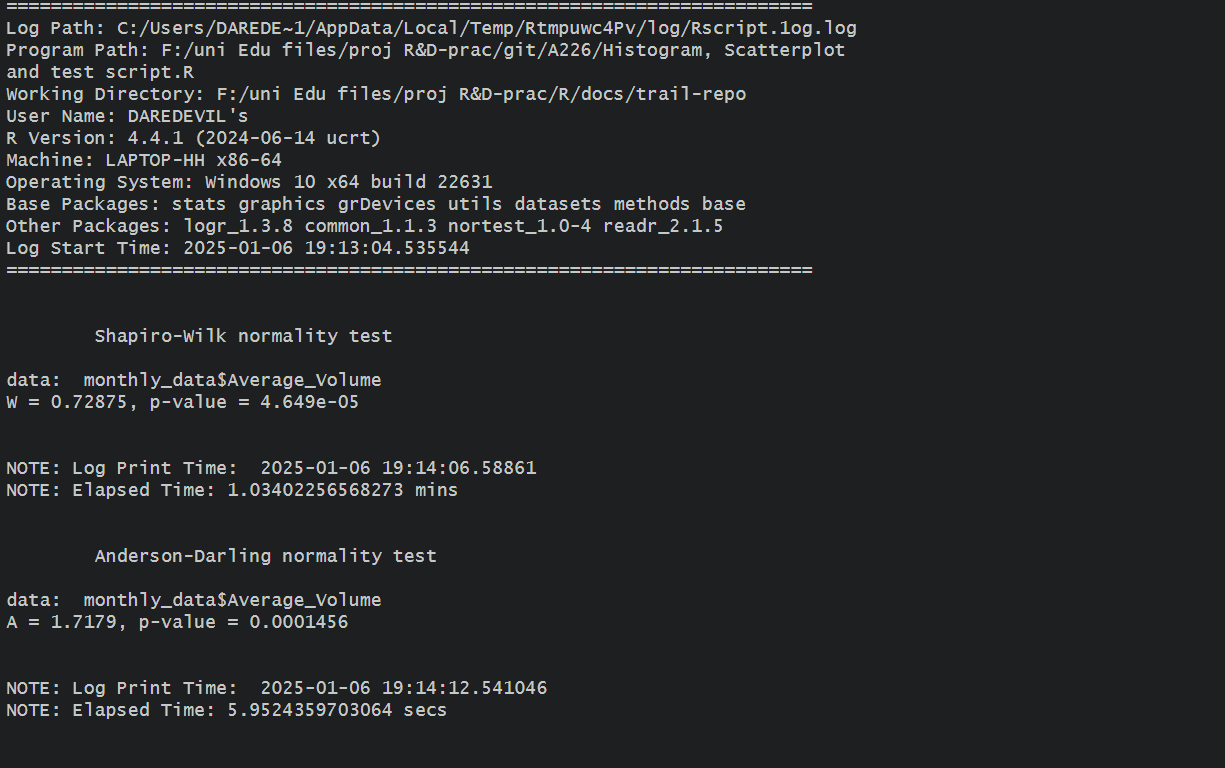


Figure 5. log file created using the package logr

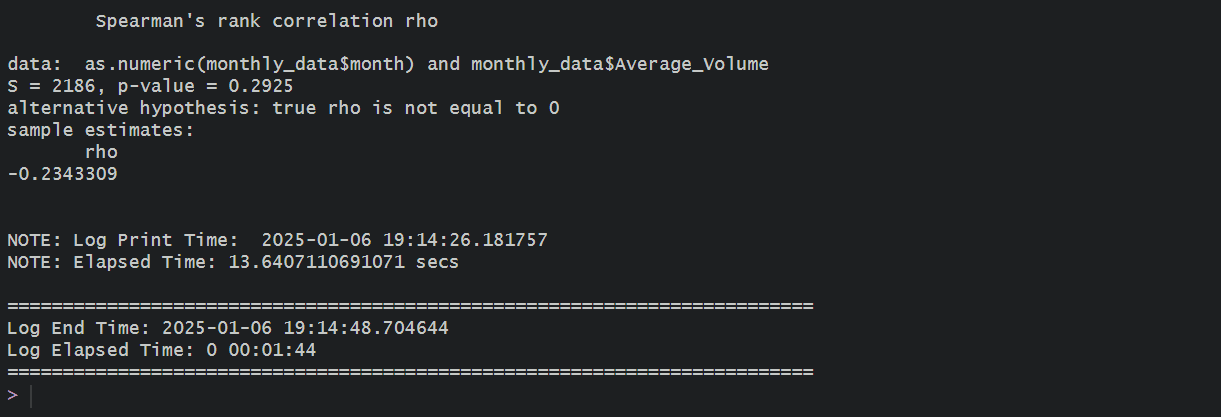


Figure 6. log file created using the package logr

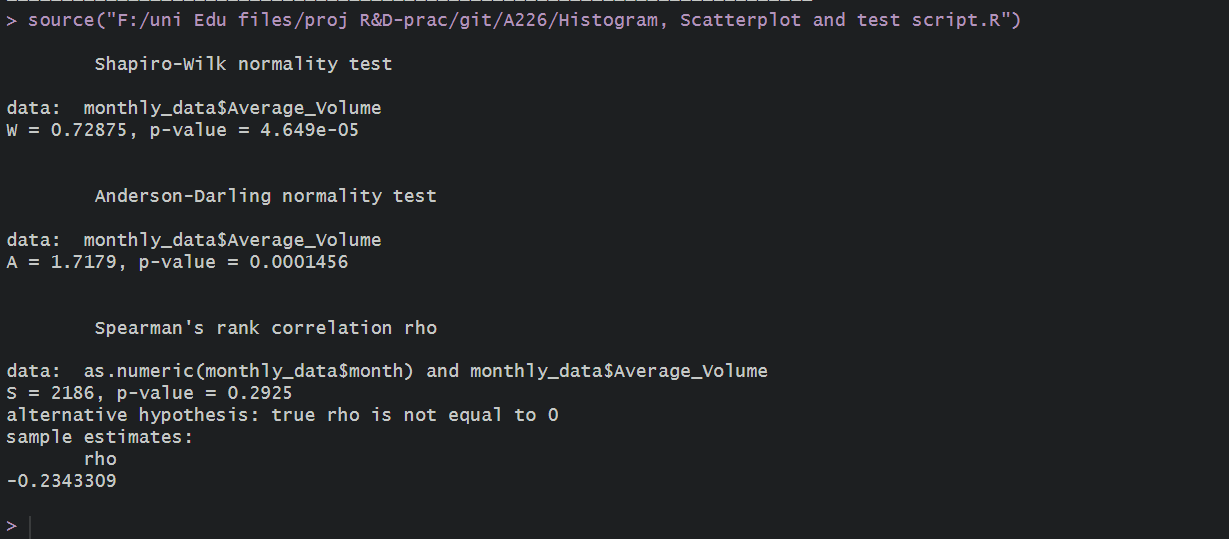


Figure 7. R code console output

**Output of R code**

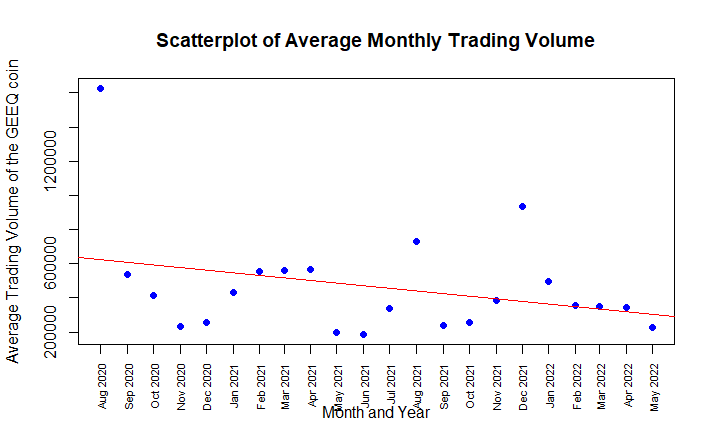


Figure 8

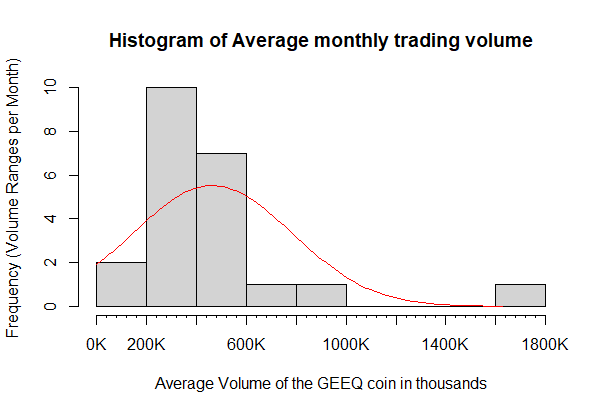


Figure 9

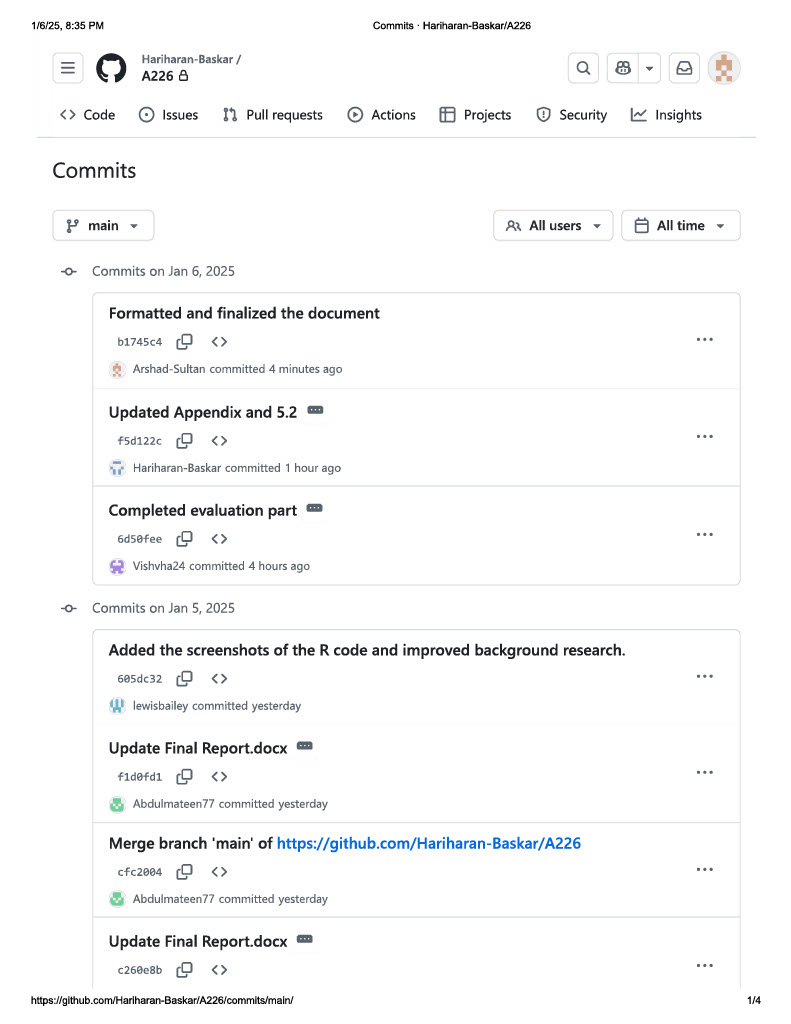
1. **GitHub log output.**

Figure 10

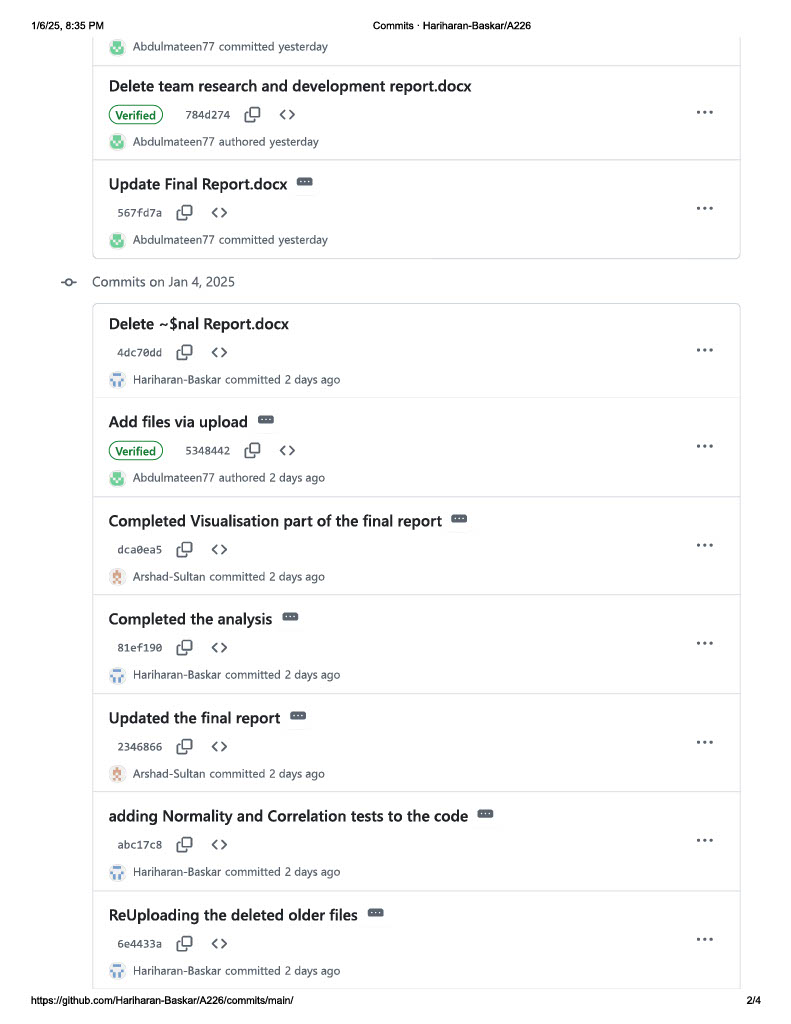
****

Figure 11

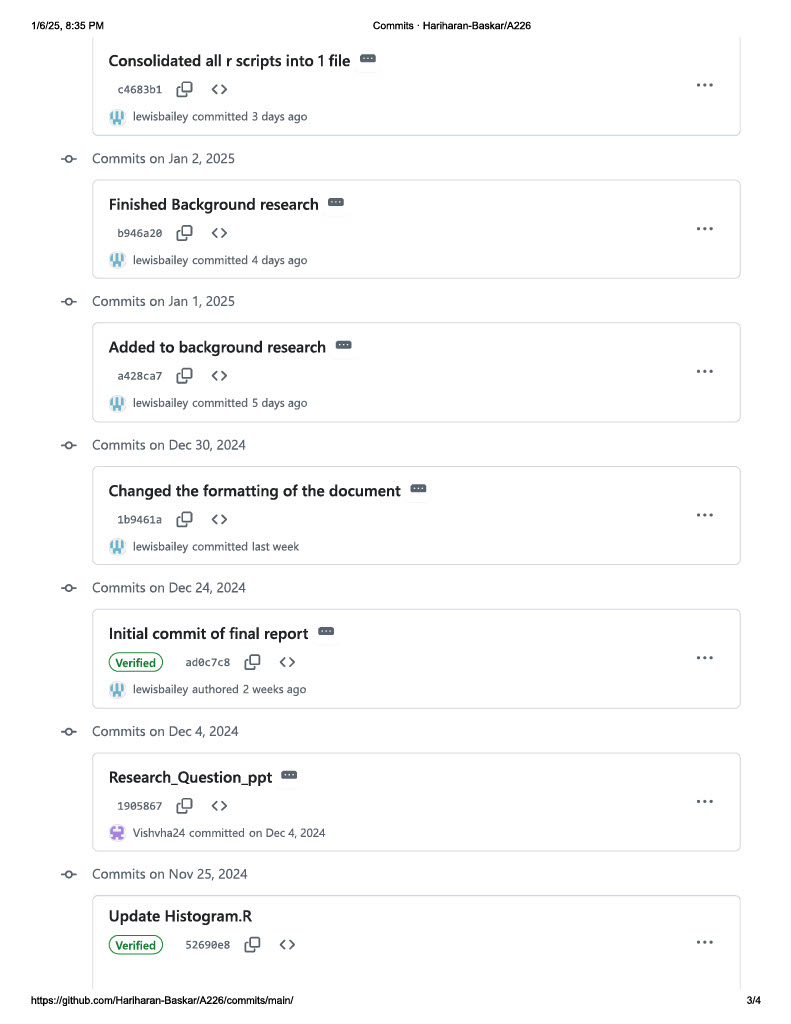
****

Figure 12

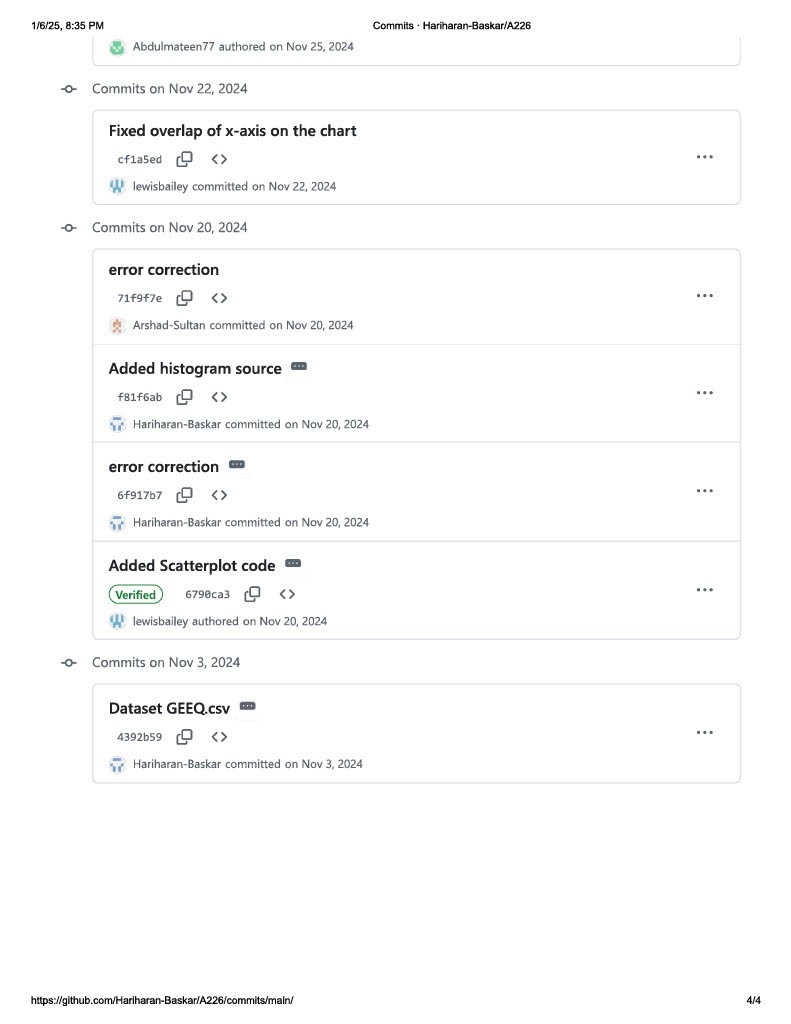
****

Figure 13