Name:	Agbo James Daniel
Matric Number:	ACE24210007
Department:	MSc. Artificial Intelligence
Course:	ATI803 Introduction to Artificial Intelligence Programming
Assignment:	**As attached**

REPORT ON LONDON WEATHER AND ENERGY CONSUMPTION ANALYSIS

This report presents an analysis of the London weather and energy consumption datasets.

The analysis aims to identify patterns and insights related to temperature trends and energy consumption during specific periods.

London Weather Data Analysis

Data Cleaning and Preprocessing:

The `london_weather.csv` dataset was loaded and cleaned by handling missing values. The most frequent value for each column was used to fill in missing entries.

Temperature Trends:

The dataset was grouped by year and month to analyze temperature trends.

The maximum, mean, and minimum temperatures were calculated for each month.

The highest maximum temperature occurred in 2020-07, 2003-08 and 2019-07.

```
202] print(grouped_by.sort_values('max_temp', ascending=False).head(10))
```

```
    Year_Month  max_temp  mean_temp  min_temp  temp_variation  year
498    2020-07      37.9  18.348387       8.7            29.2  2020
295    2003-08      37.9  21.058065       8.4            29.5  2003
486    2019-07      37.9  20.180645      10.8            27.1  2019
437    2015-06      36.7  16.843333       7.1            29.6  2015
139    1990-08      36.5  20.329032      10.4            26.1  1990
499    2020-08      36.5  20.083871       8.1            28.4  2020
330    2006-07      35.5  22.438710      12.3            23.2  2006
474    2018-07      35.0  22.309677      13.7            21.3  2018
461    2017-06      34.5  18.953333       9.4            25.1  2017
198    1995-07      34.3  20.754839      11.2            23.1  1995
```

The minimum temperatures occurred in 1981-12

```
print(grouped_by.sort_values('min_temp').head(10))
```

```
    Year_Month  max_temp  mean_temp  min_temp  temp_variation  year
35     1981-12      10.3   1.438710     -11.8            22.1  1981
36     1982-01      12.6   3.638710     -10.1            22.7  1982
85     1986-02       4.1  -0.467857      -9.6            13.7  1986
383    2010-12       9.5   1.187097      -9.4            18.9  2010
96     1987-01      10.6   1.306452      -9.1            19.7  1987
145    1991-02      12.5   2.057143      -8.9            21.4  1991
72     1985-01      13.3   1.116129      -8.0            21.3  1985
155    1991-12      13.9   4.970968      -8.0            21.9  1991
397    2012-02      16.8   4.662069      -7.6            24.4  2012
0      1979-01       9.1   0.616129      -7.5            16.6  1979
```

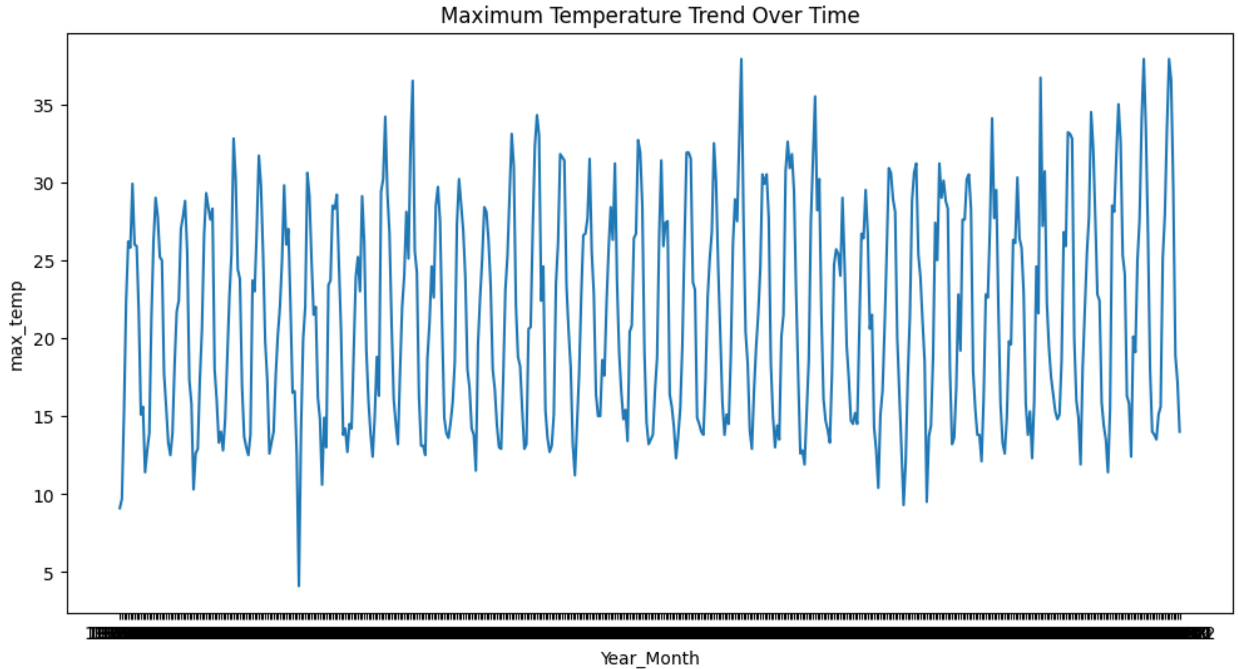
Similarly, the highest mean is in 2006-07.

```
print(grouped_by.sort_values('mean_temp', ascending=False).head(10))
```

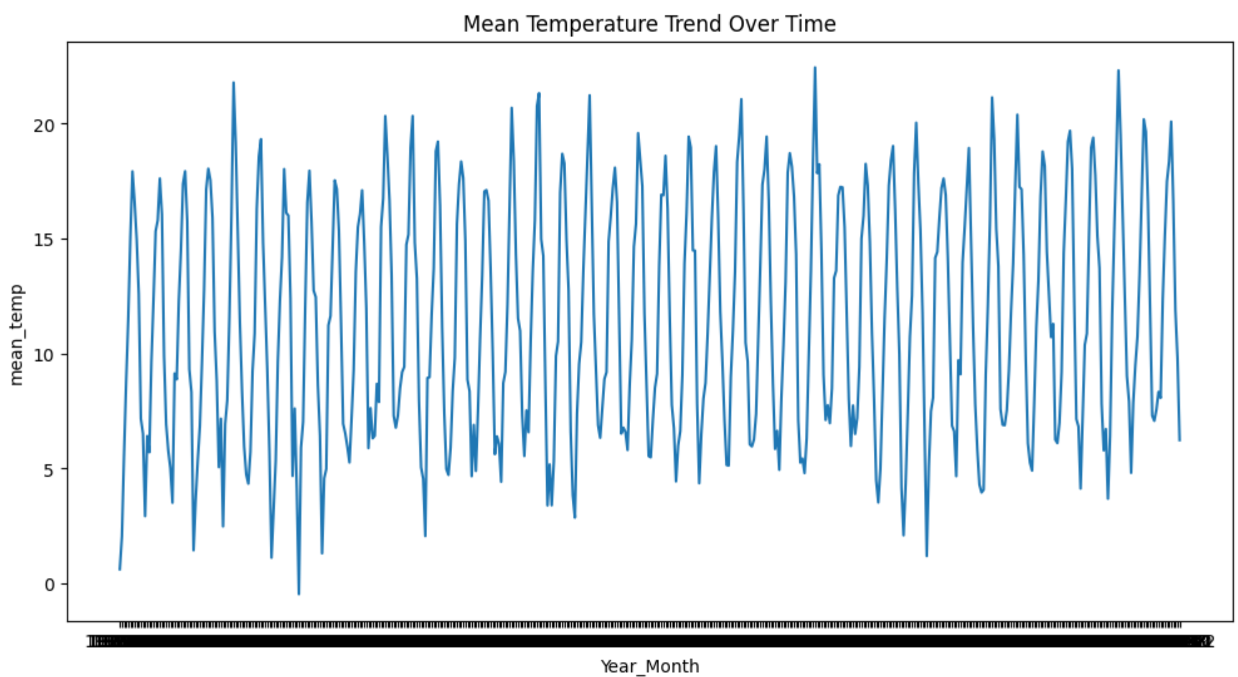
```
    Year_Month  max_temp  mean_temp  min_temp  temp_variation  year
330    2006-07      35.5  22.438710      12.3            23.2  2006
474    2018-07      35.0  22.309677      13.7            21.3  2018
54     1983-07      32.8  21.783871       8.6            24.2  1983
199    1995-08      33.0  21.322581      10.6            22.4  1995
223    1997-08      31.5  21.229032      11.6            19.9  1997
414    2013-07      34.1  21.135484      11.6            22.5  2013
295    2003-08      37.9  21.058065       8.4            29.5  2003
198    1995-07      34.3  20.754839      11.2            23.1  1995
186    1994-07      33.1  20.683871      11.9            21.2  1994
426    2014-07      30.3  20.383871      11.1            19.2  2014
```

The analysis revealed that temperature generally increased during the summer months and decreased during the winter months.

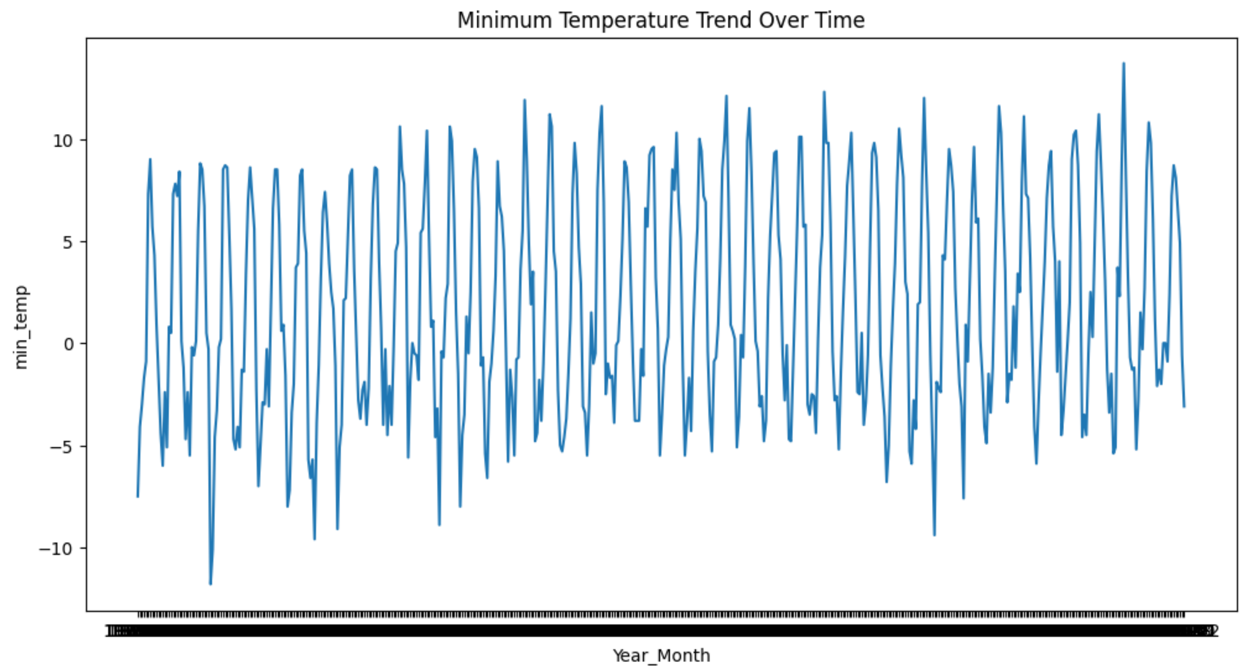
Line plots were generated to visualize the trends of maximum, mean, and minimum temperatures over time.



Maximum Temperature Trend Over Time
max_temp
35
30
25
20
15
10
5
Year_Month

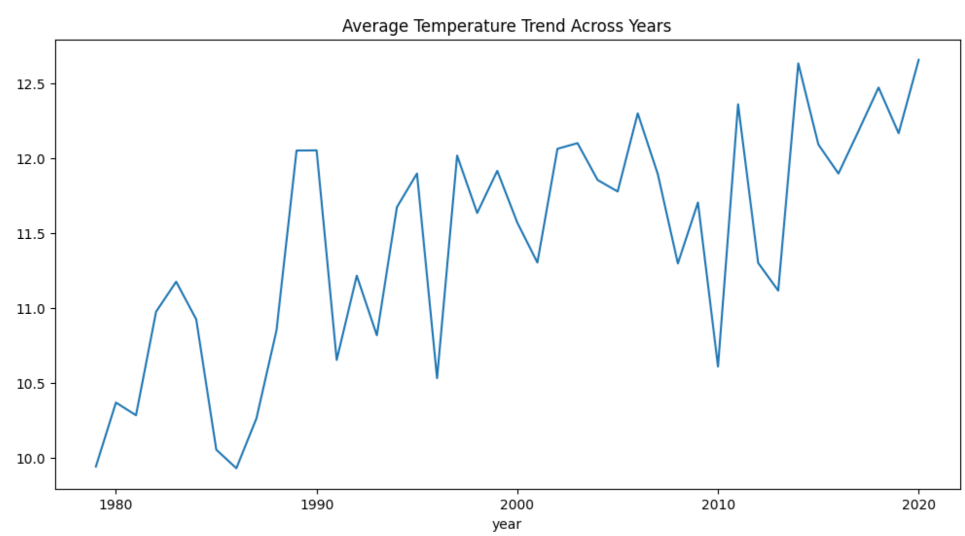


Mean Temperature Trend Over Time
mean_temp
20
15
10
5
0
Year_Month



Furthermore, the temperature variation (the difference between maximum and minimum temperatures) was calculated for each month. The months with the highest variations were 2015-06, suggesting a greater temperature range during these periods.

The average temperature across different years was computed and visualized to observe any long-term temperature changes.



Despite the fluctuations, there is a clear upward trend in average temperatures over the entire period. This could indicate a gradual warming effect, potentially linked to climate change or other environmental factors.

London Energy Consumption Analysis

Data Cleaning and Preprocessing:

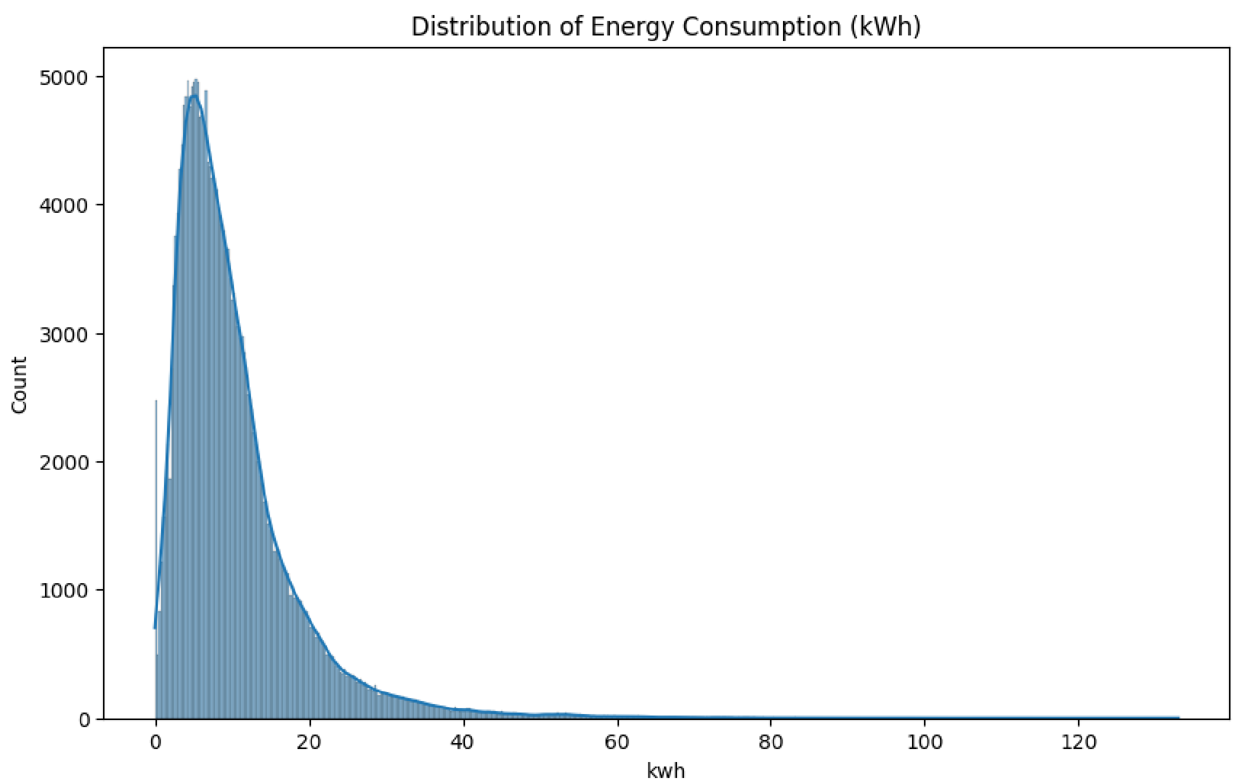
The `london_energy.csv` dataset was loaded and cleaned by handling missing values and converting the 'date' column to a datetime data type. Missing values were filled using the mode for each column.

Energy Consumption Patterns:

The dataset was merged with the `london_weather.csv` dataset based on the 'date' column to explore the relationship between weather and energy consumption.

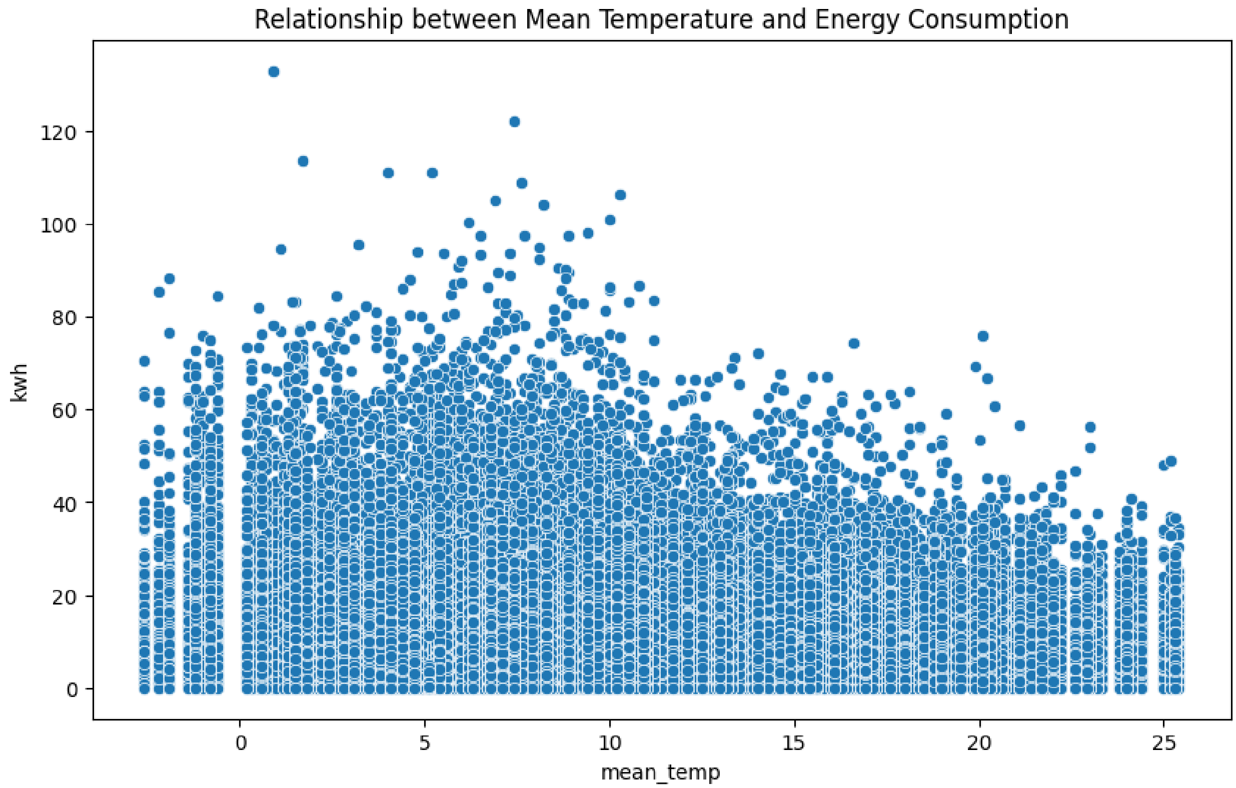
The top 10 homes with the highest total energy consumption for December 2011, December 2012, and December 2013 were identified, indicating variability in energy usage across homes.

A histogram analysis displayed the distribution of energy consumption (kWh). The distribution appears to be right-skewed(Positive).



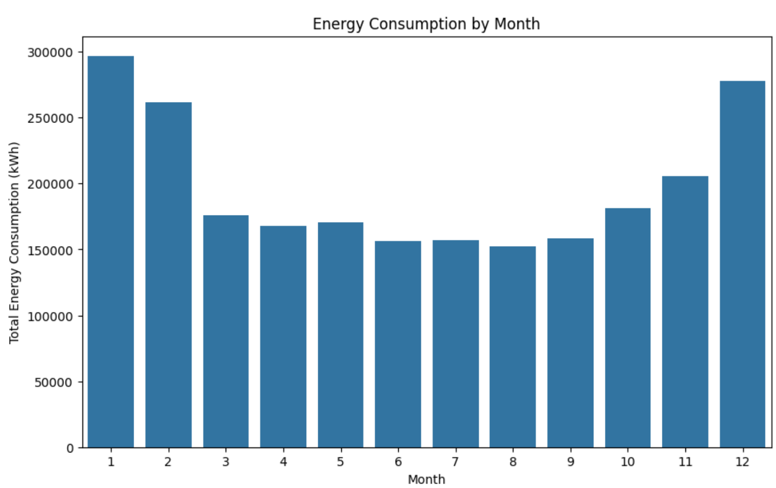
The histogram effectively illustrates that energy consumption is heavily skewed toward lower values, with a long tail representing a small number of high-energy users. This distribution provides valuable insights into energy usage patterns and can guide targeted interventions or further analysis.

A scatter plot revealed the relationship between mean temperature and energy consumption.



There is a noticeable increase in energy consumption at both very low and very high mean temperatures, suggesting that extreme temperatures (both cold and hot) lead to higher energy usage, likely due to heating or cooling needs.

Further investigation on monthly energy consumption patterns revealed that energy consumption starts increasing from September to January and decreases from February to August.



The top 10 homes with the highest total energy consumption were also identified, indicating variations in energy usage across different homes.

```
lclid
MAC000105    40441.950
MAC000049    30104.075
MAC000153    25492.898
MAC000116    25362.122
MAC000693    23550.720
MAC000096    23340.507
MAC000085    22356.677
MAC000034    22145.138
MAC000697    22017.794
MAC000040    21690.353
```

I would suggest:

- Investigate the correlation between temperature and other weather variables like sunshine and rainfall.
- Conduct further analysis on seasonal temperature variations and their impact on energy consumption.
- Compare the temperature trends with historical weather data for a more comprehensive analysis.
- Investigate factors influencing energy consumption variations across different homes.
- Develop models to predict future energy consumption based on historical data and weather patterns.

In summary, the analysis of the London weather and energy data provided valuable insights into temperature trends, energy consumption patterns, and the potential relationship between weather and energy consumption. The findings can support the development of strategies for optimizing energy consumption and promoting energy efficiency within London.

QUESTION 2 REPORT

Report on GP Prescribing Data Analysis for July 2024

This report presents an analysis of the GP prescribing dataset for the month of July 2024. The analysis aims to identify key patterns and insights related to prescription trends, costs, and GP performance. The findings can be used to inform decision-making and resource allocation within the healthcare system.

Data Overview

The dataset includes information on various aspects of GP prescriptions, including the practice (GP), the item/prescription name (AMP_NM), total quantity, total items, gross cost, actual cost, etc. Data preprocessing was performed to handle missing values, invalid entries, and data type conversions.

Analysis of Prescription Trends

Top 10, Least 10, and Average Prescriptions:

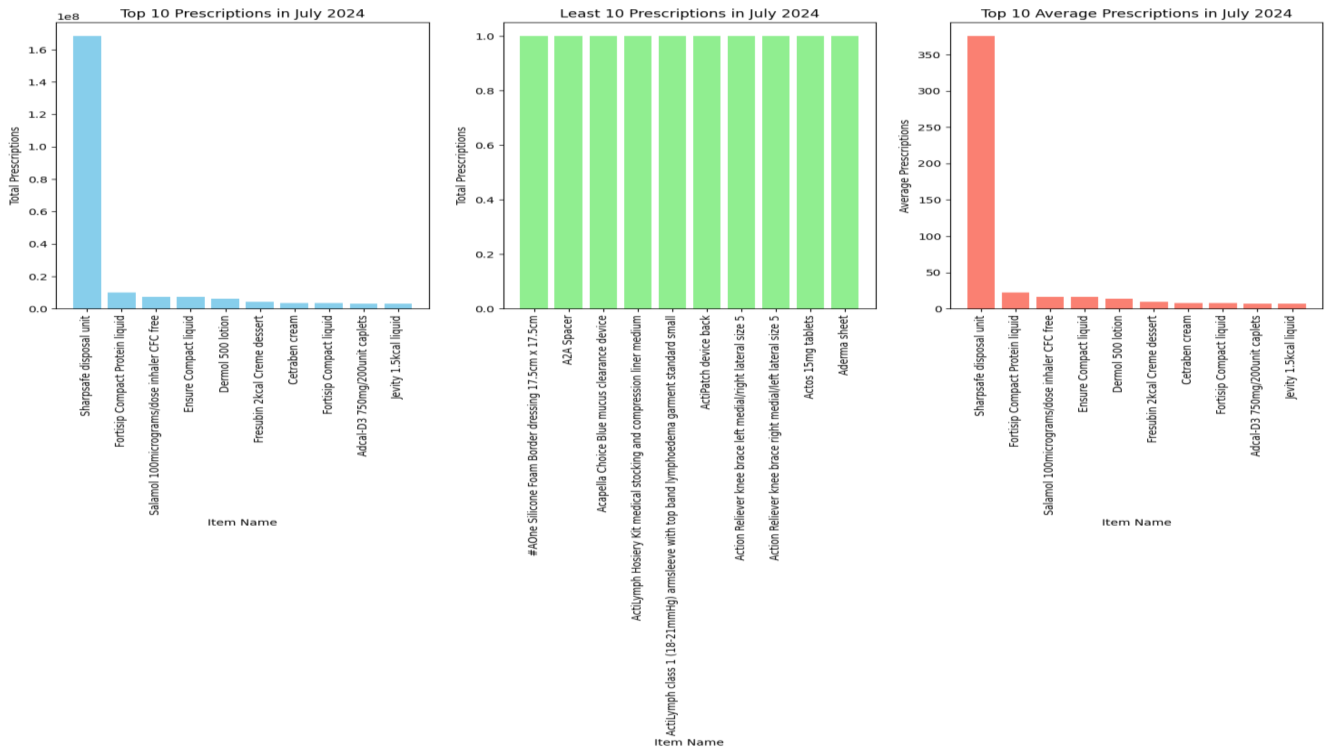


Figure 1

The bar charts above illustrate the top 10, least 10, and average prescriptions based on the total quantity prescribed during July 2024. This gives us insights into which medications are most and least commonly prescribed during the month. We can observe from the top 10 prescriptions that Shapesafe Disposal Unit was the highest prescribed. In the least 10 prescribed items, we can observe ten items that are least prescribed with total quantity of one each. The average prescriptions figure helps us understand the overall distribution of prescriptions across different items, which can help in identifying potential discrepancies or irregularities in prescription patterns

Gross Cost (£) Trend for Top 15 GPs

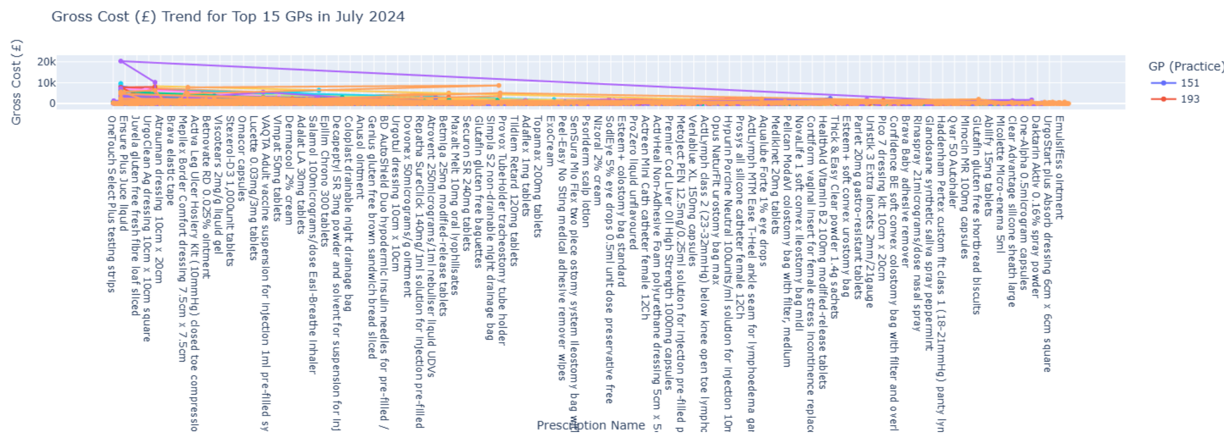


Figure 2

This line plot visualizes the Gross Cost (£) trend for the top 15 GPs based on total Gross Cost during July 2024. Each GP is represented by a distinct line, making it easy to compare their performance. We can observe from the chart that each GP demonstrates unique spending patterns, suggesting differences in prescribing habits, patient demographics, or service offerings. The hover feature allows us to see the exact Gross Cost (£) for each prescription.

Actual Cost (£) Comparison for Top 15 GPs

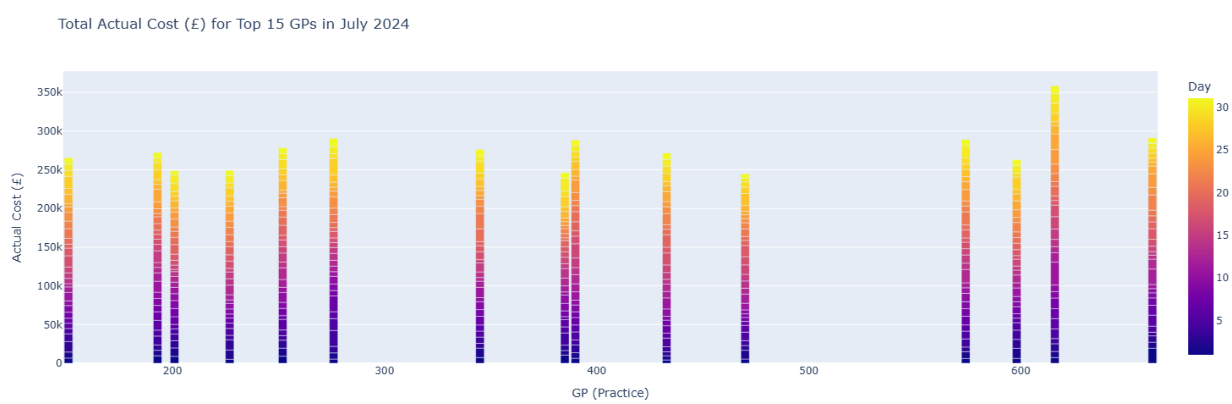


Figure 3

The bar plot visualizes the summation of Actual Cost (£) for the top 15 GPs in July 2024, with a distinct hue for each day. This provides insights into daily variations in Actual Cost for these GPs. We can observe the variability in actual costs among practices and the importance of analyzing daily spending patterns to identify trends and potential areas for cost optimization. The hover functionality enables viewing the precise Actual Cost (£) for each day.

Total Items vs. Total Quantity Scatter Plots for Randomly Selected 20 GPs

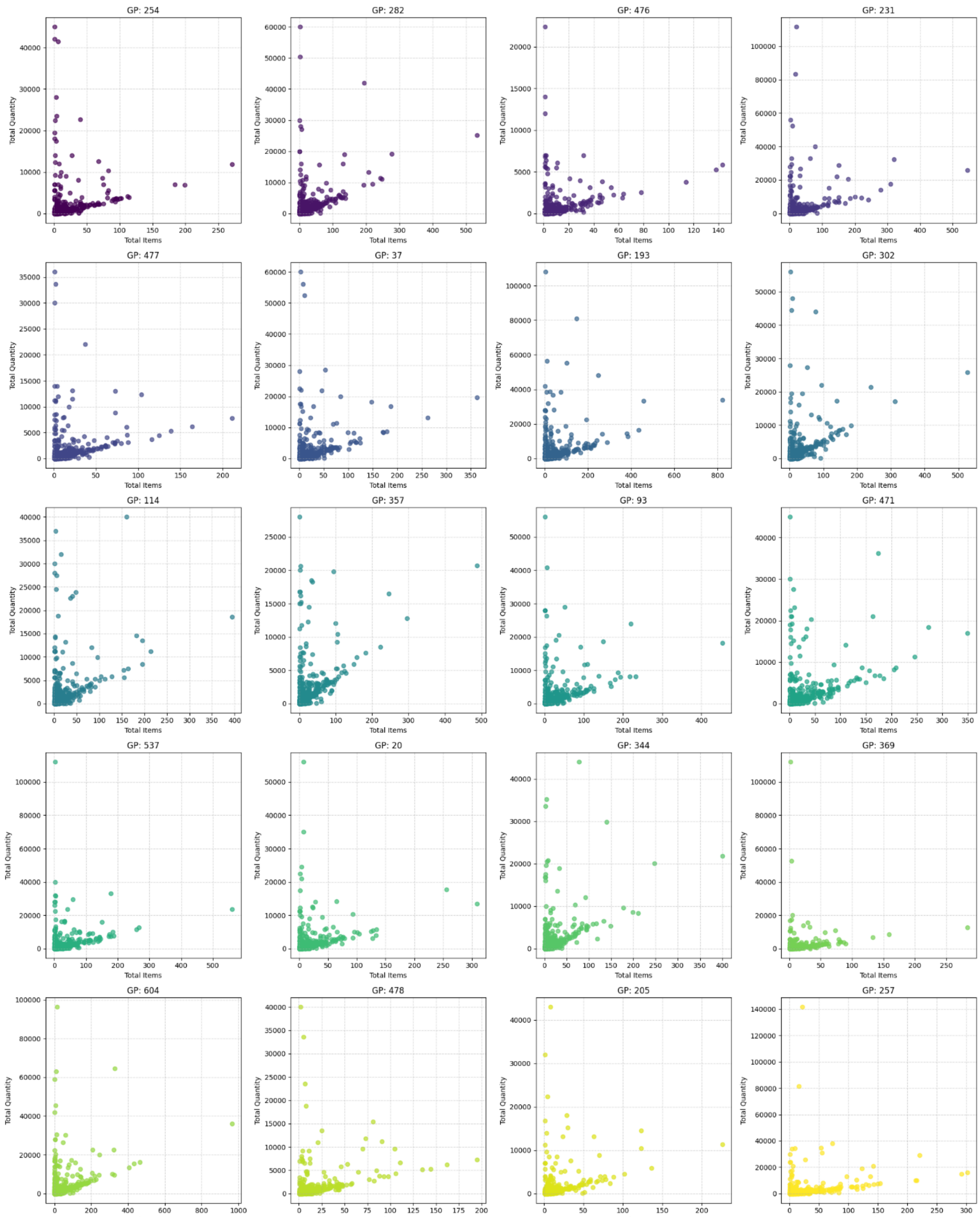


Figure 4

This grid of scatter plots compares Total Items and Total Quantity for randomly selected 20 GPs during July 2024. Each scatter plot represents a single GP. The plots allow us to assess the relationship between the total number of items prescribed and the overall quantity prescribed for each GP. Based on observations from the plots, we can identify the diversity in how GPs manage prescriptions and can inform strategies for optimizing resource allocation and identifying areas for improvement.

Conclusion

The analysis of the GP prescribing dataset for July 2024 has provided valuable insights into various aspects of prescription practices. Understanding the top and least commonly prescribed items, the cost trends for GPs, the daily variations in Actual Cost, and the relationship between Total Items and Total Quantity can be highly beneficial for healthcare stakeholders in making informed decisions regarding resource allocation, policy adjustments, and strategic planning in optimizing prescription practices.

References:

1. london_weather.csv
2. london_energy.csv
3. gp-prescribing-july-2024.csv
4. https://drive.google.com/file/d/1jq3cJinJU1FdpNm6qXlNgoiivpAiysFB/view?usp=sharing
5. https://www.data.gov.uk/dataset/176ae264-2484-4afe-a297-d51798eb8228/prescribing-by-gp-practice-presentation-level
6. https://chat.qwen.ai/