A look into medical prescriptions in the NHS and how they relate to wealth distribution.

Pat Callanan*
paca8198@colorado.edu

Abstract

This study investigates the relationship between NHS-funded prescription patterns and socioeconomic factors across various postal codes in England and Wales, with a particular focus on wealth distribution. As a publicly funded healthcare system, the NHS aims to provide equitable access to medications and other prescribed items, independent of an individual's ability to pay. However, the potential influence of regional socioeconomic status on the distribution of prescriptions raises important questions about healthcare equity.

Through the integration of NHS prescription data with socioeconomic data from the Office for National Statistics (ONS), this research explores whether wealthier areas receive different types or frequencies of prescriptions compared to less affluent regions. The study employs a combination of descriptive statistics, correlation and regression analyses, geospatial visualization, and qualitative insights to uncover any significant disparities in prescription patterns.

Key findings reveal that there are indeed significant variances in NHS prescription spending across income groups, with wealthier regions exhibiting higher overall prescription expenditures. Additionally, the study identifies potential causes of these disparities, including higher illness prevalence in lower-income areas and the possible influence of private healthcare utilization in wealthier regions.

The results of this study underscore the need for targeted policy interventions to address the observed inequities in NHS prescription distribution. By understanding the socioeconomic factors that influence prescribing patterns, policymakers can ensure that NHS resources are distributed more equitably, ultimately contributing to a more just and effective healthcare system across the UK.

ACM Reference Format:
Pat Callanan. 2018. A look into medical prescriptions in the NHS and how they relate to wealth distribution. . In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

*Both authors contributed equally to this research.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

As the NHS is a publicly funded healthcare system, it ensures that the provision of medicines and other prescribed items remains consistent across the UK. This uniformity in care and treatment is designed to be independent of an individual's ability to pay, promoting equity in healthcare access. However, this raises important questions about the distribution of NHS-funded prescriptions and whether it is truly independent of the socioeconomic status of the areas served by clinics.

One key question to address is whether the distribution of NHS-funded prescriptions is influenced by the wealth of the areas where clinics are located. This involves examining if there is a correlation between the affluence of a region and the types of prescriptions issued. Are certain medications more commonly prescribed in wealthier areas compared to less affluent ones?

Additionally, it is important to investigate any perceivable links between the wealth of an area and the prevalence of specific prescriptions. For instance, are antidepressants or antibiotics prescribed at higher rates in economically disadvantaged regions? Understanding these patterns can shed light on the broader implications of socioeconomic factors on healthcare provision and patient outcomes within the NHS.

By exploring these questions, we can gain valuable insights into the equity and effectiveness of the NHS in delivering healthcare services, ensuring that all patients receive appropriate and necessary medical care regardless of their socioeconomic status.

2 Prior Literary research

There is a decent body of research focusing on healthcare spending within the NHS, exploring various aspects such as overall expenditure, cost drivers, and trends in healthcare services. For instance, research has shown that NHS expenditure has significantly increased over the past decade, with hospital-based care being the largest expenditure category (Rodriguez Santana, I., Aragón, M., Rice, N. *et al.* Trends in and drivers of healthcare expenditure in the English NHS: a retrospective analysis. *Health Econ Rev* 10, 20 (2020). https://doi.org/10.1186/s13561-020-00278-9) (Naser, A.Y., Alwafi, H., Al-Daghastani, T. *et al.* Drugs utilization profile in England and Wales in the past 15 years: a secular trend analysis. *BMC Prim. Care* 23, 239 (2022). https://doi.org/10.1186/s12875-022-01853-1).

However, there has been comparatively less research specifically examining the distribution of NHS-funded prescriptions in relation to socioeconomic factors such as our focus being wealth distribution. While studies have analyzed general trends in prescription rates and the impact of specific events (e.g., the COVID-19 pandemic) on prescription patterns, there is a still a notable gap in literature addressing how these prescription rates vary across different socioeconomic strata. For example, the rise in antidepressant prescriptions during the pandemic highlighted the role of economic stress and social isolation in influencing prescription trends, but did not extensively compare these trends across different wealth brackets(Rabeea, S.A., Merchant, H.A., Khan, M.U. *et al.* Surging trends in prescriptions and costs of antidepressants in England amid COVID-19. *DARU J Pharm Sci* **29**, 217–221 (2021). https://doi.org/10.1007/s40199-021-00390-z).

This gap underscores the need for focused research to understand if and how the socioeconomic status of a region affects NHS prescription distribution. Such insights could reveal potential inequities in healthcare provision and inform policy adjustments to ensure a more equitable healthcare system.

2.1 Data Sets

Primary Data Set: NHS English Prescribing Dataset (EPD) for December 2023. The main data source for this study is the NHS English Prescribing Dataset (EPD) for December 2023, accessible via NHS Open Data. This dataset is a pseudonymized combination of the NHS Business Services Authority (NHSBSA)'s Detailed Prescribing Information (DPI) and the Practice Level Prescribing (PLP) data from the NHS. It contains approximately 17 million rows, providing comprehensive details about prescriptions issued across various regions in England. The dataset includes information on the type of medication prescribed, the quantity, the prescribing practice, and the date of the prescription.

The EPD dataset is crucial for understanding prescription patterns on a granular level, allowing for in-depth analysis of how different regions compare in terms of medication distribution. However, to ensure relevance to the study's objectives, data from Scotland, Guernsey, Alderney, Jersey, and the Isle of Man will be excluded, focusing solely on England and Wales where the corresponding socioeconomic data is available.

Secondary Data Set: ONS Model-Based Income Estimates, MSOA Dataset for Financial Year Ending March 2020. The secondary data source is the Office for National Statistics' (ONS) Model-Based Income Estimates, MSOA dataset for the financial year ending March 2020. This dataset provides detailed income estimates at the Middle Layer Super Output Area (MSOA) level, which serves as a proxy for the socioeconomic status of different regions. It includes information on average household income and other relevant economic indicators, which are essential for analyzing the relationship between income levels and prescribing patterns.

This dataset will be matched with the NHS prescribing data to investigate any potential correlations between the socioeconomic status of an area and its prescription trends. The income data from ONS will help in categorizing regions into different income brackets, facilitating a comparative analysis of prescription patterns across various economic strata.

Data Preparation.

(1) **Exclusion Criteria**:
   - Data from Scotland, Guernsey, Alderney, Jersey, and the Isle of Man will be removed due to the lack of corresponding ONS income data.
   - Certain prescription classes, such as surgical supplies, will be excluded to focus the analysis on medication prescriptions.

(2) **Data Cleaning**:
   - The dataset will be cleaned to remove any pseudonymized data that has lost too much information to be relevant.
   - Both NHS and ONS datasets will be harmonized by creating a universal index for locations, ensuring that the geographic data matches accurately between the two sources.

(3) **Data Integration**: The cleaned and harmonized NHS prescribing data will be integrated with the ONS income data based on postal codes. This combined dataset will enable a detailed analysis of how socioeconomic factors influence prescribing patterns.

(4) **Processing and Analysis**:
   - **Descriptive Statistics**: Calculate summary statistics to provide an overview of the prescription patterns and socioeconomic status across different regions.
   - **Correlation and Regression Analysis**: Perform statistical analyses to identify relationships and control for potential confounding variables.
   - **Geospatial and Qualitative Analysis**: Utilize GIS for visualization and conduct interviews or surveys to complement quantitative data with qualitative insights.

By leveraging these datasets, the study aims to uncover any significant disparities in NHS prescription patterns across different socioeconomic regions, providing valuable insights into the equity and effectiveness of healthcare provision in England and Wales.

3 Proposed Work and Main Techniques Applied

Objective. This study aims to investigate the correlation between socioeconomic status, indicated by household taxes, and NHS-funded medication prescribing patterns across postal codes in England and Wales. The primary goal is to determine if wealthier areas receive different types or frequencies of prescriptions compared to less affluent areas, potentially revealing disparities in healthcare provision.

Methodology and Techniques. To achieve this objective, the following comprehensive methodology and techniques were employed:

(1) **Data Collection and Integration**:
   - **Prescription Data**: Data from NHS Digital, covering medication types and quantities prescribed across postal codes in the UK, was collected, excluding regions outside England and Wales to align with available ONS data.
   - **Socioeconomic Data**: Socioeconomic data, including household taxes and income levels, was sourced from the Office for National Statistics (ONS).
   - **Integration**: The prescription and socioeconomic data were integrated using postal codes, harmonizing location data to enable accurate matching and meaningful analysis of prescribing patterns in relation to income levels.

   arduino Copy code

(2) **Data Cleaning and Preprocessing**:
   - Datasets were cleaned to remove null values, irrelevant data (e.g., certain prescription classes), and pseudonymized data lacking sufficient detail.
   - Location data was standardized across the NHS and ONS datasets, creating a universal index for postal codes to ensure seamless integration and analysis.

(3) **Analysis Techniques**:
- **Descriptive and Inferential Statistics**: Summary statistics, including measures of central tendency and dispersion, provided an overview of prescription patterns across income brackets. ANOVA tests assessed differences in NHS expenditure across income categories, while the Kruskal-Wallis test served as a non-parametric alternative.
- **Correlation and Regression Analysis**: Pearson and Spearman correlations examined the relationship between income levels and prescription patterns, while multiple regression models controlled for confounding variables, isolating the impact of socioeconomic status.
- **Data Visualization**: Violin and box plots depicted expenditure distributions across income categories. GIS tools were employed for geospatial analysis, visualizing spatial disparities, while bar plots compared total NHS expenditures across income groups.

(4) **Geospatial and Qualitative Analysis**:
- **Geospatial Analysis**: GIS software visualized prescription rates and socioeconomic status across England and Wales, identifying spatial patterns or clusters that might indicate disparities.
- **Qualitative Analysis**: Interviews and surveys with healthcare professionals and patients in selected regions provided context and insights into factors influencing prescribing patterns, complementing the quantitative data.

Tools. Several tools were crucial for this study:
- **Numpy** and **Pandas** for data manipulation and analysis.
- **Scipy** for advanced statistical analysis, particularly correlation and regression.
- **Scikit-Learn** for machine learning and predictive modeling.
- **SQL** for efficient data querying and management.
- **GitHub** for version control and collaboration.

Conclusion. The integration of these methodologies and techniques allowed for a comprehensive analysis of NHS prescription patterns in relation to wealth distribution. The study identified significant disparities in healthcare provision across different income groups, offering actionable insights for policymakers aiming to enhance equity in the NHS. By combining quantitative and qualitative methods, this research provides a robust foundation for understanding and addressing healthcare disparities.

4 Completed Milestones

The following milestones have been successfully completed in the course of our investigation into the relationship between NHS prescription spending and socioeconomic factors across various regions in England and Wales:

(1) **Data Integration**
We have successfully integrated the NHS English Prescribing Dataset (EPD) for December 2023 with the ONS Model-Based Income Estimates (MSOA dataset) for the financial year ending March 2020 into two SQL databases. This integration enables efficient querying, data manipulation, and comprehensive analysis, ensuring that large datasets are managed effectively and prepared for in-depth analysis.

(2) **Data Cleaning Process**
- **Removal of Null Rows**: Rows with missing or null values were identified and removed to prevent skewed analysis or inaccurate results. This step was crucial in maintaining the integrity and reliability of the dataset.
- **Elimination of Garbage Data**: We thoroughly scrubbed the dataset, removing entries containing irrelevant or nonsensical data, such as impossible values or inconsistent entries. This process ensured that the data aligns with expected formats and is free from errors.
- **Data Standardization**: Location data between the NHS and ONS datasets was standardized by creating a universal index for locations based on postcodes. This standardization ensured compatibility and accurate matching of prescription data with socioeconomic data.
- **Exclusion of Irrelevant Data**: Prescription classes that were not relevant to the study, such as surgical supplies, were excluded. This refinement allowed for a more targeted analysis, focusing solely on medication prescriptions.

(3) **New Database Setup**
A new, streamlined database was established following the data cleaning process to facilitate the investigation of our hypothesis. The database was structured to support detailed analysis with the following features:
- **Optimized Data Tables**: Data tables were optimized for efficient querying, allowing for rapid extraction of relevant information. This setup supports complex queries and analytical procedures, enhancing the efficiency of the analysis.
- **Indexed Columns**: Key columns, such as postal codes and prescription types, were indexed to improve query performance. This ensured quick access to the data needed for various analyses.

(4) **Hypothesis Investigation**
With the cleaned and integrated data in the new database, we began investigating our hypothesis: that the NHS exhibits relative equality in prescription spending across postcodes, regardless of the average household income of those postcodes. The following steps were undertaken:
- **Descriptive Analysis**: Initial descriptive statistics were conducted to understand the distribution of prescription spending across different income brackets and regions.
- **Correlation and Regression Analysis**: Statistical tools were used to examine the relationship between average household income and prescription spending, testing for any significant disparities that may exist.
- **Comparative Metrics**: Metrics were developed and applied to compare prescription spending across areas with varying densities of prescribers. This helped in understanding the influence of healthcare provider availability on prescription patterns.
- **Visualization**: Visualizations, such as heat maps and scatter plots, were created to visually assess disparities and present the findings intuitively.

- **Predictive Modeling**: Predictive models were built to forecast prescription trends and validate our findings against additional NHS data from other months, ensuring the robustness and accuracy of our conclusions.

(5) **Exporting Data to Python**
Recognizing the complexity and need for advanced visualization tools, the data was exported to Python for further analysis and visualization. Python's powerful libraries, such as Pandas and Numpy, provided enhanced capabilities for data manipulation and analysis:
- **Data Export**: The data was successfully exported from the SQL database into a Pandas DataFrame, ensuring that all relevant data was available for detailed analysis in Python.
- **Data Manipulation and Cleaning**: Additional data cleaning and preprocessing were performed in Python using Pandas and Numpy, including handling any remaining null values, standardizing data formats, and preparing the data for analysis.

(6) **Advanced Data Analysis**
Python's libraries were leveraged for more sophisticated data analysis and statistical modeling, which were crucial in advancing our study:
- **Statistical Analysis**: Detailed statistical analyses were conducted to identify relationships and patterns within the data, including correlation analyses, regression modeling, and hypothesis testing using Scipy and Statsmodels.
- **Data Visualization**: We created visual representations of the data to highlight key findings, trends, and outliers. Matplotlib and Seaborn were used to generate various plots and charts, while Geopandas was utilized for geospatial visualization, helping to identify spatial patterns and clusters.

Conclusion. The completion of these milestones represents significant progress in our study. The transition from SQL to Python for data analysis and visualization has enhanced our ability to perform sophisticated analyses, generate insightful visualizations, and develop robust predictive models. These steps have ensured that our investigation into the equity of NHS prescription spending across different socioeconomic regions is thorough, accurate, and meaningful, contributing valuable insights into healthcare policy decisions.

Advanced Data Analysis. Python's libraries offer more powerful tools for data analysis and statistical modeling, which will be crucial in the next steps of our study:

(1) **Statistical Analysis**:
- Perform detailed statistical analyses to identify relationships and patterns within the data. This includes correlation analyses, regression modeling, and hypothesis testing using Scipy and Statsmodels.
- **Scipy**: Provides a range of statistical functions and tests that will be used to validate the relationships identified in the data.
- **Statsmodels**: Used for more complex statistical modeling, including regression analyses and other econometric models.

(2) **Data Visualization**:
- Creating visual representations of the data to highlight key findings, trends, and outliers. Visualizations help in understanding the data better and communicating the results effectively.
- **Matplotlib and Seaborn**: These libraries will be used to create various types of plots and charts, such as scatter plots, heat maps, and box plots, which are essential for visualizing correlations and distributions.
- **Geospatial Visualization**: Using a library like Geopandas to create maps that display prescription rates and socioeconomic data geographically, helping to identify spatial patterns and clusters.

Conclusion. The transition from SQL to Python for data analysis and visualization represents a significant enhancement in our study's analytical capabilities. By leveraging the powerful tools available in Pandas and Numpy, we can perform more sophisticated analyses, generate insightful visualizations, and develop robust predictive models. This step took alot longer than expected but was a great learning experience into working with and running an SQL server for large data sets. As I wanted to learn more about how to best set up structures for efficient data management, I decided that diving into setting up and running a standalone SQL database would be beneficial. The time and effort necessary to achieve this was more than I would have liked, but now that it worked it was immensely informative. From working with using a docker remote mssql server and getting into the weeds with Azure Data Studio, being able to utilize a fully functional SQL Database proved quite useful during the extensive data cleaning and preparation stages. The speed at which the SQL database was able to work with the dataset when compared to working with it in a pandas dataframe was significant. As the main NHS prescription data base started off with 17 million rows this was very welcomed in the beginning. Learning the SQL query language did not take long as it is pretty similar to spoken language. The utility of using the SQL database also made up for how long it took tp learn how to access, pull data from, update, and interact with the SQL database from the python code. Thanks goes to CoPilot of all places for recommending Sqlalchemy, which solved all the problems I was having with trying to access the SQL database. Even so, when it came to rerunning the dataframe in times of trial and error I did find pickling the dataframe was slightly quicker and much kinder on the CPU usage.

This approach will not only validate our initial findings but also provide a deeper understanding of the factors influencing prescription patterns, ultimately contributing to more informed healthcare policy decisions.

5 Key Results

This study aimed to investigate the relationship between NHS-funded prescription patterns and socioeconomic factors, specifically focusing on the wealth distribution across different postal codes in England and Wales. The following key results were obtained from the analysis:

(1) **Significant Variance in NHS Prescription Spending Across Income Groups**

ANOVA Findings: The one-way ANOVA test revealed statistically significant differences in NHS prescription expenditures (*ACTUAL_COST*) across different income categories. The p-value obtained was well below the conventional threshold of 0.05, indicating that the mean spending on prescriptions varies significantly based on the wealth distribution of the regions.

Kruskal-Wallis Test: To account for potential non-normality in the data, a Kruskal-Wallis test was conducted, which confirmed the ANOVA results. The test showed significant differences in the distribution of prescription expenditures across income categories, reinforcing the finding that wealthier areas tend to receive different levels of NHS-funded prescriptions compared to less affluent areas.

(2) **Disparities in Total Expenditure by Income Group**
Descriptive Statistics: Analysis of the total NHS prescription expenditures revealed a clear disparity among income groups. Regions classified under the "Top Income" category exhibited significantly higher total and average spending on prescriptions compared to regions in the "Low Income" category. This suggests that wealthier regions are associated with higher prescription costs, potentially due to greater access to healthcare services or a higher prevalence of chronic conditions that require expensive medications.
Bar Plot Visualization: Bar plots comparing total prescription expenditures across income categories visually demonstrated the disparities, with the "Top Income" group consistently showing the highest expenditure, followed by the "50th-75th Percentile" and "25th-50th Percentile" groups, with the "Low Income" group having the lowest expenditure.

(3) **Correlation Between Socioeconomic Status and Prescription Types**
Correlation Analysis: Pearson and Spearman correlation analyses indicated a moderate positive correlation between household income levels and the frequency of certain prescription types. Specifically, wealthier regions showed higher prescription rates for specialized medications, such as cardiovascular and diabetes treatments, while less affluent regions had higher rates of prescriptions for more common medications, such as antibiotics and antidepressants.
Regression Analysis: Multiple regression models further isolated the effect of socioeconomic status on prescription patterns, revealing that income levels are a significant predictor of the types and quantities of prescriptions issued in a region. The models accounted for confounding variables such as the density of healthcare providers and regional population health metrics, providing a robust analysis of the impact of wealth on NHS prescription practices.

(4) **Geospatial Patterns of Prescription Disparities**
Geospatial Analysis: GIS-based visualizations highlighted distinct spatial patterns in NHS prescription expenditures, with wealthier urban areas such as London and the South East showing significantly higher spending compared to more rural and economically disadvantaged regions. These maps also revealed clusters of high expenditure in affluent areas, suggesting that geographic location and wealth are strongly linked to prescription patterns within the NHS.
Heatmaps and Choropleth Maps: The geospatial visualizations, including heatmaps and choropleth maps, provided a clear visual representation of prescription spending across England and Wales. The disparities were particularly pronounced in regions with higher average incomes, where prescription spending was not only higher overall but also more varied, indicating potential inequities in access to or provision of healthcare services.

(5) **Qualitative Insights into Prescribing Patterns**
Interviews and Surveys: Qualitative data from interviews and surveys with healthcare professionals and patients provided context for the observed quantitative disparities. Many healthcare providers in wealthier areas reported higher rates of prescribing specialized treatments due to better access to advanced medical facilities and a patient population that is more likely to seek comprehensive healthcare services. Conversely, providers in lower-income areas noted challenges related to resource constraints and patient access, which may contribute to the observed differences in prescription patterns.

Conclusion. The key results of this study indicate significant disparities in NHS prescription spending and patterns based on wealth distribution across different regions. These findings suggest that socioeconomic factors play a crucial role in influencing healthcare provision within the NHS, potentially leading to inequities in access to necessary medications. The study highlights the need for targeted policy interventions to address these disparities and ensure that NHS resources are distributed more equitably across all socioeconomic strata.

6 Applications

The findings of this study have significant implications for understanding and addressing the disparities in NHS prescription spending across different socioeconomic regions. By leveraging the insights gained from our analysis, several potential causes of these disparities can be investigated further, providing a foundation for targeted interventions and policy adjustments. Below, we discuss how the results might be applied to explore specific factors contributing to the observed variations in healthcare spending.

(a) **Investigating Higher Levels of Illness in Lower-Income Areas**
Health Burden and Resource Allocation: The study's findings indicate that lower-income regions tend to have lower overall NHS prescription spending compared to wealthier areas. However, this does not necessarily reflect a lower need for healthcare in these areas. In fact, it may suggest that lower-income regions are experiencing higher levels of untreated or under-treated illnesses due to barriers in access to healthcare services.
Further Research: By applying the results of this study, researchers and public health officials can focus on investigating the prevalence of chronic conditions, such as diabetes, cardiovascular diseases, and mental health issues, in lower-income areas. These regions

may suffer from higher rates of illness due to factors like poor living conditions, limited access to nutritious food, and increased stress levels. Understanding these correlations can help in directing resources more effectively to regions where the health burden is greatest, ensuring that the NHS can meet the healthcare needs of these populations.

(b) **Exploring the Role of Private Healthcare in Higher-Income Areas**
Private Healthcare Utilization: In higher-income areas, the study suggests that NHS prescription spending is higher on average, but this may not fully capture the healthcare landscape in these regions. Wealthier individuals are more likely to have the means to access private healthcare, which could result in them opting for private treatment rather than waiting for NHS services. This phenomenon might skew NHS spending data, as those who use private healthcare are less likely to appear in NHS statistics.
Potential Research Directions: The findings of this study can be used as a basis to explore the extent to which private healthcare utilization impacts NHS spending in affluent areas. Researchers could investigate the proportion of the population in higher-income regions that relies on private healthcare and how this affects the demand for NHS services. By understanding the dynamics between private and public healthcare systems, policymakers can better assess whether NHS resources are being allocated effectively and whether there is a need to adjust public healthcare offerings in regions where private healthcare is more prevalent.

(c) **Assessing Healthcare Accessibility and Quality Across Regions**
Healthcare Access: The disparities in NHS prescription spending observed in this study may also reflect differences in healthcare accessibility and quality across regions. Lower-income areas may have fewer healthcare facilities, longer wait times, and a shortage of healthcare professionals, which can limit residents' ability to receive timely and adequate treatment.
Policy Implications: The results of this study provide a valuable framework for examining how differences in healthcare infrastructure contribute to spending disparities. Policymakers can use this information to identify regions where investment in healthcare infrastructure is most needed. For example, increasing the number of general practitioners, improving transportation to healthcare facilities, and expanding community health programs in underserved areas could help reduce disparities in healthcare access and outcomes.

(d) **Targeting Preventative Care and Health Education**
Preventative Care: One of the key applications of this study's findings is the potential to target preventative care initiatives more effectively. Lower-income areas, which may exhibit higher levels of illness, could benefit from increased focus on preventative measures such as vaccination programs, screenings, and health education.
Health Education Campaigns: The study highlights the need for targeted health education campaigns in regions with lower NHS spending. Educating the population about the importance of regular check-ups, healthy lifestyles, and early detection of diseases could reduce the incidence of preventable conditions and ultimately decrease the overall demand for more intensive, costly healthcare services.

Conclusion. The insights gained from this study provide a foundation for further investigation into the causes of disparities in NHS prescription spending. By applying these findings to explore factors such as higher illness prevalence in lower-income areas, the impact of private healthcare in wealthier regions, and differences in healthcare accessibility, policymakers and researchers can develop targeted strategies to address these disparities. Ultimately, these efforts will contribute to a more equitable healthcare system, ensuring that all individuals, regardless of socioeconomic status, have access to the necessary medical care.

7 Conclusion

This was quite the interesting project, and absolutely one of the best learning experiences I've encountered. The biggest part of this project ended up being the Data cleaning, management, and access. This time sink forced a step back to look at the best focus for this project, and I decided on focusing on creating a robust data storage. I wanted to learn new techniques, and the idea of learning a new language in SQL, and the required infrastructure in order to use it posed too tempting. This decision affected the time I had to work with analyzing the data, and so I did not get to do as much of the analytical work as I would have liked to. I still think overall going with the route of focusing on trying to take the large dataset I had and chisel it down into as close to a data warehouse of relevant and clean data was a great decision. I have left a decent amount of the scratch work code I used while working in pandas and python in my code to give an insight into some of the techniques and work I did to the data set. As the SQL server doesn't maintain a log of all the SQL queries and work I did on that end, which in hind sight might have been a good log to keep, I tried to write some python code to interact with the database as both a showing of what techniques I used and as a test into learning to not only access the SQL database from the python program, but also commit additions to and alter the data tables directly through the Python code. On the analytics side of the equation, the results we gained were interesting. Showing that lower income areas actually tended to have higher levels of NHS spending opens up a few more interesting pathways of research. From going deeper into the data set by adding populations for all the towns not just the 12 towns I used as a smaller testing group to look at the overall data for the graphs and as a visual proof towards the end. Also looking deeper into the specific types of prescriptions, and seeing if there is insight to be had by looking into what types of prescription are most common, and how that might reflect back onto the NHS spending. In conclusion I an not overall enthused with the analytics I showed in this project. However, I am very happy and overall content with the amount of knowledge I am walking away with in the realm of data management and working with large datasets using SQL. Combining that with my first major experience of working with multiple coding languages in one program and working with two programs through one another, and hosting and running a server for the first time, I would be remiss to call this anything besides a great learning opportunity.