**Prescription Cost Analysis**

**Introduction:**

For the Capstone Project the analyst has chosen data from Prescription Cost Analysis (PCA) for the month of April 2025 and for the month-to-month comparison the author has also chosen the data for the month of March 2025. This data is collected by the National Health Service Business Service Authority (NHSBSA) and it contains the monthly administrative data. The PCA data holds evidence on all prescription’s items dispensed in the community in England monthly and submitted to the NHSBSA for repayment. PCA delivers information of the costs and volumes of all prescriptions dispensed at the community in England.

**Background:**

This data reflects the methodology used in the PCA annual national statistics publication. PCA demonstrates the total volume and cost for drugs, dressings, appliances, and medical devices that have been dispensed in the community in England. These statistics are created to assemble the total number of prescription items and spend for any individual product, at any level of the British National Formulary (BNF) hierarchy, e.g. for Paracetamol 500mg tablets, or collections of products such as antibacterial drugs, antidepressant drugs, or dressings.

Prescription data is collected by the NHS Business Services Authority (NHSBSA) for the operational drive of recompensing dispensing contractors for the costs of supplying drugs and devices, and providing essential and advanced services, to NHS patients in different Integrated Care Board (ICB) or region. The data that forms the basis of these statistics is collected as a by-product of this process. This data is used to monitor medicine uptake, to allow public scrutiny of prescribing habits, to inform local and national policy, and in academic research.

Data is collected from the submission of prescriptions by dispensing contractors to the NHSBSA. These prescriptions are issued by GPs and other authorised prescribers such as nurses, dentists, and allied health professionals. Prescriptions that are issued by hospitals can also be dispensed in the community and submitted for reimbursement. Prescriptions that are issued in hospitals and fulfilled by the hospital pharmacy or dispensary are not included in this data.

The dataset contains prescription records, with fields such as:

* Date (Year, Month)
* Region (Region Name, ICB-Name)
* Practice type (Dispenser, account, Type)
* Drug details (chemical/substance name, strength, formulation)
* Supplier
* BNF (classification: therapeutic category hierarchy)

**Aim and Objectives:**

The aim of the project is the better understanding of the PCA data and to understand the data subsequent objectives will be followed:

* Identifying the most frequently prescribed medications in each Integrated Care Board (ICB) or region.
* How does prescription frequency change from March 2025 to April 2025?
* How often are generics used instead of brand names? Are some Integrated Care Board (ICBs) better at switching to generics?
* Are some medications more prescribed in certain regions despite similar demographics?

**Methods:**

Data is gathered from the open data source and the weblink is: [https://opendata.nhsbsa.net/dataset/prescription-cost-analysis-pca-monthly-datacription Cost Analysis (PCA) Monthly Administrative Data - Datasets - Open Data Portal](https://opendata.nhsbsa.net/dataset/prescription-cost-analysis-pca-monthly-datacription%20Cost%20Analysis%20(PCA)%20Monthly%20Administrative%20Data%20-%20Datasets%20-%20Open%20Data%20Portal). The analyst has also gathered the data for the month of March. For the Data analysis mostly software tool Jupyter Notebook was mostly used. The data analysis ‘ipynb’ files will also be uploaded to have a clear understanding of the readers. The project will also be uploaded in GitHub for wider audience accessibility. The GitHub file will also include the used codes in Jupyter Notebook.

Useful tools for the analysis:

* Jupyter Notebooks (Pandas, Seaborn, Plotly, Scikit-learn)
* Excel (for the data only)

Following are the project concepts and data analysis objectives for the better understanding for the data:

**1. Top Prescribed Drugs by Region**

* Goal: Identifying the most frequently prescribed medications in each Integrated Care Board (ICB) or region.
* Insights: Top 5 prescribed medications in different parts of England
* Tools: Jupyter Notebook
* Techniques: Grouping, Aggregation, Visualization (bar plots, maps)

**A screenshot of a computer program

AI-generated content may be incorrect.**

**A graph with multiple colored bars

AI-generated content may be incorrect.**

**Table Interpretation of the bar chart:**

| **REGION\_NAME** | **BNF\_CHEMICAL\_SUBSTANCE** | **ITEMS** |
| --- | --- | --- |
| EAST OF ENGLAND | Atorvastatin | 715563 |
| EAST OF ENGLAND | Amlodipine | 388108 |
| EAST OF ENGLAND | Levothyroxine sodium | 365127 |
| EAST OF ENGLAND | Lansoprazole | 334619 |
| EAST OF ENGLAND | Omeprazole | 333471 |

**Interpretation:** According to the bar chart and the table above Atorvastatin, Amlodipine, Levothyroxine, Lansoprazole and Omeprazole are the five most prescribed medicine and are mostly prescribed in East of England region followed by London, Midlands, North East and Yorkshire, North West, South East, South West.

**2. Trend Analysis Over Time**

* Goal: How does prescription frequency change from March 2025 to April 2025?
* Insights: Exact number of items prescribed in each month, absolute and percentage change, visualisation of trends.
* Tool: Jupyter Note Book
* Techniques: Line charts, time series forecasting

**A screenshot of a computer code

AI-generated content may be incorrect.**

**A graph showing a number of green and black bars

AI-generated content may be incorrect.**

**A graph with a bar graph

AI-generated content may be incorrect.**

**Interpretation:** The green Bar Chart represents that the prescription of medications such as Atorvastatin, Bisoproplol, Lansoprazole and amlodipine has increased in the month of April compared to March. On the contrary the red Bar Chart represents that the prescription of medicines such as Amoxicillin, Clarithromycin, Salbutamol and Chloramphenicol has gone down in April compared to March.

**3. Generic vs Branded Drug Usage**

* Goal: How often are generics used instead of brand names? Are some Integrated Care Board (ICBs) better at switching to generics?
* Insights: which ICBs prescribe generics >90% of the time, which ICBs lag behind (<60%)?, opportunities to improve prescribing efficiency and reduce costs.
* Techniques: Data cleaning (matching generic/branded codes), ratios, charts

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A graph of green and black color

AI-generated content may be incorrect.**

**A graph of a number of people

AI-generated content may be incorrect.**

**Interpretation:** According to the above green bar chart, Integrated Care Board (ICB) or regions such as Nottinghamshire, Manchester, South Essex uses more generic brands compared to Surrey, Firmley, East London and other ICBs in the Red bar chart. This interpretation represents that ICBs in red bar chart can improve their prescribing efficiency and reduce costs.

**4. Regional Inequalities in Prescribing**

* Goal: Are some medications more prescribed in certain regions despite similar demographics?
* Insights: Regional prescriber habits, local policy influence, formulary preference, unmet demand or misalignment with population needs
* Techniques: Ratios, Charts

**A screenshot of a computer program

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

A graph with colorful lines

AI-generated content may be incorrect.

**Interpretation:** The bar chart Above represents that South West East of England, South East England, and London, North West, Midlands, North East and Yorkshire has almost similar ratio of prescription exhibited and has a Relative Prescription Ratio (RPR) greater than 1.5 which is above the national average, representing people on those regions have more physical health issues compared to other parts of England.