Length: 2,500 words (excluding tables, bibliographies, appendices, diagrams, pictures, graphs, and

associated captions).

Markers will look for the following features in the assignment:

1. sound understanding and demonstration of data preparation, parsing, and cleansing
2. reasonable formal documentation and technical demonstration how practices can be selected in a given geographic locality
3. feasible formal documentation and technical demonstration how practices can be selected in a given geographic locality
4. sound documentation and technical demonstration of pre-defined prescriptions and associated Costs
5. reasonable description and technical demonstration of most/least frequent and expensive medications in a selected geographic area; that includes visualisation of historical trends, variation between GP-practices, and outlier detection
6. suitable description and technical demonstration how the selected prescribing profile in a given geographic area compares to London as a reference
7. detailed discussion of the obtained results – presented using a combination of descriptive statistics and visualizations

The NHS has been challenged to make “efficiency savings” and you have been commissioned by an NHS executive to review, document and assess GP prescribing costs. Using data from the GP Practice Prescribing dataset (April 2018) address the following queries using a combination of narrative, tables, figures, and descriptive statistics:

**Questions**

1. Identify all GP practices located in London. For those practices, describe:

* the total number of patients registered
* the total number of prescriptions
* the total actual cost of these prescriptions (using the ACT COST column)
* the top 10 most frequent drugs prescribed
* the bottom 10 less frequent drugs prescribed

1. Repeat the previous instructions, this time for the city of Cambridge. Discuss and compare your findings with the answers for London in question 1 above using descriptive statistics.
2. Describe total number of prescriptions and their total actual cost (using the ACT COST column) across all practices for drugs related to:

* cardiovascular disease (British National Formulary chapter 2)
* antidepressants (British National Formulary chapter 4.3)

1. Describe the total spending and the relative costs per patient across all practices for the month of April 2018:

* visualize the monthly total spending per registered patients using a scatterplot and provide a trend line
* generate a histogram for relative spending for all practices and fit a Gaussian (normal) curve
* Using SQL, produce a table that provides the number of GP practices per city, ordered in descending order.

**Questions**

1. What was the population and the total number of deaths (from all causes, all ages) in 2010 for:

* Iceland
* Italy
* New Zealand

1. What was the distribution of deaths (all causes, all years) by age group in Italy?

* Visualise the results using a histogram.

1. What were the top five causes of death (top five ICD-10 terms) in Italy across all years for the Neoplasm ICD10-category (C00-D48)?

* Generate a table with the cause of death, the number of deaths, and the
* proportion of overall deaths.
* Generate a pie chart to visualize the proportion of deaths.

1. Are there differences by age group for deaths from Neoplasms (C00-D48) in Australia for 2010?

* Identify the top five age groups in Australia dying with a Neoplasms cause of death.

1. Compare and contrast the frequency of deaths by Neoplasms in Italy and Australia in 2010.

* Combine information on the population and deaths and describe your logic.
* Use descriptive statistics and plots.

**Introduction**

GPs in London

Get df of london GPs 'London' is not very precise. 'London postcodes' are relatively easy but would only define inner London. It would also be possible to use all 'Greater London' postcodes, but these are defined for administrative purposes of the post office and do not correlate with any administrative definition of London. 'london.gov.uk' provide a map of the area sovered by the Greater London authority, but this does not correlate with postcodes (as noted above) and would be complex to define as search criteria. Given the context, an NHS definition seems appropriate and it is possible to get a list of CCGs that are covered by the London area team of NHS England, and then search for GPs in areas covered by those CCGs.

https://files.digital.nhs.uk/5F/21E948/gp-reg-pat-prac-map.csv

https://www.london.gov.uk/sites/default/files/gla\_postcode\_map\_a3\_map1.pdf

The same approach is not possible for Cambridge as the relevant CCG is NHS Cambridgeshire and Peterborough - covering far more than Cambridge. Again there is a difference between the administrative boundary of the Cambridge and CA post codes. The most practical search criteria is the use of postcodes CB1-5, though the numbers registered with GPs in these postcodes (191931) does not correspond very well with the population of Cambridge from the ONS (125,000). The discrepancy is probably partly explained by a lack of precise alignment of the postcodes with 'Cambridge', but also patients registered with different GPs do not correspond with clearly delineated areas. Thus, administrative boundary of Cambridge != postcodes of GPs != addresses of patients. If for some reason it was necessary to look at prescriptions for people living in Cambridge, it would be necessary to get the actual addresses of individual patients, then define which postcodes corresponded to an address in Cambridge (using the full postcode). It seems unlikely that this would be either necessary or practical.

While there is no single definition of city regions, 'in essence they refer to metropolitan centres and their “functional areas”' <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/populationdynamicsofukcityregionssincemid2011/2016-10-11>

isna for London scripts

SHA 0

PCT 0

PRACTICE 0

BNF CODE 0

BNF NAME 0

ITEMS 0

NIC 0

ACT COST 0

QUANTITY 0

PERIOD 0

0

1299 GPs in London

So there are 24 GP practices we have practice data for but no spending info. All the practices we have spending data on we have practice data for.

9851208 patients

1401350 prescriptions

76586738.71 total cost

Top 10 prescribed items - though each prescription could be for a variable number of items

GlucoRX FinePoint Needles Pen Inj Screw 2374

Omeprazole\_Cap E/C 20mg 1294

Amlodipine\_Tab 10mg 1293

Amlodipine\_Tab 5mg 1293

Atorvastatin\_Tab 40mg 1292

Aspirin Disper\_Tab 75mg 1292

Cetirizine HCl\_Tab 10mg 1292

Lansoprazole\_Cap 30mg (E/C Gran) 1291

Metformin HCl\_Tab 500mg 1291

Ramipril\_Cap 5mg 1290

Top 10 prescriptions by quantity prescribed

Ensure Plus\_Milkshake Style Liq (9 Flav) 28240772

Metformin HCl\_Tab 500mg 21066834

Fortisip Bottle\_Liq (8 Flav) 18282528

Lactulose\_Soln 3.1g-3.7g/5ml 16291397

Paracet\_Tab 500mg 13089271

Ensure Compact\_Liq (4 Flav) 12145894

Dermol 500\_Lot 10758036

Fortisip Compact\_Liq (8 Flav) 9442786

Omeprazole\_Cap E/C 20mg 9332189

Fresubin 2kcal\_Drink (6 Flav) 7684352

Bottom 10 prescriptions by quantity prescribed

Metronidazole\_I/V Inf 5mg/ml 100ml Btl 1

Covonia\_Throat Spy 30ml 1

Cutimed Gel 25g Wound Dress H/Gel Ster 1

Hollister\_Karaya Pdr 1

Midazolam HCl\_Inj 1mg/ml 50ml Vl 1

Skinnies Silk Top 6-24 Months Silk Garme 1

Skinnies Silk Leggings 6-24 Months Silk 1

Dansac\_Combi Ostomy Belt 50-63mm 1

Skinnies Silk Knee Length Socks Adult Lg 1

Picato\_Gel 500mcg/g 0

**Cambridge**

17 GPs

191931 patients

1227048.96 total cost

Top 10

Fortisip Compact\_Liq (8 Flav) 544500

Nutrison Pack\_Energy 436716

Fortisip Bottle\_Liq (8 Flav) 284136

Dermol 500\_Lot 259500

Paracet\_Tab 500mg 250717

Lactulose\_Soln 3.1g-3.7g/5ml 201730

Omeprazole\_Cap E/C 20mg 199996

Metformin HCl\_Tab 500mg 185597

Nutrison Pack\_Conc Liq 156000

Epimax Crm 500g 149700

Bottom 10

Gonapeptyl Depot\_Inj 3.75mg Pfs + Dil 1

Coloplast SpeediCath Male Size 16-18 (30 1

Ciclesonide\_Inh 160mcg (60 D) CFF 1

Coloplast\_Brava Belt (For Sensura Mio) S 1

Tridestra\_Tab 1

Chlorphenamine Mal\_Inj 10mg/ml 1ml Amp 1

Xeplion\_Inj 50mg/0.5ml Pfs 1

HBVAXPRO\_Vac 10mcg/ml 0.5ml Pfs 1

Aequasyal Prote Mth Spy 40ml (App) 1

Ingenol Mebutate\_Gel 150mcg/g 0