PrAnaViz

Kishore Jagadeesan

2021-03-31

Contents

List of Figures

1 Introduction

The aim of `PrAna` is to aggregate and normalize prescription data to calculate **total prescribed quantity** of different APIs, using open source statistical software R Core Team [2018]. The name is an acronym for ***Pr****escription* ***An****alysis*

Apart, from the calculation of the total prescribed quantity of an API or a group of APIs, specified to a postcode or region, We have also developed, *an open interactive web-based tool*, `PrAnaViz` with the processed dataset for the period `2015` to `2018`.

`PrAnaViz` facilitates users **to visualise, explore and report** different spatiotemporal and long-term prescription trends for wider use.

2 Installation

PrAna can be installed as any other R package, however, since it is dependent on some other software tools some extra steps are required for the installation.

You can install the released version of PrAna from CRAN with:

```
install.packages("PrAna")
```

You can install the development version of PrAna from GitHub with:

```
# install.packages("devtools")
library(devtools)
install_github("jkkishore85/PrAna")
```

Note that you will need to have the latest R version installed in order to use this repository.

2.1 Dependencies

Software	Remarks
RStudio	Suggested IDE platform to perform the R operations
MySQL	Optional to host the generated dataset

3 Workflow

The workflow in this tutorial consists of the following steps, as in the Figure: 1:

- **Data Preparation**: Download monthly NHS prescription datasets and Dictionary of medicines and devices release files (dm+d).
- **Data Conversion**: Aggregation and conversion of the locally stored datasets into practice wise dataset achieved using the functions in `PrAna`.
- **Visualise and Analyse the data**: Visualise and analyse the processed dataset using the in-built ShinyApp `PrAnaViz`.
- **Database service**: Linking of the processed dataset to the `PrAnaViz` can be achieved by uploading the processed dataset to a local or a remote database service, for example, *MySQL.*
- **Download images and processed data**: Users can download processed data as ***.csv*** file and publication ready image ***.eps*** and ***.pdf*** files.

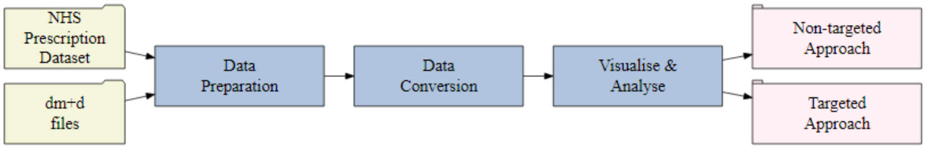


Figure 1: PrAna Workflow.

3.1 Data Preparation

Data sources

- Prescribing data and Practice information are from the monthly files published by the NHS Business Service Authority, used under the terms of the Open Government Licence.
- BNF codes and names are also from the [NHS Business Service Authority's Information Portal] NHSBSA, used under the terms of the Open Government Licence.
- dm+d weekly release data is also from the NHS Business Service Authority's Information Portal, used under the terms of the Open Government Licence.

To download NHS prescription datasets, users need to guest login and solve a captcha in NHSBSA information portal, to download the dm+d release files and *dm+d XML Transformation Tool*, users need to register an account.

Store the monthly NHS prescription data in year wise folder. For an example, store all the 2015 monthly datasets in a folder named '2015'. Extract the *dm+d release files* using *dm+d XML Transformation Tool*. The documentation to extract the release file using the tool is well explained in the readme file.

3.2 Data Conversion

Different functions were used in the aggregation and conversion workflow behind the scenes that performs the heavy lifting of a workflow step and finally return the results.

An overview of all functions involved in the data conversion is shown in the table below.

Function	*Remarks*
`csv2dat()`	Combine and convert all the comma separated value (.CSV) files in the defined file path and export it into a single R object
`importdmd()`	Import dm+d files and link it to the BNF code using the SNOMED mapping file
`practice_wise()`	Import NHS dataset files and generate processed individual GP practice prescription dataset
`runShiny()`	Run in-built Shiny Apps in the package

3.2.1 Combine NHS monthly dataset

`csv2dat()` function supports to combine different monthly NHS prescription dataset files into a single `data.frame` to process further. For an example:

```
## Load library
library(PrAna)

## Set the folder to store the combined file
setwd("C:/Datasets/2018")

## Combine and convert mulitple files in the defined folder to a dataframe
data201812 <- csv2dat("C:/Datasets/Prescription Datasets/2018/PDPI")
```

3.2.2 Import `dm+d` files

`importdmd()` function helps to import different extracted [`dm+d files`] [#dataextraction] and return multiple data objects including a `data.frame` which map each BNF code to its corresponding API(s), strength and medicinal form. Recommended to read the documentation of `importdmd()` function to know more regarding the different data objects it generate.

```
## Read the extracted dm+d files
dmdfile  <- importdmd("C:/dmdDataLoader/excel/")
```

3.2.3 Generate GP practice wise file

The final step in the data conversion is to generate prescription dataset mapped with the individual API, prescription quantity, medicinal form, and strength for the defined GP practice(s), `practice_wise()` is used to carry out this conversion.

And before execute the `practice_wise()` function it is strongly suggested to setup the destination folder as the working directory using `setwd()` function.

The `practice_wise()` function require following six parameters, as mentioned in the example below,

- Combined NHS prescription dataset, genereated using `csv2dat()` function
- A `character` vector containing GP Practices
- A `data.frame` containing BNF Code mapped to individial APIs, strength, medicinal form
- Unit of measurement with multiplication factors file
- Different medicinal forms with its corresponding codes file
- Different APIs with its corresponding codes file

```
## create a data.frame object for the BNF code mapping file
api_map <- dmdfile$api_map

## Unit of measurements
uom <- dmdfile$uomwdesc %>%
        dplyr::rename(UOM = CD)

## Medicinal form
dform <- dmdfile$dform

## ingredients list
ing <- dmdfile$ing %>%
  dplyr::rename(API_CODE = 1)

## Define the output location
setwd("C:/Datasets/Prescription Datasets/practice_wise/2018")

## Execute practice_wise() function
practice_wise(data201812, GP_practices, api_map, uom, dform, ing)
```

3.3 Visualise and Analyse the data

GP practice level files generated using the `practice_wise()` function are used to calculate the total prescribed quantity of different APIs at the particular GP practices and at a particular postcode using the in-built ShinyApp `PrAnaViz`.

`PrAnaViz` helps to visualise the total prescribed quantity of different APIs and to explore spatiotemporal trends of different APIs.

For a very quick start to `PrAnaViz`:

```
library(PrAna)
runShiny("PrAnaViz")
```

The `runShiny("PrAnaViz")` function will pop-up the `PrAnaViz` tool which will allow you to explore different spatiotemporal and long-term prescription trends with the sample dataset.

However, for a better guide to get started it is recommended to see the Section 4.

3.4 Database service

To avoid the large data loading and processing issue, authors strongly suggest to upload the processed dataset to a local or a remote database service, for example, *MySQL*, and link it to the `PrAnaViz`. More information on the linking databases to `PrAnaViz` is explained in the Section 4.2.

3.5 Download images and processed data

Users can download processed data as ***.csv*** file and publication ready image ***.eps*** and ***.pdf*** files, for their further usage and applications.

4 PrAnaViz

4.1 Introduction

The aim of this tool is to visualize and analyze the prescription data to calculate total prescribed quantity of different APIs and to explore spatio-temporal trends of different APIs for wider use.

We have created, `PrAnaViz`, in a familiar **dashboard layout** with two tabs: (1) **Targeted API**, (2) **Non-targeted API**, to visualize total quantity of different APIs at CCG region with resolution to individual postcode.

4.2 Database connection string

- The demo version of the in-built `PrAnaViz` enables users to input their database credentials to link the processed prescription dataset.
- This can added by clicking the database icon in the right side of the navigation bar, **1** as in Figure: 2.
- A drop down menu as in Figure: 3 will be generated to input the database connection strings such as, username, password and hostname

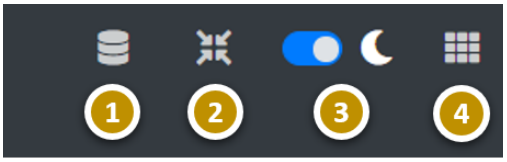


Figure 2: Right Navigation Bar

Option	*Remarks*
1	Database connection
2	Full Screen
3	Dark and Light mode
4	Customize the app colors

4.3 Targeted Approach

- In this tab, user can **input a list of APIs**, and find out the **total prescription quantity** of each API in the selected year, at the selected CCG region, as in the Figure: 4.
- This tab also visualize the total quantity of **APIs** prescribed by month, GP, Chemical form, and Medicinal form for a CCG region in a selected year, Figure: 5.

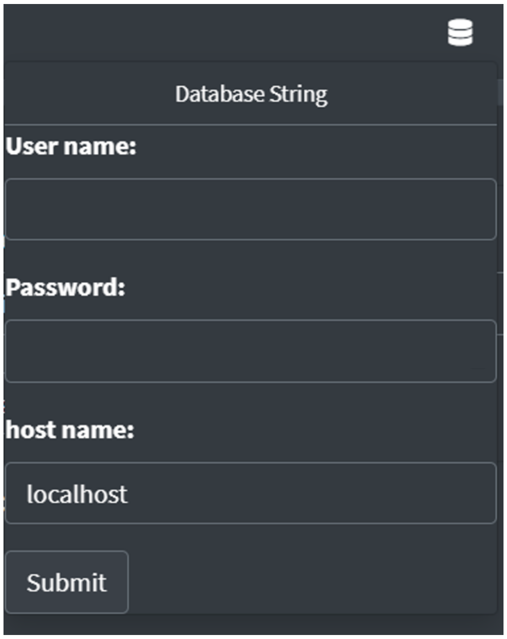
Option	*Remarks*
1	Select Targeted or Non-Targeted Approach

Option	*Remarks*
2	Upload targets i.e., APIs in `.csv` format
3	Select Year
4	Select Region
5	Select Setting Type
6	Select month-wise or year-wise plots
7	Action button to generate the plot
8	Plot generated based on user selection
9	Download buttons to download the generated plot as `.pdf` or `.eps` and data as `.csv` format
10	Maximize and minimize the tab
11	More insights on a selected API, visualize total prescribed quantity of the API per month, GP practice, postcode and medicinal form (click on a particular API on the barplot to generate these plots)
12	Download buttons to download the generated plot as `.pdf` or `.eps` and data as `.csv` format
13	Maximize and minimize the tab

4.4 Non-targeted Approach

- Total prescription quantity of an individual API at **different postcode per month** at a CCG region, can be rendered in this tab, as in the Figure: 6.
- The calculated total prescription quantity of an individual API at postcode level helps to find **the hotspots**.
- User can download data as ***.csv*** file and publication ready image ***.eps*** and ***.pdf*** files.

Option	*Remarks*
1	Select Targeted or Non-Targeted Approach
2	Select *API* in the `drop-down` menu
3	Select Region
4	Select Year
5	Select a month
6	Button to generate the plot
7	Plot generated based on user selection, click on a particular postcode to generate further plots
8	click on a particular postoode on the plot to generate these plots)
9	Visualize the total prescribed quantity of the API per month on the selected postcode
10	Download buttons to download the generated plot as `.pdf` or `.eps` and data as `.csv` format



Database String

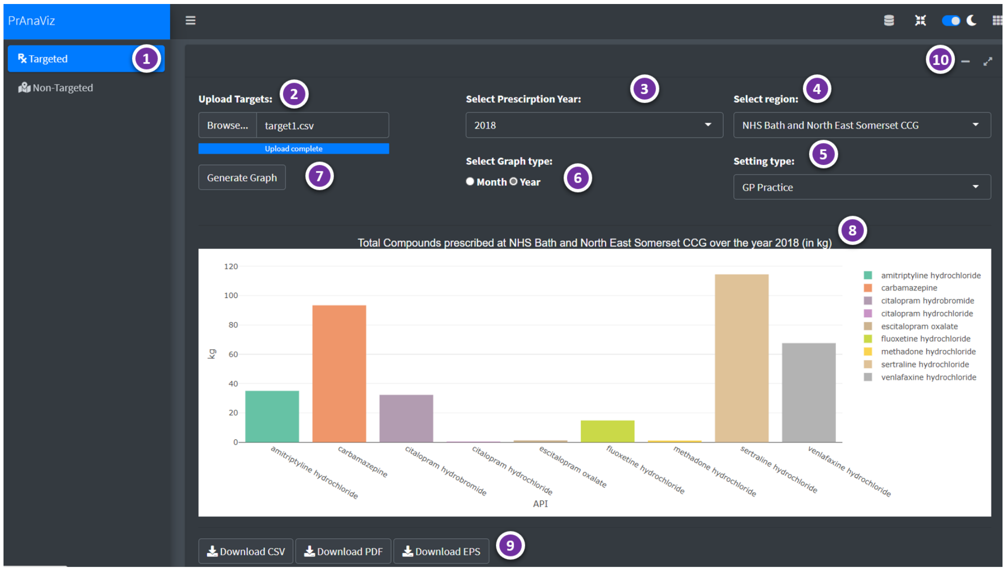
User name:

Password:

host name:

localhost

Figure 3: Database Connection String



Upload Targets: target1.csv

Upload complete

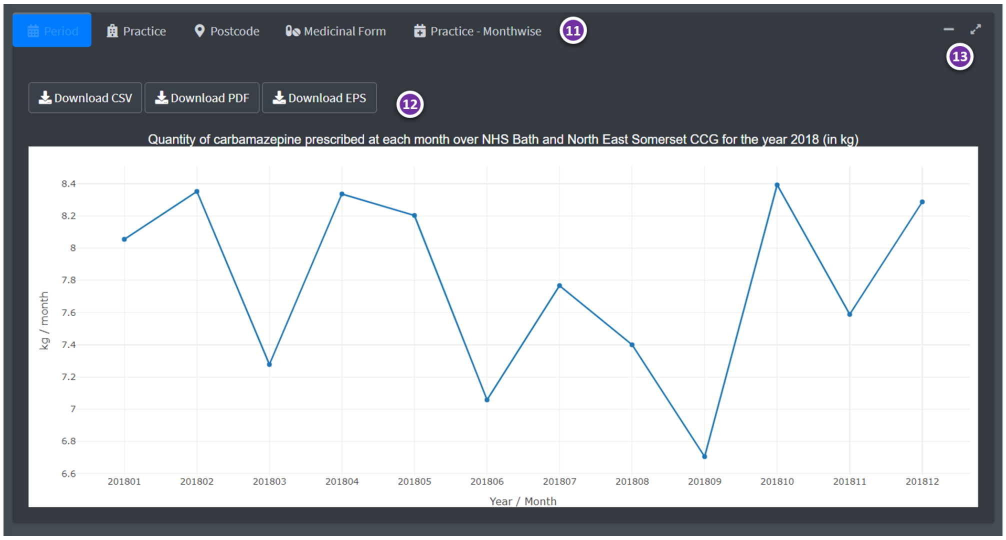
Select Prescirption Year: 2018

Select region: NHS Bath and North East Somerset CCG

Select Graph type: ● Month ○ Year

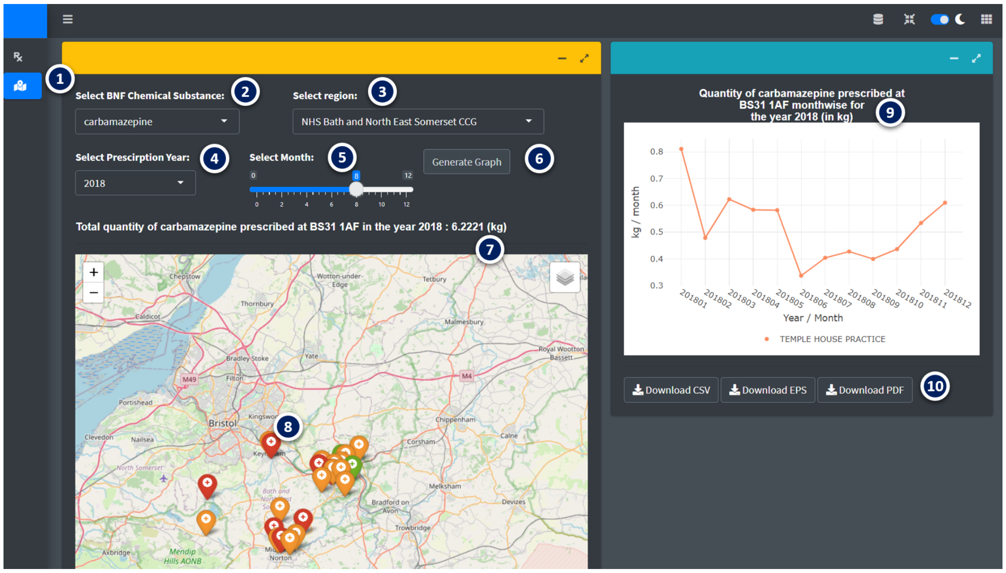
Setting type: GP Practice

Figure 4: PrAna Targeted Approach.



Quantity of carbamazepine prescribed at each month over NHS Bath and North East Somerset CCG for the year 2018 (in kg)

Figure 5: PrAna Targeted Approach.



Select BNF Chemical Substance: carbamazepine

Select region: NHS Bath and North East Somerset CCG

Select Prescirption Year: 2018

Select Month: 8

Total quantity of carbamazepine prescribed at BS31 1AF in the year 2018 : 6.2221 (kg)

Quantity of carbamazepine prescribed at BS31 1AF monthwise for the year 2018 (in kg)

Figure 6: PrAna Non-targeted Approach.

5 Acknowledgements

This package was built as a part of the **Wastewater Fingerprinting for Public Health Assessment (ENTRUST)** project funded by Wessex Water and EPSRC IAA (grant no. EP/R51164X/1).

5.1 Disclaimer

We accept no liability for any errors in the data or its publication here: use this data at your own risk. You should not use this data to make individual prescribing decisions.

References

R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2018. URL https://www.R-project.org/.