Project

AUTHOR
Aymeric Stamm

The Thyroid Disease Data Set

The data is in the `.csv` file `thyroid.csv` . It consists of $3,772$ rows (patients) on whom 28 variables were reported, namely:

`ThyroidClass`
Binary classification label indicating whether **sick = 1** or **negative=0**

`patient_age`
Age of the patient

`patient_gender`
Flag indicating gender of patient - **1=Female** and **0=Male**

`presc_thyroxine`
Flag to indicate whether thyroxine replacement prescribed **1=Thyroxine prescribed**

`queried_why_on_thyroxine`
Flag to indicate query has been actioned

`presc_anthyroid_meds`
Flag to indicate whether anti-thyroid medicine has been prescribed

`sick`
Flag to indicate sickness due to thyroxine depletion or over activity

`pregnant`
Fag to indicate whether the patient is pregnant

`thyroid_surgery`
Flag to indicate whether the patient has had thyroid surgery

`radioactive_iodine_therapyI131`
Indicates whether patient has had radioactive iodine treatment: https://www.nhs.uk/conditions/thyroid-cancer/treatment/

`query_hypothyroid`
Flag to indicate under active thyroid query https://www.nhs.uk/conditions/underactive-thyroid-hypothyroidism/

`query_hyperthyroid`

Flag to indicate over active thyroid query https://www.nhs.uk/conditions/overactive-thyroid-hyperthyroidism/

`lithium`
Lithium carbonate administered to decrease the level of thyroid hormones

`goitre`
Flag to indicate swelling of the thyroid gland https://www.nhs.uk/conditions/goitre/

`tumor`
Flag to indicate a tumor

`hypopituitarism`
Flag to indicate a diagnosed under active thyroid

`psych_condition`
Indicates whether a patient has a psychological condition

`TSH_measured`
A TSH level lower than normal indicates there is usually more than enough thyroid hormone in the body and may indicate hyperthyroidism

`TSH_reading`
The reading result of the TSH blood test

`T3_measured`
Linked to TSH reading - when free triiodothyronine rise above normal this indicates hyperthyroidism

`T3_reading`
The reading result of the T3 blood test looking for above normal levels of free triiodothyronine

`T4_measured`
Free thyroxine, also known as T4, is used with T3 and TSH tests to diagnose hyperthyroidism

`T4_reading`
The reading result of th T4 test

`thyrox_util_rate_T4U_measured`
Flag indicating the thyroxine utilisation rate https://pubmed.ncbi.nlm.nih.gov/1685967/

`thyrox_util_rate_T4U_reading`
The result of the test

`FTI_measured`
Flag to indicate measurement on the Free Thyroxine Index (FTI)https://endocrinology.testcatalog.org/show/FRTUP

`FTI_reading`
The result of the test mentioned above

`ref_src`
Indicates the referral source of the patient

Requirements

- Create a dashboard with tabs:
  - one for a recap of numbers of individual patients (allowing the user to choose the patient);
  - one with intuitive viz or data table summaries of individual variables (possibly by subgroups);
  - one for comparing groups of patients (allowing the user to choose the variable(s) by which grouping the data).
- Separately generate a HTML report providing insight into which variables might be the best preditors of thyroid disease.
- Handle missing value properly:
  - Create new variables that are copies of variables with missing values where the missing values have been imputed. Justify the method;
  - In the dashboard, when a displayed number comes from your imputation process, find a way to signal it to the user.
- Use **flexdashboard** for building the dashboard;
- Use **ggplot2** for providing visualisations;
- Use **tidyverse** packages for data wrangling;
- Exploring new packages not mentioned in class that are useful to the dashboard or the report will add a bonus to the final mark.

Useful packages

- **skimr:** helpful quick summary of data
- **janitor**: fast cleaning of data
- **naniar**: visualize missing data
- **simputation**: impute missing data
- **forcats**: handle reordering of categories of categorical variables (useful for plotting)
- interactive flexdashboards