Nordic HDM and Nopho database and Analysis App

This is a Streamlit-based web application for the analysis of blood samples in HDM (High Dose Methotrexate) treatments. The app provides a user-friendly interface to visualize and process blood sample data, perform diagnostics, and export the results.

File Structure

The application is organized into several modules to ensure a clean and maintainable codebase. The main components are:

- **constants.py** : This file contains constant values and configurations that are used across the application. Constants include database connection parameters, time intervals, and default options.
- **datasets.py** : This module defines functions for fetching and processing data from the database or other data sources. It handles data manipulation tasks, such as merging dataframes, cleaning data, and filtering records.
- **diagnostics.py** : This file implements functions for analyzing and diagnosing the data. It includes classes for managing diagnostic types and running various diagnostic tests on the data.
- **processing.py** : This module includes functions for additional data processing, filtering, and transformations. It also provides utility functions, such as creating time intervals and downloading data as a CSV file.
- **visualization.py** : This file contains functions and classes responsible for creating and displaying visualizations within the Streamlit application.
- **app.py** : This is the main Streamlit application file, where the UI components and application logic are defined. It imports necessary functions and components from the other files to create an interactive and responsive web application.

To get started, please follow the installation and usage instructions in the next sections.

Database Structure and Connection

This application uses Snowflake as its database management system to store and retrieve blood sample data. Snowflake is a cloud-based data warehouse that provides high scalability, performance, and concurrency. The database structure is organized into a series of tables, each with specific columns for storing relevant information about the blood samples and HDM treatments.

Database Tables

The main tables in the Snowflake database are:

- **HDM** : This table contains information about the HDM treatments, such as the patient ID, treatment date, and the dose of methotrexate administered.
- **Nopho** : This table contains information about the Nopho treatments, such as the patient ID, treatment date and clinical observation data.
- **Blood** : This table contains information about the blood samples, such as the patient ID, sample date, and the sample type (e.g. peripheral blood, bone marrow, etc.). Sample types are defined by IUPAC codes. The table also contains the patient ID, which is used to link the blood samples to the corresponding HDM, 6mp and Nopho treatments.
- **Temperature and blood pressure** : These tables contain information about the temperature and blood pressure measurements, such as the patient ID, measurement date, and the measurement value.

These tables are related to each other through primary and foreign keys, allowing for complex queries and data analysis.

Connecting to Snowflake

To connect the Streamlit application to Snowflake, we use the snowflake.connector library, which allows for seamless communication between the app and the database.

The connection parameters, such as account, user, password, and warehouse, are stored with the user credentials

Once the connection is established, the application can execute SQL queries to fetch data from the Snowflake database and perform data processing tasks using the functions defined in the datasets.py, diagnostics.py, and processing.py modules.