Group 9 Data Warehousing project

Version correct to 26/11/2021 15:39

Table of contents

Introduction

A brief introduction to the project

Instructions

The following content is instructions to execute the python scripts for this project succesfully

Creatting a PostgreSQL user and database

To create the database, the user MUST execute the follwing commands in the PostgreSQL shell. (PostgreSQL is available at https://www.postgresql.org/download/)

`CREATE USER g09 WITH PASSWORD 'g09' CREATEDB;` where the username and password are the choice of the user.

```
CREATE DATABASE G09_DATA_VAULT1;
```

Once this has been completed, the DBMS is ready for the scripts to be ran.

Preparing your python environment to run the code

This project was developed on Python Version 3.8. While other versions may work, it is recomended to use this version to avoid any issues.

Python 3.8 is avaiable at: https://www.python.org/downloads/release/python-380/

Required packages for this project:

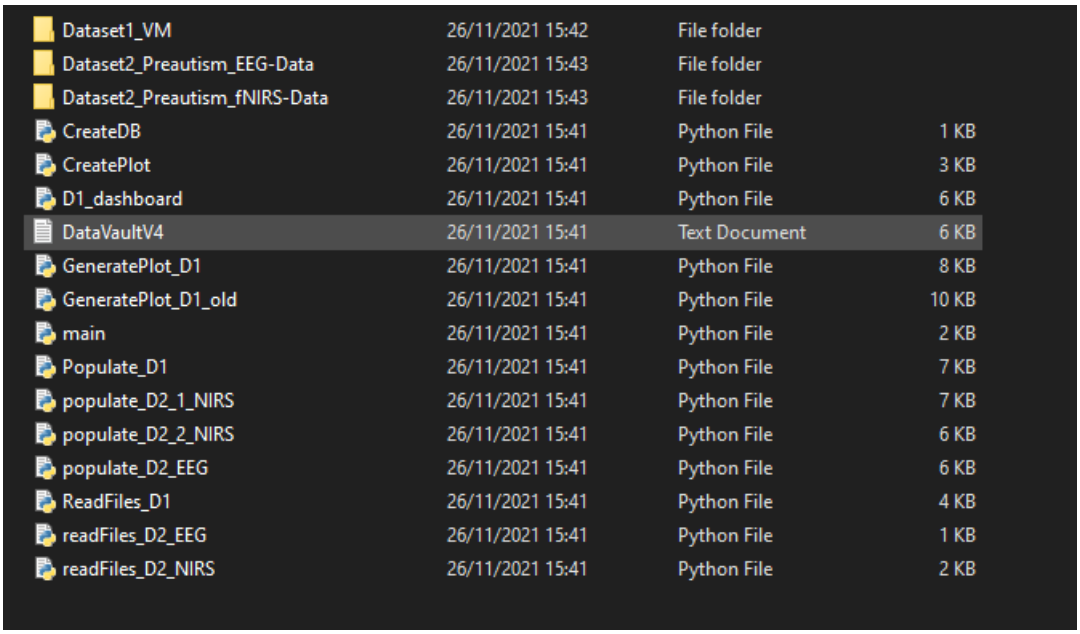
- easygui
- psycopg2
- pandas
- numpy
- plotly

Please use `pip install <package name>` in your Python environment.

Warning: do not run the code without completing this step, it will not work!

Populating the database

Before running any scripts, within the Code folder please ensure that you have pasted the correct folders for dataset 1 and 2 so that they REPLACE the folders as seen in the image below



To populate the database, the user must run `main.py`. This usually takes around 20-40 minutes to run depending on computer performance.

Running data marts

To run the data marts, just run any of the following scripts:

- Dataset 1:
  - `D1_dashboard.py`
- Dataset 2:
  - EEG: `not yet available`
  - fNIRS: `not yet available`

Understanding the data marts

Dataset 1

Some content about how to change the dash and downlaod button

Dataset 2

`not yet available`