Ed-Fi Data Views

License Agreement

We ask that all licensees DocuSign our license agreement before gaining access to our code repository. To sign view and sign our license agreement please email marcos@innovateedunyc.org

Overview

This project contains a set of scripts that can extract data from an Ed-Fi ODS (data standard v2.2) to create analytics views in Google Sheets and Google Cloud SQL. The extracted data is transformed to meet this format which we believe is an optimal format for building data dashboards in Google Data Studio.

The resulting Google Sheets or Cloud SQL tables can be connected to Google Data Studio and used to power the report library below.

Report Library

- School Profile
- Student Profile
- Student Lists

Install

This scripting library runs in a Python virtual environment. Instructions below for Windows and macOS.

Windows

We recommend installing Chocolatey, a package manager for Windows that's useful in managing multiple versions of Python on a machine.

Once Chocolatey is installed, open PowerShell as an administrator and run the commands below. This installs Git and Python version 3.6.5 on your machine. You will need pipenv installed to manage your Python virtual environment and necessary packages to run the scripts.

```
> choco feature enable -n allowGlobalConfirmation
> choco install git -y
> choco install python --version 3.6.5 -y
> pip install pipenv
> git clone git@bitbucket.org:datavizstarterpack/ed-fi-data-views.git
> cd ed-fi-data-views
> pipenv --python 3.6.5
> pipenv sync
```

macOS

We recommend installing Homebrew, a package manager for MacOS.

Once Homebrew is installed, open the terminal and run:

```
$ brew update
$ brew install git
$ brew install pyenv
$ pyenv install 3.6.5
$ git clone git@bitbucket.org:datavizstarterpack/ed-fi-data-views.git
$ cd ed-fi-data-views
$ sudo pip install pipenv
$ pyenv local 3.6.5
$ pipenv --python 3.6.5
$ pipenv sync
```

Configuration

The scripts in this repository are designed to store the extracted data in either Google Sheets or Google Cloud SQL (MySQL). Make a copy of **config-sample.ini** and name it **config.ini**. Complete the section *[SIS_ODBC]* with your Ed-Fi ODS server name, username, password, and database name.

Google Sheets

Our scripts use the Python package *pygsheets* to push data to Google Sheets. Follow the instructions located here to enable the Google Sheets API, create a service account, and download a .json file that contains the private key for account authorization. Rename the .json file **credentials.json** and copy it into the script folder.

When you created the service account, you received an email address for the user. This email is also stored in your credentials.json file. Create a folder in your Google Drive where the Sheets will be stored and share the folder with this email.

Finally, copy the folder id of that newly created Google Drive folder and store the id in your config.ini file under *folder_id*.

Cloud SQL

To push data to Cloud SQL you will want to create a MySQL instance in your Google Cloud. Once created, click on the instance name, connections, and whitelist the IP address of the server that will be running the scripts.

Edit your config.ini file to include connection details for your Cloud SQL instance.

Running Scripts

Thanks to pipenv you can easily enter your Python virtual environment and run a script via the command below (replacing *script_name* with the name of the script you'd like to run).

```
pipenv run python edfi\script_name.py
```