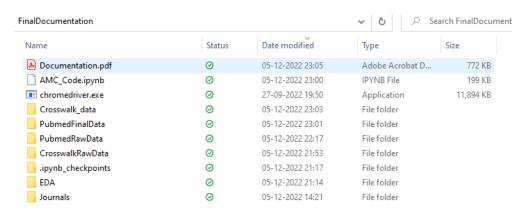
## **Documentation**

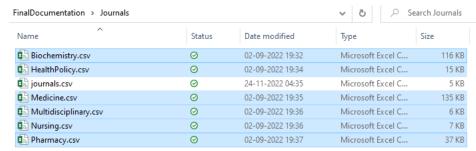
Fall 2022, Healthcare Analytics: AMC By: Bhavya Kashyap, Rida Fathima

This document provides information about the location of raw-data files, code, and resultant data. The folder locations are with respect to the current folder **FinalDocumentation**. The contents of which are:

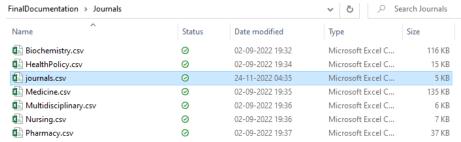


Note: AMC\_Code.ipynb is the master code file.

Extracting journals according to SJR rank:
Raw Data has been downloaded and saved in folder *Journals*.



The final data file containing the names of all required journal names has been created and saved in folder *Journals*.

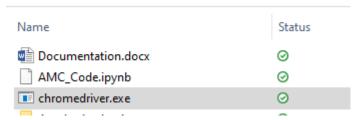


2. The *Exploratory Data Analysis* section of code imports two raw data files which are saved in the *EDA* folder as mentioned in the screenshot below.

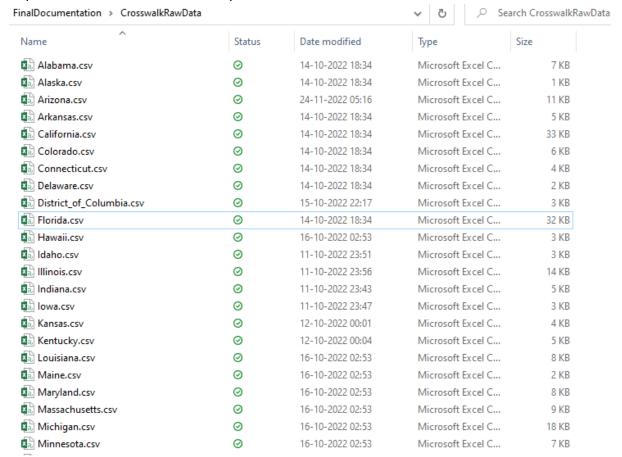


3. The *Creating crosswalk* section of the code requires a *chromedriver.exe* present which is present in the same folder as the code. Please ensure the exe file stays in the same folder as the code (this is necessary for the code to run).

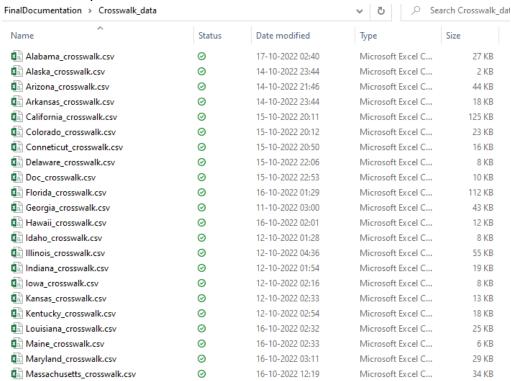
FinalDocumentation



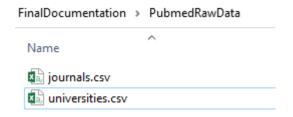
4. The *Creating crosswalk* section of the code requires raw data of each state to be imported those raw data files are present in *CrosswalkRawData* folder.



5. The *Creating crosswalk* section of the code creates crosswalk data of each state, result data files are present in *CrosswalkData* folder.



6. **Downloading bibliographical data for articles of pubmed** section of code requires raw data files of universities.csv and journals.csv which are stored in **PubmedRawData** folder.



7. **Downloading bibliographical data for articles of pubmed** section of code creates result data files for each university, these are stored in **PubmedFinalData** folder.

FinalDocumentation > PubmedFinalData

Name	^
ALBANY MI	EDICAL COLLEGE_FULLRECORD.CSV
ALBERT EIN	STEIN COLLEGE OF MEDICINE_FULLRECORD.CSV
BAYLOR UN	IIVERSITY_FULLRECORD.CSV
BOSTON UN	NIVERSITY_FULLRECORD.CSV
🖺 BROWN UN	IIVERSITY_FULLRECORD.CSV
CAMPBELL	UNIVERSITY_FULLRECORD.CSV
CASE WEST	ERN RESERVE UNIVERSITY_FULLRECORD.CSV
🖫 CENTRAL M	MICHIGAN UNIVERSITY_FULLRECORD.CSV
CITY UNIVE	RSITY OF NEW YORK_FULLRECORD.CSV
COLUMBIA	UNIVERSITY_FULLRECORD.CSV
🖫 CORNELL U	INIVERSITY_FULLRECORD.CSV
CREIGHTON	N UNIVERSITY_FULLRECORD.CSV
🗓 DARTMOUT	TH COLLEGE_FULLRECORD.CSV
🗓 DARTMOUT	TH_FULLRECORD.CSV
🗓 DREXEL UN	IVERSITY_FULLRECORD.CSV
🗓 DUKE UNIVI	ERSITY_FULLRECORD.CSV
🖫 EAST CARO	LINA UNIVERSITY_FULLRECORD.CSV
🖺 EAST TENNI	ESSEE STATE UNIVERSITY_FULLRECORD.CSV
🖺 EASTERN VI	RGINIA MEDICAL SCHOOL_FULLRECORD.CSV
🖺 EDWARD VI	A COLLEGE OF OSTEOPATHIC MEDICINE_FULLRECORD.CSV
EMORY UN	IVERSITY_FULLRECORD.CSV

NOTE: For universities for which FullRecord.csv (final csv) is not obtained, the filtered data wasn't available on pubmed.