

Assignment 4

Edward Sung
PUBH6859
4/22/24

Problem 1 Locations:

- Python Notebook:
 - https://colab.research.google.com/drive/1kY2yvWehpnQc_9d6YJveHXt0OBDspb9L?usp=sharing
- Google Drive Data:
 - https://drive.google.com/file/d/1-ThgtXtpFlHyxIBgLn6MWWAhWZ5TRjYo/view?usp=drive_link

Problem 1:

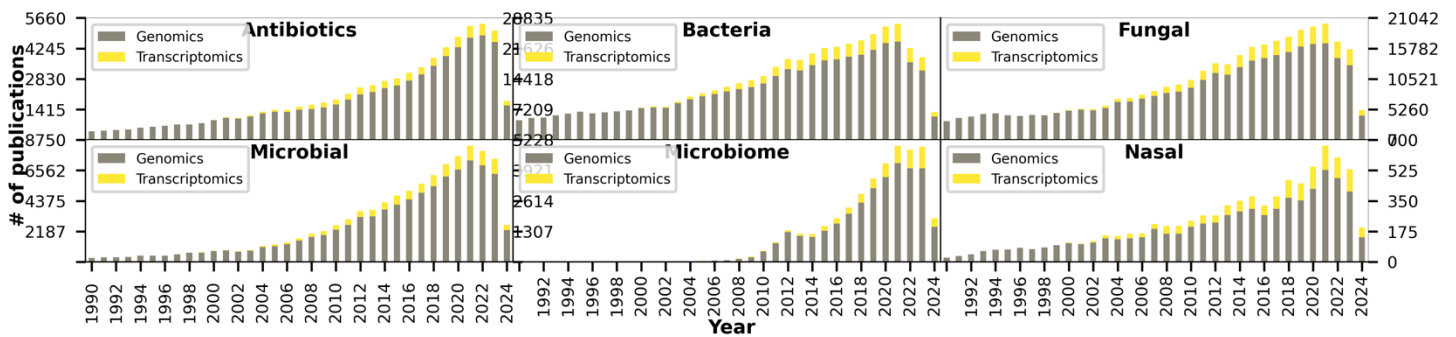
1. Go to PubMed_Visualization.ipynb github link.
 - a. https://github.com/omicsEye/pubSight/blob/main/notebook/PubMed_Visualization.ipynb
2. Can fork over the git repo to open and modify, but I instead downloaded the notebook and uploaded it into my own personal google colab account.
 - a. Renamed it as: PubMed_Visualization_EHS_Modified.ipynb
3. Go to: <https://pubmed.ncbi.nlm.nih.gov/>
 - a. Under the search bar, click Advanced
 - b. Following the table below:
 - i. Type in Main Term
 - ii. Press ADD button
 - iii. Type in AND Sub-Term
 - iv. Press ADD button
 - v. Press Search
 - vi. Under Results by Year, move slider to be from 1990 to 2024
 - vii. Download the csv using the download button
 - viii. Rename as: GROUPNAME(underline)SUBGROUPNAME.csv as instructed in notebook
 - ix. Save to a folder in google drive
 - x. Repeat for all rows in the table.

Main Term	AND Sub-Term
Bacteria	Genomics
Bacteria	Transcriptomics
Fungal	Genomics
Fungal	Transcriptomics
Antibiotics	Genomics
Antibiotics	Transcriptomics
Microbial	Genomics
Microbial	Transcriptomics
Microbiome	Genomics
Microbiome	Transcriptomics
Nasal	Genomics
Nasal	Transcriptomics

Assignment 4

Edward Sung
PUBH6859
4/22/24

4. In the python notebook, run the code:
 - a. Mount my google drive
 - b. Change: drive_path to be:
drive_path = '/content/drive/MyDrive/PUBH6859_Assignment4/'
 - c. Notebook will create png and pdf graphics of the charts



Problem 2:

- pubh6859_assignment4_boto3_create.py
- pubh6859_assignment4_boto3_delete.py
- Confirmed the scripts works using the AWS site to see it being made and deleted