Step01: 01ProteinCodingGeneList.ipynb

Upload ProteinCodingGenesRetreived01192022.csv and save it

in directory GTEX\_Intermediate\_data as pcgSymbols

Note: 01/19/2022 was the protein coding gene list on January 19, 2022

The gene\_list is provided for replication of process purpose. If you

Wish to apply a more up to date list please download the current list

From the following URL. <https://www.genenames.org/download/statistics-and-files/>

Step02: 02ConvertdotGCTtoPandas.ipynb and save it

Read: 'GTEx\_Analysis\_2017-06-05\_v8\_RNASeQCv1.1.9\_gene\_tpm.gct'

Extract the protein coding genes and save it a a csv file

Note: the unzipped file is 4.26 GB in size please download the file

And place it in the appropriate directory.

The source can be found in:

<https://gtexportal.org/home/datasets>

the file name is:

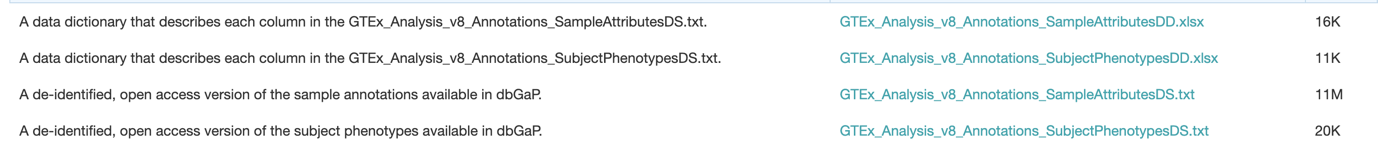
GTEx\_Analysis\_2017-06-05\_v8\_RNASeQCv1.1.9\_gene\_tpm.gct.gz

The output will be all gene samples but will include only genes that are in the intersection of the of protein coding genes the genes listed in GTEx.

Step03: 03SetupSamplteAttributes.ipynb

To run code please download the following files from the GTEx datasets

Note: For cohort1 we filter to include only death type 1 and death type 2



This code will clean sample attributes.

Step04: 04CreateDataForTissues.ipynb

Splits data into specific tissue types

Step051: 051ProcessTissue.ipynb

Preprocess at the tissue level

Step052: 052SKLearn.ipynb

Multiple linear regression at the tissue level to correct for confounding factors

Step053: 053R.R

R-code (used R studio) pre-processing prior to running WGCNA

Step054: 054R.R

Run WGCNA to per tissue

Step055: 055rrenameAsNumbers.ipynb

Step061: 061EnrichR.ipynb

Maayan Lab enricher API to determine KEGG pathways the satisfy enrichment levels per plurality vote

Step062: 062EnrichRPluralityv2.ipynb

Infer based on plurality vote from step 061 module function

Create a matrix to include modules with in one place

Step1: TissueBigMatrix1.ipynb

Step2: TissueBigMatrix2.ipynb

Build Graphs:

Found in: Dynamic19Cohort1/TissueComparisons/BigMatrix