

README for 'Modeling Liquidity Needs of Retirees'

Data Availability and Provenance:

For this project, we use Health and Retirement Study data. We use the 2020 Longitudinal file which is a cleaned version published by the RAND Center of Aging. It is publicly available (with the creation of an account) at the following link:

<https://hrsdata.isr.umich.edu/data-products/rand-hrs-longitudinal-file-2020>

The raw dataset is too large to be included in our deposit. However, we include a zipped version of the subset containing the variables and waves of interest to use, hrs_new.zip.

Computational Requirements:

For this project we use SAS, R, and Python. Note: that SAS is not required for this project as we only use it to read in the data and perform the initial variable selection. The same task could be easily performed in R; however, due to the size of the dataset the task is greatly expedited by using SAS. The file clean_data.R is the most computationally intensive and can take over an hour to run. All other code should run fairly quickly. Specific computational requirements listed below:

- SAS Version 9.4 TS1M8 (9.4 M8)
- R Version 4.3.0
 - tidyverse
 - priceR
 - Gtools
 - stringi
 - ggplot2
 - gridExtra
- Python 3.12
 - Sklearn
 - Scipy
 - Statsmodels
 - Matplotlib
 - Seaborn
 - Skbio
 - Fitter

Instructions to Replicators:

To replicate the results from this project, run the provided code in the following order. Note that the resulting dataset from the SAS code (hrs_new.zip) is provided in the deposit so running the SAS code is not necessary (if access to SAS cannot be obtained). The working directory must be adjusted in main.R and download_and_sample.sas:

1. download_and_sample.sas
2. main.R
3. Unsupervised Learning.ipynb (running this notebook is time-consuming)
4. Simulation.ipynb

Figures and Tables:

Figure/Table #	Created By
Figure 1	EDA.R
Figure 2	EDA.R
Figure 3	EDA.R
Figure 4	method1.R
Figure 5	method1.R
Figure 6	method1.R
Figure 7	method1.R
Figure 8	EDA.R
Figure 9	Unsupervised Learning.ipynb
Figure 10	Unsupervised Learning.ipynb
Figure 11	Simulation.ipynb
Figure 12	eval_sim.R
Figure 13	Simulation.ipynb
Figure 14	eval_sim.R

Figure 15	eval_sim.R
Figure 16	eval_sim.R
Figure 17	data_clean.R
Table 1	N/A
Table 2	method1.R
Table 3	method1.R
Table 4	method1.R
Table 5	method1.R
Table 6	Unsupervised Learning.ipynb
Table 7	Simulation.ipynb
Table 8	Simulation.ipynb
Table 9	Simulation.ipynb
Table 10	eval_sim.R

References

"Health and Retirement Study (HRS) Data Products." Institute for Social Research, University of Michigan. Accessed. <https://hrs.isr.umich.edu/data-products>.