Data Sources for Project

Here is a suggested list of datasets for your project.

Financial (Time Series) Data

1.   Federal Reserve Data download -- http://www.federalreserve.gov/datadownload/
     - My section maintains data on the "Financial Accounts of the United States" available at: http://www.federalreserve.gov/datadownload/Choose.aspx?rel=Z.1
     - The St. Louis Fed, economic data: https://fred.stlouisfed.org/
2.   Data Market Time Series Data Library -- https://datamarket.com/data/list/?q=provider:tsdl
3.   INFORUM -- http://inforumweb.umd.edu/econdata/econdatacontents.html *(requires Inforum's database and regression package, G, to access that data)*

General Data

1.   DATA.GOV (U.S. Government's open data) -- https://www.data.gov/
2.   NBER http://www.nber.org/data/
3.   UCI Data Repository -- *Machine learning datasets* -- http://archive.ics.uci.edu/ml/
4.   4: caesar0301/awesome-public-datasets -- https://github.com/caesar0301/awesome-public-datasets

Panel Data

R example of how to process: http://dss.princeton.edu/training/Panel101R.pdf

Big Data repositories

1.   Data Science Central -- http://www.datasciencecentral.com/profiles/blogs/big-data-sets-available-for-free
2.   Big data made simple -- http://bigdata-madesimple.com/70-websites-to-get-large-data-repositories-for-free/
3.   Amazon Web Service (AWS) -- https://aws.amazon.com/public-data-sets/

A worked example using the "Wine" dataset from UCI data repository

See github lecture3:

1: PDF version

https://github.com/wampeh1/ECOG_314/tree/master/project1/pdf/lecture3_project_guide_multivariate_data_analysis_example.pdf

2: R Markdown version:

https://github.com/wampeh1/ECOG_314/tree/master/project1/lecture3_project_guide_multivariate_data_analysis_example.rmd