

## Assignment 4

For this assignment, you'll need to open up, clean and save the following datasets, using the tools we've gone over in class. For each dataset, make sure that when you're done you have a nice, neatly labeled dataset that would be easy for you or another analyst to open and analyze. Save the result in an RData file using the names provided. You need to turn in an R script named 04-assignment\_<lastname>.Rmd that creates all of the datasets requested (lower case only for your last name, please). "Clean" only as needed.

1. Panel data for OECD countries. Save as file name `oecd.RData` <http://www.wiley.com/legacy/wileychi/baltagi/supp/Gasoline.dat>

Hint: try the `read_delim()` function.

Remember to inspect the dataframe to assure the data was read in and formatted correctly.

2. U.S. National Election Survey data from 2004. Save as file name `nes.RData`. <http://www.uta.edu/faculty/story/DataSets.htm>

Hint try: `read_spss("http://www.uta.edu/faculty/story/DataSets/NES2004.sav")`

3. General social survey, student version. Save as file name `gss.Rdata`. <http://www.uta.edu/faculty/story/DataSets.htm>

Hint:

try:  
`read_excel("GeneralSocialSurvey1996ExcelDataStudentVersion.xls", skip=0, col_names = TRUE)`

Then use `save(...)`.

4. Replication file for "STATISTICAL DISCRIMINATION OR PREJUDICE? A LARGE SAMPLE FIELD EXPERIMENT". Open up and save the `mainData.csv` file. Save it as `maindata.RData`. [Available here](#), look under the "Data and Analysis" tab.

Hint try `read_csv(...)`

5. The Lalonde dataset, covering work experiences in the Panel Study of Income Dynamics (`psid`). Save as `psid.RData` [http://users.nber.org/~rdehejia/data/psid\\_controls.txt](http://users.nber.org/~rdehejia/data/psid_controls.txt) You'll find a description of the data that you'll need [here](#).

Hint: try `-read_delim(...)`

Are the col labels missing???? If so use `names(...)`