

## Spring 2017

### Due Dates:

- **February 10, 2017** – Submit a brief write up of your economic question and possible data sources
  - This should include at least one paragraph on the background for the research question and why it is an economically interesting topic.
  - When looking for data sources some possible resources are IPUMS, Bureau of Labor Statistic, Bureau of Economic Analysis, Federal Reserve Board, and Federal Reserve Bank of St. Louis (FRED).
- **February 17, 2017** – Confirm that you have already reached out to your assigned Board economist to set up an initial discussion of your economic question and data sources.
  - Forward us an email exchange with your economist.
- **February 17, 2017** – Submit a more detailed project proposal. This should include:
  - A one page discussion of the background for your economic question. You should briefly discuss the analysis of at least one other paper that is related to your project.
  - Your economic question should be stated clearly and succinctly. It should also be stated in a way that makes clear what variables you will be trying to use. For example: *In my project I will explore whether the great recession lead to an increase or decrease in divorce rates controlling for income, race, education, and location. It is unclear what effect I would see. I might expect an increase in divorce rates since the recession was a great stress on many marriages but at the same time getting a divorce is expensive and so I might see no effect or a delayed effect.*
  - Provide some overview of the econometric methodology you intend to use in your project.
  - You should briefly discuss what data you intend to use.
- **March 3, 2017** – Submit code used for data loading, cleaning, and validation. Your code should:
  - Load in your data from a file/multiple files.
  - Deal with missing observations and variable recoding.
  - Merge any data sets that need merging
  - Produce summary statistics
  - Be easy to read - follow good style guidelines and include appropriate comments
  - Be able to be run by someone else
  - Datasets you are using must be explicitly defined and url location of downloaded data from the internet must be given in the comments.

- **March 10, 2017** - Plots and Summary Statistics. Submit four plots and/or tables, (at least one of each), that show substantive information about the data set that will help you address your economic question.
  
- **March 24, 2017** – Submit check-in presentation materials.
  - Your check-in materials should cover parts 1 and 2 (Introduction and Data Description)
  - You are expected to submit a short technical write-up that explains the central question your final paper will answer and how your dataset will help you answer that question. At a minimum it should include:
    - Your question and why it is an important question to answer
    - The data set you are using and the datasets you used to build it
    - Where you sourced your data from
    - Why this data will help you answer your question
    - Key summary statistics
    - Graphs/Tables as appropriate to help elaborate on the above
  - You will also submit **all code that you used to prepare the data**, including (but not limited to) any merges, reshapes, summarizations, etc that you performed to make your data ready for presentation. It is expected that if we download your data untouched, run it through your R code, and only your R code, the end result will be your final dataset. The person presenting your data will use this codebook to explain how you cleaned your data in R, so you will need to follow good coding standards and style including indenting your code and commenting your code.
  - Your grade for the write-up will be based on the following:
    - Codebook
      - Does the code run, once the path location of the original dataset is adjusted? Does the output of the scripts match the final output you use for the write-up?
      - Code formatting and readability. Can people read your code and easily understand what you are doing? Is your code concise? How clear and useful are your comments?
      - It is **strongly** suggested that you test your codebook, (i.e. run only the code in your codebook on the raw data), before submitting the codebook to ensure that your results are as expected.
    - Write-up
      - Formatting –Did you include graphs and R output where necessary/appropriate? Do the graphs, code, and output included

in the document add to the message and understanding or are they an unnecessary distraction?

- Content – Does your content flow logically from topic to topic, does this paper do an effective job in setting the stage for the final project by providing us with background on your question, relevant data, etc. Do you include all of the topics discussed above?
- Validity and Readability – Does the project make sense, does your dataset make sense to use in answering your question? Do your cleaning methods make sense?

- **March 31, 2017** – Check-in presentations

- You will be given the paper and R codebook from another student of our choosing. Using these files, you will prepare a PowerPoint presentation, no more than 10 minutes, which covers at least the following topics:
  - What the central question is
  - Where the data came from
  - Relevant summary statistics
  - How R was used to prepare the data set, including specific data manipulation packages and functions used
- As above, you are not expected to provide any results from analysis. This is an overview of the central question and data used to answer that question, along with specifics on how R is being used to prepare the data.
- As noted above, any graphs and/or tables you use in your presentation must be pulled straight out of the Word document that you are given. You cannot create any visual aids independently to explain the data. The only “original” material that will be in your PowerPoint presentation will be the bullet points on the slides.
- The presentation is expected to cover the data/central question and the R code in equal amounts (ie, dedicate about 5 minutes to data presentation, and 5 to the R code used).
- For your presentation, you will be graded based on how you present the information you are given (i.e. you will not lose points if you are assigned a paper that included no graphs, or incomplete code). The grade for your presentation will be made up of two categories:
  - Oral Delivery
    - Do you know the presentation and understand the material?
    - Are you able to answer basic questions from audience?
  - Written PowerPoint
    - Does your presentation flow from topic to topic?

- Does your presentation cover all the aspects of the project needed to be covered in the write-up?
  - Does the presentation make sense?
- **April 7, 2017** – Regressions. Submit the results of at least one regression relevant to your analysis with an accompanying paragraph explaining the significance and meaning of the regression and its results. This corresponds to part 3 in the above rubric.
- 
- **April 24, 2017 (note this is a Monday)** – Submit final project. This will include:
  - All sections 1-5 of the rubric from above.
  - The R-scripts containing code used for analysis with detailed comments of your code. If you are submitting multiple files you will need to provide us with a description of each file. Your scripts should contain the location where you downloaded the raw data used in your analysis. All your code must be commented and formatted so that they are easily understandable. Someone else should be able to re-run your entire analysis producing identical results to what was submitted.
  - A write up of your work. The write-up is expected to be completed in Rmarkdown and the .Rmd file as well as the knit html or pdf is required for submission. You may choose whether you prefer html or pdf.
  - A Beamer presentation of your results to give to the class. (This is the pdf type of slides that can be made in Rmarkdown). Your written presentation document should be in the form of beamer slides written in Rmarkdown. You should submit both an Rmd file as well as the pdf file for the output. The content is expected to be a summarized version of the write-up and you may use the same material in both. Your final presentation should briefly discuss your question, your data, and your data cleaning. The bulk of the presentation should be on your analysis, methods, and results. Presentations should be 15 minutes.