# Individual Data Project Guidelines
## (DRAFT for peer-review DUE April 17th; FINAL DUE April 30th)

For this project, you will work individually on analyzing a time series that we **have not looked at in class**. You will analyze the time series using the features we have discussed in class (trends, seasonality, autocorrelations) and conduct a basic validation exercise to determine how well your model forecasts. In addition to comprehensive code written in R, you must provide a **one-page summary** of your methods (double spaced, not including data visualizations), the reason why chose the model you did, and interpretation of your results (see grading rubric below for more detail). Think of this as a project that you will receive from a future employer – it is very common to receive a data set with limited information and be expected to create a model for forecasting from scratch (and be able to articulate why you chose it!). Therefore, there are not very strict guidelines for this project, but the general grading criteria is outlined below:

**Data**
- For this project you must select your own dataset
- I suggest using data from the Federal Reserve Economic Data repository (FRED)
    - https://fred.stlouisfed.org/
    - These are all financial/macroeconomic time series that follow similar behavior to what we have discussed in class
    - You can look at standard time series that we have explored (e.g., GDP), except for a country other than the U.S.
- You can use data outside of FRED, I only ask that you run the data set by me (i.e., email me a plot) if you do. There are many interesting non-financial time series out there, but not all exhibit the characteristics that are applicable to what we have discussed in class

**Grading Criteria**
- **Data Visualization (30%)**
    - Create at least 3 data visualizations:
        - Simple time plot
        - A plot that helps determine the underlying structure of the data series (e.g., ACF)
        - A plot that demonstrates the predictive ability of your model
    - Explain what you learn from each plot and how it informs model selection and/or validation
- **Model-Selection (20%)**
    - Use tools from class to determine the presence of a trend and seasonality
    - Apply a statistical test for determining stationarity. Depending on this result, what action do you take and why?
    - Specify the criteria and method that you use to determine your final model selection
- **Forecasting exercises (30%)**
    - Use your final model in a prediction exercise. You may want to use a train/validation structure or a one period ahead structure, depending on your data. Just be specific on why you choose the method you did.
    - Use valid forecasting error measures to determine how well your model performs
    - Compare your final model to at least one other **baseline** model that we have discussed in class and use a statistical test to determine if it outperforms that baseline model
- **R-Code (15%)**
    - Your code needs to be neat, easy to read, and interpretable (Hint: Use comments and indenting)
- **Peer-Review (5%)**
    - Provide feedback (one-paragraph) on an assigned student's draft (DUE April 23rd)