Mini-project 2: Unemployment

S520

Upload your draft through the Assignments tab on Canvas by 2:30 pm, Thursday 23rd February.

Background

The file `UNRATENSA.csv`, downloaded from the Federal Reserve of St. Louis' FRED data repository, contains the unadjusted U.S. monthly unemployment rate from January 1948 to January 2017. We wish to break this time series down into components, and find out what might predict the unemployment rate.

Questions

- **Specific question:** Decompose the unemployment time series into some or all of the following components:
  - Seasonal variation
  - Trend
  - Oscillations/cycles
  - Residuals

  Describe the patterns in each of the components besides the residuals. Note: You might not need all four of these components.

- **Open question:** What might predict the unemployment rate in advance? Find at least one variable that has some ability to predict changes in the unemployment rate before they happen, and describe (including graphs) the relationship between the variable(s) you choose and the unemployment rate. You do not need to make quantitative predictions (since that's beyond the scope of this course), but you should convince Jake and me that the variable does have some predictive ability. You'll need to get additional data — you may use any data available on FRED:

`https://fred.stlouisfed.org/`

Directions and tips

Work in groups of three or four. One person per group should submit a set of answers (making sure that everybody's name is on the submission.) The submission should include a PDF, your code, and the additional data you use besides the unemployment CSV file. You can use any statistical or graphical technique you want except find P-values.

The maximum length of your write-up, including graphs, is four pages. Your graphs should be of readable size. Your write-up should aim to convince someone with reasonable knowledge of statistics of your main points.

The format of the data is not entirely convenient, so you may have to do some manipulation to get it into R in a usable form. You may not be able to exactly follow the examples in the notes. That's kind of the point — real data does not always work out the same way as textbook data does.

Grading

You'll be graded on four criteria:

- **Answer to the specific question.**
- **Answer to the open question.**
- **Presentation.** This include clarity and correctness of the graphs (i.e. we should understand what you're drawing) and clarity and correctness of the writing. Spelling and grammar will be a small but non-zero proportion of the grade.
- **Reproducibility.** Jake will attempt to reproduce all your graphs and numerical results from the materials you provide. If he can, you get all the points for this; if he can't, you get no points.

We will grade your initial submission and tell you how you can improve it, then you will have a week to resubmit it.