

Data analysis project guideline

ECO 595 Applied Business Research

1. Each group will be assigned a company by the professor. Revenue data and macro data can be found in Doc sharing on ecollege.
2. Time: the presentation should be no longer than 10 minutes. One group member can record the voice-over on behalf of the whole group or all members can share the recording.
3. What to submit: the Excel file that includes all your analysis and data, and the voice-over ppt file that includes your presentation. The file names should be 'Last name 1_Last name 2_Last name 3_project name' for group work and 'Last name_First name_project name' for individual work. Only **one** copy is needed from each group.
4. Part I Time Series analysis: discuss the revenue data similar to what I have showed in the Sear example is expected: scatter plot of revenue; linear trend; moving average; structural break. Your company revenue may not exhibit strong seasonality. For practice purpose, you are required to fit a moving average to your revenue data nevertheless.

First explain what the graph tells you as a manager about the trend, seasonality, and structural break in this company data. Then conduct research to find out how management actually made the decisions in the data.

Note: you do not need to provide the graphs that cannot be obtained using Excel (eg. the four figures shown in the second to last slide in the Time Series Analysis ppt file)

5. Part II Linear regression using Macro data:
 - a. Select one variable from the macro data file that you believe can be highly correlated with the revenue data.
 - b. Run a simple linear regression using the macro variable as the independent variable and the revenue data as the dependent variable. **Report the regression results: show the excel output.**
 - c. Note that the dates for the revenue data and macro variables do not necessarily match each other. You will need to determine the date range that has available data for all variables and drop the extra data. For example, if the date for GDP ranges from 2000 to 2015 and the revenue data ranges from 1995 to 2016. Then you would only keep data from 2000 to 2015.
 - d. Comment on the significance of the slope coefficient. Briefly discuss the relationship (positive or negative) between the dependent variable and the independent variable. Is the relationship expected?
 - e. Report R-squared and briefly comment on the fitness of the model based on your opinion.

For d and e, since you cannot actually point to the numbers during a voice-over presentation, you must highlight the relevant excel output numbers on the slide about during your discussion. You can also use an animated arrow to point to the numbers if you want to.

6. Again:
 - a. make sure the font you use is big enough for the audience to read;
 - b. do not put everything you want to say on the ppt slides;
 - c. do not read the ppt slides;
 - d. rehearse before formal presentation;
7. Rehearse before formal presentation: make sure you review the general comments given to presentations earlier in the semester and avoid mistakes discussed there.