

This outline is preliminary for the Spring 2024 class. The general topics will not change, but some sections will be expanded to be more explicit regarding what is expected. You will be notified of the changes either before or as we get to them in this outline.

You must choose two companies to study. For each company, you must look up net revenue or sales data. The two companies must be chosen according to the following criteria: (a) You should feel that one company will have seasonal variation in their sales and the other will not have seasonal variation. You should write a brief (one paragraph for each company) reason for why you believe seasonal variation will or will not be present in the sales data. You should also summarize the basic structure of each company's product line(s); (b) You should make certain that you can obtain at least 10 years past sales or revenue data if it is reported quarterly or 5 years past sales or revenue data if it is reported monthly; (c) You should make sure that the data are reported consistently throughout the time series you have chosen. It is suggested that you choose companies that have fairly narrow product lines (conglomerates are out) and whose primary revenue source is within the domestic market. No two students in the class may choose the same companies; all company proposals must be approved by the instructor. The first parts of the data analysis will involve both companies; the regression parts of the analysis will involve only one. For the regression analysis, you will have to find and test "explanatory" data. You may wish to keep this in mind when choosing companies.

The following outline may be amended as circumstances warrant. However, it will serve as the basic outline for your term project, so refer to it as we go through the class. If you keep up as we go along, you will find that there will be very little last minute cramming to finish the project. I will not be collecting the portions or assigning a score for them as we go along. I will, however, assign a final score for the overall project. The dates listed below are approximations as to when you *should* have completed that portion. If you would like input or feedback from me on where you are on the project feel free to request it. It is crucial that you allow sufficient time to complete this project as an inadequate performance will almost certainly mean that you will receive a poor grade in the course.

"In Process" dates

Part I: 05 February

Part II: 11 March

Part III: 1 April

Final due date: 26 April, 12:00 (noon)

Any late final submissions will lose 10% of the possible points for each hour (or portion thereof) submitted past the deadline. (And yes, I do realize that we do not have classes on 26 April. This remains the due date & time).

Outline

- I. Statement of the problem.
 - A. State the purpose of the project.
 - B. State each company you have picked and briefly describe its purpose and product line. State whether you feel the company should have seasonal variation in its sales or net revenue. Indicate the time period over which you are studying the company and indicate and special historical events that you will want to be aware of in your subsequent analysis.
 - C. Provide a copy of your dependent data series (with appropriate citations and documentation of sources).
 - D. Explain any transformations of the data that you have made to make it more usable. Show the transformed data and explain why and how you made the transformation(s).
- II. Analysis of the dependent variables (do this for each company).
 - A. Using a plot of the data over time, indicate what you feel are the important aspects of the data regarding any jumps in the data (can you explain them?), trend, and seasonality.

- B. Use correlograms to check for trend and seasonality. Take first differences to recheck for seasonality. If data is not consistent over the entire time span you have chosen, check the data in sub-intervals within that time span.
 - C. Explicitly state and test the hypotheses needed to check for trend and seasonality.
- III. Smoothing (do this for each company).
- A. Using the analysis performed in part II, indicate the best methodology you would use to smooth your data. Indicate why other choices would not be expected to do as good a job.
 - B. Perform the smooths and indicate the one you deem to be the "best." Explain the criteria you used to determine which is "best" using both "common sense" and "goodness of fit" criteria.
 - C. If your choices of smoothing techniques are different in parts A and B, explain why you changed your mind from parts A to B.
 - D. Use your best smooth to predict your company's sales or net revenue for one year into the future (from your last data observation).
- IV. Regression (do this for ONE of your companies).
- A. Data
 - 1. Show all of your independent data series. Indicate or show all sources for the independent data. Explain why you think each independent variable will explain variation in your dependent variable. Be sure to include any appropriate dummy variables in your discussion.
 - 2. Indicate your a priori idea(s) about the sign of the coefficient that an acceptable regression should generate for each of the independent data series.
 - 3. Explain any and all appropriate transformations and "massages" of the data you made to fit it into your model.
 - 4. Check the independent data to see if you might run into multicollinearity problems.
 - 5. Indicate if you think there may be heteroskedasticity problems and make any additional transformations you think will be necessary to address that problem.
 - B. Run your least-squares analysis using the independent variables identified in part A. (Do not run any autoregressive models here.)
 - 1. In the process of running the regressions, indicate which independent variables are good explanatory variables and which are not. Explain why you excluded certain variables from further consideration
 - 2. For your best regression in this section, explain the meaning of the R-squared statistics (both regular and adjusted), the F-statistic (state the hypothesis and perform the test of the hypothesis), and the t-tests involved with each estimated coefficient (again, state each hypothesis and perform the test of each hypothesis). Also, check your regression for serial correlation (see below), heteroskedasticity and the possibility of multicollinearity.
 - 3. In addition to checking for serial correlation using the Durbin-Watson statistic, check for any pattern in the residuals by examining a correlogram. (State all appropriate hypotheses and test them!) Explain any differences that may occur between using the D-W analysis and the correlogram analysis.
 - 4. For your best regression, derive the 95% confidence band for your forecasted dependent variable. Show the actual dependent variable along with the forecasted values of the bands on a plot.