Exam 2 - Study Guide

**Regression Analysis**

What is regression and what is it used for?

What is multicollinearity? Why does it matter? How do you detect it? What should you do about it?

What is heteroskedasticity? Why does it matter? How do you detect it? What should you do about it?

What is auto-correlation? Why does it matter? How do you detect it? What should you do about it?

What are R^2 and Adjusted R^2? Why prefer one to the other?

What are AIC and BIC? How can they be used to compare models?

How do you interpret the betas / coefficients in a regression equation? B0? Bi?

What are dummy variables? How are their coefficients interpreted?

How are seasonal dummy variables constructed?

What are interaction variables?

How can concave and convex relationships be handled by linear regression?

What are residuals? How are they used in diagnostics? Why is it important that they are normally distributed?

What are the common diagnostic plots used in regression analysis?

P-Values how to interpret them with respect to regression coefficients and statistical tests

What is logistic regression? How does it differ from OLS (standard) regression? How can you interpret the regression coefficients?

What are marginal effects?

What is elasticity?

*Be able to read and explain regression analysis output.*

**Time Series Analysis:**

What is time-series data? What is cross-sectional data? What is panel data?

What is stationary vs non-stationary data?

What is differencing? How is it done?

What is “time-series decomposition”? What are the components?

What is seasonality? How do we test for stationarity?

In very basic terms, what is an ARIMA model? What are the three parameters for an ARIMA model (p, d, q)?

**Python:**

I won’t ask anything too tricky, but I will have you annotate some code again and answer a few questions about basic commands, syntax, etc. Brush up on loops, comments, functions, as well as some Pandas and Statsmodels.

Regression Analysis Slides:

[Regression](https://docs.google.com/presentation/d/1IjGVFuMNY7qtXgxAImiQAuhnR8eQXtfSR0PTYdFGti8/edit?usp=sharing)

Logistic Regression Slides:

[5760: Logistic Regression](https://docs.google.com/presentation/d/1c9WzrpH9GaFmYow9WeeKd5_0tWLoSryO6DoQKVz-4TA/edit?usp=sharing)

Time Series Analysis:

[5760: Time Series Analysis](https://docs.google.com/presentation/d/11D2LMCx1vrjxNwjieskpmcMKEKE5vRL_tqOsYqm77Zc/edit?usp=sharing)

BE SURE TO STUDY THE NOTEBOOKS ON GITHUB!!!

https://github.com/rjwrobel86/Python4Statistics/tree/main/Notebooks