

# LLO 8200:Class Outline

# Catalog Description

We have entered a time in which vast amounts of data are more widely available than ever before. At the same time, a new set of tools has been developed to analyze this data and provide decision makers with information to help them accomplish their goals. Those who engage with data and interpret it for organizational leaders have taken to calling themselves data scientists, and their craft data science. Other terms that have come into vogue are "Big Data," "Predictive Analytics" and "Data Mining." These can seem to be mysterious domains. The point of this class is to demystify much of this endeavor for individuals who will be organizational leaders.

# Key Learning Goals

Understanding Data Structures and Measurement

Understanding How Models Can Be Used to Make Predictions

Understanding How Results of Data Analysis Are Presented

# Data Structures and Measurement

How do we turn the world into numbers: variable definition?

What kinds of variables exist: continuous, categorical, etc.

How do we turn a big pile of numbers (data) into a smaller pile of numbers (summary measures)?

How do we think about the ways that numbers measure concepts?

# Understanding How Models Can Be Used to Make Predictions

Basic concepts of modeling

Linear Regression: How we predict continuous outcomes (revenues, number of students enrolling)

Classification: How we predict group membership (whether an employee will be retained; what major a student chooses)

Deep Learning: Discovering patterns when there are a very large number of outcomes (Natural Language Models)

# Understanding How Results of Data Analysis Are Presented

- What are the right numbers to present to understand patterns in the data?
- When are different types of visualizations appropriate?
- What are some common pitfalls in data visualization?
- How can visualization best support decision making?

# How We'll Get There

- Biweekly assignments: getting some practice with key concepts, building toward final assignment
- Final assignment: using a dataset that we'll provide to answer some descriptive questions or do some modeling
- Readings
- Asynchronous Content

# Data Analysis Tools

Our primary tool will be Excel with the Data Analysis ToolPak

You can use LOTS of others

Common open-source programming languages include R and Python

Common paid programming languages include Stata, SAS and SPSS

There are many, many others

So long as You can do the analysis, you can choose which to use, but the class will support Excel and the Data Analysis ToolPak