Module 1 – Linear Regression Introduction

- Families of Predictive Analytics
    - Classification
    - Regression
- An Example Model
    - Linear Regression Models as Equations
    - A House Price Linear Regression Model

Module 2 – Numeric Literacy

- Types of Data
    - Categorical vs Numeric
    - Interval vs Ratio
- Measures of Location
    - Average (Arithmetic Mean)
    - Median
- The Normal Distribution
    - Distribution Parameters
    - Standard Deviation
    - Probability of Values

Module 3 – Predictions Using the Mean

- The Mean is a Predictive Model
    - When the Mean is a Reasonable Predictive Model
    - Why the Mean is a Predictive Model
    - Issues with Prediction Using the Mean
- Our First Model
    - Predicting the Iris Dataset with the Mean
    - Is the Model Any Good?
    - Measuring Model Goodness – Total Sums of Squares
- Hands-on Lab #1

Module 4 – Correlation

- What is Correlation?
    - Data Synchronicity

- o   Types of Correlation
- o   Correlation Strength
- o   Correlation in Excel
- Lies and Statistics
  - o   Correlation is Useful, but Beware
  - o   Anscombe's Quartet
  - o   The Importance of Data Visualization
- Visualizing the Iris Dataset
- Hands-on Lab #2

Module 5 – Simple Linear Regression
- Beating the Mean
  - o   The Mean as a Baseline Predictive Model
  - o   Linear Regression vs the Mean for Better Predictions
- Simple Linear Regression in Excel
  - o   The Excel Analysis ToolPak Output
  - o   Interpreting ToolPak Output as an Equation
- Understanding the Model
  - o   Uncertainty in Linear Regression
  - o   Confidence Intervals
  - o   Explanatory Power
  - o   Correlation and Simple Linear Regression
- Hands-on Lab #3

Module 6 – Multiple Linear Regression
- When Simple Won't Do
- Categorical Data
  - o   Dummy Encoding
  - o   Interpreting Categorical Features
- The Power of Multiple Features
  - o   Multiple Features and Predictions
  - o   Better Predictions?
- The Rewards of Complexity
  - o   You Need More Complexity
  - o   Beware Complexity
- Interaction Effects

- o   Multiple Features – Chocolate and Peanut Butter
- o   Interaction Modeling
- o   Our "Final Model"
- Hands-on Lab #4

Module 7 – Is Your Model Any Good?

- Linear Regression Data Assumptions
    - o   Ordinary Least Squares (OLS) Linear Regression
    - o   OLS Makes Data Assumptions
    - o   The Six Data Assumptions
- Assumption 1 – Models Are Fully Specified
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Assumption 2 – Features Uncorrelated with Errors
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Assumption 3 – Errors Uncorrelated
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Assumption 4 – Errors Have Constant Variance
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Assumption 5 – No Multicollinearity
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Assumption 6 – Errors Normally Distributed
    - o   Intuition
    - o   Data Example
    - o   How to Validate Assumption
- Hands-on Lab #5

Module 8 – What We Did Not Cover

- Modeling Curvature
- Transforming Data
- Extrapolation
- Standardizing Data
- Models Without Intercepts

Module 9 – Additional Resources