<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
</table>
| 1      | 1    | Aug. 19 | 1. Data Organization and Ctrl Tricks  
2. Using Excel as a Calculator | Knowledge of data structure and proper recording; Knowledge of spreadsheet applications; Ability to use shortcut keys; Use of Excel as a calculator |
| 1      | 2    | Aug. 26 | 1. Describing Data  
2. What All Can the Ribbon Do? | Freezing panes; Simple functions (average, median, min, max, var, function button); Making plots; Familiarity with Excel Ribbon and its features |
| 1      | 3    | Sept. 2 | Sorting, Filtering, Cleaning, and Formatting | Sort and filter; Cleaning through column and row removal; Formatting times and dates; Cleaning individual data points |
| 1      | 4    | Sept. 9 | 1. Data Manipulation Hacks  
2. Marginal vs. Conditional Stats, and How to Calculate and Visualize That Data in Excel | Conditional formulas and calculations, the $ anchor, VLOOKUP, Other data manipulation tricks, Using data cleaning and manipulation to make graphics and draw inferences |

**End of Course 1: Introduction to Data Organization, Preparation and Analysis**  
**Beginning of Course 2: Introduction to Data Analytics**

<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
</table>
| 2      | 1    | Sept. 16 | 1. The Sea of Data Analytics and Data Mining Terminology and the Data Science Pathway Across It  
2. Types of Data  
3. Subsetting and Conditionality | Ability to discuss data science concepts using industry standard terminology; Ability to identify and describe different data classes; Knowledge of the limitations different variable classes impose on functions; Can use Excel interface to filter and subset data; Can calculate conditional statistics using basic filtering and subsetting techniques |
| 2      | 2    | Sept. 23 | 1. Marginality vs. Conditionality  
2. Tables, Pivot Tables, and Visualizing Conditionality | Can discuss the difference between marginal versus conditional statistics; Can calculate both marginal and statistical descriptive statistics; Can construct basic tables in Excel and copy to a Word document for formatting; Can construct Pivot tables in Excel; Use of graphics in Excel; Understanding of different procedures for displaying conditional differences (or lack of differences) between groups |
| 2      | 3    | Sept. 30 | 1. You mean there’s more than one type of mean? Calculating measurements of central tendency.  
2. Calculating measurements of dispersion.  
3. Advanced Excel graphics | Can discuss when it is appropriate to use different measurements of central tendency (e.g. mean, median, mode), as well as different types of means (e.g. arithmetic, harmonic, and geometric means); Can calculate common measures of central tendency and dispersion in Excel; Advanced Excel graphics |
<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
</table>
| 2      | 4    | Oct. 7 | 1. Hypothesis Testing
                2. t-Tests            | Can correctly construct null and alternative hypotheses; Can identify events and the complement of events; Understands the difference between a p-value and a Type I error ($\alpha$); Can conduct basic one population and paired t-tests; Correctly uses Excel functions to calculate test statistics, critical values, and p-values for one-tailed and two-tailed t-tests. |

**End of Course 2: Intro to Data Analytics**

**Beginning of Course 3: Data Analytics and Its Applications**

<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>Oct. 14</td>
<td>Statistical Distributions and Important Areas Under the Curve</td>
<td>Can define a statistical distribution and give examples; Can describe the binomial and normal distributions and give real-world examples of when to use each; Can define a p-value and Type I and Type II errors and draw each of these on a curve; Calculates p-value from binomial and normal examples using Excel.</td>
</tr>
</tbody>
</table>
| 3      | 2    | Oct. 21| 1. Different Descriptive Statistics and Visualizations for Different Types of Variables
                2. Tests for Binomial and Categorical Variables | Knowledge of when to calculate different types of descriptive statistics (e.g. mean vs. median vs. percentage); Can calculate various descriptive statistics in Excel; Creates appropriate graphics based on variable type and question at hand; Uses inferential statistical methods to draw conclusions based on categorical or binomial data; Ability to work with binomial and chi-square statistics. |
| 3      | 3    | Oct. 28| 1. Sampling Methods and the CLT
                2. Tests of 1 or 2 Means         | Understanding of pros and cons of different sampling methods; Use of different sampling methods to construct unbiased surveys; Application of the Central Limit Theorem; Conducts hypothesis tests involving means, including correct identification of degrees of freedom. |
| 3      | 4    | Nov. 4 | 1. Confidence Intervals
                2. Tests of Variance and Why They Are Important | Ability to conduct chi-square and Folded-F tests of variance; Construction of confidence intervals. |
| 3      | 5    | Nov. 11| Regression, Diagnostics, and Time Series Forecasting                | Simple regression; Polynomial/curvilinear regression; Linear modeling; Plots regression curves; Calculates predicted values using regression; Uses regression diagnostics to assess fit and predictability; Time series analysis; Forecasting |
| 3      | 6    | Nov. 18| 1. One-Way ANOVA and Diagnostics
                2. Factorials and Slices         | Knowledge of general linear model construction; Use of one-way ANOVA, factorials, etc. to analyze real-world experiments; Use of diagnostics to assess model fit; Use of advanced linear model techniques to interpret complex model interactions |
|        |      | Nov. 25| No Class – Happy Thanksgiving!                                     |                                                                                                                                                   |

**AY 2020-2021 Schedule**
<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>Dec. 2</td>
<td>1. RCBDs</td>
<td>Advanced experimental designs for dealing with non-independence; RCBD; Split-Plot; Knowledge of fixed vs. random factors.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Split-Plots</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>Dec. 9</td>
<td>Mixed and Generalized Linear Models</td>
<td>Advanced experimental modeling; Knowledge of fixed vs. random factors; Ability to distinguish between general and generalized linear models; Use of generalized linear models to analyze real-world data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>End of Course 3: Data Analytics and Its Applications</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Happy Holidays!!! Break from Dec. 10 to Jan. 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Beginning of Course 4: Introduction to R</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Jan. 13</td>
<td>Introduction to R</td>
<td>Basic knowledge of R syntax and workspace; Knowledge of different variable types (e.g. numeric, integer, character, logic, etc.) and the implications of different variable types when attempting to complete calculations in R.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R as a Calculator</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Jan. 20</td>
<td>Handling Different Types of Variables Lists and Data Frames</td>
<td>Knowledge of different variable types (e.g. numeric, integer, character, logic, etc.) and the implications of different variable types when attempting to complete calculations in R; Working knowledge of computational lists and data frames as well as how to utilize data contained in these objects.</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>Jan. 27</td>
<td>Dates, Times, Environments, Functions, and Packages</td>
<td>Ability to work with dates and times in R; Knowledge of different environments, including the working environment, in R; Ability to install and call packages in R; Ability to implement standard and intermediate functions in R; Ability to create basic novel functions in R.</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>Feb. 3</td>
<td>Flow Control, Looping, and plyr</td>
<td>Ability to implement set theory and conditional statements to complete coding tasks efficiently; Looping and flow control in coding; Experience working with the plyr package and data organization; Intermediate R coding experience.</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Feb. 10</td>
<td>Getting Data and Working with Your Own Stuff Syntax and Data Tidying Recap</td>
<td>Working with external datasets in R; Fluency in R language syntax; Data management, scrubbing, reorganization, merging, and tidying in R.</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>Feb. 17</td>
<td>Data Visualization</td>
<td>Data visualization in R, including basic graphics and ggplot2 graphics.</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>Feb. 24</td>
<td>Intro to Statistics and Data Analytics in R</td>
<td>Ability to conduct common statistical analyses and other data analyses in R.</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>Mar. 3</td>
<td>Random Number Generators and Basic Diagnostics</td>
<td>Can create vectors of randomly generated numbers; Use of random generated numbers to create training and validation datasets for data analysis and model validation; Use of random generated numbers to create test datasets; Ability to conduct basic model and programming diagnostics in R.</td>
</tr>
</tbody>
</table>
### End of Course 4: Introduction to R
#### Spring Break!

### Beginning of Course 5: Machine Learning and Predictive Modeling

<table>
<thead>
<tr>
<th>Course</th>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>Mar. 17</td>
<td>Introduction and R Recap</td>
<td>Ability to create simple and multiple linear regression models and output predictions; Ability to conduct logistic regression and interpret results; Creation of training and validation sets; Assessing model fit using fit statistics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Linear and Logistic Regression</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Mar. 24</td>
<td>PCA Nearest Neighbor Algorithms</td>
<td>Principal Component Analysis (PCA) and other dimension reduction techniques; Basic machine learning classifiers; kNN classification.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>Mar. 31</td>
<td>Clustering Methods</td>
<td>Knowledge of different clustering methods, including machine learning methods; Knowledge of advantages and disadvantages of different clustering methods; Ability to generate cluster assignments for real-world datasets; Optimizing cluster number; Imputation methods.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>Apr. 7</td>
<td>Decision Trees/CART Modeling</td>
<td>Creation of Decision Trees and CART models; Validation of Decision Trees and CART models; Prediction based on Decision Trees and CART models.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>Apr. 14</td>
<td>Naïve Bayes Classification</td>
<td>Knowledge and ability to use probabilistic-based machine learning classifiers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Apr. 21</td>
<td>Market Basket Analysis</td>
<td>Ability to use Market Basket Analysis to classify associative patterns and make a prediction based on the co-occurrence of other events.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>Apr. 28</td>
<td>Neural Networks and Support-Vector Machines (SVM) Classifiers</td>
<td>Ability to use complex, “black-box” methods such as neural networks to make classifications and predictions; Image classification with convolutional neural networks (CNN).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>May 5</td>
<td>Evaluating and Improving Model Performance</td>
<td>Machine learning (ML) model evaluation; ML model validation; ML model adaptation and re-evaluation.</td>
</tr>
</tbody>
</table>

**AY 2020-2021 Schedule**