Course	Week	Date	Topic	Skills
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**	
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

AY 2020-2021 Schedule

Course	Week	Date	Topic	Skills
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 (α); 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**				
3	1	Oct. 14	Statistical Distributions and Important Areas Under the Curve	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.
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
NA		Nov. 25	No Class – Happy Thanksgiving!	

AY 2020-2021 Schedule

SOUTHERN ILLINOIS UNIVERSITY
EDWARDSVILLE

Course	Week	Date	Topic	Skills
3	7	Dec. 2	1. RCBDs 2. Split-Plots	Advanced experimental designs for dealing with non-independence; RCBD; Split-Plot; Knowledge of fixed vs. random factors.
3	8	Dec. 9	Mixed and Generalized Linear Models	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.
End of Course 3: Data Analytics and Its Applications **Happy Holidays!!! Break from Dec. 10 to Jan. 12** **Beginning of Course 4: Introduction to R**				
4	1	Jan. 13	Introduction to R R as a Calculator	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.
4	2	Jan. 20	Handling Different Types of Variables Lists and Data Frames	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.
4	3	Jan. 27	Dates, Times, Environments, Functions, and Packages	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.
4	4	Feb. 3	Flow Control, Looping, and plyr	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.
4	5	Feb. 10	Getting Data and Working with Your Own Stuff Syntax and Data Tidying Recap	Working with external datasets in R; Fluency in R language syntax; Data management, scrubbing, reorganization, merging, and tidying in R.
4	6	Feb. 17	Data Visualization	Data visualization in R, including basic graphics and ggplot2 graphics.
4	7	Feb. 24	Intro to Statistics and Data Analytics in R	Ability to conduct common statistical analyses and other data analyses in R.
4	8	Mar. 3	Random Number Generators and Basic Diagnostics	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.

AY 2020-2021 Schedule

SOUTHERN ILLINOIS UNIVERSITY
EDWARDSVILLE

Course	Week	Date	Topic	Skills
			End of Course 4: Introduction to R **Spring Break!** **Beginning of Course 5: Machine Learning and Predictive Modeling**	
5	1	Mar. 17	Introduction and R Recap Linear and Logistic Regression	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.
5	2	Mar. 24	PCA Nearest Neighbor Algorithms	Principal Component Analysis (PCA) and other dimension reduction techniques; Basic machine learning classifiers; kNN classification.
5	3	Mar. 31	Clustering Methods	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.
5	4	Apr. 7	Decision Trees/CART Modeling	Creation of Decision Trees and CART models; Validation of Decision Trees and CART models; Prediction based on Decision Trees and CART models.
5	5	Apr. 14	Naïve Bayes Classification	Knowledge of and ability to use probabilistic-based machine learning classifiers.
5	6	Apr. 21	Market Basket Analysis	Ability to use Market Basket Analysis to classify associative patterns and make a prediction based on the co-occurrence of other events.
5	7	Apr. 28	Neural Networks and Support-Vector Machines (SVM) Classifiers	Ability to use complex, "black-box" methods such as neural networks to make classifications and predictions; Image classification with convolutional neural networks (CNN).
5	8	May 5	Evaluating and Improving Model Performance	Machine learning (ML) model evaluation; ML model validation; ML model adaptation and re-evaluation.

SOUTHERN ILLINOIS UNIVERSITY
EDWARDSVILLE