FULL STACK DATA SCIENCE

TOPIC

1.1

An Overview of Analytics and Data Science

This class will familiarize you with how analytics and data science help companies and businesses create value by using insights generated by data science tools and techniques

Video Lectures

9

Viewing time: 1:46 hours

Assignments

No Assignments

Learning Outcomes

- ✓ An understanding of analytics and data mining concepts
- ✓ Familiarity with some common terms in analytics
- ✓ A high level understanding of some popular algorithms and when they are used
- ✓ Examples of the application of analytics and outcomes across multiple functions and industries
- ✓ An overview of different kinds of analytics tools and their popularity

FULL STACK DATA SCIENCE

TOPIC

1.2

Analytics Methodology and Problem Solving Frameworks

In this class, you will learn about the standard methodology used in analytics projects and the key outcomes of each stage in the methodology.

Video Lectures

3

Viewing time: 0:28 hours

Assignments

No Assignments

Learning Outcomes

- ✓ Articulate the analytics methodology framework, and list the steps at each stage
- ✓ Be able to define a specific problem statement for an analytics project given a business situation and issue
- ✓ Be able to identify the most appropriate solution design for the given problem statement
- ✓ Be able to build a project plan for an analytics project with appropriate milestones based on effort estimates
- ✓ Be able to build a resource plan for an analytics project based on effort estimates

FULL STACK DATA SCIENCE

TOPIC

1.3

Models and Algorithms

This class is a short introduction to the idea of "models". You will understand how models are built and specified using analytics algorithms, and review different types of commonly used algorithms and the types of problems they are used to tackle.

Video Lectures

3

Viewing time: 0:35 hours

Assignments

1 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Understand and articulate the concept of a "model" in analytics and how it is used as a decision aid
- ✓ Demonstrate familiarity with common terminology used in analytics and the modelling process
- ✓ Be able to specify popular modelling algorithms and when they are used

FULL STACK DATA SCIENCE

TOPIC

2.1

Descriptive Statistics

This class introduces you to statistics, starting with simple descriptive statistics. You will learn how to describe data with statistics, and how to gain an understanding of patterns and insights in data with descriptive statistics

Video Lectures

8

Viewing time: 1:34 hours

Assignments

1 MCQ : Non Graded

Learning Outcomes

- ✓ Calculate values of descriptive statistics like average, median, standard deviation for data, both manually and using Excel
- ✓ Be able to use the calculated values to summarize information and trends in data
- ✓ Be able to use the calculated values to identify any potential issues with the data relative to business domain based expectations

FULL STACK DATA SCIENCE

TOPIC

2.2

Data Visualization with Excel

This topic reviews the different types of charting options available in Excel, and more importantly, when to use what type of chart. You will also learn about the multitude of formatting and labelling options available that help you generate useful and interesting data visualizations that can help with and extend data analysis

Video Lectures

3

Viewing time: 0:40 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Use Excel to generate multiple types of charts including scatter plots, line charts and column plots
- ✓ Master Excel's charting formatting options to appropriately label and create charts
- ✓ Create complex charts like the Waterfall chart

FULL STACK DATA SCIENCE

TOPIC

2.3

Data Analysis Methods: Business Intelligence and Reporting

In this topic, you will learn reporting and business intelligence functions, and how they are used for everyday essential data analysis and descriptive analytics. You will also learn how to choose the right types of visualizations for reporting

Video Lectures

8

Viewing time: 1:18 hours

Assignments

2 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Articulate a general framework for creating data visualization that is part of data analysis
- ✓ Understand the different types of charts and graphs that can be used as part of both analysis, and reporting of analysis, how to generate them in Excel, and the appropriate use of each type of visualization
- ✓ Explain the difference between data analysis, business intelligence, reporting, and analytics

FULL STACK DATA SCIENCE

TOPIC

3.1

An introduction to R

In this chapter, you will take your first steps with R. You will learn how to assign variables and learn about the basic data types in R. You will also learn how to create various data structures in R and read data from different sources.

Video Lectures

17

Viewing time: 4.61 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Understand and implement different data structures in R
- ✓ Read data from different types of sources

FULL STACK DATA SCIENCE

TOPIC

3.2

Simple Data Processing with R

This class introduces basic data manipulation on tabular data, including sorting, summarizing, and merging

Video Lectures

8

Viewing time: 1:24 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Be able to do manipulate tabular data
- ✓ Be able to subset, reorder and produce group-wise summaries
- ✓ Be able to do simple text manipulation
- ✓ Be able to merge dataframes
- ✓ Be able to use dplyr library
- ✓ Be able to work with date data

FULL STACK DATA SCIENCE

TOPIC

3.3

Data Visualization with R

This class focusses on visualization using both base R and ggplot2(). An important concept of grammar of graphics is introduced.

Video Lectures

5

Viewing time: 01:04 hours

Assignments

1 MCQ : Non Graded
3 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Be able to visualize univariate and bivariate data
- ✓ Be able to produce conditional bivariate plots
- ✓ Understand grammar of graphics and use ggplot2() library
- ✓ Be able to produce simple geo-spatial plots and understand commonly used geo-spatial data storage formats

FULL STACK DATA SCIENCE

TOPIC

4.1

Data Pre-processing – Data Exploration

This topic discusses data wrangling and exploration tasks using R. Sanity checks, missing value imputation and splitting data into test and training components are reviewed and explained in this topic

Video Lectures

4

Viewing time: 1:07 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Understand and articulate the need for a structured exploratory data analysis step at the beginning of any data analytics project
- ✓ Use the EDA framework for exploring the data and identifying any problems with the data
- ✓ Be able to identify if there are missing data issues with the dataset
- ✓ Be able to identify if there are any issues related to outliers in the dataset
- ✓ Visualize data trends and patterns that could be relevant to the analysis

FULL STACK DATA SCIENCE

TOPIC

4.2

Data Pre-processing – Data Preparation

This topic reviews how to prepare data for a modelling process, including dealing with outliers and missing values, and how to apply appropriate data transformations

Video Lectures

6

Viewing time: 01:18 hours

Assignments

1 MCQ : Non Graded
2 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Deal with missing data and outlier issues that are identified at the end of the data exploration process
- ✓ Use the data exploration output to identify important variables for the analysis and therefore the appropriate method to capture the information contained in those variables
- ✓ Transform variables as appropriate, either from non numeric types to numeric data, or other transformations that will help capture the relationship of these variables on the target variable

FULL STACK DATA SCIENCE

TOPIC

5.1

Inferential Statistics (Probability Distributions)

Inferential statistics deals with the approach of making inferences about a population given data in a sample. This class will cover the concepts that are basic building blocks of inferential statistics algorithms

Video Lectures

10

Viewing time: 1:96 hours

Assignments

1 MCQ : Non Graded

Learning Outcomes

- ✓ Demonstrate an understanding of the difference between samples and populations
- ✓ Be able to build representative samples from a population
- ✓ Calculate simple and conditional probabilities of events given data
- ✓ Articulate the concept of random variables and random outcomes
- ✓ Identify different types of distributions of random variables
- ✓ Calculate probability of random variable outcomes given a distribution, in Excel

FULL STACK DATA SCIENCE

TOPIC

5.2

Hypothesis Testing

The most popular application of probability distributions and predictive models is hypothesis testing. This class covers the framework of hypothesis testing, and reviews the types of hypothesis tests and their appropriate usage

Video Lectures

11

Viewing time: 02:03 hours

Assignments

1 MCQ : Non Graded

Learning Outcomes

- ✓ Frame a hypothesis for testing given a business problem
- ✓ Use the hypothesis testing framework to implement aa hypothesis test
- ✓ Choose an appropriate level of significance
- ✓ Calculate the probability of a random chance outcome under the null hypothesis
- ✓ Use the calculated probability to compare to the significance level and then get to an appropriate conclusion around the hypothesis being tested
- ✓ Know the different types of hypothesis tests and when to use them

FULL STACK DATA SCIENCE

TOPIC

5.3

Advanced hypothesis testing – multiple samples

This class looks at hypothesis testing when dealing with more than two samples, including ANOVA and Chi Square tests. It also reviews errors in hypothesis testing and how to deal with them

Video Lectures

9

Viewing time: 02.29 hours

Assignments

1 MCQ : Non Graded
1 Case Study : Non Graded
2 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Be able to build hypothesis tests when testing across multiple samples
- ✓ Learn how to deal with frequency data in hypothesis tests using Chi Square tests
- ✓ Implement ANOVA tests when the target variable is continuous but the independent variables are discrete
- ✓ Understand errors in hypothesis tests and how to calculate them
- ✓ Be able to decide what levels and types of errors would be appropriate given a business hypothesis test

FULL STACK DATA SCIENCE

TOPIC

6.1

Predictive Analytics -Linear Regression Models(OLS)

This topic introduces Linear Regression models. R implementation of model and assumption checks of the model have been discussed.

Video Lectures

8

Viewing time: 1:52 hours

Assignments

1 MCQ : Non Graded
1 Case Study : Non Graded

Learning Outcomes

- ✓ Identify when to use a linear regression algorithm
- ✓ Articulate how a linear regression model works, using the Ordinary Least Squares algorithm
- ✓ Implement a linear regression model in Excel and in R
- ✓ Know how to evaluate the output of a linear regression model
- ✓ Assess the overall effectiveness of the model, and decide if further iterations are required to improve fit or usability
- ✓ Translate the model output into a set of business insights and recommendations that will directly help solve the business problem originally articulated

FULL STACK DATA SCIENCE

TOPIC

6.2

Predictive Analytics - Logistic Regression Models

This topic introduces a classification algorithm in the form of logistic regression. Model validation and selection in the context of R have been discussed.

Video Lectures

10

Viewing time: 01:61 hours

Assignments

1 MCQ : Non Graded

Learning Outcomes

- ✓ Understand and articulate the logit function that links the log odds ratio of outcome probability to independent variables
- ✓ Know how to implement a binary logistic regression in R
- ✓ Know how to interpret and evaluate the output of a logistic regression
- ✓ Decide when to finalize a model based on specific evaluation criteria
- ✓ Assess model performance using measures like ROC and gain charts
- ✓ Translate the model output into actions and insights relevant to the business

FULL STACK DATA SCIENCE

TOPIC

6.3

Time Series Forecasting

In this topic you will learn about what is time series analysis and when should it be used. You will also learn about the following topics: 1. Components of Time Series analysis, 2. Various smoothening techniques and how to use them in R

Video Lectures

8

Viewing time: 1:23 hours

Assignments

2 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Know about the components of Time Series analysis
- ✓ Implement various smoothening techniques R
- ✓ Build accurate forecasts using TS techniques

FULL STACK DATA SCIENCE

TOPIC

6.4

Machine Learning - Clustering Models

This chapter introduces a machine learning algorithm-clustering. Cluster profiling along with a brief on hierarchical clustering has been discussed.

Video Lectures

13

Viewing time: 2:19 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Articulate how clustering algorithms work
- ✓ Explain the difference between the K Means algorithm and the Hierarchical clustering algorithms, and know when to use each
- ✓ Know how to prepare data for clustering including standardizing and weighting
- ✓ Know how to evaluate clustering output, and how to profile clusters
- ✓ Be able to assess the optimal number of clusters
- ✓ Be able to use cluster profiles to generate insights and recommendations to business

FULL STACK DATA SCIENCE

TOPIC

6.5

Machine Learning - Decision Trees

Classification and Regression trees are discussed in this class

Video Lectures

20

Viewing time: 1:49 hours

Assignments

1 MCQ : Non Graded

Learning Outcomes

- ✓ Articulate how decision tree algorithms work
- ✓ Understand and calculate measures used to generate splits including Chi Sq, Information Gain and Gini
- ✓ Know how to build a decision tree
- ✓ Know when to stop building the tree and when to prune output
- ✓ Be able to translate decision tree output into business rules and insights

FULL STACK DATA SCIENCE

TOPIC

7

Python for Machine Learning

This section introduces the students to Python statistical computing environment, including PANDAS and MatPlotLib

Video Lectures

21

Viewing time: 02:40 hours

Assignments

1 MCQ : Non Graded

1 MCQ : Smart Certification Test

Learning Outcomes

- ✓ An understanding of analytics and data mining concepts
- ✓ Familiarity with some common terms in analytics
- ✓ Familiarity with pandas, numpy, matplotlib
- ✓ An understanding of basic python data structures such as lists, tuples and strings

FULL STACK DATA SCIENCE

TOPIC

8

Machine Learning with Python

This topic reviews how machine learning algorithms work, and how they are different from statistical algorithms. ML algorithms covering both supervised and unsupervised learning are covered with multiple case studies

Video Lectures

42

Viewing time: 5:40 hours

Assignments

2 MCQ : Non Graded

1 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Understand how machine learning algorithms work, and how they are evaluated
- ✓ Know when and how to use supervised learning algorithms like regression, tree models or NN
- ✓ Know when and how to use unsupervised learning methods like segmentation models
- ✓ Be able to work with structured, semi-structured, and unstructured data types

FULL STACK DATA SCIENCE

TOPIC

9

Capstone Project

This topic covers an end to end case that will start with an understanding of the business problem, through the EDA and descriptive analytics process, and finally to the modelling and insights process

Video Lectures

0

Viewing time: 00:00 hours

Assignments

2 MCQ : Smart Certification Test

Learning Outcomes

- ✓ Demonstrate understanding of the analytics methodology and the steps in each stage of the process
- ✓ Demonstrate the ability to define problem statements and identify the right solutions for each
- ✓ Be able to build and implement predictive and machine learning models in tools like R, Python and SAS
- ✓ Be able to articulate and explain modelling outcomes to a non-technical business audience

FULL STACK DATA SCIENCE

TOPIC

10.1

Principles of Storytelling

To understand how to build engaging and compelling data stories using visualization design principles and best practices

Video Lectures

13

Viewing time: 02:36 hours

Assignments

No Assignment

Learning Outcomes

- ✓ How to build and design business data stories.
- ✓ Learn which charts to use for different types of data and what relationships are best conveyed using a chart type.
- ✓ Understand how visual perception and cognitive Design Principles can help create data presentations that are intuitive,clutter free and effective.
- ✓ Learn different aspects and meanings of colors and how they can be used to communicate your insights more effectively.

FULL STACK DATA SCIENCE

TOPIC

10.2

Visualization with Tableau

This will help you understand how to use tableau to build interactive and compelling visual stories.

Video Lectures

29

Viewing time: 04:31 hours

Assignments

1 MCQ : Non Graded
1 Case Study : Non Graded
1 MCQ : Smart Certification Test

Learning Outcomes

- ✓ How to connect to various data sources
- ✓ Create various kinds of charts and mapping views using the visualization types
- ✓ Learn how we can create calculated fields, table calculations and understand how to leverage the visual analytics features within tableau to explore and understand your data in an easier manner.
- ✓ How to create dashboards and stories to present results of your data analysis

FULL STACK DATA SCIENCE

TOPIC

11.1

An Introduction to the SAS language

This class covers an introduction to the language of SAS, as well as a basic understanding of how SAS works. The class introduces two essential constructs in the SAS language – DATASET and PROC. It also shows how to get data into SAS from external sources and formats using input and import codes.

Video Lectures

3

Viewing time: 0:37 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Demonstrate familiarity with the SAS windows – editor, log, list, explorer
- ✓ Be able to write SAS code for reading data from external file of the type csv as well as using a pre defined procedure
- ✓ Understand how to read the SAS log and identify if the code has executed properly or not
- ✓ Know how to look at data in SAS datasets
- ✓ Know how to locate and create temporary and permanent SAS datasets
- ✓ Know how to use simple programming concepts to subset data in SAS and to create new variables in SAS

FULL STACK DATA SCIENCE

TOPIC

11.2

Data Import into SAS

There are two types of data manipulation in SAS: across columns, for which SAS functions are used, and across rows, for which SAS procedures are used. This class focuses on functions, and reviews the multiple types of functions available in SAS. This class has extensive coding examples

Video Lectures

8

Viewing time: 1:21 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Understand how SAS processes data
- ✓ When to use functions and procedures
- ✓ Identify the right function to use based on the data types
- ✓ Build SAS code and use appropriate syntax for the function chosen
- ✓ Know how to check the end result to make sure that the output is as required
- ✓ Know how to lookup SAS help for additional help and for information on functions that you are not familiar with

FULL STACK DATA SCIENCE

TOPIC

11.3

Data Manipulation with SAS

SAS procedures are the workhorse of SAS data processing. This class covers the different types of procedures available for descriptive statistics, for summarization and aggregation, for visualization and advanced data processing. This class has extensive coding examples.

Video Lectures

8

Viewing time: 1:22 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Understand how SAS processes data for SAS procedures
- ✓ Be familiar with the syntax and options for many widely used SAS procedures
- ✓ Identify the right procedure to use based on output required
- ✓ Build SAS code and use appropriate syntax for the procedure chosen
- ✓ Know how to check the end result to make sure that the output is as required
- ✓ Know how to lookup SAS help for additional help and for information on procedures that you are not familiar with

FULL STACK DATA SCIENCE

TOPIC

11.4

Advanced Data Manipulation with SAS

This class covers how to combine data from different sources or data containing different sets of variables in SAS. There are two ways of joining datasets – appending, and merging. This class has extensive coding examples

Video Lectures

6

Viewing time: 1:17 hours

Assignments

1 MCQ : Non Graded

1 Case Study : Non Graded

Learning Outcomes

- ✓ Understand why you may need to combine datasets in SAS
- ✓ Know the difference between vertical combining and horizontal combining in SAS, and when to use each approach
- ✓ Know the syntax and application of appending data using the SET statement and the PROC APPEND procedure
- ✓ Know how to merge data using the merge statement
- ✓ Know how to control the output of the merge statement by applying conditions