

## **Data Science Syllabus**

### **Data Science with Python**

#### **Module 1: Introduction to Data Science**

- Selecting rows/observations
- Rounding Number
- Selecting columns/fields
- Merging data
- Data aggregation
- Data munging techniques

#### **Hands-on-Exercise:**



- As you gain experience with CSV and JSON files, you'll be able to process almost any data you want to analyze.
- Most online data sets can be downloaded in either or both of these. From working with these formats, you'll be able to learn other data formats as well.

#### **Module 2: Introduction to Python**

- What is Python?
- Why Python?
- Installing Python
- Python IDEs
- Jupyter Notebook Overview

#### **Hands-on-Exercise:**

- Installing Python idle for windows, Linux and
- Creating "Hello World" code

## Module 3: Python Basics

- Python Basic Data types
- Lists
- Slicing
- IF statements
- Loops
- Dictionaries
- Tuples
- Functions
- Array
- Selection by position & Labels

### Hands-on-Exercise-Constructing Operators

- Practice and Quickly learn Python necessary skills by solving simple questions and problems.
- how Python uses indentation to structure a program, and how to avoid some common indentation errors.
- You executed to make simple numerical lists, as well as a few operations you can perform on numerical lists, tuple, dictionary and set



## Module 4: Python Packages

- Pandas
- Numpy
- Sci-kit Learn
- Mat-plot library

### Hands-on-Exercise:

- Installing jupyter notebook for windows, Linux and
- Installing numpy, pandas and matplotlib

## Module 5: Importing Data

- Reading CSV files
- Saving in Python data
- Loading Python data objects
- Writing data to CSV file

### **Hands-on-Exercise:**

- To generate data sets and create visualizations of that data. You learned to create simple plots with matplotlib, and you saw how to use a scatter plot to explore random
- You learned to create a histogram with Pygal and how to use a histogram to explore the results of rolling dice of different
- Generating your own data sets with code is an interesting and powerful way to model and explore a wide variety of real-world
- As you continue to work through the data visualization projects that follow, keep an eye out for situations you might be able to model with

### **Module 6: Manipulating Data**

- Selecting rows/observations
- Rounding Number
- Selecting columns/fields
- Merging data
- Data aggregation
- Data munging techniques



### **Hands-on-Exercise:**

- As you gain experience with CSV and JSON files, you'll be able to process almost any data you want to analyze.
- Most online data sets can be downloaded in either or both of these From working with these formats, you'll be able to learn other data formats as well.

### **Module 7: Statistics Basics**

- Central Tendency
  - Mean
  - Median
  - Mode
  - Skewness
  - Normal Distribution
- Probability Basics

- What does it mean by probability?
- Types of Probability
- ODDS Ratio?
- Standard Deviation
- Data deviation & distribution
- Variance
- Bias variance Tradeoff
- Underfitting
- Overfitting
- Distance metrics
- Euclidean Distance
- Manhattan Distance
- Outlier analysis
- What is an Outlier?
- Inter Quartile Range
- Box & whisker plot
- Upper Whisker
- Lower Whisker
- Scatter plot
- Cook's Distance
- Missing Value treatment
- What is NA?
- Central Imputation
- KNN imputation
- Dummification
- Correlation
- Pearson correlation
- positive & Negative correlation



### **Hands-on-Exercise:**

- Compute probability in a situation where there are equally-likely outcomes

- Apply concepts to cards and dice
- Compute the probability of two independent events both occurring
- Compute the probability of either of two independent events occurring
- Do problems that involve conditional probabilities
- Calculate the probability of two independent events occurring
- List all permutations and combinations
- Apply formulas for permutations and combinations

## **Module 8: Error Metrics**

- Classification
  - Confusion Matrix
  - Precision
  - Recall
  - Specificity
  - F1 Score
- Regression
  - MSE
  - RMSE
  - MAPE



### **Hands-on-Exercise:**

- State why the  $z'$  transformation is necessary
- Compute the standard error of  $z$
- Compute a confidence interval on  $p$  The computation of a confidence interval
- Estimate the population proportion from sample proportions
- Apply the correction for continuity

## **Machine Learning**

### **Module 1: Supervised Learning**

- Linear Regression
  - Linear Equation
  - Slope

- Intercept
- R square value
- Logistic regression
- ODDS ratio
- Probability of success
- Probability of failure Bias Variance Tradeoff
- ROC curve
- Bias Variance Tradeoff

### **Hands-on-Exercise:**

- we've reviewed the main ways to approach the problem of modeling data using simple and definite

## **Module 2: Unsupervised Learning**

- K-Means
- K-Means ++
- Hierarchical Clustering

## **Module 3: SVM**



- Support Vectors
- Hyperplanes
- 2-D Case
- Linear Hyperplane

## **Module 4: SVM Kernel**

- Linear
- Radial
- polynomial

## **Module 5: Other Machine Learning Algorithms**

- K – Nearest Neighbour
- Naïve Bayes Classifier
- Decision Tree – CART
- Decision Tree – C50
- Random Forest

## **Hands-on-Exercise:**

- We have covered the simplest but still very practical machine learning models in an eminently practical way to get us started on the complexity
- where we will cover several regression techniques, it will be time to go and solve a new type of problem that we have not worked on, even if it's possible to solve the problem with clustering methods (regression), using new mathematical tools for approximating unknown values.
- In it, we will model past data using mathematical functions, and try to model new output based on those modeling

## **Artificial Intelligence**

### **Module 1: AI Introduction**

- Perceptron
- Multi-Layer perceptron
- Markov Decision Process
- Logical Agent & First Order Logic
- AL Applications



## **Deep Learning**

### **Module 1: Deep Learning Algorithms**

- CNN – Convolutional Neural Network
- RNN – Recurrent Neural Network
- ANN – Artificial Neural Network

## **Hands-on-Exercise:**

- We took a very important step towards solving complex problems together by means of implementing our first neural
- Now, the following architectures will have familiar elements, and we will be able to extrapolate the knowledge acquired on this chapter, to novel

### **Module 2: Introduction to NLP**

- Text Pre-processing
- Noise Removal
- Lexicon Normalization

- Lemmatization
- Stemming
- Object Standardization

### **Module 3: Text to Features(Feature Engineering)**

- Syntactical Parsing
- Dependency Grammar
- Part of Speech Tagging
- Entity Parsing
- Named Entity Recognition
- Topic Modelling
- N-Grams
- TF – IDF
- Frequency / Density Features
- Word Embedding's

### **Module 4: Tasks of NLP**



- Text Classification
- Text Matching
- Levenshtein Distance
- Phonetic Matching
- Flexible String Matching

### **Hands-on-Exercise:**

- provided, you will even be able to create new customized
- As our models won't be enough to solve very complex problems, in the following chapter, our scope will expand even more, adding the important dimension of time to the set of elements included in our generalization.

## **Tableau**

### **Module 1: Tableau Course Material**

- Start Page
- Show Me



- Connecting to Excel Files
- Connecting to Text Files
- Connect to Microsoft SQL Server
- Connecting to Microsoft Analysis Services
- Creating and Removing Hierarchies
- Bins
- Joining Tables
- Data Blending

## **Module 2: Learn Tableau Basic Reports**

- Parameters
- Grouping Example 1
- Grouping Example 2
- Edit Groups
- Set
- Combined Sets
- Creating a First Report
- Data Labels
- Create Folders
- Sorting Data
- Add Totals, Subtotals and Grand Totals to Report



### **Hands-on-Exercise:**

- Install Tableau Desktop
- Connect Tableau to various Datasets: Excel and CSV files

## **Module 3: Learn Tableau Charts**

- Area Chart
- Bar Chart
- Box Plot
- Bubble Chart
- Bump Chart

- Bullet Graph
- Circle Views
- Dual Combination Chart
- Dual Lines Chart
- Funnel Chart
- Traditional Funnel Charts
- Gantt Chart
- Grouped Bar or Side by Side Bars Chart
- Heatmap
- Highlight Table
- Histogram
- Cumulative Histogram
- Line Chart
- Lollipop Chart
- Pareto Chart
- Pie Chart
- Scatter Plot
- Stacked Bar Chart
- Text Label
- Tree Map
- Word Cloud
- Waterfall Chart



### **Hands-on-Exercise:**

- Create and use Static Sets
- Create and use Dynamic Sets
- Combine Sets into more Sets
- Use Sets as filters
- Create Sets via Formulas
- Control Sets with Parameters
- Control Reference Lines with Parameters

## Module 4: Learn Tableau Advanced Reports

- Dual Axis Reports
- Blended Axis
- Individual Axis
- Add Reference Lines
- Reference Bands
- Reference Distributions
- Basic Maps
- Symbol Map
- Use Google Maps
- Mapbox Maps as a Background Map
- WMS Server Map as a Background Map

### Hands-on-Exercise:

- Create Barcharts
- Create Area Charts
- Create Maps
- Create Interactive Dashboards
- Create Storylines
- Understand Types of Joins and how they work
- Work with Data Blending in Tableau
- Create Table Calculations
- Work with Parameters
- Create Dual Axis Charts
- Create Calculated Fields



## Module 5: Learn Tableau Calculations & Filters

- Calculated Fields
- Basic Approach to Calculate Rank
- Advanced Approach to Calculate Ra
- Calculating Running Total
- Filters Introduction

- Quick Filters
- Filters on Dimensions
- Conditional Filters
- Top and Bottom Filters
- Filters on Measures
- Context Filters
- Slicing Filters
- Data Source Filters
- Extract Filters

### **Hands-on-Exercise:**

- Creating Data Extracts in Tableau
- Understand Aggregation, Granularity, and Level of Detail
- Adding Filters and Quick Filters

### **Module 6: Learn Tableau Dashboards**

- Create a Dashboard
- Format Dashboard Layout
- Create a Device Preview of a Dashboard
- Create Filters on Dashboard
- Dashboard Objects
- Create a Story



### **Module 7: Server**

- Tableau online.
- Overview of Tableau
- Publishing Tableau objects and scheduling/subscription.

### **Hands-on-Exercise:**

- Create Data Hierarchies
- Adding Actions to Dashboards (filters & highlighting)
- Assigning Geographical Roles to Data Elements
- Advanced Data Preparation

# SQL

## Introduction to Database

- List the features of Oracle Database 11g
- Discuss the basic design, theoretical, and physical aspects of a relational database
- Categorize the different types of SQL statements
- Describe the data set used by the course
- Log on to the database using SQL Developer environment
- Save queries to files and use script files in SQL Developer

### Hands-on-Exercise:

- Prepare your environment
- Work with Oracle database tools
- Understand and work with language features

## Retrieve Data using the SQL SELECT Statement

- List the capabilities of SQL SELECT statements
- Generate a report of data from the output of a basic SQL statement
- Select All Columns
- Select Specific Columns
- Use Column Heading Defaults
- Use Arithmetic Operators
- Understand Operator Precedence
- Learn the DESCRIBE command to display the table structure



### Hands-on-Exercise

- Individual statements in SQL scripts are commonly terminated by a line break (or carriage return) and a forward slash on the next line, instead of a semicolon.
- You can create a SELECT statement, terminate it with a line break, include a forward slash to execute the statement, and save it in a script file.

## Learn to Restrict and Sort Data

- Write queries that contain a WHERE clause to limit the output retrieved
- List the comparison operators and logical operators that are used in a WHERE clause

- Describe the rules of precedence for comparison and logical operators
- Use character string literals in the WHERE clause
- Write queries that contain an ORDER BY clause to sort the output of a SELECT statement
- Sort output in descending and ascending order

### **Hands-on-Exercise:**

- Creating the queries in a compound query must return the same number of columns.
- Create corresponding columns in each query must be of compatible data types.
- ORDER BY; it is, however, permissible to place a single ORDER BY clause at the end of the compound query

### **Usage of Single-Row Functions to Customize Output**

- Describe the differences between single row and multiple row functions
- Manipulate strings with character function in the SELECT and WHERE clauses
- Manipulate numbers with the ROUND, TRUNC, and MOD functions
- Perform arithmetic with date data
- Manipulate dates with the DATE functions



### **Hands-on-Exercise:**

- Create the distinction is made between single- row functions, which execute once for each
- row in a dataset, and multiple-row functions, which execute once for all the rows in a data- set.

### **Invoke Conversion Functions and Conditional Expressions**

- Describe implicit and explicit data type conversion
- Use the TO\_CHAR, TO\_NUMBER, and TO\_DATE conversion functions
- Nest multiple functions
- Apply the NVL, NULLIF, and COALESCE functions to data
- Use conditional IF THEN ELSE logic in a SELECT

### **Hands-on-Exercise:**

- we create and discuss the NVL function, which provides a mechanism to convert null values into more arithmetic-friendly data values.

## Aggregate Data Using the Group Functions

- Use the aggregation functions in SELECT statements to produce meaningful reports
- Divide the data into groups by using the GROUP BY clause
- Exclude groups of data by using the HAVING clause

### Hands-on-Exercise:

- Group functions operate on aggregated data and return a single result per group.
- These groups usually consist of zero or more rows of data.

## Display Data from Multiple Tables Using Joins

- Write SELECT statements to access data from more than one table
- View data that generally does not meet a join condition by using outer joins
- Join a table by using a self-join

## Use Subqueries to Solve Queries

- Describe the types of problem that subqueries can solve
- Define sub-queries
- List the types of sub-queries



### Hands-on-Exercise:

- Write a query that uses subqueries in the column projection list.
- Write single-row and multiple-row subqueries

## The SET Operators

- Describe the SET operators
- Use a SET operator to combine multiple queries into a single query
- Control the order of rows returned

### Hands-on-exercise:

- Create The queries in the compound query must return the same number of columns.
- creating The corresponding columns must be of compatible data type.
- creating The set operators have equal precedence and will be applied in the order they are specified.

## Data Manipulation Statements

- Describe each DML statement

- Insert rows into a table
- Change rows in a table by the UPDATE statement
- Delete rows from a table with the DELETE statement
- Save and discard changes with the COMMIT and ROLLBACK statements
- Explain read consistency

### **Hands-on-exercise:**

- Expressions and create expose a vista of data manipulation possibilities through the interaction of arithmetic and character operators with column or literal data, or a combination of the two.

### **Use of DDL Statements to Create and Manage Tables**

- Categorize the main database objects
- Review the table structure
- List the data types available for columns
- Create a simple table
- Decipher how constraints can be created at table creation
- Describe how schema objects work



### **Other Schema Objects**

- Create a simple and complex view
- Retrieve data from views
- Create, maintain, and use sequences
- Create and maintain indexes
- Create private and public synonyms

### **Control User Access**

- Differentiate system privileges from object privileges
- Create Users
- Grant System Privileges
- Create and Grant Privileges to a Role
- Change Your Password
- Grant Object Privileges
- How to pass on privileges?



- Revoke Object Privileges

### **Hands-on-exercise:**

- create users and execute the

## **Management of Schema Objects**

- Add, Modify and Drop a Column
- Add, Drop and Defer a Constraint
- How to enable and Disable a Constraint?
- Create and Remove Indexes
- Create a Function-Based Index
- Perform Flashback Operations
- Create an External Table by Using ORACLE\_LOADER and by Using ORACLE\_DATAPUMP
- Query External Tables

### **Hands-on-exercise:**

- Create the function based index and types.

## **Manage Objects with Data Dictionary Views**



- Explain the data dictionary
- Use the Dictionary Views
- USER\_OBJECTS and ALL\_OBJECTS Views
- Table and Column Information
- Query the dictionary views for constraint information
- Query the dictionary views for view, sequence, index, and synonym information
- Add a comment to a table
- Query the dictionary views for comment information

## **Manipulate Large Data Sets**

- Use Subqueries to Manipulate Data
- Retrieve Data Using a Subquery as Source
- Insert Using a Subquery as a Target
- Usage of the WITH CHECK OPTION Keyword on DML Statements
- List the types of Multitable INSERT Statements

- Use Multitable INSERT Statements
- Merge rows in a table
- Track Changes in Data over a period of time

## **Data Management in Different Time Zones**

- Time Zones
- CURRENT\_DATE, CURRENT\_TIMESTAMP, and LOCALTIMESTAMP
- Compare Date and Time in a Session's Time Zone
- DBTIMEZONE and SESSIONTIMEZONE
- Difference between DATE and TIMESTAMP
- INTERVAL Data Types
- Use EXTRACT, TZ\_OFFSET, and FROM\_TZ
- Invoke TO\_TIMESTAMP, TO\_YMINTERVAL and TO\_DSINTERVAL

## **Retrieve Data Using Sub-queries**

- Multiple-Column Subqueries
- Pairwise and Non Pairwise Comparison
- Scalar Subquery Expressions
- Solve problems with Correlated Subqueries
- Update and Delete Rows Using Correlated Subqueries
- The EXISTS and NOT EXISTS operators
- Invoke the WITH clause
- The Recursive WITH clause



## **Regular Expression Support**

- Use the Regular Expressions Functions and Conditions in SQL
- Use Meta Characters with Regular Expressions
- Perform a Basic Search using the REGEXP\_LIKE function
- Find patterns using the REGEXP\_INSTR function
- Extract Substrings using the REGEXP\_SUBSTR function
- Replace Patterns Using the REGEXP\_REPLACE function
- Usage of Sub-Expressions with Regular Expression Support

- Implement the REGEXP\_COUNT function

### **Hands-on-exercise:**

- Expressions and create the regular columns may be aliased using the AS keyword or by leaving a space between the column or expression and the alias. In this way, both wildcard symbols can be used as either specialized or regular characters in different segments of the same character string
- 

