

Data science Fundamentals with Python

Duration:35 hrs.

Statistical techniques that a Data Scientist need Know

50 Statistical techniques--

Data Analytic Thinking

Outline the importance of data in making business decision

Describe types of data and data catagories

Describe the people and processes involved in data cycle

Compare the characteristics of small data and big data

Discuss the importance of data analytic for business decision making

Identify actions taken during the cross industry standard process of data mining

Discuss the data science links with other discipline

Data Analysis with Python

Review and delineate the evolution and purpose of Python

Describe and set up Python development environment

Practice coding with basic Python commands, operators and conditional statements

Explore and apply Python data structure concepts such as array, list, tuple, set and dictionary

Import python modules and packages

Import Python libraries such as NumPy, Pandas

Types of Data Analytics

Describe spectrum of business analytics Describe application of descriptive analytics Draw conclusions from a given set of data by using descriptive analytic techniques Describe application of diagnostics analytics Draw conclusions from a given set of data by using diagnostics analytic techniques Describe application of predictive analytics Draw conclusions from a given set of data by using predictive analytic techniques Describe application of prescriptive analytics Draw conclusions from a given set of data by using prescriptive analytic technique

Data Modelling

Discuss Cross Industry Standard Process in Data Modeling Discuss a Generic data modeling process Apply prior knowledge the address the business problems

Data Clustering

Describe concepts of clustering and visualize data

Apply K-means algorithm to cluster the data

Apply Z-score method to standardise the data

Interpret the cluster centre and create product segment

Use Dendrogram and Elbow Curve for estimating the number of clusters

Estimate the quality of clustering using Silhouette score

Hierarchical Clustering

Explain the limitations of K-means clustering

Apply hierarchical clustering to the product segmentation and the Gaussian

Distributed dataset

Describe the DBSCAN clustering technique and its benefits

Apply K-Means, Hierarchical and DBSCAN clustering to the moon dataset

Discuss the limitations of clustering algorithms and techniques to address them