

KOENIG SOLUTIONS · COURSE OUTLINE

Data Analysis and Predictive Modelling Using Python

Learn to analyse, visualise and model data in Python, from NumPy and Pandas through machine learning to deep learning.

7 Days

56 HOURS

12

MODULES

123

TOPICS

COURSE OVERVIEW

This program is well suited for both beginners or experienced developers looking to make the jump to Data Analysis, Data Science or Machine Learning. In this course, you will learn how to work with data, how to analyse data, how to visualise the data using python packages, implementing the predictive analysis concepts, implementing the predictive analysis.

WHO THIS COURSE IS FOR

Data Analysts

Data Scientists

Machine Learning Engineers

Python Developers

Business Intelligence Analysts

PREREQUISITES

- Basic knowledge of Python.

LEARNING OBJECTIVES

- Apply NumPy arrays for data manipulation and linear algebra
- Use Pandas for slicing, filtering and aggregating dataframes
- Build visualisations with Matplotlib, Seaborn and Plotly
- Perform exploratory data analysis and data pre-processing
- Implement supervised regression and classification models
- Apply feature selection and dimensionality reduction techniques
- Evaluate models using cross validation and hyperparameter tuning
- Develop clustering, NLP and deep learning models

12 Modules, 123 Topics

PART ONE

01

NumPy Package

TOPICS

- Introduction of NumPy
- Install Numpy
- Array Creation
- Array Reshaping
- Indexing
- Operation
- Sorting and Stacking
- BroadCasting
- Linear Algebra
- Saving and Loading

02

Pandas

TOPICS

Introduction · Slicing Dataframe · Filtering Dataframe · Transforming Dataframe · Advanced indexing · Stack and unstack · Groupby and aggregations

03

Matplotlib

INTRODUCTION

Introduction to Matplotlib · Exploring Data using Python · Matplotlib with Jupyter · Load data

ARCHITECTURE

How to pyplot Works · Troubleshooting issues

DIFFERENT PLOTS

Line Chart · Multi Line Plot · Fill Plot · Bar Chart · Pie Chart · Histogram · Scatterplot

PLOTTING

Themes · Scatter · Subplot · 3d plot · Grid · Save Image · Legend

CONFIGURE MATPLOTLIB

Style Sheets · Custom Style Sheets

INTERACTIVITY

Handling Event · Create An Interface With ipywidgets

MAPS

Basemap · Creating a Choropleth With Matplotlib

04

Seaborn

INTRODUCTION

- Introduction to Seaborn
- Importing Dataset
- Seaborn Vs Matplotlib
- Using Seaborn with Matplotlib

LOAD DATA

- Loading a built-in Seaborn data
- Loading a Pandas Dataframe

PLOTS

- LinePlot
- Distplot
- BarGraph
- ScatterPlot
- JointPlot (KDE)
- StripPlot
- Box
- Point Plot
- FacetGrid
- Pair Grid
- CatPlot

CUSTOMISING SEABORN PLOTS

- Changing Figure Aesthetic
- Removal of Spines
- Changing the figure Size
- Scaling the plots
- Setting the Style Temporarily

CREATING DIFFERENT TYPES OF PLOTS

- Relational Plots
- Categorical Plots
- Distribution Plots
- Regression Plots

COLOUR PALETTE

- Diverging Colour Palette
- Sequential Colour Palette
- Setting the default Colour Palette

05

Plotly

TOPICS

Installation and Setup · Line Plot · Scatter Plot · Bar Plot · Box Plot and Area Plot · 3D Plot · Spread Plot and Hist Plot · Bubble Plot and Heatmap

06

Python for Data Analytics and Pre Processing

TOPICS

- Exploratory data analysis
- Data Cleaning, handling missing values, handling categorical data
- Hot encoding
- Data transformation
- Data normalization
- Train, Test Data

07

Supervised Machine Learning Regression

TOPICS

- Concept of linear regression
- OLS, regression errors, data processing
- Multi regression, degree of freedom, adjusted R-squared
- Concept of dummy variable

08

Supervised Machine Learning Classification

TOPICS

- Logistic regression, stratification, confusion matrix
- Support vector machine, mathematical intuition, different kernels
- Decision tree
- Random forest classification
- Ensemble learning
- Bagging and Boosting
- Evaluation metrics like accuracy, precision, recall, F1, AUC ROC curve

09

Feature Selection and Dimensionality Reduction

TOPICS

- Feature selection importance
- Concept of univariate feature selection
- F-test, chi square test
- Generic univariate select
- Recursive feature elimination
- Principal component analysis
- Need to reduce dimensions and importance
- Mathematical intuition

10

Cross Validation and Hyperparameter Tuning

TOPICS

- Importance of cross validation
- Parameter and implementation of cross validation
- Concept of hyper parameter tuning
- Grid search and randomized search

11

Unsupervised Machine Learning and NLP

TOPICS

- Introduction to clustering
- Mathematical intuition behind clustering
- K-means clustering
- Elbow method
- Introduction to NLP
- Concept of tokenization, stop words, stemming, lemmatization
- Tf-Idf vector and its mathematical intuition

12

Deep Learning

TOPICS

- Need and application of deep learning
- Working of artificial neural network
- Backend(tensorflow) and frontend(keras)
- Concept of tensors
- Keras building model- construct, compile and evaluate
- Activation function and loss function
- Evaluation metrics for deep learning

NOTES

Highlight of Course

- Basic Python
- Numpy, Pandas
- Matplotlib
- Seaborn
- Plotly and Cufflinks

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training

Live Online Classes (ILO)

Fly-Me-A-Trainer (FMAT)

Flexi (Self-Paced)

1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com

✉ info@koenig-solutions.com