Data Analytics – OA25DA01

Beginner to Expert Course

Course Description

Master data analytics in this all-in-one course, designed to transform you from a beginner to an expert data analyst. Covering data manipulation, visualization, statistical analysis, and advanced techniques like machine learning and big data, this course leverages industry-standard tools such as Python, SQL, and Power BI. Perfect for 2025, it bundles four levels of learning: Beginner, Intermediate, Advanced, and Expert with Projects.

What You'll Learn

- Master data manipulation and visualization with Python and Power BI
- Perform statistical analysis and hypothesis testing
- Query and manage data with SQL
- Apply machine learning for predictive analytics
- Build and deploy data dashboards and pipelines
- Work with big data tools like Spark

Course Requirements

- No prior analytics or programming knowledge required
- A computer with Python 3.x, SQL, and Power BI Desktop installed
- Basic understanding of mathematics recommended

Who This Course Is For

- Beginners starting their journey in data analytics
- Intermediate analysts aiming to master advanced tools and techniques
- Professionals seeking to apply data analytics in business, finance, or tech

Syllabus (35 Sessions)

1. Introduction to Data Analytics

- Overview of data analytics and its applications
- Setting up tools: Python, Anaconda, Power BI, SQL
- Introduction to data types and sources
- Common analytics workflows and terminology

2. Excel for Data Analytics

- Excel interface and shortcuts
- Basic functions: SUM, AVERAGE, COUNT
- Working with tables and filters
- Creating simple charts

3. Introduction to Python for Analytics

- Installing Python and Jupyter Notebook
- Python basics: Variables, lists, dictionaries
- Installing libraries with pip (pandas, numpy)
- Writing simple data scripts

4. Data Manipulation with Pandas

- Introduction to pandas DataFrames
- Importing and exporting data (CSV, Excel)
- Filtering, sorting, and grouping data
- Handling missing values

5. Data Visualization with Python

- Introduction to matplotlib and seaborn
- Creating line, bar, and scatter plots
- Customizing visualizations
- Visualizing trends and patterns

6. SQL Basics

- Introduction to relational databases
- Setting up SQLite
- Writing SELECT, INSERT, UPDATE, DELETE queries
- Filtering and sorting data with WHERE and ORDER BY

7. Power BI Fundamentals

- Introduction to Power BI Desktop
- Connecting to data sources
- Creating basic visualizations (charts, tables)
- Building simple dashboards

8. Data Cleaning

- Identifying and handling missing data
- Removing duplicates and outliers

- Standardizing data formats
- Using Python and Excel for cleaning

9. Practice Session

- Clean and visualize a sample dataset
- Congratulations! You've built a data analytics foundation.

10. Intermediate Pandas

- Merging and joining DataFrames
- Pivot tables and cross-tabulations
- Advanced grouping and aggregation
- Working with time series data

11. Intermediate SQL

- Joins: INNER, LEFT, RIGHT, FULL
- Subqueries and nested queries
- Aggregations with GROUP BY
- Window functions (ROW_NUMBER, RANK)

12. Statistical Foundations

- Descriptive statistics: Mean, median, standard deviation
- Probability basics and distributions
- Correlation and covariance
- Using Python's scipy for stats

13. Data Visualization with Power BI

- Creating interactive dashboards
- Using DAX for calculated fields
- Working with slicers and filters
- Sharing and publishing reports

14. Introduction to APIs for Data

- Fetching data with Python's requests
- Parsing JSON data
- Integrating API data with pandas
- Handling API errors

15. Exploratory Data Analysis (EDA)

- Performing EDA with Python
- Identifying patterns and anomalies
- Visualizing relationships with seaborn
- Documenting insights

16. Practice Session

- Perform EDA on a business dataset
- Congratulations! You're an intermediate data analyst.

17. Advanced SQL

- Common Table Expressions (CTEs)
- Stored procedures and triggers
- Optimizing queries for performance
- Indexing and query execution plans

18. Hypothesis Testing

- Understanding null and alternative hypotheses
- Performing t-tests and chi-square tests
- Using scipy.stats for testing
- Interpreting p-values and significance

19. Advanced Visualization

- Creating complex charts with plotly
- Building heatmaps and pair plots
- Customizing Power BI visuals
- Storytelling with data

20. Introduction to Machine Learning

- Overview of ML with scikit-learn
- Supervised vs. unsupervised learning
- Building a simple regression model
- Evaluating model performance

21. Data Wrangling with Python

- Advanced data cleaning techniques
- Handling large datasets with pandas
- Working with dask for big data
- Automating data pipelines

22. Time Series Analysis

- Working with time series data in pandas
- Forecasting with ARIMA models
- Visualizing trends with plotly
- Applications in finance and sales

23. Version Control for Analytics

- Using Git and GitHub for data projects
- Managing Jupyter notebooks
- Collaborative workflows
- Versioning datasets

24. Practice Session

- Build a sales forecasting model
- Congratulations! You've mastered data analytics.

25. Project 1: Sales Dashboard

- Create an interactive dashboard with Power BI
- Import and clean sales data
- Visualize KPIs and trends
- Share the dashboard online

26. Project 2: Customer Segmentation

- Use pandas and scikit-learn
- Perform clustering with K-Means
- Visualize segments with seaborn
- Document insights

27. Project 3: SQL Data Pipeline

- Build a data pipeline with PostgreSQL
- Create tables and relationships
- Write complex queries for analysis
- Automate with Python

28. Project 4: Predictive Analytics

- Build a classification model with scikit-learn
- Train and test on a dataset

- Evaluate with confusion matrix
- Deploy model with joblib

29. Big Data with Spark

- Introduction to Apache Spark
- Setting up pyspark
- Processing large datasets
- Spark SQL for analytics

30. Project 5: E-Commerce Analytics

- Analyze e-commerce data with pandas
- Create visualizations with plotly
- Build a Power BI dashboard
- Derive business insights

31. Advanced Machine Learning

- Decision trees and random forests
- Hyperparameter tuning with GridSearchCV
- Feature engineering techniques
- Model deployment with Flask

32. Deploying Analytics Solutions

- Deploying Python apps to Heroku
- Hosting dashboards with Power BI Service
- Using Docker for analytics pipelines
- Setting up CI/CD with GitHub Actions

33. Project 6: Real-Time Analytics

- Build a real-time dashboard with Python
- Use APIs for live data
- Visualize with plotly and Power BI
- Deploy to Heroku

34. Big Data and Cloud Analytics

- Introduction to cloud platforms (e.g., AWS, GCP)
- Using AWS S3 for data storage
- Processing data with AWS Glue
- Analytics with Google BigQuery

35. Final Exam and Wrap-Up

- Comprehensive exam on data analytics concepts
- Review of key skills and best practices
- Guidance on data analytics career paths
- Certificate of completion