

TuxAcademy Data Analysis Program

Data Management, Data Processing and Machine Learning

Data Visualization, Ethics and Real Projects



Course Syllabus

Module 1: Data Management

1.1: Introduction to Data Analysis

- Overview of Data Analysis
- Applications of Data Analysis
- Data Science vs. Data Analytics

1.2: Data Collection and Cleaning

- Types of data: structured, unstructured, semi-structured
- Data sources: databases, web scraping, APIs
- Handling missing data
- Data normalization and standardization
- Data transformation techniques

1.3: Data Analysis

- NoSQL Databases
- SQL Databases and Writing Query
- Descriptive Statistics
- Data Visualization
- Identifying patterns and trends
- Correlation analysis

1.4: Excel for Data Analysis

- Navigation and formulas
- Sorting and filtering
- Tables
- Pivot Tables
- Charts
- VLOOKUP and XLOOKUP
- Conditional formatting

1.5: Statistical Analysis

- Probability Theory
- Inferential Statistics
- Hypothesis Testing

Module 2: Data Processing and Machine Learning

2.1: Data Manipulation with Python

- Introduction to Python for Data Analysis
- Data Manipulation with Pandas (Python)
- Data manipulation using cloud tools
- Python libraries for Big data manipulation

2.2: Data Engineering

- Data Warehousing and ETL Processes
- Building Data Pipelines
- Using SQL and NoSQL Databases

2.3: Machine Learning

- Introduction to Machine Learning
- Supervised Learning Algorithms (Linear Regression, Decision Trees, etc.)
- Unsupervised Learning Algorithms (Clustering, PCA, etc.)
- Model Evaluation and Validation

Module 3: Data Visualization and Data Ethics

3.1: Tools for Data Visualization and Presentation

- Principles of Effective Data Visualization
- Tools for Data Visualization (Matplotlib, Seaborn, ggplot2, Tableau, Cloud-Based tools)
- Creating Dashboards
- Communicating Data Insights

3.2: Data Ethics and Privacy

- Ethical considerations in Data Analysis
- Data Privacy laws and regulations (GDPR, DPDPA)
- Bias and fairness in Algorithms

Module 4: Working with Real Projects

4.1: Case Studies and Applications

- Real-World Data Analysis Case Studies
- Industry-Specific Applications (Finance, Healthcare, Retail, etc.)

4.2: Data Analysis Project

- Project Selection and Planning
- Data Collection and Preprocessing
- Model Building and Evaluation
- Presenting Results

TuxAcademy

Data Analysis Program

Data Management | Python | SQL | Machine Learning

Visualization | Data Ethics | Real World Projects