



## Data Analytics Professional Program

3-Month Comprehensive Curriculum.....	2
Month 1: Data Foundations & Statistical Analysis.....	2
Week 1: Introduction to Statistics.....	2
Week 2: Advanced Spreadsheet Modeling.....	2
Week 3: SQL for Data Analysis.....	2
Week 4: Data Cleaning & Wrangling.....	3
[Project 1: Healthcare Data Insights].....	3
Month 2: Python for Data Analysis & Visualization.....	3
Week 5: Python Fundamentals.....	3
Week 6: Data Manipulation with Pandas & NumPy.....	3
Week 7: Exploratory Data Analysis (EDA).....	4
Week 8: BI Tools & Dashboarding.....	4
[Project 2: E-Commerce Sales Dashboard].....	4
Month 3: Advanced Analytical Techniques & Career Readiness.....	4
Week 9: A/B Testing & Hypothesis Testing.....	4
Week 10: Intro to Machine Learning for Analysts.....	5
Week 11: Web Scraping & API Data Extraction.....	5
Week 12: Capstone Project & Career Readiness.....	5
[Project 3: Customer Churn Prediction (Capstone)].....	5



# Data Analytics Professional Program

## 3-Month Comprehensive Curriculum

### Month 1: Data Foundations & Statistical Analysis

**Goal:** Master data manipulation, cleansing, and core statistical methods used in decision-making.

#### Week 1: Introduction to Statistics

- **Descriptive Statistics:** Understanding Mean, Median, Variance, Standard Deviation, and identifying data distributions.
- **Probability Basics:** Exploring probability distributions, normal distribution, and utilizing z-scores for data analysis.
- **Statistical Thinking:** Applying statistical concepts to real-world business problems and recognizing common data biases.

#### Week 2: Advanced Spreadsheet Modeling

- **Advanced Functions:** Mastering Google Sheets/Excel array formulas, IF/AND/OR logic, and advanced lookups (INDEX/MATCH, XLOOKUP).
- **Data Cleaning in Spreadsheets:** Using Text-to-Columns, conditional formatting, data validation, and handling duplicates efficiently.
- **Automated Reporting:** Creating dynamic dashboards using Pivot Tables, slicers, and recording basic macros to automate repetitive tasks.

#### Week 3: SQL for Data Analysis

- **Relational Databases:** Understanding database schemas, primary and foreign keys, and basic RDBMS architecture.
- **Querying Fundamentals:** Writing complex SQL queries using SELECT, FROM, WHERE, GROUP BY, and HAVING clauses.
- **Advanced Joins:** Utilizing INNER, LEFT, RIGHT, and FULL OUTER joins to combine and analyze multiple datasets effectively.



### **Week 4: Data Cleaning & Wrangling**

- **Handling Anomalies:** Identifying and managing null values, outliers, and missing data points within datasets.
- **Data Transformation:** Changing data types, string manipulation (SUBSTRING, CONCAT), and date/time formatting directly in SQL.
- **Subqueries & CTEs:** Writing subqueries and Common Table Expressions (CTEs) for multi-step data transformations and readability.

#### **[Project 1: Healthcare Data Insights]**

- **Scope:** Clean and analyze a messy healthcare dataset using Excel and SQL to find trends in patient admission rates and demographic patterns. Present findings in a structured summary.
- 

## **Month 2: Python for Data Analysis & Visualization**

**Goal:** Utilize Python to automate analysis, manipulate large datasets, and build interactive data visual reports.

### **Week 5: Python Fundamentals**

- **Python Basics:** Setting up Jupyter Notebooks, understanding variables, data types, loops, and conditional statements.
- **Data Structures:** Working with Python lists, dictionaries, tuples, and sets for efficient data storage and retrieval.
- **Functions & Modules:** Writing reusable functions and importing essential data science and mathematical libraries.

### **Week 6: Data Manipulation with Pandas & NumPy**

- **NumPy Arrays:** Performing high-performance mathematical operations and array manipulations using NumPy.
- **Pandas DataFrames:** Loading datasets (CSV, Excel, SQL), filtering, sorting, merging, and reshaping data structures.
- **Advanced Pandas:** Grouping data, handling missing values programmatically, and utilizing apply() functions for custom transformations.



### **Week 7: Exploratory Data Analysis (EDA)**

- **Data Visualization Basics:** Using Matplotlib to create foundational charts (line, bar, scatter, pie) for data exploration.
- **Statistical Visualization:** Utilizing Seaborn to visualize data distributions, heatmaps, boxplots, and identifying variable correlations.
- **Storytelling with EDA:** Extracting actionable insights from visual patterns and preparing data narratives for stakeholder presentations.

### **Week 8: BI Tools & Dashboarding**

- **Power BI Setup:** Connecting Python and SQL data sources to Power BI for comprehensive and automated reporting.
- **Data Modeling in BI:** Creating relationships between tables (Star/Snowflake schemas) and utilizing advanced Data Analysis Expressions (DAX).
- **Interactive Dashboards:** Designing user-centric, interactive dashboards with drill-down capabilities and cross-filtering.

### **[Project 2: E-Commerce Sales Dashboard]**

- **Scope:** Perform an end-to-end Exploratory Data Analysis (EDA) on an e-commerce dataset using Python (Pandas/Seaborn) and present the final insights through an interactive Power BI dashboard.

---

## **Month 3: Advanced Analytical Techniques & Career Readiness**

**Goal:** Apply predictive concepts to structured data, gather external data, and prepare for industry roles.

### **Week 9: A/B Testing & Hypothesis Testing**

- **Hypothesis Design:** Formulating null and alternative hypotheses for business experiments and product changes.
- **Statistical Testing:** Conducting T-tests, calculating p-values, and determining statistical significance using Python libraries.
- **A/B Test Analysis:** Interpreting the results of A/B tests to make data-driven product, marketing, or operational decisions.



### **Week 10: Intro to Machine Learning for Analysts**

- **Predictive Modeling Concepts:** Understanding the core differences between regression and classification tasks in analytics.
- **Linear Regression:** Building simple and multiple linear regression models using Scikit-Learn to predict continuous business variables (e.g., sales).
- **Classification Basics:** Implementing Logistic Regression and interpreting model performance using confusion matrices and accuracy scores.

### **Week 11: Web Scraping & API Data Extraction**

- **API Integrations:** Making HTTP requests using the `requests` library to extract real-time data from public RESTful APIs.
- **JSON Handling:** Parsing complex JSON responses and flattening nested data into structured Pandas DataFrames.
- **Web Scraping:** Using BeautifulSoup to extract structured and unstructured data from HTML web pages for competitive analysis.

### **Week 12: Capstone Project & Career Readiness**

- **Portfolio Building:** Uploading completed projects to GitHub with well-documented README files and polished Jupyter Notebooks.
- **Interview Prep:** Practicing live SQL whiteboard questions, Python technical assessments, and data storytelling presentations.
- **Resume Optimization:** Crafting technical CVs and LinkedIn profiles to highlight specific data analytics tools, methodologies, and project impacts.

### **[Project 3: Customer Churn Prediction (Capstone)]**

- **Scope:** Build a predictive churn model using Python (Pandas and Scikit-Learn), identify key factors driving customer attrition, and create an executive summary presentation detailing retention strategies.