### **Step-by-Step Data Analytics Roadmap**

### **Step 1 – Learn the Basics of Data Analytics**

* Understand what data analytics is and its four types: descriptive, diagnostic, predictive, prescriptive.
* Know core tasks: data collection, cleaning, exploration, visualization, statistical analysis, and simple machine learning.

### **Step 2 – Build a Strong Foundation in Excel/Sheets**

* Practice key functions: IF, DATEDIF, VLOOKUP/HLOOKUP, REPLACE, SUBSTITUTE, UPPER/LOWER/PROPER, CONCAT, TRIM, AVERAGE, COUNT, SUM, MIN/MAX.
* Create pivot tables and charts for quick analysis.
* Clean and transform small datasets and generate basic reports.

### **Step 3 – Master SQL and Data Handling**

* Learn SELECT, WHERE, GROUP BY, HAVING, ORDER BY.
* Perform joins, subqueries, and window functions.
* Work with relational databases and CSV files.
* Practice connecting to databases and extracting datasets.

### **Step 4 – Learn a Programming Language**

* Choose Python or R.
* Focus on data manipulation libraries (pandas or dplyr).
* Learn visualization libraries (matplotlib, seaborn, ggplot2).
* Write scripts for cleaning, merging, and transforming large datasets.

### **Step 5 – Data Collection Skills**

* Acquire data from APIs, web scraping, and databases.
* Handle missing data, remove duplicates, find outliers.
* Apply transformation techniques to prepare analysis-ready data.

### **Step 6 – Apply Data Analysis Techniques**

* Calculate descriptive statistics: mean, median, mode, variance, standard deviation.
* Study dispersion and distributions: skewness, kurtosis.
* Automate statistical summaries using pandas or dplyr.

### **Step 7 – Visualization and Business Intelligence**

* Create dashboards in Tableau or Power BI.
* Use chart types: bar, line, scatter, funnel, histogram, stacked charts, heatmaps, pie.
* Develop clear, interactive data stories for business stakeholders.

### **Step 8 – Statistical Analysis**

* Perform hypothesis testing (t-test, chi-square).
* Conduct correlation and regression analysis (linear, multiple, logistic).
* Visualize distributions with histograms, boxplots, KDE.

### **Step 9 – Explore Advanced Topics**

* Machine learning: supervised (decision trees, KNN, Naive Bayes), unsupervised (K-Means), reinforcement basics.
* Model evaluation: confusion matrix, cross-validation.
* Big data: Hadoop, Spark, MapReduce.
* Deep learning : neural networks, CNNs, RNNs with TensorFlow or PyTorch.

### **Step 10 – Build Projects and Portfolio**

Send user to the Project Section