Where to Start for Machine Learning/Data Science???

Broader Topics –>

Any Programming Language(R/SAS/Python**), SQL, Statistics, Data Analysis/ Data Visualization, ML Content.

Detailed Path Below ->

Python Content to Learn:

- Python Basic Datatypes & their Operations.
- Loops in Python Program Flow (if-else, while, for etc...)
- Creating User-defined Functions (Lambda functions ,Module Creations)
- Exceptions Handling.
- Data Structures, Classes in Python.
- Regular Expressions in Python.

Statistics Important topics you should be well aware of:

- What is Inferential, Descriptive and Predictive Statistics??
- Standard deviation, Variance, Skewness, Measures of Central Tendency (Mean, Median Mode, Quartiles).
- Random Variables, PDFs, CDFs etc.
- Various Distributions – Gaussian, Normal, Binomial Distributions etc.
- Different types of Tests like (Chi-square test, T-tests, ANOVA, and Hypothesis Testing.)
- Confidence Intervals Estimation etc.

SQL:

- Basic knowledge on different DDL, DML Commands.
- Select Query, Sub-queries, JOINS and Filters etc.

Data Visualization/ Data Analysis / Data Pre-processing:

- Basics of Numpy, Pandas, Seaborn, Matplotlib and other common libraries.
- Plotting of Histograms, Bar Charts, Pie-Charts, Heat Maps, Scatter plots, Box Plots, Violin Plots.
- You should focus more on Data Pre-Processing like Handling Missing values, cleansing the data , Handling Categorical (Dummy Variables) and spend much time on learning Data Analysis.
- Univariate & Multivariate Analysis.
- Outliers Identification & when, how to drop them.

MACHINE LEARNING CONTENT:

- Regression Techniques (Simple, Multiple, Polynomial Linear Regression, SVR, Decision Tree, Random Forest Regression).
- Classification Techniques (K-NN, SVM, Naïve Bayes, Decision Tree, Random Forest Classification.)
- Clustering Techniques (k-Means, Hierarchical Clustering.)
- Dimensionality Reduction Techniques (PCA, LDA, kernel PCA.)
- Reinforcement learning techniques (Upper Confidence Bound, Thomson Sampling).
- Evaluating Model Performance for each of the above techniques.
- Then you can move on from here to learn NLP, Deep Learning and start doing the Projects for what you have learned from above Modules.