**Instructor-Led Live Coding Session on EDA**

This coaching session will be about 1.5 hours long, and it focuses on the concepts of exploratory data analysis (EDA). Learners will also be performing EDA through live coding along with the instructor. In the first hour, they will be taught the concepts and then they will perform EDA on a data set. This will be followed by doubt clearance.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Learning Objective(s)**

The learners attending this session will:

1. Be introduced to EDA and understand the fundamental concepts, and
2. Perform live coding along with the instructor on a sample data set and understand when to apply a particular concept.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Prerequisite(s)**

Learners are advised to have a good grasp on the following topics beforehand in order to make the best of this session.

* Python for Data Science
* Data Visualization in Python

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Brief Agenda**

1. Part - I: Concepts of EDA and demonstration (40 min)
2. Part - II: Performing EDA on a data set live with learner inputs (task-based) (40 min)
3. Part - III: Doubt clearance (10 min)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Detailed Lesson Plan**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component** | **Instruction Task/Learner Task/Question** | **Time (min)** | **#Qs** | **Element of Engagement** |
| **Part - I: Concepts of EDA and Demonstration** | The instructor starts with explaining what EDA is and discusses its importance. S/he then introduces the sample case study in which the learners will apply some fundamental EDA techniques using Python.  The coach walks the learners through the most important and fundamental concepts related to EDA around data cleaning (data wrangling, central tendencies, imputing, handling outliers), univariate and bivariate analyses (using different visualizations), etc, through use cases and examples. For each concept being explained, the same analysis is performed on the data set that has been chosen for the case study. | 40 | 8 | Concepts |
| **Part - II: Solving Questions** | The learners will perform some EDA tasks assigned to them by the instructor and discuss the solution for every task. Tasks can range from different topics in data cleaning to univariate/bivariate/ multivariate analysis using abstract questions. | 40 | 5 | Task support |
| **Part - III: Doubt Clearance** | The instructor clears the doubts of the students. | 10 | - | Doubt Resolution |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Additional Reading for the Week**

1. Link 1

[An Extensive Step by Step Guide to Exploratory Data Analysis](https://towardsdatascience.com/an-extensive-guide-to-exploratory-data-analysis-ddd99a03199e)

1. Link 2  
   [1. Exploratory Data Analysis](https://www.itl.nist.gov/div898/handbook/eda/eda.htm) (A good directory for concept recapitulation)
2. Link 3

[67 Types of Data Visualizations: Are You Using the Right One?](https://learn.g2.com/types-of-data-visualizations) (for visualization best practices)