EXPLORATORY DATA ANALYSIS

CODEXCUE

PURPOSE

✋ *Exploratory Data Analysis (EDA) is the cornerstone of data-driven decision-making, serving as the critical first step in uncovering patterns, anomalies, and correlations within vast datasets. It ensures that subsequent analyses are based on solid, well-understood foundations, guiding businesses and researchers towards accurate, actionable insights. By revealing the story behind the data, EDA empowers informed assumptions, hypothesis testing, and strategic planning. In a world overwhelmed with data, EDA clarifies the path forward, enabling precise predictions and impactful innovations. Essentially, EDA transforms raw data into a clear vision for informed decision-making, marking the beginning of every data analysis journey.*

PROJECTS OVERVIEW

ALLOCATED PROJECTS

In the Codex Cue Internship Program, interns will master **Exploratory Data Analysis** (EDA) through three focused projects: "Intermediate Colleges of Pakistan," "Titanic," and "Hotels Booking." These projects serve as practical arenas for you to apply EDA techniques, crucial for revealing hidden patterns and guiding further analysis.

Intermediate Colleges of Pakistan EDA project entails analyzing educational data from across Pakistan to identify performance trends and regional disparities among intermediate colleges. Interns will use EDA to dissect factors such as academic results, faculty quality, and resources, aiming to highlight areas needing improvement. The analysis seeks to inform targeted interventions to bridge educational gaps between different regions, ultimately enhancing the national education standard.

Titanic EDA you will delve into the Titanic dataset, examining socio-demographic factors like age, class, and family size to understand their impact on survival rates. Through EDA, this project aims to uncover patterns that reveal the underlying factors contributing to passenger survival. The goal is to provide a nuanced view of how different attributes influenced outcomes, offering insights into the historical event's human aspect.

Hotels Booking Focusing on hotel booking data, this project involves investigating customer booking and cancellation trends. Interns will apply EDA to pinpoint key influences such as seasonality, lead time, and customer preferences. The insights gathered will aid in understanding consumer behavior, assisting

hoteliers in optimizing marketing strategies, pricing models, and service offerings to improve occupancy rates and enhance guest experiences.

GOLDEN PROJECTS

Moving to the Golden Project Section focused on Exploratory Data Analysis (EDA), you will embark on two advanced projects: "Pakistan Suicide Bombing Attacks" and "Playstore Apps". These projects will refine your skills in data analytics, preparing you for industry-ready challenges beyond the foundational scope of NLP, emphasizing machine learning and advanced data analysis techniques.

Pakistan Suicide Bombing Attacks This project involves an in-depth analysis of data related to suicide bombing attacks in Pakistan. By applying EDA techniques, you will explore patterns, frequencies, locations, and the impact of these attacks over time. The goal is to uncover trends that could inform security policies and preventive measures. This analysis will not only highlight socio-political and demographic factors associated with these incidents but also aim to provide insights that could contribute to peacekeeping and counter-terrorism strategies.
Google PlayStore Apps In this project, you will dive into the data surrounding apps available on the Playstore. Through EDA, you will analyze user ratings, download counts, category successes, and monetization strategies to identify what drives app popularity and user engagement. This project aims to uncover key factors that app developers and marketers can leverage to enhance app visibility, improve user experience, and drive downloads.

PROJECT DETAILS

Project Deliverables

Following is a complete list of all project deliverables:

Intermediate Colleges of Pakistan EDA

Project 4	Description
IDE	Google Colab, Jupyter Notebook, Visual Studio Code
Language	Python

Dataset Source	Download Dataset From Here
Duration	07 Days

Titanic EDA

Project 4	Description
IDE	Google Colab, Jupyter Notebook, Visual Studio Code
Language	Python
Dataset Source	Use sns.load_dataset('titanic') command from Seaborn Library to use the titanic dataset
Duration	07 Days

Hotels Booking EDA

Project 4	Description
IDE	Google Colab, Jupyter Notebook, Visual Studio Code
Language	Python
Dataset Source	Download Dataset From here

Duration	07 Days

Pakistan Suicide Bombing Attacks

Project 4	Description
IDE	Google Colab, Jupyter Notebook, Visual Studio Code
Language	Python
Dataset Source	Download Dataset From Here
Duration	10 Days

Google Play Store Apps EDA

Project 4	Description
IDE	Google Colab, Jupyter Notebook, Visual Studio Code
Language	Python
Dataset Source	Download Dataset From Here
Duration	10 Days

INSTRUCTIONS

- **You have to complete least 3 Allocated Projects assigned to you to Earn Internship Completion Certificate**

- **You have to complete the atleast Golden 1 Project to get the Recommendation Letter**

- **You have to complete all allocated projects and golden projects to earn swags and more.**

MOTIVATION

Embarking on your EDA journey, we wish you success and are confident in your ability to meet the challenges ahead. Through projects like "Pakistan Suicide Bombing Attacks," "Google Play Store Apps," a "Hotels Booking," etc. This program aims to sharpen your data analytics and machine learning skills, preparing you for the competitive data science field.

These projects will equip you with the ability to derive insights from complex data and apply them in practical scenarios, making you industry-ready. Good luck on this journey of professional growth and learning.

Thank you for joining us. Your progress in these projects is set to advance your career in data analytics significantly.