# Expected Outcomes

## Project Overview and Data

- What is your project about? Briefly introduce the topic and the business problem it addresses.
- What is the data you used? Describe your dataset, its source, and the key variables it contains.
- How did you clean and prepare the data? Discuss the data manipulation steps you took, such as handling missing values, dealing with duplicates, and transforming data types.

## Exploratory Data Analysis (EDA)

- What were the key insights from your descriptive statistics? Present your findings from measures of central tendency (mean, median) and dispersion (standard deviation).
- What questions did you ask the data, and what did you discover? This is the core of your analysis. Explain the hypotheses you had and the patterns, trends, or relationships you found. For example, "We wanted to see if there was a relationship between a car's age and its price, and we discovered that..." .
- Which variables had the strongest relationships with each other? Discuss the correlations you found between different features in the dataset.

## Data Visualization

- How did you visualize your findings? Show the key charts and dashboards you created.
- Why did you choose these specific visualizations? Explain the rationale behind your choice of charts (e.g., "We used a bar chart to compare sales across different regions because it clearly shows magnitude differences").
- What story does your dashboard tell? Explain how your interactive dashboard helps a user explore the data and understand the key insights at a glance.

## Conclusion and Recommendations

- What were the main conclusions of your analysis? Summarize your most significant findings.
- What business recommendations can you make based on your findings? Translate your data insights into actionable advice. For example, "Based on the analysis, we recommend focusing marketing efforts on [specific product/region/customer segment] to increase sales."
- What are the limitations of your analysis, and what would you do next?

# Projects Pool

## 1. Project: Analyzing Student Academic Performance

- Project Description: This project aims to analyze student performance in a school or university and understand the factors that influence their grades. You can use tools like Pandas to clean and process the data, then use Matplotlib or Seaborn in Python to create insightful visualizations.
- Key Tasks: Clean the data, calculate descriptive statistics (like mean and standard deviation), and create visualizations such as pie charts and bar charts to show the relationship between variables.
- Kaggle Project Link: Students Performance in Exams

## 2. Project: E-commerce Sales Analysis

- Project Description: This is a great project for analyzing sales data and customer purchasing behavior. You can use Power BI or Tableau to build an interactive dashboard that displays key performance indicators (KPIs) like total sales, best-selling products, and sales distribution by time or region.
- Key Tasks: Prepare the data in Power BI or Tableau, and create multiple visualizations (such as maps, line charts, and bar charts) to show trends and patterns in sales.
- Kaggle Project Link: Sales Analysis Dashboard

## 3. Project: Car Dataset Analysis

- Project Description: This project focuses on exploring a dataset of cars. You can use Python and Pandas to sort and group the data, then use Seaborn to create visualizations that show the relationship between different variables like car price, model year, and engine type.
- Key Tasks: Clean the data and handle missing values, calculate descriptive statistics, and then create visualizations (like scatter plots and histograms) to explore the data's distribution and the relationships between variables.
- Kaggle Project Link: Car Dataset

## 4. Project: Real Estate Price Analysis and Comparison

- Project Description: This is an excellent project for applying your data analysis skills to explore the factors that affect house prices. You can use Python for initial analysis, then use Tableau to build a dashboard that shows price distribution by region and the impact of various factors like property size and the number of rooms.
- Key Tasks: Prepare the data, and create interactive visualizations in Tableau that allow the user to explore the data by multiple variables, such as location, price, and area.
- Kaggle Project Link: Housing Prices in CDMX

## 5. Project: Analysis of Crimes Against Women in India

- Project Description: This is an important project that analyzes data on crimes against women in India. You can use Python to analyze trends and patterns and understand the spread of this phenomenon. Then, use Power BI or Tableau to create a dashboard that visually represents the data in an easily understandable format for the public and NGOs.
- Key Tasks: Clean the data and create visualizations that show the crime rates by state or region, and the distribution of cases by type.
- Kaggle Project Link: Crimes Against Women in India

## 6. Project: Netflix Movies and TV Shows Analysis

- Project Description: This project involves analyzing the content on Netflix. The dataset includes information about movies and TV shows, such as release year, genres, directors, and ratings. A student can clean this data and perform descriptive analysis to uncover trends in content production.
- Key Tasks: Data cleaning to handle missing values and inconsistencies. Calculate the number of movies vs. TV shows over time. Identify the most prolific directors and popular genres. Visualize the data using Python (e.g., bar charts, line plots) to show trends and distribution.
- Kaggle Project Link: Netflix Movies and TV Shows

## 7. Project: Video Games Sales Analysis

- Project Description: This dataset contains a list of video games with sales data across different regions, along with other information like platform, genre, and publisher. A student can perform an in-depth exploratory analysis to understand the video game market.
- Key Tasks: Clean and prepare the data for analysis. Calculate total sales by genre, platform, and publisher. Use Python with Seaborn to visualize sales trends over time, compare sales between different platforms, and identify the most successful genres.
- Kaggle Project Link: Video Game Sales

## 8. Project: Analyzing Airbnb Listings in New York City

- Project Description: This project uses a detailed dataset of Airbnb listings in NYC. It's perfect for geospatial and statistical analysis. A student can explore how prices are distributed across different neighborhoods, what types of rooms are most common, and how factors like review scores correlate with pricing.
- Key Tasks: Data cleaning to handle missing values in key columns. Analyze the average price per neighborhood. Use Tableau to create an interactive map of NYC with listings, colored by price or room type. Build a dashboard to visualize price and availability distribution.
- Kaggle Project Link: New York City Airbnb Open Data

## 9. Project: Data Science Job Salaries

- Project Description: This project involves analyzing a dataset of data science salaries from various companies and locations. It's a great way to practice data cleaning and perform descriptive analysis to identify salary trends and factors influencing pay.
- Key Tasks: Clean the data and group salaries by job title, company size, and location. Calculate the average salary for different roles (e.g., Data Scientist vs. Data Analyst). Use Power BI or Tableau to create a dashboard that shows salary distribution by experience level and location.
- Kaggle Project Link: Data Science Job Salaries

## 10. Uber Ride-Sharing Analytics

- Project Description: This project involves analyzing Uber ride data to uncover key trends and user behavior. Students can explore the data to understand user behavior, identify peak hours, and analyze trip characteristics. This is an excellent project for practicing data cleaning and time-series analysis.
- Key Tasks: Clean the data and ensure data types are correct (e.g., converting timestamps). Analyze ride activity by time of day, day of the week, and month to find peak demand periods. Determine the most popular pickup and drop-off locations and visualize trip densities. Build a dashboard in Power BI or Tableau that allows a user to explore ride data.
- Kaggle Project Link: Uber Ride Analytics Dashboard

## 11. Mobile App Review Insights

- Project Description: This project uses a multilingual dataset of mobile app reviews to analyze user feedback and sentiment. While it's a text-based dataset, the analysis can be framed around descriptive text mining without complex predictive models.
- Key Tasks: Clean and normalize the text data by removing punctuation and converting text to a consistent case. Count the frequency of words and phrases to identify the most common topics. Group reviews by basic sentiment. Create visualizations in Python or Tableau, such as word clouds, to highlight key insights.
- Kaggle Project Link: Multilingual Mobile App Reviews Dataset

## 12. Student Stress Level Analysis

- Project Description: This project explores a dataset that combines physiological data (like heart rate and body temperature) with self-reported stress levels from students. The goal is to perform an exploratory analysis to find patterns and correlations between these different data points.
- Key Tasks: Handle missing data and use descriptive statistics to summarize key variables like average heart rate. Investigate how physiological metrics correlate with self-reported stress levels. Compare physiological data during periods of high vs. low stress. Build a dashboard that visualizes the distribution of stress levels and shows the relationship between different variables.
- Kaggle Project Link: Student Stress Monitoring Datasets