

Project Title: "E-commerce Analytics: Understanding Pakistani Market"

Project Overview:

Welcome to Week Two of the Data Analysis Bootcamp! In this project, you will dive deep into a rich dataset containing detailed information about e-commerce orders in Pakistan. You will leverage your knowledge of descriptive statistics and inferential statistics using spreadsheets to uncover insights about the e-commerce market in Pakistan. This project will challenge you to explore and analyze real-world data to make informed decisions.

Dataset Description:

You have access to a dataset with the following variables:

1. **Item ID:** Unique identifier for each product.
2. **Order Status:** The status of the order (Completed, Cancelled, Refund).
3. **Date of Order:** The date when the order was placed.
4. **SKU:** Stock Keeping Unit, a unique identifier for each product.
5. **Price:** The price of the product.
6. **Quantity:** The quantity of the product ordered in each transaction.
7. **Grand Total:** The total amount of the order.
8. **Category:** The category of the product (e.g., fashion, mobile, electronics, appliance).
9. **Payment Method:** The method of payment used (e.g., credit card, Easy-Paisa, Jazz-Cash, cash-on-delivery).
10. **Customer ID:** Unique identifier for each customer.

Project Tasks:

1. Data Exploration (2 Days):

- Load the dataset into a spreadsheet tool (e.g., Excel or Google Sheets).
- Examine the dataset's structure and missing values.
- Create summary statistics for key variables (e.g., Price, Quantity, Grand Total) to understand their distributions.

2. Descriptive Statistics (2 Days):

- Calculate descriptive statistics such as mean, median, standard deviation, and quartiles for relevant columns.
- Create visualizations (histograms, box plots, etc.) to explore the distribution of prices, quantities, and grand totals.
- Identify trends or patterns in the data, such as which product categories are most popular or which payment methods are commonly used.

3. Inferential Statistics (3 Days):

- Formulate hypotheses related to the dataset. For example:
 - Is there a significant difference in the average order total between different payment methods?
 - Do certain product categories have a higher likelihood of being canceled or refunded?
- Perform appropriate statistical tests (t-tests, ANOVA, chi-squared tests, etc.) to test your hypotheses.
- Interpret the results and provide insights and recommendations based on your findings.

4. Data Cleanup and Presentation (1 Day):

- Clean and format the dataset as needed for presentation.
- Create a visually appealing and informative spreadsheet-based presentation summarizing your findings.
- Include descriptive statistics tables, inferential statistics results, and visualizations in your presentation.

5. Final Presentation (1 Day):

- Deliver a concise presentation in form of writing, summarizing your analysis, key findings, and any actionable recommendations for e-commerce businesses operating in Pakistan.
- Be prepared to engage in discussions and answer questions about your analysis.

Deliverables:

- Excel spreadsheet containing all data and calculations.
- A presentation created in Excel or Google Sheets.
- A brief report summarizing your findings, insights, and recommendations.

This project will provide you with valuable experience in conducting data analysis using spreadsheet tools and help you gain insights into the e-commerce market in Pakistan. Happy analyzing!