GUVI

DATA SCIENCE JOB FAIR–2024

DATA / BUSINESS ANALYST

E-Commerce Analysis

Introduction:

The purpose of this document is to outline the objectives, methodologies, and expected outcomes of an e-commerce product analysis project. This analysis aims to provide insights into various aspects of product data collected from a major Indian e-commerce platform.

Data Understanding:

The dataset comprises information on products available on the e-commerce platform. It includes attributes such as product details, pricing, seller information, and more. A detailed description of each column is provided below:

Dataset: [Link](#)

Features:

Title: Title of the product.

Id: Internal ID for the product.

actual_price: The actual price of the product.

average_rating: Average rating of the product.

brand: Brand of the product.

category: Category to which the product belongs.

crawled_at: Date and time the product information was collected.

description: Description of the product.

discount: Percentage discount applied to the product.

images: URLs of images associated with the product.

out_of_stock: Indicates whether the product is out of stock.

pid: Unique product identifier.

product_details: Additional details about the product.
seller: Name of the seller offering the product.
selling_price: Current selling price of the product.
sub_category: Sub-category to which the product belongs.
url: URL of the product page.

Project Objectives:

- This e-commerce product analysis project aims to provide valuable insights into the e-commerce landscape, benefiting both businesses and consumers by informing strategic decisions and optimizations.
- Identifying popular product categories and brands.
- Analyzing pricing trends and discount strategies.
- Understanding seller behavior and performance.
- Investigating customer preferences and product satisfaction.
- Exploring correlations between various attributes such as price, rating, and brand.
- Extract valuable insights from textual data using text analysis techniques, uncovering hidden trends and patterns in product descriptions and details (Optional).

Methodology:

- Utilizing descriptive statistics to summarize key attributes and distributions.
- Employing data visualization techniques to visualize trends and patterns.
- Conducting exploratory data analysis to uncover insights and relationships.
- Utilizing statistical analysis to test hypotheses and validate findings.
- Implementing text analysis techniques (if required) to extract insights from product descriptions and reviews.
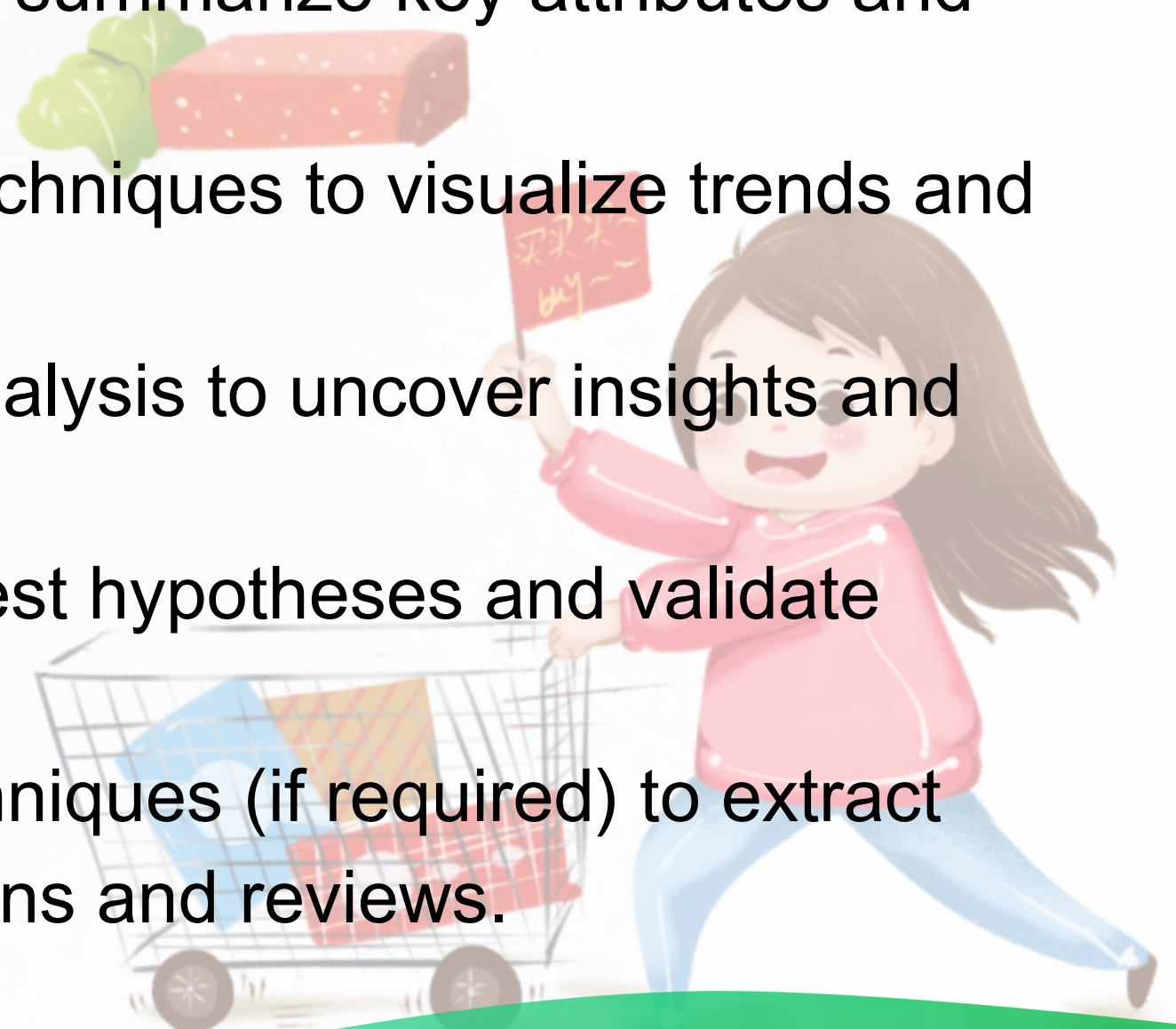
Expected Outcomes:

- Insights into popular product categories and brands.
- Understanding of pricing strategies and discount patterns.
- Evaluation of seller performance and behavior.
- Identification of customer preferences and satisfaction levels.
- Recommendations for e-commerce platform optimization and improvement.

Disclaimer:

- It's important to note that this project is open-ended, and additional analyses can be conducted based on specific requirements or preferences. The inclusion of text analysis is optional and can be implemented based on stakeholder interests and objectives.

Submission:

- Provide a well-commented notebook file (*.ipynb) containing the complete code for the project, organized into sections for data preprocessing and Analysis.
- Upload the same into github with proper Readme file.
- Presentation on the entire project, including Problem Statement, Tools Used, Approaches and Insights Found.

Evaluation metrics:

- Project evaluation will be done in the live session and have to showcase the approaches done to complete the project
- You are supposed to write a code in a modular fashion (in functional blocks)
- Maintainable: It can be maintained, even as your codebase grows.
- Portable: It works the same in every environment (operating system)
- You have to maintain your code on GitHub.(Mandatory)
- You have to keep your GitHub repo public so that anyone can check yourcode.(Mandatory)
- Proper readme file you have to maintain for any project development(Mandatory)
- Follow the coding standards: https://www.python.org/dev/peps/pep-0008/
- You should include basic workflow and execution of the entire project in the readme file on GitHub

GitHub Repo:

The attached reference document will help you use GitHub effectively. - **Link**