Super shop Data Analysis Project Overview

This project demonstrates a comprehensive analysis of sales data for a fictional super shop. The dataset comprises various attributes, including product details, pricing, sales quantities, purchase types, payment methods, and geographic information. Below is a detailed breakdown of the project, highlighting its key components and insights:

---

1. Data Overview and Preparation

- The dataset contains multiple fields such as:
    - **Order Information**: Order ID and date of purchase.
    - **Product Details**: Product category (e.g., Fries, Burgers, Beverages) and pricing.
    - **Transaction Details**: Purchase type (Online/In-store) and payment methods.
    - **Management and Geography**: Store managers and cities where the transactions occurred.
- Initial cleaning steps included:
    - Standardizing column headers.
    - Adding derived metrics such as revenue (Price × Quantity).
    - Identifying relationships between cost, price, and quantity.
    - Ensuring consistency in data types for analysis.

---

2. Data Cleaning

- Redundant or inconsistent data was removed, ensuring accuracy and readiness for analysis.
- Added calculated fields:
    - **Cost**: Manufacturing/operational cost for each product.
    - **Revenue**: Total revenue generated from each product category.
    - **Profit**: Derived using (Revenue - Cost).

---

3. Data Analysis and Insights

The analysis section is enriched with pivot tables, visualizations, and key findings:

1. **Best-Selling Products**:
    - Products like **Burgers** and **Chicken Sandwiches** are the most profitable due to their high unit price and strong sales volumes.
    - Fries and Beverages are sold in larger quantities but contribute less to total revenue.
2. **Total Revenue**:
    - The total revenue across all transactions was calculated as approximately **$136,430**.
    - Geographic breakdown shows that Madrid contributed the highest revenue, followed by Paris and Berlin.
3. **Revenue Breakdown by Payment Method**:
    - The majority of transactions used **Credit Card** payments, predominantly for in-store purchases.
    - **Gift Cards** were commonly used for online transactions.
4. **Geographic and Managerial Performance**:
    - Manager-specific and city-specific analysis highlighted **Pablo Perez** (Madrid) as the top-performing manager in terms of revenue.

---

4. Visualizations

The project includes several visualizations to support findings:

1. **Product vs. Price Analysis**:
    - A bar chart illustrates that Burgers and Chicken Sandwiches have the highest average price.
2. **Cost by Product and Purchase Type**:
    - A grouped bar chart demonstrates that in-store purchases incur higher costs compared to online or drive-thru transactions.
3. **Price vs. Revenue Correlation**:
    - A scatter plot shows a strong positive correlation between product price and total revenue, reinforcing the importance of high-priced items.

4. **Average Price Range**:
    - A box plot highlights the distribution of product prices, showing that most prices fall within a moderate range, with few outliers.

---

5. Tools and Techniques Used

- **Data Cleaning**: Performed in Excel, ensuring accurate and actionable data.
- **Pivot Tables**: Used to generate summary insights such as revenue breakdown by product, manager, and geography.
- **Visualizations**: Charts (bar, scatter, box) created in Excel to effectively communicate key patterns and trends.

---

6. Key Findings

- Burgers and Chicken Sandwiches are the most valuable products in terms of revenue.
- In-store sales dominate overall revenue generation.
- Credit Card payments are the preferred mode of transaction.
- Among the cities, Madrid leads in revenue, with Pablo Perez being the top-performing manager.

---

7. Applications of the Project

This project demonstrates proficiency in:

- Data cleaning and preparation.
- Analyzing sales data to uncover actionable insights.
- Creating visualizations for presenting findings effectively.
- Using Excel for data modeling and reporting.