1. Title: Furniture Sales Data Analysis

2. Introduction

The furniture industry is highly competitive, with various factors such as price, material, and seasonal demand influencing sales. Through this project, we aim to analyze furniture sales data to uncover important trends and insights, such as identifying popular categories, seasonal effects, and store performance. This analysis will provide a deeper understanding of customer preferences, inventory management, and profit drivers, leading to data-driven decisions that enhance sales strategies and operational efficiency.

The dataset used in this project contains detailed records of furniture sales, including product category, material, color, price, inventory levels, sales volume, discount percentage, delivery days, location, season, store type, revenue, and profit margin. The data will be stored and managed in MySQL, with two key tables created: product_info and sales_info, which will be connected using SQL joins for analysis.

3. Objectives

The primary objectives of this project are:

- Analyze the impact of various product categories, materials, and colors on sales performance.
- Identify which season sees the highest sales and which categories are in demand during different seasons.
- Investigate the effect of discounts on sales volume and revenue.
- Determine which store type (e.g., online vs. retail) has higher sales and profitability.
- Analyze inventory levels to identify potential overstock or understock situations.
- Calculate the average delivery time and its impact on customer satisfaction.
- Identify the relationship between sales volume by Store Type and Location.

4. Scope of Work

The project will involve the following tasks:

- **Data Creation and Management**: Creating a dataset using MySQL and structuring it into two tables: product_info (containing product-related data such as category, material, and color) and sales_info (containing sales-related data such as sales volume, revenue, and discounts). The two tables will be connected using SQL joins for efficient querying and analysis.
- **Data Exploration**: Gaining a detailed understanding of the dataset, including key features like category, material, price, and sales.
- **Data Preprocessing**: Cleaning the dataset by handling missing values, removing outliers, and normalizing/standardizing relevant data.
- **Feature Selection**: Identifying key variables that influence sales patterns, such as season, discount percentage, and location.

- **Data Visualization**: Using various plots and graphs to visualize relationships between key features, such as sales vs. price and season vs. inventory.
- **Interpretation of Results**: Drawing meaningful conclusions from the analysis and providing actionable insights for sales improvement and inventory management.
- **Reporting**: Documenting the findings and preparing a final report.

5. Methodology

The project will follow a structured approach:

1. **Data Collection**: The dataset contains detailed sales records, including product and transactional data.
2. **Data Creation and Management**:
    - Use MySQL to create and structure the dataset into two tables: product_info and sales_info.
    - Join the two tables using SQL joins to facilitate efficient data retrieval and analysis.
3. **Data Preprocessing**:
    - Handle missing values using appropriate imputation methods.
    - Detect and remove outliers to ensure accurate analysis.
    - Normalize or standardize data such as price and delivery days where needed.
4. **Exploratory Data Analysis (EDA)**:
    - Use descriptive statistics to summarize the dataset.
    - Create visualizations like bar plots, scatter plots, and correlation heatmaps to explore relationships between features.
5. **Visualization**:
    - Use graphs like line charts, pie charts, and histograms to present insights such as seasonal trends and category performance.
6. **Reporting**:
    - Compile the analysis, results, and insights into a comprehensive report for easy interpretation.

6. Tools and Technologies

The project will utilize the following tools and technologies:

- **Programming Language**: MySQL, Python
- **Libraries**: Pandas, NumPy, Matplotlib, Seaborn
- **Database Management**: MySQL (used for dataset creation, table structuring, and SQL joins)
- **IDE**: Jupyter Notebook or any Python-compatible Integrated Development Environment (IDE)

- **Data Source**: Kaggle, MySQL

7. Expected Outcomes

By the end of this project, we expect to achieve the following:

- Insights into seasonal sales trends and category-wise performance.
- Identification of the most popular materials and colors that drive sales.
- Analysis of the impact of discounts on sales and profit margins.
- Visualization of key metrics like inventory levels, delivery days, and store performance.
- Recommendations for optimizing inventory management and discount strategies.
- A final report summarizing key findings and providing actionable recommendations to improve sales and profitability.

8. Timeline

The project is expected to be completed within a [specific timeframe, e.g., 4 weeks], with the following milestones:

- **Week 1**: Data Collection and Preprocessing
- **Week 2**: Exploratory Data Analysis and Feature Selection
- **Week 3**: Visualization and Interpretation of Findings
- **Week 4**: Reporting and Final Submission

9. Conclusion

Based on the analysis, the key findings are:

- **Seasonal Sales**: Winter and Fall outperform Spring and Summer, particularly for categories like Chairs and Sofas. Tables, however, show poor sales except in Summer.
- **Discount Impact**: Sales fluctuate with discount percentages, peaking at both lower (5-10%) and higher (25-30%) discounts.
- **Store Performance**: Retail stores consistently outperform online stores in terms of both sales and profitability.
- **Inventory**: Beds and Chairs account for the largest inventory shares (48%), suggesting high demand, while Sofas and Tables have smaller shares.
- **Delivery**: There is no clear relationship between delivery time and sales, with high profit margins seen across varying delivery times.
- **Location**: Urban areas show decreasing sales moving from Online to Retail stores, while rural areas show a slight increase and suburban locations remain stable.