**Blinkit Sales report**

**1. Business Problem**

Blinkit, as a retail and FMCG platform, faces stiff competition in the market. To remain competitive and improve profitability, the business needs to:

* **Optimize sales and profitability** across different outlet types (e.g., grocery stores, supermarkets), item categories (e.g., snacks, dairy products), and outlet locations (e.g., Tier 1 cities, rural areas).
* Identify trends and patterns in sales data to improve **operational efficiency**, manage inventory effectively, and ensure customer satisfaction.

For Blinkit, this requires actionable insights into:

1. **Sales distribution** across outlet types, regions, and product categories.
2. Performance drivers such as **outlet location and size**.
3. Key areas where intervention (e.g., marketing campaigns, promotional discounts) can lead to improved results.

**Impact Statement**

Understanding these factors will enable Blinkit to allocate resources effectively, enhance its supply chain, improve customer satisfaction, and drive revenue growth.

**2. Data Requirement**

To address these challenges, the following data points are required:

1. **Item Data**
   * **Fat Content**: Items classified as “Low Fat” or “Regular” to understand customer preferences.
   * **Item Type**: E.g., Dairy, Fruits, Household, Snacks, etc., to identify high-performing categories.
   * **Visibility**: The extent to which an item is promoted or placed (visibility influences customer purchasing behavior).
   * **Weight**: Impacts transportation costs and packaging, especially for larger items.
   * **Sales**: Revenue generated by each item to analyze performance.
2. **Outlet Data**
   * **Outlet Identifier**: Unique ID for each outlet for trend analysis.
   * **Location Type**: Tiers (Tier 1, Tier 2, or Tier 3 cities) to segment customers based on geography.
   * **Size**: Small, Medium, or Large outlets to compare sales and operational efficiency.
   * **Type**: Supermarket (Type 1/Type 2) or Grocery to determine which format performs better.
   * **Establishment Year**: Helps identify how outlet age impacts performance.
3. **Performance Metrics**
   * **Total Sales**: Overall revenue per outlet and item category.
   * **Average Ratings**: Customer satisfaction indicators.
   * **Average Sales per Item**: Helps compare different products fairly.

**3. Data Collection and Understanding**

**Data Overview**

This dataset provides a combination of item-level details (e.g., fat content, visibility, and sales) and outlet-level attributes (e.g., location, type, and size).

Key components include:

* **Attributes**: Represent characteristics like fat content, outlet location, and type.
* **Key Metrics**: Sales and ratings that drive decision-making.
* **Business Context**: Segment outlets by size, location type, and establishment year to derive targeted insights.

**Importance of Segmentation**

Segmentation is essential because customer behavior differs based on:

* **Outlet Type**: Supermarkets in Tier 1 cities will likely attract a different customer base compared to small grocery stores in Tier 3 cities.
* **Item Type**: Dairy products may perform better in urban areas, whereas staples might dominate rural regions.

**4. Data Validation**

Before performing any analysis, the data must be validated to ensure reliability and accuracy.

**Validation Steps**

1. **Check for Bias**
   * Are certain outlet types (e.g., supermarkets) overrepresented?
   * Are sales figures skewed toward specific product categories (e.g., snacks)?
2. **Ensure Transparency**
   * Confirm metadata about data collection, such as time range and sources.
   * Verify data collection techniques (manual entries or automated systems).
3. **Validate Data Integrity**
   * **Missing Records**: E.g., missing sales values for certain items/outlets.
   * **Outliers**: Unusually high or low sales numbers that deviate significantly.
   * **Consistency**: Standardize categories like fat content (e.g., convert "LF" to "Low Fat").

**Outcome**

This step ensures the dataset is clean, balanced, and ready for analysis.

**5. Data Cleaning and Exploratory Data Analysis (EDA)**

**Data Cleaning**

* **Handling Missing Values**: Use mean/median imputation for numeric fields like sales or visibility.
* **Standardizing Categories**: Merge duplicate categories, e.g., "low fat" and "Low Fat".
* **Removing Outliers**: Remove items with visibility = 0 or sales anomalies that skew analysis.

**Key Insights from EDA**

1. **Sales Distributions**:
   * Snack Foods and Household items generate the highest sales.
   * Seafood and Breakfast items contribute the least to revenue.
2. **Outlet Insights**:
   * Medium-sized outlets outperform small-sized ones in terms of sales.
   * Supermarkets (Type 1) are the most profitable outlet type.
3. **Regional Analysis**:
   * Outlets in Tier 1 cities have the highest revenue contribution.
   * Rural (Tier 3) outlets struggle, indicating untapped potential.

**6. Tools Selection**

1. **Data Cleaning**: Python libraries like Pandas and NumPy.
2. **Data Visualization**: Tableau or Matplotlib for creating insights-driven visuals.
3. **Dashboard Development**: Tableau or Power BI to enable interactive exploration of trends.

**7. Graphs and Charts**

1. **Univariate Analysis**:
   * Distribution of sales across item categories and outlet types.
2. **Bivariate Analysis**:
   * Correlation between outlet location (e.g., Tier 1) and average sales.
3. **Multivariate Analysis**:
   * Combined effects of outlet size, type, and location on sales.

**8. Dashboard**

The dashboard is a powerful tool to visualize:

1. **Key Metrics**: Total sales, average ratings, and item counts.
2. **Segmentation**: Sales by outlet type, size, and region.
3. **Time Trends**: Sales trends based on outlet establishment year.

**9. Storytelling and Business Impact**

1. **Sales Drivers**: Fruits, snacks, and household items dominate sales.
2. **Performance by Outlet Type**: Supermarket Type 1 leads total sales.
3. **Outlet Size Insights**: Medium-sized outlets are the most profitable.

**Business Impact**

This analysis highlights where Blinkit should focus its efforts, e.g., investing more in Tier 1 locations or expanding medium-sized outlets.