**Customer Analysis Report**

**A Tableau Project Analysis for FY 2020-21**

**by: Kowsik.S**

**Table of Contents**

[1. Introduction](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#1-introduction)

[1.1 Project Overview](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#11-project-overview)

[2. Data Source Description](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#2-data-source-description)

[2.1 Key Fields](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#21-key-fields)

[2.2 Data Preparation](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#22-data-preparation)

[3. Fields and Measures](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#3-fields-and-measures)

[3.1 Dimensions](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#31-dimensions)

[3.2 Measures](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#32-measures)

[3.3 Binned Field](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#33-binned-field)

[4. Calculated Fields](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#4-calculated-fields)

[5. Visualizations: Worksheets](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#5-visualizations-worksheets)

[5.1 Worksheet 1: Age-Wise Sales Analysis](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#51-worksheet-1-age-wise-sales-analysis)

[5.2 Worksheet 2: Region-Wise Revenue Share (%)](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#52-worksheet-2-region-wise-revenue-share-)

[5.3 Worksheet 3: Month-Wise Revenue](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#53-worksheet-3-month-wise-revenue)

[5.4 Worksheet 4: Payment Method](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#54-worksheet-4-payment-method)

[5.5 Worksheet 5: Gender-Wise Sales Analysis](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#55-worksheet-5-gender-wise-sales-analysis)

[5.6 Worksheet 6: Quantity-Discount Correlation](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#56-worksheet-6-quantity-discount-correlation)

[5.7 Worksheet 7: Regionwise Qty and Total](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#57-worksheet-7-regionwise-qty-and-total)

[5.8 Worksheet 8: Revenue per State](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#58-worksheet-8-revenue-per-state)

[6. Visualizations: Dashboards](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#6-visualizations-dashboards)

[6.1 Dashboard 1: Regionwise Qty and Total with Payment Mode](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#61-dashboard-1-regionwise-qty-and-total-with-payment-mode)

[6.2 Dashboard 2: Customer Analysis](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#62-dashboard-2-customer-analysis)

[7. Key Insights](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#7-key-insights)

[8. Assumptions and Notes](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#8-assumptions-and-notes)

[8.1 Assumptions](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#81-assumptions)

[8.2 Limitations](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#82-limitations)

[8.3 Recommendations](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#83-recommendations)

[9. Conclusion](https://grok.com/chat/a6f8872e-6f3b-4d84-bdb2-6c06980baf29#9-conclusion)

**1. Introduction**

The *Customer Analysis*  project leverages Tableau to analyze sales data for the fiscal year 2020-21, focusing on customer demographics, regional performance, payment preferences, and purchasing patterns. This report provides a detailed examination of the data, visualizations, and insights derived from the Tableau worksheets and dashboards.

**1.1 Project Overview**

* **Title**: Customer Analysis
* **Purpose**: To analyze customer sales data for FY 2020-21, focusing on revenue distribution by gender, age groups, regions, payment methods, and other demographics.
* **Key Objectives**:
  + Understand revenue contributions by male and female customers.
  + Analyze customer age distribution and its impact on sales.
  + Evaluate regional sales performance and payment method preferences.
  + Identify correlations between discounts and order quantities.
* **Target Audience**: Business analysts, sales managers, and marketing teams.

**2. Data Source Description**

The dataset used in this project is a CSV file named sales\_06\_FY2020-21.csv, containing sales transaction data for the fiscal year 2020-21. It includes customer demographics, order details, and financial metrics.

**2.1 Key Fields**

* **Order Details**: Order ID, Order Date, Status, Item ID, SKU, Quantity Ordered, Price, Value, Discount Amount, Total.
* **Customer Demographics**: Customer ID, Year, Month, Reference Number, Name (Prefix, First Name, Middle Initial, Last Name, Full Name), Gender, Age, Email, Customer Since, SSN, Phone Number.
* **Location**: Place Name, County, City, State, ZIP, Region.
* **Payment**: Category, Payment Method.
* **Discount**: Discount Percent.

**2.2 Data Preparation**

Minimal cleaning was applied; the Age field was binned for analysis to create age groups for better segmentation.

**3. Fields and Measures**

**3.1 Dimensions**

* Order ID (Discrete, String): Unique identifier for each order.
* Order Date (Date): Date of the order.
* Status (String): Order status (e.g., completed, pending).
* Item ID (Discrete, String): Unique identifier for each item.
* SKU (String): Stock Keeping Unit for items.
* Category (String): Product category.
* Payment Method (String): Method of payment (e.g., credit card, cash).
* Customer ID (Discrete, String): Unique identifier for each customer.
* Year (Discrete, Number): Year of the transaction.
* Month (Discrete, Number): Month of the transaction.
* Reference Number (Discrete, String): Reference number for the transaction.
* Name Fields (String): Prefix, First Name, Middle Initial, Last Name, Full Name.
* Gender (String): Customer gender.
* Email (String): Customer email.
* Customer Since (Date): Date since the customer started transacting.
* SSN (String): Social Security Number.
* Phone Number (String): Customer phone number.
* Location Fields (String): Place Name, County, City, State, ZIP, Region.
* User Name (String): Likely the user who processed the order.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**3.2 Measures**

* Quantity Ordered (Continuous, Number): Number of items ordered.
* Price (Continuous, Number): Price per item.
* Value (Continuous, Number): Total value of the order before discounts.
* Discount Amount (Continuous, Number): Amount discounted.
* Total (Continuous, Number): Final order amount after discounts.
* Age (Continuous, Number): Customer age.
* Discount Percent (Continuous, Number): Percentage of discount applied.

**3.3 Binned Field**

* Age (bin): Automatically binned field based on Age for grouping customers into age ranges.

**4. Calculated Fields**

* **Zero Axis**:
  + **Formula**: Assumed to be a constant value of 0.
  + **Purpose**: Used as a reference line (value = 0) in charts for comparison.
* **Female Revenue**:
  + **Formula**: Assumed SUM(IF [Gender] = 'Female' THEN [Total] ELSE 0 END).
  + **Purpose**: Calculates the total revenue generated by female customers.
* **Male Revenue**:
  + **Formula**: Assumed SUM(IF [Gender] = 'Male' THEN [Total] ELSE 0 END).
  + **Purpose**: Calculates the total revenue generated by male customers.

A screenshot of a computer

AI-generated content may be incorrect.

**5. Visualizations: Worksheets**

This section describes the individual worksheets used in the Tableau project, each providing specific insights into the sales data.

**5.1 Worksheet 1: Age-Wise Sales Analysis**

* **Purpose**: Analyze sales distribution across age groups and gender.
* **Visualization**: Histogram with Age (bin) on the x-axis, SUM(Total) on the y-axis, and Gender as a color dimension (Female: orange, Male: blue).
* **Key Components**:
  + **Filters**: Category, Age (bin), Month, Region, Gender.
  + **Metrics**: Total sales (SUM(Total)).
* **Insights**:
  + Age group 30-40 has the highest sales for both genders, with females at $22.8M and males at $20.0M.
  + Males aged <20 contribute the least ($3.4M), while females in the same group contribute $3.6M.
  + Sales are relatively balanced across genders in most age groups.

A graph of blue and orange bars

AI-generated content may be incorrect.

**5.2 Worksheet 2: Region-Wise Revenue Share (%)**

* **Purpose**: Show the percentage of revenue contributed by each region.
* **Visualization**: Bubble chart with Region determining the bubble size and color, and SUM(Total) as the metric.
* **Key Components**:
  + **Filters**: Category, Age (bin), Month, Region, Gender.
  + **Metrics**: Total sales (SUM(Total)).
* **Insights**:
  + South contributes the highest revenue share (2,329,634), followed by Midwest (2,494,182), Northeast (1,097,443), and West (1,046,229).
  + The South and Midwest dominate the revenue distribution.

A diagram of a region

AI-generated content may be incorrect.

**5.3 Worksheet 3: Month-Wise Revenue**

* **Purpose**: Analyze revenue trends over months.
* **Visualization**: Histogram with Month on the x-axis and SUM(Total) on the y-axis.
* **Key Components**:
  + **Filters**: Category, Age (bin), Month, Region, Gender.
  + **Metrics**: Total sales (SUM(Total)).
* **Insights**:
  + August 2021 has the highest revenue at $2.0M.
  + November 2020 and October 2020 have the lowest revenue at $0.1M each.
  + Revenue peaks in mid-2021, with a noticeable dip in early and late months.

A screenshot of a graph

AI-generated content may be incorrect.

**5.4 Worksheet 4: Payment Method**

* **Purpose**: Evaluate the distribution of sales across different payment methods.
* **Visualization**: Histogram with Payment Method on the x-axis and SUM(Total) on the y-axis, colored by payment method.
* **Key Components**:
  + **Filters**: Payment Method, Category, Gender, State, Age (bin).
  + **Metrics**: Total sales (SUM(Total)).
* **Insights**:
  + customercredit is the most used payment method, contributing around 55M in sales.
  + mcblite and jazzvoucher have the least usage, contributing less than 5M each.
  + Common payment methods include bankalfalah, cashatdoorstep, and cod.

A graph with different colored bars

AI-generated content may be incorrect.

**5.5 Worksheet 5: Gender-Wise Sales Analysis**

* **Purpose**: Compare sales between genders across product categories.
* **Visualization**: Histogram with Category on the x-axis, Measure Values (Female Revenue, Male Revenue) on the y-axis, and Measure Names as a color dimension.
* **Key Components**:
  + **Filters**: Gender, Category, Age (bin), Month, Region, Measure Names.
  + **Metrics**: Female Revenue, Male Revenue.
* **Insights**:
  + Tablets are the top category for both genders, with females contributing ~250K and males ~220K.
  + Categories like Soghaat, Superstore, and Others have low sales for both genders.
  + Sales are generally balanced between genders across categories.

A screenshot of a graph

AI-generated content may be incorrect.

**5.6 Worksheet 6: Quantity-Discount Correlation**

* **Purpose**: Examine the relationship between discount percentage and quantity ordered.
* **Visualization**: Scatter plot with Discount Percent on the x-axis, Quantity Ordered on the y-axis, and bubble size representing SUM(Discount Percent).
* **Key Components**:
  + **Filters**: Discount Percent, Category, Age (bin), Month, Region, Gender.
  + **Metrics**: Quantity Ordered, Discount Percent.
* **Insights**:
  + Most orders have a discount percent between 0-20%, with quantities ordered typically below 10.
  + A few outliers show high quantities (e.g., 171) with discounts around 60%.
  + No strong correlation is evident; higher discounts don’t consistently lead to higher quantities ordered.

A screenshot of a computer

AI-generated content may be incorrect.

**5.7 Worksheet 7: Regionwise Qty and Total**

* **Purpose**: Compare quantity ordered and total sales across states, grouped by region.
* **Visualization**: Dual-axis bar chart with State on the x-axis, SUM(Qty Ordered) on one y-axis, and SUM(Total) on the other y-axis, colored by state.
* **Key Components**:
  + **Filters**: Region, State, Payment Method, Category, Gender, Age (bin).
  + **Metrics**: Quantity Ordered, Total sales (SUM(Total)).
* **Insights**:
  + Colorado (CO) has the highest quantity ordered (49,382) and total sales ($17.0M).
  + States like New Hampshire (NH) and Rhode Island (RI) have the lowest quantities and sales.
  + The South region (e.g., CO, TX) shows higher overall sales compared to other regions.

A screenshot of a graph

AI-generated content may be incorrect.

**5.8 Worksheet 8: Revenue per State**

* **Purpose**: Visualize revenue distribution across states in the U.S.
* **Visualization**: Map with states colored by SUM(Total), using Latitude and Longitude (generated).
* **Key Components**:
  + **Filters**: Region, City, Category, Age (bin), Month, Gender.
  + **Metrics**: Total sales (SUM(Total)).
* **Insights**:
  + Alaska (AK) and Oregon (OR) have the highest revenue at $0.0M each (likely a data error, as the map shows high values).
  + Most states have revenues between $0.1M and $0.3M.
  + The map highlights regional clusters of revenue, with the South and Midwest showing higher concentrations.

A map of the united states

AI-generated content may be incorrect.

**6. Visualizations: Dashboards**

This section describes the dashboards that combine multiple worksheets to provide a holistic view of the data.

**6.1 Dashboard 1: Region wise Qty and Total with Payment Mode**

* **Objective**: Provide a comparative view of quantities ordered and total sales by state, alongside payment method distribution.
* **Layout**:
  + **Left Section**: "Region wise Qty and Total" worksheet (dual-axis bar chart) showing quantity ordered and total sales across states.
  + **Right Section**: "Payment Mode" worksheet (histogram) showing sales distribution by payment method.
* **Interactivity**:
  + **Filters**: Region, State, Payment Method, Category, Gender.
  + Users can select specific regions or states to drill down into the data.
  + Payment method filter allows focusing on specific payment types.
* **Key Insights**:
  + Colorado (CO) leads in both quantity ordered (49,382) and total sales ($17.0M), indicating a high sales volume in the South region.
  + customercredit is the dominant payment method (55M in sales), suggesting a preference for credit-based transactions.
  + States like NH and RI show minimal activity, which may indicate low market penetration in those areas.

A screenshot of a graph

AI-generated content may be incorrect.

**6.2 Dashboard 2: Customer Analysis**

* **Objective**: Provide a comprehensive overview of customer sales data, focusing on temporal, regional, demographic, and discount trends.
* **Layout**:
  + **Top Left**: "Month-Wise Revenue" worksheet (histogram) showing revenue trends over months.
  + **Top Center**: "Revenue per State" worksheet (map) showing state-wise revenue distribution.
  + **Top Right**: "Age-Wise Sales Analysis" worksheet (histogram) showing sales by age group and gender.
  + **Bottom Left**: "Region-Wise Revenue Share (%)" worksheet (bubble chart) showing regional revenue contributions.
  + **Bottom Center**: "Gender-Wise Sales Analysis" worksheet (histogram) comparing sales by gender across categories.
  + **Bottom Right**: "Quantity-Discount Correlation" worksheet (scatter plot) analyzing the relationship between discounts and quantities ordered.
  + **Filters**: Region, Category, Month, Age (bin), displayed at the top for user interaction.
  + **Total Display**: A text box at the top right showing the total revenue ($23,36,50,194.432088).
* **Interactivity**:
  + **Filters**: Region, Category, Month, Age (bin).
  + Users can filter data by region, category, month, or age group to explore specific segments.
  + The map allows zooming into specific states for detailed analysis.
* **Key Insights**:
  + August 2021 shows a peak in revenue ($2.0M), indicating a potential seasonal trend.
  + The South region contributes the most to revenue (2,329,634), as seen in the bubble chart.
  + Tablets are the top-selling category for both genders, with females slightly outperforming males (~250K vs. ~220K).
  + Age group 30-40 is the most significant contributor to sales ($22.8M for females, $20.0M for males).
  + The scatter plot shows no strong correlation between discounts and quantities ordered, suggesting discounts may not be a primary driver for higher order volumes.
  + The map indicates potential data errors (e.g., $0.0M for AK and OR despite high visual intensity), which should be investigated.

A screenshot of a computer

AI-generated content may be incorrect.

**7. Key Insights**

* **Revenue Trends**: Sales peak in August 2021 ($2.0M), with lower activity in November and October 2020 ($0.1M each), suggesting seasonal patterns.
* **Demographic Insights**: The 30-40 age group is the most significant contributor to sales, with balanced gender contributions. Tablets are the top product category for both genders.
* **Regional Performance**: The South and Midwest regions dominate revenue, with Colorado leading in both quantity ordered and total sales.
* **Payment Preferences**: customercredit is the preferred payment method (55M in sales), indicating a reliance on credit transactions.
* **Discount Impact**: No strong correlation exists between discount percentages and quantities ordered, suggesting other factors may drive purchasing behavior.
* **Data Quality**: Unrealistically high revenue values in calculated fields (e.g., $540T for Female Revenue) and discrepancies in the map (e.g., $0.0M for AK) indicate potential data scaling or calculation errors.

**8. Assumptions and Notes**

**8.1 Assumptions**

* The Gender field contains values like "Male" and "Female" for the revenue calculations.
* The Total field represents the final revenue per order.
* No significant data transformations were applied beyond binning the Age field.

**8.2 Limitations**

* Revenue values in calculated fields are unrealistically high, suggesting a data issue (e.g., incorrect units, such as cents instead of dollars, or a calculation error in Tableau).
* The map shows discrepancies (e.g., $0.0M for AK despite high visual intensity), indicating potential data errors.
* Calculated field formulas were not provided, so assumptions were made based on field names and context.

**9. Conclusion**

The *Customer Analysis*  project provides valuable insights into sales performance for FY 2020-21. Key findings include the dominance of the 30-40 age group in sales, the South region's significant revenue contribution, and the preference for credit-based payment methods. However, data quality issues, such as unrealistic revenue values and map discrepancies, need to be addressed to ensure the reliability of the analysis. Future iterations of this project should focus on data validation and enhanced interactivity to better serve business decision-making.