OPTIMIZING ORDER LIFECYCLE AND DEMAND FORECASTING DASHBOARD

PROJECT GOALS

Our team acquired real-time data from Knowde, mapping all the zipcodes in the US, publicly available on the internet and merged it with academically accessible Order to Cash data. Using these three resources, we constructed a comprehensive dataset comprising invoice details, customer characteristics, delivery specifics, and product sales information. This dataset allowed us to delve into consumer behavior and demand trends, and provided us with the opportunity to identify areas where the company can improve its revenue streams. A common challenge for businesses is determining what areas to expand the company and which regions to increase resource allocation. We propose the use of a dashboard as a powerful tool to reveal notable patterns or issues within the company's demand. The interactiveness of a dashboard can prove highly valuable in pinpointing variations across regions, segments, and temporal intervals.

The primary goal of our dashboard is to provide demand planning executives, a tool to effectively visualize the demand and sales trends. We created three dashboards, one with a focus on customer analytics, segmentation, and demand trends, the latter is useful for product preferences. Lastly, it would also be helpful in goal setting and demand forecasting. The demand planning team would benefit from real-time demand data, which would enable them to optimize manufacturing schedules, resource allocation, and predict inventory levels based on expected demand. The executive positions would find the demand and sales trends useful for monitoring company performance. Additionally, the forecasting visualizations can provide executives with valuable insights for investing, risk, and market expansion decisions.

DATA SOURCES AND PREP

Data source 1 - Postgresql (Academic dataset- WestCoastImporters)

- 127.0.0.1 as the hostname
- “jovyan” as the PostgreSQL user
- “postgres” as the password
- 8765 as the port number
- Standard connection
- “WestCoastImporters” as the database
- “WestCoastImporters” as the display name

Customer: Contains customer information with attributes such as CustomerID, CustomerName, defaultbilltocustomerid, primarycontactpersonid etc.
CustomerCategoryMembership: Shows the association of customers with different categories, consisting of columns customerid, customercategoryid.
CustomerCategory: Contains customer category names which are the public stores where our products are sold. To name a few, Computer Store, Gift shops etc. Each of these categories have a unique categoryID.
CustomerDeal: Relates customers to specific deals ranging from 1-4 based on their purchase history and provides deal related discounts. Contains columns customerid, dealid.
SalesOrderHeader: Captures the main details of sales orders, such as Orderid, OrderDate, and links to customers, invoices, and locations.
SalesOrderLine: Details individual line items in sales orders, including StockItemid and pricing information.
InvoiceHeader: Contains invoice-level data such as InvoiceDate and links to locations and customers.
InvoiceLine: Records line-item details for invoices, connected to orders and stock items.

Data source 2 - Product Name:
The project involved scraping product names from the Knowde website specifically under the 'Appliances & Electronics' category, focusing on the 'Appliances' sub-category. These product names were then systematically

categorized and used to assist in estimating demand planning, providing valuable insights for inventory management.

Data source 3 - <u>U.S. Location Data Refinement:</u>
The 'location_us' table, formatted as CSV, includes enriched data with fields for City, State, StateID, County, Latitude, and Longitude. Enhancements include the addition of 'LocationID' as a primary key and linkage of 'CustomerID' as a foreign key for improved relational data structure. Excel's VLOOKUP function was utilized to populate County Code, Latitude, and Longitude. Data was then filtered to prioritize locations central to business operations, with a strategic emphasis on the West Coast region, reflecting the dense concentration of business activity and customer presence.
https://gist.github.com/Tucker-Eric/6a1a6b164726f21bb699623b06591389

<u>Major steps in Tableau Prep:</u>

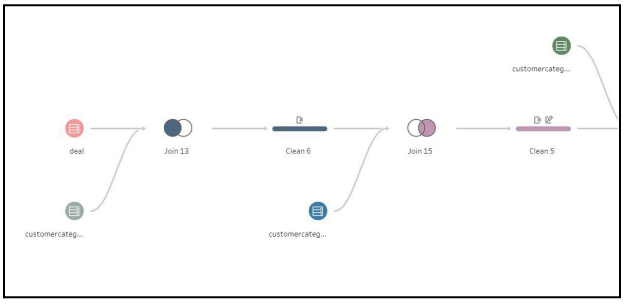
Data Sources and Joins: The workflow involves various data sources, labeled with terms like "deal", "customer category", "customer", "sales order line", "invoice line", "stock name", which are being joined together. The joining process is represented by nodes labeled "Join" followed by a number, indicating different join operations.

Cleaning Steps: There are multiple cleaning steps in the process which are meant to prepare the data for analysis in Tableau. These steps include:

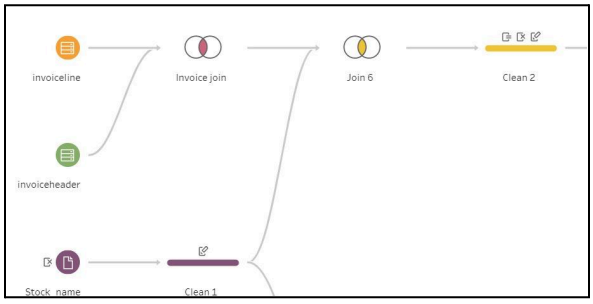
Data Type Changes: Converting the data types of certain fields to a format that is more suitable for the intended analysis or processing steps.

Trimming: Removing leading and trailing whitespace from strings in the data which can often be an artifact of data entry or extraction processes. Specifically from the "customername" field, which involves taking out unnecessary details that were included within brackets—such as location and other irrelevant information.

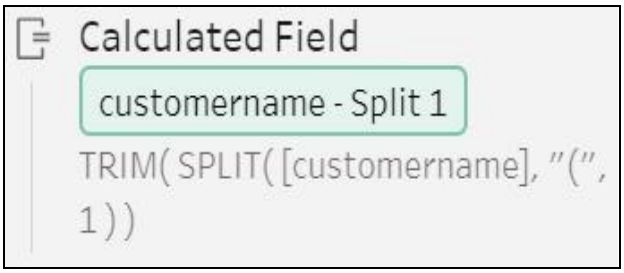
Duplicates: There are steps indicated as "Clean 6", "Clean 5", "Clean 7", which involve removing or consolidating duplicate records to ensure that each piece of data is unique within the dataset.



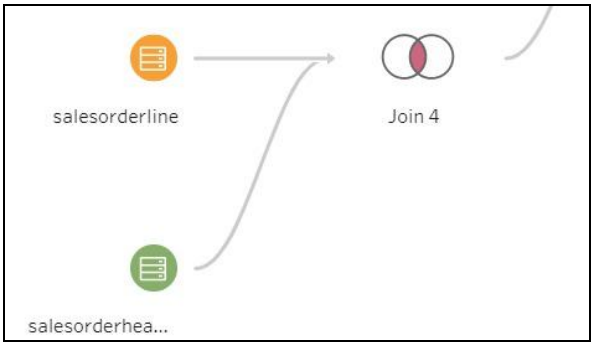
Integrating invoice data with product names (obtained through Knowde website)



Integrating customer and deal information

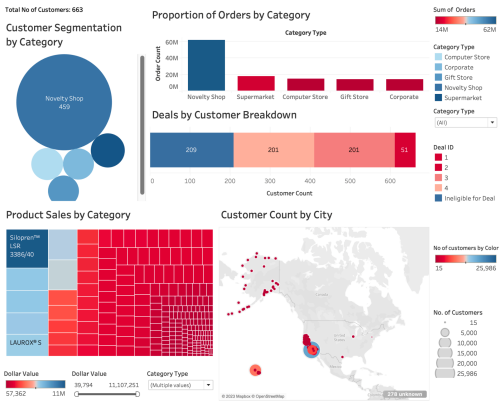


Function to drop unnecessary information from the customer names and cleaning the data



Integrating Sales and Demand data

DASHBOARD INTERACTIONS AND FINDINGS



Our first dashboard offers a detailed view of consumer analytics and demand, featuring interactive visualizations that highlight key consumer trends. The visualizations use a blue-to-red gradient to indicate performance, with red highlighting areas needing improvement.

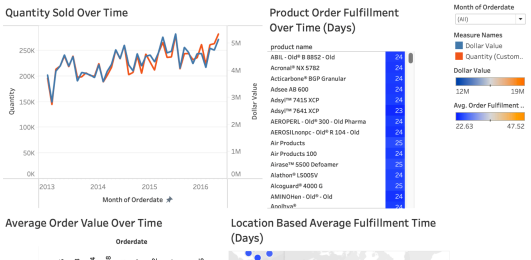
Visualization 1 utilizes a packed bubble graph to compare unique customer counts across categories. It reveals the Novelty Shop category as the most popular, with significantly more unique customers than others.

In Visualization 2, a geomap shows customer distribution by city, aiding in resource allocation and marketing. Key findings include a high concentration of consumers in Honolulu and Anaheim and notable customer bases in Standard, IL, Sugarloaf, TN, and Weimar, TX. These insights suggest new market opportunities, especially in Standard and Weimar.

Visualization 3, a bar graph, illustrates order count by category, guiding inventory and production planning. This graph interacts as a filter for the other visualizations. This interaction is extremely helpful in revealing the regions and products most in demand for specified categories.

Visualization 4, a stacked bar chart, breaks down deals by customer segment and product category, informing targeted marketing and product offerings. The interaction from order count by category allows us to see which deals are eligible in distinct customer categories. We found corporate customers, Tailspin Toys and Wingtip Toys, as primary deal participants.

Lastly, we opted to use the product sales treemap, which offers the flexibility to filter by category, enabling us to pinpoint the products with the highest demand. This actionable insight guides product development and discontinuation decisions, ensuring our product portfolio aligns with customer preferences. From this plot we discovered that the Silopren product is the best performing product id across all customer categories. We implemented a dollar value slider, to allow the user to visualize products only pertaining to a specified sales range.



The second dashboard we prepared is useful for analyzing company demand and performance. We opted to

use purple and orange gradients to cue where values are high or low. Visualization 1, tracks demand trends spanning four years by dollar value sold and number of customers over time. This visualization highlights areas with peaks and flat slopes, allowing us to discern instances where the dollar value increased while the number of customers remained constant. This feature allowed us to identify that during the year 2015 sales per customer was greater compared to the other years.

Visualization 2 tracks average order values over time, helping us monitor changes in customer spending habits. While demand fluctuates, our analysis reveals stable average order values. This interactive visualization enables users to explore monthly fluctuations by clicking on specific months. With this feature, the user is able to see which months they tend to get the highest demand and are most efficient in fulfilling orders. Visualization 3 presents a scrollable table that provides insights into order fulfillment times by product. It highlights average fulfillment durations for different products, aiding in the optimization of our fulfillment processes and enhancing customer satisfaction. Our team discovered that there were eight products severely underperforming in delivery time. To investigate if order fulfillment times vary by location, we created a fourth geomap visualization pinpointing cities with longer delivery times. This information guided us in addressing logistical challenges that may be taking place in California cities near the border of Arizona and Nevada.

Dashboard 3 is dedicated to demand forecasting, where we have employed a stacked line chart approach to present actual demand versus forecasted demand. This design choice allows for a clear and intuitive comparison of historical data with our predictions. To enhance the usability of this visualization, we have incorporated a tooltip feature that provides executive personnel and the sales team with the ability to hover over data points, revealing the number of orders placed by category on both the actual demand and forecasted demand lines. This feature permits decision-makers to view demand dynamics, aiding more informed and strategic decision-making processes.

CHANGES FROM THE DESIGN DOCUMENT

1) Data Integration (Identifying disparate data sources): Our project has successfully consolidated multiple postgreSQL and .CSV data files—such as customer, invoice, locations, and forecasting—each with unique structures and information types. We have integrated these various datasets to create a cohesive, unified view of the data and established relationships between these datasets.

2) End-User Specification: The project's design document initially identified a broad user base. This has been refined to specifically empower the demand planning team, to leverage the dashboard for resource allocation, monitoring, trend analysis, and strategic decision-making.

3) Customer Centric or Business Focused?: The project's focus is on Businesses. Dashboards highlighting Product Sales by category, customer segmentation, sales volume, and fulfillment metrics provide actionable insights for strategic growth, inventory control, customer engagement, and supply chain efficiency, ultimately driving marketing, sales forecasting, and customer satisfaction initiatives.

4) Effectiveness of Visualization: The treemap for Product Sales by Category employs size and color variations to represent sales value in dollars, allowing for immediate visual discernment of high-performing products and enabling quick identification of sales trends critical for product strategy optimization. Line chart illustrates progression of quantity and sales over time in our dashboard. The movement and direction makes it easier to spot upward or downward trends. With the use of dual axes for Demand Over Time, the line chart can effectively compare two different scales within the same context. The bar chart used in Proportion of Orders by Category chart is straightforward and effective for a categorical comparison. The length of each bar provides a clear visual indicator of the quantity, making it easy to see the differences in order counts between categories.

5) Interactions Within the Dashboard: Our product demand analytics dashboard features four key visualizations: Quantity Sold and Sales Over Time, Product Order Fulfillment Over Time, Average Order Value

Over Time, and Location-Based Average Fulfillment Time. These visualizations are crucial for discerning high-demand products based on metrics such as purchase frequency, sales value, location, and logistics efficiency. These factors also contribute to total sales.
For instance, when a user selects a product from the dashboard, the visualizations dynamically update to display that product's sales trends, including monthly sales from the previous year and the product's average demand and fulfillment timelines. This interactive functionality has been consistently implemented across all three dashboards, enhancing user engagement and data accessibility.

6) Forecasting Visualizations: The forecasting visualizations are an additional effort from our team, aimed at delivering considerable value to our targeted users by providing insights into investment strategies, risk management, and market expansion planning. This is achieved by overlaying actual sales onto historical demand and comparing these figures with projected future demand—information that is essential for strategic planning and resource allocation in business operations. These visualizations enable a visual assessment of the accuracy of demand forecasts compared to real data, allowing for necessary adjustments to forecasts and business strategies.