Project 1

IKEA

By

Mr.	Yannaphat	Romphoyen	6488070
Mr.	Techin	Wangkart	6488102
Mr.	Withsarut	Wutthisathian	6488113
Miss	Ponnapassorn	Iamborisut	6488179
Miss	Nattanicha	Sinsawet	6488190
Miss	Ravikarn	Jarungjitvittawas	6488210

A Report Submitted in Partial Fulfillment of
the Requirements for

ITCS453 Data Warehousing and Data Mining

Faculty of Information and Communication Technology
Mahidol University

2023

Executive Summary

IKEA was founded in the year 1943 and has become a globally renowned furniture store known for its Scandinavian style, high quality, and affordable prices. IKEA's efficient business operations are planned using OLAP (Online Analytical Processing) and OLTP (Online Transaction Processing) systems strategically to enhance daily transaction management, such as sales, inventory management, and extensive data analysis for business systems. This comprehensive report delves into the use of OLAP and OLTP for IKEA, focusing on systems that support IKEA and the utilization of Tableau for system operations.

OLAP and OLTP systems play a crucial role in enhancing the efficiency of IKEA by managing and analyzing transaction and business data efficiently using Tableau for clear and up-to-date data visualization. IKEA uses an OLTP system that effectively handles inventory management, sales transactions, and customer contacts. IKEA can analyze vast amounts of data, comprehend market patterns, and make detailed predictions about client needs thanks to the OLAP system. IKEA can create clear reports and dashboards with Tableau, which helps with strategic planning and improved decision-making. Customer service has dramatically improved thanks to easy access to critical information and adaptable help. By improving data analysis and data-driven decision-making, Tableau's integration with OLAP and OLTP systems also assists IKEA in cutting expenses and raising revenues. Because of improved data analysis and data-driven decision-making, IKEA is better equipped to adapt to and sustain its leadership in the quickly evolving retail sector.

In summary, OLAP and OLTP systems are instrumental in IKEA's success. Efficient data analysis, strategic planning, rapid decision-making, and using Tableau have prepared IKEA to adapt and analyze data to develop furniture that meets customer needs. This integration also helps IKEA maintain its leadership through data-driven decision-making.

Table of Contents

Project Overview

Overview

IKEA is a worldwide store that sells furniture and other things for the home. The company was started by Ingvar Kamprad in 1943 and has grown into a global business since then. IKEA's business plan is based on selling high-quality items at low prices without sacrificing style or usefulness [1]. The company's success is due to its expert use of market segmentation, which has won praise for its well-designed furniture that exudes the unique charm of Scandinavian style at prices that most people can afford.

The company's value chain includes product development, design, supply, manufacture, and sales, focusing on putting people first [2]. IKEA's digital transformation has been a significant aspect of its growth, with the company investing in IT infrastructure to support its online presence and digital assets. In addition to its retail operations, IKEA is committed to sustainability and social responsibility. The company aims to become circular and climate-positive, regenerate resources, and contribute to a fair and equal society. IKEA's business domain, therefore, is multifaceted, encompassing retail operations, design and innovation, sustainability, supply chain management, digital transformation, market expansion, corporate strategy, and employee relations [3].

Objectives

1. Utilize IKEA data to create a dimensional model such as a star schema or relational model, focusing on the inventory management system.
2. Design a comprehensive model and outline essential steps for the development of the IKEA system.
3. Analyze the designed model, enumerate findings, and establish a systematic process for system development.

Scope of the Project

1. Examining the sources of data that IKEA utilizes for its operations, including information on sales, inventory, and customer interactions.
2. Exploring the design of IKEA's data warehouse, which serves as the foundation for effective data management and analysis.

3. Investigating the Extract, Transform, Load (ETL) process employed by IKEA to ensure the accuracy and relevance of data within its systems.
4. Delving into how OLAP and OLTP systems, in conjunction with Tableau, facilitate data analysis and visualization, aiding in strategic planning and decision-making.

Business requirements for building a data warehouse for BI

1. Global Supply Chain Management:

- Data integration at the manufacturing, logistics, and inventory levels from the whole supply chain [4].
- monitoring the current stock levels in various locations.
- Analyze lead times and supplier performance to enhance procurement procedures.

2. Product Catalog Management:

- Managing a large assortment of products.
- Utilizing sales data for specific items and categories to support additional marketing initiatives and product design.
- Observing shifts in the life cycle of products and fluctuations in seasonal demand.

3. Customer Experience and Retail Analytics:

- Gathering information on customers via various channels of communication, including the company's mobile apps, in-store locations, and online.
- Recognizing the tastes, purchasing patterns, and feedback of customers.
- Analyze the success of marketing initiatives and promotions.

4. Sustainability and Corporate Responsibility:

- keeping an eye on sustainability metrics in reporting systems, such as waste management, carbon footprint, or energy usage
- Ensuring that data analysis is in line with IKEA's commitment to social responsibility and sustainability.

5. Financial Reporting and Analysis:

- Combining global financial sales data for research purposes, taking into account expenses and revenue
- Budgeting; forecasting; financial performance

6. Employee and Workforce Analytics:

- Organizing personnel information across business locations.
- Analysis of the distribution of human resources, training requirements, and worker productivity.

7. Legal and Compliance Requirements:

- Making sure that the data warehouse conforms to international data protection laws, especially in view of IKEA's global presence.
- Integrating data lineage, compliance reporting, and auditing features.

8. Scalability and Flexibility:

- The ability to scale up as IKEA continues to grow and enter new markets.
- Adaptability to new data sources and changing business requirements.

9. User Accessibility and Self-Service BI:

- supplying intuitive user interfaces for a variety of users, including executives and analysts.
- Enabling non-technical people to create reports and insights using self-service BI capabilities.

10. Advanced Analytics and AI Integration:

- Utilizing sophisticated analytics to provide prescriptive and predictive insights.
- AI and machine learning applications for automated trend analysis and decision-making assistance

How Data Warehouse important to the company

A data warehouse is crucial to IKEA for several reasons, as it helps in organizing, analyzing, and leveraging data for various aspects of its operations.

1) **Data Integration**

IKEA, being a global company, likely has data distributed across various departments, regions, and systems. A data warehouse allows for the integration of data from different sources into a centralized repository. This integration enables a holistic view of the company's operations, providing valuable insights and facilitating better decision-making.

2) **Business Intelligence**

A data warehouse serves as the foundation for business intelligence (BI) tools and analytics. IKEA can use these tools to analyze historical and current data trends, customer behavior, and market dynamics. This information is valuable for making informed strategic decisions, optimizing product offerings, and identifying areas for improvement.

3) **Supply Chain Management**

In the retail industry, efficient supply chain management is crucial. A data warehouse helps IKEA optimize its supply chain by providing real-time visibility into inventory levels, demand forecasting, and logistics. This leads to better inventory management, reduced costs, and improved overall operational efficiency.

4) **Customer Insights**

Understanding customer preferences and behavior is vital for a retailer like IKEA. A data warehouse allows the company to consolidate and analyze customer data, such as purchase history, preferences, and feedback. This information can be used to personalize marketing strategies, improve customer experience, and tailor product offerings to meet customer demands.

5) **Performance Monitoring**

IKEA can use a data warehouse to monitor the performance of various aspects of its business, including sales, marketing campaigns, and employee productivity. This enables the

company to identify trends, track key performance indicators (KPIs), and respond quickly to changing market conditions.

6) Digital Transformation

As mentioned in the overview, IKEA has invested in digital transformation. A data warehouse supports this initiative by providing a solid foundation for managing and analyzing digital data. This is crucial for e-commerce operations, online customer interactions, and other digital aspects of the business.

7) Compliance and Reporting

For a global company, adhering to various regulations and reporting standards is essential. A data warehouse facilitates the generation of accurate and timely reports, ensuring compliance with legal requirements and providing transparency in financial and operational reporting.

How BI important to the company

1) **Data-Driven Decision Making**

With the help of BI, IKEA can make data-driven decisions instead of relying solely on gut feeling or speculation. IKEA can increase productivity, streamline processes, and spot growth prospects by evaluating data on sales, customer behavior, inventory levels, and industry trends.

2) **Understanding Customer Preferences**

IKEA can learn more about the demographics, purchasing habits, and tastes of its customers thanks to BI. With the use of this data, IKEA is better able to customize its product lineup, advertising strategies, and store designs to better suit the wants and needs of its patrons.

3) **Optimizing Supply Chain Management**

By giving visibility into inventory levels, production schedules, supplier performance, and transportation logistics, business intelligence (BI) assists IKEA in optimizing its supply chain. As a result, IKEA may lower expenses, cut down on stockouts, and enhance the effectiveness of its supply chain as a whole.

4) **Monitoring Performance**

IKEA can track key performance indicators (KPIs) in a number of business domains, including sales, profitability, customer happiness, and worker efficiency, thanks to business intelligence (BI). Through real-time KPI tracking, IKEA is able to recognise patterns, anticipate any problems, and take necessary remedial action.

5) **Competitive Advantage**

IKEA may get a competitive edge in the cutthroat retail market of today by having access to fast and reliable business analytics. IKEA can outperform rivals and react swiftly to market changes by using data to analyze customer behavior and market dynamics.

6) **Product Development and Innovation**

IKEA's product development and innovation initiatives may be informed by the significant information that business intelligence (BI) can give regarding market trends, consumer feedback, and rival products. IKEA may prioritize research and development efforts by using BI data analysis

to spot new trends, comprehend consumer preferences for new items, and discover developing patterns.

7) **Personalized Marketing and Customer Experience**

With the help of BI, IKEA can divide up its clientele and tailor ads and promotions to each individual based on things like online activity, past purchases, and demographics. With a more focused strategy, IKEA can raise conversion rates, boost consumer engagement, and improve the whole shopping experience.

8) Store Performance and Merchandising

IKEA can evaluate each store's performance and improve its merchandising tactics with the use of BI. IKEA may find ways to enhance the entire shopping experience, optimise inventory levels, and improve product placement by examining sales statistics, foot traffic patterns, and the efficacy of the store layout.

9) Risk Management and Fraud Detection

For IKEA, BI may be extremely important for risk management and fraud detection. IKEA can detect possible risks and irregularities early on by evaluating data pertaining to transactions, financial performance, and security issues. This enables them to take proactive steps to reduce risks and safeguard the company's assets.

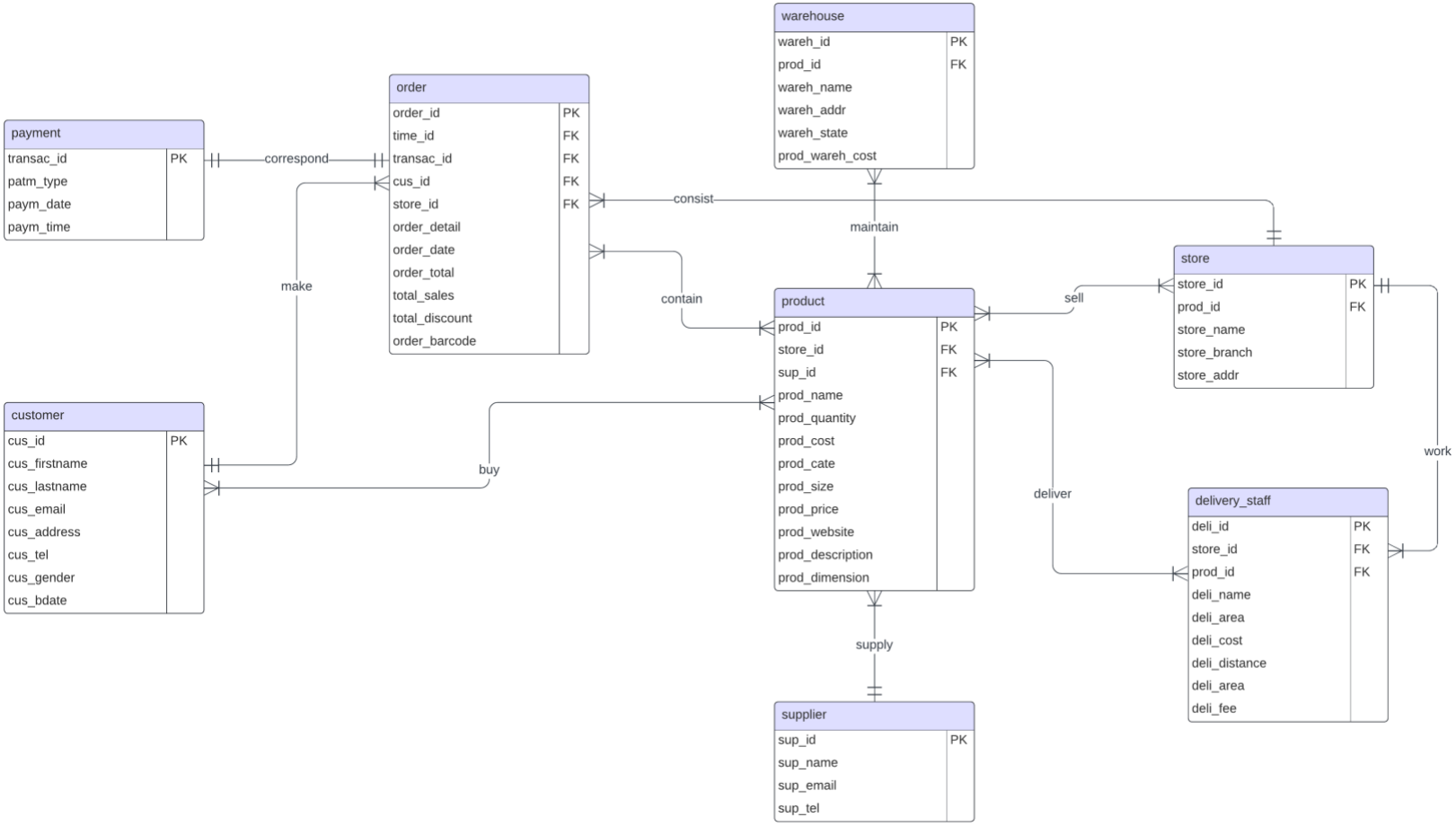
10) Global Expansion and Market Entry Strategies

BI is critical to the assessment of market prospects, competitive landscapes, and worldwide expansion plans of a multinational corporation such as IKEA. IKEA can maximize performance in new territories by customizing its product offers to local tastes, identifying viable markets for expansion, and optimizing its go-to-market strategies by utilizing BI data.

Data sources

1. **ER Diagram**

The diagram below shows the relationship between entities of the IKEA company.



This ERD seems to depict various entities involved in a retail or ecommerce business and the relationships between them. The entities include:

- order: Records details of each order, linked to a time dimension, transaction, customer, and store.
- payment: Contains information about payments, linked to specific transactions.
- customer: Stores customer information.
- warehouse: Holds details of warehouses, linked to products.
- product: Contains details about products, linked to suppliers, stores, and warehouses.
- delivery_staff: Details about delivery personnel, linked to stores and products.
- store: Contains store details, linked to products.
- supplier: Holds supplier information, linked to products.

Primary keys (PK) indicate unique identifiers for each entity, and foreign keys (FK) represent links between related entities. The ERD shows how these entities interact, such as customers making orders, orders containing products, products being supplied by suppliers, and products stored in warehouses. The relationships also indicate that payment details correspond to orders, orders involve delivery staff, and products are contained within stores and warehouses.

2. Data Dictionary

Customer

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
customer	cus_id	Customer's id	varchar(5)	Cxxxx			PK	
	cus_firstname	Customer's firstname	varchar(30)	xxxxxx				
	cus_lastname	Customer's lastname	varchar(30)	xxxxxx				
	cus_email	Customer's email	varchar(30)	xxxxxx@xxxxx				
	cus_address	Customer's address	varchar(100)	xxxx xxx xx xx, xx xxxxx				
	cus_tel	Customer's telephone number	varchar(10)	(xxx) xxx-xxxx				
	cus_gender	Customer's gender	varchar(10)	xxxxxxx	Female,Male, Other			
	cus_bdate	Customer's birth date	date	D/M/YYYY				

Location

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
location	loca_postal	Location's postal code of customer	varchar(5)	xxxxx			PK	
	loca_state	Location's state of customer	varchar(50)	xxxxx (xx)				
	loca_city	Location's city of customer	varchar(50)	xxxxxx				

Supplier

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
supplier	sup_id	Supplier's id	varchar(5)	SPExxxx			PK	
	sup_name	Supplier's name	varchar(60)	xxxxxx				
	sup_email	Supplier's email	varchar(30)	xxxxxx@xxxxx				
	sup_tel	Supplier's telephone number	varchar(10)	(xxx) xxx-xxxx				

Order

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
order	order_id	Order's id	varchar(10)	Orxxxx			PK	
	time_id	Time's id	varchar(6)	TM0001			FK	time[time_id]
	transac_id	Payment transaction id	varchar(6)	TSxxxx			FK	payment[transac_id]
	cus_id	Customer's id	varchar(5)	Cxxxx			FK	customer[cus_id]
	store_id	Store's id	varchar(10)	Sxxxx			FK	store[store_id]
	order_detail	Order's detail	varchar(500)	xxxxx				
	order_date	Order's date	date	D/M/YYYY				
	order_total	Order's total	int	xxxxxx				
	total_sales	Order's total sales	int	xxxxxx				
	total_discount	Order's total discount	int	xxxxxx				
	order_barcode	Order's barcode	varchar(13)	xxxxxxxxxxxxx				

Warehouse

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
warehouse	wareh_id	Warehouse's id	varchar(6)	WH0001			PK	
	prod_id	Product's id	varchar(6)	PDxxxx			FK	product[prod_id]
	wareh_name	Warehouse's name	varchar(5)	xxxxx Warehouse				
	wareh_addr	Warehouse's address	varchar(30)	xxxxxx				
	total_qty	Warehouse's total quantity	varchar(5)	xxxxx Warehouse				
	total_qty_sold	Warehouse's total quantity sold	varchar(30)	xxxxxx				
	wareh_state	Warehouse's state	varchar(30)	xxxxxx (xx)				

Store

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
store	store_id	Store's id	varchar(10)	Sxxxx			PK	
	prod_id	Product's id	varchar(6)	PDxxxx			FK	product[prod_id]
	store_name	Store's name	varchar(4)	IKEA				
	store_addr	Store's address	varchar(13)	United States				
	store_branch	Store's branch	varchar(30)	xxx				

Payment

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
payment	transac_id	Payment transaction id	varchar(6)	TSxxxx			PK	
	paym_type	Payment method' type	varchar(30)	xxxxxx	Cash,Debit Card,Credit Card			
	paym_date	Paymetn method' date	date	D/M/YYYY				
	paym_time	Payment method' description	time	HH:MM:SS				

Store

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
store	store_id	Store's id	varchar(10)	Sxxxx			PK	
	prod_id	Product's id	varchar(6)	PDxxxx			FK	product[prod_id]
	store_name	Store's name	varchar(4)	IKEA				
	store_addr	Store's address	varchar(13)	United States				
	store_branch	Store's branch	varchar(30)	xxx				

Payment

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
payment	transac_id	Payment transaction id	varchar(6)	TSxxxx			PK	
	paym_type	Payment method' type	varchar(30)	xxxxxx	Cash,Debit Card,Credit Card			
	paym_date	Paymetn method' date	date	D/M/YYYY				
	paym_time	Payment method' description	time	HH:MM:SS				

Time

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
time	time_id	Time's id	varchar(6)	TM0001			PK	
	wareh_id	Warehouse's id	varchar(6)	WHxxxx			FK	warehouse[wareh_id]
	prod_id	Product's id	varchar(6)	PDxxxx			FK	product[prod_id]
	sup_id	Supplier's id	varchar(6)	SPExxxx			FK	supplier[sup_id}
	time_date	Time's date	int	xx				
	time_month	Time's month	int	xx				
	time_quater	Time's quarter	varchar(2)	Qx				
	time_year	Time's year	varchar(4)	xxxx				
	total_qty	Warehouse's total quantity	int	xxx				
	total_qty_sold	Warehouse's total quantity sold	int	xxx				

Delivery Staff

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
delivery_staff	deli_id	Delivery Staff's id	varchar(36)	xxx			PK	
	store_id	Store's id	varchar(10)	Sxxxx			FK	store[store_id]
	prod_id	Product's id	varchar(6)	PDxxxx			FK	product[prod_id]
	deli_name	Delivery Staff's name	varchar(50)	xxxx				
	deli_area	The area where the delivery staff delivers the item.	varchar(50)	xxxxx (xx)				
	deli_cost	delivery's cost	double	xx.xx				
	deli_distance	delivery's distance	double	xxx.xx				
	deli_fee	delivery's fee	double	x.xx				
	deli_gps	GPS that tells delivery staff the location	varchar(500)	xxxx, -xxxx				

Product

Table name	Attribute name	Contents	Type	Format	Range	Nullable	Key	FK Reference Table
product	prod_id	Product's id	varchar(6)	PDxxxx			PK	
	store_id	Store's id	varchar(10)	Sxxxx			FK	store[store_id]
	sup_id	Supplier's id	varchar(5)	SPExxxx			FK	supplier[sup_id]
	prod_name	Product's name	varchar(100)	xxxxxx				
	prod_cate	Category's name of the product	varchar(100)	xxxxxx				
	prod_price	Product's price	double	xx.xx				
	prod_website	Product's website	varchar(800)	https://www.ikea.com/us/en/p/xxxxx				
	prod_description	Detailed description of the product	varchar(800)	xxxxxx				
	prod_depth	Product's depth	int	xxx		Y		

	prod_height	Product's height	int	xxx		Y		
	prod_width	Product's width	int	xxx		Y		
	prod_size	Product for gender	varchar(10)	xxx	Small,Medium,Big			
	prod_quantity	Product's quantity	int	xxx				
	prod_cost	Product's cost	double	xx.xx				

1. Explanation Data Sources

IKEA supplies detailed information that includes a wide range of info vital for tactical research study and also decision-making. This information abundant in client info, sales stats and also considerable item summaries functions as a keystone for getting workable understandings right into the functional and also calculated aspects of IKEA's firm. By taking apart as well as examining the numerous elements of this information we obtained a deep understanding of IKEA's client demographics, item variety, sales patterns, and also geographical market infiltration every one of which are crucial for crafting educated, data-driven company techniques.

The information outlined client info consisting of demographics plus geographical circulation, sets the phase for nuanced consumer division as well as targeted advertising initiatives. IKEA has the ability to personalize its advertising initiatives to target certain market teams plus geographical locations by evaluating consumer information, which finds valuable patterns in buying actions as well as choices. This concentrated method enhances consumer interaction as well as complete satisfaction while maximizing source appropriation to one of the most encouraging markets plus consumer sections. Additionally, the data source's comprehensive item info supplies a lens whereby to review the efficiency as well as the variety of IKEA's item profile. With the evaluation of item groups, prices strategies and also sales numbers IKEA has the ability to establish which of its items are preferred together with which might require to be changed or assessed. This product-level evaluation is crucial in leading item advancement, rates as well as stock monitoring choices making sure that IKEA's item offerings stay straightened with market need as well as consumer assumptions.

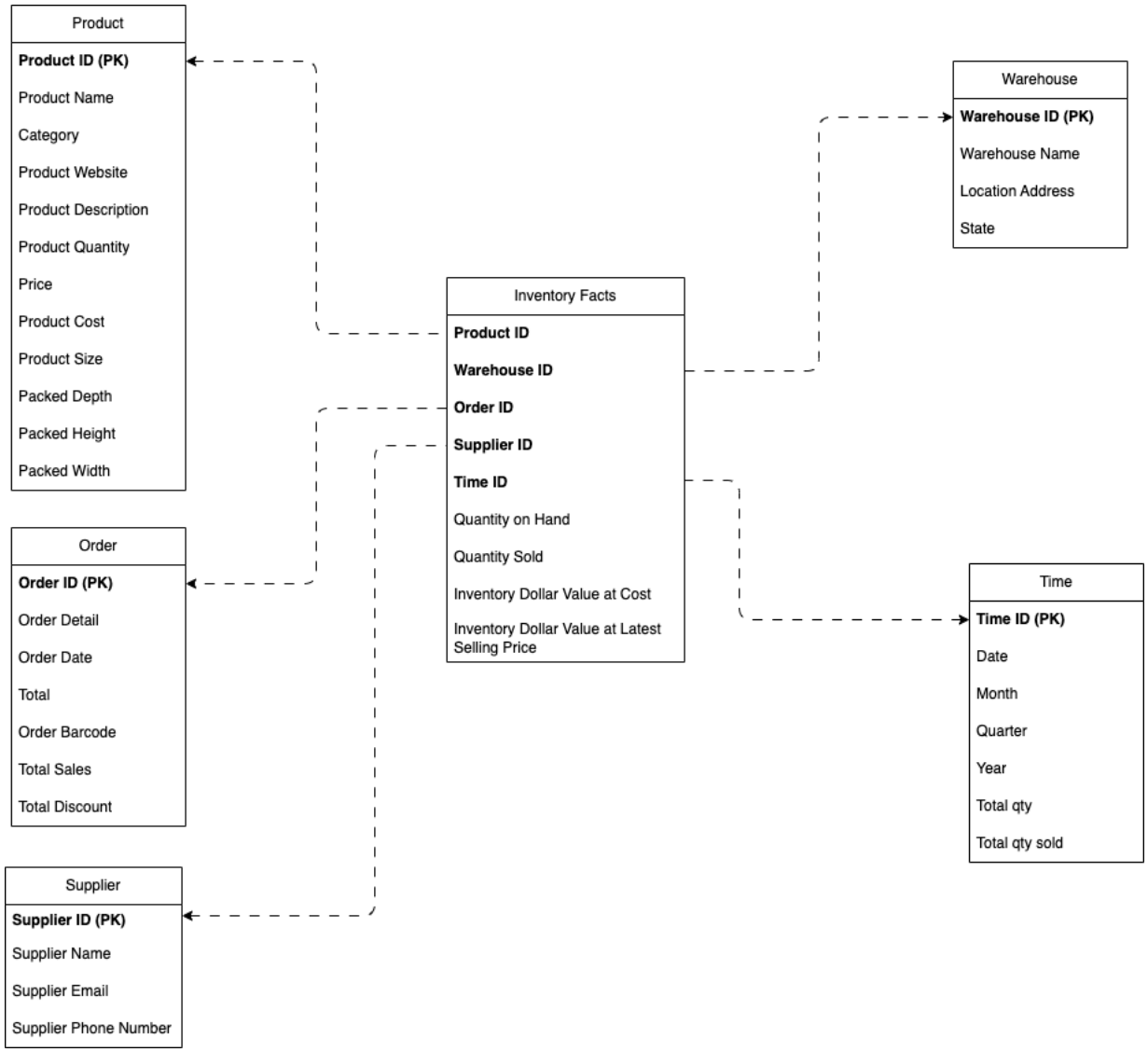
The sales information element of the data source including order IDs, days, and also amounts together with client as well as item information gives a vibrant summary of sales patterns together with cycles. This detail is very useful for recognizing seasonal changes in the item life process as well as the effect of advertising and marketing advocate sales efficiency. By associating sales information with item plus geographical info, IKEA discovered local item choices and also time-specific need patterns promoting much more reliable stock administration plus targeted marketing tasks.

The synthesis of understandings stemmed from consumer demographics, item details, plus sales information encourages IKEA to make notified tactical choices throughout different elements of its organization. From maximizing advertising methods together with improving item offerings to refining shop positionings plus online procedures the data-driven understandings amassed from the "" IKEA"" sheet make it possible for IKEA to adjust as well as flourish in the affordable retail landscape. By leveraging this extensive data source IKEA can guarantee its organization techniques are both receptive to existing market characteristics as well as anticipatory of future fads, eventually keeping its setting as a leader in the worldwide retail market.

Data warehouse design

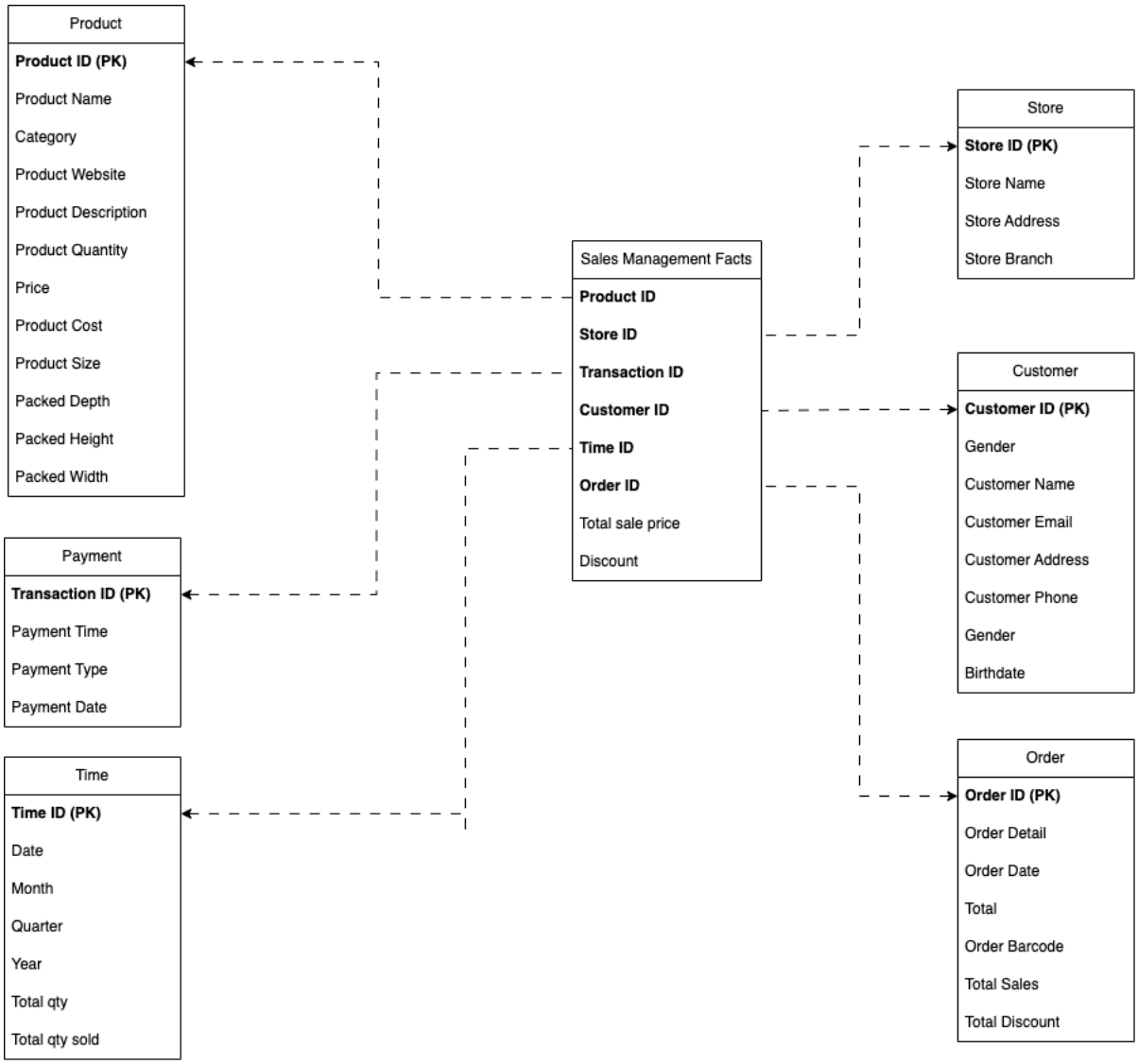
Star Schema

Inventory management



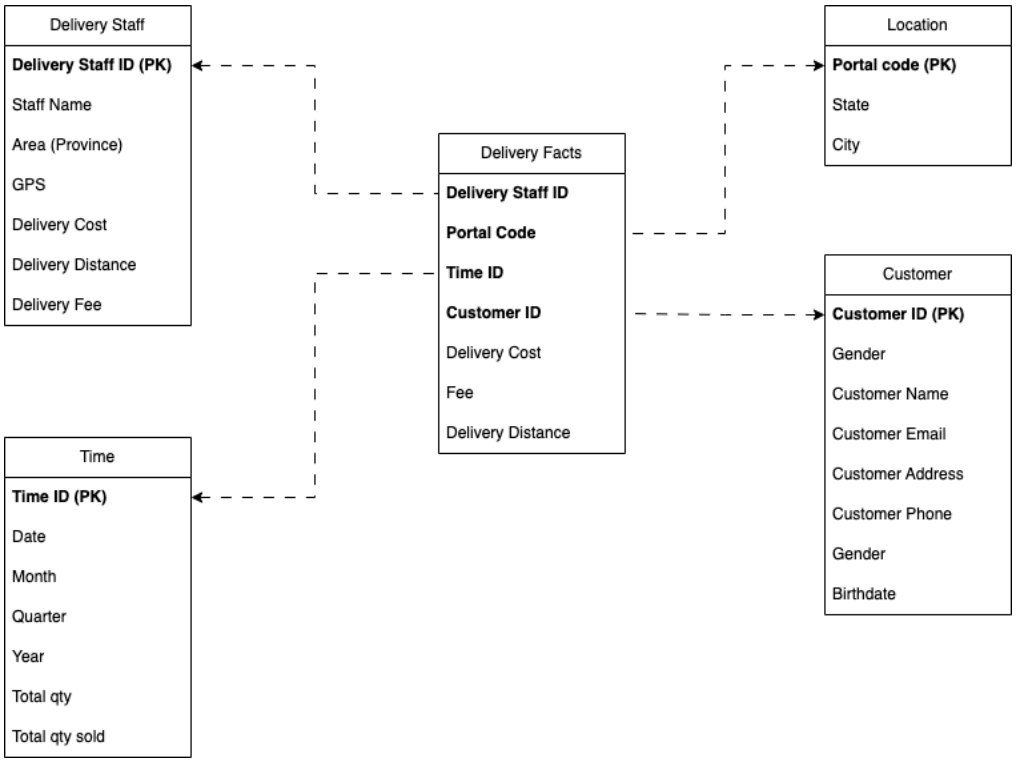
"**Inventory Facts"** containing inventory details like quantity and price, and surrounding dimension tables (e.g., "Product," "Warehouse") with descriptive attributes. This structure allows for efficient retrieval and analysis, enabling the creation of detailed reports that combine specific data points (product quantity at a warehouse) with related context (product category and supplier information). This structure will help generate insightful reports for better decision-making in the IKEA business.

Sales management



"**Sales Management Facts"** containing sales metrics like total sales and discounts. Surrounding this are dimension tables (e.g., "Customer," "Product," "Store") with additional details like customer information, product categories, and store locations. This structure allows for efficient analysis, enabling the creation of reports that combine sales figures with relevant context, such as which products sold well, in which stores, and to whom. This structure may help generate valuable reports to support strategic decisions in the IKEA business.

Delivery



"Delivery Facts" containing delivery metrics like cost and distance. Surrounding this are dimension tables (e.g., "Customer," "Product," "Location") with details like customer information, products delivered, and delivery locations. This structure allows for efficient analysis, enabling reports that combine delivery data with relevant context, such as which products were delivered to which customers by whom and at what cost. This structure helps generate valuable reports to support decision-making around delivery strategies and resource allocation in the IKEA business.

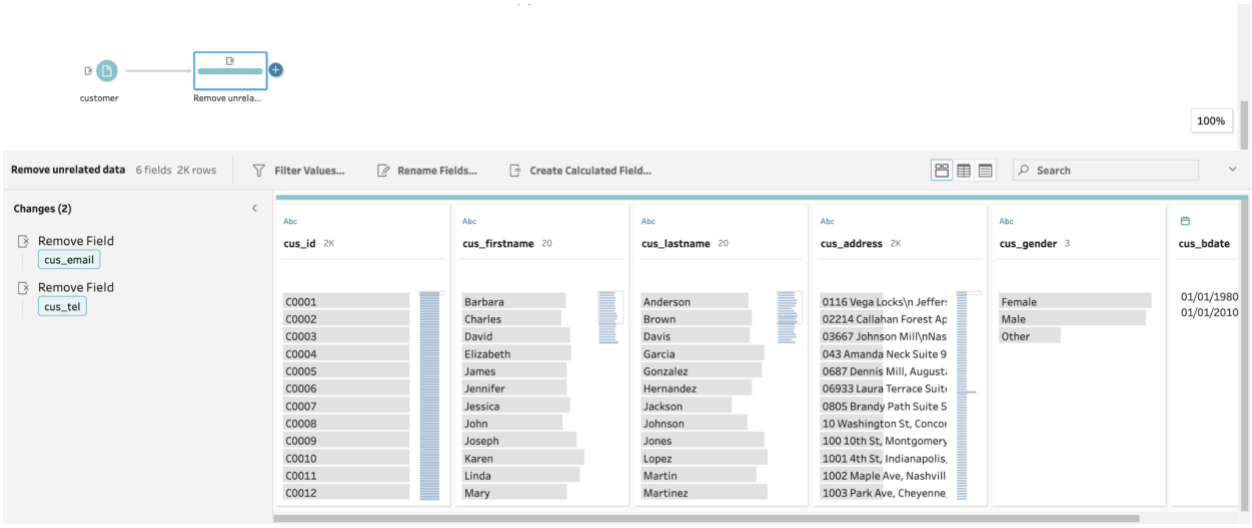
ETL process

Before we work on the data preparation and visualization, we created an assumption on the research to study and find the relationship between two or more variables. Gathering the available data source and estimating the relationship between the two or more variables produced the analytic question for visualization. The question was separated into three related parts. The first part of the question contains an assumption of the total quantity of the product, the second part is sales amount of the product, and the third part is the delivery distance from the delivery staff and customer.

A. Data Preparation

There are multiple data sources that we used in this project, and most of them are made to be ready to use in the visualization process. In this example, we select the table of the customer to illustrate the concept of our preparation process.

a. Remove unrelated data

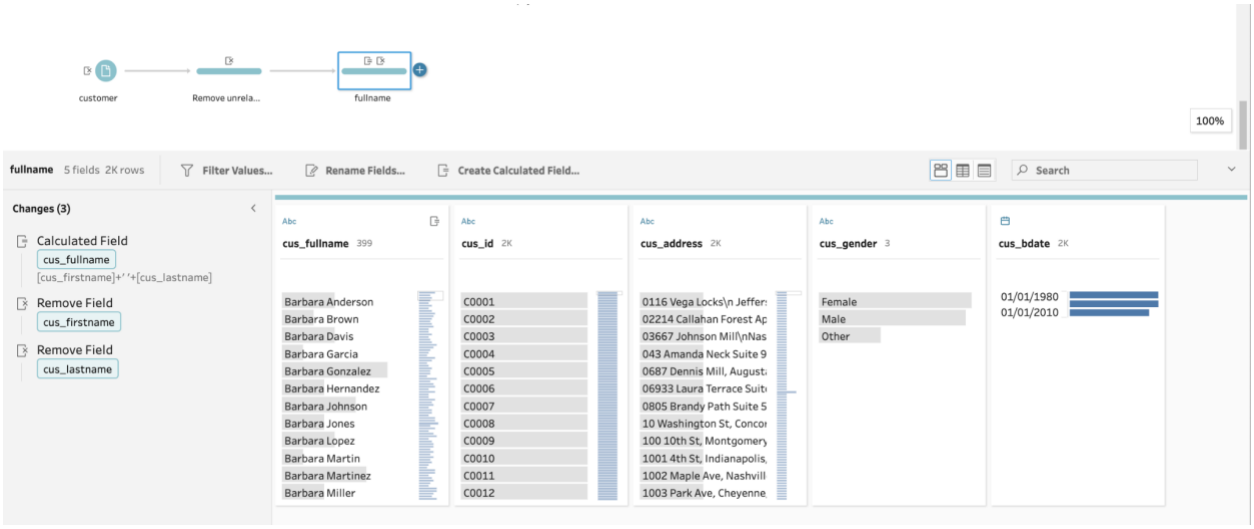


Firstly, after importing the data into Tableau Prep and choosing the data that will be used for analysis, we remove any data that is not required for the analysis or modeling being performed. This step is crucial for reducing the overall size of the dataset, improving performance, and ensuring the accuracy and validity of the analysis.

The data that will be removed are as follows:

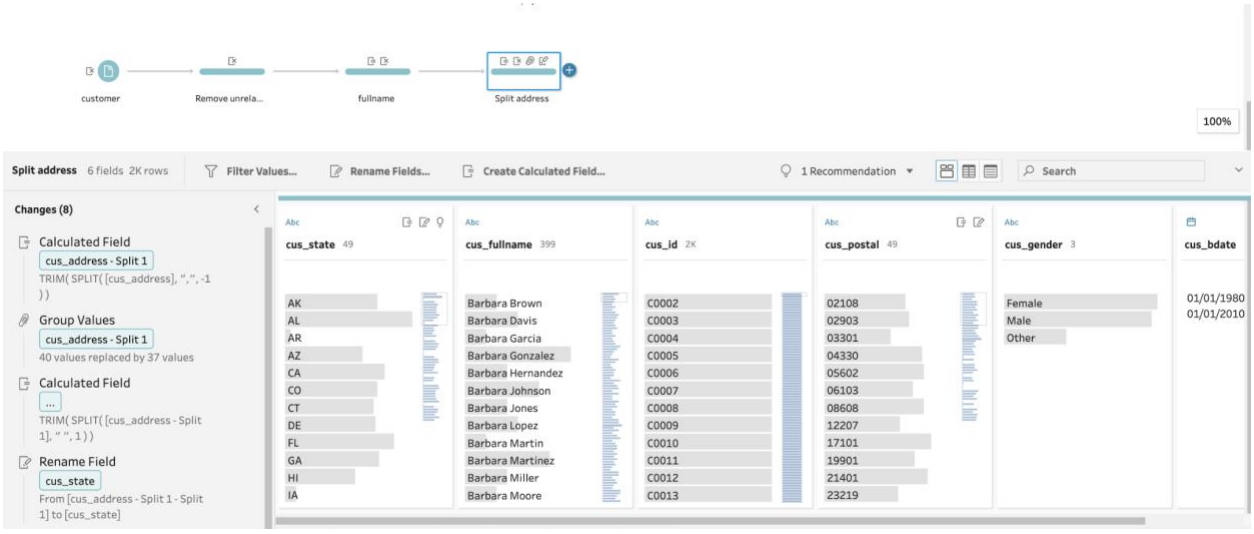
- customer: cus_tel, cus_email
- product: product description, product website

b. Cleaning data



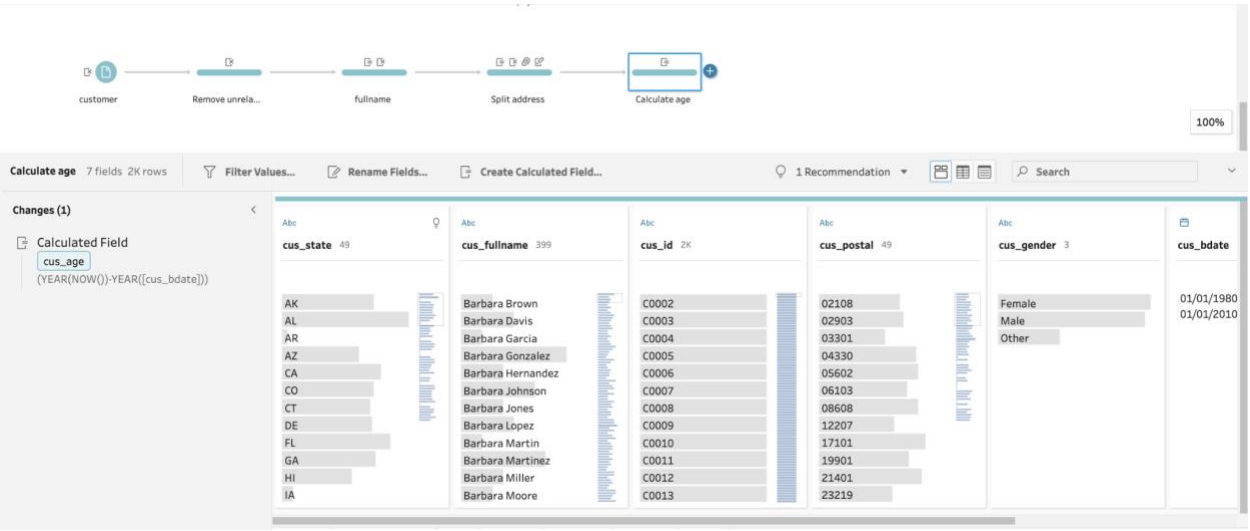
The most time-consuming step in data preparation is often cleaning up the data. Data cleaning is a step of eliminating outliers, such as punctuation from the first name and last name columns, NULL value, and wrong format data. In addition, the step of combining the related fields such as first name and last name together is acting during this step.

c. Split group of data



The next step is to split the data. To use the data in analysis, all the data must be separated into different fields in order to classify the differential. In our project, this step is used to split the address data into the fields of address with street and city, state and zip code.

d. Calculated data

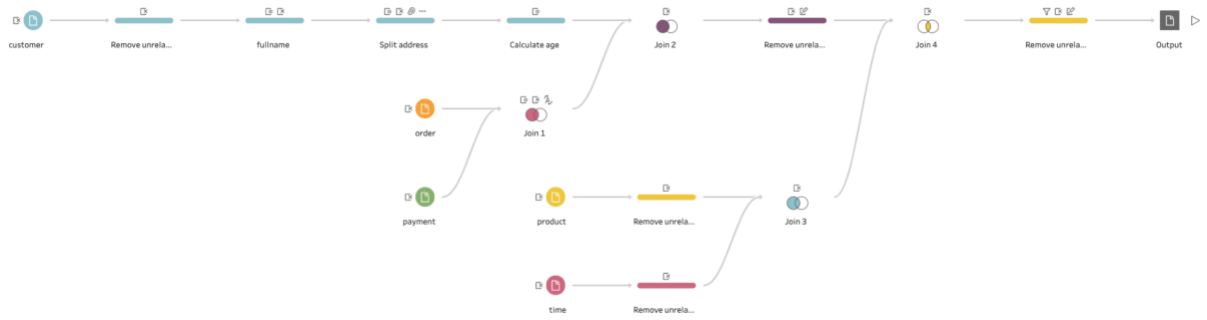


The calculated field must then be created in order to integrate the data that cannot be analyzed such as customers' birth date to age data.

e. Join data

When the data has been cleaned and verified, the following step is to join the data into a single table. Join will establish a connection with the appropriate data source.

- **Sales**



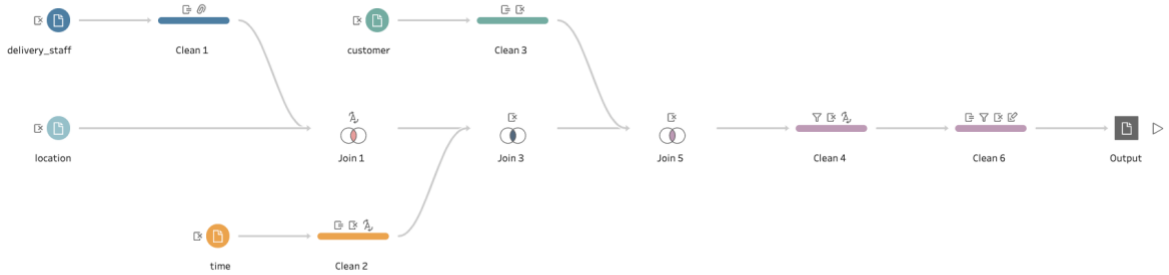
The image portrays another multi-step data processing pipeline involving a series of data manipulations and integrations. Starting from the 'customer' dataset, the first operation is to remove irrelevant or unwanted data, possibly outliers or errors. The process then proceeds to split the customer address into more manageable components and calculates the customers' age, perhaps from their date of birth.

Simultaneously, there are two additional data streams beginning with 'order' and 'payment' datasets, each undergoing a 'Remove unrelated...' step, which is likely similar to the cleaning step applied to the 'customer' dataset to ensure relevance and accuracy. These three streams merge in 'Join 1', (applied join clauses: transac_id = transac_id.) combining customer information with their corresponding orders and payments. Parallel to this, a 'product' dataset undergoes a similar unrelated data removal process and then joins the flow in 'Join 2' (applied join clauses: cus_id = cus_id).

Additionally, there's a 'time' dataset that also has unrelated data removed before it is integrated into the process at 'Join 3' (applied join clauses: prod_id = prod_id). After the various datasets are merged in 'Join 2' and 'Join 3', the combined data undergoes another removal of unrelated data in 'Remove unrelated...' before a final join operation in 'Join 4' (applied join clauses: time_id = time_id.). The process culminates with one more unrelated data removal step before the data reaches the 'Output', which would be the final, clean, and consolidated dataset ready for analysis or reporting.

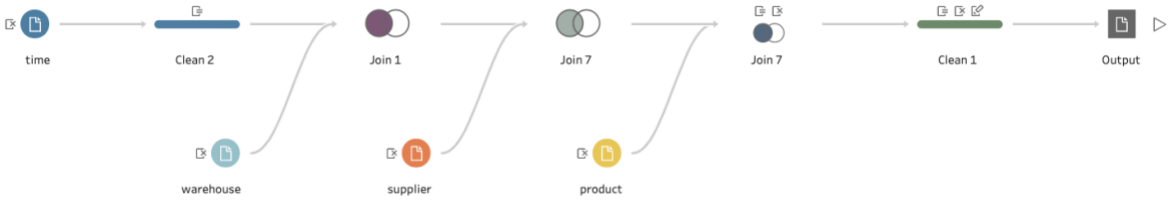
This workflow suggests a complex data environment where customer-related transactions are thoroughly cleaned and integrated, possibly for insights into customer behavior, financial processing, or inventory management.

- **Delivery**



This figure depicts a data processing workflow, which likely involves multiple stages of data cleansing and integration before arriving at a final output. Initially, there are three distinct datasets labeled 'delivery_staff', 'location', and 'time'. Each dataset undergoes an individual cleaning process, indicated by 'Clean 1', 'Clean 2', and 'Clean 3', respectively (Calculate field, Group value, remove field, and Change type). These cleaning processes are essential to ensure data quality by removing errors or inconsistencies. Following the initial cleansing, the 'delivery_staff' and 'location' datasets are merged in 'Join 1' (applied join clauses: deli_area = loca_state). The result of this merge is then subjected to another round of cleaning in 'Clean 4'. Parallelly, the 'time' data, after its initial cleaning, is combined with another dataset in 'Join 5' (applied join clauses: prod_ID = Cus_ID). This joined data is also cleaned ('Clean 6') before it finally converges with the earlier stream in 'Join 3' (applied join clauses: prod_ID = time_ID). The fully integrated and cleaned dataset is then directed towards the final output. This sequential and meticulous approach to handling data emphasizes the importance of data quality and integration in analytical processes.

- **Inventory**



This figure outlines another data processing sequence, which, like the first, involves data cleaning and merging operations leading to a final output. This process starts with four separate datasets: 'time', 'warehouse', 'supplier', and 'product'.

Initially, the 'time' data is cleaned in 'Clean 2'. Then, in 'Join 1' (applied join clauses: wareh_id = wareh_id), it is merged with the 'warehouse' data, which doesn't appear to undergo a preliminary cleaning step in this diagram. This merged data is then joined with the 'supplier' data in 'Join 2' (applied join clauses: sup_id = sup_id).

On a separate path, the 'product' data joins the flow directly into 'Join 3' (applied join clauses: prod_id = prod_id) without a prior cleaning step. After these merging operations, the combined dataset is cleaned in 'Clean 1', which could involve harmonizing the integrated data, ensuring consistency, and preparing it for final use.

The processed data then flows to the 'Output', which would be the end product of this data processing pipeline, likely a unified dataset ready for analysis, reporting, or application in business processes. This workflow suggests a focus on synchronizing data related to time, logistics, and supply chain management before it is considered clean and complete for the intended purpose.

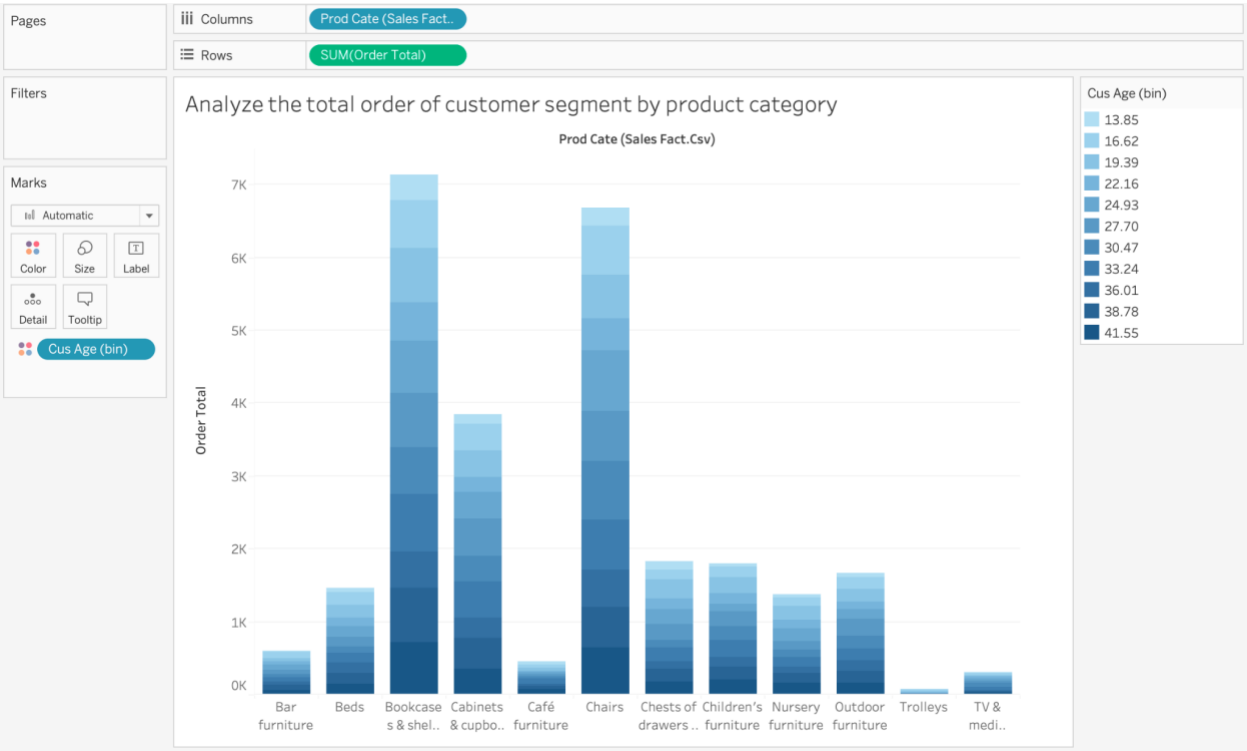
Analysis and Visualization

A. Data Analysis

The analysis encompasses an evaluation of multiple dimensions of sales data, including the profitability of different products, customer buying patterns, payment methods used for each product category, time periods with high sales, and locations with high sales for each product. Furthermore, we will investigate the relationship between location, payment methods, and the number of online and offline orders from each branch. The topic that will be analyzed are as follows:

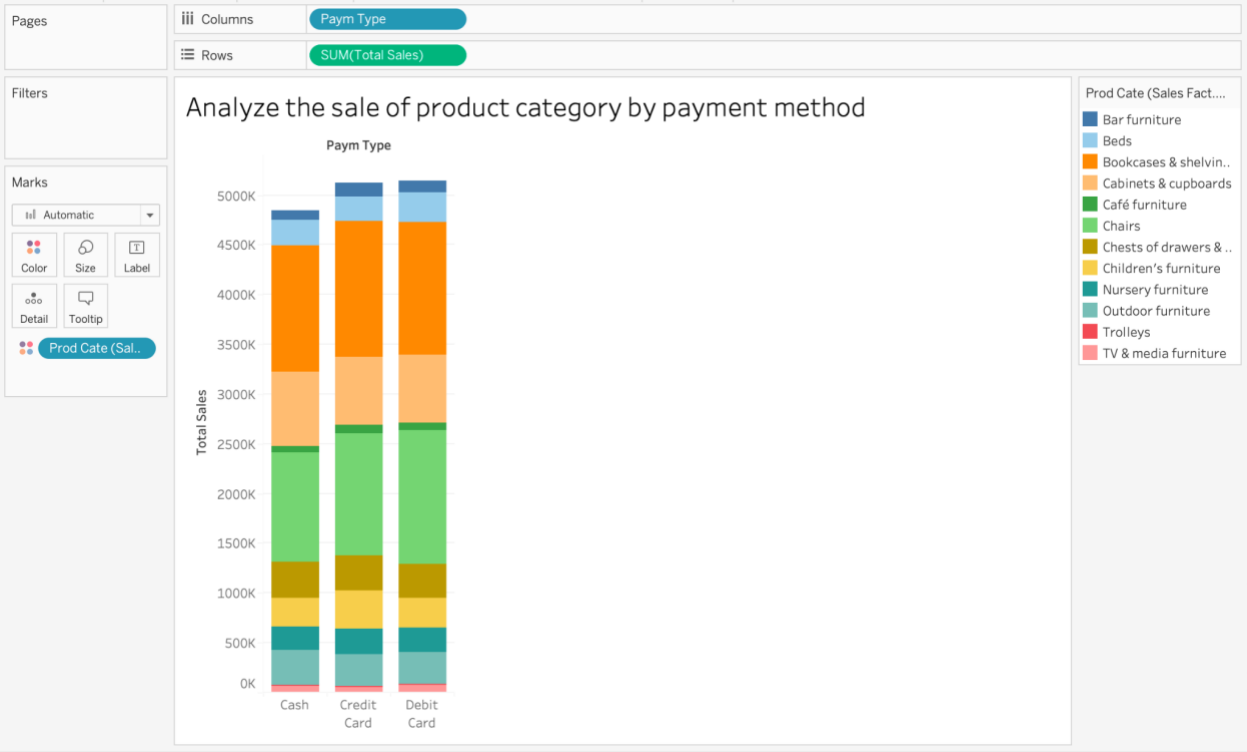
1. Analyze the total order of customer segment by product category
2. Analyze the sale of product category by payment method
3. Analyze the sale of product categories by order location-based
4. Analyze the total quantity of product categories by time series
5. Analyze the delivery distance by order location-based
6. Analyze the delivery cost of delivery area per month

1. **Analyze the total order of customer segment by product category.**



This Bar chart from Tableau Desktop, which presents a comparison of total customer orders by product category. The categories include a range of furniture items, such as bar furniture, beds, bookcases, cabinets, and more. The bars in the chart are color-coded to represent different customer age groups, as indicated by the legend labeled "Cus Age (bin)" on the right side of the chart. This age-based color coding suggests a segmentation of the order totals by customer age, allowing for a visual representation of which age groups are purchasing which types of furniture. The chart reveals, for example, that categories like Beds and Chairs have notably higher order totals compared to others like Café furniture and Trolleys. Such data is crucial for businesses to understand purchasing patterns across different age demographics, which can aid in optimizing marketing strategies and managing inventory according to customer preferences and buying behavior. However, the exact age ranges for the bins are not specified, and there is no interactive capability provided in the image to explore the data further.

2. Analyze the sale of product category by payment method.

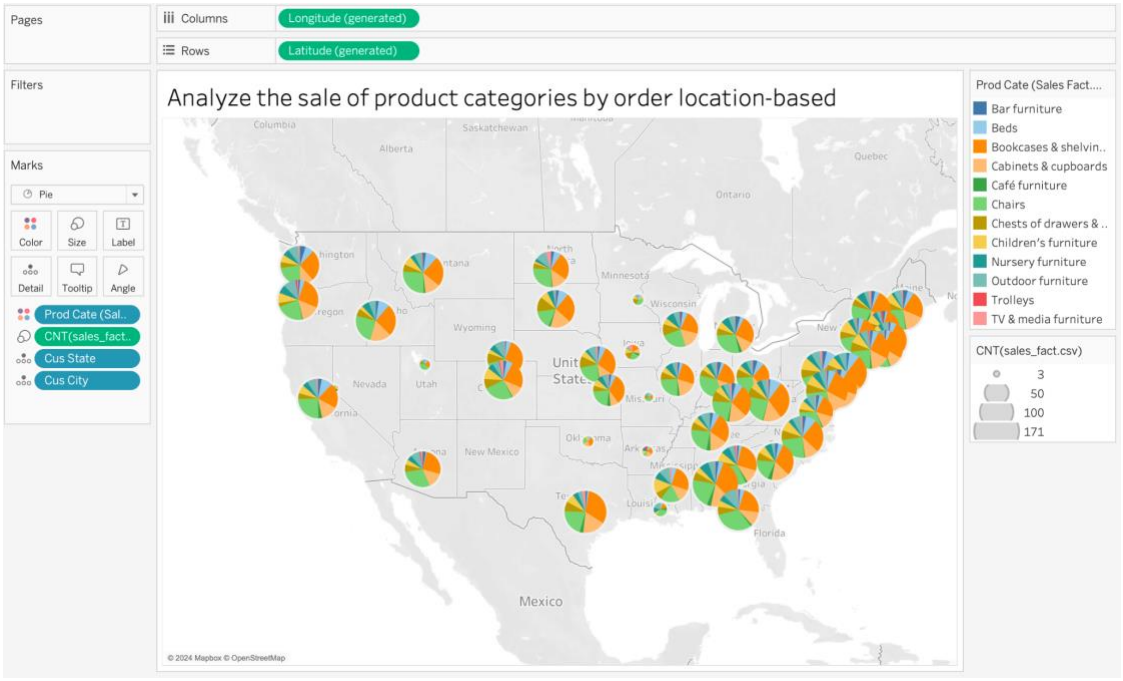


This stacked bar chart from Tableau Desktop, detailing the sales of different product categories segmented by payment method. Each bar represents a payment method, specifically cash, credit card, and debit card. The bars are subdivided into colored segments that correspond to various types of furniture, such as bar furniture, beds, bookcases & shelving, and so on, with each product category having a unique color as indicated in the legend on the right side of the chart.

This visualization is designed to analyze sales distribution across different payment methods for each product category. For instance, one could determine which product category has the highest sales for credit card payments or compare which payment method is most popular for purchasing nursery furniture. The total height of each bar reflects the total sales for that payment method, and the color segments show how much each product category contributes to this total.

The stacked bar chart format allows for a quick comparison of the total sales for each payment type while also providing insight into the sales composition. Such a chart would be useful for understanding consumer payment preferences and strategizing business decisions like stock inventory, payment processing options, and targeted promotions.

3. Analyze the sale of product categories by order location based.

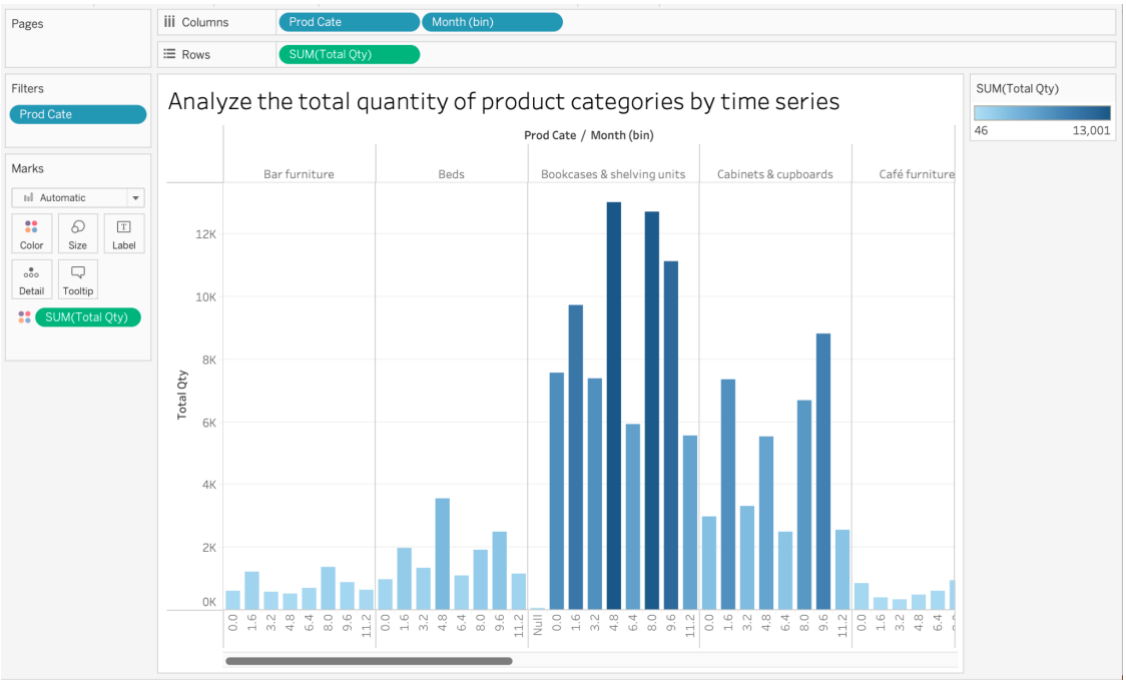


This figure showcases a geographical sales data visualization, specifically a pie chart map overlay, which is used to analyze the sale of different product categories across various locations. This type of chart is particularly effective for displaying how sales of products are distributed geographically and which products are popular in specific regions.

Each pie chart is placed on the map at the location corresponding to the data point, likely a city or a state within the United States. The size of each pie chart appears to represent the volume of sales (the larger the pie, the higher the sales), as indicated by the legend titled "CNT(sales_fact.csv)" with varying sizes for the numbers 3, 50, 100, and 171. The slices within each pie chart are color-coded to match the product categories listed in the legend on the right side of the image, such as bar furniture, beds, bookcases & shelving, etc.

This visualization allows stakeholders to quickly assess which product categories are performing well in which areas, and to identify regional market trends. For example, one might observe that outdoor furniture sells better in certain states, possibly those with warmer climates, while other categories like bookcases & shelving may perform well across the board. This information could be invaluable for strategic planning in marketing, distribution, and inventory management.

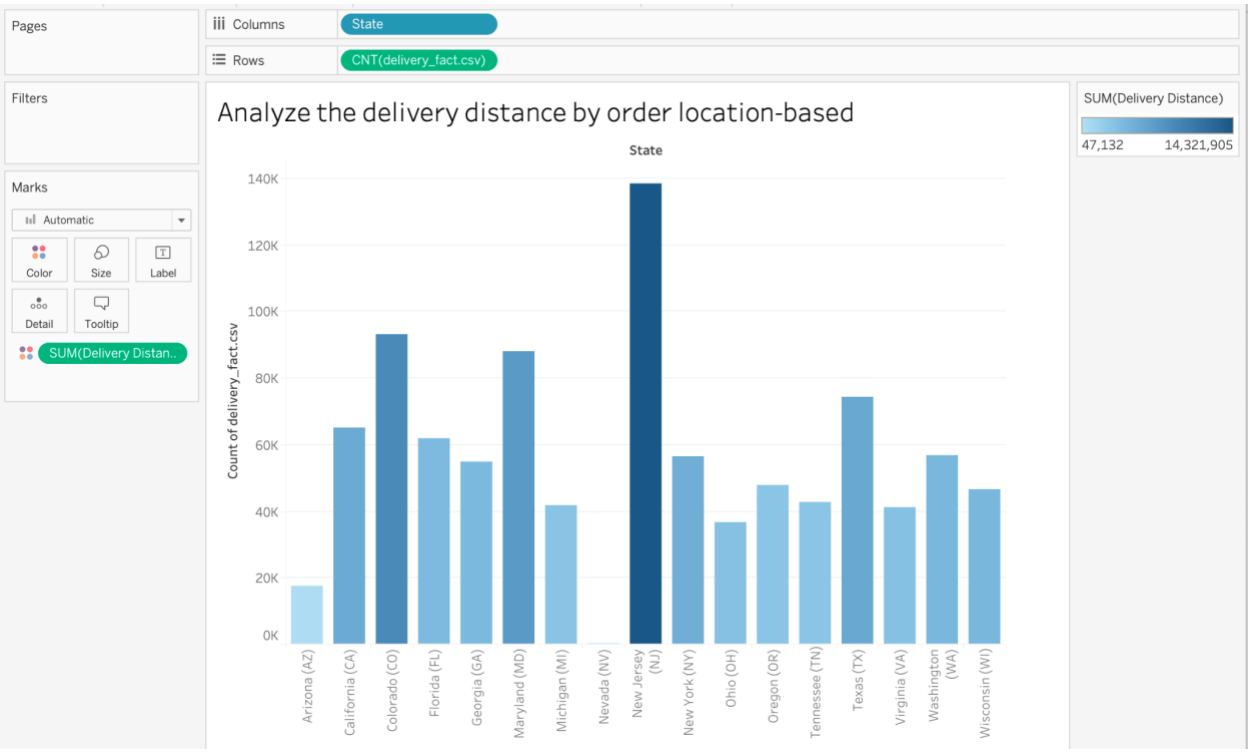
4. Analyze the total quantity of product categories by time series.



A bar chart designed to analyze the total quantity of various product categories sold over time, as seen in a data visualization tool. The x-axis of the chart is divided into segments that likely represent time periods (months), although the specific time intervals are not clear from the image. The y-axis indicates the total quantity of items sold, with values ranging from 0 to what appears to be over 12,000.

Each bar on the chart represents the total quantity of a particular product category sold during each time bin. The product categories are likely labeled at the bottom of each cluster of bars, with visible categories including bar furniture, beds, bookcases & shelving units, cabinets & cupboards, and café furniture. The color of the bars is uniform, indicating the chart is not segmented by another variable such as customer demographic or sales region. On the right, there is a legend indicating the range of the total quantity sold, from 46 to 13,001, which likely corresponds to the scale used on the y-axis. This type of visualization helps businesses track sales performance over time and identify trends or patterns, such as seasonal fluctuations or the popularity of certain product categories. It can inform inventory management, marketing strategies, and product development decisions. However, without more context or interactive capabilities, further detailed analysis is not possible from the static image alone.

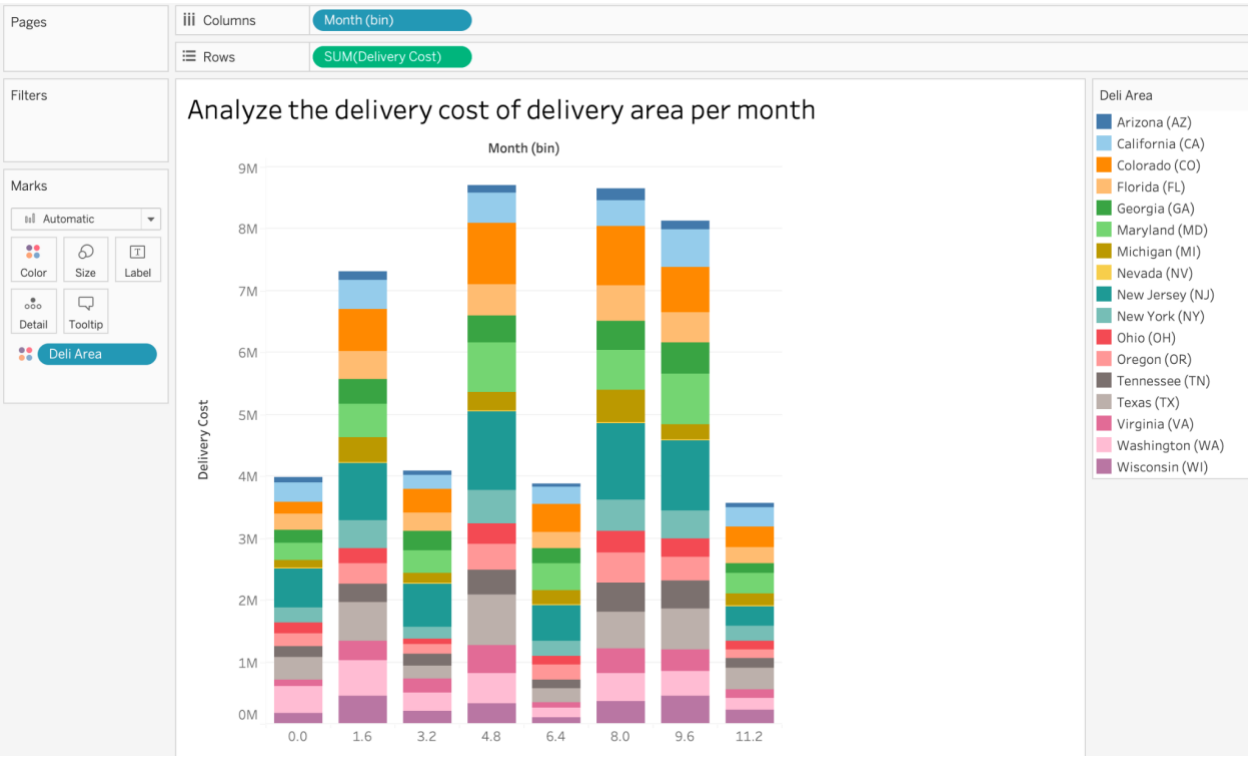
5. **Analyze the delivery distance by order location based.**



A bar chart that is being used to analyze delivery distances for orders across various states. The x-axis represents different states in the United States, and the y-axis indicates the cumulative distance of deliveries in those states. Each bar on the chart corresponds to a state and shows the total delivery distance for that particular state.

The chart includes a range of states from Arizona (AZ) to Wisconsin (WI), with one state (New York, NY) noticeably having a much higher total delivery distance compared to the others. This could suggest that New York has a higher volume of orders, a wider distribution range, or a combination of both. On the right side of the chart, there's a legend indicating the scale of 'SUM(Delivery Distance)' which ranges from 47,132 to 14,321,905. These numbers likely represent the range of total delivery distances across all states shown. This kind of data visualization is useful for logistics and distribution analysis, helping a business understand where it is shipping its products the most and potentially indicating which states may require more efficient distribution strategies or infrastructure to accommodate the high delivery distances.

6. Analyze the delivery cost of delivery area per month.



A stacked bar chart which is designed to analyze delivery costs by delivery area on a month-by-month basis. The x-axis is labeled "Month (bin)," suggesting that each bar represents a different month, though the specific months are not labeled. The y-axis shows delivery costs, which range from 0 to 9 million.

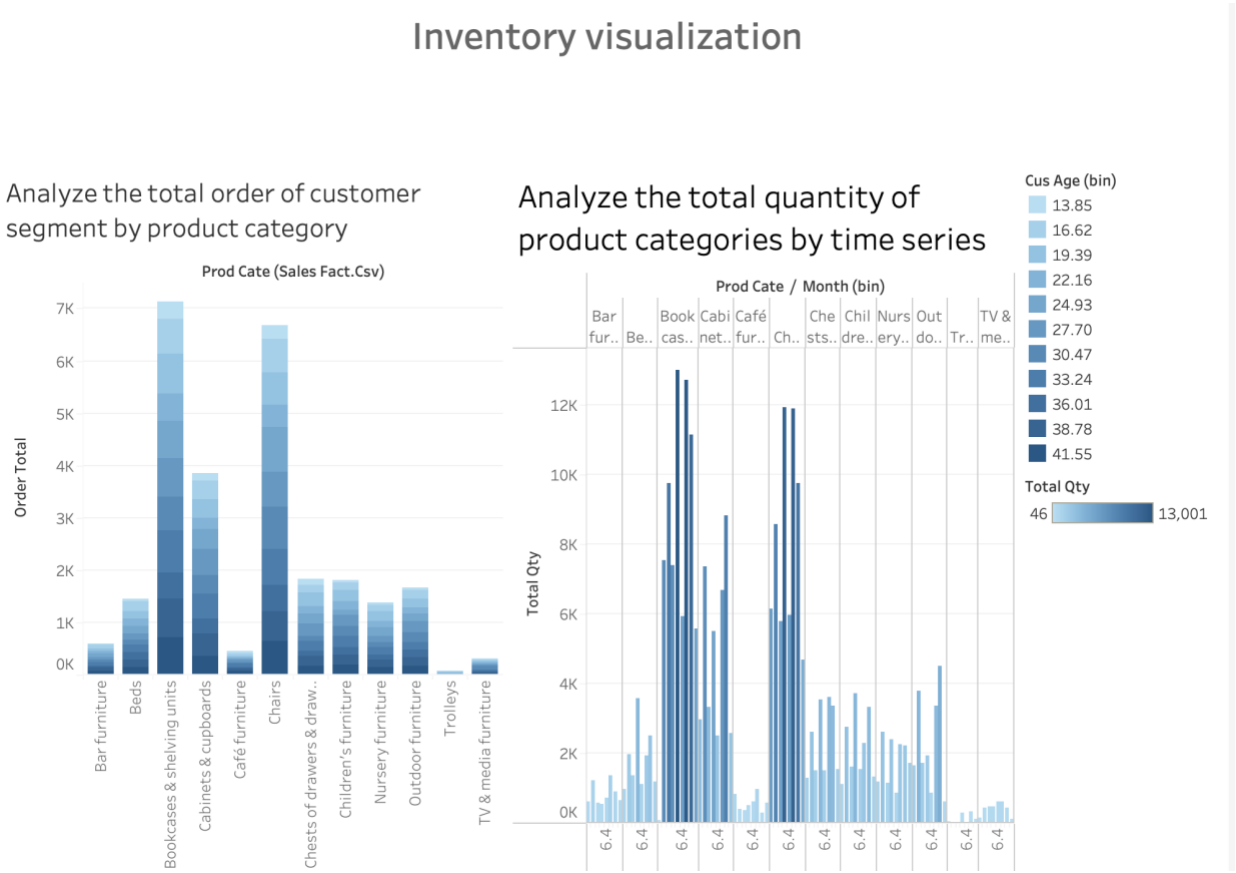
Each bar is composed of segments that are color-coded to correspond with various delivery areas, most likely different states as indicated by the legend on the right. The legend lists several U.S. states with associated colors, such as Arizona (AZ), California (CA), and New York (NY).

The chart allows viewers to compare the delivery costs incurred for each state across different months. For instance, it could show that the delivery costs for California are higher in certain months, which might suggest seasonal trends or other state-specific factors that affect delivery expenses.

The chart could be useful for a logistics manager or business analyst looking to optimize delivery routes, negotiate carrier rates, or manage budget allocations for logistics. It also helps to identify which states are the most expensive to deliver to and may benefit from cost reduction strategies.

B. Data visualization Dashboard

1. Inventory visualization



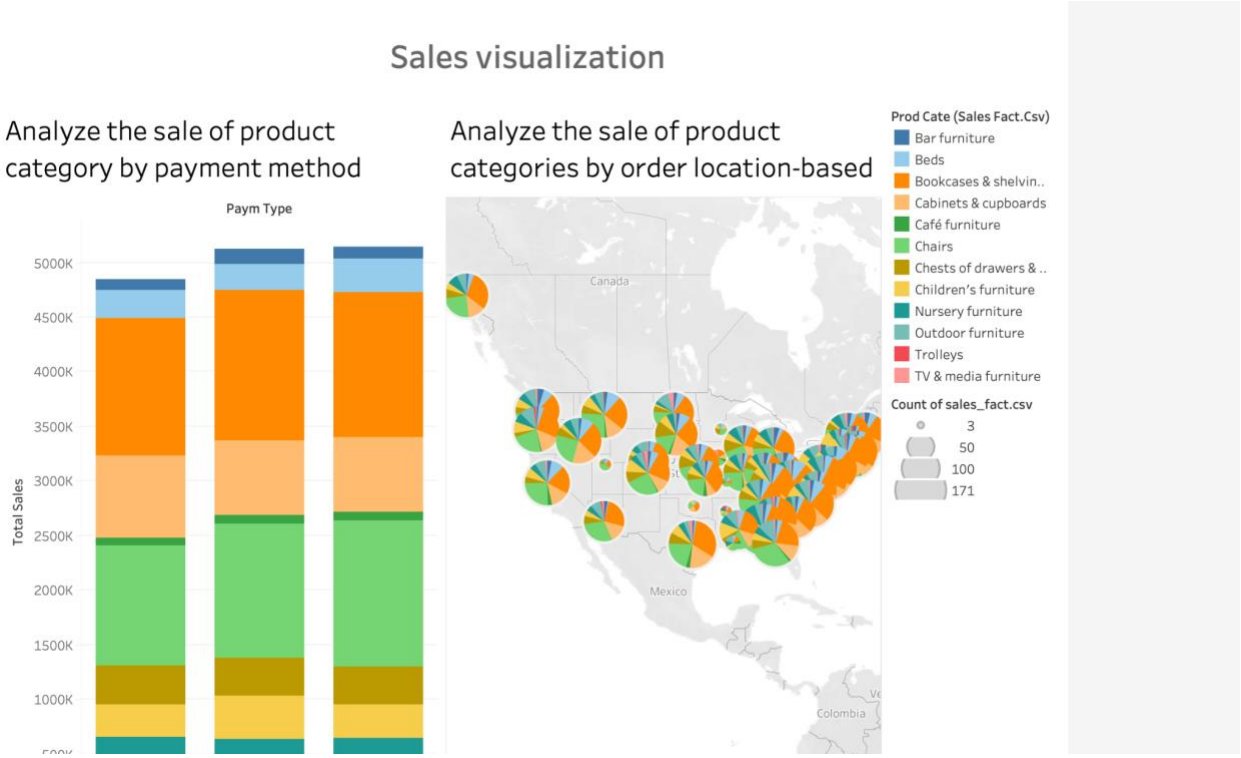
a composite of two bar charts titled "Inventory visualization" and is used for analyzing different aspects of product sales.

The first chart on the left is titled "Analyze the total order of customer segments by product category." It shows a vertical bar graph with product categories on the x-axis, which include bar furniture, beds, bookcases & shelving units, cabinets & cupboards, café furniture, and several others. The y-axis represents the order total, which reaches up to 7K (presumably 7,000 orders or a currency amount). The bars indicate the total orders per product category, with beds and chairs showing the highest totals.

The second chart on the right is titled "Analyze the total quantity of product categories by time series." This is also a vertical bar graph, but it shows product categories over time, with the x-axis labeled with product categories and broken down into time increments (possibly months). The y-axis represents the total quantity, stretching up to 12K (12,000 units or a currency amount). The bars are color-coded, which may correlate to a legend not fully visible in the image, suggesting

the data is segmented further, potentially by customer age group, as hinted by the legend on the right side of the chart.

2. Sales visualization



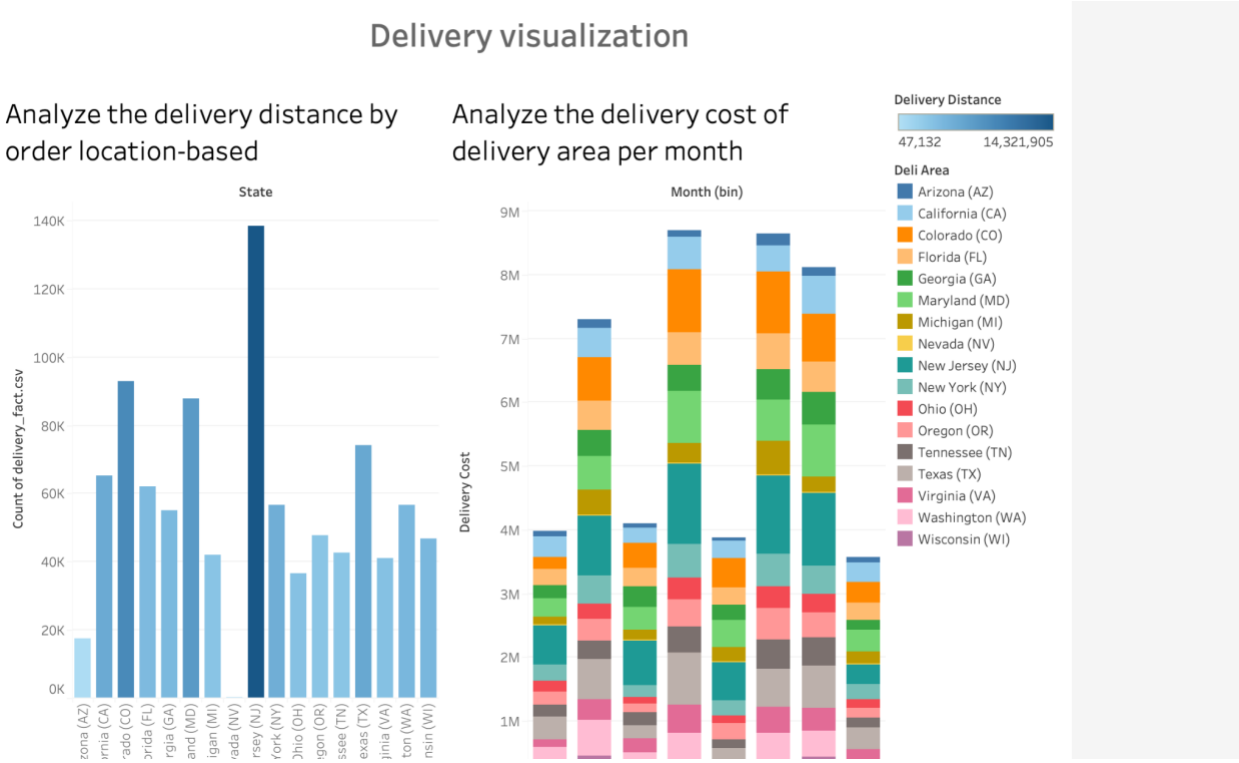
The figure depicts two sales data visualizations. On the left is a stacked bar chart titled "Analyze the sale of product categories by payment method." This chart displays total sales, segmented by payment type, which are indicated on the x-axis. Each bar is divided into colored sections that represent different product categories, such as bar furniture, beds, and bookcases & shelving, with each category's sales contribution stacked upon one another. This allows for a comparison of how much each product category contributes to the sales made via cash, credit card, or debit card.

On the right is a map of the United States titled "Analyze the sale of product categories by order location based." Overlaid on the map are pie charts placed at various locations, likely representing states or cities. Each pie chart's size and the slice sizes within it represent the volume and distribution of sales among different product categories in that particular location. The pie

charts vary in size, indicating the total number of sales, as outlined by the legend "Count of sales_fact.csv," with different sizes corresponding to sales counts of 3, 50, 100, and 171.

Together, these visualizations provide a comprehensive overview of sales performance by product category across different payment methods and geographical locations. This could help a business to identify which categories are popular in which areas and how customers prefer to pay for them, information that can be used to inform strategic decisions about inventory, marketing, and sales tactics.

3. Delivery visualization



Two bar charts related to delivery, categorized under "Delivery visualization." On the left, the chart titled "Analyze the delivery distance by order location-based" depicts a series of vertical bars, each representing a different state in the U.S., as seen on the x-axis. The y-axis measures the 'Count of delivery_fact.csv,' which likely corresponds to the total delivery distance associated with orders from each state. One state, in particular, stands out with a significantly higher delivery distance count than the others.

The chart on the right, "Analyze the delivery cost of delivery area per month," presents a stacked bar chart with each bar divided into colored segments. These segments correspond to different states, as listed in the legend on the right, and are stacked to represent the total delivery cost per month, with the x-axis denoting the time period in months and the y-axis showing delivery costs. This visualization illustrates the monthly distribution of delivery costs across different states, providing insight into which states have higher delivery costs and how those costs vary monthly.

Together, these charts offer a multifaceted view of delivery logistics, highlighting where and when delivery distances and costs are greatest, which could be vital for optimizing delivery routes, managing logistics expenses, and improving overall supply chain efficiency.

Discussion and conclusion

Discussion

In the rapidly evolving retail sector, IKEA's strategic adoption of data warehousing and data mining technologies represents a paradigm shift towards a more data-driven approach in enhancing operational efficiency and customer satisfaction. By leveraging Online Analytical Processing (OLAP) and Online Transaction Processing (OLTP) systems, IKEA has significantly improved its ability to manage sales data, monitor inventory levels, and analyze customer behavior in real-time. This integration facilitates a seamless flow of information across various departments, enabling a holistic view of the business operations and fostering an environment where data-driven decisions thrive.

The cornerstone of IKEA's strategy lies in its innovative use of data warehousing to consolidate disparate data sources into a single repository. This approach not only simplifies data management but also enhances the quality and accessibility of the data, allowing for more sophisticated analysis and insights. Through the application of data mining techniques, IKEA can uncover hidden patterns, correlations, and trends within its vast datasets, translating these insights into actionable intelligence that informs strategic decisions, from product development to marketing campaigns.

Moreover, the employment of advanced data visualization tools like Tableau has revolutionized the way IKEA communicates data insights within the organization. By transforming complex data sets into intuitive, interactive visualizations, stakeholders at all levels can easily comprehend critical business metrics, enabling faster and more informed decision-making processes. This visual approach to data analysis has not only increased operational transparency but also facilitated a culture of continuous improvement and innovation across the company.

Despite the clear benefits, the journey towards a fully integrated data-driven organization is fraught with challenges. Issues such as ensuring data privacy, managing the complexity of data integration, and keeping pace with the rapid advancements in data processing technologies require constant vigilance and adaptability. Nevertheless, IKEA's proactive and forward-thinking approach in

addressing these challenges exemplifies its commitment to leveraging cutting-edge technologies to maintain its competitive edge in the market.

In conclusion, IKEA's strategic focus on data warehousing and data mining underscores the company's recognition of the transformative power of data analytics in the retail industry. By prioritizing the integration of these technologies into its business model, IKEA not only enhances its operational efficiencies and customer engagement strategies but also sets a benchmark for innovation, demonstrating the critical role of data-driven insights in achieving sustainable business growth and success in today's digital age.

Conclusion

The conclusion of this comprehensive report on the utilization of OLAP (Online Analytical Processing) and OLTP (Online Transaction Processing) systems at IKEA, combined with the deployment of Tableau for enhanced system operations, encapsulates the transformative impact of integrating advanced data management and analysis technologies. Through meticulous investigation and analysis, the report has illuminated how these technologies fortify IKEA's strategic and operational frameworks, propelling the global retail giant toward even greater efficiency, customer satisfaction, and market adaptability.

At the heart of IKEA's success is its adept use of OLAP and OLTP systems to streamline and optimize its inventory management, sales transactions, and customer interactions. This technological backbone enables IKEA to process and analyze vast quantities of data, facilitating deep insights into market trends, customer preferences, and operational efficiencies. Such capabilities are crucial for IKEA, enabling the company to maintain its competitive edge in the fast-paced retail industry by ensuring that decision-making is both data-driven and agile.

The integration of Tableau elevates IKEA's data analysis and visualization processes, offering clear, actionable insights through dynamic reports and dashboards. This tool has been instrumental in enhancing strategic planning and operational decision-making across the company. By leveraging Tableau's powerful visualization capabilities, IKEA can swiftly interpret complex datasets, identify trends, and make informed decisions that align with its business objectives and customer needs.

Moreover, the report highlights the critical role of a well-designed data warehouse in supporting these systems. Through a detailed exploration of IKEA's data warehouse design, ETL (Extract, Transform, Load) processes, and the strategic application of BI (Business Intelligence), it becomes evident how these elements collectively support a robust infrastructure for data management and analysis. This foundation is indispensable for sustaining IKEA's commitment to delivering high-quality, affordable products while also pursuing sustainability and social responsibility goals.

In essence, the convergence of OLAP and OLTP systems with Tableau at IKEA exemplifies the transformative power of integrating advanced data management and analytics technologies into

corporate strategy and operations. This integration not only enhances operational efficiencies and decision-making but also supports IKEA's mission to provide exceptional customer experiences and drive sustainable business growth. As IKEA continues to navigate the complexities of the global retail market, its forward-thinking approach to data management and analysis will undoubtedly remain a key factor in its enduring success and innovation.

References

[1] IKEA, “IKEA.com - International homepage,” IKEA, 2023. https://www.ikea.com/

[2] V. Prasannakumar, “IKEA - A Case Study,” www.linkedin.com, Jul. 19, 2023. https://www.linkedin.com/pulse/ikea-case-study-vivek-prasannakumar/

[3] T. Stackpole, “Inside IKEA’s digital transformation,” Harvard Business Review, Jun. 04, 2021. https://hbr.org/2021/06/inside-ikeas-digital-transformation

[4] B. Behzad, M. Farzad, and R. Davidrajuh, “Understanding the IKEA Warehouse Processes and Modeling Using Modular Petri Nets,” International Journal of Simulation Systems Science & Technology, Mar. 2020, doi: https://doi.org/10.5013/ijssst.a.21.02.03.

Appendix

VDO Clip demonstration on how to use Tableau Prep to implement the ETL process to integrate data into the data warehouse and how to use Tableau Desktop to implement analysis and visualization reports:
https://drive.google.com/file/d/1Gr43Rkg2_hQrNDvZlXi01vriNmo87tBb/view?usp=sharing