



**Project 1: Data Warehousing for Business Intelligence
BMW**

By

Mr. Thanawath Huayhongthong	6388016
Miss Chalumphu Atjarit	6388020
Miss Pattanan Korkiattrakool	6388022
Mr. Naphat Sookjitsumran	6388059
Miss Vipavee Ngamyingsurat	6388094

**Submitted to
Asst. Prof. Dr. Songsri Tangsripairoj**

**A Report Submitted in Partial Fulfillment of
The Requirements for ITCS 453 Data Warehousing and Data Mining
Faculty of Information and Communication Technology
Mahidol University**

2022

Table of contents

Executive Summary	2
Project overview	3
Business Domain	3
Objectives and Scope	5
Business Requirements	5
Data Warehouse and BI Importance to the company	6
Data sources	7
ER diagrams	7
Data dictionaries of the operational database (OLTP)	9
Star schema diagrams	11
ETL process	17
Discussion and conclusion on the benefits of data warehouse and BI to the company	26
Discussion on the benefits of data warehouse and BI	26
Conclusion on the benefits of data warehouse and BI	26
Link for the VDO presentation:	27
https://loom.com/share/e12f4fbdd28f48cba00f149b7ad30d71	27
Reference	28

Executive Summary

The purpose of this data warehouse project is to provide BMW, an automobile company, with a comprehensive overview of its business and to implement business intelligence solutions to solve problems that can be addressed through data warehousing. This project involves analyzing key business processes and identifying opportunities for improvement through the use of data warehousing and business intelligence solutions. The project will include designing and building operational transactional databases (OLTP) and online analytical processing (OLAP) databases, as well as defining appropriate data sources for the data warehouse components used. The project will also involve designing the extraction, transform, and load (ETL) process, providing comprehensive data analysis and visualization and presenting the expected benefits of data warehousing and business intelligence solutions for BMW.

In creating BMW's data warehouse, careful consideration has been given to the company's business requirements, which include capacity, delivery, safety, production processes, distribution, inventory management, workflow, sustainability, integration with booking and sales systems, product, marketing, financial requirements, and data visualization. The implementation of business intelligence solutions will help BMW to gain insights about their customers, optimize pricing and ordering strategies, improve supply chain management, and enhance overall operational efficiency and profitability. By utilizing BMW's data warehouse and business intelligence solutions, the company can gain valuable insights into its operations, make informed decisions, and manage and plan business operations to reduce costs and increase net profits for the company.

Project overview



Business Domain

BMW, short for Bayerische Motoren Werke AG, is a German multinational corporation that designs, manufactures, and sells luxury vehicles, including cars, motorcycles, and bicycles. BMW is headquartered in Munich, Germany, and operates in over 150 countries worldwide. The company's business domain is primarily focused on the production and sale of premium quality vehicles that offer innovative technologies, outstanding performance, and exceptional design. BMW's portfolio includes various models of cars, motorcycles, and bicycles, catering to a diverse range of customers globally.

In addition to the production of vehicles, BMW also offers financial services, such as leasing, financing, and insurance for its customers. The company has a strong focus on sustainability and has committed to reducing carbon emissions and increasing the use of renewable energy in its operations. BMW is a well-established and respected brand in the automobile industry, known for its engineering excellence, cutting-edge technology, and luxurious designs. The company's business domain encompasses various aspects, including design, production, marketing, sales, and services, all aimed at providing exceptional experiences to its customers.

- Design and Production: BMW's design and production process is focused on creating vehicles that combine innovation, quality, and sustainability. The company has several design studios around the world that work on developing new concepts and designs for their vehicles. Once a design is finalized, BMW uses advanced manufacturing techniques to produce high-quality vehicles that meet the company's rigorous standards.
- Marketing and Sales: BMW markets and sells its products through a global network of dealerships and showrooms. The company uses a range of marketing strategies, including advertising, events, and sponsorships, to promote its products and brand. BMW also has an online presence, offering customers the ability to configure and order their vehicles online.
- Services: BMW offers a range of services to its customers, including maintenance and repair, roadside assistance, and extended warranty programs. The company also provides financial services, such as leasing and financing, to help customers purchase their vehicles. BMW's services are designed to provide customers with a seamless and convenient ownership experience.
- Sustainability: BMW is committed to sustainability and has set several goals to reduce its carbon emissions and environmental impact. The company has invested in renewable energy sources, such as wind and solar power, to power its production facilities. BMW also uses sustainable materials and production methods to reduce waste and promote environmental stewardship.

Overall, BMW's business domain encompasses several aspects of the automotive industry, including design, production, marketing, sales, and services. The company's focus on innovation, quality, and sustainability has helped it become a leading brand in the luxury vehicle market.

Objectives and Scope

The objective of this project is to build a data warehouse that can consolidate data from various sources such as sales, marketing, customer support, and production. The data warehouse will provide a unified view of the company's data and enable better decision-making by the management team. The scope of the project is to design, develop, and implement a data warehouse that meets the company's business requirements.

Business Requirements

To build an effective data warehouse, the following business requirements need to be considered:

- Data Integration: BMW needs a data warehouse that integrates data from various sources, including sales, marketing, production, and customer service. The data should be structured and standardized, to ensure consistency and accuracy.
- Data Quality: The data in the data warehouse should be of high quality, with minimal errors and inconsistencies. Data cleansing and validation processes should be implemented to ensure the accuracy and completeness of the data.
- Data Governance: BMW requires a robust data governance framework that ensures data security, privacy, and compliance. The data warehouse should comply with relevant laws and regulations, such as GDPR, and access to the data should be controlled and monitored.
- Scalability and Flexibility: BMW needs a data warehouse that is scalable and flexible, to accommodate changing business needs and data volumes. The data warehouse should be able to handle large amounts of data and support the integration of new data sources and technologies.
- User-Friendly Interface: The data warehouse should have a user-friendly interface that enables easy access to the data by authorized users. The interface should be intuitive and user-friendly, with features such as search, filtering, and sorting.

In summary, BMW requires a data warehouse that integrates data from various sources, ensures data quality and governance, supports reporting and analysis, is scalable and

flexible, and has a user-friendly interface. These business requirements will enable BMW to leverage its data assets to drive business value, gain insights into customer behavior, and make informed decisions.

Data Warehouse and BI Importance to the company

A data warehouse and business intelligence (BI) are important to BMW for the following reasons:

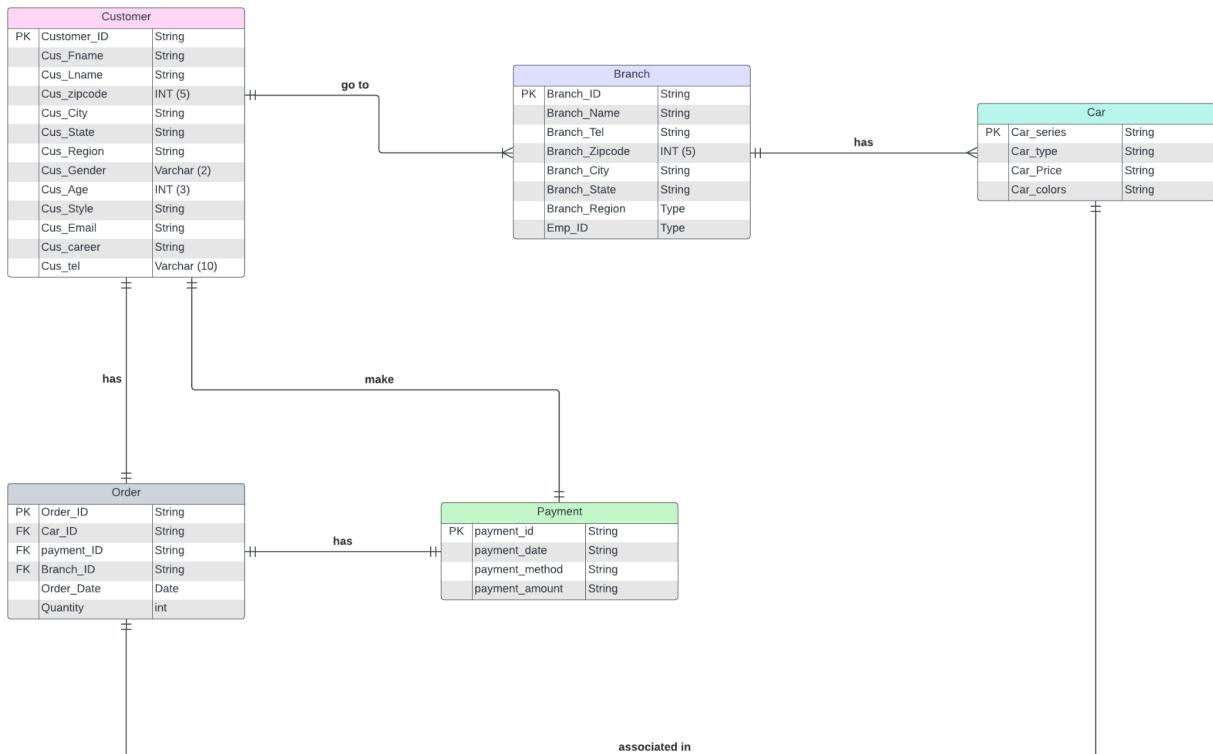
- Data-Driven Decision Making: Data warehouse and BI enable BMW to make data-driven decisions by providing access to accurate, timely, and relevant data. This helps BMW to identify patterns, trends, and insights that can inform business decisions, improve operational efficiency, and enhance customer experience.
- Improved Customer Experience: Data warehouse and BI help BMW to gain a deeper understanding of its customers, including their preferences, behavior, and needs. This enables BMW to personalize its offerings, improve customer satisfaction, and build customer loyalty.
- Improved Customer Experience: Data warehouse and BI help BMW to gain a deeper understanding of its customers, including their preferences, behavior, and needs. This enables BMW to personalize its offerings, improve customer satisfaction, and build customer loyalty.
- Better Collaboration: Data warehouse and BI facilitate better collaboration among teams by providing a shared platform for accessing and analyzing data. This helps BMW to break down silos, improve communication, and foster a culture of collaboration and innovation.
- Increased Efficiency: A data warehouse can help BMW streamline its operations by providing accurate and timely data to business users and decision-makers.

In summary, data warehouses and BI are important to BMW company for their ability to enable data-driven decision making, improve customer experience, provide a competitive advantage, optimize operations, and facilitate better collaboration. By leveraging these technologies, BMW can stay ahead of industry trends, respond quickly to market changes, and deliver exceptional value to its customers.

Data sources

The data sources of our selected company consist of two diagrams which are ER diagram and Data dictionary. We have provided the explanation of each diagram.

ER diagrams



Link for ER diagram:

https://lucid.app/lucidchart/698cfab7-4e51-4396-a160-180780cad0cb/edit?viewport_loc=6%2C-7%2C2786%2C1605%2C0_0&invitationId=inv_11496eb2-29cd-4319-b783-4e00d80283ce

The ER diagram represents the database schema for the Data warehouse of an automobile company, BMW. The diagram provides a comprehensive overview of BMW's corporate-customer trading system in which the diagram represents the core entities and its attributes with the relationship between entities. The diagram consists of five entities which are Customer, Branch, Car, Order, and Payment. Each entity related to others followed business rules and constraints on each relationship.

The examples of relationships and business rule of the ER diagram:

- One customer has one order at a time, and it will store in Customer and Order entity, Customer_ID and Order_ID as its primary key.
- One order can has one payment at a time too, and information about payment it will stored in entity of payment
- One branch can have many cars in branch
- One customer can make one payment
- One customer can go to many branch

Data dictionaries of the operational database (OLTP)

Entity: Customer					
Attribute name	Constraints	Data Type	Field Length	Description	Data Format
Customer_ID	PK	String	13	10 unique digits to identify the customer	x-xxx-xxxx-xx
Cus_Fname	Not null	String	20	Firstname of customer	'Amanada'
Cus_Lname	Not null	String	20	Lastname of customer	'Johnson'
Cus_zipcode	Not null	Int	5	5 unique digits to identify customer zipcode	10230
Cus_City	Not null	String	20	City name of customer	Newyork
Cus_State	Not null	String	20	State name of customer	Newyork City
Cus_Region	Not null	String	20	Region name of customer	East
Cus_Salary	Not null	Int	8	Salary of customer	70,000
Cus_Gender	Not null	Varchar	2	Gender of customer (M, F)	F
Cus_Age	Not null	Int	3	Age of customer	38
Cus_Style	Not null	String	500	Driving style of customer	Slow
Cus_Email	Not null	String	50	Email of customer	xxxxxx@gmail.com
Cus_career	Not null	String	100	Career of customer	Lawyer
Cus_tel	Not null	Varchar	13	Mobile number of customer	x-xxx-xxx-xxx
Entity: Car					
Attribute name	Constraints	Data Type	Field Length	Description	Data Format
Car_id	PK	String	12	12 unique digits identify a car	xxx-xxx-xxxx
Car_series	Not null	String	50	Identify series of car	BMW-series 3
Car_type	Not null	String	50	Identify type of car	SUV, sedan, Sport
Car_Price	Not null	String	10	Price of car	\$200,000
Car_colors	Not null	String	10	Color of car	yellow, white, black
Entity: Branch					
Attribute name	Constraints	Data Type	Field Length	Description	Data Format
Branch_ID	PK	String	12	12 unique digits to identify the branch	xxx-xxx-xxxx
Branch_Name	Not null	String	50	Name of Branch	Beverly Hills BMW
Branch_Tel	Not null	String	10	Mobile number of branch	888-888-4290
Branch_Zipcode	Not null	String	5	Zipcode of branch	90036
Branch_City	Not null	String	50	City name of branch	Los angeles
Branch_State	Not null	String	50	State name of branch	California
Branch_Region	Not null	String	50	Region name of branch	west
Entity: Payments					
Attribute name	Constraints	Data Type	Field Length	Description	Data Format
payment_id	PK	String	12	12 unique digits to identify the payment	PM2-405-079Y
Car_id	FK	String	12	12 unique digits identify a car	xxx-xxx-xxxx
payment_date	Not null	Date	dd/mm/yy	date of payment	25/12/22
payment_method	Not null	String	50	type of payment method	pay by installments/ paid in full
payment_amount	Not null	String	10	total amount of payment	40 k
Entity: Order					
Attribute name	Constraints	Data Type	Field Length	Description	Data Format
Order_ID	PK	String	12	12 unique digits to indentify the order	xxx-xxx-xxxx
Car_ID	FK	String	12	12 unique digits to identify the car	xxx-xxx-xxxx
payment_id	FK	String	12	12 unique digits to identify the payment	PM2-405-079Y
Branch_ID	FK	String	12	12 unique digits to identify the branch	xxx-xxx-xxxx
Order_Date	not null	Date	dd/mm/yy	Date of order	25/12/22
Quantity	not null	int	1	quantity of order	1
options	not null	String	10	option for receive car	Delivery, Pickup

Link for Data dictionaries:

https://docs.google.com/spreadsheets/d/1zltVwwiHWtLyLwFk7_LHOFa0KObqes-enSjsWvPXpAo/edit#gid=0

The data dictionary is a description of data in business terms from the above figure is a business term of the automobile company, BMW by the data dictionaries that describe the content, data format, and structure of the data warehouse which are Customer, Car, Branch, Payment, and Order.

The Customer Entity represents the customer of BMW that contains several attributes which are Customer_ID, Cus_Fname, Cus_Lname, Cus_Zipcode, Cus_City, Cus_State, Cus_Region, Cus_Salary, Cus_Gender, Cus_Age, Cus_Style, Cus_Email, Cus_career, Cus_Tel.

The Car Entity represents the car of BMW that contains several attributes which are Car_id, Car_series, Car_type, Car_Price, and Car_colors.

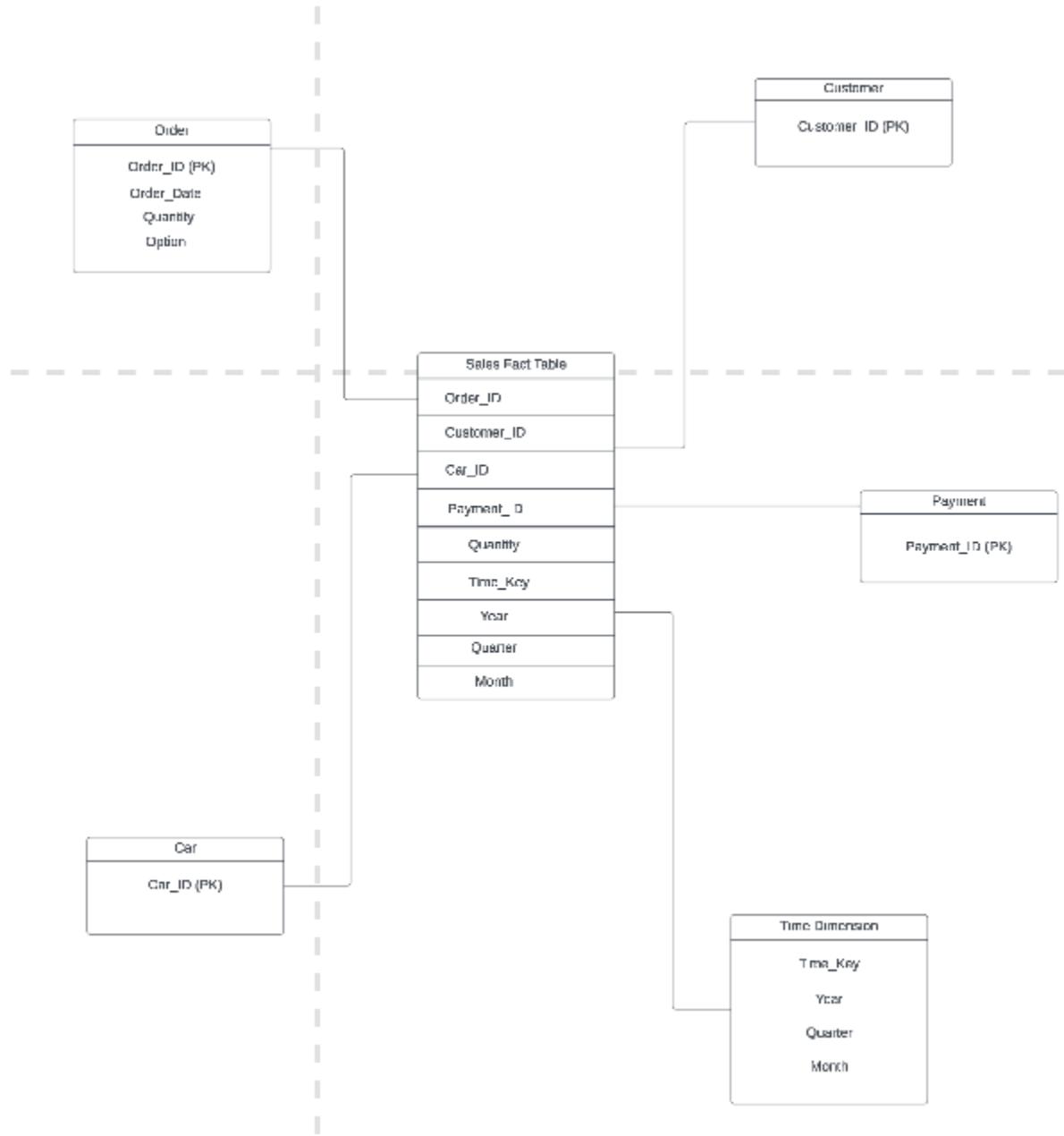
The Branch Entity represents the branches of BMW that contain several attribute which is Branch_ID, Branch_Name, Branch_Tel, Branch_Zipcode, Branch_City, Branch_State, Branch_Region

The Payments Entity represents the Payment of BMW that contains several attributes which are payment_id, Car_id, payment_date, payment_method, and payment_amount.

The Order Entity represents the Order of BMW that contains several attributes which are Order_ID, Car_ID, payment_id, Branch_ID, Order_Date, Quantity, and options.

Data warehouse design

Star schema diagrams



Link for Star Schema Sale Process:

https://lucid.app/lucidchart/320772c5-8d6f-4e93-b661-159317df6d9d/edit?viewport_loc=-2307%2C-1037%2C5490%2C2657%2C0_0&invitationId=inv_774e416f-b32e-46df-a779-deab30b29b51

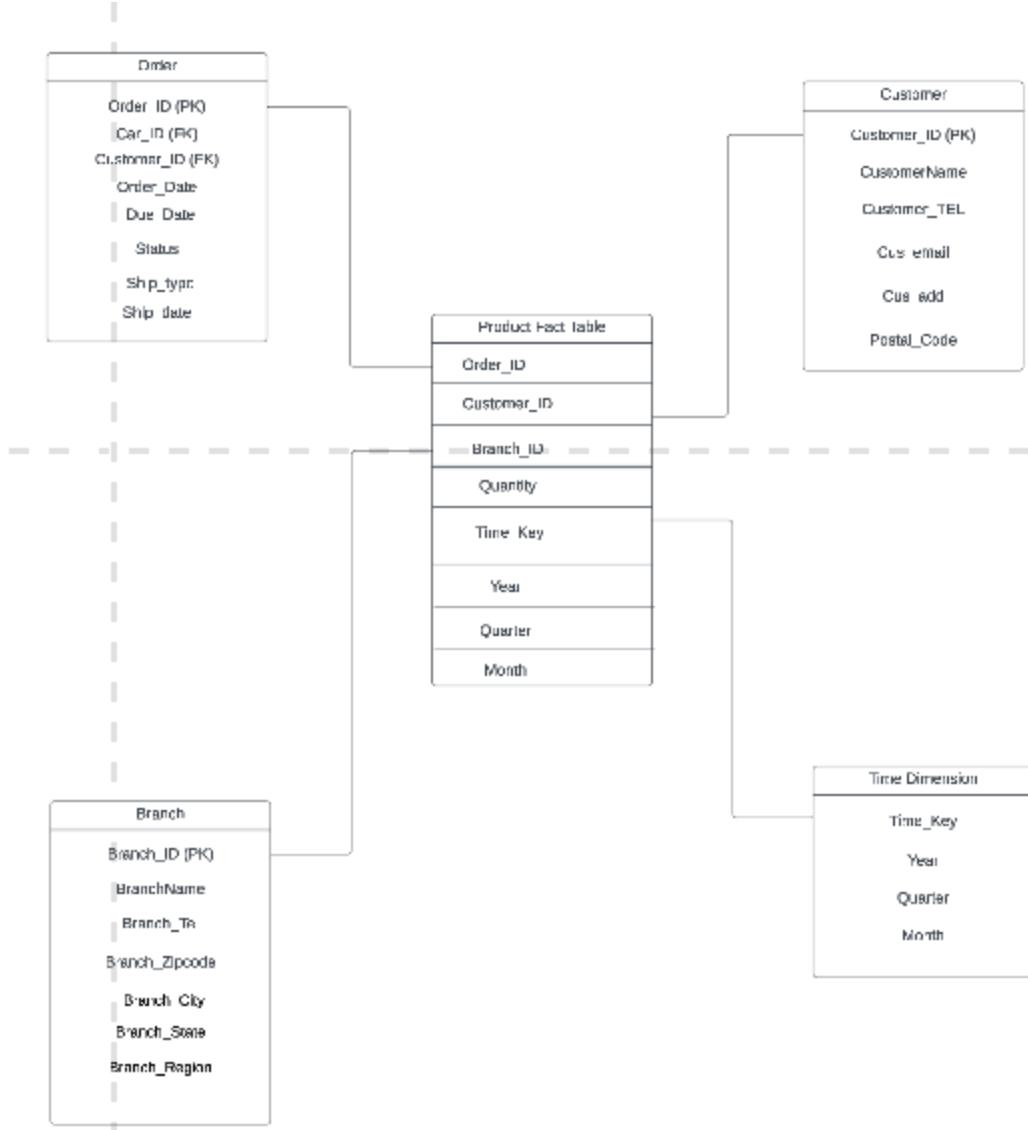
Sale Process

The automobile, BMW in Sale process is focus on sale a product which is car that located around the USA so in the fact table name Sale Fact Table will contain Order_ID, Customer_ID, Car_ID and payment_id which are the foreign key and numerics measure will contain a year, quarter, and month to analysis the sales over past 5 year from 2018-2022 all of this information will analysis the sale process in BMW.

The “Time” dimension table that contains measure includes attributes of “Year”, “Quarter” and “Month” for analysis of sales process from customer, order, payment, and car. The Customer table includes “Cus_ID” as the primary key to get information about customers to analyze, whether it is their name, salary, address or driving style. The Order table includes “Order_ID” is the primary key for analyzing order date and quantity of order and customers can choose options to receive cars for pickup and delivery. The Car table includes information of BMW cars such as series, type, colors and price for analysis of the sale process. The Payment table related to customer and order products that contain order_ID and Customer_ID to analyze a sale process of a BMW car.

Focusing on customer data from specific states and regions, BMW can bring information about gender or driving style that can be used to analyze the customers that will make the most profit for the company. For example, BMW has the largest branch in San Antonio there are more customers than other states. This information can be used by BMW to analyze the target of customers to increase their profit. So, BMW can analyze their sales process based on customer information.

The BMW sale process star schema focused on the view of customer and car information which is information that can be used for analysis and visualize.



Link for Star Schema Product Process:

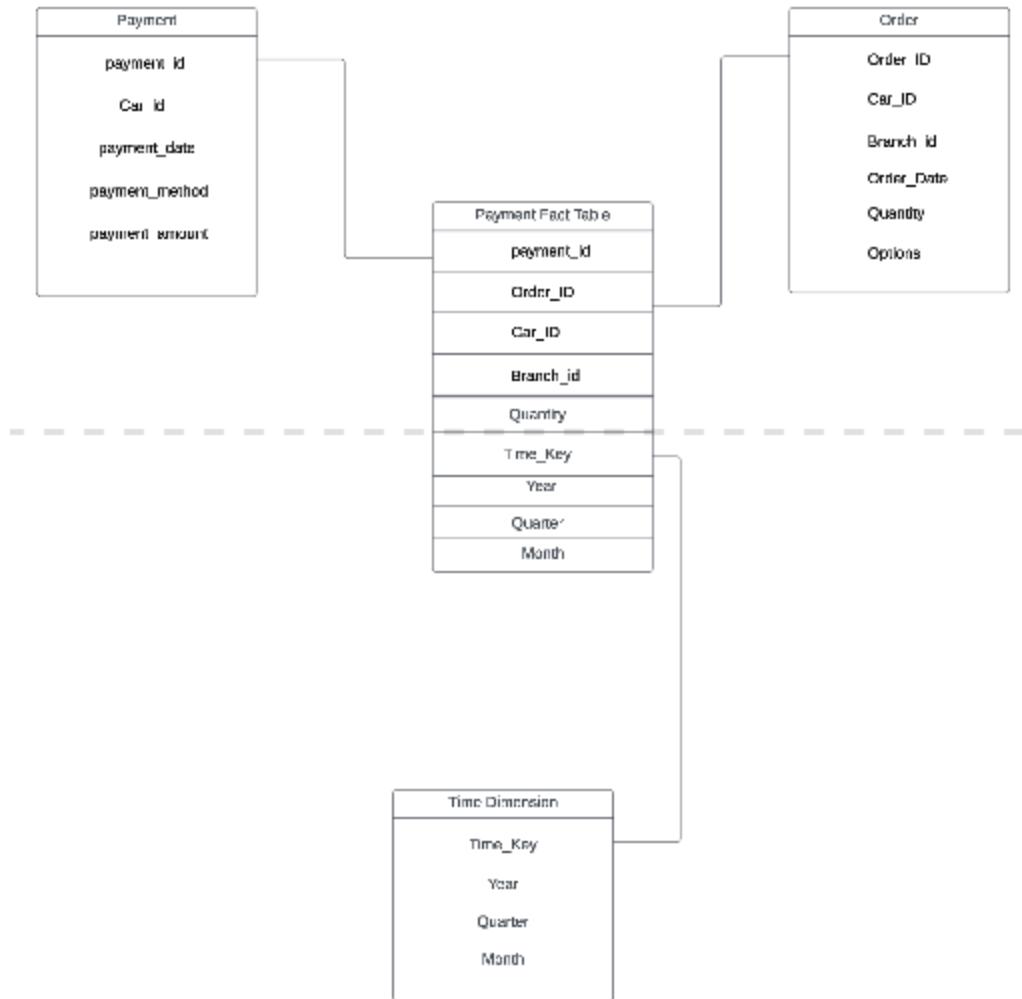
https://lucid.app/lucidchart/320772c5-8d6f-4e93-b661-159317df6d9d/edit?viewport_loc=-2307%2C-1037%2C5490%2C2657%2C0_0&invitationId=inv_774e416f-b32e-46df-a779-deab30b29b51

Product Process

The automobile, BMW in Product process is focus on the sale a product which is the car that located around the USA which is 367 branches so in the fact table name Process Fact Table will contain Order_ID, Customer_ID, Branch_ID, and Car_ID which are the foreign key and numerics measure will contain a year, quarter, and month to analysis the product over past 5 years from 2018-2022 that BMW have how much of car that stock in their warehouse.

The “Time” dimension table that contains measure includes attributes of “Year”, “Quarter” and “Month” for analysis of product process from customer, order, and branch. The Customer table includes “Cus_ID” as the primary key to get information about customers to analyze, whether it is their name, salary, address or driving style. The Order table includes “Order_ID” is the primary key for analyzing order date and quantity of order and customers can choose options to receive cars for pickup and delivery. The Branch table includes information of that branch such as location of this branch in the form of state and region. So, this process including the order so, will be able to tell how many orders each branch has. Finally, order related to the customer and go to the payment process.

This star schema helps BMW understand order and total sales of each branch which will be able to analyze profits and revenues precisely because there is clear information and can be visualized to show more about the business if they focus on order product of each branch in each state and region.



Link for Star Schema Payment Process:

https://lucid.app/lucidchart/320772c5-8d6f-4e93-b661-159317df6d9d/edit?viewport_loc=-2307%2C-1037%2C5490%2C2657%2C0_0&invitationId=inv_774e416f-b32e-46df-a779-deab30b29b51

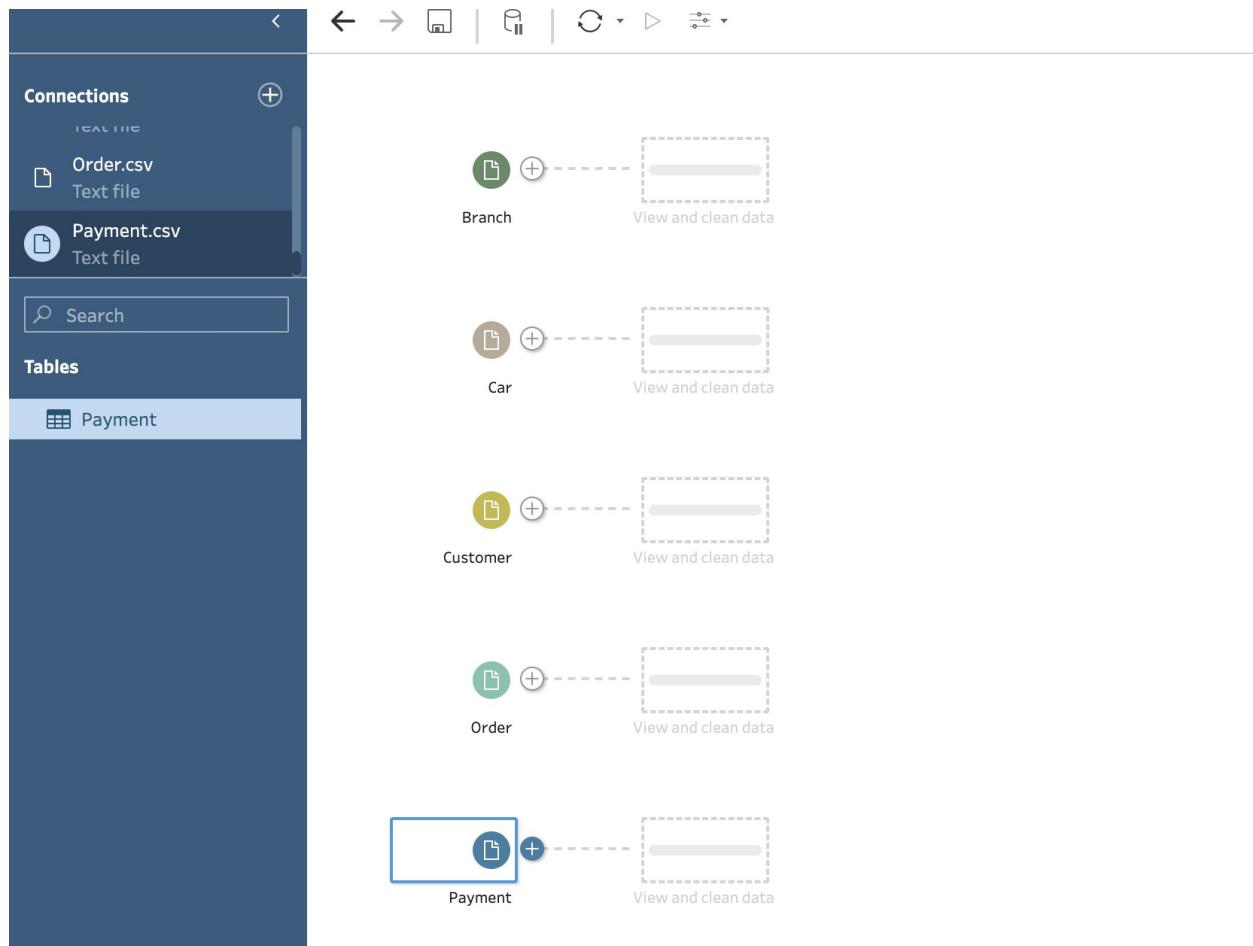
Payment Process

The automobile, BMW in Product process is focus on the sale a product which is the car that located around the USA which is 367 branches so in the fact table name Process Fact Table will contain Order_ID, Customer_ID, Branch_ID, and Car_ID which are the foreign key and numerics measure will contain a year, quarter, and month to analysis the product over past 5 years from 2018-2022 that BMW have how much of car that stock in their warehouse.

This star schema helps BMW to understand their customer and payment process that include payment, order and time. There will be two payment methods which are paid in full and pay by installment. Each order will have a different method. Therefore, BMW can analyze by bringing customer data and state or region to analyze which methods that customers pay more based on their career and salary to suit each customer's expenses. So, BMW can analyze customers by payment method for profits and they can improve their payment process to have a new target customer which brings income to BMW business.

ETL process

1. Connecting the data which are Branch, Car, Customer, Order, and Payment.

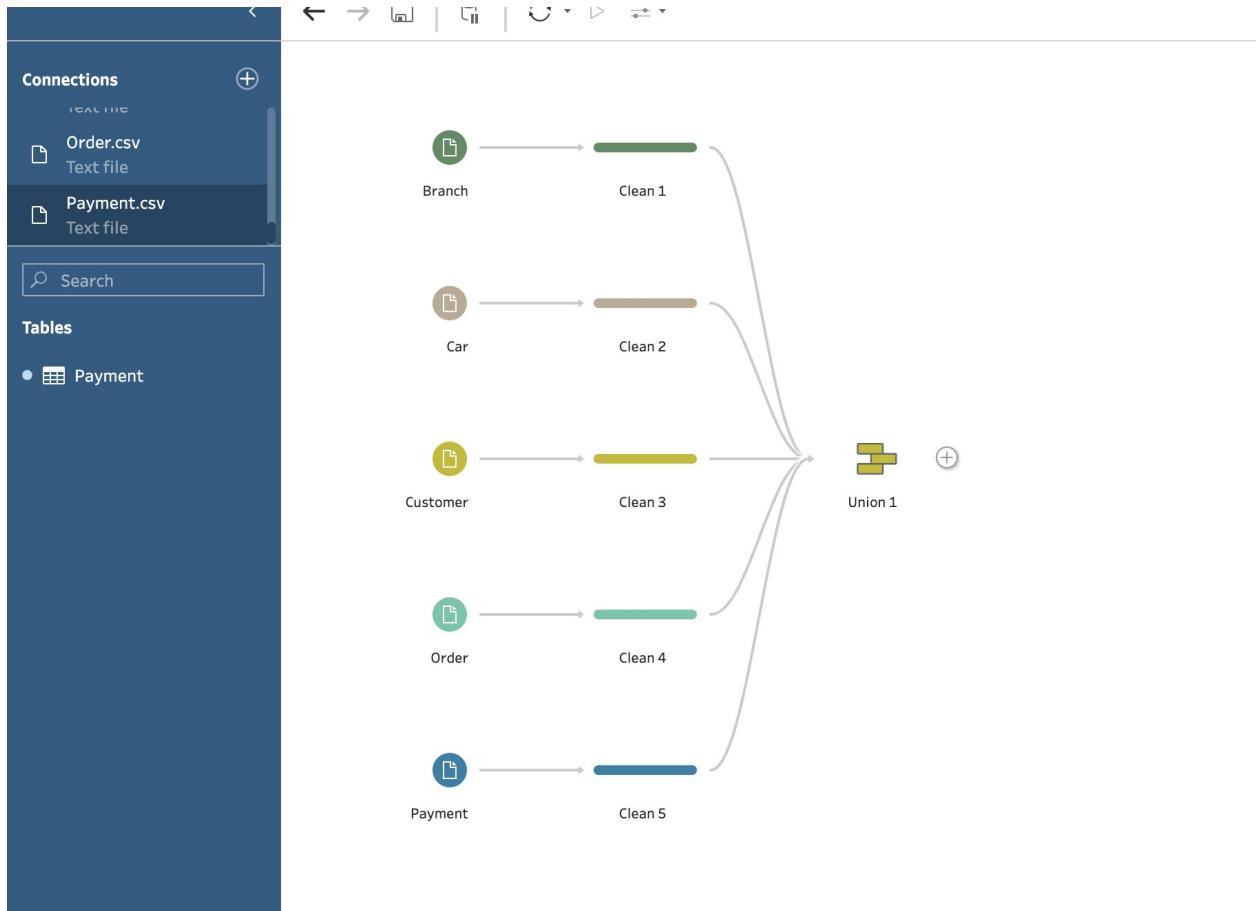


2. Clear the data and make the data in format. Such as removing unused fields and converting the null value to zero.

The screenshot shows the Alteryx Data Designer interface with the following details:

- Connections:** Order.csv (Text file) and Payment.csv (Text file) are connected to their respective 'Clean' steps.
- Tables:** Payment is selected.
- Workflow:** Five 'Clean' steps (Clean 1 through Clean 5) are connected to the tables Branch, Car, Customer, Order, and Payment respectively.
- Clean 5 Step:** Shows 38 fields and 1K rows. A message says "No data changes. Clean and transform your data and see a list of your changes here." It also displays 27 recommendations.
- Data View:** A preview of the data is shown with columns: Customer ID, Fname, Lname, Postal Code, and City. The data includes records for Sage, Emery, Bridger, Braeden, Kendrick, and Whitaker.

3. After cleaning and formatting the data, put all the file union together.

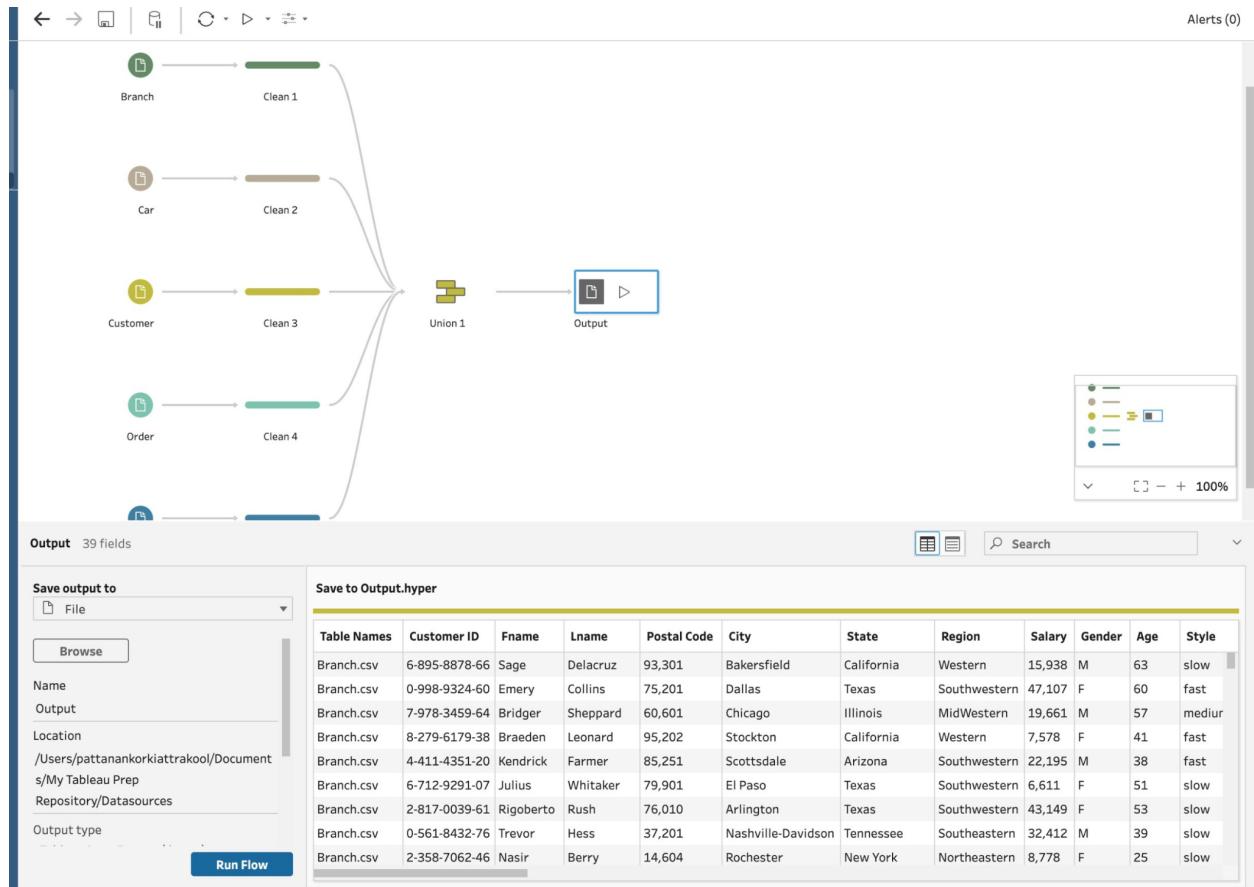


Then, the file is in union together.

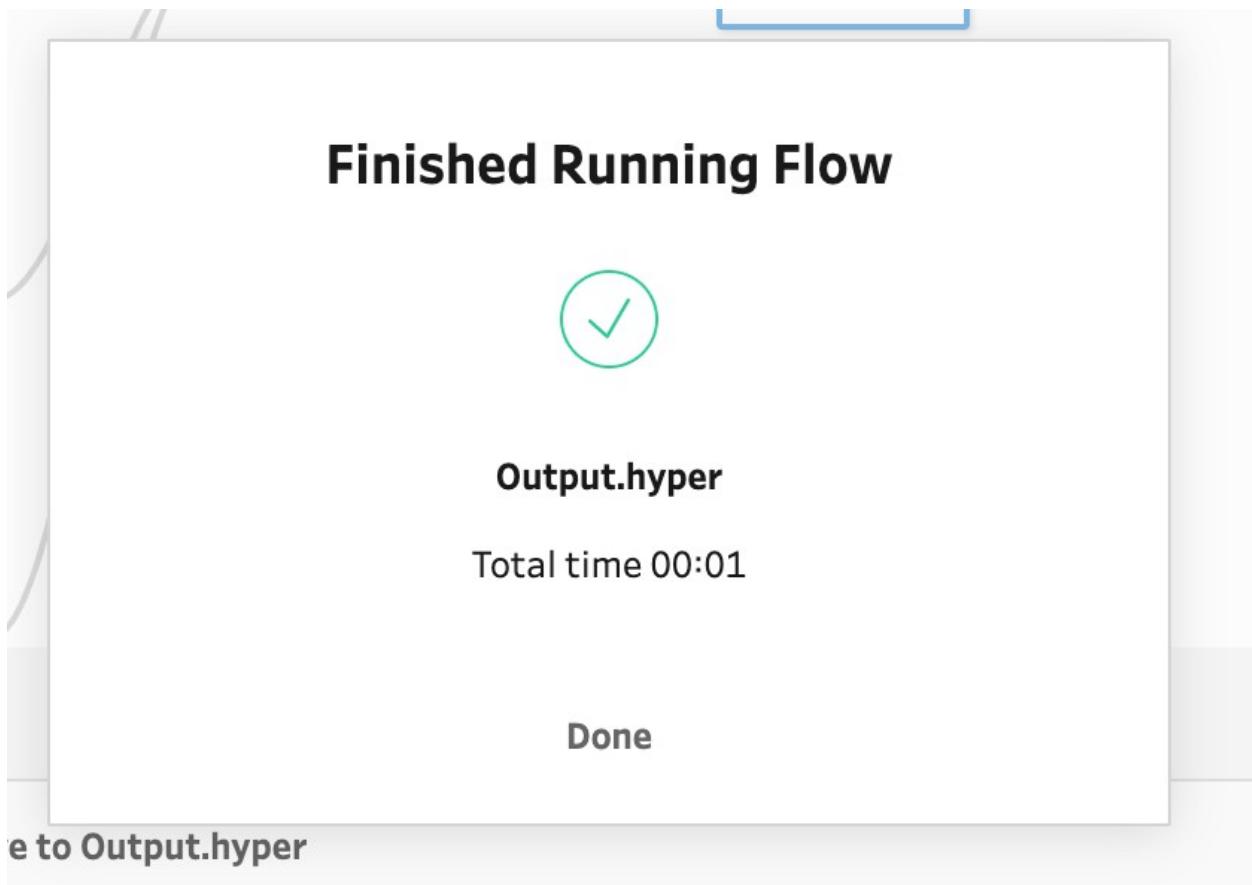
The screenshot shows a data processing interface with the following elements:

- Union 1** 39 fields 7K rows
- Filter Values...** button
- Settings** tab (selected)
- Changes (0)**
- Inputs** section:
 - Clean 1
 - Clean 2
 - Clean 3
 - Clean 4
 - Clean 5

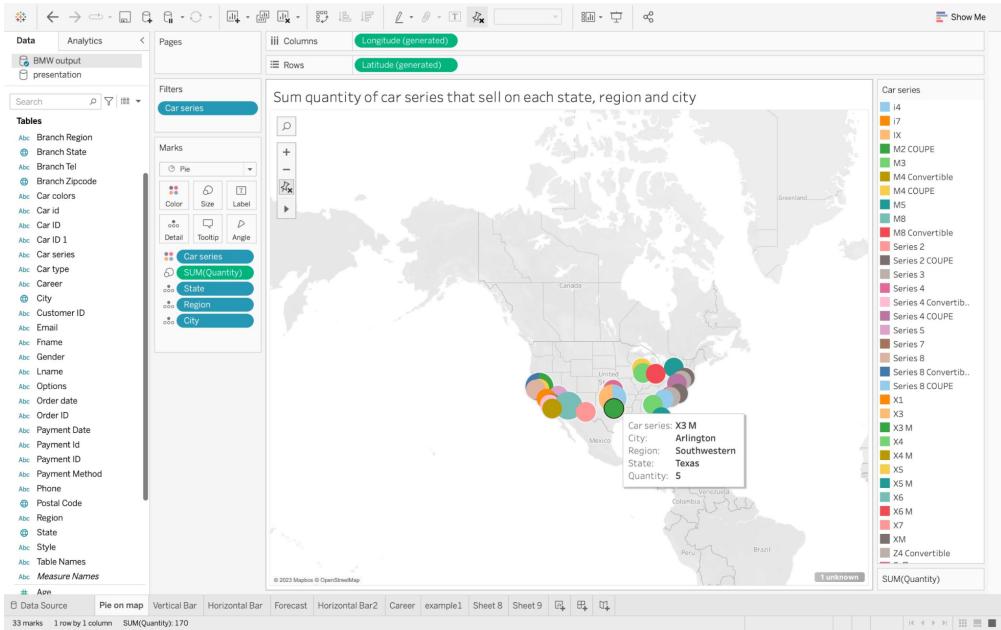
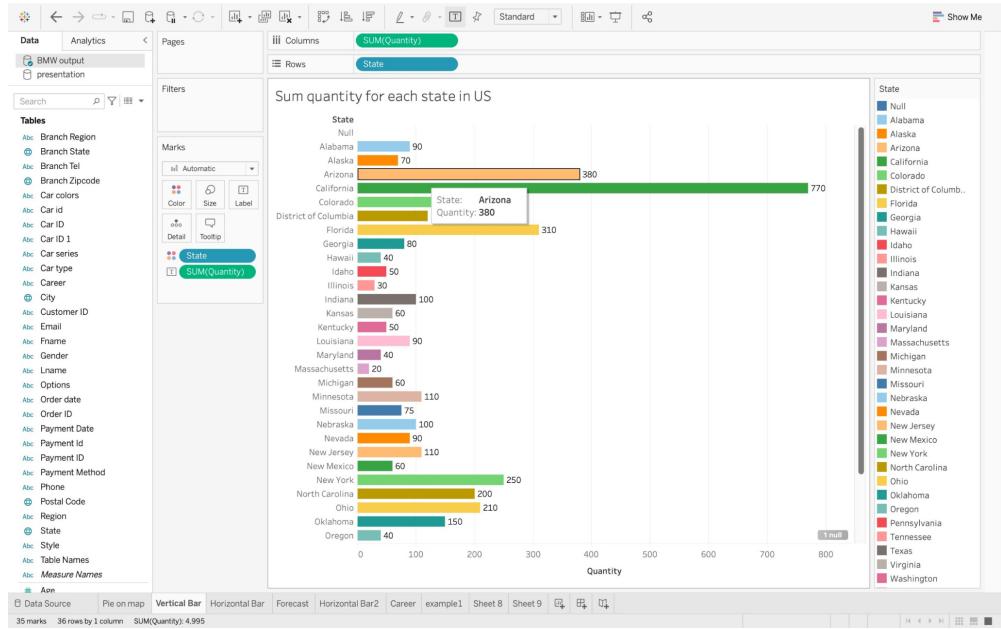
4. Last step then press the button run flow, we will get the file ready to use in Tableau Desktop for visualizing the data.

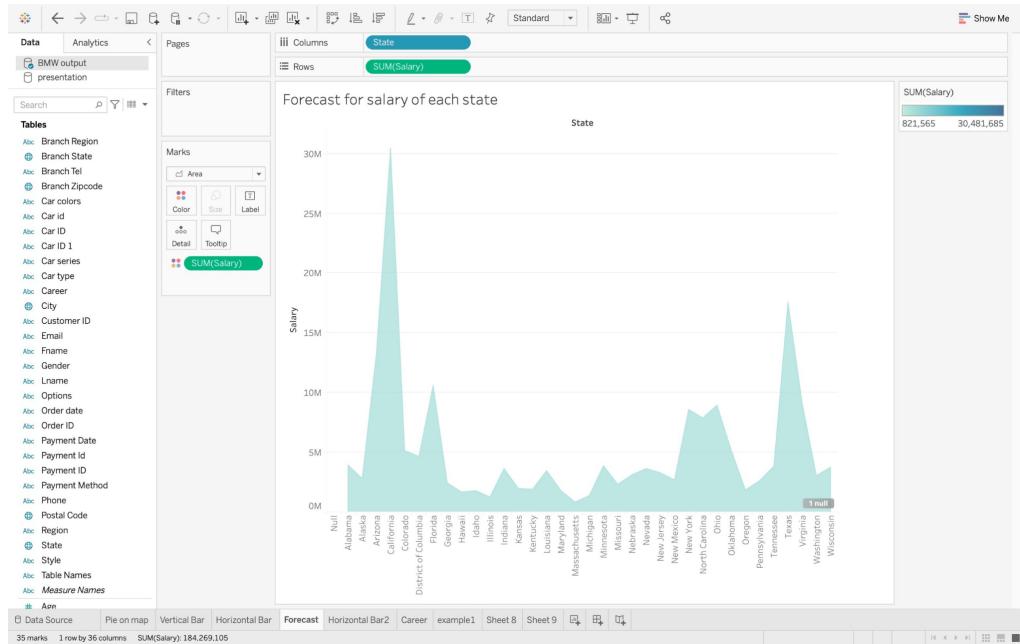
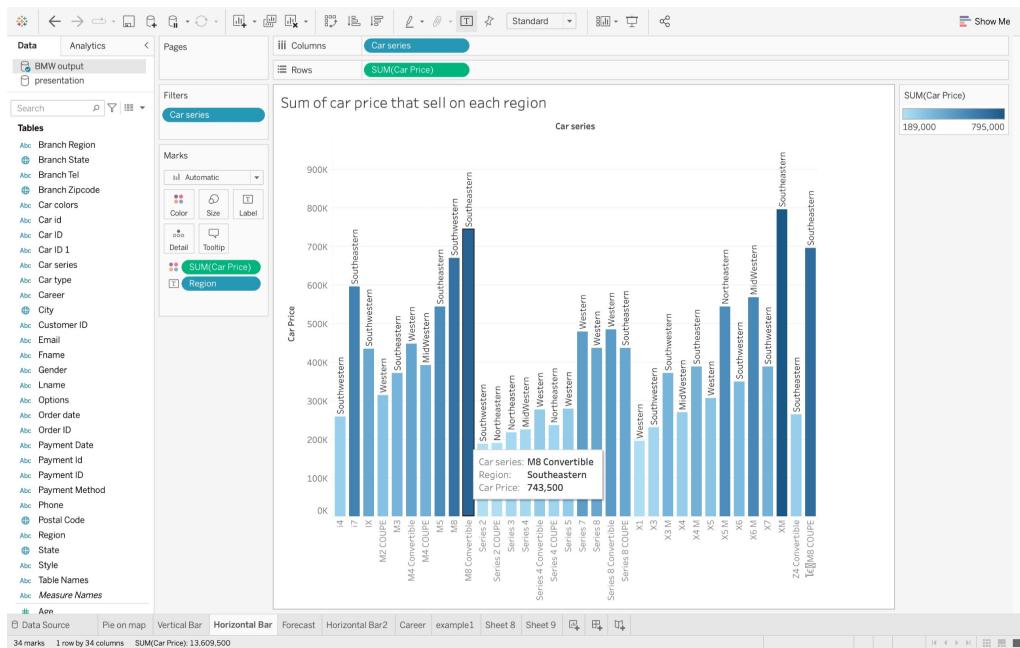


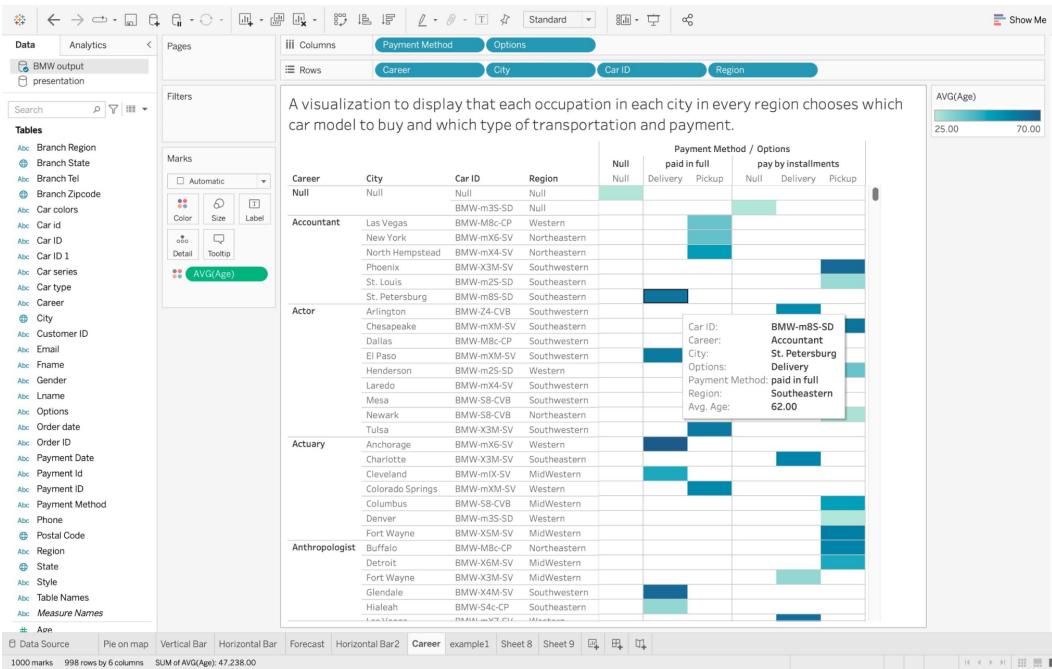
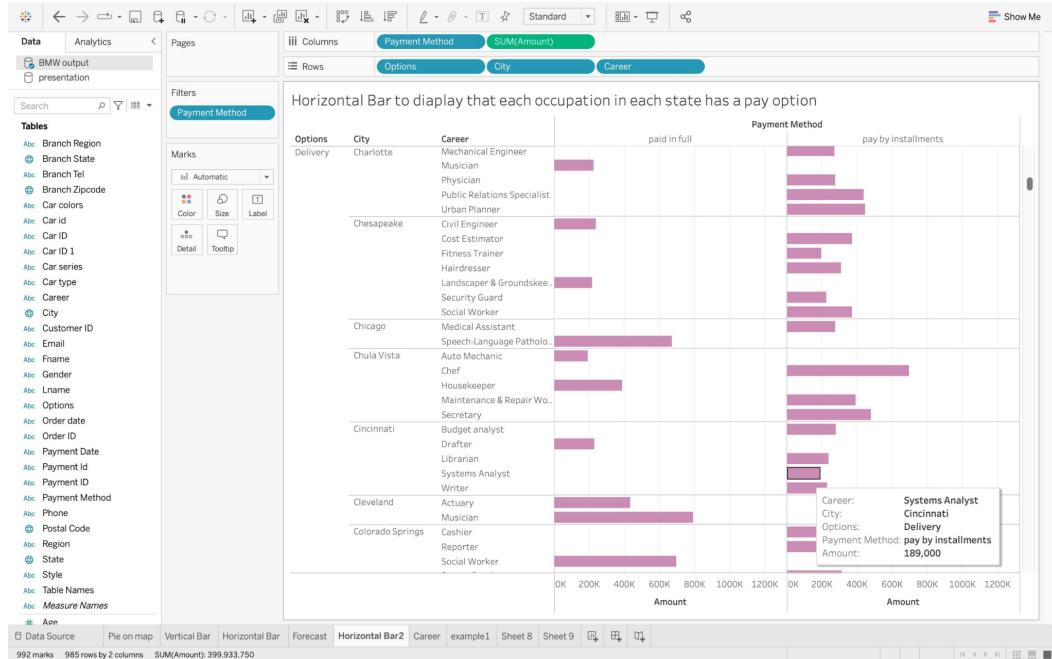
5. Finished



Analysis and Visualization reports







Discussion and conclusion on the benefits of data warehouse and BI to the company

Discussion on the benefits of data warehouse and BI

Data warehouse and business intelligence (BI) tools offer numerous benefits to BMW company. These benefits include:

- Improved Decision Making: Data warehouse and BI tools enable BMW to access accurate, timely, and relevant data from various sources. This data can be analyzed and transformed into actionable insights that inform decision making, from identifying market trends to optimizing operations.
- Enhanced Customer Experience: Data warehouse and BI tools enable BMW to gain a deeper understanding of its customers, including their preferences, behavior, and needs. This information can be used to personalize offerings, improve customer satisfaction, and build customer loyalty.
- Competitive Advantage: Data warehouse and BI tools enable BMW to stay ahead of industry trends, respond quickly to market changes, and optimize its operations for maximum efficiency and profitability. This gives BMW a competitive advantage in the global automotive industry.
- Improved Operational Efficiency: Data warehouse and BI tools enable BMW to identify areas for improvement, streamline processes, and reduce costs. This helps BMW to operate more efficiently, deliver products and services faster, and improve overall business performance.

Conclusion on the benefits of data warehouse and BI

In conclusion, data warehouses and BI tools are critical to the success of the BMW company. By leveraging these technologies, BMW can gain a competitive advantage, improve customer satisfaction, optimize its operations, and make informed decisions based on data-driven insights. By investing in a data warehouse and BI tools, BMW can enhance its brand identity and reputation as a leading luxury automotive manufacturer in the global market.

Link for the VDO presentation:

<https://loom.com/share/e12f4fbdd28f48cba00f149b7ad30d71>

Reference

1. BMW Group, [Online]. Available from:
<https://www.bmwgroup.com/en.html> [Accessed: March 1, 2023]
2. Home, [Online]. Available from:
<https://www.101domain.com/bmw.htm> [Accessed: March 1, 2023]
3. BMW Dealer Careers, "BMW Dealer Locations - Washington," BMW Dealer Careers, 2023. [Online]. Available:
https://www.bmwdealmcareers.com/dealer-locations/#state_Washington.
[Accessed: Feb 27, 2023].