

# **Spartans Sporting Goods**

## **PROJECT REPORT**



**MSIS 2621: Business Intelligence and Data Warehousing**



**Guided by:**

**Professor Sreeram Chavali**

**Submitted By:**

**Ramya Mamidipaka**

**Jai Ambekar**

**Mukesh Ganesh**

<b>Table of Contents</b>	
<b>Introduction.....</b>	<b>3</b>
<b>System Benefits .....</b>	<b>3</b>
<b>Key features of sports equipment management software .....</b>	<b>3</b>
<b>Mind Map .....</b>	<b>3</b>
<b>Part-1 Business Analysis.....</b>	<b>4</b>
<b>Scenario.....</b>	<b>4</b>
<b>Parsing Scenario to Identify key Information.....</b>	<b>4</b>
<b>Business Data.....</b>	<b>4</b>
<b>Swimlane Diagram.....</b>	<b>5</b>
<b>Interpreting the Tasks and Goals.....</b>	<b>5</b>
<b>Analytical Measures .....</b>	<b>6</b>
<b>Lead Measures .....</b>	<b>6</b>
<b>Lag Measures .....</b>	<b>6</b>
<b>Descriptive Analytics .....</b>	<b>6</b>
<b>Predictive Analytics .....</b>	<b>7</b>
<b>Business questions.....</b>	<b>7</b>
<b>Part- 2 Data Model Design .....</b>	<b>8</b>
<b>OLTP Schema .....</b>	<b>8</b>
<b>ODS Schema:.....</b>	<b>8</b>
<b>OLAP Schema .....</b>	<b>9</b>
<b>Part- 3 ETL Implementation .....</b>	<b>10</b>
<b>Part 1: Extracting data from different data sources: .....</b>	<b>11</b>
<b>Part 2: Load data from ODS Tables to Dimension Tables .....</b>	<b>16</b>
<b>Part- 4 Tableau Implementation .....</b>	<b>22</b>
<b>Dashboard 1 - Lag Measures:.....</b>	<b>22</b>
<b>Dashboard 2 - Lead Measures:.....</b>	<b>23</b>
<b>Dashboard 3 - Descriptive Analysis: .....</b>	<b>23</b>
<b>Dashboard 4 – Predictive Analysis:.....</b>	<b>24</b>
<b>Conclusion .....</b>	<b>24</b>

# Introduction

Spartan Sporting goods management system ties together the use **point of sale (POS)** coupled to a powerful **inventory management** engine in addition to a fully integrated, real-time **accounting system** which automatically makes postings as we complete **day-to-day transactions**.

Technology in maintaining sporting goods enables managers to keep everything organized. By having quick access to data on the database, improved response time for customer queries, optimized order tracking technique and equipment maintenance with which managers can enhance daily sports activities for improved performance.

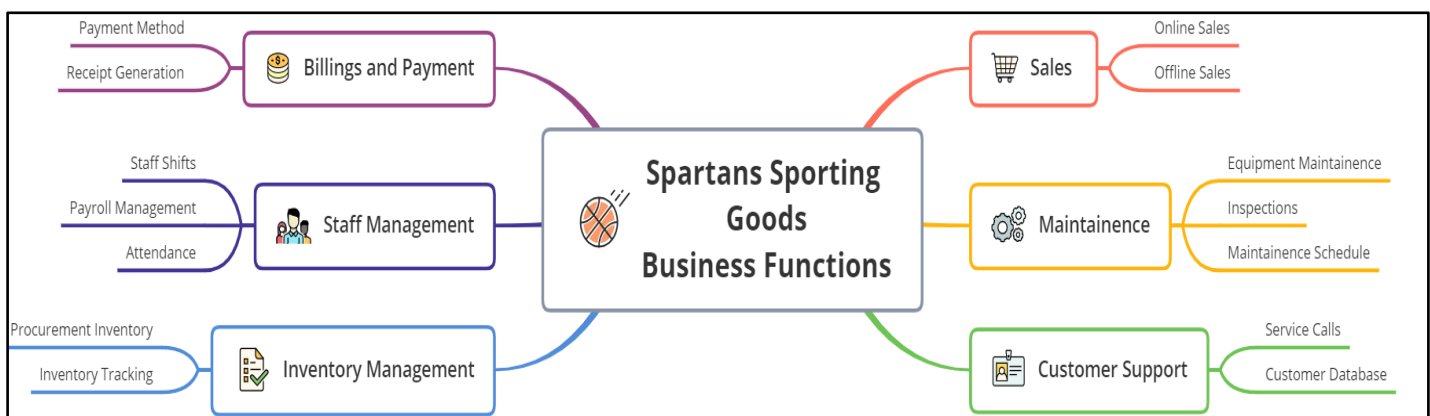
## System Benefits

1. Eliminate manual work errors
2. Benefit from real-time insights
3. Enforce security
4. Reduce Report Generation time
5. Seamless Customer Experience

## Key features of sports equipment management software

1. Sales Management
2. Inventory management
3. Payments Processing
4. Customer Support
5. Staff Management
6. Maintenance Scheduling

## Mind Map



# Part-1 Business Analysis

The Business Scenario here includes the details Spartan Sporting goods management, the process and operation flow, swim lane diagram.

## Scenario

A **Customer** enters the store to buy a new pair of shoes and tennis racket to start with a new hobby. The **Salesman** guides him through the aisles and assists him with the correct fit and model for the shoes. The Customer navigates through the store and finally gets his fit and size for shoes. Satisfied with the shoes, Customer now proceeds towards the tennis racket rack and seeks the help of Salesman for a racket for beginners. Salesman has a dialog with the customer regarding requirements for the racket types or any parameters the customer is looking for. Then Salesman shows him the existing rackets in different price ranges and according to the parameters of the customer, which are best suitable for beginners. Customer chooses his pick and proceeds towards payment. Salesman informs the **Inventory clerk** about the customer's pick and the Inventory clerk replaces the purchased items in the rack. The **Cashier** hands over the goods and receipt to the Customer. The receipt contains the bill amount for the order along with the prices for each item and the tax applied. The Customer then selects his Feedback on the screen based on his experience at the store.

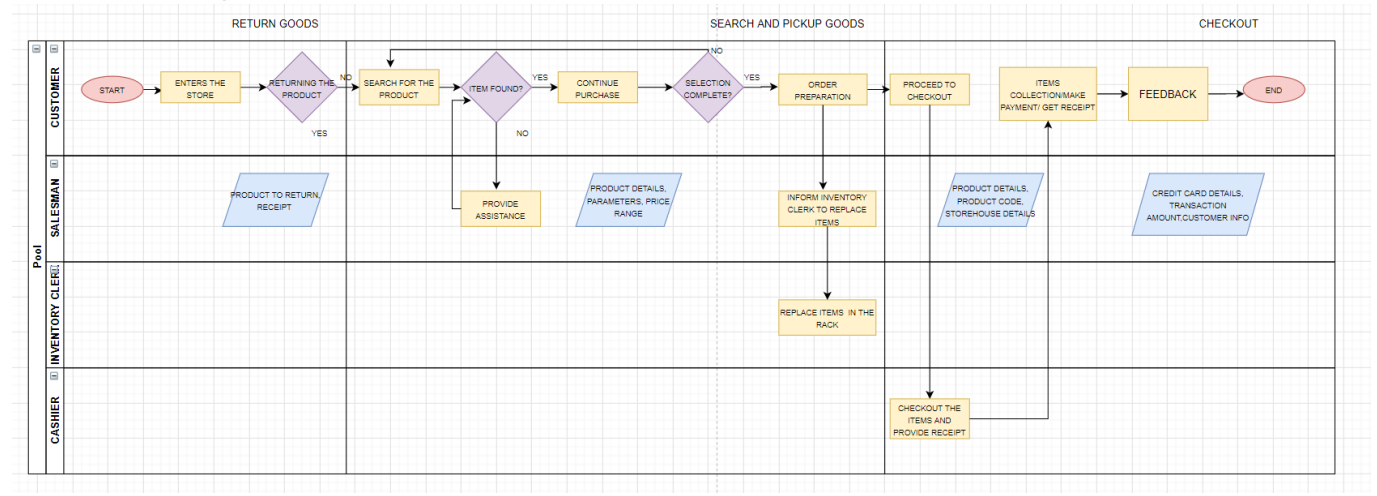
## Parsing Scenario to Identify key Information

- **Stakeholders:** Customer, Salesman, Inventory Clerk, Cashier
- **Products:** Shoes, Tennis Racket
- **Infrastructure:** Scanner, Computer, Card Reader, Receipt Printer

## Business Data

Type of Data	Entity	Data
Business Data	Order Receipt	Sale Summary Sale Item Details Sale Amount Store Detail
Inventory Data	Inventory Checklist	Items Purchased Sale Summary

## Swimlane Diagram



## Interpreting the Tasks and Goals

	Customer	Staff
<b><u>Tasks</u></b>	<ul style="list-style-type: none"> <li>● Enter Store</li> <li>● Look for items</li> <li>● Ask for items</li> <li>● Pick up items</li> <li>● Proceed to checkout counter</li> <li>● Swipe credit card (pay cash)</li> <li>● Take Receipt</li> <li>● Exit store</li> <li>● Leave in car</li> </ul>	<ul style="list-style-type: none"> <li>● Welcome the Customer</li> <li>● Help them through the racks</li> <li>● Guide them the best product according to their use</li> <li>● Explain the loyalty programs and offers</li> <li>● Proceed through the checkout process</li> <li>● Receive the feedback from the customer</li> </ul>
<b><u>Goals</u></b>	Get the best product for the most reasonable price	Enable smooth experience through the process to get the customers visit the store again

## **Analytical Measures**

### **Lead Measures**

1. Number of items in an order
2. Total orders in a day
3. Number of New customers daily
4. Inventory at the end of Day
5. Feedback score at the end of the end
6. Peak business hours for each day
7. New Customers signing up for the loyalty program Daily

### **Lag Measures**

1. Profit for the Quarter/Year
2. Frequent Customers
3. Most sold items in each category of goods every quarter/year
4. Price Dynamics based on the Orders/ Market
5. ROI/ Revenue growth
6. Expenses/ Costs incurred in a Quarter/Year
7. Stock for the coming season
8. How well was the Loyalty program received by the customers
9. Goods returned the most

### **Descriptive Analytics**

This is a reflective analysis of user data and is meant to provide insight into historical patterns of behaviors and performance in online learning environments. Descriptive analytics is leveraged when a business needs to understand the overall performance of the company at an aggregate level and describe the various aspects.

1. Trends of revenue in the total fiscal year
2. Comparing sales for the past years
3. Determining the most common payment type(cash/debit/credit)
4. Most bought goods for the customers for each region
5. Validate if a promotional campaign was successful or not
6. Customer Clusters for the promotions
7. Sales across various branches
8. Goods which were out of stock the most
9. Reviews

## **Predictive Analytics**

The subsequent step in data reduction is predictive analytics. Analyzing past data patterns and trends can accurately inform a business about what could happen in the future. This helps in setting realistic goals for the business, effective planning and restraining expectations. Predictive analytics is used by businesses to study the data and ogle into the crystal ball to find answers to the question “What could happen in the future based on previous trends and patterns?”. This is mainly undertaken to gather data on customers and predict next actions based on historical behavior.

1. Improving operations to better manage inventory and other resources
2. To set prices based on things like seasonality
3. Launch promotions that better targeted your customers
4. Introducing new goods into the chain based on inputs from the market

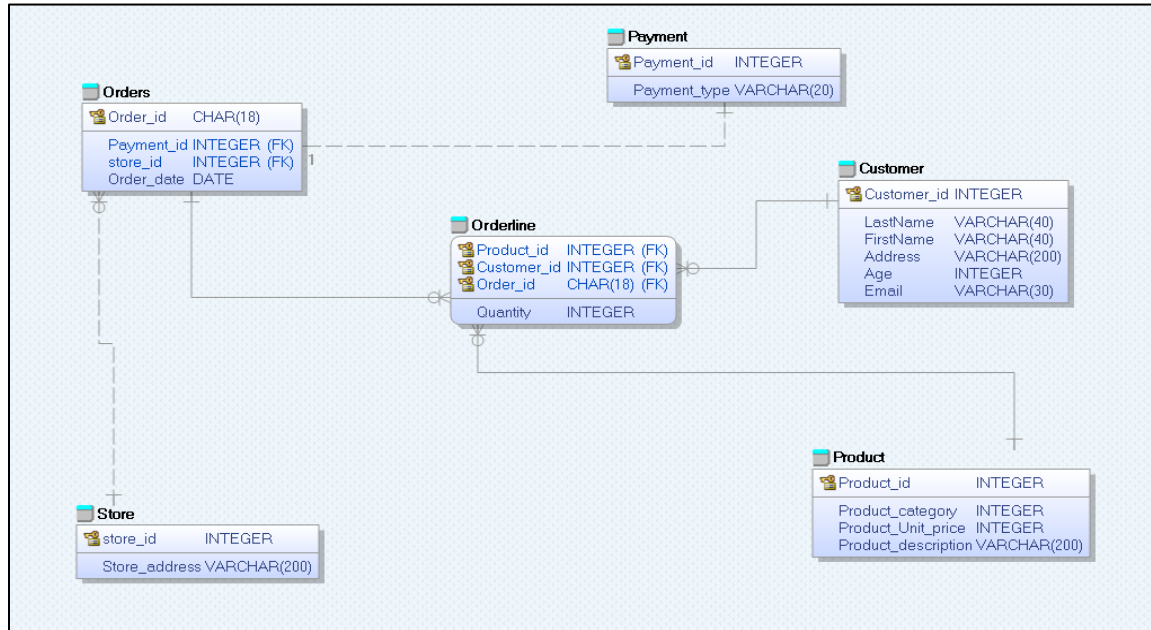
## **Business questions**

1. How many transactions are made per day/month/year?
2. What is the average sale per transaction?
3. Which product/item is the most trending in the store?
4. What is the peak time at the store?
5. What is the maximum/minimum total bill amount for day/month/year?
6. Which were the category products that were sold the most/least?
7. What is the frequency of customers which came to the store?
8. Are the items on sale being purchased more by customers?
9. What percentage of customers are returning orders?
10. Which item/category are returned the most?
11. How many transactions were made by Credit Card/Debit Card?

# Part- 2 Data Model Design

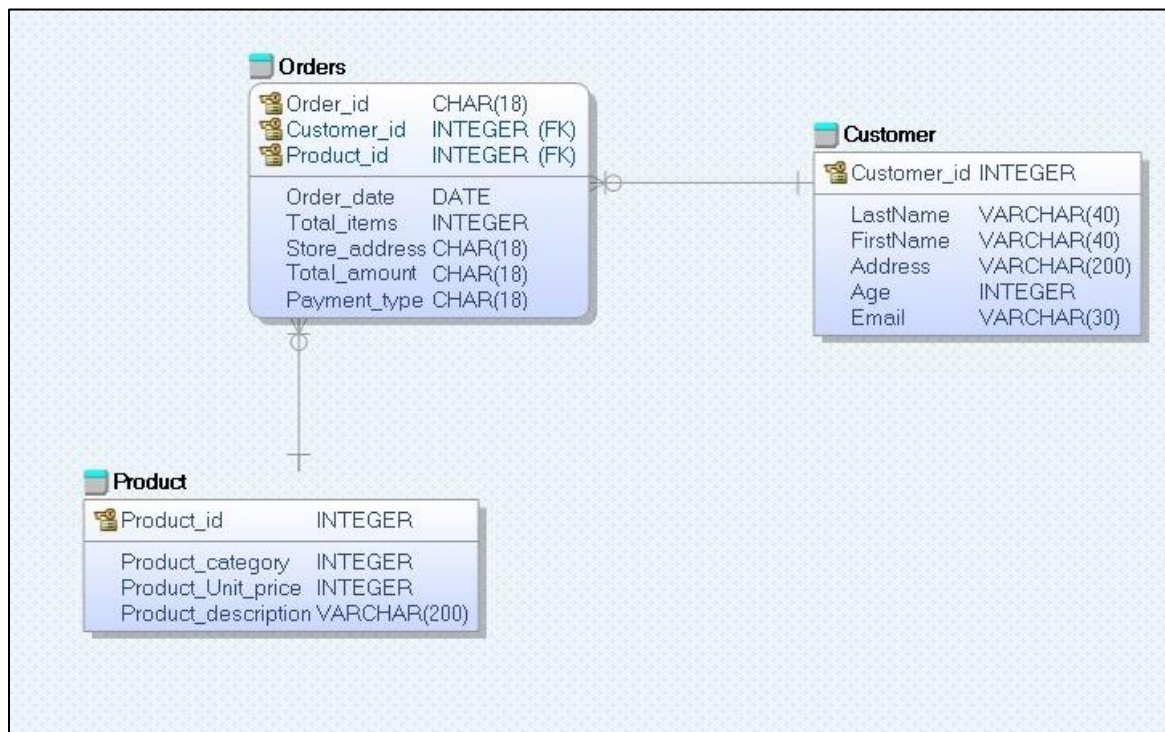
## OLTP Schema

The OLTP schema is designed in ERWIN to have 6 tables Orders, Payment, Order line, Store, Customer, Product.



## ODS Schema:

It is an interim area for a data warehouse sitting between the data sources and the data warehouse. An ODS deals exclusively with current operational data and basic status-level reporting, because an ODS continuously overwrites data. Here we merged Store details, payment, order line details into order table

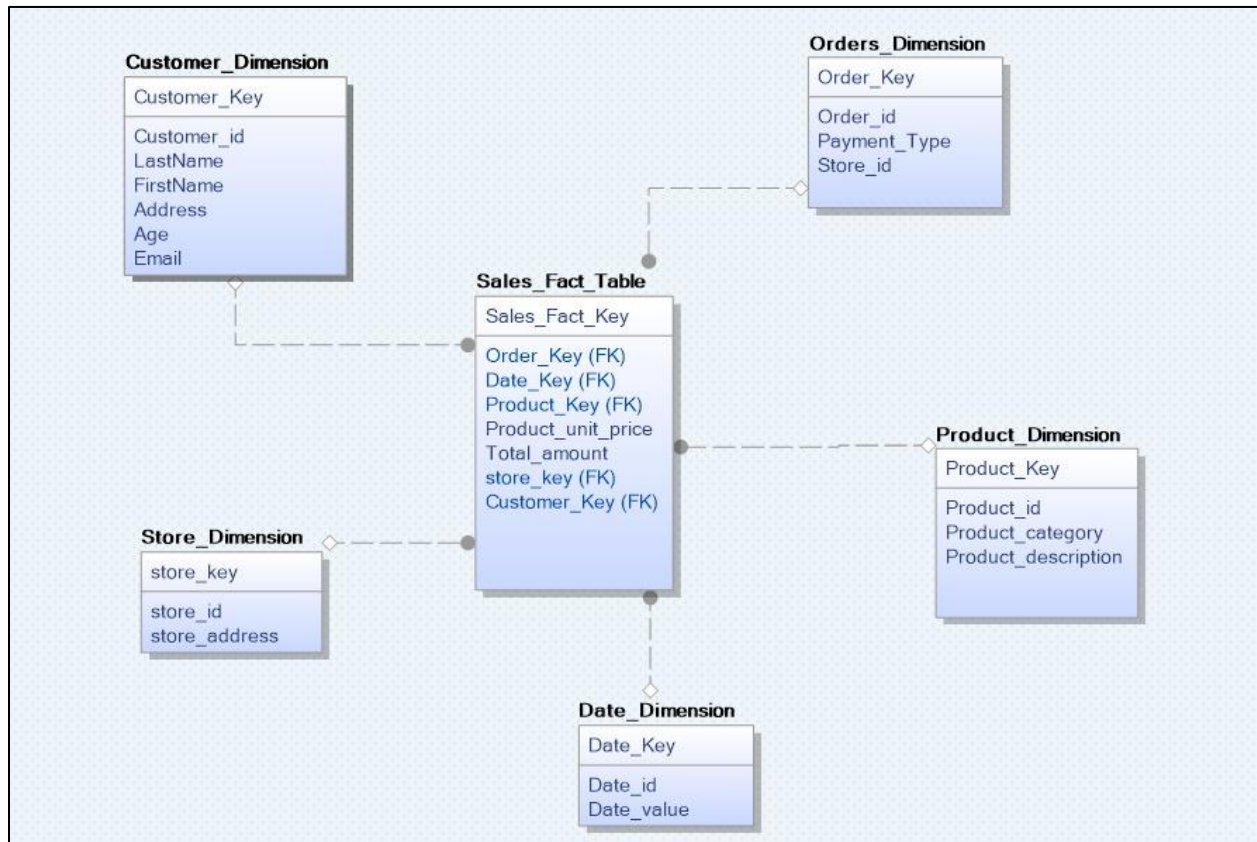




## OLAP Schema

Designed the OLAP schema for the Spartans sports management in ERWIN. There are five Dimension tables and one fact table. The Dimension tables are:

1. Customer
2. Date
3. Store
4. Product
5. Orders



# Part- 3 ETL Implementation

## Data Sources used for implementation:

1. MySQL Database
2. Json
3. CSV
4. Excel

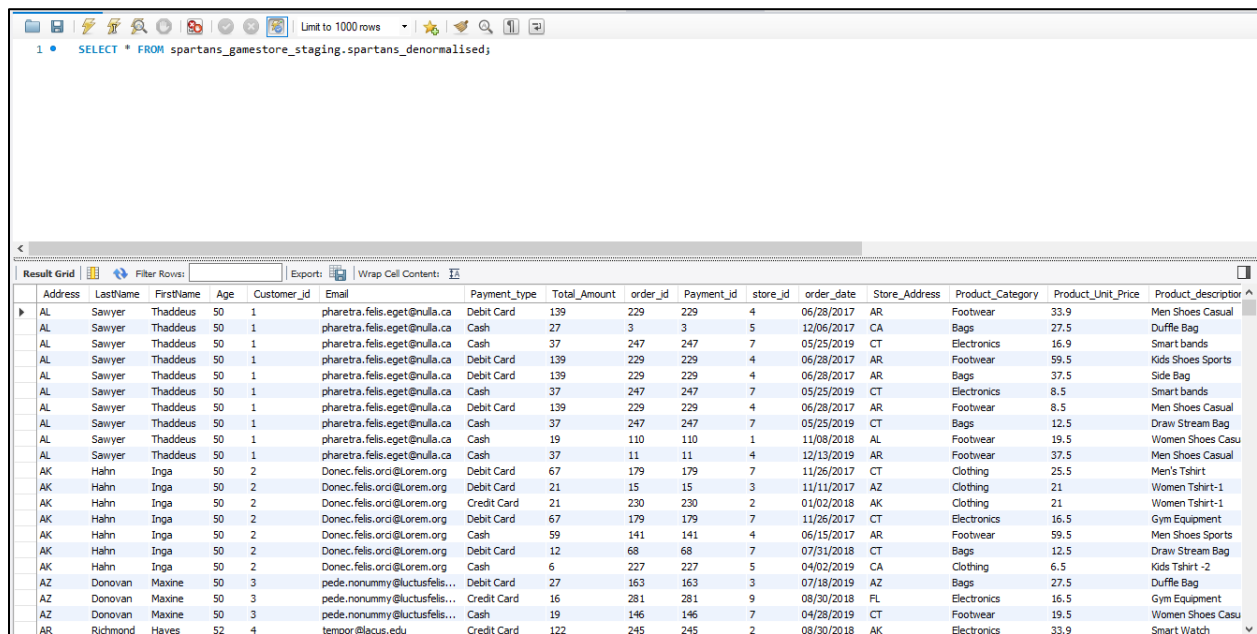
**Data Sets:** Data are manually generated from generatedata.com

**Website:** Generatedata.com

## Procedure Followed

1. Loaded the data from multiple sources mentioned above into the spartans\_gamestore database
2. Created a staging denormalized schema from the OLTP database
3. Created dimension table from the database with Surrogate keys
4. Loaded the dimension tables into spartans\_gamestore\_dw database
5. And then loaded the Fact table into the DW database

## Denormalized Table: Staging Denormalized Schema



The screenshot shows a SQL query result in a database client. The query is: `SELECT * FROM spartans_gamestore_staging.spartans_denormalised;`. The result is a table with 15 columns: Address, LastName, FirstName, Age, Customer\_id, Email, Payment\_type, Total\_Amount, order\_id, Payment\_id, store\_id, order\_date, Store\_Address, Product\_Category, Product\_Unit\_Price, and Product\_descriptor. The table contains 20 rows of data, including customers like Sawyer, Thaddeus, Hahn, Inga, Donovan, Maxine, and Richmond, with various payment types and product descriptions.

Address	LastName	FirstName	Age	Customer_id	Email	Payment_type	Total_Amount	order_id	Payment_id	store_id	order_date	Store_Address	Product_Category	Product_Unit_Price	Product_descriptor
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Debit Card	139	229	229	4	06/28/2017	AR	Footwear	33.9	Men Shoes Casual
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	27	3	3	5	12/06/2017	CA	Bags	27.5	Duffie Bag
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	37	247	247	7	05/25/2019	CT	Electronics	16.9	Smart bands
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Debit Card	139	229	229	4	06/28/2017	AR	Footwear	59.5	Kids Shoes Sports
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Debit Card	139	229	229	4	06/28/2017	AR	Bags	37.5	Side Bag
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	37	247	247	7	05/25/2019	CT	Electronics	8.5	Smart bands
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Debit Card	139	229	229	4	06/28/2017	AR	Footwear	8.5	Men Shoes Casual
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	37	247	247	7	05/25/2019	CT	Bags	12.5	Draw Stream Bag
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	19	110	110	1	11/08/2018	AL	Footwear	19.5	Women Shoes Casu
AL	Sawyer	Thaddeus	50	1	pharetra.felis.eget@nula.ca	Cash	37	11	11	4	12/13/2019	AR	Footwear	37.5	Men Shoes Casual
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Debit Card	67	179	179	7	11/26/2017	CT	Clothing	25.5	Men's Tshirt-1
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Debit Card	21	15	15	3	11/11/2017	AZ	Clothing	21	Women Tshirt-1
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Credit Card	21	230	230	2	01/02/2018	AK	Clothing	21	Women Tshirt-1
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Debit Card	67	179	179	7	11/26/2017	CT	Electronics	16.5	Gym Equipment
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Cash	59	141	141	4	06/15/2017	AR	Footwear	59.5	Men Shoes Sports
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Debit Card	12	68	68	7	07/31/2018	CT	Bags	12.5	Draw Stream Bag
AK	Hahn	Inga	50	2	Donec.felis.ori@lorem.org	Cash	6	227	227	5	04/02/2019	CA	Clothing	6.5	Kids Tshirt -2
AZ	Donovan	Maxine	50	3	pede.nonummy@luctusfelis...	Debit Card	27	163	163	3	07/18/2019	AZ	Bags	27.5	Duffie Bag
AZ	Donovan	Maxine	50	3	pede.nonummy@luctusfelis...	Credit Card	16	281	281	9	08/30/2018	FL	Electronics	16.5	Gym Equipment
AZ	Donovan	Maxine	50	3	pede.nonummy@luctusfelis...	Cash	19	146	146	7	04/28/2019	CT	Footwear	19.5	Women Shoes Casu
AR	Richmond	Hayes	52	4	tempor@lacus.edu	Credit Card	122	245	245	2	08/30/2018	AK	Electronics	33.9	Smart Watch

## Part 1: Extracting data from different data sources:

This is Phase 1 of ETL transformations, where we extracted data from various data sources and stored it in MySQL DB using **Pentaho**

### 1. Customers data: Excel file to Database

**Execution Results**

#	Customer_id	LastName	FirstName	Address	Age	Email
1	1.0	Sawyer	Thaddeus	AL	50.0	pharetra.felis.eget@nulla.ca
2	2.0	Hahn	Inga	AK	50.0	Donec.felis.orci@Lorem.org
3	3.0	Donovan	Maxine	AZ	50.0	pede.nonummy@luctusfelispurus.net
4	4.0	Richmond	Hayes	AR	52.0	tempor@lacus.edu
5	5.0	Mosley	Natalie	CA	23.0	ornare.libero.at@Fuscedolorquam.com
6	6.0	Mcclain	Bruno	CO	36.0	a.auctor@faucibus.net
7	7.0	Nicholson	Shaine	CT	28.0	amet.risus@velarcuCurabitur.net
8	8.0	Nunez	Tallulah	DE	47.0	urna@semperaurorMauris.com
9	9.0	Hubbard	Kay	FL	29.0	ligula.consectetur@etmalesuada.org
10	10.0	Fowler	Hedley	GA	56.0	augue.ac@gravidamauris.net
11	11.0	Solomon	Quail	HI	57.0	odio.sagittis.semper@atnisi.net
12	12.0	Mason	Vera	ID	43.0	ornare@ornare.com
13	13.0	Koch	Neil	IL	50.0	et.libero.Proin@duiCum.net
14	14.0	Hudson	Yolanda	IN	44.0	sem.egestas.blandit@sociisnatoque.net
15	15.0	Gibbs	Gareth	IA	60.0	volutpat@semvitae.co.uk
16	16.0	Hahn	Gwendolyn	KS	24.0	Pellentesque.ut.ipsum@non.co.uk
17	17.0	Mcgowan	Ariel	OH	37.0	malesuada@Phasellusnulla.com

## MySQL Results

**SQL Query:** `SELECT * FROM spartans_gamestore.customer;`

Customer_id	LastName	FirstName	Address	Age	Email
1	Sawyer	Thaddeus	AL	50	pharetra.felis.eget@nulla.ca
2	Hahn	Inga	AK	50	Donec.felis.orci@Lorem.org
3	Donovan	Maxine	AZ	50	pede.nonummy@luctusfelispurus.net
4	Richmond	Hayes	AR	52	tempor@lacus.edu
5	Mosley	Natalie	CA	23	ornare.libero.at@Fuscedolorquam.com
6	Mcclain	Bruno	CO	36	a.auctor@faucibus.net
7	Nicholson	Shaine	CT	28	amet.risus@velarcuCurabitur.net
8	Nunez	Tallulah	DE	47	urna@semperaurorMauris.com
9	Hubbard	Kay	FL	29	ligula.consectetur@etmalesuada.org
10	Fowler	Hedley	GA	56	augue.ac@gravidamauris.net
11	Solomon	Quail	HI	57	odio.sagittis.semper@atnisi.net
12	Mason	Vera	ID	43	ornare@ornare.com
13	Koch	Neil	IL	50	et.libero.Proin@duiCum.net
14	Hudson	Yolanda	IN	44	sem.egestas.blandit@sociisnatoque.net
15	Gibbs	Gareth	IA	60	volutpat@semvitae.co.uk
16	Hahn	Gwendolyn	KS	24	Pellentesque.ut.ipsum@non.co.uk
17	Mcgowan	Ariel	OH	37	malesuada@Phasellusnulla.com

## 2. Order line data: Excel file to Database

The screenshot shows a data integration tool interface. At the top, there's a toolbar with various icons and a list of tables: Welcome!, Payment\_Table\_jso..., Product\_dim\_to\_DW..., Customer\_dim\_to\_D..., Order\_dim\_to\_DW, Date\_dim\_to\_DW, Fact\_table\_DW, store\_dim\_to\_DW, and Product\_table\_csv... Below the toolbar, a workflow diagram shows an input file 'input-Orderline\_table\_excel\_to\_DB' connected to an output database 'Output-Spartans\_DB'. The 'Execution Results' section at the bottom displays a table with 13 rows and 4 columns: Customer\_id, product\_id, order\_id, and a fourth column with values ranging from 1.0 to 13.0.

	Customer_id	product_id	order_id	
1	99.0	46.0	1.0	
2	70.0	15.0	2.0	
3	1.0	3.0	3.0	
4	95.0	13.0	4.0	
5	5.0	34.0	5.0	
6	34.0	24.0	6.0	
7	26.0	25.0	7.0	
8	67.0	8.0	8.0	
9	79.0	43.0	9.0	
10	54.0	3.0	10.0	
11	1.0	42.0	11.0	
12	52.0	15.0	12.0	
13	31.0	42.0	13.0	

## Mysql Results

The screenshot shows a MySQL query result. The query is 'SELECT \* FROM spartans\_gamestore.order\_line;'. The result is displayed in a table with 4 columns: Customer\_id, product\_id, order\_id, and Quantity. The table contains 17 rows of data.

Customer_id	product_id	order_id	Quantity
99	46	1	3
70	15	2	4
1	3	3	4
95	13	4	4
5	34	5	2
34	24	6	2
26	25	7	3
67	8	8	3
79	43	9	1
54	3	10	4
1	42	11	2
52	15	12	2
31	42	13	4
87	43	14	1
2	11	15	1
51	4	16	4
7	34	17	2

### 3. Payment data: Json file to Database

The screenshot shows a data integration tool interface. At the top, there's a workflow diagram with two nodes: 'input-Payment\_table\_JSON' and 'Output-Spartans\_D8', connected by an arrow. Below the diagram is the 'Execution Results' section, which includes tabs for 'Logging', 'Execution History', 'Step Metrics', 'Performance Graph', 'Metrics', and 'Preview data'. The 'Preview data' tab is selected, showing a table with 13 rows of payment data.

#	Payment_id	Payment_type	Total_Amount
1	1	Debit Card	180
2	2	Credit Card	42
3	3	Cash	27
4	4	Credit Card	37
5	5	Cash	29
6	6	Cash	49
7	7	Cash	49
8	8	Debit Card	18
9	9	Debit Card	4
10	10	Debit Card	27
11	11	Cash	37
12	12	Cash	42
13	13	Credit Card	37

### Mysql Results

The screenshot shows a MySQL query editor. At the top, there's a toolbar with various icons and a 'Limit to 1000 rows' dropdown. Below the toolbar is a text area containing the SQL query: `1 • SELECT * FROM spartans_gamestore.payment;`. At the bottom, there's a 'Result Grid' section showing the results of the query in a table format. The table has 17 rows and 3 columns: 'Payment\_id', 'Payment\_type', and 'Total\_Amount'.

Payment_id	Payment_type	Total_Amount
1	Debit Card	180
2	Credit Card	42
3	Cash	27
4	Credit Card	37
5	Cash	29
6	Cash	49
7	Cash	49
8	Debit Card	18
9	Debit Card	4
10	Debit Card	27
11	Cash	37
12	Cash	42
13	Credit Card	37
14	Credit Card	4
15	Debit Card	21
16	Cash	33
17	Cash	29

## 4. Product data: CSV file to Database

The screenshot shows a data integration tool interface. At the top, there's a workflow diagram with two nodes: 'Input-Product\_Table-CSV' and 'Output-Spartans\_DB', connected by an arrow. Below this, the 'Execution Results' tab is active, displaying a table of data. The table has columns: Product\_ID, Product\_Category, Product\_Unit\_Price, and Product\_description. The data is as follows:

#	Product_ID	Product_Category	Product_Unit_Price	Product_description
1	1	Clothing	25.5	Men's Tshirt
2	2	Footwear	33.9	Men Shoes Casual
3	3	Bags	27.5	Duffle Bag
4	4	Electronics	33.9	Smart Watch
5	5	Clothing	33.9	Women Tshirt-2
6	6	Footwear	76.5	Women Shoes Casual
7	7	Bags	42.5	Side Bag
8	8	Footwear	18.5	Kids Shoes Casual
9	9	Bags	18.5	Back pack
10	10	Electronics	16.9	Smart bands
11	11	Clothing	21.0	Women Tshirt-1
12	12	Footwear	21.0	Kids Shoes Sports
13	13	Bags	37.5	Hiking backpack

## Mysql Results

The screenshot shows a MySQL query result. The query is: `SELECT * FROM spartans_gamestore.products;`. The result is displayed in a grid format with columns: Product\_ID, Product\_Category, Product\_Unit\_Price, and Product\_description. The data is as follows:

Product_ID	Product_Category	Product_Unit_Price	Product_description
1	Clothing	25.5	Men's Tshirt
2	Footwear	33.9	Men Shoes Casual
3	Bags	27.5	Duffle Bag
4	Electronics	33.9	Smart Watch
5	Clothing	33.9	Women Tshirt-2
6	Footwear	76.5	Women Shoes Casual
7	Bags	42.5	Side Bag
8	Footwear	18.5	Kids Shoes Casual
9	Bags	18.5	Back pack
10	Electronics	16.9	Smart bands
11	Clothing	21	Women Tshirt-1
12	Footwear	21	Kids Shoes Sports
13	Bags	37.5	Hiking backpack
14	Electronics	16.5	Gym Equipment
15	Clothing	42.5	Men's Tshirt-2
16	Footwear	49.5	Women Shoes Sports
17	Bags	99.5	Duffle Bag

## 5. Stores data: Excel file to Database

The screenshot displays the Apache Spark SQL IDE interface. At the top, a diagram shows a transformation from 'input-Store\_Table\_excel' to 'output-Spartans\_DB', both marked with green checkmarks. Below this, the 'Execution Results' panel is active, showing a log of events:

- 2020/11/25 18:53:27 - Spoon - Using legacy execution engine
- 2020/11/25 18:53:27 - Spoon - Transformation opened.
- 2020/11/25 18:53:27 - Spoon - Launching transformation [Store\_Table\_excel\_to\_DB]...
- 2020/11/25 18:53:27 - Spoon - Started the transformation execution.
- 2020/11/25 18:53:28 - Store\_Table\_excel\_to\_DB - Dispatching started for transformation [Store\_Table\_excel\_to\_DB]
- 2020/11/25 18:53:28 - input-Store\_Table\_excel.0 - Finished processing (I=10, O=0, R=0, W=10, U=0, E=0)
- 2020/11/25 18:53:28 - output-Spartans\_DB.0 - Finished processing (I=10, O=10, R=10, W=10, U=0, E=0)
- 2020/11/25 18:53:28 - Spoon - The transformation has finished!!

## Mysql Results

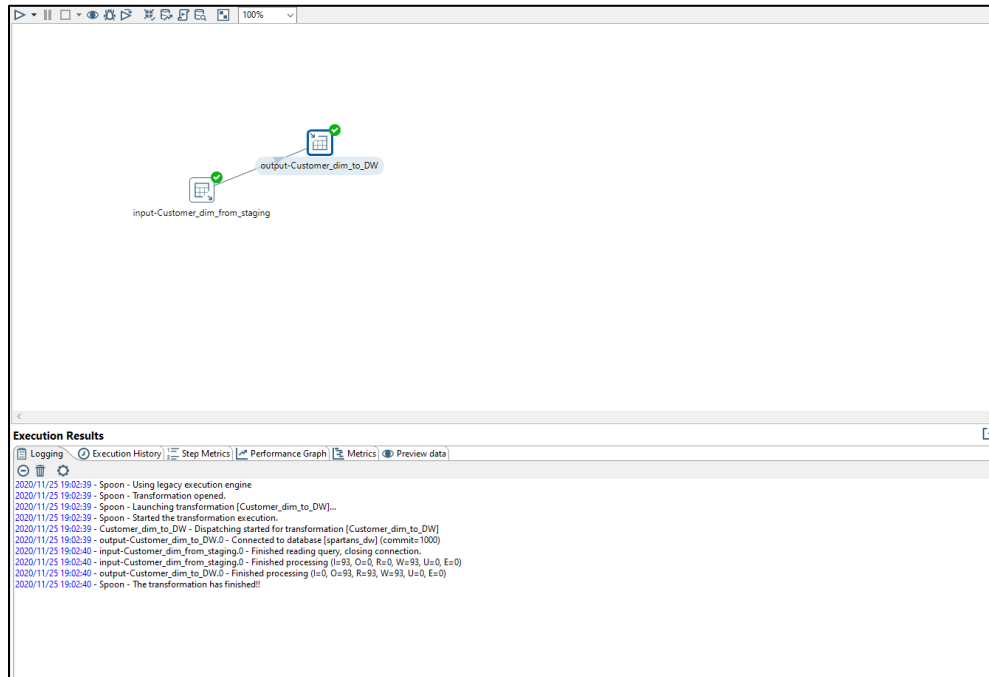
The screenshot shows the MySQL Workbench interface. A query is executed: `SELECT * FROM spartans_gamestore.store;`. The 'Result Grid' panel displays the following data:

Store_ID	Store_Address
1	AL
2	AK
3	AZ
4	AR
5	CA
6	CO
7	CT
8	DE
9	FL
10	GA

## Part 2: Load data from ODS Tables to Dimension Tables

In Phase 2 of ETL Transformations, we created Staging schema from all the input tables and then we loaded data in dimensions tables and the fact table using pentaho.

### 1. Customer dimension data to Data Warehouse



### Mysql Results

<



## 2. Date dimension data from excel to Data Warehouse

The screenshot shows a data integration tool interface. At the top, a workflow diagram illustrates the process: 'input-dim\_date\_from\_excel' (represented by a document icon) connects to 'output-Date\_dim\_to\_DW' (represented by a database icon). Below the workflow, the 'Execution Results' tab is active, displaying a table of data. The table has two columns: 'date\_key' and 'datevalue'. The data represents dates from 01/01/2017 to 01/13/2017.

	date_key	datevalue
1	1	01/01/2017
2	2	01/02/2017
3	3	01/03/2017
4	4	01/04/2017
5	5	01/05/2017
6	6	01/06/2017
7	7	01/07/2017
8	8	01/08/2017
9	9	01/09/2017
10	10	01/10/2017
11	11	01/11/2017
12	12	01/12/2017
13	13	01/13/2017

## Mysql Results

The screenshot shows a MySQL query result. The query is: `SELECT * FROM spartans_gamestore_dw.dim_date;`. The result is displayed in a table with two columns: 'date\_key' and 'datevalue'. The data represents dates from 01/01/2017 to 01/17/2017.

	date_key	datevalue
1	1	01/01/2017
2	2	01/02/2017
3	3	01/03/2017
4	4	01/04/2017
5	5	01/05/2017
6	6	01/06/2017
7	7	01/07/2017
8	8	01/08/2017
9	9	01/09/2017
10	10	01/10/2017
11	11	01/11/2017
12	12	01/12/2017
13	13	01/13/2017
14	14	01/14/2017
15	15	01/15/2017
16	16	01/16/2017
17	17	01/17/2017

### 3. Order dimension data to Data Warehouse

The screenshot shows a data integration workflow in a tool like Talend. A job icon with a green checkmark is connected to a table icon, also with a green checkmark. The workflow is labeled 'input-order\_from\_staging' and 'output-dim\_orders\_to\_DW'. Below the workflow, the 'Execution Results' section is visible, showing a table of data.

#	order_id	Payment_type	Quantity	Product_Unit_Price	Order_key
1	229	Debit Card	4	33.9	1
2	3	Cash	4	27.5	2
3	247	Cash	4	16.9	3
4	229	Debit Card	3	59.5	4
5	229	Debit Card	3	37.5	5
6	247	Cash	2	8.5	6
7	229	Debit Card	1	8.5	7
8	247	Cash	3	12.5	8
9	110	Cash	4	19.5	9
10	11	Cash	2	37.5	10
11	179	Debit Card	2	25.5	11
12	15	Debit Card	1	21.0	12
13	230	Credit Card	1	21.0	13

### Mysql Results

The screenshot shows a MySQL query editor with a query window at the top containing the statement: `SELECT * FROM spartans_gamestore_dw.dim_orders;`. Below the query window, the 'Result Grid' is displayed, showing the results of the query in a table format.

Order_key	order_id	Payment_type	Quantity	Product_Unit_Price
1	229	Debit Card	4	33.9
2	3	Cash	4	27.5
3	247	Cash	4	16.9
4	229	Debit Card	3	59.5
5	229	Debit Card	3	37.5
6	247	Cash	2	8.5
7	229	Debit Card	1	8.5
8	247	Cash	3	12.5
9	110	Cash	4	19.5
10	11	Cash	2	37.5
11	179	Debit Card	2	25.5
12	15	Debit Card	1	21
13	230	Credit Card	1	21
14	179	Debit Card	3	16.5
15	141	Cash	3	59.5
16	68	Debit Card	2	12.5
17	227	Cash	4	6.5

## 4. Product dimension data to Data Warehouse

The screenshot shows a data pipeline execution results window. At the top, a flow diagram illustrates the process: 'Input-product\_dim\_from\_staging' (represented by a green icon) flows into 'Output-dim\_product\_to\_DW' (represented by a blue icon). Below this, the 'Execution Results' section is active, displaying a table of product data. The table has columns for product\_id, product\_category, product\_description, Product\_Unit\_Price, and Product\_Key. The data is organized into 13 rows, each representing a different product category and its details.

#	product_id	product_category	product_description	Product_Unit_Price	Product_Key
1	2	Footwear	Men Shoes Casual	33.9	1
2	3	Bags	Duffle Bag	27.5	2
3	10	Electronics	Smart bands	16.9	3
4	18	Footwear	Kids Shoes Sports	59.5	4
5	29	Bags	Side Bag	37.5	5
6	30	Electronics	Smart bands	8.5	6
7	32	Footwear	Men Shoes Casual	8.5	7
8	33	Bags	Draw Stream Bag	12.5	8
9	38	Footwear	Women Shoes Casual	19.5	9
10	42	Footwear	Men Shoes Casual	37.5	10
11	1	Clothing	Men's Tshirt	25.5	11
12	11	Clothing	Women Tshirt-1	21.0	12
13	14	Electronics	Gym Equipment	16.5	13

## Mysql Results

The screenshot shows a MySQL query results window. The query executed is `SELECT * FROM spartans_gamestore_dw.dim_product;`. The results are displayed in a table with columns: Product\_Key, product\_id, product\_category, product\_description, and Product\_Unit\_Price. The table contains 17 rows of data, representing various product categories and their details.

Product_Key	product_id	product_category	product_description	Product_Unit_Price
1	2	Footwear	Men Shoes Casual	33.9
2	3	Bags	Duffle Bag	27.5
3	10	Electronics	Smart bands	16.9
4	18	Footwear	Kids Shoes Sports	59.5
5	29	Bags	Side Bag	37.5
6	30	Electronics	Smart bands	8.5
7	32	Footwear	Men Shoes Casual	8.5
8	33	Bags	Draw Stream Bag	12.5
9	38	Footwear	Women Shoes Casual	19.5
10	42	Footwear	Men Shoes Casual	37.5
11	1	Clothing	Men's Tshirt	25.5
12	11	Clothing	Women Tshirt-1	21
13	14	Electronics	Gym Equipment	16.5
14	26	Footwear	Men Shoes Sports	59.5
15	45	Clothing	Kids Tshirt -2	6.5
16	4	Electronics	Smart Watch	33.9
17	25	Clothing	Women Tshirt-2	49.5

## 5. Store dimension data to Data Warehouse

The screenshot shows a data integration workflow with two steps: 'input-store\_dim\_from\_staging' and 'output-store\_dim\_to\_DW'. Below the workflow is the 'Execution Results' section, which includes tabs for Logging, Execution History, Step Metrics, Performance Graph, Metrics, and Preview data. The 'First rows' tab is selected, displaying a table with 10 rows of data.

#	store_id	store_address	store_key
1	1	AL	1
2	2	AK	2
3	3	AZ	3
4	4	AR	4
5	5	CA	5
6	6	CO	6
7	7	CT	7
8	8	DE	8
9	9	FL	9
10	10	GA	10

## Mysql Results

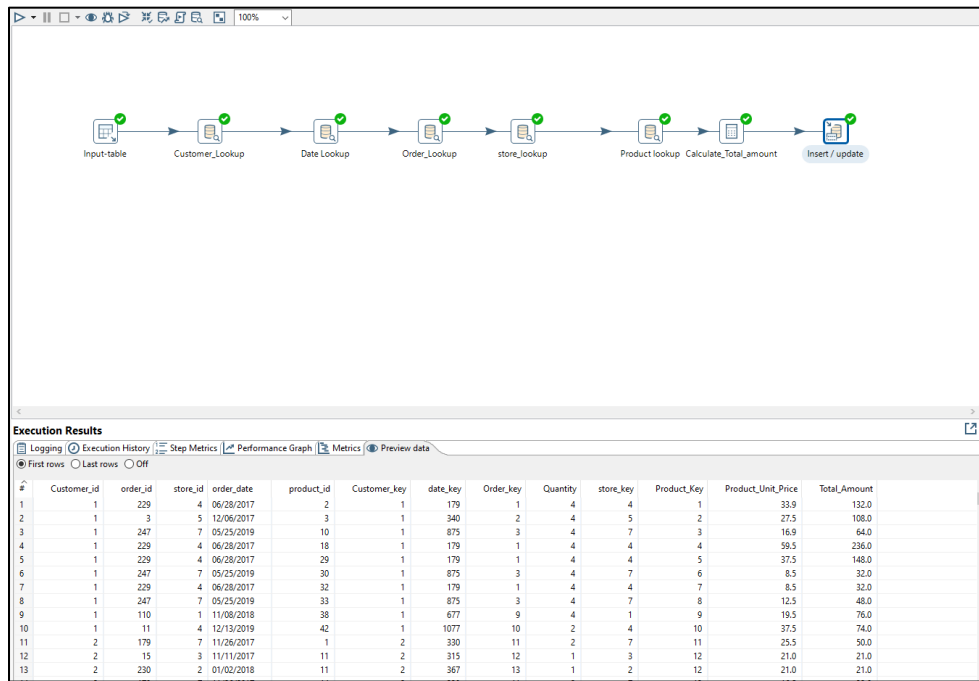
The screenshot shows a MySQL query editor with the following SQL statement:

```
SELECT * FROM spartans_gamestore_dw.dim_store;
```

Below the query editor is the 'Result Grid' section, which displays the results of the query in a table format. The table has three columns: 'store\_key', 'store\_id', and 'store\_address'. The results are as follows:

store_key	store_id	store_address
1	1	AL
2	2	AK
3	3	AZ
4	4	AR
5	5	CA
6	6	CO
7	7	CT
8	8	DE
9	9	FL
10	10	GA

## 6. Sales Fact data to Data Warehouse



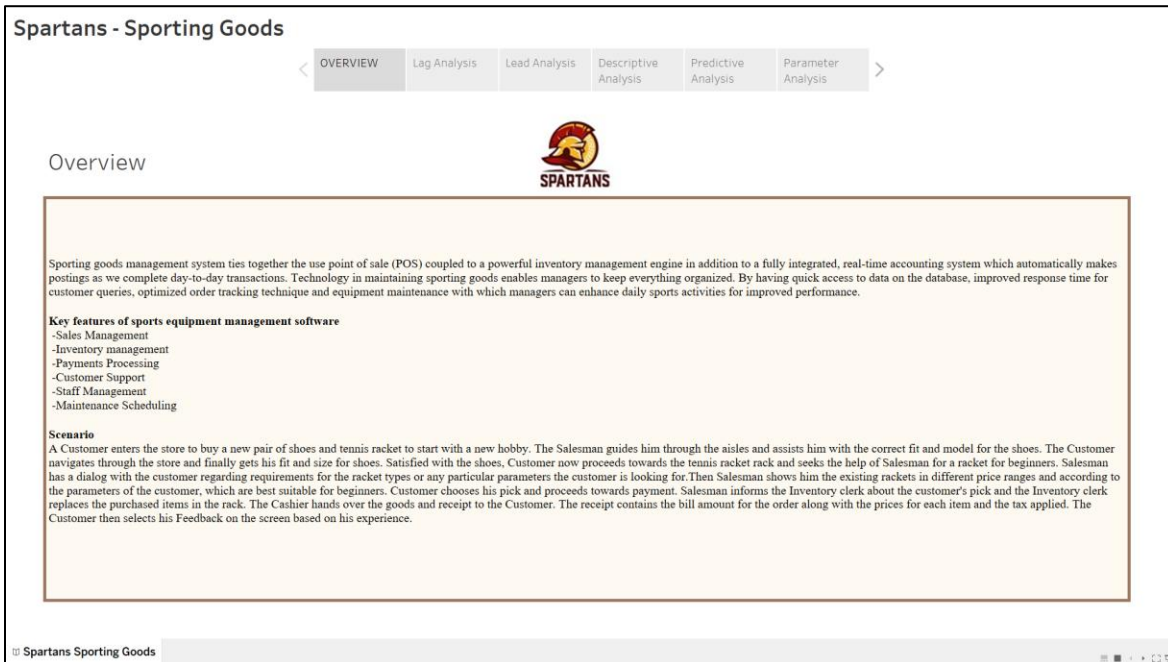
## Mysql Results

The screenshot shows a MySQL Workbench interface with a SQL query executed: `SELECT * FROM spartans_gamestore_dw.sales_fact_dw;`. The results are displayed in a table with 13 rows and 13 columns: Customer\_key, date\_key, Order\_key, Product\_key, store\_key, Quantity, Product\_Unit\_Price, and Total\_Amount.

	Customer_key	date_key	Order_key	Product_key	store_key	Quantity	Product_Unit_Price	Total_Amount
1	179	1	1	4	4	34		132
1	340	2	2	5	4	28		108
1	875	3	3	7	4	17		64
1	179	1	4	4	4	60		236
1	179	1	5	4	4	38		148
1	875	3	6	7	4	9		32
1	179	1	7	4	4	9		32
1	875	3	8	7	4	13		48
1	677	9	9	1	4	20		76
1	1077	10	10	4	2	38		74
2	330	11	11	7	2	26		50
2	315	12	12	3	1	21		21
2	367	13	12	2	1	21		21
2	330	11	13	7	2	17		32
2	166	15	14	4	3	60		177
2	577	16	8	7	2	13		24
2	822	17	15	5	4	7		24

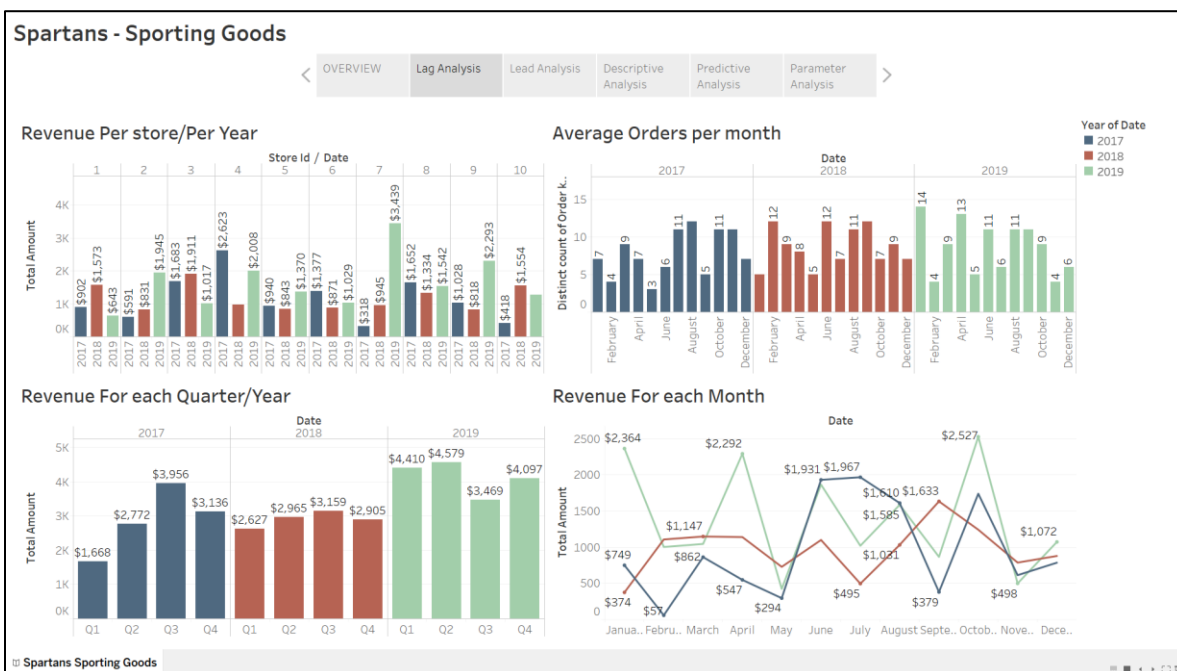
# Part- 4 Tableau Implementation

The Tableau file, has lead, lag, prescriptive and descriptive measures in a story. The main metrics which are highly insightful and impact the business decisions are taken and included in dashboards.



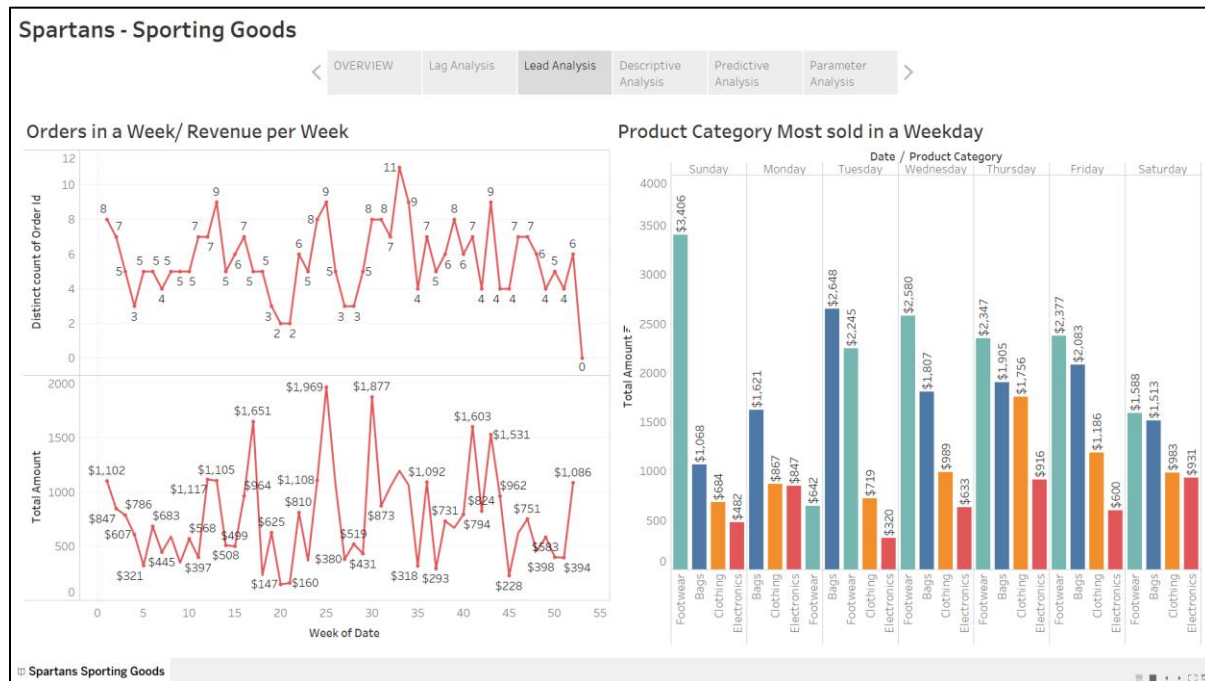
## Dashboard 1 - Lag Measures:

1. Revenue per store/per year
2. Revenue each Quarter
3. Average Orders per month
4. Revenue each month



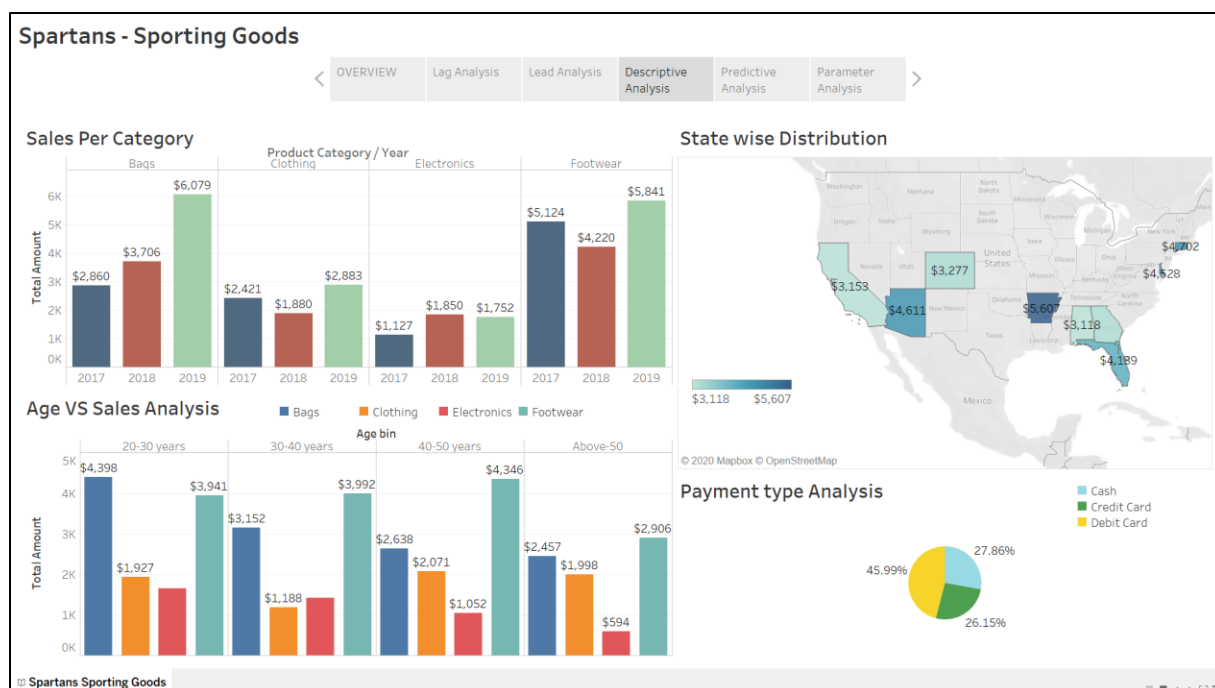
## Dashboard 2 - Lead Measures:

1. Orders in week
2. Revenue per week
3. Product most sold in each weekday



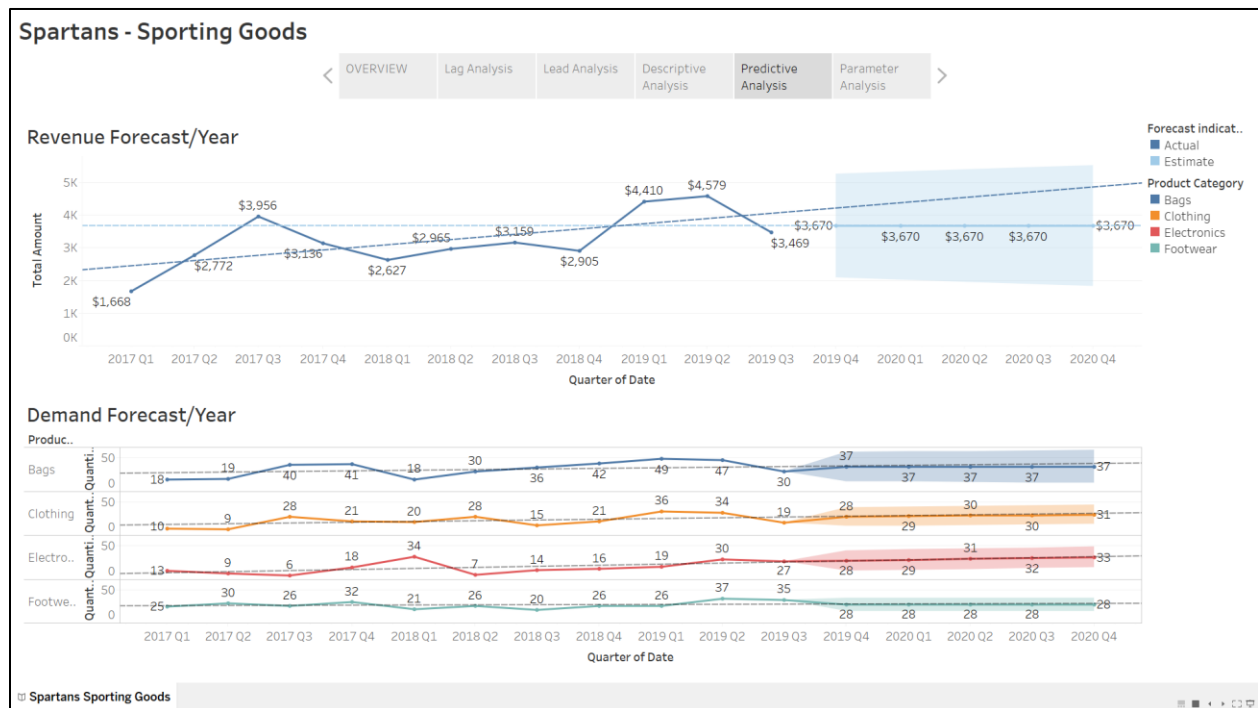
## Dashboard 3 - Descriptive Analysis:

1. Sales Vs Category
2. Age Vs category
3. Payment Type Analysis
4. State wise distribution



## Dashboard 4 – Predictive Analysis:

1. Revenue Forecast
2. Demand Forecast



## Conclusion

In summary, BI makes it possible to combine data from multiple sources, analyze the information into a digested format, and then disseminate the information to relevant stakeholders. This allows companies to see the big picture and make smart business decisions. By analyzing all these metrics and reports business can make critical decisions about the inventory, which stores are lucrative, which products need improvement, product unit price optimization.

### Key Learnings:

1. Learn to use the language of the business and ask questions.
2. Know the audience, data, and right tools.
3. BI makes it easier for people to see and understand their data without the technical knowledge.
4. Without comparison values, numbers on a dashboard are meaningless for the users.
5. Having Single truth value across organization.

Link to tableau:

[https://public.tableau.com/profile/ramya.mamidipaka#!/vizhome/Final\\_Submission/SpartansSportingGoods](https://public.tableau.com/profile/ramya.mamidipaka#!/vizhome/Final_Submission/SpartansSportingGoods)