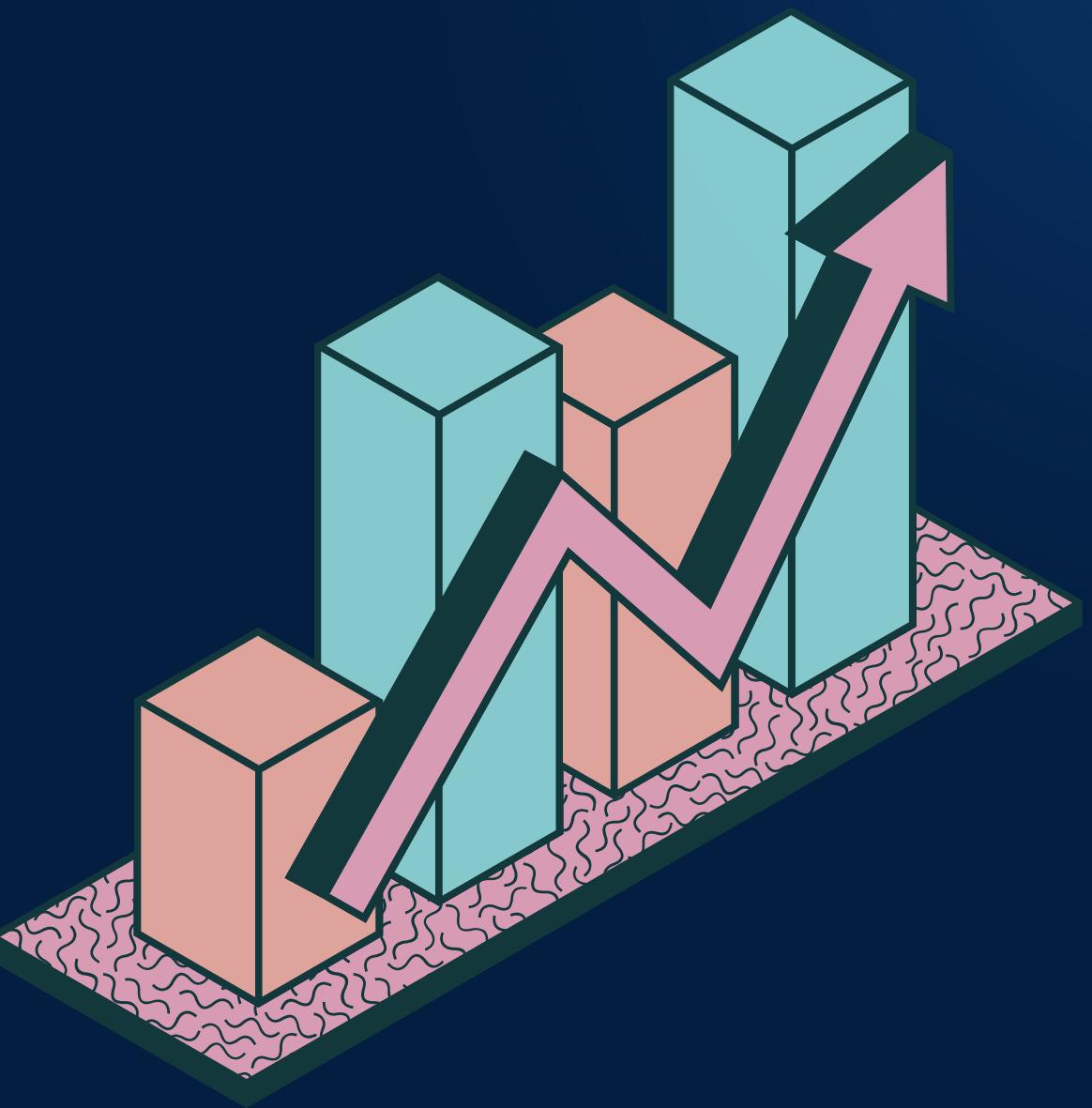


# SUPERSTORE SALES PROJECT



# OVERVIEW

01

Project Scope

02

Project Steps

03

Data Wrangling

04

Data Analysis

05

Data Visualization

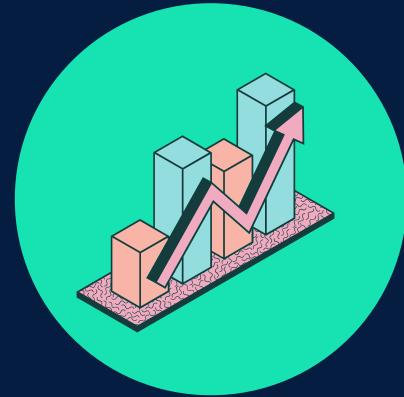
06

objectives

07

Our Team

# 01 PROJECT SCOPE



## SUPERSTORE SALES

The Superstore Sales project aims to analyze sales data to identify trends, customer behaviors, and product performance. It will involve data wrangling, analysis, and visualization to provide actionable insights for optimizing inventory and improving sales strategies. The ultimate goal is to enhance overall business performance and drive growth.

02

# PROJECT STEPS

## 1. DATA WRANGLING:

This is the process of cleaning and transforming raw data into a usable format. It involves tasks such as removing inconsistencies, handling missing values, and restructuring datasets to prepare them for analysis.

## 2. DATA ANALYSIS:

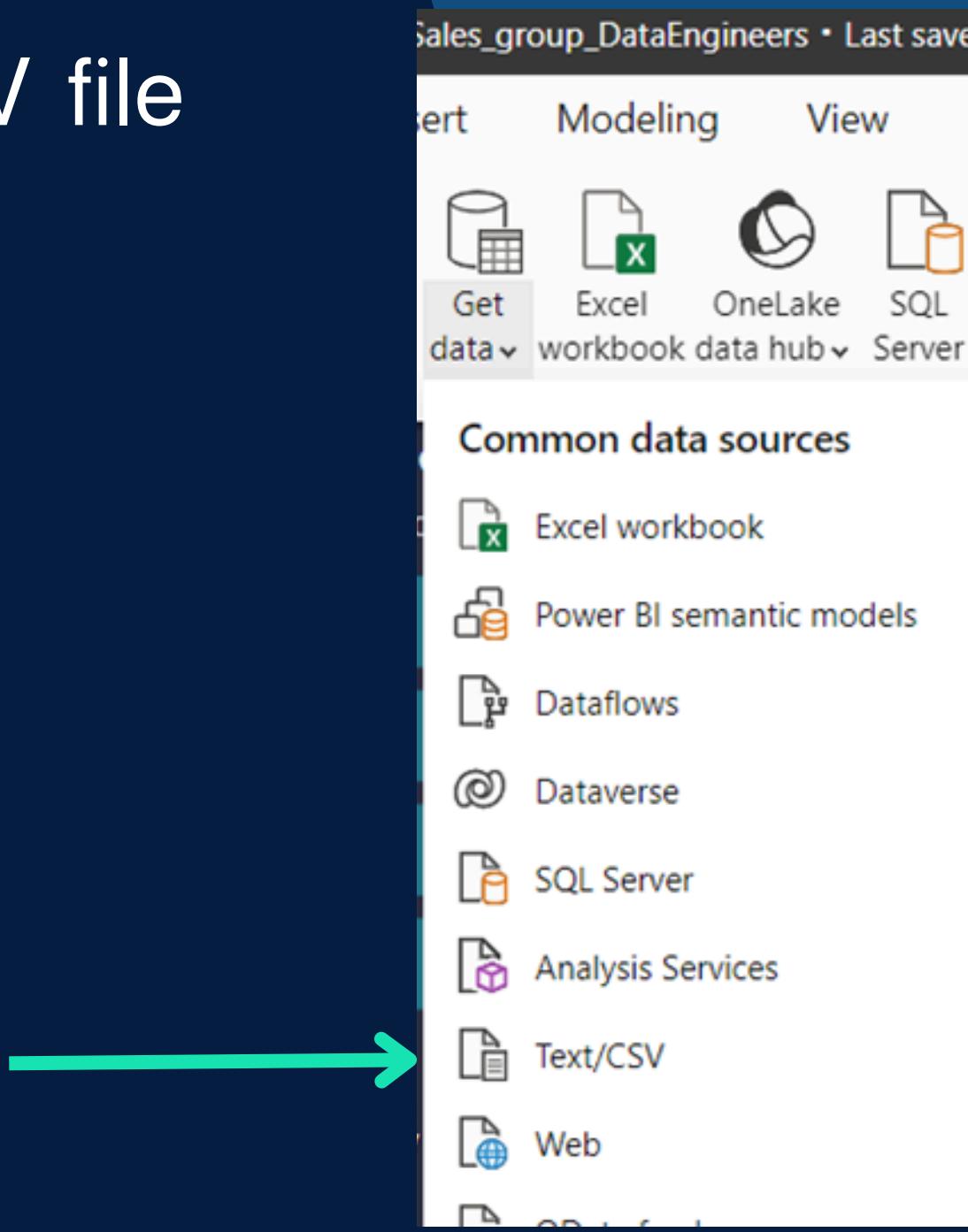
This step involves examining and interpreting the cleaned data to uncover patterns, trends, and insights. Techniques can include statistical analysis or exploratory data analysis (EDA) to answer specific research questions.

## 3. DATA VISUALIZATION:

This is the presentation of data in graphical or visual formats, such as charts, graphs, or dashboards. Effective visualization helps to communicate insights clearly and allows stakeholders to easily understand the findings and make informed decisions.

## 03 DATA WRANGLING:

### 01 Importing Data from CSV file



02

## Data cleaning & preparation

- All columns have the correct data type except order date and ship date.
- There is ambiguity in the order date.
- Power bi converts the order and ship dates to text.

Power BI Data Editor interface showing the Fact\_Orders table and its properties and applied steps.

**PROPERTIES**

- Name: Fact\_Orders
- All Properties

**APPLIED STEPS**

- Source
- Promoted Headers
- Changed Type
- Removed Columns
- Merged Queries
- Expanded Dim\_ShipMode
- Merged Queries1
- Merged Queries2
- Expanded Dim\_places.1

	A <sup>B</sup> <sub>C</sub> Order ID	A <sup>B</sup> <sub>C</sub> Order Date	A <sup>B</sup> <sub>C</sub> Ship Date	A <sup>B</sup> <sub>C</sub> Ship Mode	A <sup>B</sup> <sub>C</sub> Customer ID	A <sup>B</sup> <sub>C</sub> Customer Name	A <sup>B</sup> <sub>C</sub> Segment
1	CA-2017-152156	08/11/2017	11/11/2017	Second Class	CG-12520	Claire Gute	Consumer
2	CA-2017-152156	08/11/2017	11/11/2017	Second Class	CG-12520	Claire Gute	Consumer
3	CA-2017-138688	12/06/2017	16/06/2017	Second Class	DV-13045	Darrin Van Huff	Corporate
4	US-2016-108966	11/10/2016	18/10/2016	Standard Class	SO-20335	Sean O'Donnell	Consumer
5	US-2016-108966	11/10/2016	18/10/2016	Standard Class	SO-20335	Sean O'Donnell	Consumer
6	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
7	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
8	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
9	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
10	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
11	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
12	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer
13	CA-2018-114412	15/04/2018	20/04/2018	Standard Class	AA-10480	Andrew Allen	Consumer

03

## Correct the data type of the date

- Using M language convert date order&ship(text) to correct date format

The screenshot shows the Power BI Data Editor interface. On the left is a table grid with 15 rows of data. The columns are labeled: Customer ID, Product ID, Sales, Ship ID, Country ID, and OrderDate. The 'OrderDate' column contains dates like '11/8/2017' and '12/7/2016'. At the top of the screen, there is an M code editor bar with the following formula:

```
= Table.AddColumn(#"Removed Other Columns", "OrderDate", each Date.FromText([Order Date], "en-GB"))
```

To the right of the table, there are several panes: 'Query Settings' (with 'Name' set to 'Fact\_Orders'), 'PROPERTIES' (showing 'Name' is 'Fact\_Orders' and 'All Properties'), and 'APPLIED STEPS' (listing various steps such as 'Source', 'Promoted Headers', 'Changed Type', etc.).

	Customer ID	Product ID	Sales	Ship ID	Country ID	OrderDate
1	CG-12520	FUR-BO-10001798	261.96	1	101	11/8/2017
2	CG-12520	FUR-CH-10000454	731.94	1	101	11/8/2017
3	DK-12985	OFF-AP-10001271	152.94	2	101	12/7/2016
4	DK-12985	FUR-CH-10002647	283.92	2	101	12/7/2016
5	DV-13045	OFF-LA-10000240	14.62	1	102	6/12/2017
6	JS-15685	OFF-AR-10004930	20.1	2	102	9/17/2017
7	JS-15685	TEC-PH-10004093	73.584	2	102	9/17/2017
8	JS-15685	OFF-PA-10000304	6.48	2	102	9/17/2017
9	KB-16240	FUR-CH-10000595	190.72	1	102	11/7/2016
10	SO-20335	FUR-TA-10000577	957.5775	2	103	10/11/2016
11	SO-20335	OFF-ST-10000760	22.368	2	103	10/11/2016
12	BH-11710	FUR-FU-10001487	48.86	2	104	6/9/2015
13	BH-11710	OFF-AR-10002833	7.28	2	104	6/9/2015
14	BH-11710	TEC-PH-10002275	907.152	2	104	6/9/2015
15	BH-11710	OFF-BI-10003910	18.504	2	104	6/9/2015

04

# Data Transformation

- Reshape Date
- Merging queries
- Custom transformations

## -Reshape data

- .Create star schema model
- .Facts table : Measurements, metrics, or facts about an organization
- .Links to dimension tables for more details
- .Dimension table :Dimensions/attributes about a process
- .Holds reference data Dimension tables add more detail to fact table

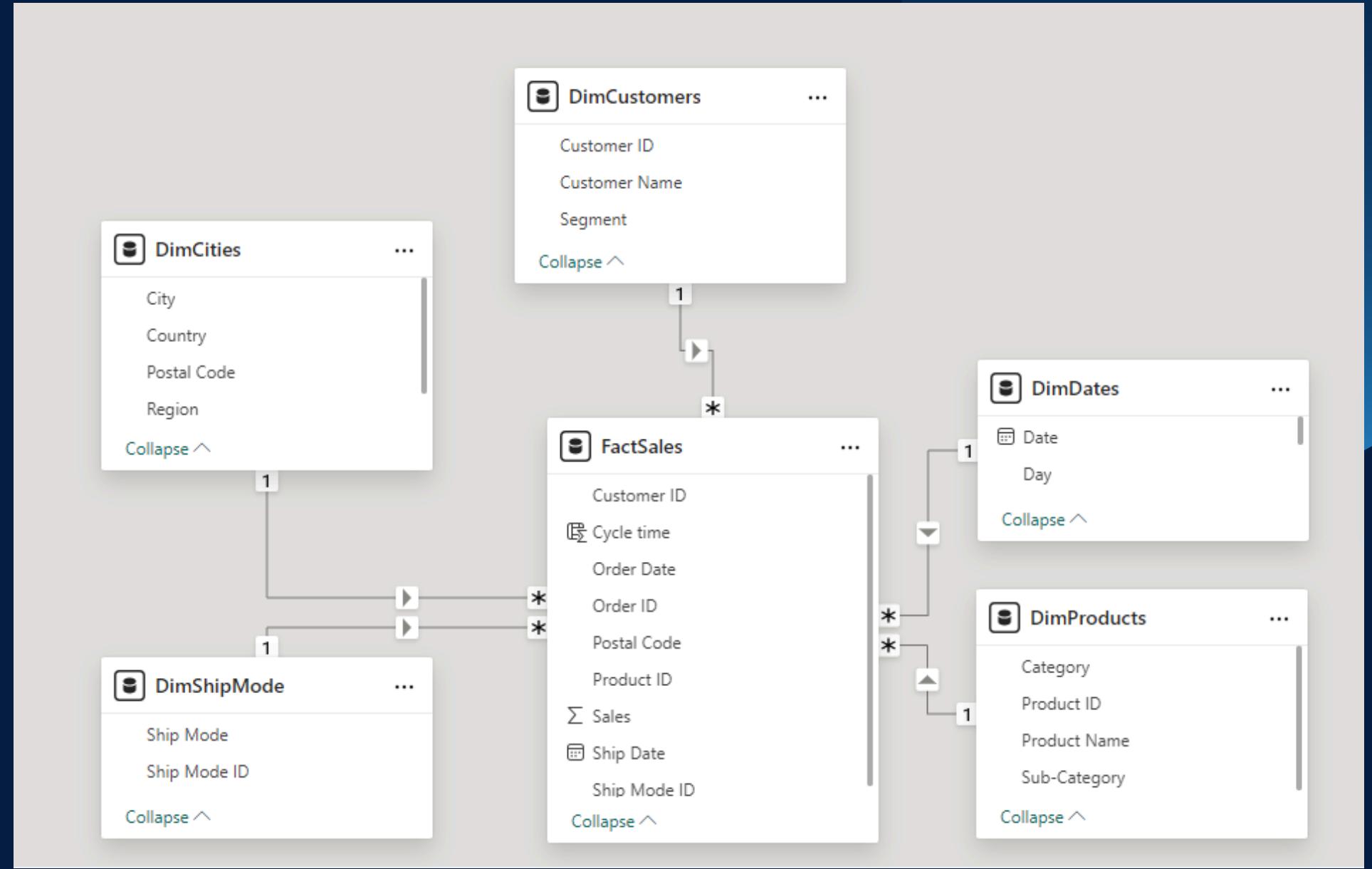
Date	Year	Month	Day	Month Name	Day of Week	Day Name	Quarter	DateKey
Friday, July 1, 2016	2016	7	1	July	5	Friday	3	20160701
Saturday, July 2, 2016	2016	7	2	July	6	Saturday	3	20160702
Sunday, July 3, 2016	2016	7	3	July	0	Sunday	3	20160703
Monday, July 4, 2016	2016	7	4	July	1	Monday	3	20160704
Tuesday, July 5, 2016	2016	7	5	July	2	Tuesday	3	20160705
Wednesday, July 6, 2016	2016	7	6	July	3	Wednesday	3	20160706
Thursday, July 7, 2016	2016	7	7	July	4	Thursday	3	20160707
Friday, July 8, 2016	2016	7	8	July	5	Friday	3	20160708
Saturday, July 9, 2016	2016	7	9	July	6	Saturday	3	20160709
Sunday, July 10, 2016	2016	7	10	July	0	Sunday	3	20160710

Fact table

Order ID	Product ID	Ship ID	Country ID	OrderDateKey	ShipDateKey	Customer ID	Sales	DiffShipOrderDate
CA-2016-134894	OFF-AP-10001271	2	101	20161207	20161211	DK-12985	152.94	4
CA-2016-134894	FUR-CH-10002647	2	101	20161207	20161211	DK-12985	283.92	4
CA-2015-123260	TEC-AC-10002323	2	104	20150826	20150830	FM-14290	176.8	4
CA-2017-130162	OFF-ST-10001328	2	104	20171028	20171101	JH-15910	93.06	4
CA-2017-130162	TEC-PH-10002563	2	104	20171028	20171101	JH-15910	302.376	4
US-2018-129441	FUR-FU-10000448	2	104	20180907	20180911	JC-15340	47.94	4
CA-2016-110457	FUR-TA-10001768	2	106	20160302	20160306	DK-13090	787.53	4
US-2016-118983	OFF-AP-10002311	2	107	20161122	20161126	HP-14815	68.81	4
US-2016-118983	OFF-BI-10000756	2	107	20161122	20161126	HP-14815	2.544	4
CA-2017-137330	OFF-AR-10000246	2	111	20171209	20171213	KB-16585	19.46	4

Fact table

## Star schema



## Create Dimension tables

- Dim\_Customers, Dim\_places, Dim\_ProductDetails, Dim\_ShipMode, Dim\_date
- Make primary key in each dim table and foreign keys in fact table refer to each primary key in dim tables

## 2. Merging queries

- Create merging queries between the dim table where you define the surrogate key and the fact table to create a foreign key column in it.

The screenshot shows the Power BI desktop interface. On the left, there's a 'Recent Sources' pane. The main area features the 'Merge' dialog box, which is prompting the user to select tables and matching columns for merging. Below the dialog, the 'Query' tab of the ribbon is selected, showing various data transformation tools like 'Advanced Editor', 'Manage Parameters', 'Choose Columns', 'Remove Columns', 'Keep Rows', 'Remove Rows', 'Split Column', 'Group By', 'Replace Values', 'Merge Queries', 'Append Queries', 'Combine Files', and 'Combine'. To the right of the merge dialog, the 'Query Settings' pane is open, showing properties for the current query named 'Fact\_Orders'. The 'APPLIED STEPS' section lists several steps taken during the query's creation, including 'Merged Queries1' (which is currently selected), 'Merged Queries2', 'Expanded Dim\_places.1', 'Expanded Dim\_places', 'Renamed Columns', 'Removed Columns1', 'Removed Other Columns', 'Added Custom', 'Added Custom1', 'Changed Type1', 'Removed Columns2', 'Added Custom4', 'Renamed Columns1', 'Added Custom2', 'Added Custom3', 'Removed Columns3', and 'Changed Type2'. The overall background of the Power BI interface is dark blue.

### 3. Custom transformations

- .Create difference between order date and ship date in fact table
- .Using function Duration.Days([ShipDate] - [OrderDate])

The screenshot shows the Microsoft Power BI Query Editor interface. The top navigation bar includes 'Steps', 'Query' (selected), 'Manage Columns', 'Reduce Rows', 'Sort', 'Transform' (selected), 'Combine', and 'AI Insights'. The right side features 'Query Settings' and 'APPLIED STEPS' sections.

The main area displays a table with the following columns:

Sales	Ship ID	Country ID	OrderDate	ShipDate	diff
261.96	1	101	11/8/2017	11/11/2017	3
731.94	1	101	11/8/2017	11/11/2017	3
152.94	2	101	12/7/2016	12/11/2016	4
283.92	2	101	12/7/2016	12/11/2016	4
14.62	1	102	6/12/2017	6/16/2017	4
20.1	2	102	9/17/2017	9/22/2017	5
73.584	2	102	9/17/2017	9/22/2017	5
6.48	2	102	9/17/2017	9/22/2017	5
190.72	1	102	11/7/2016	11/9/2016	2
957.5775	2	103	10/11/2016	10/18/2016	7
22.368	2	103	10/11/2016	10/18/2016	7
48.86	2	104	6/9/2015	6/14/2015	5
7.28	2	104	6/9/2015	6/14/2015	5
907.152	2	104	6/9/2015	6/14/2015	5

The query editor displays the M code for this transformation:

```
= Table.AddColumn(#"Removed Columns2", "diff", each Duration.Days([ShipDate] - [OrderDate]))
```

# 04 DATA ANALYSIS:

## 1. DAX MEASURES:

01

Total Sales = SUM(FactSales[Sales])

02

Total Orders = Count(FactSales[Order ID])

03

Total Products = COUNT(DimProducts[Product ID])

04

Total Customers = count(DimCustomers[Customer ID])

05

Avg. orders Sales = SUM(FactSales[Sales]) / COUNT(FactSales[Order ID])

06

Avg. Cycle time = SUM(FactSales[Cycle time])/COUNT(FactSales[Order ID])

### DAXMeasures

-  Avg. Cycle time
-  Avg. orders Sales
-  Total Customers
-  Total Orders
-  Total Products
-  Total Sales

## 2. POSSIBLE INSIGHTS:

### 1. CARDS

01

Total Sales

02

Total Orders

03

Total Products

04

Total Customers

05

Avg. Sales by Orders

06

Avg. Cycle time

## 2. POSSIBLE INSIGHTS:

### 2. SALES:

01

Sales by date. (Line chart)

02

Sales by state. (Map)

03

Sales by segments. (Donut chart or Pi chart)

04

Sales by ship mode (Donut chart or Pi chart)

05

Sales by categories , subcategories and products. ( Tree map or Decomposition tree)

## 2. POSSIBLE INSIGHTS:

### 3. ORDERS:

01

Orders by date. (Line chart)

02

Orders by state. (Map)

03

Orders by segments. (Donut chart, Pi chart)

04

Orders by ship mode (Donut chart or Pi chart or  
Decomposition tree)

05

Orders by categories , subcategories and products. ( Tree map  
or Decomposition tree)

## 2. POSSIBLE INSIGHTS:

### 4. OTHERS:

01

Average cycle time by state. (Bar chart)

02

Orders by segments and customers.  
(Decomposition tree)

03

Total Sales, Total Orders, Avg. Sales by Orders and  
Avg. Cycle Time. (Tooltip)

04

Filters by city, state, customer, segment, product,  
year and etc.....

# 05

# DATA VISUALIZATION

## 01

## cards

- Card visuals in Power BI are used to display single values or key metrics prominently, providing a quick and clear summary of important data points.

4922  
*Total Orders*

230.77  
*Avg. Order Value*

2.26M  
*Total Sales*

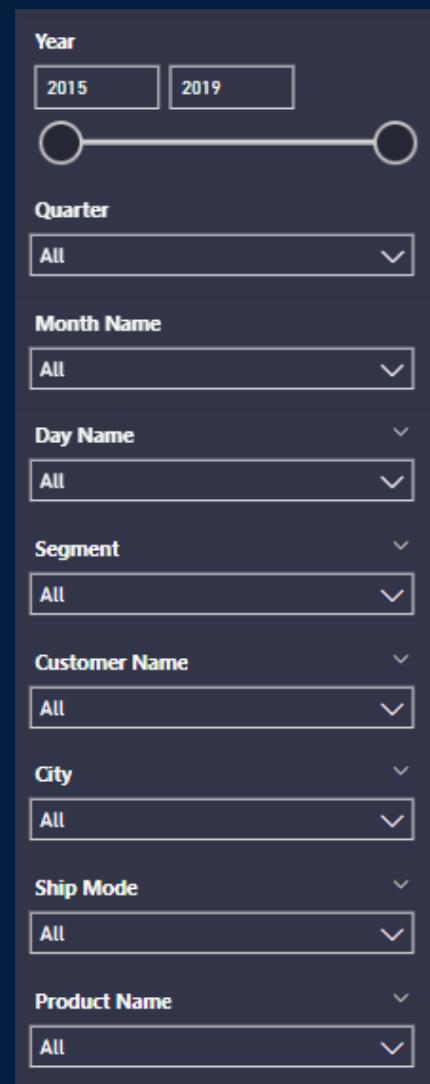
793  
*Total Customers*

4  
*Avg. Cycle Time*

## 02

## Filters

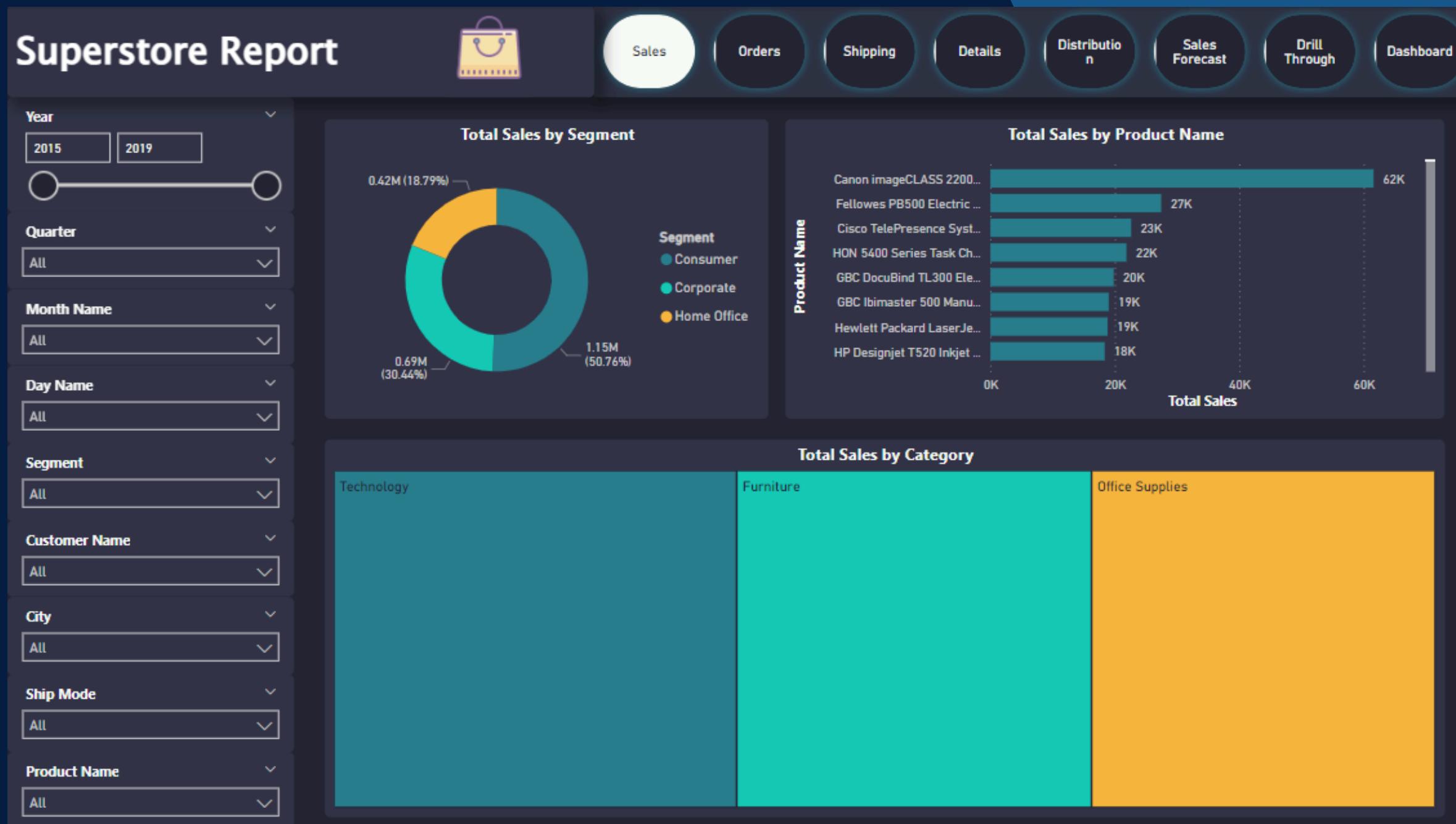
- Filters in Power BI allow users to refine and limit the data displayed in reports and visuals based on specific criteria.



03

## Sales Visuals

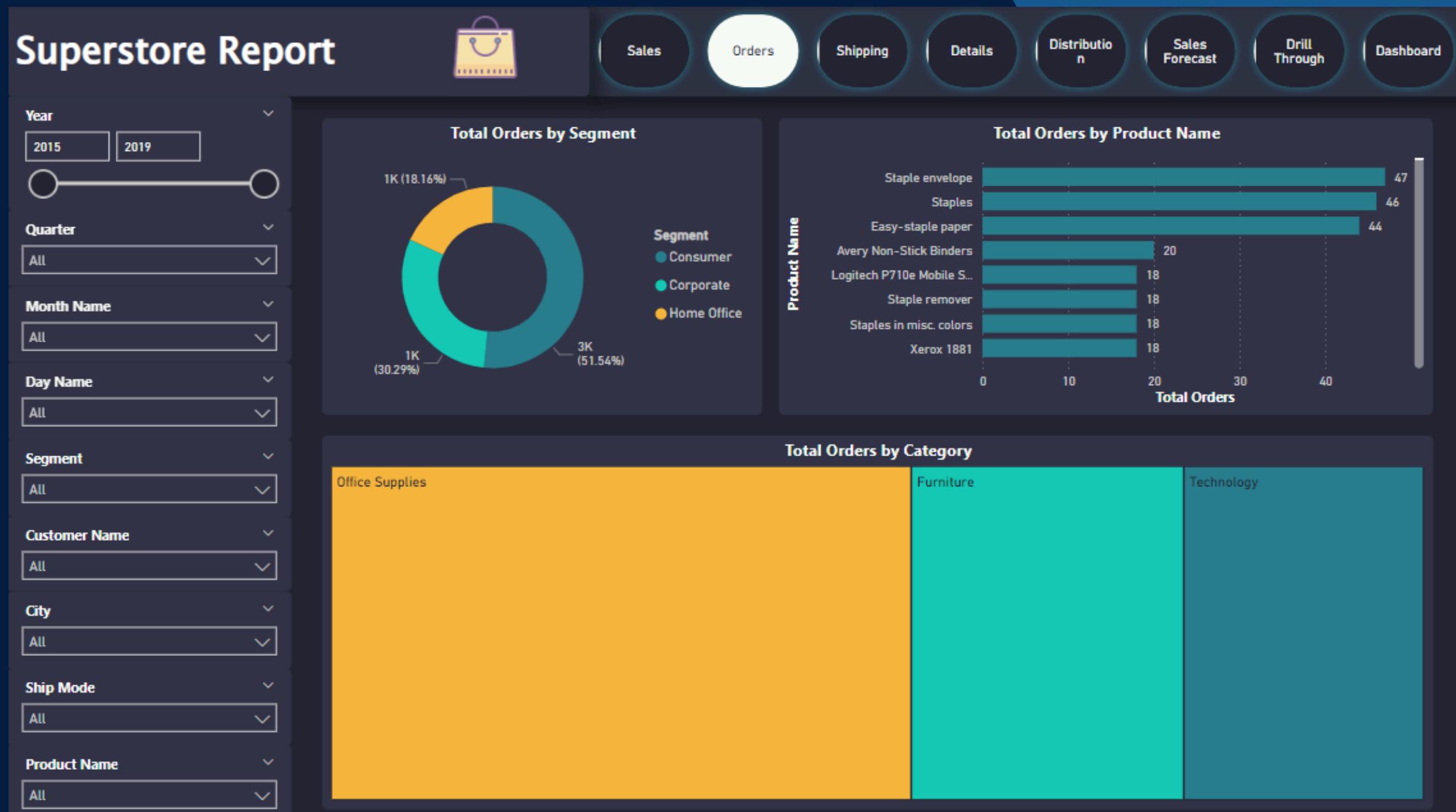
- Consumer segment has the most sales by 51%
- Canon image class 2200 Advanced Copier has the most sales 62,000\$.
- Technology has the most sales.



## 04

## Orders Visuals

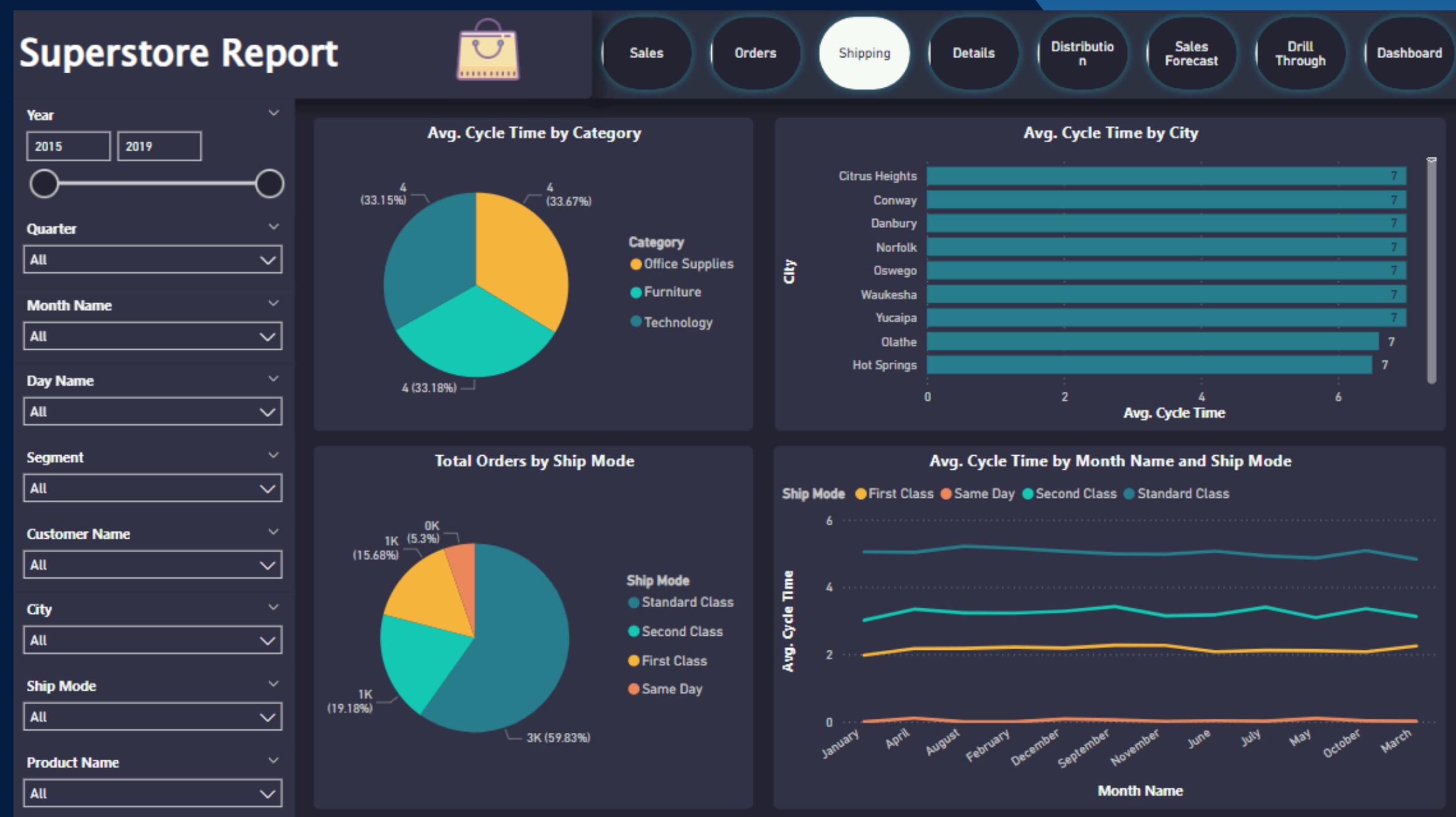
- California has the most orders
- office Supplies has Max orders & Technology has Min orders
- Consumer has the most sales by 51%.



## 05

## Details Visuals

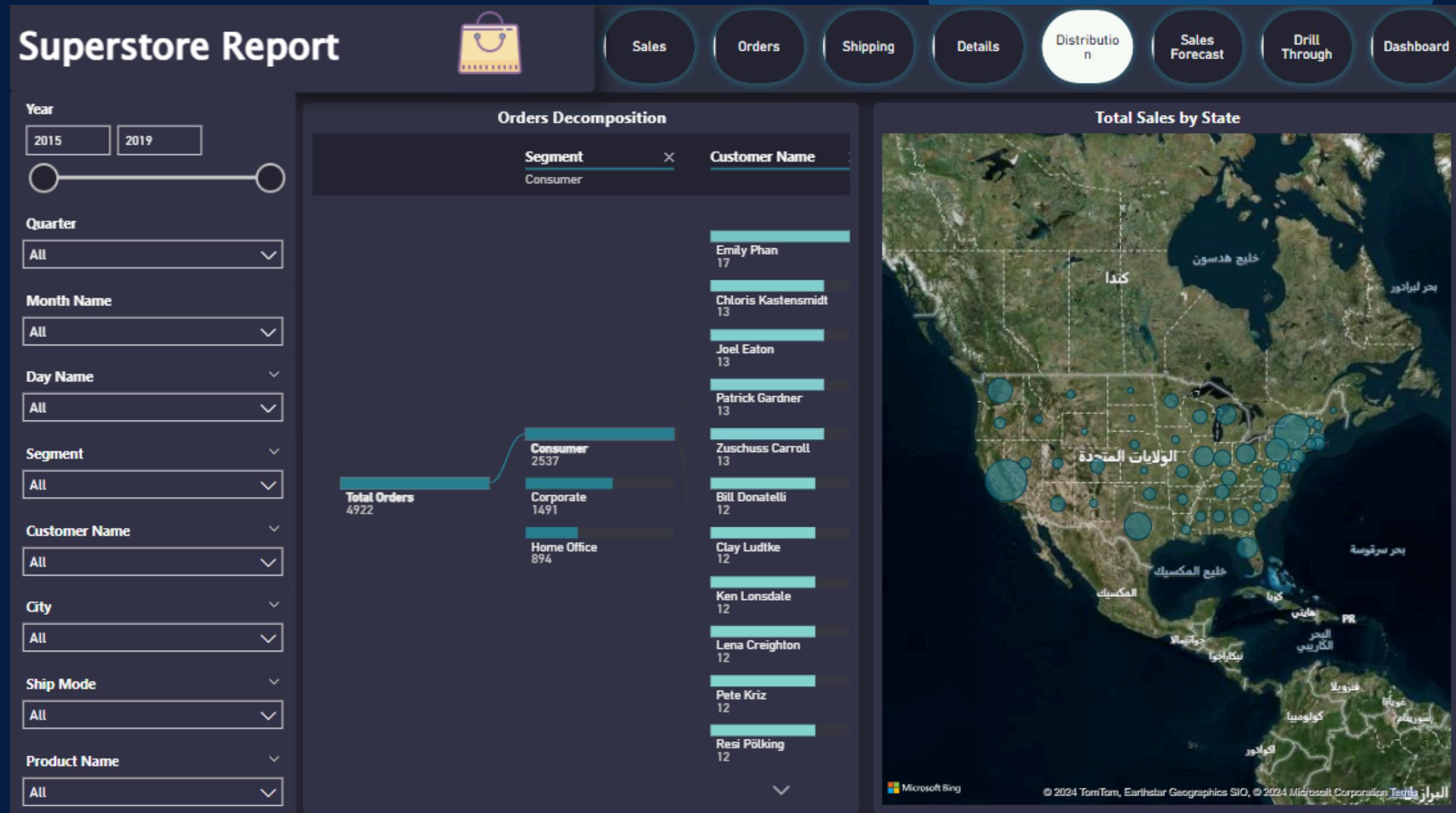
- Max C.T is 7 days at District of Columbia.
- Standard Class contribute of 59.8 % of Sales.
- Sales Increases at the end of the year.



## 06

## Map Visuals

- Heat map shows sales distribution.
- Customer Emily Phan has the most orders.



07

## Tooltip

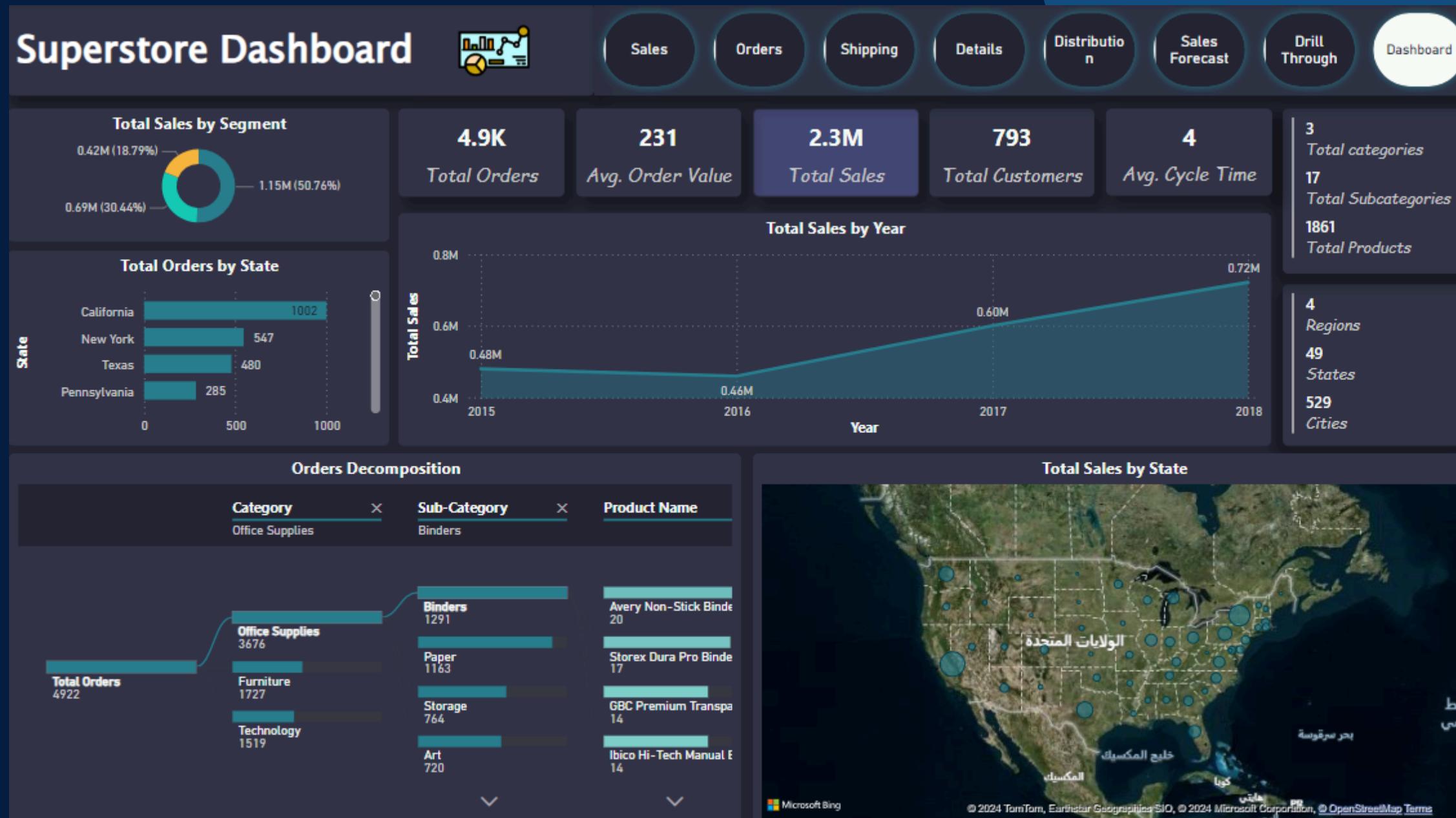
- Shows These data by State (Sales, Orders, Avg. sales by order and C.T)

Sales by City				
City	Total Sales	Total Orders	Avg. Order Value	Avg. Cycle Time
New York City	252,462.55	439	283.35	4
Los Angeles	173,420.18	378	238.21	4
Seattle	116,106.32	210	272.55	4
<b>Total</b>	<b>2,261,536.78</b>	<b>4922</b>	<b>230.77</b>	<b>4</b>

## 08

## Dashboard

- Shows Most orders by Category are office supplies.
- Also shows the most important insights of these we have explained.



## 06 OBJECTIVES

### Growth

refers to the increase in size, value, or importance over time, often measured in terms of revenue, market share, or user base. It can result from factors such as improved strategies, market expansion, or innovation. Sustainable growth balances short-term gains with long-term stability and profitability.

### Customer Satisfaction

reflects how well the shopping experience and products meet customer expectations, driving loyalty and repeat business.



07

## OUR TEAM



**MOHAMED  
ESSAM**

**Role**

**Data Wrangling**



**MAHMOUD  
ERAQI**

**Role**

**Data Analysis**



**MAHMOUD  
ABD  
ELKHALEQ**

**Role**

**Data Visualization**

# THANK YOU

