

# Superstore

## Relational Database Design



# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

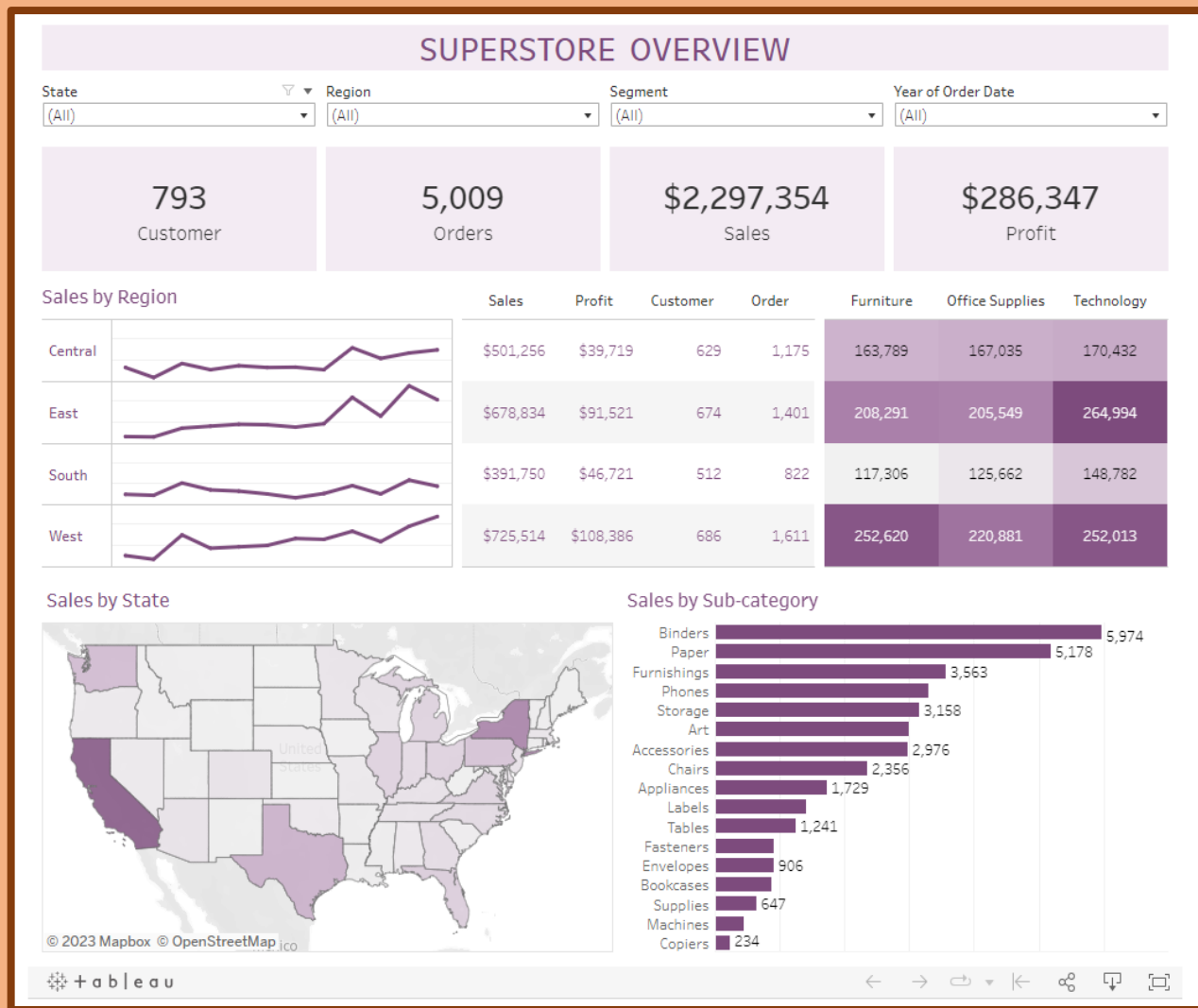
4. Creating Relational DB

5. 3NF Normalization

6. Further Implementation

**Business  
Background  
Story**

# Business Background



◆ Superstore is an e-commerce website, covering a wide range of categories of products across the United States' geography.

◆ Superstore shares their data online and the dataset is famous for data visualization and data analysis practice.

◆ As their business volume gradually expands, the existing data storage method can no longer meet their needs.

# Business Objectives

## Demands:

- ◆ A Superstore Giant is seeking help to redesign its data storage structure.
- ◆ They need more efficient storage and retrieval system as volumes surge.
- ◆ They need easy manipulation methods for daily use and further analysis.

## Objectives: Building a relational database for Superstore

Meets 3NF Standards

Retains Original Information

Simplifies Data Structure

Implements Flexible Queries

# Business Background

➤ Each row indicates one item within an order.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	Postal C	City	State	Region	Country	Custom	Custom	Segmen	Product	Product	Manufac	Sub-Cat	Category	Order ID	Order Date	Ship Date	Ship Mo	Order De	Discoun	Quantity	Profit	Sales
1	73120	Oklahoma	Oklahoma	Central	United States	C00709	Aaron Berg	Consumer	L01390	Samsung C	Samsung	Phones	Technology	CA-2013-1	2013/11/11	2013/11/13	First Class	CA-2013-1	0%	2	\$62	\$222
2	73120	Oklahoma	Oklahoma	Central	United States	C00709	Aaron Berg	Consumer	L01430	Sauder Fac	Sauder	Bookcases	Furniture	CA-2013-1	2013/11/11	2013/11/13	First Class	CA-2013-1	0%	2	\$55	\$342
3	76017	Arlington	Texas	Central	United States	C00709	Aaron Berg	Consumer	L00116	Akro Stacki	Akro	Storage	Office Supp	CA-2011-1	2011/2/19	2011/2/25	Standard C	CA-2011-1	20%	2	(\$3)	\$13
4	98103	Seattle	Washington	West	United States	C00709	Aaron Berg	Consumer	L00412	Carina 42"	Carina	Storage	Office Supp	CA-2011-1	2011/3/7	2011/3/8	First Class	CA-2011-1	0%	3	\$5	\$243
5	98103	Seattle	Washington	West	United States	C00709	Aaron Berg	Consumer	L00763	Global Pust	Global	Chairs	Furniture	CA-2011-1	2011/3/7	2011/3/8	First Class	CA-2011-1	20%	1	\$5	\$49
6	98103	Seattle	Washington	West	United States	C00709	Aaron Berg	Consumer	L01143	Newell 330	Newell	Art	Office Supp	CA-2011-1	2011/3/7	2011/3/8	First Class	CA-2011-1	0%	3	\$5	\$18
7	10035	New York	New York	East	United States	C00462	Aaron Hawl	Corporate	L00219	Avery 51	Avery	Labels	Office Supp	CA-2011-1	2011/12/31	2012/1/1	First Class	CA-2011-1	0%	3	\$9	\$19
8	10035	New York	New York	East	United States	C00462	Aaron Hawl	Corporate	L01595	Verbatim 2"	Verbatim	Accessories	Technology	CA-2011-1	2011/12/31	2012/1/1	First Class	CA-2011-1	0%	3	\$11	\$35
9	12180	Troy	New York	East	United States	C00462	Aaron Hawl	Corporate	L00068	ACCOHIDE	ACCOHIDE	Binders	Office Supp	CA-2011-1	2011/4/22	2011/4/24	Second Cla	CA-2011-1	20%	3	\$3	\$10
10	12180	Troy	New York	East	United States	C00462	Aaron Hawl	Corporate	L01499	Staples	Other	Envelopes	Office Supp	CA-2011-1	2011/4/22	2011/4/24	Second Cla	CA-2011-1	0%	8	\$121	\$248
11	19134	Philadelphia	Pennsylvania	East	United States	C00462	Aaron Hawl	Corporate	L00244	Avery Fluor	Avery	Art	Office Supp	CA-2014-1	2014/12/19	2014/12/23	Standard C	CA-2014-1	20%	7	\$2	\$19
12	39503	Gulport	Mississippi	South	United States	C00462	Aaron Hawl	Corporate	L00577	Eldon Imag	Eldon	Furnishings	Furniture	CA-2013-1	2013/3/21	2013/3/26	Second Cla	CA-2013-1	0%	7	\$38	\$86
13	90004	Los Angeles	California	West	United States	C00462	Aaron Hawl	Corporate	L00542	EcoTones	Other	Paper	Office Supp	CA-2011-1	2011/5/13	2011/5/19	Standard C	CA-2011-1	0%	2	\$4	\$8
14	90004	Los Angeles	California	West	United States	C00462	Aaron Hawl	Corporate	L00906	Iceberg Nex	Iceberg	Chairs	Furniture	CA-2011-1	2011/5/13	2011/5/19	Standard C	CA-2011-1	20%	6	\$21	\$279
15	94109	San Francisco	California	West	United States	C00462	Aaron Hawl	Corporate	L00714	GBC VeloB	GBC	Binders	Office Supp	US-2011-1	2011/10/25	2011/10/28	First Class	US-2011-1	20%	4	\$19	\$49
16	94122	San Francisco	California	West	United States	C00462	Aaron Hawl	Corporate	L00660	Fellowes St	Fellowes	Storage	Office Supp	CA-2012-1	2012/12/27	2012/12/31	Standard C	CA-2012-1	0%	2	\$61	\$323
17	94122	San Francisco	California	West	United States	C00462	Aaron Hawl	Corporate	L00731	Geemarc A	Other	Phones	Technology	CA-2012-1	2012/12/27	2012/12/31	Standard C	CA-2012-1	20%	9	\$75	\$668
18	10035	New York	New York	East	United States	C00523	Aaron Sma	Corporate	L01743	Xerox 1956	Xerox	Paper	Office Supp	US-2011-1	2011/7/27	2011/8/2	Standard C	US-2011-1	0%	11	\$32	\$66
19	22204	Arlington	Virginia	South	United States	C00523	Aaron Sma	Corporate	L00788	Hanging Pe	Other	Storage	Office Supp	CA-2013-1	2013/3/29	2013/4/1	First Class	CA-2013-1	0%	2	\$8	\$31
20	28540	Jacksonville	North Carolina	South	United States	C00523	Aaron Sma	Corporate	L00177	Avery 3 1/2"	Avery	Binders	Office Supp	CA-2014-1	2014/1/3	2014/1/5	First Class	CA-2014-1	70%	5	(\$13)	\$16
21	28540	Jacksonville	North Carolina	South	United States	C00523	Aaron Sma	Corporate	L00265	Avery Recy	Avery	Binders	Office Supp	CA-2014-1	2014/1/3	2014/1/5	First Class	CA-2014-1	70%	6	(\$21)	\$29
22	28540	Jacksonville	North Carolina	South	United States	C00523	Aaron Sma	Corporate	L00428	Cisco CP-7	Cisco	Machines	Technology	CA-2014-1	2014/1/3	2014/1/5	First Class	CA-2014-1	50%	2	(\$28)	\$696
23	78745	Austin	Texas	Central	United States	C00523	Aaron Sma	Corporate	L00274	Bady BDG10	Other	Machines	Technology	CA-2014-1	2014/8/2	2014/8/8	Standard C	CA-2014-1	40%	3	(\$264)	\$1,440
24	78745	Austin	Texas	Central	United States	C00523	Aaron Sma	Corporate	L01788	Xerox 1998	Xerox	Paper	Office Supp	CA-2014-1	2014/8/2	2014/8/8	Standard C	CA-2014-1	20%	7	\$13	\$36
25	91104	Pasadena	California	West	United States	C00523	Aaron Sma	Corporate	L00978	Lesro Sheff	Lesro	Tables	Furniture	CA-2014-1	2014/10/4	2014/10/9	Standard C	CA-2014-1	20%	3	(\$6)	\$171
26	94110	San Francisco	California	West	United States	C00523	Aaron Sma	Corporate	L00168	Atlantic Met	Atlantic	Bookcases	Furniture	CA-2013-1	2013/9/26	2013/9/28	First Class	CA-2013-1	15%	2	\$84	\$478
27	97756	Redmond	Oregon	West	United States	C00523	Aaron Sma	Corporate	L00269	Avery Trap	Avery	Binders	Office Supp	US-2014-1	2014/9/5	2014/9/6	First Class	US-2014-1	70%	7	(\$59)	\$88
28	3301	Concord	New Hampshire	East	United States	C00160	Adam Bella	Home Office	L00076	Acme Forgr	Acme	Supplies	Office Supp	CA-2013-1	2013/8/30	2013/9/4	Standard C	CA-2013-1	0%	3	\$8	\$28
29	10009	New York	New York	East	United States	C00160	Adam Bella	Home Office	L00049	Acco Hangi	Acco	Binders	Office Supp	CA-2013-1	2013/9/2	2013/9/4	First Class	CA-2013-1	20%	1	\$1	\$3
30	10009	New York	New York	East	United States	C00160	Adam Bella	Home Office	L00682	GBC Docu	GBC	Binders	Office Supp	CA-2013-1	2013/9/2	2013/9/4	First Class	CA-2013-1	20%	4	\$1,415	\$4,355
31	10009	New York	New York	East	United States	C00160	Adam Bella	Home Office	L01369	Sabrent 4-F	Other	Accessories	Technology	CA-2013-1	2013/9/2	2013/9/4	First Class	CA-2013-1	0%	1	\$2	\$7
32	10009	New York	New York	East	United States	C00160	Adam Bella	Home Office	L01661	Xerox 1881	Xerox	Paper	Office Supp	CA-2013-1	2013/9/2	2013/9/4	First Class	CA-2013-1	0%	2	\$12	\$25
33	10009	New York	New York	East	United States	C00160	Adam Bella	Home Office	L01661	Xerox 1881	Xerox	Paper	Office Supp	CA-2013-1	2013/9/2	2013/9/4	First Class	CA-2013-1	0%	4	\$23	\$49
34	19143	Philadelphia	Pennsylvania	East	United States	C00160	Adam Bella	Home Office	L01827	Xerox 231	Xerox	Paper	Office Supp	CA-2014-1	2014/9/17	2014/9/22	Standard C	CA-2014-1	20%	4	\$7	\$21
35	22980	Waynesboro	Virginia	South	United States	C00160	Adam Bella	Home Office	L00610	Eureka The	Eureka	Appliances	Office Supp	US-2013-1	2013/3/14	2013/3/19	Standard C	US-2013-1	0%	2	\$45	\$160
36	22980	Waynesboro	Virginia	South	United States	C00160	Adam Bella	Home Office	L00874	Howard Mill	Howard Mill	Furnishings	Furniture	US-2013-1	2013/3/14	2013/3/19	Standard C	US-2013-1	0%	2	\$41	\$128
37	22980	Waynesboro	Virginia	South	United States	C00160	Adam Bella	Home Office	L00892	Ibico Cover	Ibico	Binders	Office Supp	US-2013-1	2013/3/14	2013/3/19	Standard C	US-2013-1	0%	4	\$21	\$46
38	46142	Greenwood	Indiana	Central	United States	C00160	Adam Bella	Home Office	L00129	Ampad Gol	Ampad	Paper	Office Supp	CA-2014-1	2014/11/6	2014/11/8	First Class	CA-2014-1	0%	1	\$2	\$4
39	46142	Greenwood	Indiana	Central	United States	C00160	Adam Bella	Home Office	L01499	Staples	Other	Paper	Office Supp	CA-2014-1	2014/11/6	2014/11/8	First Class	CA-2014-1	0%	3	\$82	\$168
40	46142	Greenwood	Indiana	Central	United States	C00160	Adam Bella	Home Office	L01515	Strathmore	Strathmore	Paper	Office Supp	CA-2014-1	2014/11/6	2014/11/8	First Class	CA-2014-1	0%	10	\$31	\$68
41	90004	Los Angeles	California	West	United States	C00160	Adam Bella	Home Office	L01009	Logitech G	Logitech	Accessories	Technology	CA-2014-1	2014/5/8	2014/5/13	Standard C	CA-2014-1	0%	1	\$29	\$80
42	98105	Seattle	Washington	West	United States	C00160	Adam Bella	Home Office	L00816	Hon 2090	Hon	Chairs	Furniture	CA-2014-1	2014/11/7	2014/11/14	Standard C	CA-2014-1	20%	2	(\$73)	\$450
43	98105	Seattle	Washington	West	United States	C00160	Adam Bella	Home Office	L00825	Hon 5100	Hon	Tables	Furniture	CA-2014-1	2014/11/7	2014/11/14	Standard C	CA-2014-1	0%	7	\$367	\$2,037
44	98105	Seattle	Washington	West	United States	C00160	Adam Bella	Home Office	L01404	SanDisk Cr	SanDisk	Accessories	Technology	CA-2014-1	2014/11/7	2014/11/14	Standard C	CA-2014-1	0%	3	\$33	\$109
45	98198	Des Moines	Washington	West	United States	C00160	Adam Bella	Home Office	L01651	Wirebound	Wirebound	Paper	Office Supp	CA-2012-1	2012/9/18	2012/9/24	Standard C	CA-2012-1	0%	2	\$9	\$19

# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

4. Creating Relational DB

5. 3NF Normalization

6. Further Implementation

**Entity  
Relationship  
Model**

# Creating ER Model -- Entities

- Each row should include 5 clusters of information.
- Each cluster could be a table within the database.

**What**

**Product**

**When**

**Order**

**Who**

**Customer**

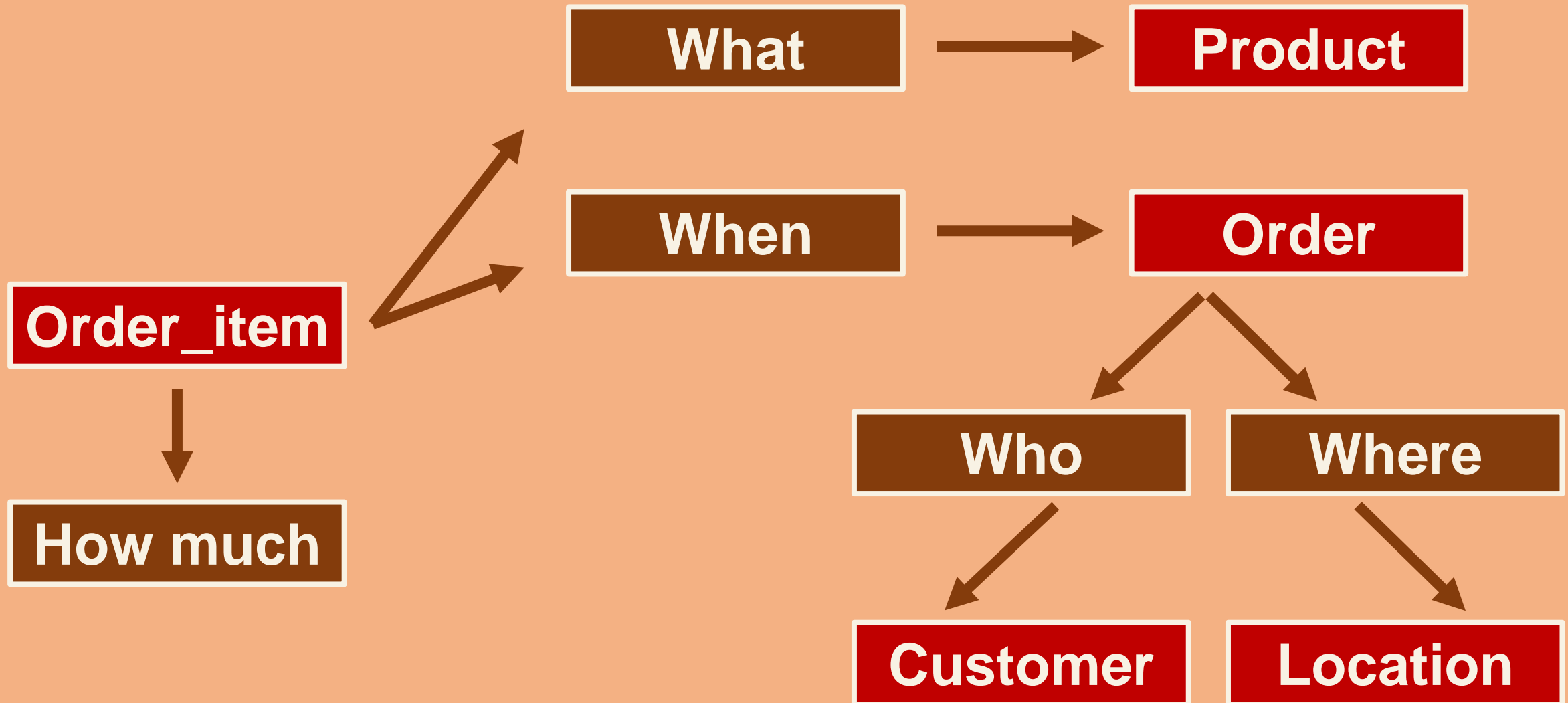
**Where**

**Location**

**How much**

**Order\_item**

# Creating ER Model – Entities Connection





# Creating ER Model – Relationship1

ProductID	Product Name	Manufacturer	Sub Category	Category
L01086	Message Book	Message Book	Paper	Office Supplies
L00217	Avery 508	Avery	Labels	Office Supplies
L00707	GBC Standard	GBC	Binders	Office Supplies
L01371	SAFCO Boltless	SAFCO	Storage	Office Supplies
L00251	Avery Hi-Liter	Avery	Art	Office Supplies

OrderItemID	ProductID	OrderID	Discount	Quantity	Profit	Sales
CA-2011-103	L01086	CA-2011-103	20%	2	\$6	\$16
CA-2011-112	L00217	CA-2011-112	20%	3	\$4	\$12
CA-2011-112	L00707	CA-2011-112	80%	2	(\$5)	\$4
CA-2011-112	L01371	CA-2011-112	20%	3	(\$65)	\$273
CA-2011-141	L00251	CA-2011-141	20%	3	\$5	\$20

## Product to Order\_item Relationship:

A Product may belong to one or more Orders\_items.  
Each Order\_item must have one and only one Product.

**Product.ProductID (PK) --> Order\_item.ProductID (FK)**

# Creating ER Model – Relationship2

OrderID	CustomerID	PostalCode	OrderDate	ShipDate	ShipMode
CA-2011-103	C00184	77095	2011/1/4	2011/1/8	Standard Cl
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Cl
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Cl
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Cl
CA-2011-141	C00484	19143	2011/1/6	2011/1/13	Standard Cl

OrderItemID	ProductID	OrderID	Discount	Quantity	Profit	Sales
CA-2011-103	L01086	CA-2011-103	20%	2	\$6	\$16
CA-2011-112	L00217	CA-2011-112	20%	3	\$4	\$12
CA-2011-112	L00707	CA-2011-112	80%	2	(\$5)	\$4
CA-2011-112	L01371	CA-2011-112	20%	3	(\$65)	\$273
CA-2011-141	L00251	CA-2011-141	20%	3	\$5	\$20

## Order to Order\_item Relationship:

An Order may have one or more Orders\_items.  
Each Order\_item must belong to one and only one Order.

**Order.OrderID (PK) --> Order\_item.OrderID (FK)**

# Creating ER Model – Relationship3

CustomerID	Customer Name	Segment
C00184	Darren Powe	Consumer
C00477	Phillina Ober	Home Office
C00477	Phillina Ober	Home Office
C00477	Phillina Ober	Home Office
C00484	Mick Brown	Consumer

OrderID	CustomerID	PostalCode	OrderDate	ShipDate	ShipMode
CA-2011-103	C00184	77095	2011/1/4	2011/1/8	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-141	C00484	19143	2011/1/6	2011/1/13	Standard Class

## Customer to Order Relationship:

A Customer may have one or more Orders.  
Each Order must belong to one and only one Customer.

**Customer.CustomerID (PK) --> Order.CustomerID (FK)**

# Creating ER Model – Relationship4

PostalCode	City	State	Region	Country
77095	Houston	Texas	Central	United States
60540	Naperville	Illinois	Central	United States
60540	Naperville	Illinois	Central	United States
60540	Naperville	Illinois	Central	United States
19143	Philadelphia	Pennsylvania	East	United States

OrderID	CustomerID	PostalCode	OrderDate	ShipDate	ShipMode
CA-2011-103	C00184	77095	2011/1/4	2011/1/8	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-112	C00477	60540	2011/1/5	2011/1/9	Standard Class
CA-2011-141	C00484	19143	2011/1/6	2011/1/13	Standard Class

## Location to Order Relationship:

A PostalCode may belong to one or more Orders.  
Each Order must have one and only one PostalCode.

**Location.PostalCode (PK) --> Order.PostalCode (FK)**

# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

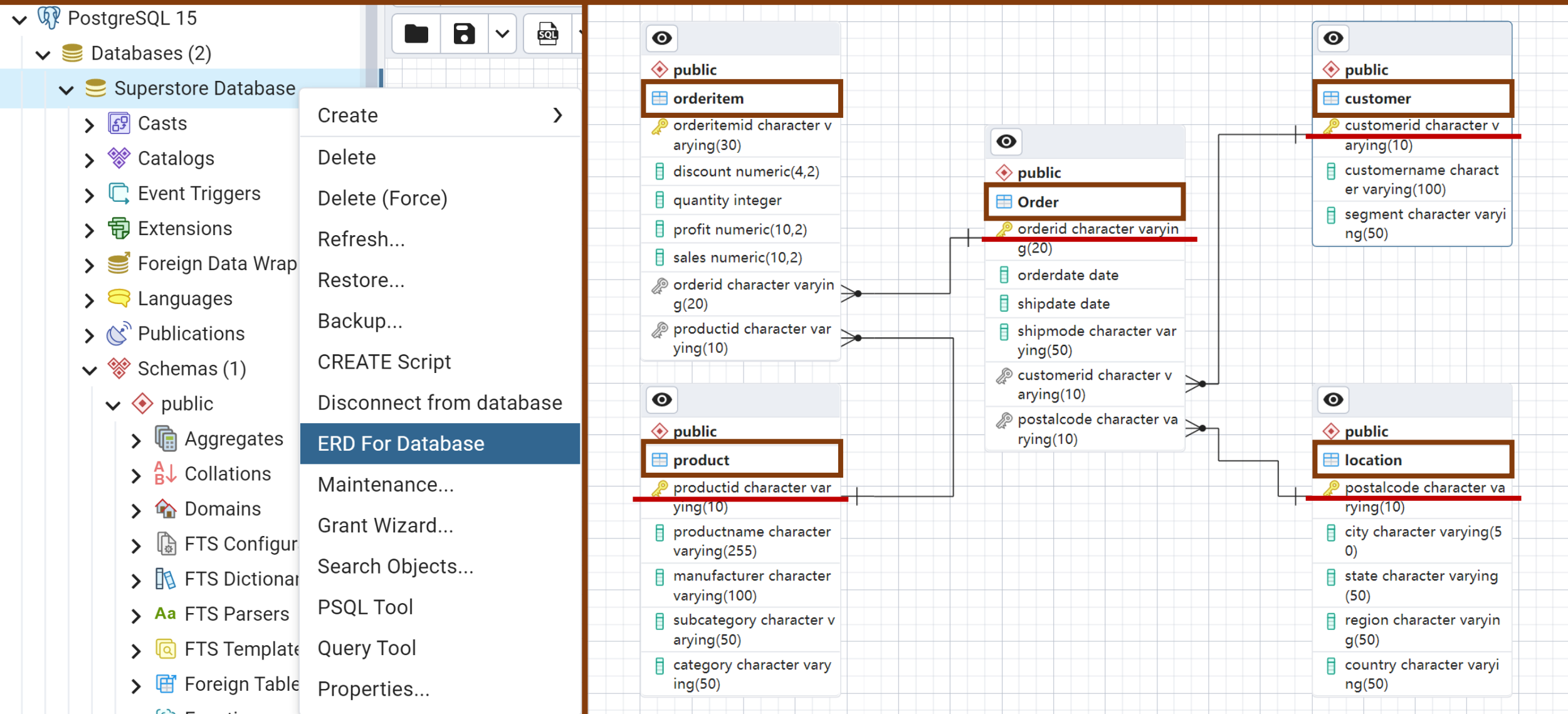
4. Creating Relational DB

5. 3NF Normalization

6. Further Implementation

**Entity  
Relationship  
Diagram**

# Creating ERD



# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

4. Creating Relational DB

5. 3NF Normalization

6. Further Implementation

**Relational  
Database  
Design**

# Creating Relational DB

Query		Query History	
1	SELECT	*	FROM public.orderitem
2	ORDER BY	orderitemid	ASC LIMIT 100
Data Output			
Messages			
Notifications			
	orderitemid	discount	quantity
	[PK] character	numeric (4,2)	integer
1	CA-2011-...	0.00	3
2	CA-2011-...	0.20	3
3	CA-2011-...	0.20	6
4	CA-2011-...	0.20	6
5	CA-2011-...	0.20	1

Query		Query History	
1	SELECT	*	FROM public.product
2	ORDER BY	productid	ASC LIMIT 100
Data Output			
Messages			
Notifications			
	productid	productname	manufacturer
	[PK] character	character varying	character vary
1	L00001	While you Wer...	Other
2	L00002	#10- 4 1/8 x 9 ...	Other
3	L00003	#10- 4 1/8 x 9 ...	Other
4	L00004	#10 Gummed ...	Other
5	L00005	#10 Self-Seal ...	Other

Query		Query History	
1	SELECT	*	FROM public."order"
2	ORDER BY	orderid	ASC LIMIT 100
Data Output			
Messages			
Notifications			
	orderid	orderdate	shipdate
	[PK] character	date	date
1	CA-2011-...	2011-09-07	2011-09-13
2	CA-2011-...	2011-07-08	2011-07-12
3	CA-2011-...	2011-03-14	2011-03-18
4	CA-2011-...	2011-01-29	2011-02-04
5	CA-2011-...	2011-04-08	2011-04-15

Query		Query History	
1	SELECT	*	FROM public.customer
2	ORDER BY	customerid	ASC LIMIT 100
Data Output			
Messages			
Notifications			
	customerid	customername	segment
		character varying (100)	character varying (50)
		William Brown	Consumer
		Paul Prost	Home Office
		Matt Abelman	Home Office
		John Lee	Consumer
		Seth Vernon	Consumer

Query		Query History	
1	SELECT	*	FROM public.location
2	ORDER BY	postalcode	ASC LIMIT 100
Data Output			
Messages			
Notifications			
	postalcode	city	state
	[PK] character varying	character varying	character varying
1	10009	New York City	New York
2	10011	New York City	New York
3	10024	New York City	New York
4	10035	New York City	New York
5	1040	Holyoke	Massachusetts



# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

4. Creating Relational DB

5. 3NF Normalization

6. Further Implementation

**Third  
Normal  
Form**

# 3NF Normalization

- ✓ Customer, Location, Order, and OrderItem relations are in **3NF**
- ✓ There is **no partial functional** and **no transitive functional dependencies**.

Customer (CustomerID, CustomerName, Segment)

◆ FD1: CustomerID → CustomerName, Segment

Location (PostalCode, City, State, Region, Country)

◆ FD1: PostalCode → City, State, Region, Country

Order (OrderID, OrderDate, ShipDate, ShipMode, CustomerID(FK), PostalCode(FK))

◆ FD1: OrderID → OrderDate, ShipDate, ShipMode, CustomerID(FK), PostalCode(FK)

OrderItem (OrderItemID, Discount, Quantity, Profit, Sales, OrderID(FK), ProductID(FK))

◆ FD1: OrderItemID → Discount, Quantity, Profit, Sales, OrderID(FK), ProductID(FK)

# 3NF Normalization

- ✓ Product is in 2NF and there is no partial functional dependencies.
- ✓ Product is not in 3NF since there is transitive functional dependencies.

Product (ProductID, ProductName, Manufacturer, SubCategory, Category)  
FD1: ProductID → ProductName, Manufacturer, SubCategory, Category  
FD2: SubCategory → Category

Create a new relation to put **SubCategory → Category**

Category (SubCategory, Category)

✓ FD1: SubCategory → Category

Product (ProductID, ProductName, Manufacturer, SubCategory)

✓ FD1: ProductID → ProductName, Manufacturer, SubCategory

# Project Agenda

1. Business Background

2. Creating ER Model

3. Creating ERD

4. Creating Relational DB

5. 3NF Normalization

6. Summarization

# Applications

# Summarization

## Final Relational Model in 3NF

Customer (CustomerID, CustomerName, Segment)

◆ FD1: CustomerID → CustomerName, Segment)

Location (PostalCode, City, State, Region, Country)

◆ FD1: PostalCode → City, State, Region, Country

Order (OrderID, OrderDate, ShipDate, ShipMode, CustomerID(FK), PostalCode(FK))

◆ FD1: OrderID → OrderDate, ShipDate, ShipMode, CustomerID(FK), PostalCode(FK)

OrderItem (OrderItemID, Discount, Quantity, Profit, Sales, OrderID(FK), ProductID(FK))

◆ FD1: OrderItemID → Discount, Quantity, Profit, Sales, OrderID(FK), ProductID(FK)

Category (SubCategory, Category)

◆ FD1: SubCategory → Category

Product (ProductID, ProductName, Manufacturer, SubCategory)

◆ FD1: ProductID → ProductName, Manufacturer, SubCategory

# Calculate the average sales per order in Chicago in the year 2013, grouped by customer segment.

Query Query History

```
1 SELECT
2     Segment,
3     AVG(sum_sales) AS AvgSalesPerOrder
4 FROM (
5     SELECT
6         o.OrderID,
7         c.Segment,
8         SUM(oi.Sales) AS sum_sales
9     FROM
10         "order" o
11     JOIN
12         OrderItem oi ON o.OrderID = oi.OrderID
13     JOIN
14         Customer c ON o.CustomerID = c.CustomerID
15     JOIN
16         Location l ON o.PostalCode = l.PostalCode
17     WHERE
18         l.City = 'Chicago'
19         AND EXTRACT(YEAR FROM o.OrderDate) = 2013
20     GROUP BY
21         o.OrderID, c.Segment
22 ) subquery
23 GROUP BY
24     Segment;
```

Data Output Messages Graph Visualiser X Notifications

	segment character varying (50)	avgsalesperorder numeric
1	Consumer	411.5714285714285714
2	Corporate	76.06666666666666667
3	Home Office	547.3333333333333333

✓ The result is correct!

✓ The structure is understandable!

✓ The queries are easily adaptable!

# Thanks

for  
Watching

