MSIS 2403:

DATABASE MANAGEMENT SYSTEMS

BLENDA HOANG

WINTER 2022

E-commerce Retailer Scenario

Alex navigates to www.Nordstrom.com (an upscale department store) to find some new pieces for the winter season at 8:45pm. He first clicks through to the Men's section and subsequently explores the "Bestsellers" page. Since shopping for clothes online can be a hit or miss, he filters the page to sort by customer rating to see what items customers loved the most. Nothing catches his eye, so he decides to search for "men's slippers" in the search bar on the upper righthand side of the page. He proceeds to filter on shoe size = "8" and sorts by price "low to high". He finds a pair of Ugg's indoor slippers in olive green and checks if any nearby stores carry this item for pickup. All the nearby stores are out of stock in this item, so he proceeds to add the item to his cart. From there, he clicks to his cart to do a guest checkout. He selects his shipping preference as "Standard Shipping" and adds in his shipping address, contact information (email and phone number), and payment information (credit card and billing address). He then clicks "Place Order" and completes his transaction at 9:14pm.

Table Creation

```
13  create table product_data(
14  ItemNumber varchar(250) primary key,
15  Brand text,
16  Description text,
17  Price double);
18
19  create table customer_data(
20  Email varchar(250) primary key,
21  PaymentMethod text,
22  BillTo text,
23  BillingAddress text,
24  BillingCity text,
25  BillingState text,
26  BillingZipCode text,
27  ShipTo text,
28  ShippingName text,
29  ShippingAddress text,
30  ShippingCity text,
31  ShippingState text,
32  ShippingZipCode text);
33
34  create table order_data(
35  OrderNumber int primary key,
36  OrderDateTime text,
37  OrderStatus text);
38
39  create table shipping_data(
40  ShippingMethod varchar(250) primary key,
41  ShippingCost int);
42
43  create table transaction_data(
44  RecordNumber varchar(250) primary key,
45  OrderNumber int,
46  Color text,
47  Size text,
48  Qty int,
49  ItemNumber varchar(250),
50  EstimatedTax double,
51  OrderTotal double,
52  ShipTo text,
53  ShippingMethod varchar(250),
54  Email varchar(250),
55  constraint FK_1 foreign key (OrderNumber) references order_data (OrderNumber),
56  constraint FK_2 foreign key (ItemNumber) references product_data (ItemNumber),
57  constraint FK_3 foreign key (ShippingMethod) references shipping_data (ShippingMethod),
58  constraint FK_4 foreign key (Email) references customer_data (Email));
```

```sql
insert into product_data
(ItemNumber, Brand, Description, Price)
select distinct(ItemNumber), Brand, Description, Price from HW.sql_import;

insert into customer_data
(Email, PaymentMethod, BillTo, BillingAddress, BillingCity, BillingState, BillingZipCode,
ShipTo, ShippingName, ShippingAddress, ShippingCity, ShippingState, ShippingZipCode)
select distinct(Email), PaymentMethod, BillTo, BillingAddress, BillingCity, BillingState,
BillingZipCode, ShipTo, ShippingName, ShippingAddress, ShippingCity, ShippingState,
ShippingZipCode from HW.sql_import;

insert into order_data
(OrderNumber, OrderDateTime, OrderStatus)
select distinct(OrderNumber), OrderDateTime, OrderStatus from HW.sql_import;

insert into shipping_data
(ShippingMethod, ShippingCost)
select distinct(ShippingMethod), ShippingCost from HW.sql_import;

insert into transaction_data
(RecordNumber, OrderNumber, Color, Size, Qty, ItemNumber, EstimatedTax, OrderTotal, ShipTo,
ShippingMethod, Email)
select RecordNumber, OrderNumber, Color, Size, Qty, ItemNumber, EstimatedTax, OrderTotal, ShipTo,
ShippingMethod, Email from HW.sql_import;
```

```sql
select * from shipping_data;
```

100% 23:81

Result Grid Filter Rows: Search Edit: Export/Import:

ShippingMethod	ShippingCost
Express Shipping	10
Standard Shipping	0
NULL	NULL

```sql
select * from transaction_data;
```

100% 23:71

Result Grid Filter Rows: Search Edit: Export/Import:

RecordNumber	OrderNumber	Color	Size	Qty	ItemNumber	EstimatedTax	OrderTotal	ShipTo	ShippingMethod	Email
1	343194572	BURNT OLIVE SUEDE	8 M	1	#31405	4.92	52.92	Alexander Chin	Standard Shipping	aredmondchin@gmail.com
10	343194579	Volcano	One Size	1	#6572401	1.53	18.03	Julian Dixon	Standard Shipping	jdd1992@yahoo.com
11	343194580	Dove	One Size	1	#128771	10.2	120.45	Rachel Alvarez	Standard Shipping	r_margarita@yahoo.com
12	343194580	Soft Camel	One Size	1	#128771	10.2	120.45	Rachel Alvarez	Standard Shipping	r_margarita@yahoo.com
13	343194581	Rhodium	6	1	#5850638	13.88	163.88	Chris Ha	Standard Shipping	cycling_rocks@hotmail.com
2	343194573	Black	One Size	1	#6205127	110.08	1300.08	Stephanie Hoang	Standard Shipping	shoaaang@yahoo.com
3	343194573	Cognac Patent	6	1	#6192417	11.1	131.05	Stephanie Hoang	Standard Shipping	shoaaang@yahoo.com
4	343194573	Intense	One Size	1	#5597577	2.78	32.78	Stephanie Hoang	Standard Shipping	shoaaang@yahoo.com
5	343194574	Core Black/White/Grey	10 M	1	#6173014	13.88	163.88	Cassandra Chu	Standard Shipping	cassiecc@aol.com
6	343194575	Pinecone Brown	One Size	2	#5943770	5.37	73.37	Amy Capener	Express Shipping	dmoreno20@gmail.com
7	343194576	N/A	1 oz	1	#6227770	6.85	80.85	Dennis Chin	Standard Shipping	d_chin_5@yahoo.com
8	343194577	Clear	One Size	1	#6579846	3.82	45.07	Amy Sun	Standard Shipping	sunnyflowers@gmail.com
9	343194578	Yellow	One Size	1	#6047771	6.11	72.11	James Tsai	Standard Shipping	james_tsai95@hotmail.com
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

```
86 •  select * from product_data;
```

100% 39:80

Result Grid Filter Rows: Search Edit: Export/Import:

ItemNumber	Brand	Description	Price
#128771	BAREFOOT DREAMS	CozyChic Throw Blanket	110.25
#31405	UGG	UGG Scuff Slipper (Men)	48
#5597577	Chanel	LES BEIGES HEALTHY GLOW LIP BALM	30
#5850638	NADRI	Set of 5 Stacking Rings	150
#5943770	The North Face	Dock Worker Recycled Beanie	58
#6047771	OHOM	Ui Mug & Warmer Set	66
#6173014	ADIDAS	NMD R1 Primeblue Sneaker	150
#6192417	Steve Madden	Lynden Block Heel Bootie	119.95
#6205127	Balenciaga	Hourglass Leather Wallet on a Chain	1190
#6227770	JO MALONE LONDON	Blossoms Nashi Blossom Cologne	74
#6572401	CAPRI BLUE	Volcano Petite Jar Candle	16.5
#6579846	FERM LIVING	Still Carafe	41.25
NULL	NULL	NULL	NULL

```
87 •  select * from customer_data;
```

100% 29:79

Result Grid Filter Rows: Search Edit: Export/Import:

Email	PaymentMethod	BillTo	BillingAddress	BillingCity	BillingState	BillingZipCode	ShipTo	ShippingName	Ship
aredmondchin@gmail.com	Visa ending in 0297	Alexander Chin	1000 Capitola Court	Seattle	WA	10009	Alexander Chin	Alexander Chin	103
cassiecc@aol.com	Gift card	Cassandra Chu	1000 Capitola Court	Seattle	WA	10009	Cassandra Chu	Cassandra Chu	100
cycling_rocks@hotmail.com	Visa ending in 5666	Chris Ha	111 Sandy Dune Way	Santa Barbara	CA	92341	Chris Ha	Chris Ha	111
d_chin_5@yahoo.com	Visa ending in 5274	Dennis Chin	301 Nickel Lane	Milpitas	CA	95677	Dennis Chin	Dennis Chin	301
dmoreno20@gmail.com	Visa ending in 1189	Dominic Chin	2943 Pescadero Ter	Fremont	CA	94538	Amy Capener	Amy Capener	157
james_tsai95@hotmail.com	Visa ending in 9070	James Tsai	470 Fremont Blvd	Fremont	CA	94537	James Tsai	James Tsai	470
jdd1992@yahoo.com	Visa ending in 7700	Julian Dixon	1 Hacker Way	Menlo Park	CA	94035	Julian Dixon	Julian Dixon	1 H
r_margarita@yahoo.com	Nordstrom Note	Rachel Alvarez	1490 Cherry Lane	Sacramento	CA	98921	Rachel Alvarez	Rachel Alvarez	149
shoaaang@yahoo.com	Visa ending in 0693	Stephanie Hoang	2906 Stanhope Drive	San Jose	CA	95121	Stephanie Hoang	Stephanie Hoang	290
sunnyflowers@gmail.com	Visa ending in 8860	Amy Sun	5967 Santos Lane	Walnut Creek	CA	97824	Amy Sun	Amy Sun	596
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

```
88 •  select * from order_data;
```

100% 38:82

Result Grid Filter Rows: Search Edit: Export/Import:

OrderNumber	OrderDateTime	OrderStatus
343194572	1/16/22 5:29	In Process
343194573	1/14/22 0:33	Shipped
343194574	1/7/22 6:35	Shipped
343194575	1/16/22 21:10	Shipped
343194576	2/1/22 17:06	Shipped
343194577	2/4/22 14:48	In Process
343194578	2/18/22 17:16	In Process
343194579	2/23/22 8:54	In Process
343194580	3/11/22 20:58	Shipped
343194581	3/17/22 8:52	In Process
NULL	NULL	NULL

Business Questions

1. What are the prices of items purchased sorted from highest to lowest price?

- Used a left outer join to retain all the transaction data and merge in product related information

- Used the 'order by desc' command to sort prices from high to low

```
74 •  select RecordNumber, OrderNumber, transaction_data.ItemNumber, Description, Price
75        from transaction_data left outer join product_data on
76        transaction_data.ItemNumber = product_data.ItemNumber
77        order by Price desc;
```

100% 26:70

Result Grid Filter Rows: Search Export:

RecordNumber	OrderNumber	ItemNumber	Description	Price
2	343194573	#6205127	Hourglass Leather Wallet on a Chain	1190
13	343194581	#5850638	Set of 5 Stacking Rings	150
5	343194574	#6173014	NMD R1 Primeblue Sneaker	150
3	343194573	#6192417	Lynden Block Heel Bootie	119.95
11	343194580	#128771	CozyChic Throw Blanket	110.25
12	343194580	#128771	CozyChic Throw Blanket	110.25
7	343194576	#6227770	Blossoms Nashi Blossom Cologne	74
9	343194578	#6047771	Ui Mug & Warmer Set	66
6	343194575	#5943770	Dock Worker Recycled Beanie	58
1	343194572	#31405	UGG Scuff Slipper (Men)	48
8	343194577	#6579846	Still Carafe	41.25
4	343194573	#5597577	LES BEIGES HEALTHY GLOW LIP...	30
10	343194579	#6572401	Volcano Petite Jar Candle	16.5

2. How many orders are currently in process versus completed (i.e., shipped)?

- Used the 'count' function to count all the orders

- Used the group by function to split the order count by in process versus completed orders

```
88 •  select OrderStatus, count(*) from order_data
89    group by OrderStatus;
```

00% 32:80

Result Grid Filter Rows: Search Export:

OrderStatus	count(*)
In Process	5
Shipped	5

3. What is the average order amount by customer?

- Used the 'avg' function to find customer average total order amount, rounded to 2 digits via the 'round' function
- Used the group by function to return average order amount by customer

```
91 •  select ShipTo, round(avg(OrderTotal),2) as Avg_Order_Amount from transaction_data
92    group by ShipTo;
```

100% 29:82

Result Grid | Filter Rows: Search | Export:

ShipTo	Avg_Order_Amount
Alexander Chin	52.92
Julian Dixon	18.03
Rachel Alvarez	120.45
Chris Ha	163.88
Stephanie Hoang	487.97
Cassandra Chu	163.88
Amy Capener	73.37
Dennis Chin	80.85
Amy Sun	45.07
James Tsai	72.11

4. What is the split of customers who use Visa cards, gift cards, or Nordstrom Notes as payment on their order?

- Used a case statement to find count of all Visa, Gift Card, and Nordstrom Note payment types and a wildcard operator since the Visa records had the last 4 digits of the CC which made each Visa record unique

```
97 •  select
98        count(case when PaymentMethod like '%Visa%' then 1 else null end) as Visa,
99        count(case when PaymentMethod like '%Gift%' then 1 else null end) as Gift_Card,
100       count(case when PaymentMethod like '%Nordstrom%' then 1 else null end) as Nordstrom_Note
101   from customer_data;
```

100% 17:92

Result Grid | Filter Rows: Search | Export:

Visa	Gift_Card	Nordstrom_Note
8	1	1

5. What is the ratio of one size items versus items that come in multiple sizes?

- Used a case statement and a ratio formula to calculate the ratio of one sized items which are easier to stock versus multiple size items (which require more inventory space and can be difficult to sell through)

```
113 •  select sum(case when Size = 'One Size' then 1 else 0 end) / count(size) as Ratio_OneSize_Items
114    from transaction_data;
```

100% 20:108

Result Grid | Filter Rows: Search | Export:

Ratio_OneSize_Items
0.6154

6. What is the total average order amount?

- Used the 'sum' and 'count' functions to calculate the overall total order average for the entire dataset

```
116 •  select sum(OrderTotal) / count(OrderTotal) as Total_Avg_Order
117    from transaction_data;
```

100% 29:106

Result Grid | Filter Rows: Search | Export:

Total_Avg_Order
182.68615384615384

7. What is the split of High Tier, Mid-Tier, and Low Tier items?

- Used a case statement to create a new column identifying the ranking of each item based on price
- Used the 'order by' function to sort the prices from how to low

```
119 •   select Brand, Description, Price,
120     case
121         when price > 500 then 'High Tier'
122         when (price < 500) and (price > 100) then 'Mid Tier'
123         else 'Low Tier'
124     end as Ranking
125     from product_data
126     order by Price desc;
```

100% 23:114

Result Grid Filter Rows: Search Export:

Brand	Description	Price	Ranking
Balenciaga	Hourglass Leather Wallet on a Chain	1190	High Tier
NADRI	Set of 5 Stacking Rings	150	Mid Tier
ADIDAS	NMD R1 Primeblue Sneaker	150	Mid Tier
Steve Madden	Lynden Block Heel Bootie	119.95	Mid Tier
BAREFOOT DREAMS	CozyChic Throw Blanket	110.25	Mid Tier
JO MALONE LONDON	Blossoms Nashi Blossom Cologne	74	Low Tier
OHOM	Ui Mug & Warmer Set	66	Low Tier
The North Face	Dock Worker Recycled Beanie	58	Low Tier
UGG	UGG Scuff Slipper (Men)	48	Low Tier
FERM LIVING	Still Carafe	41.25	Low Tier
Chanel	LES BEIGES HEALTHY GLOW LIP...	30	Low Tier
CAPRI BLUE	Volcano Petite Jar Candle	16.5	Low Tier

8. What is the distribution of customers that use Gmail, AOL, Hotmail, or Yahoo?

- Used a case statement and wildcard operator to identify customers by email provider
- Used the 'count' function and a group by statement to find the total count per email type

```
128 •   select count(Email),
129     case
130         when email like "%gmail%" then "Gmail_User"
131         when email like "%aol%" then "AOL_User"
132         when email like "%hotmail%" then "Hotmail_User"
133         when email like "%yahoo%" then "Yahoo_User"
134         else "N/A"
135         end as Email_List
136         from Customer_data
137         group by Email_List;
```

100% 18:125

Result Grid Filter Rows: Search Export:

count(Email)	Email_List
3	Gmail_User
1	AOL_User
2	Hotmail_User
4	Yahoo_User

9. What items were shipped to CA (include Brand & Item Number)?

- Used two inner joins to pull the relevant column data required from customer, transaction, and product data

- Used the 'having' function to filter the data to only CA records

```
172 •   select ShippingState, Brand, pd.ItemNumber
173     from customer_data as cd
174     inner join transaction_data as td
175     on cd.email = td.email
176     inner join product_data as pd
177     on pd.ItemNumber = td.ItemNumber
178     having ShippingState = 'CA'
179     order by 1 asc;
```

100% 19:167

Result Grid Filter Rows: Search Export:

ShippingState	Brand	ItemNumber
CA	UGG	#31405
CA	ADIDAS	#6173014
CA	NADRI	#5850638
CA	JO MALONE LONDON	#6227770
CA	OHOM	#6047771
CA	CAPRI BLUE	#6572401
CA	BAREFOOT DREAMS	#128771
CA	BAREFOOT DREAMS	#128771
CA	Balenciaga	#6205127
CA	Steve Madden	#6192417
CA	Chanel	#5597577
CA	FERM LIVING	#6579846

10. What items being shipped to CA are still "in Process"?

- Used 2 inner joins to retrieve transaction data along with order and customer data

- Used 'where' command to filter data to only 'In Process' orders

```
181 •   select ShippingState, OrderStatus
182     from customer_data as cd
183     inner join transaction_data as td
184     on cd.email = td.email
185     inner join order_data as od
186     on od.OrderNumber = td.OrderNumber
187     where OrderStatus = 'In Process'
188     order by 1 asc;
```

100% 16:188

Result Grid Filter Rows: Search Export:

ShippingState	OrderStatus
CA	In Process
CA	In Process
CA	In Process
CA	In Process
CA	In Process

11. Which customers ordered items from Barefoot Dreams (return list of up to 10 customers)?

- Used 'where' function to join two data sets to retrieve transaction information along with product information
- Used 'and' function to filter data to only the Barefoot Dreams brand items
- Used the 'limit' function to restrict the data to max of 10 records

```
190 •   select ShipTo, td.OrderNumber, td.ItemNumber, Brand, Description
191     from transaction_data td, product_data pd
192     where td.ItemNumber = pd.ItemNumber
193     and Brand = 'BAREFOOT DREAMS'
194     limit 10;
```

100% 16:179

Result Grid Filter Rows: Search Export:

ShipTo	OrderNumber	ItemNumber	Brand	Description
Rachel Alvarez	343194580	#128771	BAREFOOT DREAMS	CozyChic Throw Blanket
Rachel Alvarez	343194580	#128771	BAREFOOT DREAMS	CozyChic Throw Blanket

12. Which orders were paid for with Visas or Nordstrom Notes that had Billing State of 'CA'?

- Used an inner join to return transaction data along with customer information
- Used 'where' function to restrict Payment Method types
- Used a subquery to restrict the Billing State to only CA state

```
196 •   select OrderNumber, PaymentMethod
197     from customer_data as cd
198     inner join transaction_data as td
199     on cd.email = td.email
200     where (PaymentMethod like 'Visa%' or PaymentMethod like 'Note%')
201     and BillingState in (
202         select BillingState
203         from customer_data
204         where BillingState = 'CA'
205     );
```

100% 36:192

Result Grid Filter Rows: Search Export:

OrderNumber	PaymentMethod
343194581	Visa ending in 5666
343194576	Visa ending in 5274
343194575	Visa ending in 1189
343194578	Visa ending in 9070
343194579	Visa ending in 7700
343194573	Visa ending in 0693
343194573	Visa ending in 0693
343194573	Visa ending in 0693
343194577	Visa ending in 8860

13. What orders were placed in February 2022?
- Used 2 inner joins to return transaction data, order data, and product data
- Used 'where' and 'date' functions to return only transactions in February

```
207 •  select OrderDateTime, od.OrderNumber, Brand, Color
208    from order_data od
209    inner join transaction_data td on od.OrderNumber = td.OrderNumber
210    inner join product_data pd on td.ItemNumber = pd.ItemNumber
211    where month(STR_TO_DATE(`OrderDateTime`, '%m/%d/%Y')) = 2;
212
```

100% 21:202

Result Grid Filter Rows: Search Export:

OrderDateTime	OrderNumber	Brand	Color
2/1/22 17:06	343194576	JO MALONE LONDON	N/A
2/4/22 14:48	343194577	FERM LIVING	Clear
2/18/22 17:16	343194578	OHOM	Yellow
2/23/22 8:54	343194579	CAPRI BLUE	Volcano

14. For Express Shipping, update the shipping cost from $10 to $8. **Dataset Modification*

```
103 •  update shipping_data
104    set ShippingCost = 8
105    where ShippingCost = 10;
106 •  select * from shipping_data;
107
```

100% 20:101

Result Grid Filter Rows: Search Edit: Export/Import:

ShippingMethod	ShippingCost
Express Shipping	8
Standard Shipping	0
NULL	NULL

15. For the FERM LIVING brand, discount all items 25% due to upcoming promotion. **Dataset Modification*

```
108 •  update product_data
109    set Price = round(Price * 0.75,2)
110    where Brand like '%Ferm%';
111 •  select * from product_data;
112
```

100% 21:104

Result Grid Filter Rows: Search Edit: Export/Import:

ItemNumber	Brand	Description	Price
#128771	BAREFOOT DREAMS	CozyChic Throw Blanket	110.25
#31405	UGG	UGG Scuff Slipper (Men)	48
#5597577	Chanel	LES BEIGES HEALTHY GLOW LIP BALM	30
#5850638	NADRI	Set of 5 Stacking Rings	150
#5943770	The North Face	Dock Worker Recycled Beanie	58
#6047771	OHOM	Ui Mug & Warmer Set	66
#6173014	ADIDAS	NMD R1 Primeblue Sneaker	150
#6192417	Steve Madden	Lynden Block Heel Bootie	119.95
#6205127	Balenciaga	Hourglass Leather Wallet on a Chain	1190
#6227770	JO MALONE LONDON	Blossoms Nashi Blossom Cologne	74
#6572401	CAPRI BLUE	Volcano Petite Jar Candle	16.5
#6579846	FERM LIVING	Still Carafe	41.25
NULL	NULL	NULL	NULL

Views

1. Create view for customers that have billing city as “San Jose”.

```
158 •  create view SanJose_Customers as
159    select email, billingaddress, shippingaddress
160    from customer_data
161    where billingcity like '%San Jose%';
162 •  select * from SanJose_Customers;
```

100% 33:162

Result Grid Filter Rows: Search Export:

email	billingaddress	shippingaddress
shoaaang@yahoo.com	2906 Stanhope Drive	2906 Stanhope Drive

2. Create view for “Low Tier Items” (items priced below $100).

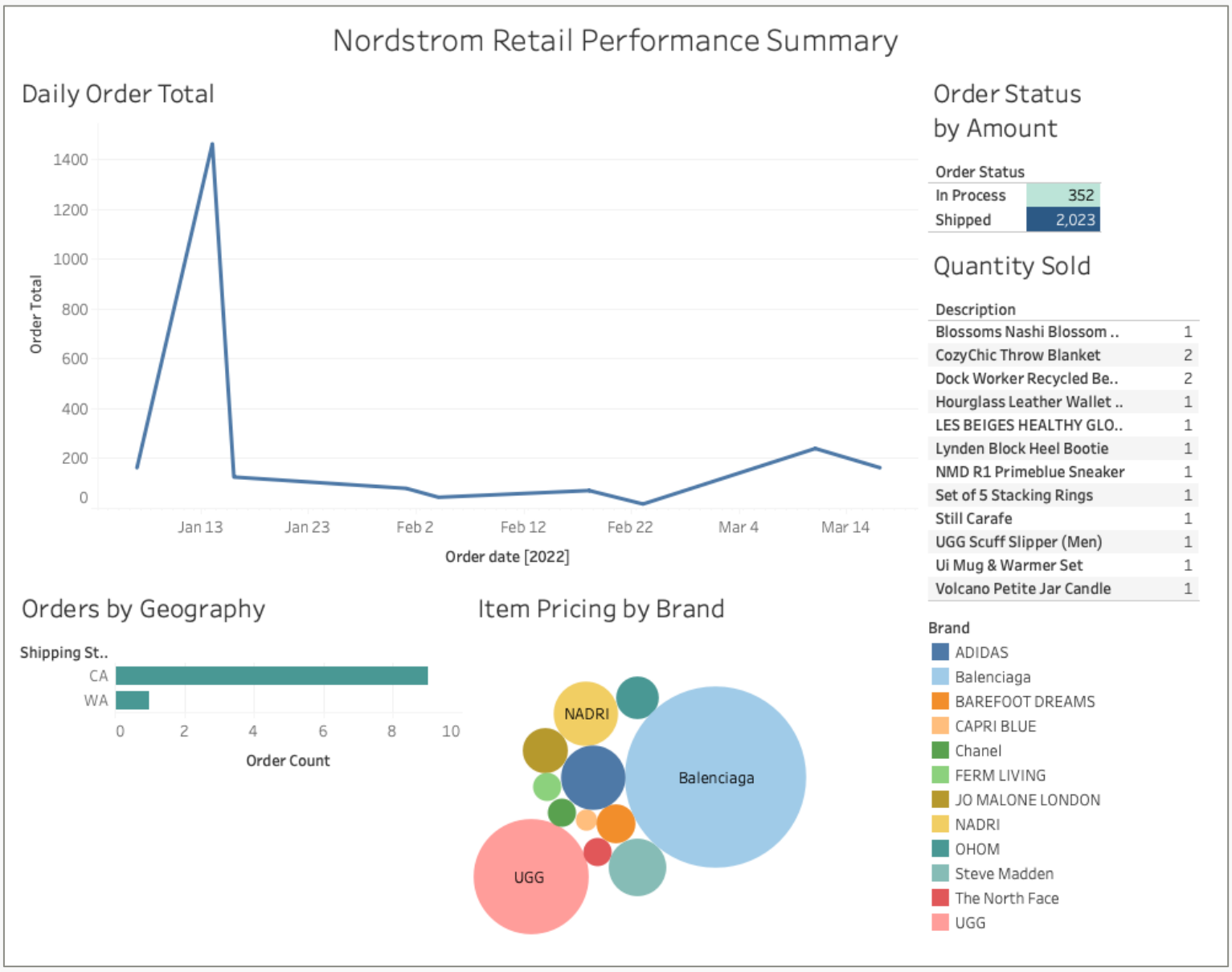
```
164 •  create view Low_Tier_Items as
165    select ItemNumber, Brand, Description, Price
166    from product_data
167    where Price < 100;
168 •  select * from Low_Tier_Items;
169
```

100% 1:157

Result Grid Filter Rows: Search Export:

ItemNumber	Brand	Description	Price
#31405	UGG	UGG Scuff Slipper (Men)	48
#5597577	Chanel	LES BEIGES HEALTHY GLOW LIP BALM	30
#5943770	The North Face	Dock Worker Recycled Beanie	58
#6047771	OHOM	Ui Mug & Warmer Set	66
#6227770	JO MALONE LONDON	Blossoms Nashi Blossom Cologne	74
#6572401	CAPRI BLUE	Volcano Petite Jar Candle	16.5
#6579846	FERM LIVING	Still Carafe	41.25

Tableau Report



Conclusion

For this project, I explored the data through the lens of a retailer, pulling key patterns and insights such as customer demographics, sales trends, and what portion of sales are currently still in process versus completed. The thought process behind the structure of my queries was to solve for questions a retailer may be concerned about to improve upon their business and sales metrics. For example, if there isn't a large queue of orders in process, it indicates lagging sales and there may be a need for an accelerant to boost sales revenue.

Given this project was completed on a limited scope of manufactured data, in hindsight, having a larger dataset may have been beneficial to identify thematic, overarching trends more easily. Although the amount of data was sufficient for various types of analyses, a more robust dataset might have allowed for different business questions to be asked. Other opportunities to improve upon the scenario was to have data from the customer's interaction within the website (i.e., what products they clicked, if they read reviews, etc.) versus primarily transactional data.