MIS 633-B Final Presentation

Developing a Data Warehouse & Business Intelligence system For Uptown Brew

Group-10 Members:

- *Richard Nyhus*
- *HarshaVardhan Reddy Putta*
- *Paras pandhree*
- *Venkata Jaya Chandra Mullapudi*
- *Aishway Pauranik*

Introduction

- Across US coffee culture, however, one establishment bucks the trend and has persisted in spite of the third wave of coffee. This is the diner, which is now irrevocably intertwined with how people perceive coffee in the US.

- Since the founding of the US' coffeehouses before the American Revolution, coffee shop culture in the country has evolved significantly. From war rations to the third wave, it's clear that coffee has been an inextricable aspect of US culture for centuries.

- According to the overview led by the National Coffee Association in the United States, over 70% of the customers like at-home espresso readiness, and 59% of the espresso consumed everyday is a connoisseur espresso. Subsequently, a steady shift has been seen from soda pops to espresso drinks among buyers in the locale as of late.

- The penetration rates of single-serve brewers in households in the largest cities in the United States is approximately 23% to 75%. In the United States, single-cup brewers are found in 41% of homes and 28% of offices, according to the National Coffee Drinking Trends survey

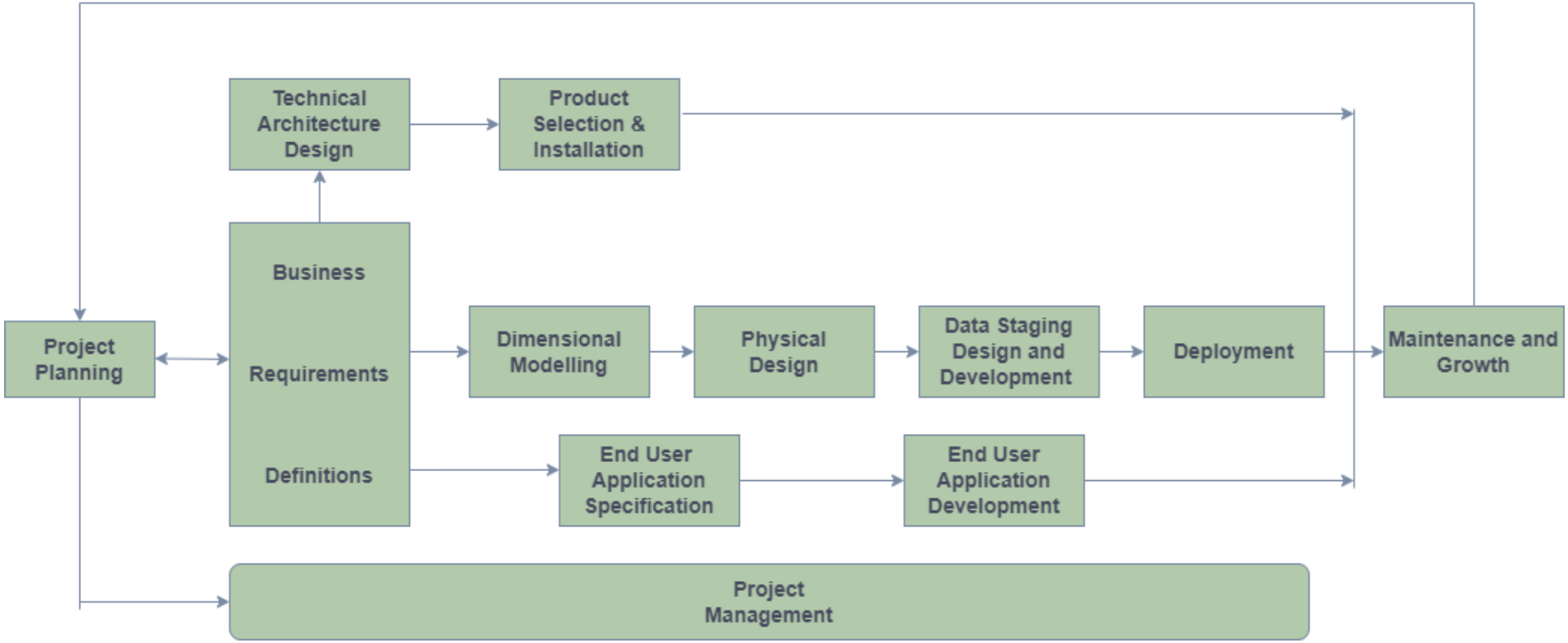
Company Overview: Uptown Brew

Company Overview: Uptown Brew

- Uptown Brew is a New York-based coffee chain with 1 warehouse and 8 brick-and-mortar locations in New York.
- Uptown Brew sells perishable and non-perishable goods such as bakery goods, coffee, and merchandise.
- In recent years, it has consistently generated an annual income of approx. $8.75 million.
- It is currently stated goal for its first quarter of this decade is to expand operations to bay areas and mid-Atlantic states and begin marketing its products and merchandise to new customers.
- Hence, the Coffee chain still possesses a highly outdated legacy in the Bi team, with the MongoDB server as the database but disjoined reporting and lack of data knowledge across the groups of the organization.

Systems Problem/Solution

Business Lifecycle Approach



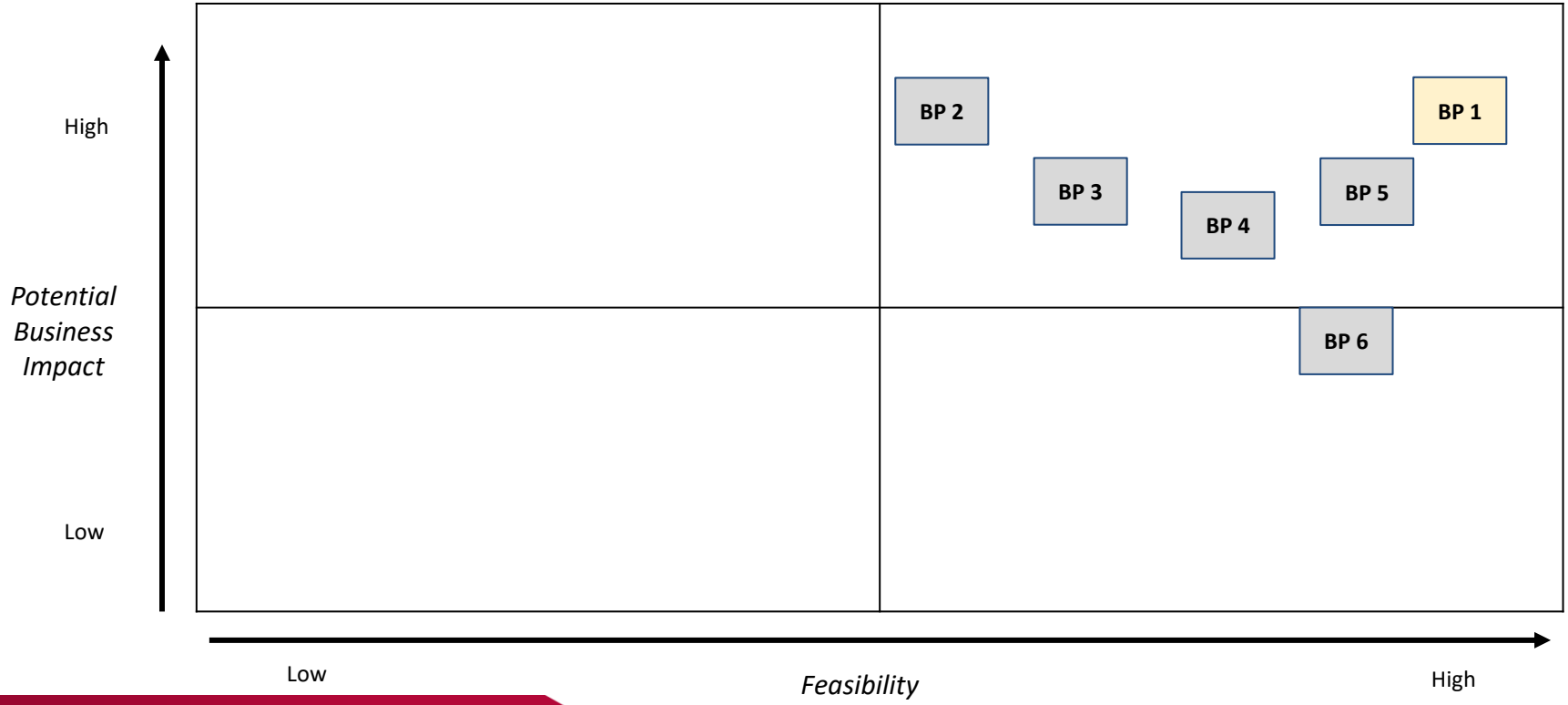
Opportunity Matrix

Business Process / Event	Sales	Logistics / Shipping	Customer Service / Marketing	Operations / Finance
1. In-store Sales Transactions			X	X
2. Shipping Sales Transactions		X		
3. Processing In-store Transactions	X			
4. Processing Supplier Transactions	X			
5. Refresh Inventory		X		X
6. Priority Customer Identification			X	

Bus Architecture Matrix

Business Process / Event	Date	Product	Payment Method	Customer	Staff	Sales Outlet	Shipping
1. In-store Sales Transactions	X	X	X	X	X	X	
2. Shipping Sales Transactions	X	X	X		X		X
3. Processing In-store Transactions	X	X	X		X	X	
4. Processing Shipping Transactions	X	X	X		X	X	X
5. Refresh Inventory		X	X				X
6. Priority Customer Identification				X		X	

Prioritization Grid



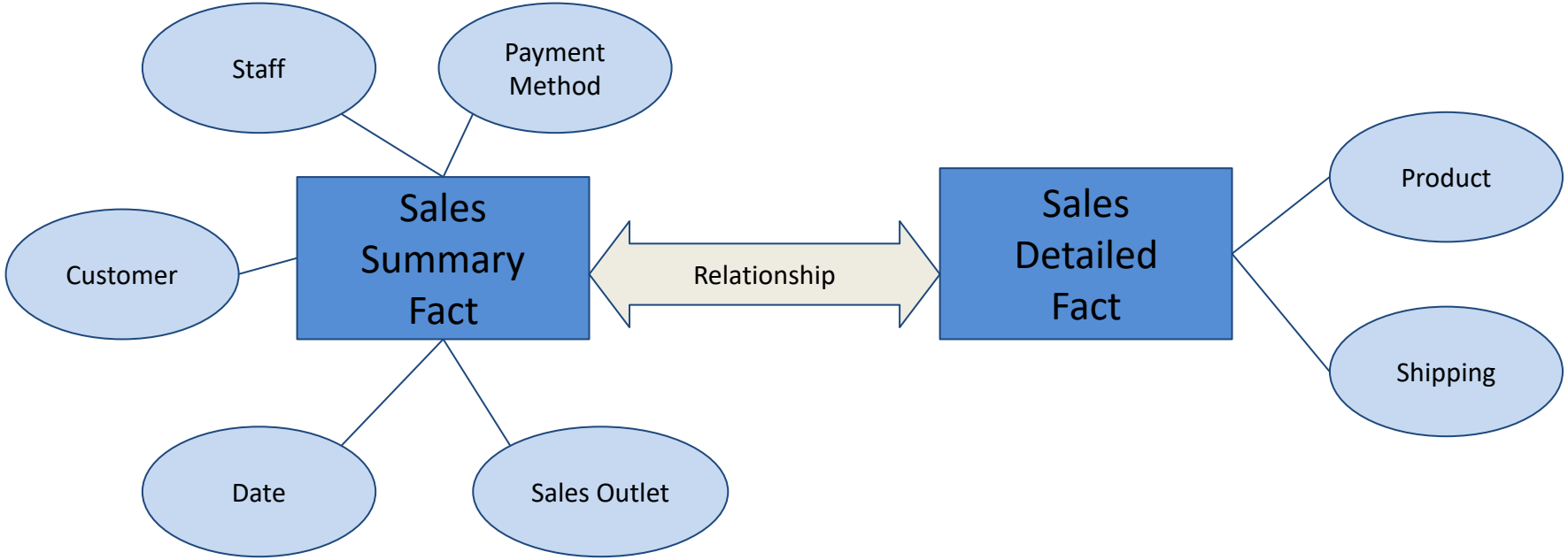
Data Transformation

After extraction, the information should be changed. For example, bringing together information types, managing spelling blunders, wiping out information uncertainty, and parsing into standard configurations. Information change is normally the most intricate part. It is additionally the longest advance in ETL improvement. The scope of information change is extremely wide. The interaction is to change the information from basic information type to very intricate information cleaning innovation.

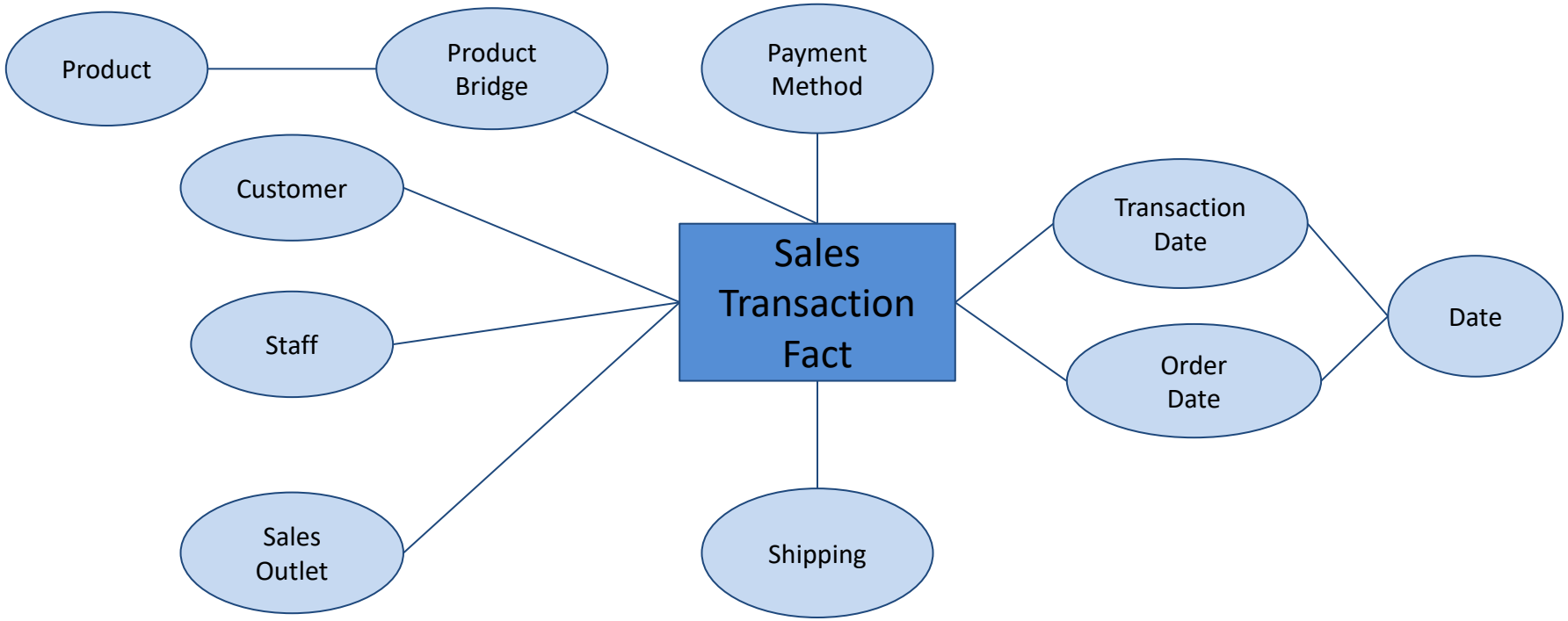
(7) Data Transformation Rules:

1. Only specific data columns are loaded
2. Unified coding
3. Reordering improves query performance
4. Merge data sources and remove duplicates
5. Row column Transformation
6. Merge duplicate columns
7. Data validation

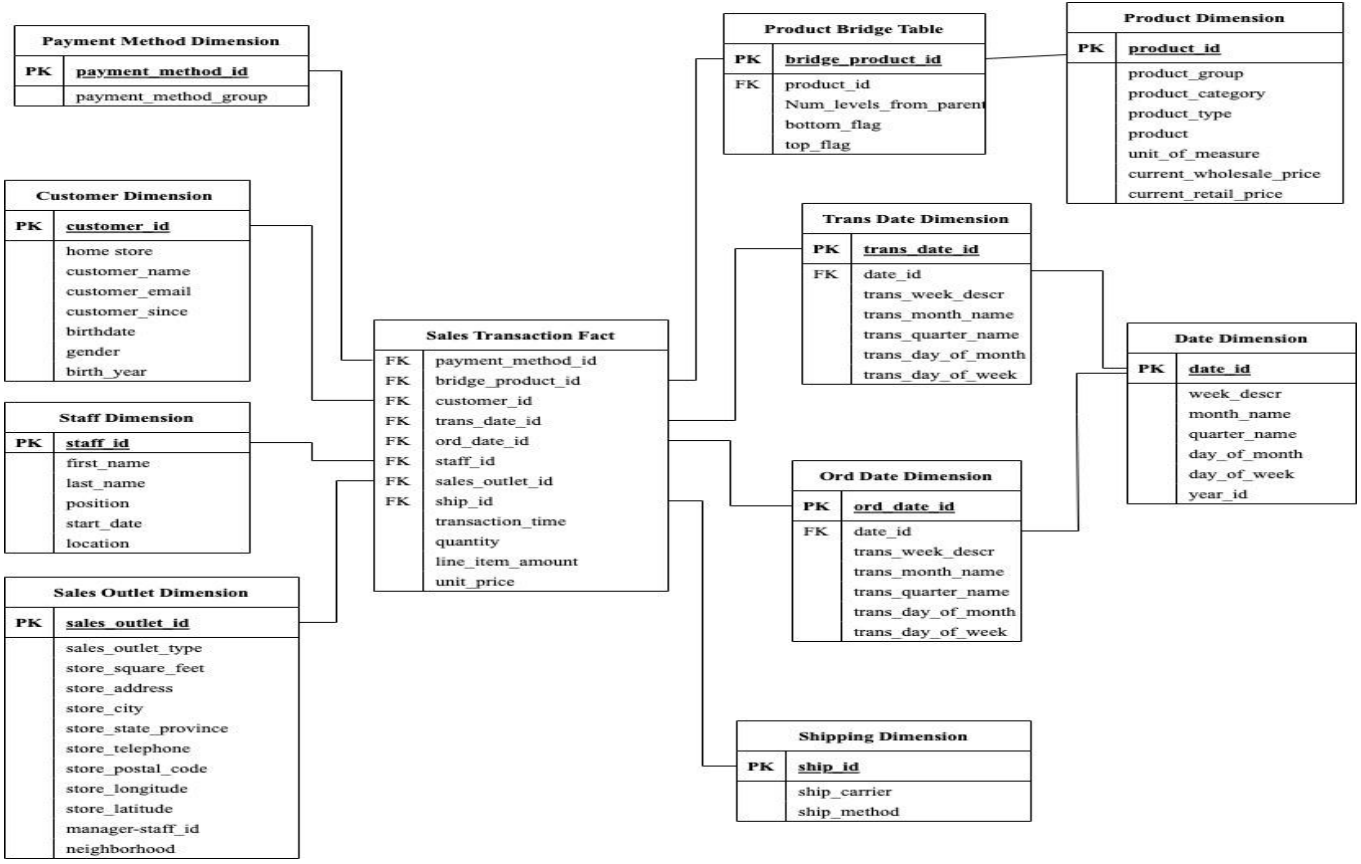
Logical Fact Diagram



High End Data Model



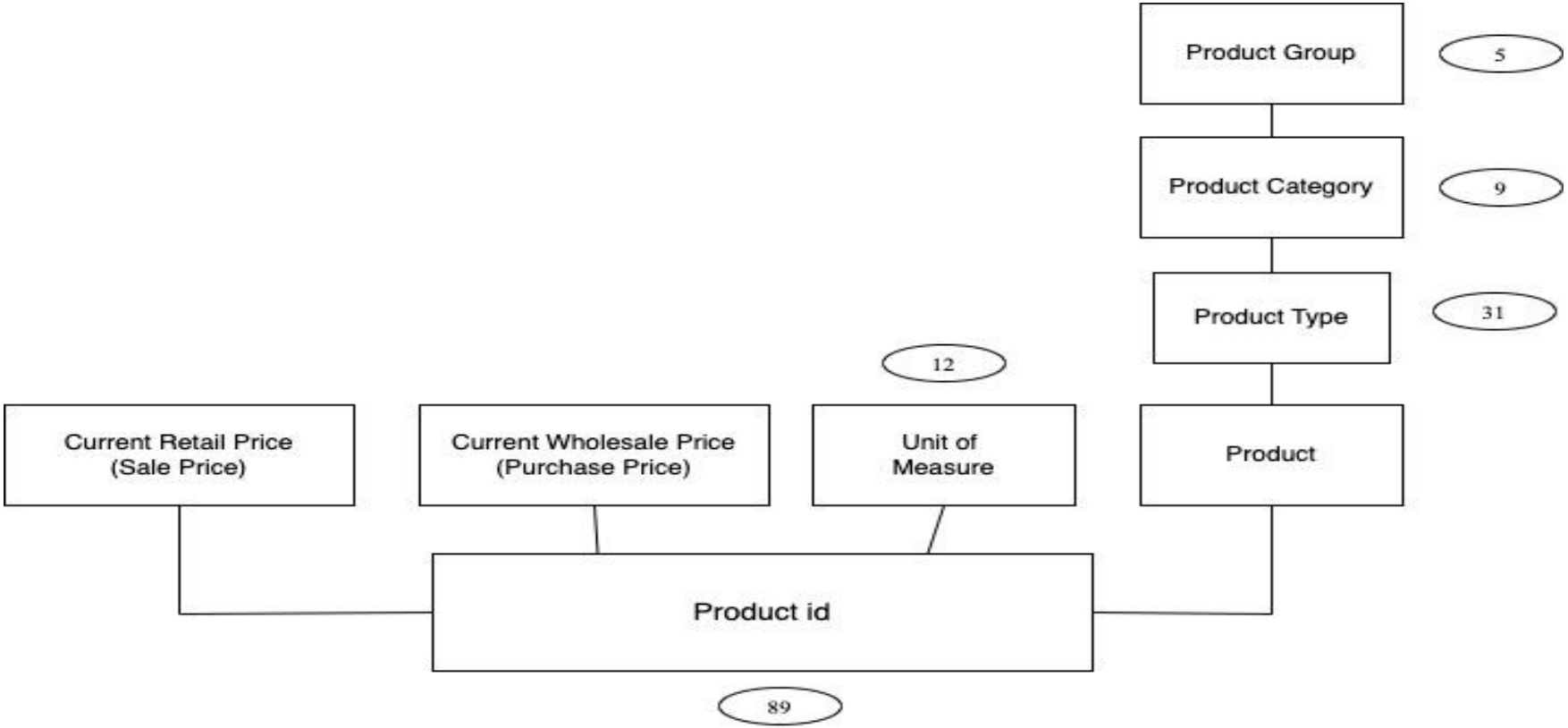
Star Schema


Payment Method Dimension
PK payment_method_id
payment_method_group
Customer Dimension
PK customer_id
home store
customer_name
customer_email
customer_since
birthdate
gender
birth_year
Staff Dimension
PK staff_id
first_name
last_name
position
start_date
location
Sales Outlet Dimension
PK sales_outlet_id
sales_outlet_type
store_square_feet
store_address
store_city
store_state_province
store_telephone
store_postal_code
store_longitude
store_latitude
manager-staff_id
neighborhood
Sales Transaction Fact
FK payment_method_id
FK bridge_product_id
FK customer_id
FK trans_date_id
FK ord_date_id
FK staff_id
FK sales_outlet_id
FK ship_id
transaction_time
quantity
line_item_amount
unit_price
Product Bridge Table
PK bridge_product_id
FK product_id
Num_levels_from_parent
bottom_flag
top_flag
Product Dimension
PK product_id
product_group
product_category
product_type
product
unit_of_measure
current_wholesale_price
current_retail_price
Trans Date Dimension
PK trans_date_id
FK date_id
trans_week_descr
trans_month_name
trans_quarter_name
trans_day_of_month
trans_day_of_week
Date Dimension
PK date_id
week_descr
month_name
quarter_name
day_of_month
day_of_week
year_id
Ord Date Dimension
PK ord_date_id
FK date_id
trans_week_descr
trans_month_name
trans_quarter_name
trans_day_of_month
trans_day_of_week
Shipping Dimension
PK ship_id
ship_carrier
ship_method

Dimension Attribute Description: Product Dimension

Name	Description	Cardinality	Slowly Changing Dimension Policy	Sample Values
Product id	Unique value used to identify each product.	89	No Update	1, 2, 3, 4, 5, 6, 7, 8, 9… 89.
Current Wholesale Price	Purchase cost for each item (per unit).	x	Type 2	x
Current Retail Price	Sale price for each item (per unit).	x	Type 2	x
Unit of Measurement	Unit of measure for each item sold.	12	Type 2	1 lb, 24 oz, 16 oz, 12 oz, 8 oz, 6 oz, 1 oz, etc.
Product Type	Type within a product category that a product is associated with.	31	Type 1	Organic Beans, Herbal tea, Drinking Chocolate, etc.
Product Category	Category within a product group that a product is associated with.	9	Type 1	Coffee beans, Loose Tea Coffee, Tea, Bakery, etc.
Product Group	A general group of products that a product is associated with.	5	Type 1	Whole Bean/Teas, Food, Add-ons, Beverages, Merchandise.

Dimension Table Diagram: Product Dimension



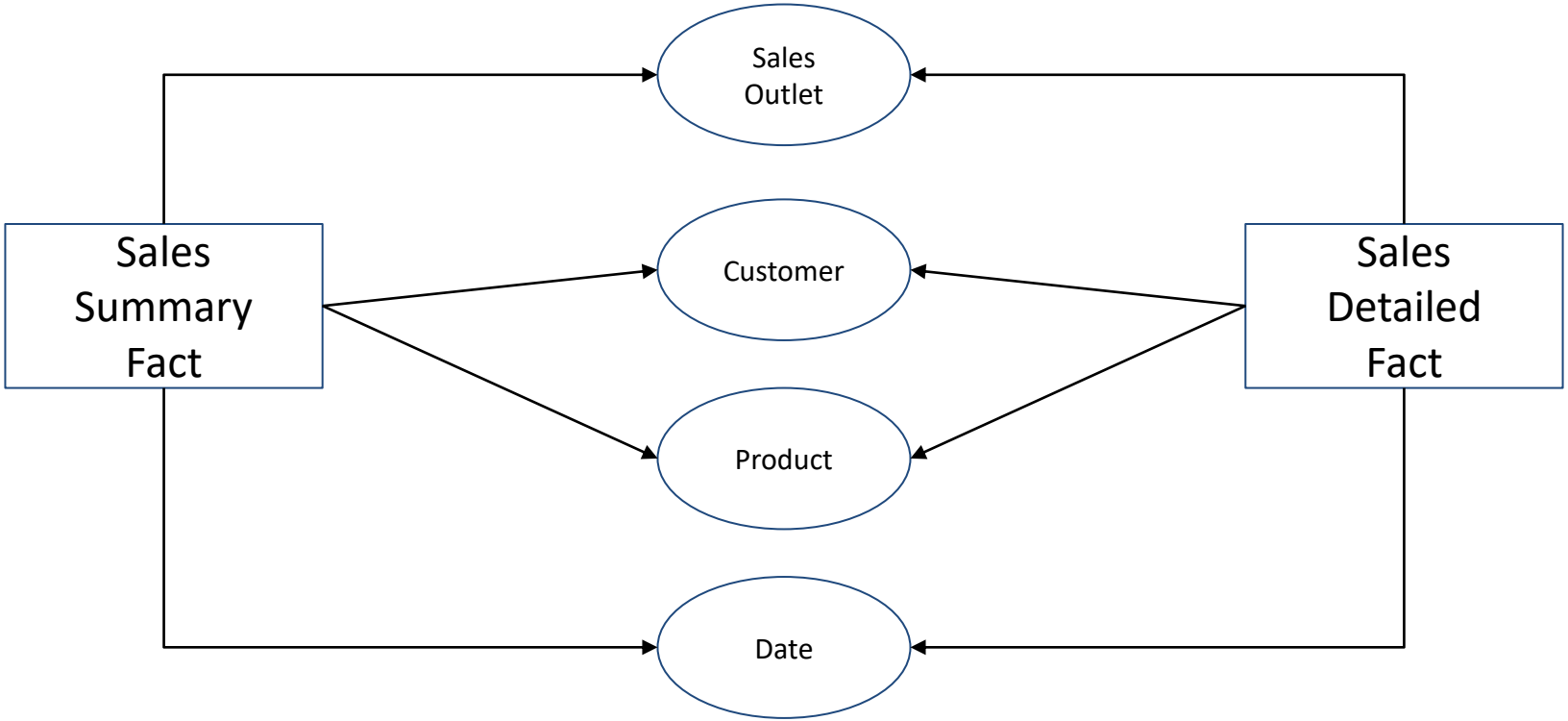
(3) Delivery Steps of Conformed Dimension:

1. **Standardization:** The purpose of standardization is to make the data coding methods and data formats of different data sources from different branches to the same, to lay a foundation for the next data matching.
2. **Matching and duplication:** Two aspects of data matching: one is to match the identification of different data sources and different attributes of the same thing to make the data more perfect; The other is to identify the same data from different data sources as duplicates, to lay a foundation for filtering in the next step
3. **Surviving:** Main purpose of data filtering is to select the conformed dimension as the master data, the final delivered conformed dimension data

(4) Conformed Dimensions in Uptown Brew's DW

1. Date Dimension
2. Customer Dimension
3. Product Dimension
4. Sales Outlet Dimension

Conformed Dimensions (cont'd.)

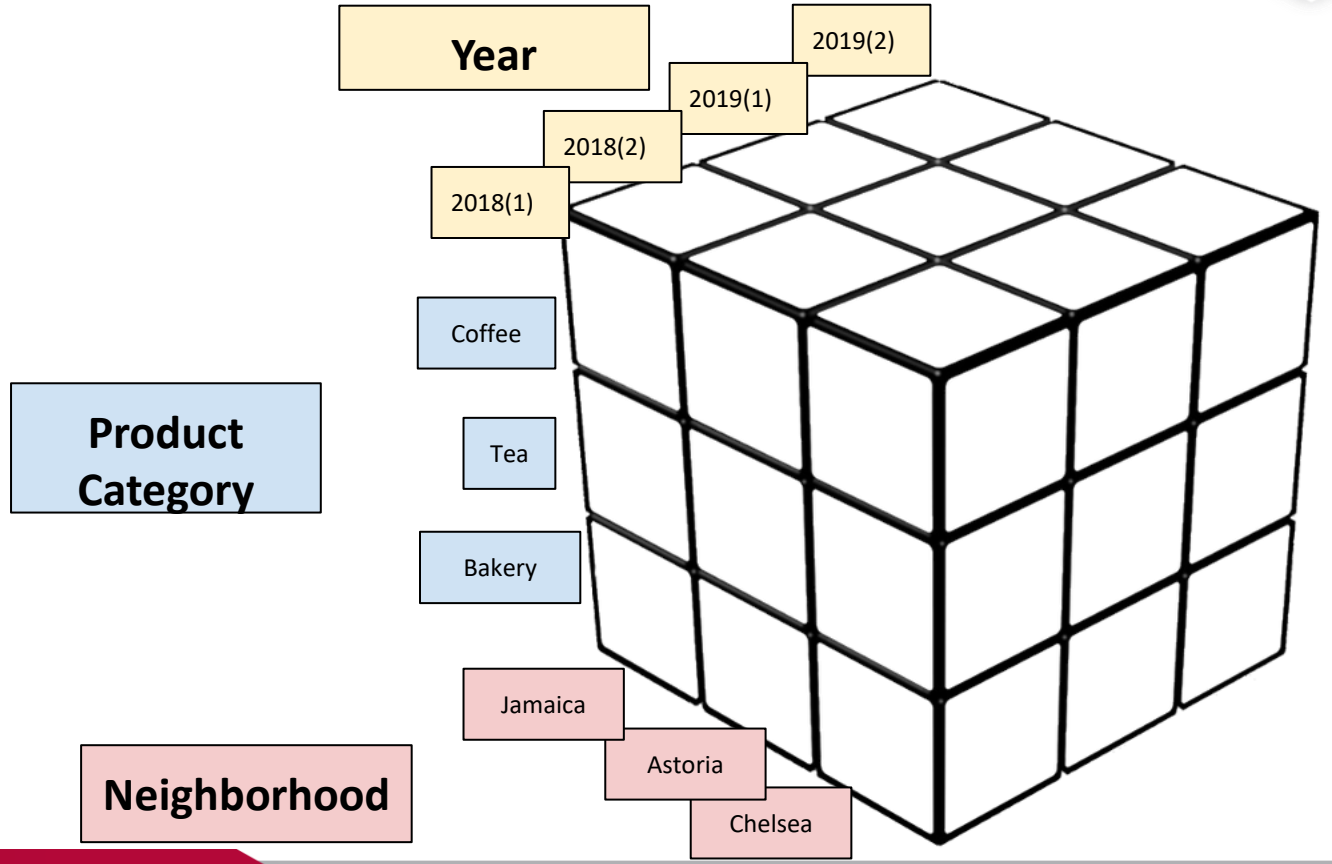


Two Way Aggregate Table

Product ID	Quantity	Product Category	Order Date	Transaction Date	Customer Name	Store Outlet
52	1	Tea	3/26/2019	4/1/2019	Melissa Johnson	3
47	1	Tea	3/24/2019	4/2/2019	Sean	5
74	1	Food	3/28/2019	4/8/2019	Teegan	5
83	1	Merchandise	3/30/2019	4/11/2019	Brady	8

OLAP MDDB Cube

Product Category,
Neighborhood,
and Year
are all Dense Dimensions



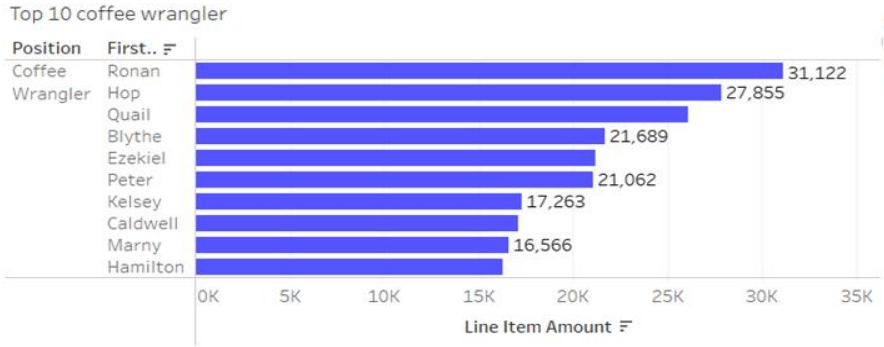
What is Business Intelligence?

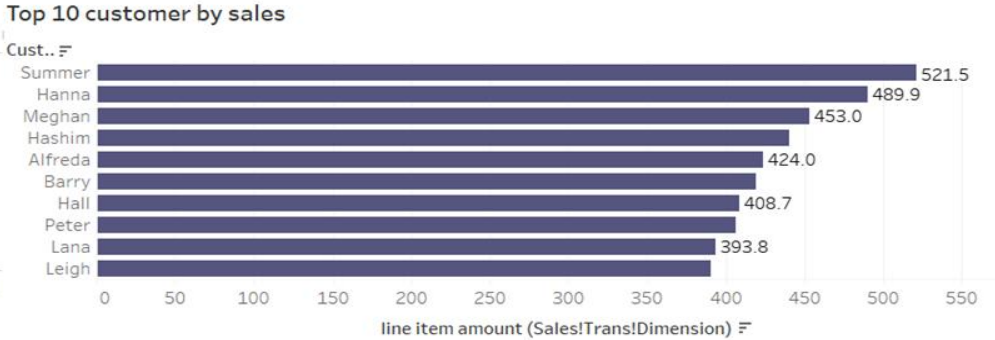
- Business intelligence (BI) is the process of transforming data into actionable insights that help an organization make strategic and tactical choices. To offer users with detailed insight about the state of the business, BI tools access and analyze data sets and show analytical findings in reports, summaries, dashboards, graphs, charts, and maps.

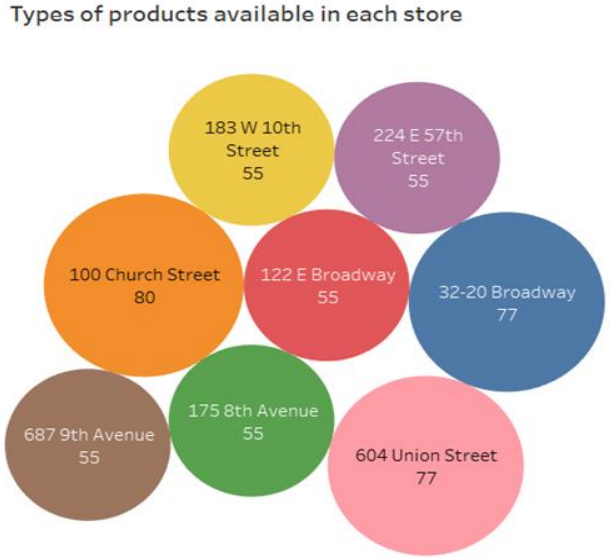
- The phrase "business intelligence" is also used to describe a set of technologies that enable quick, easy-to-understand access to data-driven insights about an organization's current status.

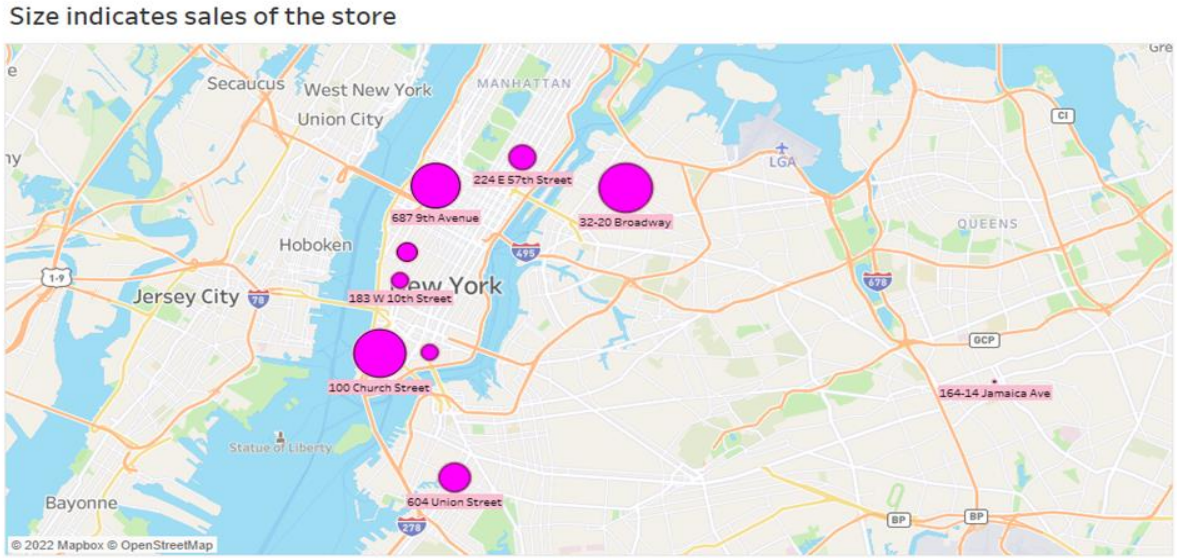
- For example, an organization that wants to better manage its supply chain needs BI capabilities to discover where delays are occurring and where variances exist within the shipping process. That organization might also utilize its business intelligence capabilities to figure out which products are the most frequently delayed, as well as which forms of transportation are the most frequently implicated in delays.

Executive Dashboard



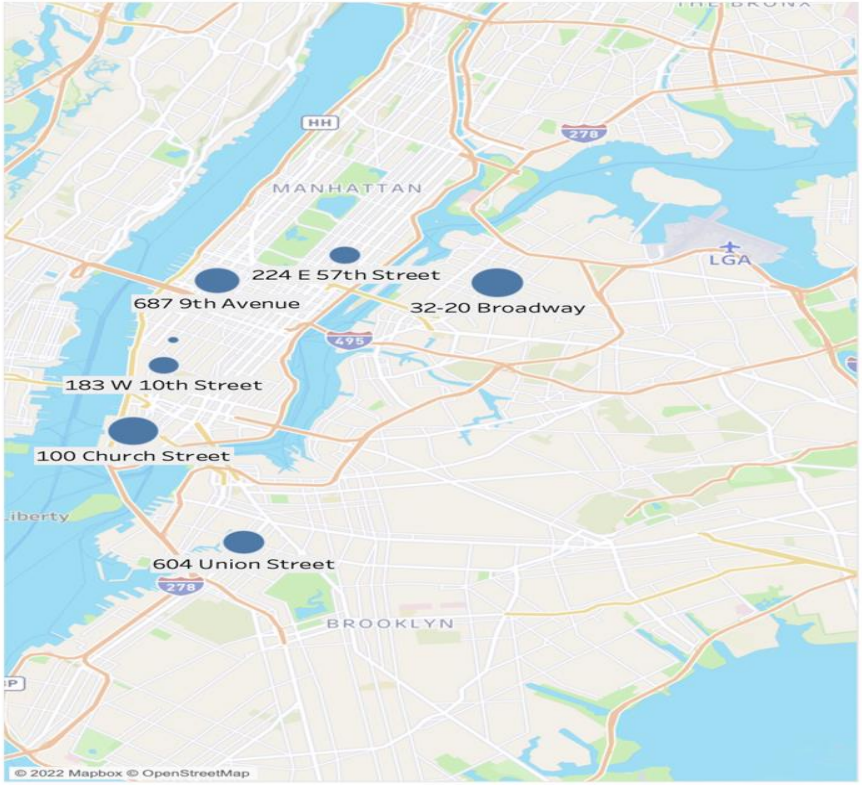




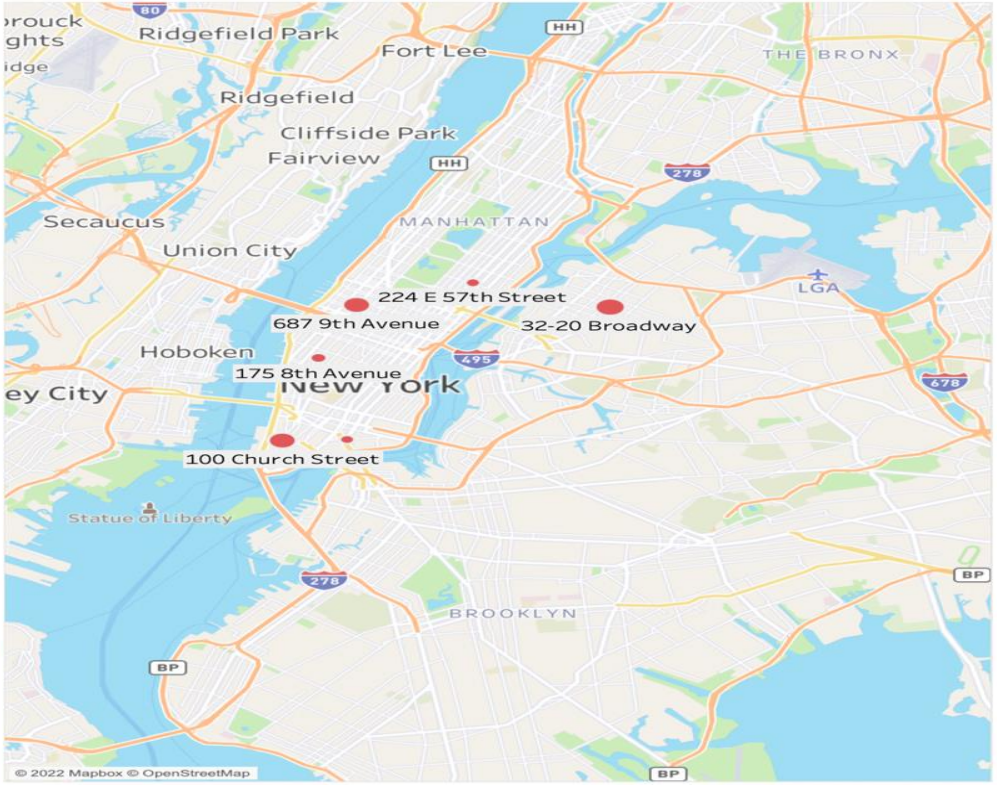


Sales Comparison Between Wednesdays and Sundays

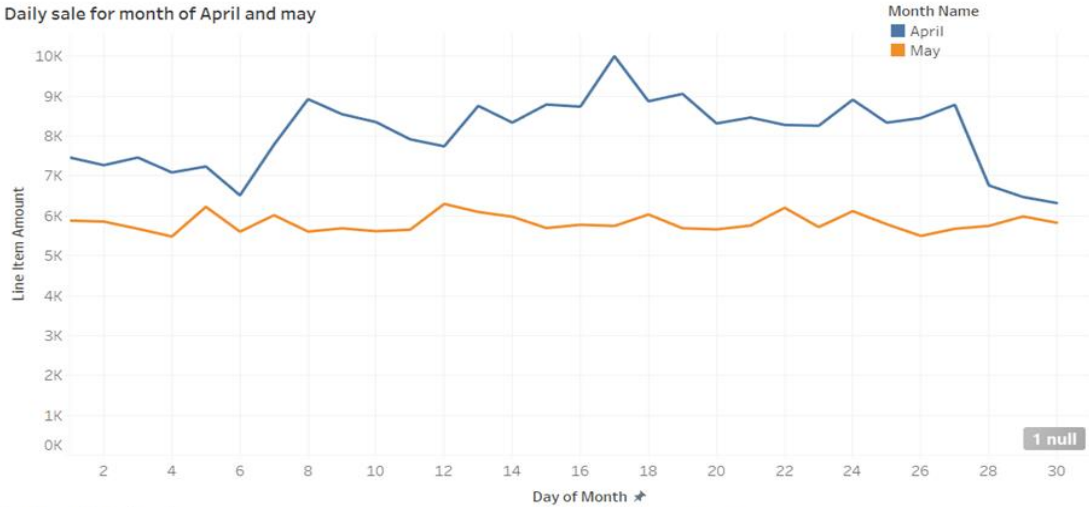
Size indicates sales of each store on wednesday

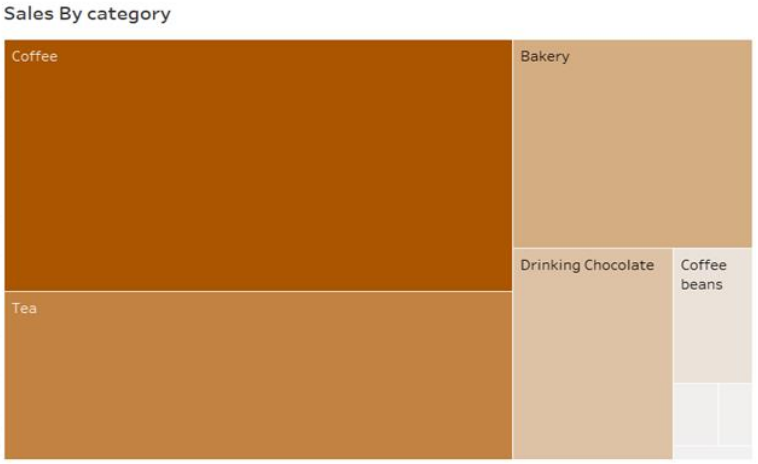


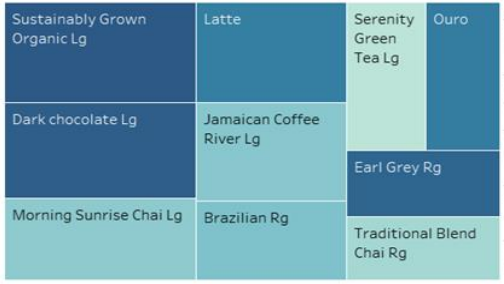
Size indicates sales of each store on sunday

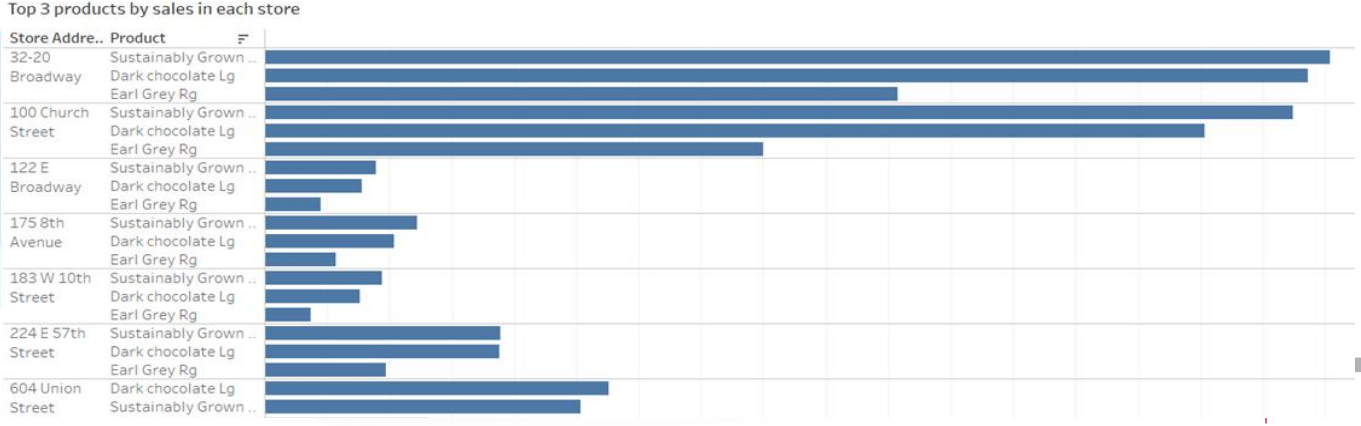


Sales Manager Dashboard

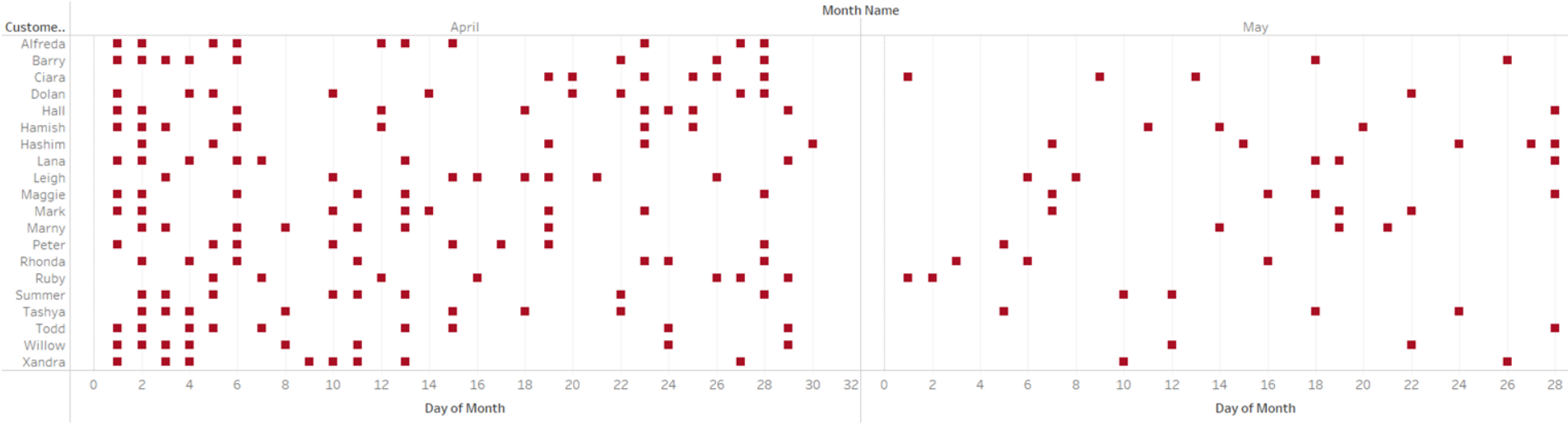




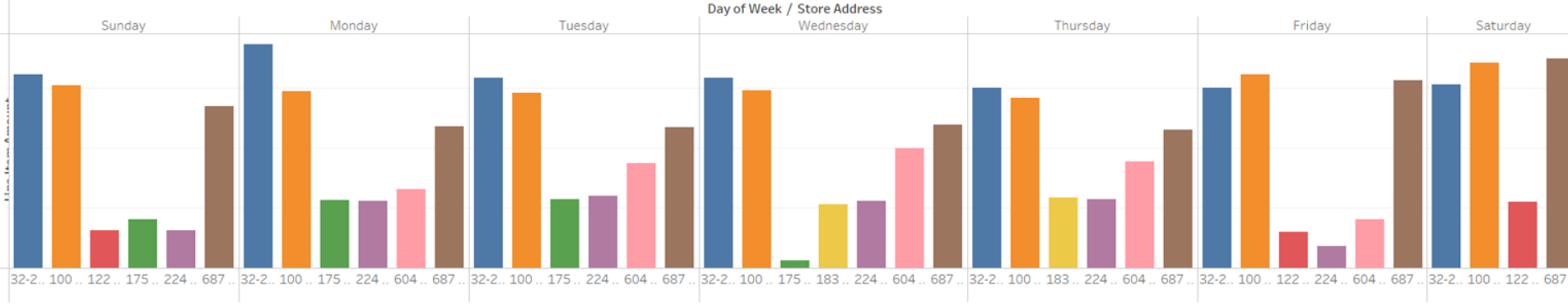




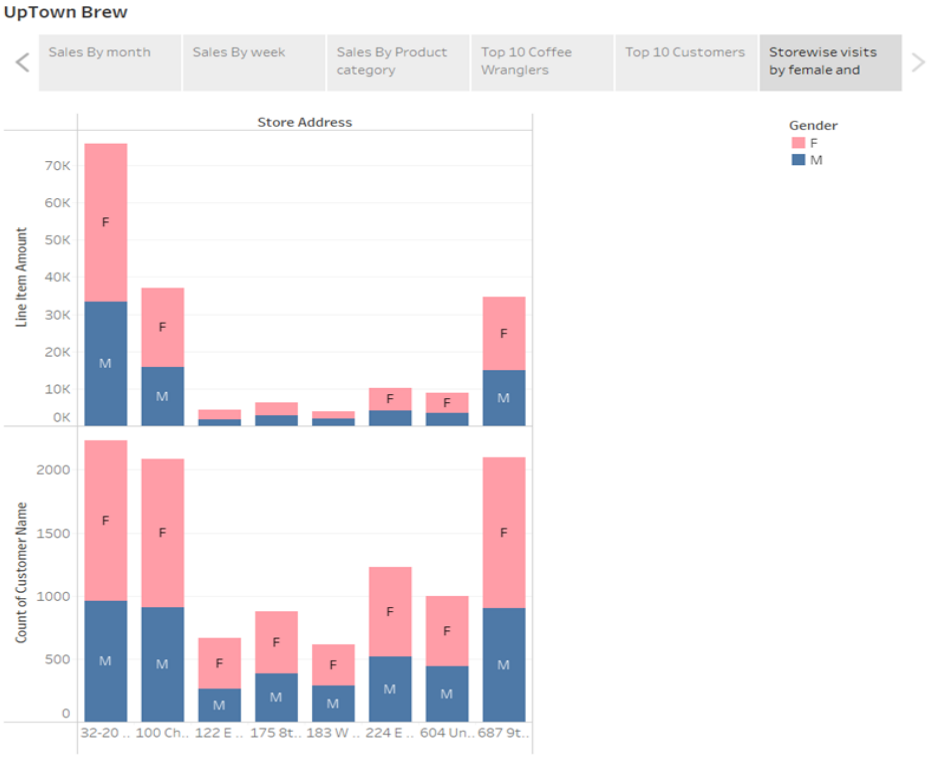
Visits by customers



Sales By week



Storyboard


UpTown Brew
Sales By month
Sales By week
Sales By Product category
Top 10 Coffee Wranglers
Top 10 Customers
Storewise visits by female and
Store Address
Gender
F
M
Line Item Amount
70K
60K
50K
40K
30K
20K
10K
0K
Count of Customer Name
2000
1500
1000
500
0
32-20 ..
100 Ch..
122 E ..
175 8t..
183 W ..
224 E ..
604 Un..
687 9t..

Q&A