



The Coffee Shop

PHÂN TÍCH DỮ LIỆU CỦA CHUỖI CỬA HÀNG CÀ PHÊ TẠI TIỂU BANG NEW YORK

Thảo Vy, Mỹ Kim, Như Quỳnh



MENU

I. Tổng quan

II. Quá trình ETL

III. Phân tích dữ liệu

IV. Đề xuất



1

TỔNG QUAN



Giới thiệu

Một chuỗi cửa hàng cà phê hư cấu đang vận hành tám cửa hàng tại tiểu bang New York. Chuỗi cửa hàng cà phê này được thành lập bởi Amber và Sandeep. Tuy nhiên, trong bảng dữ liệu thống kê giao dịch chỉ có thông tin từ ba cửa hàng được ghi nhận.



Tổng quan



ETL



Phân tích



Đề xuất





Thực trạng

Dữ liệu giao dịch từ 3 trong số 8 cửa hàng của chuỗi cà phê hư cấu tại New York vào tháng 4/2019 đã cho thấy kết quả tích cực, vượt qua mục tiêu về doanh thu và lợi nhuận. Tuy nhiên, sự suy giảm doanh thu ở cuối tháng đòi hỏi sự chú ý đặc biệt. Với sự mở rộng của chuỗi, việc tối ưu hóa hoạt động và tăng cường hiệu suất trở nên cần thiết, cùng với việc sử dụng dữ liệu từ các cửa hàng khác nhau để định hình chiến lược tương lai và đảm bảo thành công liên tục.



Tổng quan



ETL



Phân tích



Đề xuất





Bài toán

- Thiếu dữ liệu: Thiếu thông tin từ 5 cửa hàng và các tháng hoạt động trước đó gây khó khăn trong việc phân tích hiệu quả kinh doanh và đưa ra quyết định chiến lược.
- Cạnh tranh: Ngành cà phê tại New York cạnh tranh gay gắt, đòi hỏi thu hút và giữ chân khách hàng là một bài toán quan trọng.
- Tối ưu hóa hoạt động: Cần tối ưu hóa hiệu quả hoạt động của 3 cửa hàng hiện có để tăng lợi nhuận và giảm chi phí.
- Mở rộng thị trường: Mở rộng thị phần và tăng số lượng cửa hàng là mục tiêu quan trọng để phát triển bền vững.



Tổng quan



ETL



Phân tích



Đề xuất





Mục tiêu

Mục tiêu chung: chuỗi cửa hàng cà phê hư cấu này sẽ trở thành một thương hiệu cà phê uy tín và được yêu thích tại New York, với hệ thống cửa hàng rộng khắp, doanh thu cao và lợi nhuận bền vững.

Mục tiêu cụ thể:

- Tăng doanh thu.
- Mở rộng thị trường
- Nâng cao nhận diện thương hiệu
- Tăng tỷ lệ khách hàng trung thành
- Giảm chi phí vận hành



Tổng quan



ETL



Phân tích



Đề xuất



A cartoon illustration of a woman with dark, curly hair, wearing a yellow jacket over a white shirt. She is sitting on an orange couch, holding a brown paper coffee cup. The background is a light beige color with abstract brown and tan organic shapes.

2

Quá trình ETL

ETL

Nhóm thu thập dữ liệu gốc từ Kaggle. Dữ liệu gồm 9 file csv riêng lẻ, mỗi file chứa một bảng thông tin.

==> Vì dữ liệu nằm ở các file riêng lẻ khó, nhóm tiến hành tổng hợp các file riêng đó vào 1 file chung 9 sheet, mỗi sheet chứa dữ liệu của 1 file. Nhóm tiến hành tiền xử lý trên file vừa tổng hợp được.



Tổng quan



ETL



Phân tích



Đề xuất



ETL

A	B	C	D	E	F	G	H	I	J
1	transaction_date	Date_ID	Week_ID	Week_Des	Month_ID	Month_Na	Quarter_ID	Quarter_N	Year_ID
2	04/01/2019	20190401	14	Week 14	4	April	2	Q2	2019
3	04/02/2019	20190402	14	Week 14	4	April	2	Q2	2019
4	04/03/2019	20190403	14	Week 14	4	April	2	Q2	2019
5	04/04/2019	20190404	14	Week 14	4	April	2	Q2	2019
6	04/05/2019	20190405	14	Week 14	4	April	2	Q2	2019
7	04/06/2019	20190406	14	Week 14	4	April	2	Q2	2019
8	04/07/2019	20190407	14	Week 14	4	April	2	Q2	2019
9	04/08/2019	20190408	15	Week 15	4	April	2	Q2	2019
10	04/09/2019	20190409	15	Week 15	4	April	2	Q2	2019
11	04/10/2019	20190410	15	Week 15	4	April	2	Q2	2019
12	04/11/2019	20190411	15	Week 15	4	April	2	Q2	2019
13	04/12/2019	20190412	15	Week 15	4	April	2	Q2	2019
14	4/13/2019	20190413	15	Week 15	4	April	2	Q2	2019
15	4/14/2019	20190414	15	Week 15	4	April	2	Q2	2019
16	4/15/2019	20190415	16	Week 16	4	April	2	Q2	2019
17	4/16/2019	20190416	16	Week 16	4	April	2	Q2	2019
18	4/17/2019	20190417	16	Week 16	4	April	2	Q2	2019
19	4/18/2019	20190418	16	Week 16	4	April	2	Q2	2019

Đầu tiên, nhóm chuyển kiểu dữ liệu của tất cả các cột id về dạng ký tự (text).

Sau đó chỉnh format cột transaction_date (sheet sales, pastry inventory, date)



Tổng quan



ETL



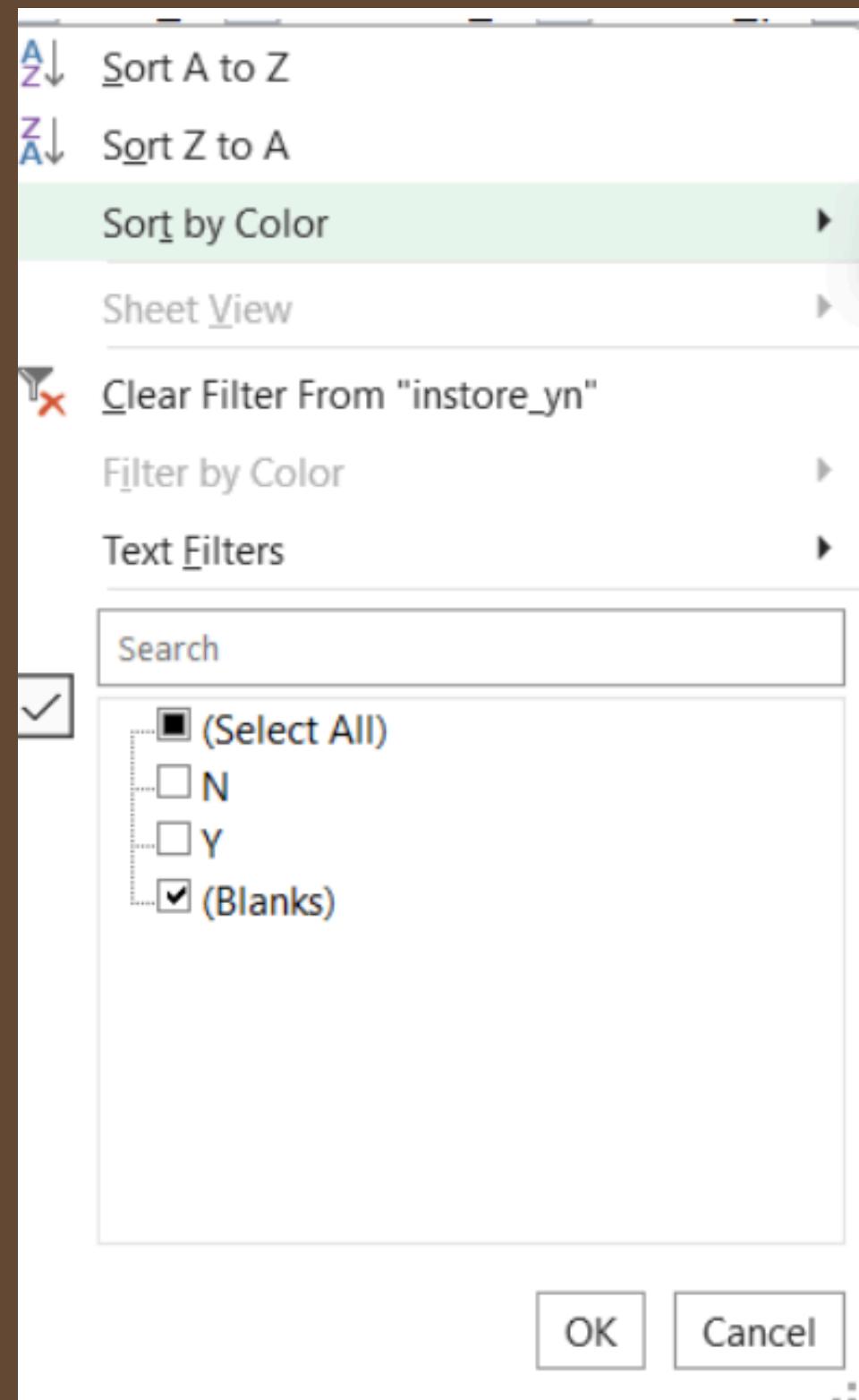
Phân tích



Đề xuất



ETL



cột `instore_yn` có dữ liệu bị thiếu ở dạng khoảng trắng không chứa giá trị(blank), tiến hành drop 295 rows này.



Tổng quan



ETL



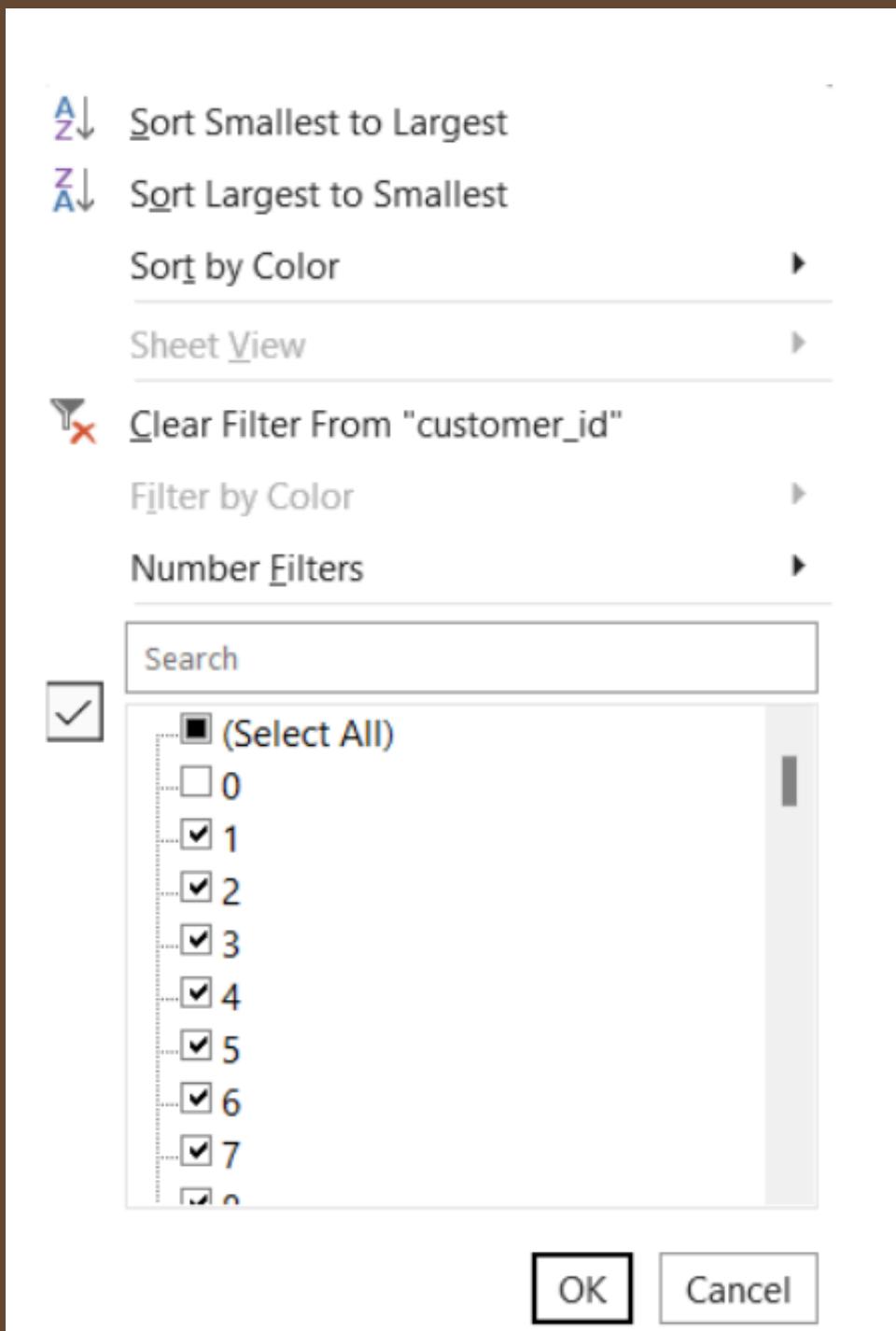
Phân tích



Đề xuất



ETL



Drop customer_id = 0



Tổng quan



ETL



Phân tích



Đề xuất



ETL

	D	E	F	G	H	I	J	K	L	M
1	product_type	product	product_description	unit_of_measure	current_wholesale_price	current_retail_price	current_retail_price	tax_exempt_yn	promo_yn	new_product_yn
2	Organic Beans	Brazilian - Organ	It's like Carnival in a cup.	12 oz	14.4	\$18.00	18	Y	N	N
3	House blend	Bear Our Old Time Dir	Out packed blend of bear	12 oz	14.4	\$18.00	18	Y	N	N
4	Espresso Beans	Espresso Roast	Our house blend for a go-	1 lb	11.8	\$14.75	14.75	Y	N	N
5	Espresso Beans	Primo Espresso F	Our premium single sourc	1 lb	16.36	\$20.45	20.45	Y	N	N
6	Gourmet Beans	Columbian Medi	A smooth cup of coffee	1 lb	12	\$15.00	15	Y	N	N
7	Gourmet Beans	Ethiopia	From the home of coffee	1 lb	16.8	\$21.00	21	Y	N	N
8	Premium Beans	Jamaican Coffee Ya man,	it will start your	1 lb	15.8	\$19.75	19.75	Y	N	N
9	Premium Beans	Civet Cat	The most expensive coffe	.5 lb	36	\$45.00	45	Y	N	N
10	Organic Beans	Organic Decaf Bl	Our blend of hand picked	1 lb	18	\$22.50	22.5	Y	N	N
11	Green beans	Guatemalan Sust	Green beans you can roa	1 lb	8	\$10.00	10	Y	N	N
12	Herbal tea	Lemon Grass	You will think you are Th	.9 oz	7.16	\$8.95	8.95	Y	N	N
13	Herbal tea	Peppermint	Cool and refreshing to he	.9 oz	7.16	\$8.95	8.95	Y	N	N
14	Black tea	English Breakfast	The traditional cup to sta	.9 oz	7.16	\$8.95	8.95	Y	N	N
15	Black tea	Earl Grey	A full leaf of Orange Pek	.9 oz	7.16	\$8.95	8.95	Y	N	N
16	Green tea	Serenity Green T	Mountain grown and har	.1 oz	7.4	\$9.25	9.25	Y	N	N
17	Chai tea	Traditional Blend	A traditional blend.	.9 oz	7.16	\$8.95	8.95	Y	N	N
18	Chai tea	Morning Sunrise	Fair trade and organic an	.9 oz	7.6	\$9.50	9.5	Y	N	N
19	Chai tea	Spicy Eye Opener	A spicier blend to awak	.9 oz	8.76	\$10.95	10.95	Y	N	N
20	Drinking Chocolate	Dark chocolate	This drinking chocolate is	1 lb	5.12	\$6.40	6.4	Y	N	N
21	Organic Chocolate	Sustainably Grov	Certified organic contain	1 lb	6.08	\$7.60	7.6	Y	N	N
22	Drinking Chocolate	Chili Mayan	Fragrant with spices, this	1 lb	10.66	\$13.33	13.33	Y	N	N
23	Drip coffee	Our Old Time Dir	An honest cup a coffee.	8 oz	0.4	\$2.00	2	Y	N	N
24	Drip coffee	Our Old Time Dir	An honest cup a coffee.	16 oz	0.5	\$2.50	2.5	Y	N	N
25	Drip coffee	Our Old Time Dir	An honest cup a coffee.	24 oz	0.6	\$3.00	3	Y	N	N
26	Organic brewed c	Brazilian Sm	It's like Carnival in a cup.	8 oz	0.44	\$2.20	2.2	Y	N	N
27	Organic brewed c	Brazilian Rg	It's like Carnival in a cup.	16 oz	0.6	\$3.00	3	Y	N	N
28	Organic brewed c	Brazilian Lg	It's like Carnival in a cup.	24 oz	0.7	\$3.50	3.5	Y	N	N
29	Gourmet brewed	Columbian Medi	A smooth cup of coffee	.8 oz	0.4	\$2.00	2	Y	N	N

cột current_retail_price đang có ký tự \$ nên nó ở định dạng text, nhóm tiến hành xóa ký tự \$ và đưa về dạng số để đảm bảo khả năng tính toán cho cột dữ liệu.



Tổng quan



ETL



Phân tích



Đề xuất



ETL

	D	E	F	G	H	I	J	K	L	M
1	product_type	product	product_description	unit_of_measure	current_wholesale_price	current_retail_price	current_retail_price	tax_exempt_yn	promo_yn	new_product_yn
2	Organic Beans	Brazilian - Organ	It's like Carnival in a cup.	12 oz	14.4	\$18.00	18	Y	N	N
3	House blend	Bear Our Old Time Dir	Out packed blend of bear	12 oz	14.4	\$18.00	18	Y	N	N
4	Espresso Beans	Espresso Roast	Our house blend for a go-	1 lb	11.8	\$14.75	14.75	Y	N	N
5	Espresso Beans	Primo Espresso F	Our premium single sourc	1 lb	16.36	\$20.45	20.45	Y	N	N
6	Gourmet Beans	Columbian Medi	A smooth cup of coffee	1 lb	12	\$15.00	15	Y	N	N
7	Gourmet Beans	Ethiopia	From the home of coffee	1 lb	16.8	\$21.00	21	Y	N	N
8	Premium Beans	Jamaican Coffee Ya man,	it will start your	1 lb	15.8	\$19.75	19.75	Y	N	N
9	Premium Beans	Civet Cat	The most expensive coffe	.5 lb	36	\$45.00	45	Y	N	N
10	Organic Beans	Organic Decaf Bl	Our blend of hand picked	1 lb	18	\$22.50	22.5	Y	N	N
11	Green beans	Guatemalan Sust	Green beans you can roa	1 lb	8	\$10.00	10	Y	N	N
12	Herbal tea	Lemon Grass	You will think you are Th	.9 oz	7.16	\$8.95	8.95	Y	N	N
13	Herbal tea	Peppermint	Cool and refreshing to he	.9 oz	7.16	\$8.95	8.95	Y	N	N
14	Black tea	English Breakfast	The traditional cup to sta	.9 oz	7.16	\$8.95	8.95	Y	N	N
15	Black tea	Earl Grey	A full leaf of Orange Pek	.9 oz	7.16	\$8.95	8.95	Y	N	N
16	Green tea	Serenity Green T	Mountain grown and har	.1 oz	7.4	\$9.25	9.25	Y	N	N
17	Chai tea	Traditional Blend	A traditional blend.	.9 oz	7.16	\$8.95	8.95	Y	N	N
18	Chai tea	Morning Sunrise	Fair trade and organic an	.9 oz	7.6	\$9.50	9.5	Y	N	N
19	Chai tea	Spicy Eye Opener	A spicier blend to awak	.9 oz	8.76	\$10.95	10.95	Y	N	N
20	Drinking Chocolate	Dark chocolate	This drinking chocolate is	1 lb	5.12	\$6.40	6.4	Y	N	N
21	Organic Chocolate	Sustainably Grov	Certified organic contain	1 lb	6.08	\$7.60	7.6	Y	N	N
22	Drinking Chocolate	Chili Mayan	Fragrant with spices, this	1 lb	10.66	\$13.33	13.33	Y	N	N
23	Drip coffee	Our Old Time Dir	An honest cup a coffee.	8 oz	0.4	\$2.00	2	Y	N	N
24	Drip coffee	Our Old Time Dir	An honest cup a coffee.	16 oz	0.5	\$2.50	2.5	Y	N	N
25	Drip coffee	Our Old Time Dir	An honest cup a coffee.	24 oz	0.6	\$3.00	3	Y	N	N
26	Organic brewed c	Brazilian Sm	It's like Carnival in a cup.	8 oz	0.44	\$2.20	2.2	Y	N	N
27	Organic brewed c	Brazilian Rg	It's like Carnival in a cup.	16 oz	0.6	\$3.00	3	Y	N	N
28	Organic brewed c	Brazilian Lg	It's like Carnival in a cup.	24 oz	0.7	\$3.50	3.5	Y	N	N
29	Gourmet brewed	Columbian Medi	A smooth cup of coffee	.8 oz	0.4	\$2.00	2	Y	N	N

cột current_retail_price đang có ký tự \$ nên nó ở định dạng text, nhóm tiến hành xóa ký tự \$ và đưa về dạng số để đảm bảo khả năng tính toán cho cột dữ liệu.



Tổng quan



ETL



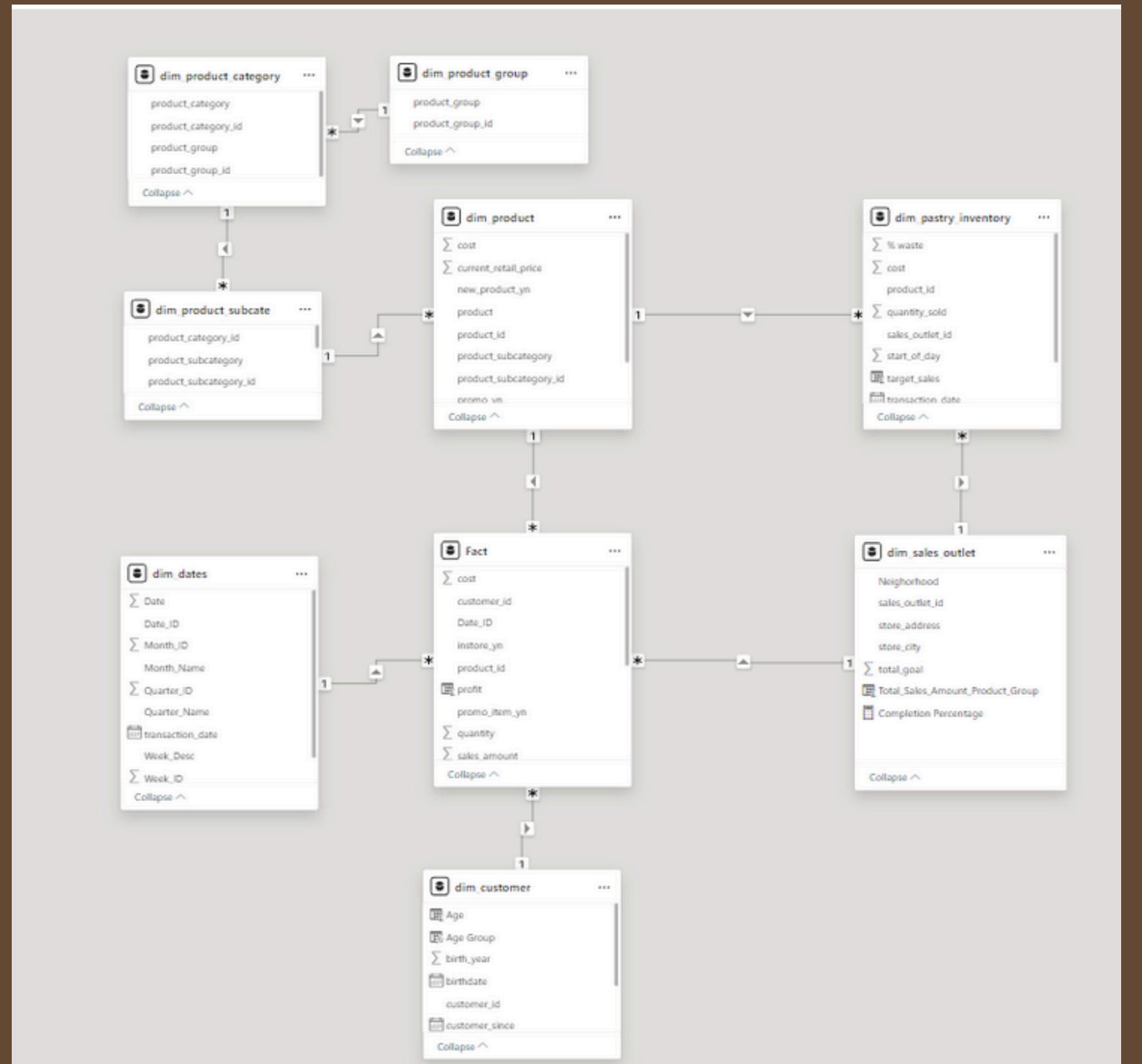
Phân tích



Đề xuất



Dimensional Modeling



Tổng quan



ETL



Phân tích



Đề xuất



3

Phân tích dữ liệu





66K

Total Target for Outlets in Fac

114.12K

Sum of sales_amount

6.58

Average of current_retail_price

36K

Sum of quantity

72.44K

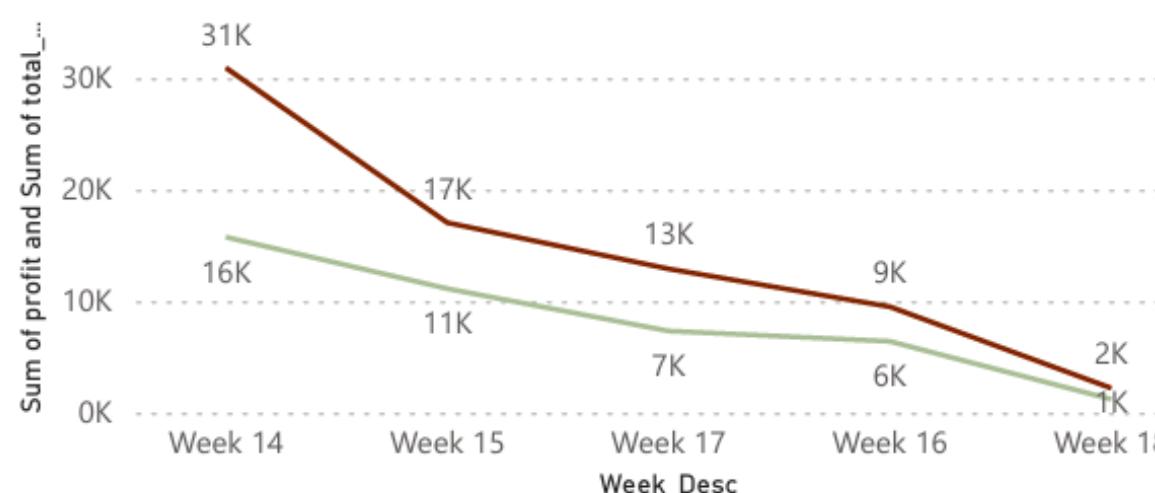
Sum of profit

41.68K

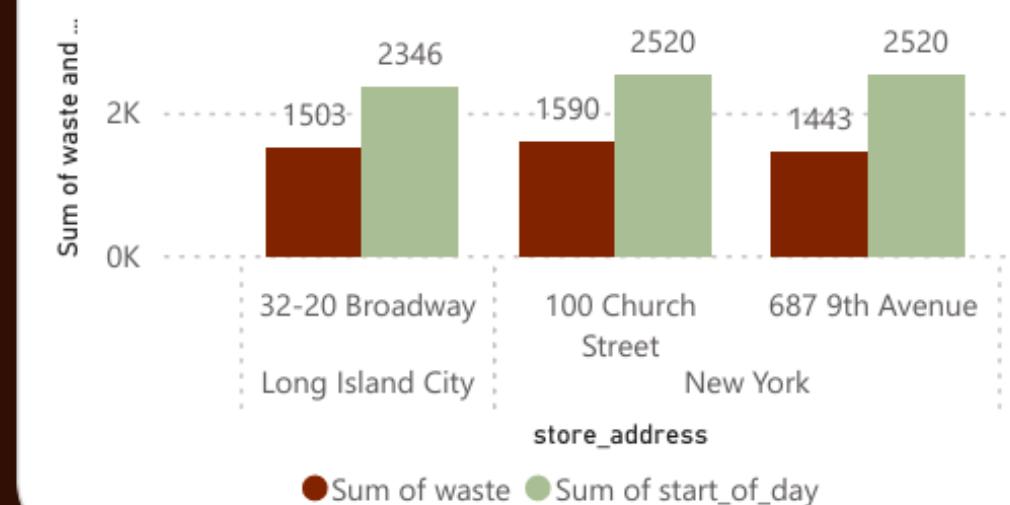
Sum of total_cost

Sum of profit and Sum of total_cost by Week_Desc

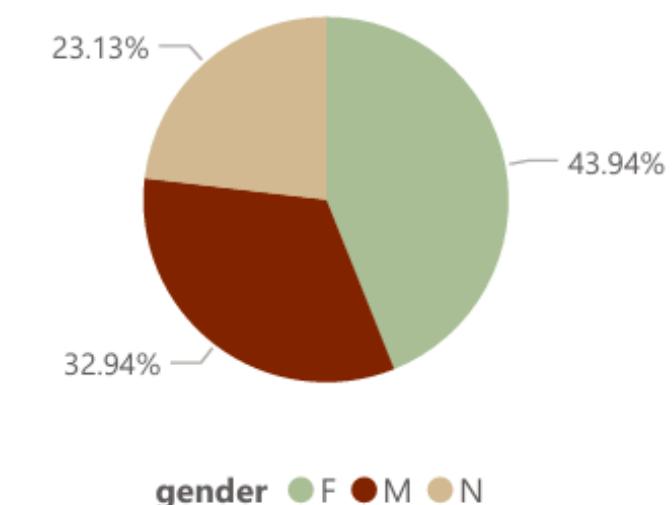
● Sum of profit ● Sum of total_cost



Sum of waste and Sum of start_of_day by store_city and store_address

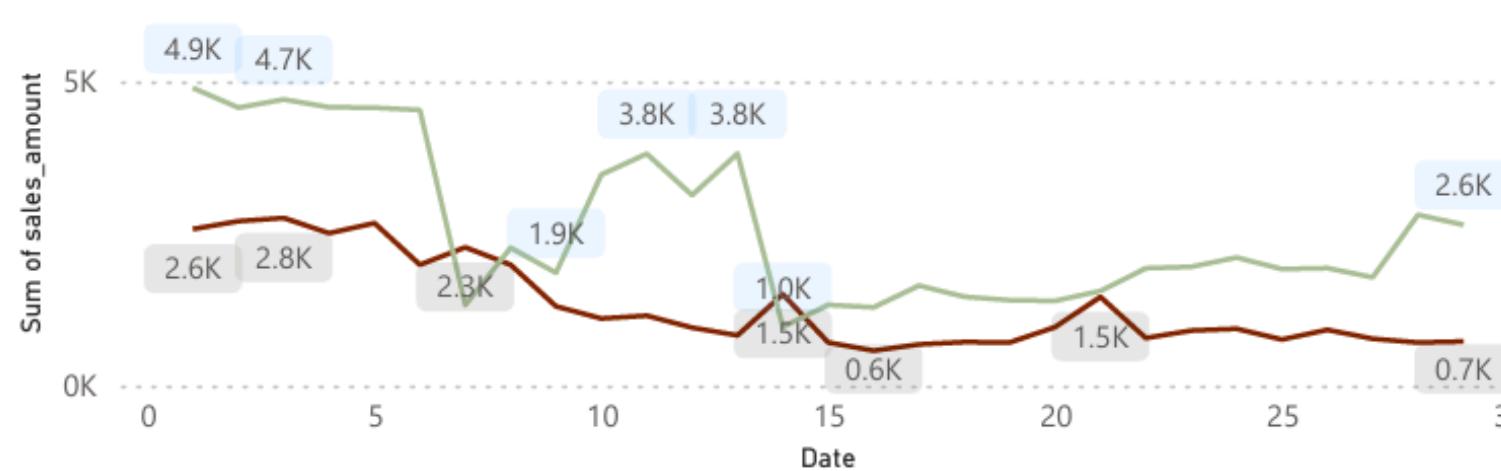


Sum of sales_amount by gender



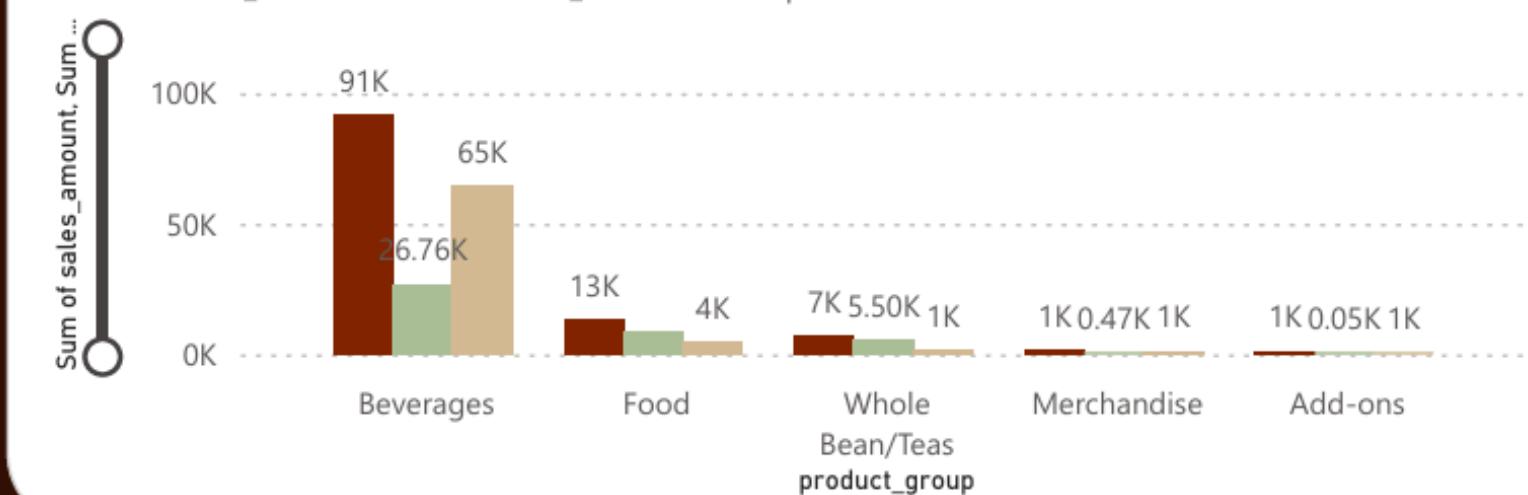
Sum of sales_amount by Date and store_city

store_city ● Long Island City ● New York



Sum of sales_amount, Sum of total_cost and Sum of profit by product_group

● Sum of sales_amount ● Sum of total_cost ● Sum of profit





THE
COFFEE
SHOP

Analysis by Time



360

Max of sales_amount

0.04

Min of cost

72.44K

Sum of profit

41.68K

Sum of total_cost

product_group,
product_category,
product_subcategory

All

Week_ID, Date

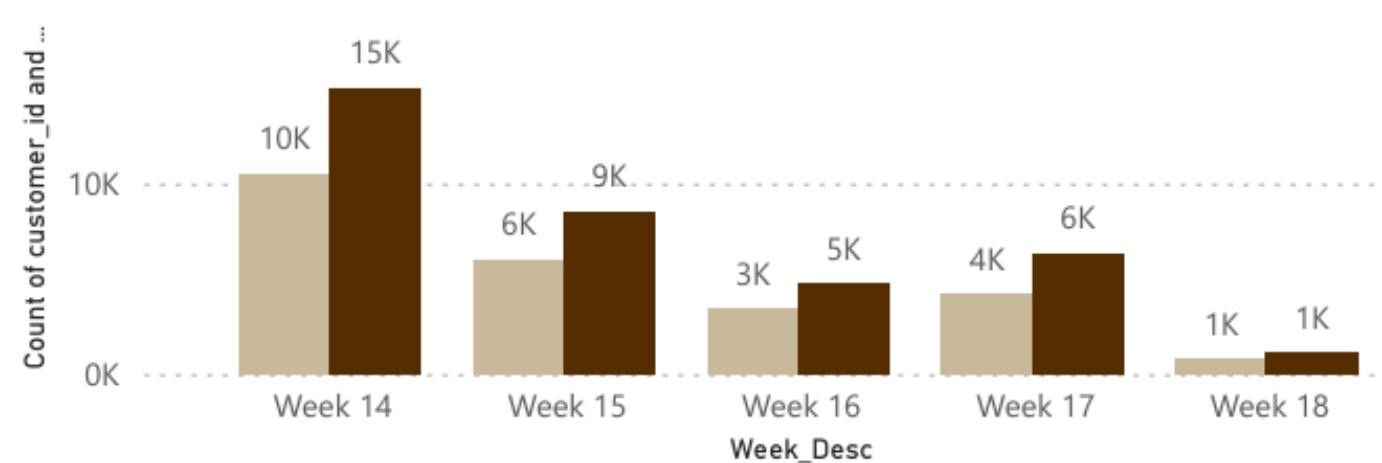
All

Age Group

All

Count of customer_id and Sum of quantity by Week_Desc

● Count of customer_id ● Sum of quantity



Sum of quantity by transaction_time (bins)

3601 3598

2477

2498

83

804

7.5K

7.2K

1.8K

1.9K

1.8K

3.6K

3.2K

1.4K

4.6K

1.9K

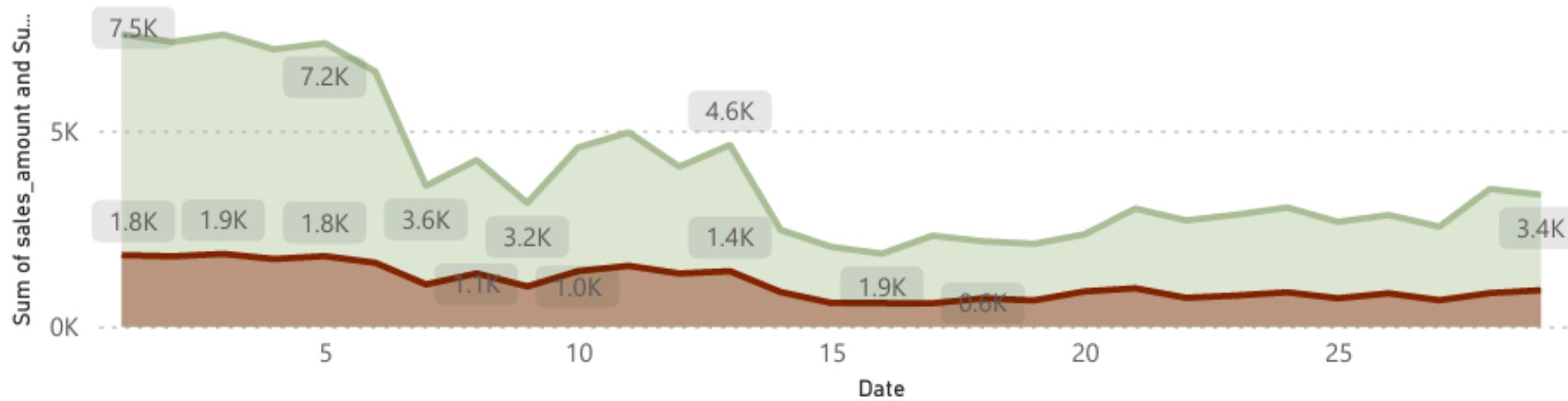
1.0K

0.6K

3.4K

Sum of sales_amount and Sum of cost by Date

● Sum of sales_amount ● Sum of cost



transaction_time (bins)



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Product



45.00

Max of current_retail_price

0.80

Min of current_retail_price

All Products

SelectedProduct

(Blank)

ProductGroupOfProduct

80

Count of product_id

product_group, product...

All

Week_ID, Date

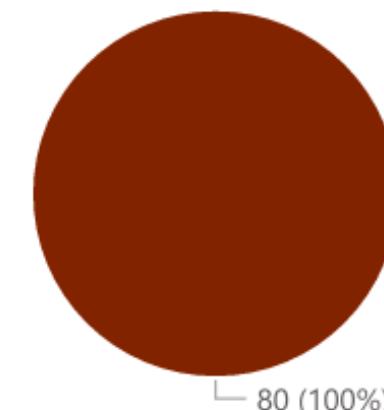
All

Age Group

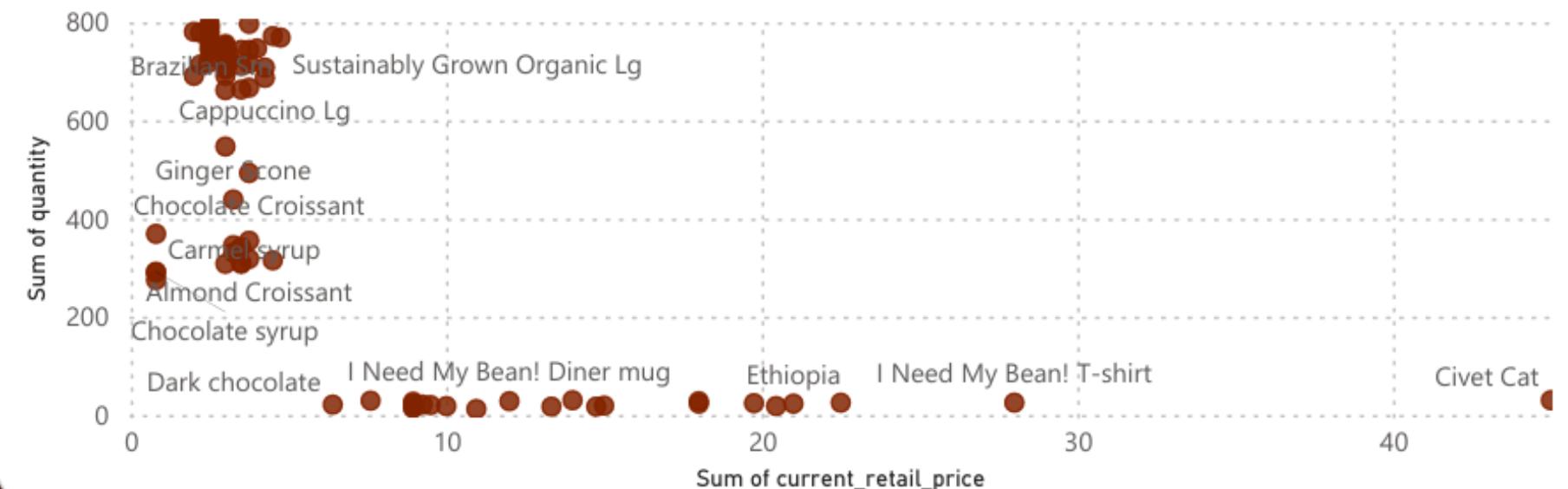
All

Count of product_id by promo_yn

new_product_yn ● N

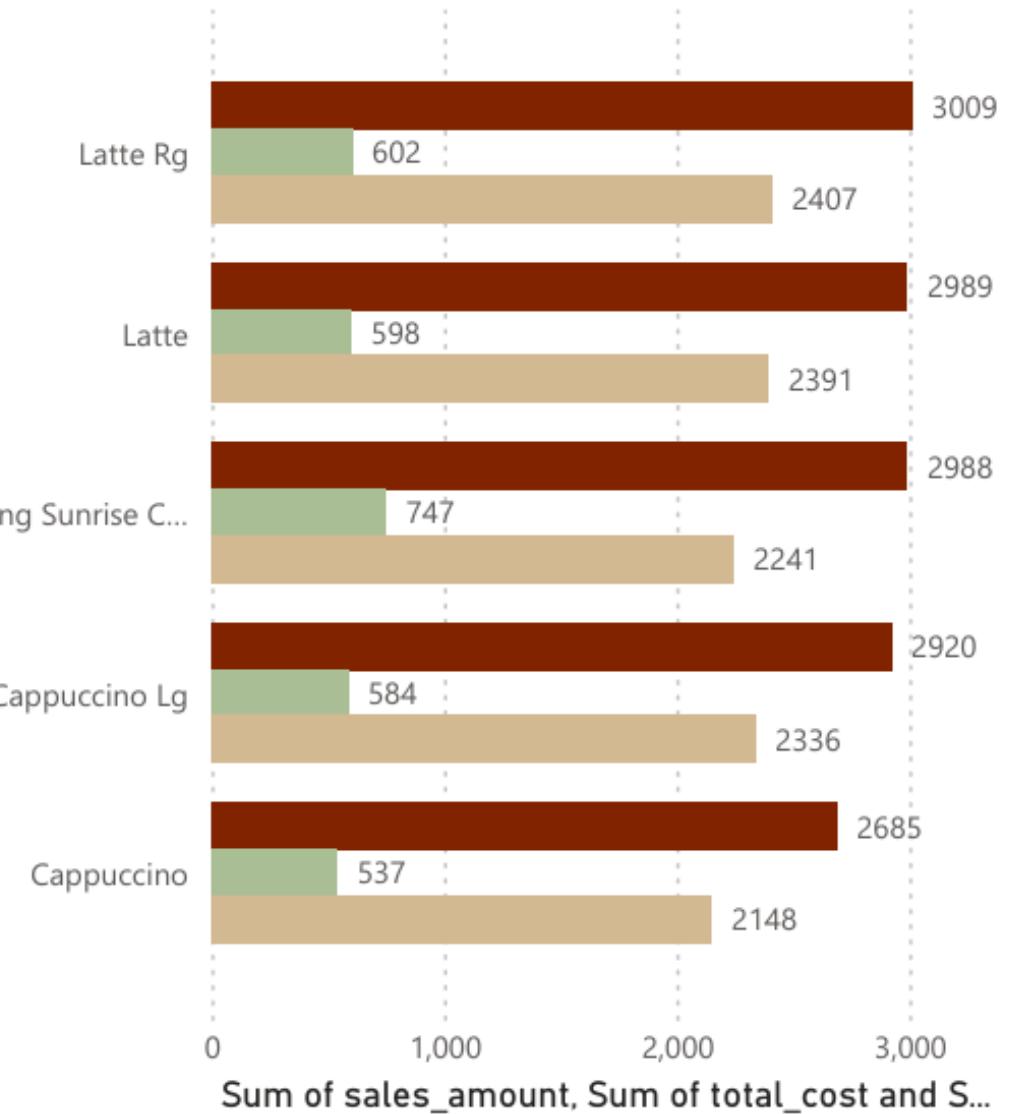


Sum of current_retail_price and Sum of quantity by product



Sum of sales_amount, Sum of total_cost and Sum of profit by product

● Sum of sales_amount ● Sum of total_cost ● Sum of profit



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Product Group



Average current_retail_price of Add-ons

0.80

Average current_retail_price of Beverages

3.21

Average current_retail_price of Food

3.47

Average current_retail_price of Merchandise

19.25

Average current_retail_price of Whole Bean/Teas

14.58

product_group, product_group

All

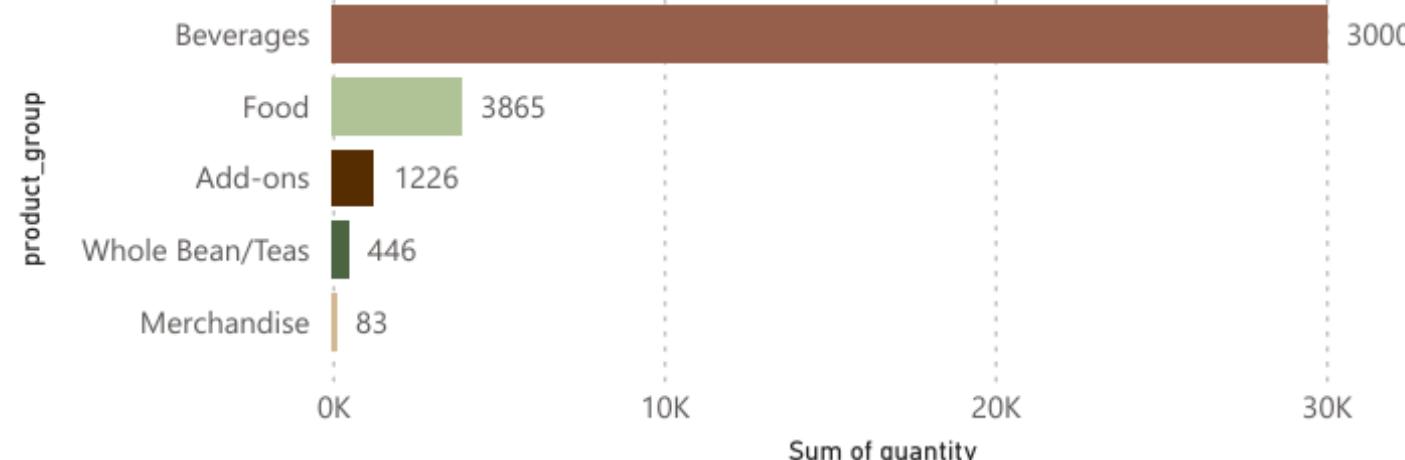
Week_ID, Date

All

Age Group

All

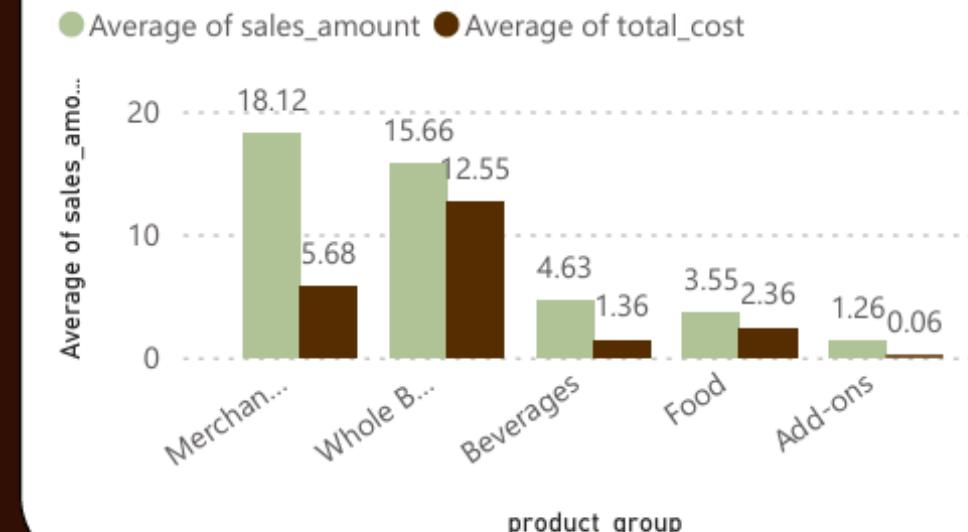
Sum of quantity by product_group



Count of product by product_group



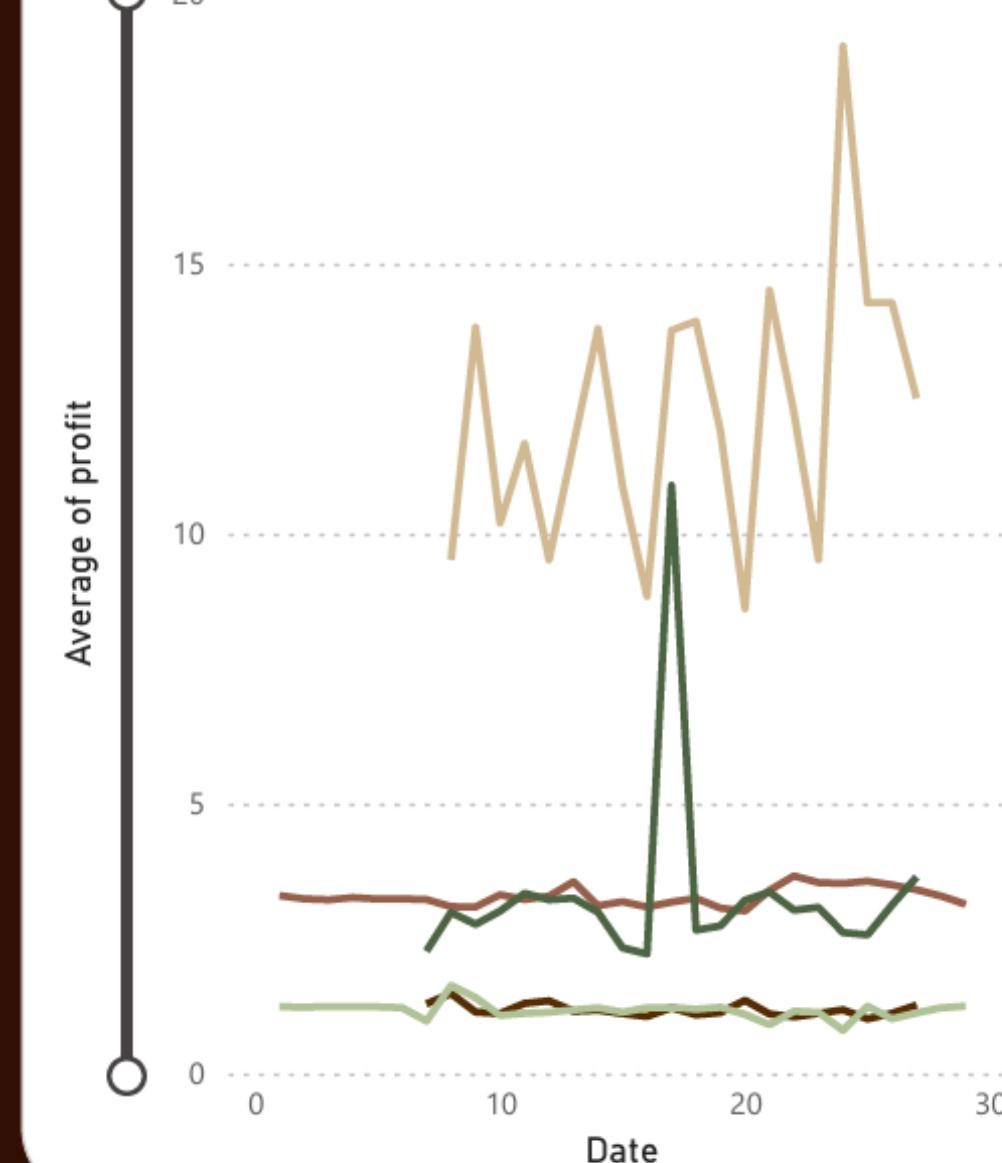
Average of sales_amount and Average of total_cost by product_group



Average of profit by Date and product_group

product_group

● Add-ons ● Beverages ● Food



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Product Category



1

Min of customer_id

73

Max of product_id

4203

Max of transaction_id

product_group, product...

All

Week_ID, Date

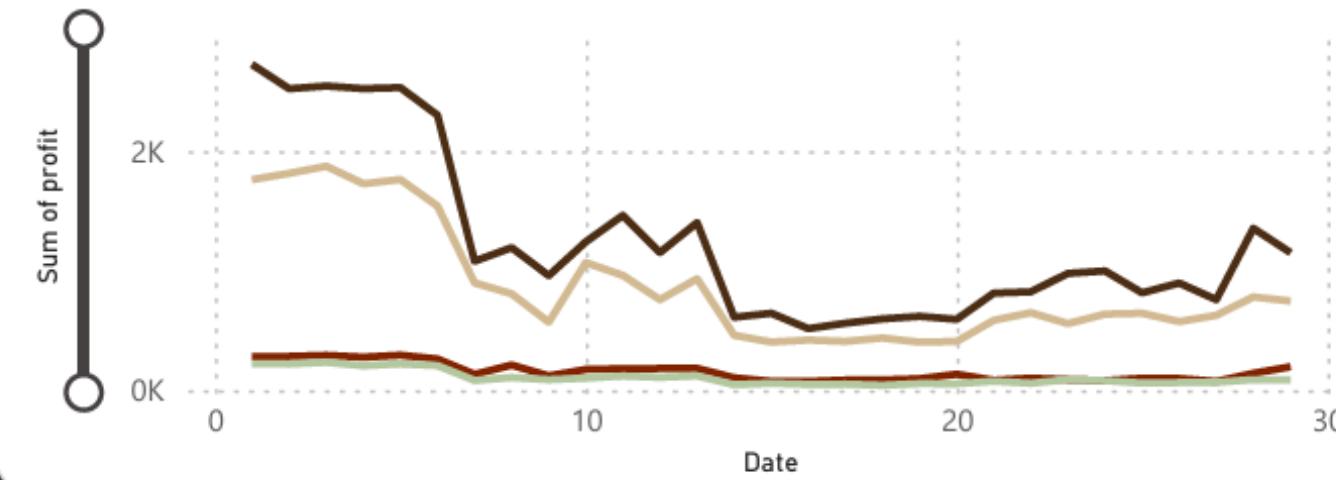
All

Age Group

All

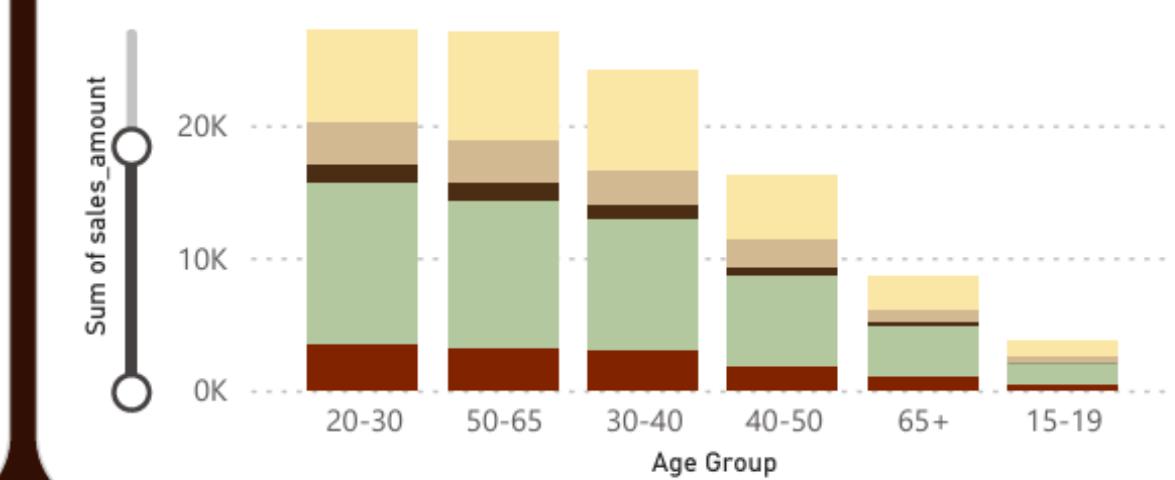
Sum of profit by Date and product_category

product_category ● Bakery ● Coffee ● Drinking Chocolate ● Tea



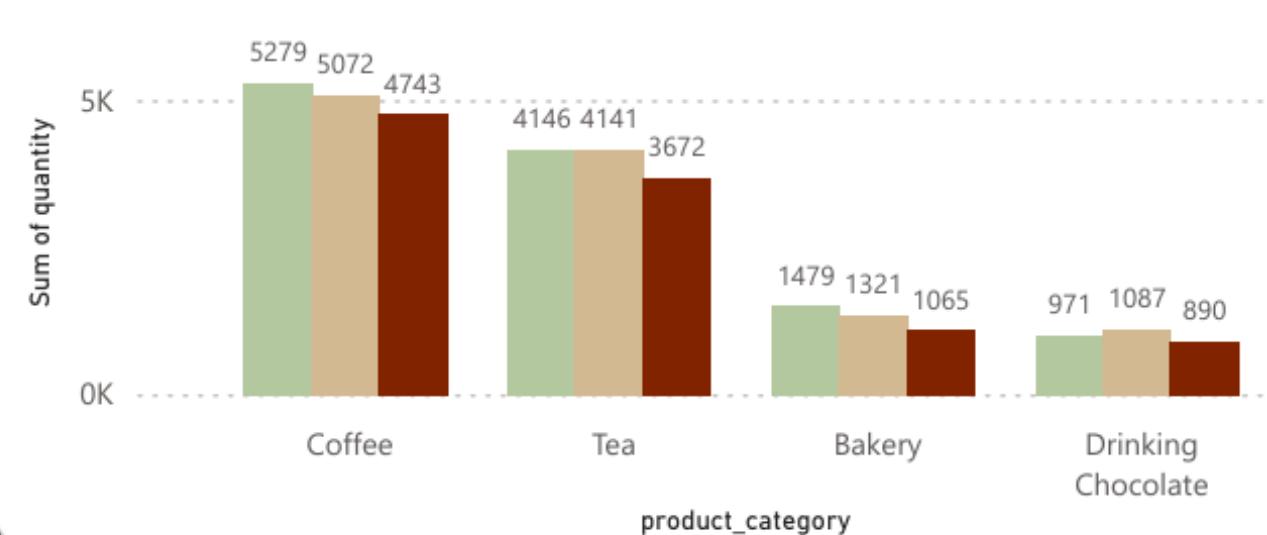
Sum of sales_amount by Age Group and product_category

product_category ● Bakery ● Coffee ● Coffee beans ● Drinking Choco... ● Tea



Sum of quantity by product_category and store_address

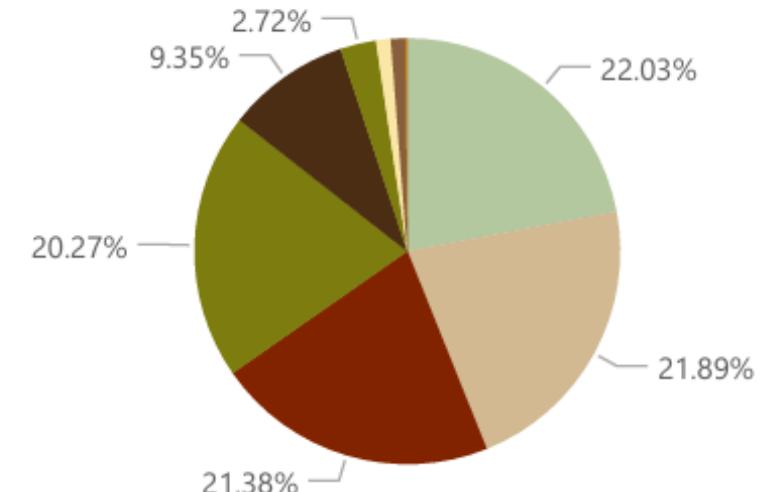
store_address ● 100 Church Street ● 32-20 Broadway ● 687 9th Avenue



Sum of total_cost by product_category

product_category

- Drinking Chocol...
- Coffee
- Bakery
- Tea
- Coffee beans
- Loose Tea
- Branded
- Packaged Choco...



Tổng quan



ETL



Phân tích



Đề xuất



Analysis by Product Subcategory



53.40%

Top5SubcategoryPercentage

product_group, product_subcategory

All

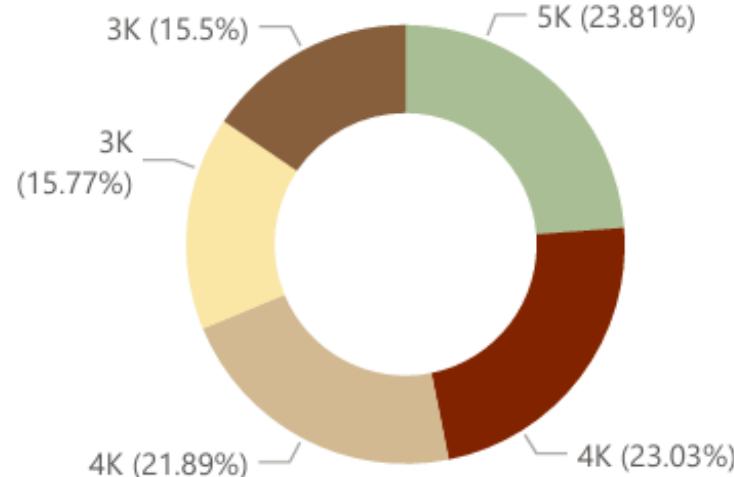
Week_ID, Date

All

Age Group

All

Sum of quantity of Top 5 product_subcategory by quantity



product_subcategory

- Brewed Chai tea
- Gourmet brewed c...
- Barista Espresso
- Brewed Black tea
- Hot chocolate

19023

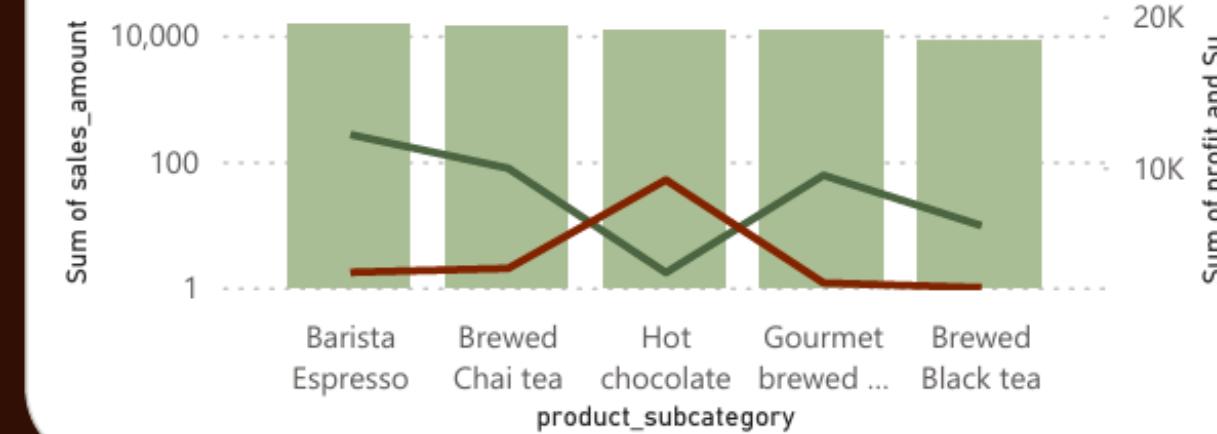
Top5SubcategoryQuantity

116

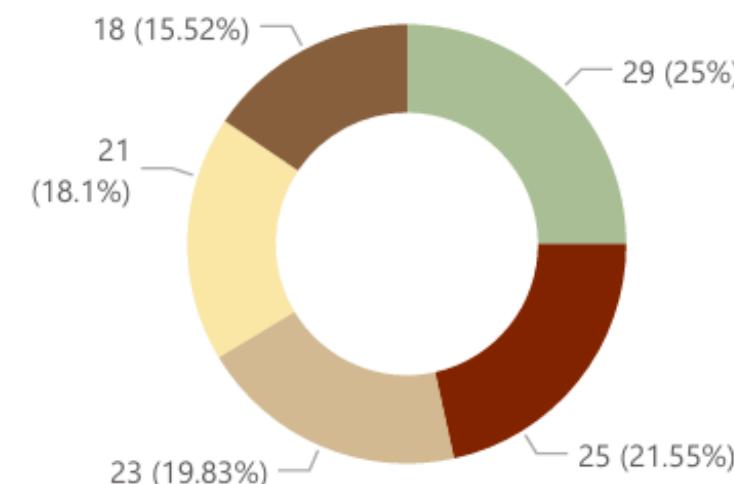
Bottom5SubcategoryQuantit...

Sum of sales_amount, Sum of profit and Sum of total_cost by product_subcategory

● Sum of sales_amount ● Sum of profit ● Sum of total_cost



Sum of quantity of 5 lowest subcategory by quantity

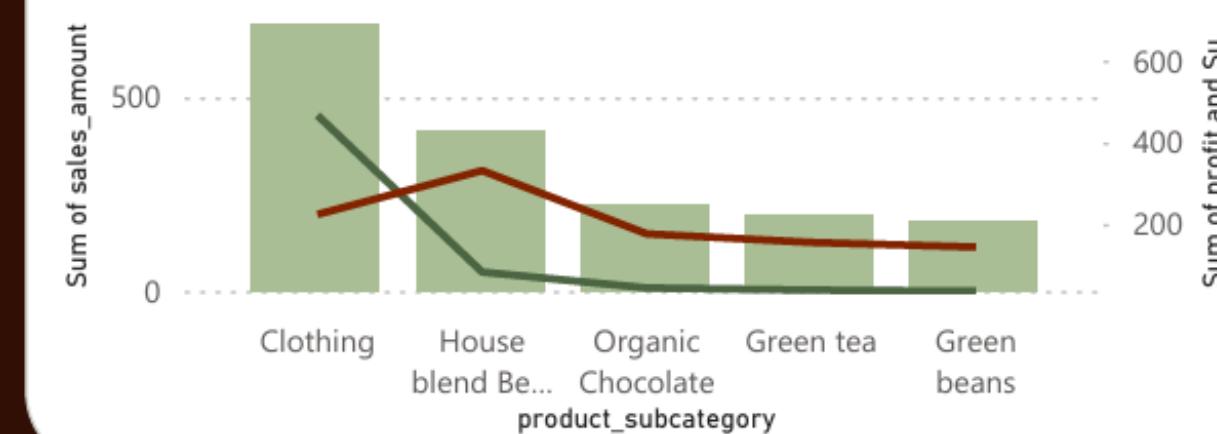


product_subcategory

- Organic Chocolate
- Clothing
- House blend Beans
- Green tea
- Green beans

Sum of sales_amount, Sum of profit and Sum of total_cost by product_subcategory

● Sum of sales_amount ● Sum of profit ● Sum of total_cost



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Sales outlet



114.12K

Sum of sales_amount

72.44K

Sum of profit

8

Max of quantity

72.00

Max of profit

product_group,
product_category,
product_subcategory

All

Week_ID, Date

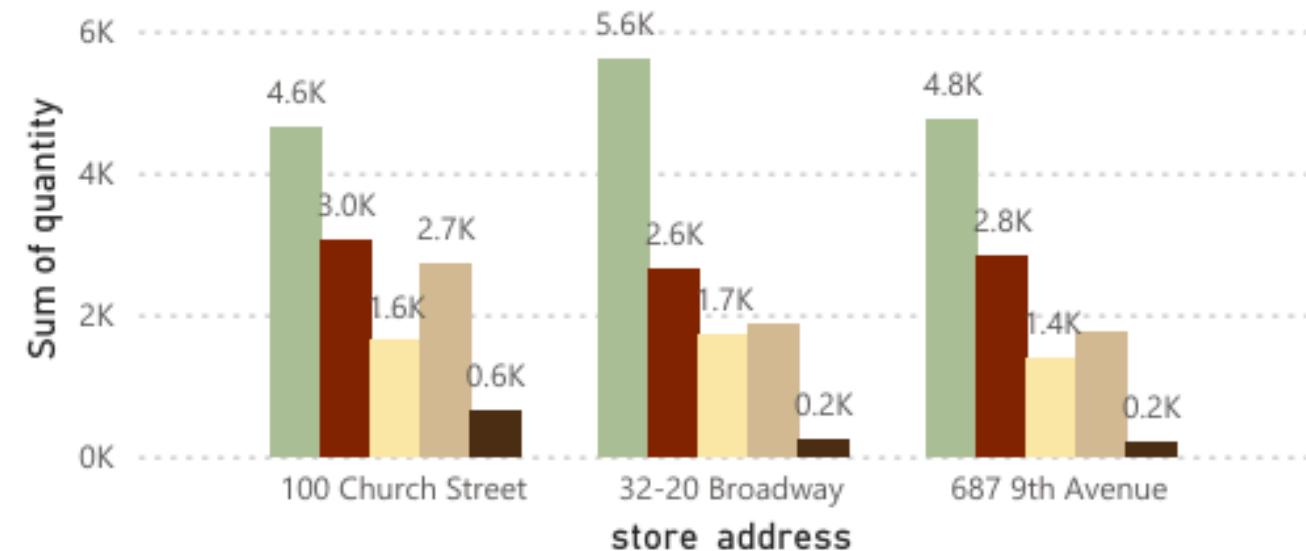
All

Age Group

All

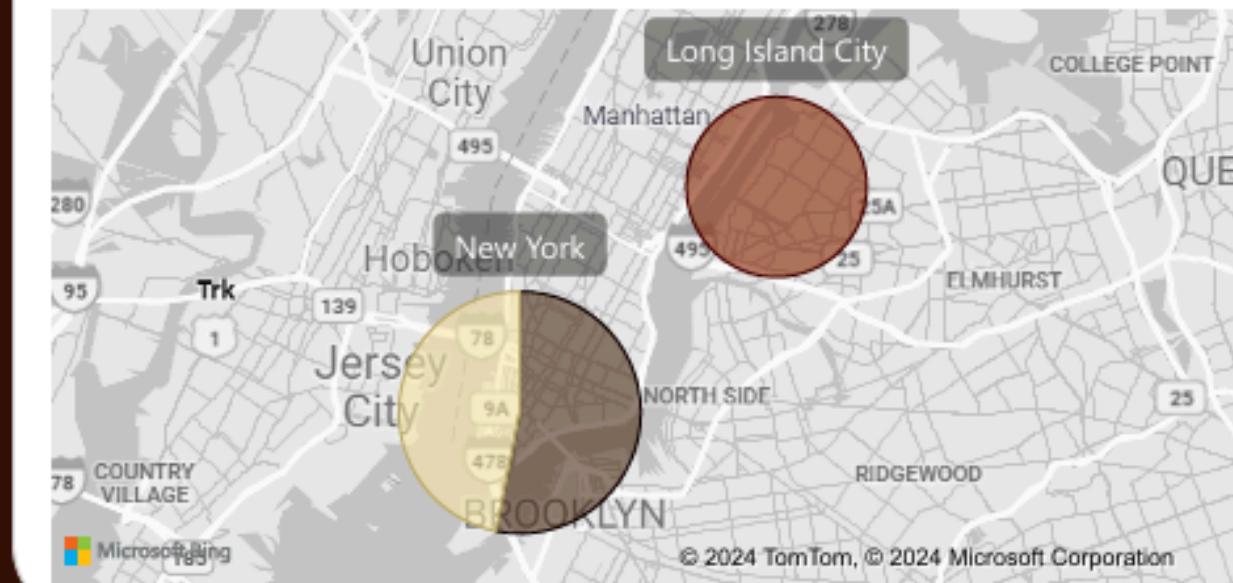
Sum of quantity by store_address and Week_Desc

Week_Desc ● Week 14 ● Week 15 ● Week 16 ● Week 17 ● Week 18



Sum of profit by store_city and store_address

store_address ● 100 Church Street ● 32-20 Broadway ● 687 9th Avenue



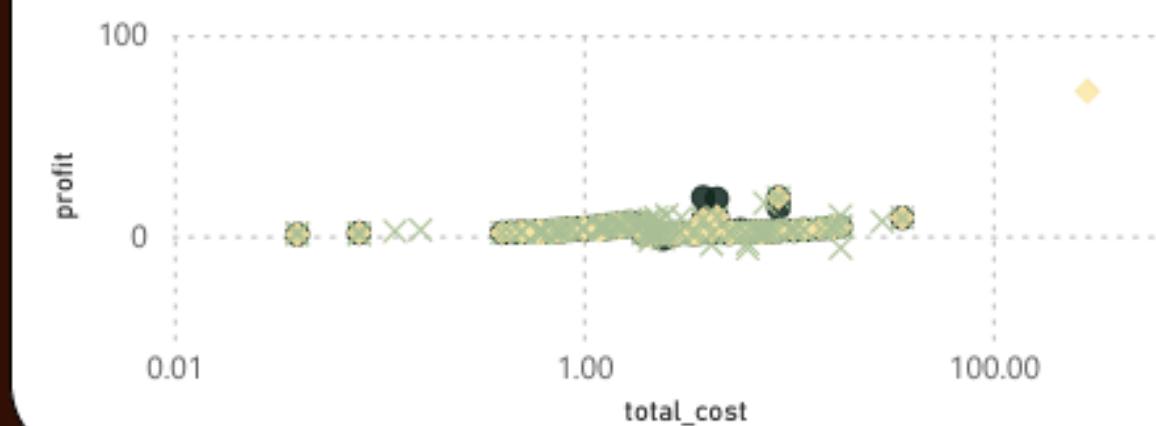
sales_outlet_id store_address

Sum of sales_amount total_goal Completion Percentage

3	32-20 Broadway	38,554.09	18000	214.19%
5	100 Church Street	39,953.87	25000	159.82%
8	687 9th Avenue	35,613.77	22500	158.28%
Total		114,121.73		456.49%

Neighborhood, total_cost and profit

Neighborhood ● Astoria ● Hell's Kitchen ● Lower Manhattan



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Pastry inventory



2854

Sum of quantity_sold

10.06K

Sum of waste_value

58.76%

Average of % waste

product_group,
product_category,
product_subcategory

All

Week_ID, Date

All

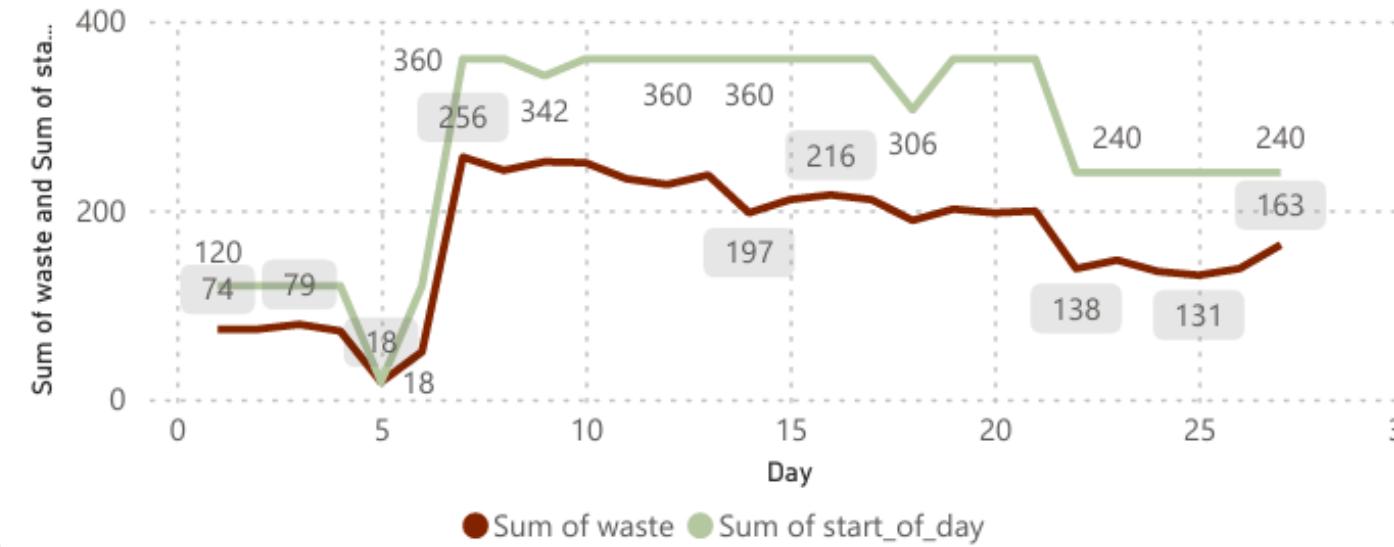
Age Group

All

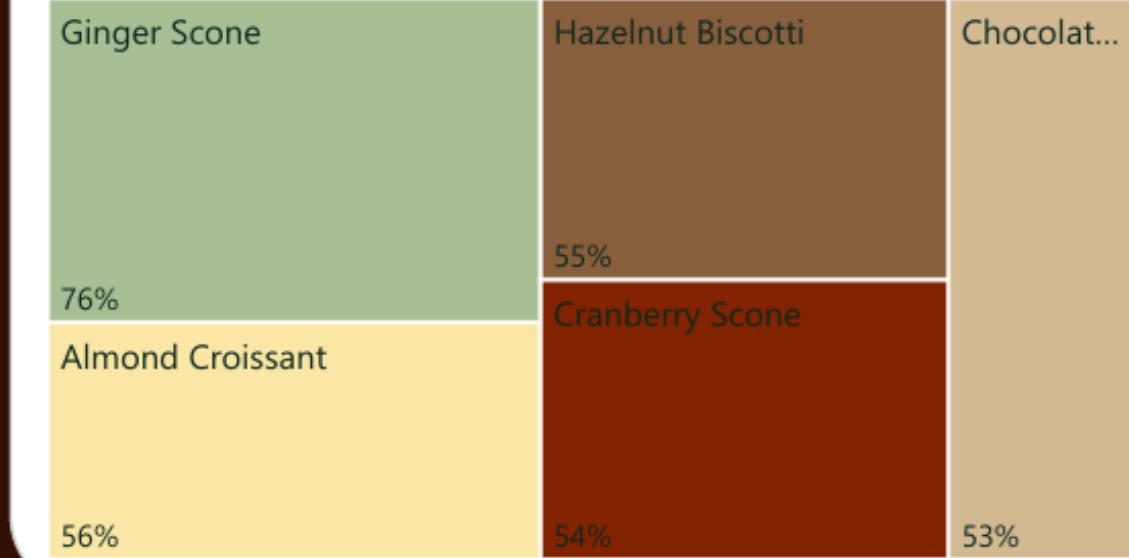
store_city

All

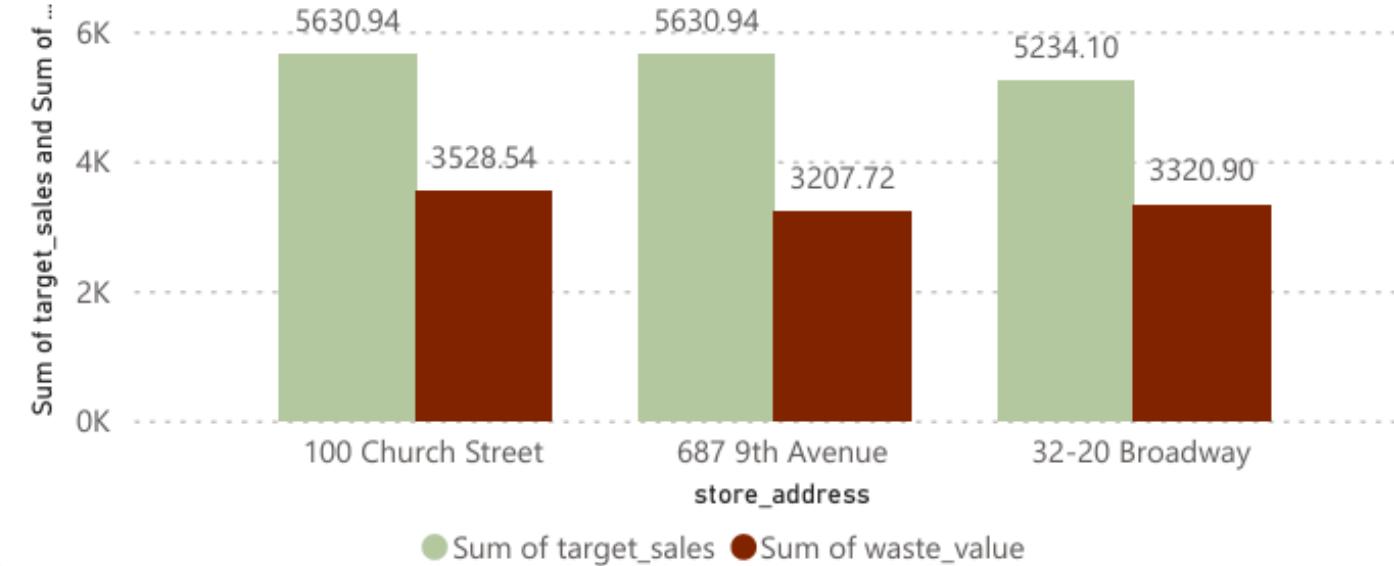
Sum of waste and Sum of start_of_day by Day



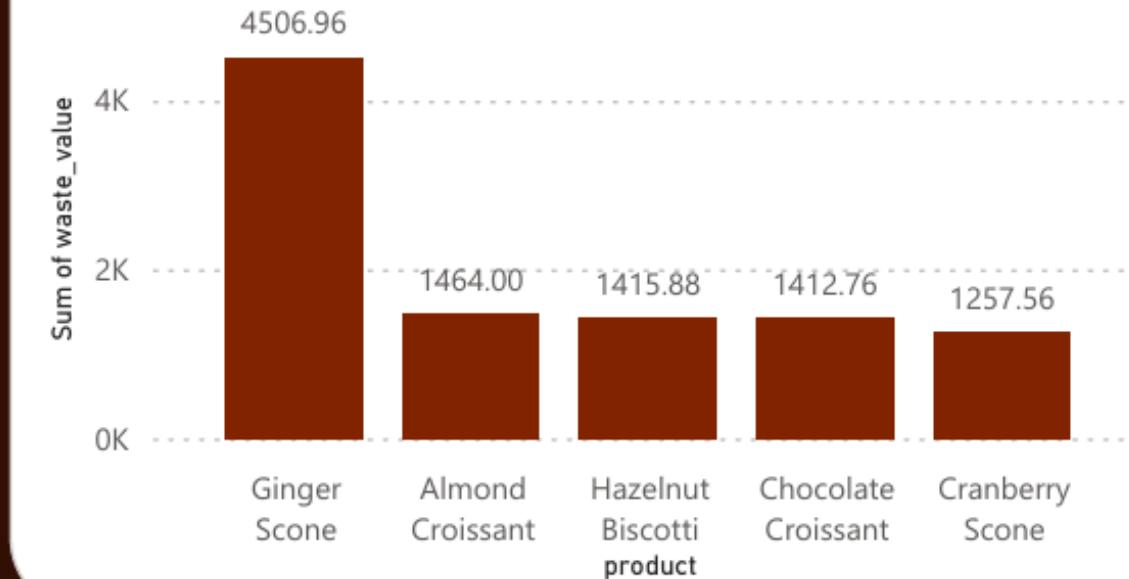
Average of % waste by product



Sum of target_sales and Sum of waste_value by store_address



Sum of waste_value by product



Tổng quan



ETL



Phân tích



Đề xuất



THE
COFFEE
SHOP

Analysis by Customer

41

Average of Age

1

Average of Engaged_years

2246

Count of customer_id

3257

Count of transaction_id

product_group,
product_category,
product_subcategory

All

Week_ID, Date

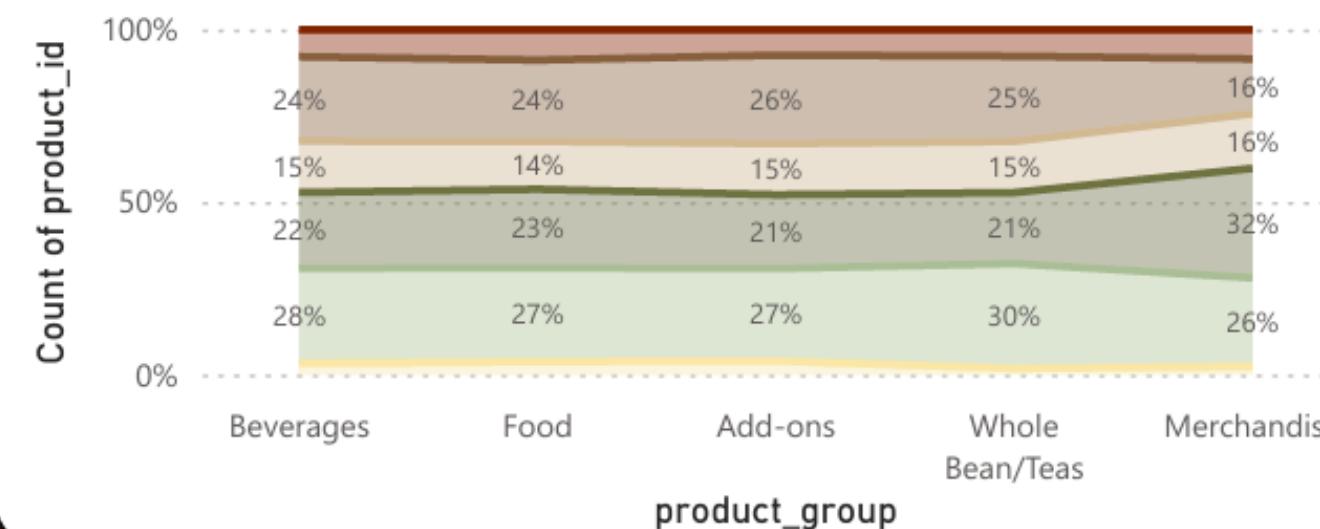
All

Age Group

All

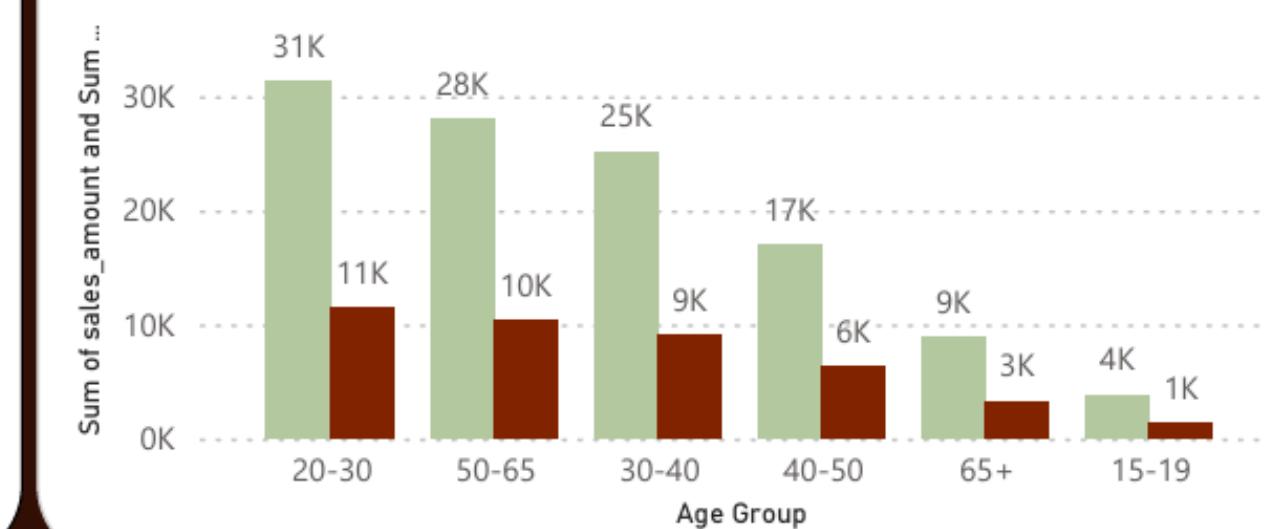
Count of product_id by product_group and Age Group

Age Group ● 15-19 ● 20-30 ● 30-40 ● 40-50 ● 50-65 ● 65+

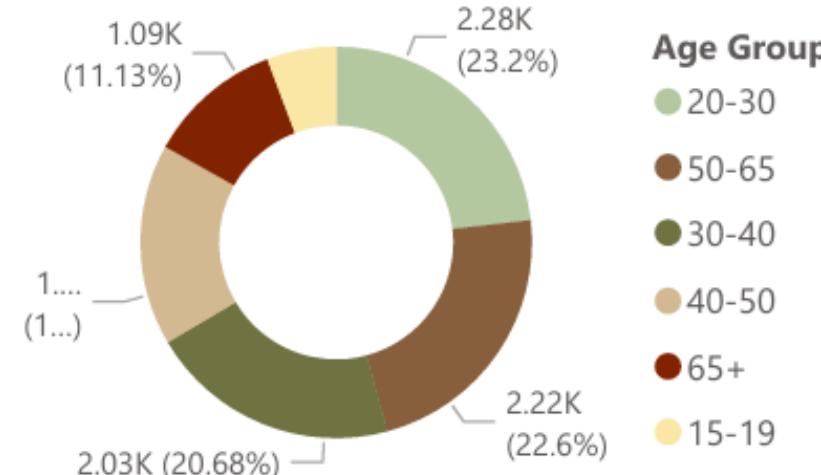


Sum of sales_amount and Sum of total_cost by Age Group

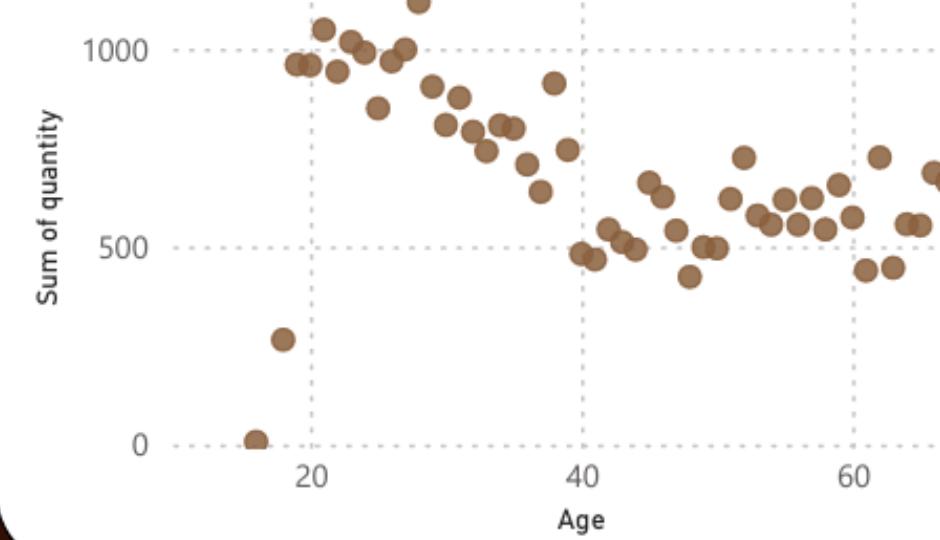
● Sum of sales_amount ● Sum of total_cost



Count of transaction_id by Age Group

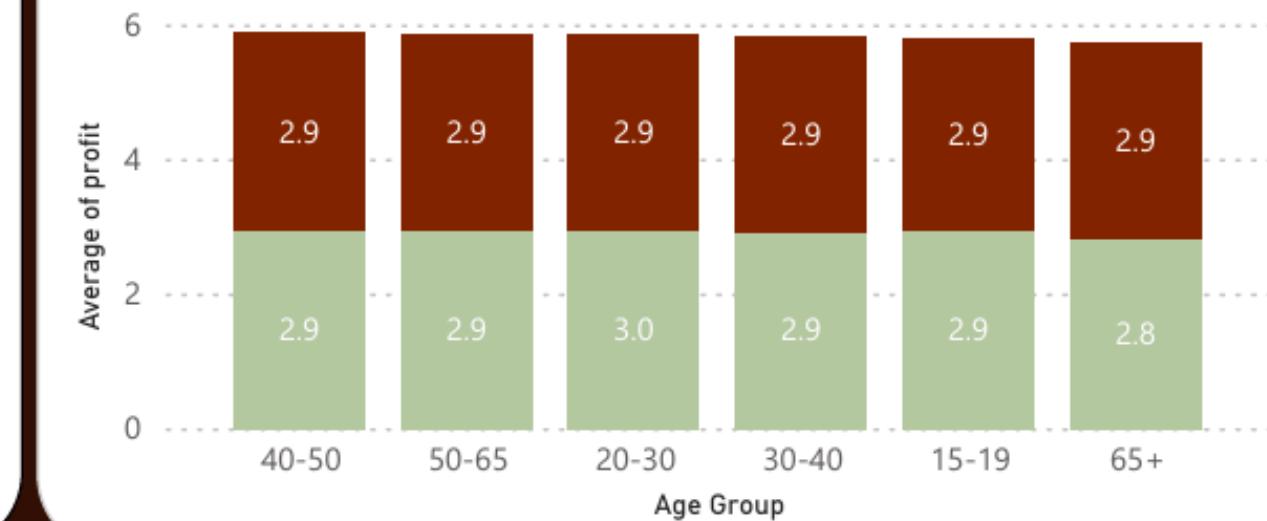


Sum of quantity by Age



Average of profit by Age Group and gender

gender ● F ● M



Tổng quan



ETL



Phân tích



Đề xuất

4

Giải pháp





Mục tiêu

- Áp dụng các chương trình khuyến mãi
- Nâng cao hiệu quả quảng bá sản phẩm mới
- Thăm dò ý kiến khách hàng
- Tạo chiến lược marketing
- Cải thiện chất lượng sản phẩm/dịch vụ
- Quản lý kho hiệu quả



Tổng quan



ETL



Phân tích



Đề xuất

Thank you for listening

