

DA Assignment – 1

The growth of supermarkets in most populated cities is increasing and market competitions are also high. The dataset is one of the historical sales of supermarket company which has recorded in 3 different branches for 3 months data. Predictive data analytics methods are easy to apply to this dataset.

Attribute information:

Invoice id: Computer-generated sales slip invoice identification number

Branch: Branch of supercenter (3 branches are available identified by A, B and C)

City: Location of supercenters

Customer type: Type of customers, recorded by Members for customers using member cards and Normal for those without member cards

Gender: Gender type of customer

Product line: General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel

Unit price: The price of each product in \$

Quantity: Number of products purchased by the customer

Tax: 5% tax fee for customers buying

Total: Total price including tax

Date: Date of purchase (Record available from January 2019 to March 2019)

Time: Purchase time (10 am to 9 pm)

Payment: Payment used by the customer for the purchase (3 methods are available – Cash, Credit card and Ewallet)

COGS: Cost of goods sold

Gross margin percentage: Gross margin percentage

Gross income: Gross income

Rating: Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

Challenge:

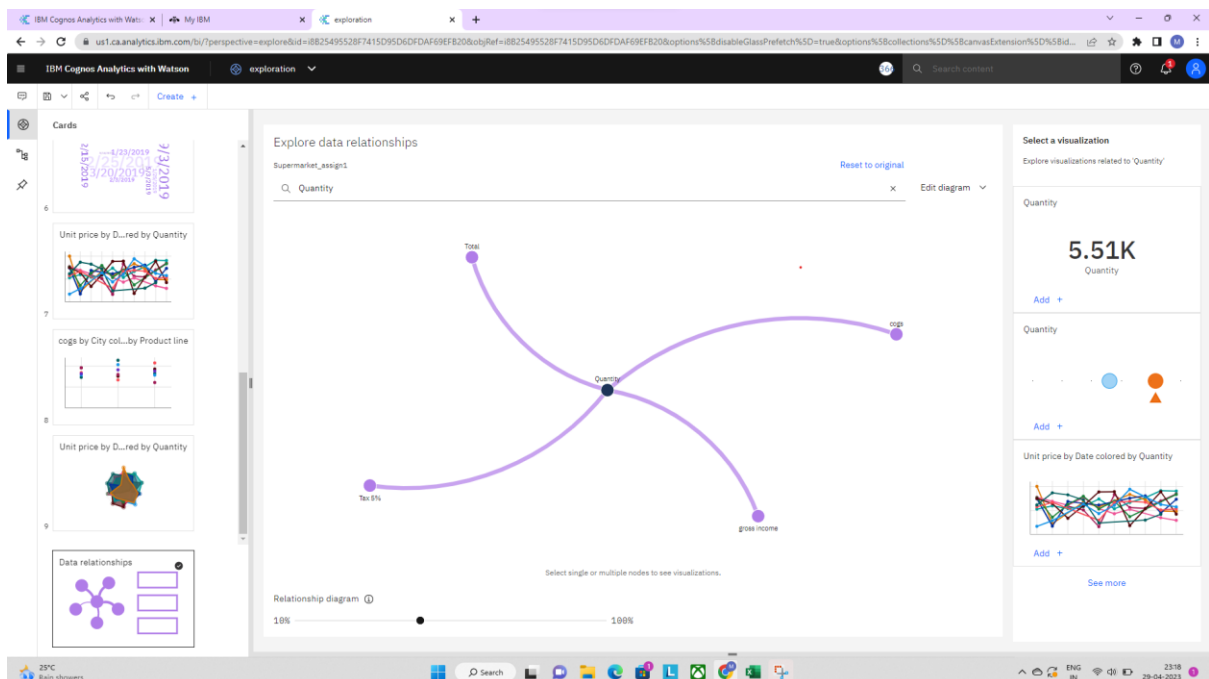
Upload the dataset to Cognos Analytics, delete the unnecessary columns, create a data module, explore and visualize the dataset

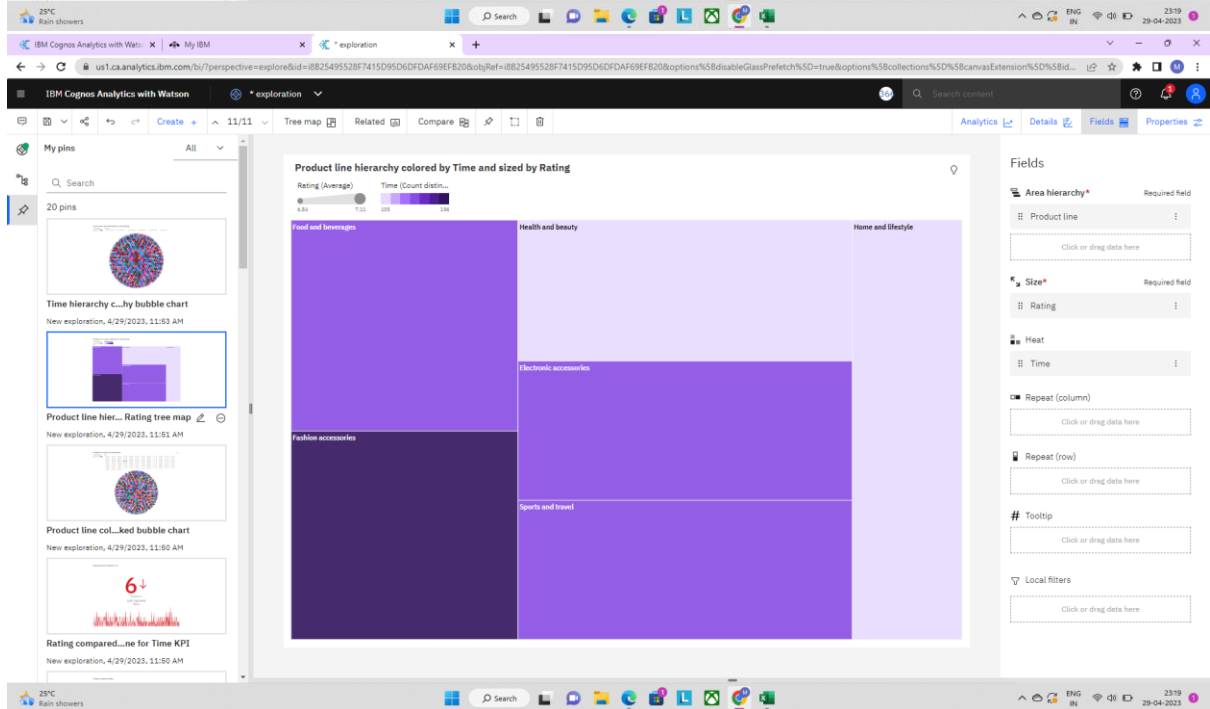
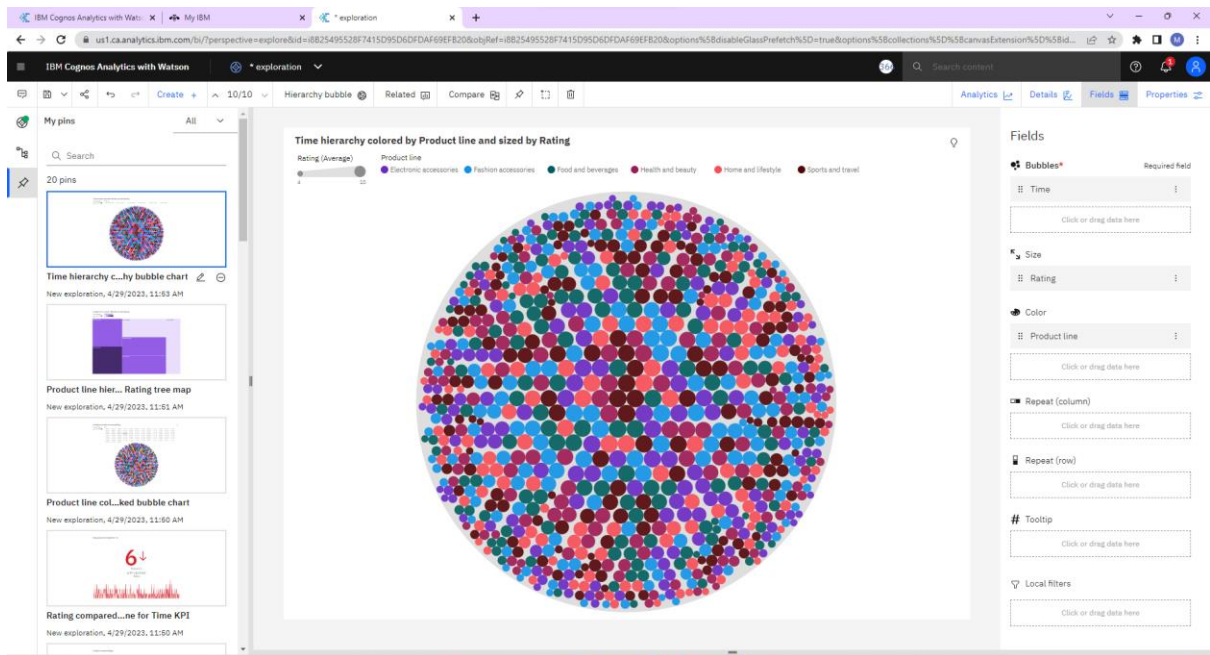
Data Module:

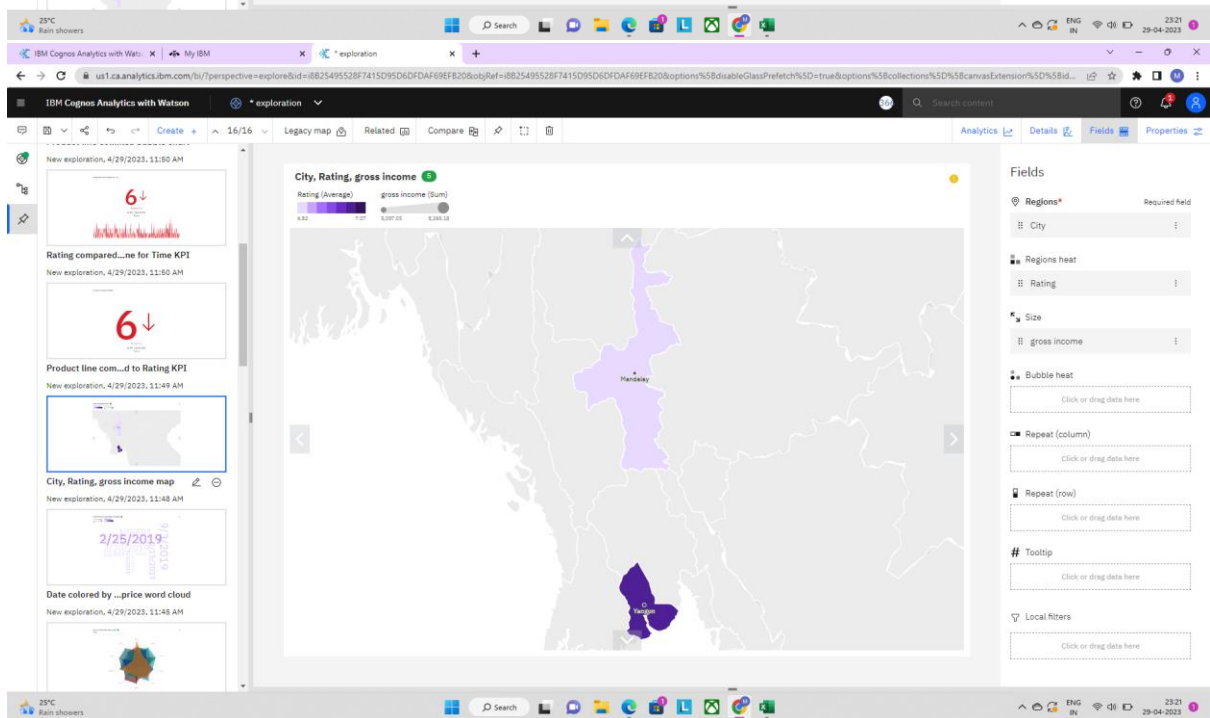
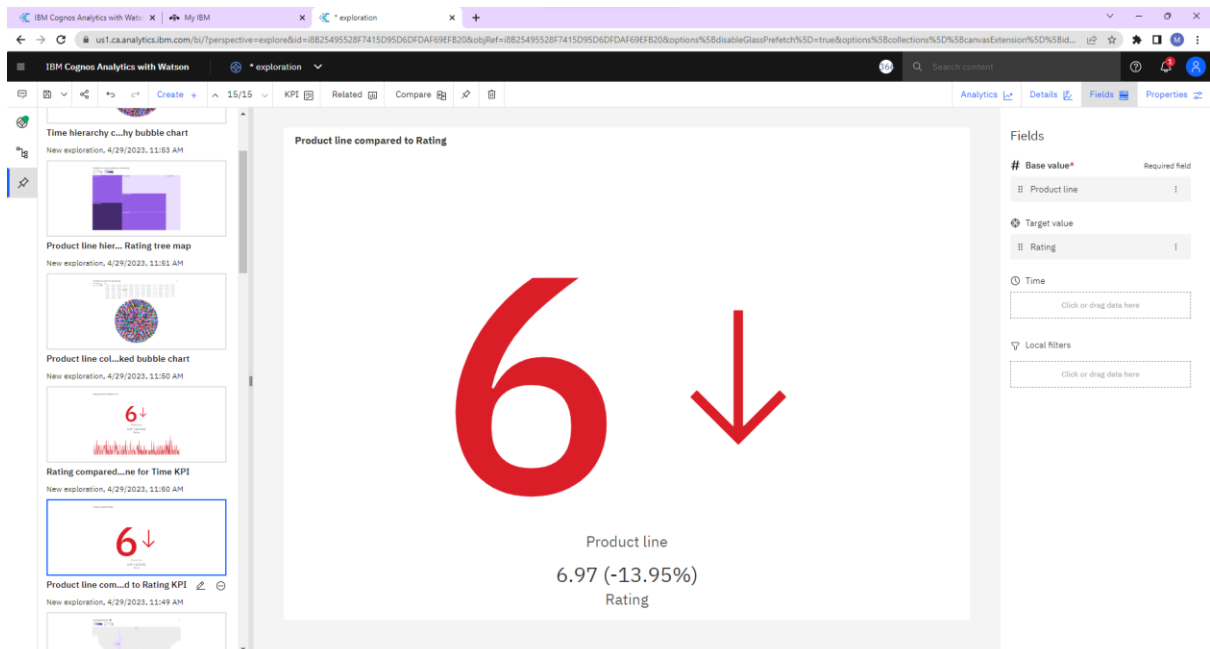
The screenshot shows the IBM Cognos Analytics Data Module interface. On the left, a 'Data module' sidebar lists various fields: Row Id, Invoice ID, Branch, City, Customer type, Gender, Product line, Unit price, Quantity, Tax 5%, Total, Date, Time, Payment, costs, gross mar..., gross income, and Rating. The main area displays a data grid with 19 rows and 10 columns. The columns are: Row Id, Invoice ID, Branch, City, Customer type, Gender, Product line, Unit price, and Quantity. The data represents supermarket transactions across different branches and customer types.

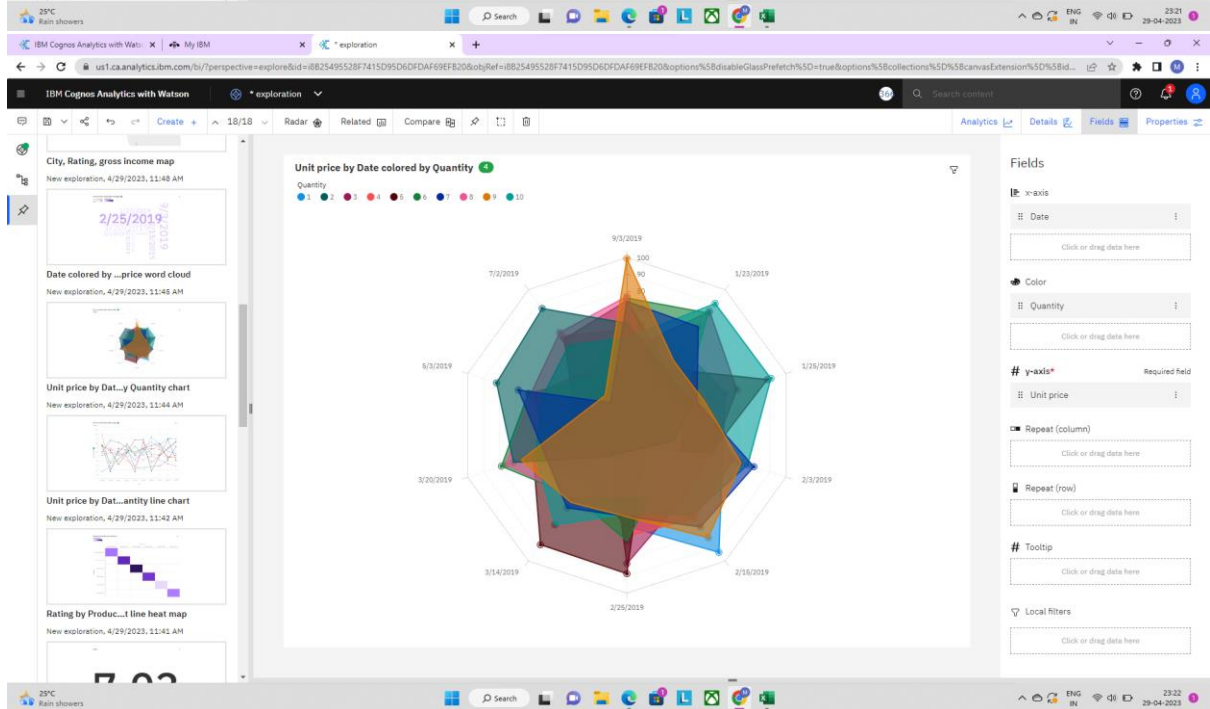
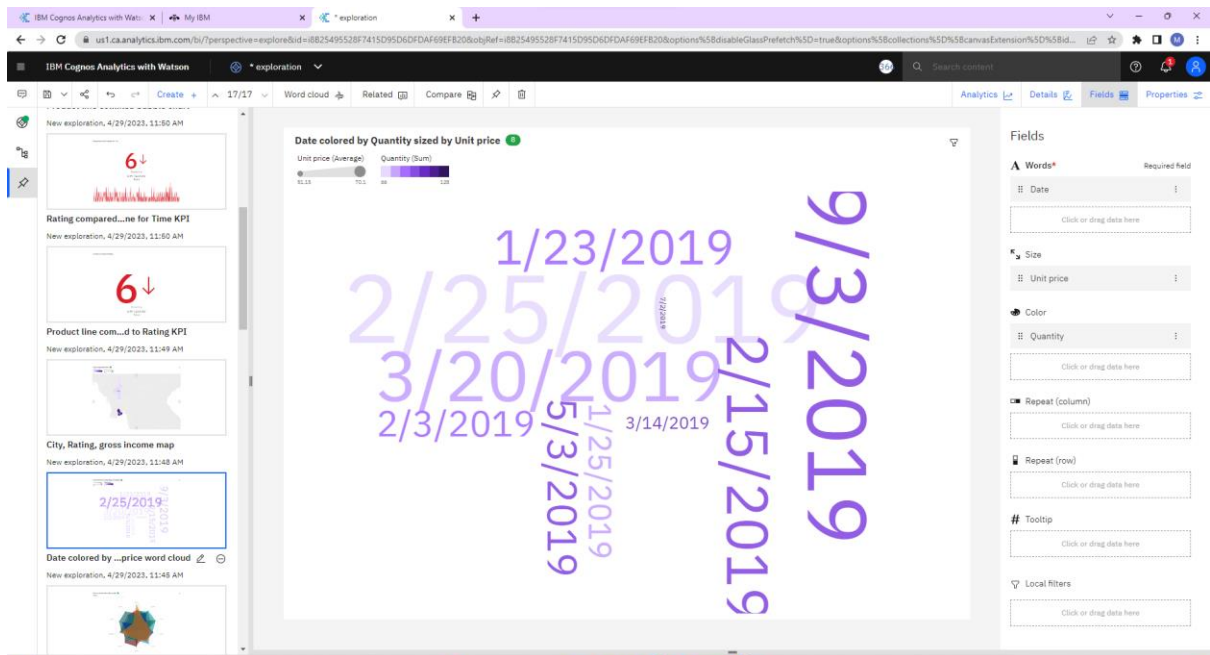
Row Id	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity
1	750-67-0420	A	Yangon	Member	Female	Health and beauty	74.69	7
2	226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5
3	631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7
4	123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8
5	373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7
6	699-14-3026	C	Naypyitaw	Normal	Male	Electronic accessories	85.39	7
7	355-63-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6
8	315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10
9	665-32-9167	A	Yangon	Member	Female	Health and beauty	36.26	2
10	692-92-5582	B	Mandalay	Member	Female	Food and beverages	54.84	3
11	351-62-0822	B	Mandalay	Member	Female	Fashion accessories	14.48	4
12	529-56-3974	B	Mandalay	Member	Male	Electronic accessories	25.51	4
13	365-64-0515	A	Yangon	Normal	Female	Electronic accessories	46.95	5
14	252-56-2699	A	Yangon	Normal	Male	Food and beverages	43.19	10
15	829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10
16	299-46-1805	B	Mandalay	Member	Female	Sports and travel	93.72	6
17	656-95-9349	A	Yangon	Member	Female	Health and beauty	68.93	7
18	765-26-6951	A	Yangon	Normal	Male	Sports and travel	72.61	6
19	329-62-1586	A	Yangon	Normal	Male	Food and beverages	54.67	3

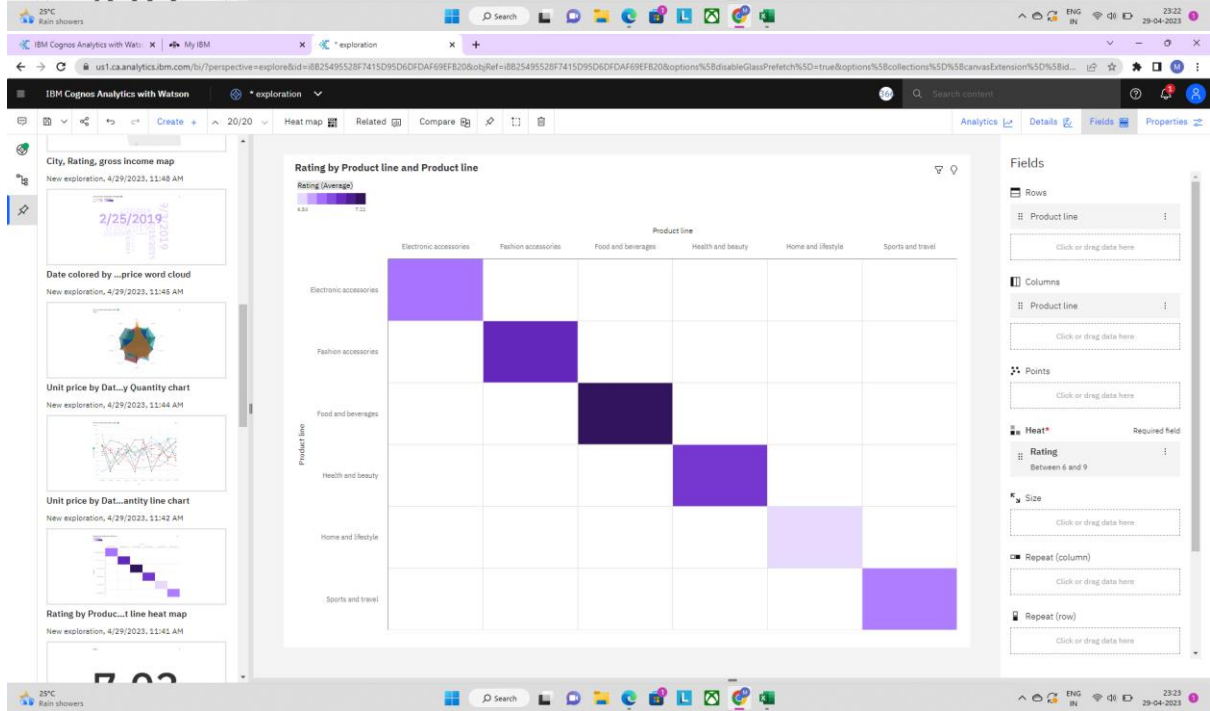
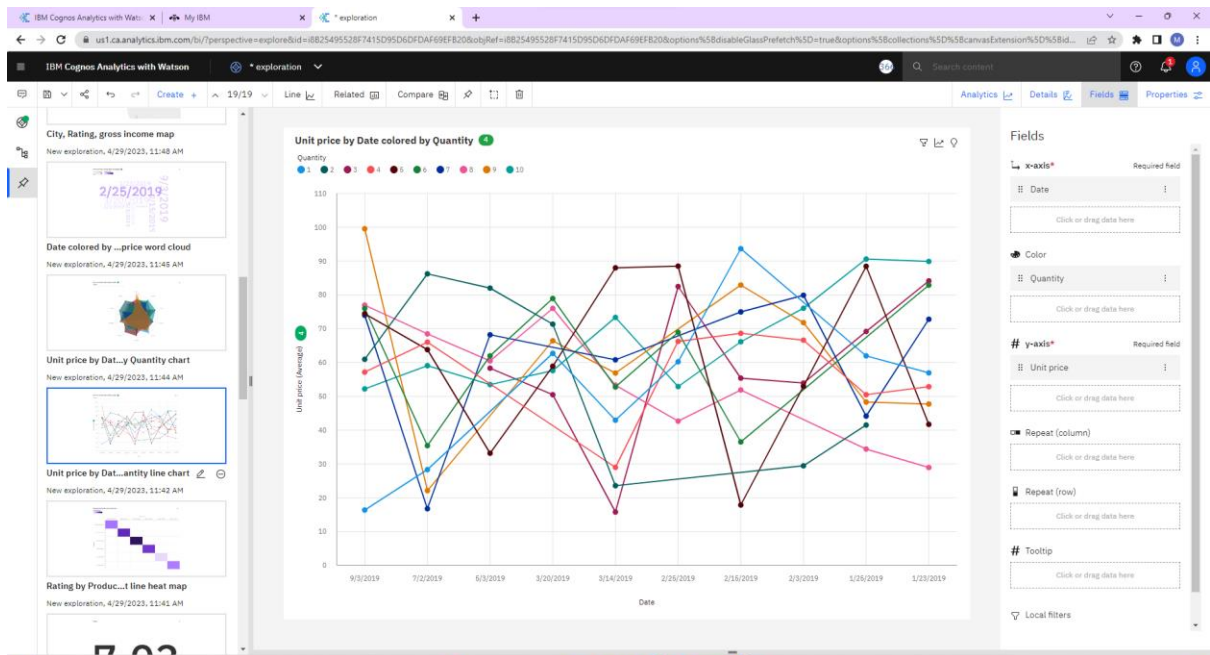
Data Exploration:

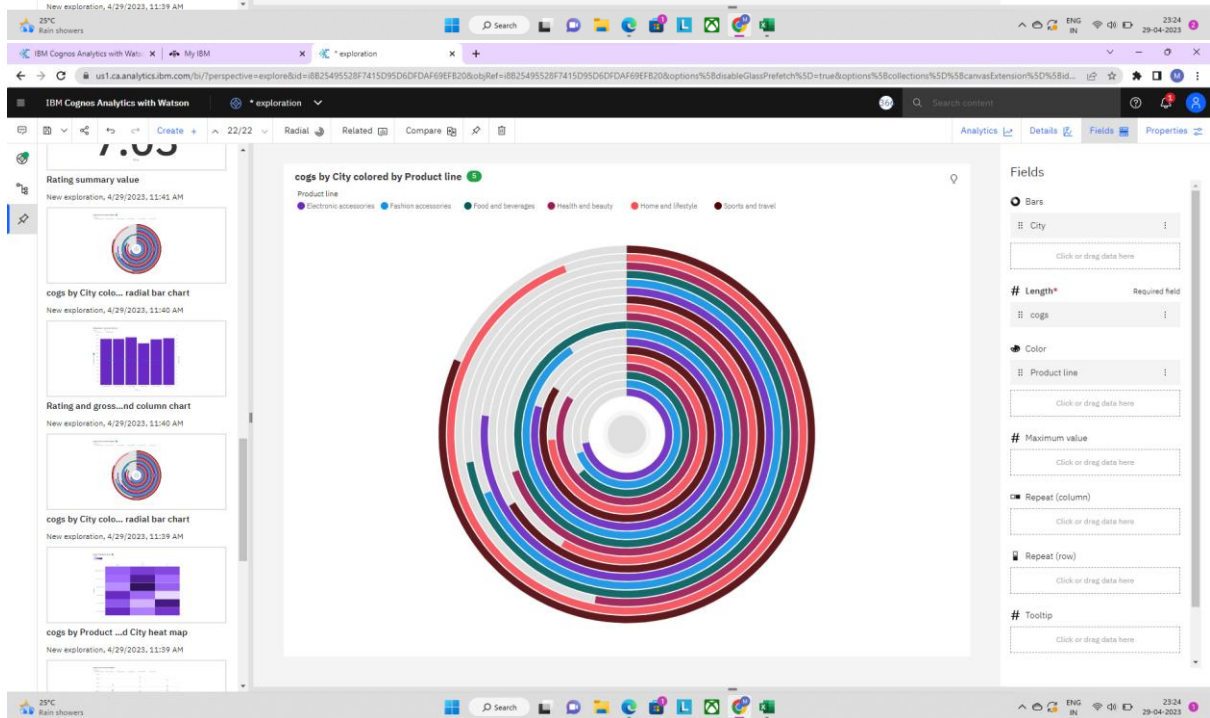
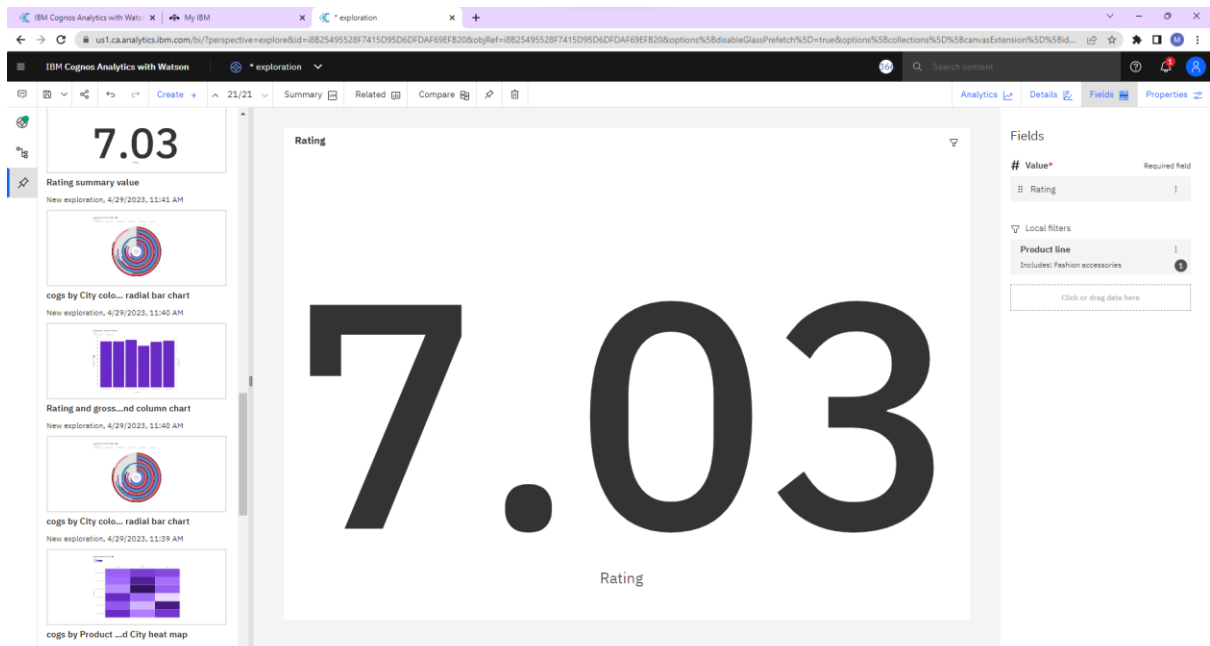


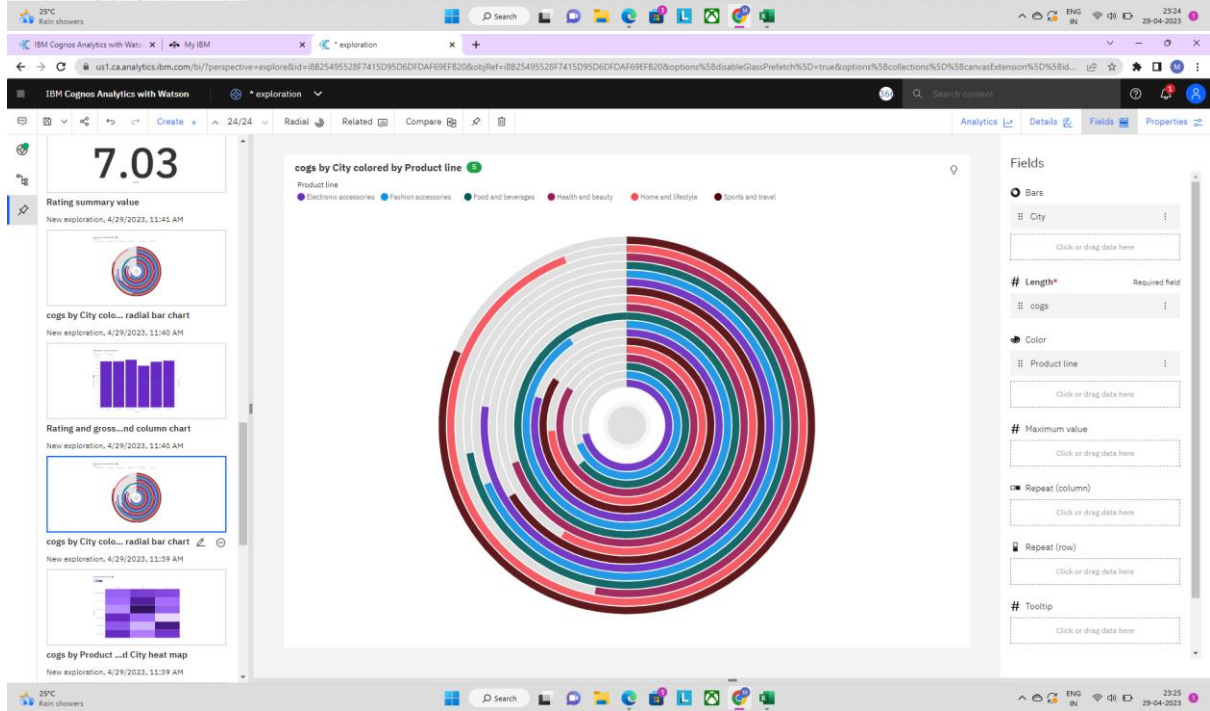
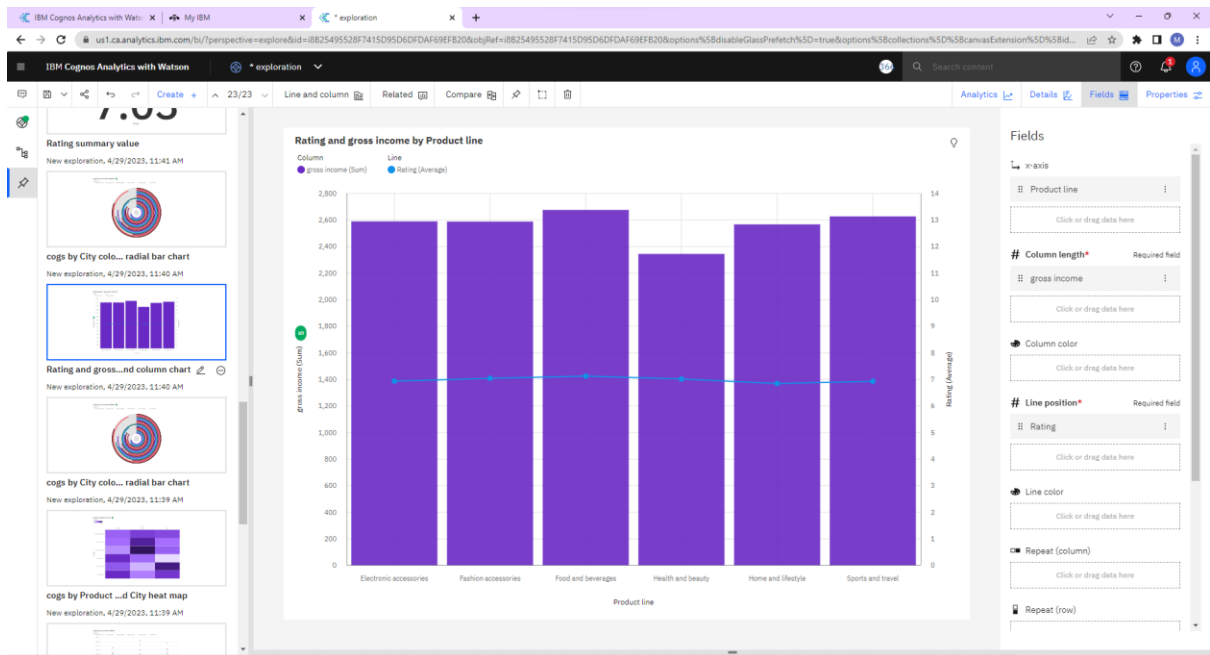


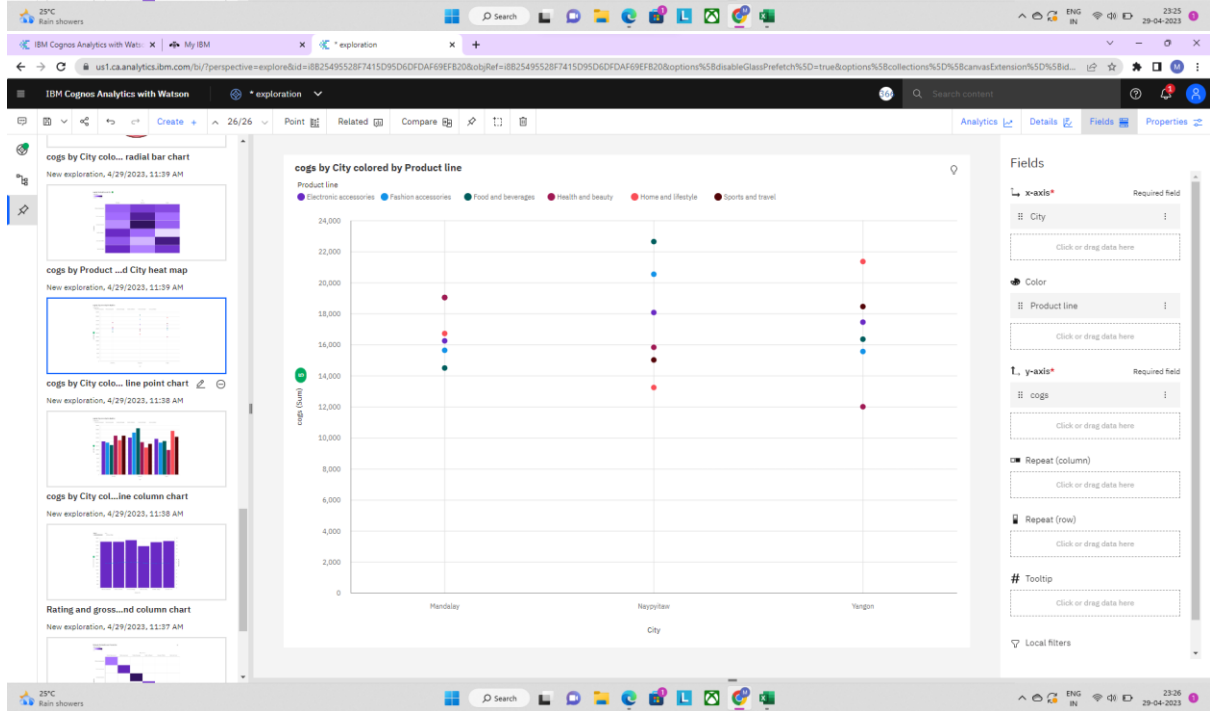
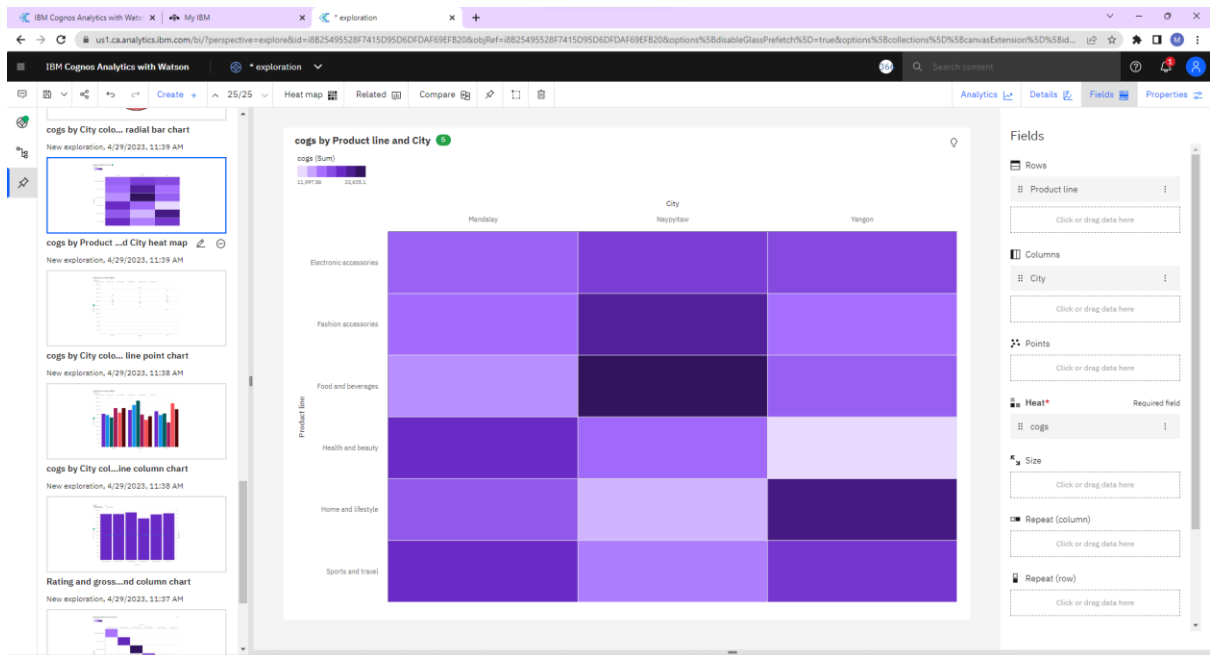


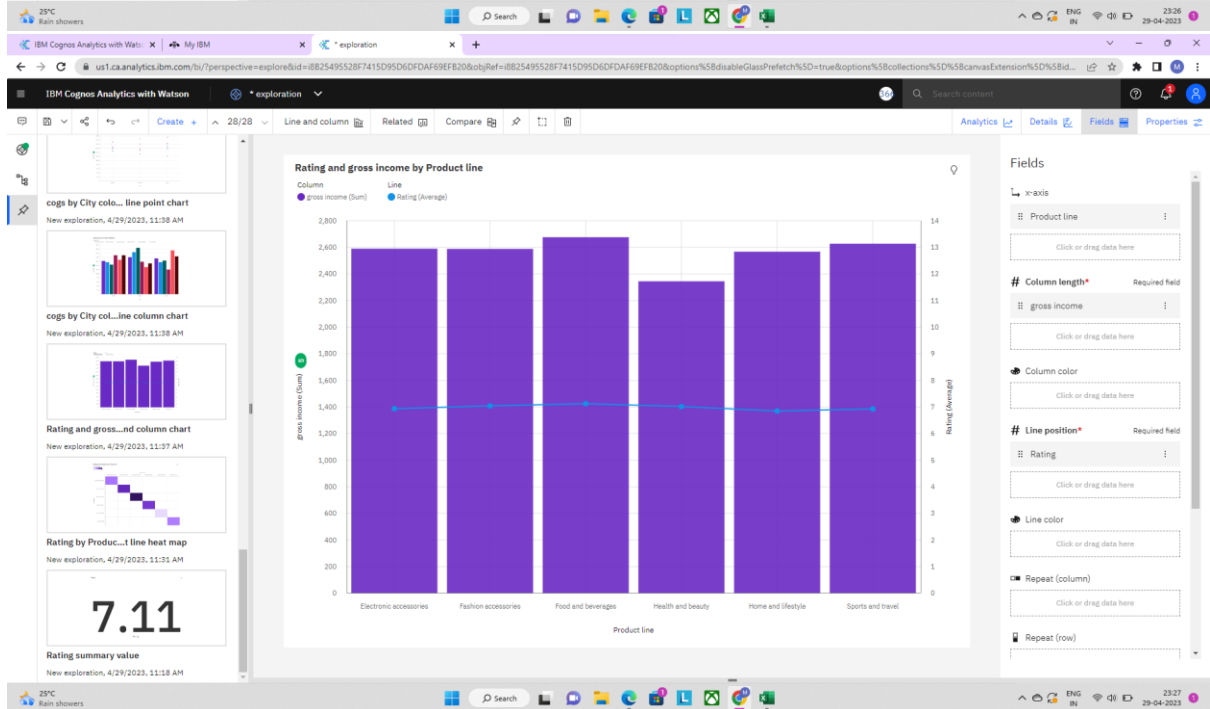
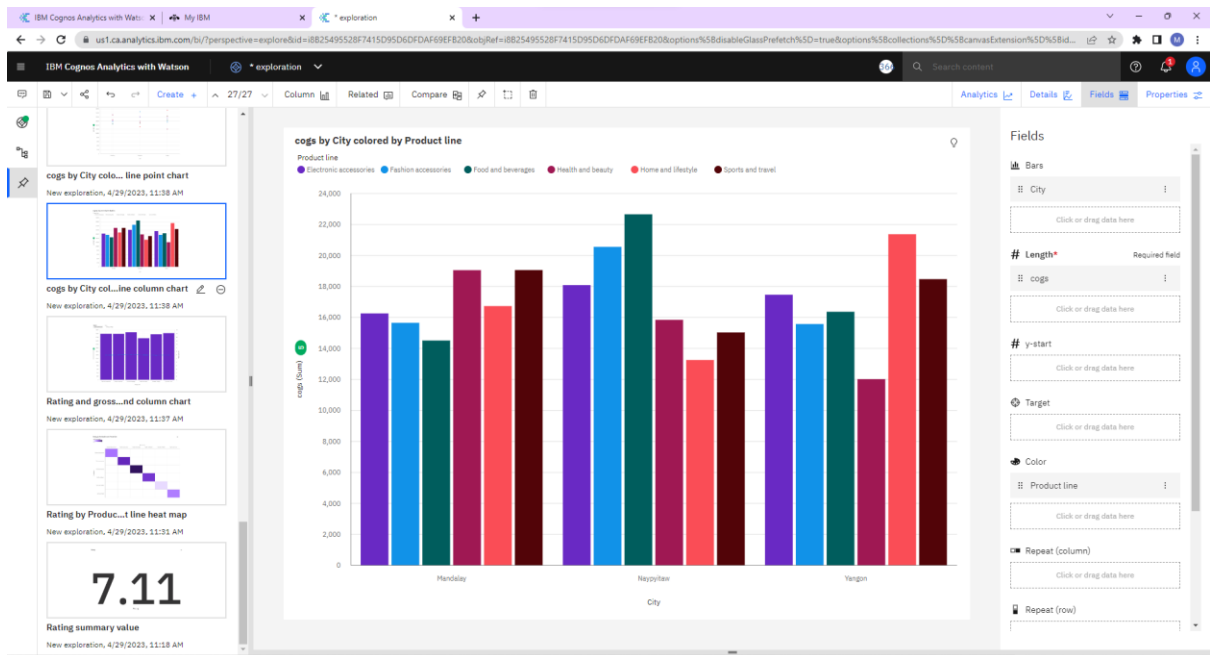


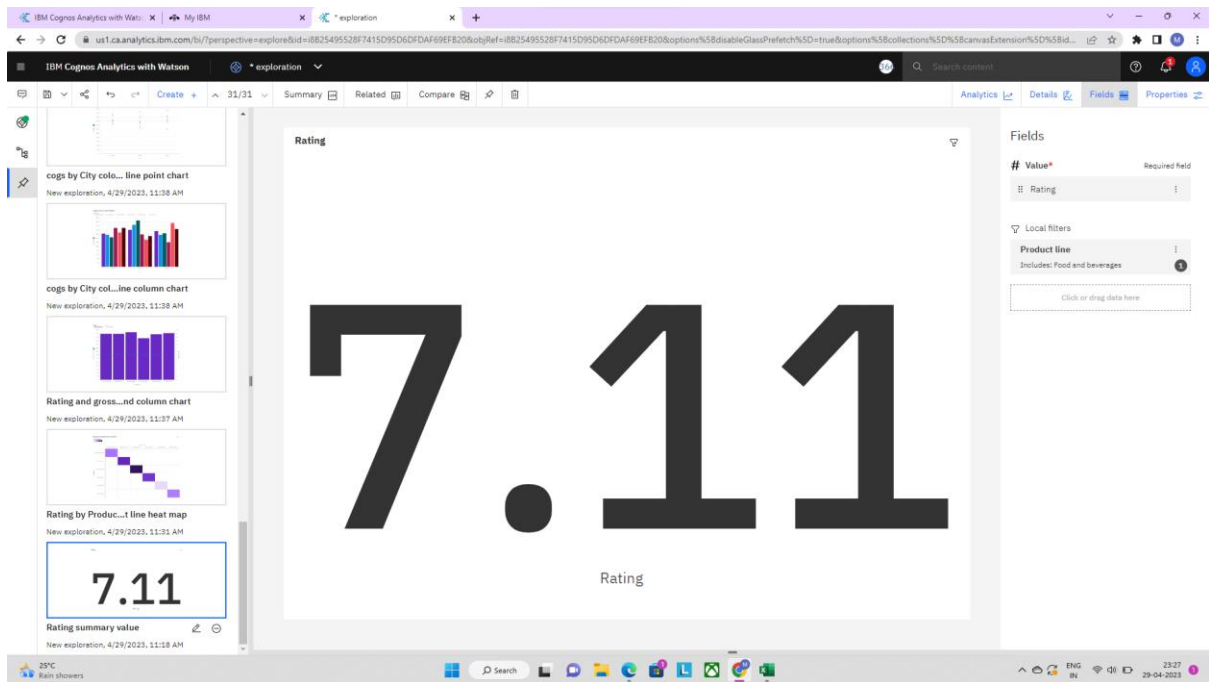












Data Visualization:

