

# Data Analytics with IBM Cognos Analytics

## Assignment 1

Name : Korukonda Pradeep

Registration Number : 21BCE5255

Email: [korukonda.pradeep2021@vitstudent.ac.in](mailto:korukonda.pradeep2021@vitstudent.ac.in)

VIT Chennai

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 a supermarket company which has been recorded in 3 different branches for 3 months. 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

The screenshot shows the IBM Cognos Analytics interface. On the left, there is a 'Data module' pane with a search bar and a list of fields: # Row Id, Invoice ID, Branch, City (with sub-items Mandalay, Naypyitaw, Yangon), Customer type, Gender, Product line, Unit price, Quantity, Tax 5%, Total, Date, Time, and Payment. The main area displays a grid of data with columns: Row Id, Invoice ID, Branch, City, Customer type, Gender, and Product line. The grid contains 13 rows of data.

Row Id	Invoice ID	Branch	City	Customer type	Gender	Product line
1	750-67-8428	A	Yangon	Member	Female	Health and be
2	226-31-3081	C	Naypyitaw	Normal	Female	Electronic acc
3	631-41-3108	A	Yangon	Normal	Male	Home and lifes
4	123-19-1176	A	Yangon	Member	Male	Health and be
5	373-73-7910	A	Yangon	Normal	Male	Sports and tra
6	699-14-3026	C	Naypyitaw	Normal	Male	Electronic acc
7	355-53-5943	A	Yangon	Member	Female	Electronic acc
8	315-22-5665	C	Naypyitaw	Normal	Female	Home and lifes
9	665-32-9167	A	Yangon	Member	Female	Health and be
10	692-92-5582	B	Mandalay	Member	Female	Food and beve
11	351-62-0822	B	Mandalay	Member	Female	Fashion access
12	529-56-3974	B	Mandalay	Member	Male	Electronic acc
13	365-64-0515	A	Yangon	Normal	Female	Electronic acc

## Removing Invoice ID

The screenshot shows the IBM Cognos Analytics interface. On the left, the 'Data module' pane lists various fields including 'Invoice ID', 'Branch', 'City', 'Customer type', 'Gender', 'Product line', 'Unit price', 'Quantity', 'Tax 5%', 'Total', 'Date', 'Time', 'Payment', and 'cogs'. The main area displays a data grid with columns: Row Id, Invoice ID, Branch, City, Customer type, Gender, and Product line. A context menu is open over the 'Invoice ID' column, with the 'Remove' option selected.

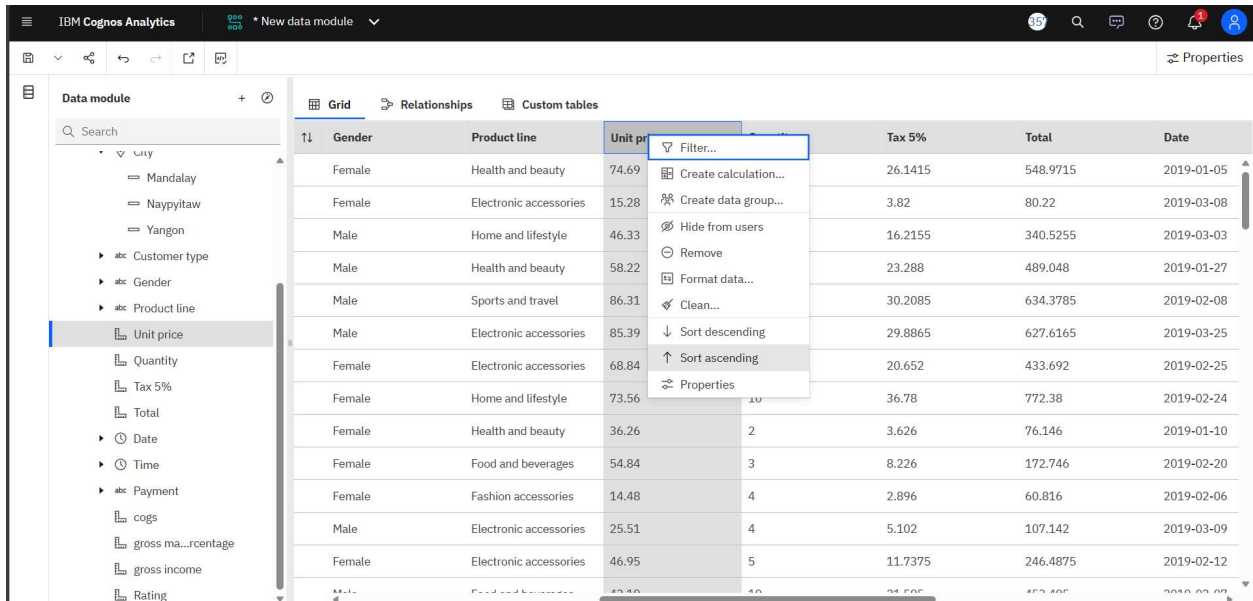
Row Id	Invoice ID	Branch	City	Customer type	Gender	Product line
1	750-67-8428		Yangon	Member	Female	Health and be
2	226-31-3081		Naypyitaw	Normal	Female	Electronic acc
3	631-41-3108		Yangon	Normal	Male	Home and lifes
4	123-19-1176		Yangon	Member	Male	Health and be
5	373-73-7910		Yangon	Normal	Male	Sports and trav
6	699-14-3026		Naypyitaw	Normal	Male	Electronic acc
7	355-53-5943		Yangon	Member	Female	Electronic acc
8	315-22-5665		Naypyitaw	Normal	Female	Home and lifes
9	665-32-9167		Yangon	Member	Female	Health and be
10	692-92-5582	B	Mandalay	Member	Female	Food and beve
11	351-62-0822	B	Mandalay	Member	Female	Fashion access
12	529-56-3974	B	Mandalay	Member	Male	Electronic acc
13	365-64-0515	A	Yangon	Normal	Female	Electronic acc

## Setting the decimal places of gross margin percentage to 3

The screenshot shows the 'Data format' dialog box in IBM Cognos Analytics. The 'Column' is set to 'gross margin percentage'. The 'Format type' is 'Number'. The 'Number of decimal places' is set to 3. Other options like 'Negative sign symbol', 'Use thousands separator', 'Negative sign position', and 'Missing value characters' are set to their default values. The background shows a data grid with columns 'gross income' and 'Rating'.

gross income	Rating
26.1415	9.1
3.82	9.6
16.2155	7.4
23.288	8.4
30.2085	5.3
29.8865	4.1
20.652	5.8
36.78	8
3.626	7.2
8.226	5.9
2.896	4.5
5.102	6.8
11.7375	7.1

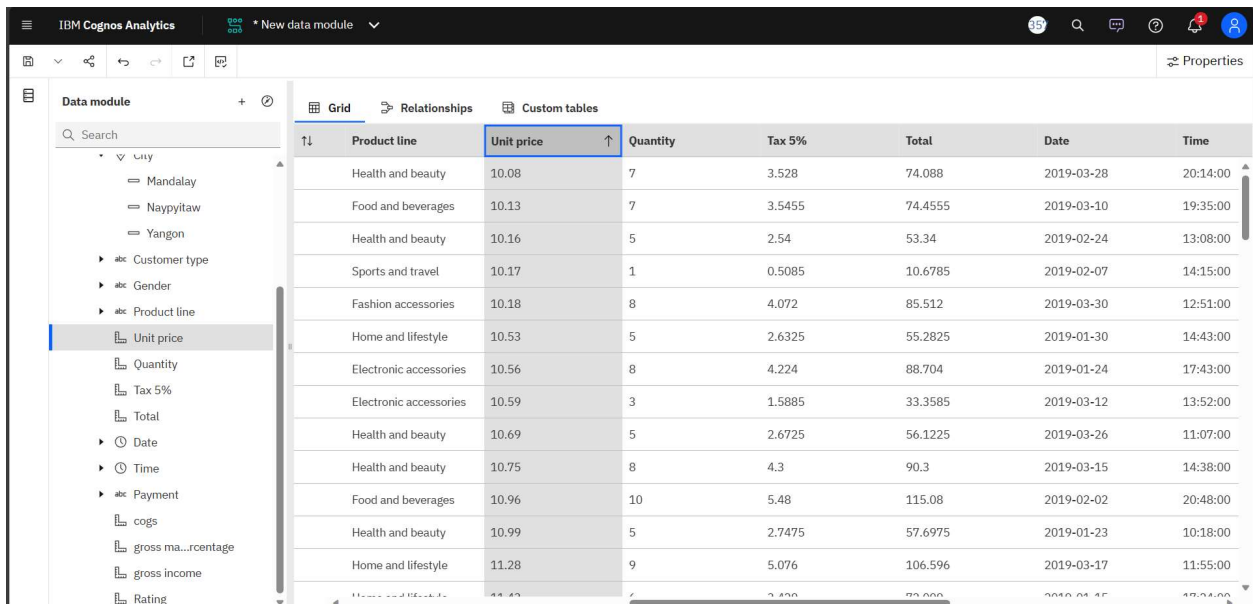
## Sorting Unit price in ascending order



IBM Cognos Analytics interface showing a data grid. The left sidebar displays the Data module hierarchy, including City (Mandalay, Naypyitaw, Yangon), Customer type, Gender, Product line, Unit price, Quantity, Tax 5%, Total, Date, Time, Payment, cogs, gross margin percentage, gross income, and Rating. The main grid shows data with columns: Gender, Product line, Unit price, Tax 5%, Total, and Date. A context menu is open over the 'Unit price' column, showing options: Filter..., Create calculation..., Create data group..., Hide from users, Remove, Format data..., Clean..., Sort descending, Sort ascending, and Properties. The 'Sort ascending' option is highlighted.

Gender	Product line	Unit price	Tax 5%	Total	Date
Female	Health and beauty	74.69	26.1415	548.9715	2019-01-05
Female	Electronic accessories	15.28	3.82	80.22	2019-03-08
Male	Home and lifestyle	46.33	16.2155	340.5255	2019-03-03
Male	Health and beauty	58.22	23.288	489.048	2019-01-27
Male	Sports and travel	86.31	30.2085	634.3785	2019-02-08
Male	Electronic accessories	85.39	29.8865	627.6165	2019-03-25
Female	Electronic accessories	68.84	20.652	433.692	2019-02-25
Female	Home and lifestyle	73.56	36.78	772.38	2019-02-24
Female	Health and beauty	36.26	3.626	76.146	2019-01-10
Female	Food and beverages	54.84	8.226	172.746	2019-02-20
Female	Fashion accessories	14.48	2.896	60.816	2019-02-06
Male	Electronic accessories	25.51	5.102	107.142	2019-03-09
Female	Electronic accessories	46.95	11.7375	246.4875	2019-02-12

After result:



IBM Cognos Analytics interface showing the same data grid after sorting 'Unit price' in ascending order. The left sidebar remains the same. The main grid shows data with columns: Product line, Unit price, Quantity, Tax 5%, Total, Date, and Time. The 'Unit price' column is highlighted, and the data rows are sorted in ascending order of unit price.

Product line	Unit price	Quantity	Tax 5%	Total	Date	Time
Health and beauty	10.08	7	3.528	74.088	2019-03-28	20:14:00
Food and beverages	10.13	7	3.5455	74.4555	2019-03-10	19:35:00
Health and beauty	10.16	5	2.54	53.34	2019-02-24	13:08:00
Sports and travel	10.17	1	0.5085	10.6785	2019-02-07	14:15:00
Fashion accessories	10.18	8	4.072	85.512	2019-03-30	12:51:00
Home and lifestyle	10.53	5	2.6325	55.2825	2019-01-30	14:43:00
Electronic accessories	10.56	8	4.224	88.704	2019-01-24	17:43:00
Electronic accessories	10.59	3	1.5885	33.3585	2019-03-12	13:52:00
Health and beauty	10.69	5	2.6725	56.1225	2019-03-26	11:07:00
Health and beauty	10.75	8	4.3	90.3	2019-03-15	14:38:00
Food and beverages	10.96	10	5.48	115.08	2019-02-02	20:48:00
Health and beauty	10.99	5	2.7475	57.6975	2019-01-23	10:18:00
Home and lifestyle	11.28	9	5.076	106.596	2019-03-17	11:55:00

## Splitting the date

**Split column - Date**

Select the items that will be used to create new columns.

Year: 2019, Month: 1, Day: 5, ☒ Include the day of the week

**Preview**

Date	Year	Month	Day	Day of the week
2019-01-05	2019	1	5	Saturday
2019-02-06	2019	2	6	Wednesday
2019-03-29	2019	3	29	Friday
2019-03-11	2019	3	11	Monday
2019-03-22	2019	3	22	Friday
2019-02-10	2019	2	10	Sunday
2019-03-27	2019	3	27	Wednesday
2019-03-12	2019	3	12	Tuesday
2019-02-14	2019	2	14	Thursday
2019-01-24	2019	1	24	Thursday

Cancel Next

## Splitting time

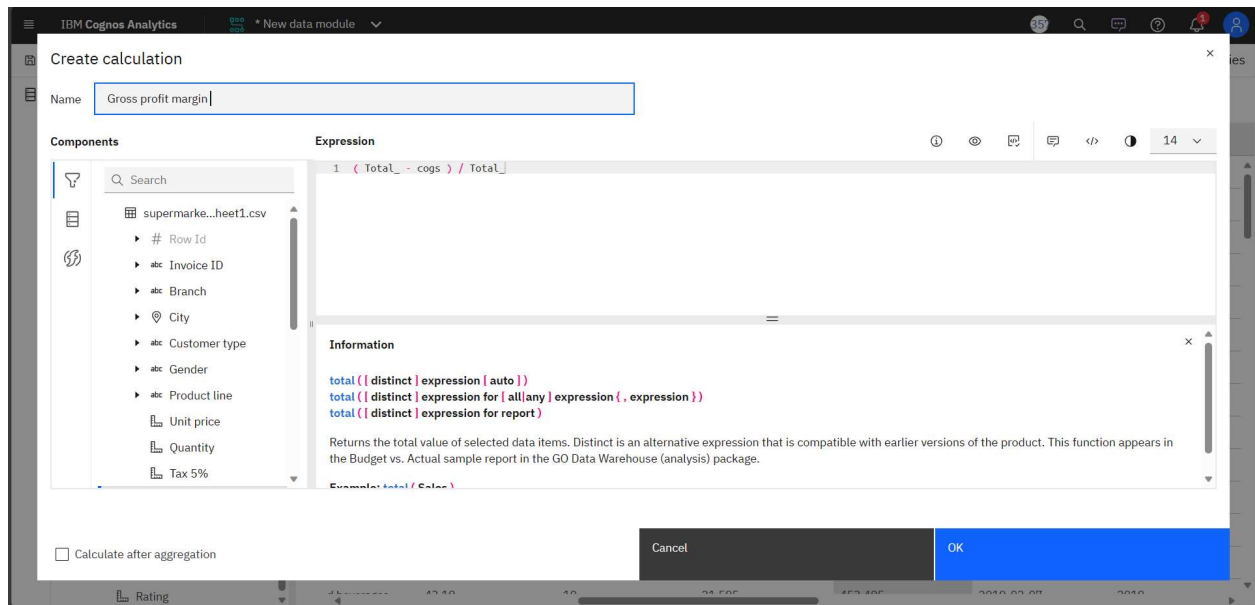
**Split column - Time**

Review new columns. You can change the column names, and deselect the columns that you don't need.

Time	<input checked="" type="checkbox"/> Hour	<input checked="" type="checkbox"/> Minute
13:23:00	13	23
10:25:00	10	25
15:30:00	15	30
11:24:00	11	24
19:20:00	19	20
15:36:00	15	36
10:12:00	10	12
13:24:00	13	24
18:45:00	18	45
10:11:00	10	11
19:01:00	19	1
11:28:00	11	28
15:55:00	15	55
20:36:00	20	36

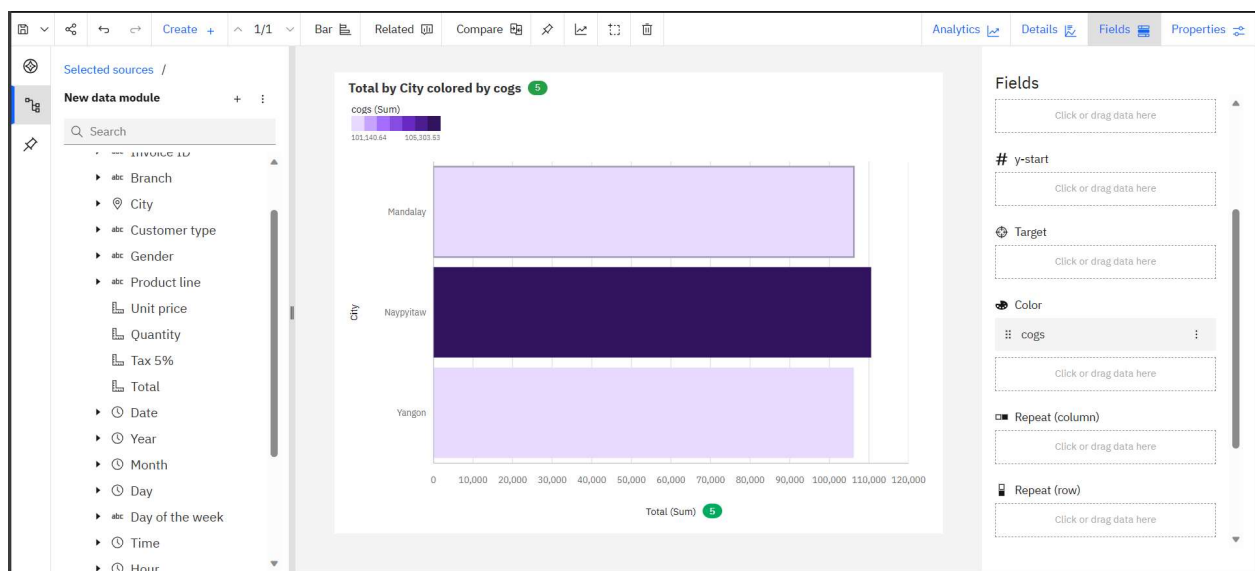
Previous Cancel OK

## Initializing new column



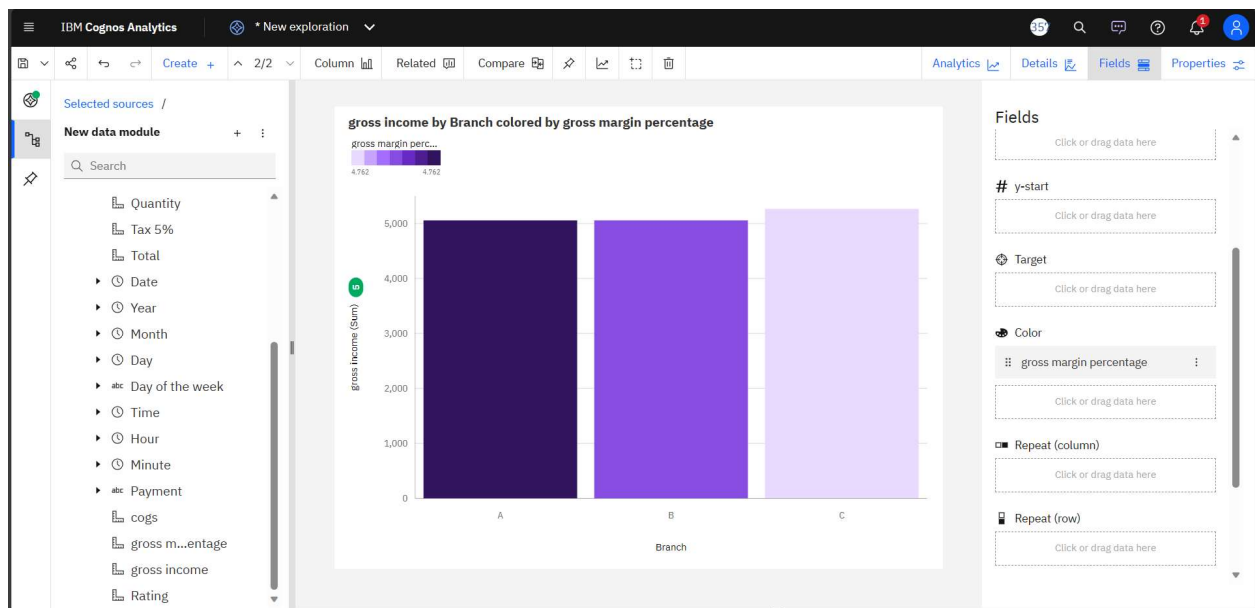
## Creating Visualizations

### Total revenue by each city colored by cogs - Bar Graph



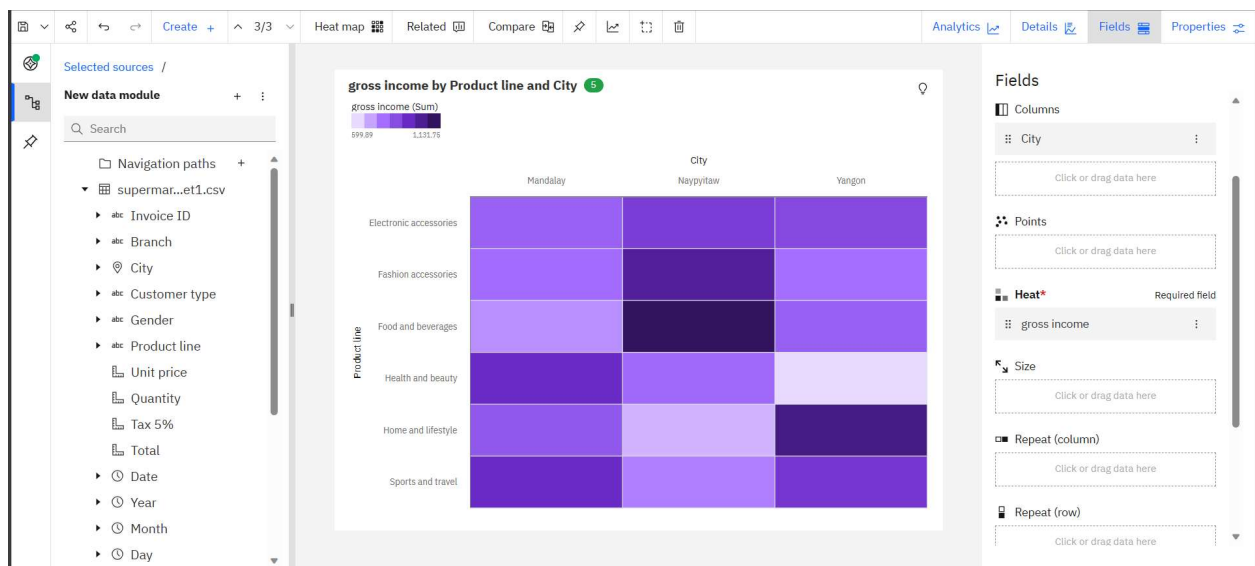
## Gross income by branch colored by gross margin percentage

### Column Graph

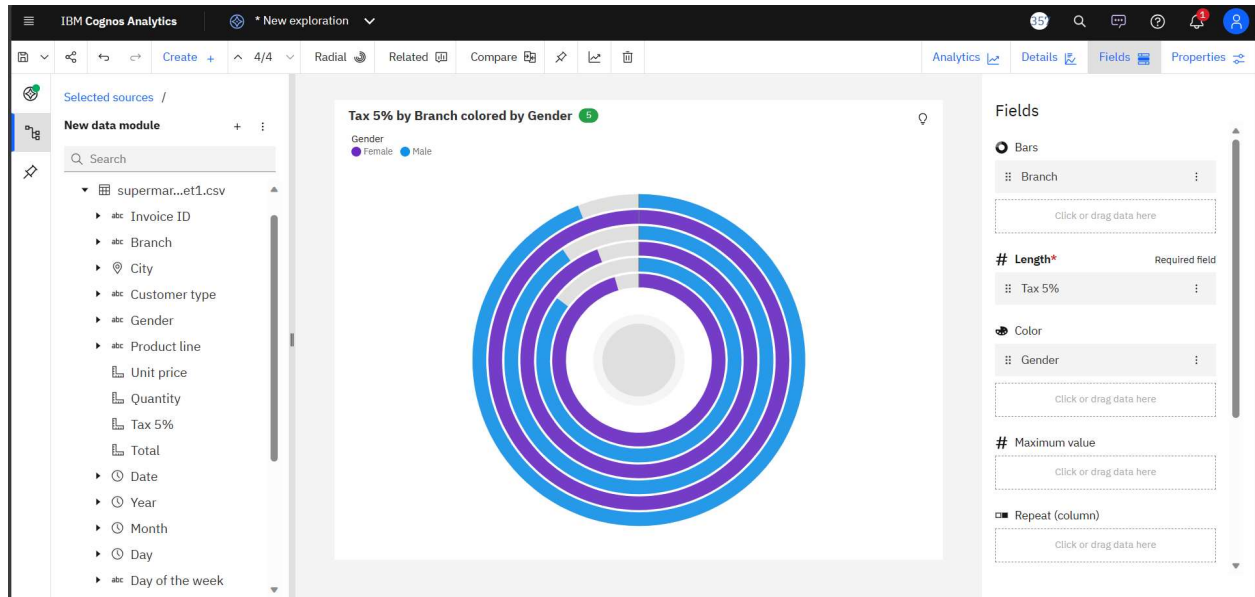


## Gross income by product line and city

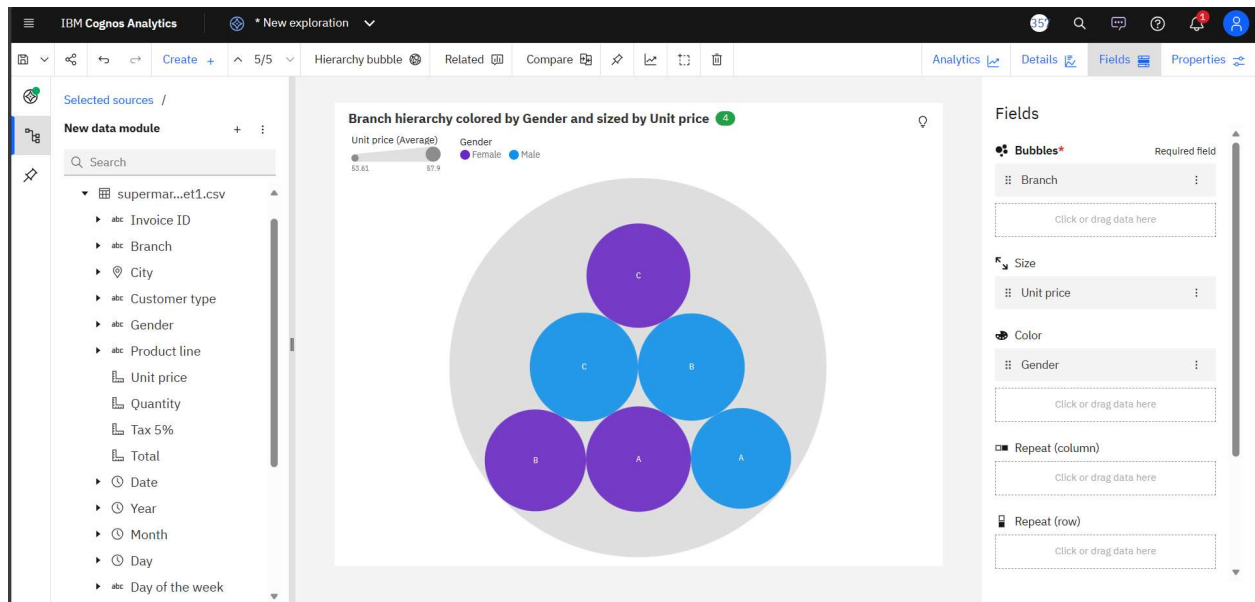
### Heat Map



## Tax by branch colored by gender Radial Graph

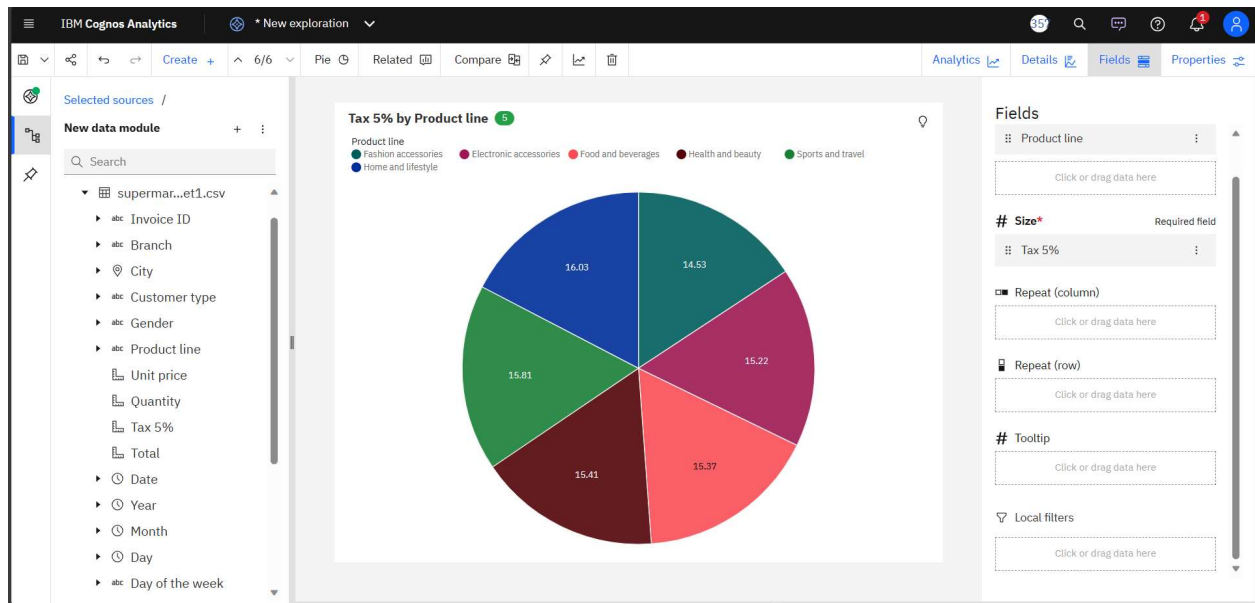


## Branch hierarchy colored by Gender and sized by Unit price Hierarchical bubble

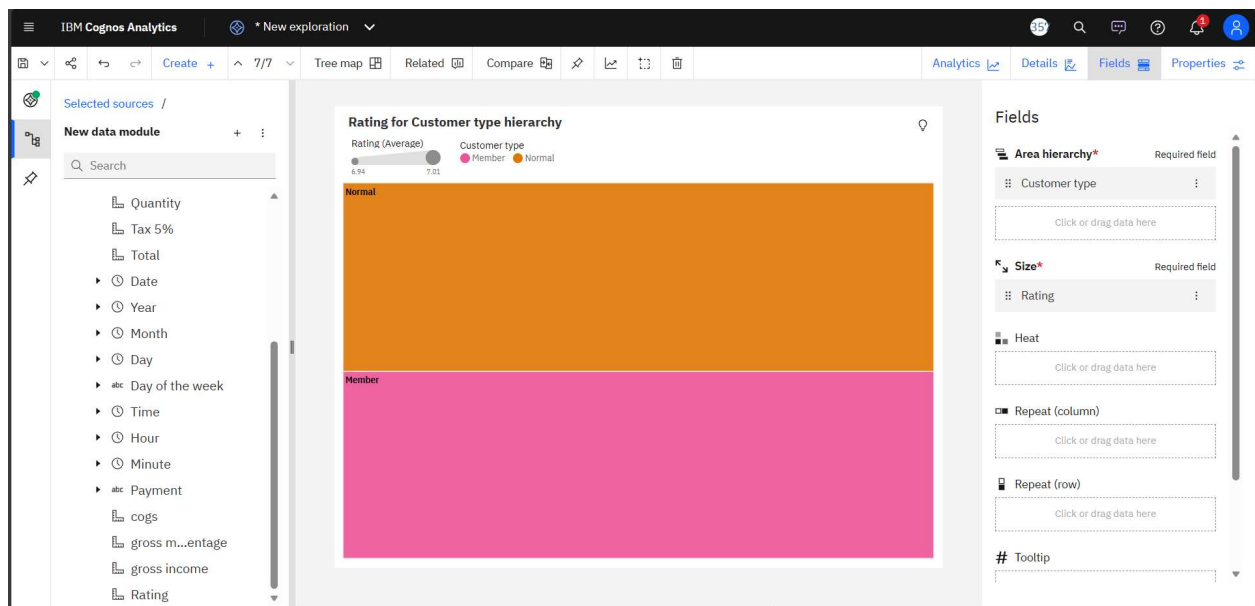




## Product line by tax in pie chart

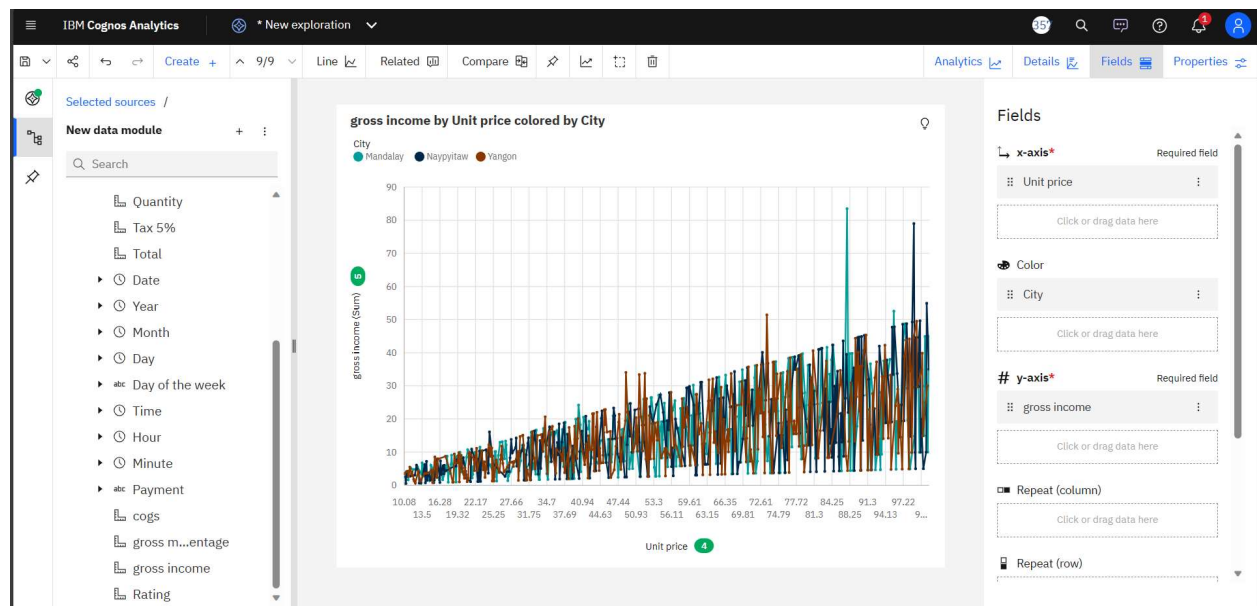


## Rating for customer type hierarchy Tree Map



## Gross income by unit price colored by city

### Line graph



## Quantity and tax with gross income colored by gender for branch

### Bubble Chart

