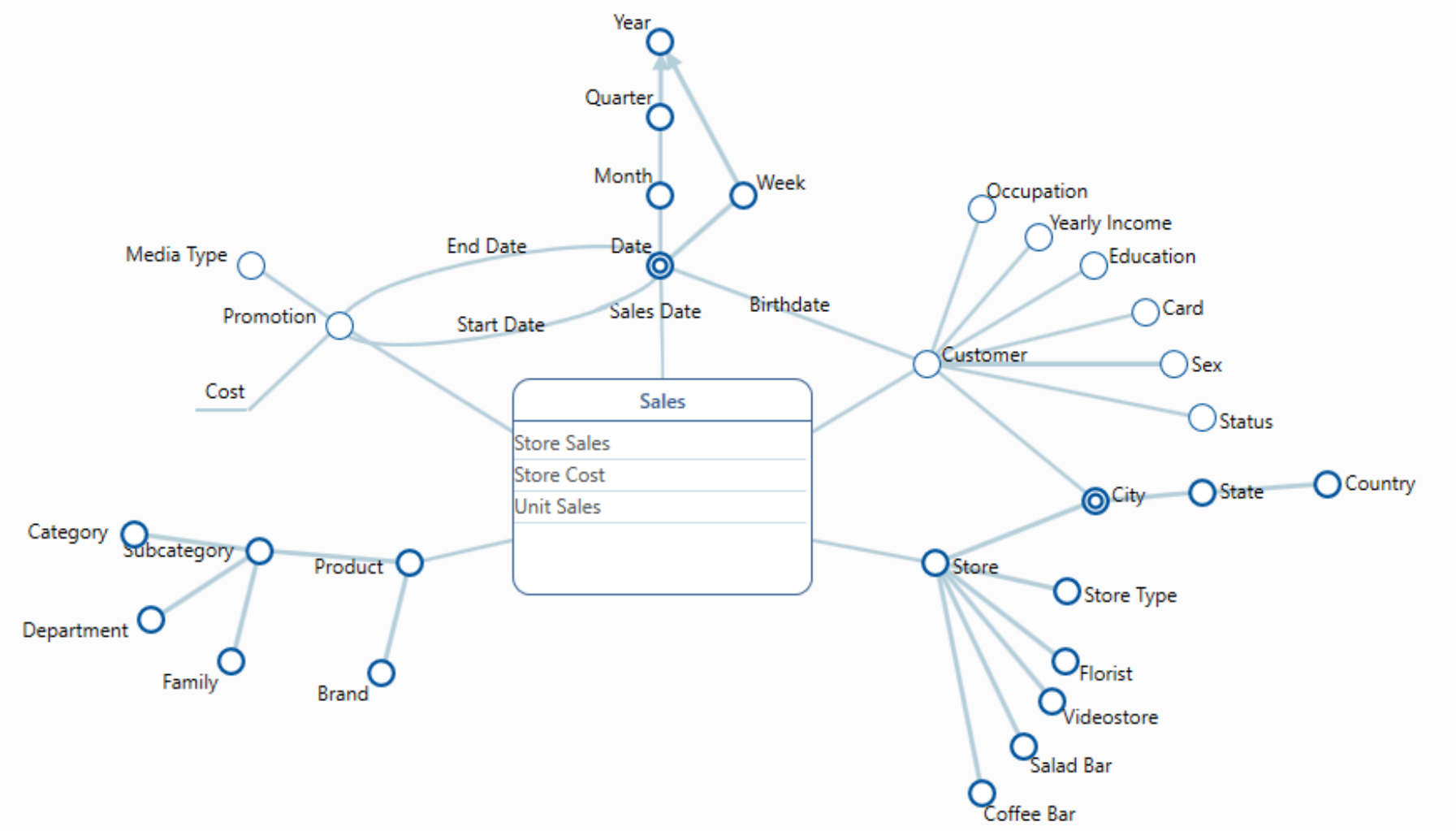


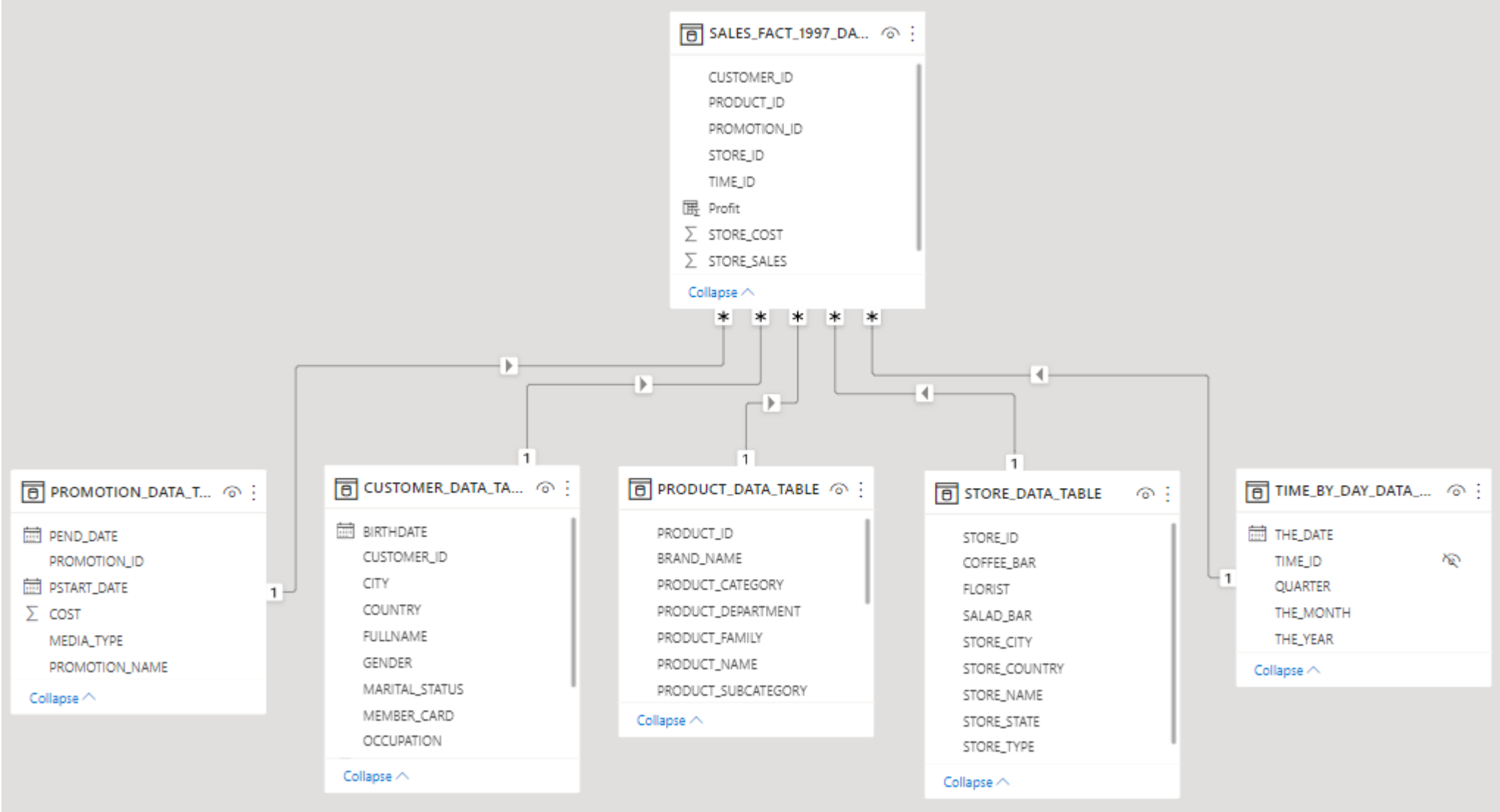
Dai dati alla business intelligence attraverso la trasformazione digitale

Ravenna, 09-10/02/2024

DFM – Foodmart



Database Structure – Foodmart (Sales)



Database Structure – Foodmart (Sales)

	PROMOTION_ID	PROMOTION_NAME	MEDIA_TYPE	COST	PSTART_DATE	PEND_DATE
1	532	Weekend Markdown	Bulk Mail	14997	1998-03-10 00:00:00	1998-03-13 00:00:00
2	533	Shelf Clearing Days	Sunday Paper, Radio, TV	10196	1998-03-24 00:00:00	1998-03-27 00:00:00
3	534	Mystery Sale	Daily Paper, Radio	6884	1998-04-07 00:00:00	1998-04-10 00:00:00
4	535	Mystery Sale	Street Handout	8527	1998-04-22 00:00:00	1998-04-23 00:00:00
5	536	Bye Bye Baby	Daily Paper, Radio, TV	5816	1998-05-06 00:00:00	1998-05-09 00:00:00
6	537	Big Time Discounts	Daily Paper, Radio, TV	14641	1998-05-21 00:00:00	1998-05-23 00:00:00
7	538	Saving Days	Sunday Paper, Radio, TV	6711	1998-06-04 00:00:00	1998-06-05 00:00:00
8	539	High Roller Savings	Sunday Paper, Radio, TV	9555	1998-06-17 00:00:00	1998-06-19 00:00:00
9	540	Mystery Sale	Cash Register Handout	8553	1998-07-02 00:00:00	1998-07-04 00:00:00
10	541	Price Winners	Sunday Paper, Radio	12251	1998-07-17 00:00:00	1998-07-18 00:00:00
11	542	Green Light Special	Street Handout	12358	1998-07-31 00:00:00	1998-08-01 00:00:00
12	543	Fantastic Discounts	Radio	13746	1998-08-12 00:00:00	1998-08-14 00:00:00
13	544	Saving Days	Product Attachment	8326	1998-08-28 00:00:00	1998-08-29 00:00:00
14	545	Two for One	Daily Paper	6887	1998-08-11 00:00:00	1998-08-14 00:00:00

PROMOTION_DATA_T...

- PEND_DATE
- PROMOTION_ID
- PSTART_DATE
- COST
- MEDIA_TYPE
- PROMOTION_NAME

[Collapse ^](#)

CUSTOMER_DATA_TA...

- BIRTHDATE
- CUSTOMER_ID
- CITY
- COUNTRY
- FULLNAME
- GENDER
- MARITAL_STATUS
- MEMBER_CARD
- OCCUPATION

[Collapse ^](#)

PRODUCT_DATA_TABLE

- PRODUCT_ID
- BRAND_NAME
- PRODUCT_CATEGORY
- PRODUCT_DEPARTMENT
- PRODUCT_FAMILY
- PRODUCT_NAME
- PRODUCT_SUBCATEGORY

[Collapse ^](#)

STORE_DATA_TABLE

- STORE_ID
- COFFEE_BAR
- FLORIST
- SALAD_BAR
- STORE_CITY
- STORE_COUNTRY
- STORE_NAME
- STORE_STATE
- STORE_TYPE

[Collapse ^](#)

TIME_BY_DAY_DATA...

- THE_DATE
- TIME_ID
- QUARTER
- THE_MONTH
- THE_YEAR

[Collapse ^](#)

Database Structure – Foodmart (Sales)

CUSTOMER_ID	CITY	STATE...	COUNTRY	BIRTHDATE	MARITAL...	YEARLY_INCOME	GENDER	MEMBER_CARD	OCCUPATION	FULLNAME	POPULATION	
1	107	Cliffside	BC	Canada	09-MAG-21	M	\$30K - \$50K	M	Bronze	Manual	Barney Velasquez	38000000
2	108	Shawnee	BC	Canada	13-MAG-31	M	\$90K - \$110K	M	Bronze	Management	Kenneth Dubois	38000000
3	109	Newton	BC	Canada	01-AGO-29	S	\$50K - \$70K	M	Bronze	Professional	Kate Maestas	38000000
4	110	West Covina	CA	USA	03-FEB-42	S	\$10K - \$30K	M	Normal	Manual	Josie Underwood	328000000
5	111	Palo Alto	CA	USA	21-OTT-65	M	\$70K - \$90K	F	Bronze	Management	Ramon Strain	328000000
6	112	Santa Monica	CA	USA	24-NOV-48	M	\$30K - \$50K	F	Bronze	Skilled Manual	Pat Azari	328000000
7	113	Sooke	BC	Canada	20-MAG-64	M	\$110K - \$130K	F	Bronze	Professional	Bob Dabit	38000000
8	114	Sooke	BC	Canada	16-APR-43	S	\$150K +	F	Silver	Professional	Marty Carmona	38000000
9	115	Altadena	CA	USA	09-GIU-77	S	\$90K - \$110K	F	Silver	Professional	Gina Saxton	328000000
10	116	Burnaby	BC	Canada	11-OTT-34	S	\$90K - \$110K	M	Bronze	Management	Juan McLaughlin	38000000
11	117	Albany	OR	USA	04-FEB-54	M	\$90K - \$110K	F	Bronze	Professional	Carol Eyster	328000000
12	118	Bremerton	WA	USA	04-OTT-13	S	\$50K - \$70K	M	Bronze	Professional	Nancy Henry	328000000
13	119	N. Vancouver	BC	Canada	26-LUG-57	S	\$10K - \$30K	M	Normal	Manual	Chris Barros	38000000
14	120	Milwaukie	OR	USA	28-GEN-57	M	\$30K - \$50K	F	Bronze	Skilled Manual	Christopher Groome	328000000
15	121	Santa Monica	CA	USA	17-FEB-12	S	\$30K - \$50K	F	Silver	Manual	Anna Hill	328000000

PROMOTION_DATA_T...

- PEND_DATE
- PROMOTION_ID
- PSTART_DATE
- COST
- MEDIA_TYPE
- PROMOTION_NAME

Collapse ^

CUSTOMER_DATA_TA...

- BIRTHDATE
- CUSTOMER_ID
- CITY
- COUNTRY
- FULLNAME
- GENDER
- MARITAL_STATUS
- MEMBER_CARD
- OCCUPATION

Collapse ^

PRODUCT_DATA_TABLE

- PRODUCT_ID
- BRAND_NAME
- PRODUCT_CATEGORY
- PRODUCT_DEPARTMENT
- PRODUCT_FAMILY
- PRODUCT_NAME
- PRODUCT_SUBCATEGORY

Collapse ^

STORE_DATA_TABLE

- STORE_ID
- COFFEE_BAR
- FLORIST
- SALAD_BAR
- STORE_CITY
- STORE_COUNTRY
- STORE_NAME
- STORE_STATE
- STORE_TYPE

Collapse ^

TIME_BY_DAY_DATA...

- THE_DATE
- TIME_ID
- QUARTER
- THE_MONTH
- THE_YEAR

Collapse ^

Database Structure – Foodmart (Sales)

	PRODUCT_ID	BRAND_NAME	PRODUCT_NAME	PRODUCT_SUBCATEGORY	PRODUCT_CATEGORY	PRODUCT_DEPARTMENT	PRODUCT_FAMILY
1	1	Washington	Washington Berry Juice	Juice	Pure Juice Beverages	Beverages	Drink
2	2	Washington	Washington Mango Drink	Flavored Drinks	Drinks	Beverages	Drink
3	3	Washington	Washington Strawberry Drink	Flavored Drinks	Drinks	Beverages	Drink
4	4	Washington	Washington Cream Soda	Soda	Carbonated Beverages	Beverages	Drink
5	5	Washington	Washington Diet Soda	Soda	Carbonated Beverages	Beverages	Drink
6	6	Washington	Washington Cola	Soda	Carbonated Beverages	Beverages	Drink
7	7	Washington	Washington Diet Cola	Soda	Carbonated Beverages	Beverages	Drink
8	8	Washington	Washington Orange Juice	Juice	Pure Juice Beverages	Beverages	Drink
9	9	Washington	Washington Cranberry Juice	Juice	Pure Juice Beverages	Beverages	Drink
10	10	Washington	Washington Apple Juice	Juice	Pure Juice Beverages	Beverages	Drink
11	11	Washington	Washington Apple Drink	Flavored Drinks	Drinks	Beverages	Drink
12	12	Jeffers	Jeffers Oatmeal	Cereal	Breakfast Foods	Breakfast Foods	Food
13	13	Jeffers	Jeffers Corn Puffs	Cereal	Breakfast Foods	Breakfast Foods	Food
14	14	Jeffers	Jeffers Wheat Puffs	Cereal	Breakfast Foods	Breakfast Foods	Food
15	15	Jeffers	Jeffers Grits	Cereal	Breakfast Foods	Breakfast Foods	Food

PROMOTION_DATA_T...

- PEND_DATE
- PROMOTION_ID
- PSTART_DATE
- COST
- MEDIA_TYPE
- PROMOTION_NAME

Collapse ^

CUSTOMER_DATA_TA...

- BIRTHDATE
- CUSTOMER_ID
- CITY
- COUNTRY
- FULLNAME
- GENDER
- MARITAL_STATUS
- MEMBER_CARD
- OCCUPATION

Collapse ^

PRODUCT_DATA_TABLE

- PRODUCT_ID
- BRAND_NAME
- PRODUCT_CATEGORY
- PRODUCT_DEPARTMENT
- PRODUCT_FAMILY
- PRODUCT_NAME
- PRODUCT_SUBCATEGORY

Collapse ^

STORE_DATA_TABLE

- STORE_ID
- COFFEE_BAR
- FLORIST
- SALAD_BAR
- STORE_CITY
- STORE_COUNTRY
- STORE_NAME
- STORE_STATE
- STORE_TYPE

Collapse ^

TIME_BY_DAY_DATA...

- THE_DATE
- TIME_ID
- QUARTER
- THE_MONTH
- THE_YEAR

Collapse ^

Database Structure – Foodmart (Sales)

STORE_ID	STORE_TYPE	STORE_NAME	STORE_CITY	STORE_STATE	STORE_COUNTRY	COFFEE_BAR	VIDEO_STORE	SALAD_BAR	FLORIST
1	1 Supermarket	Store 1	Acapulco	Guerrero	Mexico	0	0	0	0
2	2 Small Grocery	Store 2	Bellingham	WA	USA	1	0	0	0
3	3 Supermarket	Store 3	Bremerton	WA	USA	0	0	1	0
4	4 Gourmet Supermarket	Store 4	Camacho	Zacatecas	Mexico	1	0	1	1
5	5 Small Grocery	Store 5	Guadalajara	Jalisco	Mexico	1	0	0	0
6	6 Gourmet Supermarket	Store 6	Beverly Hills	CA	USA	1	1	1	1
7	7 Supermarket	Store 7	Los Angeles	CA	USA	0	0	0	1
8	8 Deluxe Supermarket	Store 8	Merida	Yucatan	Mexico	1	1	1	1
9	9 Mid-Size Grocery	Store 9	Mexico City	DF	Mexico	0	0	0	0
10	10 Supermarket	Store 10	Orizaba	Veracruz	Mexico	0	0	1	0
11	11 Supermarket	Store 11	Portland	OR	USA	0	0	0	0
12	12 Deluxe Supermarket	Store 12	Hidalgo	Zacatecas	Mexico	1	1	1	1
13	13 Deluxe Supermarket	Store 13	Salem	OR	USA	1	1	1	1
14	14 Small Grocery	Store 14	San Francisco	CA	USA	1	0	0	0
15	15 Supermarket	Store 15	Seattle	WA	USA	1	0	0	0
16	16 Supermarket	Store 16	Spokane	WA	USA	0	0	0	0

PROMOTION_DATA_T...

- PEND_DATE
- PROMOTION_ID
- PSTART_DATE
- COST
- MEDIA_TYPE
- PROMOTION_NAME

Collapse ^

CUSTOMER_DATA_TA...

- BIRTHDATE
- CUSTOMER_ID
- CITY
- COUNTRY
- FULLNAME
- GENDER
- MARITAL_STATUS
- MEMBER_CARD
- OCCUPATION

Collapse ^

PRODUCT_DATA_TABLE

- PRODUCT_ID
- BRAND_NAME
- PRODUCT_CATEGORY
- PRODUCT_DEPARTMENT
- PRODUCT_FAMILY
- PRODUCT_NAME
- PRODUCT_SUBCATEGORY

Collapse ^

STORE_DATA_TABLE

- STORE_ID
- COFFEE_BAR
- FLORIST
- SALAD_BAR
- STORE_CITY
- STORE_COUNTRY
- STORE_NAME
- STORE_STATE
- STORE_TYPE

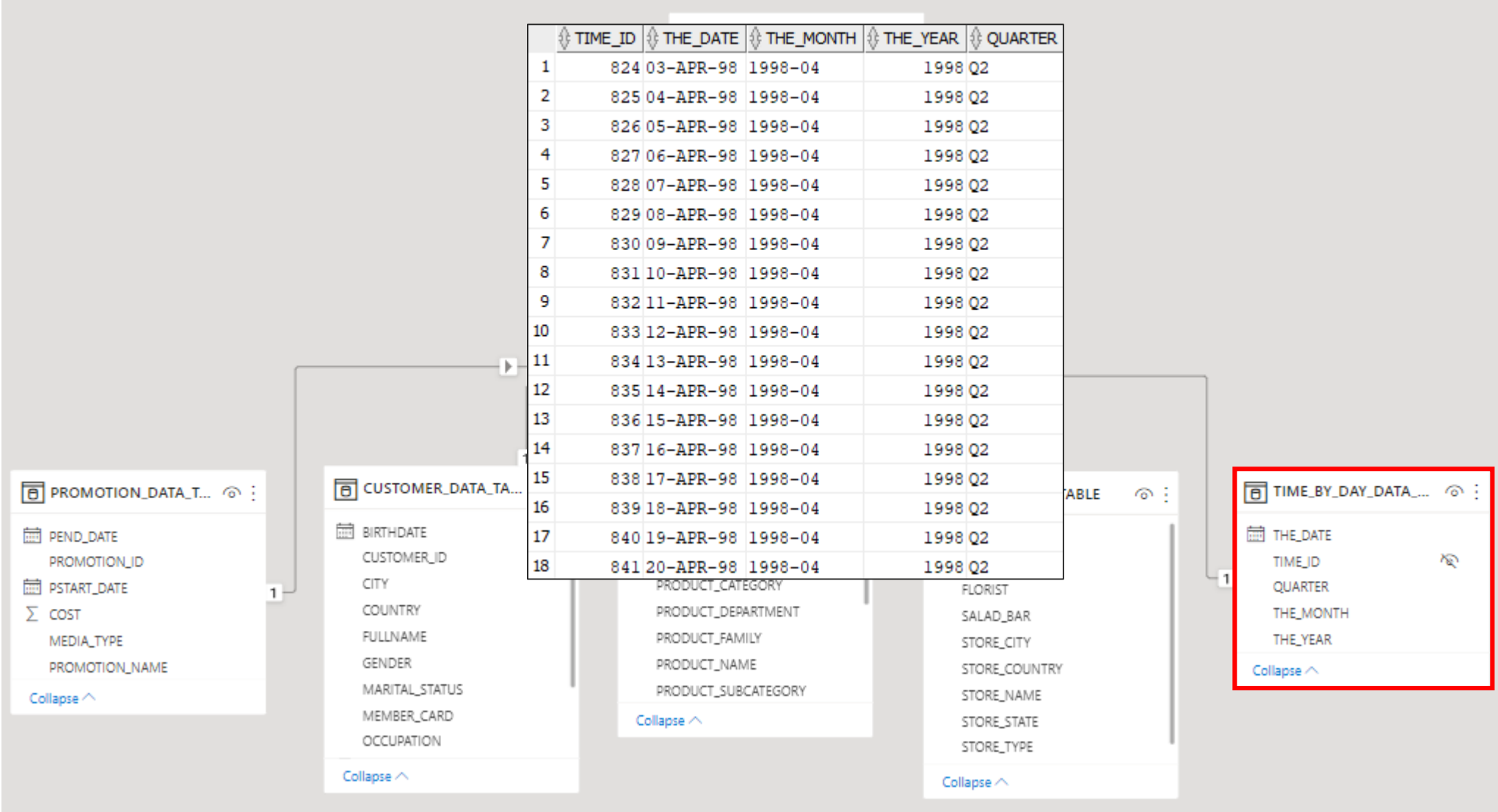
Collapse ^

TIME_BY_DAY_DATA...

- THE_DATE
- TIME_ID
- QUARTER
- THE_MONTH
- THE_YEAR

Collapse ^

Database Structure – Foodmart (Sales)



Database Structure – Foodmart (Sales)

The screenshot displays a data visualization interface with a central table and three floating panels. The central table has columns: PRODUCT_ID, TIME_ID, CUSTOMER_ID, PROMOTION_ID, STORE_ID, STORE_SALES, STORE_COST, and UNIT_SALES. The floating panels show field lists for different data sources.

SALES_FACT_1997_DA... Panel:

- CUSTOMER_ID
- PRODUCT_ID
- PROMOTION_ID
- STORE_ID
- TIME_ID
- Profit
- Σ STORE_COST
- Σ STORE_SALES

PROMOTION_DATA_T... Panel:

- PEND_DATE
- PROMOTION_ID
- PSTART_DATE
- Σ COST
- MEDIA_TYPE
- PROMOTION_NAME

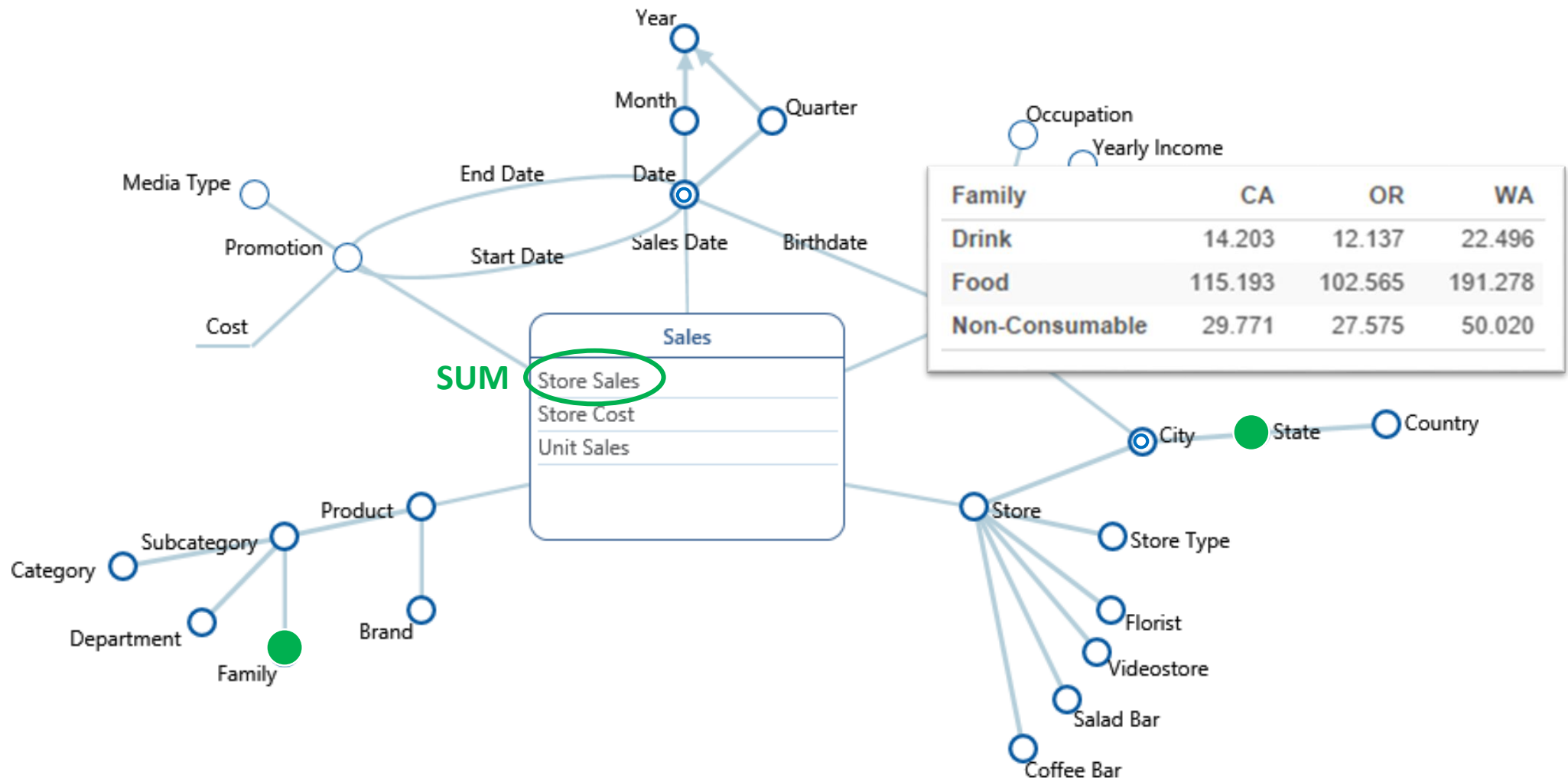
TIME_BY_DAY_DATA... Panel:

- THE_DATE
- TIME_ID
- QUARTER
- THE_MONTH
- THE_YEAR

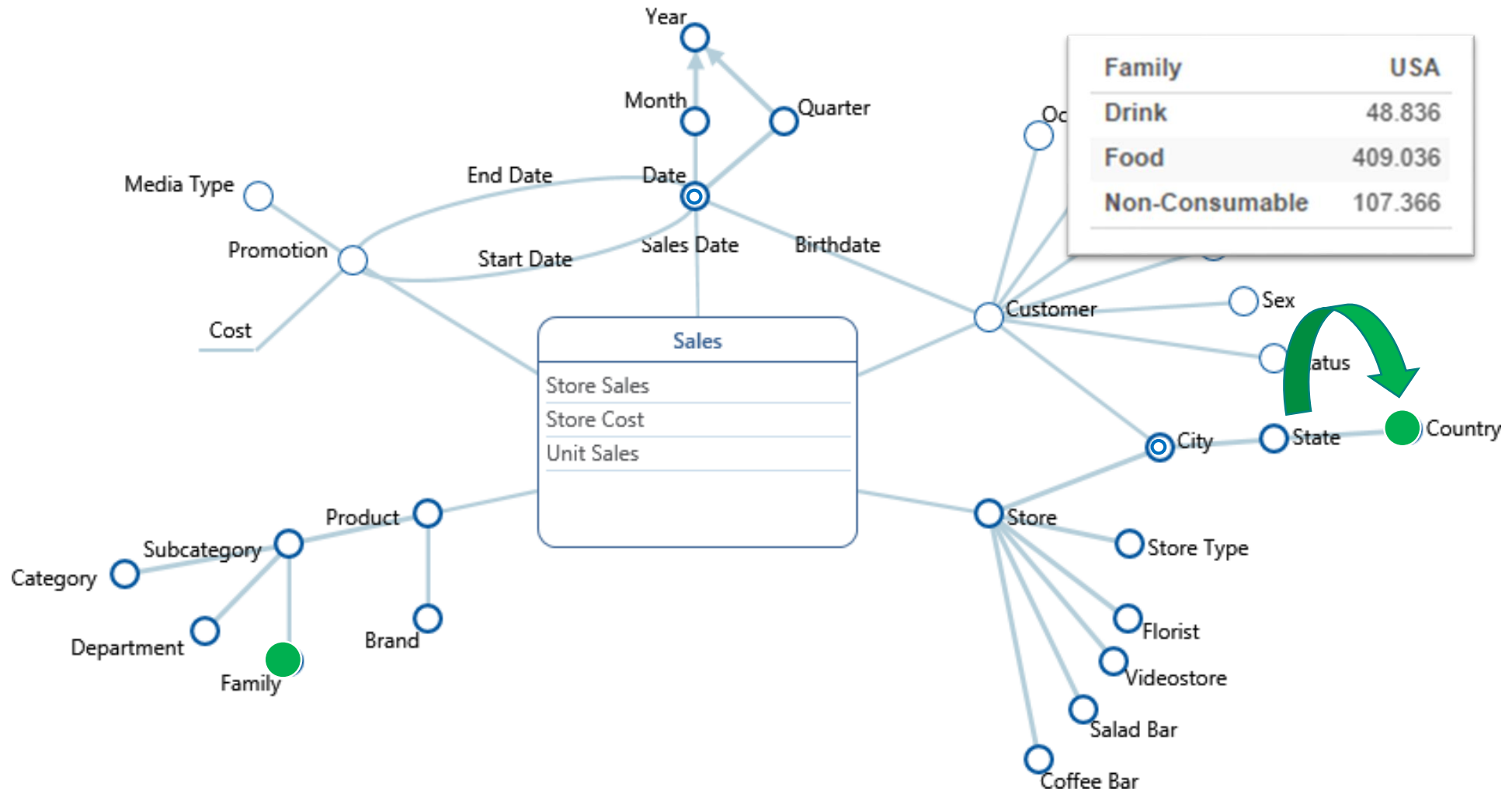
Central Table Data:

	PRODUCT_ID	TIME_ID	CUSTOMER_ID	PROMOTION_ID	STORE_ID	STORE_SALES	STORE_COST	UNIT_SALES
1	1	369	4728	501	7	11,4	3,99	4
2	1	377	9788	1547	13	8,55	4,0185	3
3	1	414	6666	34	17	8,55	4,1895	3
4	1	440	5313	413	24	8,55	3,762	3
5	1	463	916	302	7	11,4	4,902	4
6	1	474	4461	1839	11	8,55	2,9925	3
7	1	489	1312	162	3	8,55	3,6765	3
8	1	500	9169	1435	23	11,4	5,358	4
9	1	529	5607	501	6	11,4	4,902	4
10	1	534	456	828	15	11,4	4,332	4
11	1	570	923	30	15	8,55	2,736	3
12	1	574	9358	1097	15	8,55	4,275	3
13	1	576	7704	486	3	5,7	2,508	2
14	1	590	3441	131	3	8,55	3,42	3
15	1	594	6248	1860	24	11,4	3,876	4
16	1	596	5929	496	15	14,25	5,5575	5
17	1	616	1565	116	24	8,55	4,1895	3
18	1	617	638	1038	11	8,55	2,9925	3

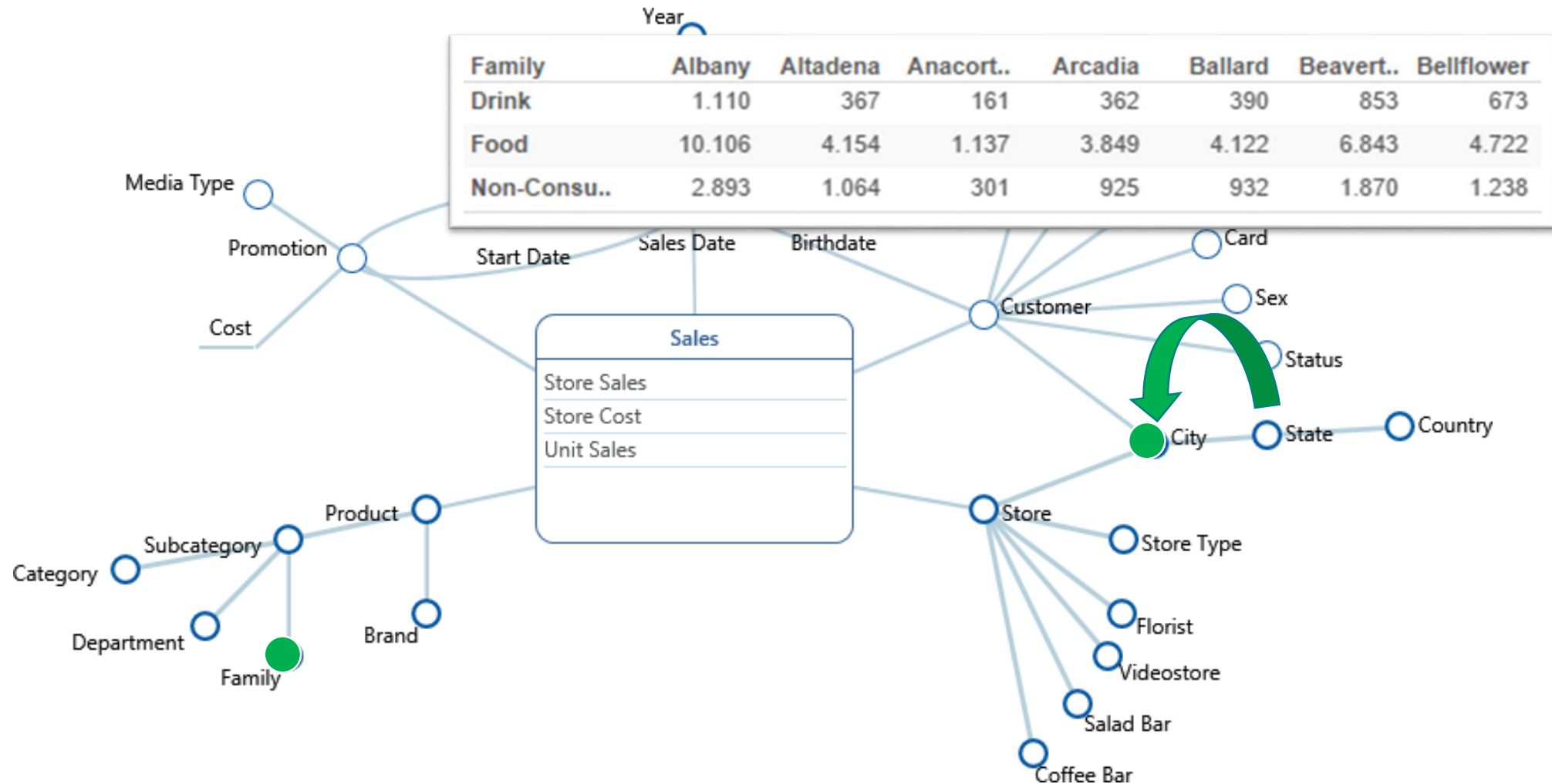
OLAP query



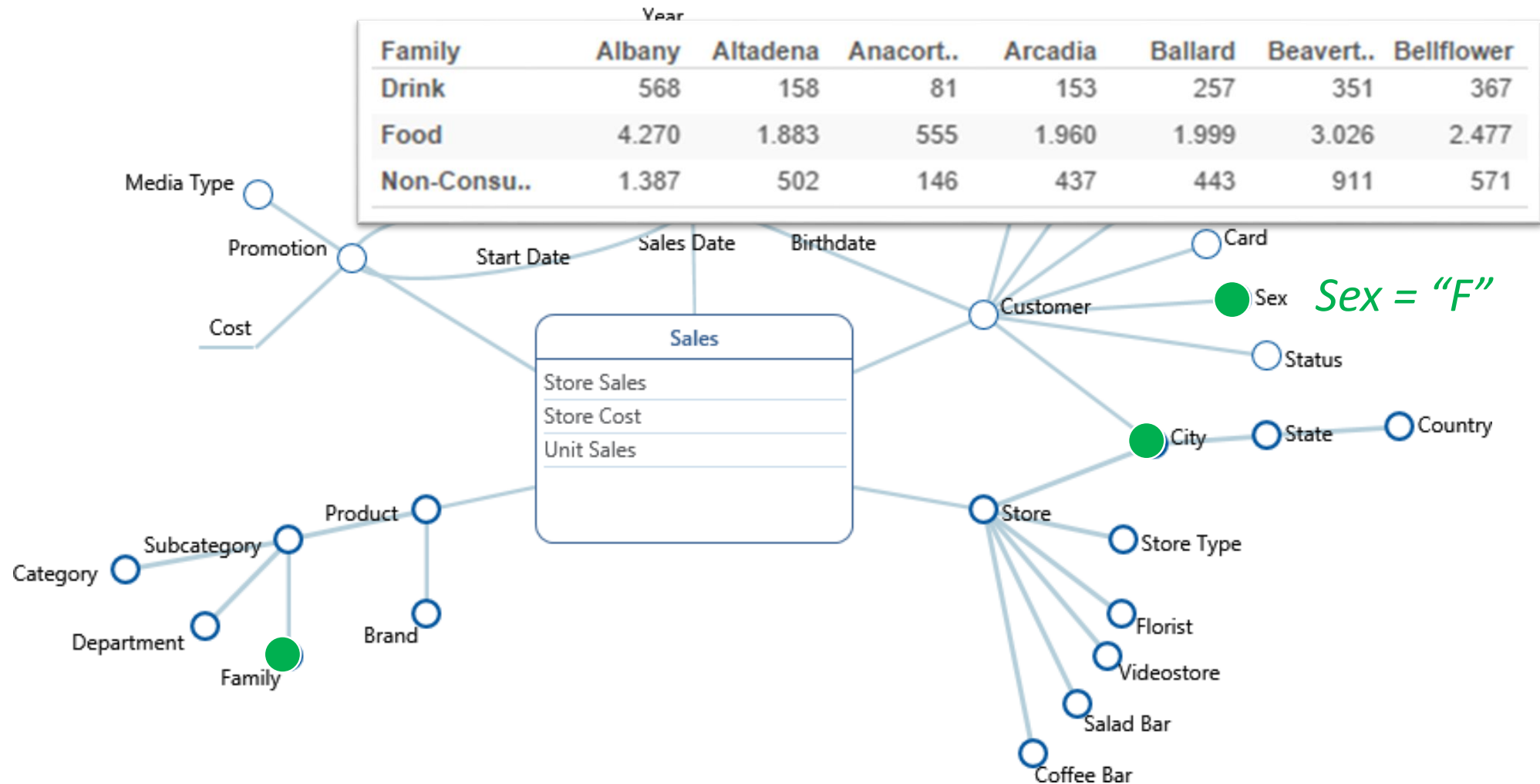
OLAP Operators: Roll-Up



OLAP Operators: Drill-down



OLAP Operators: Slice & Dice

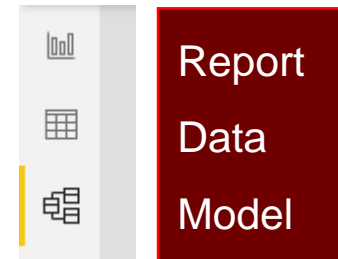


What is Power BI

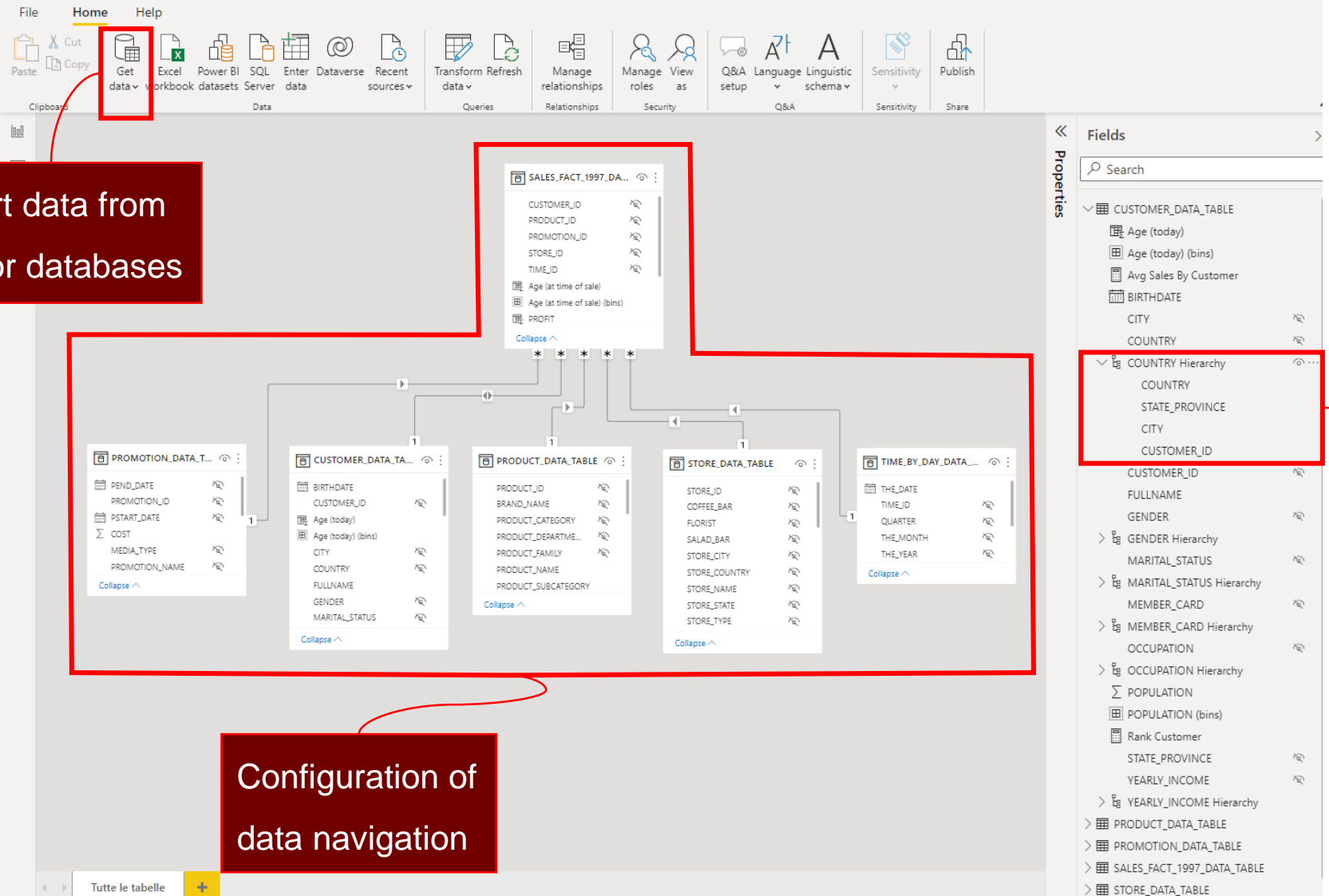
- Power BI is a commercial tool for data visualization and analysis that is part of the Microsoft Power Platform
- It provides a simple GUI to formulate queries on any kind of data source
 - Not exactly an OLAP tool
 - Besides advanced BI software (e.g., Oracle, SAP), similar commercial alternatives are Tableau and Qlik
 - Open-source alternatives
 - [Saiku](#)
 - [Apache Superset](#)
 - Power BI is easy(-ish) to use and provides appealing visualizations
 - Desktop version is free

Power BI main windows

- Report
 - Create charts and dashboards
- Data
 - Change data types
- Model
 - Setup connection to data source
 - Create hierarchies



Power BI main windows: Model



Import data from files or databases

Hierarchy

Configuration of data navigation

Power BI main windows: Data

The screenshot displays the Power BI Desktop interface with the Data view selected. The ribbon shows 'Table tools' and 'Column tools'. The 'Table tools' ribbon has 'Name' (BIRTHDATE), 'Data type' (Date), and 'Format' (mercoledì 14 marz...). The 'Column tools' ribbon has 'Summarization' (Don't summarize) and 'Data category' (Uncategorized). A table with columns CUSTOMER_ID, CITY, STATE_PROVINCE, COUNTRY, BIRTHDATE, MARITAL_STATUS, YEARLY_INCOME, GENDER, MEMBER_CARD, OCCUPATION, FULLNAME, and POPULATION is shown. The BIRTHDATE column is highlighted in yellow. Two red callout boxes are present: one pointing to the 'Data type' dropdown with the text 'Change data type (and format)', and another pointing to the 'Summarization' dropdown with the text 'Change default aggregation function'. The 'Fields' pane on the right shows the data model structure.

CUSTOMER_ID	CITY	STATE_PROVINCE	COUNTRY	BIRTHDATE	MARITAL_STATUS	YEARLY_INCOME	GENDER	MEMBER_CARD	OCCUPATION	FULLNAME	POPULATION
193	Los Angeles	CA	USA	martedì 7 aprile 1942	S	\$30K - \$50K	M	Bronze	Skilled Manual	Jeannette Eldridge	32800
44	Bellflower	CA	USA	domenica 20 maggio 1917	S	\$30K - \$50K	M	Bronze	Skilled Manual	Sandra Brunner	32800
529	Chula Vista	CA	USA	giovedì 15 febbraio 1951	S	\$30K - \$50K	M	Bronze	Skilled Manual	Kyley Arbelaez	32800
			USA	mercoledì 21 marzo 1945	S	\$30K - \$50K	M	Bronze	Skilled Manual	William Richter	32800
			USA	martedì 4 luglio 1967	S	\$30K - \$50K	M	Bronze	Skilled Manual	Helen Meyer	32800
			USA	lunedì 11 maggio 1959	S	\$30K - \$50K	M	Bronze	Skilled Manual	Dianne Slattengren	32800
			USA	mercoledì 26 novembre 1919	S	\$30K - \$50K	M	Bronze	Skilled Manual	Wanda Vernon	32800
			USA	lunedì 3 agosto 1970	S	\$30K - \$50K	M	Bronze	Skilled Manual	Sharon Lynn	32800
			USA	mercoledì 10 ottobre 1962	S	\$30K - \$50K	M	Bronze	Skilled Manual	Lillian Heckman	32800
			USA	sabato 20 ottobre 1917	S	\$30K - \$50K	M	Bronze	Skilled Manual	Melvin Carreras	32800
			USA	mercoledì 10 maggio 1911	S	\$30K - \$50K	M	Bronze	Skilled Manual	Audrey Topping	32800
1274	West Covina	CA	USA	lunedì 23 marzo 1959	S	\$30K - \$50K	M	Bronze	Skilled Manual	French Wilson	32800
984	Colma	CA	USA	venerdì 7 marzo 1969	S	\$30K - \$50K	M	Bronze	Skilled Manual	Natalie Barber	32800
1006	Palo Alto	CA	USA	lunedì 4 aprile 1932	S	\$30K - \$50K	M	Bronze	Skilled Manual	Margaret Lewis	32800
1123	San Diego	CA	USA	giovedì 10 maggio 1973	S	\$30K - \$50K	M	Bronze	Skilled Manual	Richard Yoshimura	32800
1127	Burbank	CA	USA	lunedì 8 luglio 1957	S	\$30K - \$50K	M	Bronze	Skilled Manual	Tony Irvin	32800
1207	Richmond	CA	USA	domenica 19 novembre 1961	S	\$30K - \$50K	M	Bronze	Skilled Manual	Tamar Hubbard	32800
1218	Chula Vista	CA	USA	lunedì 10 marzo 1952	S	\$30K - \$50K	M	Bronze	Skilled Manual	Kadija Proctor	32800
1616	Chula Vista	CA	USA	lunedì 10 luglio 1944	S	\$30K - \$50K	M	Bronze	Skilled Manual	Shirley Colvin	32800
1521	Novato	CA	USA	venerdì 5 marzo 1937	S	\$30K - \$50K	M	Bronze	Skilled Manual	Carla Stevenson	32800
1522	Novato	CA	USA	sabato 18 marzo 1911	S	\$30K - \$50K	M	Bronze	Skilled Manual	Bonnie Afana	32800
1549	Newport Beach	CA	USA	martedì 11 ottobre 1955	S	\$30K - \$50K	M	Bronze	Skilled Manual	April McLallen	32800
1826	West Covina	CA	USA	sabato 11 settembre 1920	S	\$30K - \$50K	M	Bronze	Skilled Manual	Peter Burr	32800
2521	Arcadia	CA	USA	martedì 9 febbraio 1937	S	\$30K - \$50K	M	Bronze	Skilled Manual	William Jones	32800
2146	Oakland	CA	USA	giovedì 8 marzo 1951	S	\$30K - \$50K	M	Bronze	Skilled Manual	Dale Ucti	32800
2565	Palo Alto	CA	USA	venerdì 3 luglio 1942	S	\$30K - \$50K	M	Bronze	Skilled Manual	Gonzalo Piant	32800
2592	Pomona	CA	USA	domenica 3 marzo 1946	S	\$30K - \$50K	M	Bronze	Skilled Manual	Valerie Thomas	32800
3188	Lakewood	CA	USA	domenica 14 febbraio 1937	S	\$30K - \$50K	M	Bronze	Skilled Manual	Christopher Young	32800
3285	Bellflower	CA	USA	venerdì 14 aprile 1916	S	\$30K - \$50K	M	Bronze	Skilled Manual	Dorothy Lay	32800
4008	Palo Alto	CA	USA	giovedì 3 ottobre 1929	S	\$30K - \$50K	M	Bronze	Skilled Manual	Vandella Bancroft	32800
4077	Los Angeles	CA	USA	lunedì 3 novembre 1930	S	\$30K - \$50K	M	Bronze	Skilled Manual	Hermina Lay	32800
3844	Woodland Hills	CA	USA	mercoledì 26 giugno 1935	S	\$30K - \$50K	M	Bronze	Skilled Manual	Vera Crosby	32800
3936	Beverly Hills	CA	USA	giovedì 22 agosto 1968	S	\$30K - \$50K	M	Bronze	Skilled Manual	Judy Wood	32800
3939	Torrance	CA	USA	lunedì 12 marzo 1979	S	\$30K - \$50K	M	Bronze	Skilled Manual	Brandi Wilson	32800
4115	Newport Beach	CA	USA	sabato 15 luglio 1911	S	\$30K - \$50K	M	Bronze	Skilled Manual	Brandi Wilson	32800
4184	Daly City	CA	USA		S	\$30K - \$50K	M	Bronze	Skilled Manual	Brandi Wilson	32800

Power BI main windows: Report

Foodmart \$565K
STORE_SALES

STORE_SALES by STORE_TYPE

STORE_TYPE	STORE_SALES	Percentage
Supermarket	\$319K	56.47%
Deluxe Supermarket	\$162K	28.68%
Gourmet Supermarket	\$46K	8.09%
Mid-Size Grocery	\$24K	4.3%
Small Grocery	\$16K	2.86%

STORE_SALES by STATE_PROVINCE

STATE_PROVINCE	STORE_SALES
WA	\$0.2M
CA	\$0.15M
OR	\$0.1M

STORE_SALES by THE_DATE and PRODUCT_FAMILY

PRODUCT_FAMILY	STORE_SALES
Vegetables	\$71.383
Snack Foods	\$67.504
Dairy	\$37.567
Meat	\$30.575
Fruit	\$29.131
Jams and Jellies	\$23.224
Bread	\$16.455
Electrical	\$16.165
Canned Soup	\$15.966
Breakfast Foods	\$15.555

Visualizations

Fields

Start by selecting a visualization

Then fill in the properties (that depend on the chosen visualization) by drag&dropping fields from the right panel

Filters

Visualizations

Fields

X-axis

THE_YEAR Hierarchy

THE_DATE

Y-axis

STORE_SALES

Secondary y-axis

Add data fields here

Legend

PRODUCT_FAMILY

Small multiples

Add data fields here

Tooltips

Add data fields here

Power BI main windows: Report

The screenshot displays the Power BI Report View interface. At the top, there is a ribbon with tabs: File, Home, Insert, Modeling, View, Help, Format, Data / Drill, Table tools, and Column tools. Below the ribbon are several toolbars for Name, Format, Summarization, Data category, Sort, Data groups, Manage relationships, and New column. The main area contains three visualizations: a pie chart titled 'Foodmart \$565K STORE_SALES by STORE_TYPE', a bar chart titled 'STORE_SALES by STATE_PROVINCE', and a line chart titled 'STORE_SALES by THE_DATE and PRODUCT_FAMILY'. On the right side, there are three panes: Filters, Visualizations, and Fields. The Filters pane shows filters for 'PRODUCT_FAMILY is (All)', 'STORE_SALES is (All)', and 'THE_DATE is (All)'. The Visualizations pane shows various chart types and a 'Build visual' section. The Fields pane shows a list of data sources and fields, with a context menu open over the 'PRODUCT_FAMILY' field. The context menu includes options like 'Check', 'Create hierarchy', 'Add to hierarchy', 'New measure', 'New column', 'New quick measure', 'Rename', 'Delete from model', 'Hide', 'View hidden', 'Unhide all', 'Collapse all', 'Expand all', 'New group', 'Add to filters', and 'Add to drill through'. A red box highlights the context menu, and a red arrow points from a text box to it. Another red box highlights the Filters pane, and a red arrow points from a text box to it. A third red box highlights the 'Add data fields here' button in the Filters pane, and a red arrow points from a text box to it.

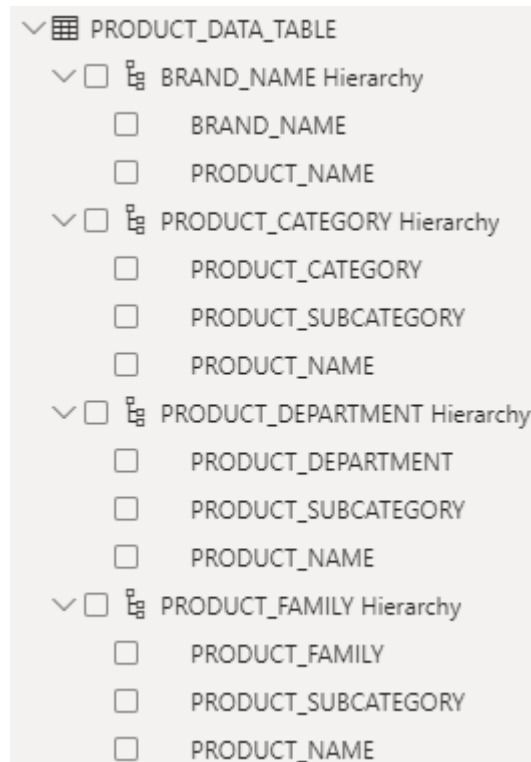
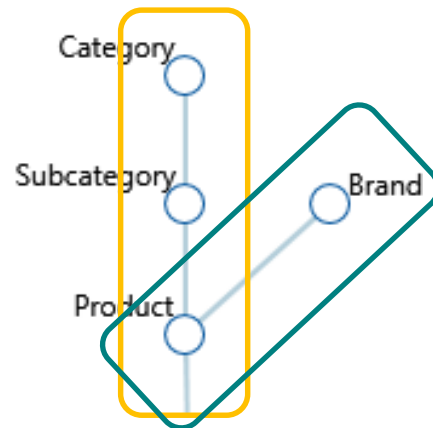
Filters can be set here

Right-click a field to access some functionalities, e.g.:

- Create hierarchy
- New measure
- New column
- New group

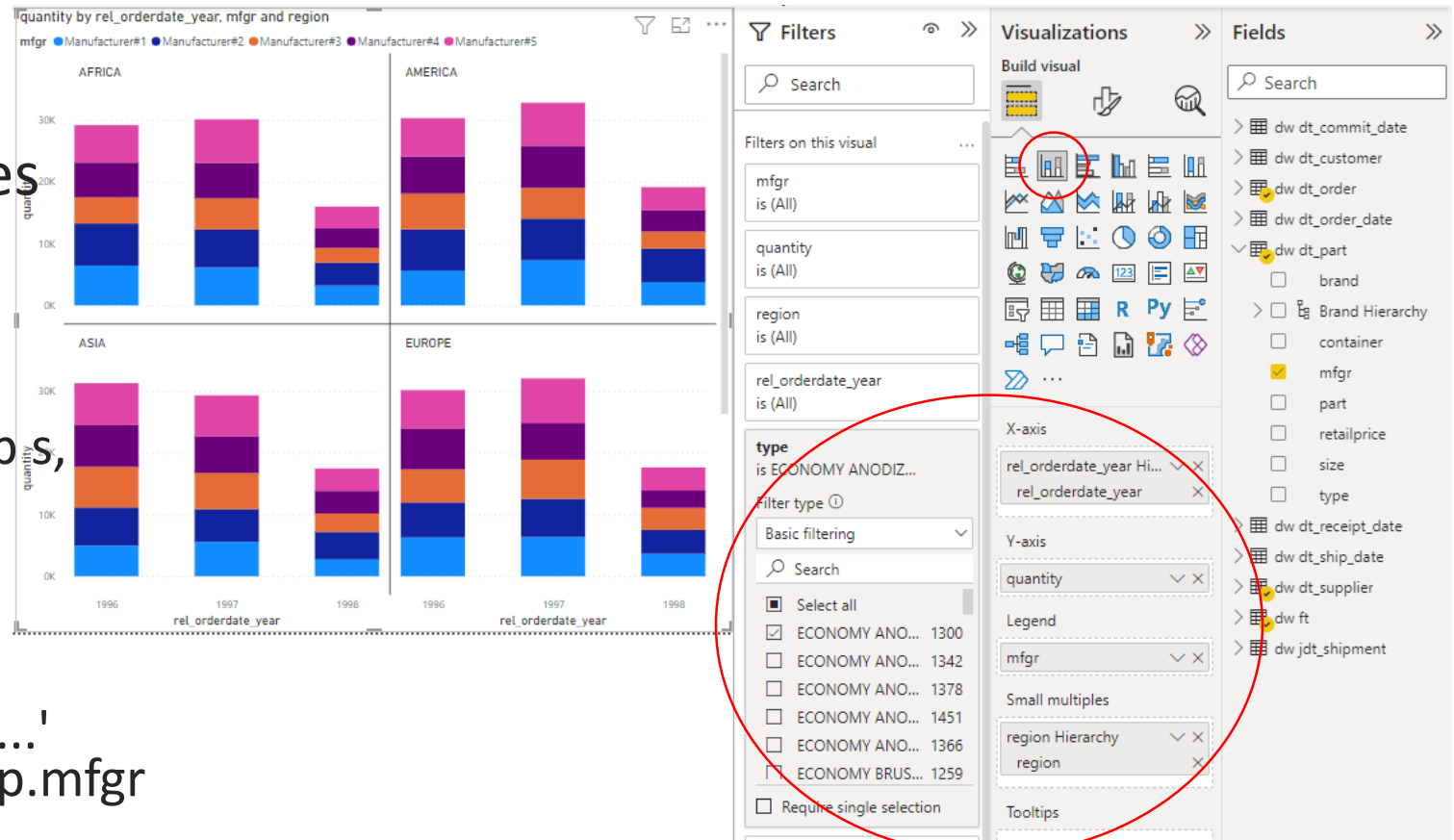
Hierarchies

- Each path from leaves to root becomes a different hierarchy; shared attributes are duplicated



OLAP basics

- Queries are formulated by
 - 1) choosing a visualization
 - 2) drag&dropping attributes into the visualization properties
- ```
SELECT d.year, s.region,
 p.mfgr, sum(ft.quantity)
FROM ft, dt_part p, dt_supp s,
 dt_order o, dt_date d
WHERE ft.idpart = p.id
 AND ft.idsupp = s.id
 AND ft.idorder = o.id
 AND o.iddate = d.id
 AND p.type = 'ECONOMY ...'
GROUP BY d.year, s.region, p.mfgr
```



# OLAP basics

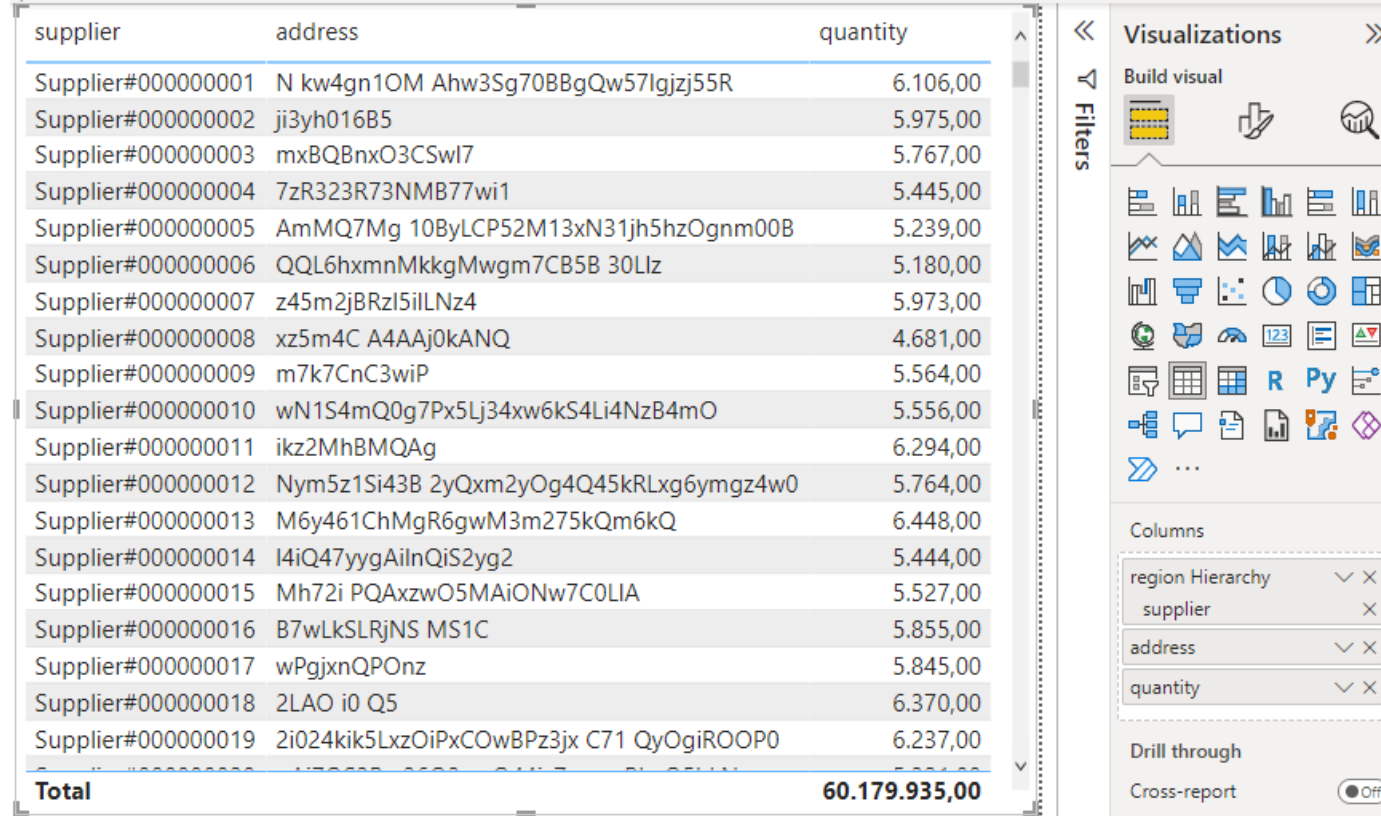
- The same query can be issued in different ways, where only the visualization changes (but the data is the same)
  - `SELECT d.year, s.region, p.mfgr, sum(ft.quantity) FROM ft, dt_part p, dt_supp s, dt_order o, dt_date d WHERE ft.idpart = p.id AND ft.idsupp = s.id AND ft.idorder = o.id AND o.iddate = d.id AND p.type = 'ECONOMY ...' GROUP BY d.year, s.region, p.mfgr`

The screenshot displays a data visualization tool interface. On the left, a table shows sales data grouped by region (AMERICA, EUROPE, ASIA, MIDDLE EAST, AFRICA) and manufacturer, with columns for years 1996, 1997, 1998, and a total. On the right, a configuration panel for a pivot table visualization is shown. The 'Filters' section includes 'mfgr is (All)', 'quantity is (All)', 'region is (All)', and 'rel\_orderdate\_year is (All)'. The 'Rows' section has 'region Hierarchy' and 'mfgr'. The 'Columns' section has 'rel\_orderdate\_year Hi...' and 'rel\_orderdate\_year'. The 'Values' section has 'quantity'. A red circle highlights the 'Basic filtering' section and the 'Rows' and 'Columns' sections.

| region             | 1996              | 1997              | 1998             |                   |
|--------------------|-------------------|-------------------|------------------|-------------------|
| <b>AMERICA</b>     | <b>30.289,00</b>  | <b>32.789,00</b>  | <b>19.125,00</b> | <b>82.203,00</b>  |
| Manufacturer#2     | 6.675,00          | 6.723,00          | 5.422,00         | 18.820,00         |
| Manufacturer#5     | 6.190,00          | 6.981,00          | 3.721,00         | 16.892,00         |
| Manufacturer#1     | 5.620,00          | 7.309,00          | 3.761,00         | 16.690,00         |
| Manufacturer#4     | 6.012,00          | 6.824,00          | 3.451,00         | 16.287,00         |
| Manufacturer#3     | 5.792,00          | 4.952,00          | 2.770,00         | 13.514,00         |
| <b>EUROPE</b>      | <b>30.141,00</b>  | <b>32.025,00</b>  | <b>17.613,00</b> | <b>79.779,00</b>  |
| Manufacturer#5     | 6.201,00          | 7.201,00          | 3.726,00         | 17.128,00         |
| Manufacturer#1     | 6.289,00          | 6.357,00          | 3.637,00         | 16.283,00         |
| Manufacturer#2     | 5.674,00          | 6.199,00          | 3.865,00         | 15.738,00         |
| Manufacturer#4     | 6.633,00          | 5.949,00          | 2.807,00         | 15.389,00         |
| Manufacturer#3     | 5.344,00          | 6.319,00          | 3.578,00         | 15.241,00         |
| <b>ASIA</b>        | <b>31.273,00</b>  | <b>29.277,00</b>  | <b>17.456,00</b> | <b>78.006,00</b>  |
| Manufacturer#5     | 6.811,00          | 6.606,00          | 3.613,00         | 17.030,00         |
| Manufacturer#4     | 6.735,00          | 5.947,00          | 3.681,00         | 16.363,00         |
| Manufacturer#2     | 6.193,00          | 5.291,00          | 4.427,00         | 15.911,00         |
| Manufacturer#3     | 6.578,00          | 5.851,00          | 2.972,00         | 15.401,00         |
| Manufacturer#1     | 4.956,00          | 5.582,00          | 2.763,00         | 13.301,00         |
| <b>MIDDLE EAST</b> | <b>30.819,00</b>  | <b>31.645,00</b>  | <b>15.356,00</b> | <b>77.820,00</b>  |
| Manufacturer#5     | 6.623,00          | 7.158,00          | 3.624,00         | 17.405,00         |
| Manufacturer#2     | 6.546,00          | 7.289,00          | 3.355,00         | 17.190,00         |
| Manufacturer#1     | 5.963,00          | 6.223,00          | 3.140,00         | 15.326,00         |
| Manufacturer#4     | 6.078,00          | 5.756,00          | 2.748,00         | 14.582,00         |
| Manufacturer#3     | 5.609,00          | 5.219,00          | 2.489,00         | 13.317,00         |
| <b>AFRICA</b>      | <b>29.142,00</b>  | <b>30.098,00</b>  | <b>15.960,00</b> | <b>75.200,00</b>  |
| Manufacturer#2     | 6.795,00          | 6.115,00          | 3.717,00         | 16.627,00         |
| Manufacturer#5     | 6.026,00          | 7.050,00          | 3.508,00         | 16.584,00         |
| Manufacturer#1     | 6.441,00          | 6.140,00          | 3.198,00         | 15.779,00         |
| Manufacturer#4     | 5.650,00          | 5.740,00          | 3.164,00         | 14.554,00         |
| Manufacturer#3     | 4.230,00          | 5.053,00          | 2.373,00         | 11.656,00         |
| <b>Total</b>       | <b>151.664,00</b> | <b>155.834,00</b> | <b>85.510,00</b> | <b>393.008,00</b> |

# OLAP basics

- Descriptive attributes can be used only in combination with the attribute that they describe



The screenshot displays a data visualization interface. On the left, a table lists 19 suppliers with their IDs, addresses, and quantities. The total quantity is 60,179,935.00. On the right, a sidebar titled 'Visualizations' offers various chart types and options. The 'Columns' section shows 'region Hierarchy', 'supplier', 'address', and 'quantity' are selected. The 'Drill through' and 'Cross-report' options are also visible.

| supplier           | address                                  | quantity             |
|--------------------|------------------------------------------|----------------------|
| Supplier#000000001 | N kw4gn1OM Ahw3Sg70BBgQw57lgjzj55R       | 6.106,00             |
| Supplier#000000002 | ji3yh016B5                               | 5.975,00             |
| Supplier#000000003 | mxBQBnxO3CSwl7                           | 5.767,00             |
| Supplier#000000004 | 7zR323R73NMB77wi1                        | 5.445,00             |
| Supplier#000000005 | AmMQ7Mg 10ByLCP52M13xN31jh5hzOgnm00B     | 5.239,00             |
| Supplier#000000006 | QQL6hxmnmkkgMwgm7CB5B 30Llz              | 5.180,00             |
| Supplier#000000007 | z45m2jBRzI5iILNz4                        | 5.973,00             |
| Supplier#000000008 | xz5m4C A4AAj0kANQ                        | 4.681,00             |
| Supplier#000000009 | m7k7CnC3wiP                              | 5.564,00             |
| Supplier#000000010 | wN1S4mQ0g7Px5Lj34xw6kS4Li4NzB4mO         | 5.556,00             |
| Supplier#000000011 | ikz2MhBMQAq                              | 6.294,00             |
| Supplier#000000012 | Nym5z1Si43B 2yQxm2yOg4Q45kRLxg6ymgz4w0   | 5.764,00             |
| Supplier#000000013 | M6y461ChMgR6gwM3m275kQm6kQ               | 6.448,00             |
| Supplier#000000014 | I4iQ47yygAilnQiS2yg2                     | 5.444,00             |
| Supplier#000000015 | Mh72i PQAxzwO5MAiONw7C0LIA               | 5.527,00             |
| Supplier#000000016 | B7wLkSLRjNS MS1C                         | 5.855,00             |
| Supplier#000000017 | wPgjxnQPOnz                              | 5.845,00             |
| Supplier#000000018 | 2LAO i0 Q5                               | 6.370,00             |
| Supplier#000000019 | 2i024kik5LxzOiPxCOwBPz3jx C71 QyOgiROOP0 | 6.237,00             |
| <b>Total</b>       |                                          | <b>60.179.935,00</b> |

# OLAP basics

- Navigation of hierarchies is enabled when hierarchies are used in some (not all) visualizations
  - Drill-up: corresponds to roll-up
  - Drill-down: interactive mode to slice-and-drill
  - Next-level: drills-down and replaces the current level with the finer one in the hierarchy
  - Expand-all: drills-down and keeps the current level

The screenshot displays a BI tool interface. On the left, a table shows data for regions and a total. The table has two columns: 'region' and 'quantity'. The rows are: AFRICA (11.777.289,00), AMERICA (12.252.085,00), ASIA (12.044.451,00), EUROPE (11.969.295,00), MIDDLE EAST (12.136.815,00), and a Total row (60.179.935,00). Above the table, a toolbar contains navigation icons: an up arrow (drill-up), a down arrow (drill-down), a double down arrow (next-level), and a refresh icon. These four icons are circled in red. On the right, a visualization pane shows a grid of icons at the top. Below, the 'Rows' section contains a hierarchy: 'region Hierarchy' (with a dropdown arrow and close button), 'region', 'nation', and 'supplier'. The 'Columns' section is empty with the text 'Add data fields here'. The 'Values' section contains 'quantity' (with a dropdown arrow and close button).

| region       | quantity             |
|--------------|----------------------|
| AFRICA       | 11.777.289,00        |
| AMERICA      | 12.252.085,00        |
| ASIA         | 12.044.451,00        |
| EUROPE       | 11.969.295,00        |
| MIDDLE EAST  | 12.136.815,00        |
| <b>Total</b> | <b>60.179.935,00</b> |



# OLAP basics

- Slice: open the filters panel
  - Either add an attribute/measure (e.g., *mfgr*)
  - Or expand an attribute/measure already used in the query (e.g., *region*)

The screenshot displays a BI tool interface with three main components:

- Data Table:** A table with columns 'region' and 'quantity'. The rows are: AFRICA (2.380.171,00), AMERICA (2.458.263,00), ASIA (2.392.870,00), EUROPE (2.405.254,00), MIDDLE EAST (2.421.242,00), and a Total row (12.057.800,00).
- Filters Panel:** Shows a filter for 'mfgr' with the value 'is Manufacturer#1'. The filter type is 'Basic filtering'. A list of manufacturers is shown with checkboxes: 'Select all' (checked), 'Manufacturer#1 40084' (checked), 'Manufacturer#2 39636' (unchecked), 'Manufacturer#3 40304' (unchecked), 'Manufacturer#4 39841' (unchecked), and 'Manufacturer#5 40135' (unchecked). There is also a 'Require single selection' checkbox (unchecked).
- Visualizations Panel:** Shows a 'Build visual' section with a grid of visualization icons. Below it, the 'Rows' section contains a hierarchy: 'region Hierarchy' (expanded), 'region', 'nation', and 'supplier'. The 'Columns' section is empty with the text 'Add data fields here'. The 'Values' section contains 'quantity'. The 'Drill through' section is empty.

# OLAP basics

- Slice: open the filters panel
  - On attributes: select one or more categorical values
  - On measures: select a range of values, either pre- or post-aggregation
    - Post-aggregation (top-right figure)
      - "SELECT [...] GROUP BY [...] HAVING SUM(quantity) > 12M"
      - Done when filtering on the attribute that is already in the
    - Pre-aggregation (bottom-right figure)
      - "SELECT [...] WHERE quantity > 10 [...]"
      - Requires drag&dropping *quantity* again from the Fields p
- Filter type "Top N" al

The screenshot displays the Power BI interface with two data tables and their corresponding filter panels. The top table shows a summary of quantities by region, and the bottom table shows a filtered view of the same data. The filter panels show the configuration for these views, including the filter type and the values being filtered.

| region       | quantity             |
|--------------|----------------------|
| AMERICA      | 12.252.085,00        |
| ASIA         | 12.044.451,00        |
| MIDDLE EAST  | 12.136.815,00        |
| <b>Total</b> | <b>36.433.351,00</b> |

Filters on this visual: quantity is greater than 12000...

region is (All)

Filters on this page: Add data fields here

Filters on all pages: Add data fields here

Rows: region Hierarchy, region

Columns: Add data fields here

Values: quantity

| region       | quantity             |
|--------------|----------------------|
| AFRICA       | 11.266.647,00        |
| AMERICA      | 11.723.317,00        |
| ASIA         | 11.522.041,00        |
| EUROPE       | 11.454.417,00        |
| MIDDLE EAST  | 11.610.049,00        |
| <b>Total</b> | <b>57.576.471,00</b> |

Filters: Search

Filters on this visual: quantity is (All)

quantity is greater than 10

Filter type: Advanced filtering

Show items when the value is greater than 10

And Or

Apply filter

region is (All)

Visualizations: Build visual

Rows: region Hierarchy, region

Columns: Add data fields here

Values: quantity

# OLAP basics

- Ordering
  - Click on an attribute in the visualization
  - Or click the three dots in the top-right corner

| region       | quantity             |
|--------------|----------------------|
| AMERICA      | 12.252.085,00        |
| MIDDLE EAST  | 12.136.815,00        |
| ASIA         | 12.044.451,00        |
| EUROPE       | 11.969.295,00        |
| AFRICA       | 11.777.289,00        |
| <b>Total</b> | <b>60.179.935,00</b> |

| year         | quantity             |
|--------------|----------------------|
| <b>1996</b>  | <b>23.337.248,00</b> |
| AFRICA       | 4.572.784,00         |
| AMERICA      | 4.741.595,00         |
| ASIA         | 4.671.953,00         |
| EUROPE       | 4.646.000,00         |
| MIDDLE EAST  | 4.704.916,00         |
| <b>1997</b>  | <b>23.194.818,00</b> |
| AFRICA       | 4.516.368,00         |
| AMERICA      | 4.729.382,00         |
| ASIA         | 4.644.887,00         |
| EUROPE       | 4.606.912,00         |
| MIDDLE EAST  | 4.697.269,00         |
| <b>1998</b>  | <b>13.647.869,00</b> |
| AFRICA       | 2.688.137,00         |
| AMERICA      | 2.781.108,00         |
| ASIA         | 2.727.611,00         |
| EUROPE       | 2.716.383,00         |
| MIDDLE EAST  | 2.734.630,00         |
| <b>Total</b> | <b>60.179.935,00</b> |

# More functionalities: visualization

- Visualizations > Legend
  - Use to break the marks (e.g., bars, trend lines, pie slices) into multiple sub-marks
  - [see OLAP basics]
- Visualizations > Small multiples
  - Use to break the chart into multiple sub-charts
  - [see OLAP basics]
- Data/Drill > Visual table
  - Use to get a table-like visualization of the chart
- Data/Drill > Data point table
  - Useful to export to check the original values before the aggregation

# More functionalities: new fields

- Right-click on an attribute > Create > ...
  - > Group (discrete attribute)
    - Manually take some members from an attribute and put them in a new attribute
  - > Group (numeric attribute)
    - Discretize numerical into bins (e.g., the account balance of customers)
    - Support only for equi-width binning
- Table tools > ...
  - New column
    - Create a new (virtual column) based on a custom formula
  - New measure
    - Define a function to aggregate data (to be used in place of SUM(), AVG(), etc.)

# More functionalities: columns

- Create a new (virtual column) based on a custom formula
- Examples in FT
  - *profit = 'dw ft'[extendedprice] \* 'dw ft'[quantity]*
  - *taxed = IF('dw ft'[tax]>0,"taxed","not taxed")*
- In DT\_CUSTOMER
  - *order\_age = DATEDIFF('dw dt\_order'[rel\_orderdate\_date], TODAY(), YEAR)*

# More functionalities: measures

- Define a function to aggregate data (to be used in place of SUM(), AVG(), etc.)
- Examples in FT:
  - *sum\_of\_profit = SUM('dw ft'[profit])*
  - *avg\_profit\_by\_customer = SUM('dw ft'[profit]) / DISTINCTCOUNT('dw dt\_customer'[custkey])*

# More functionalities: advanced measures

- Nested aggregation

- Further aggregate the results of a query
- For instance:
  - Calculate the sum(Quantity) by Brand and Nation (right)
  - Calculate the average of the result by Brand (left)

- *AvgNationSales* = *AVERAGEX*(  
*SUMMARIZE*(  
     'dw ft',  
     'dw dt\_customer'[nation],  
     'dw dt\_part'[brand],  
     "NationBrandSales",  
     *SUM*('dw ft'[quantity])  
 ),  
 [NationBrandSales]  
 )  
 • *AvgNationSales2* =  
*SUM*('dw ft'[quantity])/  
*DISTINCTCOUNT*('dw dt\_customer'[nation])

| Rows           | Columns              | Values |
|----------------|----------------------|--------|
| brand          |                      |        |
|                | Add data fields here |        |
| AvgNationSales |                      |        |

| brand    | AvgNationSales |
|----------|----------------|
| Brand#11 | 94.903,84      |
| Brand#12 | 98.313,88      |
| Brand#13 | 96.028,72      |
| Brand#14 | 96.489,28      |
| Brand#15 | 96.576,28      |
| Brand#21 | 95.807,92      |
| Brand#22 | 95.674,56      |
| Brand#23 | 95.195,36      |
| Brand#24 | 95.526,36      |
| Brand#25 | 94.173,64      |
| Brand#31 | 95.635,64      |
| Brand#32 | 95.836,28      |
| Brand#33 | 97.503,56      |
| Brand#34 | 96.464,04      |
| Brand#35 | 99.319,32      |
| Brand#41 | 96.421,32      |

| nation         | Brand#11            | Brand#12            | Brand#13            |
|----------------|---------------------|---------------------|---------------------|
| ALGERIA        | 95.170,00           | 99.180,00           | 93.247,00           |
| ARGENTINA      | 95.578,00           | 93.795,00           | 95.578,00           |
| BRAZIL         | 97.671,00           | 97.624,00           | 95.578,00           |
| CANADA         | 95.273,00           | 98.789,00           | 95.578,00           |
| CHINA          | 97.149,00           | 98.115,00           | 95.578,00           |
| EGYPT          | 92.476,00           | 91.385,00           | 95.578,00           |
| ETHIOPIA       | 93.693,00           | 97.359,00           | 95.578,00           |
| FRANCE         | 99.013,00           | 105.375,00          | 95.578,00           |
| GERMANY        | 91.970,00           | 96.545,00           | 95.578,00           |
| INDIA          | 94.867,00           | 100.471,00          | 95.578,00           |
| INDONESIA      | 93.452,00           | 102.805,00          | 95.692,00           |
| IRAN           | 92.635,00           | 96.600,00           | 98.583,00           |
| IRAQ           | 90.990,00           | 97.213,00           | 93.228,00           |
| JAPAN          | 94.168,00           | 95.491,00           | 95.122,00           |
| JORDAN         | 98.498,00           | 101.384,00          | 96.241,00           |
| KENYA          | 93.894,00           | 97.680,00           | 93.988,00           |
| MOROCCO        | 95.503,00           | 96.707,00           | 97.608,00           |
| MOZAMBIQUE     | 96.091,00           | 101.409,00          | 99.871,00           |
| PERU           | 95.535,00           | 95.422,00           | 97.404,00           |
| ROMANIA        | 96.347,00           | 97.373,00           | 98.774,00           |
| RUSSIA         | 95.737,00           | 102.122,00          | 95.312,00           |
| SAUDI ARABIA   | 91.905,00           | 97.534,00           | 95.573,00           |
| UNITED KINGDOM | 94.468,00           | 96.375,00           | 98.938,00           |
| UNITED STATES  | 95.970,00           | 101.097,00          | 94.739,00           |
| VIETNAM        | 94.543,00           | 99.997,00           | 95.621,00           |
| <b>Total</b>   | <b>2.372.596,00</b> | <b>2.457.847,00</b> | <b>2.400.718,00</b> |

| Rows   | Columns | Values   |
|--------|---------|----------|
| nation |         |          |
|        | brand   |          |
|        |         | quantity |



# More functionalities: advanced measures

- Nested aggregation

- Further aggregate the results of a query

- For instance:

- Calculate the sum(Quantity) by Brand
- Calculate the average of the result by

- $AvgNationSales = AVERAGEX($   
 $SUMMARIZE($   
 $'dw ft',$   
 $'dw dt\_customer'[nation],$   
 $'dw dt\_part'[brand],$   
 $"NationBrandSales",$   
 $SUM('dw ft'[quantity]))$   
 $[NationBrandSales]$   
 $)$   
 $AvgNationSales2 =$   
 $SUM('dw ft'[quantity])/$   
 $DISTINCTCOUNT('dw dt\_customer'[nation])$

- SUMMARIZE computes an intermediate result:
  - Aggregates data from 'dw ft'
  - Group by nation and brand
  - The computed measure is called NationBrandSales, defined as SUM(qty)
- AVERAGEX takes the result of the SUMMARIZE to compute the average of NationBrandSales
  - The fact that this average is done by brand is driven by the query visualization properties

| nation   | Brand#11  | Brand#12       | Brand#13            |
|----------|-----------|----------------|---------------------|
| Brand#15 | 96.576,28 | JORDAN         | 98.498,00           |
| Brand#21 | 95.807,92 | KENYA          | 93.894,00           |
| Brand#22 | 95.674,56 | MOROCCO        | 95.503,00           |
| Brand#23 | 95.195,36 | MOZAMBIQUE     | 96.091,00           |
| Brand#24 | 95.526,36 | PERU           | 95.535,00           |
| Brand#25 | 94.173,64 | ROMANIA        | 96.347,00           |
| Brand#31 | 95.635,64 | RUSSIA         | 95.737,00           |
| Brand#32 | 95.836,28 | SAUDI ARABIA   | 91.905,00           |
| Brand#33 | 97.503,56 | UNITED KINGDOM | 94.468,00           |
| Brand#34 | 96.464,04 | UNITED STATES  | 95.970,00           |
| Brand#35 | 99.319,32 | VIETNAM        | 94.543,00           |
| Brand#41 | 96.421,32 | <b>Total</b>   | <b>2.372.598,00</b> |

# More functionalities: advanced measures

- Use coarser data at a finer aggregation level

- For instance:

- Calculate the sum(Quantity) by Nation (and Region)
- Divide the result by the sum(Quantity) by Region
- $PercWithinRegion = SUM('dw ft'[quantity]) /$

*CALCULATE(*  
*SUM('dw ft'[quantity]),*  
*REMOVEFILTERS('dw dt\_customer'[nation])*  
*)*

CALCULATE computes the  $SUM(qty)$  by removing *nation* from the group by defined in the visualization properties

CALCULATE returns just a value, whereas SUMMARIZE returns a table

| region         | quantity             | PercWithinRegion |
|----------------|----------------------|------------------|
| <b>AFRICA</b>  | <b>12.002.988,00</b> | <b>1,00</b>      |
| ALGERIA        | 2.389.575,00         | 0,20             |
| ETHIOPIA       | 2.374.729,00         | 0,20             |
| KENYA          | 2.391.042,00         | 0,20             |
| MOROCCO        | 2.394.780,00         | 0,20             |
| MOZAMBIQUE     | 2.452.862,00         | 0,20             |
| <b>AMERICA</b> | <b>11.981.946,00</b> | <b>1,00</b>      |
| ARGENTINA      | 2.395.653,00         | 0,20             |
| BRAZIL         | 2.424.241,00         | 0,20             |
| CANADA         | 2.399.893,00         | 0,20             |
| PERU           | 2.367.599,00         | 0,20             |
| UNITED STATES  | 2.394.560,00         | 0,20             |
| <b>ASIA</b>    | <b>12.068.349,00</b> | <b>1,00</b>      |
| CHINA          | 2.433.209,00         | 0,20             |
| INDIA          | 2.388.296,00         | 0,20             |
| INDONESIA      | 2.443.400,00         | 0,20             |
| JAPAN          | 2.397.002,00         | 0,20             |
|                | 2.406.442,00         | 0,20             |
|                | <b>12.184.563,00</b> | <b>1,00</b>      |
|                | 2.470.715,00         | 0,20             |
|                | 2.412.224,00         | 0,20             |
|                | 2.453.943,00         | 0,20             |
|                | 2.454.745,00         | 0,20             |
|                | 2.392.936,00         | 0,20             |
|                | <b>11.942.089,00</b> | <b>1,00</b>      |
|                | 2.367.813,00         | 0,20             |
|                | 2.402.245,00         | 0,20             |
|                | 2.367.257,00         | 0,20             |
|                | 2.454.503,00         | 0,21             |
| SAUDI ARABIA   | 2.350.271,00         | 0,20             |
| <b>Total</b>   | <b>60.179.935,00</b> | <b>1,00</b>      |

Rows

- region
- nation

Columns

Add data fields here

Values

- quantity
- PercWithinRegion

# References

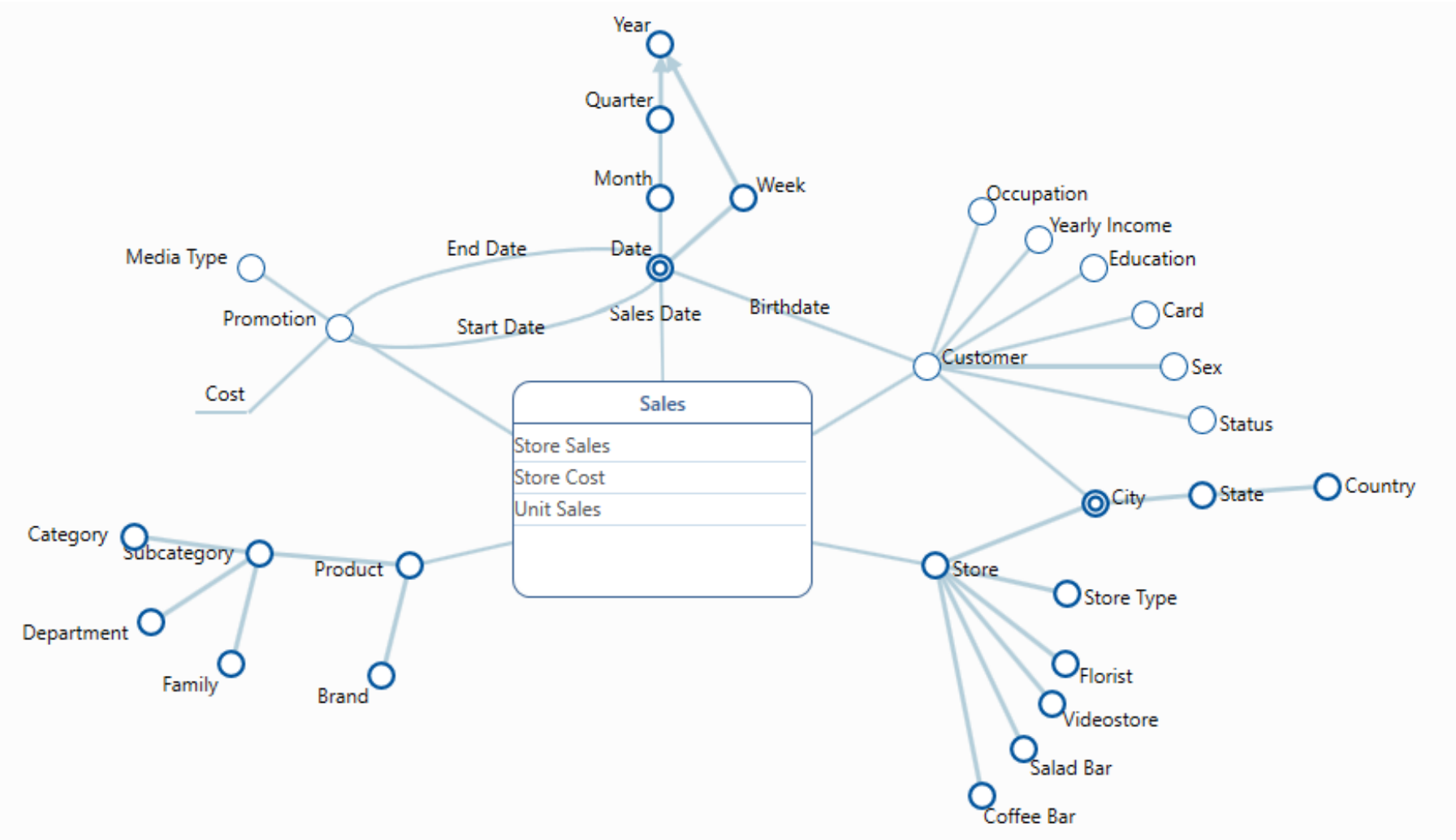
Doc: <https://docs.microsoft.com/en-us/power-bi/fundamentals/>

A *lot* of YouTube videos



# Foodmart Sales - Exercises

# DFM – Foodmart



# Exercise 1

- Use a bar chart to plot the total sum of *STORE\_SALES* for each *STORE\_STATE*
  - Which one is the state with the highest sales?
- Apply a drill-down operation to show the sales at the *STORE\_CITY* level
  - Are there cities whose sales are much lower than the others'?
- How many stores are there in each *STORE\_STATE*? In each *STORE\_CITY*?
  - Color the bars based on the *Count(Distinct)* summarization function over the *STORE\_NAME* attribute
  - Would it be reasonable to say that cities with fewer stores also have lower total sales?

# Exercise 2

- Use a bar chart to plot the total sum of *STORE\_SALES* by *STORE\_CITY* and assign the *STORE\_TYPE* to the Legend property
  - Can you notice any interesting pattern?
- Use a bar chart to plot the total sum of *STORE\_SALES* by *STORE\_TYPE*
  - Assign the number of stores to the color property
  - Is the result surprising/expected?

# Exercise 3

- Use a line chart to plot the monthly sales trend
  - Any interesting pattern?
- Split the previous chart by *STORE\_STATE*
  - Put the *STORE\_STATE* in the Legend
  - Does the previous pattern hold for each state?
- Visualise the impact of each *PRODUCT\_FAMILY* on the total sales while still showing the monthly trends
  - Use a Stacked area chart, where the *STORE\_STATE* is in the Small multiples and the *PRODUCT\_FAMILY* in the Legend



# Exercise 4

- Analyze sales by *STORE\_TYPE* (sorted by descending order)
- Drill-down to the stores
- Add the number of customers
  - Use the *Count(Distinct)* summarization function
  - In case of wrong calculation (i.e., if you get the same value in all rows):
    - Go back to the Model
    - Double-click the relationships between CUSTOMER and SALES
    - Set the *Cross filter direction* to *Both*
- Add the average sales per customer
  - Create a new measure, calculated by dividing the sum of store sales by the count of distinct customers

# Exercise 5

- Create a table to visualize the sales for each *OCCUPATION* (*Customer* dimension)
- Exclude (i.e., filter out) the tuples where the value of *STORE\_SALES* is lower than 5
- Apply another filter (in addition to the previous one) to exclude all occupations where the total sales is lower than 80K

# Exercise 6

- Create a table to visualize with the top ten customers by total sales
  - Show both *CUSTOMER\_ID* and *FULLNAME*
- Add the *Occupation* field
- Turn it into a matrix (without the *FULLNAME*)
- Add a measure on the Customer table calculating a ranking of customers
  - First, declare a new measure simply calculating the sum of *STORE\_SALES*
  - Then, declare a new measure calculating the RANKX, where
    - The 1st parameter is the attribute that we want to order, i.e., the *CUSTOMER\_ID*
    - The 2nd parameter is the measure to be used for ordering, i.e., the one declared above
- Take the first ten customers for each occupation by filtering on the rank

# Exercise 7

- Create a histogram of StoreSales
  - Right-click on *STORE\_SALES* > New group > Create bins of size 2
  - Create a bar chart showing the count of records for each bin
- Use the same binning to plot a bar chart with the average *STORE\_COST* for each bin
  - Do you see a correlation in the data?
- Plot the same result as a scatter chart
  - Find the chart in the list of visuals
  - Put *STORE\_COST* and *STORE\_SALES* on X and Y axis, respectively (without summarizing)

# Exercise 8

- Create a new column calculating the profits
  - $PROFIT = STORE\_SALES - STORE\_COST$
- Create a line chart showing the monthly trend of profits, sales, and costs

# Exercise 9

- Check distribution of profits with respect to different attributes
  - Try some combinations of attributes (e.g., with the stacked bar chart)
    - E.g., check the distribution of profits with respect to different combinations of occupation and gender
  - Is there any categorical values that sticks out? Or are profits mainly driven by the number of customers?
- Create bins where necessary (e.g., population)
  - Calculate the age of customers from their birthdate
    - A new column must be defined
    - `<newColumn> = DATEDIFF(<date1>, <date2>, YEAR)`
  - Calculate bins of customer ages and check the number of customers in each bin and distribution of profits

# Foodmart Inventory - Exercises

# Exercise A.1

- Visualise, through a line chart, the monthly trend of the average *Supply Time* of each state (use the state of the warehouse). Show one line for each state and place them on the same chart, that is, the lines must share the same  $x$  and  $y$  axes.



# Exercise A.2

- Create a matrix to visualise the sum of *Units Shipped* for each combination of *Warehouse* and *Store*

# Exercise A.3

- Find the state whose warehouses have the highest total *Store Invoice*. Perform a drill-down on such state and filter-out all the others. Visualize, for each city of the selected state, the total sum of *Store Invoice* as a bar chart, and color the bars by the average *Supply Time*.

# Exercise A.4

- Visualise through a bar chart the sum of *Units Shipped* for each *Warehouse* in the period going from 1997-06-01 to 1997-08-31. Split the bars by *Product family* and add the average *Supply Time* to the tooltips. Also, keep only the three warehouses with the maximum amount of *Units Shipped*.

# Exercise A.5

- In the product table, create a *units\_per\_pallet* column that multiplies the number of *units\_per\_case* by the number of *cases\_per\_pallet*
- Create a report with multiple pie charts, where you show the proportion of *Units ordered* by either:
  - A reasonable grouping for the *units\_per\_pallet*
  - A reasonable grouping for the *gross\_weight*
  - *recyclable\_package*