นาย DAYUTH THY 66122420126

ใบงานที่ 3
Data Transformation- data aggregation
ให้นักศึกษาดำเนินการ Transformation จากชุดข้อมูล (Dataset)
Sales_data.xlsx โดยทำการ import data เข้ามาไว้ที่ ของ local
Repository และดำเนินการดังต่อไปนี้
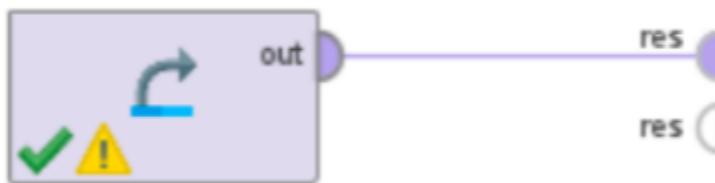
1. ชุดข้อมูลนี้ประกอบไปด้วยกี่ attribute อะไรบ้าง และมีจำนวน
   examples set ทั้งหมดเท่าไหร่
   Answer : 6 Attributes
   - **Transaction ID**
   - **Customer Name**
   - **Region**
   - **Sale Amount**
   - **Purchase Date**
   - **Category**

   Example :
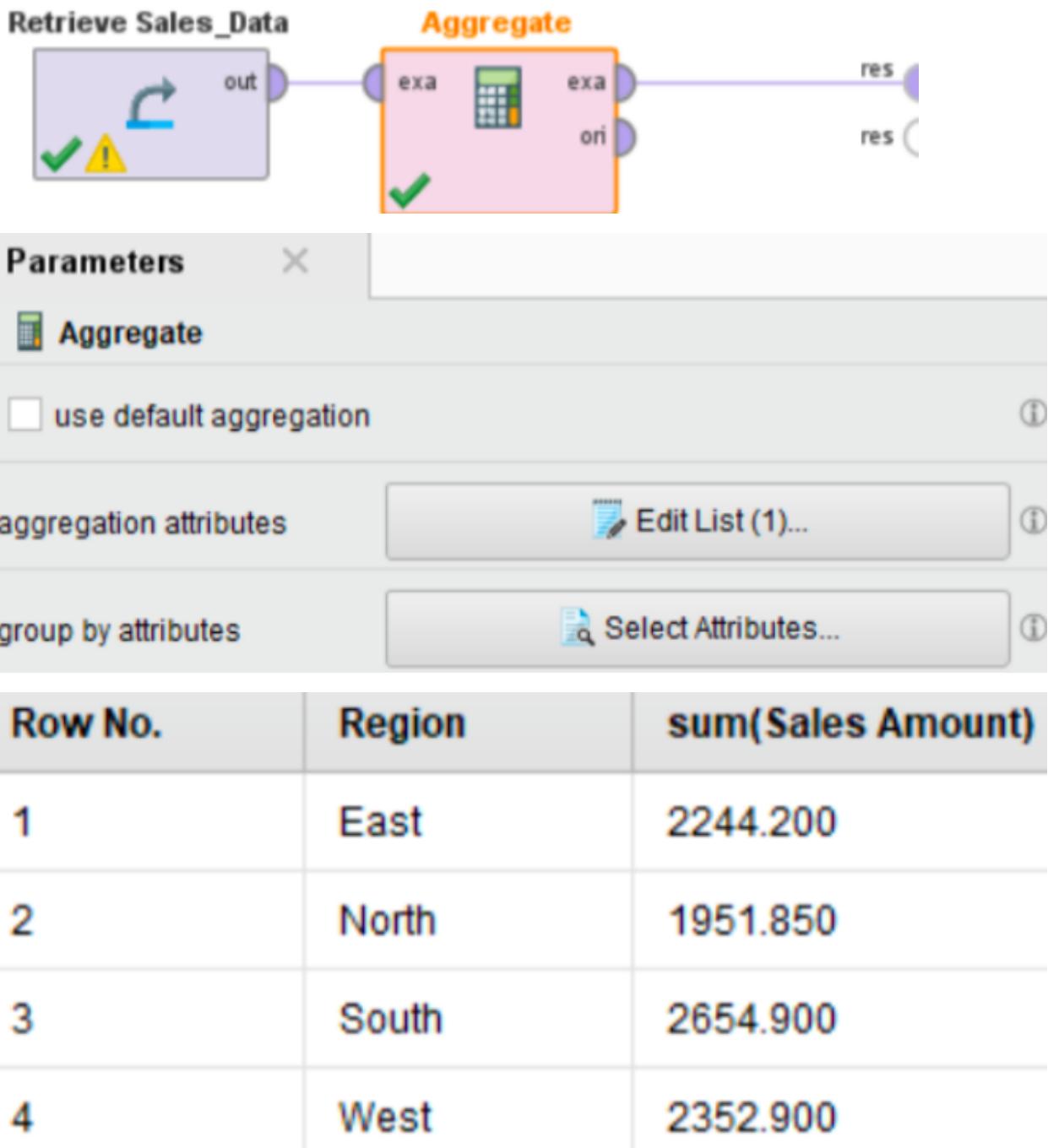
Name		Type	Missing	Statistics			Filter (6 / 6 attributes):	Search for Attributes	
∨ **Transaction ID**		Nominal	0	Least T030 (1)	Most T001 (1)	Values T001 (1), T002 (1), ...[28 more]			
∨ **Customer Name**		Nominal	0	Least Zachary Peach (1)	Most Alice Smith (1)	Values Alice Smith (1), Anna Lavender (1), ...[28 more]			
∨ **Region**		Nominal	0	Least West (7)	Most North (8)	Values North (8), South (8), ...[2 more]			
∨ **Sales Amount**		Real	0	Min 120.500	Max 500.900	Average 306.795			
∨ **Purchase Date**		Nominal	0	Least 2024-01-30 (1)	Most 2024-01-01 (1)	Values 2024-01-01 (1), 2024-01-02 (1), ...[28 more]			
∨ **Category**		Nominal	0	Least Groceries (10)	Most Electronics (10)	Values Electronics (10), Furniture (10), ...[1 more]			

Row No.	Transaction ID	Customer Name	Region	Sales Amount	Purchase Date	Category
1	T001	Alice Smith	North	250	2024-01-01	Electronics
2	T002	Bob Johnson	South	120.500	2024-01-02	Groceries
3	T003	Charlie Brown	East	300.750	2024-01-03	Furniture
4	T004	Diana Prince	West	450.200	2024-01-04	Electronics
5	T005	Evan Green	North	200	2024-01-05	Groceries
6	T006	Fiona Black	South	500.900	2024-01-06	Furniture
7	T007	George White	East	340.300	2024-01-07	Electronics
8	T008	Hannah Blue	West	280.450	2024-01-08	Groceries
9	T009	Ivan Gray	North	150.650	2024-01-09	Furniture
10	T010	Julia Gold	South	360.800	2024-01-10	Electronics
11	T011	Kevin Red	East	410.250	2024-01-11	Groceries
12	T012	Laura Silver	West	230.100	2024-01-12	Furniture
13	T013	Michael Yellow	North	310.400	2024-01-13	Electronics
14	T014	Nancy Orange	South	400	2024-01-14	Groceries
15	T015	Oliver Purple	East	210.500	2024-01-15	Furniture
16	T016	Patricia Cyan	West	330.350	2024-01-16	Electronics
17	T017	Quincy Magenta	North	270.200	2024-01-17	Groceries
18	T018	Rachel Pink	South	190.750	2024-01-18	Furniture
19	T019	Steve Teal	East	320.800	2024-01-19	Electronics
20	T020	Tina Amber	West	410.900	2024-01-20	Groceries
21	T021	Uma Indigo	North	290	2024-01-21	Furniture
22	T022	Victor Navy	South	380.400	2024-01-22	Electronics
23	T023	Wendy Maroon	East	240.700	2024-01-23	Groceries
24	T024	Xander Olive	West	350.600	2024-01-24	Furniture
25	T025	Yasmine Lime	North	220.100	2024-01-25	Electronics

ExampleSet (30 examples,0 special attributes,6 regular attributes)
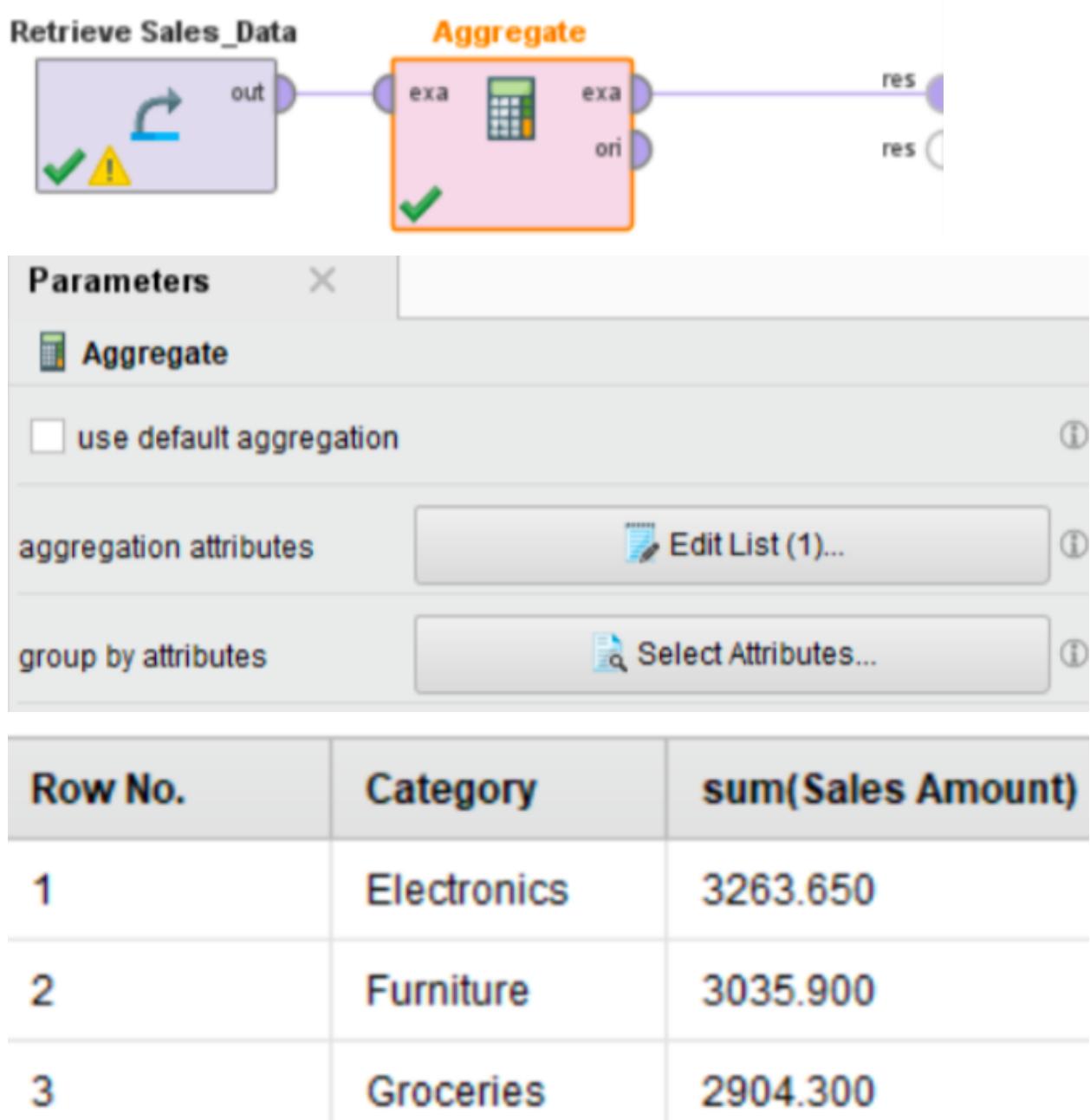
26	T026	Zachary Peach	South	310.800	2024-01-26	Groceries
27	T027	Anna Lavender	East	420.900	2024-01-27	Furniture
28	T028	Ben Emerald	West	300.300	2024-01-28	Electronics
29	T029	Chris Sapphire	North	260.500	2024-01-29	Groceries
30	T030	Dana Ruby	South	390.750	2024-01-30	Furniture

2. ให้ดำเนินการหาผลรวมของ sales amount โดย group by Region

Answer :

Retrieve Sales_Data   **Aggregate**

Parameters ✕

🖩 **Aggregate**

☐ use default aggregation ⓘ

aggregation attributes   📝 Edit List (1)... ⓘ

group by attributes   🔍 Select Attributes... ⓘ

Row No.	Region	sum(Sales Amount)
1	East	2244.200
2	North	1951.850
3	South	2654.900
4	West	2352.900
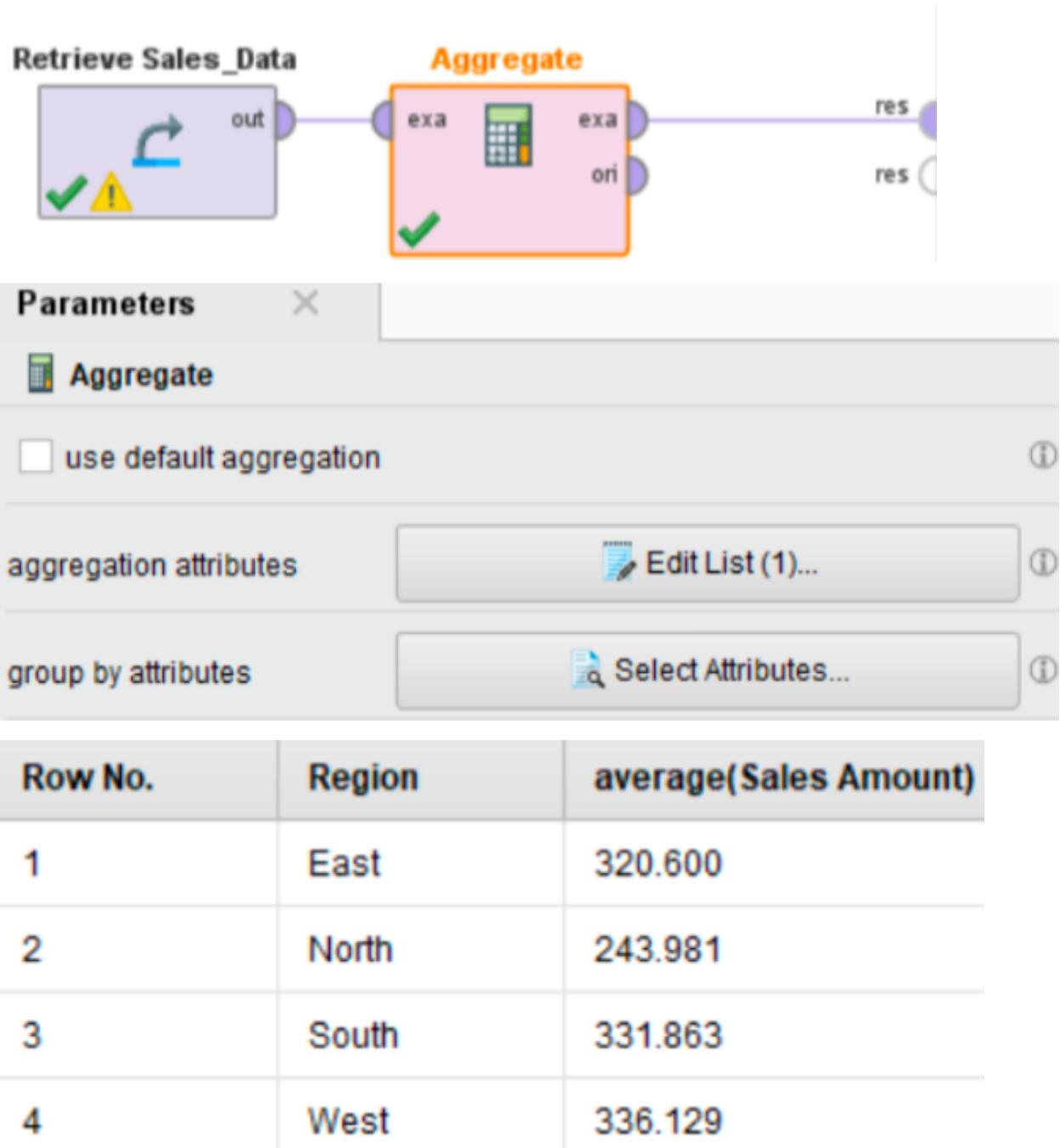
3. ให้ดำเนินการหาผลรวมของ sales amount โดย group by Category

Answer :

Retrieve Sales_Data

Aggregate

Parameters ×

Aggregate

☐ use default aggregation

| aggregation attributes | Edit List (1)... |
| group by attributes | Select Attributes... |

Row No.	Category	sum(Sales Amount)
1	Electronics	3263.650
2	Furniture	3035.900
3	Groceries	2904.300

4. ให้ดำเนินการหาค่าเฉลี่ยของ sales amount โดย group by
   Region
   Answer :

Retrieve Sales_Data

out

Aggregate

exa

exa

ori

res

res

Parameters ✕

🖩 **Aggregate**

☐ use default aggregation ⓘ

aggregation attributes | 📝 Edit List (1)... | ⓘ

group by attributes | 🔍 Select Attributes... | ⓘ

Row No.	Region	average(Sales Amount)
1	East	320.600
2	North	243.981
3	South	331.863
4	West	336.129

5. ให้ดำเนินการหาค่าเฉลี่ยของ sales amount โดย group by Category

ให้นักศึกษา capture หน้า process และ result ของแต่ละข้อส่ง

Answer :

Retrieve Sales_Data   **Aggregate**

Parameters

Aggregate

☐ use default aggregation

aggregation attributes   Edit List (1)...

group by attributes   Select Attributes...

Category	
Electronics	3263.65
Furniture	3035.90
Groceries	2904.30