

นาย DAYUTH THY 66122420126

ใบงานที่ 2

Data Transformation- data discretization

ให้นักศึกษาดำเนินการ Transformation จากชุดข้อมูล (Dataset) **Loan_data.xlsx** โดยทำการ import data เข้ามาไว้ที่ของ local Repository และดำเนินการดังต่อไปนี้

1. ชุดข้อมูลนี้ประกอบไปด้วยกี่ attribute อะไรบ้าง และมีจำนวน examples set ทั้งหมดเท่าไหร่

Answer : 5 Attributes

- ID
- Age
- Income
- Loan Approval Status
- Credit score

Example :

Retrieve Loan_Data



Parameters



Retrieve Loan_Data (Retrieve)

repository entry

//Local Repository/data/Loan_Data



Row No.	ID	Age	Income	Loan Approval Status	Credit Score
1	1	32	80300	Approved	520
2	2	61	10062	Approved	844
3	3	76	110671	Approved	364
4	4	19	118108	Approved	615
5	5	69	34114	Approved	613
6	6	61	59750	Approved	750
7	7	39	30271	Approved	421
8	8	75	65437	Approved	733
9	9	44	16199	Rejected	605
10	10	48	75512	Rejected	831
11	11	41	72949	Rejected	450
12	12	22	85559	Rejected	310
13	13	58	37130	Rejected	794
14	14	32	100102	Approved	549
15	15	49	36399	Rejected	512
16	16	57	20399	Approved	464
17	17	39	67106	Rejected	583
18	18	79	16925	Approved	396
19	19	25	23668	Rejected	584
20	20	64	54605	Approved	442
21	21	52	67183	Rejected	337
22	22	18	51991	Approved	787
23	23	52	37246	Rejected	353
24	24	27	108302	Approved	409
25	25	59	60951	Rejected	663

Name	Type	Missing	Statistics		
			Min	Max	Average
▼ ID	Integer	0	1	25	13
▼ ▲ Age	Integer	0	18	79	47.920
▼ Income	Integer	0	10062	118108	57637.560
▼ Loan Approval Status	Nominal	0	Least Rejected (11)	Most Approved (14)	Values Approved (14), Rejected (11)
▼ Credit Score	Integer	0	310	844	557.160

2. ให้ดำเนินการปรับชุดข้อมูลให้อยู่ในกลุ่มดังนี้

Age:

Interval	Group name
18-30	Young adult
31-50	Middle-aged adult
51-80	Senior adult

Income:

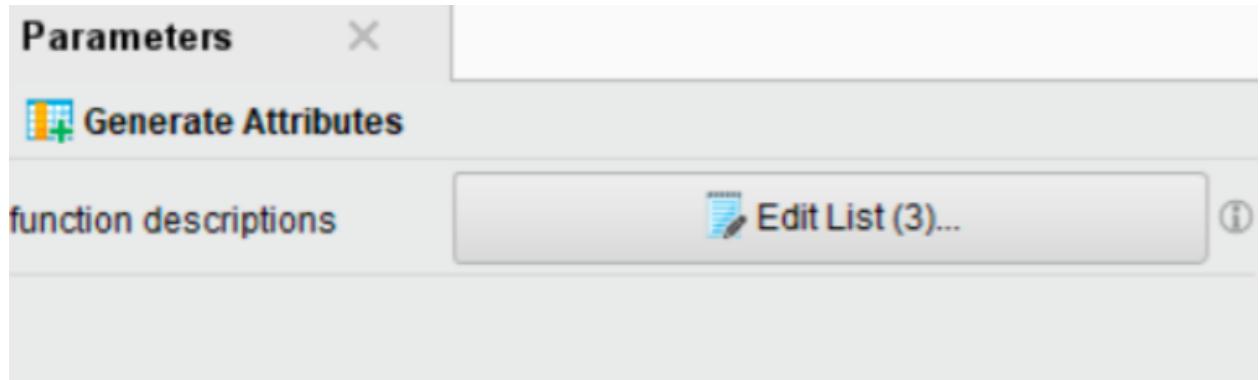
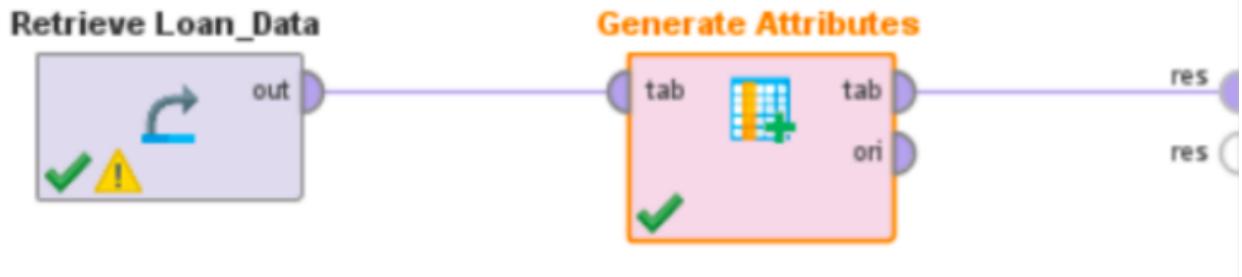
Interval	Group name
<30000	Low income
30001-60000	Middle income
>60000	High income

Credit Score:

Interval	Group name
300-500	Poor credit
501-700	Average Credit
701-850	Excellent Credit

ให้นักศึกษา capture หน้า process และ result ของแต่ละข้อส่ง

Answer :



Row No.	ID	Age	Income	Loan Approval Status	Credit Score	Age_Group	Income_Group	Credit_Score_Group
1	1	32	80300	Approved	520	Middle-aged adult	High income	Average Credit
2	2	61	10062	Approved	844	Senior adult	Low income	Excellent Credit
3	3	76	110671	Approved	364	Senior adult	High income	Poor credit
4	4	19	118108	Approved	615	Young adult	High income	Average Credit
5	5	69	34114	Approved	613	Senior adult	Middle income	Average Credit
6	6	61	59750	Approved	750	Senior adult	Middle income	Excellent Credit
7	7	39	30271	Approved	421	Middle-aged adult	Middle income	Poor credit
8	8	75	65437	Approved	733	Senior adult	High income	Excellent Credit
9	9	44	16199	Rejected	605	Middle-aged adult	Low income	Average Credit
10	10	48	75512	Rejected	831	Middle-aged adult	High income	Excellent Credit
11	11	41	72949	Rejected	450	Middle-aged adult	High income	Poor credit
12	12	22	85559	Rejected	310	Young adult	High income	Poor credit
13	13	58	37130	Rejected	794	Senior adult	Middle income	Excellent Credit
14	14	32	100102	Approved	549	Middle-aged adult	High income	Average Credit
15	15	49	36399	Rejected	512	Middle-aged adult	Middle income	Average Credit
16	16	57	20399	Approved	464	Senior adult	Low income	Poor credit
17	17	39	67106	Rejected	583	Middle-aged adult	High income	Average Credit
18	18	79	16925	Approved	396	Senior adult	Low income	Poor credit
19	19	25	23668	Rejected	584	Young adult	Low income	Average Credit
20	20	64	54605	Approved	442	Senior adult	Middle income	Poor credit
21	21	52	67183	Rejected	337	Senior adult	High income	Poor credit
22	22	18	51991	Approved	787	Young adult	Middle income	Excellent Credit
23	23	52	37246	Rejected	353	Senior adult	Middle income	Poor credit
24	24	27	108302	Approved	409	Young adult	High income	Poor credit
25	25	59	60951	Rejected	663	Senior adult	High income	Average Credit