
Unit 3

Power BI

Exploring PowerBI



Power BI

SUBMITTED BY

VISHAL JATIA - 2021UCS1589

RISHNAK KUMAR - 2021UCS1588

NABEEL AHMED - 2021UCS1578

CSE-2

Explore Power View, Power Query

- Create a table Employee(empid, gender, department, salary, country, year_of_joining)
- Connect to the Employee data file.
- Remove missing gender and department values.
- Extract year_of_joining column and visualise number of employees w.r.t year of experience in the company.
- Perform self-join using Power Query.
- Aggregate salary with gender and Visualise using Pie chart.

Create a table Employee(empid, gender, department, salary, country, year_of_joining)

1 ² ₃ empid	A ^B _C gender	A ^B _C department	1 ² ₃ salary	A ^B _C country	1 ² ₃ year_of_joining
1	Male	Sales	50000	USA	2010
2	Female	HR	60000	Canada	2012
3	Male	IT	70000	UK	2015
4	Female	Finance	55000	Germany	2011
5	Male	Marketing	48000	France	2013
6	Female	Operations	65000	USA	2014
7	Male	Sales	52000	Canada	2016
8	Female	HR	61000	UK	2017
9	Male	IT	72000	Germany	2018
10	Female	Finance	56000	France	2019
11	Male	Marketing	49000	USA	2010
12	Female	Operations	66000	Canada	2012
13	Male	Sales	53000	UK	2013
14	Female	HR	62000	Germany	2014
15	Male	IT	73000	France	2015
16	Female	Finance	57000	USA	2016
17	Male	Marketing	50000	Canada	2017
18	Female	Operations	67000	UK	2018
19	Male	Sales	54000	Germany	2019
20	Female	HR	63000	France	2010
21	Male	IT	74000	USA	2011
22	Female	Finance	58000	Canada	2012
23	Male	Marketing	51000	UK	2013
24	Female	Operations	68000	Germany	2014
25	Male	Sales	55000	France	2015
26	Female	HR	64000	USA	2016
27	Male	IT	75000	Canada	2017
28	Female	Finance	59000	UK	2018
29	Male	Marketing	52000	Germany	2019
30	Female	Operations	69000	France	2010

Remove missing gender and department values.

= Table.SelectRows("#Changed Type", each [gender] <> null and [gender] <> "")						
i ² ₃ empid	A ^B _C gender	A ^B _C department	i ² ₃ salary	A ^B _C country	i ² ₃ year_of_joining	
1	1 Male	Sales	50000	USA	2010	
2	2 Female	HR	60000	Canada	2012	
3	3 Male	IT	70000	UK	2015	
4	4 Female	Finance	55000	Germany	2011	
5	5 Male	Marketing	48000	France	2013	
6	6 Female	Operations	65000	USA	2014	
7	7 Male	Sales	52000	Canada	2016	
8	8 Female	HR	61000	UK	2017	
9	9 Male	IT	72000	Germany	2018	
10	10 Female	Finance	56000	France	2019	
11	11 Male	Marketing	49000	USA	2010	
12	12 Female	Operations	66000	Canada	2012	
13	13 Male	Sales	53000	UK	2013	
14	14 Female	HR	62000	Germany	2014	
15	15 Male	IT	73000	France	2015	
16	16 Female	Finance	57000	USA	2016	
17	17 Male	Marketing	50000	Canada	2017	
18	18 Female	Operations	67000	UK	2018	
19	19 Male	Sales	54000	Germany	2019	
20	20 Female	HR	63000	France	2010	
21	21 Male	IT	74000	USA	2011	
22	22 Female	Finance	58000	Canada	2012	
23	23 Male	Marketing	51000	UK	2013	
24	24 Female	Operations	68000	Germany	2014	
25	25 Male	Sales	55000	France	2015	
26	26 Female	HR	64000	USA	2016	
27	27 Male	IT	75000	Canada	2017	
28	28 Female	Finance	59000	UK	2018	
29	29 Male	Marketing	52000	Germany	2019	
30	30 Female	Operations	69000	France	2010	

= Table.SelectRows("#Filtered Rows", each [department] <> null and [department] <> "")						
123 empid	A8C gender	A8C department	123 salary	A8C country	123 year_of_joining	
1	Male	Sales	50000	USA	2010	
2	Female	HR	60000	Canada	2012	
3	Male	IT	70000	UK	2015	
4	Female	Finance	55000	Germany	2011	
5	Male	Marketing	48000	France	2013	
6	Female	Operations	65000	USA	2014	
7	Male	Sales	52000	Canada	2016	
8	Female	HR	61000	UK	2017	
9	Male	IT	72000	Germany	2018	
10	Female	Finance	56000	France	2019	
11	Male	Marketing	49000	USA	2010	
12	Female	Operations	66000	Canada	2012	
13	Male	Sales	53000	UK	2013	
14	Female	HR	62000	Germany	2014	
15	Male	IT	73000	France	2015	
16	Female	Finance	57000	USA	2016	
17	Male	Marketing	50000	Canada	2017	
18	Female	Operations	67000	UK	2018	
19	Male	Sales	54000	Germany	2019	
20	Female	HR	63000	France	2010	
21	Male	IT	74000	USA	2011	
22	Female	Finance	58000	Canada	2012	
23	Male	Marketing	51000	UK	2013	
24	Female	Operations	68000	Germany	2014	
25	Male	Sales	55000	France	2015	
26	Female	HR	64000	USA	2016	
27	Male	IT	75000	Canada	2017	
28	Female	Finance	59000	UK	2018	
29	Male	Marketing	52000	Germany	2019	
30	Female	Operations	69000	France	2010	

Extract year_of_joining column and visualise number of employees w.r.t year of experience in the company.

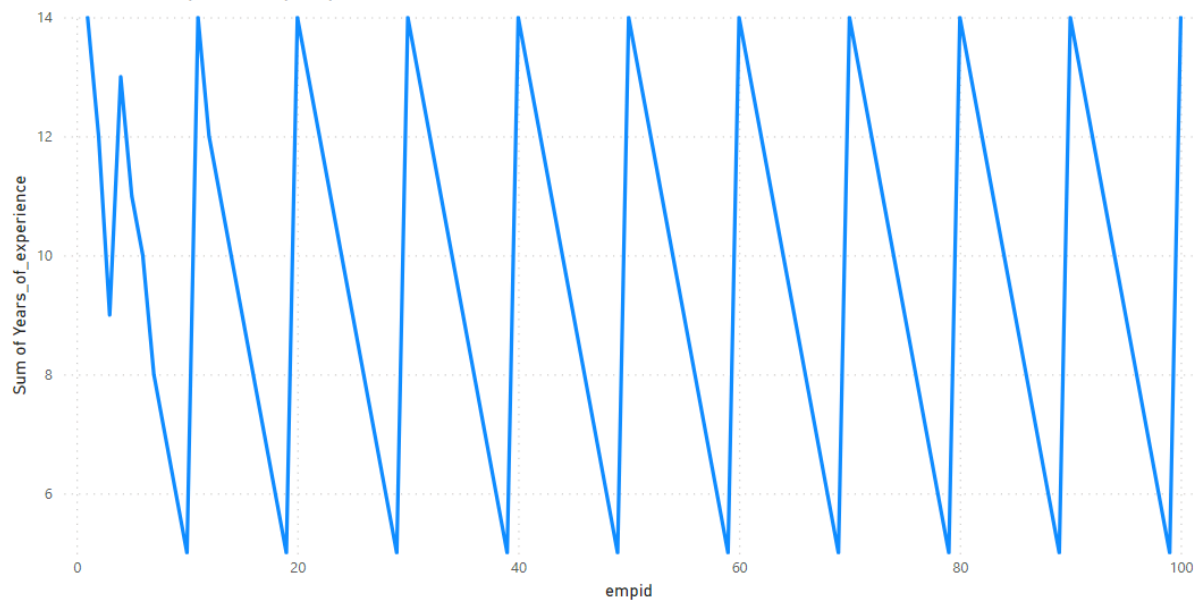
Formula to create year of experience column:

Date.Year(DateTime.LocalNow()) - [Year of Joining]

Column Added

= Table.AddColumn("#Filtered Rows1", "Years of Experience", each Date.Year(DateTime.LocalNow()) - [year_of_joining])							
	gender	department	salary	country	year_of_joining	Years of Experience	
1	Male	Sales	50000	USA	2010	14	
2	Female	HR	60000	Canada	2012	12	
3	Male	IT	70000	UK	2015	9	
4	Female	Finance	55000	Germany	2011	13	
5	Male	Marketing	48000	France	2013	11	
6	Female	Operations	65000	USA	2014	10	
7	Male	Sales	52000	Canada	2016	8	
8	Female	HR	61000	UK	2017	7	
9	Male	IT	72000	Germany	2018	6	
10	Female	Finance	56000	France	2019	5	
11	Male	Marketing	49000	USA	2010	14	
12	Female	Operations	66000	Canada	2012	12	
13	Male	Sales	53000	UK	2013	11	
14	Female	HR	62000	Germany	2014	10	
15	Male	IT	73000	France	2015	9	
16	Female	Finance	57000	USA	2016	8	
17	Male	Marketing	50000	Canada	2017	7	
18	Female	Operations	67000	UK	2018	6	
19	Male	Sales	54000	Germany	2019	5	
20	Female	HR	63000	France	2010	14	
21	Male	IT	74000	USA	2011	13	
22	Female	Finance	58000	Canada	2012	12	
23	Male	Marketing	51000	UK	2013	11	
24	Female	Operations	68000	Germany	2014	10	
25	Male	Sales	55000	France	2015	9	
26	Female	HR	64000	USA	2016	8	
27	Male	IT	75000	Canada	2017	7	
28	Female	Finance	59000	UK	2018	6	
29	Male	Marketing	52000	Germany	2019	5	
30							

Sum of Years_of_experience by empid



Perform self-join using Power Query.

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Keep Rows Remove Rows Split Column Group By Use First Row as Headers Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Text Analytics Vision Azure Machine Learning

Queries [2] emp_data Employee_SelfJoin

Table.AddColumn

emp_data

empid	gender	department	salary	country	year_of_joining	Years_of_experience
1	Male	Sales	50000	USA	2010	14
2	Female	HR	60000	Canada	2012	12
3	Male	IT	70000	UK	2015	9
4	Female	Finance	55000	Germany	2011	13
5	Male	Marketing	48000	France	2013	11

Employee_SelfJoin

empid	gender	department	salary	country	year_of_joining	Years_of_experience
1	Male	Sales	50000	USA	2010	14
2	Female	HR	60000	Canada	2012	12
3	Male	IT	70000	UK	2015	9
4	Female	Finance	55000	Germany	2011	13
5	Male	Marketing	48000	France	2013	11

Join Kind

Left Outer (all from first, matching from second)

☐ Use fuzzy matching to perform the merge

Fuzzy matching options

☒ The selection matches 100 of 100 rows from the first table.

OK Cancel

7 COLUMNS, 100 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 12:15

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Keep Rows Remove Rows Split Column Group By Use First Row as Headers Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Text Analytics Vision Azure Machine Learning

Queries [2] emp_data Employee_SelfJoin

Table.NestedJoin(#"Added Custom", {"department"}, emp_data, {"department"}, "emp_data", JoinKind.LeftOuter)

emp_data

empid	gender	department	salary	country	year_of_joining	Years_of_experience
1	Male	Sales	50000	USA	2010	14
2	Female	HR	60000	Canada	2012	12
3	Male	IT	70000	UK	2015	9
4	Female	Finance	55000	Germany	2011	13
5	Male	Marketing	48000	France	2013	11

Expand

(Select All Columns)

☒ empid

☒ gender

☒ department

☒ salary

☒ country

☒ year_of_joining

☒ Years_of_experience

☒ Use original column name as prefix

OK Cancel

8 COLUMNS, 100 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 12:21

Self Joined table:

Untitled - Power BI Desktop

File Home Help **Table tools** Share

Name Employee_SelfJoin

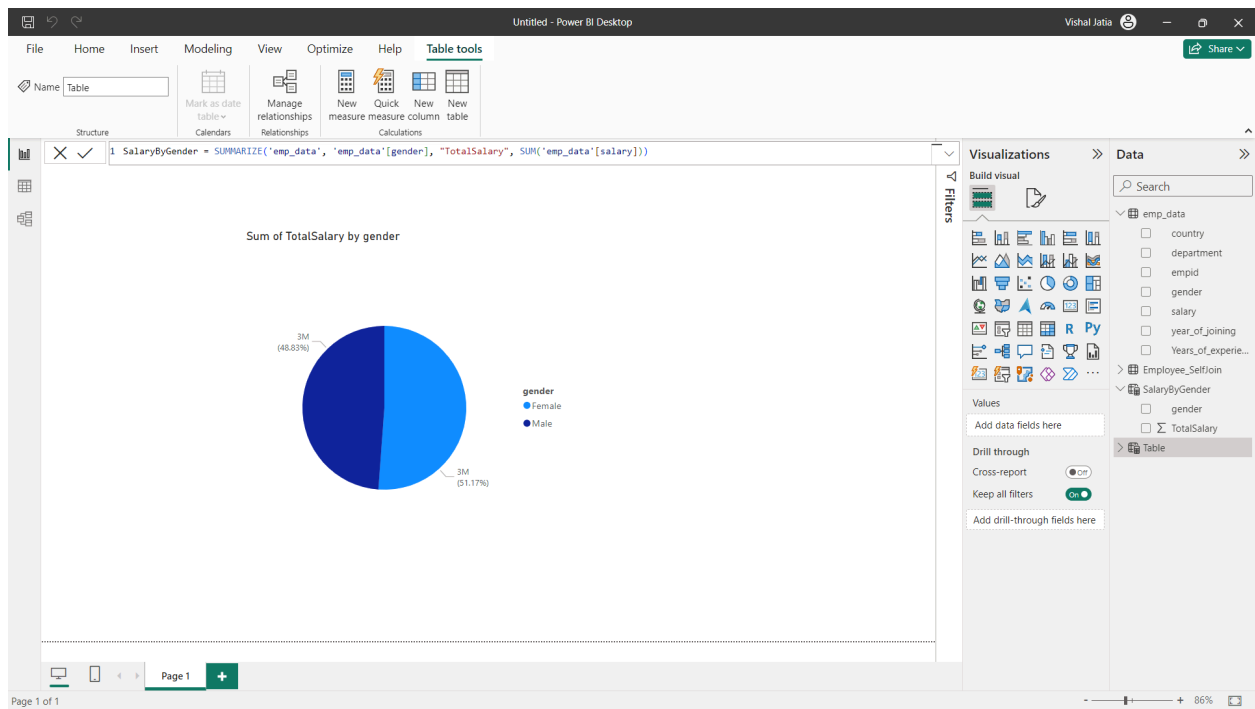
Structure

Mark as date table
Manage relationships
New measure
Quick measure column
New table

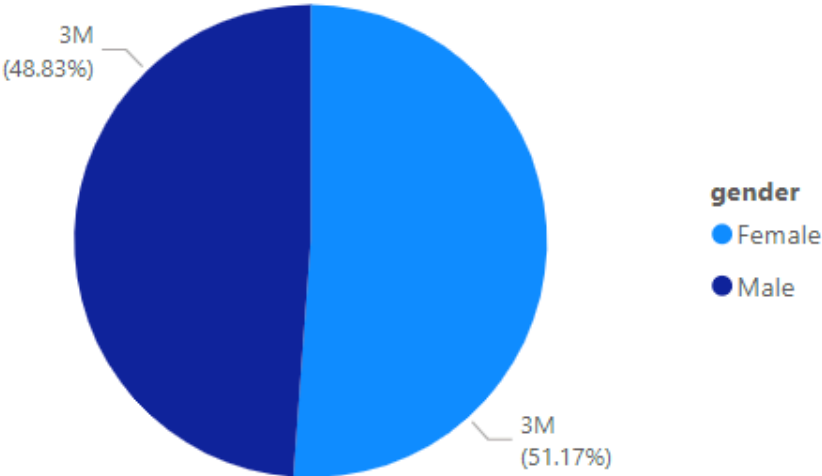
empid	gender	department	salary	country	year_of_joining	Years_of_experience	emp_data.gender	emp_data.department	emp_data.salary	emp_data.Years_of_experience
2	Female	HR	60000	Canada	2012	12	Female	HR	60000	12
2	Female	HR	60000	Canada	2012	12	Female	HR	61000	7
2	Female	HR	60000	Canada	2012	12	Female	HR	62000	10
2	Female	HR	60000	Canada	2012	12	Female	HR	63000	14
2	Female	HR	60000	Canada	2012	12	Female	HR	64000	8
2	Female	HR	60000	Canada	2012	12	Female	HR	65000	12
2	Female	HR	60000	Canada	2012	12	Female	HR	66000	6
2	Female	HR	60000	Canada	2012	12	Female	HR	67000	10
2	Female	HR	60000	Canada	2012	12	Female	HR	68000	14
2	Female	HR	60000	Canada	2012	12	Female	HR	69000	8
2	Female	HR	60000	Canada	2012	12	Female	HR	70000	12
2	Female	HR	60000	Canada	2012	12	Female	HR	71000	6
2	Female	HR	60000	Canada	2012	12	Female	HR	72000	10
2	Female	HR	60000	Canada	2012	12	Female	HR	73000	14
2	Female	HR	60000	Canada	2012	12	Female	HR	74000	8
2	Female	HR	60000	Canada	2012	12	Female	HR	75000	12
2	Female	HR	60000	Canada	2012	12	Female	HR	76000	6
32	Female	HR	65000	Canada	2012	12	Female	HR	60000	12
32	Female	HR	65000	Canada	2012	12	Female	HR	61000	7
32	Female	HR	65000	Canada	2012	12	Female	HR	62000	10
32	Female	HR	65000	Canada	2012	12	Female	HR	63000	14
32	Female	HR	65000	Canada	2012	12	Female	HR	64000	8
32	Female	HR	65000	Canada	2012	12	Female	HR	65000	12
32	Female	HR	65000	Canada	2012	12	Female	HR	66000	6
32	Female	HR	65000	Canada	2012	12	Female	HR	67000	10
32	Female	HR	65000	Canada	2012	12	Female	HR	68000	14
32	Female	HR	65000	Canada	2012	12	Female	HR	69000	8
32	Female	HR	65000	Canada	2012	12	Female	HR	70000	12
32	Female	HR	65000	Canada	2012	12	Female	HR	71000	6

Table: Employee_SelfJoin (1,668 rows)

Aggregate salary with gender and Visualise using Pie chart.



Sum of TotalSalary by gender



Visualise the result of any Machine Learning algorithm on any dataset of your choice in PowerBI.

Introduction:

"This model leverages a comprehensive set of factors, including gender, marital status, number of dependents, education level, self-employment status, applicant income, co-applicant income, loan amount, loan term, credit history, and property location, to accurately predict the likelihood of loan approval. Our goal extends beyond prediction and into the realm of visualisation, where Power BI will be employed to present these findings in a compelling and insightful manner."

Here's why this version is stronger:

- Highlights range of factors: Emphasises the multi-dimensional nature of the model.
- Concisely states the goal: Clearly mentions the goal of loan status prediction.
- Enhances 'visualisation' aspect: Underscores the use of Power BI to communicate results effectively.

Visualisations:

1. Accuracy and Roc score

Accuracy Score: 81.82%

Roc Score: 66.67%

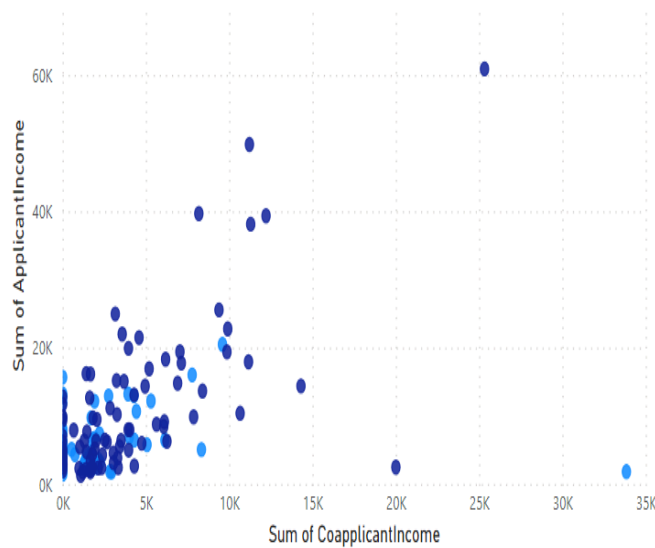
2. Confusion matrix

Predicted	0	1	All
True			
0	7	14	21
1	0	56	56
All	7	70	77

3. Finding the relationship between loan amount (real and predicted) and the income of the family.

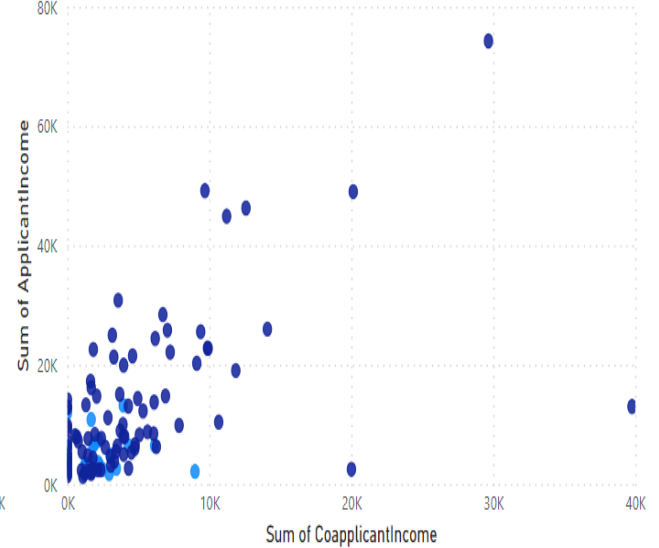
Sum of CoapplicantIncome and Sum of ApplicantIncome by LoanAmount and Loan_Status

Loan_Status ● 0 ● 1



Sum of CoapplicantIncome and Sum of ApplicantIncome by LoanAmount and Predicted_Loan_Status

Predicted_Loan_Status ● 0 ● 1

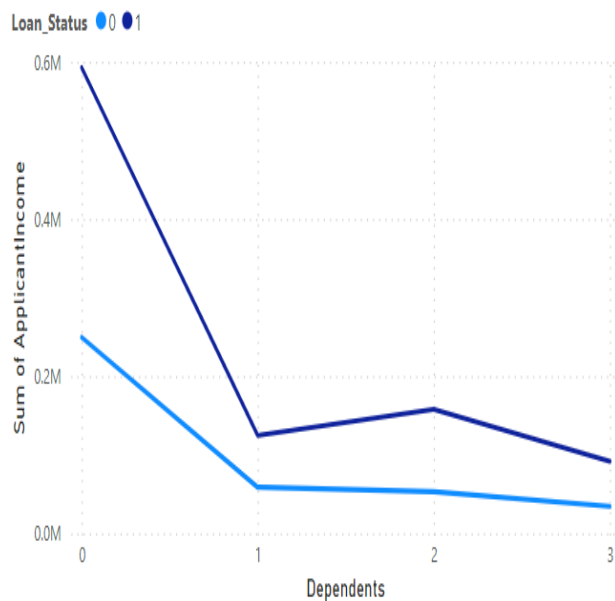


- Matrix visualisation for the loan amount for property value (normalised) vs credit history

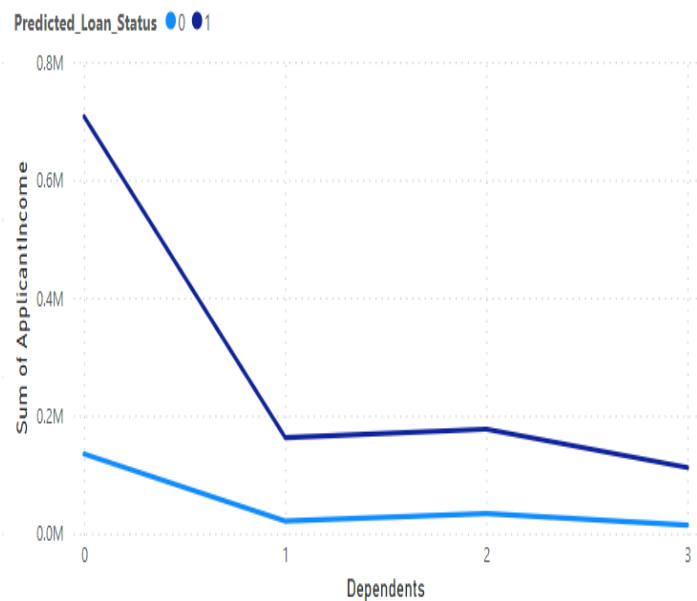
Property_Area	0	1	Total
0	2212	9476	11688
1	1719	10679	12398
3	2232	13682	15914
Total	6163	33837	40000

- The loan status (actual and predicted) based on the income of the house and number of dependents

Sum of ApplicantIncome by Dependents and Loan_Status



Sum of ApplicantIncome by Dependents and Predicted_Loan_Status



6. Loan Status and Predicted_Loan_Status by ApplicantIncome

Sum of Loan_Status and Sum of Predicted_Loan_Status by ApplicantIncome

