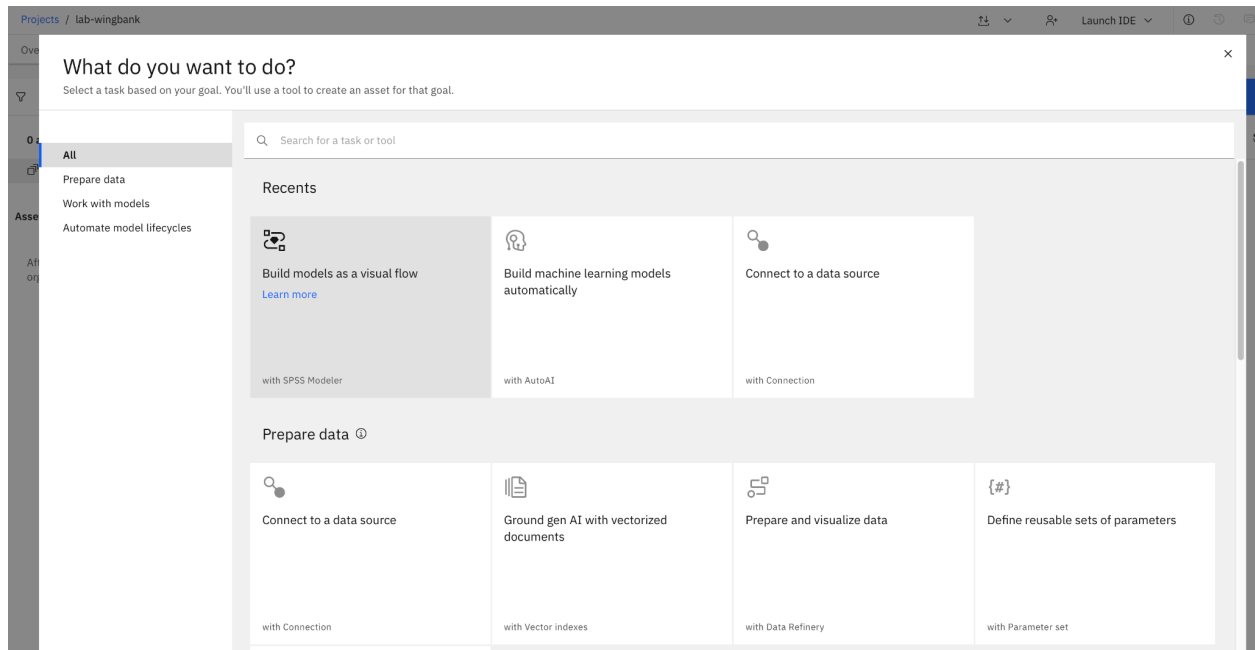
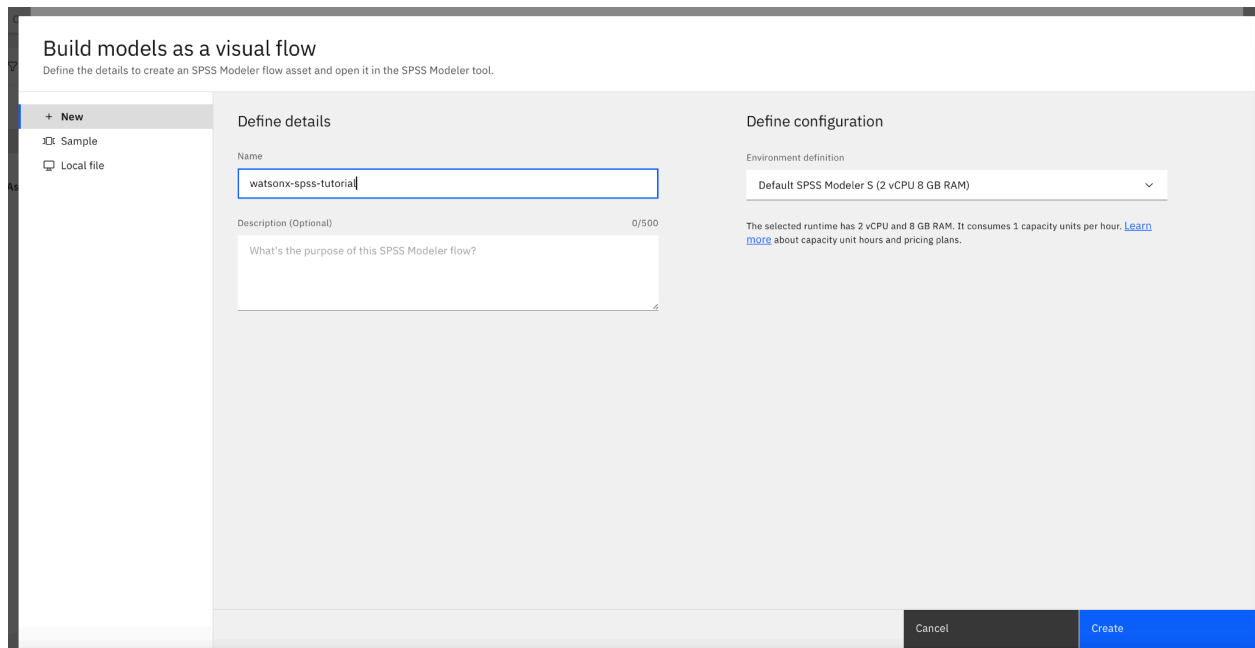


Lab 4 SPSS Modeler

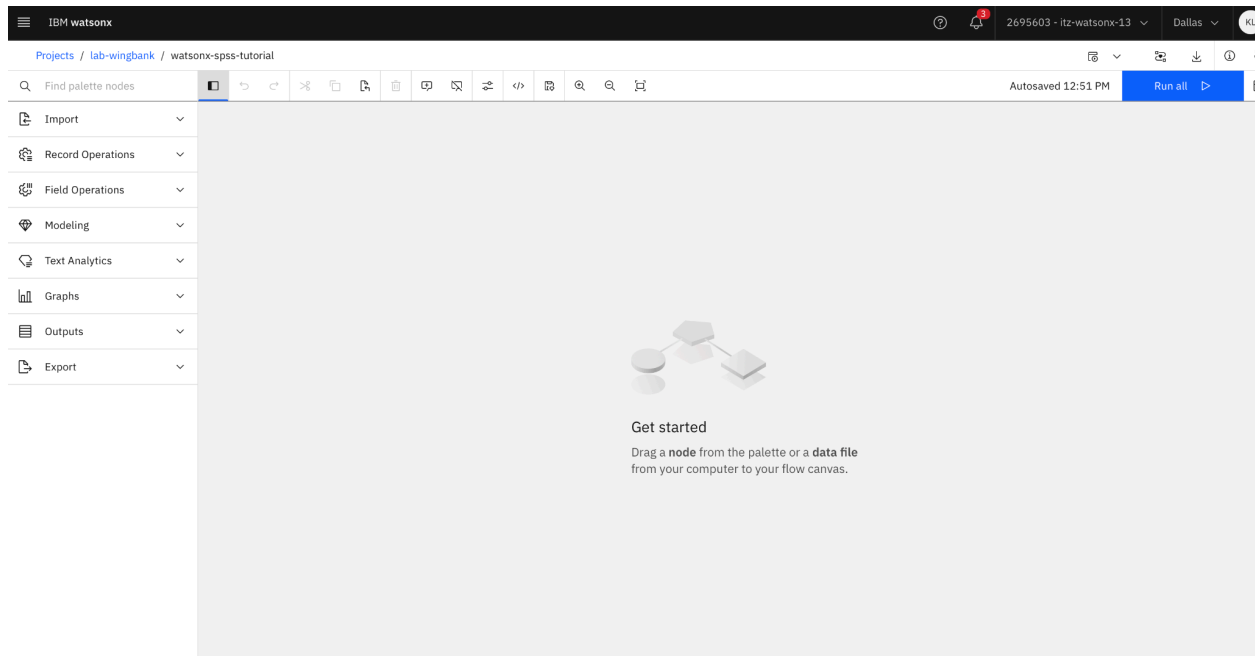
Please select the SPSS Modeler.



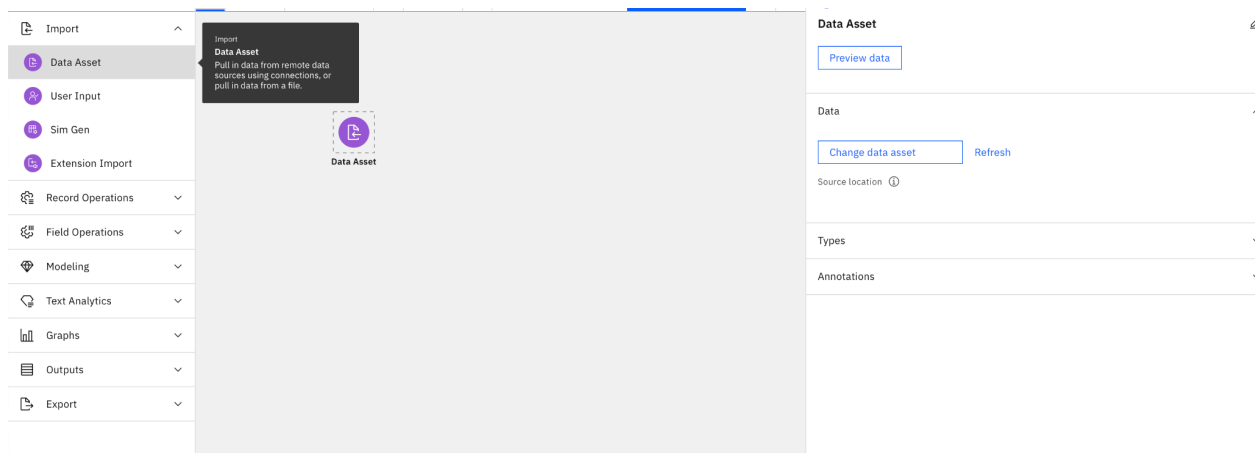
Type in the name of the visual flow. This will be the name of your spss modeler asset.



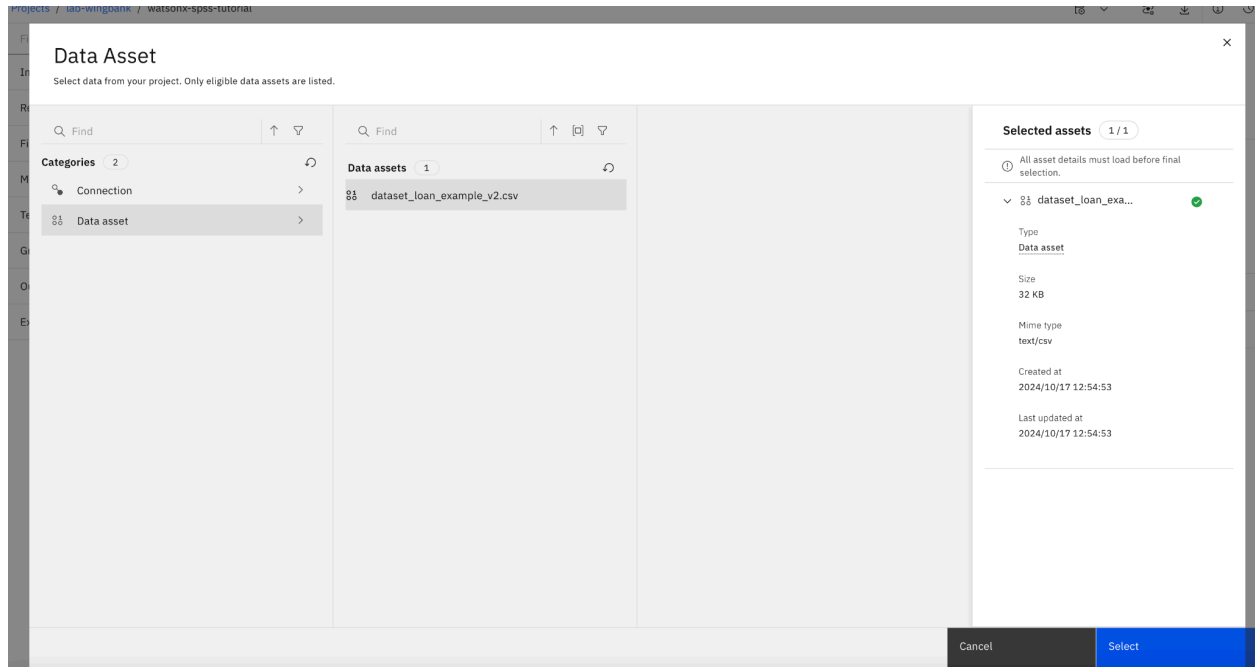
Once the resource is created, you should see this empty canvas.



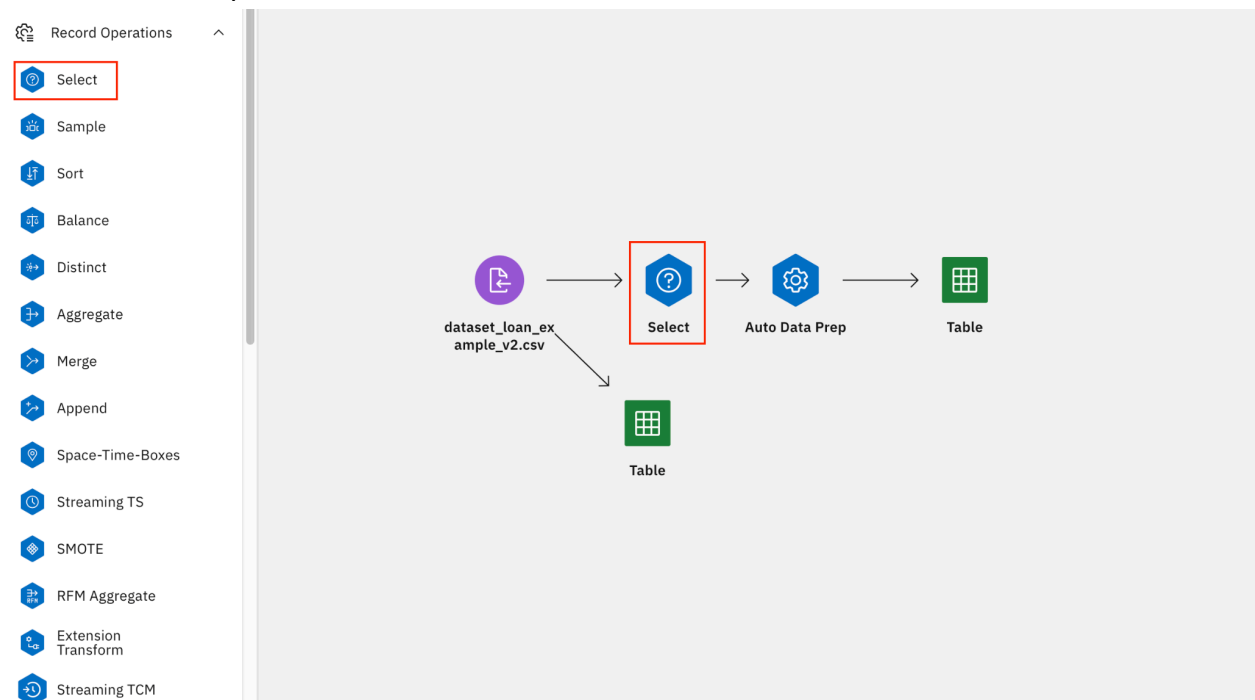
The first step is to create a data asset. Here, please press on the change data asset.

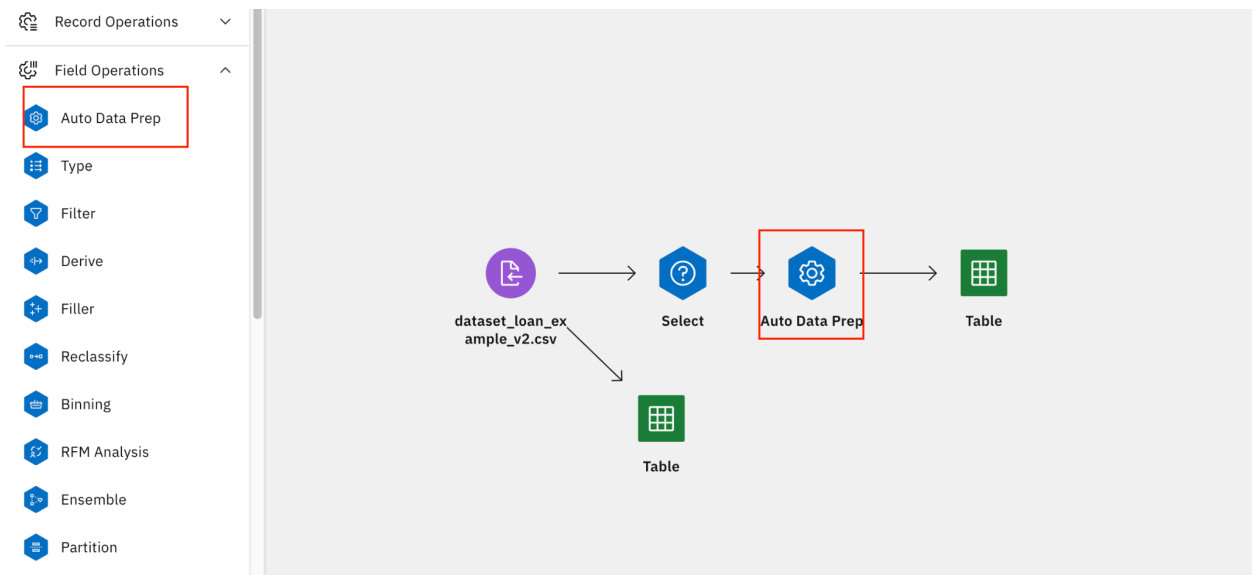


Please locate your `dataset_loan_example_v2.csv`.

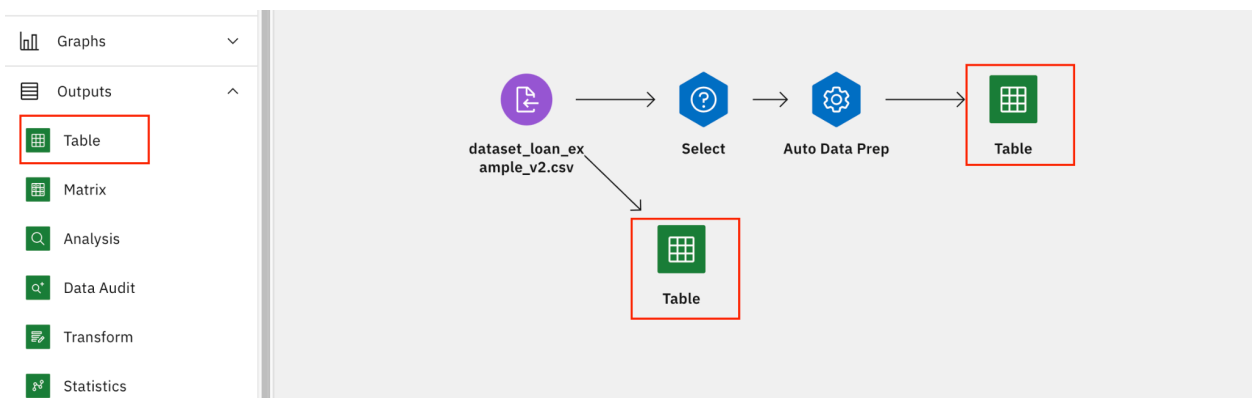


Please drag and drop the select icon under 'record operations' to select the columns you want for Auto Data Preparation.





Once, selected auto data prep, we can add tables to visualize the preprocessed and un-preprocessed data. Please open each table and rename the table icon accordingly to avoid confusion.



Autosaved 1:09 PM Run selected ▶

dataset_loan_example_v2.csv

Select

Auto Data Prep

Table

Table

Preprocessed_data

Run

Settings

Highlight records where ⓘ

Validate

1

Annotations

Cancel Save

Find palette nodes Autosaved 1:11 PM Run all ▶

dataset_loan_example_v2.csv

Select

Auto Data Prep

Preprocessed_data

original_data

Outputs

Models

Outputs are available during your SPSS Modeler session. After your session ends, the outputs are no longer available.


All results

- original_data (12 fields, 61... Just now
- Preprocessed_data (12 fiel... Just now
- Table (12 fields, 614... 2 minutes ago
- Table (12 fields, 614 recor... 3 minutes ago

Press on run all, once ran you should get to see the preprocessed data and the original data table.

View Output: original_data (12 fields, 614 records)


Compare  

Select all data 

	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Y
1	Male	No	0	Graduate	No	5849	0.000	null	360	1	Urban	Y
2	Male	Yes	1	Graduate	No	4583	1508.000	128	360	1	Rural	N
3	Male	Yes	0	Graduate	Yes	3000	0.000	66	360	1	Urban	Y
4	Male	Yes	0	Not Graduate	No	2583	2358.000	120	360	1	Urban	Y
5	Male	No	0	Graduate	No	6000	0.000	141	360	1	Urban	Y
6	Male	Yes	2	Graduate	Yes	5417	4196.000	267	360	1	Urban	Y
7	Male	Yes	0	Not Graduate	No	2333	1516.000	95	360	1	Urban	Y
8	Male	Yes	3	Graduate	No	3036	2504.000	158	360	0	Semiurban	N
9	Male	Yes	2	Graduate	No	4006	1526.000	168	360	1	Urban	Y
10	Male	Yes	1	Graduate	No	12841	10968.000	349	360	1	Semiurban	N
11	Male	Yes	2	Graduate	No	3200	700.000	70	360	1	Urban	Y
12	Male	Yes	2	Graduate		2500	1840.000	109	360	1	Urban	Y
13	Male	Yes	2	Graduate	No	3073	8106.000	200	360	1	Urban	Y
14	Male	No	0	Graduate	No	1853	2840.000	114	360	1	Rural	N
15	Male	Yes	2	Graduate	No	1299	1086.000	17	120	1	Urban	Y
16	Male	No	0	Graduate	No	4950	0.000	125	360	1	Urban	Y
17	Male	No	1	Not Graduate	No	3596	0.000	100	240	null	Urban	Y
18	Female	No	0	Graduate	No	3510	0.000	76	360	0	Urban	N
19	Male	Yes	0	Not Graduate	No	4887	0.000	133	360	1	Rural	N
20	Male	Yes	0	Graduate		2600	3500.000	115	null	1	Urban	Y
21	Male	Yes	0	Not Graduate	No	7660	0.000	104	360	0	Urban	N

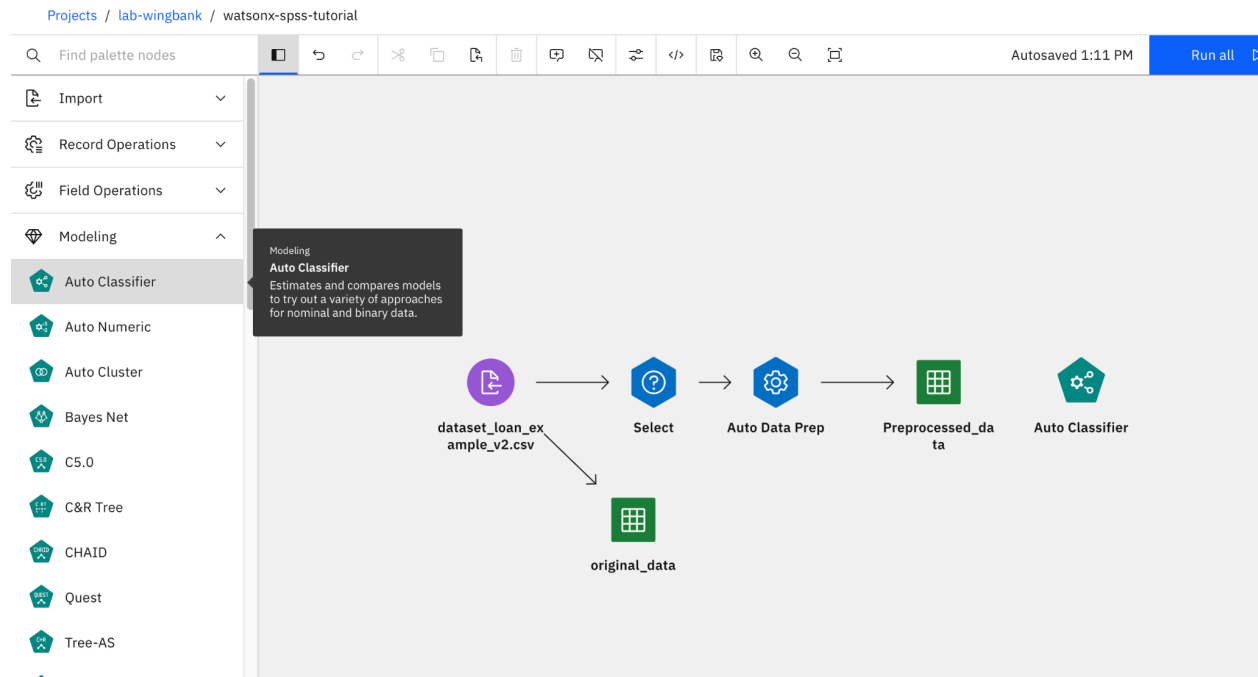
View Output: Preprocessed_data (12 fields, 614 records)

Compare  

Select all data 

	Dependents_transformed	Credit_History_transformed	ApplicantIncome_transformed	CoapplicantIncome_transformed	LoanAmount_transformed	Loan_Amount_Term_transformed	Gender_transfor
1	0	1	0.073	-0.554	0.000	0.280	2
2	1	1	-0.134	-0.039	-0.219	0.280	2
3	0	1	-0.393	-0.554	-0.957	0.280	2
4	0	1	-0.462	0.252	-0.314	0.280	2
5	0	1	0.098	-0.554	-0.064	0.280	2
6	2	1	0.002	0.880	1.435	0.280	2
7	0	1	-0.503	-0.036	-0.612	0.280	2
8	3	0	-0.388	0.302	0.138	0.280	2
9	2	1	-0.229	-0.033	0.257	0.280	2
10	1	1	1.217	3.194	2.411	0.280	2
11	2	1	-0.361	-0.315	-0.909	0.280	2
12	2	1	-0.475	0.075	-0.445	0.280	2
13	2	1	-0.381	2.216	0.638	0.280	2
14	0	1	-0.581	0.416	-0.386	0.280	2
15	2	1	-0.672	-0.183	-1.540	-3.449	2
16	0	1	-0.074	-0.554	-0.255	0.280	2
17	1	null	-0.296	-0.554	-0.552	-1.585	2
18	0	0	-0.310	-0.554	-0.838	0.280	1
19		1	-0.085	-0.554	-0.160	0.280	2
20	0	1	-0.459	0.642	-0.374	0.000	2
21	0	0	0.369	-0.554	-0.505	0.280	2

Next we will be creating an Auto AI Classifier. Drag and drop a line between auto data prep and auto classifier. The “type” icon will automatically be created.



Here, please change the field type to target for “Y_transform”.

Autosaved 3:03 PM

Run selected

Type

Type

Preview data

Settings

Read values to instantiate data

Read values Clear values Clear all values

Find in column Field

Field	Me...	R...	Val...	Values
LoanAmo	Continu	Inpu	Instan	-1.635129...
Loan_Am	Continu	Inpu	Instan	-5.126413...
Gender_tr	Nomine	Inpu	Instan	0, 1, 2
Married_t	Nomine	Inpu	Instan	0, 1, 2
Education	Flag	Inpu	Instan	0, 1
Self_Empl	Nomine	Inpu	Instan	0, 1, 2
Property_	Nomine	Inpu	Instan	0, 1, 2
Y_transfoi	Flag	Targi	Instan	0, 1

Default mode

Read metadata Pass (do not scan)

dataset_loan_example_v2.csv

Select

Auto Data Prep

Type

Y_transformed

original_data

Preprocessed_data

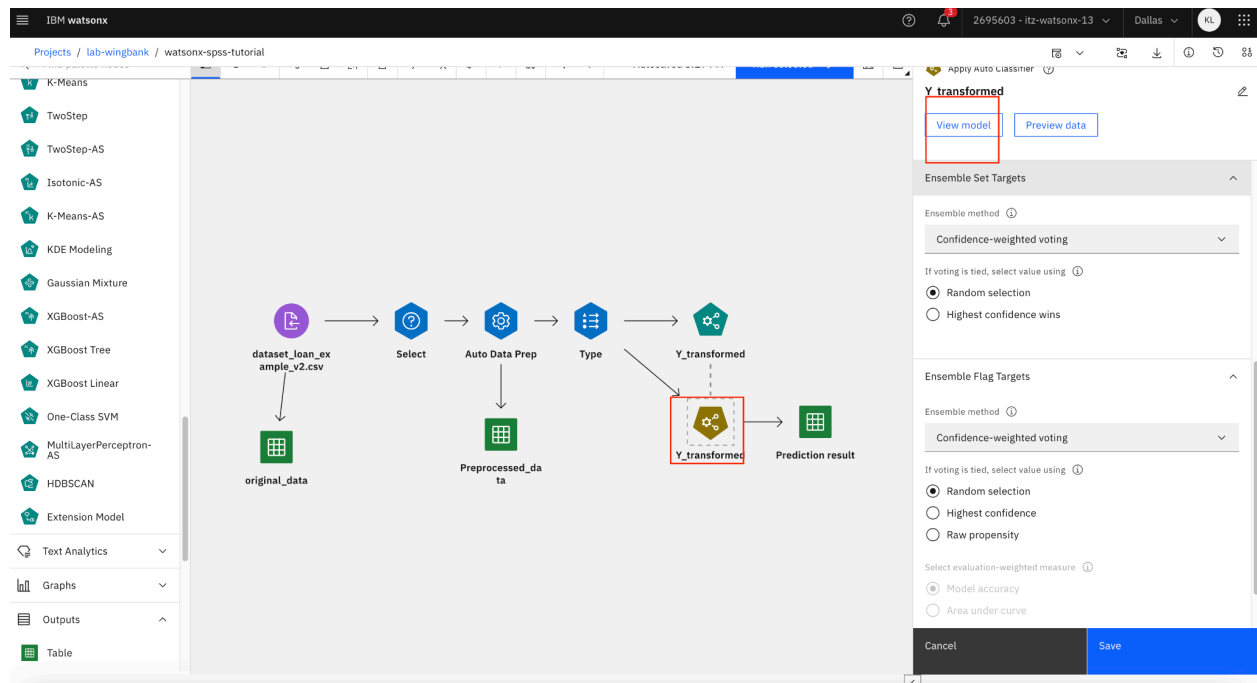
Please select the target as “Y_transformed” and the rest of the columns as inputs.

The screenshot shows the IBM Watsonx SPSS Modeler interface. On the left, a palette of modeling nodes is visible, including Auto Classifier, Auto Numeric, Auto Cluster, Bayes Net, C5.0, C&R Tree, CHAID, Quest, Tree-AS, Random Trees, Random Forest, Decision List, Time Series, and TCM. The main workspace displays a workflow: 'dataset_loan_example_v2.csv' is imported as 'original_data', which flows into a 'Select' node, then an 'Auto Data Prep' node (outputting 'Preprocessed_data'), then a 'Type' node, and finally an 'Auto Classifier' node (outputting 'Y_transformed'). A 'Messages' window is open, showing an error for the 'Random Forest' node: 'Target field is required.' and 'Execution was interrupted.' The 'Auto Classifier' node's configuration panel on the right shows 'Y_transformed' selected as the 'Target' and a list of input fields including 'LoanAmount_transformed', 'Loan_Amount_Term_transformed', 'Gender_transformed', 'Married_transformed', 'Education_transformed', and 'Full_Featured_transformed'. The 'Splits' section is currently empty.

Here we run all again, to run the auto ai classifier

This screenshot shows the 'Run details' and 'Outputs' panels of the IBM Watsonx SPSS Modeler. The 'Run details' panel indicates that the project 'lab-wingbank' and flow 'watsonx-spss-tutorial' have been executed successfully, with 3 branches and an elapsed time of 00:00:48. The 'Outputs' panel shows a list of results, including 'original_data (12 fields, 61...)' and 'Preprocessed_data (12 fiel...)', both of which were generated 'Just now'. The main workspace shows the same workflow as the previous screenshot, but with the 'Auto Classifier' node now highlighted in green, indicating successful execution.

Now we can view the auto AI pipeline by clicking View Model.



Prediction result

View Output: Prediction result (14 fields, 614 records) [Compare](#)

Select all data

	ormed	Gender_transformed	Married_transformed	Education_transformed	Self_Employed_transformed	Property_Area_transformed	Y_transformed	\$XF-Y_transformed	\$XFC-Y_transformed
1	0.280	2	1	1	2	1	1	1	0.729
2	0.280	2	2	1	2	0	0	1	0.731
3	0.280	2	2	1	1	1	1	1	0.767
4	0.280	2	2	0	2	1	1	1	0.777
5	0.280	2	1	1	2	1	1	1	0.730
6	0.280	2	2	1	1	1	1	1	0.782
7	0.280	2	2	0	2	1	1	1	0.759
8	0.280	2	2	1	2	2	0	0	0.849
9	0.280	2	2	1	2	1	1	1	0.754
10	0.280	2	2	1	2	2	0	1	0.462
11	0.280	2	2	1	2	1	1	1	0.622
12	0.280	2	2	1	0	1	1	1	0.827
13	0.280	2	2	1	2	1	1	1	0.765
14	0.280	2	1	1	2	0	0	1	0.551
15	-3.449	2	2	1	2	1	1	1	0.795
16	0.280	2	1	1	2	1	1	1	0.731
17	-1.585	2	1	0	2	1	1	1	0.716
18	0.280	1	1	1	2	1	0	0	0.907
19	0.280	2	2	0	2	0	0	1	0.562
20	0.000	2	2	1	0	1	1	1	0.819
21	0.280	2	2	0	2	1	0	0	0.900

Here is the model trained and ranked.

View Model: Y_transformed

Auto Classifier ③

Models

Auto Classifier - Models ③

TARGET : Y_TRANSFORMED

USE	MODEL NAME	ESTIMATOR	BUILD TIME (MINS)	NO. FIELDS USED	ACCURACY	ACCUMULATED ACCURACY	AREA UNDER CURVE	ACCUMULATED AUC	RECALL	PRECISION	ACTIONS
<input checked="" type="checkbox"/>	Random Trees 1	Random Trees	< 1	11	89.577	89.577	0.956	0.956	0.912	0.934	
<input checked="" type="checkbox"/>	Logistic regression 1	Nominal Regression	< 1	11	72.801	72.801	0.784	0.784	0.981	0.795	
<input checked="" type="checkbox"/>	LSVM 1	Linear SVM	< 1	11	80.945	80.945	0.794	0.794	0.983	0.790	
<input checked="" type="checkbox"/>	Tree-AS 1	CHAID	< 1	2	80.945	80.945	0.758	0.758	0.983	0.790	
<input checked="" type="checkbox"/>	Bayesian Network 1	BayesNet	< 1	11	74.756	74.756	0.826	0.826	0.979	0.818	

[1]

Auto Classification Node evaluation measure may be different from model detail, since it comes from different analysis stage, with different computation method, on different test datasets.