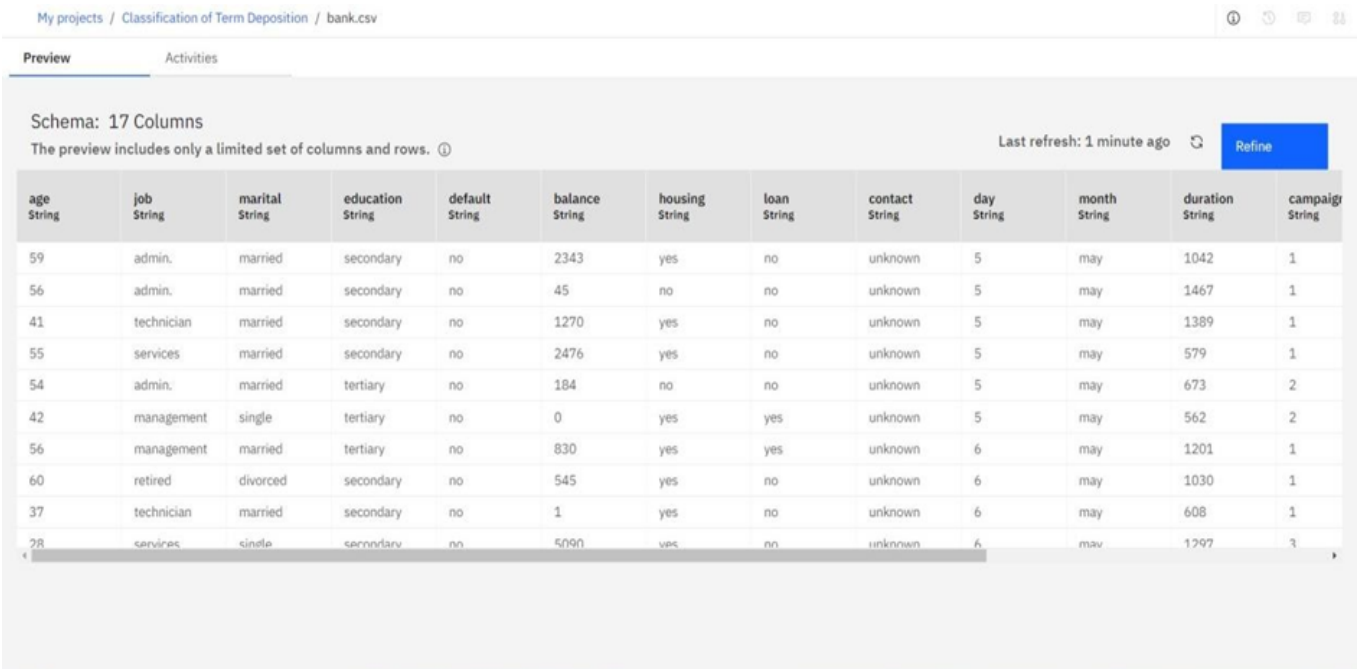
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

DATA ANALYTICS

Assignment-3

NAME: Dhinesh.B
REG NO: 19BEC0684
CAMPUS: Vellore

Classification of term deposit buyers for retail banking



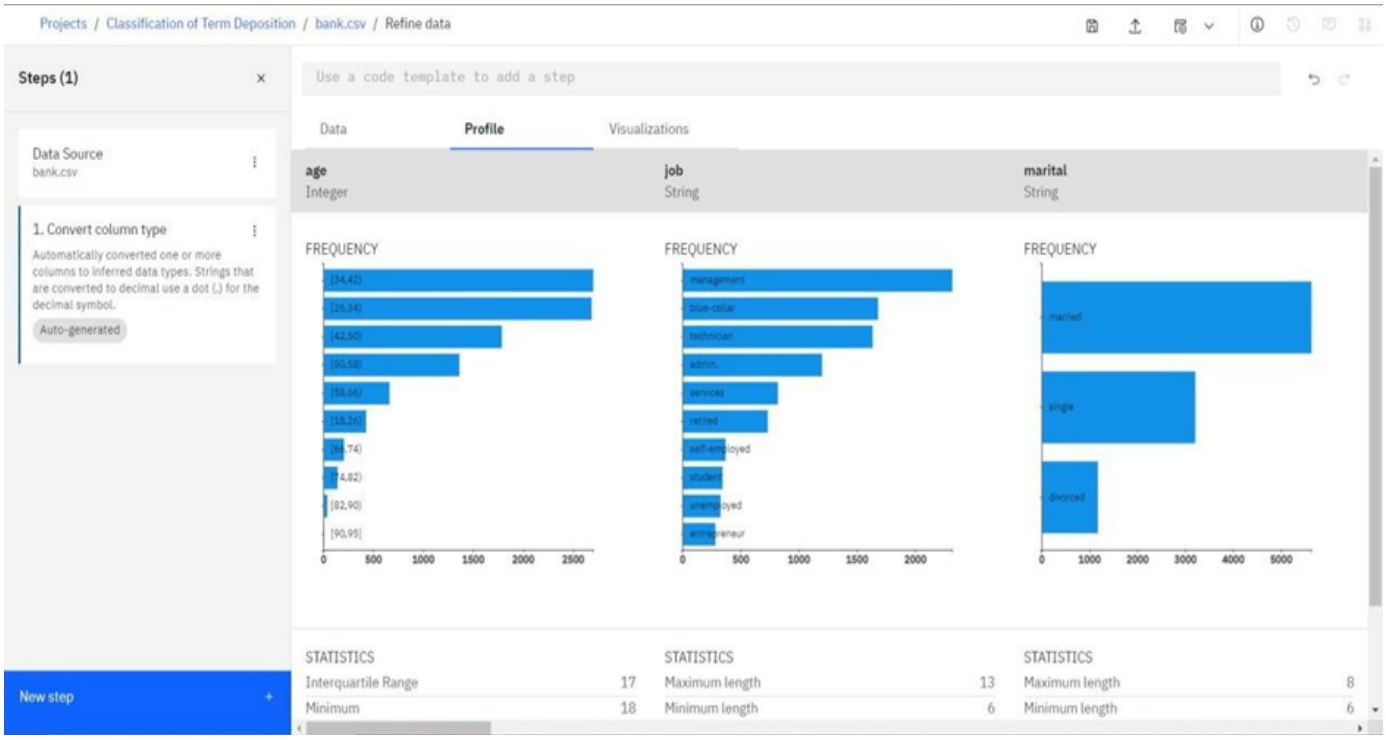
My projects / Classification of Term Deposition / bank.csv

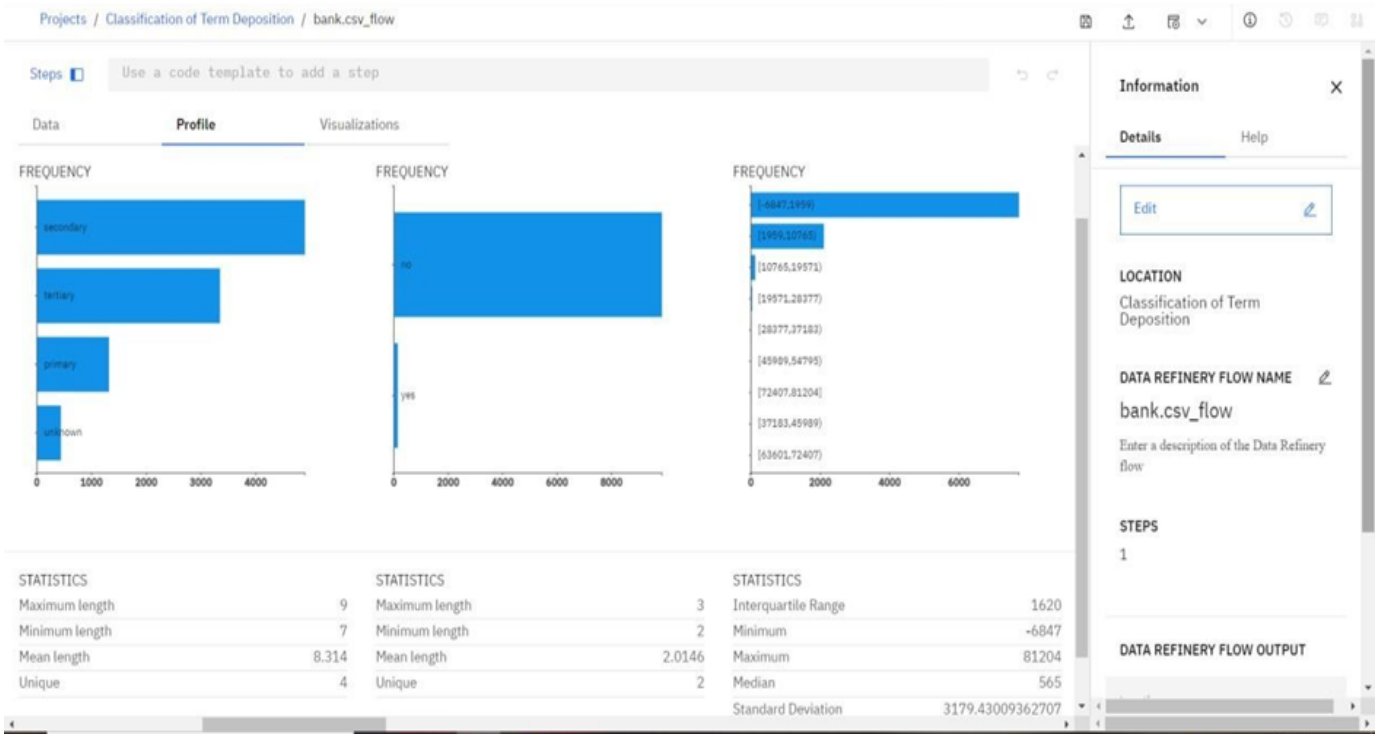
Preview | Activities

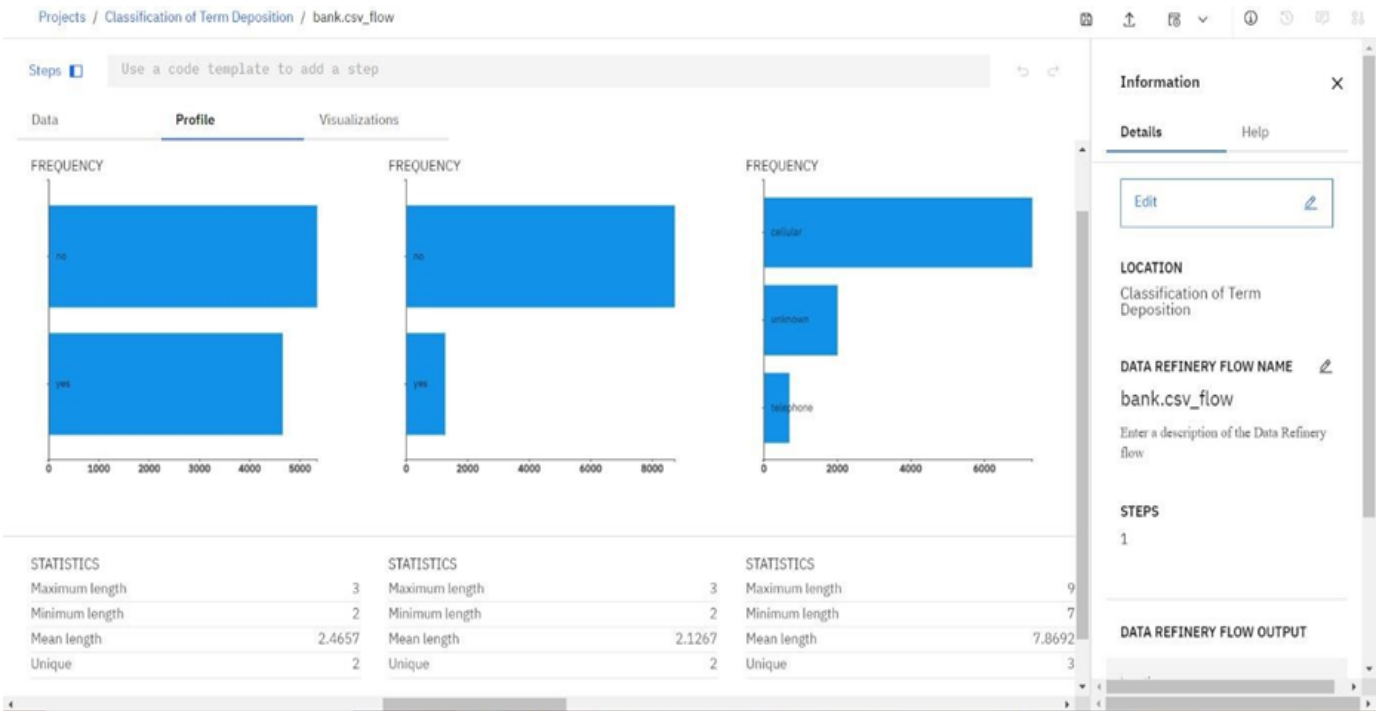
Schema: 17 Columns
The preview includes only a limited set of columns and rows.

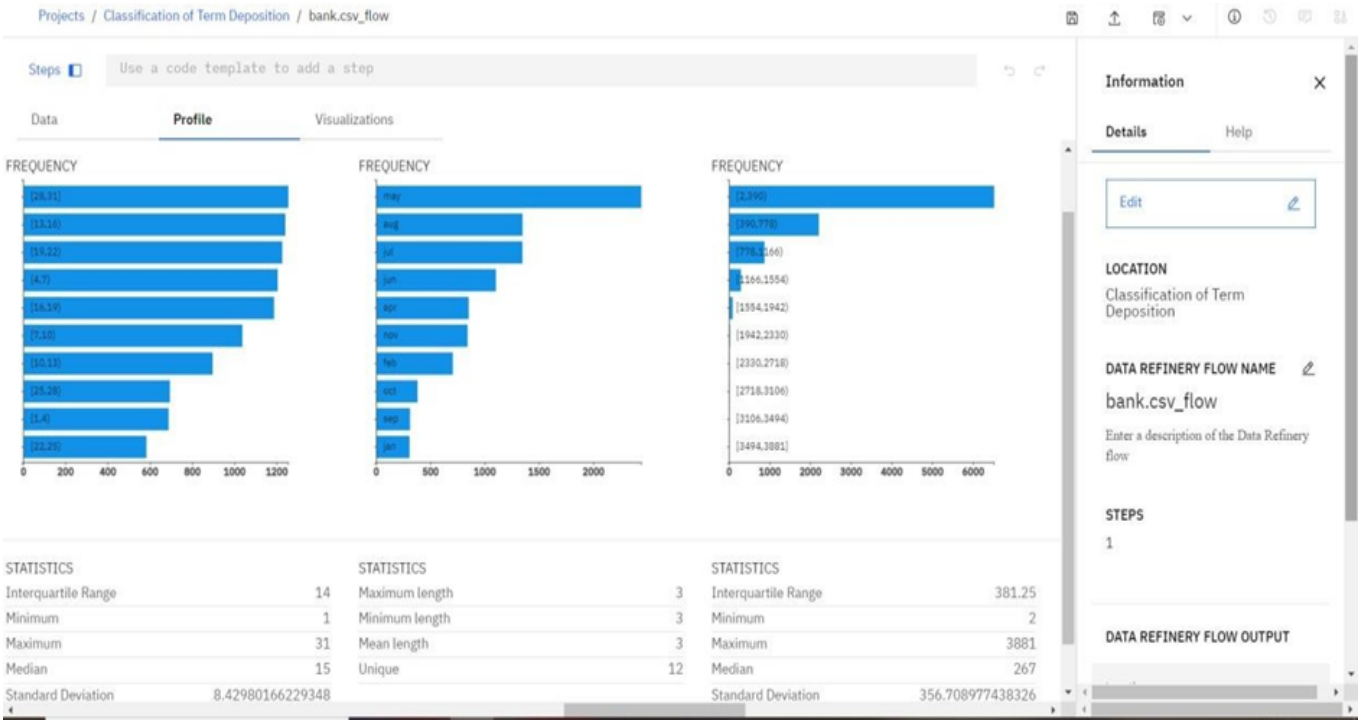
Last refresh: 1 minute ago | Refine

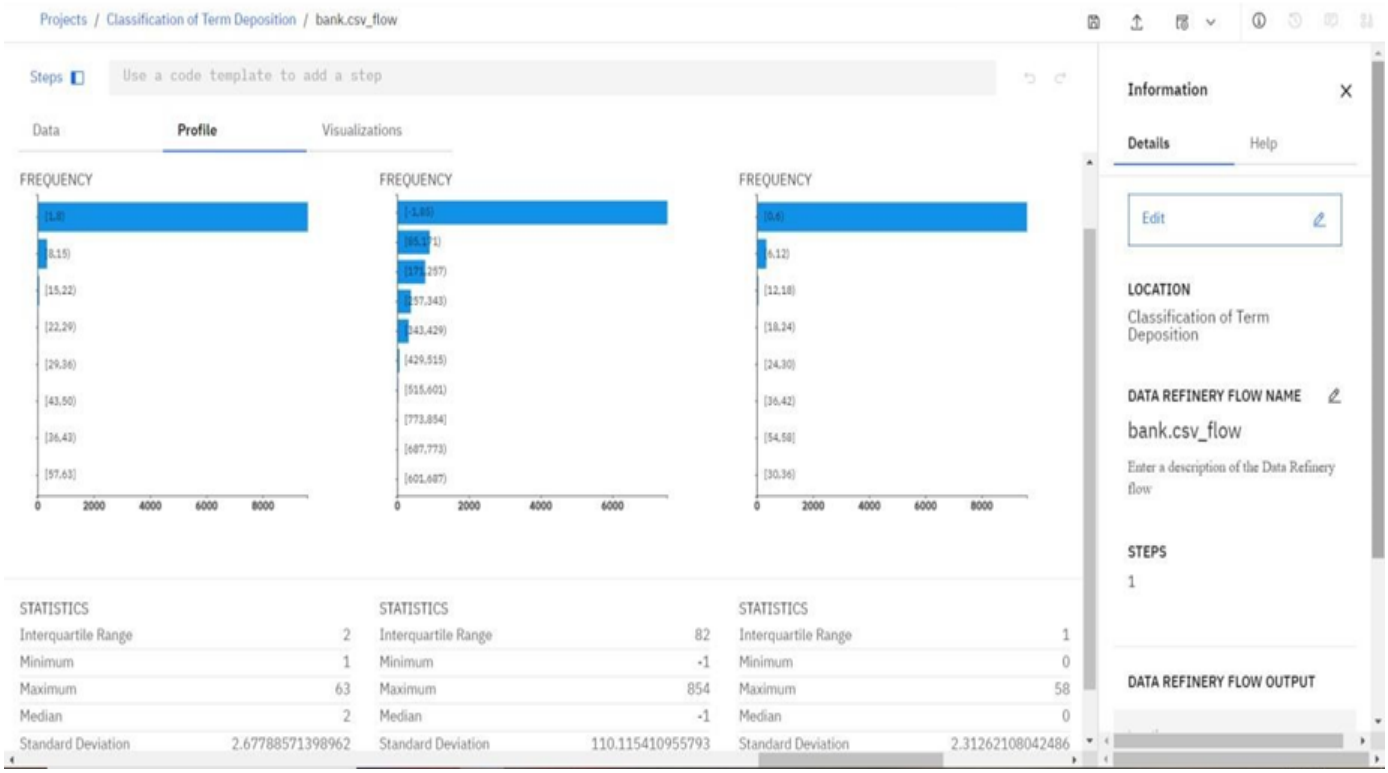
age String	job String	marital String	education String	default String	balance String	housing String	loan String	contact String	day String	month String	duration String	campaig String
59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1
56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1
41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1
55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1
54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2
42	management	single	tertiary	no	0	yes	yes	unknown	5	may	562	2
56	management	married	tertiary	no	830	yes	yes	unknown	6	may	1201	1
60	retired	divorced	secondary	no	545	yes	no	unknown	6	may	1030	1
37	technician	married	secondary	no	1	yes	no	unknown	6	may	608	1
28	services	single	secondary	no	5090	yes	no	unknown	6	may	1297	3

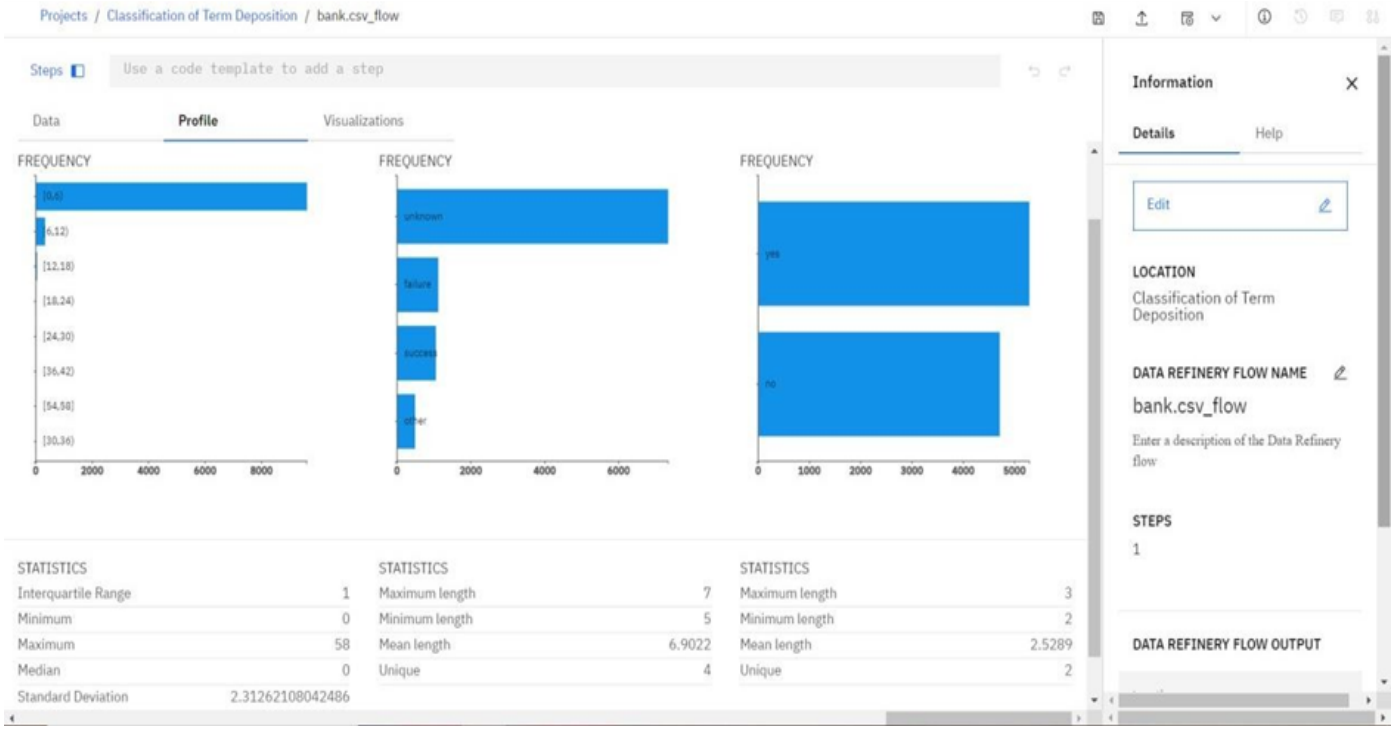

Projects / Classification of Term Deposition / bank.csv / Refine data
Steps (1)
Use a code template to add a step
Data Source
bank.csv
1. Convert column type
Automatically converted one or more columns to inferred data types. Strings that are converted to decimal use a dot (.) for the decimal symbol.
Auto-generated
Data
Profile
Visualizations
age
Integer
job
String
marital
String
FREQUENCY
management
blue-collar
technician
admin.
services
retired
self-employed
student
unemployed
entrepreneur
married
single
divorced
STATISTICS
Interquartile Range 17
Minimum 18
Maximum length 13
Minimum length 6
Maximum length 8
Minimum length 6
New step

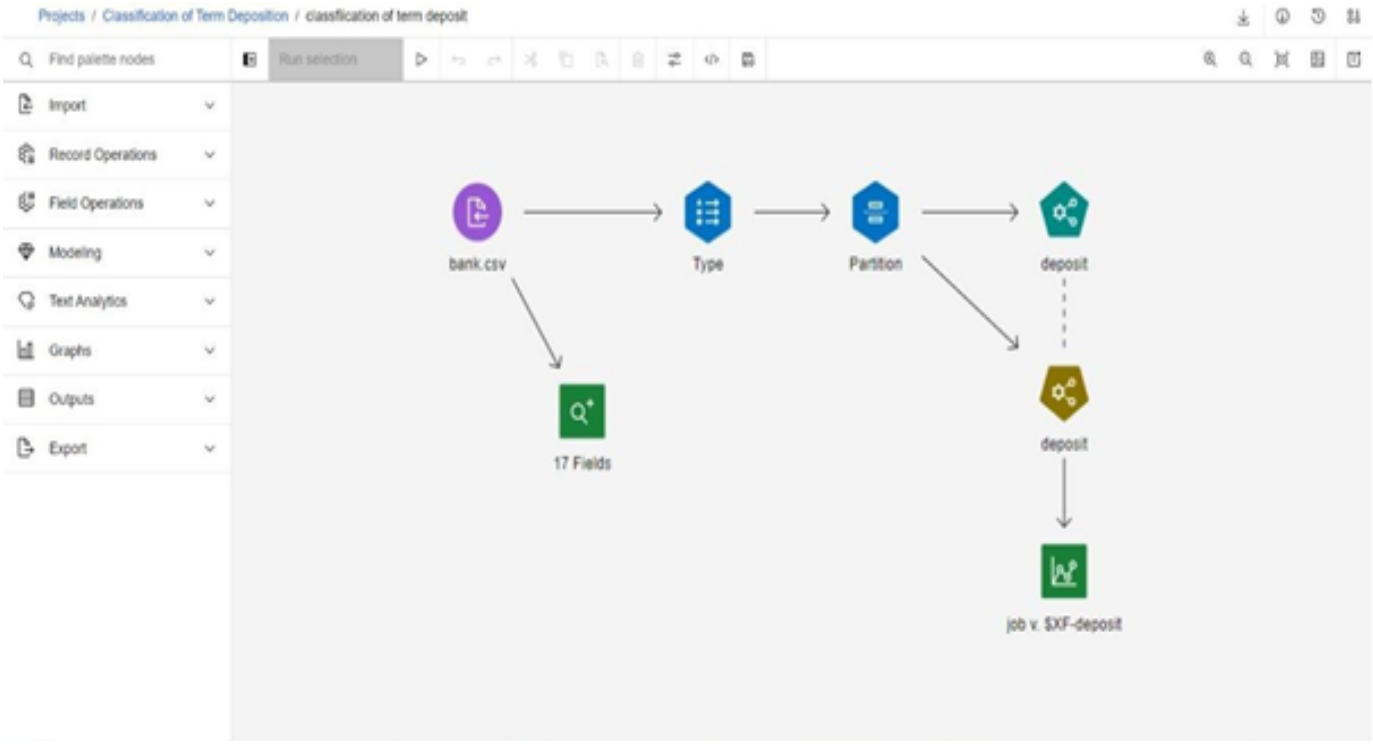

Projects / Classification of Term Deposition / bank.csv_flow
Steps
Use a code template to add a step
Data
Profile
Visualizations
FREQUENCY
secondary
tertiary
primary
unknown
yes
STATISTICS
Maximum length 9
Minimum length 7
Mean length 8.314
Unique 4
Maximum length 3
Minimum length 2
Mean length 2.0146
Unique 2
Interquartile Range 1620
Minimum -6847
Maximum 81204
Median 565
Standard Deviation 3179.43009362707
Information
Details
Help
Edit
LOCATION
Classification of Term Deposition
DATA REFINERY FLOW NAME
bank.csv_flow
Enter a description of the Data Refinery flow
STEPS
1
DATA REFINERY FLOW OUTPUT

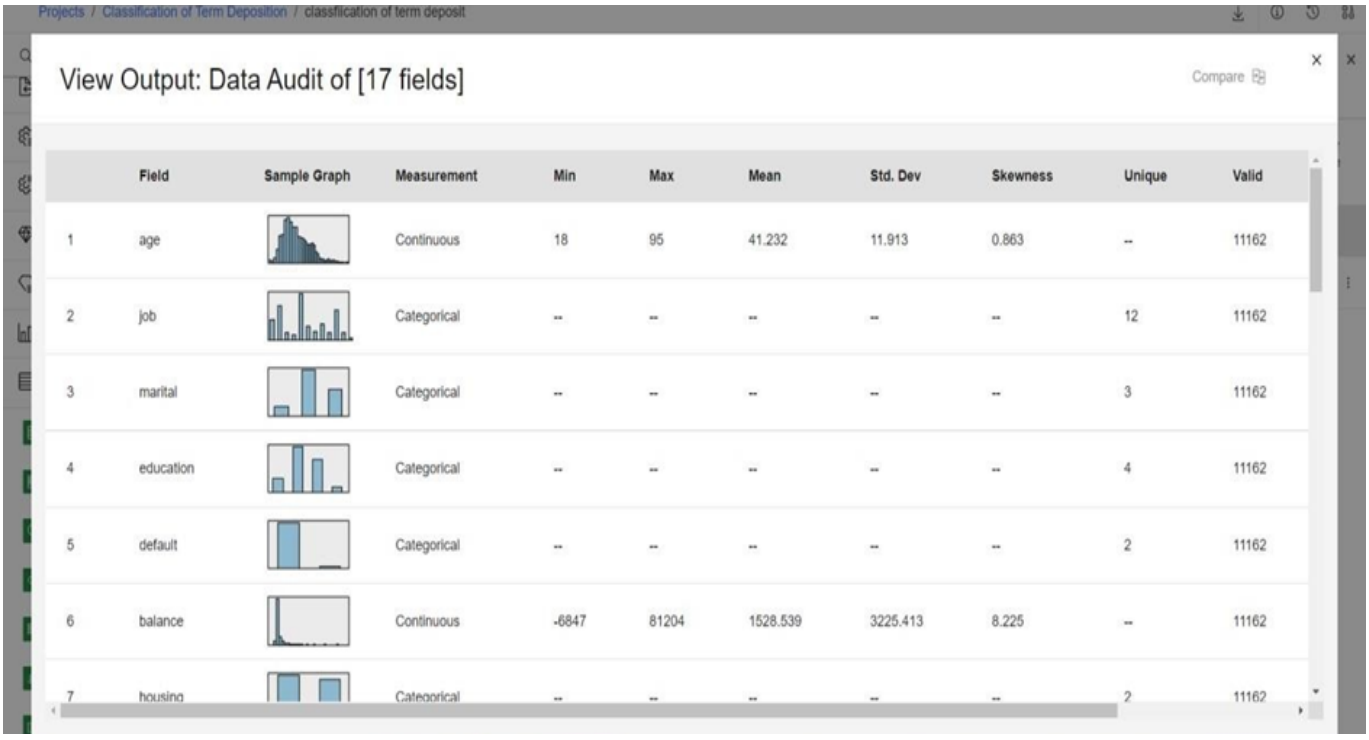

Projects / Classification of Term Deposition / bank.csv_flow
Steps
Use a code template to add a step
Data
Profile
Visualizations
FREQUENCY
no
yes
0 1000 2000 3000 4000 5000
FREQUENCY
no
yes
0 2000 4000 6000 8000
FREQUENCY
cellular
unknown
telephone
0 2000 4000 6000
STATISTICS
Maximum length 3
Minimum length 2
Mean length 2.4657
Unique 2
STATISTICS
Maximum length 3
Minimum length 2
Mean length 2.1267
Unique 2
STATISTICS
Maximum length 9
Minimum length 7
Mean length 7.8692
Unique 3
Information
Details
Help
Edit
LOCATION
Classification of Term Deposition
DATA REFINERY FLOW NAME
bank.csv_flow
Enter a description of the Data Refinery flow
STEPS
1
DATA REFINERY FLOW OUTPUT


Projects / Classification of Term Deposition / bank.csv_flow
Steps
Use a code template to add a step
Data
Profile
Visualizations
FREQUENCY
[28,31]
[13,16]
[19,22]
[4,7]
[16,19]
[7,10]
[10,13]
[25,28]
[1,4]
[22,25]
0 200 400 600 800 1000 1200
FREQUENCY
may
aug
jul
jun
apr
nov
feb
oct
sep
jan
0 500 1000 1500 2000
FREQUENCY
[2,390]
[390,778]
[778,1166]
[1166,1554]
[1554,1942]
[1942,2330]
[2330,2718]
[2718,3106]
[3106,3494]
[3494,3881]
0 1000 2000 3000 4000 5000 6000
STATISTICS
Interquartile Range 14
Minimum 1
Maximum 31
Median 15
Standard Deviation 8.42980166229348
STATISTICS
Maximum length 3
Minimum length 3
Mean length 3
Unique 12
STATISTICS
Interquartile Range 381.25
Minimum 2
Maximum 3881
Median 267
Standard Deviation 356.708977438326
Information
Details
Help
Edit
LOCATION
Classification of Term Deposition
DATA REFINERY FLOW NAME
bank.csv_flow
Enter a description of the Data Refinery flow
STEPS
1
DATA REFINERY FLOW OUTPUT

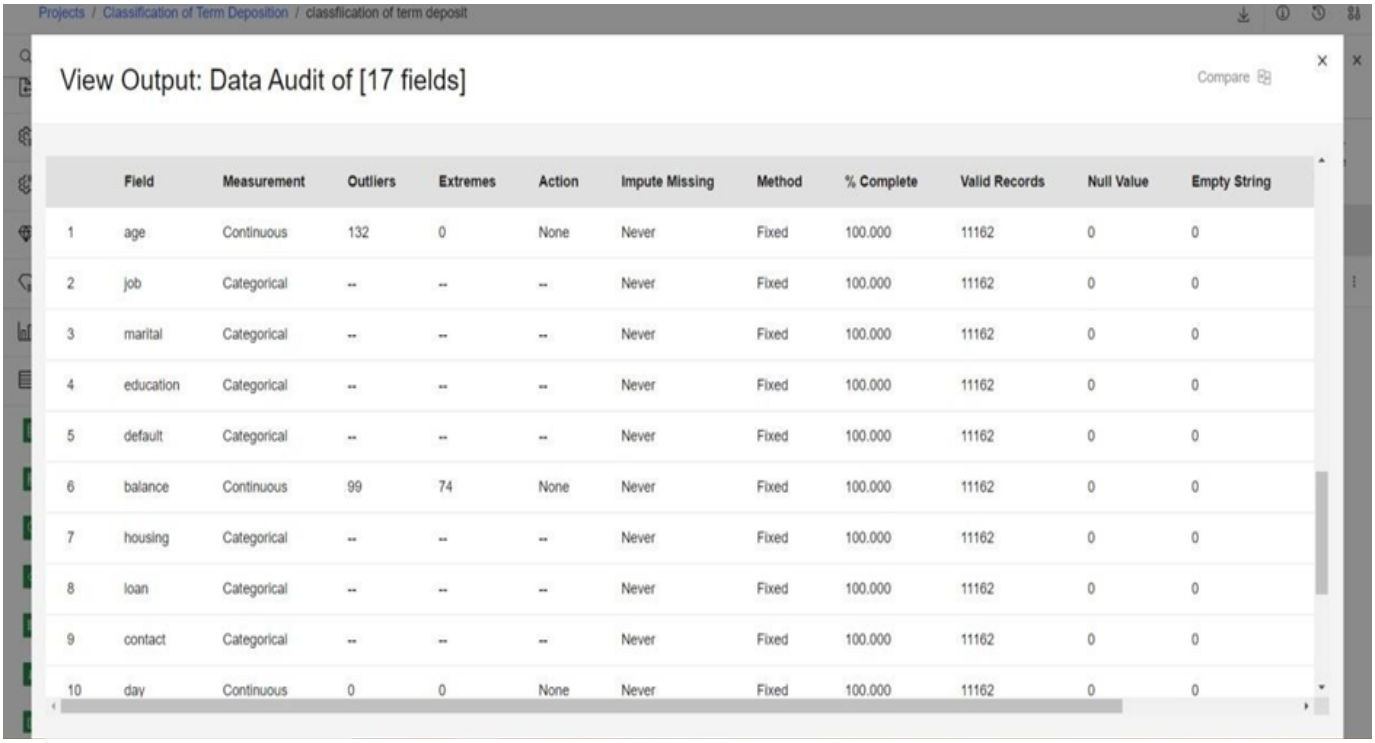

Projects / Classification of Term Deposition / bank.csv_flow
Steps
Use a code template to add a step
Data
Profile
Visualizations
FREQUENCY
STATISTICS
Interquartile Range 2
Minimum 1
Maximum 63
Median 2
Standard Deviation 2.67788571398962
STATISTICS
Interquartile Range 82
Minimum -1
Maximum 854
Median -1
Standard Deviation 110.115410955793
STATISTICS
Interquartile Range 1
Minimum 0
Maximum 58
Median 0
Standard Deviation 2.31262108042486
Information
Details
Help
Edit
LOCATION
Classification of Term Deposition
DATA REFINERY FLOW NAME
bank.csv_flow
Enter a description of the Data Refinery flow
STEPS
1
DATA REFINERY FLOW OUTPUT


Projects / Classification of Term Deposition / bank.csv_flow
Steps
Use a code template to add a step
Data
Profile
Visualizations
FREQUENCY
STATISTICS
Interquartile Range 1
Minimum 0
Maximum 58
Median 0
Standard Deviation 2.31262108042486
STATISTICS
Maximum length 7
Minimum length 5
Mean length 6.9022
Unique 4
STATISTICS
Maximum length 3
Minimum length 2
Mean length 2.5289
Unique 2
Information
Details
Help
Edit
LOCATION
Classification of Term Deposition
DATA REFINERY FLOW NAME
bank.csv_flow
Enter a description of the Data Refinery flow
STEPS
1
DATA REFINERY FLOW OUTPUT





	Field	Sample Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1	age		Continuous	18	95	41.232	11.913	0.863	--	11162
2	job		Categorical	--	--	--	--	--	12	11162
3	marital		Categorical	--	--	--	--	--	3	11162
4	education		Categorical	--	--	--	--	--	4	11162
5	default		Categorical	--	--	--	--	--	2	11162
6	balance		Continuous	-6847	81204	1528.539	3225.413	8.225	--	11162
7	housing		Categorical	--	--	--	--	--	2	11162



	Field	Measurement	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value	Empty String
1	age	Continuous	132	0	None	Never	Fixed	100.000	11162	0	0
2	job	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
3	marital	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
4	education	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
5	default	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
6	balance	Continuous	99	74	None	Never	Fixed	100.000	11162	0	0
7	housing	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
8	loan	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
9	contact	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0
10	day	Continuous	0	0	None	Never	Fixed	100.000	11162	0	0

Type

Type

Read Values | Clear All Values

Find in column Field

Field	Measure	Role	Value mode	Values	Check
# age	Continuous	None	Instantiated	18, 95	None
abc job	Nominal	Input	Instantiated	admin., blue-collar...	None
abc marital	Nominal	Input	Instantiated	divorced, married, ...	None
abc education	Nominal	None	Instantiated	primary, secondar...	None
abc default	Flag	None	Instantiated	no, yes	None
# balance	Continuous	None	Instantiated	-6847, 81204	None
abc housing	Flag	None	Instantiated	no, yes	None
abc loan	Flag	None	Instantiated	no, yes	None

Cancel | Save

Type

Type

Find in column Field

Field	Measure	Role	Value mode	Values	Check
abc contact	Nominal	None	Instantiated	cellular, telephone...	None
# day	Continuous	None	Instantiated	1, 31	None
abc month	Nominal	None	Instantiated	apr, aug, dec, feb, ...	None
# duration	Continuous	None	Instantiated	2, 3881	None
# campaign	Continuous	None	Instantiated	1, 63	None
# pdays	Continuous	None	Instantiated	-1, 854	None
# previous	Continuous	None	Instantiated	0, 58	None
abc poutcome	Nominal	None	Instantiated	failure, other, succ...	None

Format

Type

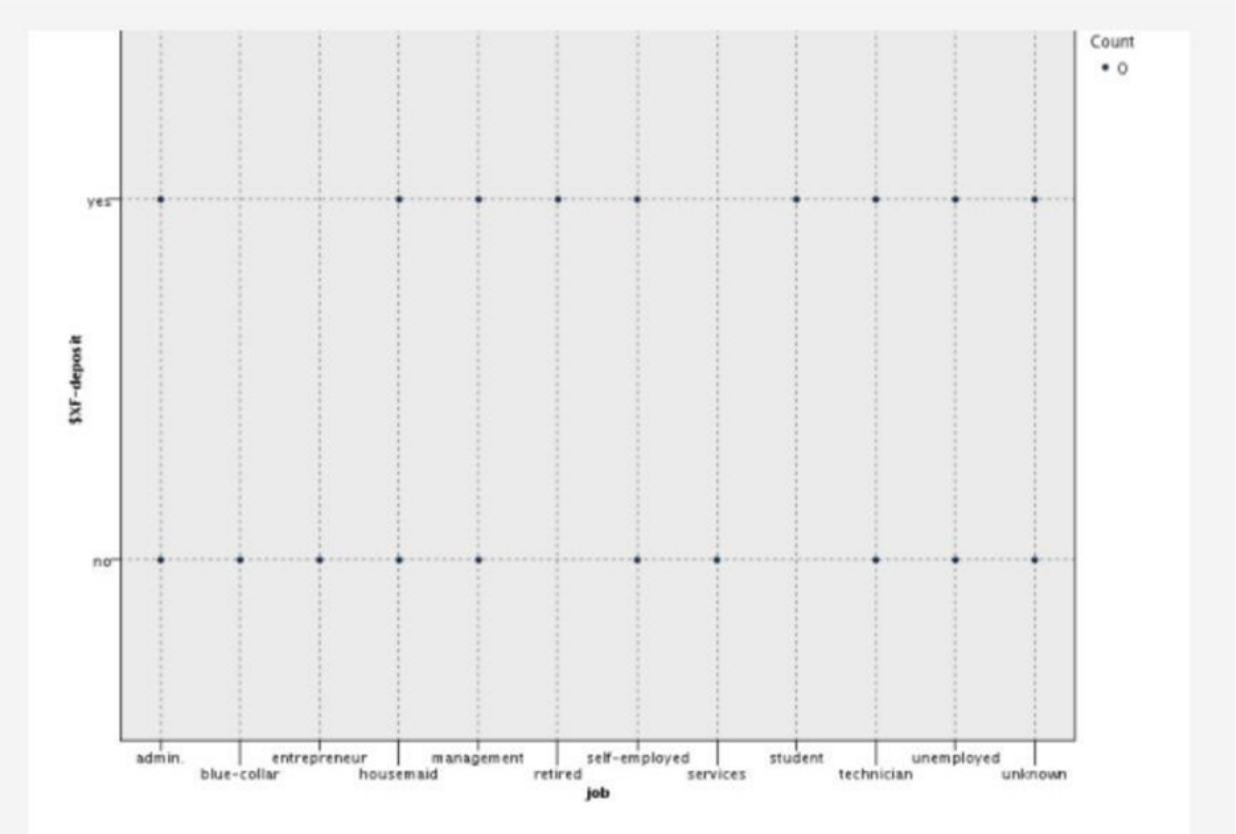
Type

Find in column Field

Field	Measure	Role	Value mode	Values	Check
# day	Continuous	None	Instantiated	1, 31	None
abc month	Nominal	None	Instantiated	apr, aug, dec, feb, ...	None
# duration	Continuous	None	Instantiated	2, 3881	None
# campaign	Continuous	None	Instantiated	1, 63	None
# pdays	Continuous	None	Instantiated	-1, 854	None
# previous	Continuous	None	Instantiated	0, 58	None
abc poutcome	Nominal	None	Instantiated	failure, other, succ...	None
abc deposit	Flag	Target	Instantiated	no, yes	None

Format

Cancel Save



View Model: deposit

Auto Classifier ⓘ

Models

Auto Classifier - Models ⓘ

TARGET : DEPOSIT

USE	MODEL NAME	ESTIMATOR	BUILD TIME (MINS)	NO. FIELDS USED	ACCURACY	ACCUMULATED ACCURACY	AREA UNDER CURVE	ACCUMULATED AUC	RECALL	PRECISION	ACTIONS
☑	Random Trees 1	Random Trees	< 1	2	57.295	57.295	0.595	0.595	0.581	0.548	
☑	LSVM 1	Linear SVM	< 1	2	57.414	57.414	0.609	0.609	0.408	0.573	
☑	Tree-AS 1	CHAID	< 1	2	57.622	57.622	0.605	0.605	0.378	0.583	
☑	CHAID 1	CHAID	< 1	2	57.622	57.622	0.605	0.605	0.378	0.583	
☑	C5 1	C5.0	< 1	2	57.444	57.444	0.587	0.587	0.388	0.578	