

Data Analytics Externship Program

Assignment-3

Name: Tamanampudi Naveen Sai
Email ID: naveen.19bcd7003@vitap.ac.in

Classification of Term Deposit Buyers for Retail Banking

Dataset Used: bank.csv

Creating the Project

The screenshot shows the 'New project' creation interface in IBM Watson Studio. The URL in the browser is dataplatform.cloud.ibm.com/projects/create-project?context=cpdaas. The form fields are as follows:

- Name:** Assignment - 3
- Description:** Classification of Term Deposit Buyers for retail banking
- Storage:** Cloud Object Storage+ur
- Choose project options:** A checkbox labeled "Restrict who can be a collaborator" is present.
- Project includes integration with Cloud Object Storage for storing project assets.**

At the bottom right, there are "Cancel" and "Create" buttons.

Uploading Data:

The screenshot shows the IBM Watson Studio interface with the 'Assets' tab selected. On the left, there's a sidebar for 'Asset types' with 'Data' selected, showing one data asset. The main area displays a table of assets with columns for 'Name' and 'Last modified'. A single row is shown for 'bank.csv', which is a CSV file uploaded by 'Naveen Sai Tamanampudi (You)' at 'Now'. To the right, a sidebar titled 'Data in this project' has a section for uploading files, with a message 'Drop data files here or browse for files to upload'.

Creating a Data Refinery:

The screenshot shows the 'Select data from project' dialog in IBM Watson Studio. On the left, under 'Assignment - 3', there are sections for 'Connection' and 'Data asset'. In the center, under 'Data assets', there is a list with 'bank.csv' selected. On the right, a 'Selected assets' panel shows 'bank.csv' with its details: Asset name is 'bank.csv', Asset type is 'Data asset', Size is '897 KB', Last modified is '2022/04/25 12:26:30', and Created on is '2022/04/25 12:26:30'. At the bottom right of the dialog are 'Cancel' and 'Select' buttons.

Viewing the output of refinery:

Service Details - IBM Cloud

IBM Watson Studio

Projects / Assignment - 3 / bank.csv / Refine data

Steps (1)

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

New step

Information

LOCATION: Assignment - 3

DATA REFINERY FLOW NAME: bank.csv_flow

Enter a description of the Data Refinery flow

STEPS: 1

DATA REFINERY FLOW OUTPUT

Location: Assignment - 3/Data assets

SOURCE FILE: bank.csv SAMPLE SIZE: First 10000 rows

Viewing Profile:

Service Details - IBM Cloud

IBM Watson Studio

Projects / Assignment - 3 / bank.csv / Refine data

Steps

Use a code template to add a step

Data

Profile

Visualizations

age Integer

job String

marital String

education String

FREQUENCY

[54,42]	2500
[26,34]	2000
[42,50]	1800
[60,68]	1500
[58,66]	1200
[16,24]	1000
[66,74]	800
[14,22]	700
[82,90]	500
[90,98]	300

FREQUENCY

management	2200
blue-collar	1800
technician	1600
admin.	1200
services	1000
retired	800
self-employed	700
student	600
unemployed	500
entrepreneur	400

FREQUENCY

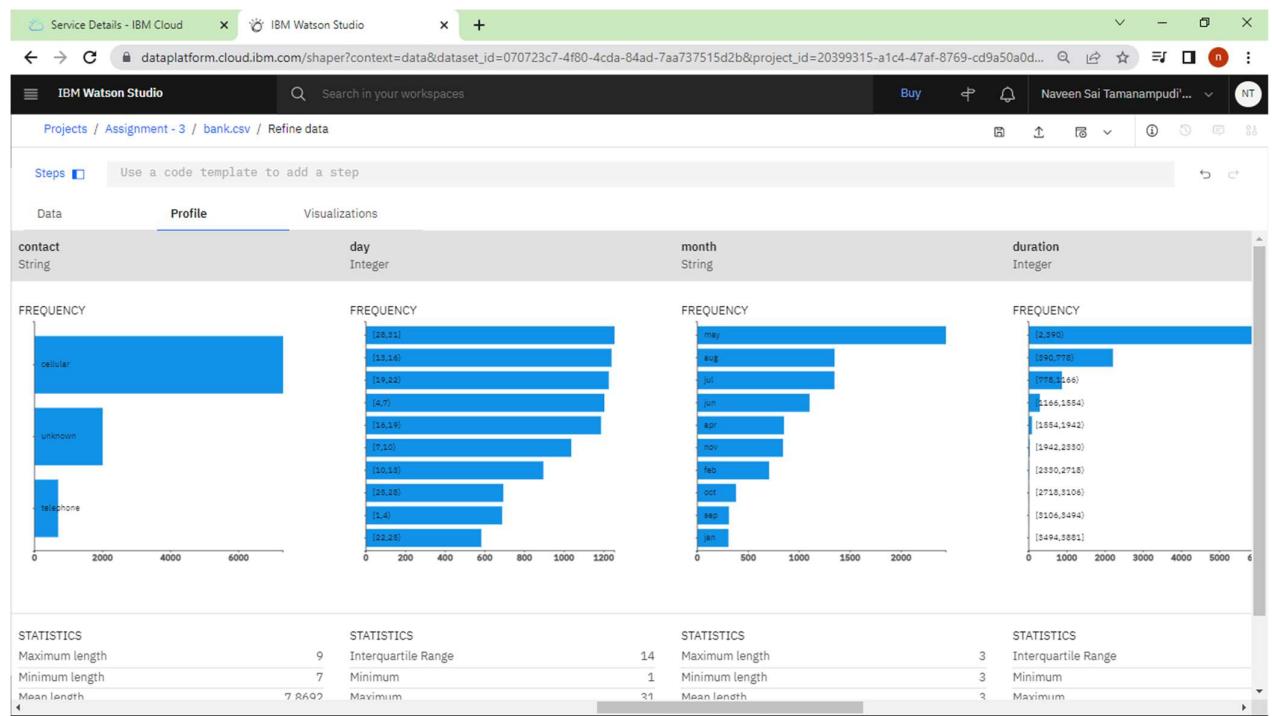
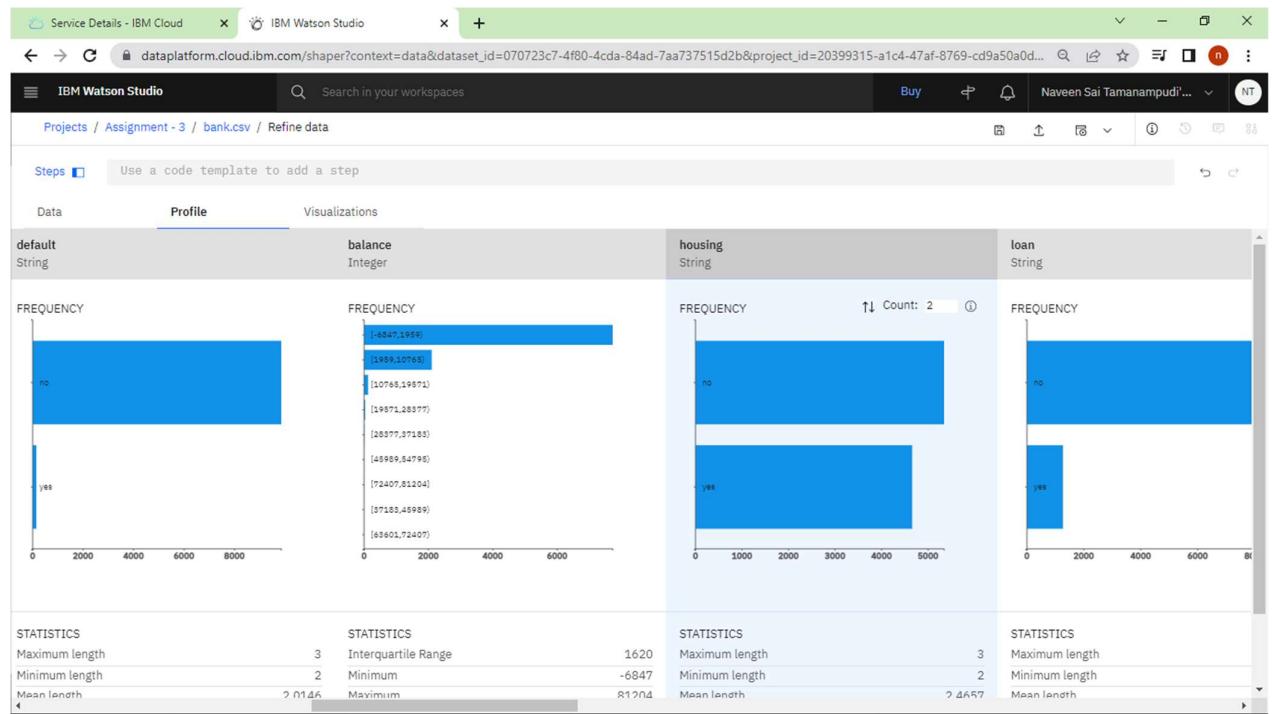
married	5200
single	3500
divorced	1000

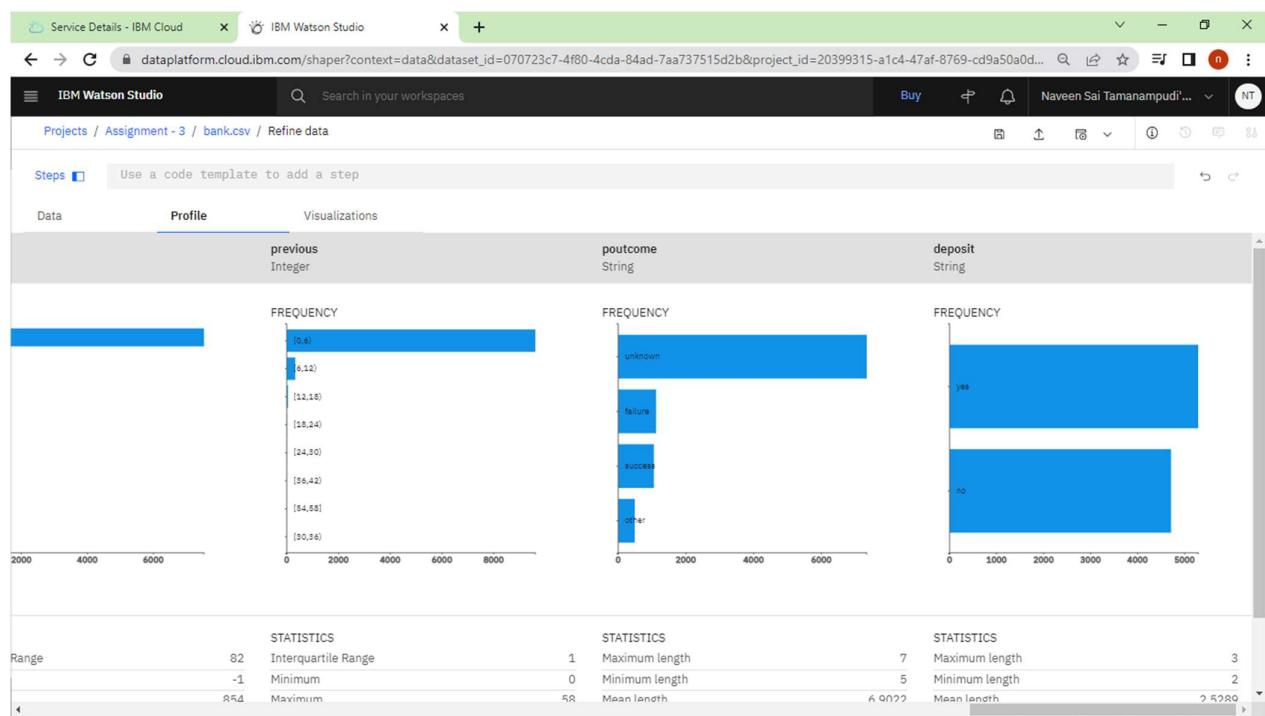
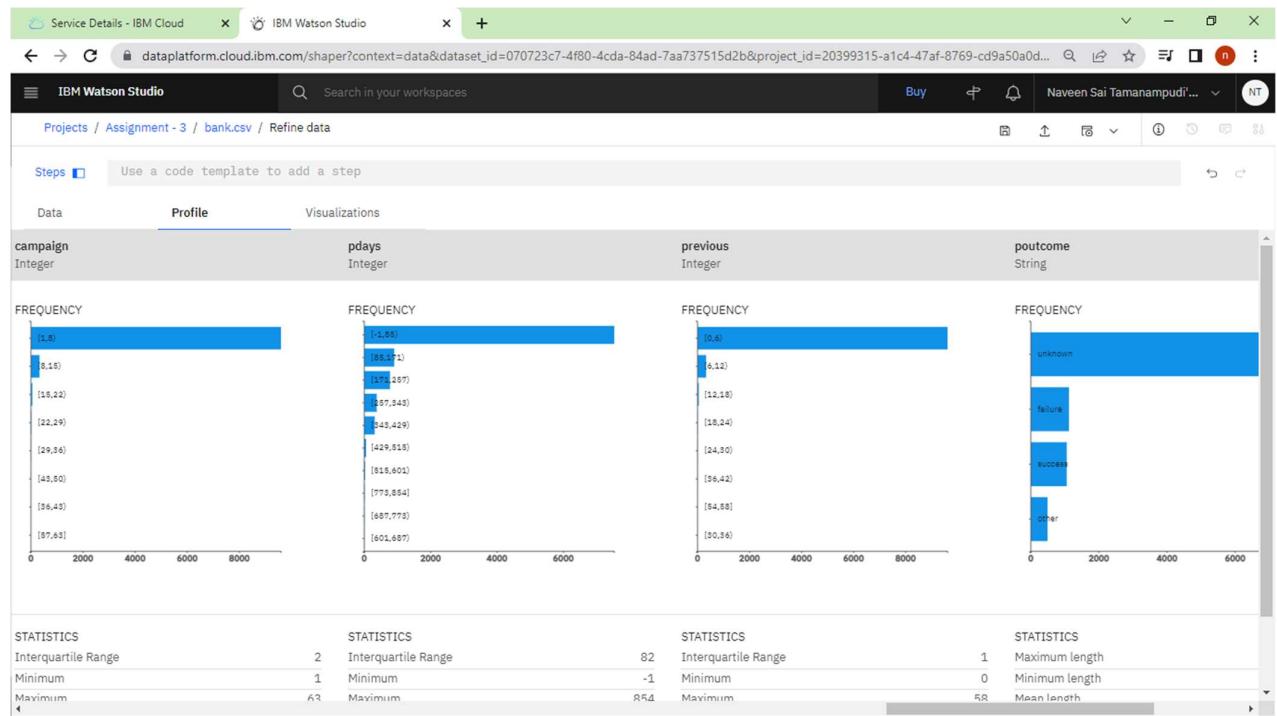
FREQUENCY

secondary	4200
tertiary	3200
primary	1200
unknown	500

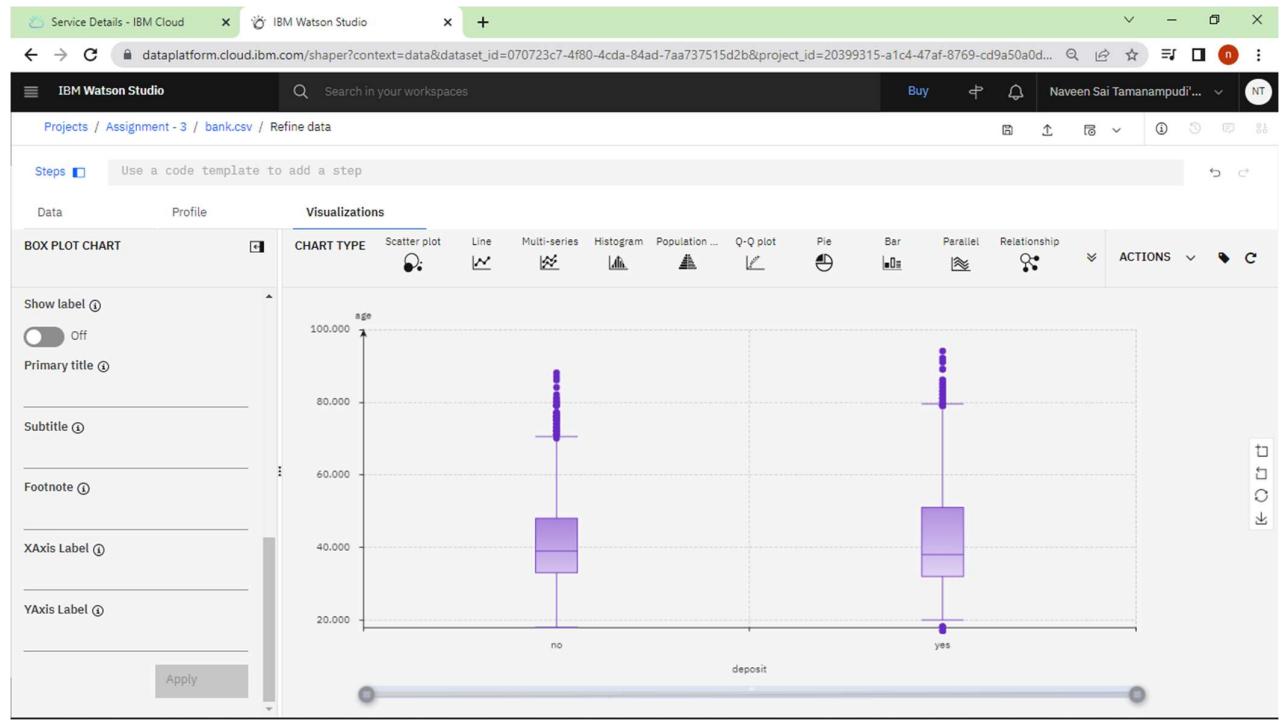
STATISTICS

Interquartile Range	17	Maximum length	13	Maximum length	8	Maximum length
Minimum	18	Minimum length	6	Minimum length	6	Minimum length
Maximum	95	Mean length	9.3258	Mean length	6.7966	Mean length





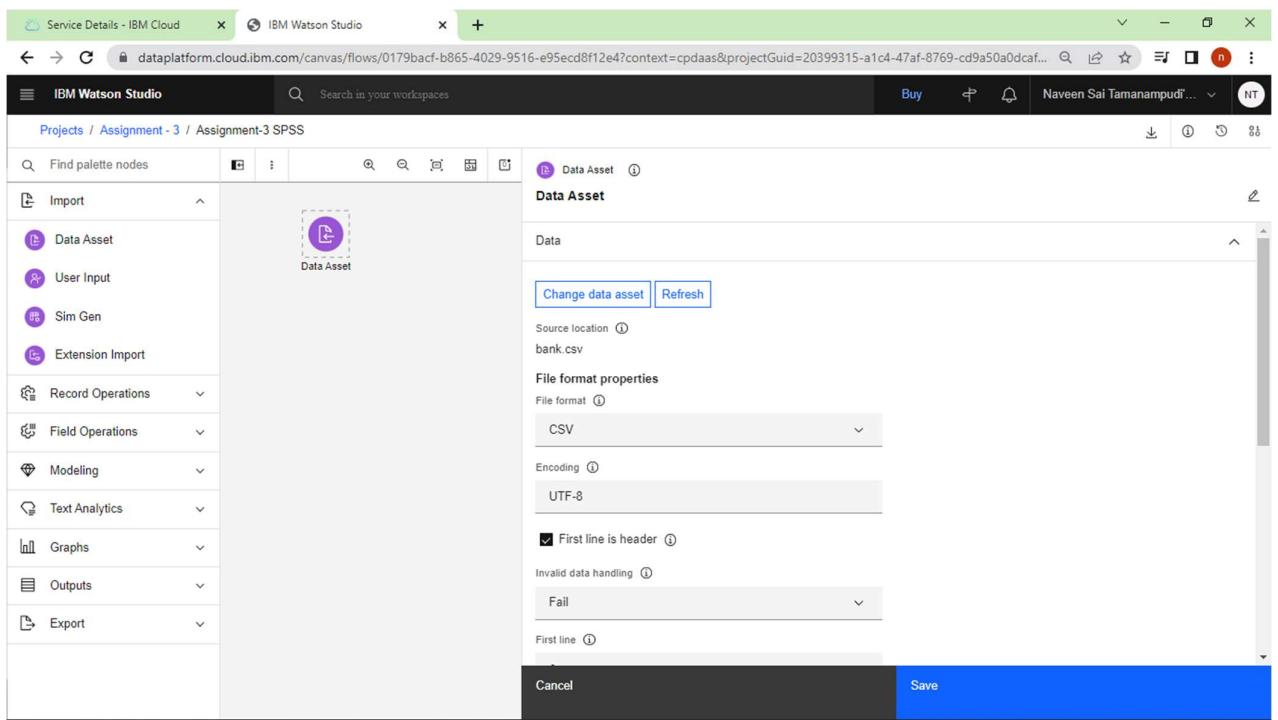
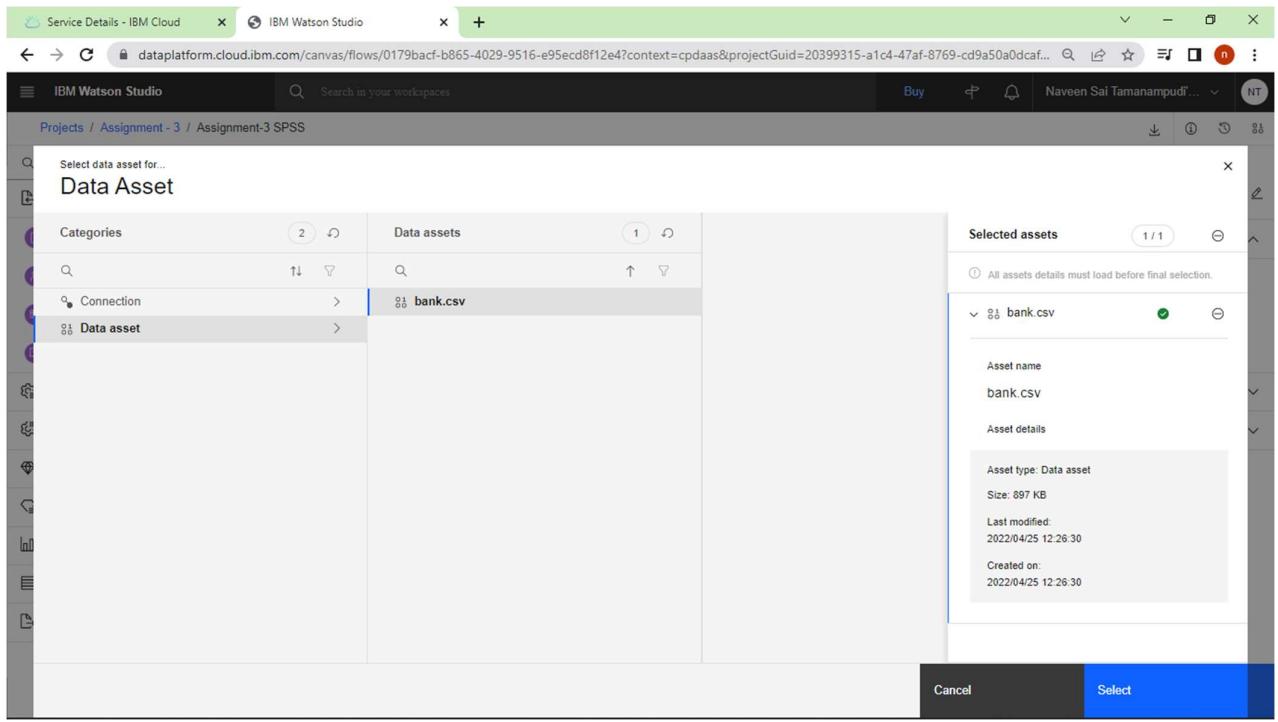
Boxplot of Age vs Deposit:



Creating an SPSS Modeler:

The screenshot shows the IBM Watson Studio interface. The top navigation bar includes tabs for 'Service Details - IBM Cloud' and 'IBM Watson Studio'. The main workspace title is 'Projects / Assignment - 3'. A sub-header 'Create an SPSS Modeler flow' is visible. On the left, a sidebar under 'New asset' shows options: '+ New' (selected), 'Gallery sample', and 'Local file'. The main area is titled 'Create an SPSS Modeler flow'. It has two main sections: 'Define details' and 'Define configuration'. In 'Define details', the 'Name' field is filled with 'Assignment-3 SPSS'. In 'Define configuration', the 'Environment definition' dropdown is set to 'Default SPSS Modeler S (2 vCPU 8 GB RAM)'. Below these fields is a note: 'To create additional runtime environments, view options in the Environments tab.' At the bottom right are 'Cancel' and 'Create' buttons, with 'Create' being highlighted in blue.

Choosing Dataset in Data Asset Node:



Output from Data Audit:

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

dataplatform.cloud.ibm.com/canvas/flows/0179bacf-b865-4029-9516-e95ecd8f12e4?context=cpdaas&project_id=20399315-a1c4-47af-8769-cd9a50a0dcfa

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudi...

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: Data Audit of [17 fields]

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

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

dataplatform.cloud.ibm.com/canvas/flows/0179bacf-b865-4029-9516-e95ecd8f12e4?context=cpdaas&project_id=20399315-a1c4-47af-8769-cd9a50a0dcfa

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudi...

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: Data Audit of [17 fields]

Field	Sample Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
8 loan		Categorical	--	--	--	--	--	2	11162
9 contact		Categorical	--	--	--	--	--	3	11162
10 day		Continuous	1	31	15.658	8.421	0.111	--	11162
11 month		Categorical	--	--	--	--	--	12	11162
12 duration		Continuous	2	3881	371.994	347.128	2.144	--	11162
13 campaign		Continuous	1	63	2.508	2.722	5.546	--	11162
14 pdays		Continuous	-1	854	51.330	108.758	2.450	--	11162
15 previous		Continuous	0	50	0.022	2.202	7.225	--	11162

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudi...

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: Data Audit of [17 fields]

15	previous		Continuous	0	58	0.833	2.292	7.335	--	11162	
16	poutcome		Categorical	--	--	--	--	--	4	11162	
17	deposit		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

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudi...

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: Data Audit of [17 fields]

5	default	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
6	balance	Continuous	99	74	None	Never	Fixed	100.000	11162	0	0	0	0
7	housing	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
8	loan	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
9	contact	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
10	day	Continuous	0	0	None	Never	Fixed	100.000	11162	0	0	0	0
11	month	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
12	duration	Continuous	179	22	None	Never	Fixed	100.000	11162	0	0	0	0
13	campaign	Continuous	136	74	None	Never	Fixed	100.000	11162	0	0	0	0
14	pdays	Continuous	149	27	None	Never	Fixed	100.000	11162	0	0	0	0
15	previous	Continuous	163	57	None	Never	Fixed	100.000	11162	0	0	0	0
16	poutcome	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0
17	deposit	Categorical	--	--	--	Never	Fixed	100.000	11162	0	0	0	0

Creating a Type Node and setting Inputs, Targets:

Type

Settings

Read values Clear values

Field	Measure	Role	Value mode	Values
# age	Continuous	Input	Read	
abc job	Categorical	Input	Read	
abc marital	Categorical	Input	Read	
abc education	Categorical	Input	Read	
abc default	Categorical	Input	Read	
# balance	Continuous	Input	Read	
abc housing	Categorical	Input	Read	
abc loan	Categorical	Input	Read	

Cancel Save

Type

Find in column Field

Field	Measure	Role	Value mode	Values
# day	Continuous	None	Read	
abc month	Categorical	None	Read	
# duration	Continuous	Input	Read	
# campaign	Continuous	Input	Read	
# pdays	Continuous	Input	Read	
# previous	Continuous	Input	Read	
abc poutcome	Categorical	Input	Read	
abc deposit	Flag	Target	Read	

Default mode ⓘ
Read metadata Pass (do not scan)

Set unique fields to None

Cancel Save

Changing the Outliers and Extremes to 0:

The screenshot shows the IBM Watson Studio Data Audit interface. On the left, there's a sidebar with options like Import, Record Operations, Field Operations, Modeling, Text Analytics, Graphs, Outputs, and Export. The main area shows a flow diagram where 'bank.csv' is being processed by a 'Type' node. To the right, under '17 Fields', there are sections for 'Calculate' (with checkboxes for 'Count of records with valid values' and 'Breakdown count of records with invalid values'), 'Detection Method' (with radio buttons for 'Standard deviation from mean' and 'Interquartile ranges from upper/lower quartiles'), 'Outliers' (set to 0), and 'Extremes' (set to 0). At the bottom are 'Cancel' and 'Save' buttons.

Running the model again and viewing the Data Audit Output:

The screenshot shows the 'View Output: 17 Fields' report in IBM Watson Studio. The table provides detailed statistics for each field:

	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	
8	loan		Categorical	--	--	--	--	--	2	11162	
9	contact		Categorical	--	--	--	--	--	3	11162	
10	day		Continuous	1	31	15.658	8.421	0.111	--	11162	
11	month		Categorical	--	--	--	--	--	12	11162	
12	duration		Continuous	2	3881	371.994	347.128	2.144	--	11162	
13	campaign		Continuous	1	63	2.508	2.722	5.546	--	11162	

13	campaign		Continuous	1	63	2.508	2.722	5.546	--	11162	
14	pdays		Continuous	-1	854	51.330	108.758	2.450	--	11162	
15	previous		Continuous	0	58	0.833	2.292	7.335	--	11162	
16	poutcome		Categorical	--	--	--	--	--	4	11162	
17	deposit		Categorical	--	--	--	--	--	2	11162	
Field	Measurement	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value		
1 age	Continuous	0	11162	None	Never	Fixed	100.000	11162	0		

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: 17 Fields

2	job	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
3	marital	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
4	education	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
5	default	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
6	balance	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
7	housing	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
8	loan	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
9	contact	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
10	day	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
11	month	Categorical	--	--	--	Never	Fixed	100.000	11162	0						

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

Projects / Assignment - 3 / Assignment-3 SPSS

View Output: 17 Fields

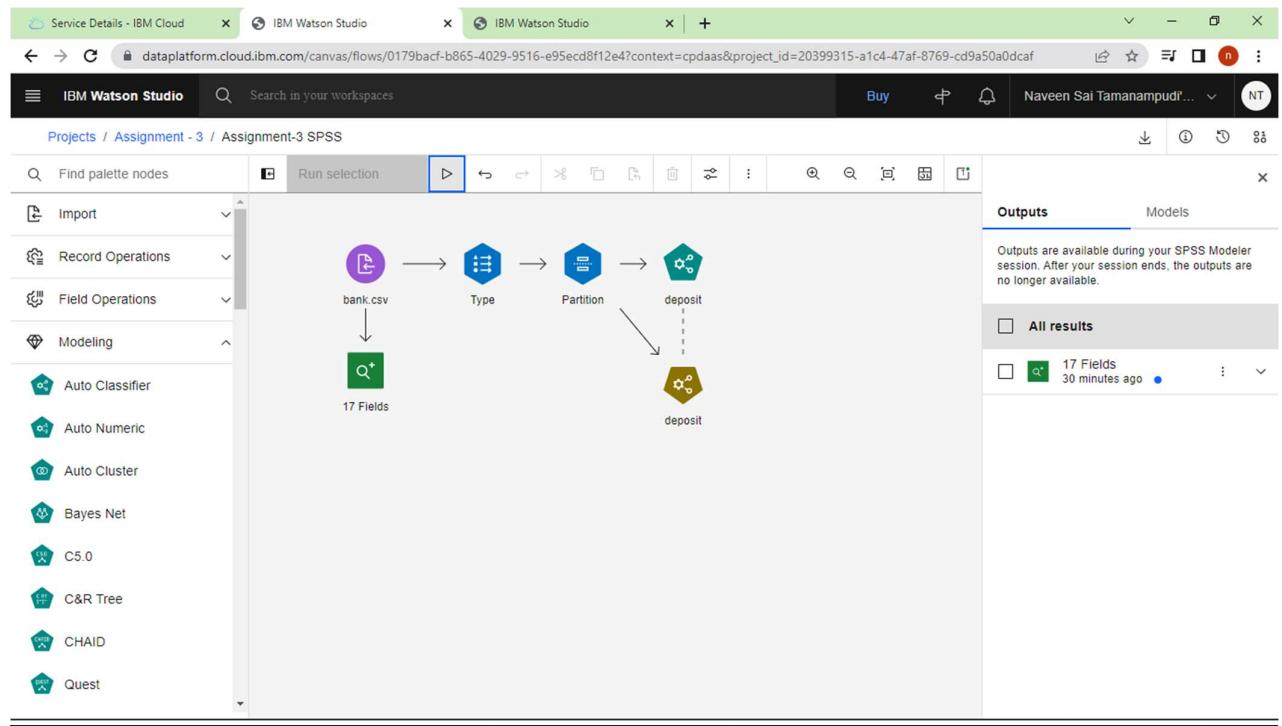
o	idall	Categorical	--	--	--	never	fixed	100.000	11162	0						
9	contact	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
10	day	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
11	month	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
12	duration	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
13	campaign	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
14	pdays	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
15	previous	Continuous	0	11162	None	Never	Fixed	100.000	11162	0						
16	poutcome	Categorical	--	--	--	Never	Fixed	100.000	11162	0						
17	deposit	Categorical	--	--	--	Never	Fixed	100.000	11162	0						

Creating a Partition Node with 80:20 split:

The screenshot shows the IBM Watson Studio interface with a flow editor. On the left, a palette lists nodes: Derive, Filler, Reclassify, Binning, RFM Analysis, Ensemble, Partition, SetToFlag, Restructure, Transpose, Field Reorder, and History. The main workspace contains a flow starting with a 'bank.csv' input node, followed by a 'Type' node, and then a 'Partition' node. The 'Partition' node is highlighted with a dashed border. To the right of the flow is a configuration panel titled 'Partition'. It includes settings for 'Derived Field Name' (set to 'Partition'), 'Training Partition(%)' (set to 80), 'Testing Partition(%)' (set to 20), and options for 'Create validation partition' (unchecked) and 'Repeatable partition assignment' (checked). A seed value '1234567' is also specified. At the bottom are 'Cancel' and 'Save' buttons.

Creating an Auto Classifier Model:

The screenshot shows the IBM Watson Studio interface with a flow editor. On the left, a palette lists nodes under 'Modeling': Import, Record Operations, Field Operations, Auto Classifier, Auto Numeric, Auto Cluster, Bayes Net, C5.0, C&R Tree, CHAID, and Quest. The main workspace contains a flow starting with a 'bank.csv' input node, followed by a 'Type' node, a 'Partition' node (highlighted with a dashed border), and finally an 'Auto Classifier' node (also highlighted with a dashed border). Below the 'Auto Classifier' node is a label 'deposit'. The configuration panel on the right is not fully visible but appears to be for the 'Auto Classifier' node.



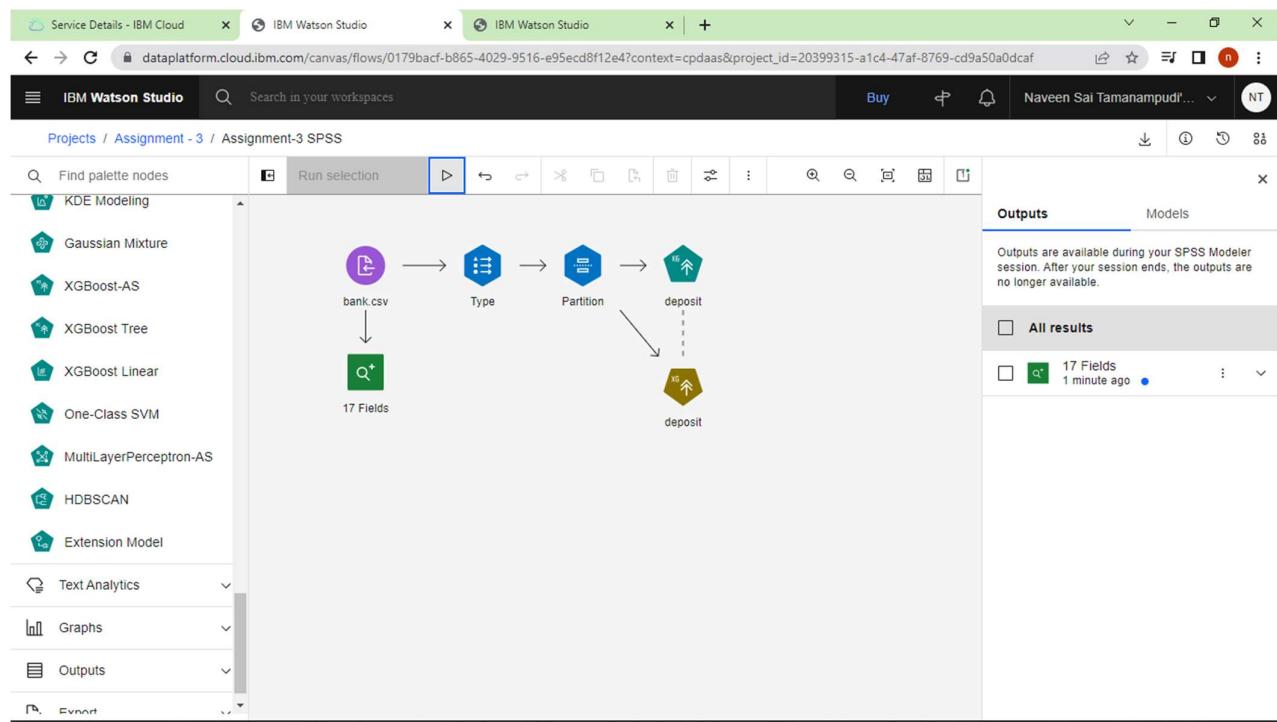
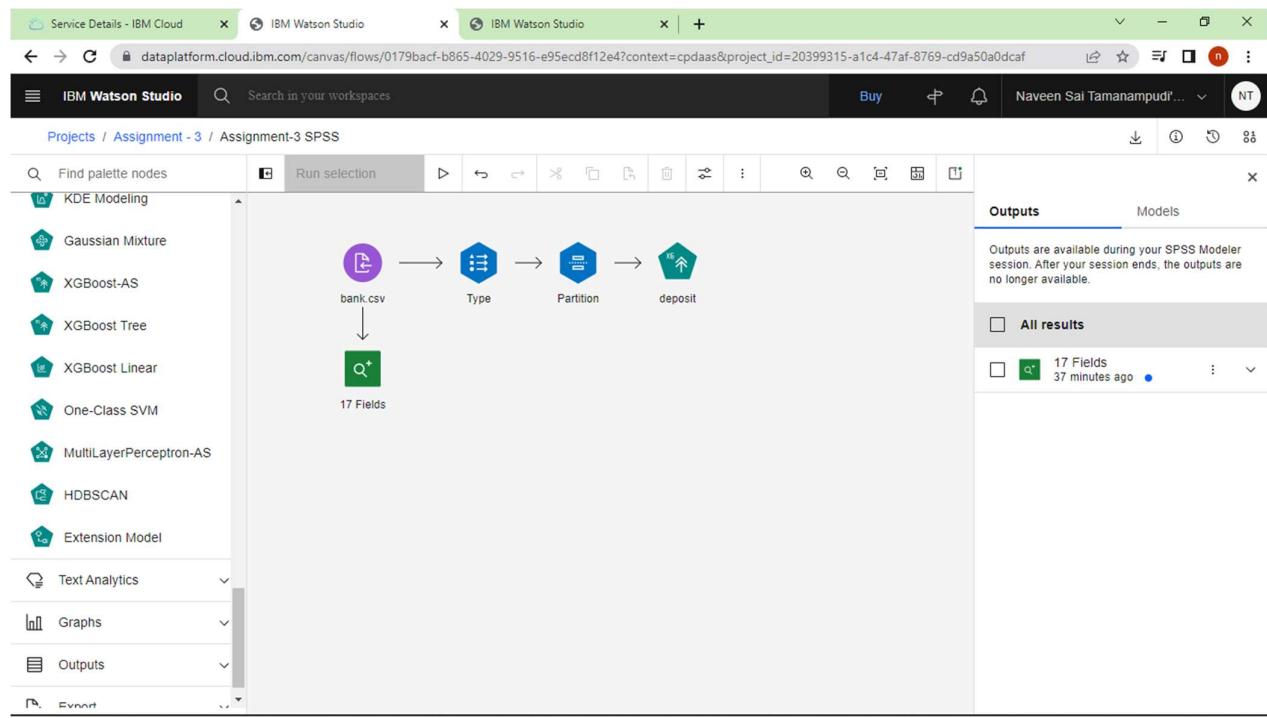
Output:

The screenshot shows the 'View Model: deposit' window in IBM Watson Studio. The left sidebar shows the 'Auto Classifier' node selected. The main area displays a table of models under the heading 'Auto Classifier - Models'. The table includes columns for USE, MODEL NAME, ESTIMATOR, BUILD TIME (MINS), NO. FIELDS USED, ACCURACY, ACCUMULATED ACCURACY, AREA UNDER CURVE, ACCUMULATED AUC, RECALL, PRECISION, and ACTIONS. The XGBoost Tree model has the highest accuracy of 80.423%.

USE	MODEL NAME	ESTIMATOR	BUILD TIME (MINS)	NO. FIELDS USED	ACCURACY	ACCUMULATED ACCURACY	AREA UNDER CURVE	ACCUMULATED AUC	RECALL	Precision	ACTIONS
✓	XGBoost Linear 1	XGBoost Linear	1	14	77.948	77.948	0.868	0.868	0.736	0.786	
✓	XGBoost Tree 1	XGBoost Binary Classification Model	1	14	80.423	80.423	0.889	0.889	0.813	0.784	
✓	Logistic regression 1	Nominal Regression	1	14	78.758	78.758	0.872	0.872	0.749	0.793	
✓	LSVM 1	Linear SVM	1	14	77.768	77.768	0.871	0.871	0.721	0.793	
✓	C5.0 1	C5.0	1	13	80.018	80.018	0.876	0.876	0.801	0.783	

Here, XGBoost Tree is having highest Accuracy. So, replacing Auto classifier with XGBoost Tree.

Creating an XGBoost Tree Model:



Viewing Output:

The screenshot shows the IBM Watson Studio interface with a tab titled "Service Details - IBM Cloud". The main content area displays "View Model: deposit". On the left, a sidebar menu includes "Model viewer", "Model information", "Feature summary", "Evaluation", "Model evaluation", and "Confusion matrix". The "Model information" section is selected. It contains a table with the following details:

Prediction column	deposit
Algorithm	XGBoost Binary Classification Model
Number of features	29
Number of evaluation instances	8940
NumberofTrees	10
MaximumDepth	6

The screenshot shows the IBM Watson Studio interface with a tab titled "Service Details - IBM Cloud". The main content area displays "View Model: deposit". On the left, a sidebar menu includes "Model viewer", "Model information", "Feature summary", "Evaluation", "Model evaluation", and "Confusion matrix". The "Feature summary" section is selected. It contains a table with the following data:

All features	Search feature or transformer names
poutcome_success	None 12.62% █
contact_unknown	None 10.46% █
duration	None 8.65% █
housing	None 5.34% █
pdays	None 4.28% █
loan	None 4.13% █
poutcome_other	None 4.04% █
campaign	None 4.04% █
Total importance	100%

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

dataplatform.cloud.ibm.com/canvas/flows/0179bacf-b865-4029-9516-e95ecd8f12e4?context=cpdaas&project_id=20399315-a1c4-47af-8769-cd9a50a0dcfa

IBM Watson Studio Search in your workspaces Buy +

Projects / Assignment - 3 / Assignment-3 SPSS

View Model: deposit

Model viewer
Model information
Feature summary
Evaluation
Model evaluation
Confusion matrix

feature	value	percentage
campaign	None	4.04%
job_blue-collar	None	4.01%
education_tertiary	None	3.93%
contact_telephone	None	3.80%
marital_married	None	3.62%
contact_cellular	None	3.15%
age	None	3.13%
balance	None	2.66%

Items per page: 15 ▾ 1–15 of 29 items 1 ▾ of 2 pages ▶

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

dataplatform.cloud.ibm.com/canvas/flows/0179bacf-b865-4029-9516-e95ecd8f12e4?context=cpdaas&project_id=20399315-a1c4-47af-8769-cd9a50a0dcfa

IBM Watson Studio Search in your workspaces Buy +

Projects / Assignment - 3 / Assignment-3 SPSS

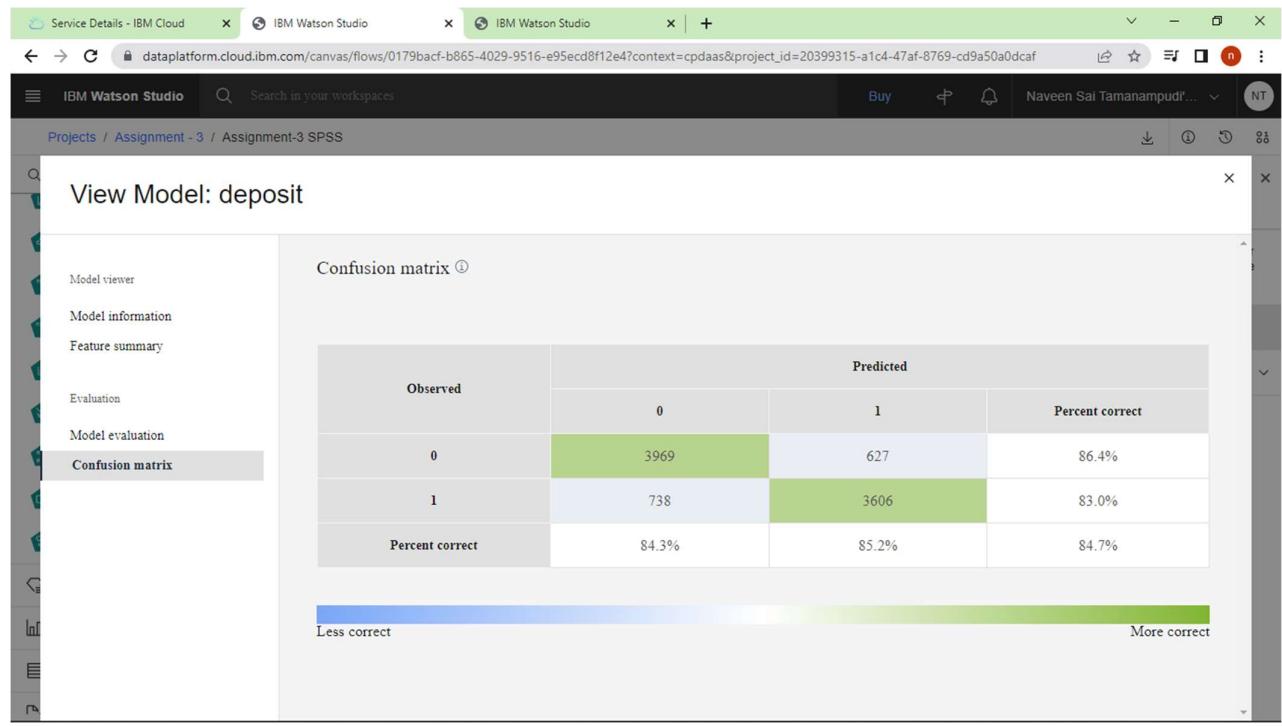
View Model: deposit

Model viewer
Model information
Feature summary
Evaluation
Model evaluation
Confusion matrix

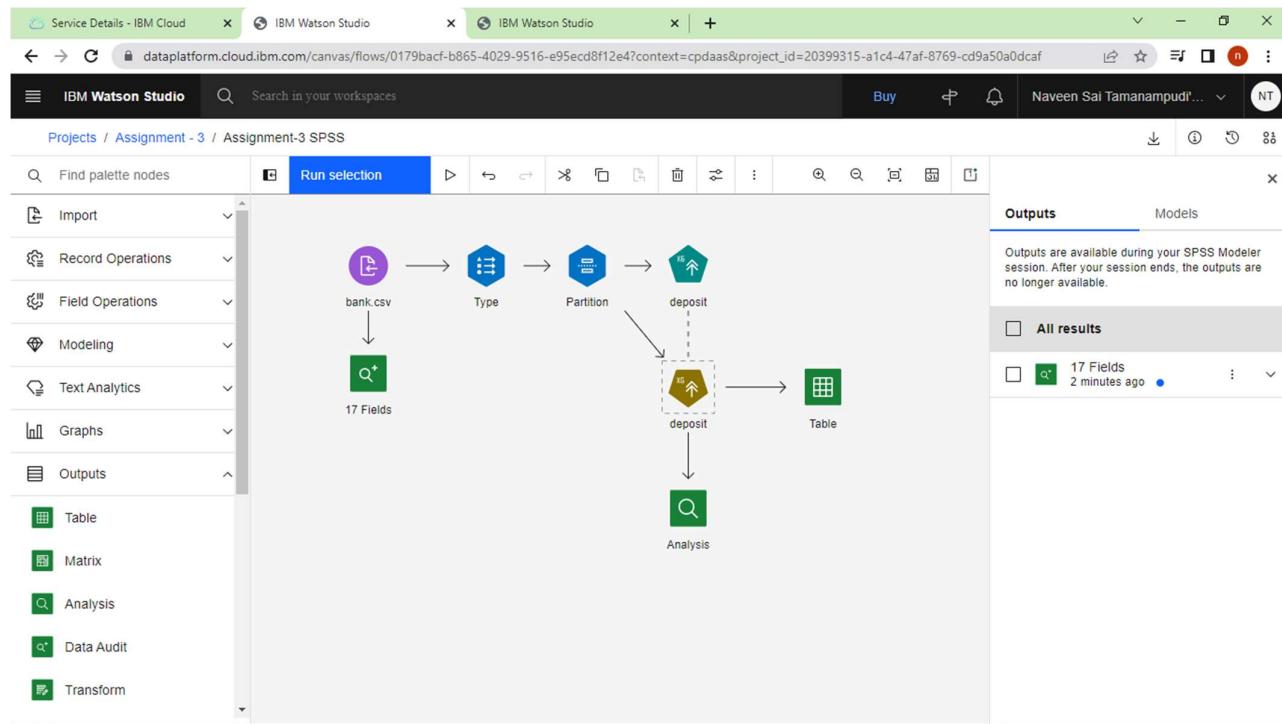
Model evaluation ①

Model evaluation measure

metric	value
Accuracy	0.847
Precision	0.830
Recall	0.852
F1	0.388



Creating Table and Analysis Nodes:



Output of Analysis:

The screenshot shows the IBM Watson Studio interface with a project titled "Assignment - 3 / Assignment-3 SPSS". A central panel displays the output for the "deposit" field, specifically comparing \$XGT-deposit with deposit. It includes a table showing the count and percentage of correct and wrong predictions for training and testing partitions.

'Partition'	1_Training	2_Testing
Correct	7,575	84.73%
Wrong	1,365	15.27%
Total	8,940	2,222

Output of Table:

The screenshot shows the IBM Watson Studio interface with a project titled "Assignment - 3 / Assignment-3 SPSS". A central panel displays a table titled "View Output: Table (20 fields, 11,162 records)". The table contains 13 rows of data, each with 16 columns corresponding to the fields: age, job, marital, education, default, balance, housing, loan, contact, day, month, duration, campaign, pdays, and previous. The data represents various demographic and financial information for 11,162 records.

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous
1	59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1	-1	
2	56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1	-1	
3	41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1	-1	
4	55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1	-1	
5	54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2	-1	
6	42	management	single	tertiary	no	0	yes	yes	unknown	5	may	562	2	-1	
7	56	management	married	tertiary	no	830	yes	yes	unknown	6	may	1201	1	-1	
8	60	retired	divorced	secondary	no	545	yes	no	unknown	6	may	1030	1	-1	
9	37	technician	married	secondary	no	1	yes	no	unknown	6	may	608	1	-1	
10	28	services	single	secondary	no	5090	yes	no	unknown	6	may	1297	3	-1	
11	38	admin.	single	secondary	no	100	yes	no	unknown	7	may	786	1	-1	
12	30	blue-collar	married	secondary	no	309	yes	no	unknown	7	may	1574	2	-1	
13	29	management	married	tertiary	no	199	yes	yes	unknown	7	may	1689	4	-1	

Service Details - IBM Cloud IBM Watson Studio IBM Watson Studio

IBM Watson Studio Search in your workspaces Buy +

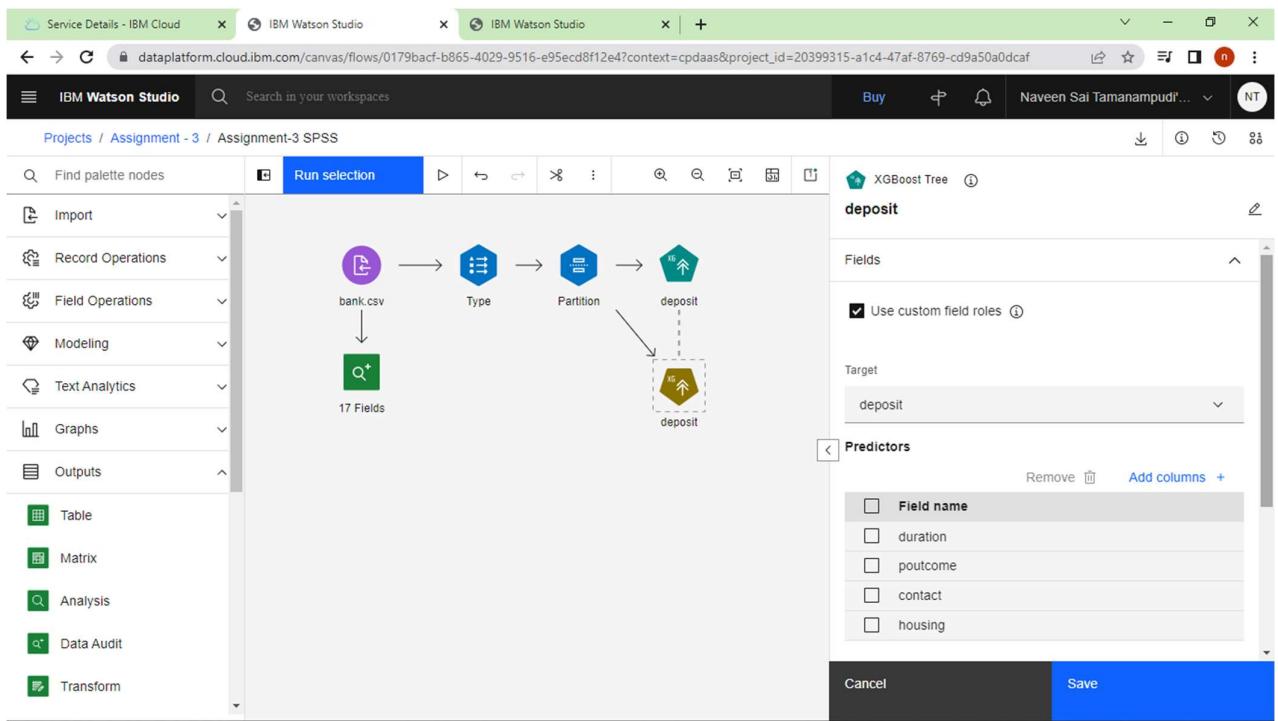
Projects / Assignment - 3 / Assignment-3 SPSS

View Output: Table (20 fields, 11,162 records) Compare

housing loan contact day month duration campaign pdays previous poutcome deposit Partition \$XGT-deposit \$XGTC-deposit

housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	deposit	Partition	\$XGT-deposit	\$XGTC-deposit
yes	no	unknown	5	may	1042	1	-1	0	unknown	yes	1_Training	yes	0.842
no	no	unknown	5	may	1467	1	-1	0	unknown	yes	1_Training	yes	0.823
yes	no	unknown	5	may	1389	1	-1	0	unknown	yes	1_Training	yes	0.856
yes	no	unknown	5	may	579	1	-1	0	unknown	yes	2_Testing	no	0.582
no	no	unknown	5	may	673	2	-1	0	unknown	yes	1_Training	yes	0.715
yes	yes	unknown	5	may	562	2	-1	0	unknown	yes	1_Training	yes	0.639
yes	yes	unknown	6	may	1201	1	-1	0	unknown	yes	1_Training	yes	0.839
yes	no	unknown	6	may	1030	1	-1	0	unknown	yes	1_Training	yes	0.841
yes	no	unknown	6	may	608	1	-1	0	unknown	yes	1_Training	yes	0.547
yes	no	unknown	6	may	1297	3	-1	0	unknown	yes	1_Training	yes	0.839
yes	no	unknown	7	may	786	1	-1	0	unknown	yes	1_Training	yes	0.786
yes	no	unknown	7	may	1574	2	-1	0	unknown	yes	1_Training	yes	0.831
yes	yes	unknown	7	may	1689	4	-1	0	unknown	yes	2_Testing	yes	0.879

Using Custom Field Roles



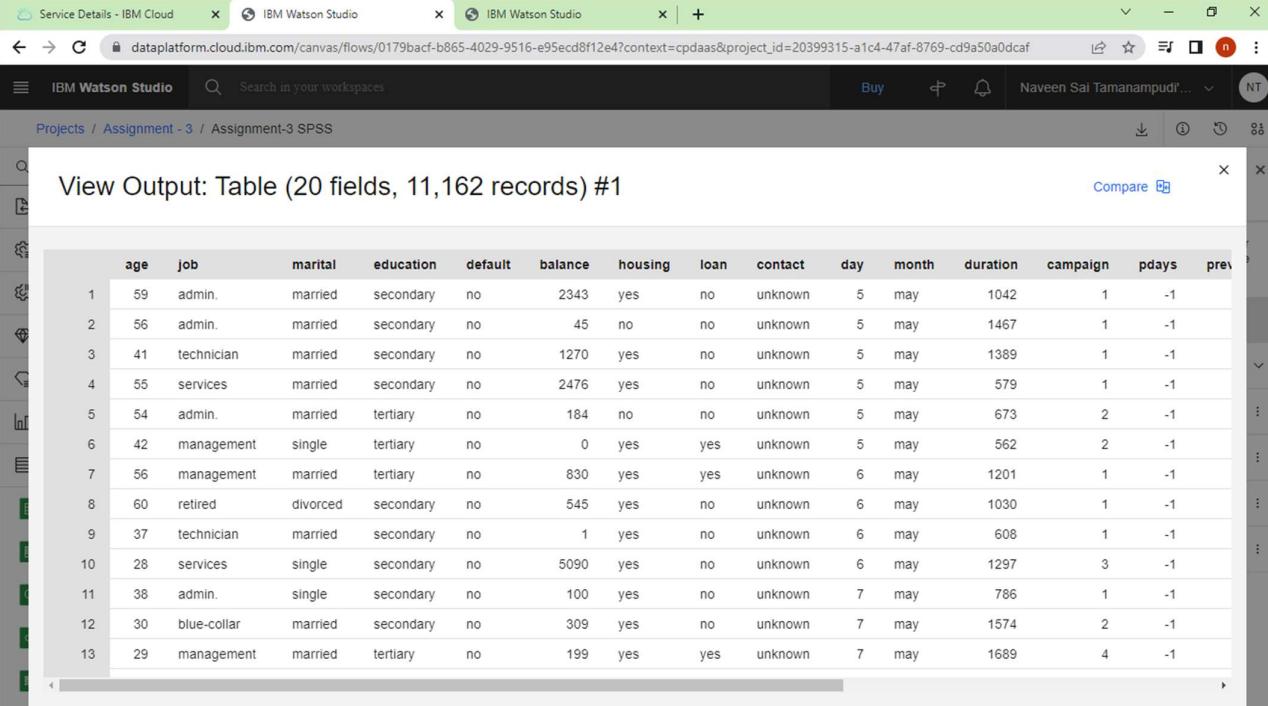
Selecting the top 4 attributes as Input:

The screenshot shows the IBM Watson Studio interface with a project titled "Assignment - 3 / Assignment-3 SPSS". On the left, a palette lists various operations like Import, Record Operations, Field Operations, Modeling, Text Analytics, Graphs, Outputs, Table, Matrix, Analysis, Data Audit, and Transform. In the center, a flow diagram starts with an "Import" node reading "bank.csv", followed by a "Type" node and a "Partition" node. A callout from the "Partition" node points to a "Select Fields for deposit" dialog box. This dialog box contains a search bar for "Field name" and a filter "abc". It lists 17 fields with checkboxes, where "housing", "duration", "poutcome", and "deposit" are checked. The "deposit" field is highlighted with a yellow dashed border. At the bottom are "Cancel" and "OK" buttons.

Building the Model again:

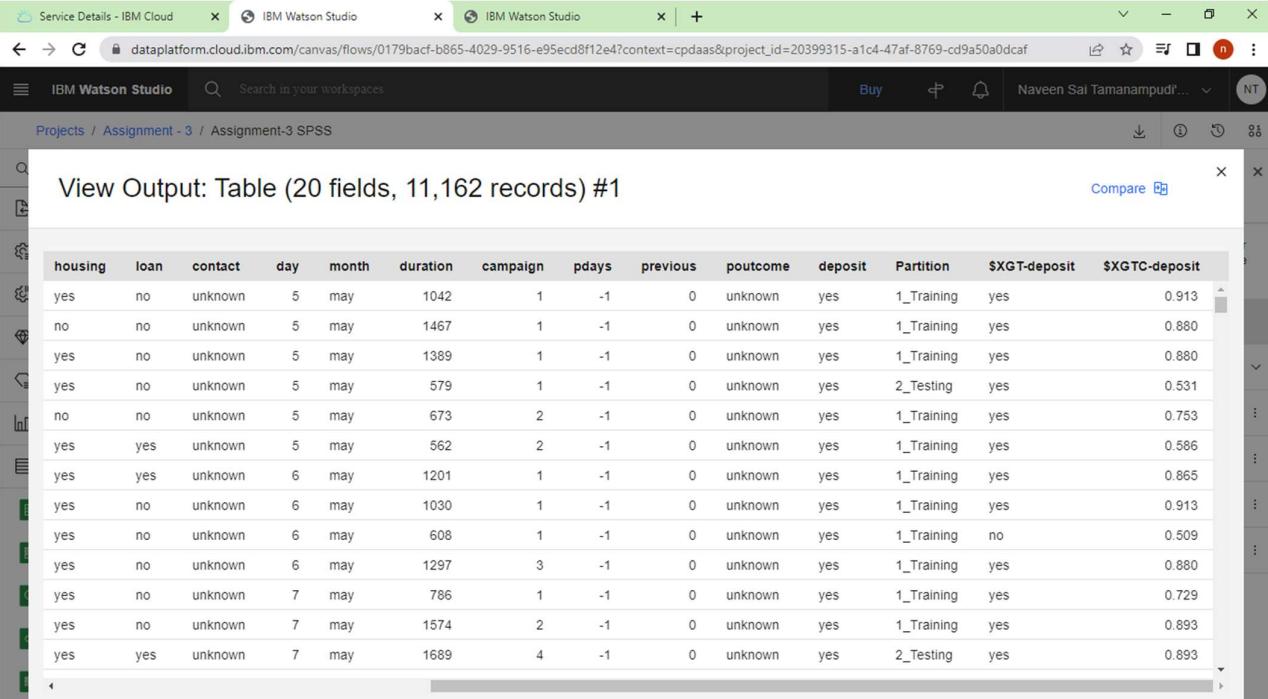
The screenshot shows the IBM Watson Studio interface with the same project structure. The flow diagram now includes a "deposit" node with a yellow dashed border, indicating it is the target variable. Arrows point from the "Partition" node to the "deposit" node, and from the "deposit" node to a "Table" node and an "Analysis" node. To the right, the "Outputs" panel is open, showing a list of available outputs: "17 Fields Just now", "Table (20 fields, 11,162 re... Just now", "Analysis of [deposit] #1 Just now", "Analysis of [deposit] 6 minutes ago", and "Table (20 fields, 11,162 re... 6 minutes ago".

Viewing output of Table:



The screenshot shows a table titled "View Output: Table (20 fields, 11,162 records) #1". The table has 20 columns and 11,162 rows. The columns are labeled: age, job, marital, education, default, balance, housing, loan, contact, day, month, duration, campaign, pdays, previous, poutcome, deposit, Partition, \$XGT-deposit, and \$XGTC-deposit. The data includes various categorical and numerical values for each record.

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	deposit	Partition	\$XGT-deposit	\$XGTC-deposit
1	59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1	-1						
2	56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1	-1						
3	41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1	-1						
4	55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1	-1						
5	54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2	-1						
6	42	management	single	tertiary	no	0	yes	yes	unknown	5	may	562	2	-1						
7	56	management	married	tertiary	no	830	yes	yes	unknown	6	may	1201	1	-1						
8	60	retired	divorced	secondary	no	545	yes	no	unknown	6	may	1030	1	-1						
9	37	technician	married	secondary	no	1	yes	no	unknown	6	may	608	1	-1						
10	28	services	single	secondary	no	5090	yes	no	unknown	6	may	1297	3	-1						
11	38	admin.	single	secondary	no	100	yes	no	unknown	7	may	786	1	-1						
12	30	blue-collar	married	secondary	no	309	yes	no	unknown	7	may	1574	2	-1						
13	29	management	married	tertiary	no	199	yes	yes	unknown	7	may	1689	4	-1						



The screenshot shows a table titled "View Output: Table (20 fields, 11,162 records) #1". The table has 20 columns and 11,162 rows. The columns are labeled: housing, loan, contact, day, month, duration, campaign, pdays, previous, poutcome, deposit, Partition, \$XGT-deposit, and \$XGTC-deposit. The data includes various categorical and numerical values for each record.

housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	deposit	Partition	\$XGT-deposit	\$XGTC-deposit
yes	no	unknown	5	may	1042	1	-1	0	unknown	yes	1_Training	yes	0.913
no	no	unknown	5	may	1467	1	-1	0	unknown	yes	1_Training	yes	0.880
yes	no	unknown	5	may	1389	1	-1	0	unknown	yes	1_Training	yes	0.880
yes	no	unknown	5	may	579	1	-1	0	unknown	yes	2_Testing	yes	0.531
no	no	unknown	5	may	673	2	-1	0	unknown	yes	1_Training	yes	0.753
yes	yes	unknown	5	may	562	2	-1	0	unknown	yes	1_Training	yes	0.586
yes	yes	unknown	6	may	1201	1	-1	0	unknown	yes	1_Training	yes	0.865
yes	no	unknown	6	may	1030	1	-1	0	unknown	yes	1_Training	yes	0.913
yes	no	unknown	6	may	608	1	-1	0	unknown	yes	1_Training	no	0.509
yes	no	unknown	6	may	1297	3	-1	0	unknown	yes	1_Training	yes	0.880
yes	no	unknown	7	may	786	1	-1	0	unknown	yes	1_Training	yes	0.729
yes	no	unknown	7	may	1574	2	-1	0	unknown	yes	1_Training	yes	0.893
yes	yes	unknown	7	may	1689	4	-1	0	unknown	yes	2_Testing	yes	0.893

Viewing Output of Analysis:

The screenshot shows the IBM Watson Studio interface with a browser tab for Service Details - IBM Cloud and two tabs for IBM Watson Studio. The main workspace shows a project titled 'Assignment - 3 / Assignment-3 SPSS'. A specific analysis titled 'View Output: Analysis of [deposit] #1' is displayed. The analysis tree shows 'Results for output field deposit' and 'Comparing \$XGT-deposit with deposit'. A table provides a detailed comparison of training and testing data for the 'XGT-deposit' model.

'Partition'	1_Training	2_Testing	
Correct	7,239	80.97%	1,752
Wrong	1,701	19.03%	470
Total	8,940		2,222