

Data Analytics Externship Program

Assignment-2

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

Risk Analysis of Space Shuttle using Multiple Linear Regression

Dataset Used: challengers.csv

Creating the Project

The screenshot shows the 'New project' form in the IBM Watson Studio web interface. The browser tabs at the top include 'Service Details - IBM Cloud' and 'IBM Watson Studio'. The address bar shows the URL 'dataplatfom.cloud.ibm.com/projects/create-project?context=cpdaas'. The page header includes the 'IBM Watson Studio' logo, a search bar, a 'Buy' button, a notification bell, and the user's name 'Naveen Sai Tamanampudi' with a profile icon labeled 'NT'.

The form is titled 'New project' and is divided into two main sections: 'Define details' and 'Storage'.

Define details

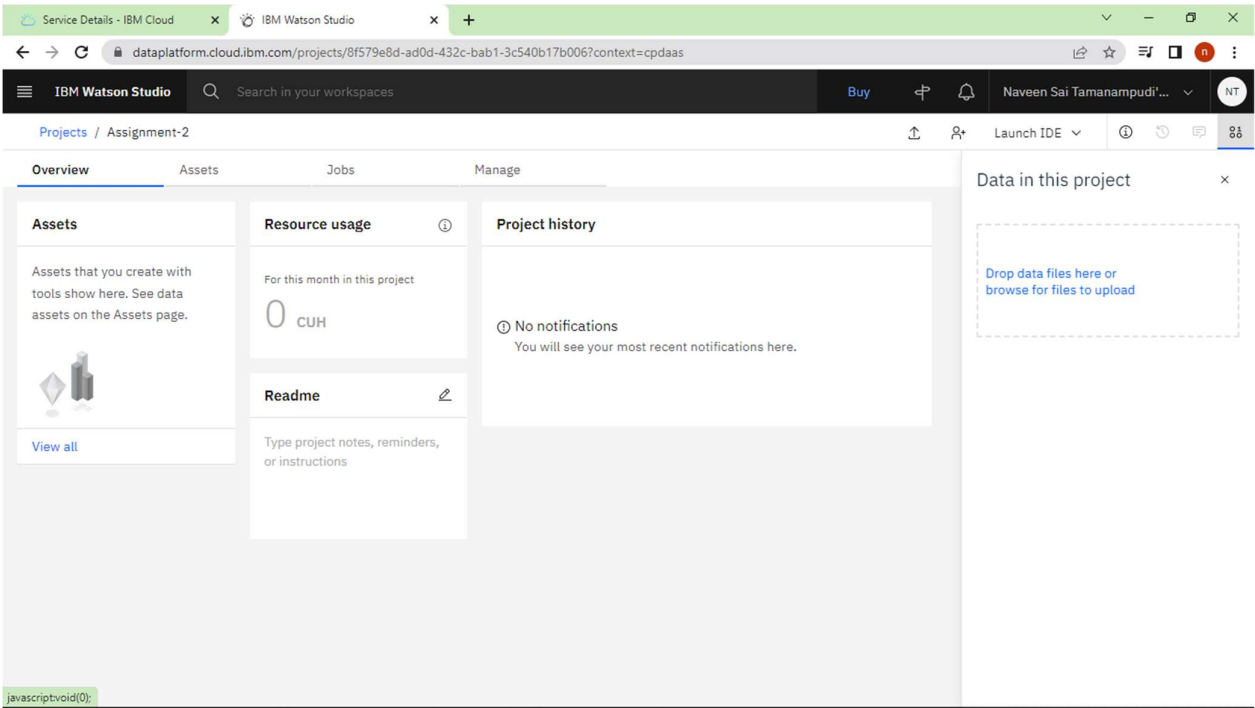
- Name:** A text input field containing 'Assignment-2'.
- Description:** A text area containing 'Risk Analysis of Space Shuttle using Multiple Linear Regression'.
- Choose project options:** A checkbox labeled 'Restrict who can be a collaborator' is unchecked. Below it, a note states: 'Project includes integration with [Cloud Object Storage](#) for storing project assets.'

Storage

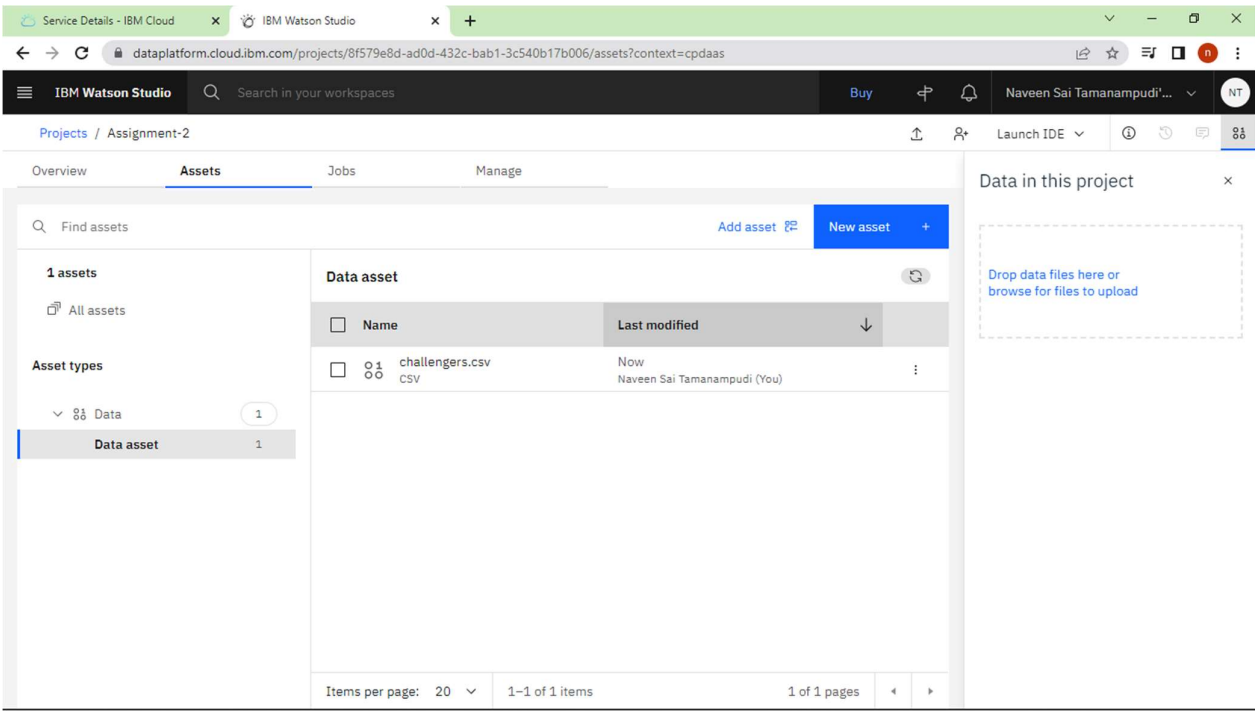
- Storage:** A text input field containing 'Cloud Object Storage-ur'.

At the bottom right of the form, there are two buttons: 'Cancel' (dark gray) and 'Create' (blue).

Project Overview:



Uploading Data:



Viewing the Data:

Service Details - IBM Cloud

IBM Watson Studio

← → ↻

dataplatfom.cloud.ibm.com/projects/8f579e8d-ad0d-432c-bab1-3c540b17b006/data-assets/0bca57ae-0123-4c7a-aac7-13ca2333637e/preview?context=cpdaas&...

☆

☰

🔍

NT

IBM Watson Studio

Search in your workspaces

Buy

🔔

Naveen Sai Tamanampudi'...

NT

My projects / Assignment-2 / challengers.csv

🔍

🕒

🗨

⚙

Preview

Activities

Schema: 5 Columns

The preview includes only a limited set of columns and rows. ⓘ

Last refresh: 15 seconds ago ↻

Refine

o_ring_ct String	o_ring_failures String	temperature String	pressure String	launch_id String
6	0	66	50	1
6	1	70	50	2
6	0	69	50	3
6	0	68	50	4
6	0	67	50	5
6	0	72	50	6
6	0	73	100	7
6	0	70	100	8

Information

×

Data asset

challengers.csv

Description

No description available for this asset

Tags

No tags available for this asset

Creator

Naveen Sai Tamanampudi

Usage

Created on Apr 25, 2022, 04:16 PM

Size

388 Bytes

Creating a Data Refinery:

Service Details - IBM Cloud

IBM Watson Studio

← → ↻

dataplatfom.cloud.ibm.com/projects/8f579e8d-ad0d-432c-bab1-3c540b17b006/assets?context=cpdaas

☆

☰

🔍

NT

IBM Watson Studio

Search in your workspaces

Buy

🔔

Naveen Sai Tamanampudi'...

NT

Projects / Assignment-2

⬆

👤

Launch IDE

🔍

🕒

🗨

⚙

Select data from project

Assignment-2 2 ↻

Data assets 1 ↻

🔍

🔍

🔗 Connection >

📄 Data asset >

📄 challengers.csv

Selected assets 1/1 ⓘ

ⓘ All assets details must load before final selection.

✓ 📄 challengers.csv ⓘ

Asset name

challengers.csv

Asset details

Asset type: Data asset

Size: 388 B

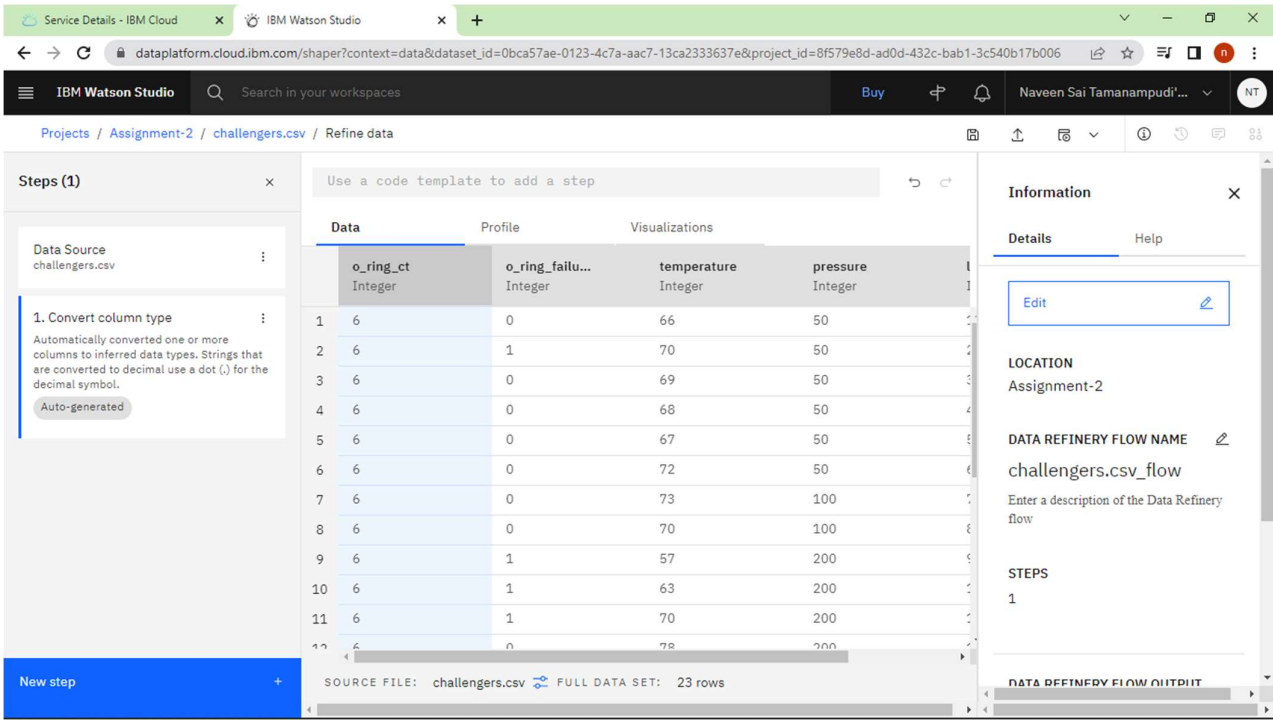
Last modified: 2022/04/25 10:46:19

Created on: 2022/04/25 10:46:19

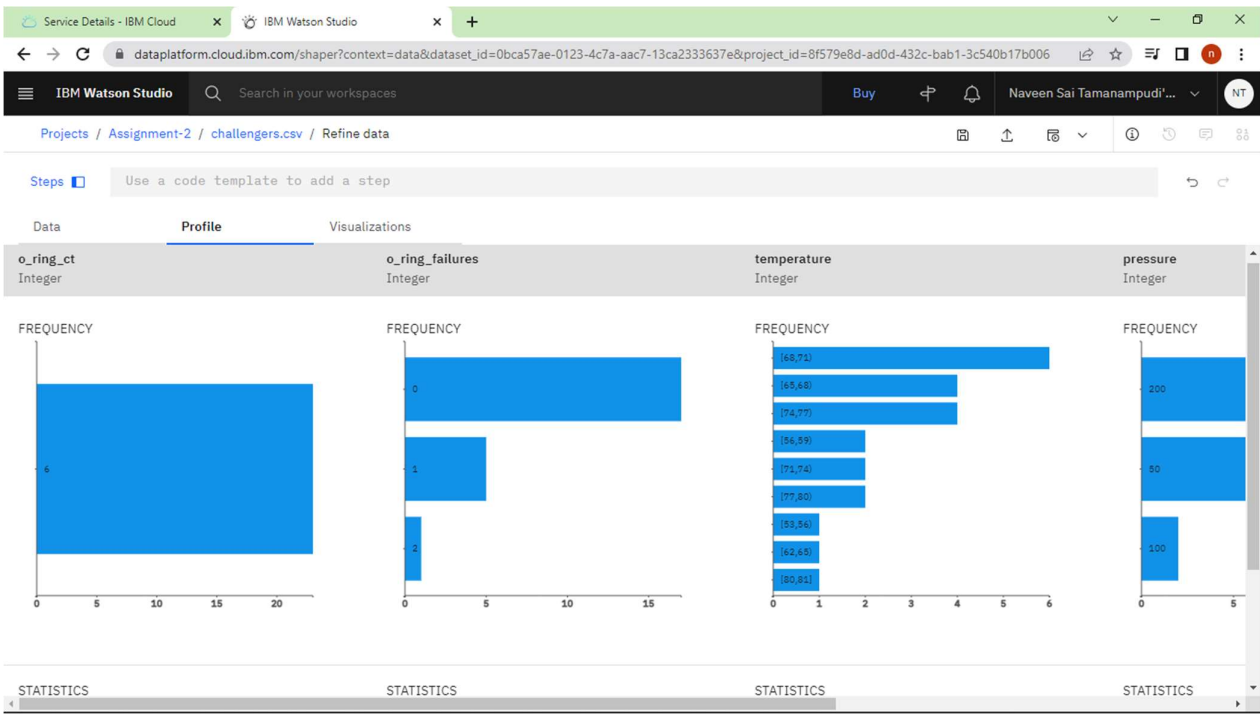
Cancel

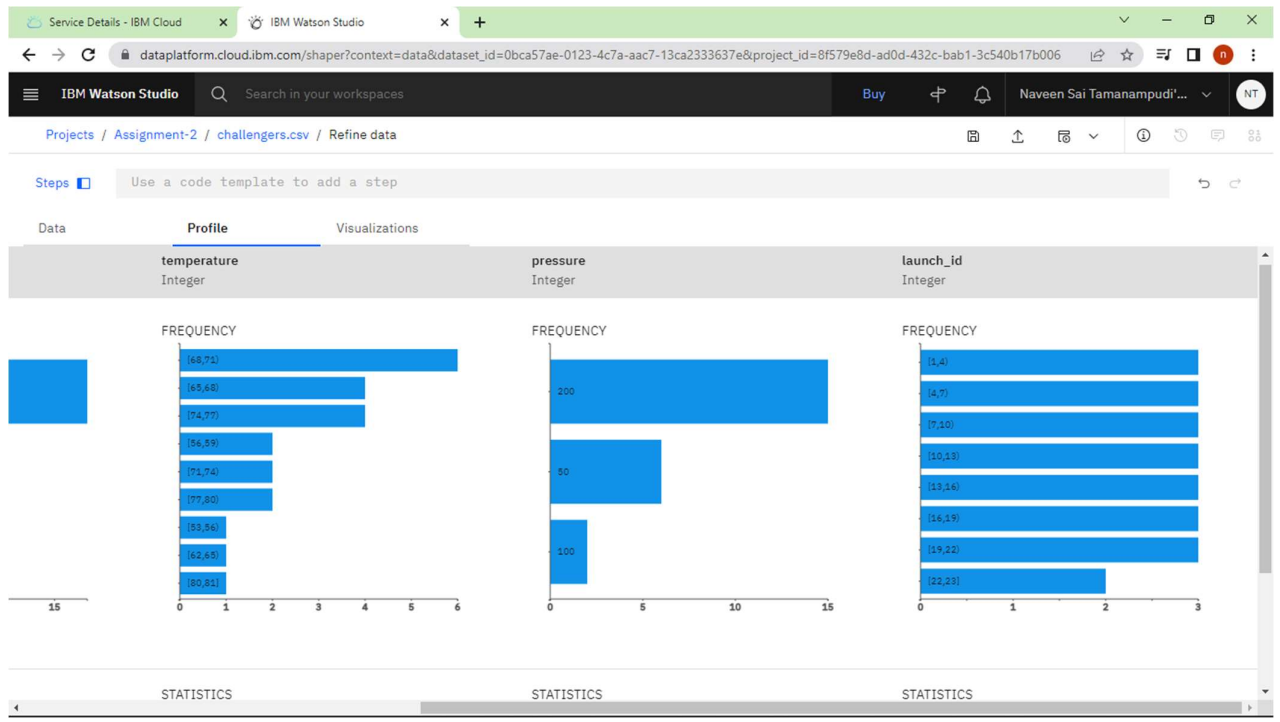
Select

Viewing the output of refinery:

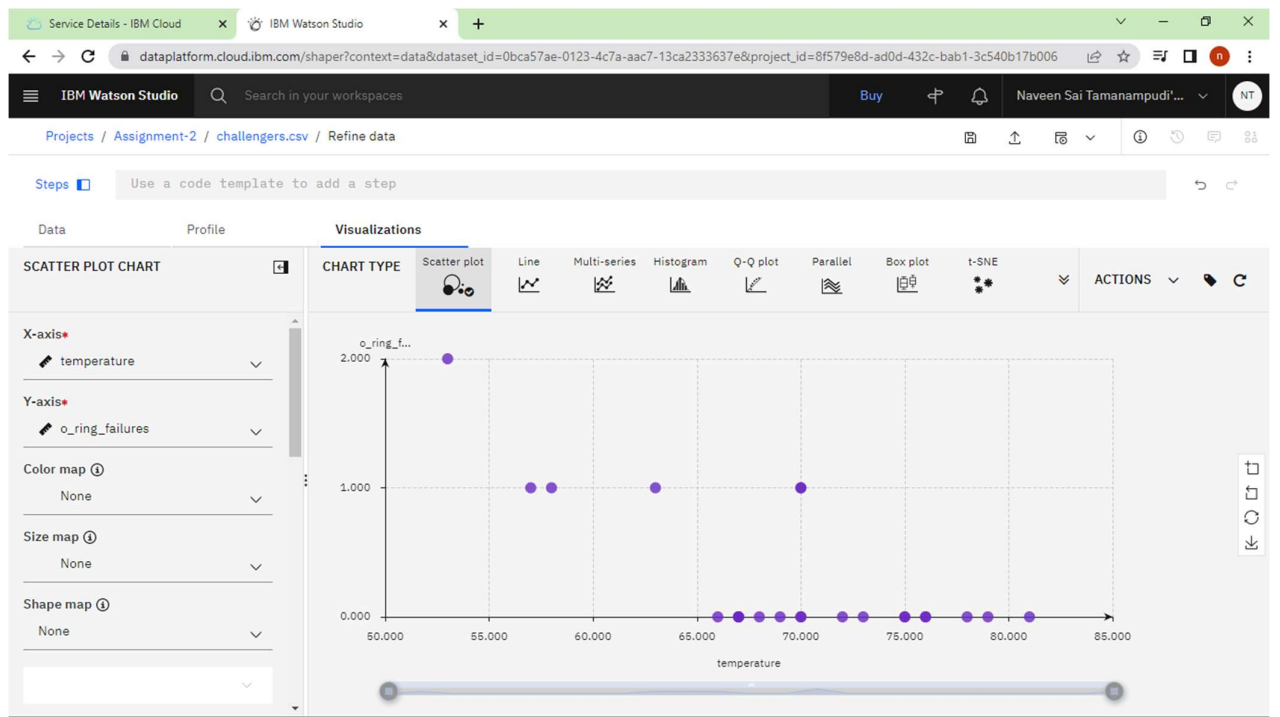


Viewing Profile:

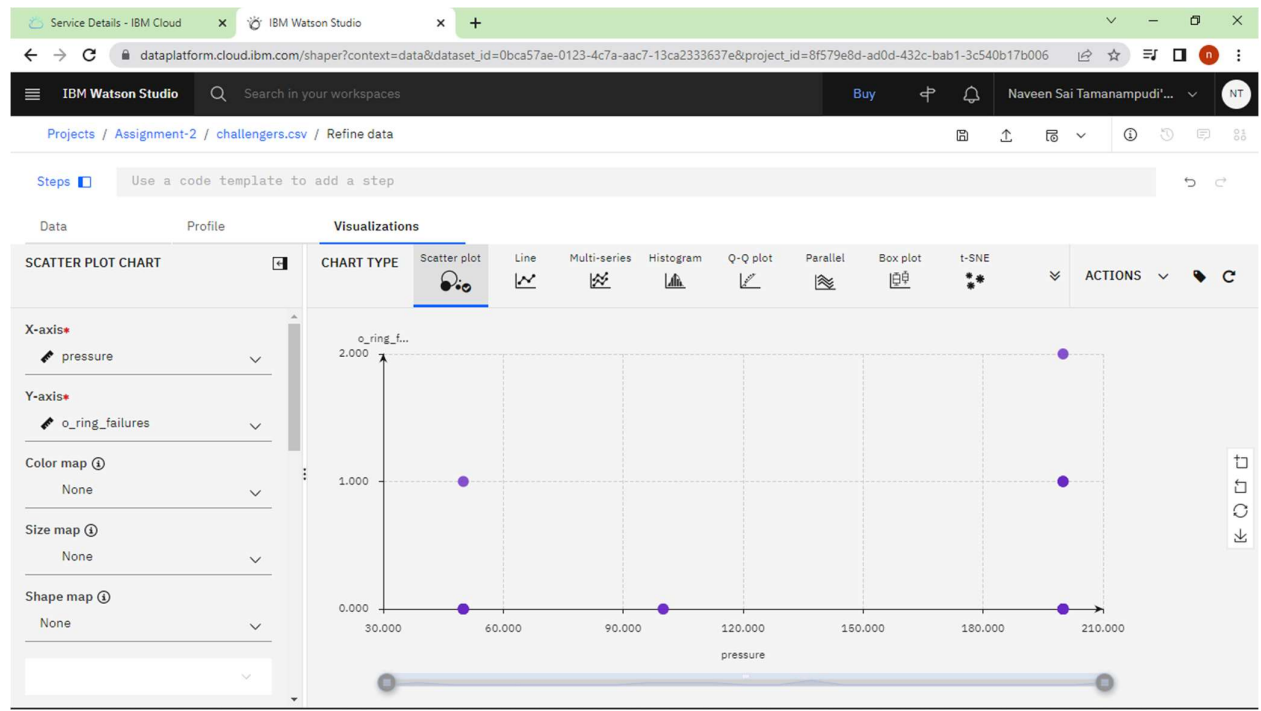




Scatterplot of Temperature vs o_ring_failures:



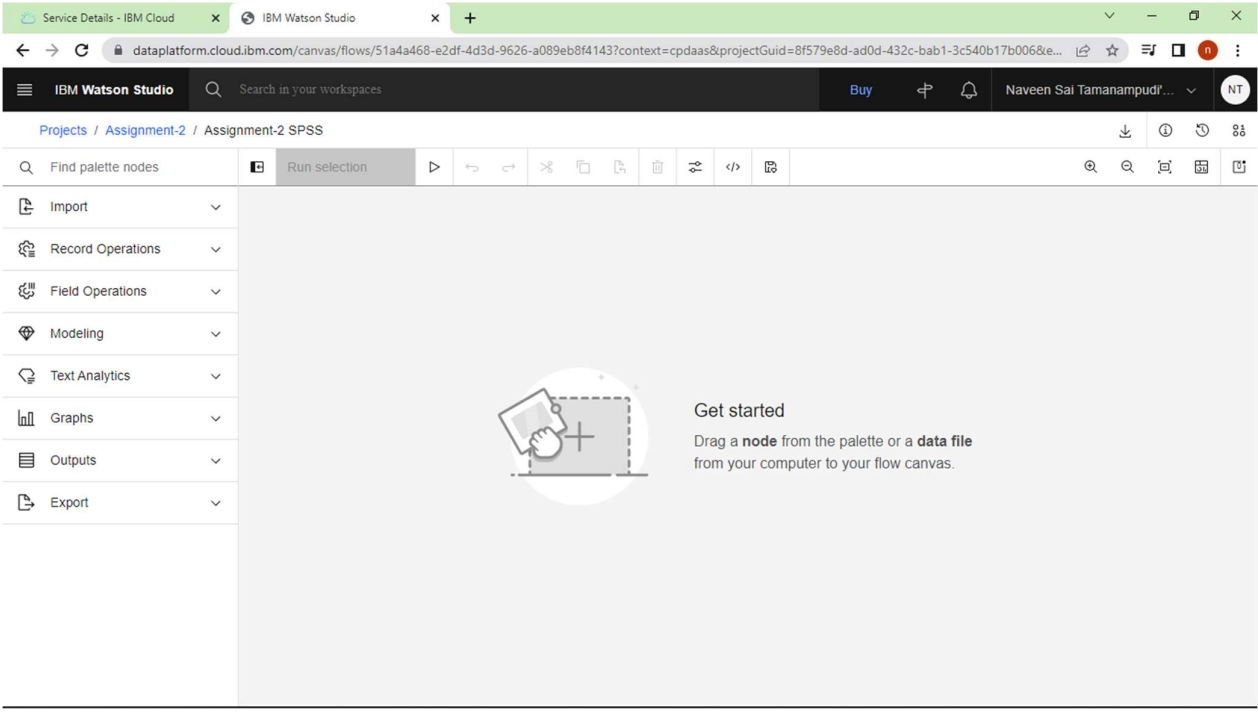
Scatterplot of Pressure vs o ring failures



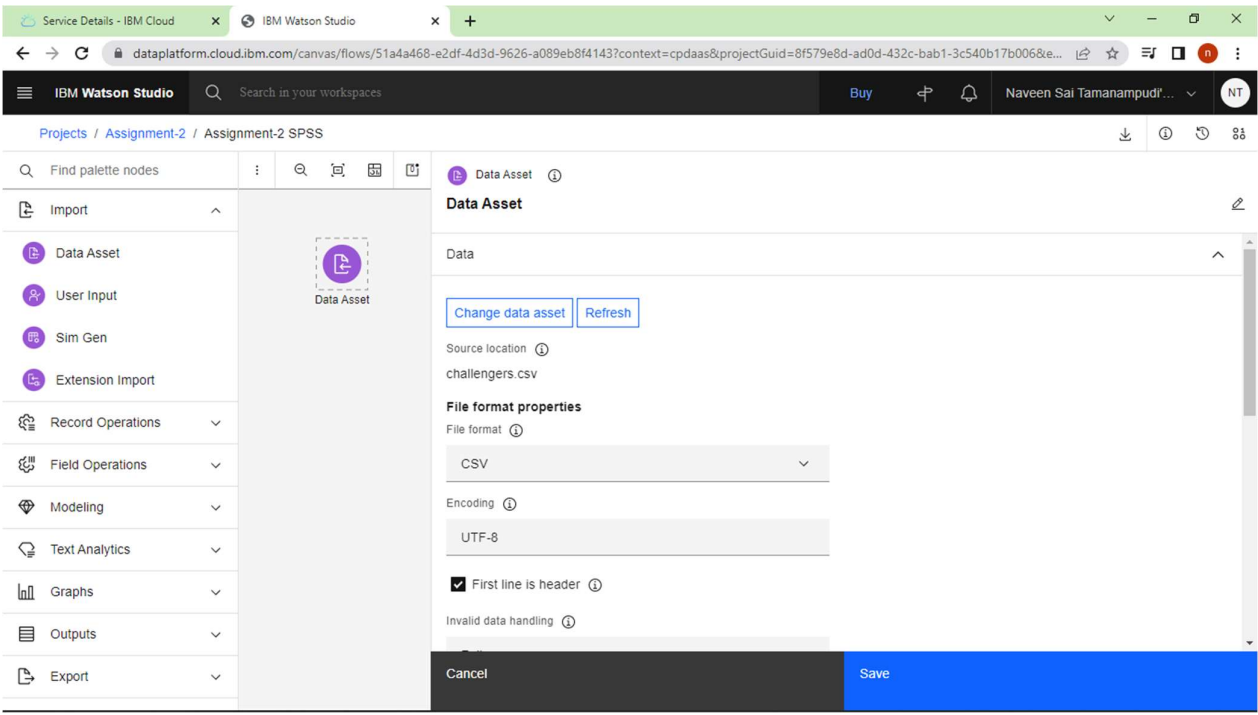
Creating an SPSS Modeler:

The screenshot shows the 'Create an SPSS Modeler flow' form in IBM Watson Studio. The form is divided into two main sections: 'Define details' and 'Define configuration'. The 'Define details' section has a 'Name' field with the value 'Assignment-2 SPSS' and a 'Description (Optional)' field with the text 'What's the purpose of this SPSS Modeler flow?'. The 'Define configuration' section has an 'Environment definition' dropdown menu with the value 'Default SPSS Modeler S (2 vCPU 8 GB RAM)'. Below this, there's a note: 'To create additional runtime environments, view options in the Environments tab.' At the bottom right, there are 'Cancel' and 'Create' buttons.

Created Workspace:



Choosing Dataset in Data Asset Node:



Creating a Data Audit Node:

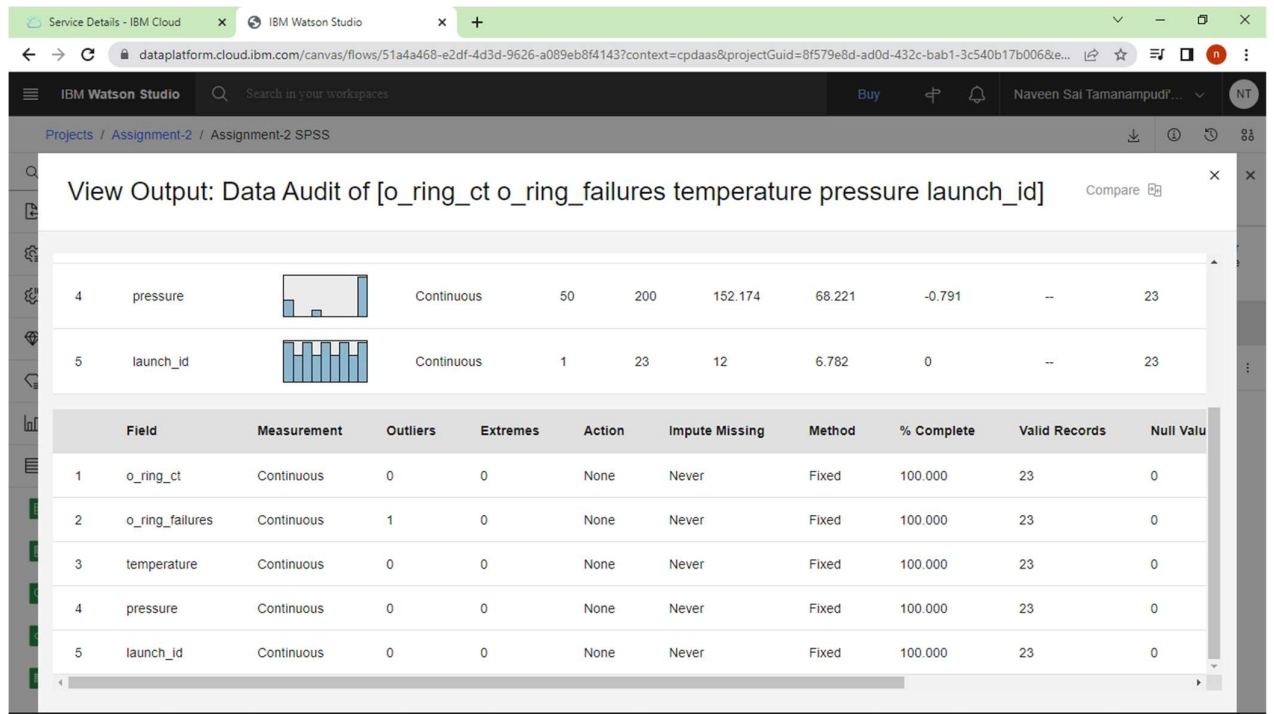
The screenshot shows the IBM Watson Studio interface. On the left, a sidebar lists various nodes: Import, Record Operations, Field Operations, Modeling, Text Analytics, Graphs, Outputs, Table, Matrix, Analysis, Data Audit, and Transform. The 'Data Audit' node is highlighted. In the center workspace, a 'Data Asset' node is connected to a 'Data Audit' node, which is labeled '5 Fields'. On the right, a 'Messages' panel shows a 'Success' message: 'Run was successful'. Below it, an 'Outputs' panel shows a list of outputs, including 'Data Audit of [o_ring_ct o...'.

Output from Data Audit:

The screenshot shows the 'View Output: Data Audit of [o_ring_ct o_ring_failures temperature pressure launch_id]' window. It displays a table with statistical data for five fields. Each row includes a field name, a small histogram, a measurement type, and various statistical metrics.

	Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1	o_ring_ct		Continuous	6	6	6	0	--	--	23
2	o_ring_failures		Continuous	0	2	0.304	0.559	1.735	--	23
3	temperature		Continuous	53	81	69.565	7.057	-0.654	--	23
4	pressure		Continuous	50	200	152.174	68.221	-0.791	--	23
5	launch_id		Continuous	1	23	12	6.782	0	--	23

Below the table, there is a row of tabs: Field, Measurement, Outliers, Extremes, Action, Impute Missing, Method, % Complete, Valid Records, and Null Value.



Creating a Type Node and setting Inputs, Targets:

Service Details - IBM Cloud x IBM Watson Studio x +

dataplatfom.cloud.ibm.com/canvas/flows/51a4a468-e2df-4d3d-9626-a089eb8f4143?context=cpdaas&projectGuid=8f579e8d-ad0d-432c-bab1-3c540b17b006&e...

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudr'... NT

Projects / Assignment-2 / Assignment-2 SPSS

Find palette nodes

- Import
- Record Operations
- Field Operations
- Auto Data Prep
- Type
- Filter
- Derive
- Filler
- Reclassify
- Binning
- RFM Analysis
- Ensemble

Data Asset → Type

5 Fields

Type

Settings

Read values Clear values

Find in column Field

Field	Measure	Role	Value mode	Values
# o_ring_ct	Continuous	None	Read	
# o_ring_failures	Continuous	Target	Read	
# temperature	Continuous	Input	Read	
# pressure	Continuous	Input	Read	
# launch_id	Continuous	None	Read	

Default mode

☒ Read metadata ☐ Pass (do not scan)

Set unique fields to None

Cancel Save

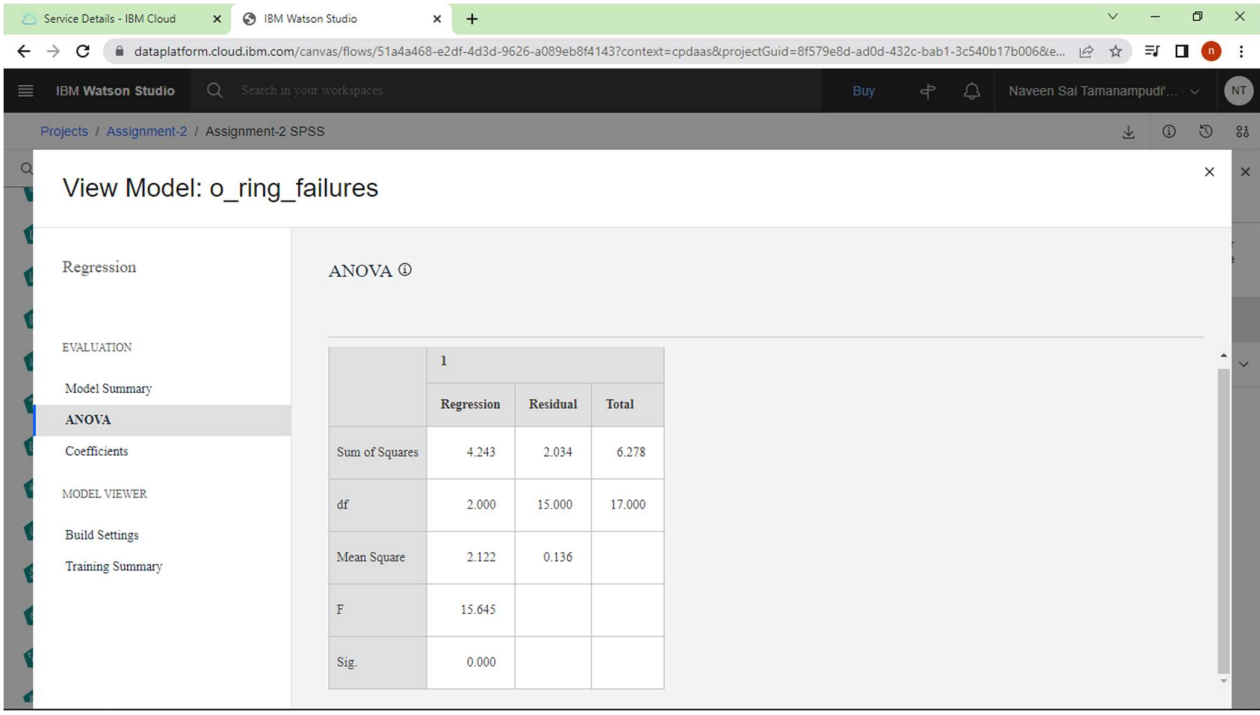
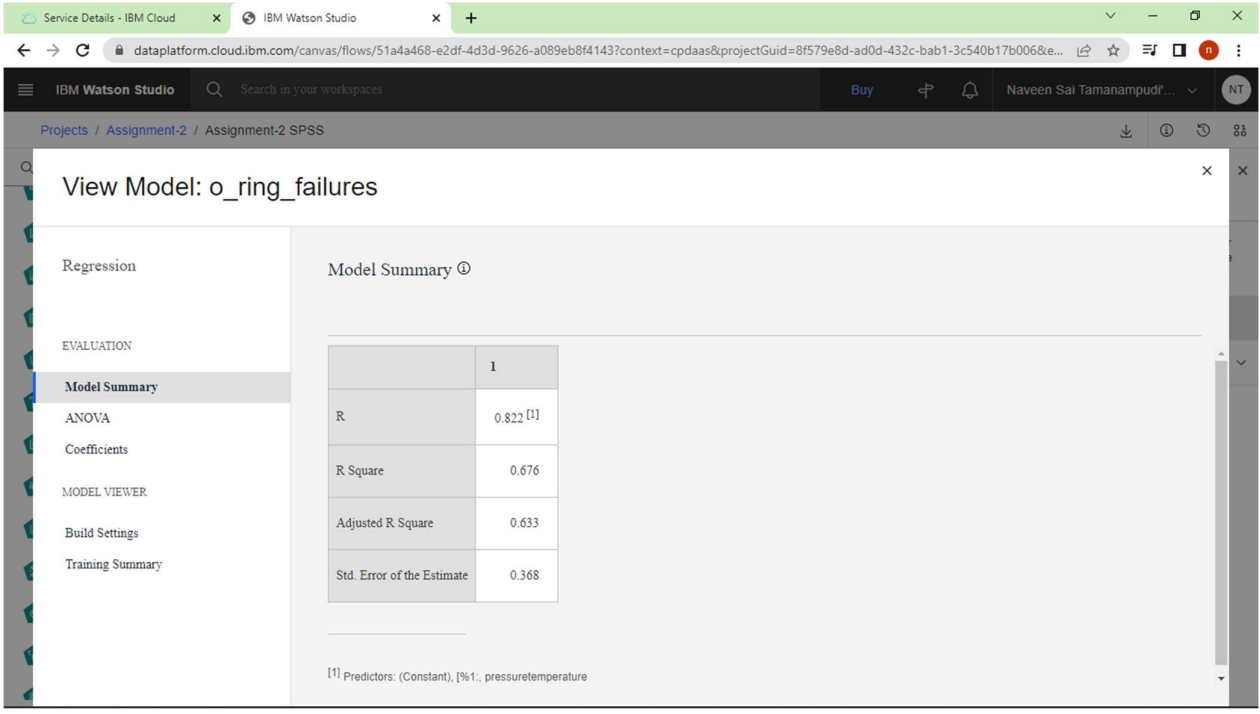
Creating a Partition Node with 80:20 split:

The screenshot shows the IBM Watson Studio interface. On the left, a palette of nodes is visible, including Derive, Filler, Reclassify, Binning, RFM Analysis, Ensemble, Partition, SetToFlag, Restructure, Transpose, Field Reorder, and History. The main workspace displays a flowchart: a 'Data Asset' node points to a '5 Fields' node, which then points to a 'Type' node, and finally to a 'Partition' node. The 'Partition' node is highlighted with a dashed box. On the right, the 'Partition' node's settings panel is open. It shows 'Derived Field Name' as 'Partition', 'Training Partition(%)' as '70', and 'Testing Partition(%)' as '30'. The 'Create validation partition' checkbox is unchecked, and the 'Repeatable partition assignment' checkbox is checked. The 'Seed' field contains '1234567'. At the bottom of the settings panel are 'Cancel' and 'Save' buttons.

Creating a Regression Model:

The screenshot shows the IBM Watson Studio interface with a different flowchart. The 'Data Asset' node points to '5 Fields', which points to 'Type', and then to 'Partition'. The 'Partition' node points to a 'Linear' regression model node. The 'Linear' node has an output labeled 'o_ring_failures'. Below the 'Linear' node is a 'Data Audit of [o_rin...]' node. On the right, the 'Outputs' panel is visible, showing a list of outputs with a checkbox for 'All results' and a dropdown for 'Data Audit of [o_rin...]'.

Output:



Service Details - IBM Cloud x IBM Watson Studio x +

dataplatfom.cloud.ibm.com/canvas/flows/51a4a468-e2df-4d3d-9626-a089eb8f4143?context=cpdaas&projectGuid=8f579e8d-ad0d-432c-bab1-3c540b17b006&e...

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudr... NT

Projects / Assignment-2 / Assignment-2 SPSS

View Model: o_ring_failures

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

MODEL VIEWER

Build Settings

Training Summary

Coefficients

		1		
		(Constant)	temperature	pressure
Unstandardized Coefficients	B	4.181	-0.060	0.003
	Std. Error	0.832	0.012	0.001
Standardized Coefficients				
	Beta		-0.763	0.298
t		5.025	-5.194	2.027
Sig.		0.000	0.000	0.061

Service Details - IBM Cloud x IBM Watson Studio x +

dataplatfom.cloud.ibm.com/canvas/flows/51a4a468-e2df-4d3d-9626-a089eb8f4143?context=cpdaas&projectGuid=8f579e8d-ad0d-432c-bab1-3c540b17b006&e...

IBM Watson Studio Search in your workspaces Buy Naveen Sai Tamanampudr... NT

Projects / Assignment-2 / Assignment-2 SPSS

View Model: o_ring_failures

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

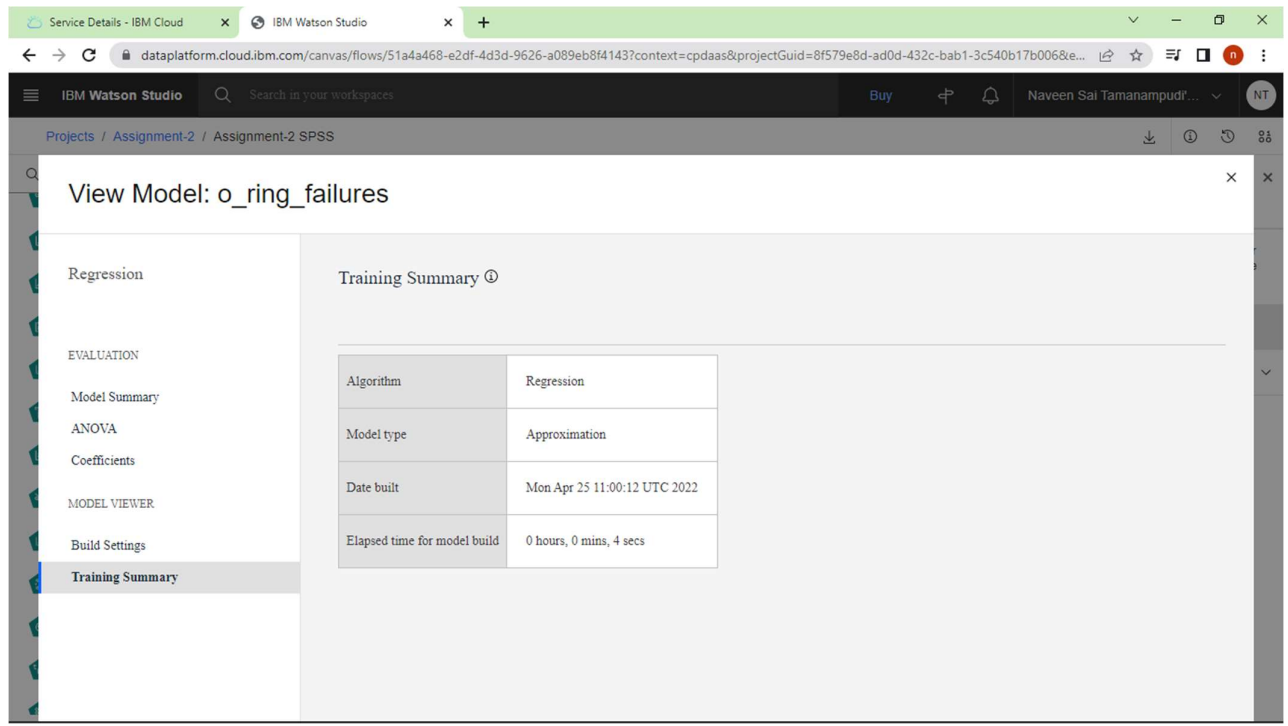
MODEL VIEWER

Build Settings

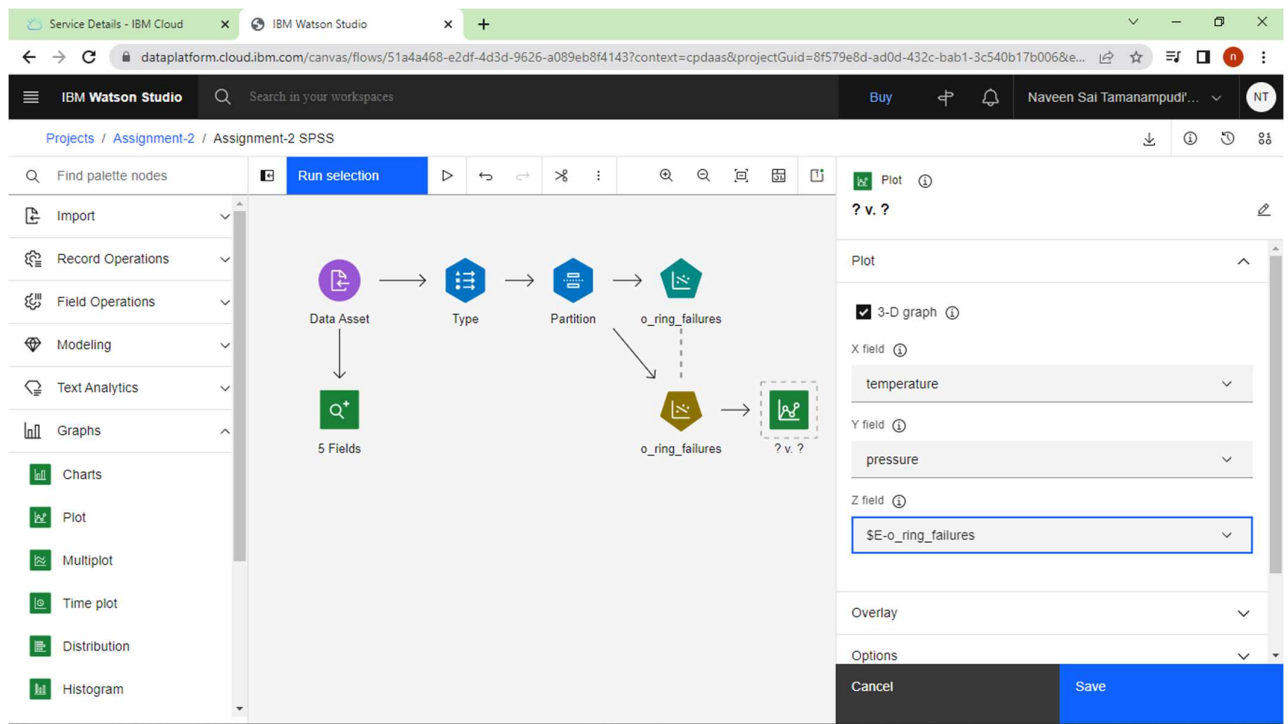
Training Summary

Build Settings

Use partitioned data	true
Calculate predictor importance	true
Method	Enter
Include constant in equation	true
Use weight	false
Mode	Simple



Creating a Plot Node:



Output:

Service Details - IBM Cloud x IBM Watson Studio x +

dataplatfom.cloud.ibm.com/canvas/flows/51a4a468-e2df-4d3d-9626-a089eb8f4143?context=cpdaas&projectGuid=8f579e8d-ad0d-432c-bab1-3c540b17b006&e...

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

Projects / Assignment-2 / Assignment-2 SPSS

Find palette nodes Run selection

Import Record Operations Field Operations Modeling Text Analytics Graphs Charts Plot Multiplot Time plot Distribution Histogram

Data Asset Type Partition o_ring_failures 5 Fields o_ring_failures ? v. ?

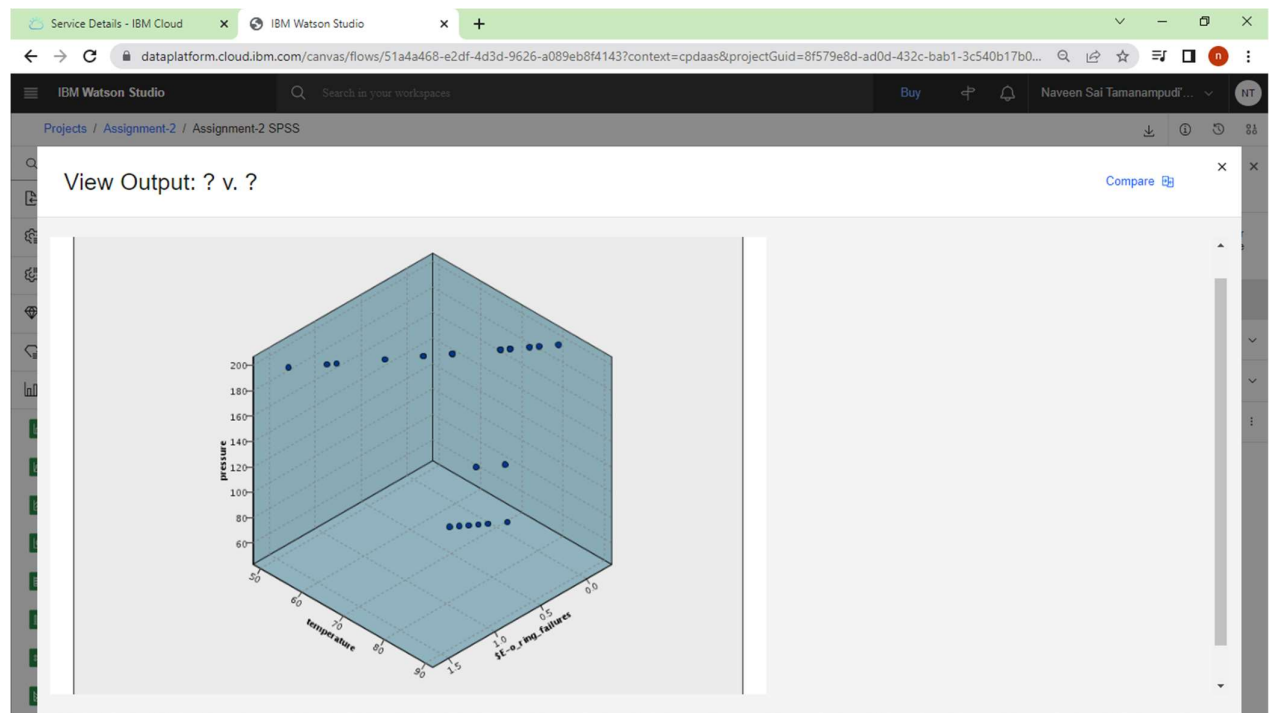
Outputs Models

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

All results

Data Audit of [o_rin... Just now

? v. ? Just now



Service Details - IBM Cloud x IBM Watson Studio x +

dataplatfom.cloud.ibm.com/canvas/flows/51a4a68-e2df-4d3d-9626-a089eb8f4143?context=cpdaas&projectGuid=8f579e8d-ad0d-432c-bab1-3c540b17b0...

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

Projects / Assignment-2 / Assignment-2 SPSS

Find palette nodes Run selection

Import Record Operations Field Operations Modeling Text Analytics Graphs

Charts Plot Multiplot Time plot Distribution Histogram Collection Web

Data Asset Type Partition o_ring_failures

5 Fields o_ring_failures ? v. ?

Plot ? v. ?

3-D graph

X field temperature

Y field pressure

Z field o_ring_failures

Overlay Options Appearance

Cancel Save

