

Data Analytics

Assignment-2

Name: G.Naga Sai Praneeth

Regd.No:19MIM10044

Risk Analysis using Multiple Linear Regression

Dataset: Challengers.csv

Create Project:

The screenshot displays the IBM Watson Studio web interface. The browser address bar shows the URL: eu-gb.dataplatform.cloud.ibm.com/projects/88d4b70f-742e-43d7-8e01-b3e8ee8d3212/assets?context=cpdaas#. The top navigation bar includes the IBM Watson Studio logo, a search bar, and user account information for 'Praneeth Gadipudi's Accou...'. The main navigation tabs are Overview, Assets (selected), Jobs, and Manage. On the left sidebar, under 'Find assets', there is a section for '1 assets' and 'Asset types' with a filter for 'Data' (1 asset). The main content area shows a table of data assets. The table has columns for 'Name' and 'Last modified'. A single asset is listed: 'challengers.csv' (CSV format), modified '15 minutes ago' by 'Praneeth Gadipudi (You)'. There are buttons for 'Add asset' and 'New asset' in the top right of the asset list.

Name	Last modified
challengers.csv CSV	15 minutes ago Praneeth Gadipudi (You)

Create Data Refinery and add dataset:

eu-gb.dataplatform.cloud.ibm.com/shaper?context=data&dataset_id=a51244d2-9765-47a6-8ed3-5b34077b3d46&project_id=88d4b70f-742e-43d7-8e01-b3e8ee8d3212

IBM Watson Studio

Search in your workspaces

Buy

Praneeth Gadipudi's Accou...

PG

Projects / Challengers / challengers.csv / Refine data

Steps

Use a code template to add a step

Data

Profile

Visualizations

	o_ring_ct Integer	o_ring_failu... Integer	temperature Integer	pressure Integer	launch_id Integer
1	6	0	66	50	1
2	6	1	70	50	2
3	6	0	69	50	3
4	6	0	68	50	4
5	6	0	67	50	5
6	6	0	72	50	6
7	6	0	73	100	7
8	6	0	70	100	8
9	6	1	57	200	9
10	6	1	63	200	10
11	6	1	70	200	11
12	6	0	78	200	12
13	6	0	67	200	13
14	6	2	53	200	14

SOURCE FILE: challengers.csv FULL DATA SET: 23 rows

Profile:

eu-gb.dataplatform.cloud.ibm.com/shaper?context=data&dataset_id=a51244d2-9765-47a6-8ed3-5b34077b3d46&project_id=88d4b70f-742e-43d7-8e01-b3e8ee8d3212

IBM Watson Studio

Search in your workspaces

Buy

Praneeth Gadipudi's Accou...

PG

Projects / Challengers / challengers.csv / Refine data

Steps

Use a code template to add a step

Data

Profile

Visualizations

o_ring_ct
Integer

FREQUENCY

STATISTICS

Interquartile Range	6
Minimum	6

o_ring_failures
Integer

FREQUENCY

STATISTICS

Interquartile Range	0
Minimum	0

temperature
Integer

FREQUENCY

STATISTICS

Interquartile Range	0.5
Minimum	0

pressure
Integer

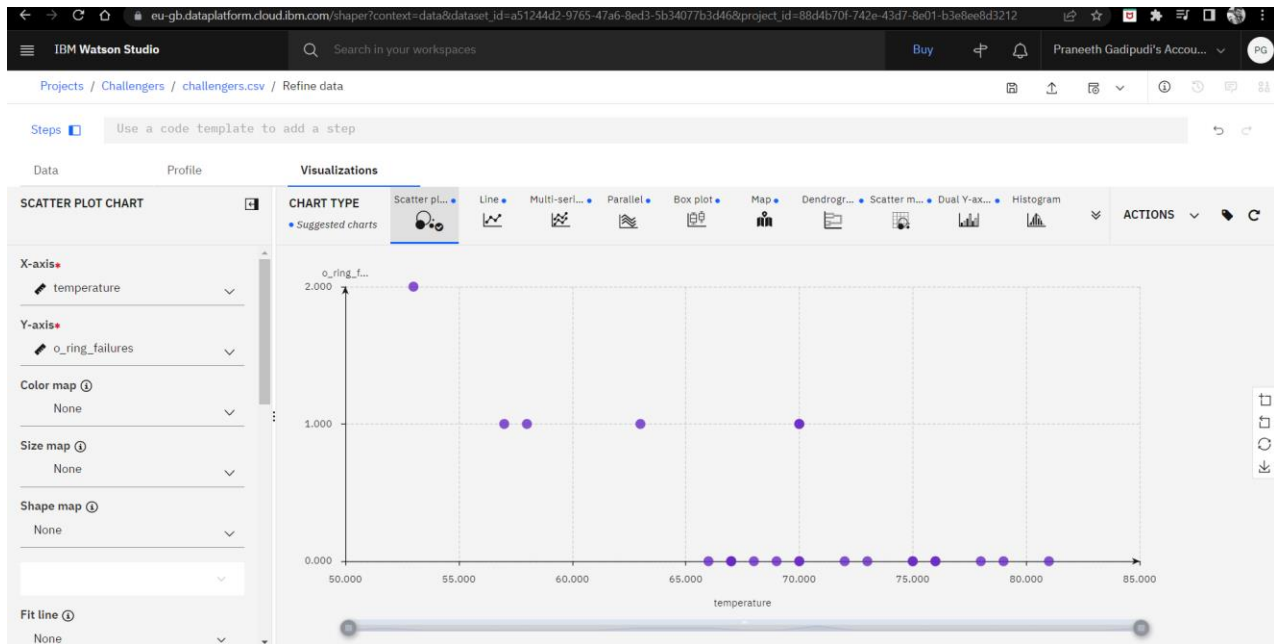
FREQUENCY

STATISTICS

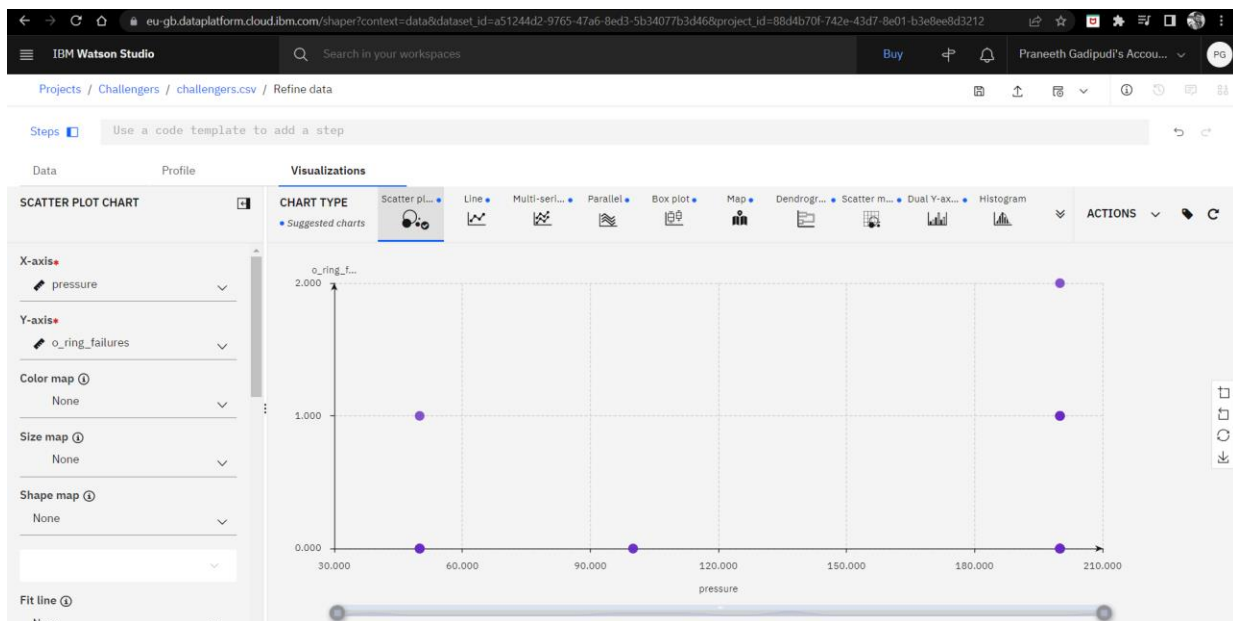
Interquartile Range	8
Minimum	53

Visualization:

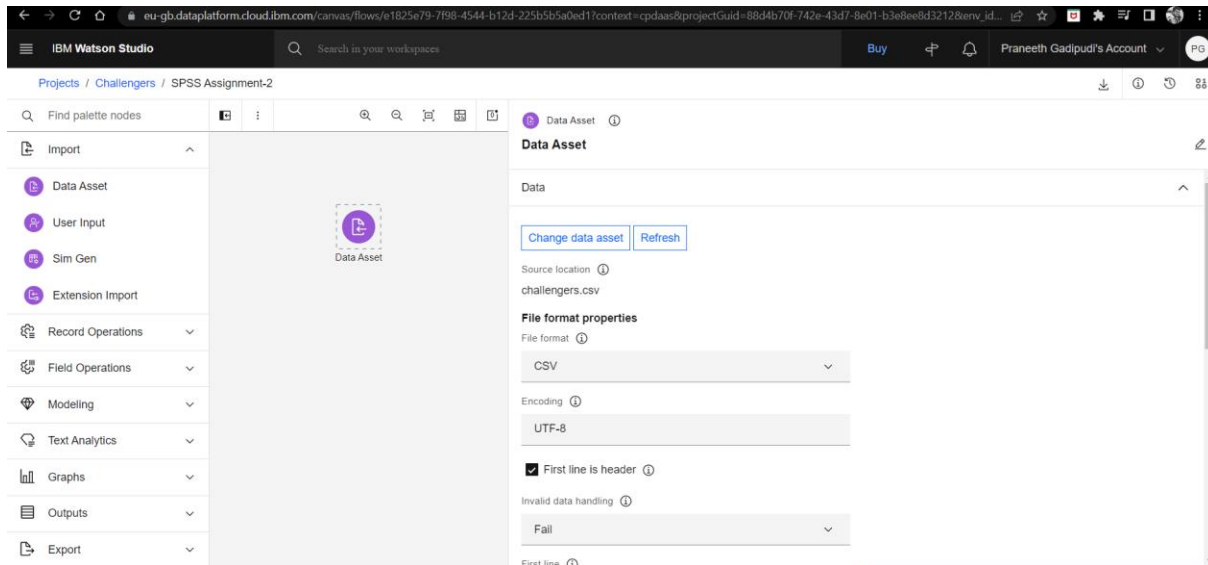
Scatter Plot of temp vs o ring failures



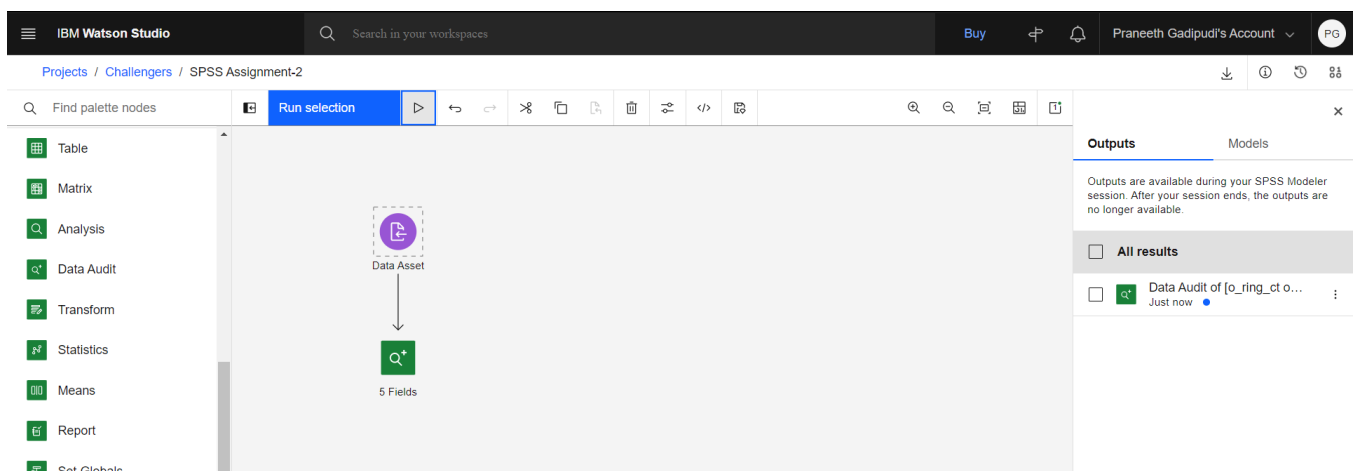
Pressure vs o ring failures:



Create a SPSS Modeler and add dataset:








Create Data Audit:



Output Data Audit:

View Output: Data Audit of [o_ring_ct o_ring_failures temperature pressure launch_id]

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

	Field	Measurement	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value	Empty String
1	o_ring_ct	Continuous	0	0	None	Never	Fixed	100.000	23	0	0
2	o_rino_failures	Continuous	1	0	None	Never	Fixed	100.000	23	0	0

Type Node:

Find palette nodes

Import

Record Operations

Field Operations

Auto Data Prep

Type

Filter

Derive

Filler

Reclassify

Binning

RFM Analysis

Ensemble

Partition

Data Asset

5 Fields

Type

Settings

Read values Clear values

Find in column Field

Field	Measure	Role	Value mode	Values
# o_ring_ct	Continuous	Input	Read	
# o_ring_failures	Continuous	Input	Read	
# temperature	Continuous	Input	Read	
# pressure	Continuous	Input	Read	
# launch_id	Continuous	Input	Read	

Default mode

☒ Read metadata ☐ Pass (do not scan)

Set unique fields to None

Set categorical fields to None if they exceed this many values

Set continuous integer field to ordinal if range less than or equal to

Cancel Save

Partition Node:

The screenshot shows the IBM Watson Studio interface for configuring a Partition node. The workflow consists of three nodes: Data Asset, Type, and Partition. The Data Asset node is connected to the Type node, which is connected to the Partition node. The Partition node is selected, and its settings are displayed on the right. The settings include:

- Derived Field Name: Partition
- Training Partition(%): 70
- Testing Partition(%): 30
- Create validation partition: ☐
- Repeatable partition assignment: ☒
- Seed: 1234567
- Use unique field to assign partitions: ☐

The Partition node is highlighted with a dashed border, and the 'Run selection' button is visible in the top toolbar.

Creating a Regression Model:

The screenshot shows the IBM Watson Studio interface for creating a Regression Model. The workflow consists of four nodes: Data Asset, Type, Partition, and o_ring_failures. The Data Asset node is connected to the Type node, which is connected to the Partition node, which is connected to the o_ring_failures node. The o_ring_failures node is selected, and its settings are displayed on the right. The settings include:

- Derived Field Name: o_ring_failures
- Training Partition(%): 70
- Testing Partition(%): 30
- Create validation partition: ☐
- Repeatable partition assignment: ☒
- Seed: 1234567
- Use unique field to assign partitions: ☐

The o_ring_failures node is highlighted with a dashed border, and the 'Run selection' button is visible in the top toolbar. The 'Outputs' tab is selected on the right, showing the results of the model training.

Output:

View Model: o_ring_failures

Regression	ANOVA ⓘ
EVALUATION	
Model Summary	
ANOVA	
Coefficients	
MODEL VIEWER	
Build Settings	
Training Summary	

	1		
	Regression	Residual	Total
Sum of Squares	3.686	2.592	6.278
df	1.000	16.000	17.000
Mean Square	3.686	0.162	
F	22.758		
Sig.	0.000		

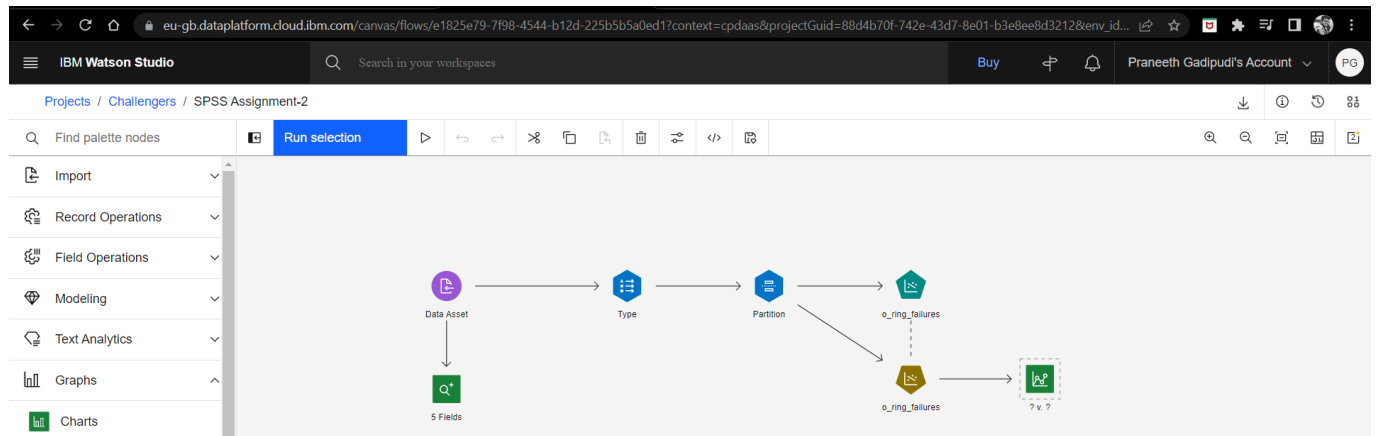
View Model: o_ring_failures

Regression	Model Summary ⓘ
EVALUATION	
Model Summary	
ANOVA	
Coefficients	
MODEL VIEWER	
Build Settings	
Training Summary	

	1
R	0.766 [1]
R Square	0.587
Adjusted R Square	0.561
Std. Error of the Estimate	0.402

[1] Predictors: (Constant), [%1:, temperature

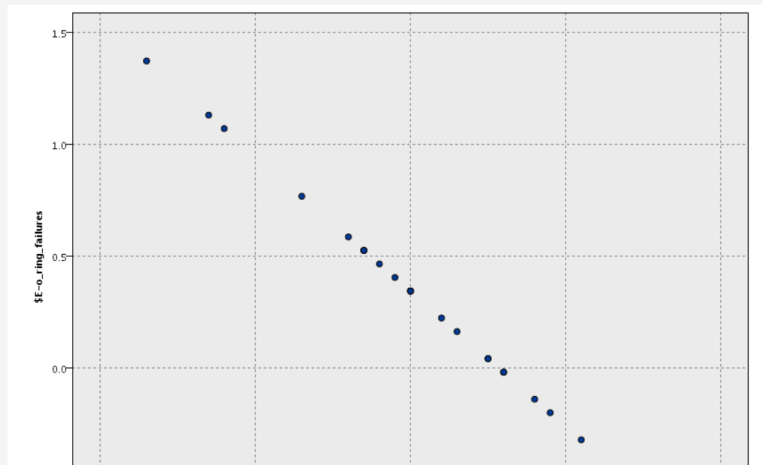
Plot Node:



2D-Output:

View Output: ? v. ?

[Compare](#)



3D Output:

