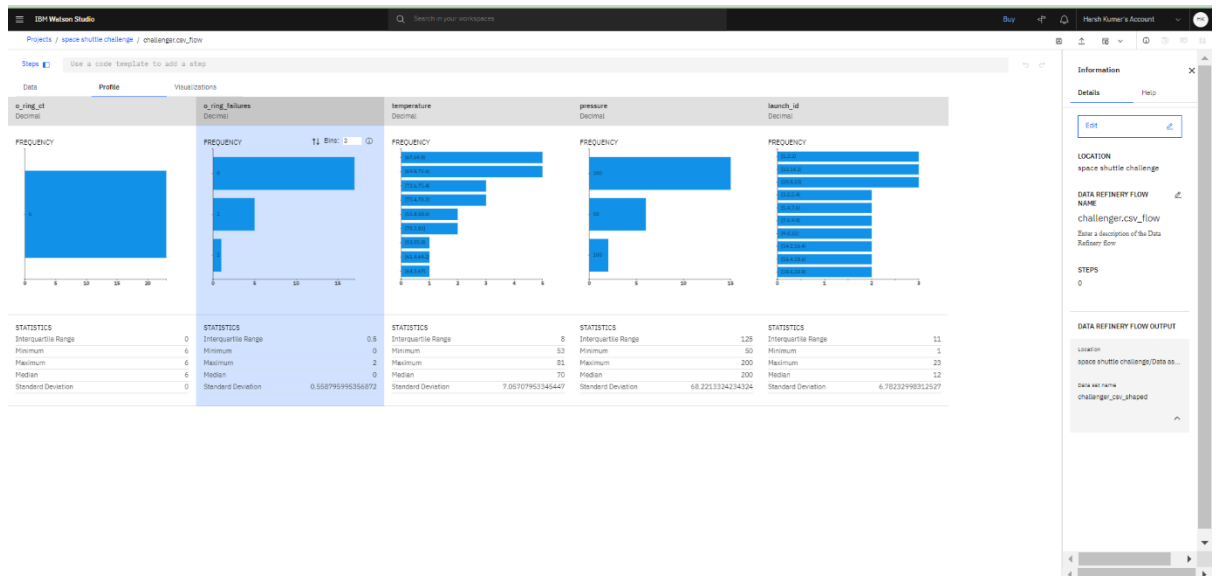


Data Analytics

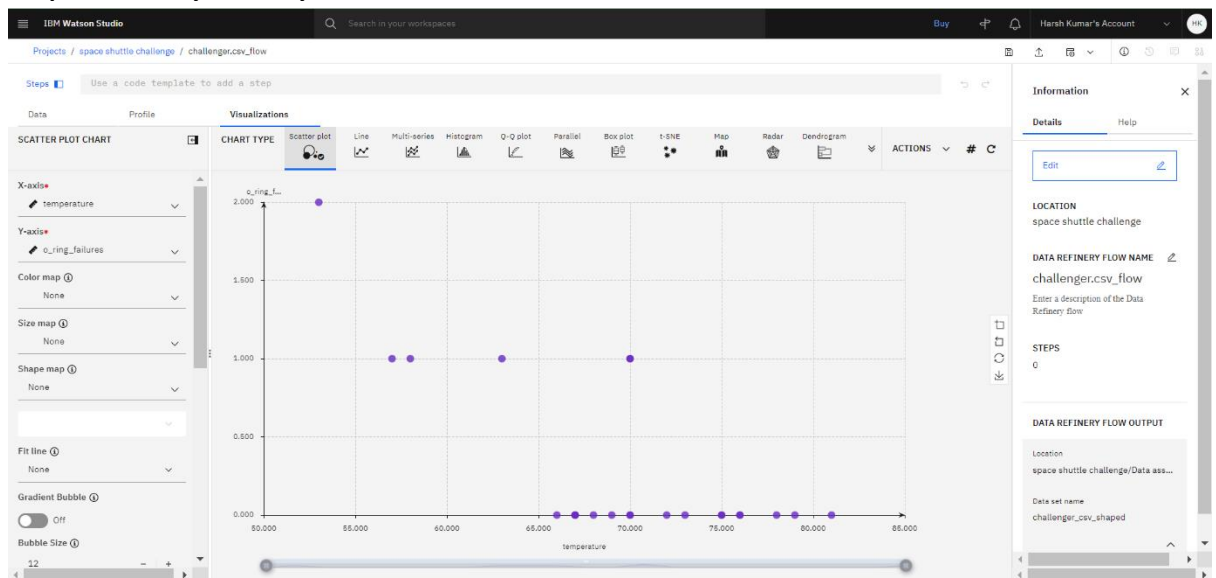
Case Study: Risk Analysis of Space Shuttle, Multiple Linear Regression analysis.

1. Data Refinery

- Descriptive analytics of data



- Exploratory Analytics



x-axis = temperature ,y-axis – o_ring_failures

- At Lower temperature failure are more,while at higher temperature failure are very less.

IBM Watson Studio

Search in your workspace

Buy Harsh Kumar's Account

Projects / space shuttle challenge / schedule launch

Find palette nodes

Run selection

Import

Record Operations

Field Operations

Modeling

Text Analytics

Graphs

Outputs

Export

Data asset

Type

5 Fields

Type

Settings

Read values Clear values

Find in column Field

Field	Measure	Role	Value mode	Values
<input type="checkbox"/> # o_ring_ct	Continuous	None	Instantiated	6, 6
<input type="checkbox"/> # o_ring_failures	Continuous	Target	Instantiated	0, 2
<input type="checkbox"/> # temperature	Continuous	Input	Instantiated	53, 81
<input type="checkbox"/> # pressure	Continuous	Input	Instantiated	50, 200
<input type="checkbox"/> # launch_id	Continuous	None	Instantiated	1, 23

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
 - Training – 70%
 - Testing – 30%

IBM Watson Studio

Search in your workspace

Buy Harsh Kumar's Account

Projects / space shuttle challenge / schedule launch

Find palette nodes

Run selection

Import

Record Operations

Field Operations

Modeling

Text Analytics

Graphs

Outputs

Export

Data asset

Type

Partition

5 Fields

Partition

Settings

Derived Field Name

Partition

Training Partition(%)

70

Testing Partition(%)

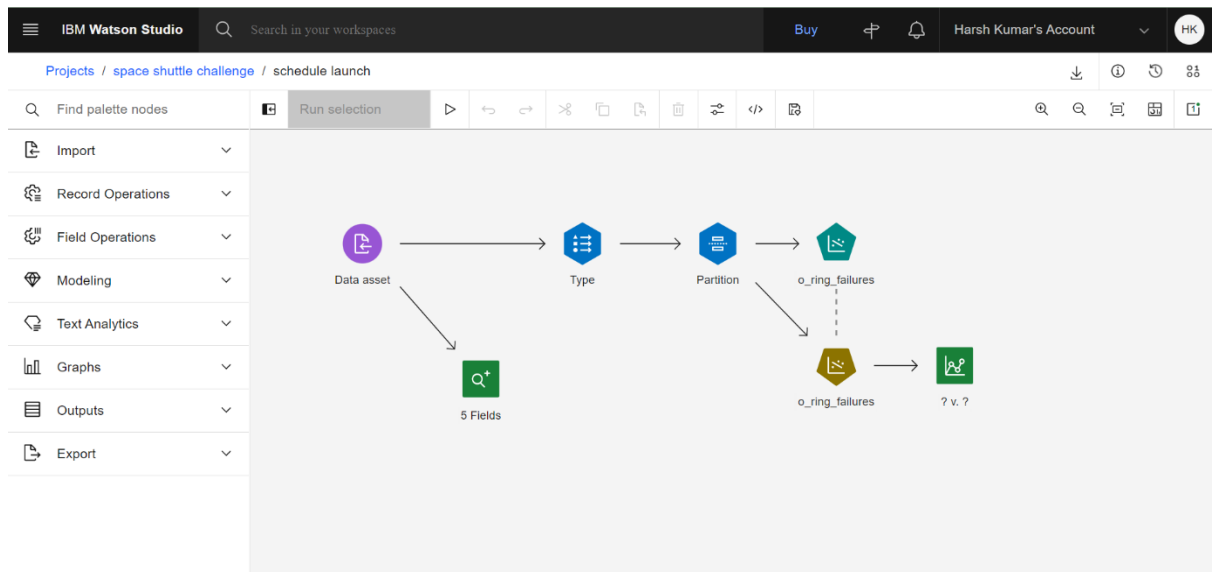
30

☐ Create validation partition

☒ Repeatable partition assignment

Cancel Save

- Model



- **Output**
 - **View Model**

IBM Watson Studio

Search in your workspaces

Buy

Harsh Kumar's Account

Projects / space shuttle challenge / schedule launch

View Model: o_ring_failures

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

MODEL VIEWER

Build Settings

Training Summary

Model Summary

Model Summary

	1
R	0.822 [1]
R Square	0.676
Adjusted R Square	0.633
Std. Error of the Estimate	0.368

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

IBM Watson Studio

Search in your workspaces

Buy

Harsh Kumar's Account

Projects / space shuttle challenge / schedule launch

View Model: o_ring_failures

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

MODEL VIEWER

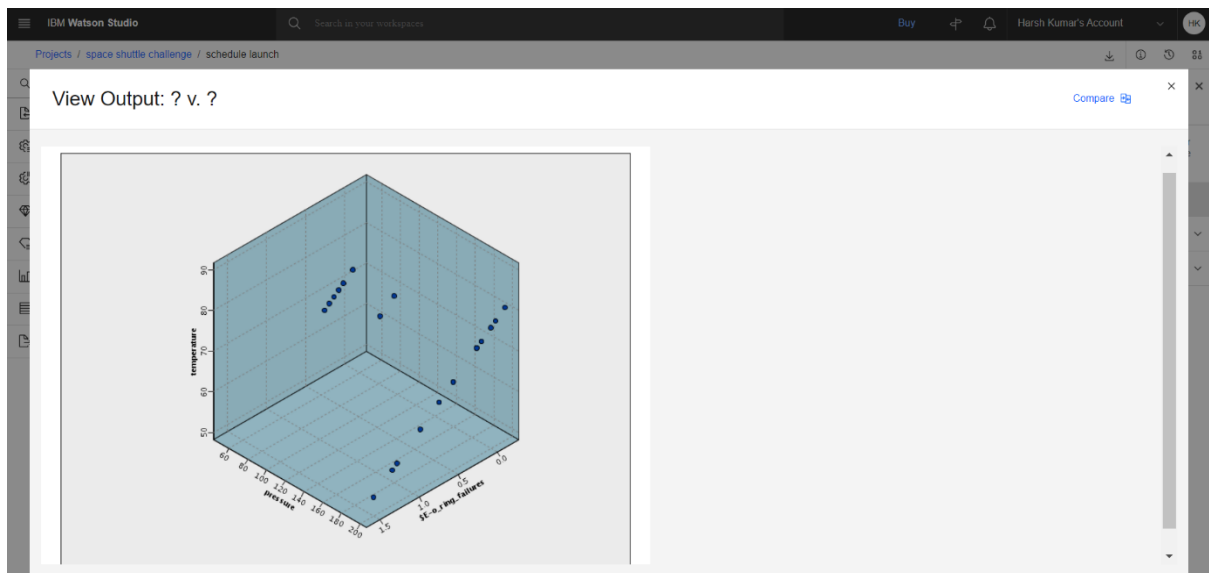
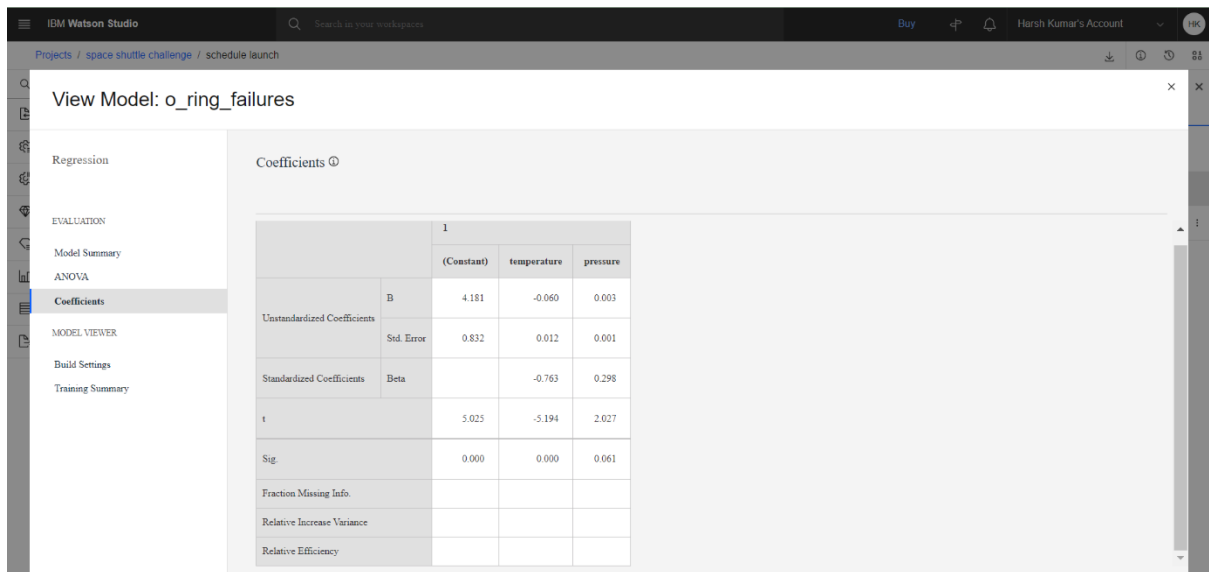
Build Settings

Training Summary

ANOVA

ANOVA

	1		
	Regression	Residual	Total
Sum of Squares	4.243	2.034	6.278
df	2.000	15.000	17.000
Mean Square	2.122	0.136	
F	15.645		
Sig.	0.000		



Data Audit

IBM Watson Studio






Search in your workspaces

Buy Harsh Kumar's Account

Projects / space shuttle challenge / schedule launch

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

Compare

	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	7.067	-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	White Space	Blank Value
1	o_ring_ct	Continuous	0	0	None	Never	Fixed	100.000	23	0	0	0	0
2	o_ring_failures	Continuous	1	0	None	Never	Fixed	100.000	23	0	0	0	0
3	temperature	Continuous	0	0	None	Never	Fixed	100.000	23	0	0	0	0
4	pressure	Continuous	0	0	None	Never	Fixed	100.000	23	0	0	0	0
5	launch_id	Continuous	0	0	None	Never	Fixed	100.000	23	0	0	0	0