

Deepak Raj Gupta

19BCE2620

VIT – Vellore

Assignment 02

Problem Statement:

Visualize the following data in IBM cloud using Suitable Models.

1) Bank.csv

IBM Watson Studio interface showing the 'Bank.csv' dataset in the 'Data' tab. The table displays 16 rows of data with columns: age, job, marital, education, default, balance, housing, and loan. A right-hand sidebar shows the 'Details' tab with fields for 'LOCATION' (Bank), 'DATA REFINERY FLOW NAME' (bank.csv_flow), and 'DATA REFINERY FLOW OUTPUT'.

	age	job	marital	education	default	balance	housing	loan
1	59	admin.	married	secondary	no	2343	yes	no
2	56	admin.	married	secondary	no	45	no	no
3	41	technician	married	secondary	no	1270	yes	no
4	55	services	married	secondary	no	2476	yes	no
5	54	admin.	married	tertiary	no	184	no	no
6	42	management	single	tertiary	no	0	yes	yes
7	56	management	married	tertiary	no	830	yes	yes
8	60	retired	divorced	secondary	no	545	yes	no
9	37	technician	married	secondary	no	1	yes	no
10	28	services	single	secondary	no	5090	yes	no
11	38	admin.	single	secondary	no	100	yes	no
12	30	blue-collar	married	secondary	no	309	yes	no
13	29	management	married	tertiary	no	199	yes	yes
14	46	blue-collar	single	tertiary	no	460	yes	no
15	31	technician	single	tertiary	no	703	yes	no
16	35	management	divorced	tertiary	no	3837	yes	no

Fig: Data Refining

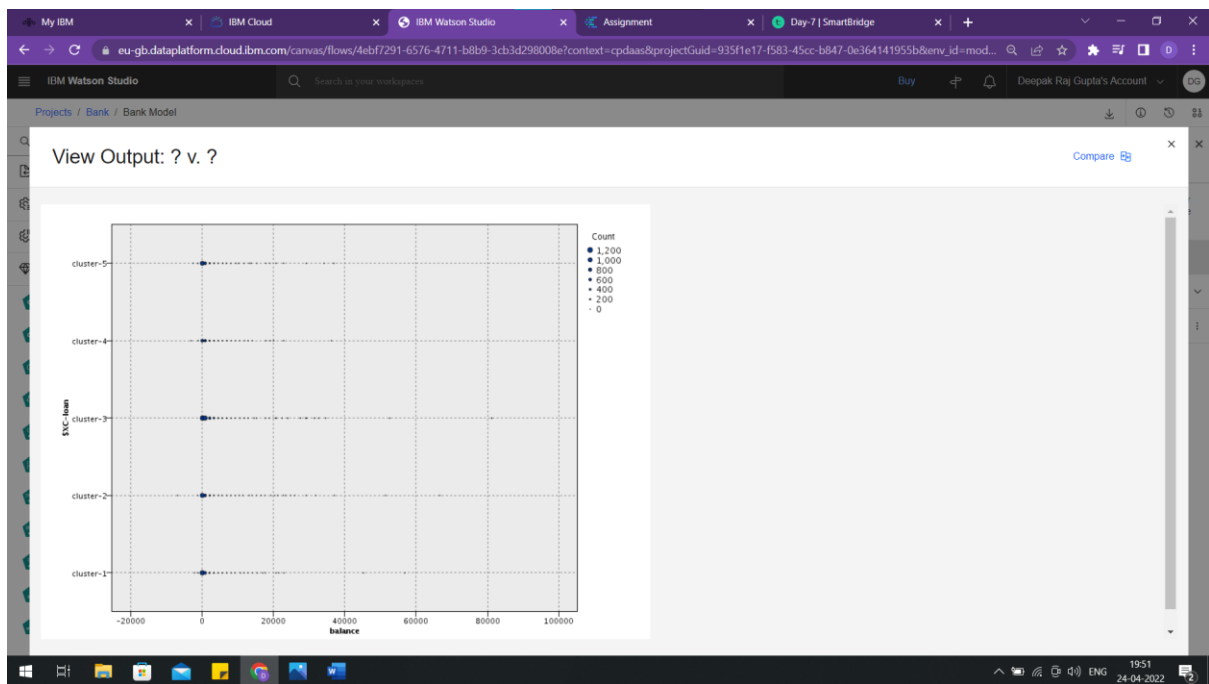


Fig: Cluster Visualization in graph

2) Challenger.csv

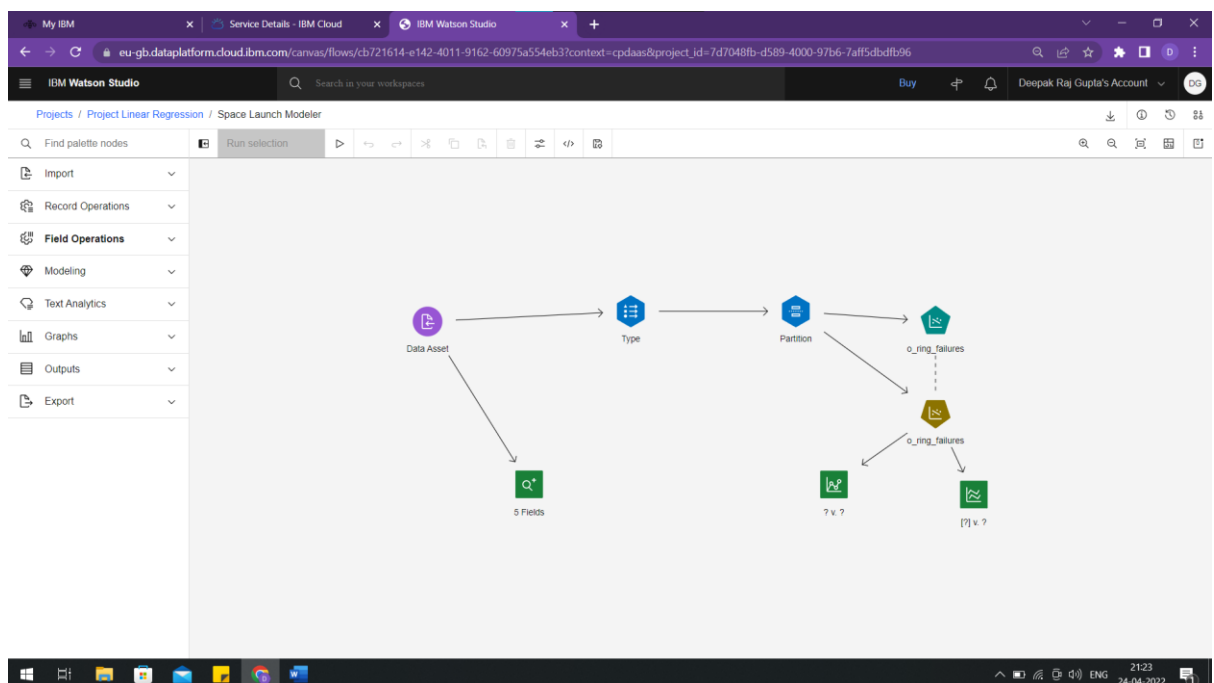


Fig: Regression Model creation

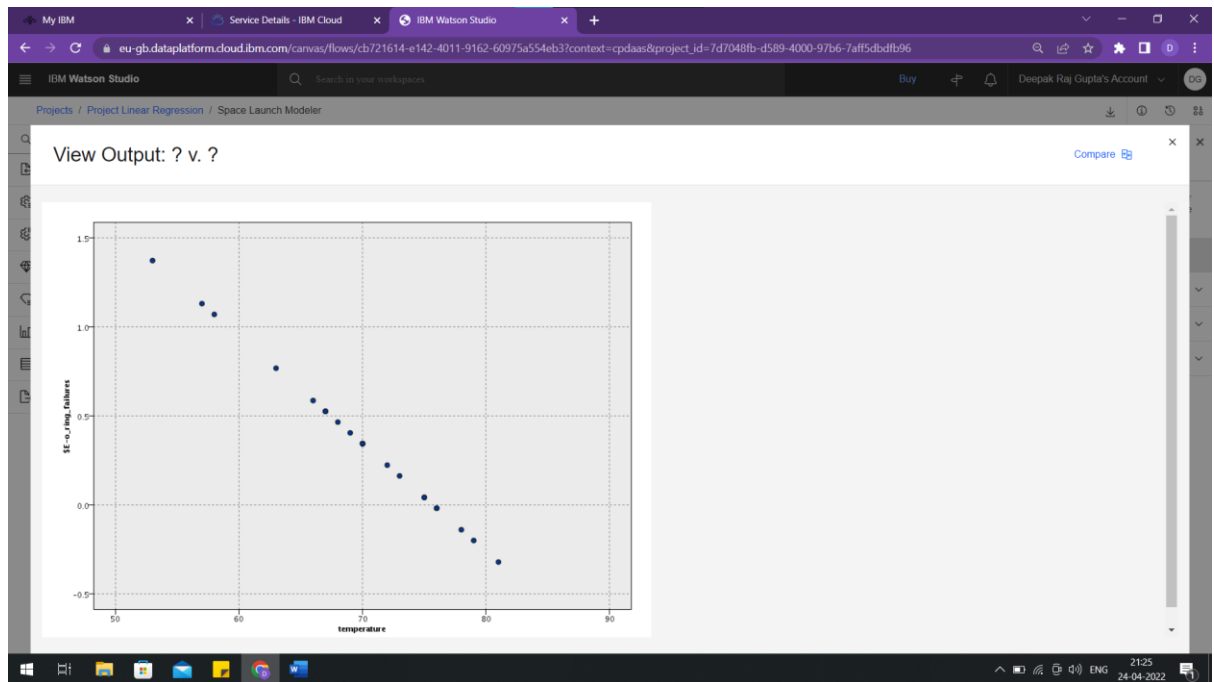


Fig: Linear Regression (Temperature vs O_ring_failures)

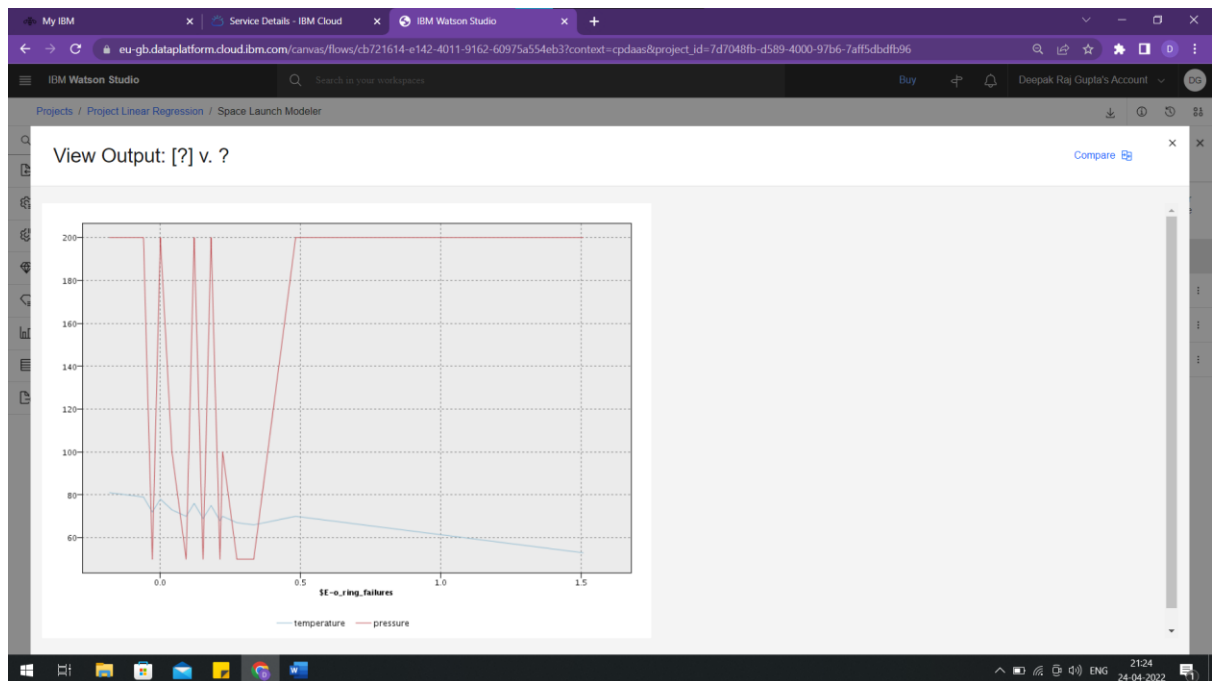


Fig: Multiple Linear Regression (Temperature and Pressure vs O_ring_failures)