

Deepak Raj Gupta

19BCE2620

VIT - Vellore

## Assignment 03

### Problem Statement:

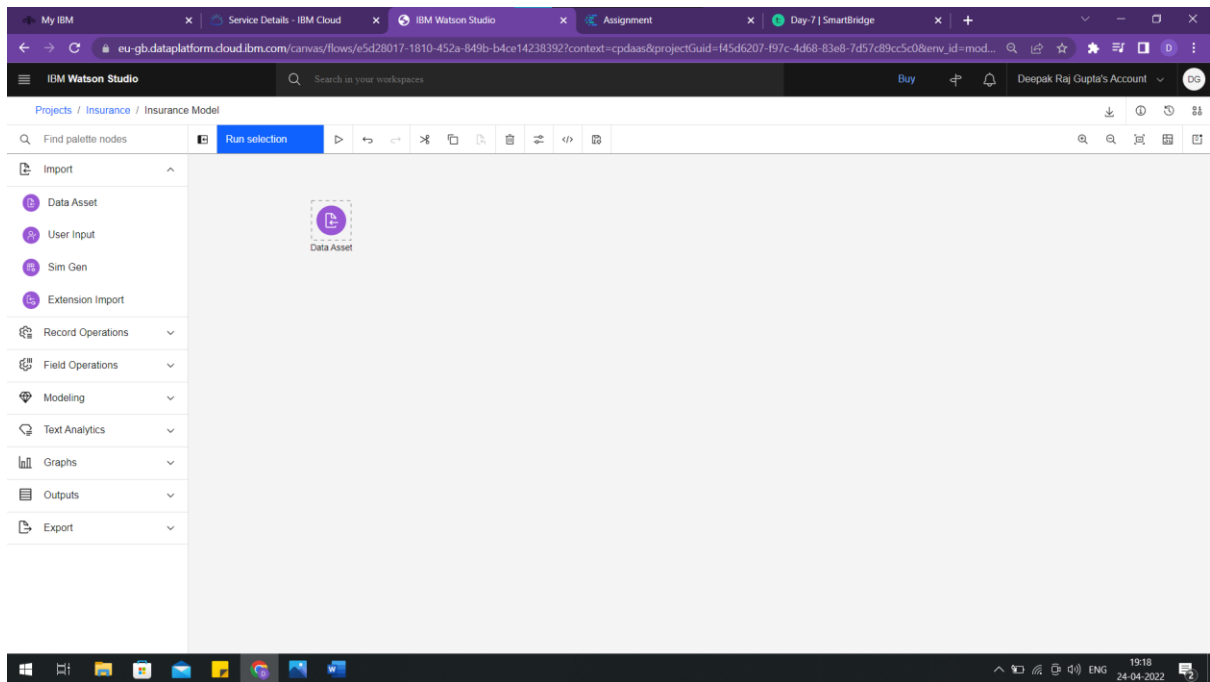
Visualize the following data in IBM cloud using Suitable Models.

#### 1) Insurance

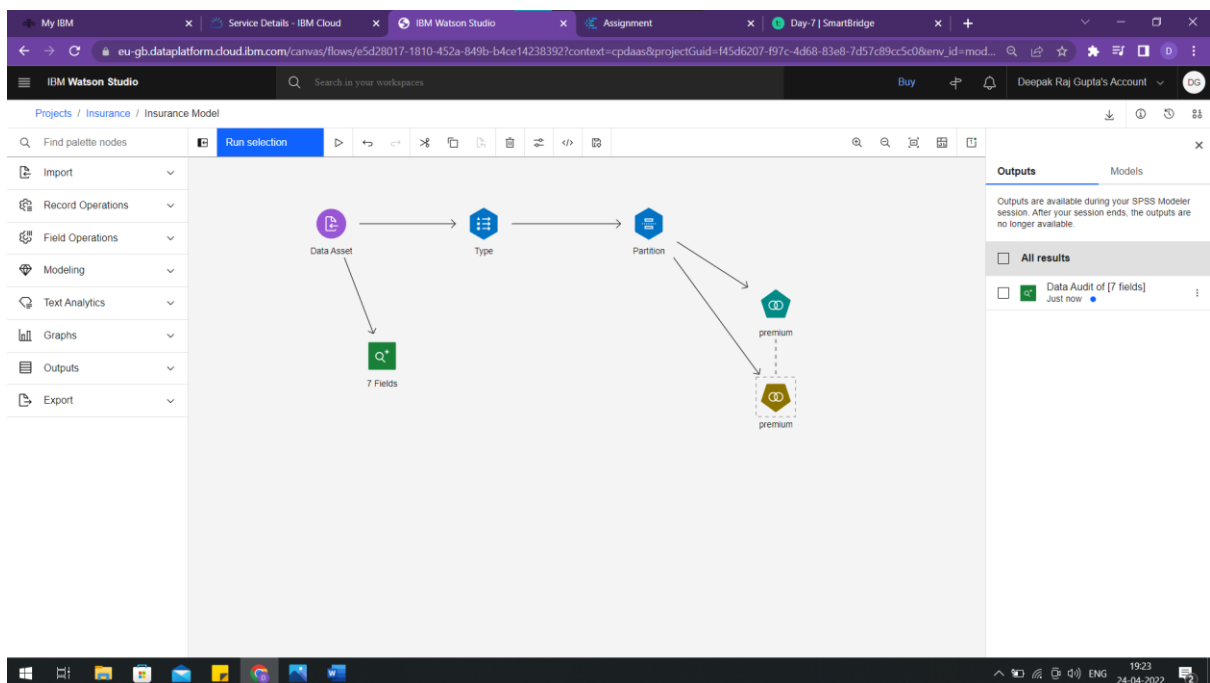
The screenshot displays the IBM Watson Studio interface for refining data. The main window shows a table with columns: age, sex, bmi, children, smoker, and region. The data is sourced from 'insurance.csv' and contains 1338 rows. The interface includes a 'Steps' panel on the left with a step titled '1. Convert column type' and an 'Auto-generated' button. The right panel shows the 'Information' tab with details about the data refinery flow, including the location 'Insurance' and the flow name 'insurance.csv\_flow'.

	age	sex	bmi	children	smoker	region
1	19	female	27.9	0	yes	southwest
2	18	male	33.77	1	no	southeast
3	28	male	33	3	no	southeast
4	33	male	22.705	0	no	northwest
5	32	male	28.88	0	no	northwest
6	31	female	25.74	0	no	southeast
7	46	female	33.44	1	no	southeast
8	37	female	27.74	3	no	northwest
9	37	male	29.83	2	no	northeast
10	60	female	25.84	0	no	northwest
11	25	male	26.22	0	no	northeast
12	62	female	26.29	0	yes	southeast
13	23	male	34.4	0	no	southwest
14	56	female	39.82	0	no	southeast
15	27	male	42.13	0	yes	southeast
16	19	male	24.6	1	no	southwest

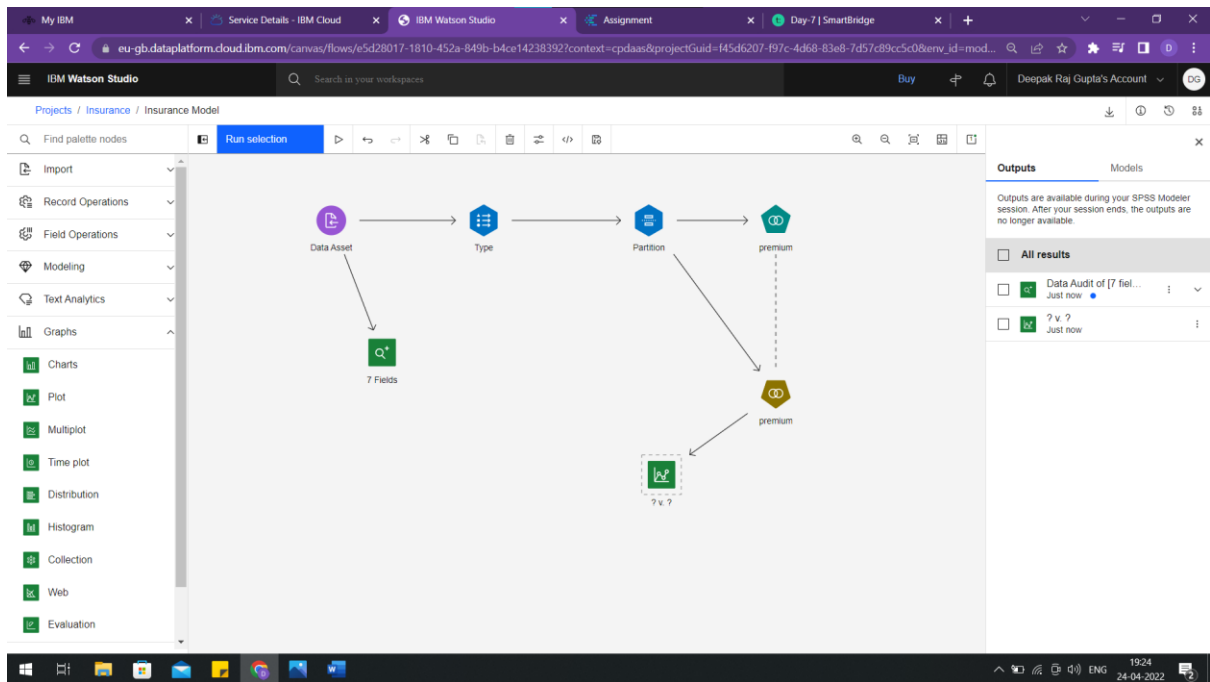
Fig: Refining Data



*Fig: Importing data into model*



*Fig: Creating clustering Model*



*Fig: plotting values on graph*

The screenshot displays the 'View Model: premium' dialog box in IBM Watson Studio. The dialog shows a table of models for the 'premium' variable. The table has columns: USE, MODEL NAME, ESTIMATOR, GRAPH, SILHOUETTE, BUILD TIME (MINS), NUMBER OF CLUSTERS, SMALLEST CLUSTER (N), SMALLEST CLUSTER (%), LARGEST CLUSTER (N), LARGEST CLUSTER (%), SMALLEST/LARGEST, and IMPORTANCE. Three models are listed: TwoStep 1, Kohonen 1, and K-means 1.

USE	MODEL NAME	ESTIMATOR	GRAPH	SILHOUETTE	BUILD TIME (MINS)	NUMBER OF CLUSTERS	SMALLEST CLUSTER (N)	SMALLEST CLUSTER (%)	LARGEST CLUSTER (N)	LARGEST CLUSTER (%)	SMALLEST/LARGEST	IMPORTANCE
•	TwoStep 1	TwoStep		0.705	< 1	2	374	0.402	557	0.598	0.671	0.671
○	Kohonen 1	Kohonen		0.561	< 1	12	34	0.037	221	0.237	0.154	0.154
○	K-means 1	KMeans		0.643	< 1	5	148	0.159	214	0.230	0.692	0.692

*Fig: Model View*

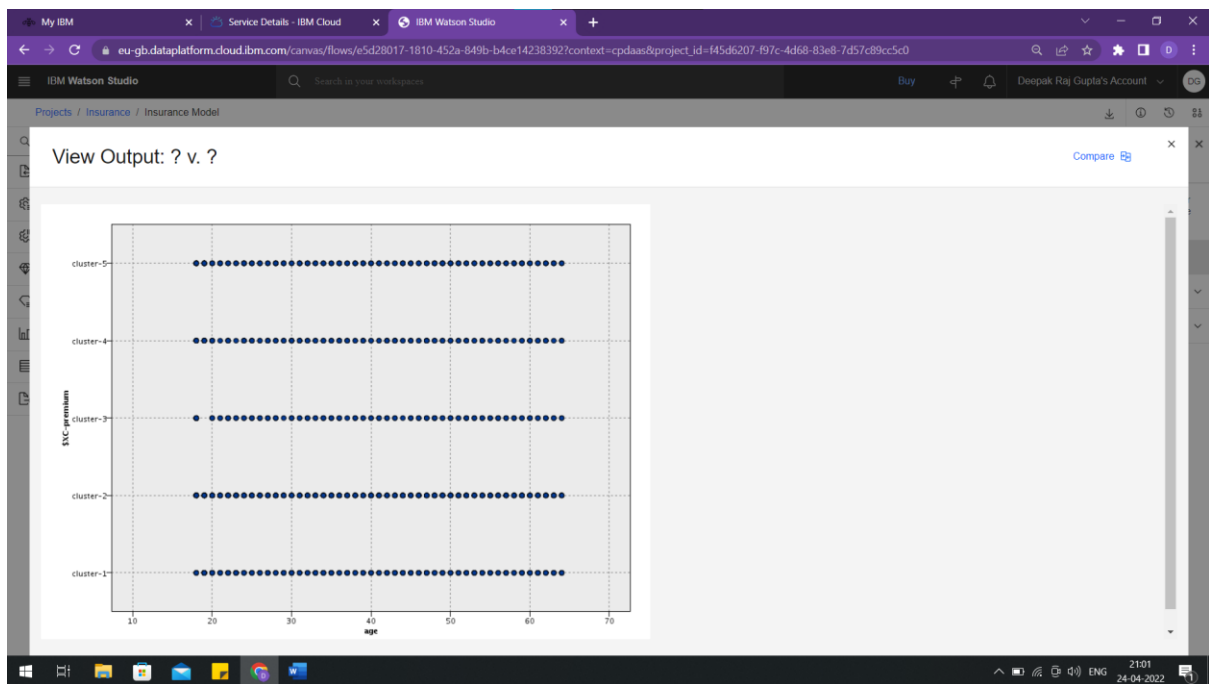


Fig: Data Visualization in graph

## 2) Diabetes

The figure shows the IBM Watson Studio interface for a data refining step. The main table displays the first 50 rows of the 'diabetes.csv' dataset. The table has columns for Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigree, and Age. The sidebar on the right shows details for the 'diabetes.csv\_flow' data refining step, including its location and output.

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigree	Age
1	6	148	72	35	0	33.6	0.627	50
2	1	85	66	29	0	26.6	0.351	31
3	8	183	64	0	0	23.3	0.672	32
4	1	89	66	23	94	28.1	0.167	21
5	0	137	40	35	168	43.1	2.288	33
6	5	116	74	0	0	25.6	0.201	30
7	3	78	50	32	88	31	0.248	26
8	10	115	0	0	0	35.3	0.134	29
9	2	197	70	45	543	30.5	0.158	53
10	8	125	96	0	0	0	0.232	54
11	4	110	92	0	0	37.6	0.191	30
12	10	168	74	0	0	38	0.537	34
13	10	139	80	0	0	27.1	1.441	57
14	1	189	60	23	846	30.1	0.398	59
15	5	166	72	19	175	25.8	0.587	51
16	7	100	0	0	0	30	0.484	32

Fig: Data Refining

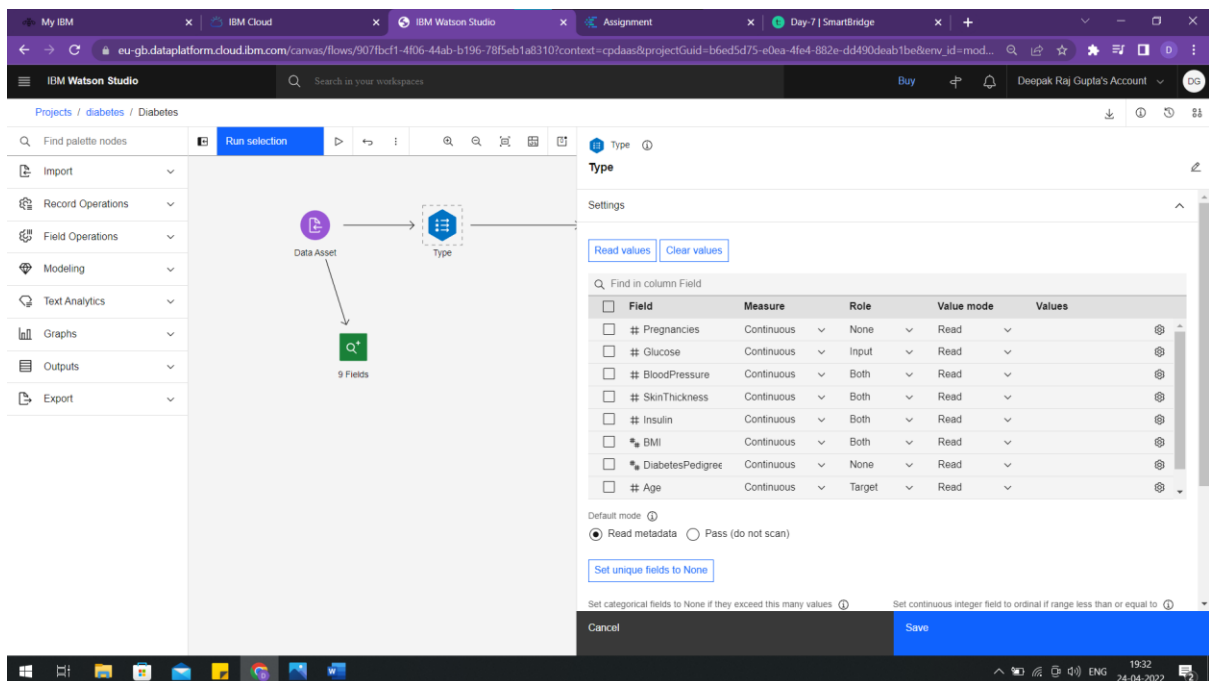


Fig: Defining type of data

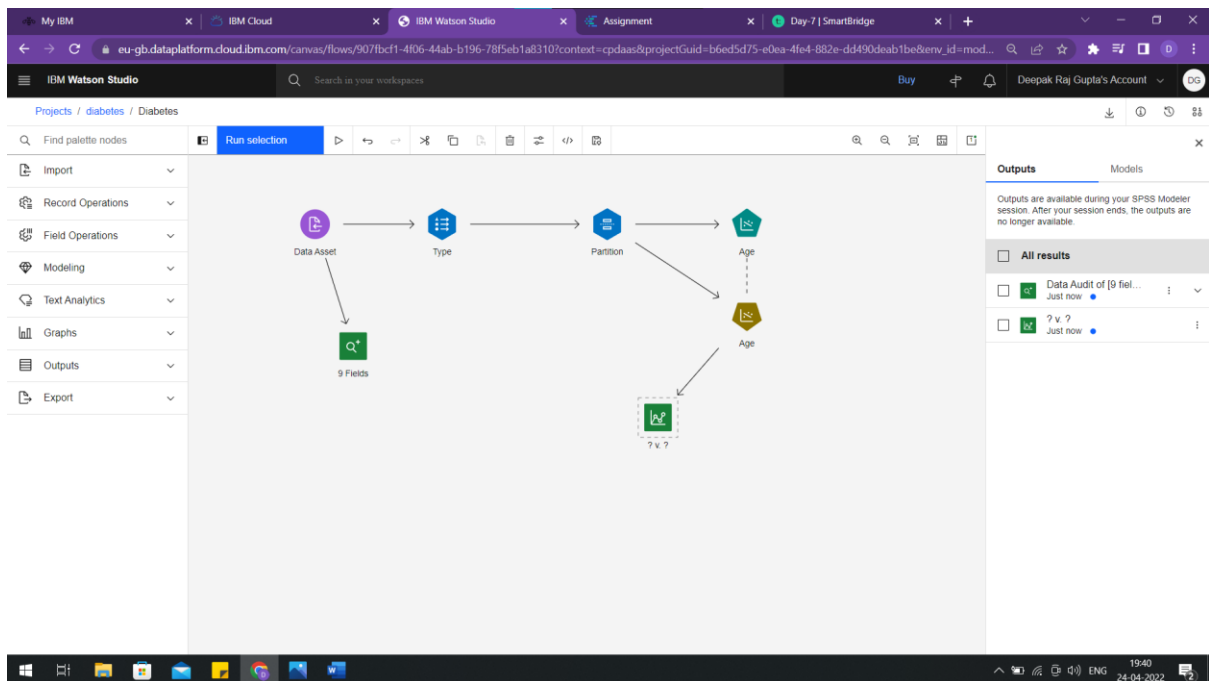
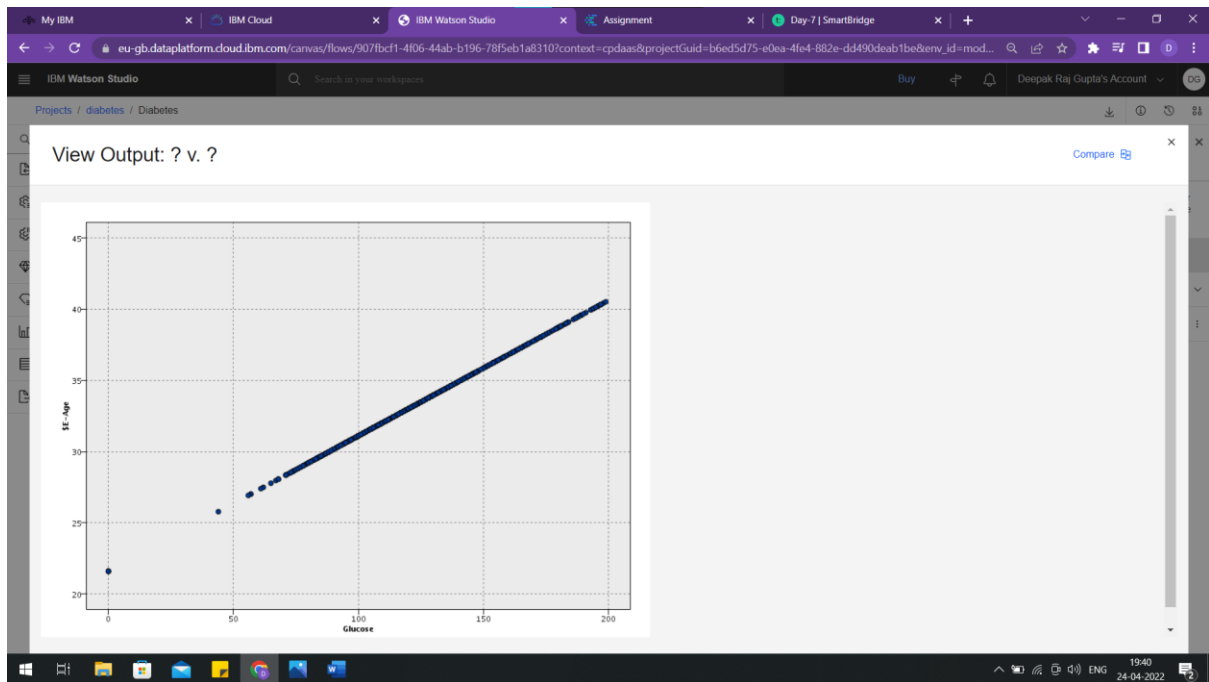
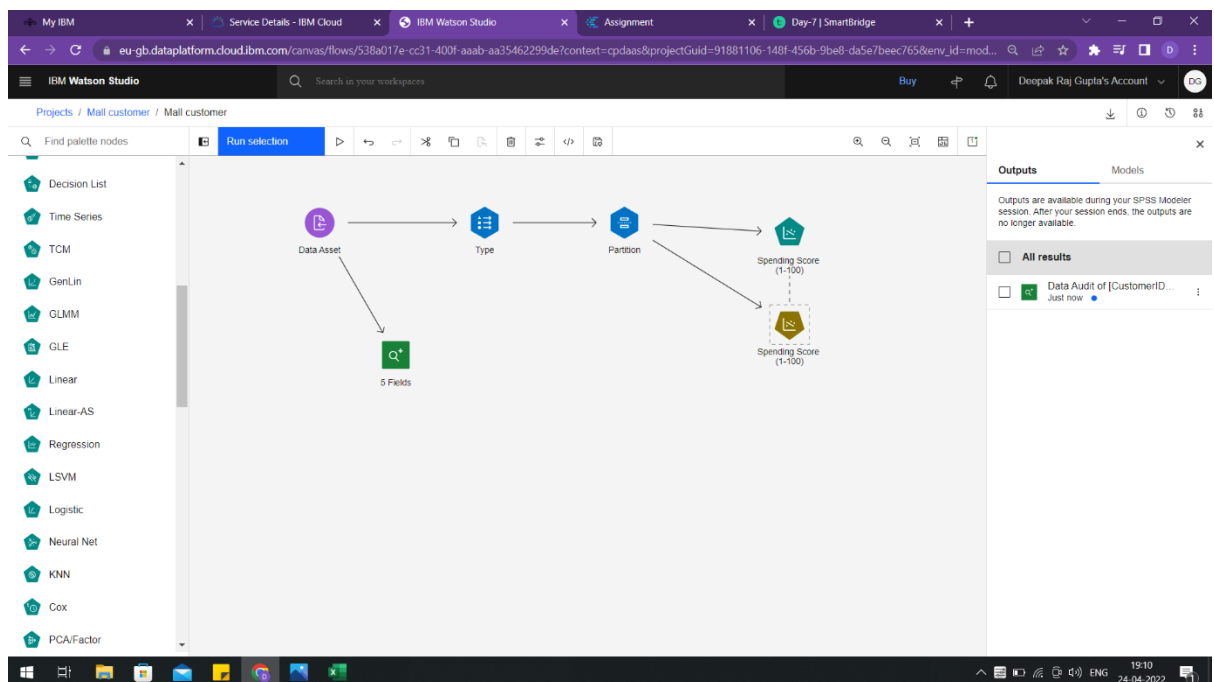


Fig: Design of Regression Model



*Fig: Data Visualization in graph*

### 3) Mall Customer



*Fig: Regression Model Creation*

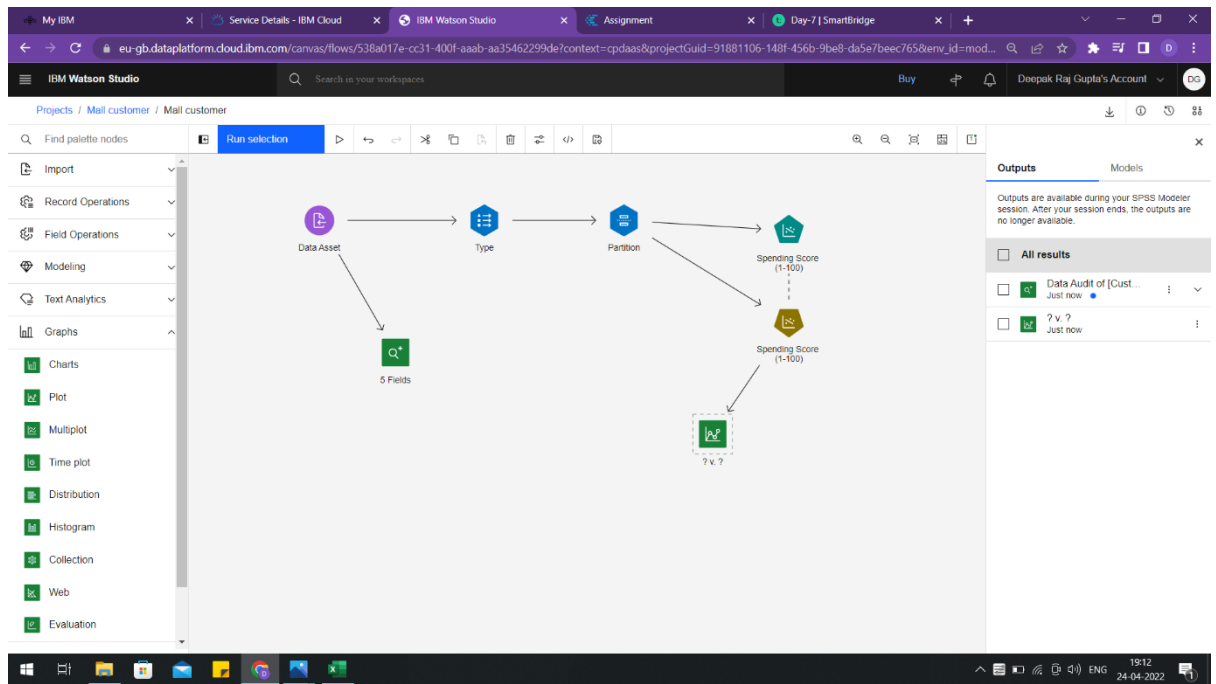


Fig: Executing model

View Model: Spending Score (1-100)

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

MODEL VIEWER

Build Settings

Training Summary

Coefficients

		1	
		(Constant)	Annual Income (k\$)
Unstandardized Coefficients	B	49.478	0.036
	Std. Error	5.684	0.089
Standardized Coefficients	Beta		0.035
t		8.705	0.400
Sig.		0.000	0.690
Fraction Missing Info.			
Relative Increase Variance			
Relative Efficiency			

My IBM Service Details - IBM Cloud IBM Watson Studio Assignment Day-7 | SmartBridge

eu-gb.dataplatform.cloud.ibm.com/canvas/flows/538a017e-cc31-400f-aaab-aa35462299de?context=cpdaas&projectGuid=91881106-148f-456b-9be8-da5e7bec765&env\_id=mod...

IBM Watson Studio Search in your workspaces Buy Deepak Raj Gupta's Account

Projects / Mail customer / Mail customer

### View Model: Spending Score (1-100)

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

My IBM Service Details - IBM Cloud IBM Watson Studio Assignment Day-7 | SmartBridge

eu-gb.dataplatform.cloud.ibm.com/canvas/flows/538a017e-cc31-400f-aaab-aa35462299de?context=cpdaas&projectGuid=91881106-148f-456b-9be8-da5e7bec765&env\_id=mod...

IBM Watson Studio Search in your workspaces Buy Deepak Raj Gupta's Account

Projects / Mail customer / Mail customer

### View Model: Spending Score (1-100)

Regression

EVALUATION

Model Summary

ANOVA

Coefficients

MODEL VIEWER

**Build Settings**

**Training Summary**

#### Training Summary

Algorithm	Regression
Model type	Approximation
Date built	Sun Apr 24 13:39:45 UTC 2022
Elapsed time for model build	0 hours, 0 mins, 3 secs



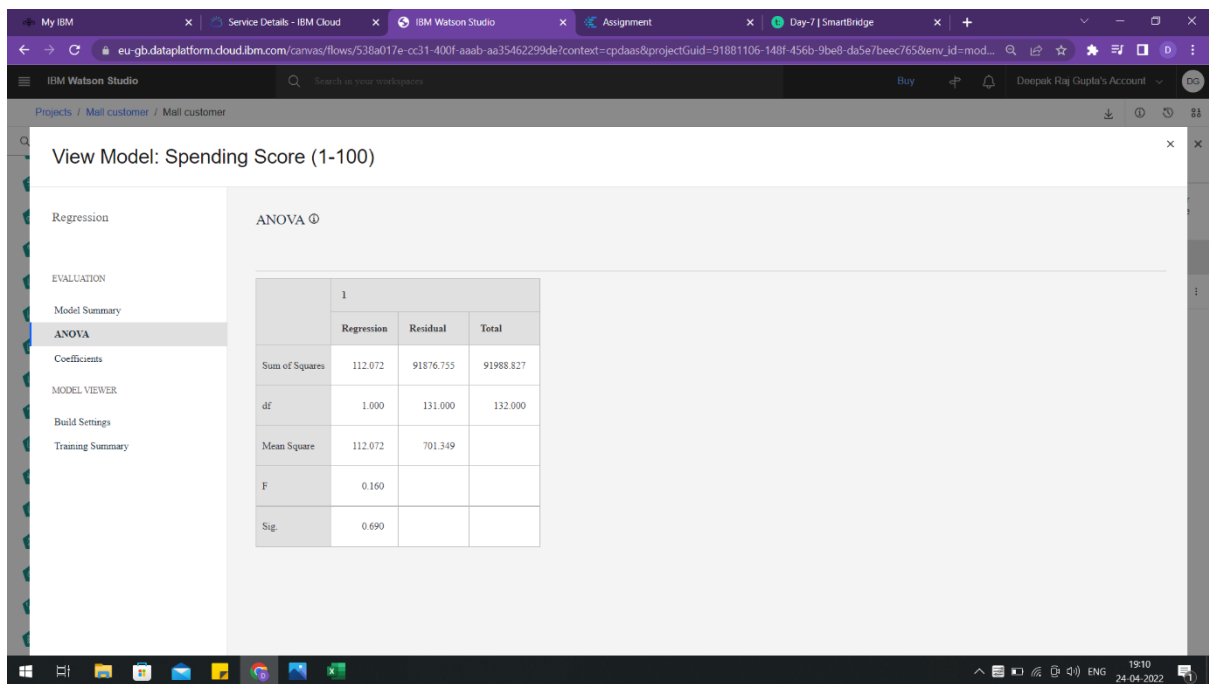
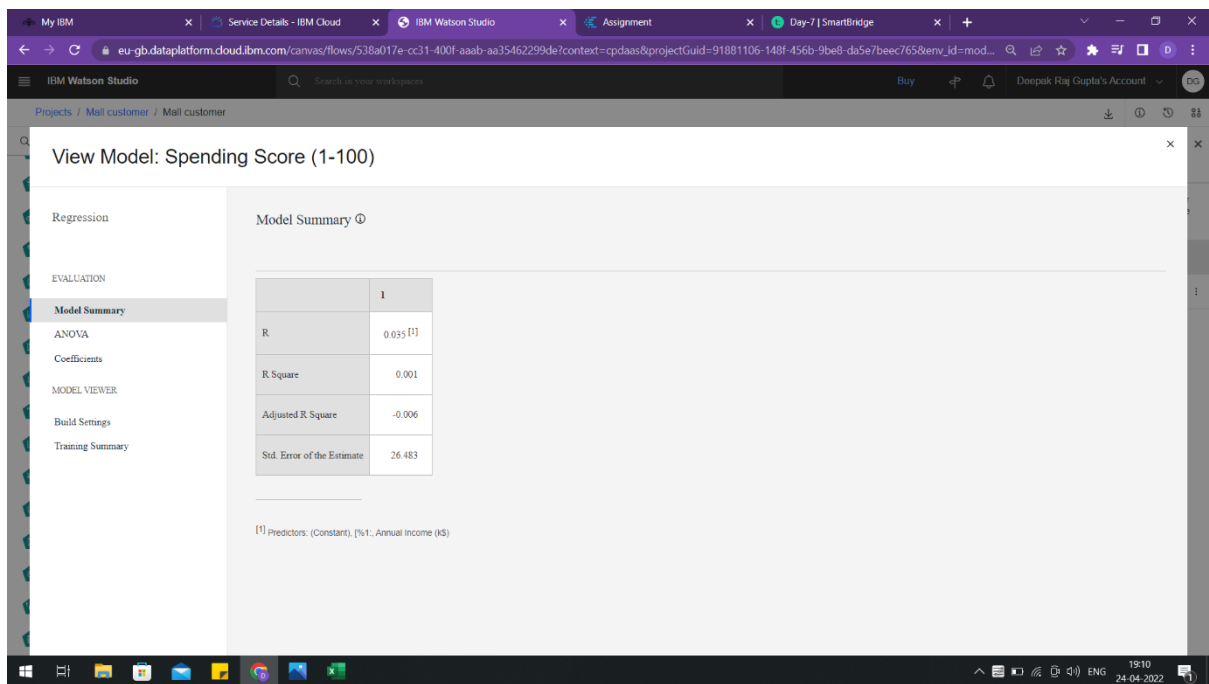
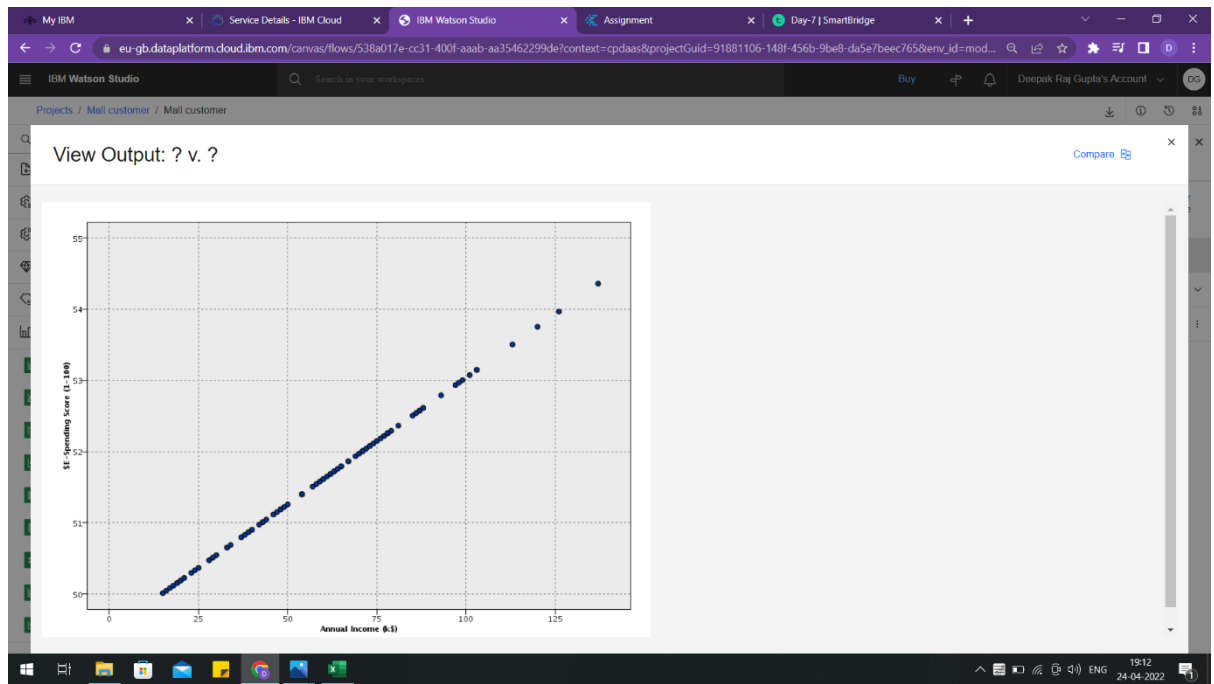


Fig: Detailed Information of the model



*Fig: Data Visualization in graph*