Pranjal Gupta

# Assignment 03

**Data set: Diabetes.csv**

**Link: https://dataplatform.cloud.ibm.com/projects/047cfd7a-4874-4934-a9d1-a9f5d8071774/assets?context=cpdaas**

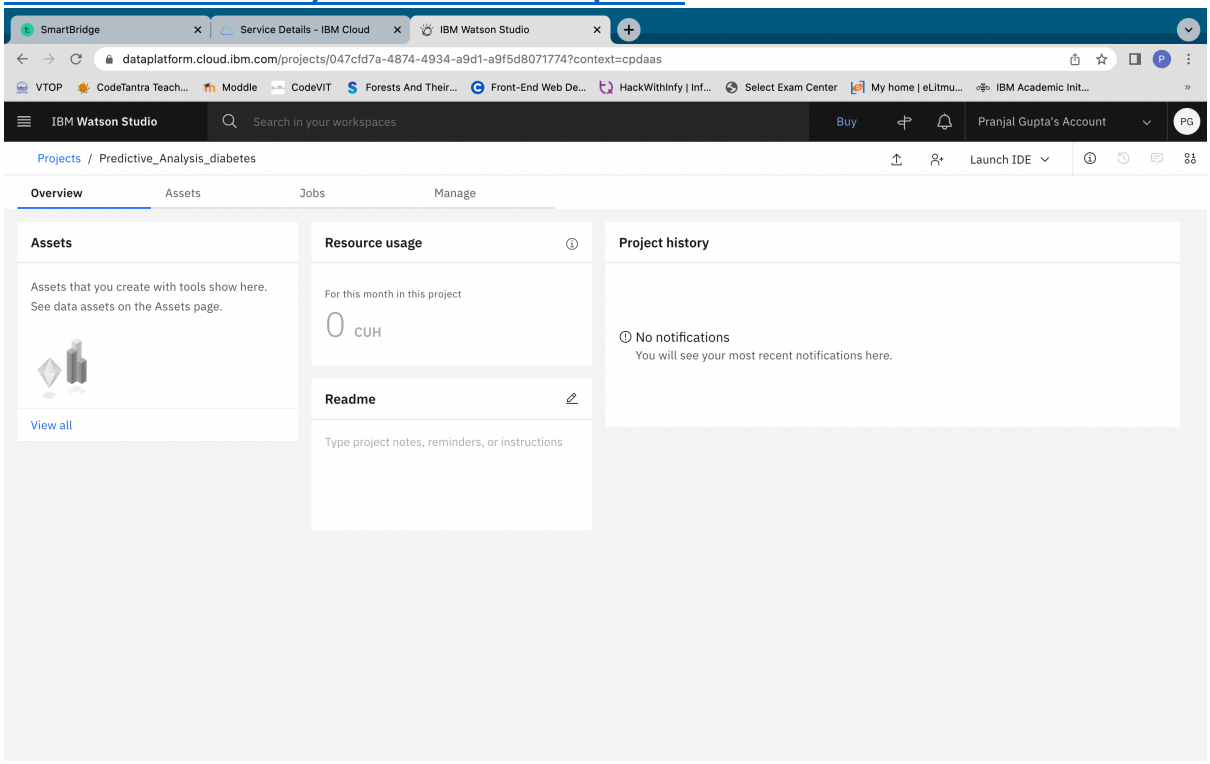



Figure: Created the project in Watson Studio

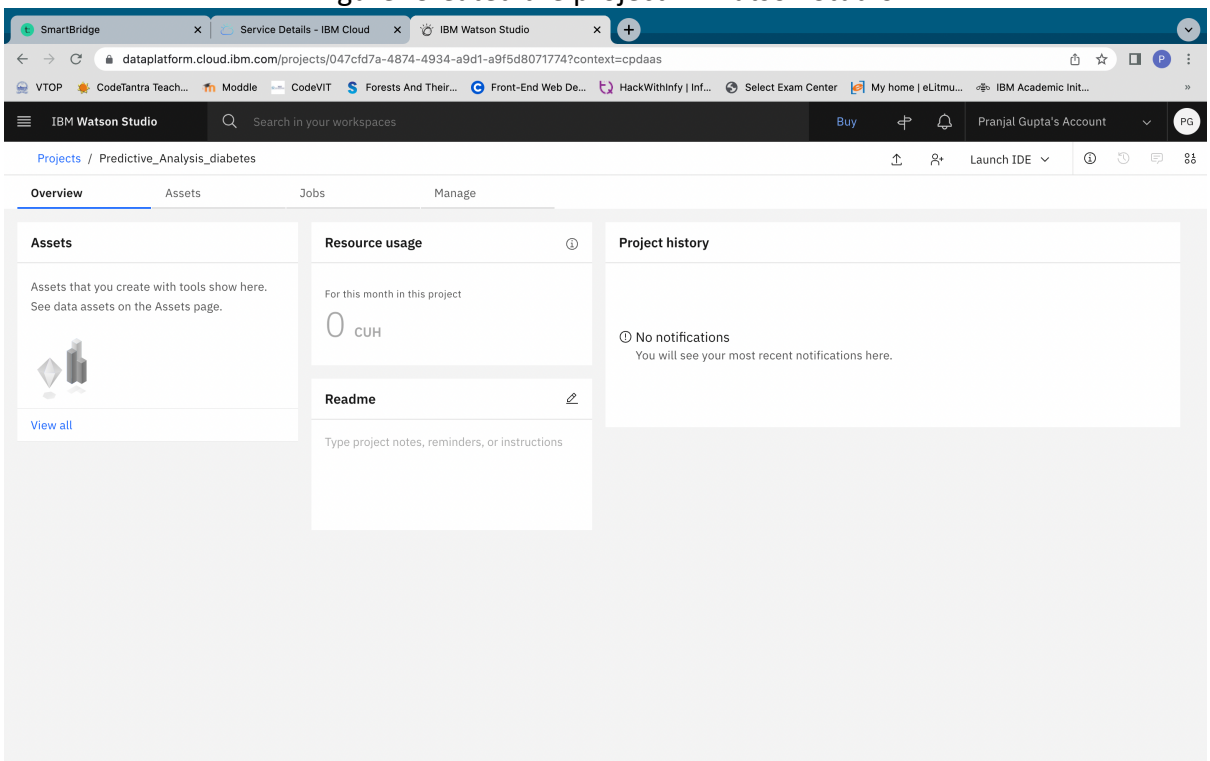
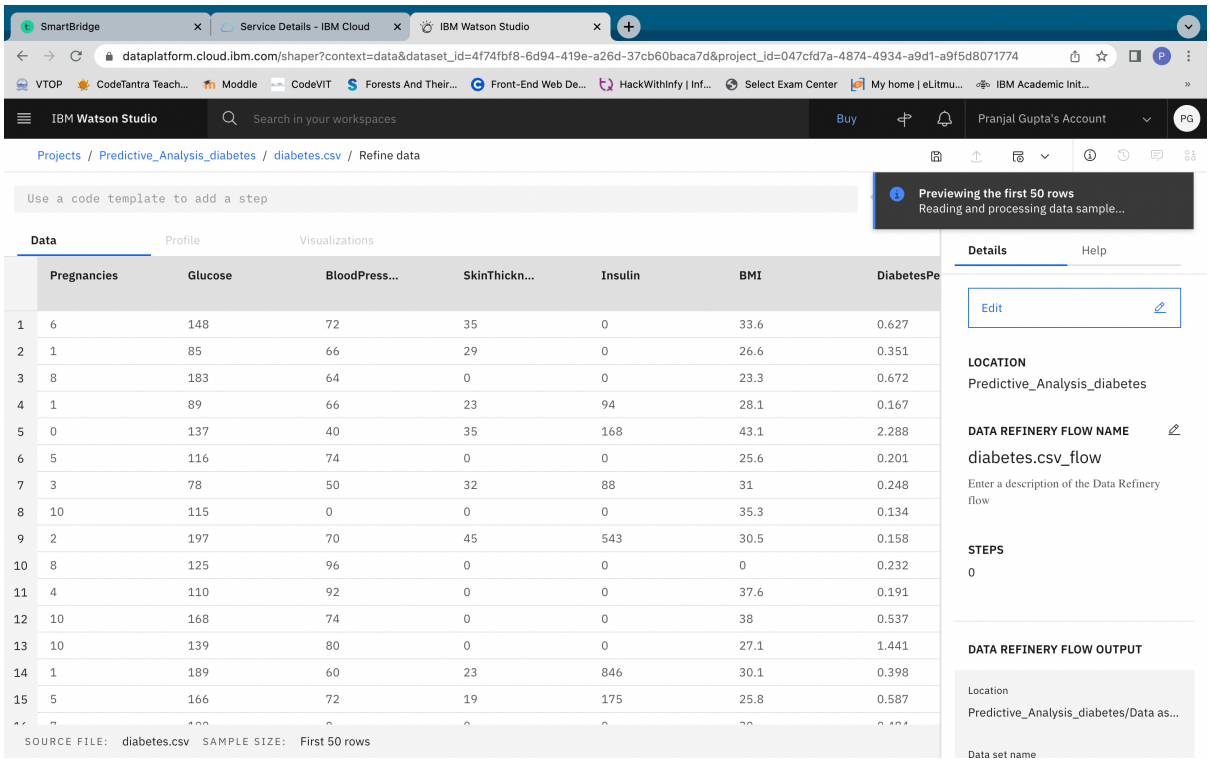



Figure: Adding the dataset

Pranjal Gupta




| | Pregnancies | Glucose | BloodPress... | SkinThickn... | Insulin | BMI | DiabetesPe |
|---|---|---|---|---|---|---|---|
| 1 | 6 | 148 | 72 | 35 | 0 | 33.6 | 0.627 |
| 2 | 1 | 85 | 66 | 29 | 0 | 26.6 | 0.351 |
| 3 | 8 | 183 | 64 | 0 | 0 | 23.3 | 0.672 |
| 4 | 1 | 89 | 66 | 23 | 94 | 28.1 | 0.167 |
| 5 | 0 | 137 | 40 | 35 | 168 | 43.1 | 2.288 |
| 6 | 5 | 116 | 74 | 0 | 0 | 25.6 | 0.201 |
| 7 | 3 | 78 | 50 | 32 | 88 | 31 | 0.248 |
| 8 | 10 | 115 | 0 | 0 | 0 | 35.3 | 0.134 |
| 9 | 2 | 197 | 70 | 45 | 543 | 30.5 | 0.158 |
| 10 | 8 | 125 | 96 | 0 | 0 | 0 | 0.232 |
| 11 | 4 | 110 | 92 | 0 | 0 | 37.6 | 0.191 |
| 12 | 10 | 168 | 74 | 0 | 0 | 38 | 0.537 |
| 13 | 10 | 139 | 80 | 0 | 0 | 27.1 | 1.441 |
| 14 | 1 | 189 | 60 | 23 | 846 | 30.1 | 0.398 |
| 15 | 5 | 166 | 72 | 19 | 175 | 25.8 | 0.587 |

Figure: Created a data refinery asset

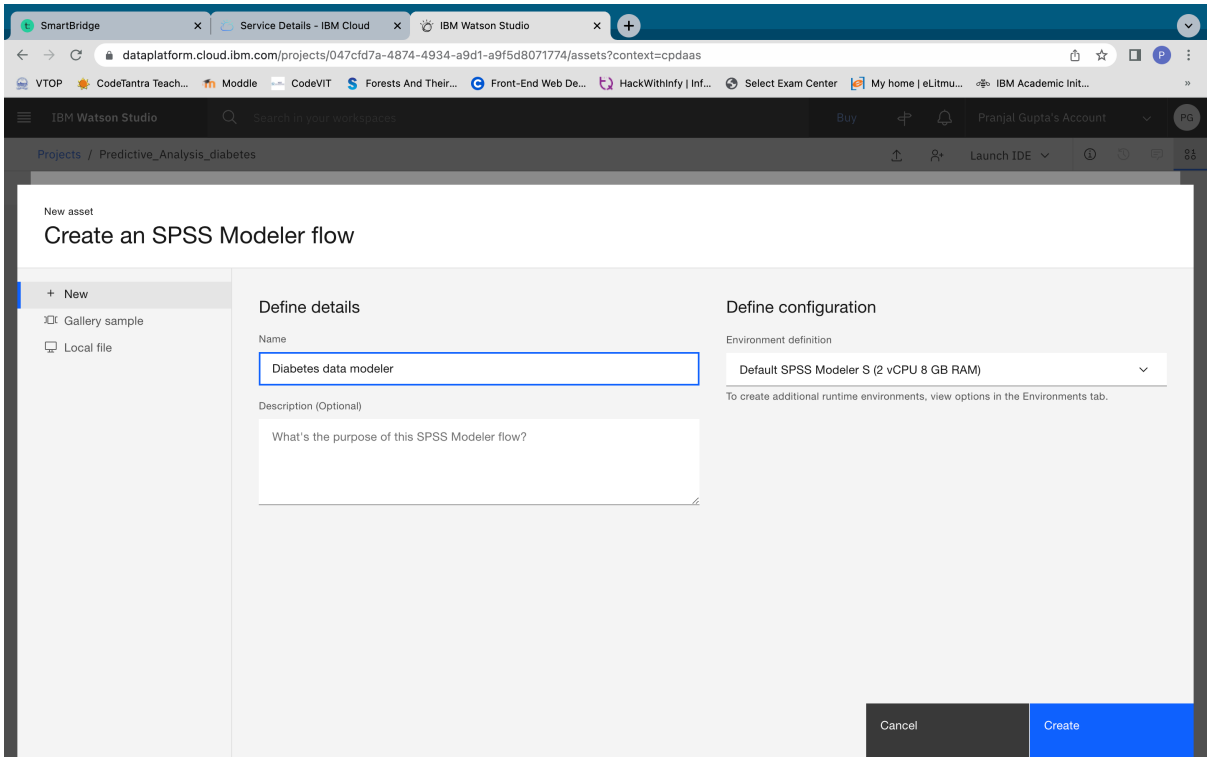



Figure: Creating SPSS modeler

Pranjal Gupta

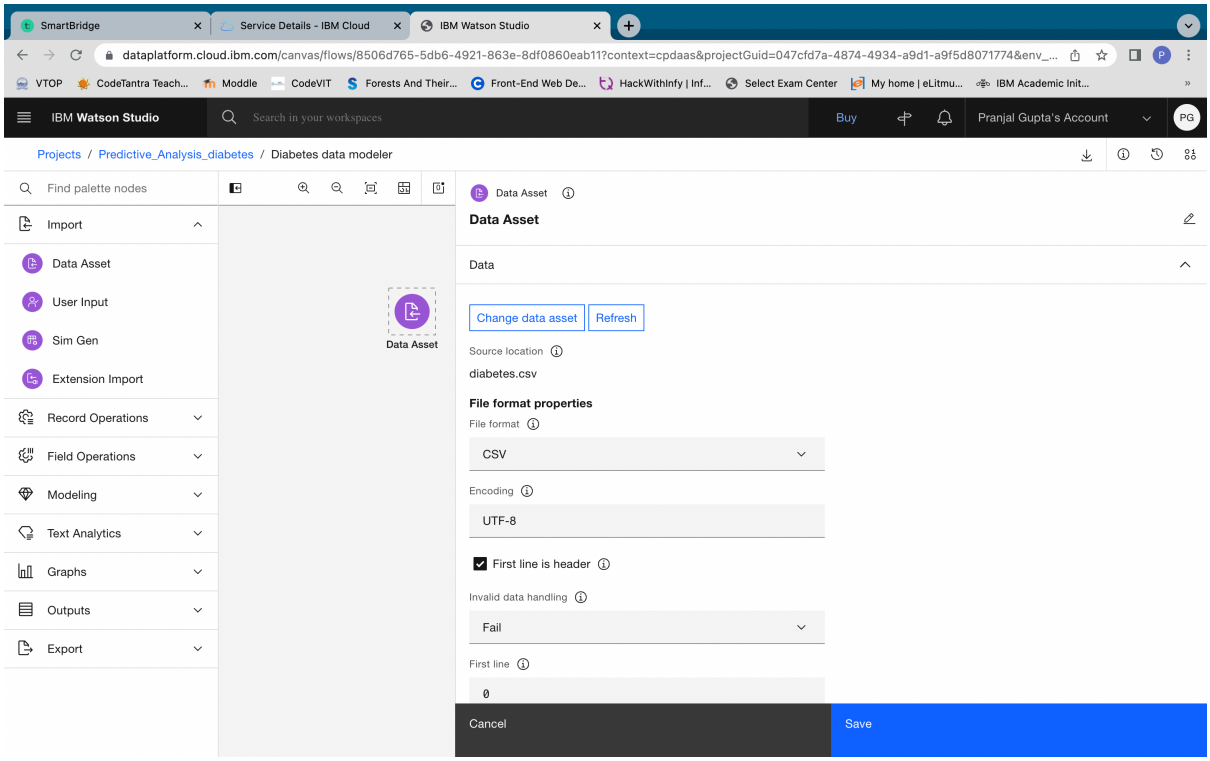



Figure: Imported Data asset

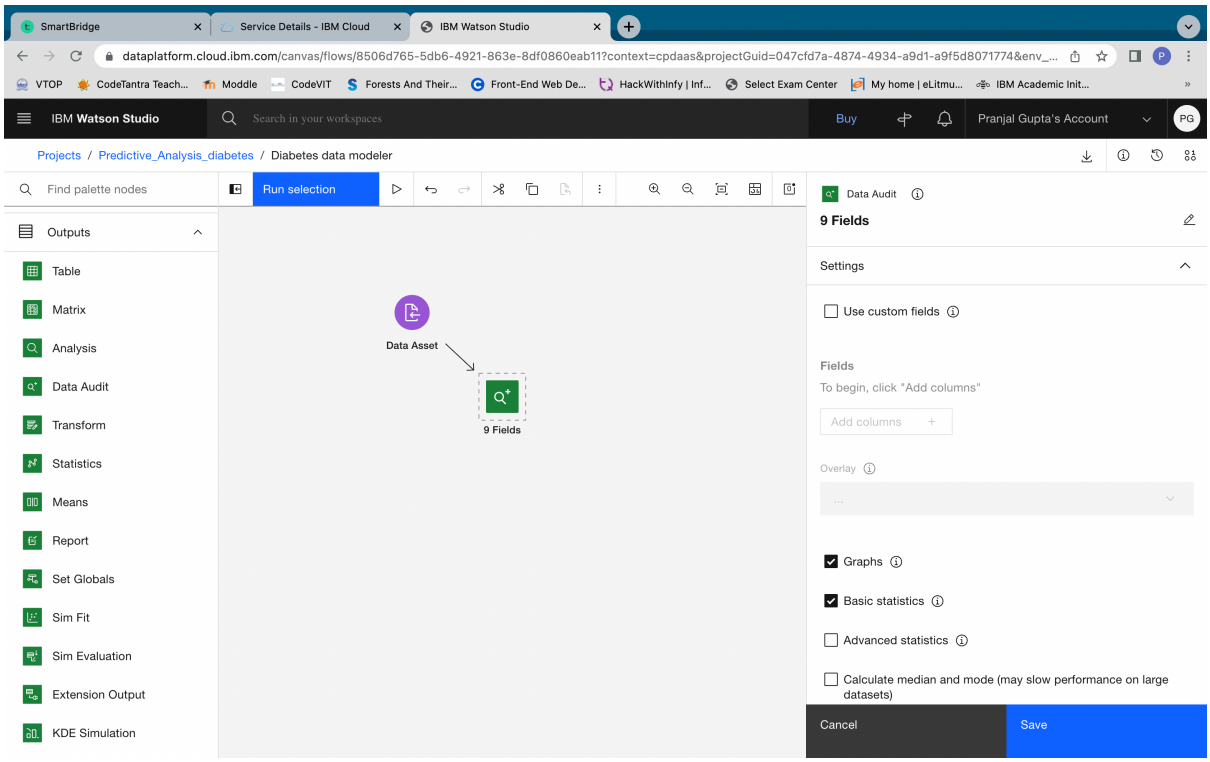



Figure: Performed Data audit

Pranjal Gupta

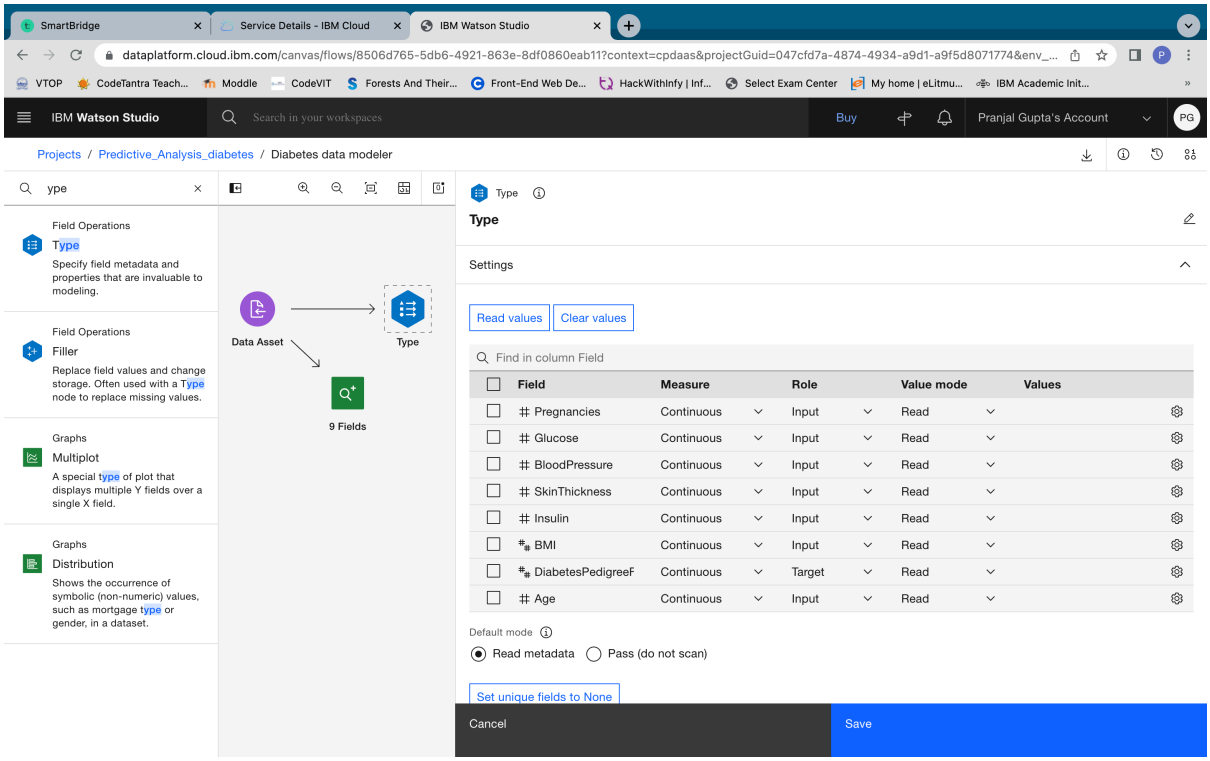



Figure: Type of attributes specified

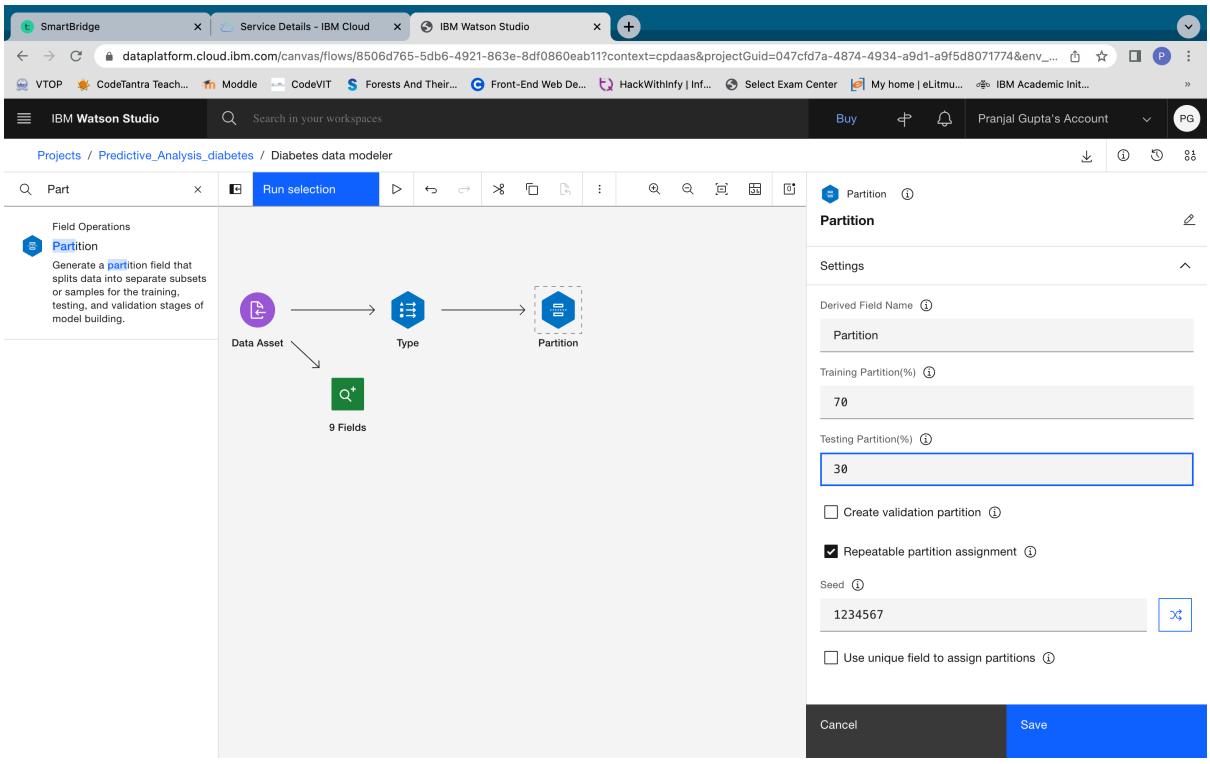



Figure: Performing data partitioning

Pranjal Gupta

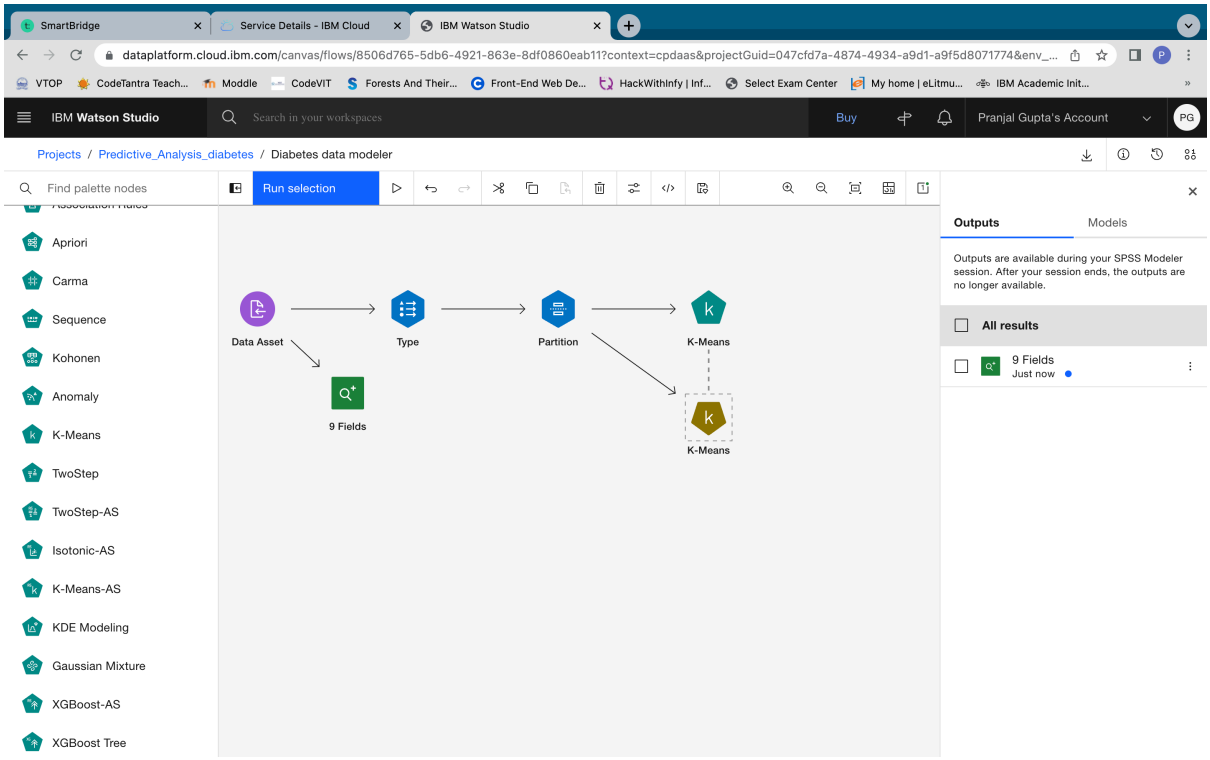



Figure: Performed K-Means Clustering

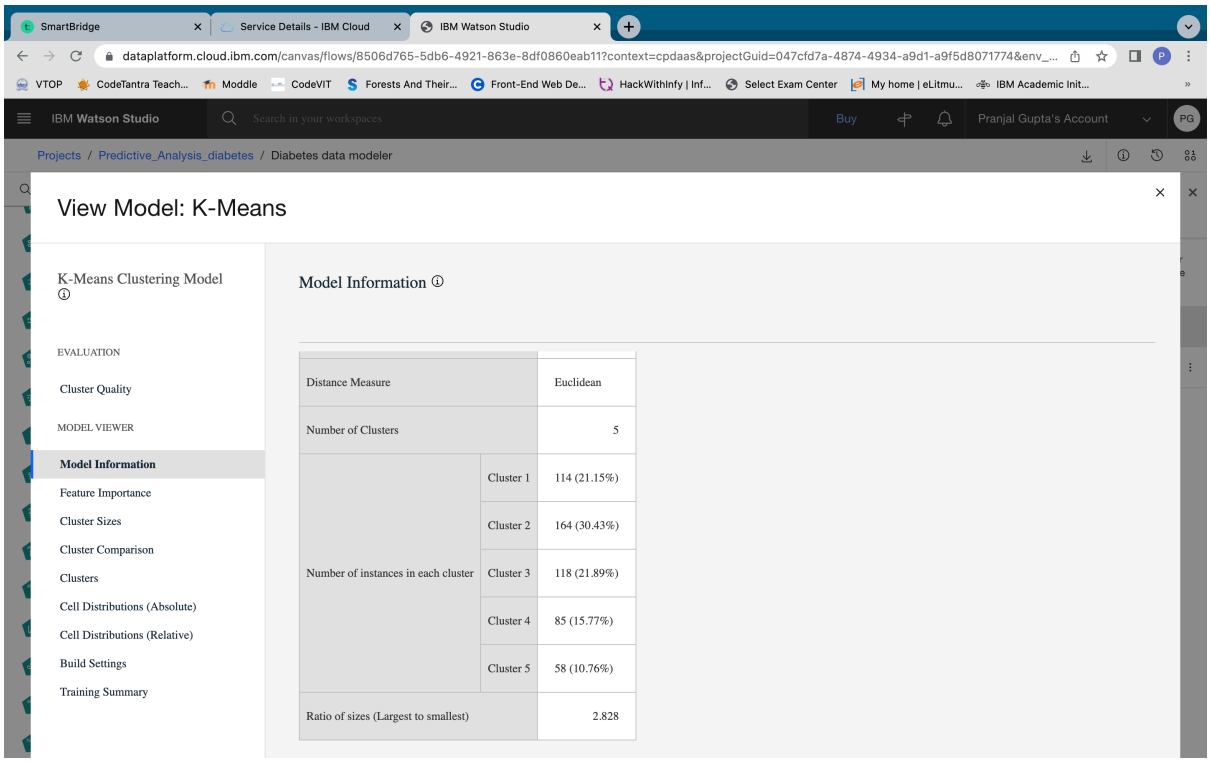



Figure: Model Information

Pranjal Gupta

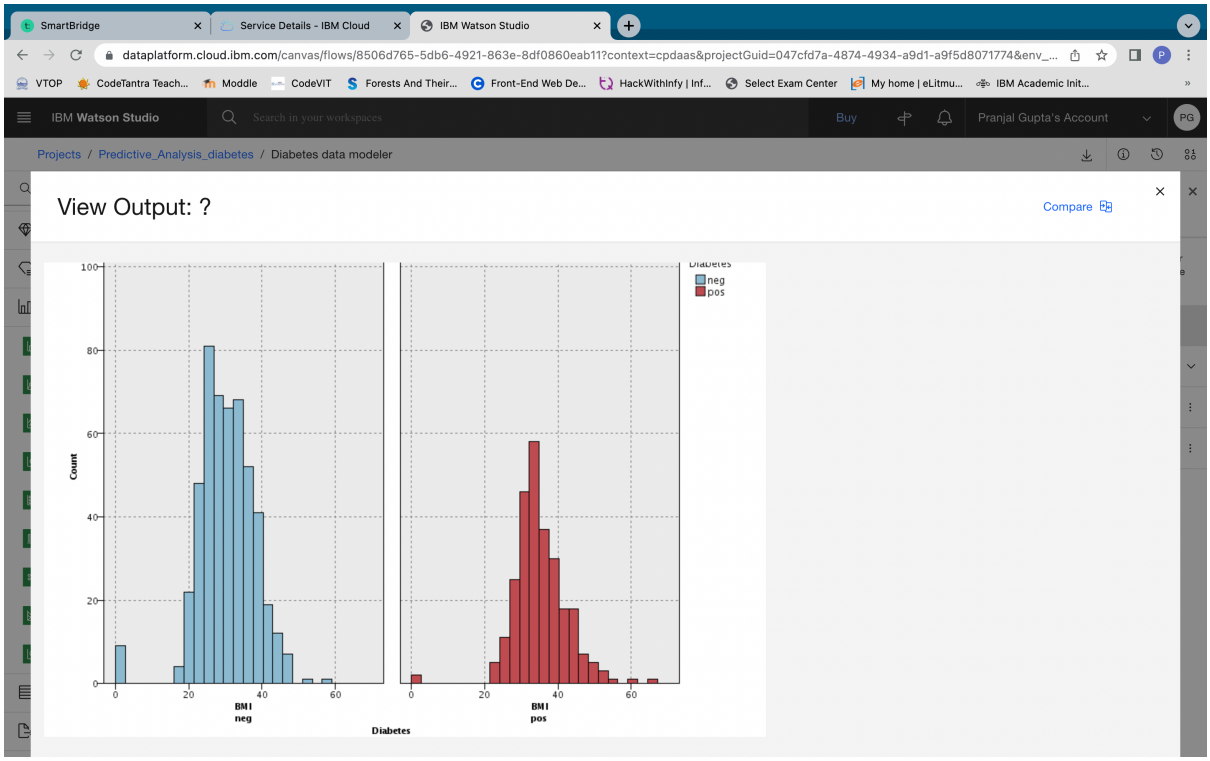



Figure: Histogram analysis

Pranjal Gupta

# Data Set: Insurance.csv

**Link:** https://dataplatform.cloud.ibm.com/projects/dc4c292b-f4c9-47bd-a802-c854178b2449/assets?context=cpdaas

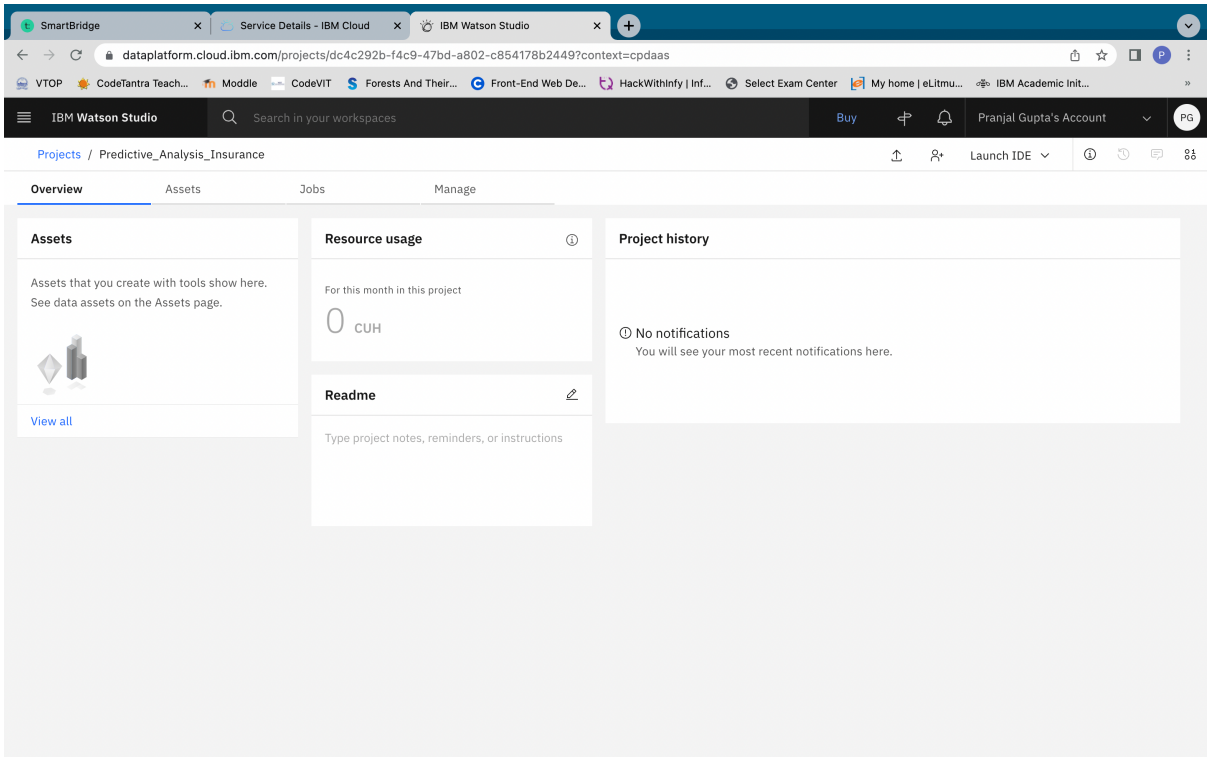



Figure: Created project for Insurance data set

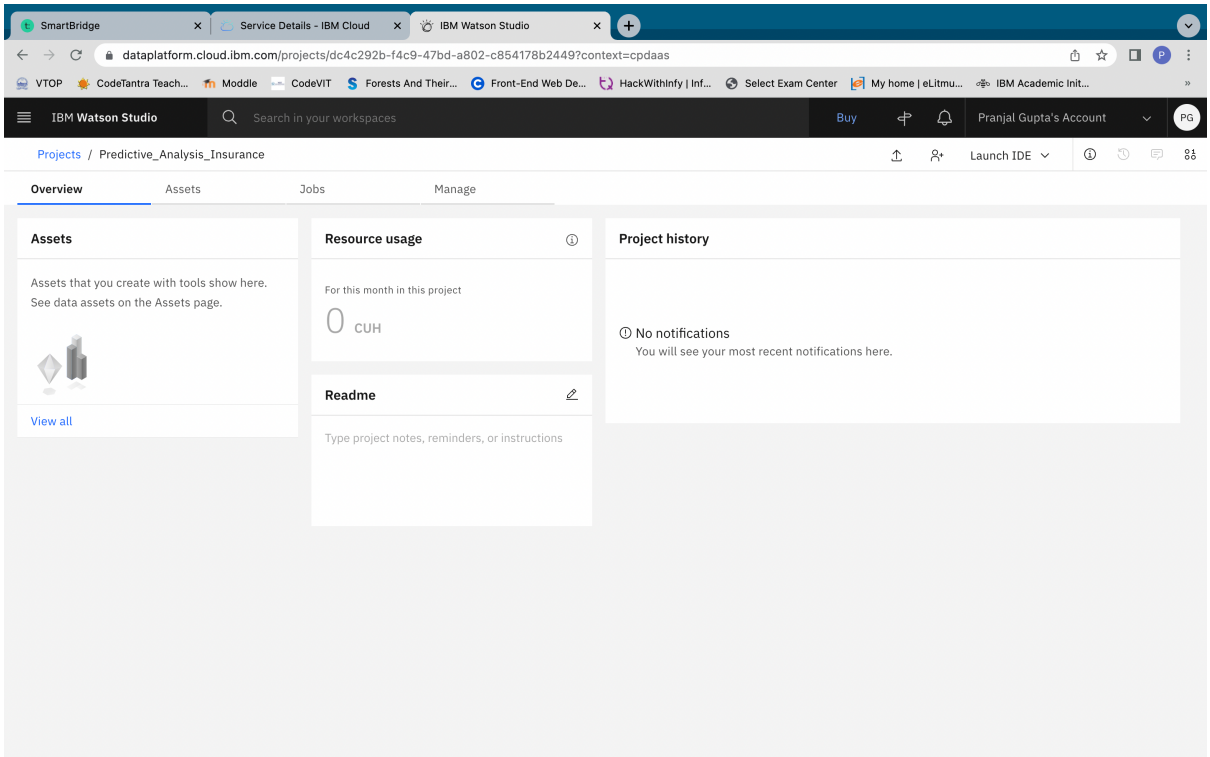



Figure: Adding insurance data set

Pranjal Gupta

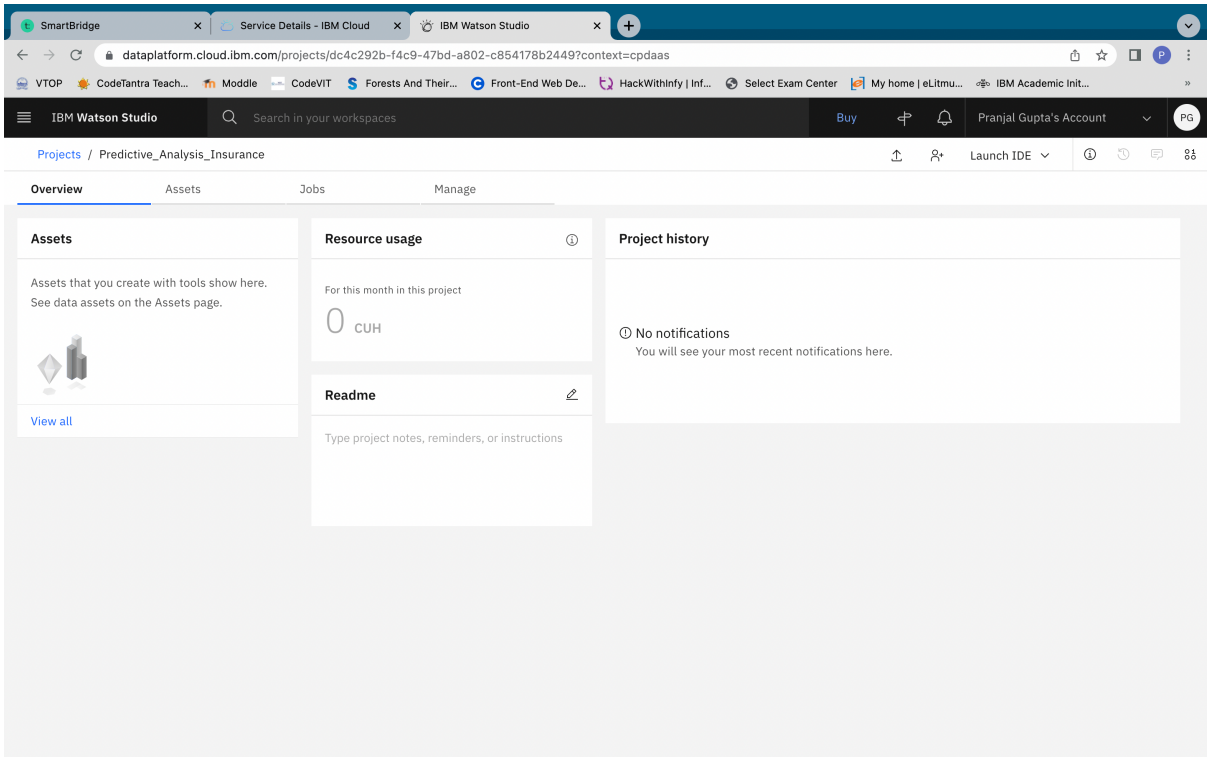



Figure: Created a data refinery flow

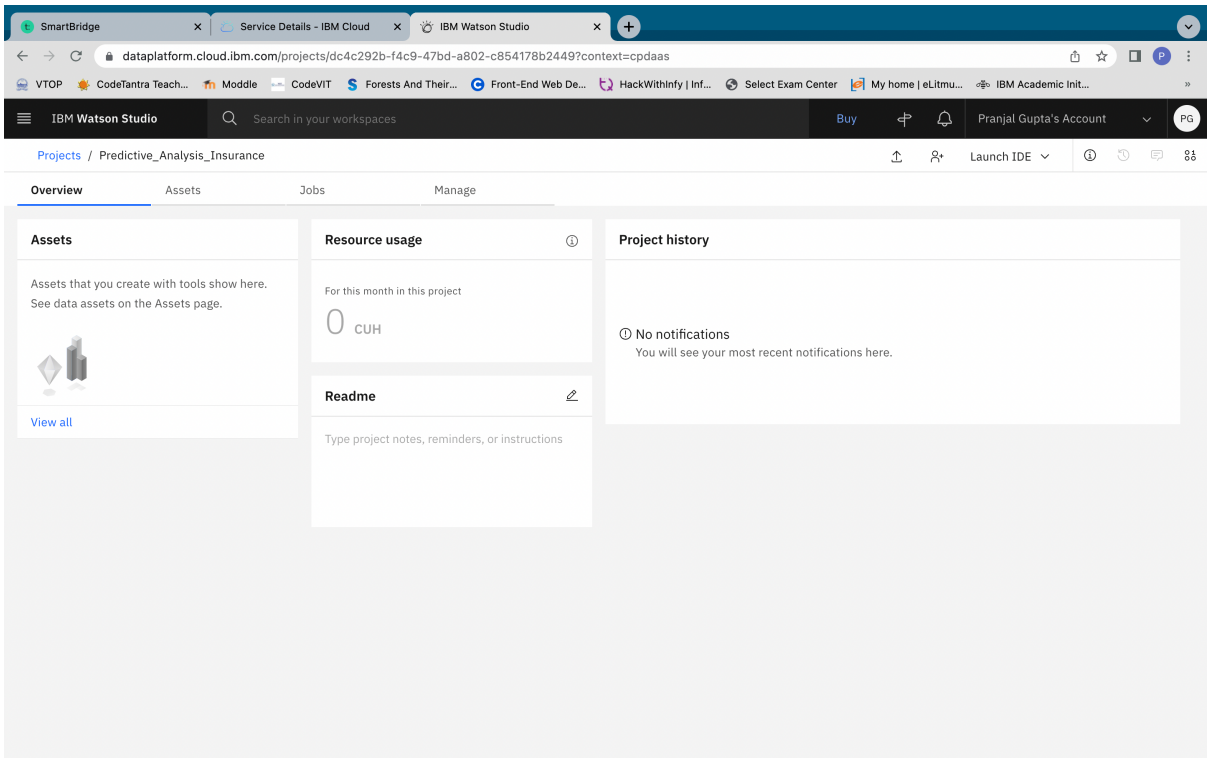



Figure: Creating a SPSS modeler for insurance data set

Pranjal Gupta

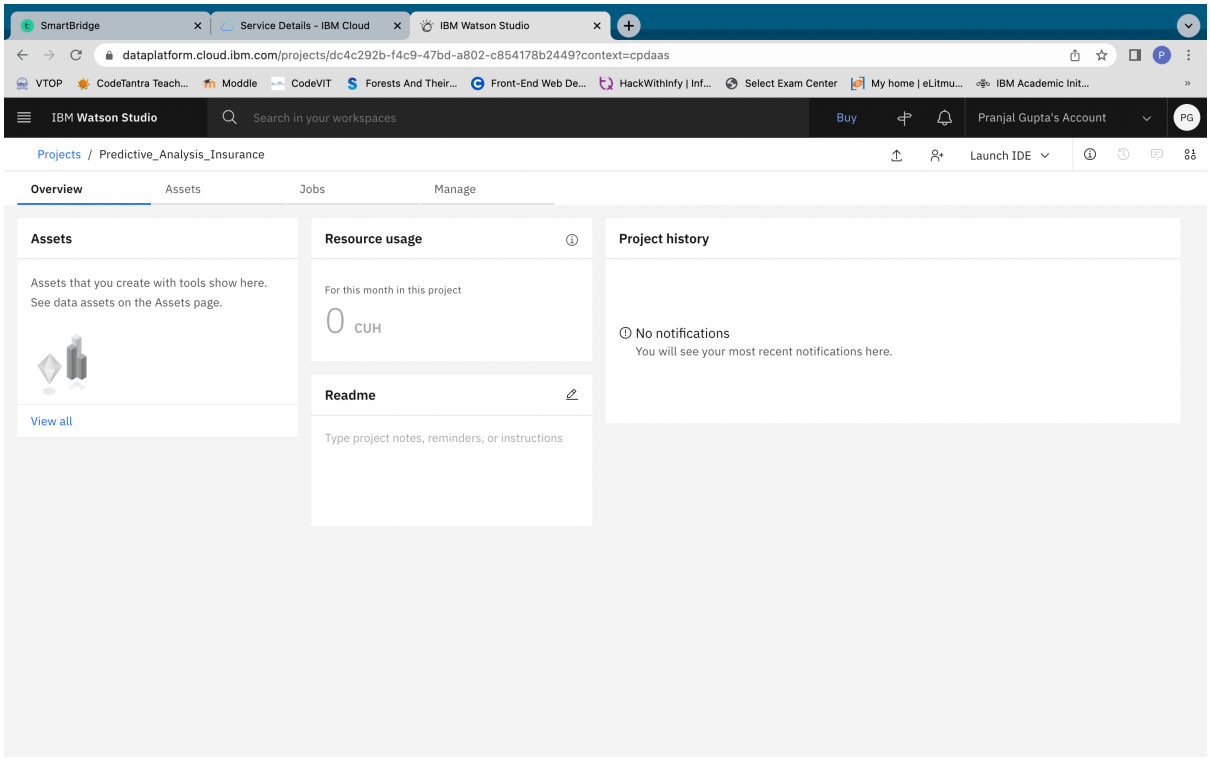



Figure: Importing data to the data asset node

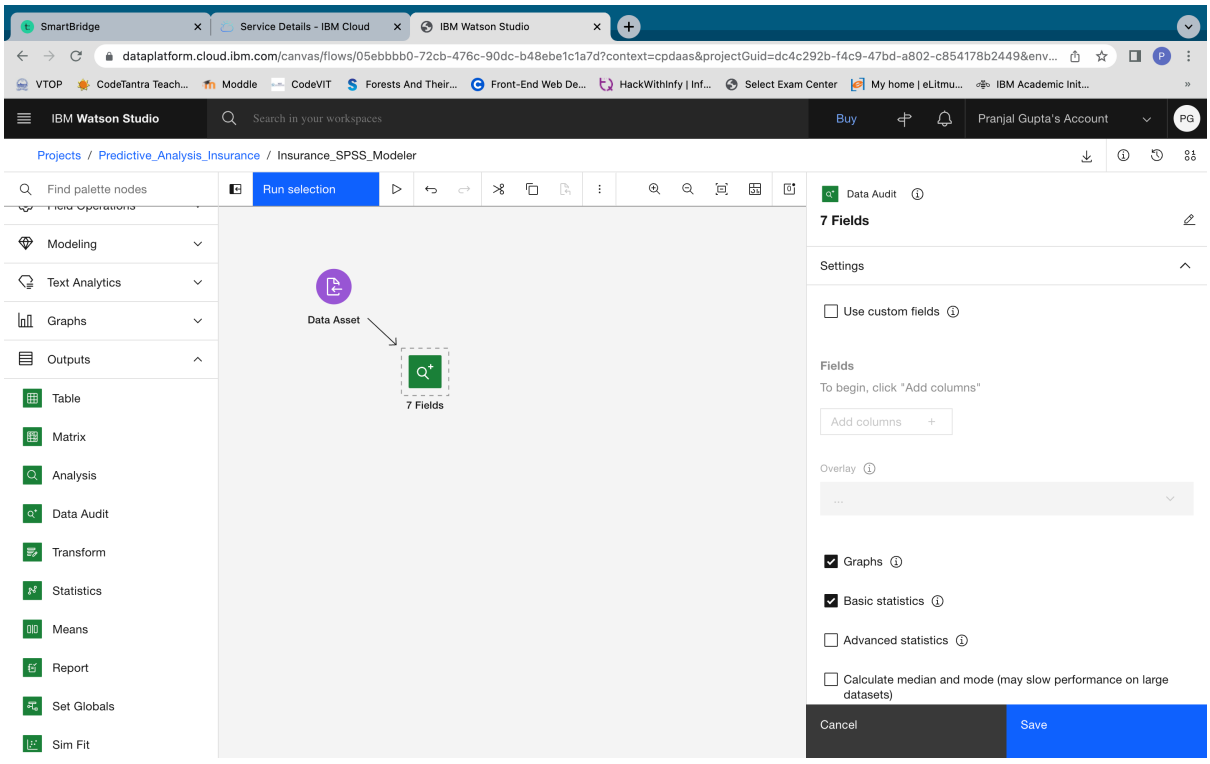



Figure: Performed data audit

Pranjal Gupta

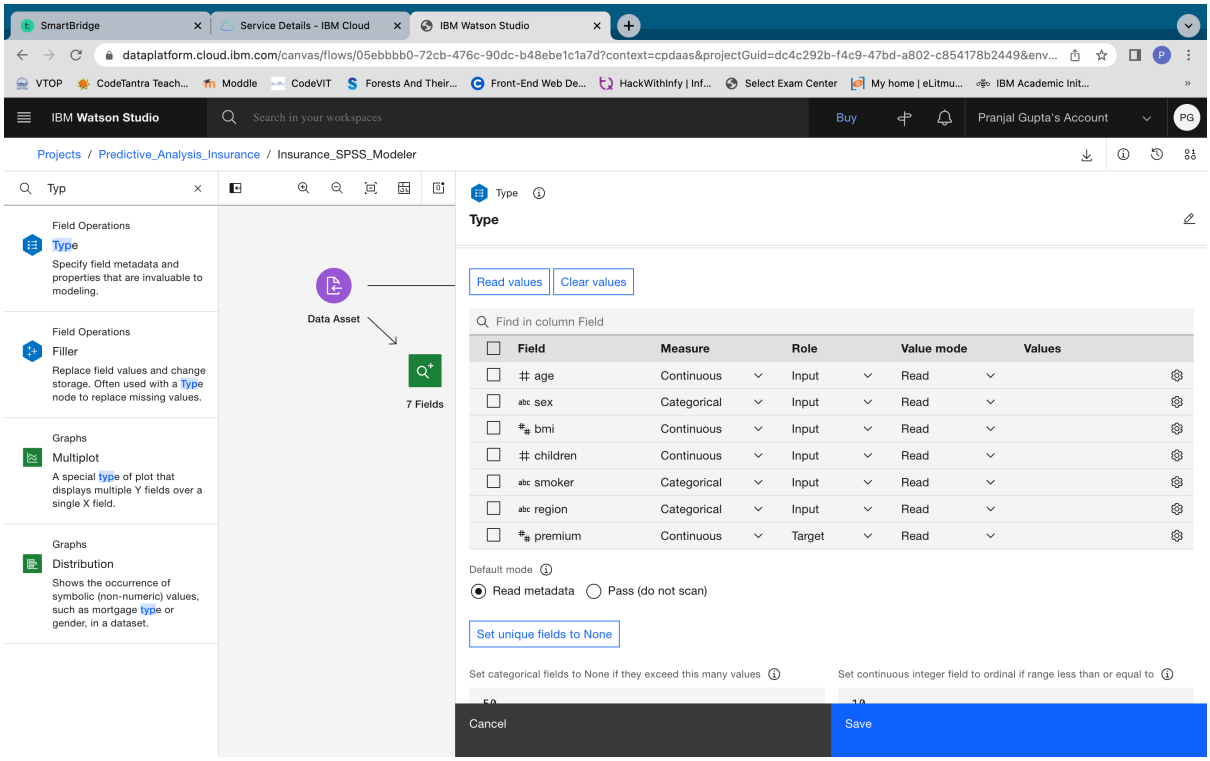



Figure: Specified the target attribute

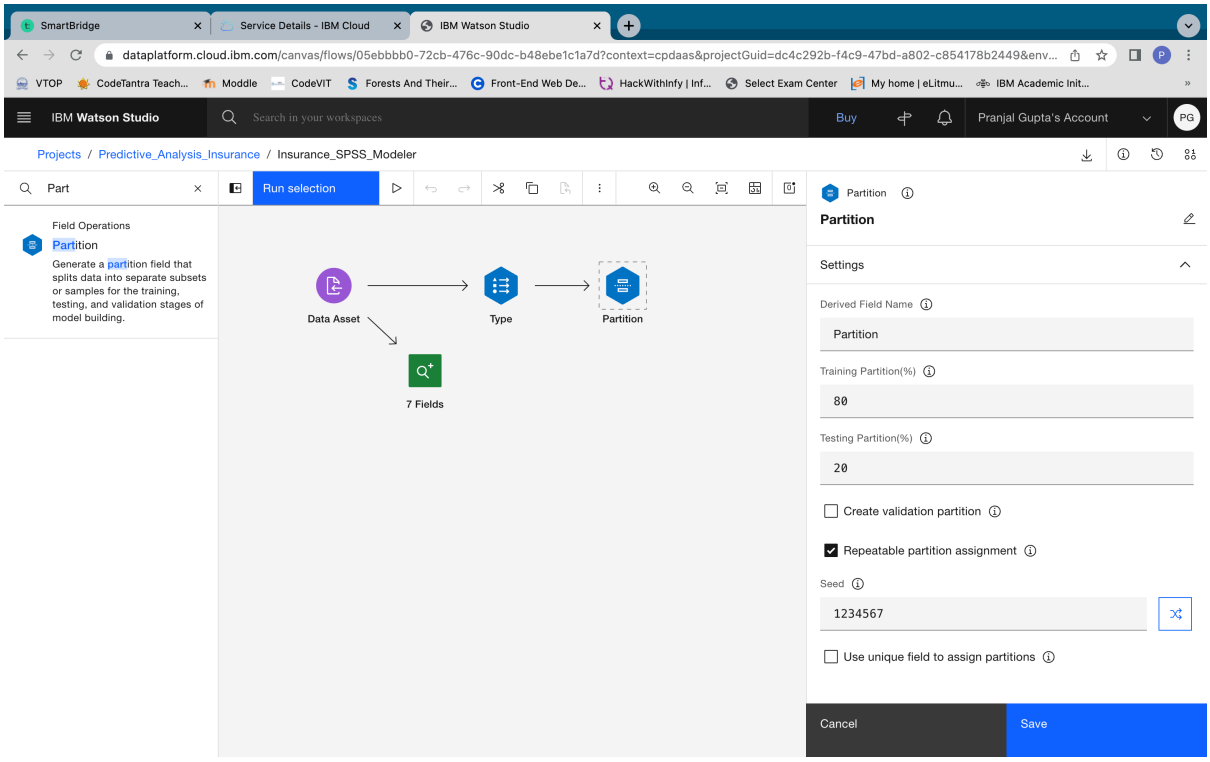



Figure: Fixed the test and training data partition

Pranjal Gupta

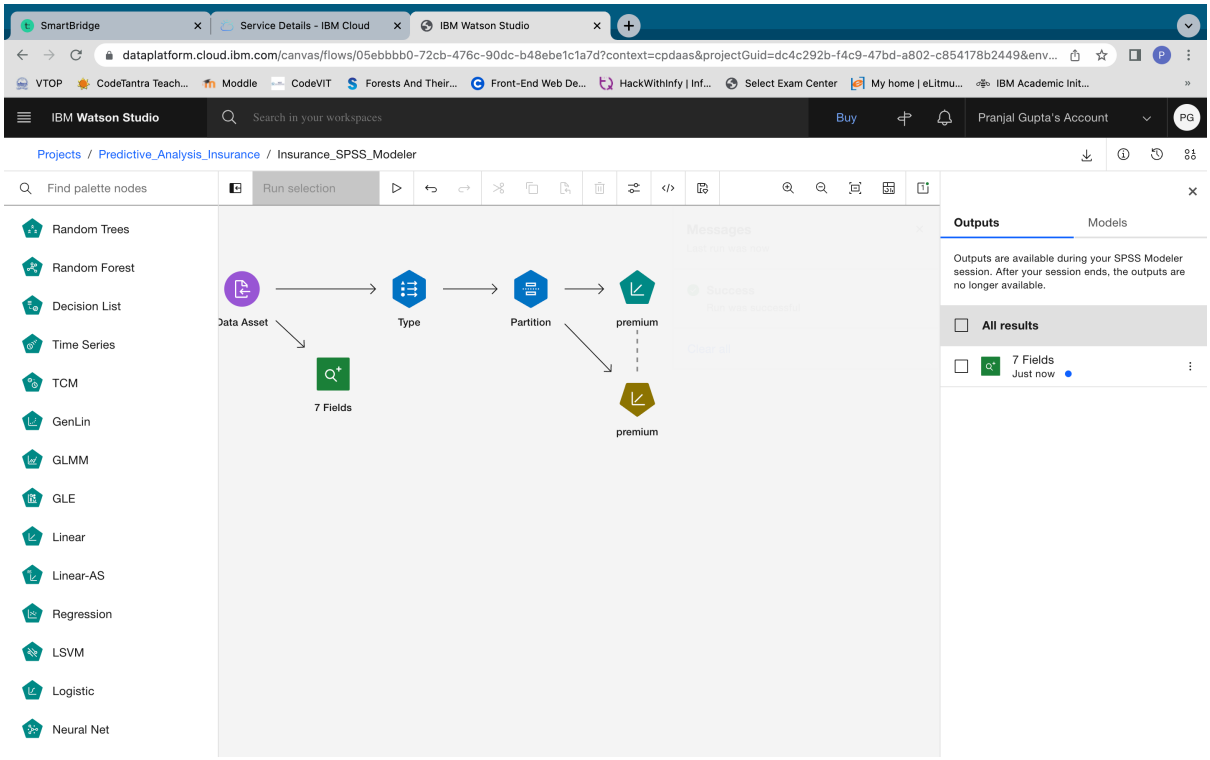



Figure: Performed linear regression


Figure: Linear plot for modelized premium over region

Pranjal Gupta

# Data Set: MallCustomers.csv

Link: https://dataplatform.cloud.ibm.com/projects/430ec4a6-234b-4189-b26c-9a1d4178eb09/assets?context=cpdaas

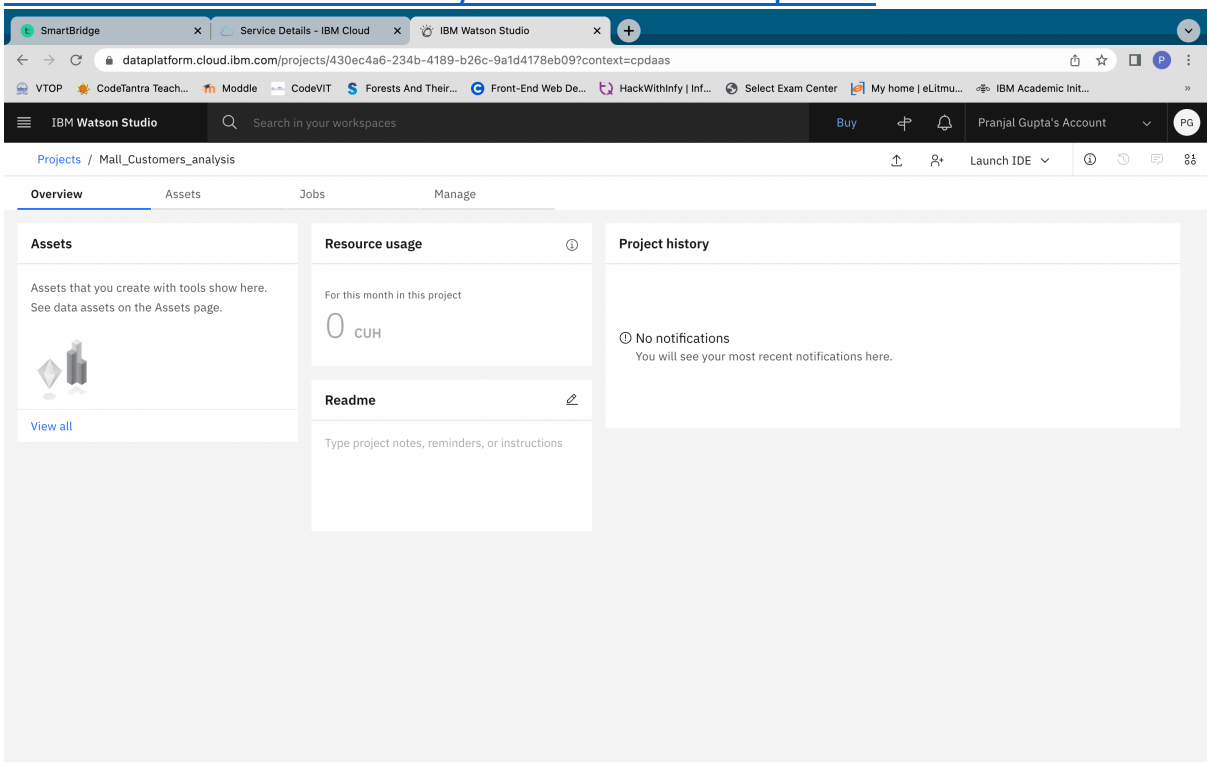



Figure: Created project for MallCustomers.csv

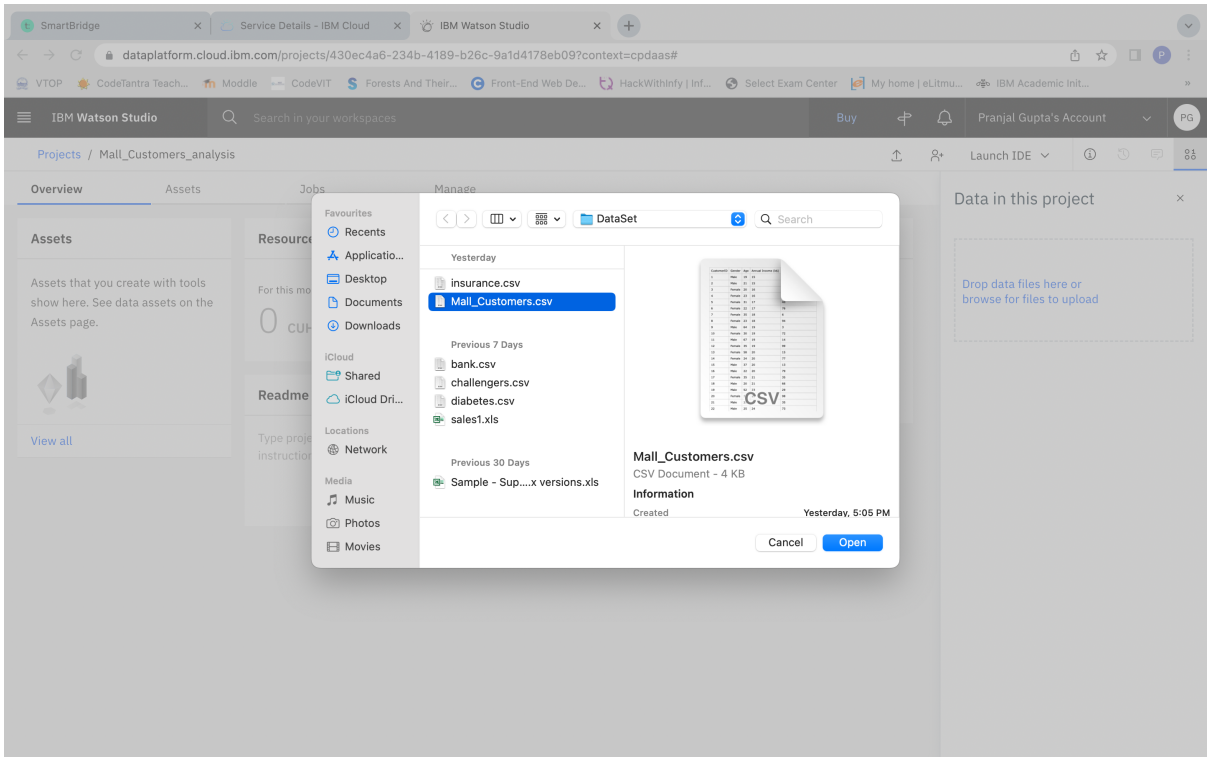



Figure: Adding data set

Pranjal Gupta

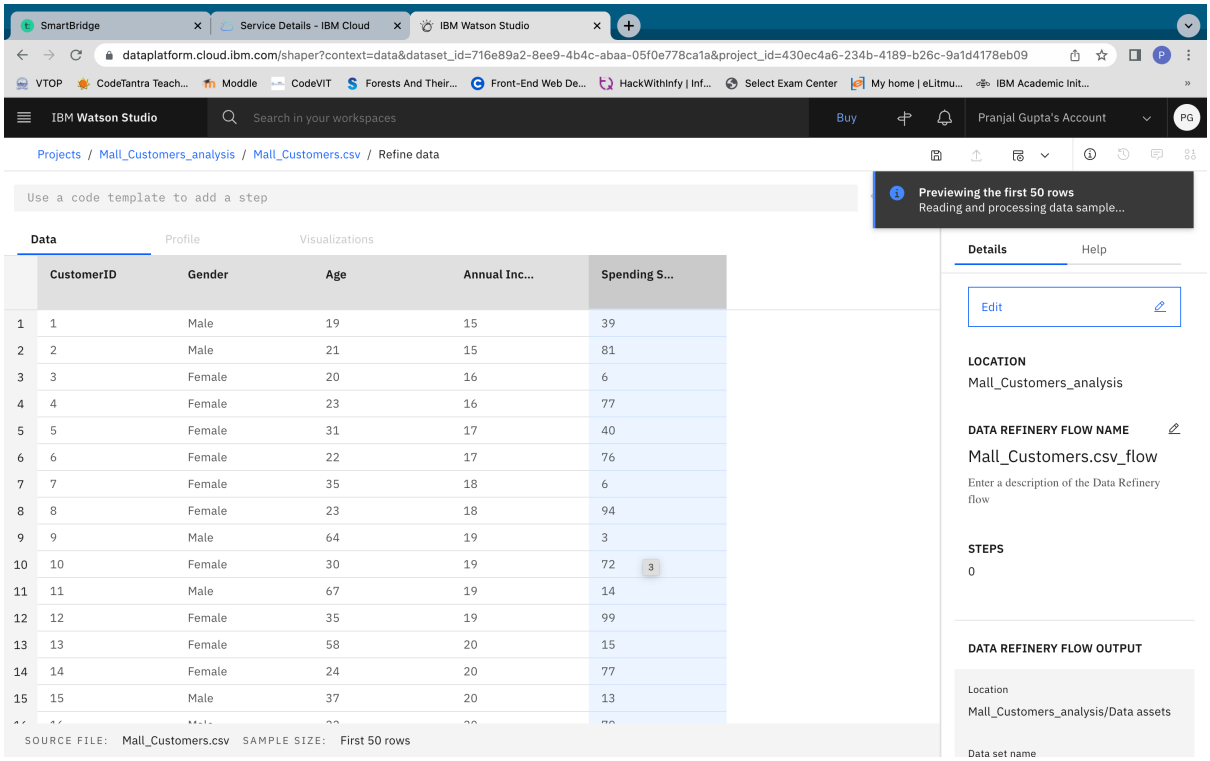



Figure: Creating data refinery flow

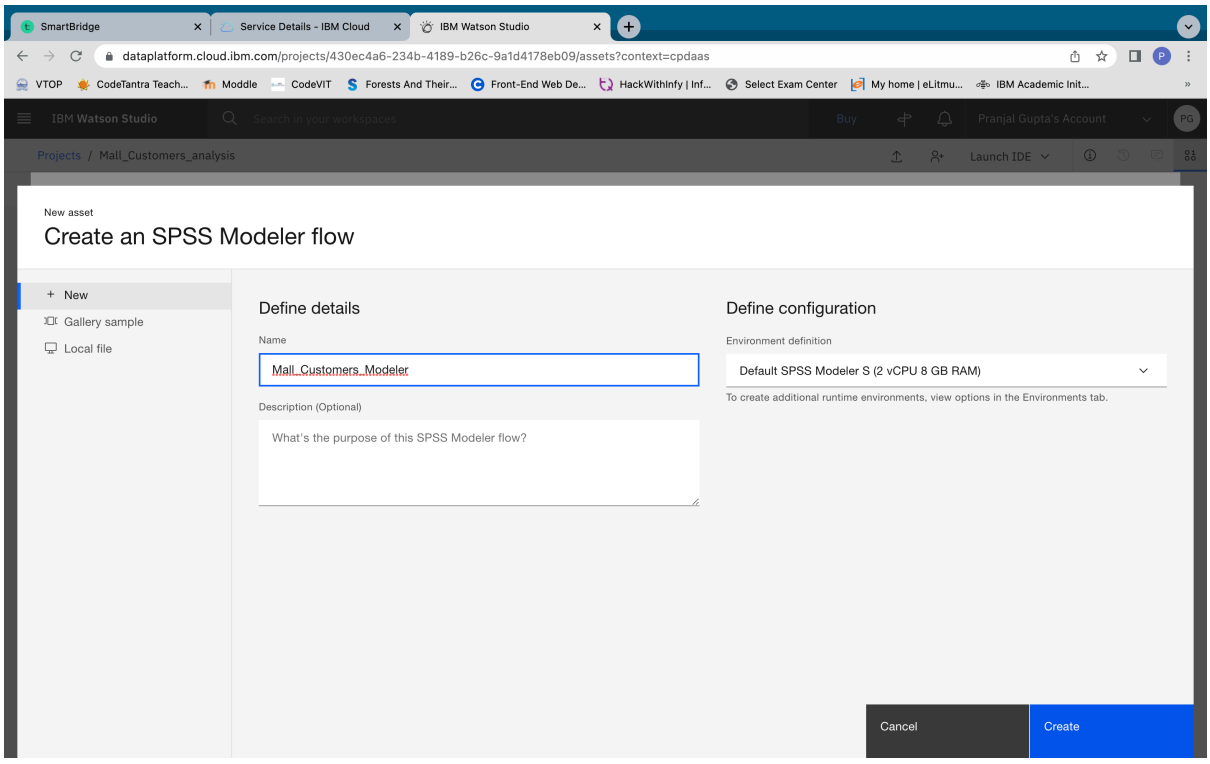



Figure: Creating SPSS modeler

Pranjal Gupta

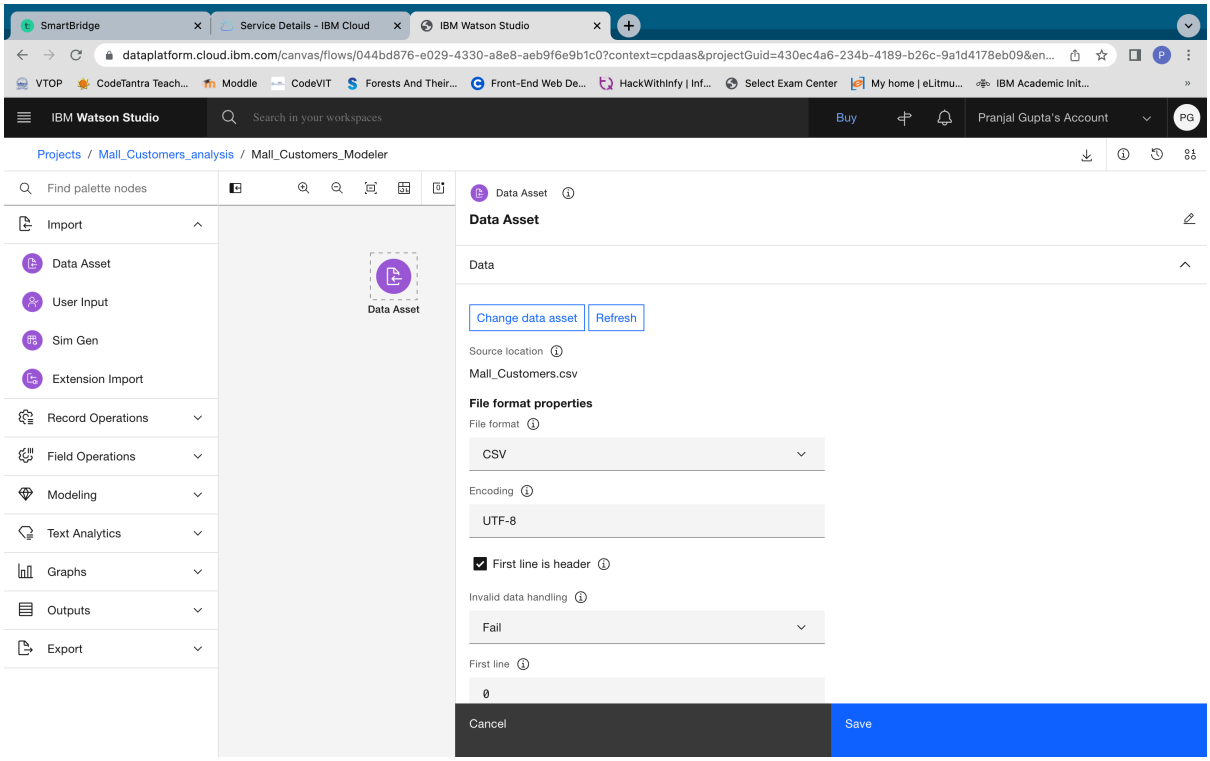



Figure: Imported data asset

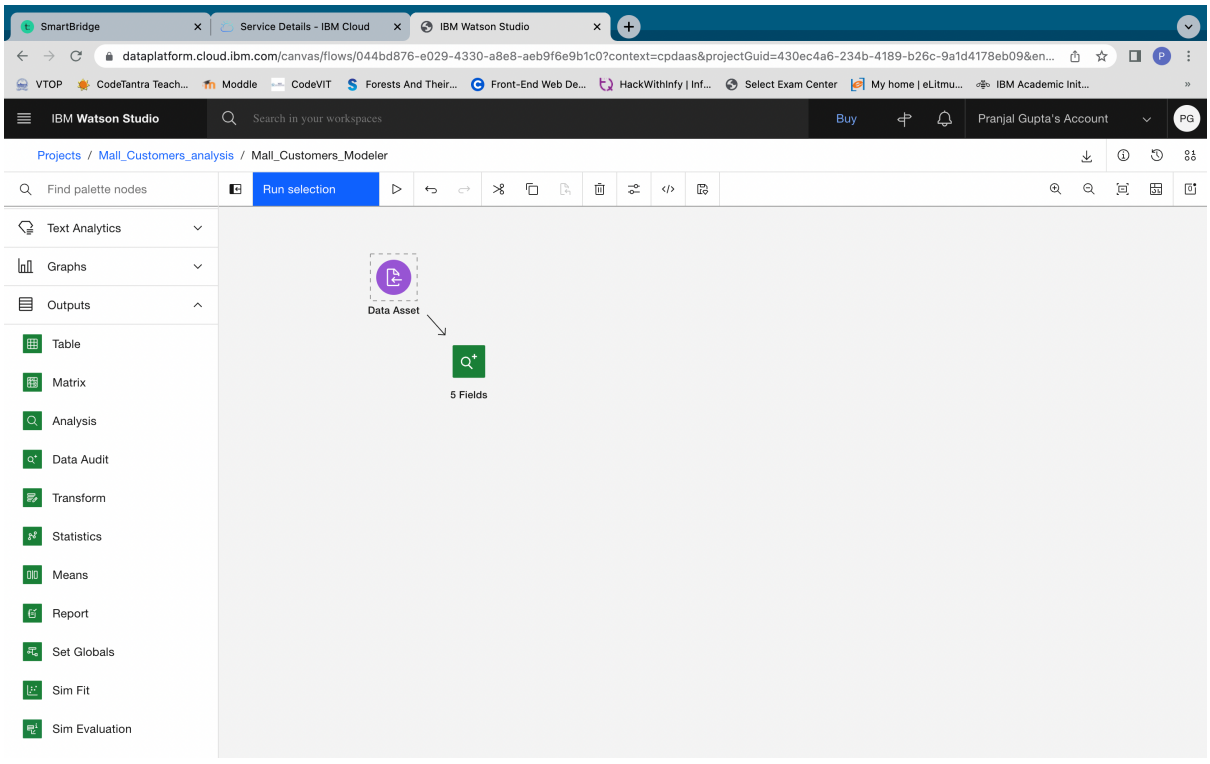



Figure: Performed data audit

Pranjal Gupta

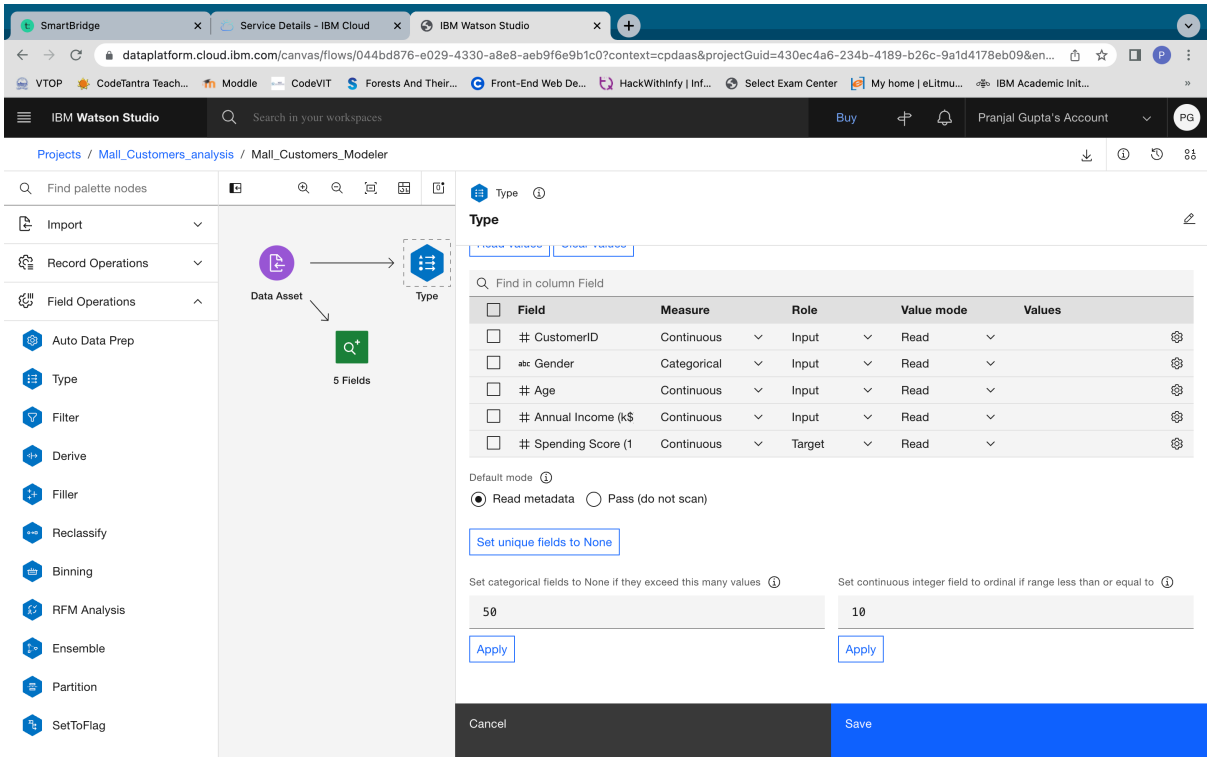



Figure: Specified the target attribute

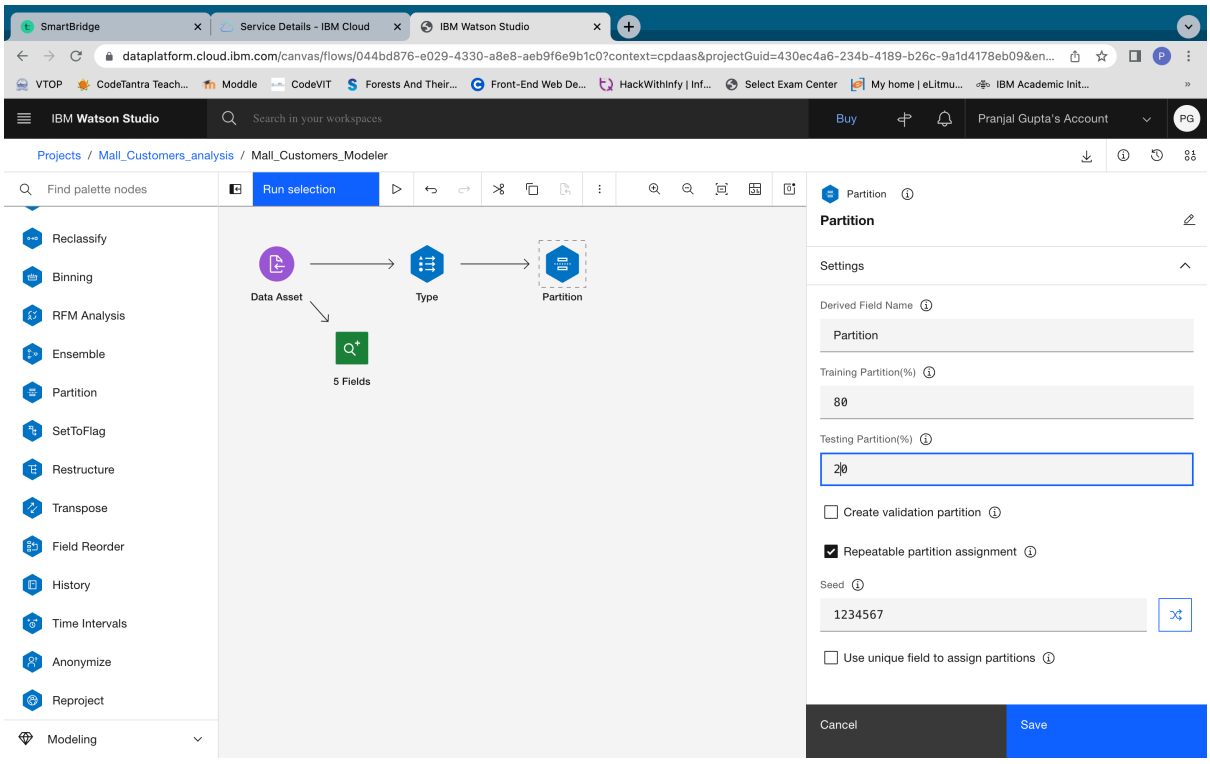



Figure: Fixed the training and test data

Pranjal Gupta

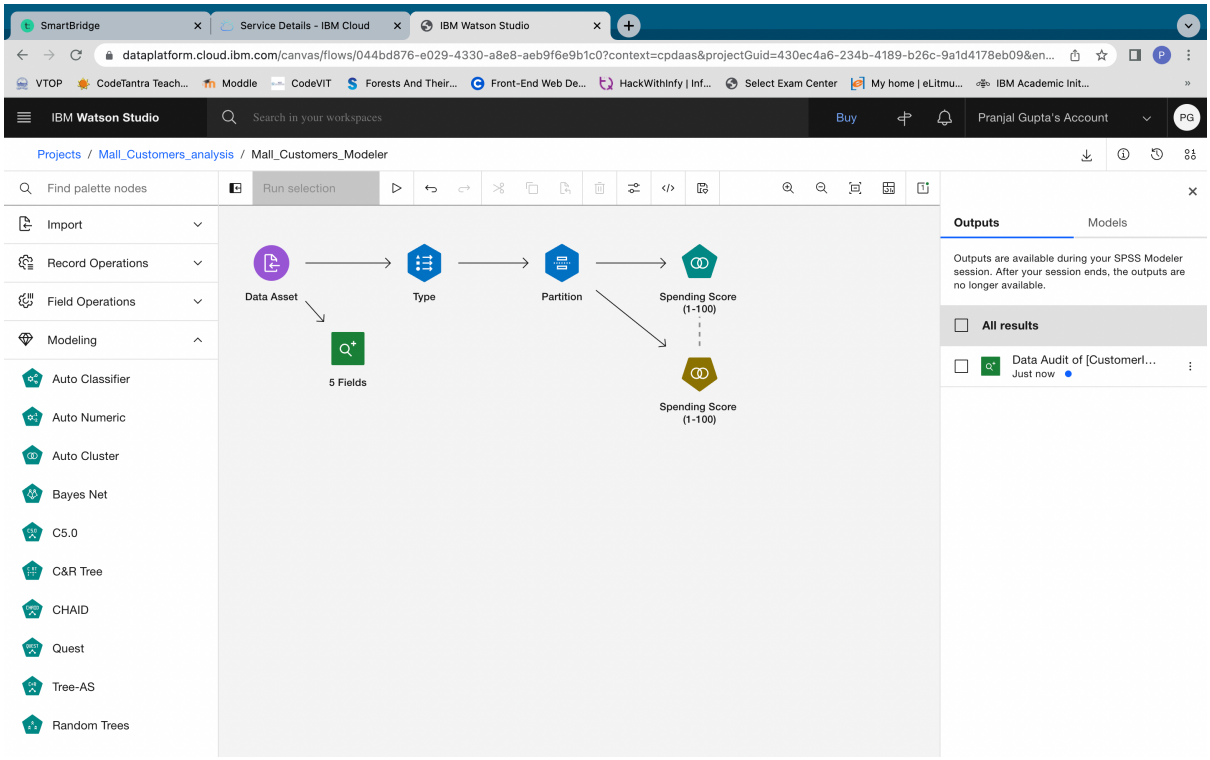



Figure: Performed auto clustering

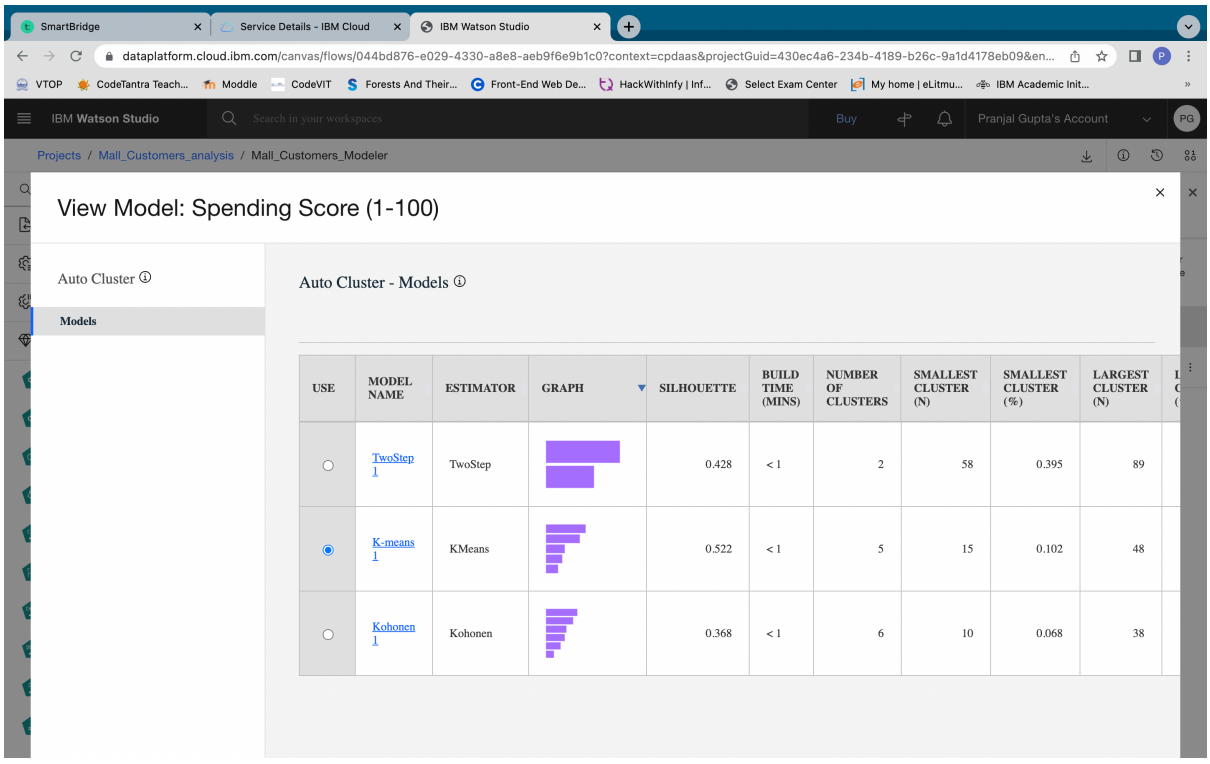



Figure: Model View

Pranjal Gupta

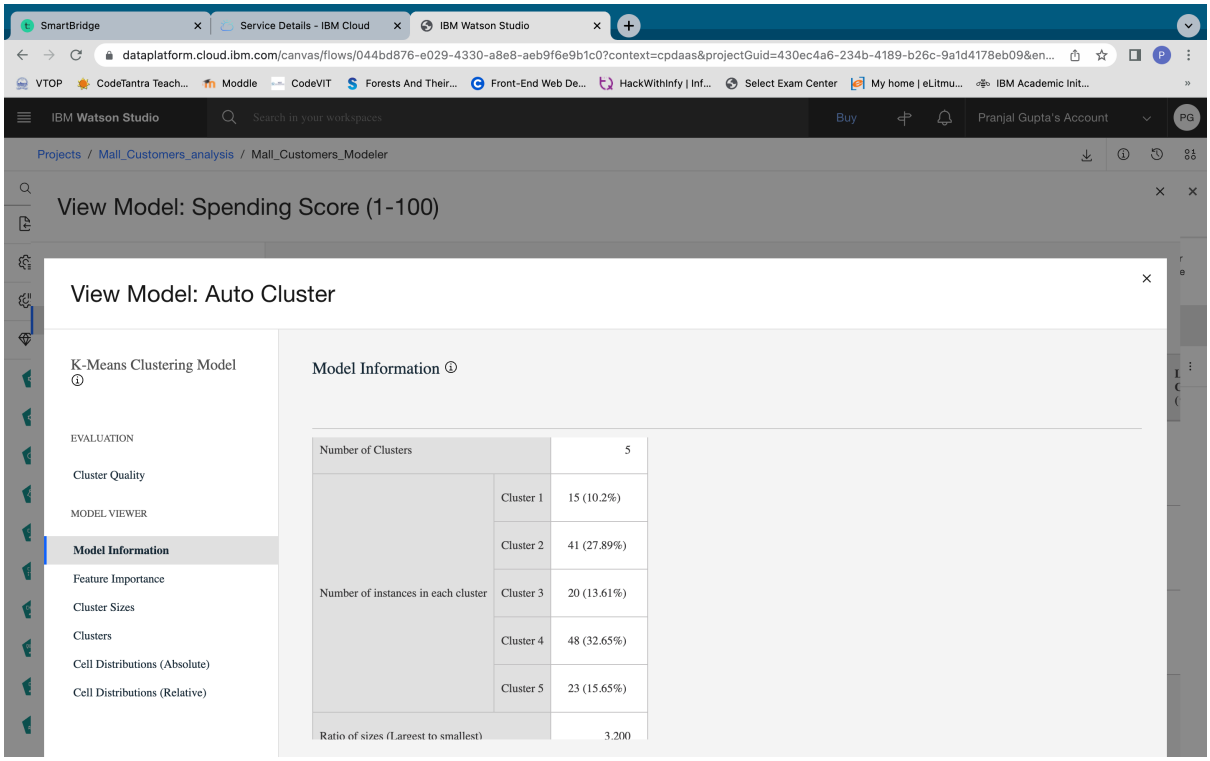



Figure: Model information for K-means clustering

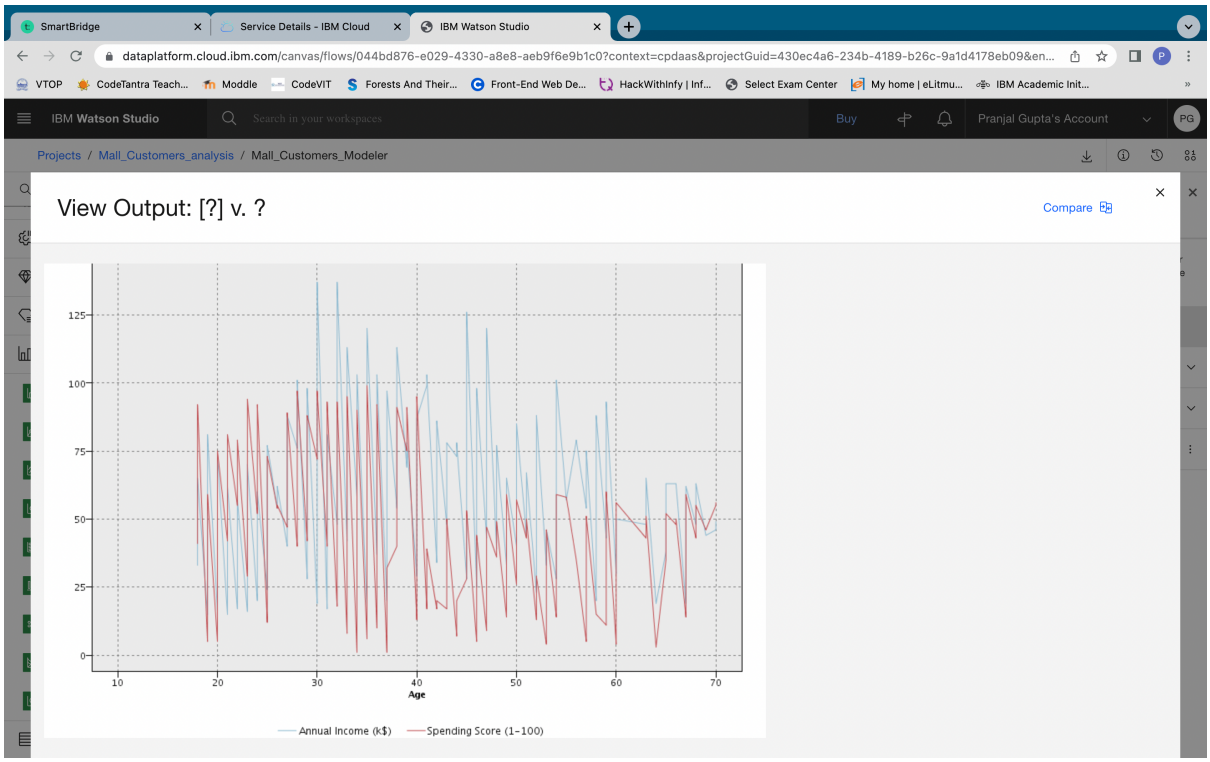



Figure: Linear plot for Age over annual income over spending score