

Practice Assignment

Name: Thiriveedhi dheeraj

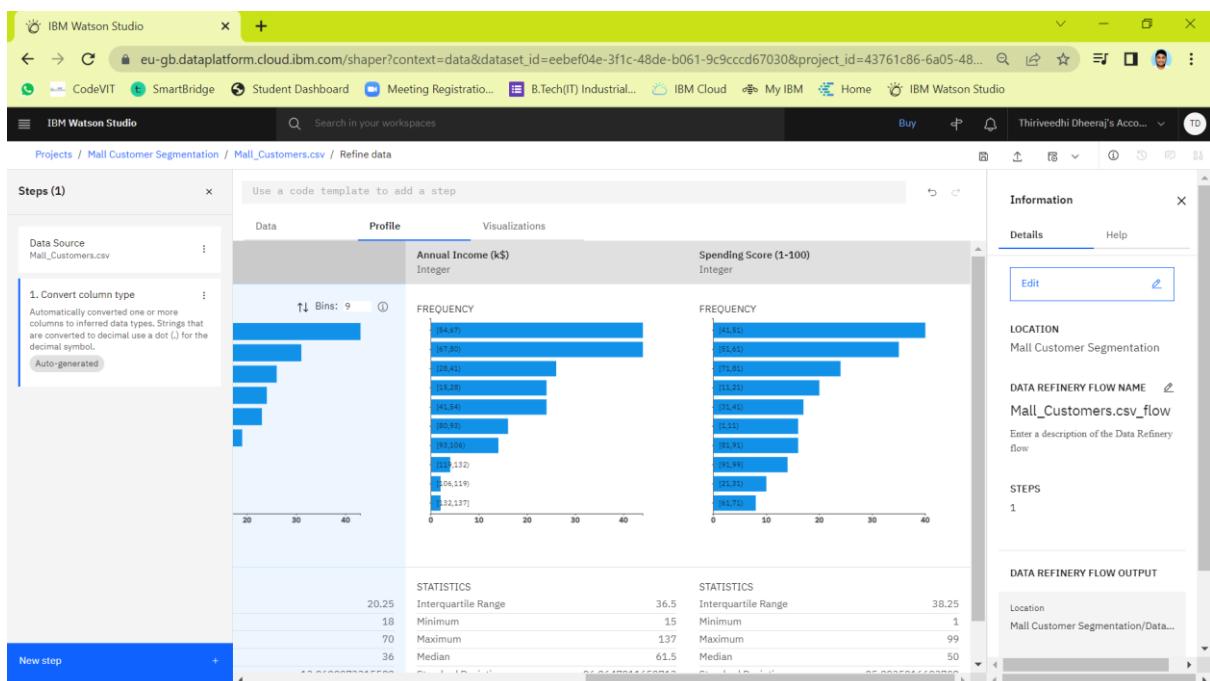
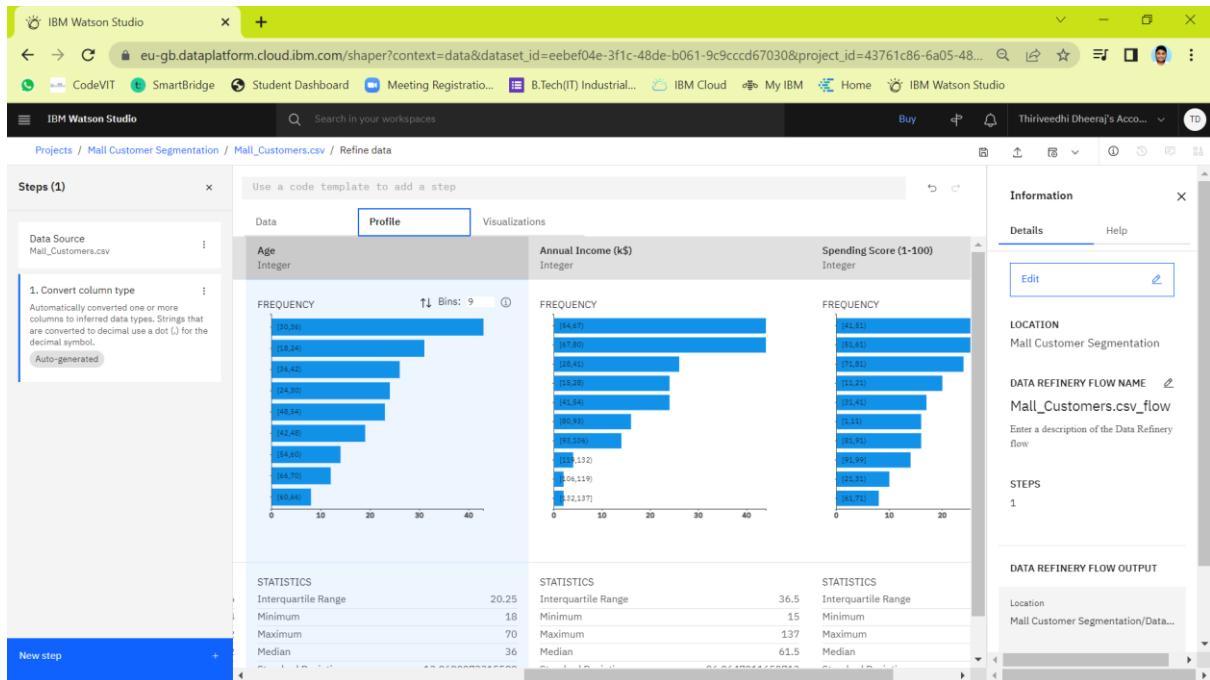
Reg No: 19BIT0066

Campus: VIT, Vellore

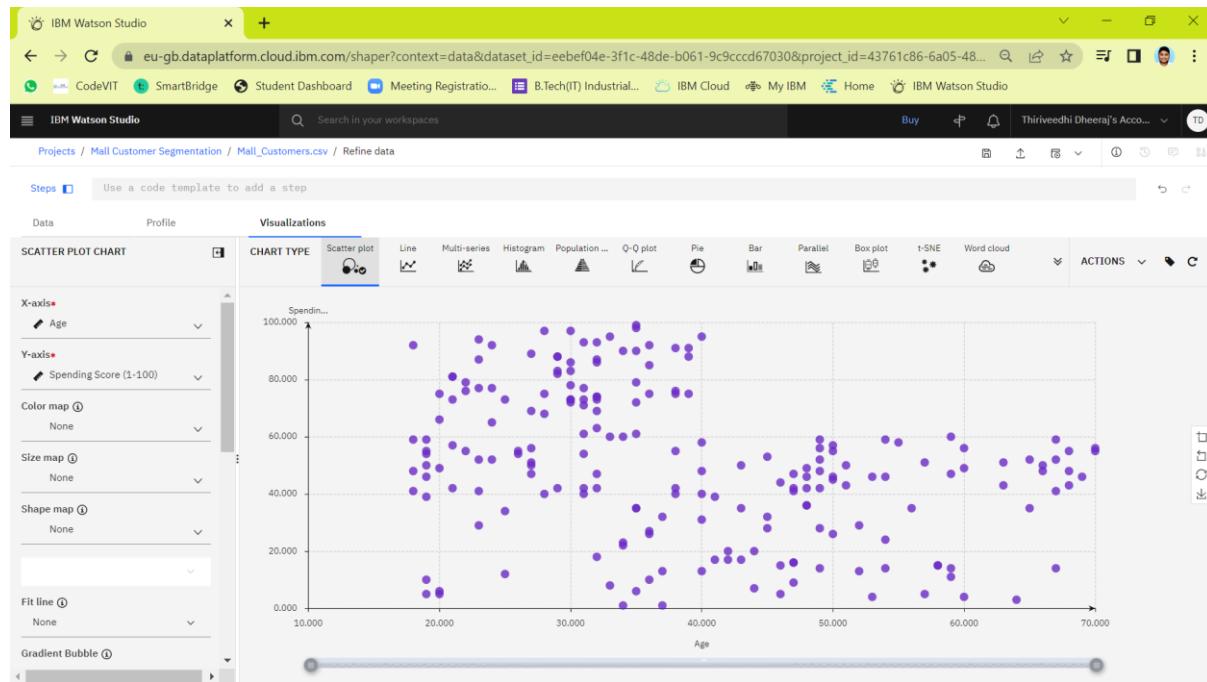
Mall customer segmentation:

This screenshot shows the IBM Watson Studio interface for a project titled "Mall Customer Segmentation". The main area displays a table of data from "Mail_Customers.csv" with 200 rows. The columns are CustomerID (Integer), Gender (String), Age (Integer), Annual Inc... (Integer), and Spending S... (Integer). The "Data" tab is selected. On the right, the "Information" panel shows the data refinery flow name as "Mall_Customers.csv_flow" and the location as "Mall Customer Segmentation".

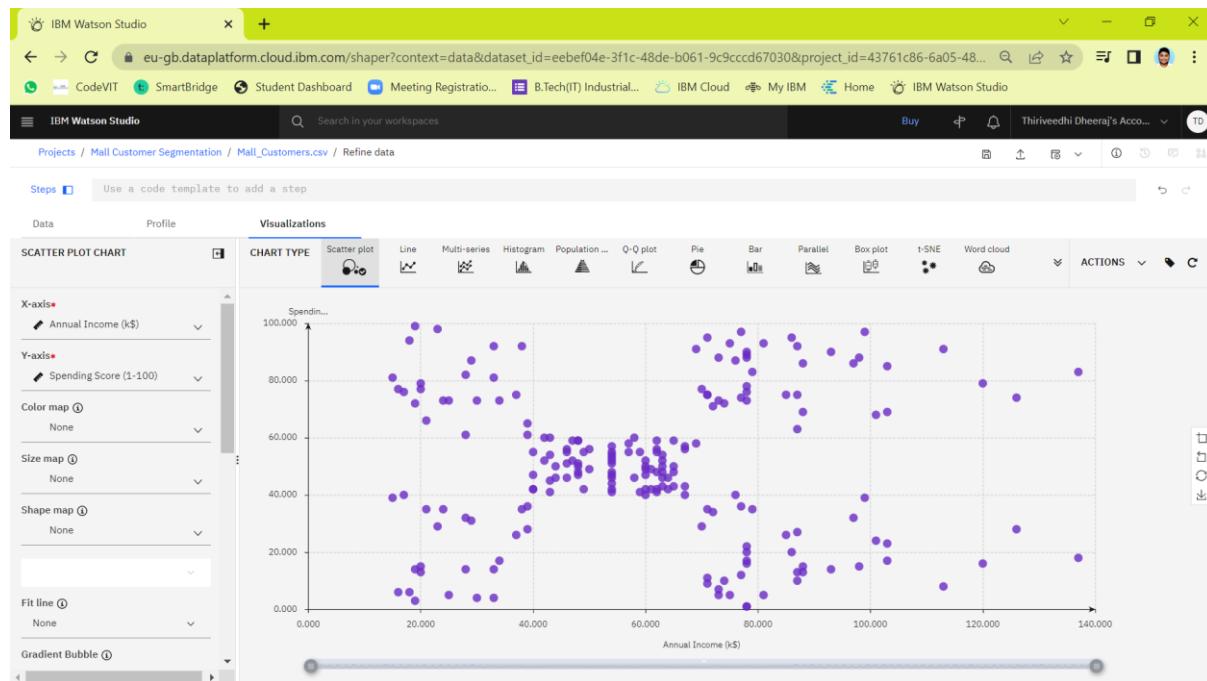
This screenshot shows the same project in IBM Watson Studio, but the "Profile" tab is now selected in the main data view. It displays three bar charts for CustomerID, Gender, and Age. The "CustomerID" chart shows frequency distribution across ranges [1,2], [3,4], [4,5], [5,6], [6,7], [7,8], [8,9], [9,10], [10,11], and [11,12]. The "Gender" chart shows frequencies for Female (~85) and Male (~15). The "Age" chart shows frequencies for age groups [10,16], [16,24], [24,32], [32,40], [40,48], [48,56], [56,64], [64,72], [72,80], and [80,88]. Below the charts, statistical summary tables are provided for each column.



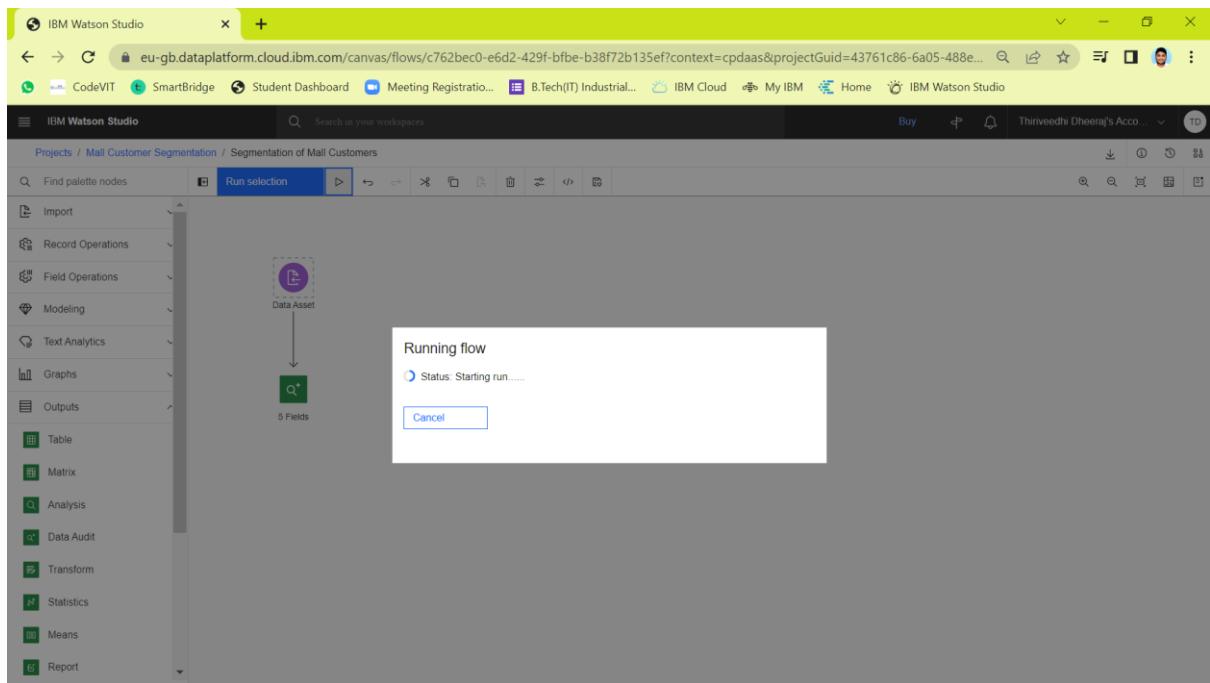
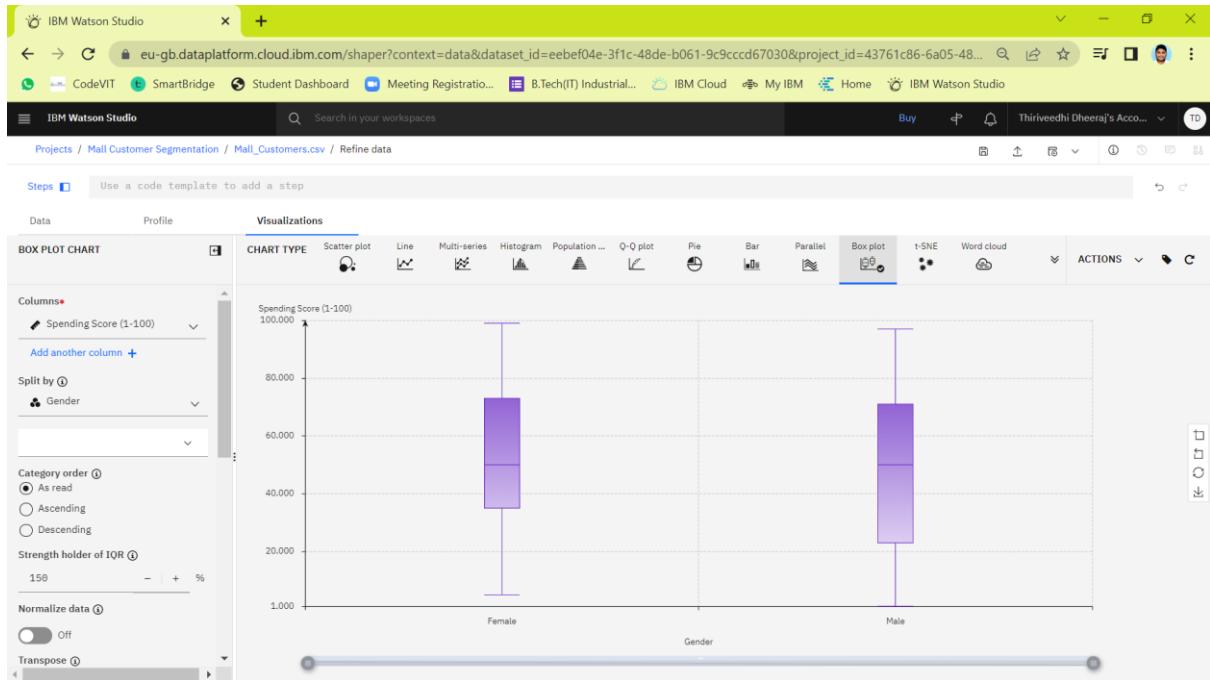
Age vs Spending score:



Annual Income vs spending score:



Spending score based on gender using box plot:



IBM Watson Studio

eu-gb.dataplatfrom.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

IBM Watson Studio

Projects / Mall Customer Segmentation / Segmentation of Mall Customers

Run selection

Find palette nodes

Import Record Operations Field Operations Modeling Text Analytics Graphs Outputs Table Matrix Analysis Data Audit Transform Statistics Means Report

Data Asset

5 Fields

Messages

Last run was now

Success Run was successful

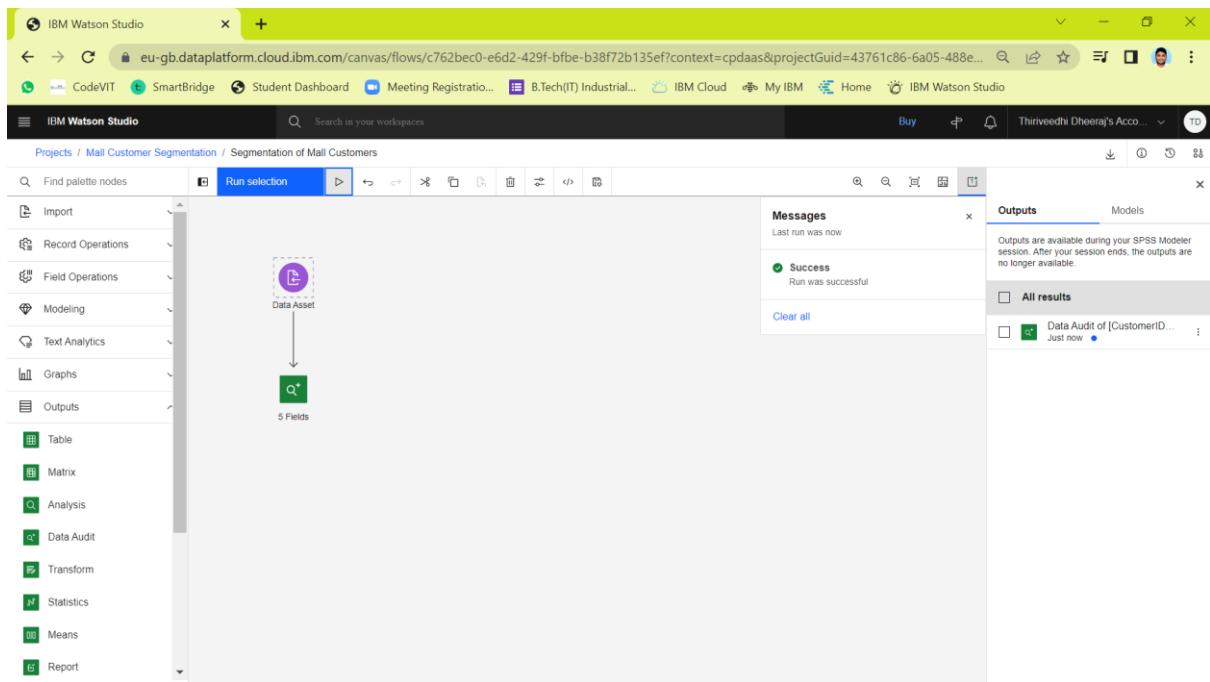
Clear all

Outputs Models

Outputs are available during your SPSS Modeler session. After your session ends, the outputs are no longer available.

All results

Data Audit of [CustomerID... Just now



IBM Watson Studio

eu-gb.dataplatfrom.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

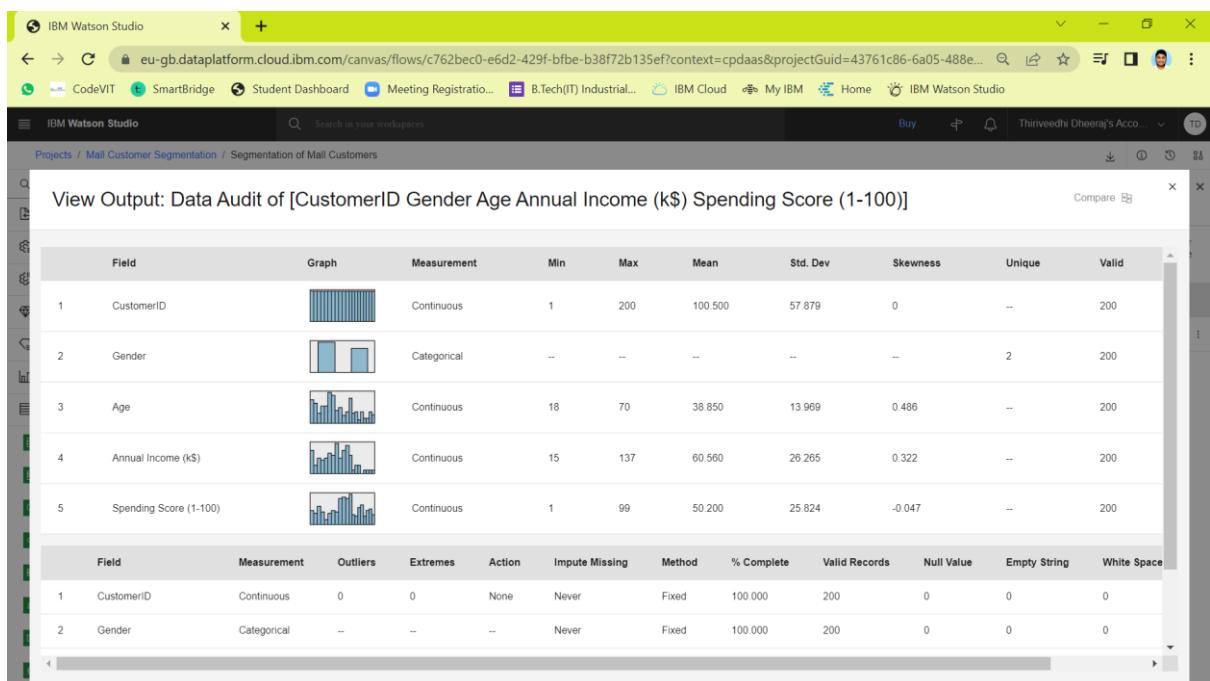
IBM Watson Studio

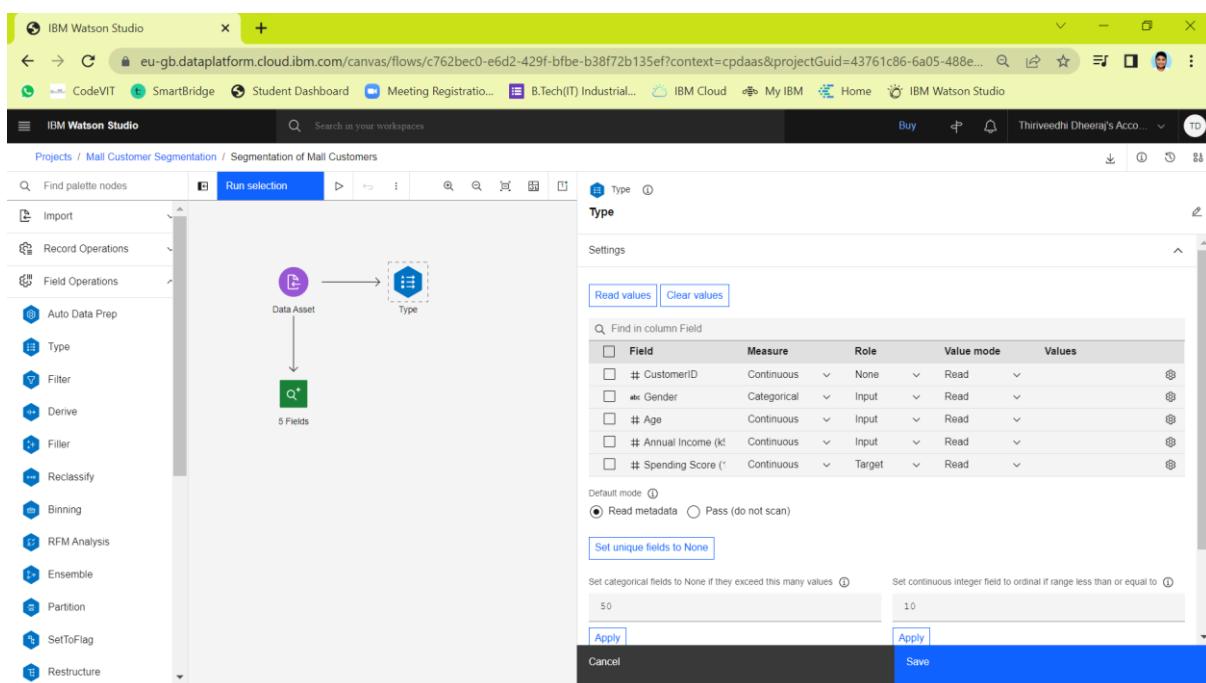
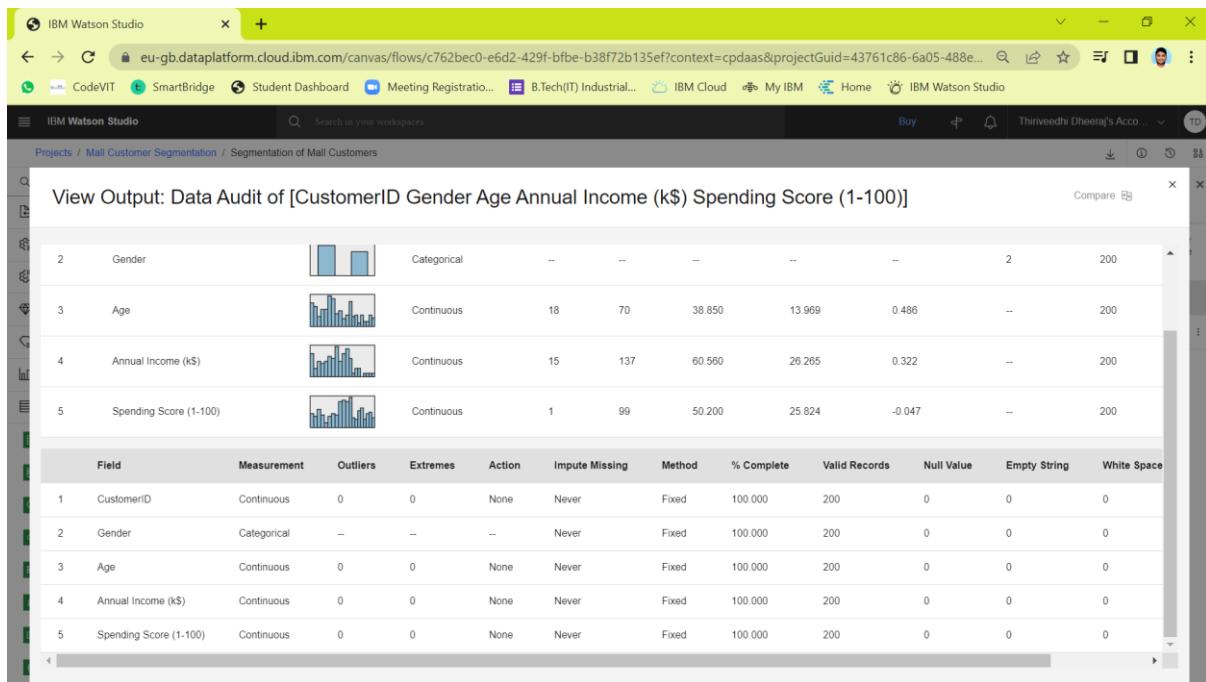
Projects / Mall Customer Segmentation / Segmentation of Mall Customers

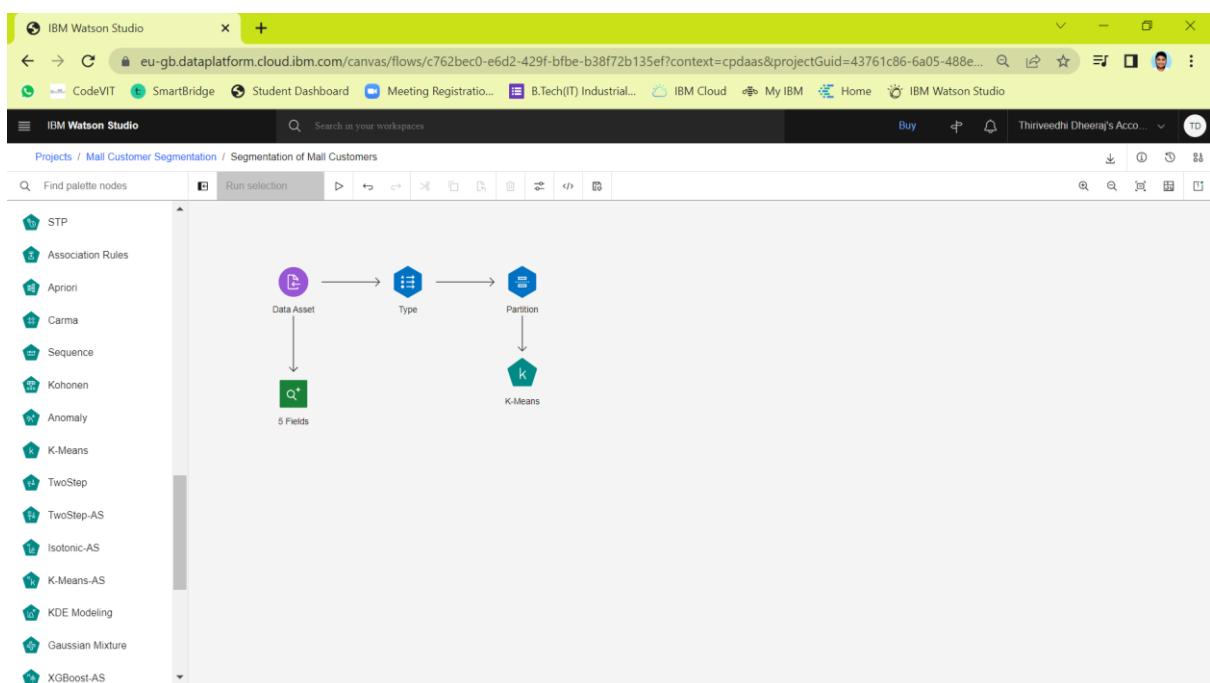
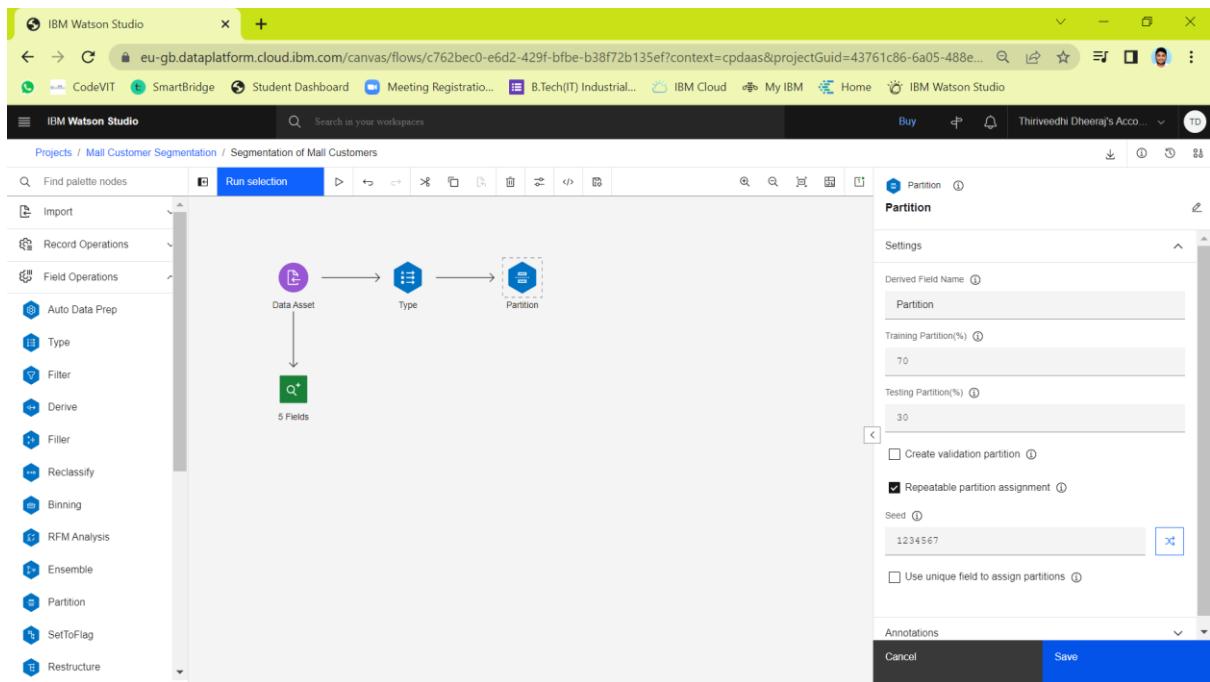
View Output: Data Audit of [CustomerID Gender Age Annual Income (k\$) Spending Score (1-100)]

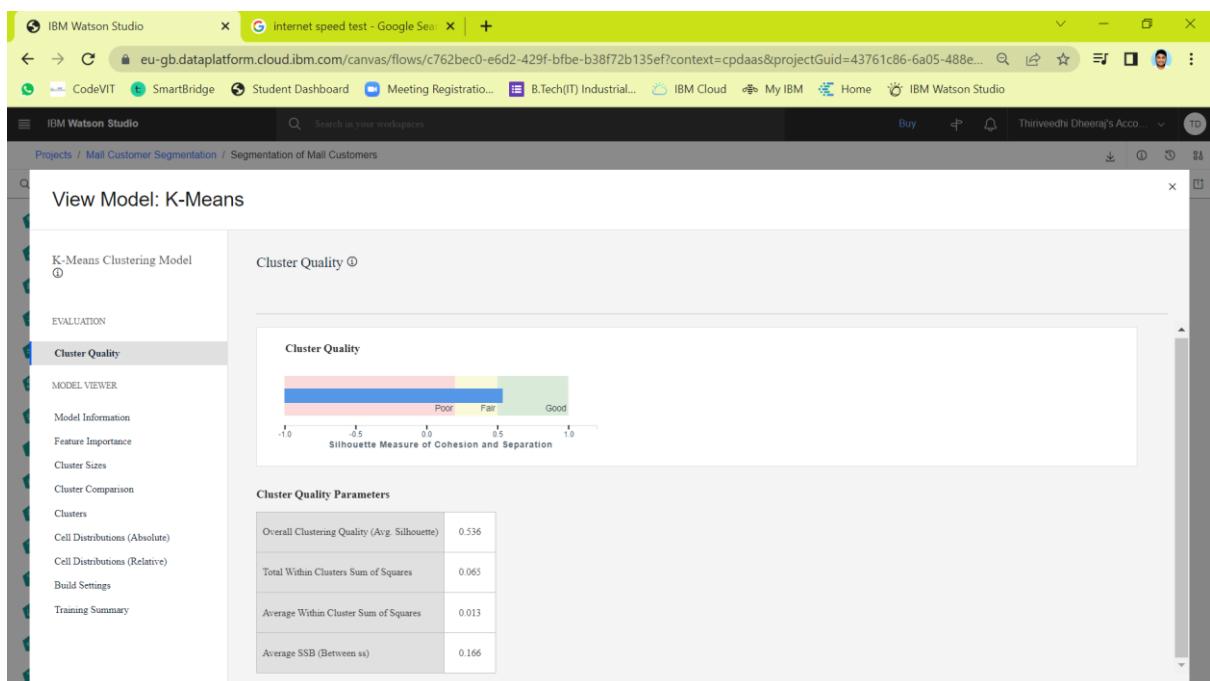
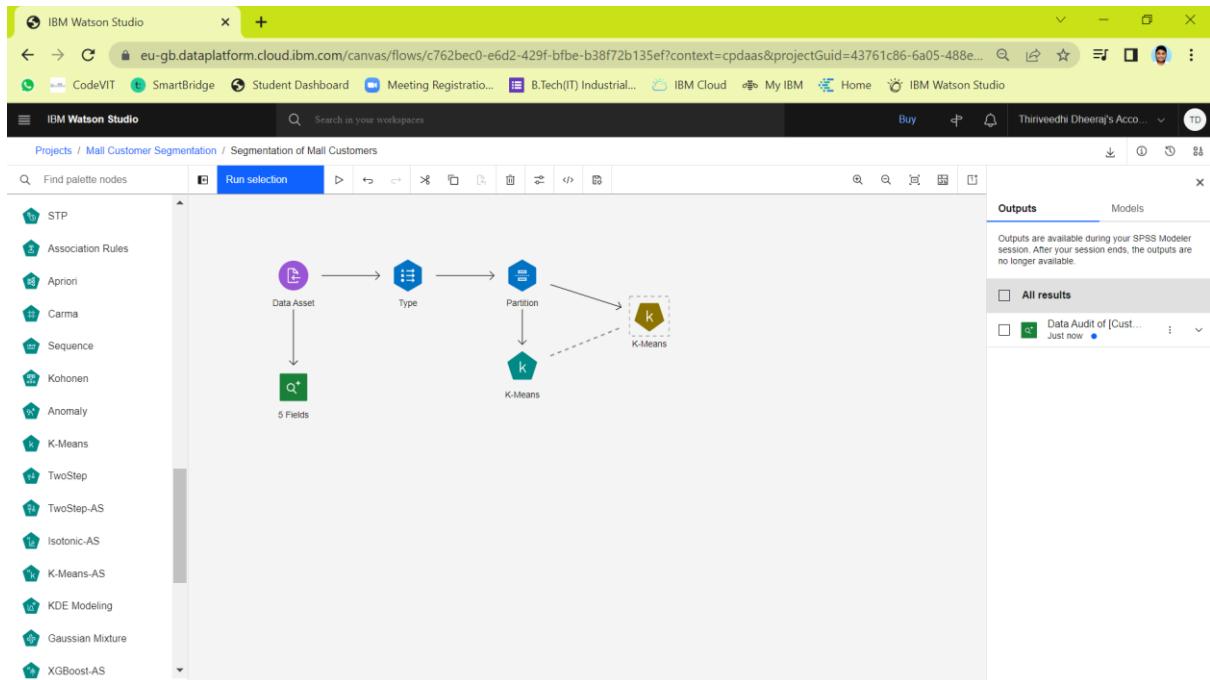
Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1 CustomerID		Continuous	1	200	100.500	57.879	0	--	200
2 Gender		Categorical	--	--	--	--	--	2	200
3 Age		Continuous	18	70	38.850	13.969	0.486	--	200
4 Annual Income (k\$)		Continuous	15	137	60.560	26.265	0.322	--	200
5 Spending Score (1-100)		Continuous	1	99	50.200	25.824	-0.047	--	200

Field	Measurement	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value	Empty String	White Space
1 CustomerID	Continuous	0	0	None	Never	Fixed	100.000	200	0	0	0
2 Gender	Categorical	--	--	--	Never	Fixed	100.000	200	0	0	0









IBM Watson Studio internet speed test - Google Se... +

eu-gb.dataplatfrom.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

IBM Watson Studio Buy Search in your workspace Thiruveedhi Dheeraj's Acco... 33

View Model: K-Means

K-Means Clustering Model

EVALUATION Cluster Quality

MODEL VIEWER

Model Information

Feature Importance Cluster Sizes Cluster Comparison Clusters Cell Distributions (Absolute) Cell Distributions (Relative) Build Settings Training Summary

Model Information

Algorithm	K-Means
Model Class	Center Based
Number of Features	3
Distance Measure	Euclidean
Number of Clusters	5
Cluster 1	13 (9.77%)
Cluster 2	33 (19.83%)
Cluster 3	19 (14.29%)
Cluster 4	19 (14.29%)
Cluster 5	29 (21.8%)
Ratio of sizes (Largest to smallest)	4.077

IBM Watson Studio internet speed test - Google Se... +

eu-gb.dataplatfrom.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

IBM Watson Studio Buy Search in your workspace Thiruveedhi Dheeraj's Acco... 33

Projects / Mail Customer Segmentation / Segmentation of Mail Customers

View Model: K-Means

K-Means Clustering Model

EVALUATION Cluster Quality

MODEL VIEWER

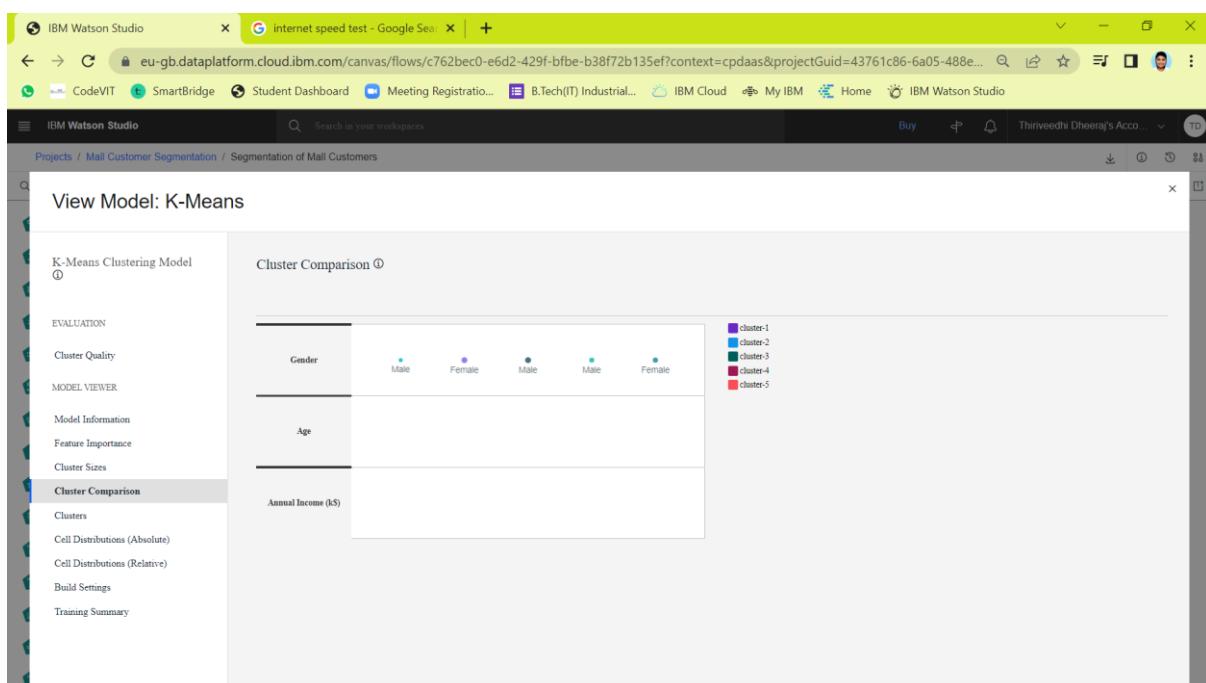
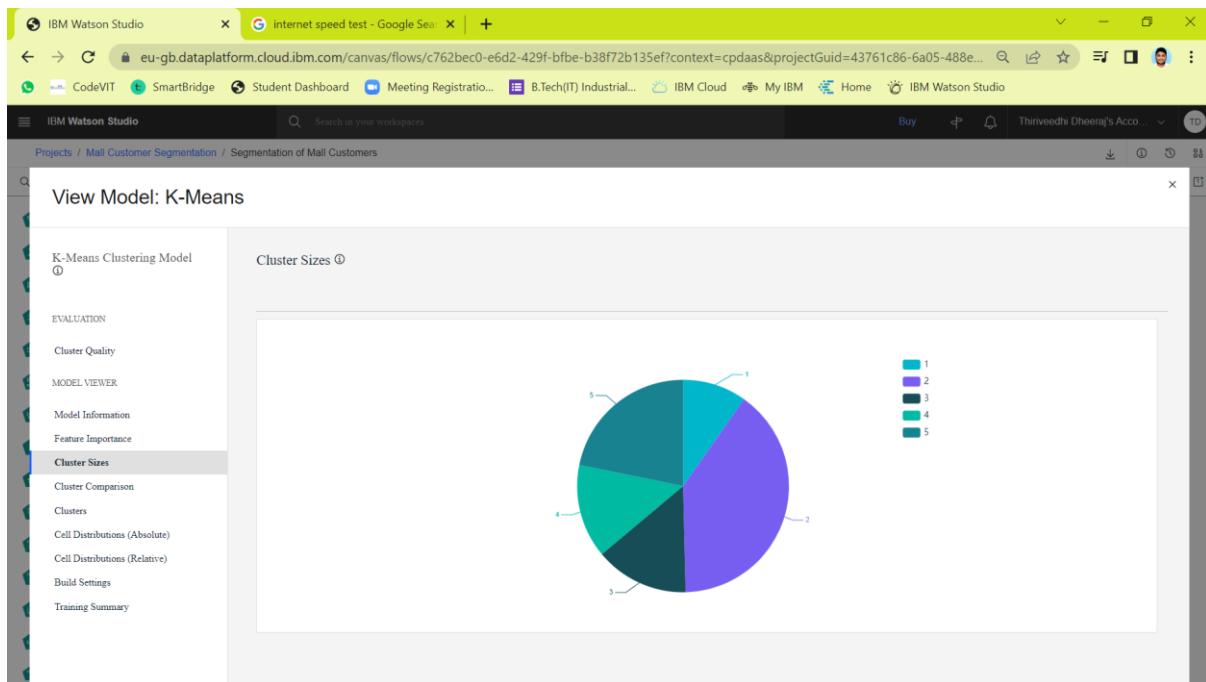
Model Information

Feature Importance

Cluster Sizes Cluster Comparison Clusters Cell Distributions (Absolute) Cell Distributions (Relative) Build Settings Training Summary

Feature Importance

Feature	Importance
Gender	1.00
Age	0.89
Annual Income (k\$)	0.56



IBM Watson Studio internet speed test - Google Sea... +

eu-gb.dataplatform.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

IBM Watson Studio Search in your workspace

Projects / Mail Customer Segmentation / Segmentation of Mail Customers

View Model: K-Means

K-Means Clustering Model

EVALUATION Cluster Quality

MODEL VIEWER Model Information Feature Importance Cluster Sizes Cluster Comparison Clusters Cell Distributions (Absolute) Cell Distributions (Relative) Build Settings Training Summary

Clusters

Clusters

Input Importance

Cluster	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5
Size	Gender Male (100.00%)	Gender Female (100.00%)	Gender Male (100.00%)	Gender Male (100.00%)	Gender Female (100.00%)
Inputs	Age 27.12	Age 45.59	Age 36.24	Age 26.83	Age 27.23
	Annual Income (k\$) 37.12	Annual Income (k\$) 70.88	Annual Income (k\$) 88.47	Annual Income (k\$) 61.17	Annual Income (k\$) 47.23

Count

Gender

Female

Male

IBM Watson Studio internet speed test - Google Sea... +

eu-gb.dataplatform.cloud.ibm.com/canvas/flows/c762bec0-e6d2-429f-bfbe-b38f72b135ef?context=cpdaas&projectGuid=43761c86-6a05-488e...

CodeVIT SmartBridge Student Dashboard Meeting Registration... B.Tech(IT) Industrial... IBM Cloud My IBM Home IBM Watson Studio

IBM Watson Studio Search in your workspace

Projects / Mail Customer Segmentation / Segmentation of Mail Customers

View Model: K-Means

K-Means Clustering Model

EVALUATION Cluster Quality

MODEL VIEWER Model Information Feature Importance Cluster Sizes Cluster Comparison Clusters Cell Distributions (Absolute) Cell Distributions (Relative) Build Settings Training Summary

Cell Distributions (Absolute)

Cell Distributions (Absolute)

Cluster

cluster_1 cluster_2 cluster_3 cluster_4 cluster_5

Size

Gender

Age

Annual Income (k\$)

Count

Gender

Female

Male

The screenshot shows the IBM Watson Studio interface with a project titled "Mail Customer Segmentation". The left sidebar lists various model evaluation metrics: Cluster Quality, Model Information, Feature Importance, Cluster Size, Cluster Comparison, Clusters, Cell Distributions (Absolute), and Cell Distributions (Relative). The "Cell Distributions (Relative)" option is currently selected.

The main area displays the "View Model: K-Means" results. It includes a heatmap showing cell distributions across five clusters for variables like Sex, Gender, Age, and Annual Income. To the right, there is a bar chart titled "Gender" showing the count of females and males.

	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5
Sex	High	Low	Low	Low	High
Gender	Medium	Low	High	Medium	Low
Age	Low	High	Medium	High	Low
Annual Income (k\$)	Low	Medium	High	Low	Medium

Bar Chart Data:

Gender	Count
Female	100
Male	25

The screenshot shows the IBM Watson Studio interface with a project titled "Mall Customer Segmentation". The left sidebar has a tree view with nodes like "K-Means Clustering Model", "EVALUATION", "Cluster Quality", "MODEL VIEWER", "Model Information", "Feature Importance", "Cluster Sizes", "Cluster Comparison", "Clusters", "Cell Distributions (Absolute)", "Cell Distributions (Relative)", and "Build Settings" (which is selected). The main panel displays the "Build Settings" configuration for a K-Means model. It includes a table with the following rows:

Use partitioned data	true
Calculate raw propensity scores	false
Calculate adjusted propensity scores	false
Number of clusters	5
Generate distance field	false
Cluster label	String
Label prefix	cluster
Optimize	Memory
Mode	Simple

