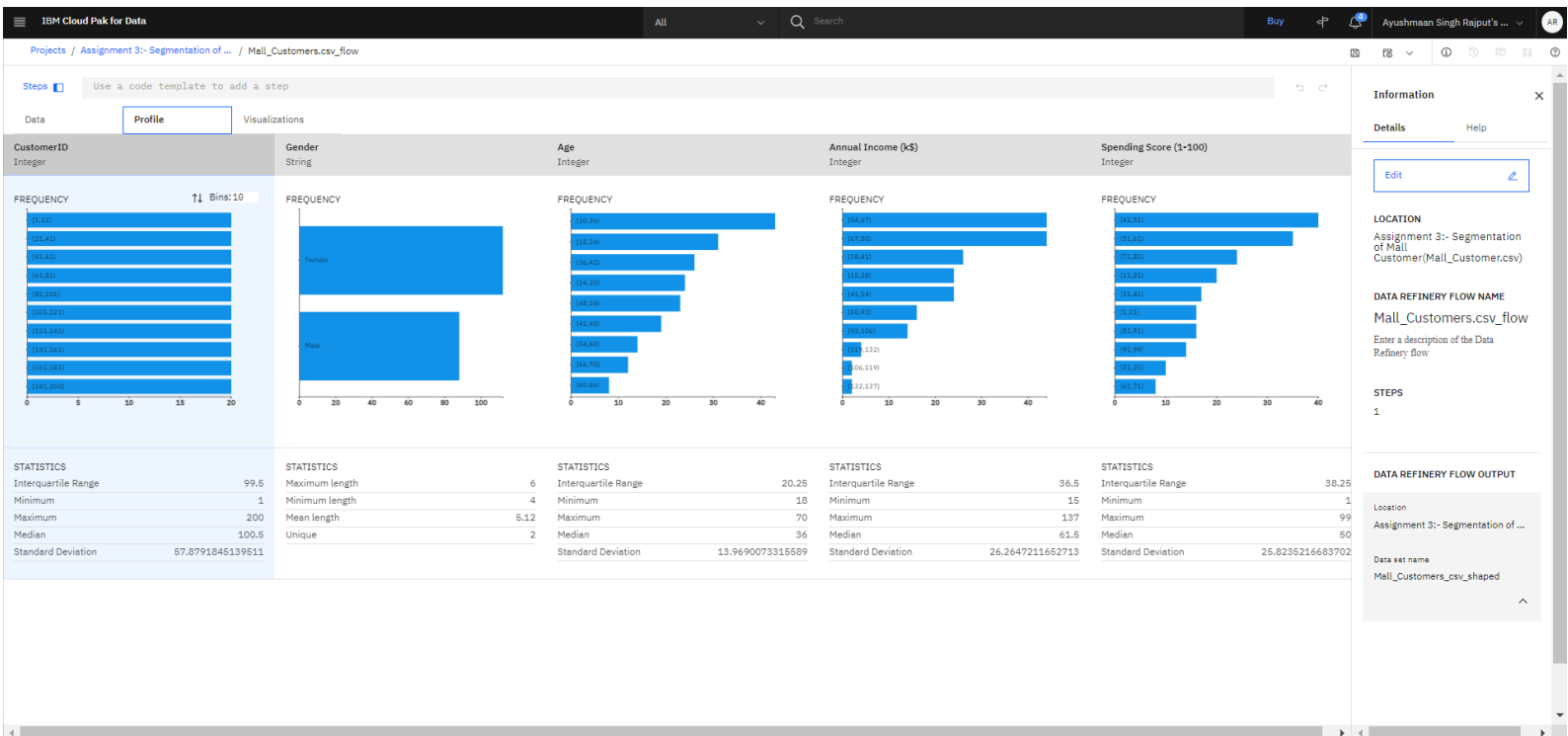


## ASSIGNMENT-3\_--SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306

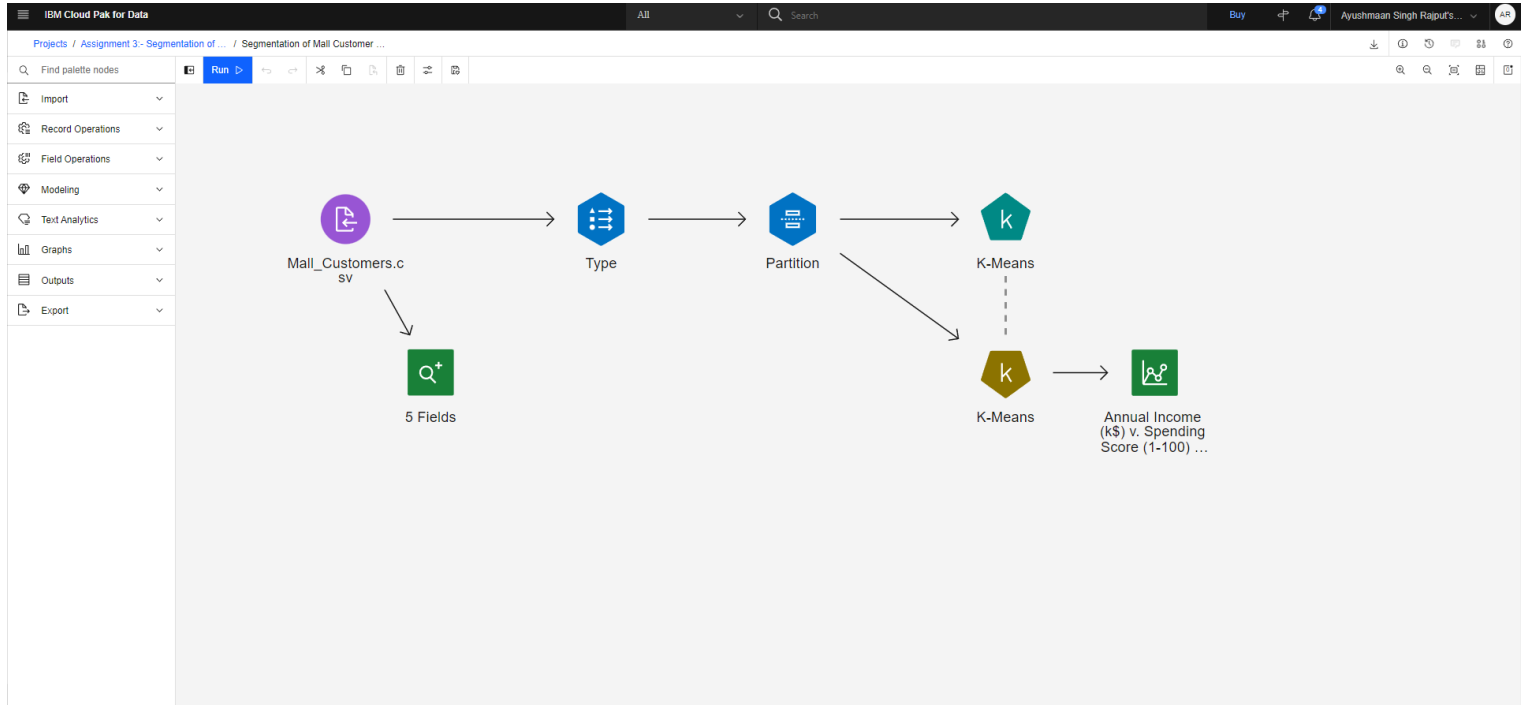
# Data Refinery Flow



## ASSIGNMENT-3\_--SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306

# SPSS Modular Flow



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

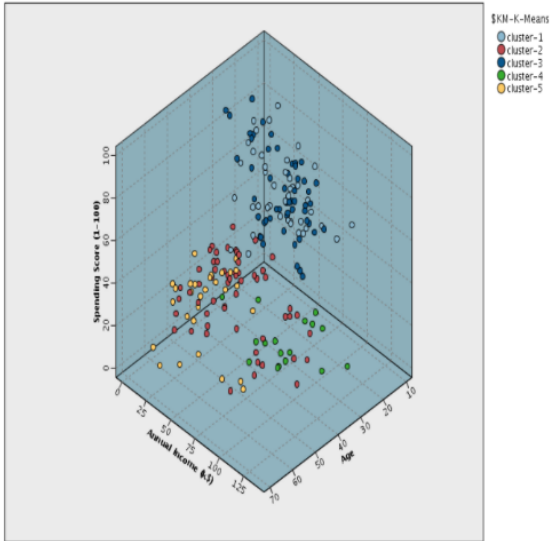
Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1 CustomerID		Continuous	1	200	100.500	57.879	0	--	200
2 Gender		Categorical	--	--	--	--	--	2	200
3 Age		Continuous	18	70	38.850	13.969	0.486	--	200
4 Annual Income (k\$)		Continuous	15	137	60.560	26.265	0.322	--	200
5 Spending Score (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	Blank Value
1 CustomerID	Continuous	0	0	None	Never	Fixed	100.000	200	0	0	0	0
2 Gender	Categorical	--	--	--	Never	Fixed	100.000	200	0	0	0	0
3 Age	Continuous	0	0	None	Never	Fixed	100.000	200	0	0	0	0
4 Annual Income (k\$)	Continuous	0	0	None	Never	Fixed	100.000	200	0	0	0	0
5 Spending Score (1-100)	Continuous	0	0	None	Never	Fixed	100.000	200	0	0	0	0

## ASSIGNMENT-3\_--SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306

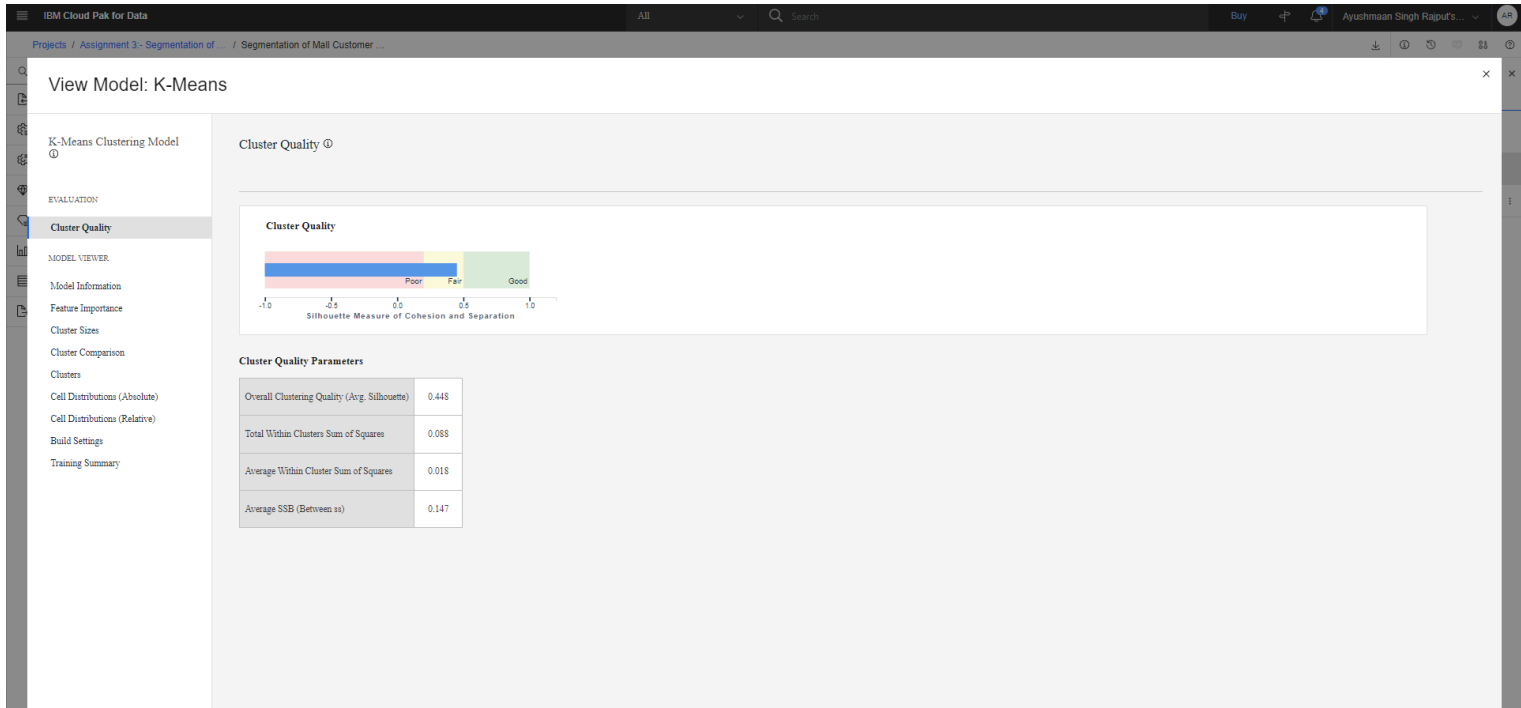
View Output: Annual Income (k\$) v. Spending Score (1-100) v. Age



## ASSIGNMENT-3--SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306

# View-Model



IBM Cloud Pak for Data

Projects / Assignment 3 - Segmentation of ... / Segmentation of Mall Customer ...

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

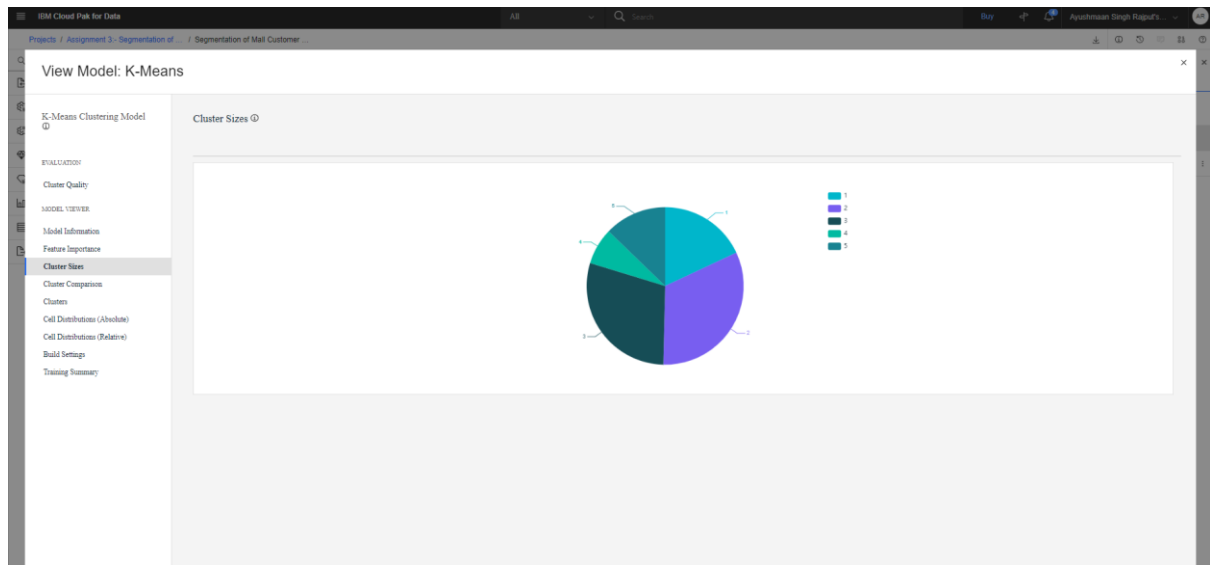
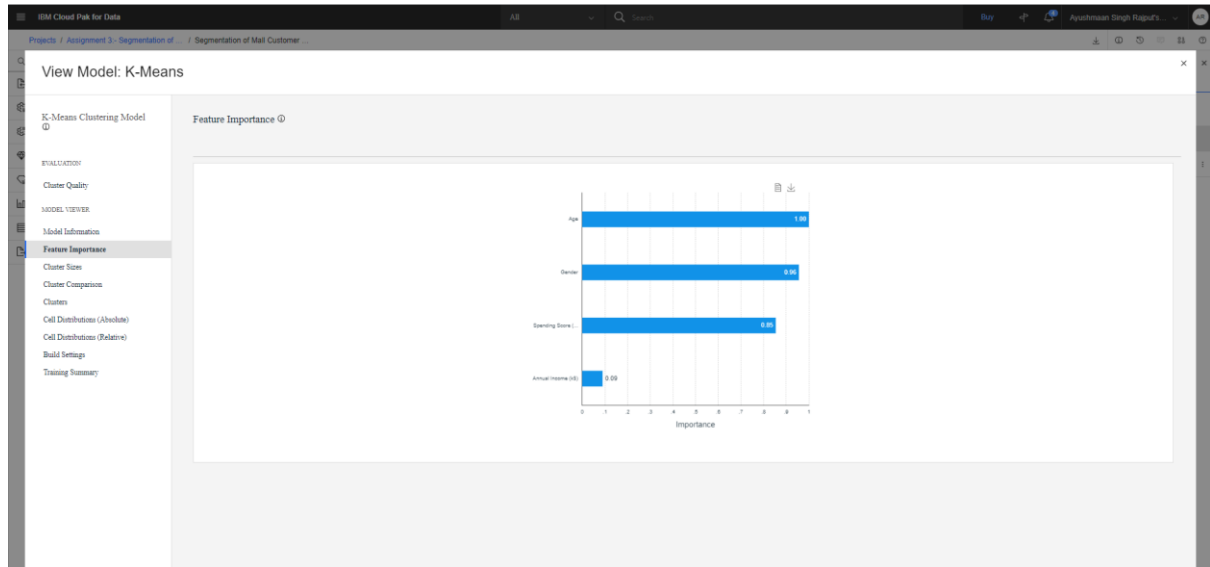
Training Summary

Model Information

Algorithm	K-Means	
Model Class	Center Based	
Number of Features	4	
Distance Measure	Euclidean	
Number of Clusters	5	
Number of instances in each cluster	Cluster 1	24 (18.05%)
	Cluster 2	43 (32.33%)
	Cluster 3	39 (29.32%)
	Cluster 4	10 (7.52%)
	Cluster 5	17 (12.78%)
Ratio of sizes (Largest to smallest)	4.300	

## ASSIGNMENT-3\_\_SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306



## ASSIGNMENT-3\_--SEGMENTATION-OF-MALL-CUSTOMER(MALL\_CUSTOMER.CSV)

Ayushmaan Singh Rajput  
19BCE10306

