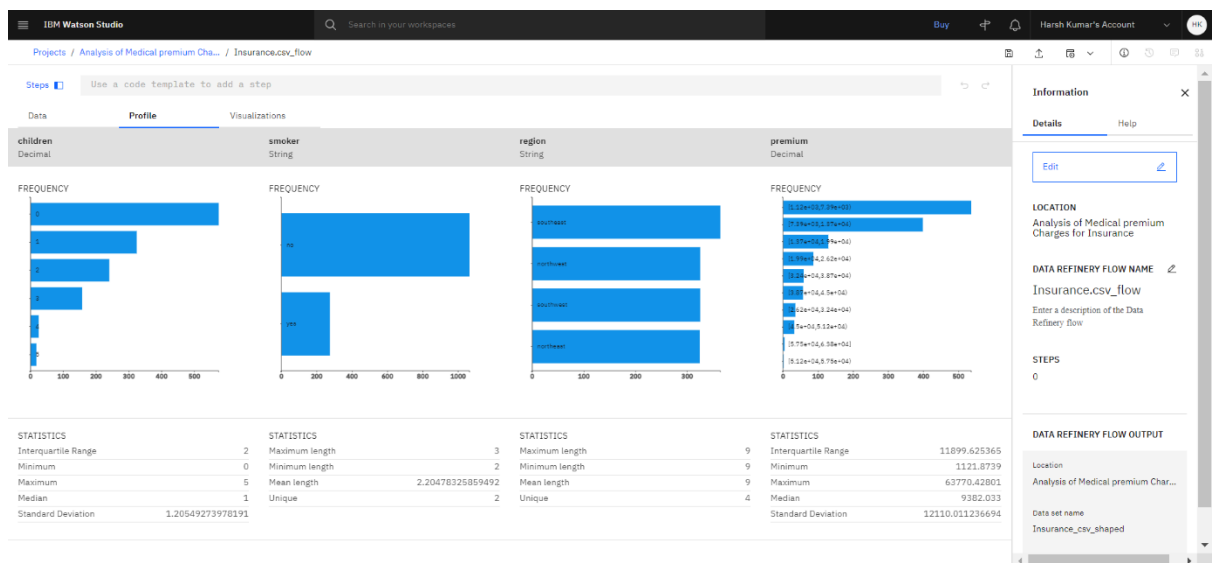
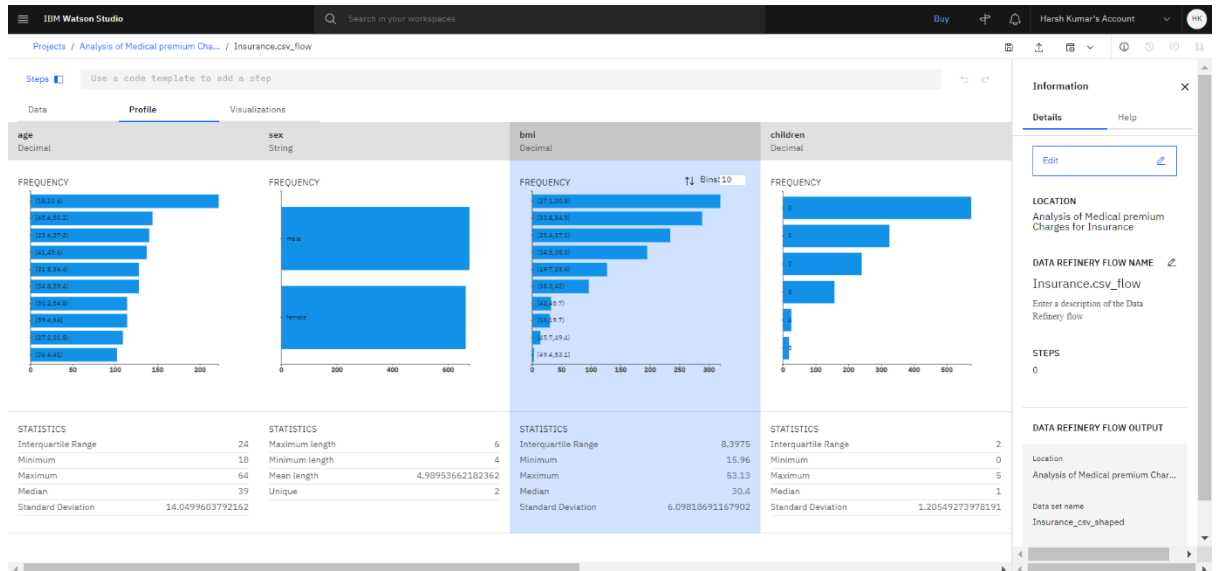


Data Analytics

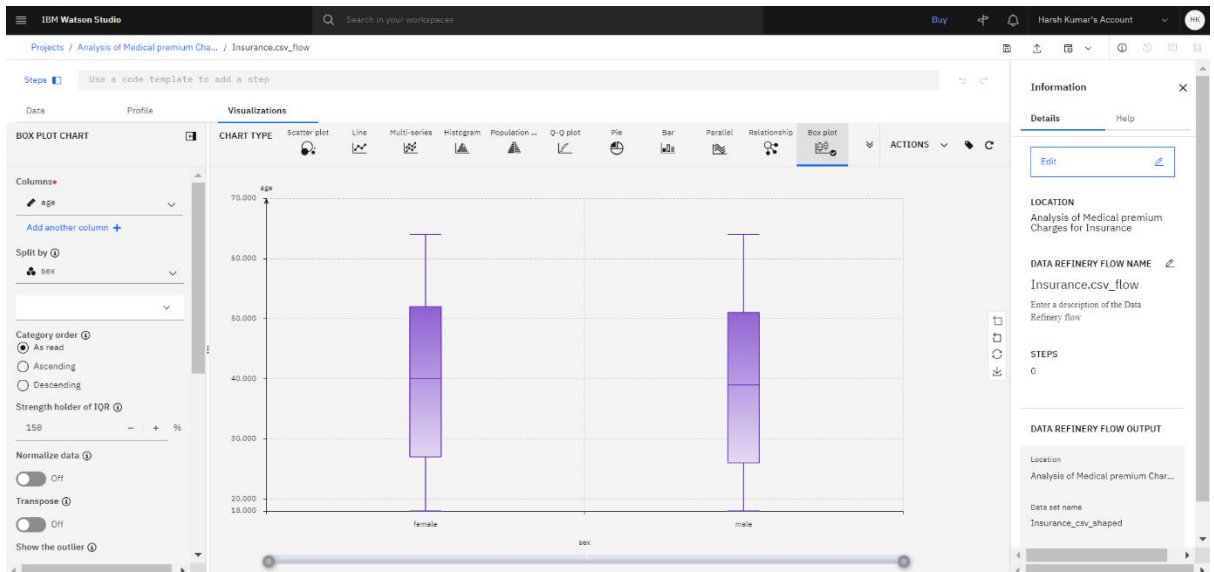
Case Study: Analysis of Medical Premium charges for insurers

Data Refinery

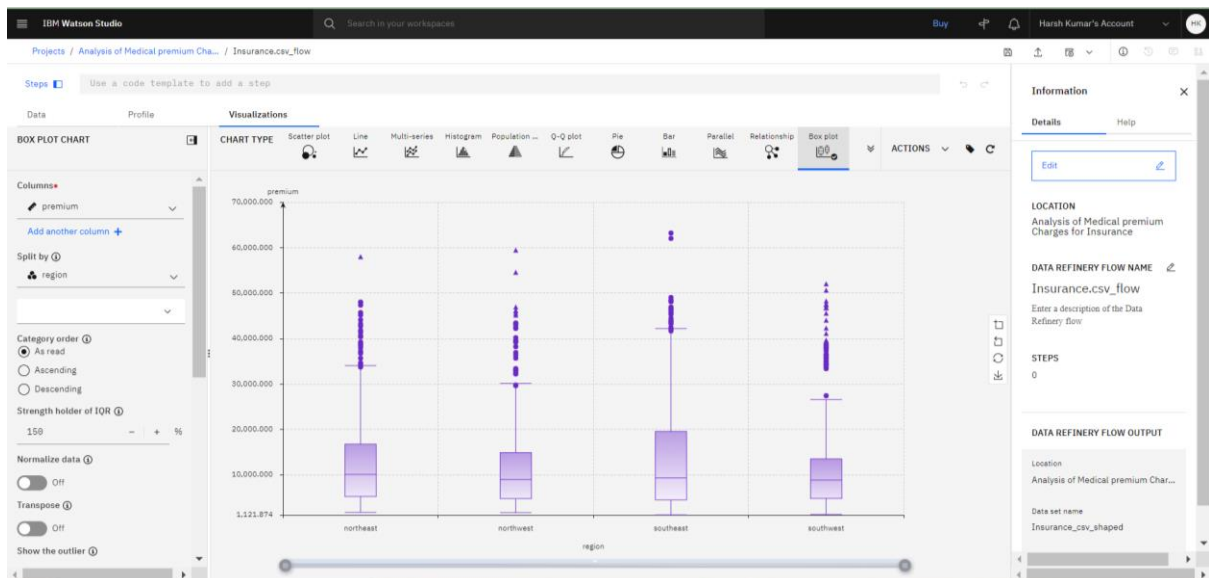
- Descriptive analytics of data



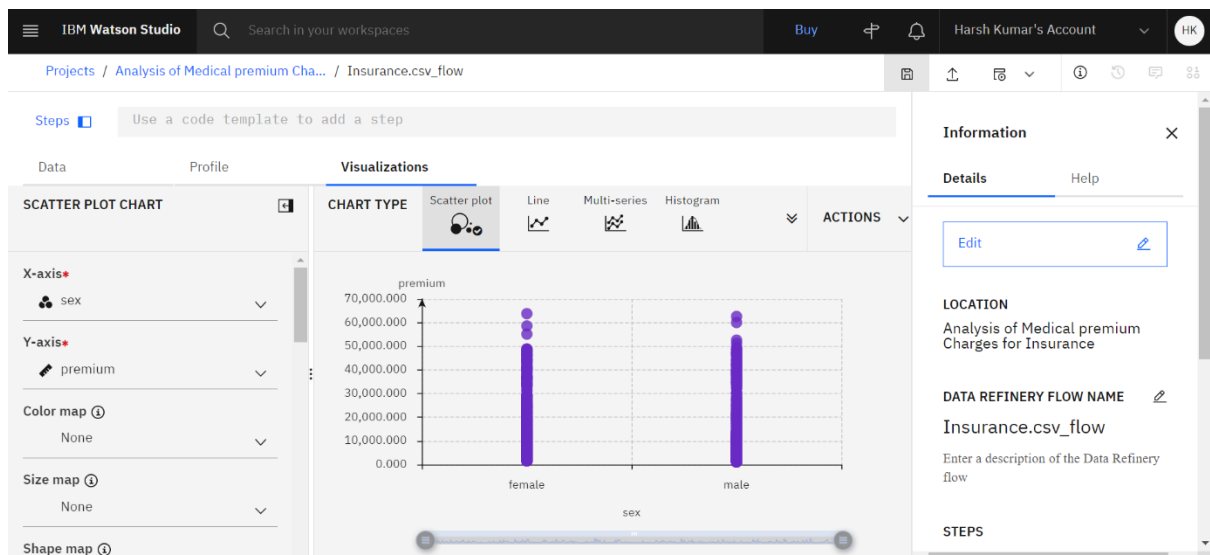
- Exploratory Analytics



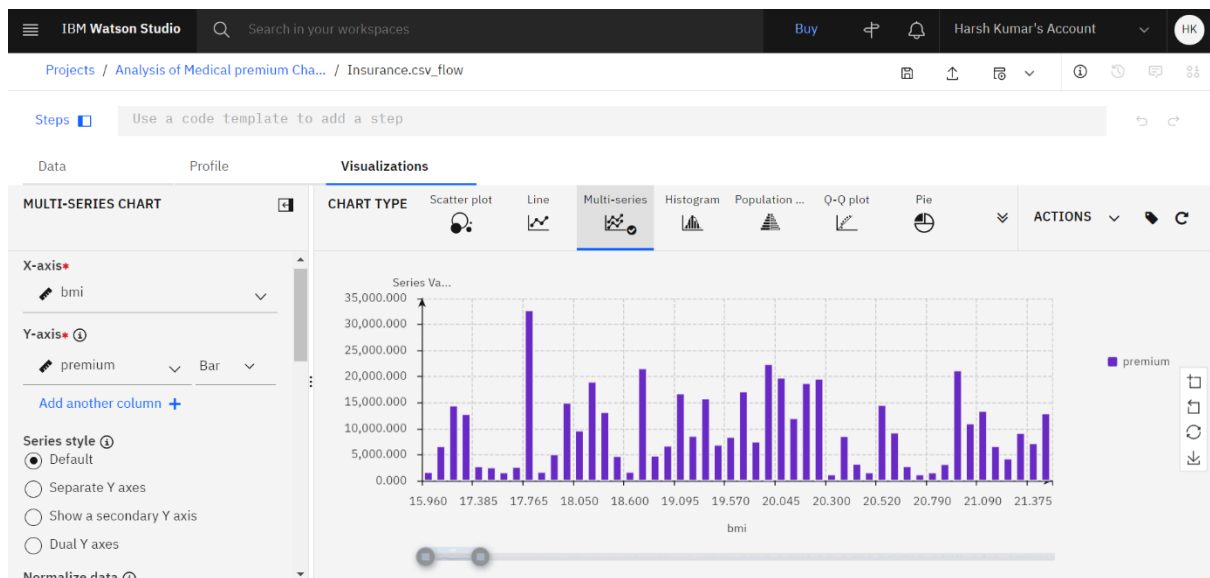
Columns-age ,split by-sex



Columns-premium ,split by-resion



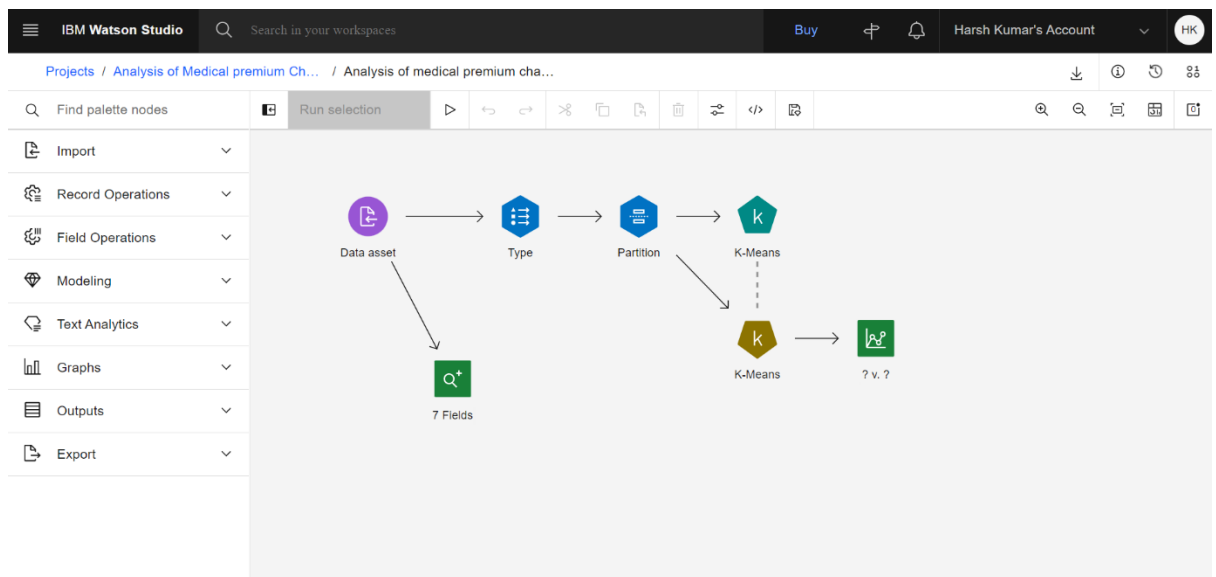
X-axis = sex , Y-axis = Premium



X-axis = bmi , Y-axis = Premium

2.SPSS Modeler

- Model



- **Type node**

BMI as none

premium as Output or target

other as input

The 'Type' node configuration panel in IBM Watson Studio shows a table of fields. The 'premium' field is highlighted as the target. The table lists the following fields:

Field	Measure	Role	Value mode	Values
# age	Continuous	Input	Instantiated	18, 64
sex	Flag	Input	Instantiated	female, male
bmi	Continuous	None	Instantiated	15.96, 53.13
# children	Continuous	Input	Instantiated	0, 5
smoker	Flag	Input	Instantiated	no, yes
region	Nominal	Input	Instantiated	northeast, northwest...
premium	Continuous	Target	Instantiated	1121 8739, 63770 4...

Default mode: ☒ Read metadata ☐ Pass (do not scan)

Set unique fields to None

Set categorical fields to None if they exceed this many values: 50

Set continuous integer field to ordinal if range less than or equal to: 10

Buttons: Cancel, Save

- **Partition node**

Training – 70%

Testing – 30%

IBM Watson Studio

Search in your workspaces

Buy

Harsh Kumar's Account

Projects / Analysis of Medical premium Ch... / Analysis of medical premium cha...

Find palette nodes

Run selection

Import

Record Operations

Field Operations

Modeling

Text Analytics

Graphs

Outputs

Export

Data asset

Type

Partition

7 Fields

Partition

Settings

Derived Field Name

Partition

Training Partition(%)

70

Testing Partition(%)

30

Create validation partition

Repeatability partition assignment

Cancel

Save

Model

IBM Watson Studio

Search in your workspaces

Buy

Harsh Kumar's Account

Projects / Analysis of Medical premium Ch... / Analysis of medical premium cha...

Find palette nodes

Run selection

Import

Record Operations

Field Operations

Modeling

Text Analytics

Graphs

Outputs

Export

Data asset

Type

Partition

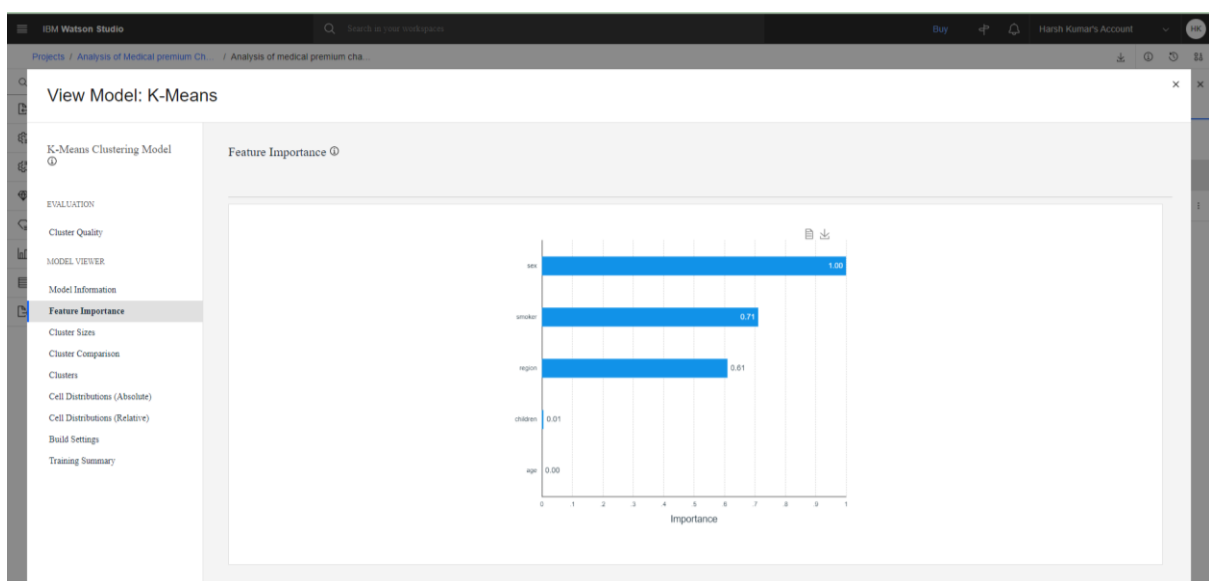
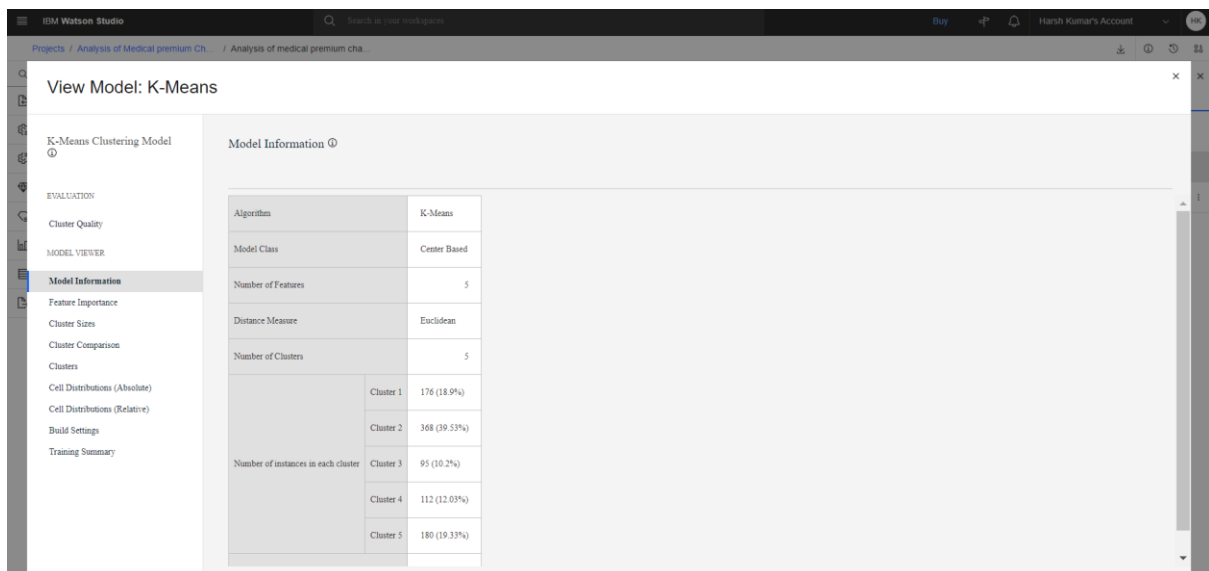
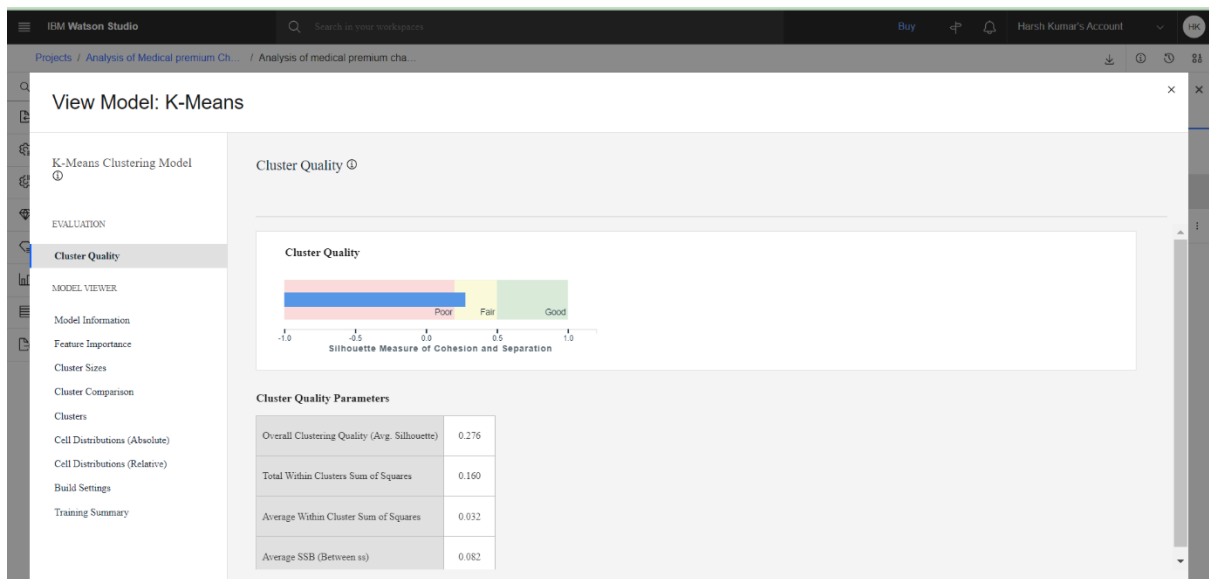
K-Means

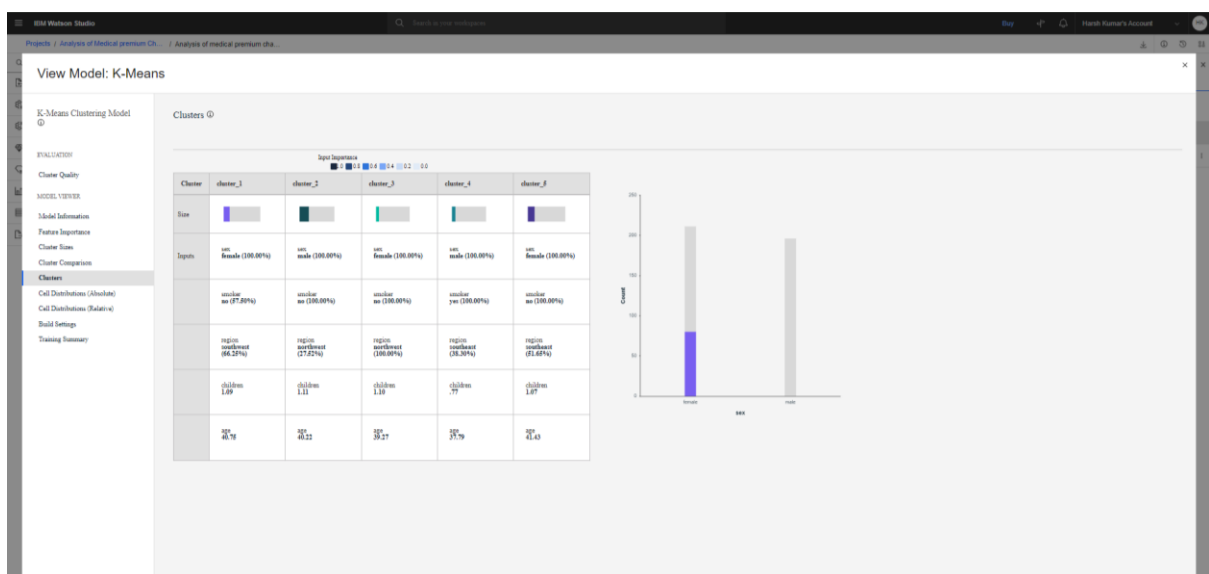
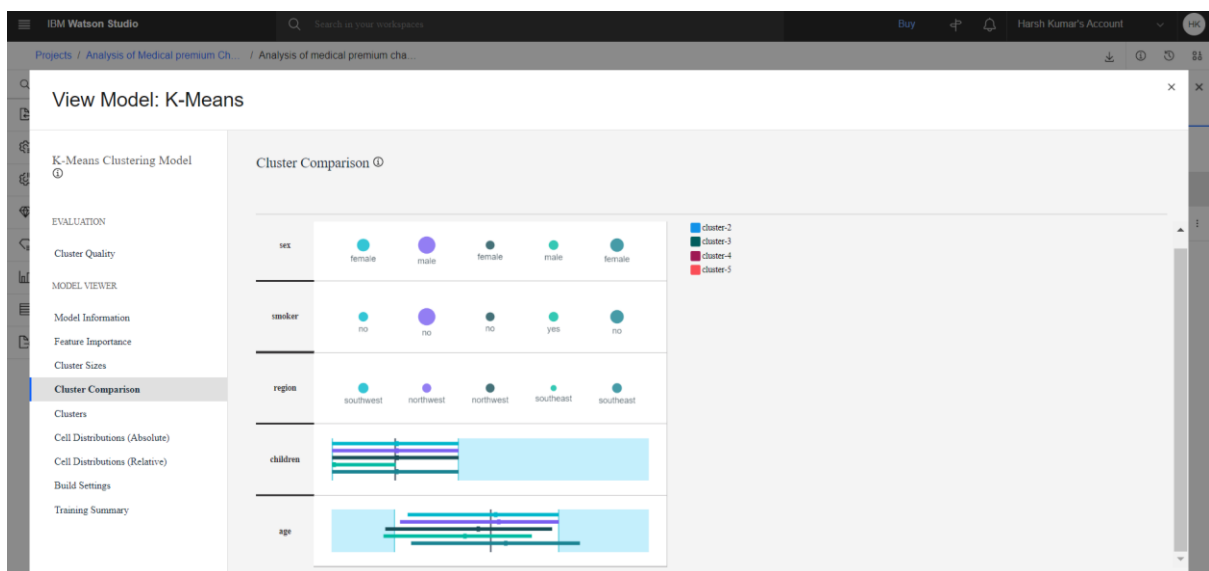
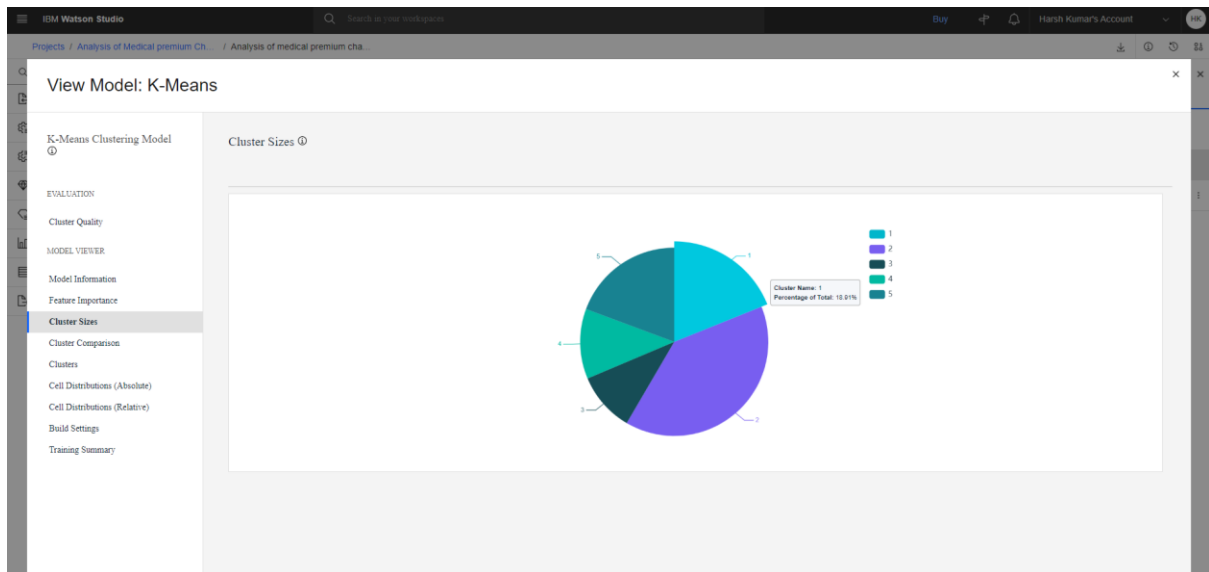
K-Means

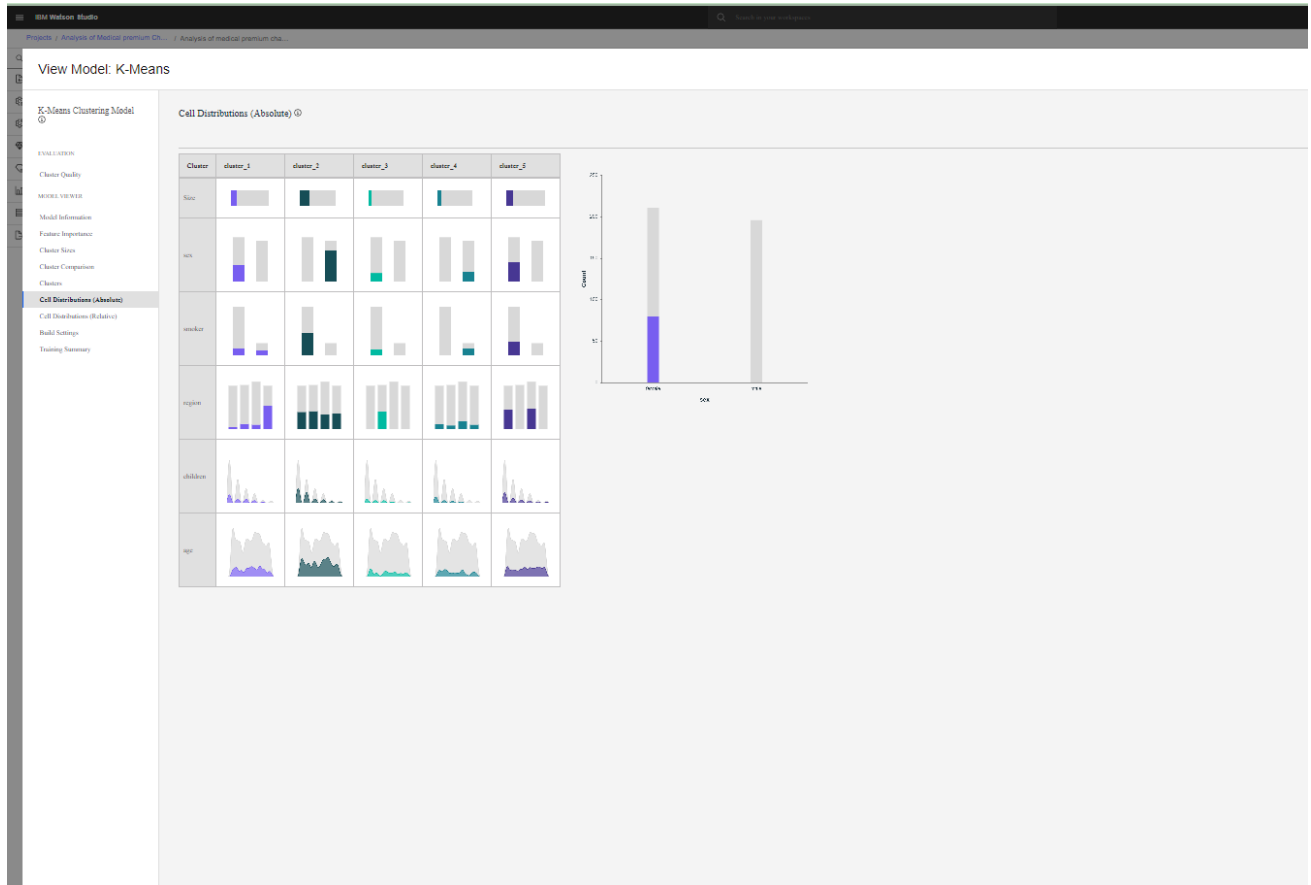
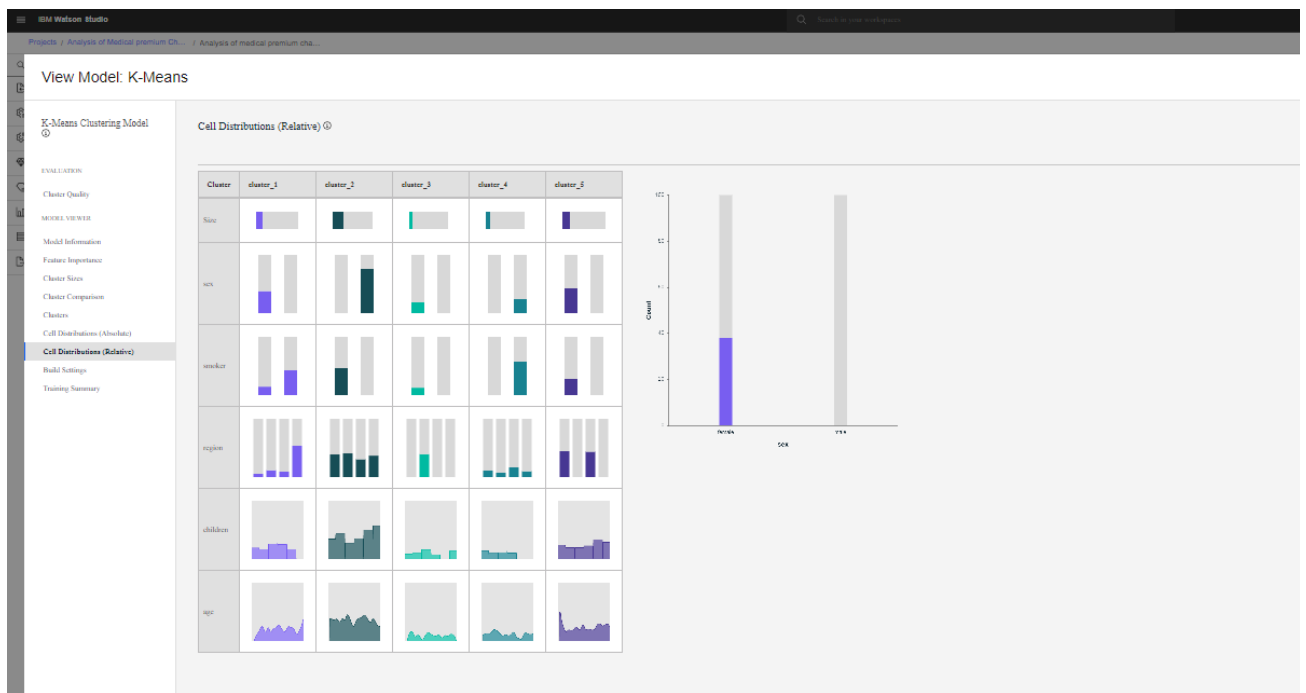
7 v. ?

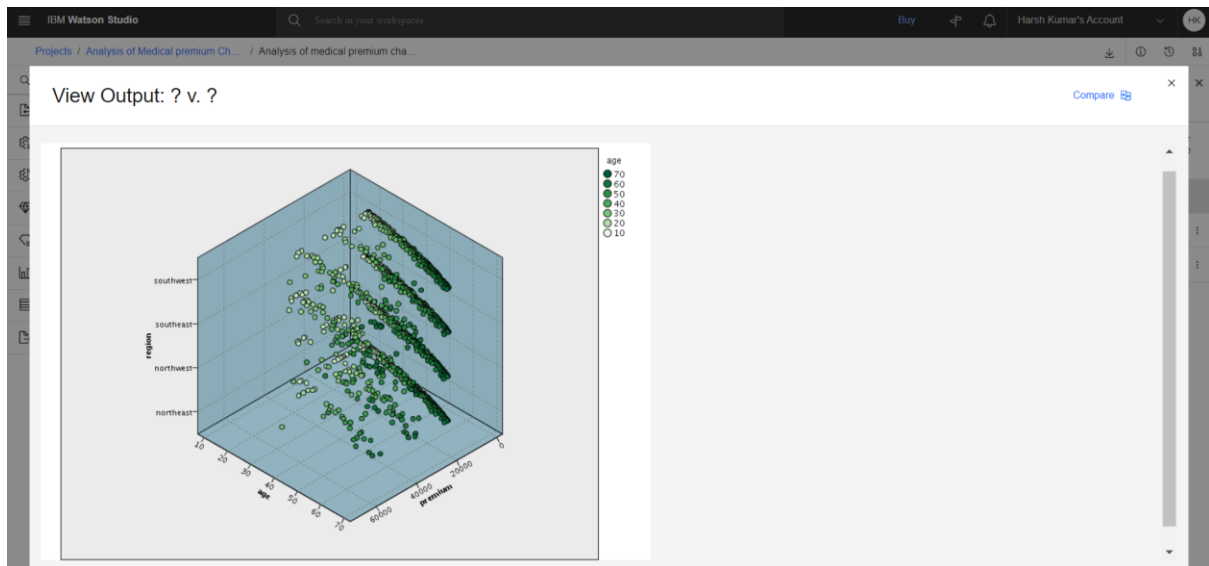
7 Fields

Output

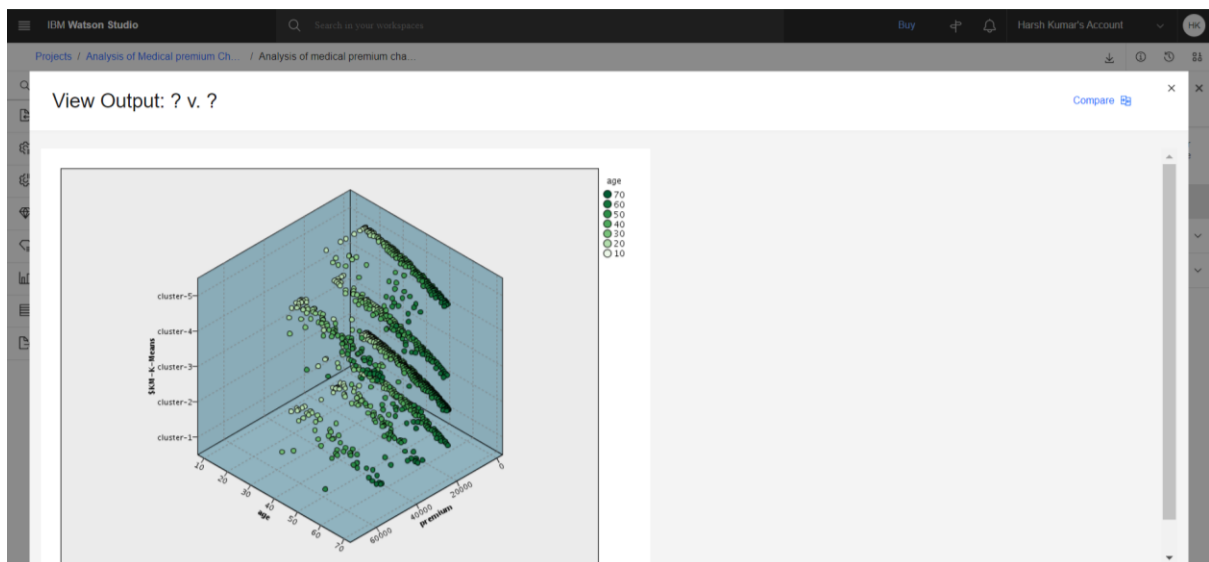




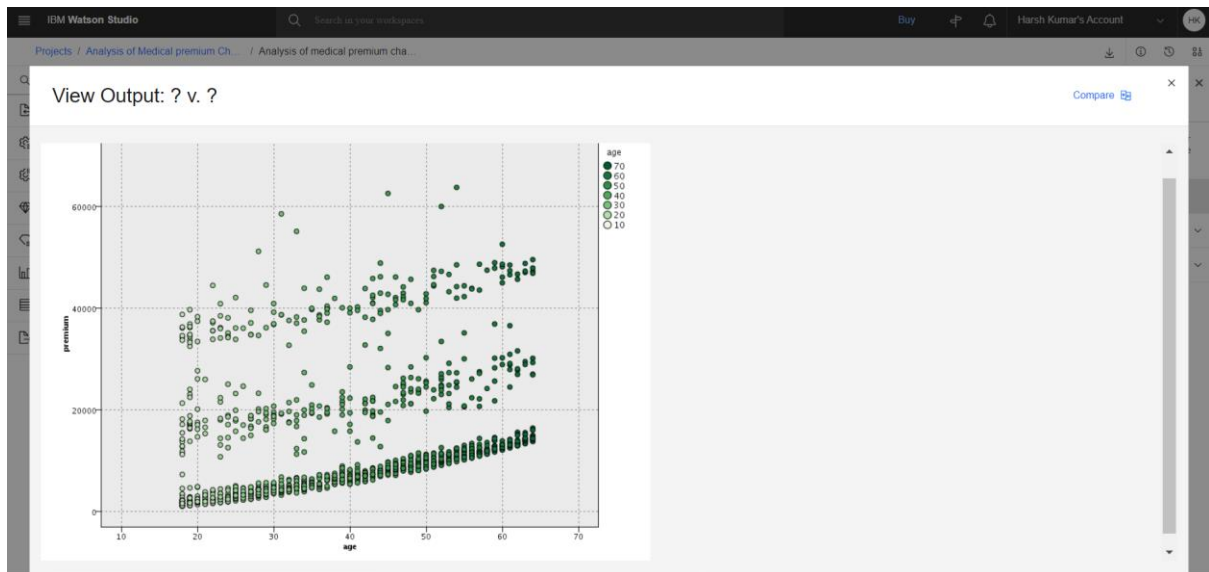




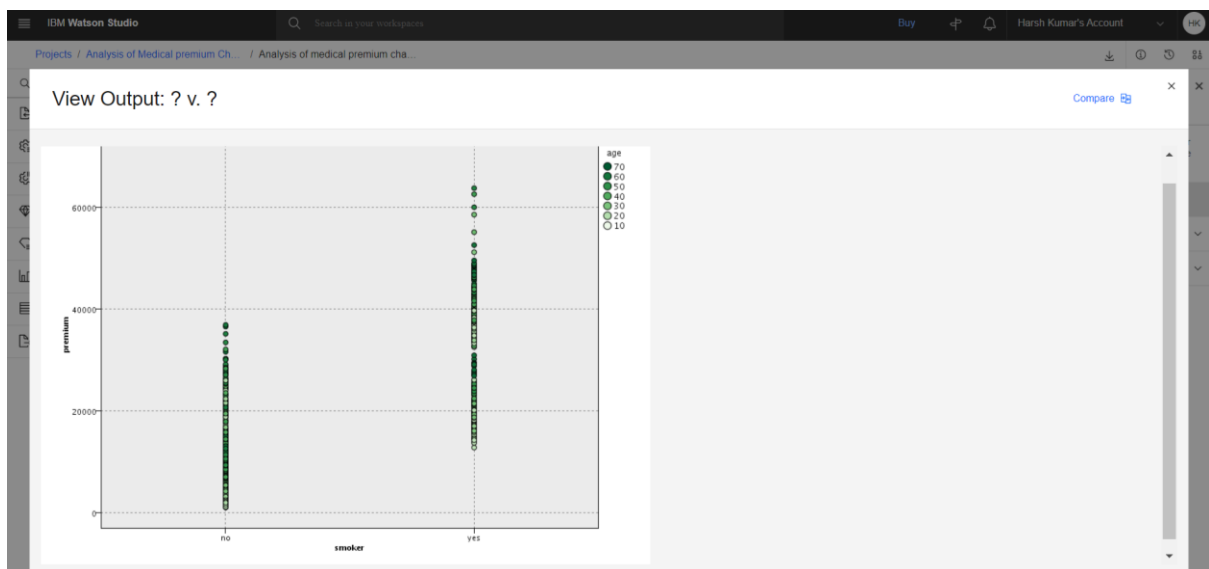
x-axis=age , y-axis=region ,z -axis=premium



x-axis=age , y-axis=\$KM-k-Means ,z -axis=premium



x-axis=age , y-axis=premium



x-axis=smoker , y-axis=premium

○ Data Audit

IBM Watson Studio

Search in your workspace

Buy

Harsh Kumar's Account

Projects / Analysis of Medical premium Ch... / Analysis of medical premium cha...

View Output: Data Audit of [7 fields] #4

Compare

	Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1	age		Continuous	18	64	39.207	14.050	0.056	--	1338
2	sex		Categorical	--	--	--	--	--	2	1338
3	bmi		Continuous	15.960	53.130	30.663	6.098	0.284	--	1338
4	children		Continuous	0	5	1.095	1.205	0.938	--	1338
5	smoker		Categorical	--	--	--	--	--	2	1338
6	region		Categorical	--	--	--	--	--	4	1338
7	premium		Continuous	1121.874	63770.428	13270.422	12110.011	1.516	--	1338

IBM Watson Studio

Search in your workspace

Buy

Harsh Kumar's Account

Projects / Analysis of Medical premium Ch... / Analysis of medical premium cha...

View Output: Data Audit of [7 fields] #4

Compare

	Field	Measurement	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value	Empty String	White Space	Blank Value
1	age	Continuous	0	0	None	Never	Fixed	100.000	1338	0	0	0	0
2	sex	Categorical	--	--	--	Never	Fixed	100.000	1338	0	0	0	0
3	bmi	Continuous	4	0	None	Never	Fixed	100.000	1338	0	0	0	0
4	children	Continuous	18	0	None	Never	Fixed	100.000	1338	0	0	0	0
5	smoker	Categorical	--	--	--	Never	Fixed	100.000	1338	0	0	0	0
6	region	Categorical	--	--	--	Never	Fixed	100.000	1338	0	0	0	0
7	premium	Continuous	7	0	None	Never	Fixed	100.000	1338	0	0	0	0