Q What assets are you looking for?

Data assets

New Data asset +

0assets selected.

Name	Type	Created by	Last modified
Csv insurance .csv	Data Asset	Sanyam Jain	Dec 22, 2021, 12:35 PM

Modeler flows

New Modeler flow +

Name	Type	Created by	Last modified
Medical Premium Charges	5PSSModeler	Sanyam Jain	Dec 22, 2021, 01 :57 PM

DataRefinery flows

New Data Refinery flow +

Name	Type	Created by	Last modified
insurance .csv flow	Data Refinery flow	Sanyam Jain	Dec 22, 2021, 01:16 PM

Steps @ Use a cccle I ei\\p at e ten ndd a ste|3

Data Profile Visualizations

	age String	sex String	bmi String	children String	smoker String	region String	premium String
1	19	female	27.9	0	yes	southwest	16884.924
2	18	male	33.77	1	no	southeast	1725.5523
3	28	male	33	3	no	southeast	4449.462
4	33	male	22.705	0	no	northwest	21984.47061
5	32	male	28.88	0	no	northwest	3866.8552
6	31	female	25.74	0	no	southeast	3756.6216
7	46	female	33.44	1	no	southeast	8240.5896
B	37	female	27.74	3	no	northwest	7281.5056
9	37	male	29.83	2	no	northeast	6406.4107
10	60	female	2584	0	no	northwest	28923.13692
11	25	male	26Z2	0	no	northeast	2721.3208
12	6Z	female	26Z9	0	yes	southeast	27808.7251
13	23	male	34.4	0	no	southwest	1826.843
14	56	female	39.82	0	no	southeast	11090.7178

Information

Details Help

Edit

LOCATION

I'fledical Premium Charges

DATA REFINERY FLOW NANE

insurance .csv_flow

Enter a description of the Dsta Refinery

STEPS

0

DATA REFINERY FLOw ourpur

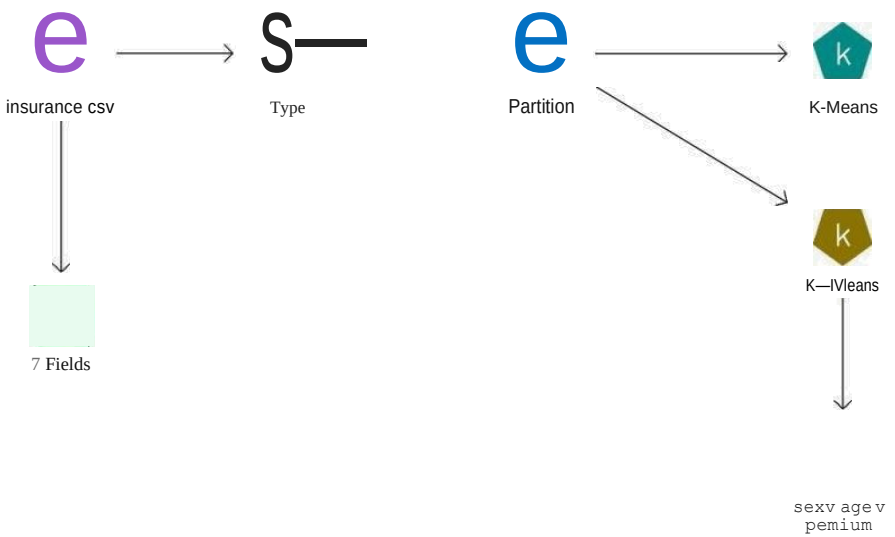
Location

Projects / Medical Premium Charges / Medical Premium Charges

Q Find palette nodes

Run

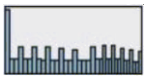
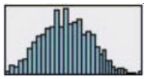
- Import v
- Record Operations v
- Field Operations v
- @ Modeling
- Text Analytics v
- @ Graphs v
- @ Outputs v
- Export v



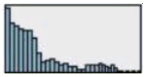
@ Find in column Field

	Field	Measure	Role		Value mode		Values	Check	
☐	# age	Continuous	Input	v	Instantiated	v	18, 64	None	v
☐	abc sex	Flag	Input		Instantiated		female, male	None	@
☐	#₋# bmi	Continuous	Input		Instantiated		15.96, 55.13	None	@
☐	@ children	Continuous	Input		Instantiated		0, 5	None	@
☐	a ..smoker	Flag	Input		Instantiated		no, yes	None	@
☐	a region	Nominal	Input		Instantiated		northeast, northwe...	None	
☐	premium	Continuous	Input		Instantiated		1121.8739, 63770....	None	@

View Output: Data Audit of [7 fields]

	Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1	age		Gontinuous	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								4	1338

View Output: Data Audit of [7 fields]

6	region		Categorical						4	1338
7	premium		Continuous	1121.874	63770.428	13270.422	12110.011	1.516		1338

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

View Model: K-Means

K-Means Clustering Model

EVALUATION

Cluster Quality

hIODEL \'IEVUR

Model Information

Feature Importance

Cluster Sizes

Cluster Comparison

Clusters

Cell Distribuhons (Absolute)

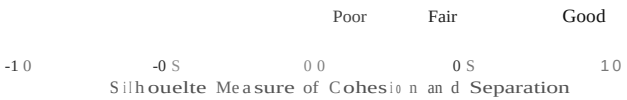
Cell Distribtihons (Relative)

Build Settings

TrauipSusuig

Cluster Quality O

ClusterQuality



Cluster Quality Parameters

Overall Clustering Quality (Avg. Silhouette)	0.249
Total Within Cluatera Sum of Squares	0.132
Average Within Cluster Simi of squares	0.026

View Model: K-Means

K-Means Clustering Model
ⓘ

EVALUATION

Cluster Quality

ODEL\IEIVER

Model Information

Feature Importance

Cluster Sizes

Cluster Comparison

Clusters

Cell Diatribuhons (Absolute)

Cell Distribuhons (Relative)

Build Settings

Tra g Summary

Cluster Quality O

Poor Fair Good

-1 0 -0 5 0 0 0.S 1 0

Silhoue1te Measure of Cohesi0 n and Separation

Cluster Quality Parameters

Overall Clustering Quality (Avg. Silhouette)	0.249
Tod tMnCufineSxDO Lqum€B	0132
Average Within Cluster Simi of squares	0 026
AvemgeSSB m)	0.075

View Model: K-Means

K-Means Clustering Model

EVALUATION

Cluster Quality

NIODEL A'mUUR

Model Information

Feature Importance

Cluster Sizes

Clriater Comparison

Clusters

Cell Diatribiihons (Abeolute)

Cell Distribuhons (Relative)

Build Settings

Training Summary

Model Information O

igodflun	K-Menus
Model Class	Center Based
Hianber of Features	7
Distance Measure	Euclidean
Ninnber of Clusters	5

Cluster 1	81 (8.7%)
Cluster 2	368 (39.53%)

View Model: K-Means

K-Means Clustering Model

ⓘ

EVALUATION

Clriater Quality

ODELADAIR

Model Information

Feattu< Importance

Clriater Sizes

Cluster Comparison

Clusters

Cell Distribtihons (Absolute)

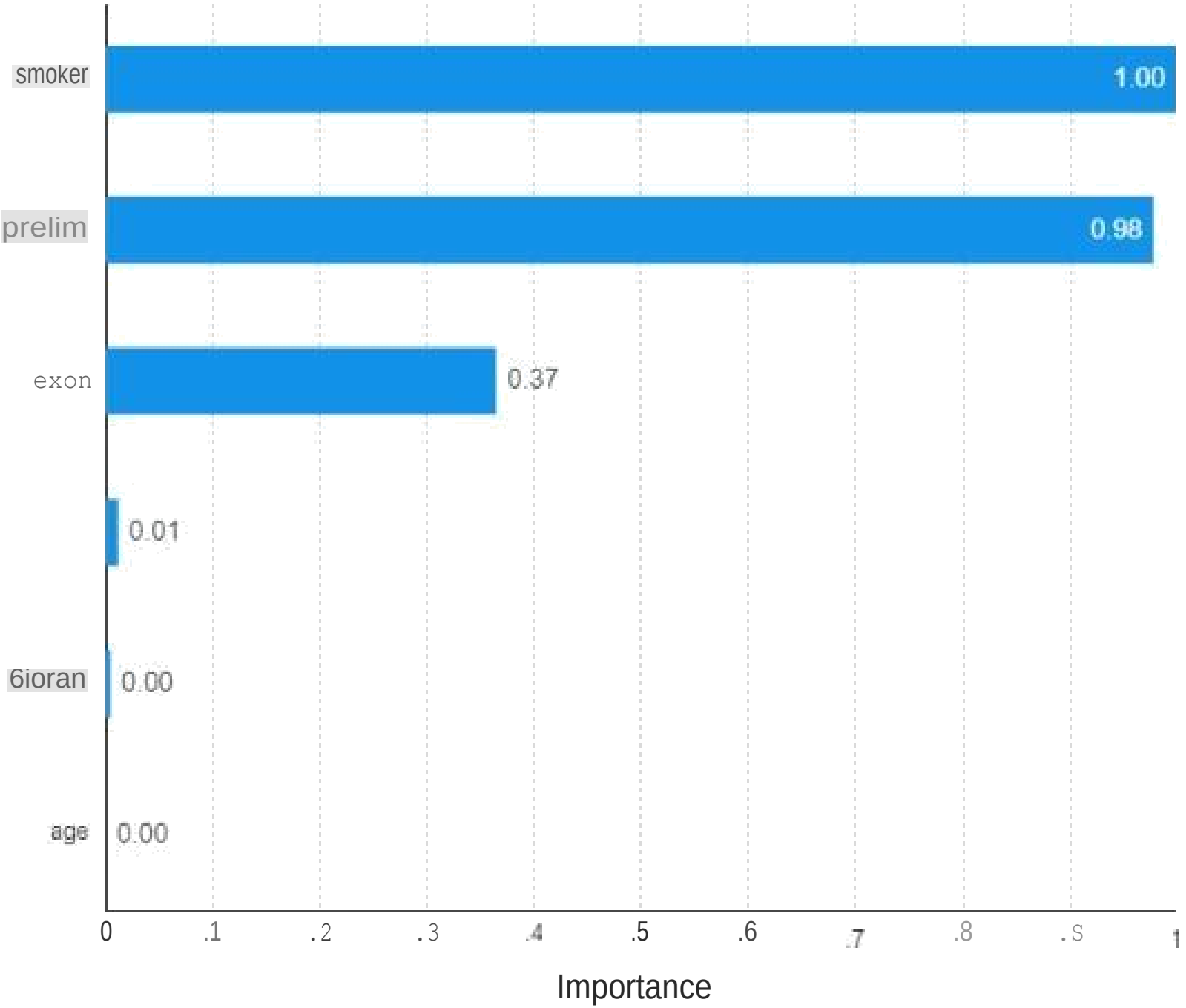
Cell Distribtihons (Relative)

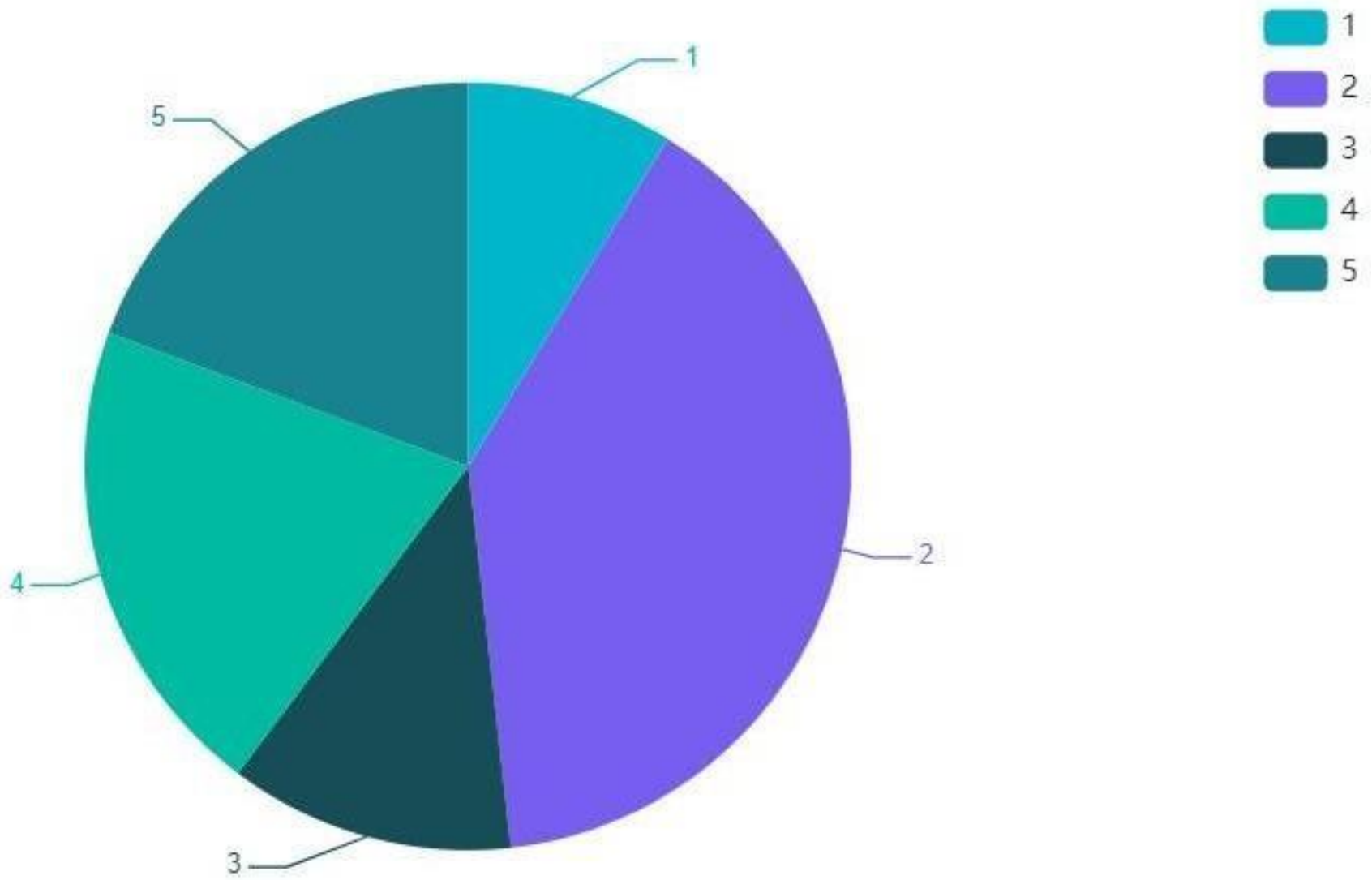
Build Setting

Training Summary

Model Information O

Ninnber of Clusters		5
Ninnber of instances in each cluster	Cluster 1	81 (8.7%)
	Cluster 2	368 (39.53%)
	Cluster 3	112 (12.03%)
	Cluster 4	190 (20.41%)
	Cluster 5	180 (19.33%)
Ratio of sizes (Largest to smallest)		4.543


smoker
1.00
prelim
0.98
exon
0.37
0.01
6ioran
0.00
age
0.00
0
.1
.2
.3
.4
.5
.6
.7
.8
.S
1
Importance


1
2
3
4
5
1
2
3
4
5

View Model: K-Means

K-Means Clustering Model
ⓘ

EVALUATION

Cluster Quality

ODEL\JI?IR

Morlel Information

Feattu< Importance

Cluster Sizes

Cluster Cozaparison

Clusters

Cell Distributions (Absolute)

Cell Distribtihoiis (Relatix e)

Build Settings

Cluster Comparison O

View Model: K-Means

K-Means Clustering Model
ⓘ

EVALUATION

Clri8ter Quality

N4ODEL \WU K

Model Information

FmhueI‹ npo2ax

Clri8ter Sizes

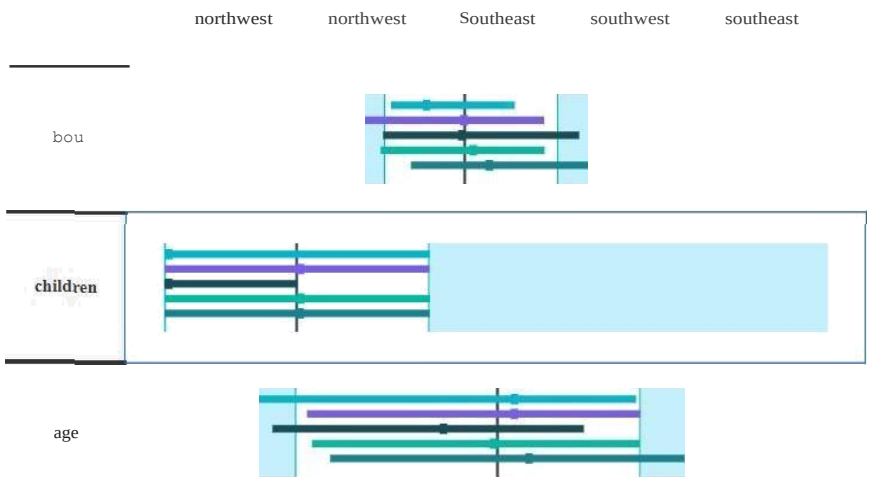
ClRsterComparLoo

Clusters

Cell Distribtihons (Absolute)

Cell Distribtihom (Relative)

Cluster Comparison O



View Model: K-Means

K-Means Clustering Model
0

Clusters O

EVALUATION

Cluster Quality

hIODEL \W \\OR

Model Information

Feature Importance

Cluster Sizes

Cluster Comparison

Clusters

Cell Distributions (Absolute)

Cell Distributions (Relative)

Build Settings

&< Inn °
.0 0.8 0.6 0.4 0.2 0.0

Cluster	c1uster_1	clnster_2	**cluster 3**	**cluster 4**	cluster_5
Size					
Inputs	SQL female (100.00%)	SQL male (100.04%)	SQL male (100.00%)	SQL female (100.00%)	SQL female (100.00%)
	snooker *Yes (100.00%)	smoker no (100.00%)	SInOkL *'es (100.00%)	smoker **no (100.00%)**	smoker **no (100.00%)**

250
200
450
i00

View Model: K-Means

K-Means Clustering Model
ⓘ

Clusters O

EA LCATION

4nput Importance

0 08 06 04 02 00

Cluster Quality

NIODEL AHVUR

h4odel Information

Feature Importance

Cluster Sizes

Cluster Comparison

Clusters

Cell Distribtihons (Absolute)

Cell Distribtihons (Relative)

Build Settings

Cluster	**cluster_1**	cInster_2	**clnster_3**	**cluster_4**	**cluster_5**
	premium 30921.55	premium 8866.62	p/ j \jm 31714.55	premium 8191.76	iiwrniuin 8842.25
	region northwest (35.29%)	repion szDrtbvest (27.52%)	region southeast (38J0%)	region southwest **(53.49%)**	region southeast (51.65%)
	bmi 29.95	bmi 30.52	bmi 30.70	bmi 30.39	bmi 31.77

250
200
150
100
C

View Model: K-Means

K-Means Clustering Model
0

Clusters O

EVALUATION

Cluster Quality

ñlODEL \WWT-R

Model Inforiiiation

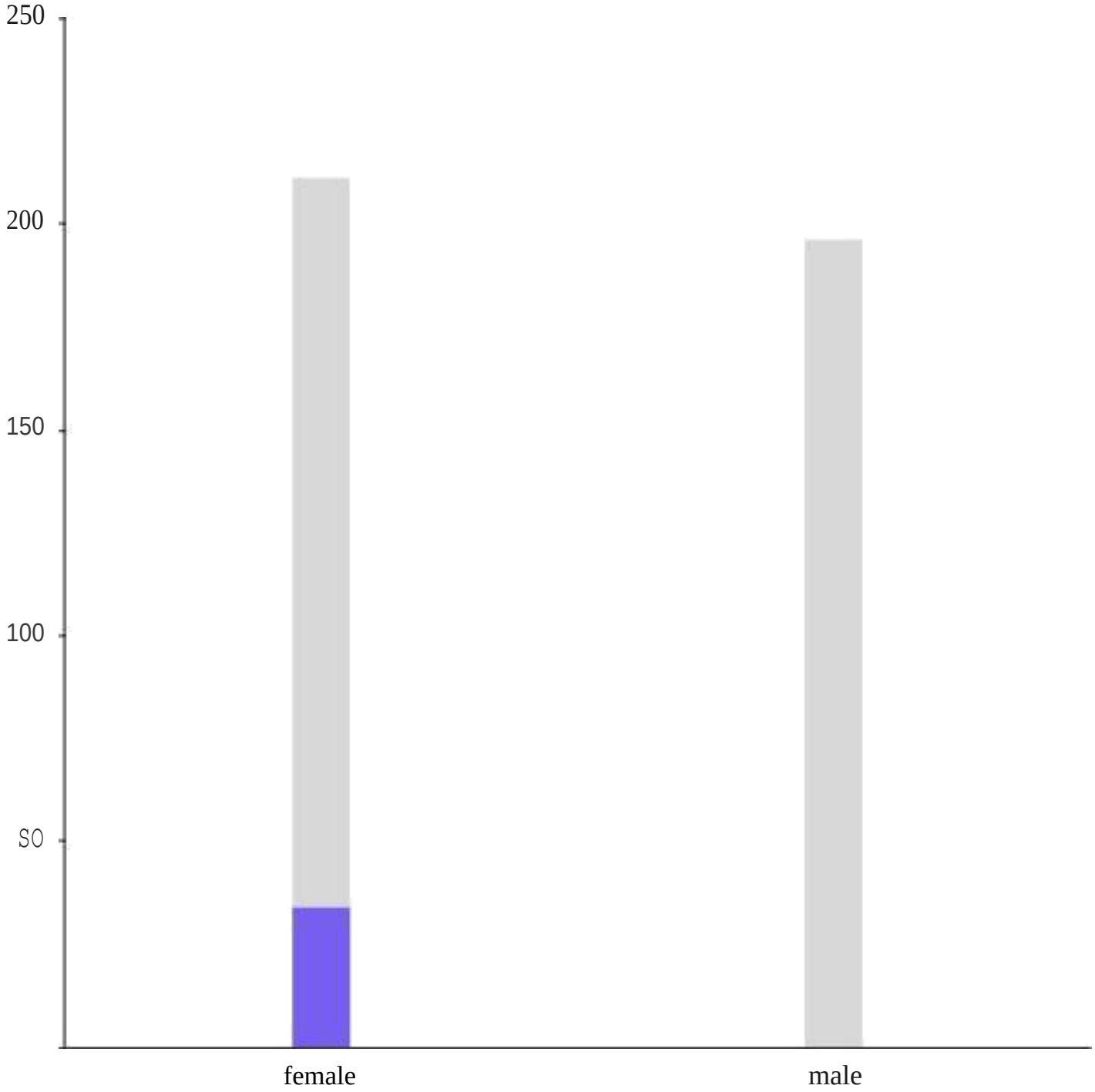
feature Importance

Cluster Sizes

Cluster Comparison

Cell Distribuhous (Absolute)

	bmi 29.95	bmi 30J2	bmi 30.70	bmi 30J9	bA 3*77
	children .91	children **1.11**	children .77	children **1.16**	children 1.07
	ape 39.65	ape 40.22	age 37.79	ape 40.50	age 41.43


250
200
150
100
50
female
male

View Model: K-Means

K-Means Clustering Model

Cluster Quality

MODEL VIEWER

Model liifoniiatioii

Feature liuportance

Cluster Sizes

Cluster Comparison

Clusters

Cell Distributions (Absolute)

Cell Distributions (Relative)

Build Settings

Cell Distributions (Absolute) O

View Model: K-Means

K-Means Clustering NIode1

Cell Distributions (Absolute) O

Cluster Quality

NIODEL \W\\T-R

Model Information

Featiue Importance

Cluster Sizes

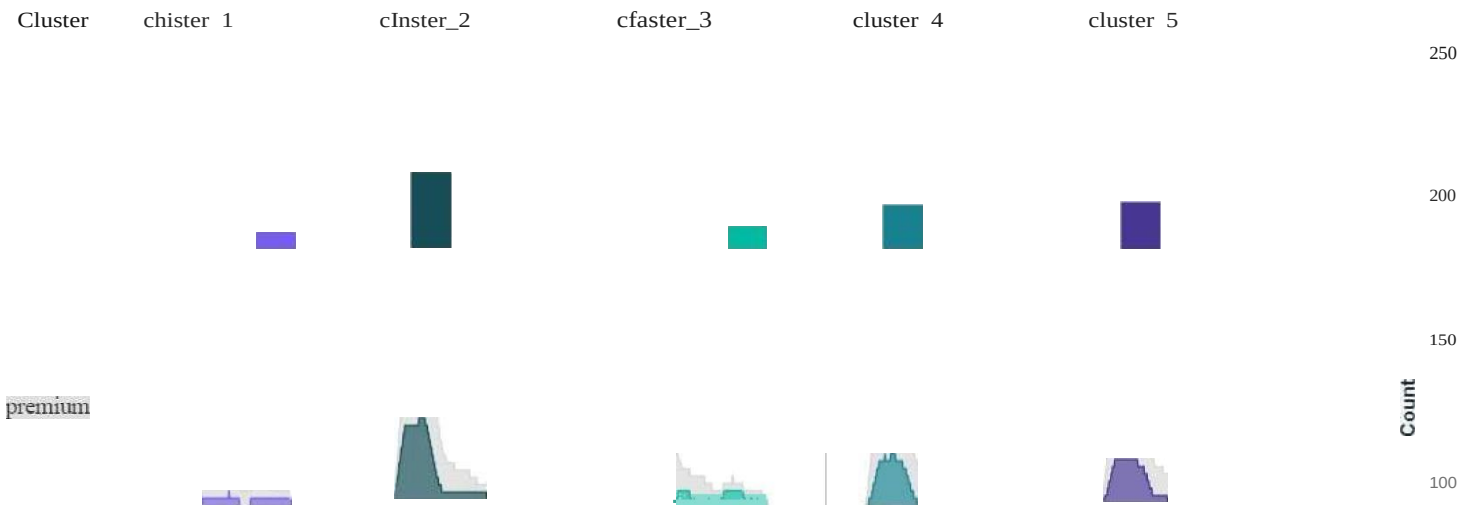
Cluster Comparison

Clusters

Cell Distributions (Absolute)

Cell Distributors (Relatix e)

Build Settings



View Model: K-Means

K-Gleans C mustering klodel

tell Distributions (Absolute) O

Cluster Quality

?1ODEL\'O Il

h4orlel In oiiiiation

FeaRu< Iriiportance

Cluster Sizes

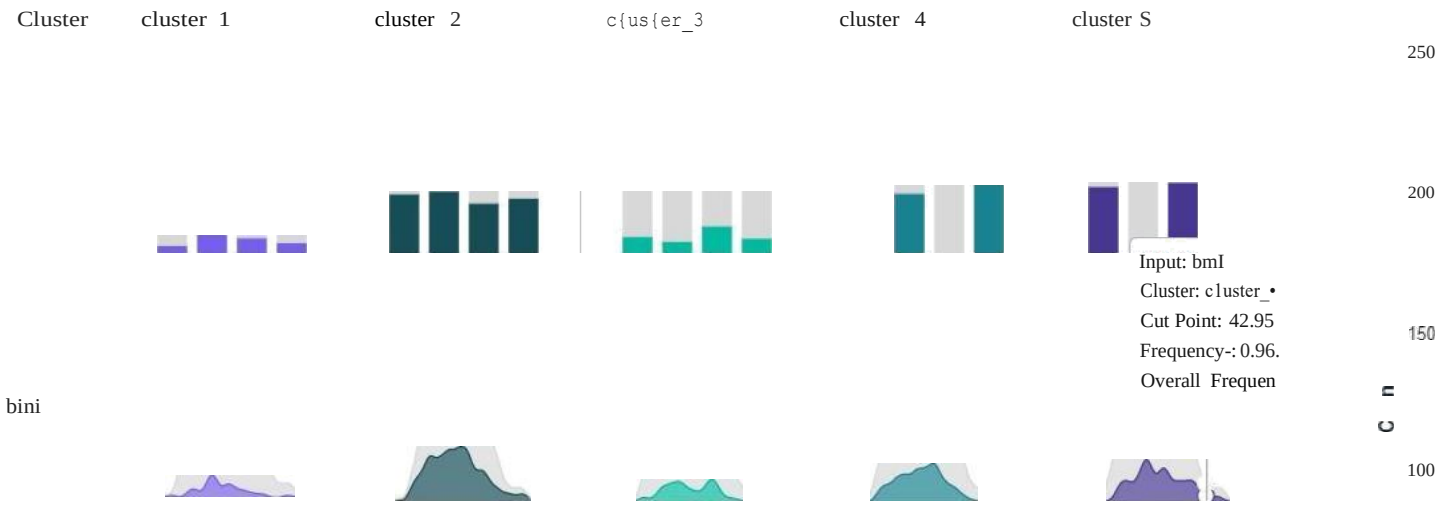
Cluster Comparison

Clusters

Cell Distributions (Absolute)

Cell Distribiitioiis (Relative)

Build Settings



View Model: K-Means

K-Means Clustering Model

tell Distributions (Absolute) O

Cluster Quality

ODELAIE Il

Morlel Information

Feattu< Importance

Cluster Sizes

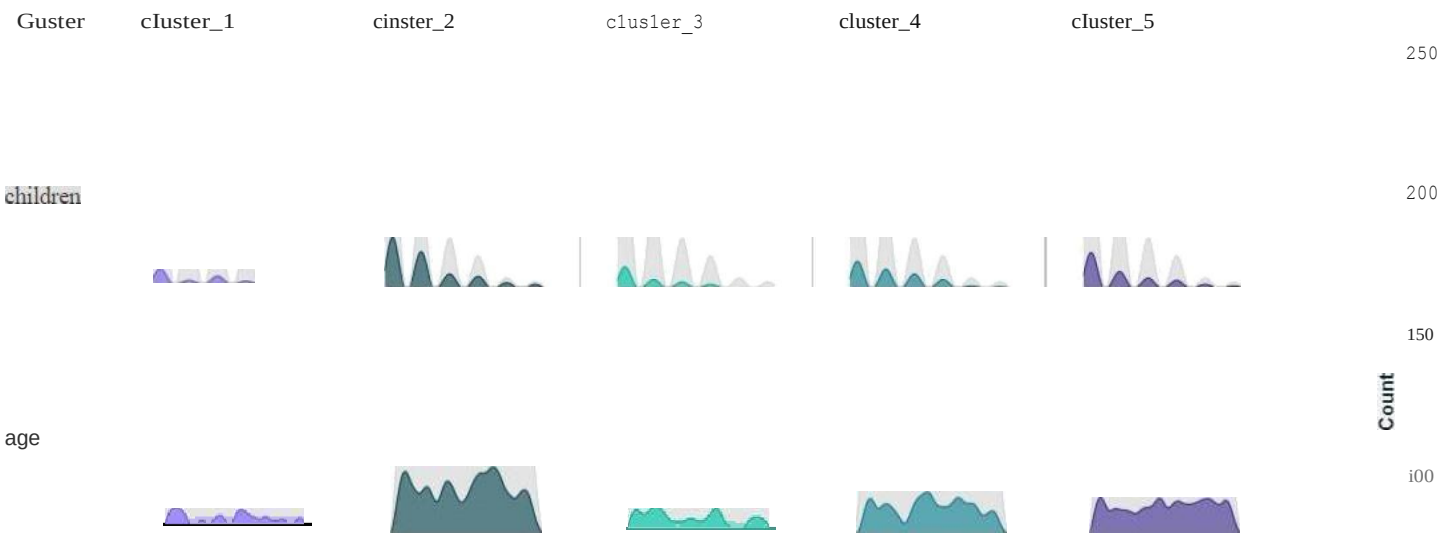
Cluster Comparison

Clusters

Cell Distributions (Absolute)

Cell Distribiihons (Relative)

Build Settings



View Model: K-Means

K-Means Clustering Model

Cell Distributions (Relative) O

Cluster Quality

hIODEL \WVYR

Model Information

Feature Importance

Cluster Sizes

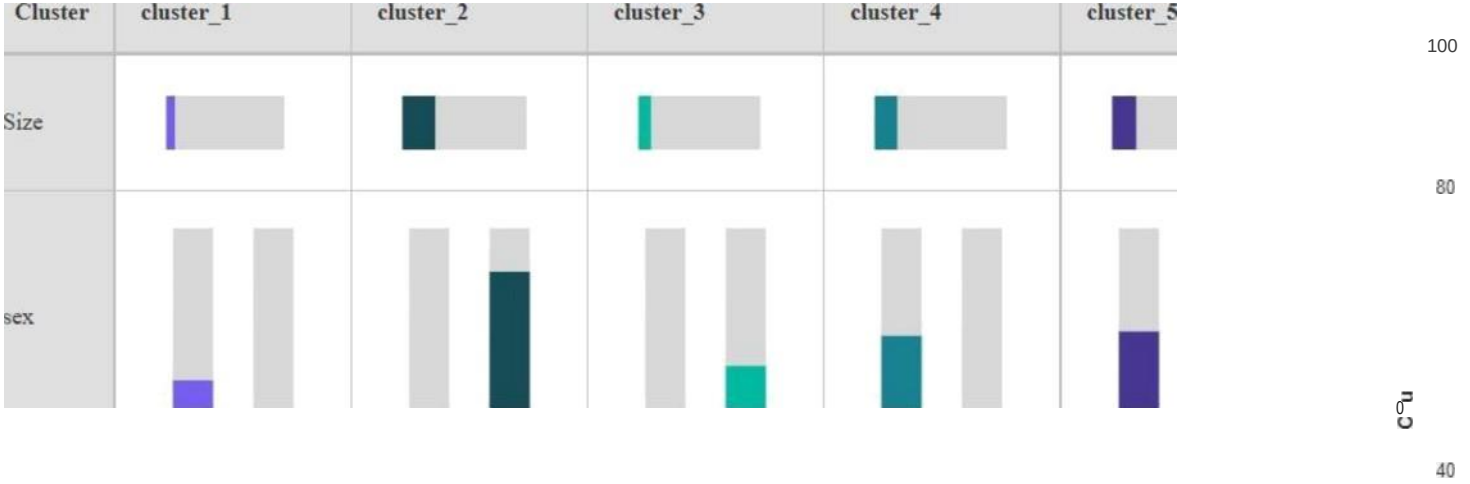
Cluster Comparison

Clusters

Cell Distribufions (Absolute)

Cell Distributions (Relative)

Build Settings



View Model: K-Means

K-Means Clustering Nlodel
ⓘ

Cluster Quality

ODELTTEñIR

Model Information

Feanu< Importance

Cluster Sizes

Cluster Comparison

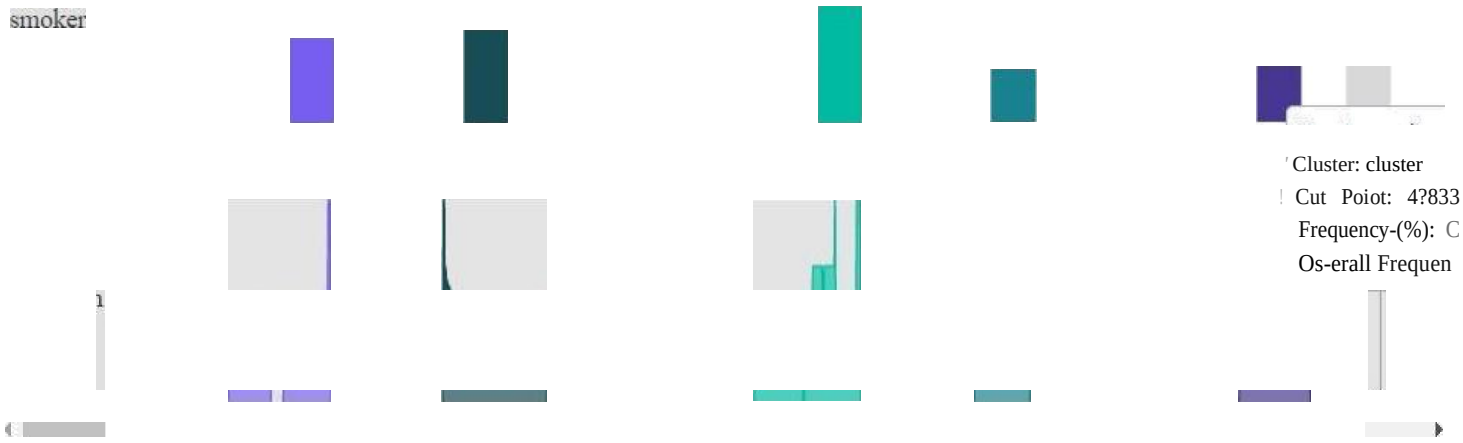
Clusters

Cell Distributions (Absolute)

Cell Distributions (Relath•e)

Build Setting

Cell Distributions (Relative) O



View Model: K-Means

K-Means Clustering Model

tell Distributions (Relative) O

Cluster Quality

MODEL VIEWER

Model Information

Feature Importance

Cluster Sizes

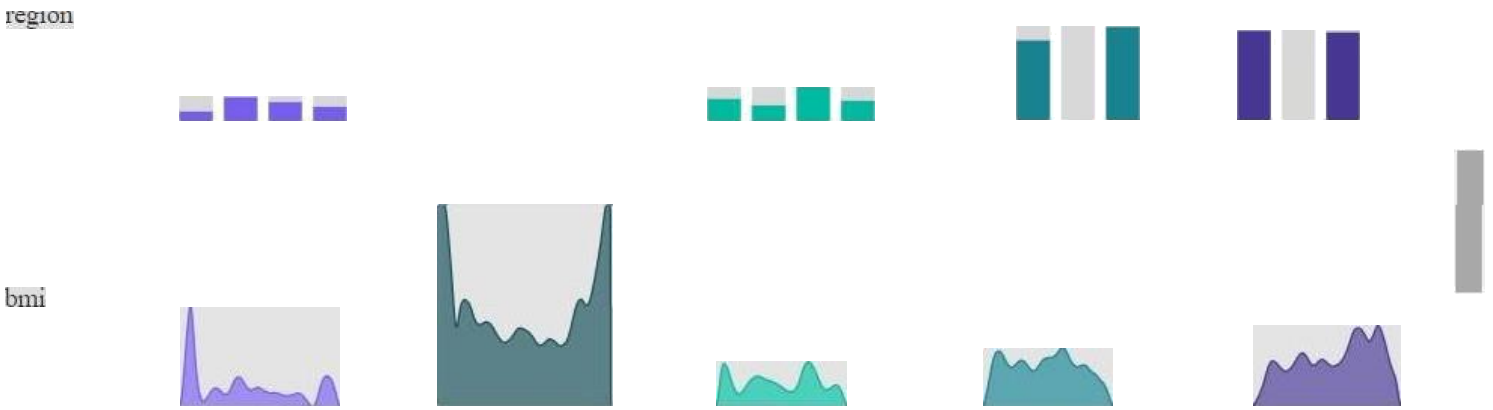
Cluster Comparison

Clusters

Cell Distribuhoiis (Abeolute)

Cell Distributions (Relative)

Build Settings



View Model: K-Means

K-Means Clustering Model

Cell Distributions (Relative) O

Cluster Quality

MODEL VIEWER

Model Information

Fearing Importance

Cluster Sizes

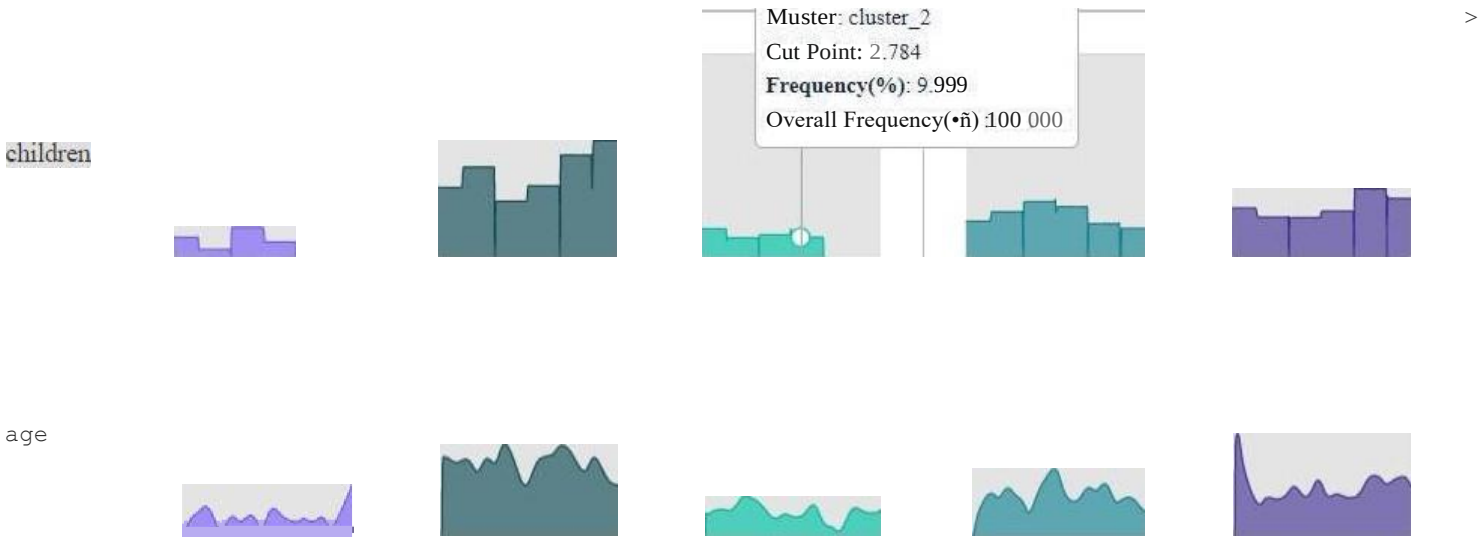
Cluster Comparison

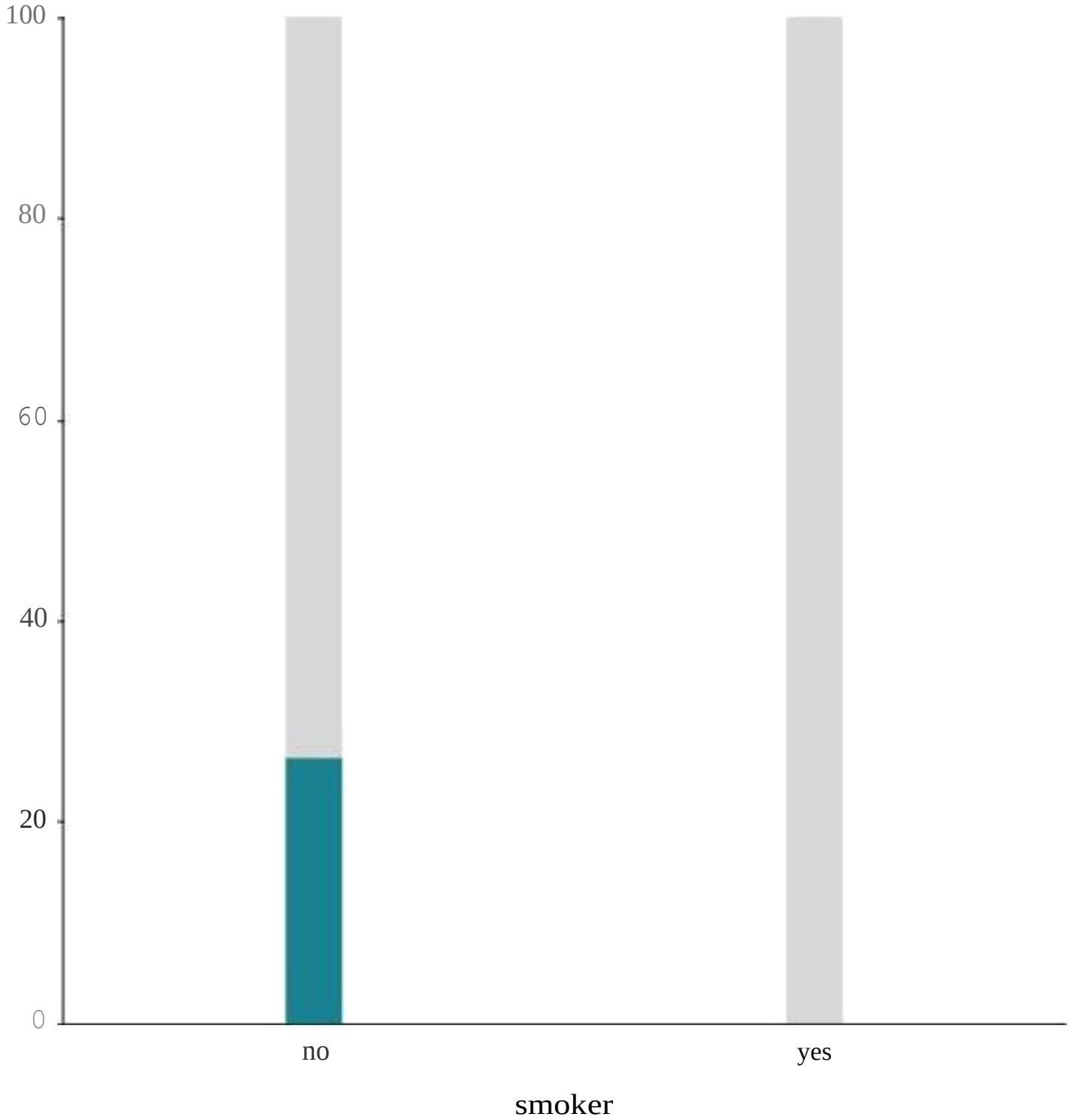
Clusters

Cell Distributions (Abeolute)

Cell Distributions (Relative)

Build Settings




100
80
60
40
20
0
no
yes
smoker

View Model: K-Means

K-Means Clustering Model

Build Settings O

Cluster Quality

hIODEL \WVT-R

Model Information

Feature Importance

Cluster Sizes

Cluster Comp6Fi8Oll

Clusters

Cell Distributions (Absolute)

Cell Distributions (Relatix c)

Use partitioned data	auc
Calculate raw propensity scores	fidsc
Calculate adjusted propmsity acores	false
Nimibcr ofclusters	5
Generate distance field	false
Cluster label	Striap

View Model: K-Means

K-Means Clustering Model

Training Summary O

Cluster Quehty

ODEL\TEwIR

Model Information

Feature Importance

Cluster Sizes

Cluster Comparison

Clusters

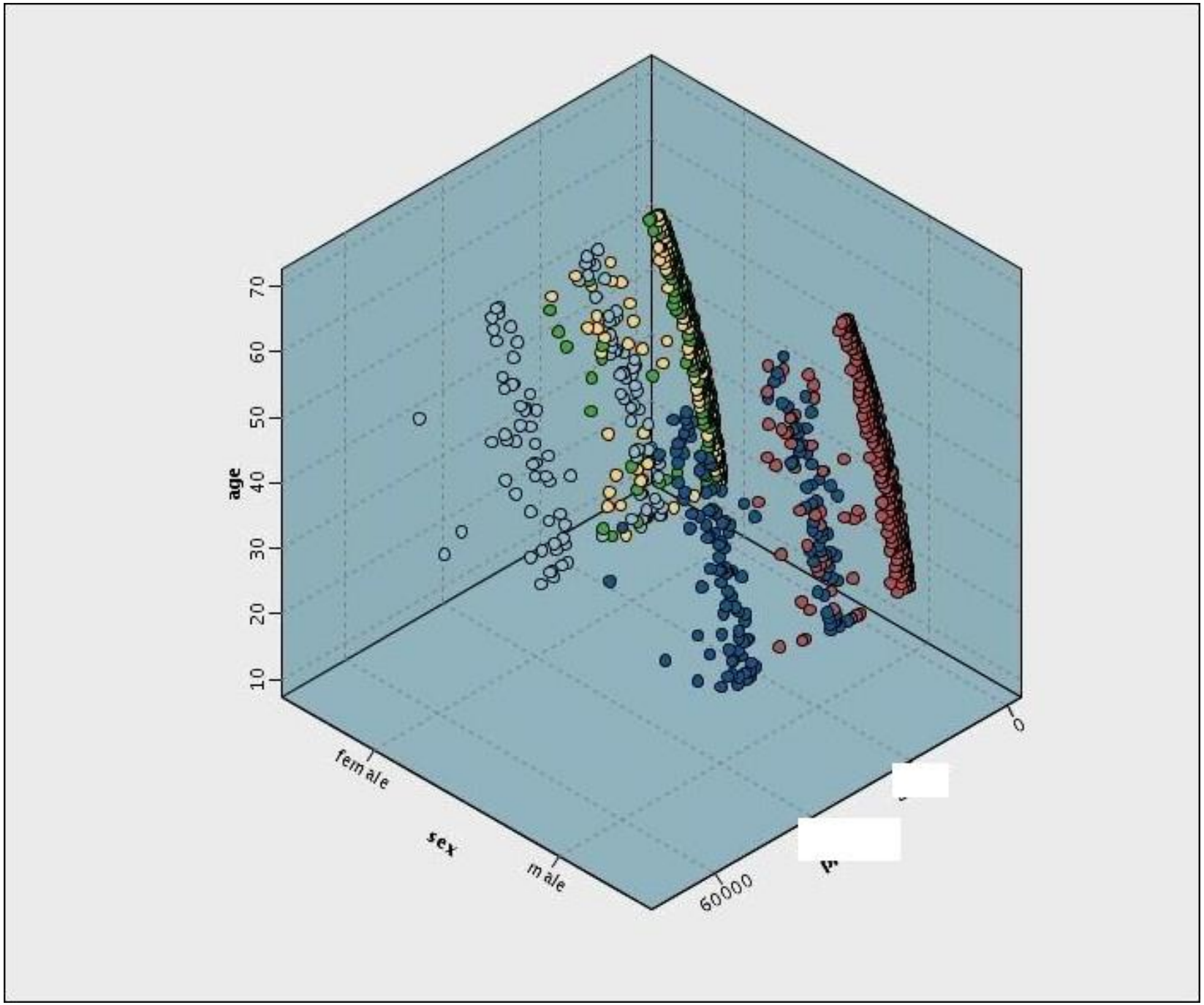
Cell Distribuhons (Absoliñc)

Cell DiaHJbuhons (Rela6ve)

Build Settings

Training Summary

Algorithm	K-means
Model iype	Clustering
Date built	Tue Dec 28 10:43:11 UTC 2021
Elapsed time ter model build	0 lioun, 0 ruins, 0 secs


$ KM —K—M eans
@ cluster— 1
@ cluster—2
@ cluster—3
M cluster—4
@ cluster—5
age
70
60
50
40
30
20
10
female
sex
male
60000
0