

Cluster Analysis

Multivariate Solutions

What is Cluster Analysis?

- Cluster analysis is an exploratory data analysis tool for solving classification problems. That is, it is a rather loose collection of statistical methods that can be used to assign cases to groups (clusters).
- Its object is to sort cases (people, things, events, etc) into groups, or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters.
- Each cluster describes, in terms of the data collected, the class to which its members belong.

When to Use Cluster Analysis

- Identify segments for marketing
 - Cluster consumers into purchase intent segments
 - Are there identifiable groups of television shows that may attract similar audiences?
- Cluster events to forecast political change
 - Economic downturn
 - Political scandal
 - Major terrorist attack
- Cluster cities based on disease (e.g. cancer clusters)

Types of Cluster Analysis

- Q-Factor (Factor Analysis)
- K-Means
- Means Based
- Hierarchical
- Individual Based
- Other Methods

Q-Factor Cluster Analysis

When Buying Gasoline, How Important Is....

	Dependability/Protection	Price	Environment	Performance
That the fuel protects your engine	0.84	0.07	0.13	0.16
That the fuel has a consistent and reliable quality	0.82	0.13	0.18	0.22
That the fuel prolongs the engine life of your car	0.76	0.15	0.19	0.13
That the fuel adds power to the engine	0.66	0.08	0.02	0.37
That the fuel have a higher octane level	0.63	0.25	0.16	0.06
That the fuel have a competitive price	0.23	0.76	0.13	-0.01
That the fuel has a low price	0.01	0.74	0.08	0.11
That the fuel gives the best mileage	0.19	0.52	0.06	0.33
That the fuel reduces other costs in the operation of your car	0.43	0.52	0.08	0.35
That the fuel does not pollute the air	0.20	0.10	0.92	0.10
That the use of the fuel does not damage the environment	0.22	0.17	0.89	0.16
That the fuel helps your car to accelerate better	0.35	0.08	0.02	0.75
That the fuel helps your car run more smoothly	0.10	0.17	0.27	0.71

Q-Factor Cluster Analysis

Using the Output to Create Market Segments

	Respondents in Category	Percentage of Sample
Dependability/Protection	77	12.8
Price	407	67.8
Environment	55	9.2
Performance	61	10.2
Total	600	100