TRANSPORTATION PREFERENCES

30/01/2024

Professor Furio Camillo

Professor Matteo Farnè

Utku Yavuz

Tolga Sümer

Nyongtakubai Ashu Ampue

Mohamed Amine Amri

1. Introduction

In modern society, transportation has a very critical role in daily life, enabling mobility and access to any resource helping the daily life. From big cities to the urban areas, individuals rely on various modes of transportation to manage their daily routines, whether it's commuting to work, running errands, or socializing. Understanding the preferences, patterns, and behaviors associated with daily transportation choices has a critical role.

As globalization increases the choices and reasons behind those decisions change. It’s observed that most countries, especially Europe, are experiencing economic and demographic changes, this diversifies the income / budget, mobility needs, priorities, and preferences regarding daily transportation. Additionally, these economic and demographic changes affect urbanization and motorization which have primary connection to daily transportation preferences.

Analyzing transportation preferences help to understand how people get around cities and suburbs. This insight is crucial for designing better transportation systems that are efficient, sustainable, and fair to everyone.

By knowing why people choose different transportation options, improve transportation quality, transportation diversity, and environmental conservation. The primary purpose of this study is to analyze and understand what influences people to make different transportation habits and preferences.

1.1 Survey Description

The survey has been carried from Google Forms and the results have been analyzed through the SAS program. SAS (Statistical Analysis System) is a software suite used for advanced analytics, business intelligence, data management, and predictive analytics. It provides tools for data manipulation, statistical analysis, and data visualization. SAS is widely used in various industries for tasks such as data mining, forecasting, optimization, and reporting.

In total the survey had 189 respondents from many demographic backgrounds. The survey was published and shared in English; therefore the survey was limited to the people who were proficient in English. It was composed of 4 demographic questions, 5 socio demographic questions, 11 behavioral questions and 15 quantitative / attribute questions.

The survey questions, divided into the relative sections, are the followings:

Demographic Questions:

How old are you? *

Mark only one oval.

- <18
- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 65>

Which gender do you identify as? *

Mark only one oval.

- Male
- Female
- Prefer not to say

Where are you from? *

Mark only one oval.

- Afghanistan
- Albania
- Algeria
- Andorra
- Angola

Where do you live? *

Mark only one oval.

- Afghanistan
- Albania
- Algeria
- Andorra
- Angola

Socio Demographic Questions:

What is your occupation status? *

Mark only one oval.

- Unemployed
- Retired
- Self employed
- Employed
- Student
- Prefer not to say

What is your primary mode of transportation for daily commuting?

Mark only one oval.

- Car
- Public transit (bus, train, subway)
- Bicycle
- Walking
- Motorcycle

How far is your typical daily commute in terms of miles or kilometers?

Mark only one oval.

- Less than 8 km (5 miles)
- 8-16 km (5-10 miles)
- 16-32 km (10-20 miles)
- 32-48 km (20-30 miles)
- More than 48 km (30 miles)
- I do not have a daily commute

On average, how much time do you spend commuting each day?

Mark only one oval.

- Less than 15 minutes
- 15-30 minutes
- 30-60 minutes
- 60-90 minutes
- More than 90 minutes
- I do not have a daily commute

Do you own a personal vehicle (car, motorcycle, bicycle)? If so, what type?

Mark only one oval.

- Car
- Motorcycle
- Bicycle
- I do not own a personal vehicle

Behavioral Questions:

What is your preferred mode of transportation for longer distances (e.g., intercity travel)?

Mark only one oval.

- ◯ Car
- ◯ Train
- ◯ Bus
- ◯ Airplane

What factors influence your choice of transportation (e.g., convenience, cost, environmental impact)?

Mark only one oval.

- ◯ Convenience
- ◯ Cost
- ◯ Environmental impact
- ◯ Time efficiency
- ◯ Safety

In your opinion, in what areas can the improvements be made to the transportation infrastructure in your area?

Mark only one oval.

- ◯ # of lines can be increased
- ◯ # of vehicles (transport frequency) can be increased
- ◯ Types of transport can be increased

Have you considered using electric or hybrid vehicles for your daily commute?

Mark only one oval.

- ◯ Yes
- ◯ No

Would you be willing to pay a premium for environmentally friendly transportation options (e.g., electric vehicles, eco-friendly fuels)?

Mark only one oval.

- Yes
- No

What factors would encourage you to use public transportation more frequently?

Mark only one oval.

- Improved reliability
- Lower cost
- Enhanced safety measures
- Better connectivity

Have you ever participated in a transportation-related survey or focus group before this one?

Mark only one oval.

- Yes
- No

How does the cost of transportation impact your mode of choice? Are you willing to pay a premium for specific benefits?

Mark only one oval.

- Willing to pay a premium
- Neutral
- Cost is a consideration but not a deciding factor
- Prefer lower-cost options
- Cost is the primary consideration

Have you ever used a bike-sharing or scooter-sharing service?

Mark only one oval.

- Yes
- No

Have you ever participated in carpooling or ride-sharing programs?

Mark only one oval.

- ◯ Yes, regularly
- ◯ Yes, occasionally
- ◯ No, never

Do you prefer a single mode of transportation or a combination (e.g. walking + public transit)?

Mark only one oval.

- ◯ Single mode
- ◯ Combination of modes
- ◯ It depends on the circumstances

Attribute Questions:

How frequently do you use public transportation (bus, train, subway)?

Mark only one oval.

- ◯ Daily
- ◯ 2-3 times a week
- ◯ Once a week
- ◯ Rarely
- ◯ Never

How satisfied are you with the current transportation options available in your area?

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Very satisfied

Would you be open to trying new transportation options if they were introduced in your area?

Mark only one oval.

- ◯ Yes, definitely
- ◯ Yes, probably
- ◯ Maybe
- ◯ Probably not
- ◯ No, definitely not

Would you be willing to use alternative transportation methods for environmental reasons (e.g., walking, biking)?

Mark only one oval.

- Definitely
- Probably
- Neutral
- Probably not
- Definitely not

Have you ever encountered challenges with your chosen mode of transportation (e.g., delays, breakdowns)?

Mark only one oval.

- Frequently
- Occasionally
- Rarely
- Never

How would you rate the accessibility of transportation options for individuals with mobility challenges in your area?

Mark only one oval.

- Very accessible
- Somewhat accessible
- Neutral
- Not very accessible
- Not accessible at all

On a scale of 1 to 10, how important is environmental sustainability in your transportation choices?

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not i											Very important

How often do you engage in active transportation (e.g., walking, biking) for leisure or exercise?

Mark only one oval.

- Daily
- 2-3 times a week
- Once a week
- Rarely
- Never

In your opinion, how well does the current transportation infrastructure accommodate cyclists and pedestrians?

Mark only one oval.

- Very well
- Adequately
- Neutral
- Poorly
- Very poorly

Do you use any mobile apps or technology to plan or optimize your daily commute?

Mark only one oval.

- Daily
- Regularly
- Occasionally
- Rarely
- Never

Do you think there is a need for increased investment in public transportation infrastructure in your area?

Mark only one oval.

	1	2	3	4	5	
Stror	◯	◯	◯	◯	◯	Strongly agree

How important is safety in influencing your transportation decisions?

Mark only one oval.

- Very Important
- Important
- Neutral
- Not Important
- Not at all Important

How do you perceive public transportation in terms of comfort, convenience, and overall experience?

Mark only one oval.

- Very Positive
- Positive
- Neutral
- Negative
- Very Negative

How many times per week do you use your personal vehicle for non-commuting purposes?

Mark only one oval.

- Daily
- 2-3 times a week
- Once a week
- Rarely
- Never

How important is parking availability when choosing your mode of transportation?

Mark only one oval.

- Very important
- Important
- Neutral
- Not important
- Not applicable (I do not own a personal vehicle)

2. Preliminary Analysis

2.1 Library and Labeling

2.1.1 Library

The SAS code begins with specifying a library named 'c', which points to the directory 'C:\Users\bbsstudent\Desktop'. This library location is utilized throughout the code for data import, manipulation, and output. It provides a convenient way to organize and access data files and results within the specified directory.

The 'proc import' procedure is utilized to import data from an Excel file named 'survey-new.xlsx' located in the specified directory. The data is imported into a SAS dataset named 'c.transport', using the XLSX engine for Excel files. The 'replace' option is used to replace any existing dataset with the same name.

After importing the data, the 'proc contents' procedure is employed to examine the contents of the 'c.transport' dataset. This procedure provides a summary of the dataset's attributes, including variable names, types, lengths, and formats. It helps in understanding the structure of the dataset and verifying the successful import of data.

2.1.2 Labeling

The 'label' statement is used to assign descriptive labels to variables in the dataset. Each variable is labeled to provide meaningful interpretations of its content. For example, variables such as 'new_category1' to 'new_category10' are labeled with descriptive names indicating different aspects related to transportation options, environmental reasons, sustainability, and other factors.

2.2 Frequency Procedure

The Frequency Procedure, implemented through the 'proc freq' statement, is a fundamental analysis tool used to examine the distribution and frequencies of categorical variables within the dataset. It provides valuable insights into the prevalence of different categories within each variable and facilitates the exploration of relationships between variables.

2.2.1 Analysis of Categorical Variables

In the provided SAS code, the Frequency Procedure, implemented through the 'proc freq' statement, is utilized to analyze the distribution and frequencies of categorical variables within the dataset. Let's relate the components of the Frequency Procedure to the code provided:

```
proc freq data=c.transport_1; *age 1; table cluster*age / expected chisq; run;
```

Here, the 'proc freq' procedure is applied to the dataset 'c.transport_1'. The statement 'table cluster*age' specifies the variables for which frequency analysis is conducted. This code segment examines the distribution of the 'age' variable within each 'cluster' category.

2.3 Means Procedure

The Means Procedure, facilitated by the 'proc means' statement, is a statistical analysis tool used to compute descriptive statistics such as means, minimums, maximums, and other measures of central tendency and variability for numerical variables in the dataset. This procedure offers valuable insights into the distribution and characteristics of quantitative data, aiding in data exploration and hypothesis testing.

Let's relate the components of the Frequency Procedure to the code provided:

proc means data=c.coord_adj;
var prin1-prin4;

Here, the 'proc means' procedure is applied to the dataset 'c.coord_adj'. The 'var' statement specifies the numerical variables ('prin1' to 'prin4') for which descriptive statistics are computed. These variables represent principal components derived from the principal component analysis (PCA) performed earlier in the code. The first 4 principle component are chosen according to the Kaiser's Criterion, which indicates that the values should be bigger than 1.

Before Size effect:

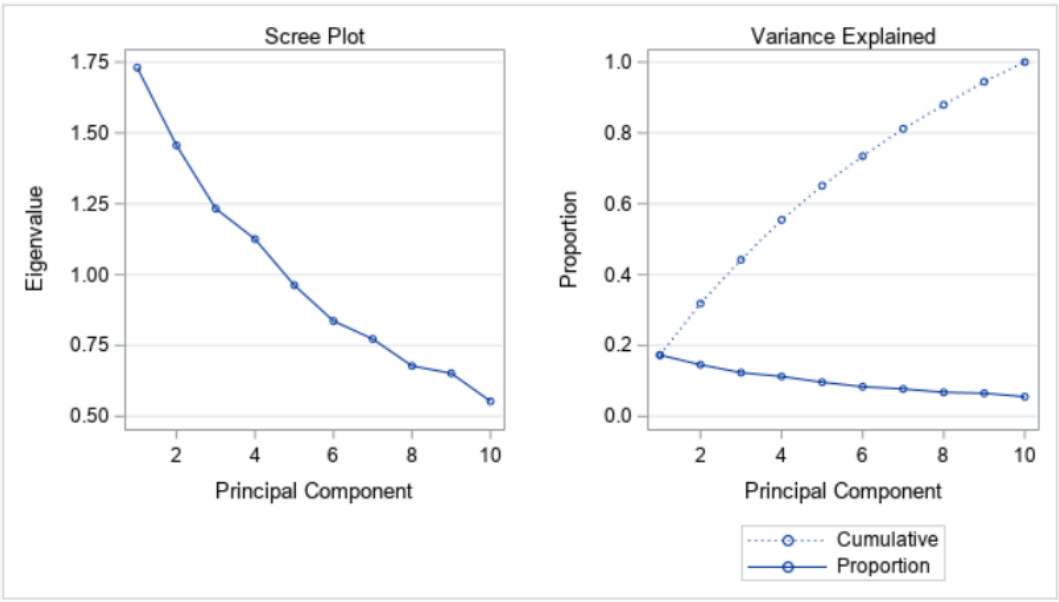
Eigenvectors																
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin9	Prin10	Prin11	Prin12	Prin13	Prin14	Prin15
Q1	Q1	0.387556	0.026530	0.285290	0.233756	0.090869	-.076783	0.114431	0.038231	-.029330	0.471596	0.212467	-.255128	0.375560	0.358484	-.285137
Q2	Q2	0.126986	0.230119	-.433321	-.310139	-.024716	-.016903	-.269267	0.170085	-.365054	0.247622	-.249217	0.239924	0.472424	-.001389	0.084840
Q3	Q3	-.049784	0.085975	0.483281	0.123051	0.320137	0.215212	-.267808	-.487677	0.028120	0.046830	-.294662	0.139794	0.280316	-.249209	0.181769
Q4	Q4	-.189650	0.423295	0.209197	-.199899	-.307036	0.105160	-.140633	-.078795	0.050247	-.282470	-.296585	-.204901	0.003021	0.580881	-.141853
Q5	Q5	0.399690	0.157763	-.076202	-.190630	0.005803	0.219294	0.243723	-.285745	0.201941	0.186769	-.091935	0.546029	-.351260	0.120065	-.252933
Q6	Q6	-.182443	0.378929	0.139824	-.236386	-.406812	-.141117	0.169404	0.015667	0.403641	0.391912	0.107815	-.106927	0.100025	-.416077	0.103157
Q7	Q7	0.182380	0.144069	0.394707	0.080271	-.286337	0.127939	0.424803	0.225307	-.564783	-.223375	-.075015	0.133634	0.015192	-.247400	0.019156
Q8	Q8	0.436660	0.194967	-.119318	0.059914	-.096052	0.041132	0.027694	-.149048	0.196959	-.361450	0.373870	0.013766	0.209938	0.138061	0.589108
Q9	Q9	0.185272	0.197480	0.012953	0.570718	-.189063	0.057362	-.440837	0.436164	0.274062	-.023615	-.155750	0.196731	-.115060	-.119835	-.104641
Q10	Q10	-.327443	-.188103	0.176233	-.102753	0.134662	0.432589	0.195335	0.459831	0.250819	0.076107	0.092224	0.343504	0.235051	0.265152	0.200394
Q11	Q11	-.056586	0.344822	-.048896	0.159892	0.477909	-.491085	0.407651	0.166512	0.131407	-.074425	-.358089	0.048582	-.013749	0.090825	0.156971
Q12	Q12	-.246066	0.376477	0.167925	-.065392	0.226938	-.235608	-.243803	0.010270	-.192792	-.109869	0.612343	0.351887	-.053194	0.014322	-.218225
Q13	Q13	0.191207	0.317306	-.076536	-.241865	0.445155	0.471739	-.019936	0.248303	0.056807	-.110917	0.071334	-.437253	-.106290	-.260284	-.157463
Q14	Q14	-.254575	0.117306	-.385917	0.355857	-.075342	0.216129	0.301813	-.245383	0.094439	-.237328	0.059752	0.046695	0.433430	-.125161	-.409454
Q15	Q15	-.267571	0.288210	-.203779	0.374925	-.003606	0.306646	0.063074	-.138976	-.310611	0.415561	0.088684	-.109259	-.329887	0.166198	0.347115

Eigenvectors											
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin9	Prin10
Q1	Q1	0.182953	0.402129	-.169664	0.620351	-.183707	-.249146	-.297009	0.396007	-.148217	0.168443
Q2	Q2	0.485557	0.220803	-.036847	-.192968	0.251192	-.206822	-.072384	0.175995	0.583399	-.440959
Q3	Q3	0.451849	0.117948	-.057018	-.484926	0.223527	-.064622	0.189485	0.279602	-.520464	0.323128
Q4	Q4	0.069192	0.487683	0.217084	0.034553	0.307597	0.668920	-.319358	-.215961	0.014987	0.134667
Q5	Q5	0.048224	0.069123	0.632786	0.385434	0.274471	-.123710	0.586617	0.053432	0.045157	0.065320
Q6	Q6	0.082418	-.038062	-.652481	0.293232	0.185393	0.410616	0.512627	0.048533	0.039289	-.101883
Q7	Q7	0.303887	-.109359	0.278092	-.047776	-.661766	0.466815	0.118956	0.327298	0.007616	-.197210
Q8	Q8	0.479053	0.094035	-.086213	0.064003	-.346760	-.151063	0.146831	-.681973	0.140847	0.317160
Q9	Q9	0.213998	-.590097	0.029653	0.129603	0.203764	0.137269	-.239480	0.217781	0.363336	0.539592
Q10	Q10	0.380861	-.403173	0.087650	0.289589	0.221749	0.005717	-.262078	-.250075	-.459432	-.457544

After size effect:

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	1.73050606	0.27424072	0.1731	0.1731
2	1.45626534	0.22318593	0.1456	0.3187
3	1.23307941	0.10762041	0.1233	0.4420
4	1.12545900	0.16266588	0.1125	0.5545
5	0.96279312	0.12651221	0.0963	0.6508
6	0.83628091	0.06336589	0.0836	0.7344
7	0.77291502	0.09504068	0.0773	0.8117
8	0.67787435	0.02605133	0.0678	0.8795
9	0.65182301	0.09881923	0.0652	0.9447
10	0.55300378		0.0553	1.0000

Eigenvectors											
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin9	Prin10
new_category1	transportation_options	0.381422	-.029549	-.356147	0.033758	0.488239	-.028182	0.572506	0.379435	0.014796	0.120579
new_category2	environmental_reasons	0.434865	0.262437	0.102212	-.359696	0.026915	0.122320	0.232402	-.702974	0.168957	0.098865
new_category3	sustainability	0.325665	0.360362	0.181488	-.142094	-.461449	0.396106	-.053716	0.532581	0.177087	0.152418
new_category4	active_transportation	0.212561	-.524699	0.086182	-.067328	-.324481	-.512476	0.033731	0.043202	0.443599	0.315983
new_category5	mobile_apps	-.068801	-.449344	0.276799	-.201929	0.462550	0.526781	-.255822	0.040103	0.165889	0.299900
new_category6	investment	0.052721	0.133872	-.709428	0.138987	0.044830	0.112878	-.464515	-.105265	0.460277	0.056857
new_category7	cost	-.035712	-.027386	0.239825	0.776525	-.062347	0.277063	0.306324	-.168232	0.362823	-.055129
new_category8	safety	0.310944	0.324884	0.284053	0.363331	0.302467	-.331650	-.396905	0.003352	-.203033	0.429911
new_category9	personal_vehicle	-.552818	0.226683	-.149346	-.050620	-.116929	0.025224	0.283591	-.065467	-.030006	0.719877
new_category10	parking	-.321073	0.384145	0.286242	-.217025	0.343027	-.295094	0.034537	0.176236	0.574419	-.230982



3. Principal Component Analysis

3.1 Size Effect: Recognition and Elimination

3.1.1 Recognition

3.1.1.1 Cluster Analysis

The 'proc cluster' procedure is utilized to identify clusters within the dataset based on the variables specified in the analysis. However, the size of each cluster can vary, potentially influencing the interpretation of results.

3.1.1.2 Group Comparison

The code conducts group comparisons using statistical tests such as t-tests between different clusters. The size of each cluster can impact the statistical power and significance of the results obtained from these comparisons. Large clusters may have more statistical power than smaller ones, potentially biasing the results.

First, we provided the whole questions, but after size effect verification we removed some questions according to the magnitudes of the first eigenvector of the principal components from the data's covariance matrix, in which they all have positive eigenvalues.

3.1.2 Elimination

To address the size effect, the analysis includes procedures to eliminate its influence. This is achieved through the use of statistical tests (e.g., t-tests) that compare different clusters to a reference group (cluster 6), allowing for the identification of significant differences while mitigating the impact of cluster size.

3.1.2.1 Statistical Tests

The code conducts t-tests to compare different clusters. While the t-tests can identify significant differences between clusters, they may not fully account for variations in cluster sizes. However, the comparison to a reference group (cluster 6) helps mitigate the impact of cluster size differences to some extent.

3.1.2.2 Consideration of Cluster Composition

Hierarchical clustering methods, such as the 'method=ward' option used in 'proc cluster', take into account the composition and structure of clusters. This helps in recognizing the inherent differences between clusters and informs the interpretation of results.

3.1.2.3 Data Normalization

Although not explicitly implemented in the provided code, data normalization techniques could potentially be applied to standardize variables and mitigate the impact of differences in variable scales or magnitudes across clusters. Normalization ensures that variables contribute equally to the analysis, irrespective of their initial scales.

4. Clustering

Clustering analysis is a fundamental technique used to group data points into clusters based on similarity or proximity. In the provided SAS code, clustering analysis is conducted to identify natural groupings or patterns within the dataset. Here's an overview of the key components related to clustering analysis:

4.1 Clustering Method

In the provided SAS code, Ward's method is utilized as the clustering algorithm. Here's how it relates to the code and its significance:

Ward's Method: This hierarchical clustering algorithm operates by iteratively merging clusters to minimize the within-cluster variance. In the context of the provided code, the 'method=ward' option within the 'proc cluster' procedure indicates the use of Ward's method for hierarchical clustering. Ward's method is chosen for its ability to produce compact and well-separated clusters by minimizing the variance within clusters. It tends to create balanced and interpretable dendrograms, making it suitable for identifying meaningful patterns in the data.

4.2 Cluster Formation

The provided SAS code employs hierarchical clustering, specifically Ward's method, to form clusters. Hierarchical clustering iteratively merges data points or clusters based on their similarity or dissimilarity, creating a hierarchical structure of nested clusters.

Procedure Execution: In the code, the 'proc cluster' procedure is utilized to execute the clustering algorithm. By specifying the 'method=ward' option, the procedure employs Ward's method to iteratively merge clusters while minimizing the within-cluster variance.

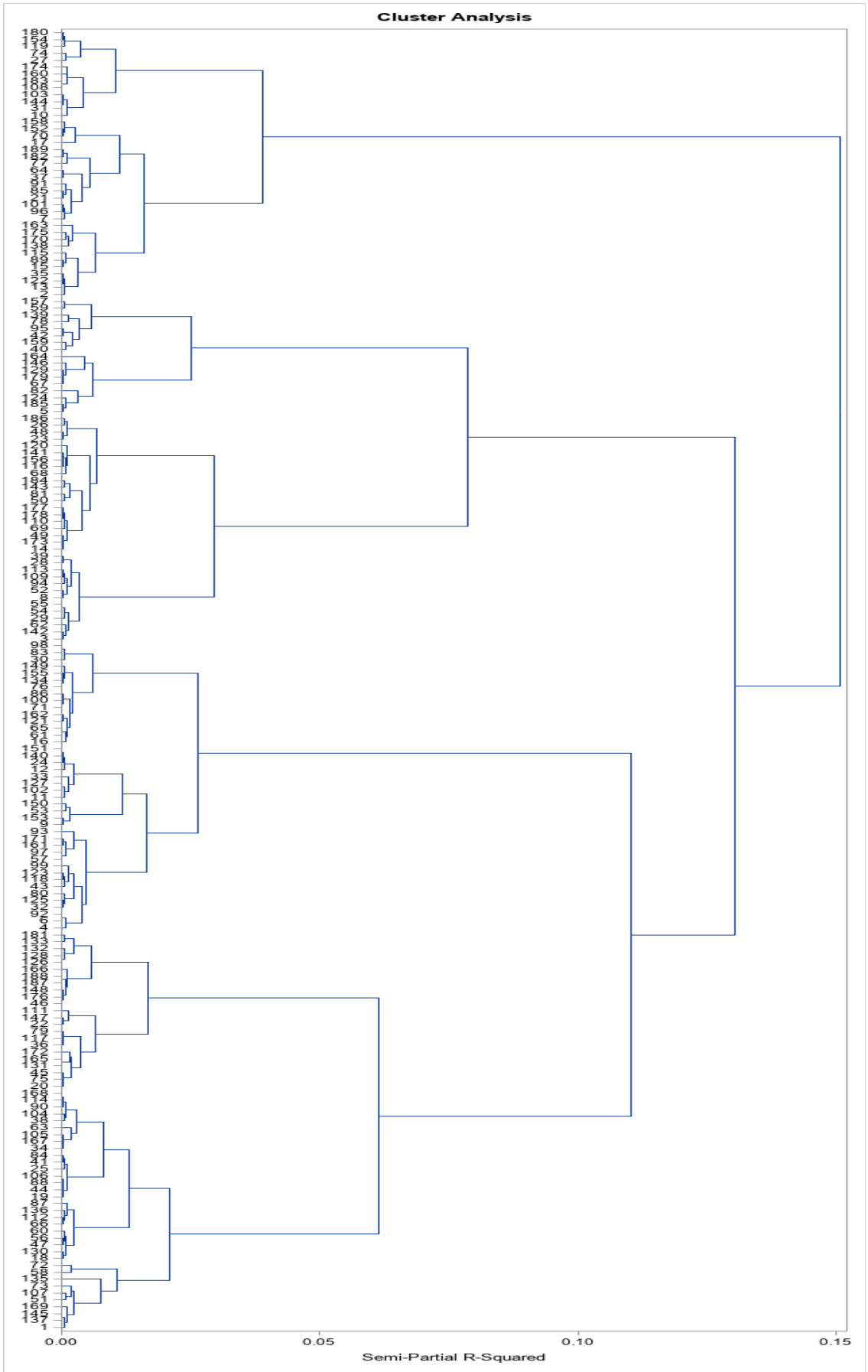
The clustering algorithm iteratively merges clusters based on the chosen distance measure until a predefined stopping criterion is met. The resulting clusters form a hierarchical structure, visualized through a dendrogram, which illustrates the relationships between clusters at different levels of granularity.

4.3 Dendrogram and Cluster Identification

In the provided SAS code, dendrogram visualization is achieved through the 'proc template' and 'proc sgrender' procedures. These procedures generate a hierarchical tree-like diagram (dendrogram) that illustrates the relationships between clusters. Each branch of the dendrogram represents a cluster, and the length of the branches indicates the distance or dissimilarity between clusters.

Clusters closer to the root of the dendrogram represent larger and more generalized groupings, while clusters closer to the leaves represent smaller and more specific groupings. We choose to cut the dendrogram at a certain height of 4 to define clusters.

Our dendrogram is shown below:


Cluster Analysis
0.00
0.05
0.10
0.15
Semi-Partial R-Squared

4.4 T-Test

In the provided SAS code, multiple t-tests are conducted to compare the means of various variables across different clusters. The 'proc ttest' procedure is utilized to perform these statistical comparisons, with each t-test focusing on a specific variable of interest.

T-tests are conducted to compare each cluster against a reference group (cluster 6) or other clusters of interest. This comparison allows analysts to identify clusters with distinct characteristics or behaviors compared to the reference group, providing valuable insights into cluster differentiation.

4.5 Cluster Analysis

Positive t-value indicates that the mean of the first group is higher than the mean of the second group and negative t-value indicates that the mean of the first group is lower than the mean of the second group. Also pr>|t| < .0001 indicates that the probability of observing a t-value as extreme as, or more extreme than, the one obtained in the sample, assuming the null hypothesis is true, is less than 0.0001. With a p-value much smaller than 0.05 (in this case, less than 0.0001), the evidence against the null hypothesis is very strong. In the clusters below it can be seen that the rules above are accurate.

4.5.1 Cluster 1

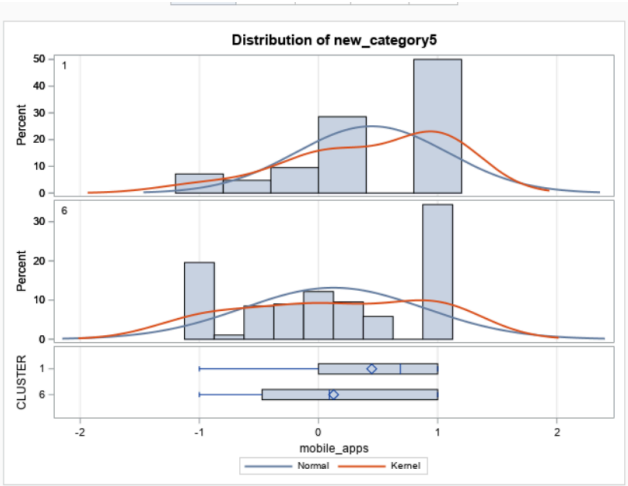
Variable: new_category5 (mobile_apps)

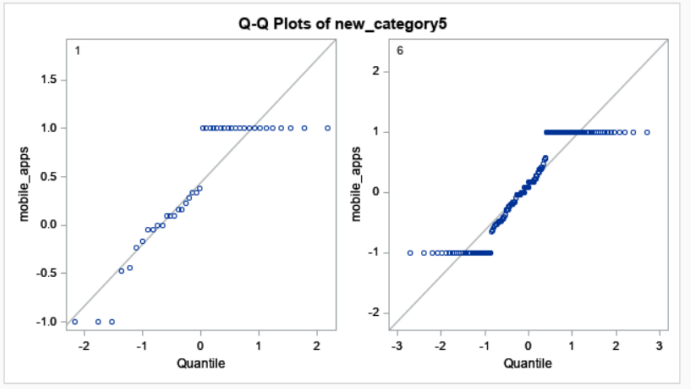
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	0.4465	0.6384	0.0985	-1.0000	1.0000
6		189	0.1267	0.7588	0.0552	-1.0000	1.0000
Diff (1-2)	Pooled		0.3198	0.7387	0.1260		
Diff (1-2)	Satterthwaite		0.3198		0.1129		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		0.4465	0.2476	0.6455	0.6384	0.5253	0.8141
6		0.1267	0.0178	0.2356	0.7588	0.6892	0.8441
Diff (1-2)	Pooled	0.3198	0.0715	0.5681	0.7387	0.6768	0.8131
Diff (1-2)	Satterthwaite	0.3198	0.0945	0.5450			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	2.54	0.0118
Satterthwaite	Unequal	69.29	2.83	0.0061

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	1.41	0.1907





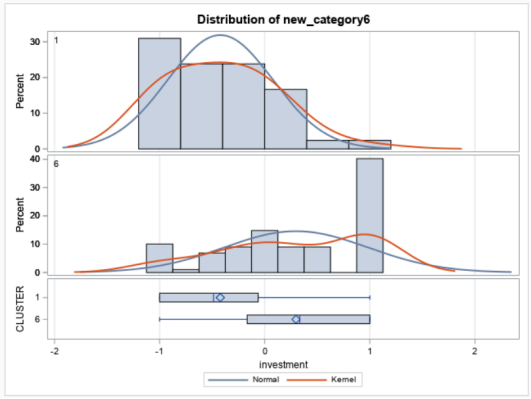
Variable: new_category6 (investment)

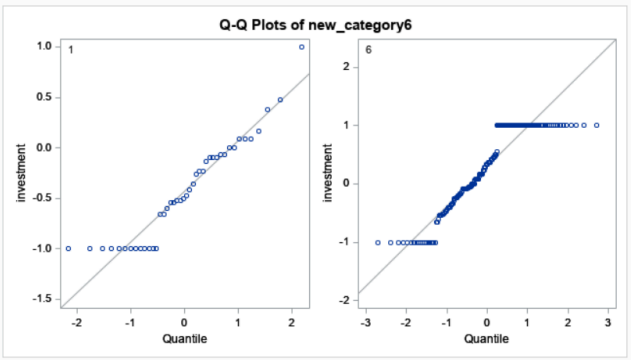
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	-0.4226	0.5005	0.0772	-1.0000	1.0000
6		189	0.2977	0.6835	0.0497	-1.0000	1.0000
Diff (1-2)	Pooled		-0.7203	0.6545	0.1117		
Diff (1-2)	Satterthwaite		-0.7203		0.0919		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		-0.4226	-0.5786	-0.2666	0.5005	0.4118	0.6383
6		0.2977	0.1996	0.3958	0.6835	0.6208	0.7603
Diff (1-2)	Pooled	-0.7203	-0.9403	-0.5003	0.6545	0.5997	0.7205
Diff (1-2)	Satterthwaite	-0.7203	-0.9031	-0.5375			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	-6.45	<.0001
Satterthwaite	Unequal	79.06	-7.84	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	1.86	0.0199





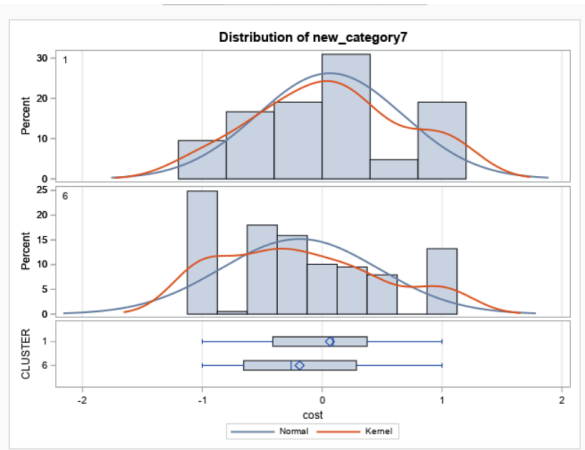
Variable: new_category7 (cost)

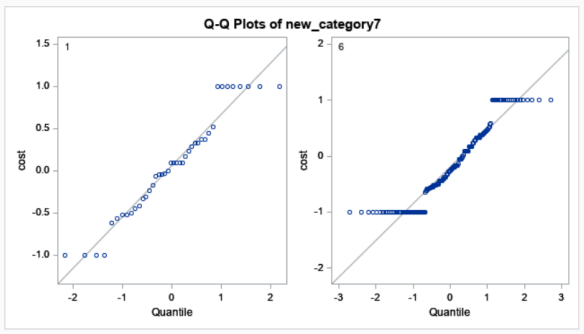
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	0.0645	0.6082	0.0938	-1.0000	1.0000
6		189	-0.1884	0.6571	0.0478	-1.0000	1.0000
Diff (1-2)	Pooled		0.2528	0.6486	0.1106		
Diff (1-2)	Satterthwaite		0.2528		0.1053		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		0.0645	-0.1251	0.2540	0.6082	0.5004	0.7756
6		-0.1884	-0.2826	-0.0941	0.6571	0.5969	0.7310
Diff (1-2)	Pooled	0.2528	0.0348	0.4708	0.6486	0.5943	0.7140
Diff (1-2)	Satterthwaite	0.2528	0.0424	0.4632			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	2.28	0.0232
Satterthwaite	Unequal	64.089	2.40	0.0193

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	1.17	0.5687





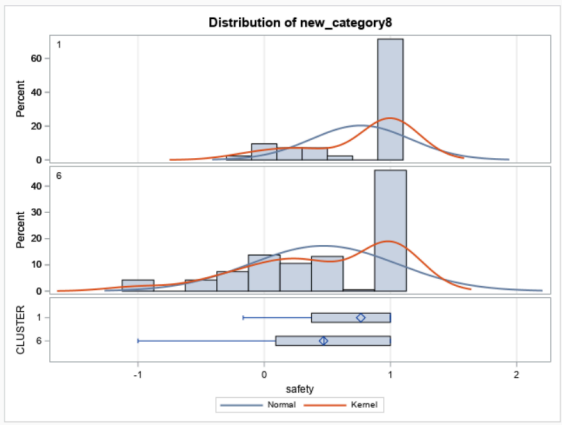
Variable: new_category8 (safety)

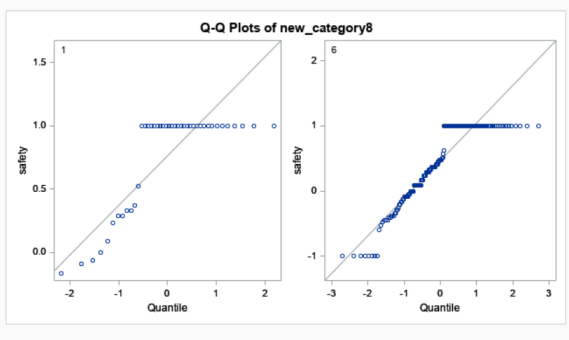
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	0.7652	0.3923	0.0605	-0.1667	1.0000
6		189	0.4696	0.5783	0.0421	-1.0000	1.0000
Diff (1-2)	Pooled		0.2956	0.5497	0.0938		
Diff (1-2)	Satterthwaite		0.2956		0.0737		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		0.7652	0.6430	0.8874	0.3923	0.3228	0.5002
6		0.4696	0.3866	0.5526	0.5783	0.5253	0.6433
Diff (1-2)	Pooled	0.2956	0.1109	0.4804	0.5497	0.5036	0.6051
Diff (1-2)	Satterthwaite	0.2956	0.1491	0.4422			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	3.15	0.0018
Satterthwaite	Unequal	85.802	4.01	0.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	2.17	0.0042





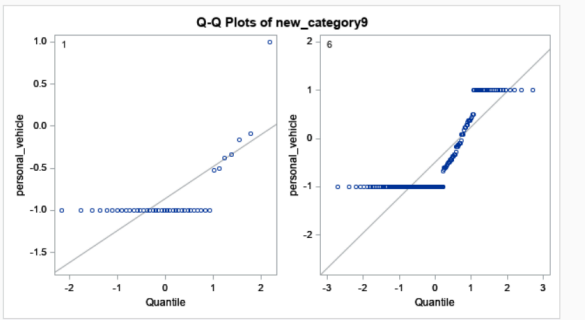
Variable: new_category9 (personal_vehicle)

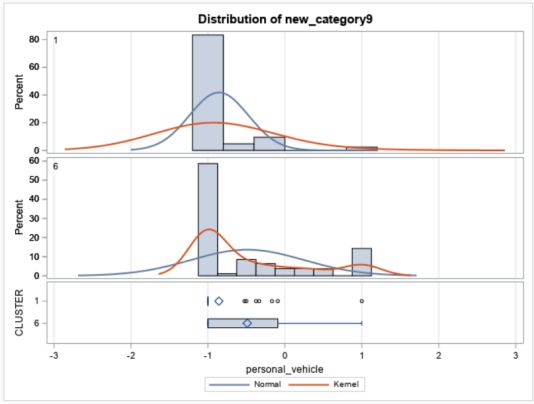
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	-0.8569	0.3817	0.0589	-1.0000	1.0000
6		189	-0.4875	0.7331	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		-0.3694	0.6836	0.1166		
Diff (1-2)	Satterthwaite		-0.3694		0.0794		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		-0.8569	-0.9758	-0.7380	0.3817	0.3140	0.4867
6		-0.4875	-0.5927	-0.3823	0.7331	0.6659	0.8155
Diff (1-2)	Pooled	-0.3694	-0.5992	-0.1396	0.6836	0.6263	0.7525
Diff (1-2)	Satterthwaite	-0.3694	-0.5267	-0.2121			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	-3.17	0.0017
Satterthwaite	Unequal	118.42	-4.65	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	3.69	<.0001





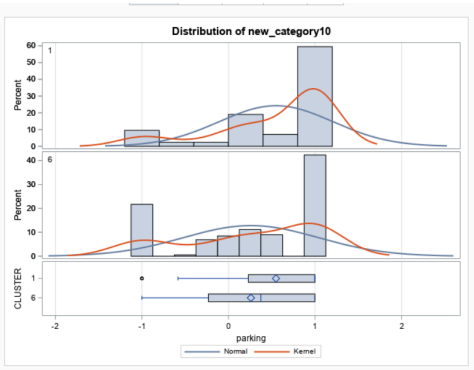
Variable: new_category10 (parking)

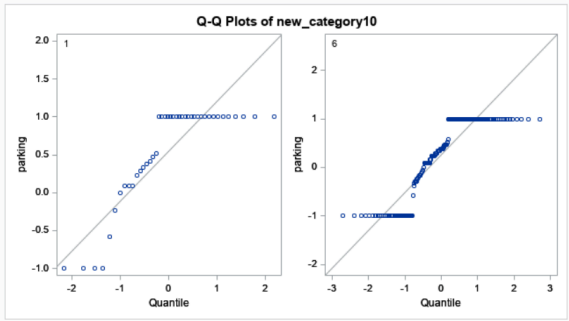
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		42	0.5498	0.6585	0.1016	-1.0000	1.0000
6		189	0.2596	0.7799	0.0567	-1.0000	1.0000
Diff (1-2)	Pooled		0.2902	0.7596	0.1296		
Diff (1-2)	Satterthwaite		0.2902		0.1164		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		0.5498	0.3446	0.7550	0.6585	0.5418	0.8396
6		0.2596	0.1477	0.3715	0.7799	0.7084	0.8676
Diff (1-2)	Pooled	0.2902	0.0349	0.5455	0.7596	0.6959	0.8362
Diff (1-2)	Satterthwaite	0.2902	0.0581	0.5224			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	229	2.24	0.0261
Satterthwaite	Unequal	69.086	2.49	0.0150

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	41	1.40	0.1996





4.5.2 Cluster 2

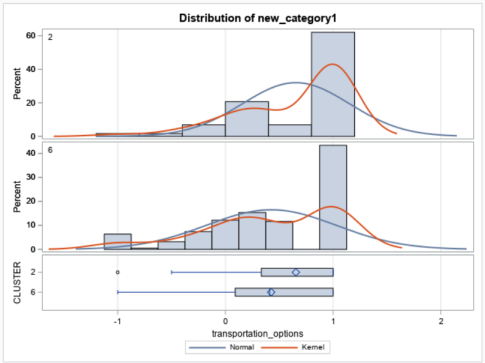
Variable: new_category1 (transportation_options)

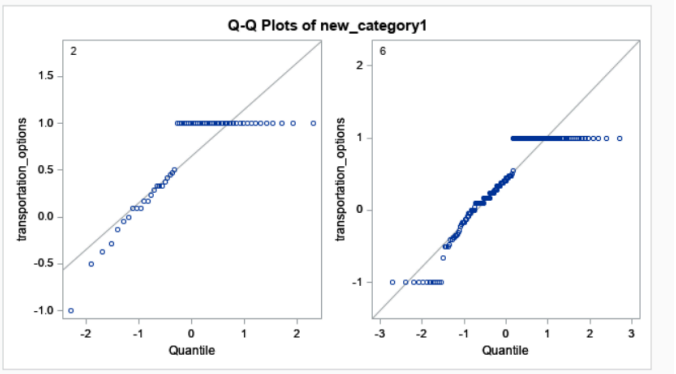
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	0.6550	0.4981	0.0654	-1.0000	1.0000
6		189	0.4254	0.6048	0.0440	-1.0000	1.0000
Diff (1-2)	Pooled		0.2296	0.5817	0.0873		
Diff (1-2)	Satterthwaite		0.2296		0.0788		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		0.6550	0.5240	0.7859	0.4981	0.4211	0.6098
6		0.4254	0.3386	0.5122	0.6048	0.5493	0.6727
Diff (1-2)	Pooled	0.2296	0.0576	0.4016	0.5817	0.5344	0.6382
Diff (1-2)	Satterthwaite	0.2296	0.0734	0.3857			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	2.63	0.0091
Satterthwaite	Unequal	113.22	2.91	0.0043

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	57	1.47	0.0875





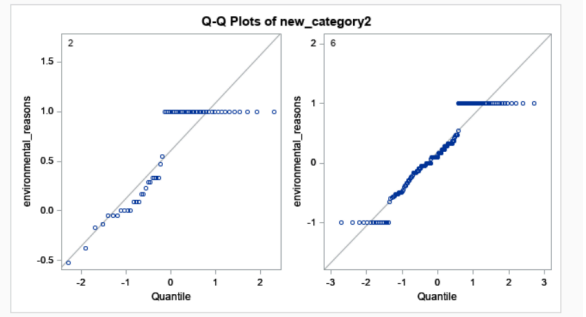
Variable: new_category2 (environmental_reasons)

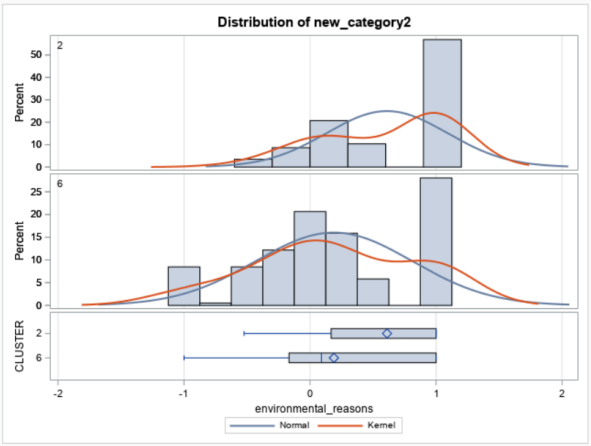
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	0.6107	0.4793	0.0629	-0.5238	1.0000
6		189	0.1894	0.6226	0.0453	-1.0000	1.0000
Diff (1-2)	Pooled		0.4213	0.5923	0.0889		
Diff (1-2)	Satterthwaite		0.4213		0.0775		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		0.6107	0.4847	0.7368	0.4793	0.4052	0.5869
6		0.1894	0.1001	0.2788	0.6226	0.5655	0.6926
Diff (1-2)	Pooled	0.4213	0.2461	0.5964	0.5923	0.5442	0.6499
Diff (1-2)	Satterthwaite	0.4213	0.2678	0.5748			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	4.74	<.0001
Satterthwaite	Unequal	121.42	5.43	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	57	1.69	0.0223





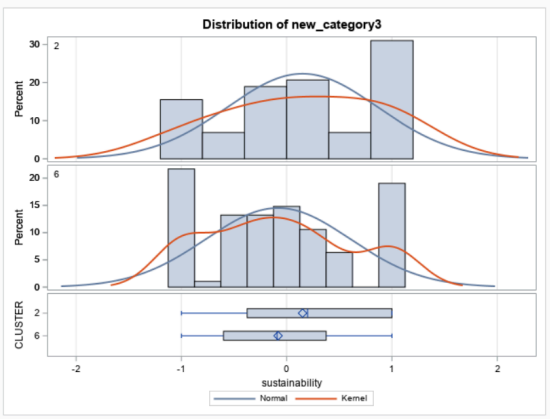
Variable: new_category3 (sustainability)

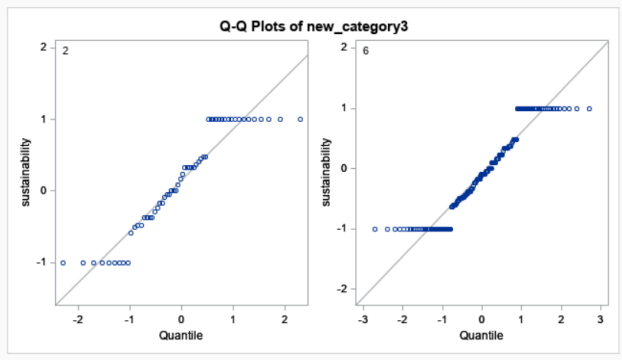
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	0.1512	0.7143	0.0938	-1.0000	1.0000
6		189	-0.0815	0.6868	0.0500	-1.0000	1.0000
Diff (1-2)	Pooled		0.2327	0.6933	0.1041		
Diff (1-2)	Satterthwaite		0.2327		0.1063		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		0.1512	-0.0367	0.3390	0.7143	0.6039	0.8745
6		-0.0815	-0.1801	0.0170	0.6868	0.6239	0.7640
Diff (1-2)	Pooled	0.2327	0.0277	0.4377	0.6933	0.6370	0.7606
Diff (1-2)	Satterthwaite	0.2327	0.0216	0.4437			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	2.24	0.0263
Satterthwaite	Unequal	91.697	2.19	0.0311

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	57	188	1.08	0.6844





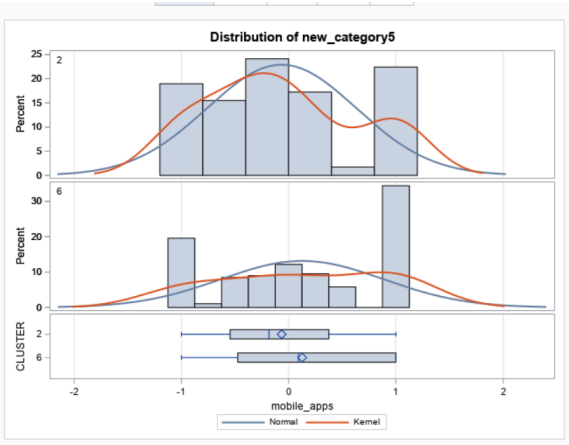
Variable: new_category5 (mobile_apps)

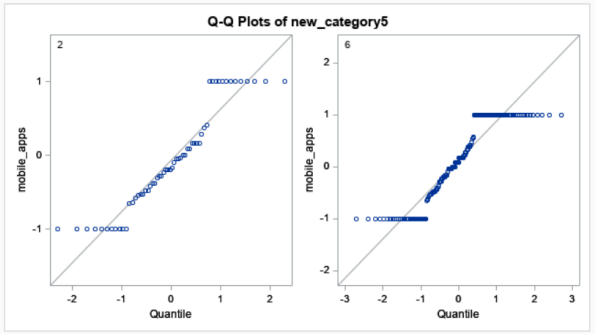
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	-0.0648	0.6968	0.0915	-1.0000	1.0000
6		189	0.1267	0.7588	0.0552	-1.0000	1.0000
Diff (1-2)	Pooled		-0.1916	0.7448	0.1118		
Diff (1-2)	Satterthwaite		-0.1916		0.1069		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		-0.0648	-0.2481	0.1184	0.6968	0.5891	0.8532
6		0.1267	0.0178	0.2356	0.7588	0.6892	0.8441
Diff (1-2)	Pooled	-0.1916	-0.4118	0.0287	0.7448	0.6843	0.8172
Diff (1-2)	Satterthwaite	-0.1916	-0.4035	0.0204			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	-1.71	0.0879
Satterthwaite	Unequal	101.93	-1.79	0.0760

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	57	1.19	0.4568





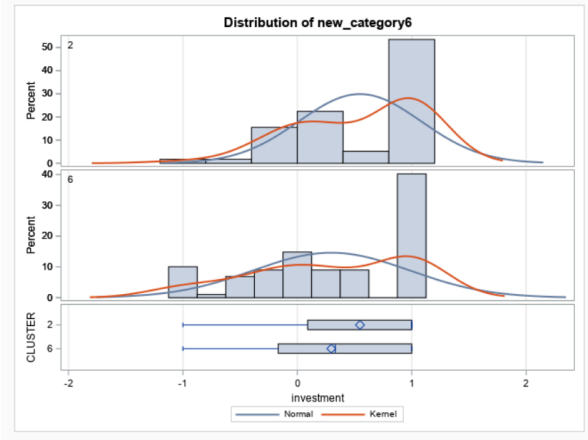
Variable: new_category6 (investment)

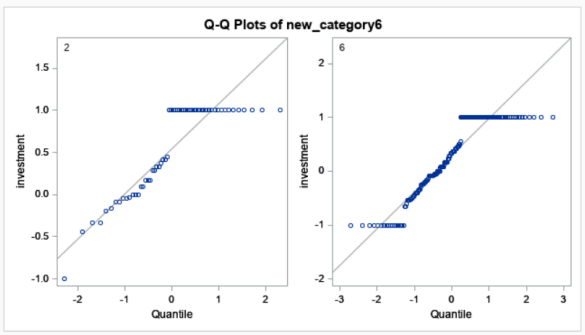
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	0.5479	0.5342	0.0701	-1.0000	1.0000
6		189	0.2977	0.6835	0.0497	-1.0000	1.0000
Diff (1-2)	Pooled		0.2502	0.6518	0.0978		
Diff (1-2)	Satterthwaite		0.2502		0.0860		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		0.5479	0.4074	0.6883	0.5342	0.4516	0.6540
6		0.2977	0.1996	0.3958	0.6835	0.6208	0.7603
Diff (1-2)	Pooled	0.2502	0.0574	0.4429	0.6518	0.5989	0.7151
Diff (1-2)	Satterthwaite	0.2502	0.0799	0.4204			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	2.56	0.0112
Satterthwaite	Unequal	119.52	2.91	0.0043

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	57	1.64	0.0309





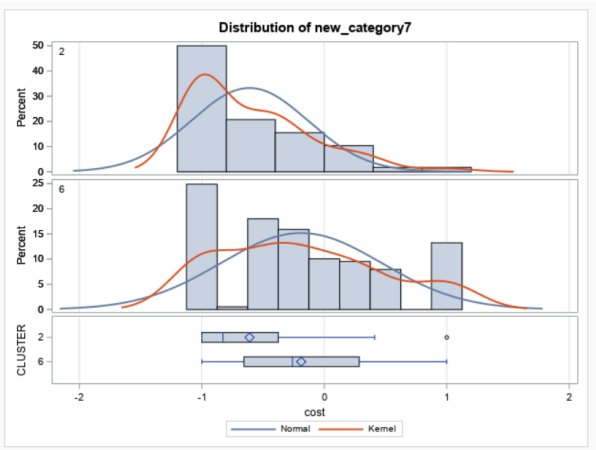
Variable: new_category7 (cost)

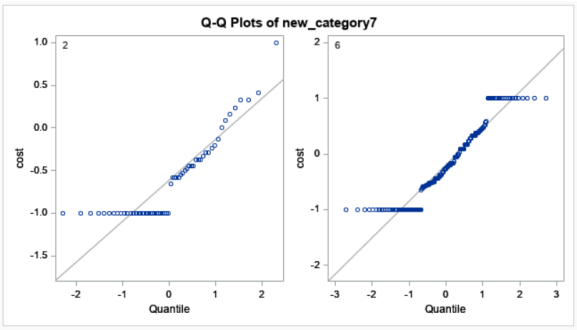
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
2		58	-0.6102	0.4798	0.0630	-1.0000	1.0000
6		189	-0.1884	0.6571	0.0478	-1.0000	1.0000
Diff (1-2)	Pooled		-0.4218	0.6204	0.0931		
Diff (1-2)	Satterthwaite		-0.4218		0.0791		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
2		-0.6102	-0.7363	-0.4840	0.4798	0.4057	0.5875
6		-0.1884	-0.2826	-0.0941	0.6571	0.5969	0.7310
Diff (1-2)	Pooled	-0.4218	-0.6052	-0.2384	0.6204	0.5700	0.6807
Diff (1-2)	Satterthwaite	-0.4218	-0.5783	-0.2653			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	245	-4.53	<.0001
Satterthwaite	Unequal	128.58	-5.33	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	57	1.88	0.0064





4.5.3 Cluster 3

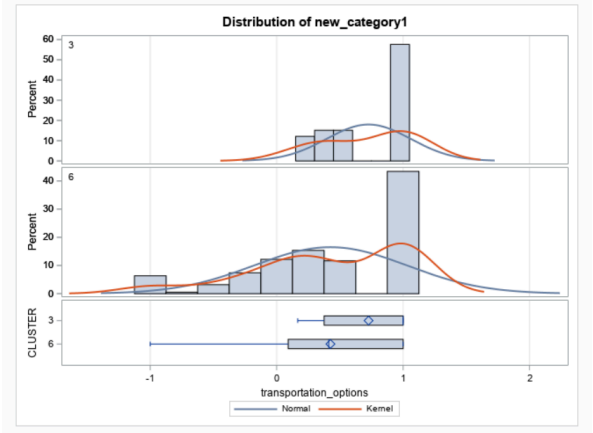
Variable: new_category1 (transportation_options)

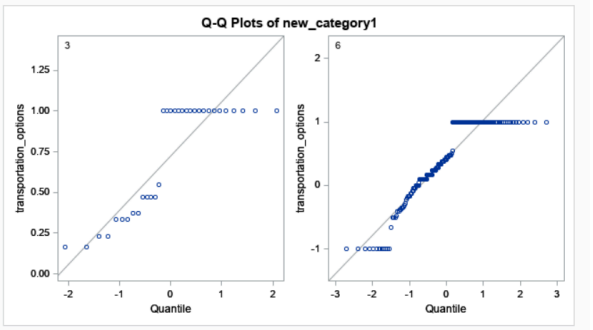
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3		33	0.7268	0.3325	0.0579	0.1667	1.0000
6		189	0.4254	0.6048	0.0440	-1.0000	1.0000
Diff (1-2)	Pooled		0.3014	0.5733	0.1082		
Diff (1-2)	Satterthwaite		0.3014		0.0727		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3		0.7268	0.6089	0.8447	0.3325	0.2674	0.4399
6		0.4254	0.3386	0.5122	0.6048	0.5493	0.6727
Diff (1-2)	Pooled	0.3014	0.0883	0.5146	0.5733	0.5243	0.6323
Diff (1-2)	Satterthwaite	0.3014	0.1566	0.4463			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	220	2.79	0.0058
Satterthwaite	Unequal	75.349	4.15	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	32	3.31	0.0002





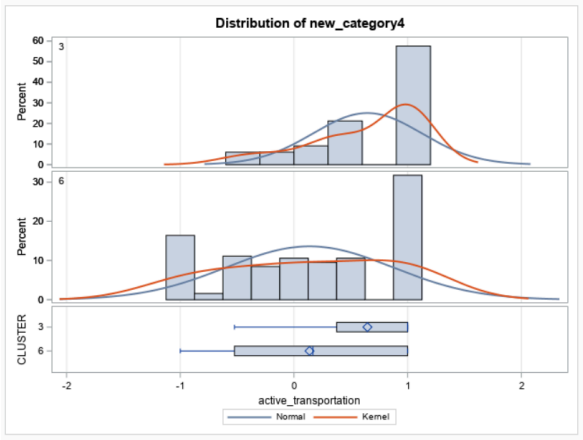
Variable: new_category4 (active_transportation)

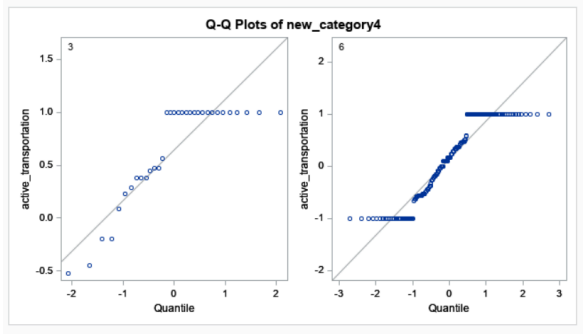
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3		33	0.6461	0.4779	0.0832	-0.5238	1.0000
6		189	0.1366	0.7325	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		0.5095	0.7013	0.1323		
Diff (1-2)	Satterthwaite		0.5095		0.0988		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3		0.6461	0.4767	0.8155	0.4779	0.3843	0.6321
6		0.1366	0.0315	0.2418	0.7325	0.6654	0.8149
Diff (1-2)	Pooled	0.5095	0.2487	0.7702	0.7013	0.6414	0.7735
Diff (1-2)	Satterthwaite	0.5095	0.3120	0.7069			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	220	3.85	0.0002
Satterthwaite	Unequal	61.875	5.16	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	32	2.35	0.0054





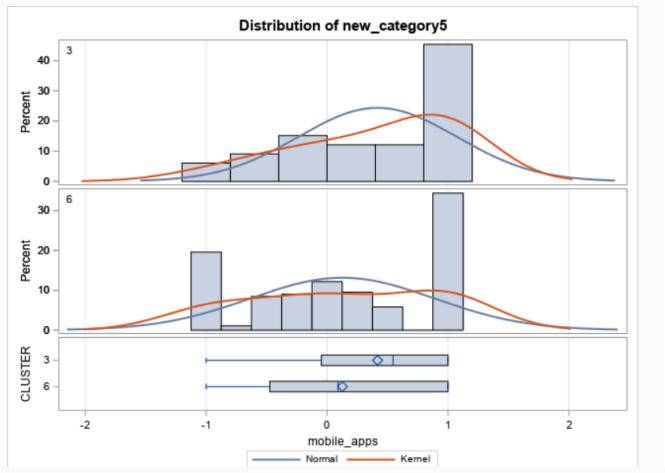
Variable: new_category5 (mobile_apps)

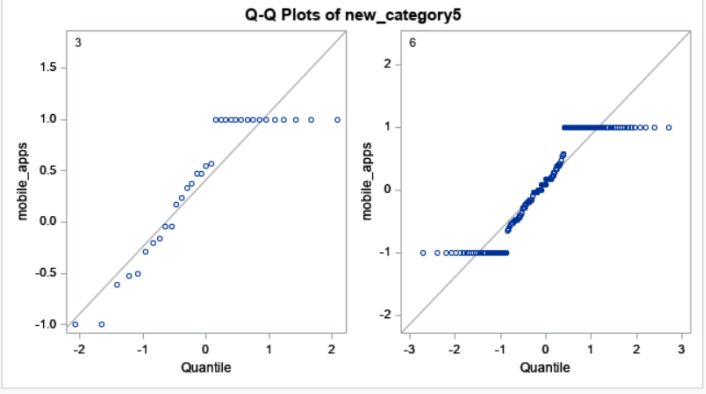
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3		33	0.4175	0.6544	0.1139	-1.0000	1.0000
6		189	0.1267	0.7588	0.0552	-1.0000	1.0000
Diff (1-2)	Pooled		0.2908	0.7445	0.1405		
Diff (1-2)	Satterthwaite		0.2908		0.1266		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3		0.4175	0.1854	0.6495	0.6544	0.5263	0.8656
6		0.1267	0.0178	0.2356	0.7588	0.6892	0.8441
Diff (1-2)	Pooled	0.2908	0.0139	0.5676	0.7445	0.6810	0.8212
Diff (1-2)	Satterthwaite	0.2908	0.0363	0.5452			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	220	2.07	0.0396
Satterthwaite	Unequal	48.334	2.30	0.0260

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	32	1.34	0.3226





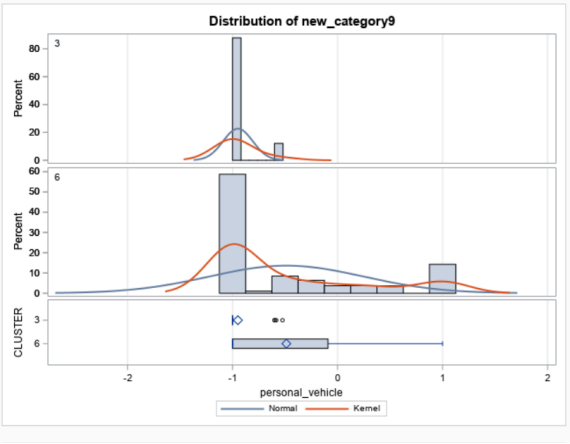
Variable: new_category9 (personal_vehicle)

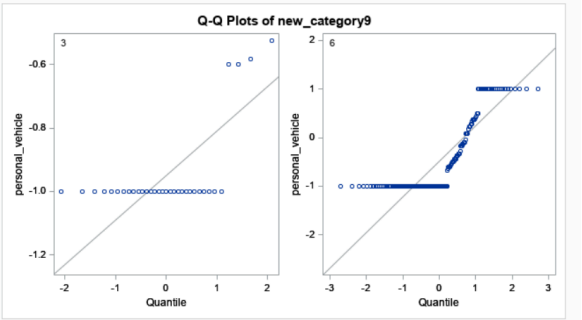
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3		33	-0.9487	0.1407	0.0245	-1.0000	-0.5238
6		189	-0.4875	0.7331	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		-0.4612	0.6798	0.1283		
Diff (1-2)	Satterthwaite		-0.4612		0.0587		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3		-0.9487	-0.9986	-0.8988	0.1407	0.1132	0.1861
6		-0.4875	-0.5927	-0.3823	0.7331	0.6659	0.8155
Diff (1-2)	Pooled	-0.4612	-0.7140	-0.2084	0.6798	0.6218	0.7498
Diff (1-2)	Satterthwaite	-0.4612	-0.5768	-0.3455			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	220	-3.60	0.0004
Satterthwaite	Unequal	218.54	-7.86	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	32	27.14	<.0001





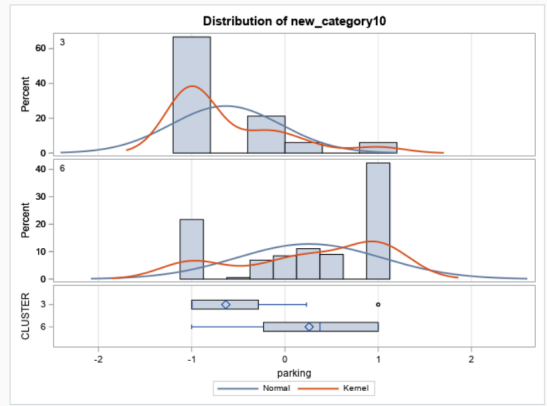
Variable: new_category10 (parking)

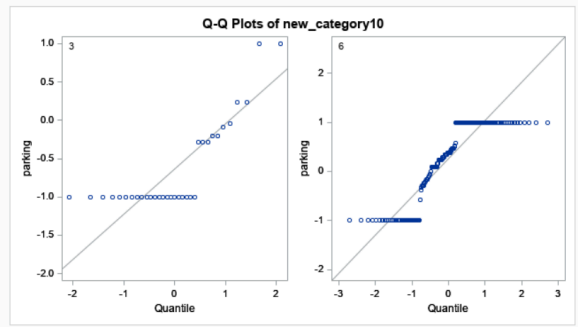
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3		33	-0.6344	0.5906	0.1028	-1.0000	1.0000
6		189	0.2596	0.7799	0.0567	-1.0000	1.0000
Diff (1-2)	Pooled		-0.8940	0.7553	0.1425		
Diff (1-2)	Satterthwaite		-0.8940		0.1174		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3		-0.6344	-0.8438	-0.4249	0.5906	0.4750	0.7812
6		0.2596	0.1477	0.3715	0.7799	0.7084	0.8676
Diff (1-2)	Pooled	-0.8940	-1.1748	-0.6131	0.7553	0.6909	0.8332
Diff (1-2)	Satterthwaite	-0.8940	-1.1294	-0.6585			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	220	-6.27	<.0001
Satterthwaite	Unequal	53.607	-7.61	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	32	1.74	0.0637





4.5.4 Cluster 4

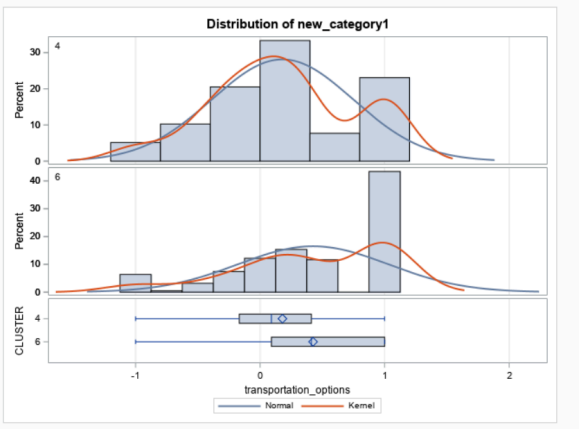
Variable: new_category1 (transportation_options)

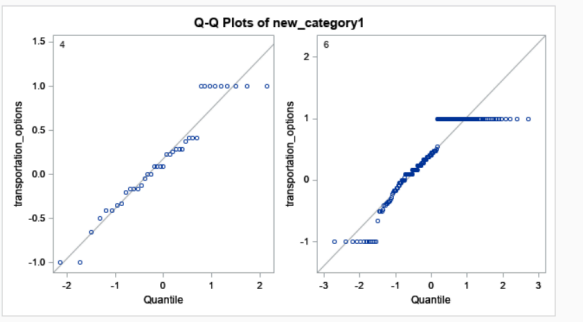
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	0.1797	0.5675	0.0909	-1.0000	1.0000
6		189	0.4254	0.6048	0.0440	-1.0000	1.0000
Diff (1-2)	Pooled		-0.2457	0.5987	0.1053		
Diff (1-2)	Satterthwaite		-0.2457		0.1010		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		0.1797	-0.00431	0.3636	0.5675	0.4638	0.7314
6		0.4254	0.3386	0.5122	0.6048	0.5493	0.6727
Diff (1-2)	Pooled	-0.2457	-0.4532	-0.0382	0.5987	0.5482	0.6594
Diff (1-2)	Satterthwaite	-0.2457	-0.4479	-0.0436			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	-2.33	0.0205
Satterthwaite	Unequal	57.257	-2.43	0.0181

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	38	1.14	0.6595





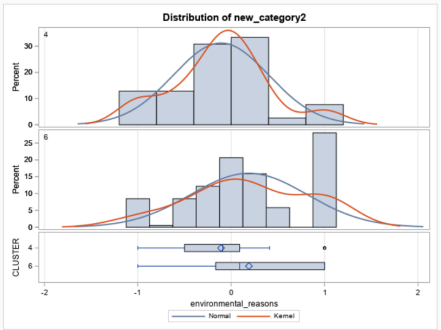
Variable: new_category2 (environmental_reasons)

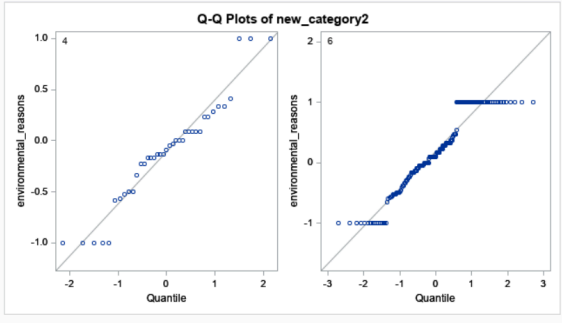
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	-0.1090	0.5119	0.0820	-1.0000	1.0000
6		189	0.1894	0.6226	0.0453	-1.0000	1.0000
Diff (1-2)	Pooled		-0.2984	0.6054	0.1065		
Diff (1-2)	Satterthwaite		-0.2984		0.0936		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		-0.1090	-0.2749	0.0570	0.5119	0.4183	0.6597
6		0.1894	0.1001	0.2788	0.6226	0.5655	0.6926
Diff (1-2)	Pooled	-0.2984	-0.5082	-0.0886	0.6054	0.5543	0.6668
Diff (1-2)	Satterthwaite	-0.2984	-0.4855	-0.1113			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	-2.80	0.0055
Satterthwaite	Unequal	63.544	-3.19	0.0022

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	38	1.48	0.1528





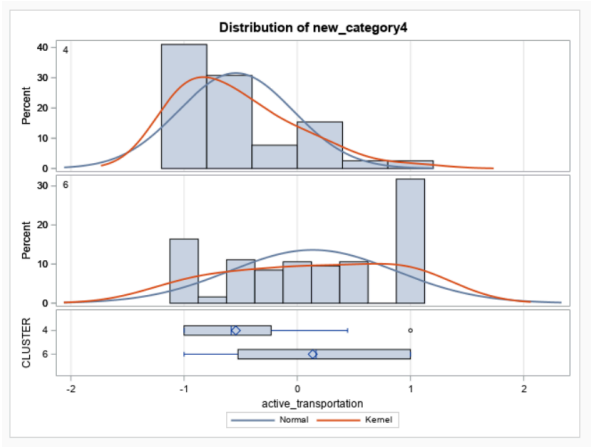
Variable: new_category4 (active_transportation)

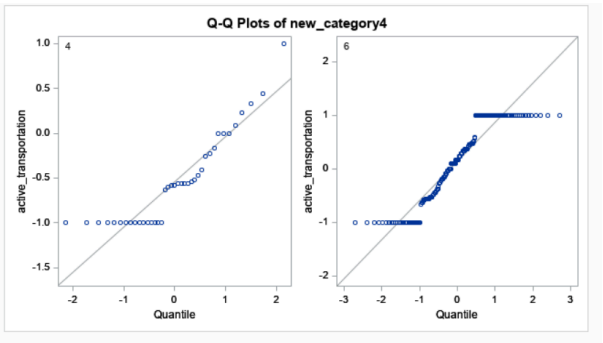
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	-0.5428	0.5055	0.0809	-1.0000	1.0000
6		189	0.1366	0.7325	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		-0.6794	0.6995	0.1230		
Diff (1-2)	Satterthwaite		-0.6794		0.0969		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		-0.5428	-0.7066	-0.3789	0.5055	0.4131	0.6514
6		0.1366	0.0315	0.2418	0.7325	0.6654	0.8149
Diff (1-2)	Pooled	-0.6794	-0.9219	-0.4370	0.6995	0.6406	0.7706
Diff (1-2)	Satterthwaite	-0.6794	-0.8725	-0.4864			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	-5.52	<.0001
Satterthwaite	Unequal	75.221	-7.01	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	38	2.10	0.0081





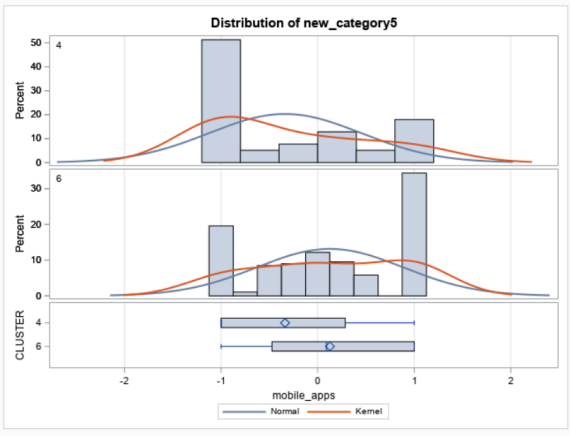
Variable: new_category5 (mobile_apps)

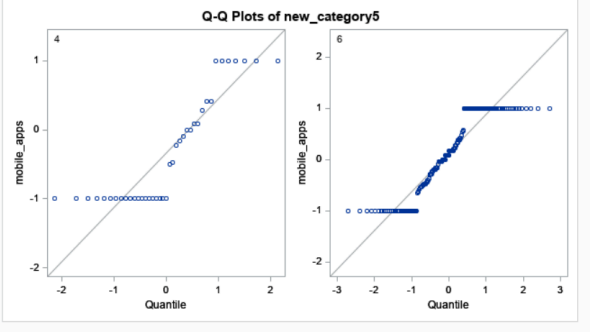
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	-0.3377	0.7873	0.1261	-1.0000	1.0000
6		189	0.1267	0.7588	0.0552	-1.0000	1.0000
Diff (1-2)	Pooled		-0.4644	0.7636	0.1343		
Diff (1-2)	Satterthwaite		-0.4644		0.1376		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		-0.3377	-0.5929	-0.0825	0.7873	0.6434	1.0146
6		0.1267	0.0178	0.2356	0.7588	0.6892	0.8441
Diff (1-2)	Pooled	-0.4644	-0.7291	-0.1998	0.7636	0.6993	0.8412
Diff (1-2)	Satterthwaite	-0.4644	-0.7404	-0.1885			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	-3.46	0.0007
Satterthwaite	Unequal	53.566	-3.37	0.0014

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	38	188	1.08	0.7260





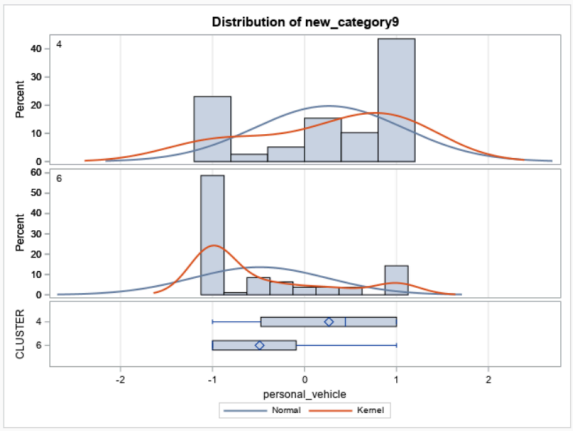
Variable: new_category9 (personal_vehicle)

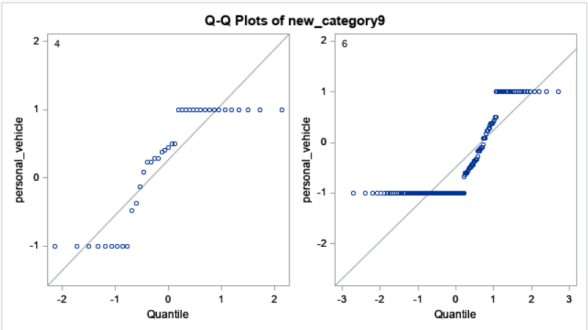
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	0.2661	0.8096	0.1296	-1.0000	1.0000
6		189	-0.4875	0.7331	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		0.7536	0.7465	0.1313		
Diff (1-2)	Satterthwaite		0.7536		0.1402		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		0.2661	0.00361	0.5285	0.8096	0.6616	1.0434
6		-0.4875	-0.5927	-0.3823	0.7331	0.6659	0.8155
Diff (1-2)	Pooled	0.7536	0.4949	1.0123	0.7465	0.6836	0.8223
Diff (1-2)	Satterthwaite	0.7536	0.4722	1.0349			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	5.74	<.0001
Satterthwaite	Unequal	51.647	5.38	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	38	188	1.22	0.3897





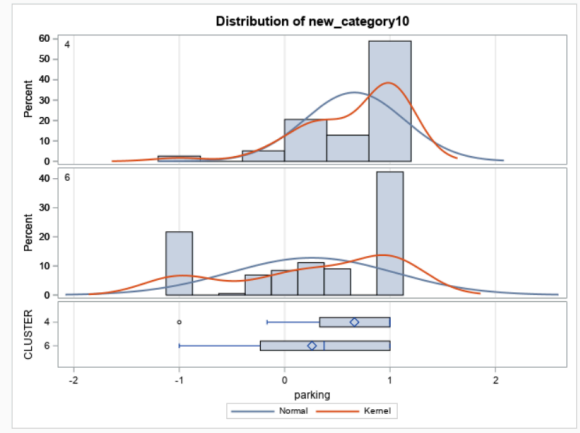
Variable: new_category10 (parking)

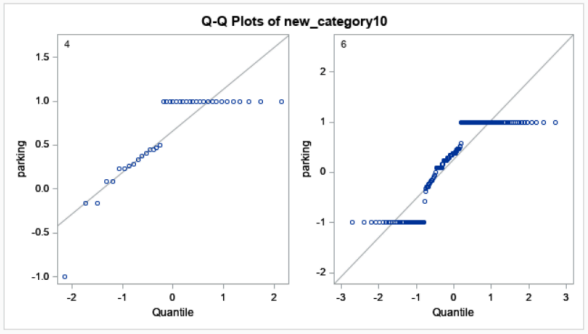
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
4		39	0.6625	0.4730	0.0757	-1.0000	1.0000
6		189	0.2596	0.7799	0.0567	-1.0000	1.0000
Diff (1-2)	Pooled		0.4029	0.7373	0.1297		
Diff (1-2)	Satterthwaite		0.4029		0.0946		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
4		0.6625	0.5092	0.8158	0.4730	0.3865	0.6096
6		0.2596	0.1477	0.3715	0.7799	0.7084	0.8676
Diff (1-2)	Pooled	0.4029	0.1474	0.6584	0.7373	0.6751	0.8122
Diff (1-2)	Satterthwaite	0.4029	0.2148	0.5910			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	226	3.11	0.0021
Satterthwaite	Unequal	87.066	4.26	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	38	2.72	0.0005





4.5.5 Cluster 5

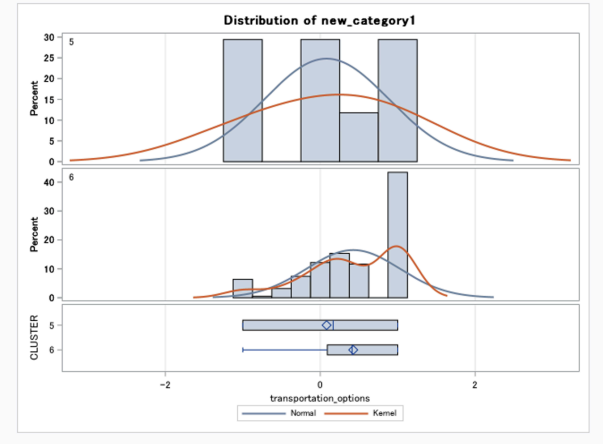
Variable: new_category1 (transportation_options)

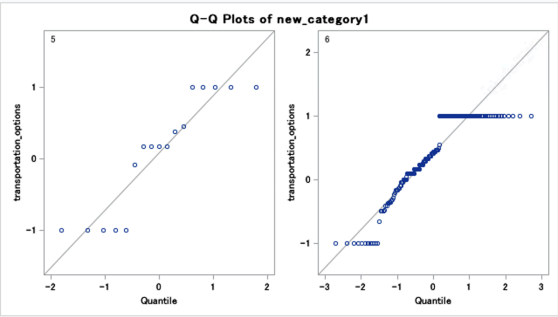
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	0.0821	0.8040	0.1950	-1.0000	1.0000
6		189	0.4254	0.6048	0.0440	-1.0000	1.0000
Diff (1-2)	Pooled		-0.3433	0.6227	0.1577		
Diff (1-2)	Satterthwaite		-0.3433		0.1999		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		0.0821	-0.3313	0.4955	0.8040	0.5988	1.2237
6		0.4254	0.3386	0.5122	0.6048	0.5493	0.6727
Diff (1-2)	Pooled	-0.3433	-0.6542	-0.0324	0.6227	0.5677	0.6896
Diff (1-2)	Satterthwaite	-0.3433	-0.7639	0.0772			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-2.18	0.0306
Satterthwaite	Unequal	17.666	-1.72	0.1034

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.77	0.0764





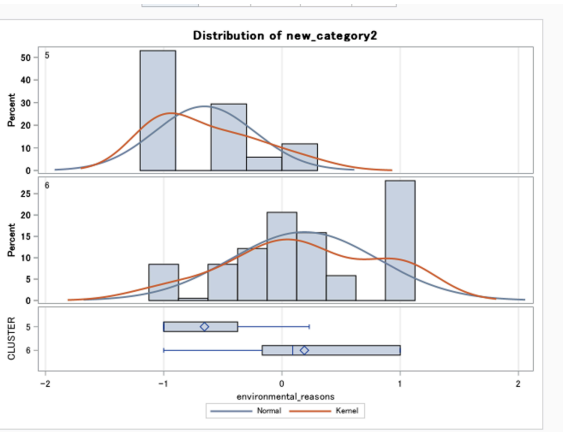
Variable: new_category2 (environmental_reasons)

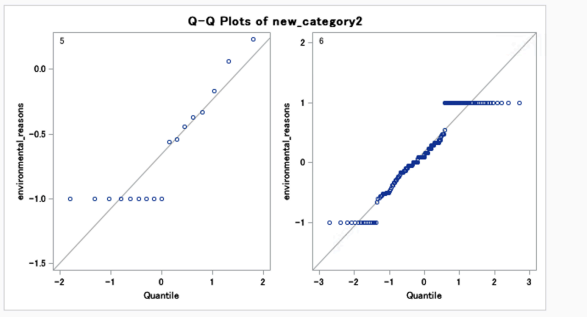
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	-0.6551	0.4224	0.1025	-1.0000	0.2308
6		189	0.1894	0.6226	0.0453	-1.0000	1.0000
Diff (1-2)	Pooled		-0.8446	0.6093	0.1543		
Diff (1-2)	Satterthwaite		-0.8446		0.1120		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		-0.6551	-0.8723	-0.4379	0.4224	0.3146	0.6429
6		0.1894	0.1001	0.2788	0.6226	0.5655	0.6926
Diff (1-2)	Pooled	-0.8446	-1.1487	-0.5404	0.6093	0.5554	0.6747
Diff (1-2)	Satterthwaite	-0.8446	-1.0764	-0.6127			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-5.47	<.0001
Satterthwaite	Unequal	22.788	-7.54	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	16	2.17	0.0749





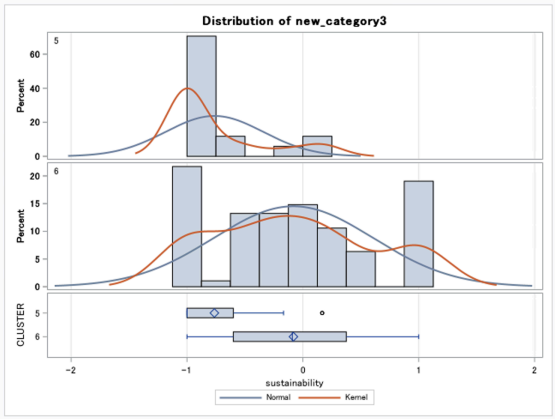
Variable: new_category3 (sustainability)

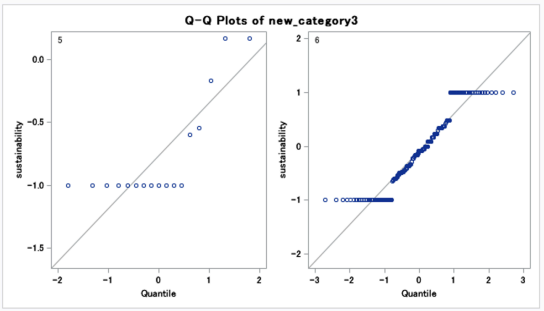
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	-0.7635	0.4207	0.1020	-1.0000	0.1667
6		189	-0.0815	0.6868	0.0500	-1.0000	1.0000
Diff (1-2)	Pooled		-0.6819	0.6698	0.1696		
Diff (1-2)	Satterthwaite		-0.6819		0.1136		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		-0.7635	-0.9798	-0.5471	0.4207	0.3133	0.6403
6		-0.0815	-0.1801	0.0170	0.6868	0.6239	0.7640
Diff (1-2)	Pooled	-0.6819	-1.0163	-0.3475	0.6698	0.6106	0.7417
Diff (1-2)	Satterthwaite	-0.6819	-0.9162	-0.4477			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-4.02	<.0001
Satterthwaite	Unequal	24.47	-6.00	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	16	2.66	0.0267





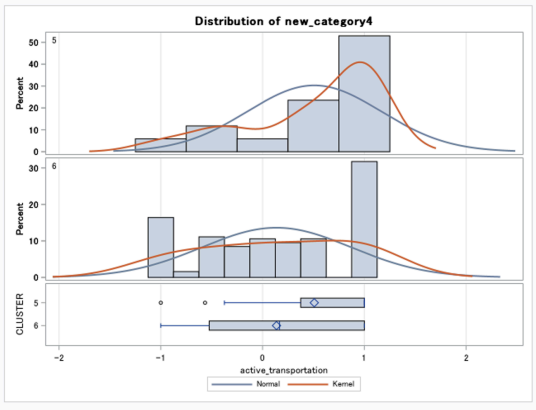
Variable: new_category4 (active_transportation)

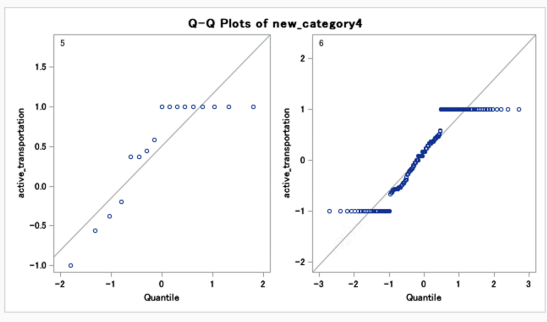
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	0.5081	0.6581	0.1596	-1.0000	1.0000
6		189	0.1366	0.7325	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		0.3715	0.7270	0.1841		
Diff (1-2)	Satterthwaite		0.3715		0.1683		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		0.5081	0.1697	0.8465	0.6581	0.4902	1.0017
6		0.1366	0.0315	0.2418	0.7325	0.6654	0.8149
Diff (1-2)	Pooled	0.3715	0.00851	0.7344	0.7270	0.6628	0.8051
Diff (1-2)	Satterthwaite	0.3715	0.0201	0.7228			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	2.02	0.0449
Satterthwaite	Unequal	19.744	2.21	0.0393

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	16	1.24	0.6471





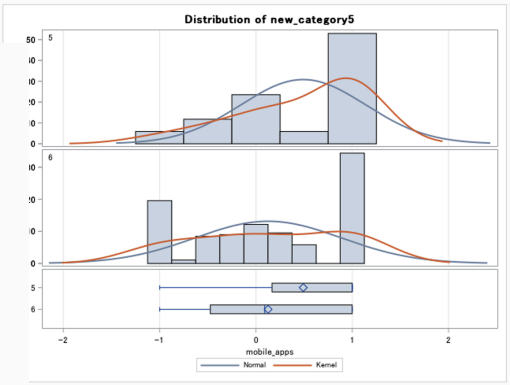
Variable: new_category5 (mobile_apps)

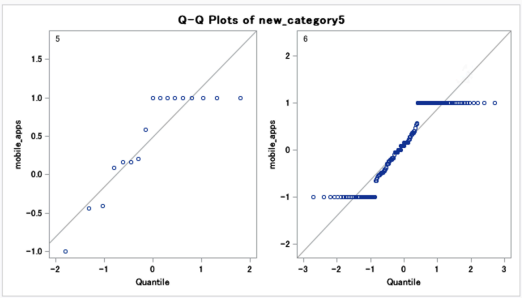
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	0.4913	0.6475	0.1570	-1.0000	1.0000
6		189	0.1267	0.7588	0.0552	-1.0000	1.0000
Diff (1-2)	Pooled		0.3645	0.7507	0.1901		
Diff (1-2)	Satterthwaite		0.3645		0.1665		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		0.4913	0.1583	0.8242	0.6475	0.4823	0.9855
6		0.1267	0.0178	0.2356	0.7588	0.6892	0.8441
Diff (1-2)	Pooled	0.3645	-0.0102	0.7393	0.7507	0.6843	0.8313
Diff (1-2)	Satterthwaite	0.3645	0.0175	0.7116			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	1.92	0.0565
Satterthwaite	Unequal	20.17	2.19	0.0405

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	188	16	1.37	0.4740





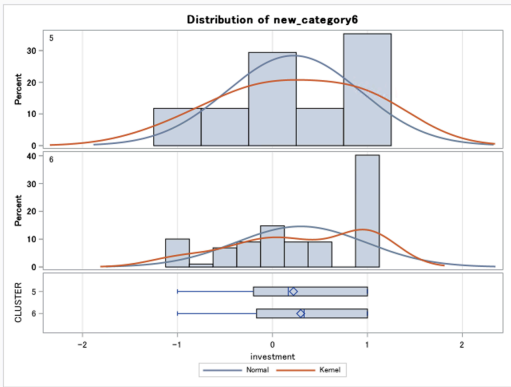
Variable: new_category6 (investment)

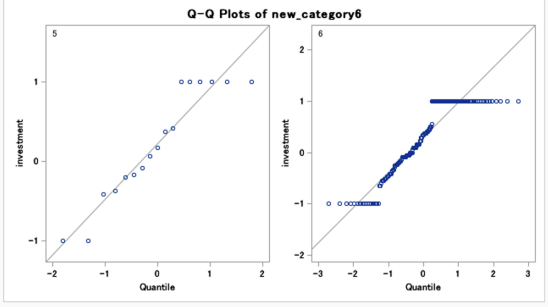
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	0.2219	0.7023	0.1703	-1.0000	1.0000
6		189	0.2977	0.6835	0.0497	-1.0000	1.0000
Diff (1-2)	Pooled		-0.0758	0.6850	0.1734		
Diff (1-2)	Satterthwaite		-0.0758		0.1774		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		0.2219	-0.1392	0.5829	0.7023	0.5230	1.0688
6		0.2977	0.1996	0.3958	0.6835	0.6208	0.7603
Diff (1-2)	Pooled	-0.0758	-0.4178	0.2661	0.6850	0.6245	0.7586
Diff (1-2)	Satterthwaite	-0.0758	-0.4474	0.2958			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-0.44	0.6624
Satterthwaite	Unequal	18.831	-0.43	0.6739

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.06	0.8014





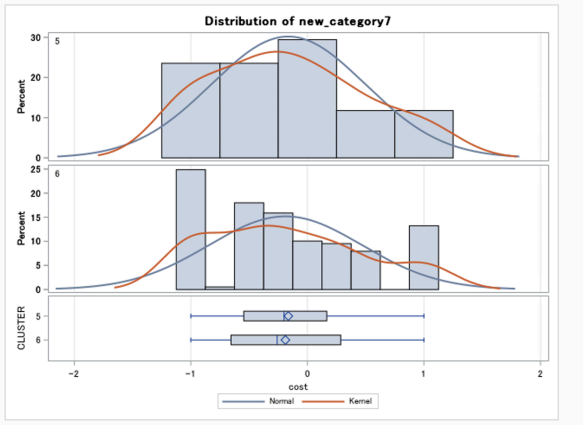
Variable: new_category7 (cost)

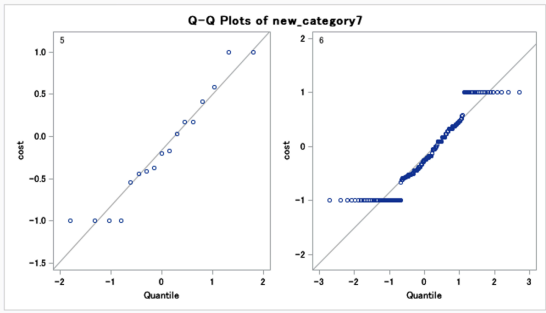
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	-0.1637	0.6607	0.1602	-1.0000	1.0000
6		189	-0.1884	0.6571	0.0478	-1.0000	1.0000
Diff (1-2)	Pooled		0.0247	0.6574	0.1665		
Diff (1-2)	Satterthwaite		0.0247		0.1672		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		-0.1637	-0.5034	0.1760	0.6607	0.4921	1.0055
6		-0.1884	-0.2826	-0.0941	0.6571	0.5969	0.7310
Diff (1-2)	Pooled	0.0247	-0.3035	0.3529	0.6574	0.5993	0.7280
Diff (1-2)	Satterthwaite	0.0247	-0.3254	0.3747			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	0.15	0.8823
Satterthwaite	Unequal	18.961	0.15	0.8843

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.01	0.8941





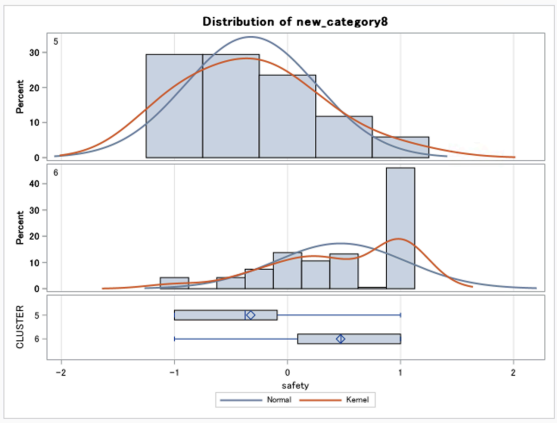
Variable: new_category8 (safety)

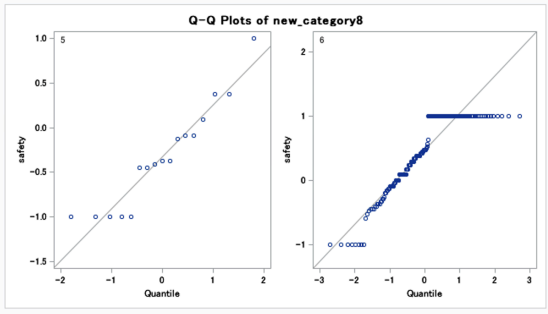
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	-0.3248	0.5796	0.1406	-1.0000	1.0000
6		189	0.4696	0.5783	0.0421	-1.0000	1.0000
Diff (1-2)	Pooled		-0.7944	0.5784	0.1465		
Diff (1-2)	Satterthwaite		-0.7944		0.1467		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		-0.3248	-0.6228	-0.0268	0.5796	0.4317	0.8821
6		0.4696	0.3866	0.5526	0.5783	0.5253	0.6433
Diff (1-2)	Pooled	-0.7944	-1.0832	-0.5056	0.5784	0.5273	0.6406
Diff (1-2)	Satterthwaite	-0.7944	-1.1016	-0.4873			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-5.42	<.0001
Satterthwaite	Unequal	18.981	-5.41	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.00	0.9077





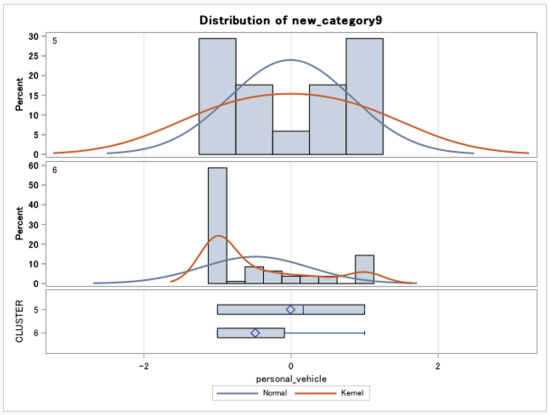
Variable: new_category9 (personal_vehicle)

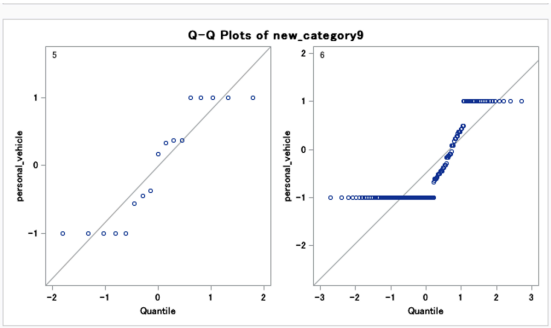
CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	-0.00792	0.8320	0.2018	-1.0000	1.0000
6		189	-0.4875	0.7331	0.0533	-1.0000	1.0000
Diff (1-2)	Pooled		0.4796	0.7413	0.1877		
Diff (1-2)	Satterthwaite		0.4796		0.2087		

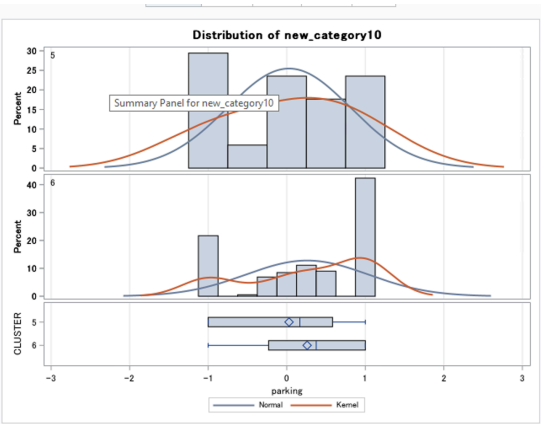
CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		-0.00792	-0.4357	0.4199	0.8320	0.6197	1.2663
6		-0.4875	-0.5927	-0.3823	0.7331	0.6659	0.8155
Diff (1-2)	Pooled	0.4796	0.1095	0.8497	0.7413	0.6758	0.8210
Diff (1-2)	Satterthwaite	0.4796	0.0416	0.9176			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	2.55	0.0113
Satterthwaite	Unequal	18.305	2.30	0.0336

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.29	0.4162







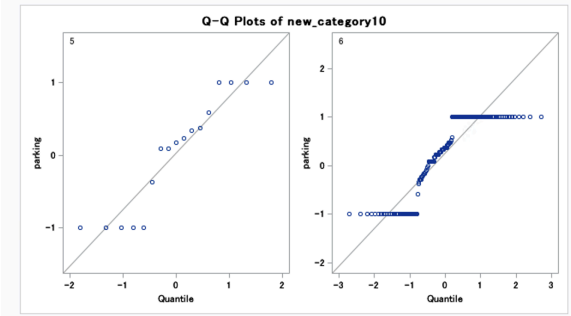
Variable: new_category10 (parking)

CLUSTER	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
5		17	0.0292	0.7833	0.1900	-1.0000	1.0000
6		189	0.2596	0.7799	0.0567	-1.0000	1.0000
Diff (1-2)	Pooled		-0.2304	0.7802	0.1975		
Diff (1-2)	Satterthwaite		-0.2304		0.1983		

CLUSTER	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5		0.0292	-0.3735	0.4319	0.7833	0.5833	1.1921
6		0.2596	0.1477	0.3715	0.7799	0.7084	0.8676
Diff (1-2)	Pooled	-0.2304	-0.6199	0.1591	0.7802	0.7113	0.8640
Diff (1-2)	Satterthwaite	-0.2304	-0.6454	0.1846			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	204	-1.17	0.2448
Satterthwaite	Unequal	18.968	-1.16	0.2595

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	16	188	1.01	0.8990



4.6 Chi-Square Test

Within the provided SAS code, categorical variables such as transportation options, environmental reasons, and demographic characteristics are likely subjected to the chi-square test. These variables represent factors that may influence cluster memberships and behaviors within each cluster.

table cluster*age / expected chisq;

The 'expected' and 'chisq' options in the 'proc freq' statement instruct SAS to calculate expected counts and perform a chi-square test of independence between the 'cluster' and 'age' variables. The chi-square test assesses whether there is a statistically significant association between the 'cluster' and 'age' variables.

The chi-square test yields a p-value that indicates the probability of observing the observed association between variables and clusters under the null hypothesis. A low p-value (< 0.05) suggests that the association is statistically significant, indicating that the variables are not independent of cluster memberships.

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by age

	age(age)						
CLUSTER	18-25	26-35	36-45	46-55	56-65	<18	Total
1	21 20.667 11.11 50.00 22.58	12 12.889 6.35 28.57 20.69	0 2 0.00 0.00 0.00	4 2.4444 2.12 9.52 36.36	5 3.7778 2.65 11.90 29.41	0 0.2222 0.00 0.00 0.00	42 22.22
2	30 28.54 15.87 51.72 32.26	15 17.799 7.94 25.86 25.86	3 2.7619 1.59 5.17 33.33	2 3.3757 1.06 3.45 18.18	8 5.2169 4.23 13.79 47.06	0 0.3069 0.00 0.00 0.00	58 30.69
3	17 16.238 8.99 51.52 18.28	13 10.127 6.88 39.39 22.41	1 1.5714 0.53 3.03 11.11	1 1.9206 0.53 3.03 9.09	1 2.9683 0.53 3.03 5.88	0 0.1746 0.00 0.00 0.00	33 17.46
4	17 19.19 8.99 43.59 18.28	13 11.968 6.88 33.33 22.41	3 1.8571 1.59 7.69 33.33	3 2.2698 1.59 7.69 27.27	3 3.5079 1.59 7.69 17.65	0 0.2063 0.00 0.00 0.00	39 20.63
5	8 8.3651 4.23 47.06 8.60	5 5.2169 2.65 29.41 8.62	2 0.8095 1.06 11.76 22.22	1 0.9894 0.53 5.88 9.09	0 1.5291 0.00 0.00 0.00	1 0.0899 0.53 5.88 100.00	17 8.99
Total	93 49.21	58 30.69	9 4.76	11 5.82	17 8.99	1 0.53	189 100.00

Statistics for Table of CLUSTER by age

Statistic	DF	Value	Prob
Chi-Square	20	23.6110	0.2598
Likelihood Ratio Chi-Square	20	21.4303	0.3722
Mantel-Haenszel Chi-Square	1	0.0834	0.7727
Phi Coefficient		0.3534	
Contingency Coefficient		0.3332	
Cramer's V		0.1767	

WARNING: 63% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by sex

CLUSTER	sex(sex) Female	Male	Total
1	16 17.556 8.47 38.10 20.25	26 24.444 13.76 61.90 23.64	42 22.22
2	23 24.243 12.17 39.66 29.11	35 33.757 18.52 60.34 31.82	58 30.69
3	15 13.794 7.94 45.45 18.99	18 19.206 9.52 54.55 16.36	33 17.46
4	18 16.302 9.52 46.15 22.78	21 22.698 11.11 53.85 19.09	39 20.63
5	7 7.1058 3.70 41.18 8.86	10 9.8942 5.29 58.82 9.09	17 8.99
Total	79 41.80	110 58.20	189 100.00

Statistics for Table of CLUSTER by sex

Statistic	DF	Value	Prob
Chi-Square	4	0.8344	0.9338
Likelihood Ratio Chi-Square	4	0.8333	0.9339
Mantel-Haenszel Chi-Square	1	0.4532	0.5008
Phi Coefficient		0.0664	
Contingency Coefficient		0.0663	
Cramer's V		0.0664	

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by continent

CLUSTER	continent(continent) Africa	Asia	Europe	North America	Total
1	6 8.8889 3.17 14.29 15.00	8 9.7778 4.23 19.05 18.18	25 21.778 13.23 59.52 25.51	3 1.5556 1.59 7.14 42.86	42 22.22
2	15 12.275 7.94 25.86 37.50	15 13.503 7.94 25.86 34.09	26 30.074 13.76 44.83 26.53	2 2.1481 1.06 3.45 28.57	58 30.69
3	6 6.9841 3.17 18.18 15.00	5 7.6825 2.65 15.15 11.36	21 17.111 11.11 63.64 21.43	1 1.2222 0.53 3.03 14.29	33 17.46
4	8 8.254 4.23 20.51 20.00	12 9.0794 6.35 30.77 27.27	18 20.222 9.52 46.15 18.37	1 1.4444 0.53 2.56 14.29	39 20.63
5	5 3.5979 2.65 29.41 12.50	4 3.9577 2.12 23.53 9.09	8 8.8148 4.23 47.06 8.16	0 0.6296 0.00 0.00 0.00	17 8.99
Total	40 21.16	44 23.28	98 51.85	7 3.70	189 100.00

Statistics for Table of CLUSTER by continent

Statistic	DF	Value	Prob
Chi-Square	12	8.9929	0.7035
Likelihood Ratio Chi-Square	12	9.4005	0.6684
Mantel-Haenszel Chi-Square	1	1.8912	0.1691
Phi Coefficient		0.2181	
Contingency Coefficient		0.2131	
Cramer's V		0.1259	

WARNING: 35% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by work

CLUSTER	work(work) Employed	Prefer not to say	Retired	Student	Unemployed	Total
1	17 16.222 8.99 40.48 23.29	1 1.5556 0.53 2.38 14.29	3 2 1.59 7.14 33.33	17 19.556 8.99 40.48 19.32	4 2.6667 2.12 9.52 33.33	42 22.22
2	22 22.402 11.64 37.93 30.14	2 2.1481 1.06 3.45 28.57	3 2.7619 1.59 5.17 33.33	27 27.005 14.29 46.55 30.68	4 3.6825 2.12 6.90 33.33	58 30.69
3	13 12.746 6.88 39.39 17.81	1 1.2222 0.53 3.03 14.29	0 1.5714 0.00 0.00 0.00	16 15.365 8.47 48.48 18.18	3 2.0952 1.59 9.09 25.00	33 17.46
4	15 15.063 7.94 38.46 20.55	2 1.4444 1.06 5.13 28.57	2 1.8571 1.06 5.13 22.22	19 18.159 10.05 48.72 21.59	1 2.4762 0.53 2.56 8.33	39 20.63
5	6 6.5661 3.17 35.29 8.22	1 0.6296 0.53 5.88 14.29	1 0.8095 0.53 5.88 11.11	9 7.9153 4.76 52.94 10.23	0 1.0794 0.00 0.00 0.00	17 8.99
Total	73 38.62	7 3.70	9 4.76	88 46.56	12 6.35	189 100.00

Statistics for Table of CLUSTER by work

Statistic	DF	Value	Prob
Chi-Square	16	6.5189	0.9815
Likelihood Ratio Chi-Square	16	9.2052	0.9047
Mantel-Haenszel Chi-Square	1	0.0126	0.9105
Phi Coefficient		0.1857	
Contingency Coefficient		0.1826	
Cramer's V		0.0929	

WARNING: 60% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by primary_transportation

CLUSTER	primary_transportation(primary_transportation) Bicycle	Car	Motorcycle	Public transit (bus, train, subway)	Walking	Total
1	2 1.5556 1.06 4.76 28.57	13 11.333 6.88 30.95 25.49	5 2.2222 2.65 11.90 50.00	20 24.222 10.58 47.62 18.35	2 2.6667 1.06 4.76 16.67	42 22.22
2	2 2.1481 1.06 3.45 28.57	15 15.651 7.94 25.86 29.41	2 3.0688 1.06 3.45 20.00	34 33.45 17.99 58.62 31.19	5 3.6825 2.65 8.62 41.67	58 30.69
3	1 1.2222 0.53 3.03 14.29	10 8.9048 5.29 30.30 19.61	1 1.746 0.53 3.03 10.00	19 19.032 10.05 57.58 17.43	2 2.0952 1.06 6.06 16.67	33 17.46
4	1 1.4444 0.53 2.56 14.29	9 10.524 4.76 23.08 17.65	1 2.0635 0.53 2.56 10.00	25 22.492 13.23 64.10 22.94	3 2.4762 1.59 7.69 25.00	39 20.63
5	1 0.6296 0.53 5.88 14.29	4 4.5873 2.12 23.53 7.84	1 0.8995 0.53 5.88 10.00	11 9.8042 5.82 64.71 10.09	0 1.0794 0.00 0.00 0.00	17 8.99
Total	7 3.70	51 26.98	10 5.29	109 57.67	12 6.35	189 100.00

Statistics for Table of CLUSTER by primary_transportation

Statistic	DF	Value	Prob
Chi-Square	16	8.9606	0.9150
Likelihood Ratio Chi-Square	16	9.3296	0.8992
Mantel-Haenszel Chi-Square	1	0.7231	0.3951
Phi Coefficient		0.2177	
Contingency Coefficient		0.2128	
Cramer's V		0.1089	

WARNING: 64% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by distance

CLUSTER	distance(distance)						
	16-32 km (10-20 miles)	32-48 km (20-30 miles)	8-16 km (5-10 miles)	I do not have a daily commute	Less than 8 km (5 miles)	More than 48 km (30 miles)	Total
1	8 5.7778 4.23 19.05 30.77	2 3.1111 1.06 4.76 14.29	10 11.556 5.29 23.81 19.23	5 2.8889 2.65 11.90 38.46	15 16 7.94 35.71 20.83	2 2.6667 1.06 4.76 16.67	42 22.22
2	7 7.9788 3.70 12.07 26.92	4 4.2963 2.12 6.90 28.57	16 15.958 8.47 27.59 30.77	3 3.9894 1.59 5.17 23.08	22 22.095 11.64 37.93 30.56	6 3.6825 3.17 10.34 50.00	58 30.69
3	2 4.5397 1.06 6.06 7.69	4 2.4444 2.12 12.12 28.57	7 9.0794 3.70 21.21 13.46	1 2.2698 0.53 3.03 7.69	16 12.571 8.47 48.48 22.22	3 2.0952 1.59 9.09 25.00	33 17.46
4	5 5.3651 2.65 12.82 19.23	3 2.8889 1.59 7.69 21.43	13 10.73 6.88 33.33 25.00	3 2.6825 1.59 7.69 23.08	14 14.857 7.41 35.90 19.44	1 2.4762 0.53 2.56 8.33	39 20.63
5	4 2.3386 2.12 23.53 15.38	1 1.2593 0.53 5.88 7.14	6 4.6772 3.17 35.29 11.54	1 1.1693 0.53 5.88 7.69	5 6.4762 2.65 29.41 6.94	0 1.0794 0.00 0.00 0.00	17 8.99
Total	26 13.76	14 7.41	52 27.51	13 6.88	72 38.10	12 6.35	189 100.00

Statistics for Table of CLUSTER by distance

Statistic	DF	Value	Prob
Chi-Square	20	14.5251	0.8029
Likelihood Ratio Chi-Square	20	15.4623	0.7494
Mantel-Haenszel Chi-Square	1	0.6473	0.4211
Phi Coefficient		0.2772	
Contingency Coefficient		0.2671	
Cramer's V		0.1386	

WARNING: 60% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by time

CLUSTER	time(time)						
	15-30 minutes	30-60 minutes	60-90 minutes	I do not have a daily commute	Less than 15 minutes	More than 90 minutes	Total
1	7 8.6667 3.70 16.67 17.95	13 15.333 6.88 30.95 18.84	8 7.3333 4.23 19.05 24.24	6 3.3333 3.17 14.29 40.00	3 2.8889 1.59 7.14 23.08	5 4.4444 2.65 11.90 25.00	42 22.22
2	10 11.968 5.29 17.24 25.64	26 21.175 13.76 44.83 37.68	9 10.127 4.76 15.52 27.27	4 4.6032 2.12 6.90 26.67	3 3.9894 1.59 5.17 23.08	6 6.1376 3.17 10.34 30.00	58 30.69
3	7 6.8095 3.70 21.21 17.95	14 12.048 7.41 42.42 20.29	5 5.7619 2.65 15.15 15.15	0 2.619 0.00 0.00 0.00	3 2.2698 1.59 9.09 23.08	4 3.4921 2.12 12.12 20.00	33 17.46
4	11 8.0476 5.82 28.21 28.21	9 14.238 4.76 23.08 13.04	8 6.8095 4.23 20.51 24.24	3 3.0952 1.59 7.69 20.00	3 2.6825 1.59 7.69 23.08	5 4.127 2.65 12.82 25.00	39 20.63
5	4 3.5079 2.12 23.53 10.26	7 6.2063 3.70 41.18 10.14	3 2.9683 1.59 17.65 9.09	2 1.3492 1.06 11.76 13.33	1 1.1693 0.53 5.88 7.69	0 1.7989 0.00 0.00 0.00	17 8.99
Total	39 20.63	69 36.51	33 17.46	15 7.94	13 6.88	20 10.58	189 100.00

Statistics for Table of CLUSTER by time

Statistic	DF	Value	Prob
Chi-Square	20	13.9215	0.8345
Likelihood Ratio Chi-Square	20	18.0129	0.5866
Mantel-Haenszel Chi-Square	1	1.0239	0.3116
Phi Coefficient		0.2714	
Contingency Coefficient		0.2619	
Cramer's V		0.1357	

WARNING: 53% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by vehicle

CLUSTER	vehicle(vehicle)				
	Bicycle	Car	I do not own a personal vehicle	Motorcycle	Total
1	4 4.6667 2.12 9.52 19.05	11 13.556 5.82 26.19 18.03	23 22.222 12.17 54.76 23.00	4 1.5556 2.12 9.52 57.14	42 22.22
2	6 6.4444 3.17 10.34 28.57	21 18.72 11.11 36.21 34.43	30 30.688 15.87 51.72 30.00	1 2.1481 0.53 1.72 14.29	58 30.69
3	4 3.6667 2.12 12.12 19.05	10 10.651 5.29 30.30 16.39	18 17.46 9.52 54.55 18.00	1 1.2222 0.53 3.03 14.29	33 17.46
4	5 4.3333 2.65 12.82 23.81	13 12.587 6.88 33.33 21.31	20 20.635 10.58 51.28 20.00	1 1.4444 0.53 2.56 14.29	39 20.63
5	2 1.8889 1.06 11.76 9.52	6 5.4868 3.17 35.29 9.84	9 8.9947 4.76 52.94 9.00	0 0.6296 0.00 0.00 0.00	17 8.99
Total	21 11.11	61 32.28	100 52.91	7 3.70	189 100.00

Statistics for Table of CLUSTER by vehicle

Statistic	DF	Value	Prob
Chi-Square	12	6.4668	0.8908
Likelihood Ratio Chi-Square	12	6.1164	0.9101
Mantel-Haenszel Chi-Square	1	1.3396	0.2471
Phi Coefficient		0.1850	
Contingency Coefficient		0.1819	
Cramer's V		0.1068	

WARNING: 45% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by preference

CLUSTER	preference(preference)					
	Convenience	Cost	Environmental impact	Safety	Time efficiency	Total
1	11 9.5556 5.82 26.19 25.58	8 12.889 4.23 19.05 13.79	1 4 0.53 2.38 5.56	1 1.1111 0.53 2.38 20.00	21 14.444 11.11 50.00 32.31	42 22.22
2	10 13.196 5.29 17.24 23.26	22 17.799 11.64 37.93 37.93	5 5.5238 2.65 8.62 27.78	0 1.5344 0.00 0.00 0.00	21 19.947 11.11 36.21 32.31	58 30.69
3	10 7.5079 5.29 30.30 23.26	9 10.127 4.76 27.27 15.52	4 3.1429 2.12 12.12 22.22	2 0.873 1.06 6.06 40.00	8 11.349 4.23 24.24 12.31	33 17.46
4	7 8.873 3.70 17.95 16.28	13 11.968 6.88 33.33 22.41	6 3.7143 3.17 15.38 33.33	1 1.0317 0.53 2.56 20.00	12 13.413 6.35 30.77 18.46	39 20.63
5	5 3.8677 2.65 29.41 11.63	6 5.2169 3.17 35.29 10.34	2 1.619 1.06 11.76 11.11	1 0.4497 0.53 5.88 20.00	3 5.8466 1.59 17.65 4.62	17 8.99
Total	43 22.75	58 30.69	18 9.52	5 2.65	65 34.39	189 100.00

Statistics for Table of CLUSTER by preference

Statistic	DF	Value	Prob
Chi-Square	16	18.9823	0.2696
Likelihood Ratio Chi-Square	16	20.9379	0.1809
Mantel-Haenszel Chi-Square	1	2.9998	0.0833
Phi Coefficient		0.3169	
Contingency Coefficient		0.3021	
Cramer's V		0.1585	

WARNING: 40% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by improvement

CLUSTER	improvement(improvement)				
	# of lines can be increased	# of vehicles (transport frequency) can be increased	Public transports as bus should be free for students	Types of transport can be increased	Total
1	9 10.667 4.76 21.43 18.75	17 17.333 8.99 40.48 21.79	0 0.2222 0.00 0.00 0.00	16 13.778 8.47 38.10 25.81	42 22.22
2	16 14.73 8.47 27.59 33.33	21 23.937 11.11 36.21 26.92	0 0.3069 0.00 0.00 0.00	21 19.026 11.11 36.21 33.87	58 30.69
3	6 8.381 3.17 18.18 12.50	17 13.619 8.99 51.52 21.79	0 0.1746 0.00 0.00 0.00	10 10.825 5.29 30.30 16.13	33 17.46
4	13 9.9048 6.88 33.33 27.08	16 16.095 8.47 41.03 20.51	1 0.2063 0.53 2.56 100.00	9 12.794 4.76 23.08 14.52	39 20.63
5	4 4.3175 2.12 23.53 8.33	7 7.0159 3.70 41.18 8.97	0 0.0899 0.00 0.00 0.00	6 5.5767 3.17 35.29 9.68	17 8.99
Total	48 25.40	78 41.27	1 0.53	62 32.80	189 100.00

Statistics for Table of CLUSTER by improvement

Statistic	DF	Value	Prob
Chi-Square	12	8.8727	0.7138
Likelihood Ratio Chi-Square	12	8.2443	0.7658
Mantel-Haenszel Chi-Square	1	1.1017	0.2939
Phi Coefficient		0.2167	
Contingency Coefficient		0.2118	
Cramer's V		0.1251	

WARNING: 30% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by sharing			
	sharing(sharing)		
CLUSTER	No	Yes	Total
1	23 20 12.17 54.76 25.56	19 22 10.05 45.24 19.19	42 22.22
2	26 27.619 13.76 44.83 28.89	32 30.381 16.93 55.17 32.32	58 30.69
3	17 15.714 8.99 51.52 18.89	16 17.286 8.47 48.48 16.16	33 17.46
4	16 18.571 8.47 41.03 17.78	23 20.429 12.17 58.97 23.23	39 20.63
5	8 8.0952 4.23 47.06 8.89	9 8.9048 4.76 52.94 9.09	17 8.99
Total	90 47.62	99 52.38	189 100.00

Statistics for Table of CLUSTER by sharing

Statistic	DF	Value	Prob
Chi-Square	4	1.9230	0.7499
Likelihood Ratio Chi-Square	4	1.9270	0.7492
Mantel-Haenszel Chi-Square	1	0.6610	0.4162
Phi Coefficient		0.1009	
Contingency Coefficient		0.1004	
Cramer's V		0.1009	

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by mode				
	mode(mode)			
CLUSTER	Combination of modes	It depends on the circumstances	Single mode	Total
1	14 12.444 7.41 33.33 25.00	19 19.111 10.05 45.24 22.09	9 10.444 4.76 21.43 19.15	42 22.22
2	21 17.185 11.11 36.21 37.50	26 26.392 13.76 44.83 30.23	11 14.423 5.82 18.97 23.40	58 30.69
3	9 9.7778 4.76 27.27 16.07	16 15.016 8.47 48.48 18.60	8 8.2063 4.23 24.24 17.02	33 17.46
4	8 11.556 4.23 20.51 14.29	15 17.746 7.94 38.46 17.44	16 9.6984 8.47 41.03 34.04	39 20.63
5	4 5.037 2.12 23.53 7.14	10 7.7354 5.29 58.82 11.63	3 4.2275 1.59 17.65 6.38	17 8.99
Total	56 29.63	86 45.50	47 24.87	189 100.00

Statistics for Table of CLUSTER by mode

Statistic	DF	Value	Prob
Chi-Square	8	9.0378	0.3391
Likelihood Ratio Chi-Square	8	8.5372	0.3828
Mantel-Haenszel Chi-Square	1	3.0722	0.0796
Phi Coefficient		0.2187	
Contingency Coefficient		0.2136	
Cramer's V		0.1546	

Sample Size = 189

Frequency
Expected
Percent
Row Pct
Col Pct

Table of CLUSTER by environment			
	environment(environment)		
CLUSTER	No	Yes	Total
1	18 19.333 9.52 42.86 20.69	24 22.667 12.70 57.14 23.53	42 22.22
2	19 26.698 10.05 32.76 21.84	39 31.302 20.63 67.24 38.24	58 30.69
3	18 15.19 9.52 54.55 20.69	15 17.81 7.94 45.45 14.71	33 17.46
4	20 17.952 10.58 51.28 22.99	19 21.048 10.05 48.72 18.63	39 20.63
5	12 7.8254 6.35 70.59 13.79	5 9.1746 2.65 29.41 4.90	17 8.99
Total	87 46.03	102 53.97	189 100.00

Statistics for Table of CLUSTER by environment

Statistic	DF	Value	Prob
Chi-Square	4	9.8057	0.0438
Likelihood Ratio Chi-Square	4	9.9772	0.0408
Mantel-Haenszel Chi-Square	1	5.6073	0.0179
Phi Coefficient		0.2278	
Contingency Coefficient		0.2221	
Cramer's V		0.2278	

Sample Size = 189

5.1 Cluster 1: Safety First Commuters

This group consists of 42 individuals which is close to 22% of the total population. The cluster primarily includes younger adults aged 18-25, with a significant minority aged 26-35. The younger adults 18-25 is 50% of the group with 21 individuals and a significant minority aged 26-35 has 12 individuals which corresponds to roughly 29% of the populations of Cluster 1. The gender distribution favors males as males have 26 individuals which is 62% of the total group and females have 16 individuals with 38% of the total group. Geographically, they live in Europe, Africa, and Asia with a high contribution from Europe with 60%. In terms of employment, there's a split between students, employed individuals, and some unemployed. The students and employed individuals are equal to 40.5% of the total group and unemployed people are covering the 10% of the group. They heavily rely on public transportation where public transport users are covering 48% of the group and personal car users are covering 31% of the group. The individuals of this group are reflecting a conscious effort towards environmentally friendly

commuting where 57% of the group will pay more for the transportation if the transportation options are improved to be more environmentally safe. This group is usually between 30-60 minutes which corresponds to 31% of the group, then 60-90 minutes with 19% followed by 15-30 minutes with 17% of the group. This group usually travels small distances less than 8 km with 36%, then 8-16 km with 24% followed by 16-32 km with 19% of the group. Individuals of this cluster usually do not have a personal vehicle with 55% and the individuals that do have personal vehicle have a car with 26% of the group. This group decides on the mode of daily transportation mainly looking at the optimal time with 50%, then valuing convenience with 26% followed by valuing cost with 19% of the group. They usually have the option to choose between one or more modes of transportation with 45% but 33% of the group uses multiple modes of transportation and 22% uses single modes of transportation. There is a significant difference between ride sharing users who are 45% and non-users of ridesharing with 45% for this group. This group wants to see the number of vehicles increased (40.5%) or types of transport increased (38%) with a small minority that wants the number of lines increased (21.5%). Main differences of this group are caring for transportation safety with t value 3.15 and usage of mobile apps with 2.54. That's why we call this group “Safety first commuters”.

5.2 Cluster 2: Environmentalists

With 58 individuals, this cluster is the largest group and the age skews towards younger adults aged 18-25 (52%), though there's representation from the 26-35 age group (26%). The male population outnumbers females where males are 60% of the group and 40% are females. They live in Europe, Africa, and Asia, representing diverse continents, where Europe outnumbers (45%) Africa (26%) and Asia (26%). Most are students (47%) or employed individuals (38%) with a very small minority of unemployed people (7%). Public transportation is their preferred with 59%, followed by car travel 26% of the group. They usually travel less than 8 km with 38%, followed by 8-16 km with 28% and a small minority that travels 16-32 km with 12%. They usually spend 30-60 minutes on daily transportation by 45%, followed by 15-30 minutes by 17% and 60-90 minutes by 16%. They usually do not have a vehicle by 52% and the ones that do have a vehicle have a car by 36%. This group prefers low-cost options with 38%, then transportation options with optimal time with 36% and a small minority prefers convenience over all with 17%. They are usually open to ridesharing by 55% of the population and 45% non- users of ridesharing. They usually have the option to go with one or more modes of transport (45%) but a significant 36% use multiple modes of transport and 19% use a single mode of transport. They are also willing to pay more for environmental solutions with 67%, showcasing some environmental consciousness alongside their cost-conscious behavior and 33% are against paying more. This group wants to see improvements made on the number of vehicles to be increased and types of transportation increased equally by 36% and 28% wants to see the lines increased. Main differences of this group are caring for transportation options with t value 2.63 and environment caring with 4.74. That's why we call this group “Environmentalists”.

5.3 Cluster 3: Active Travelers

This group consists of 33 individuals, primarily aged between 18-25 by 52% and 26-35 by 40%. This group consists of males by 55% and females by 45%. They are geographically diverse,

living in Europe, Africa, and Asia but Europe dominates by 64% followed by Africa (18%) and Asia (15%). The group mainly includes students (49%) and employed individuals (40%). They usually use public transport by 58% but 30% travel by car. They usually travel less than 8 km (49%) followed by 8-18 km (21%) and 32-48 km (12%). They usually spend 30-60 minutes on transportation (43%) followed by 15-30 minutes (21%) and 60-90 minutes (15%). They prioritize time and convenience in transportation choices, with a notable reliance on public transportation. Unlike the previous clusters, they exhibit less inclination to pay extra for environmental solutions, perhaps indicating a higher priority on personal convenience over environmental concerns. They usually do not have a vehicle by 55% but 30% own a car. They are almost equally distributed by ridesharing with ridesharing users with 52% and non-users with 48%. This group has the option to choose between one or more modes of transport by 48% but users of multiple modes of transport (27%) and single mode of transport (25%) are very close. They are not environmentally conscious as many individuals will not pay more for an environmentally better solution by 55%. They prefer convenience by 30%, then low-cost by 27% and optimal time by 24%. For improvement, many individuals want the number of vehicles to increase, then number of transportation options increased by 30% and number of lines increased by 18%. Main differences of this group are caring for active transportation with t value 2.79 and environment caring with 3.85. That's why we call this group “Active Travelers”.

5.4 Cluster 4: Car Addicts

This group comprises 39 individuals with the ages of 18-25 (44%) and 26-35 (33%). Males are 54% and females are 46%. They live in Europe, Africa, and Asia where Europe is 46% which is lower than the other groups followed by Asia (31%) and Africa (21%). The population includes mainly students and employed people with 49% students and 38% employed people. They use primarily public transportation 64% and cars 23%. They usually travel less than 8 km (36%) and 8-16 km (33%) with a small minority of 16-32 km (13%). They have a close distribution of transportation of 1-30 minutes, 30-60 minutes, and 60-90 minutes with 28%, 23% and 21% respectively. They usually do not have a vehicle (51%) but 33% have a car. They usually use ridesharing by 59% and non-users with 41%. They usually use a combination of multiple modes of transportation with 41%, followed by 38% of individuals who have the option to choose between one or more modes of transportation with a minority of single mode users (21%). This group is split between paying more for a better environmental solution by 51% to 49% not in favor of paying. They prefer low-cost options for transportation with 33% followed by optimal time (30%) and convenience (18%). This group wants to see improvements made on types of transportation with 49% followed by the number of vehicles increased by 41% and the number of lines increased with 33%. Main differences of this group are caring for parking with t value 3.11 and personal vehicle usage with 5.74. That's why we call this group “Car Addicts”.

5.5 Cluster 5: Active Personal vehicle Lovers

With 17 individuals, this cluster is the smallest group with individuals aged 18-25 (47%) and 26-35 (29%). This group has 10 males and 7 females (59% and 41%). They live in Europe (47%), Africa (29%), and Asia (24%). Cluster 5 includes students (53%) and employed individuals (35%). They heavily use public transportation with 65% and use a car with 24%.

They usually travel 8-16 km with 35% followed by less than 8 km (29%) and 16-32 km (23%). They usually spend 30-60 minutes for transportation (41%) followed by 13-30 minutes (24%) and 60-90 minutes (18%). They usually do not have a vehicle by 53% but the car users are also 35%. There is a small gap between ride sharing users and non-users favoring users for 53% to 47%. They usually have the option to select between one or more different transportation methods (59%) but 24% use multiple modes of transportation and 18% use a single mode of transportation. They have the most favored statistic considering all the questions for all of the clusters on paying more for a better environmental solution. They are not willing to pay more for a better environmental solution with 71% to 29%. They prefer cost over all (35%) followed by preferring convenience 29% and time by 18%. This group wants to see improvements in the number of vehicles increased by 41% followed by types of transport increased by 35% and number of lines increased by 24%. Main differences of this group are caring for active transportation with t value 2.02 and personal vehicle usage with 2.55. That's why we call this group “Active Personal vehicle Lovers”.

6. Conclusion

The thorough examination of the population's transportation choices provides important new information about the range of priorities and behaviors influencing daily transportation habits. Different patterns surface among the five clusters that have been identified, indicating different levels of importance given to elements like cost-effectiveness, environmental sustainability, safety, and convenience.

The "Environmentalists" group emphasizes a strong understanding towards environmentally conscious choices alongside financial concerns, whereas the "Safety First Commuters" group prioritizes safety and environmentally friendly solutions and shows an intentional effort towards sustainable commuting methods. Conversely, the "Active Travelers" exhibit a predilection for active modes of transportation and place a higher value on time and convenience. Conversely, the "Car Addicts" exhibit a strong bond with their personal cars, indicating a preference for parking accessibility and convenience. Last but not least, the "Active Personal Vehicle Lovers" group strongly prefers personal vehicles and shows resistance to environmental issues when using public transportation.

These distinct clusters highlight the diverse nature of transportation preferences, influenced by demographic factors, geographical location, and individual priorities. Decision-makers and urban planners must have an understanding of these preferences in order to create and improve transportation networks. Through the integration of the analysis's findings, decision-makers can endeavor to develop transportation solutions that are sustainable, safe, diverse, and meet the changing demands of communities, all the while encouraging environmental preservation and guaranteeing fair access to mobility choices.