Analysis Report

This report is structured as follows.

Contents

Cluster Analysis 2

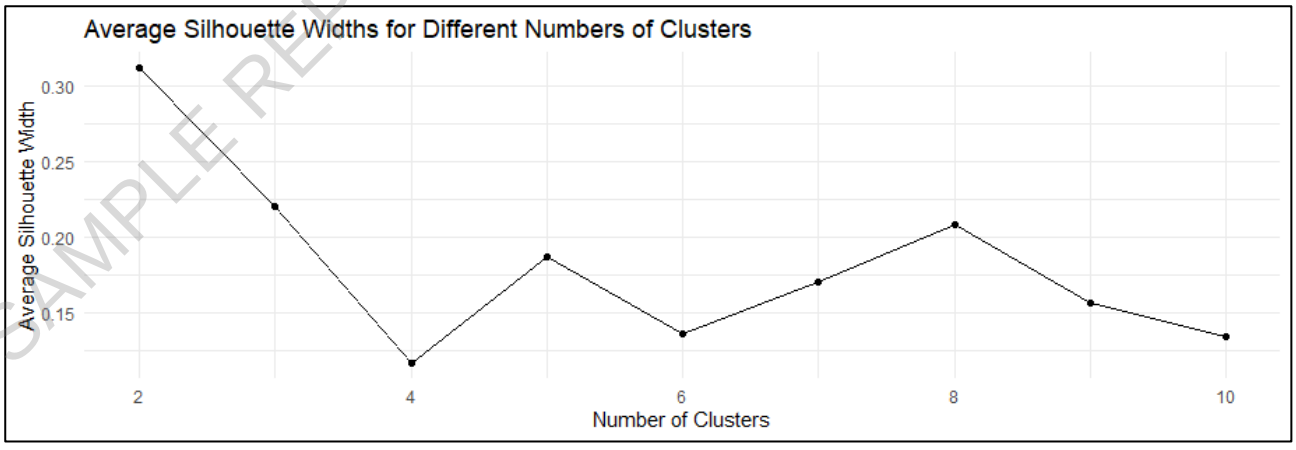
Cluster Profiling 3

Cluster Analysis

Cluster analysis is a powerful tool in data science for grouping a set of objects in such a way that objects in the same group (or cluster) are more similar to each other than to those in other clusters. The technique used in this analysis is known as k-modes clustering, which is particularly suitable for categorical data. K-modes, an extension of the k-means algorithm, replaces means with modes in the clustering process and uses a frequency-based method to update the modes. This makes it ideal for handling categorical variables, as it doesn't rely on mean-based distance metrics, which are less meaningful for categorical data.

Before deciding on the number of clusters to use in k-modes, a silhouette analysis was performed. Silhouette analysis is a method to find the optimal number of clusters and interpret the clustering's quality. It measures how similar an object is to its own cluster compared to other clusters. The silhouette value ranges from -1 to +1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. If most objects have a high value, then the clustering configuration is appropriate. If many points have a low or negative value, then the clustering configuration may have too many or too few clusters.

In this specific analysis, the silhouette widths for different numbers of clusters were calculated. The results were as follows: for 2 clusters, the average silhouette width was 0.3149, indicating moderate separation between clusters. However, as the number of clusters increased, the silhouette width generally decreased, suggesting that additional clusters may not provide meaningful or distinct groupings. Notably, there was a slight increase in silhouette width at 6 clusters (0.2052), 7 clusters (0.2093), and 9 clusters (0.2195), which suggests some potential for these configurations. However, none of these values approached the relatively higher silhouette width observed for 2 clusters. The graph below shows a line plot of the silhouette widths.



The same information is presented in the form of a table below.

Number_of_Clusters	Average_Silhouette_Width
2	**0.3149**
3	0.1696
4	0.1352
5	0.1651
6	0.2052
7	0.2093
8	0.2016
9	0.2195
10	0.1872

Based on these results, using two clusters could be considered the most appropriate for this particular dataset, as it provides a reasonable balance between having a sufficient number of clusters and ensuring that these clusters are distinct from each other. This decision aligns with the primary goal of cluster analysis: to discover the inherent structure of the data in a way that enhances our understanding of it. The k-modes algorithm, given its suitability for categorical data, complements this goal by effectively grouping data points based on their category similarities, thus providing valuable insights into the dataset's categorical features.

Cluster Profiling

After the number of clusters was defined, the k-modes algorithm was executed again and the cluster number were assigned to each case in the dataset. One cluster, hereby defined as Cluster 1 contained 177 individuals, while Cluster 2 was composed by 154 individuals. The second phase of the project involved using chi-square tests to test if clusters were significantly different across all variables in the data. Results showed that the following categories were not significantly different between clusters ($p > 0.05$):

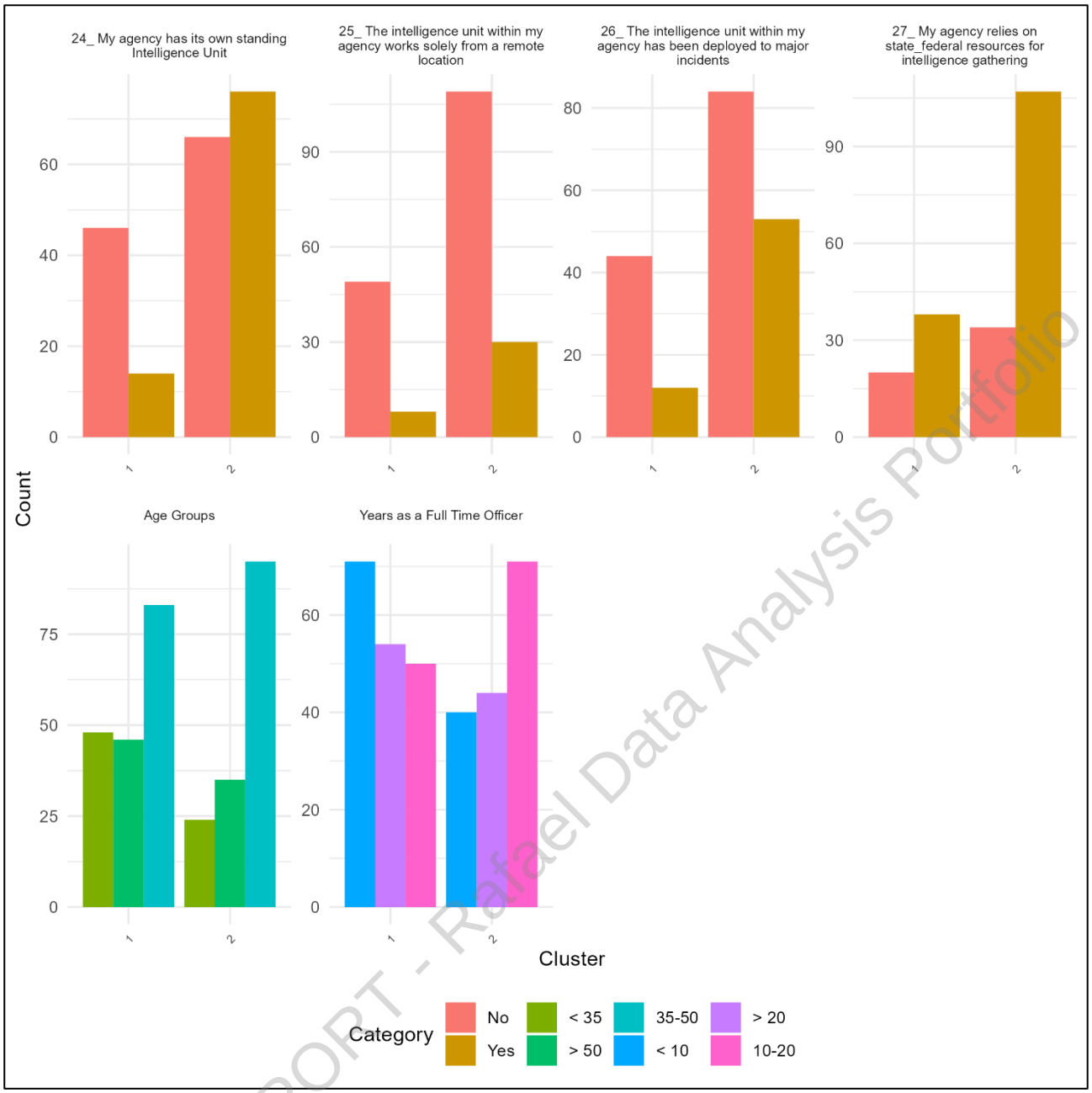
- '23 My agency does not frequently provide its member with intelligence reports;
- 25 The intelligence unit within my agency works solely from a remote location;
- 27 My agency relies on state federal resources for intelligence gathering;
- 6 What type of geographic area is under your agencyâ€™s jurisdiction;
- Gender.

This means that the method was not able to differentiate individuals regarding those variables since the sample was too homogeneous.

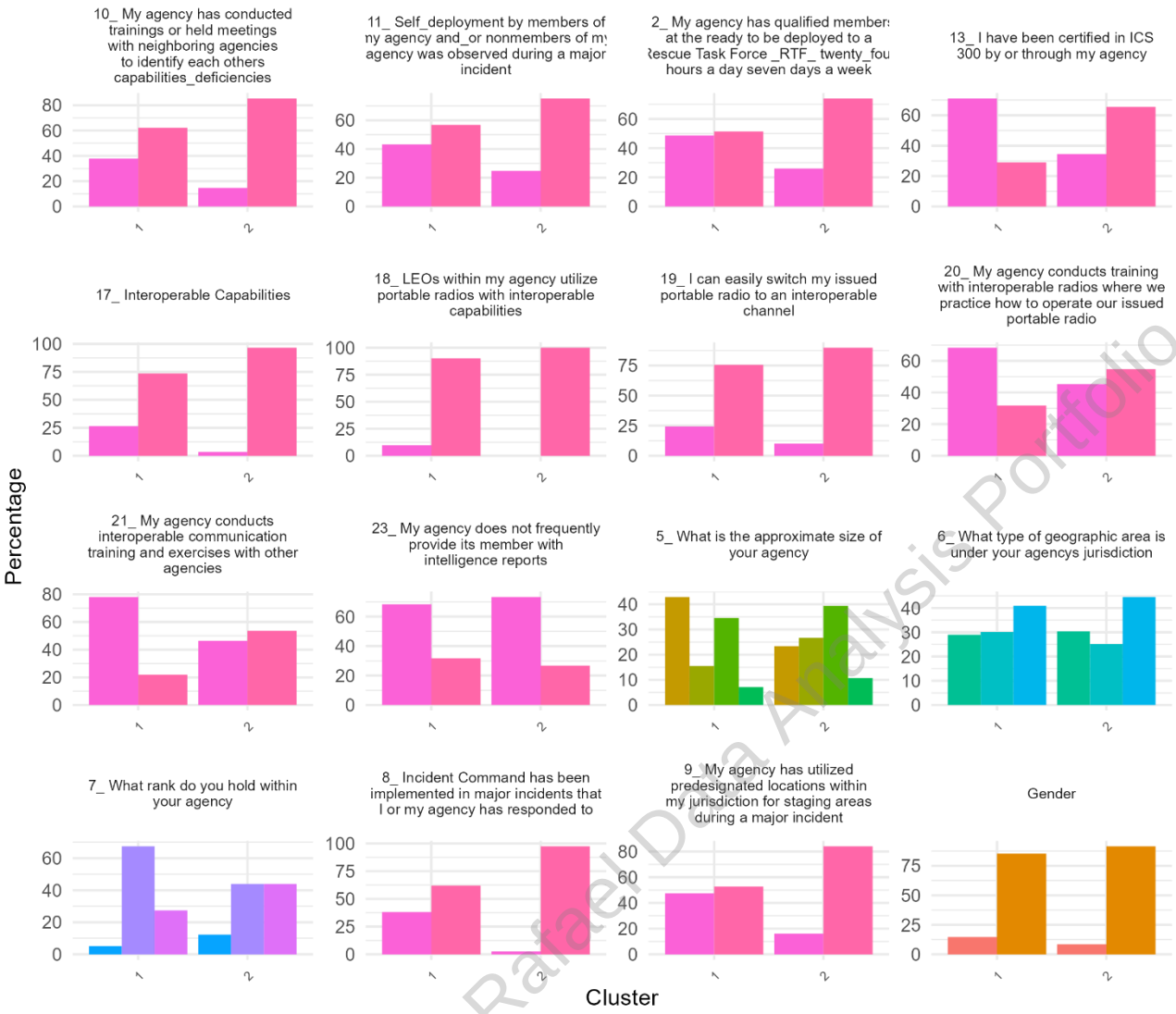
All the other measures were different between clusters according to the Chi-square tests. The table below presents the results.

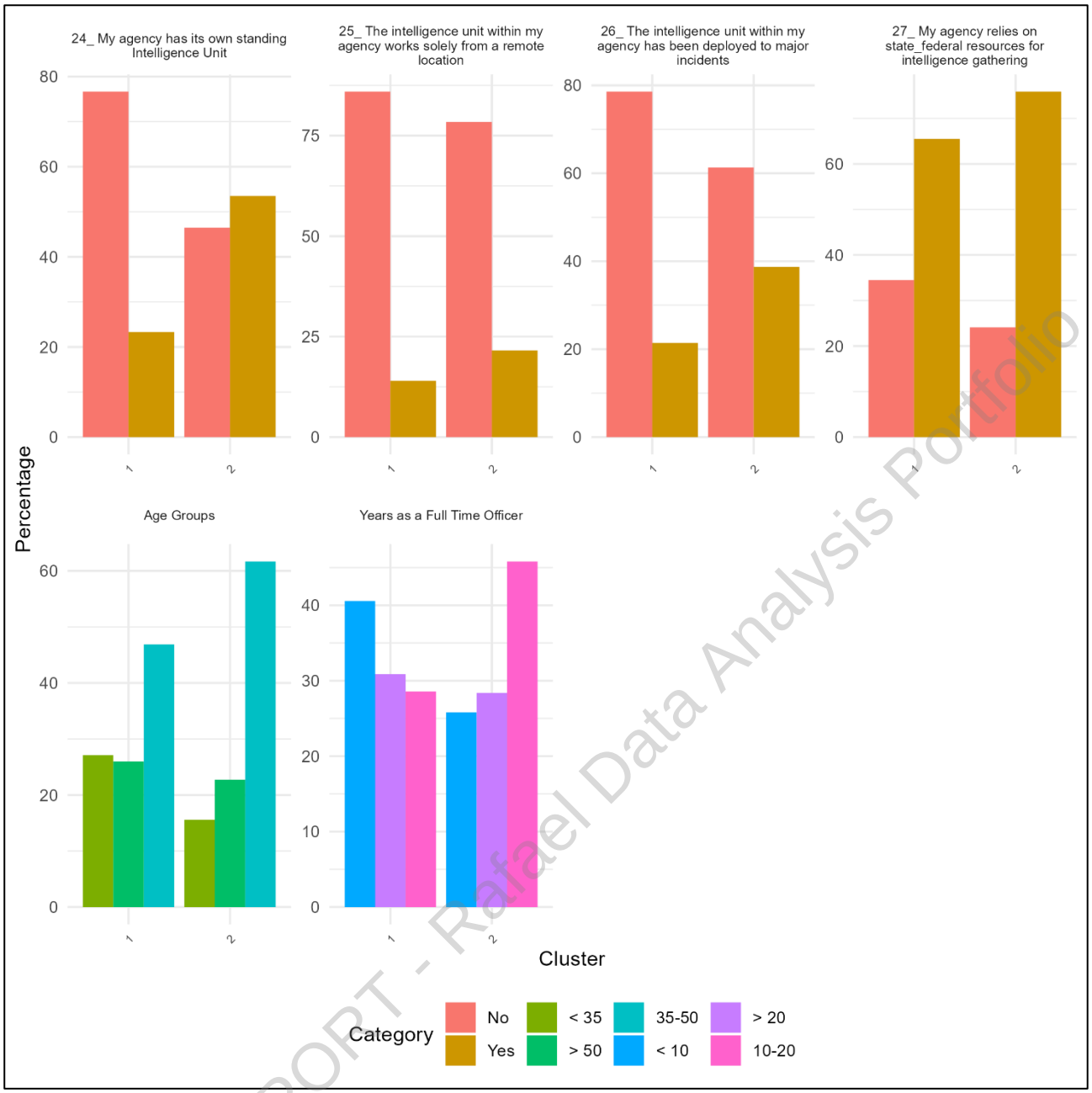
Variable	Category	Cluster 1		Cluster 2		Chi Square	p value
		Count	%	Count	%		
10 My agency has conducted trainings or held meetings with neighboring agencies to identify each other's capabilities deficiencies	No	14	37.8	22	14.7	8.814	0.003
	Yes	23	62.2	128	85.3	8.814	0.003
11 Self deployment by members of my agency and or nonmembers of my agency was observed during a major incident	No	16	43.2	37	24.8	4.069	0.044
	Yes	21	56.8	112	75.2	4.069	0.044
12 My agency has qualified members at the ready to be deployed to a Rescue Task Force RTF twenty four hours a day seven days a week	No	18	48.6	39	26.0	6.155	0.013
	Yes	19	51.4	111	74.0	6.155	0.013
13 I have been certified in ICS 300 by or through my agency	No	27	71.1	51	34.5	15.159	0.000
	Yes	11	28.9	97	65.5	15.159	0.000
17 Interoperable Capabilities	No	22	26.5	5	3.4	24.933	0.000
	Yes	61	73.5	141	96.6	24.933	0.000
18 LEOs within my agency utilize portable radios with interoperable capabilities	No	4	9.8	0	0.0	9.592	0.002
	Yes	37	90.2	137	100.0	9.592	0.002
19 I can easily switch my issued portable radio to an interoperable channel	No	10	24.4	14	10.2	4.286	0.038
	Yes	31	75.6	123	89.8	4.286	0.038
20 My agency conducts training with interoperable radios where we practice how to operate our issued portable radio	No	28	68.3	62	45.3	5.810	0.016
	Yes	13	31.7	75	54.7	5.810	0.016
21 My agency conducts interoperable communication training and exercises with other agencies	No	32	78.0	63	46.3	11.508	0.001
	Yes	9	22.0	73	53.7	11.508	0.001
23 My agency does not frequently provide its member with intelligence reports	No	41	68.3	104	73.2	0.288	0.591
	Yes	19	31.7	38	26.8	0.288	0.591
24 My agency has its own standing Intelligence Unit	No	46	76.7	66	46.5	14.361	0.000
	Yes	14	23.3	76	53.5	14.361	0.000
25 The intelligence unit within my agency works solely from a remote location	No	49	86.0	109	78.4	1.030	0.310
	Yes	8	14.0	30	21.6	1.030	0.310
26 The intelligence unit within my agency has been deployed to major incidents	No	44	78.6	84	61.3	4.556	0.033
	Yes	12	21.4	53	38.7	4.556	0.033
27 My agency relies on state federal resources for intelligence gathering	No	20	34.5	34	24.1	1.741	0.187
	Yes	38	65.5	107	75.9	1.741	0.187
5 What is the approximate size of your agency	< 50	72	42.9	35	23.3	15.374	0.002
	> 1500	26	15.5	40	26.7	15.374	0.002
	50 - 500	58	34.5	59	39.3	15.374	0.002
	501 - 1500	12	7.1	16	10.7	15.374	0.002
6 What type of geographic area is under your agency's jurisdiction	Rural	51	29.0	47	30.3	1.029	0.598
	Suburban	53	30.1	39	25.2	1.029	0.598
	Urban	72	40.9	69	44.5	1.029	0.598
7 What rank do you hold within your agency	Executive Level	9	5.1	19	12.3	19.319	0.000
	Non-Supervisory	118	67.4	68	43.9	19.319	0.000
	Supervisory	48	27.4	68	43.9	19.319	0.000
8 Incident Command has been implemented in major incidents that I or my agency has responded to	No	67	38.1	4	2.6	59.513	0.000
	Yes	109	61.9	151	97.4	59.513	0.000
9 My agency has utilized predesignated locations within my jurisdiction for staging areas during a major incident	No	18	47.4	24	16.0	15.435	0.000
	Yes	20	52.6	126	84.0	15.435	0.000
Age Groups	< 35	48	27.1	24	15.6	8.747	0.013
	> 50	46	26.0	35	22.7	8.747	0.013
	35-50	83	46.9	95	61.7	8.747	0.013
Gender	Female	26	14.7	13	8.4	2.521	0.112
	Male	151	85.3	141	91.6	2.521	0.112
Years as a Full Time Officer	10-20	50	28.6	71	45.8	12.155	0.002
	< 10	71	40.6	40	25.8	12.155	0.002
	> 20	54	30.9	44	28.4	12.155	0.002

To help visualizing the differences between cluster, the following visualizations were developed. They show bar plots with the different counts observed for each variable disaggregated by Clusters 1 and 2.



Besides the count plots, percentage plots were also created to allow better comparison. They show the percentage of respondents to each question, relative to the total size of each cluster.





The following conclusions can be drawn based on the results.

In Cluster 1, a smaller proportion of individuals (37.8%) reported that their agency had conducted trainings or held meetings with neighboring agencies to identify each other's capabilities and deficiencies, compared to a higher proportion (85.3%) in Cluster 2. This suggests that Cluster 2 may be more engaged in collaborative practices and shared learning experiences.

Regarding the observation of self-deployment during major incidents, Cluster 2 (75.2%) had a higher percentage of individuals observing this phenomenon than Cluster 1 (56.8%), indicating a possible difference in operational procedures or experiences between the clusters. Similarly, a higher percentage of individuals in Cluster 2 (74.0%) reported having qualified members ready to be deployed to a Rescue Task Force (RTF) 24/7, compared to Cluster 1 (51.4%).

A notable difference was observed in ICS 300 certification, where a significantly larger proportion of Cluster 2 (65.5%) had been certified, compared to only 28.9% in Cluster 1. This difference could reflect variations in training or professional development opportunities between the clusters.

When considering interoperable capabilities, a striking 96.6% of individuals in Cluster 2 reported having such capabilities, compared to 73.5% in Cluster 1. This difference was further underscored by the use of portable radios with interoperable capabilities, where 100% of Cluster 2 and 90.2% of Cluster 1 reported usage, highlighting a near-universal adoption in Cluster 2.

The proportion of individuals whose agencies conduct interoperable communication training and exercises with other agencies was higher in Cluster 2 (53.7%) compared to Cluster 1 (22.0%). Interestingly, the proportion of agencies with their own standing Intelligence Unit was significantly higher in Cluster 2 (53.5%) compared to Cluster 1 (23.3%).

In terms of agency size, significant differences were observed across clusters. Cluster 1 had a larger proportion of individuals from smaller agencies (with less than 50 members) and larger agencies (more than 1500 members), while Cluster 2 had a more even distribution across different agency sizes.

Differences in rank held within the agency were also noteworthy. A higher proportion of individuals in Cluster 2 held supervisory roles (43.9%) compared to Cluster 1 (27.4%), and a smaller proportion in non-supervisory roles (43.9% in Cluster 2 compared to 67.4% in Cluster 1).

The implementation of Incident Command in major incidents was reported by a vast majority in Cluster 2 (97.4%) compared to Cluster 1 (61.9%), indicating a possible difference in operational protocols or experiences in handling major incidents.

In conclusion, the cluster analysis reveals notable distinctions and some similarities between Cluster 1 and Cluster 2, reflecting varying operational, training, and resource characteristics within the agency. Cluster 2 distinctly stands out in terms of higher engagement in collaborative practices, more widespread implementation of interoperable capabilities, and a greater prevalence of advanced training, as evidenced by the higher rates of ICS 300 certification and ongoing training exercises with other agencies. This cluster also has a higher representation of supervisory roles and a more balanced distribution across different agency sizes. In contrast, Cluster 1 is characterized by a lesser degree of these attributes, indicating variability in resources and operational approaches, with a notable representation from both smaller and larger agencies. Despite these differences, both clusters show a commitment to operational readiness and adaptability, albeit at varying levels of resource allocation and collaborative integration. This analysis underscores the importance of understanding the diverse operational contexts and resource environments in which these individuals function within their respective agencies.