



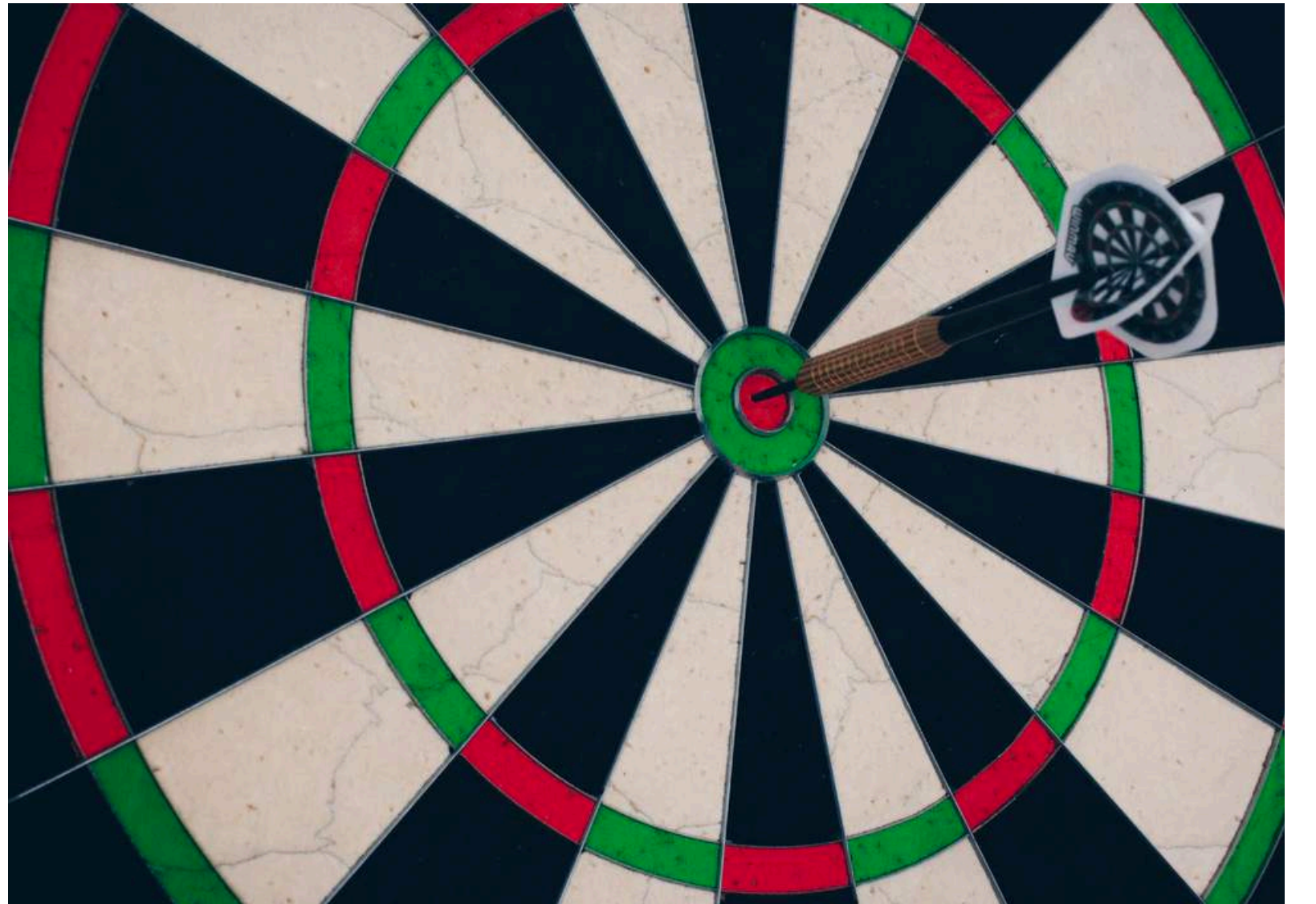
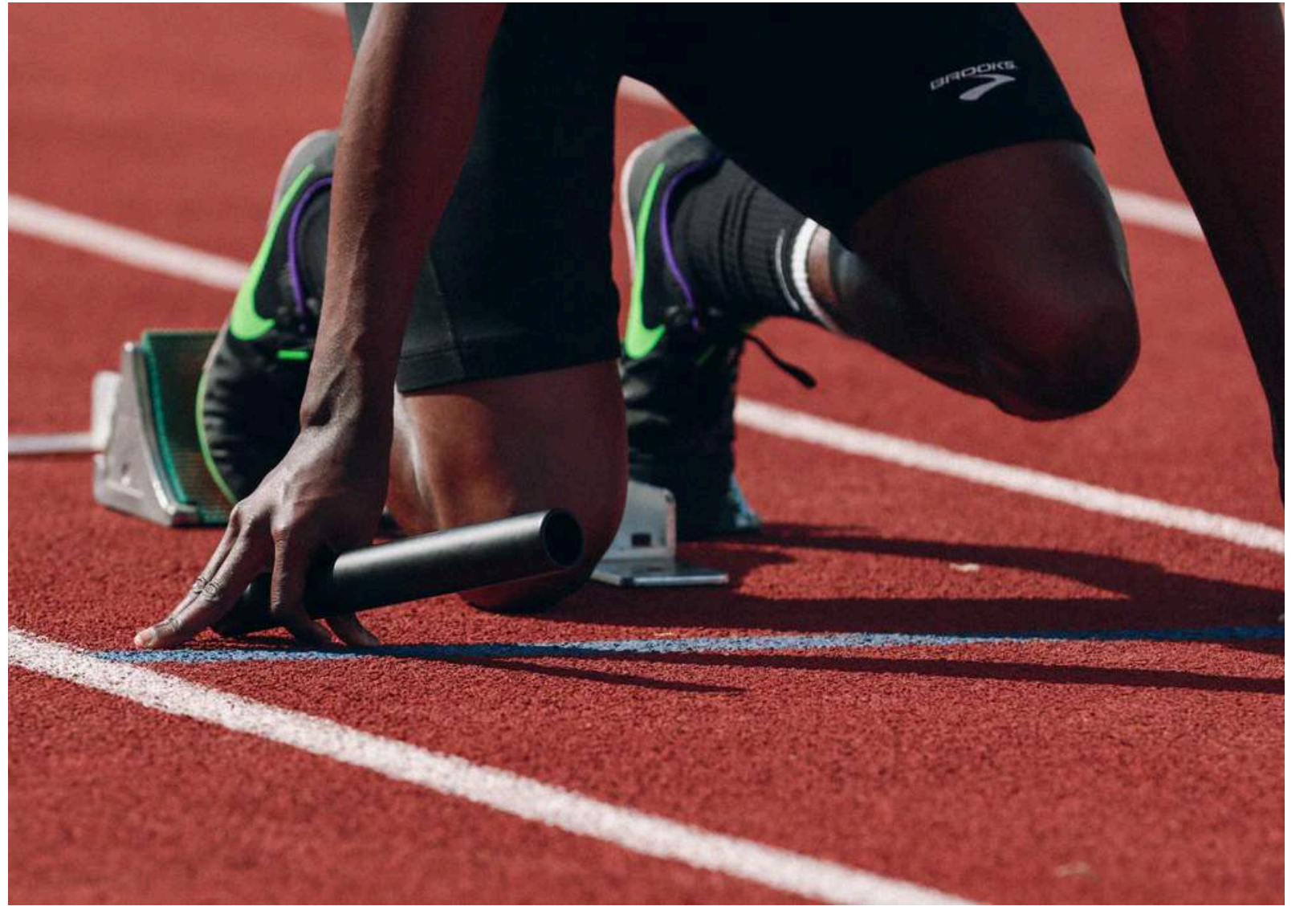
Cluster Analysis

Youpele Michael

Overview

- Objective
- Data Exploration
- Data Cleaning and Preprocessing
- Cluster Analyses + Comments
- Conclusion



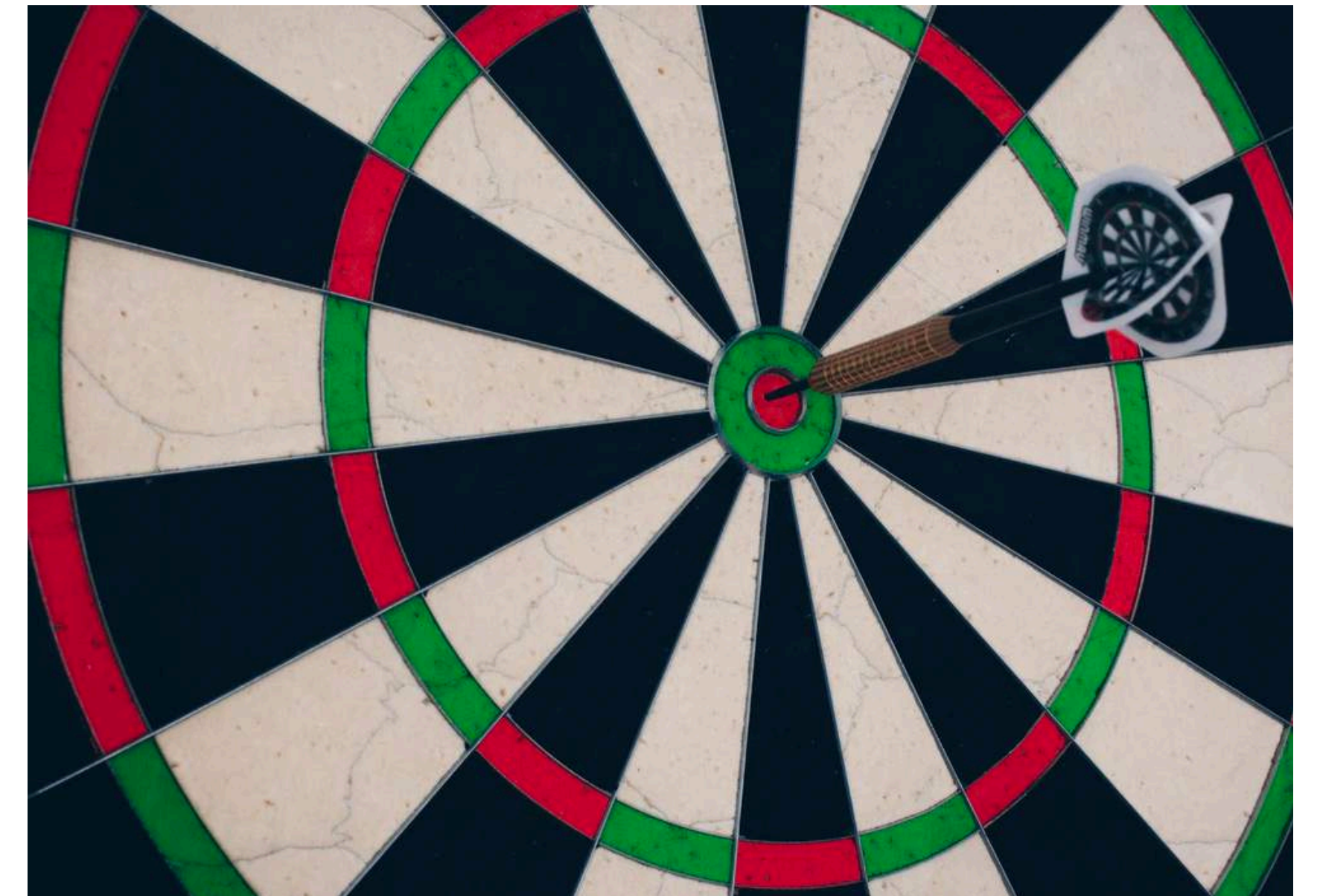


Objective



Objective

The aim of the analysis is to identify distinct archetypes within the respondents based on their answers.





Objective

The 3 main questions in the questionnaire:

- Whether people noticed any changes to the London mobility landscape (5 options + the negative answer (“no changes noticed”)?
- Whether people support further actions (5 options + the negative answer (“none of the above”)?
- Whether people knew the cities mobility policy (and it was adequate, too slow, too quick) or did not know the policy at all?



Data Exploration





Data Exploration

- Dataset contains 301 entries (rows), with 34 features (columns).
- Some of the features are: UserID, Age, Created, Gender, Home location, Statement on Mobility Policy, Changes noticed during the Covid, Policies Supported etc

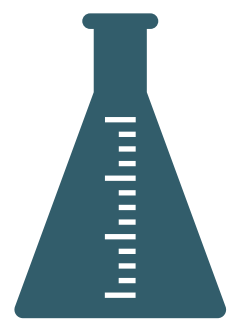


Data Exploration

- Dataset has a lot of nan (not a number) values
- Distinct values:
 - User ID has 301, ie 301 different persons participated in the survey.
 - Age was split into 6 age groups
 - 2 Genders
 - 6 values in Statement on Mobility Policy
 - Etc

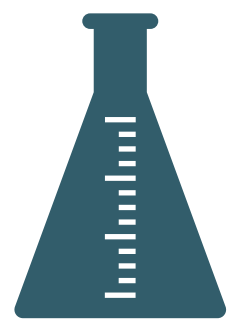


Data Cleaning and Preprocessing



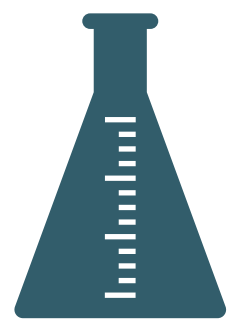
Data Cleaning and Preprocessing

- Removed columns that were not relevant to the objective, like Gender, Age, UserID etc
- Converted all nan values to zero.
- Replaced all the ticked answers 'x' to 1.



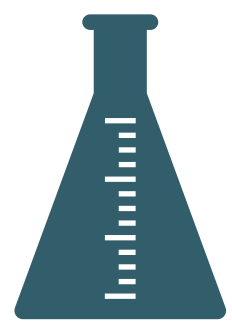
Data Cleaning and Preprocessing

- Encoded the values in the 'Statement on Mobility Policy'.
 - From 0 (not in support or I don't know...) to 5 (in support and city pace too slow...)
- Grouped the sub-columns of "I noticed these changes during COVID" and "I support these policies" columns into their respective main columns.
- By summing the values of the sub-columns to form the value of the main columns



Data Cleaning and Preprocessing

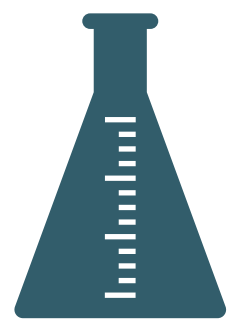
- Created an additional column "I noticed these changes during COVID and I support these policies"
- The value of this column is the sum of the two main columns.
- Created a new dataset that only contains these preprocessed columns (engineered features).



Data Cleaning and Preprocessing

Raw dataset

	User ID	Created	Age	Gender	Home location	Device type	Region	Statement on Mobility Policy	Cycle Lanes	Low Traffic	...	Driving Bans Introduced ("I noticed these changes during COVID")	Change of PT Schedule ("I noticed these changes during COVID")	New Modes Added ("I noticed these changes during COVID")	None of the above ("I noticed these changes during COVID")
0	MqkKEL6WeI91	2021-06-21T20:22:14+00:00	18-20	Female	Greater London	iOS	UK	I know the mobility strategy my city is pursui...	NaN	Introduction of traffic restricted areas (pede...	...	NaN	NaN	NaN	NaN
1	qVlp98JWL7A5	2021-06-21T20:22:19+00:00	21-24	Male	Greater London	iOS	UK	Other (please comment)	Expansion of cycle lanes	NaN	...	NaN	NaN	NaN	NaN
2	rYkvNLbKGkmx	2021-06-21T20:22:22+00:00	21-24	Male	Greater London	iOS	UK	I know the mobility strategy my city is pursui...	NaN	Introduction of traffic restricted areas (pede...	...	NaN	NaN	NaN	NaN
3	0xk2d3195IAw	2021-06-21T20:22:24+00:00	18-20	Female	Greater London	iOS	UK	I know the mobility strategy my city is pursui...	Expansion of cycle lanes	NaN	...	NaN	NaN	NaN	NaN
4	PLIX23Gvdkpv	2021-06-21T20:22:31+00:00	21-24	Male	Greater London	iOS	UK	I know the mobility strategy my city is pursui...	Expansion of cycle lanes	NaN	...	x	NaN	NaN	NaN



Data Cleaning and Preprocessing

Preprocessed dataset

	Statement on Mobility Policy	I noticed these changes during COVID	I support these policies	I noticed these changes during COVID and I support these policies
0	3	1	1	2
1	0	1	2	3
2	1	1	1	2
3	3	1	1	2
4	5	2	1	3
5	3	4	3	7
6	0	3	2	5
7	0	1	2	3
8	3	4	4	8
9	0	3	1	4



Cluster Analyses



Cluster Analyses

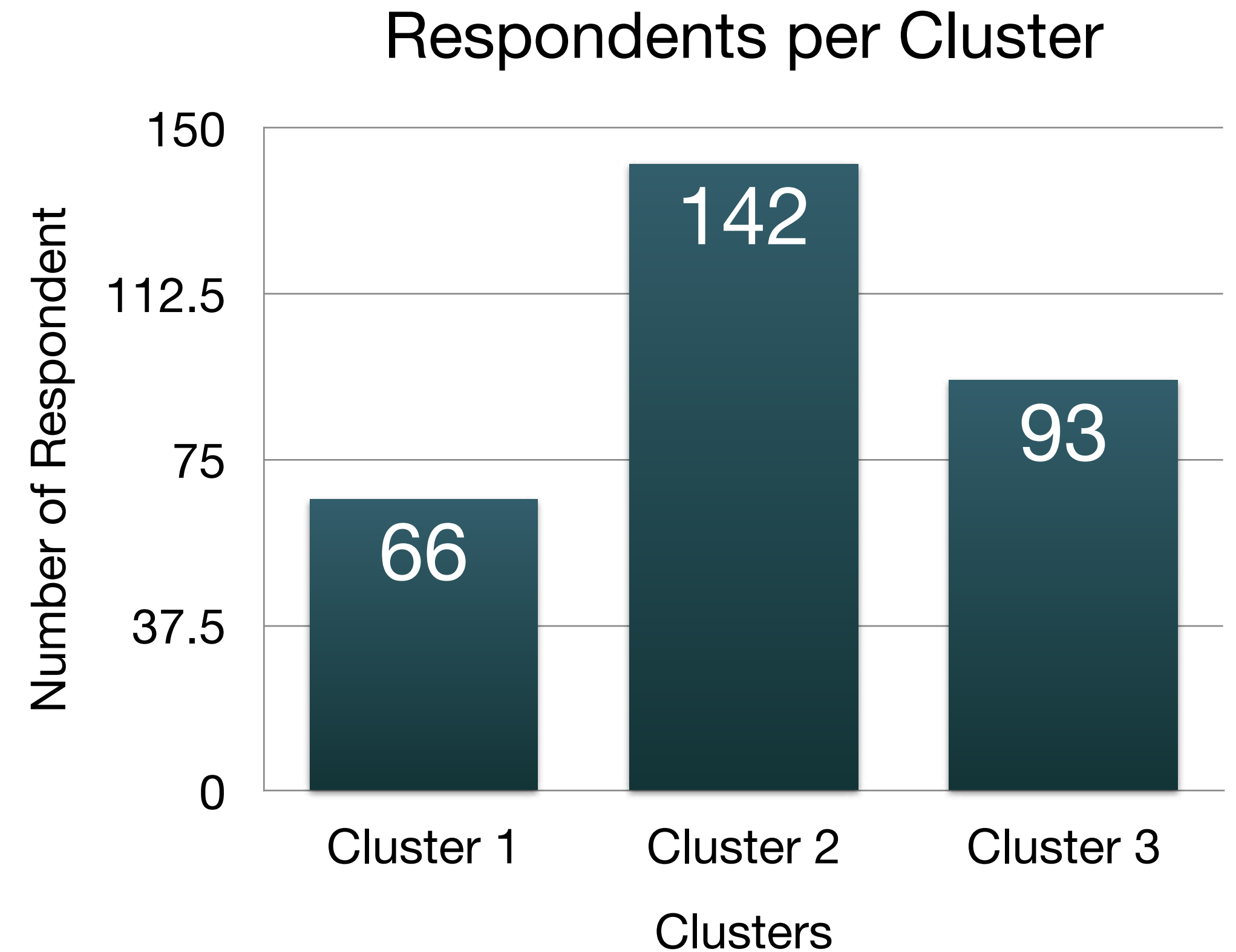
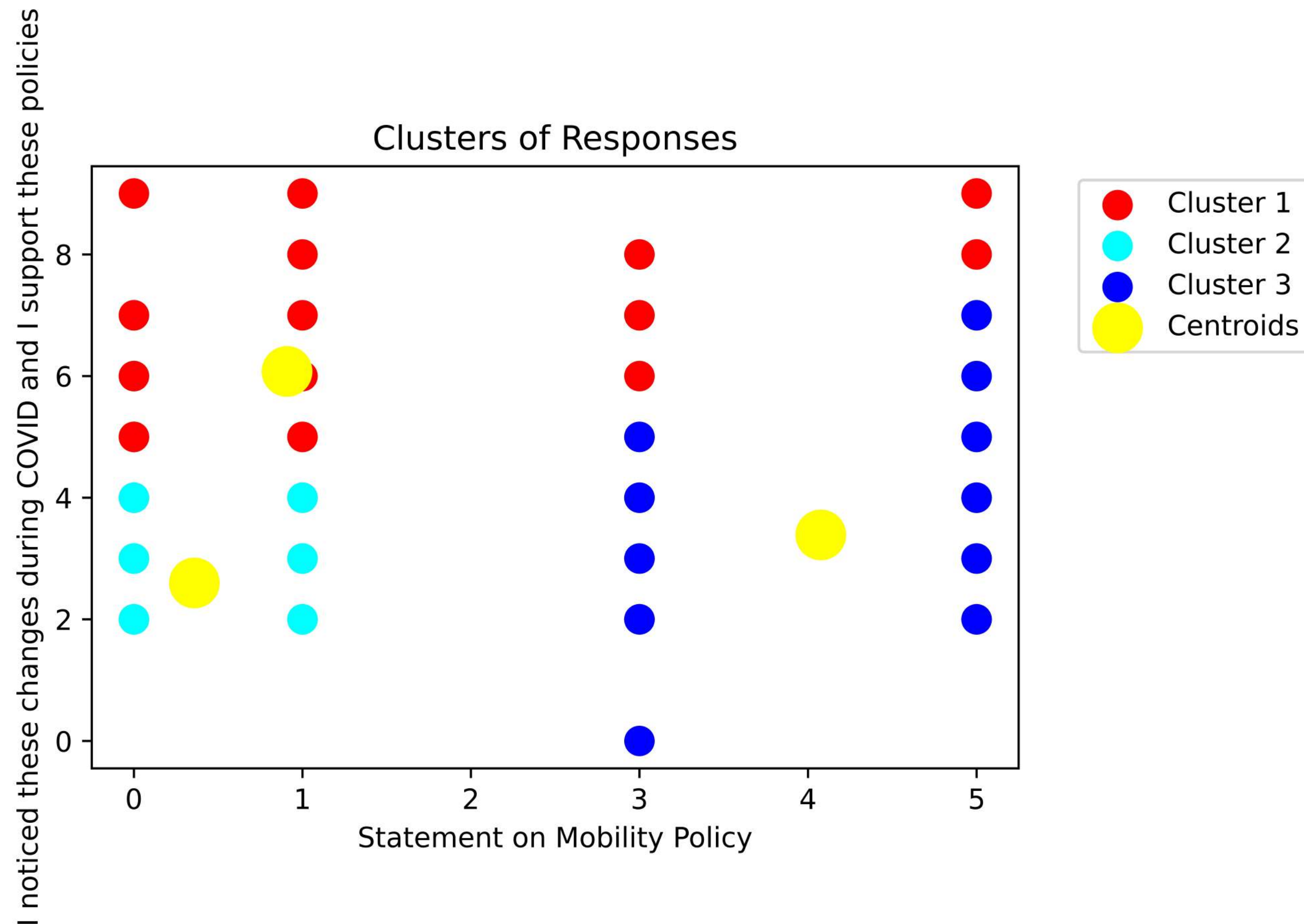
5 cluster analyses were done on different combination of at least 2 columns in the preprocessed dataset:

1. "Statement on Mobility Policy" VS "I noticed these changes during COVID and I support these policies".
2. "Statement on Mobility Policy" VS "I noticed these changes during COVID".
3. "Statement on Mobility Policy" VS "I support these policies"
4. "I noticed these changes during COVID" VS "I support these policies"
5. 'Statement on Mobility Policy' VS 'I noticed these changes during COVID' VS "I support these policies"



Cluster Analysis 1

"Statement on Mobility Policy" VS "I noticed these changes during COVID and I support these policies"





Cluster Analysis 1



"Statement on Mobility Policy" VS "I noticed these changes during COVID and I support these policies"

- Cluster 1 — This is the group that thinks the local government is on the right track and going at the right speed
- Cluster 2 — This is the group that do not support the actions being taken and wish they would cease.
- Cluster 3 — This is the group that strongly support London becoming more 'mobility-friendly' but do not believe progress is quick enough.



Cluster Analysis 1 — Comments



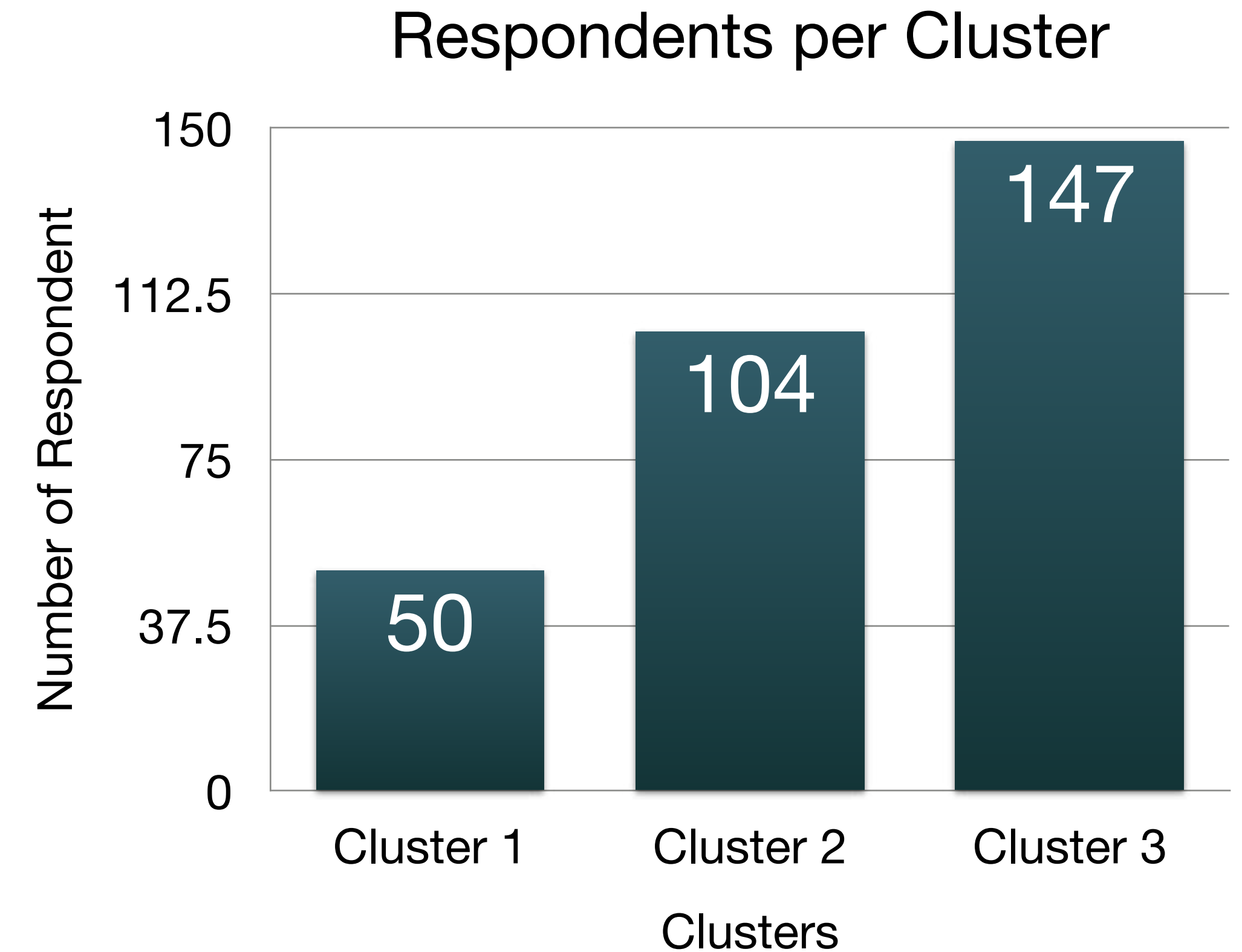
"Statement on Mobility Policy" VS "I noticed these changes during COVID and I support these policies"

- If the survey mirrors the London populace by this clustering analysis:
- Having a total of 159 respondents (cluster 1+3) who are in favour of the mobility policy, it's almost a 52/48% chance that the populace might embrace an alternative or unorthodox mobility service or product.
- A company with an alternative or unorthodox mobility service or product must look for inventive ways to win some of the 142 respondents (ie 47% of the population) who are against the policy. If they want to survive in this space.



Cluster Analysis 2

"Statement on Mobility Policy" VS "I noticed these changes during COVID"





Cluster Analysis 2



"Statement on Mobility Policy" VS "I noticed these changes during COVID and I support these policies"

- Cluster 1 — This is the group that thinks the local government is on the right track and going at the right speed.
- Cluster 2 — This is the group that strongly support London becoming more 'mobility-friendly' but do not believe progress is quick enough.
- Cluster 3 — This is the group that do not support the actions being taken and wish they would cease.



Cluster Analysis 2 – Comments



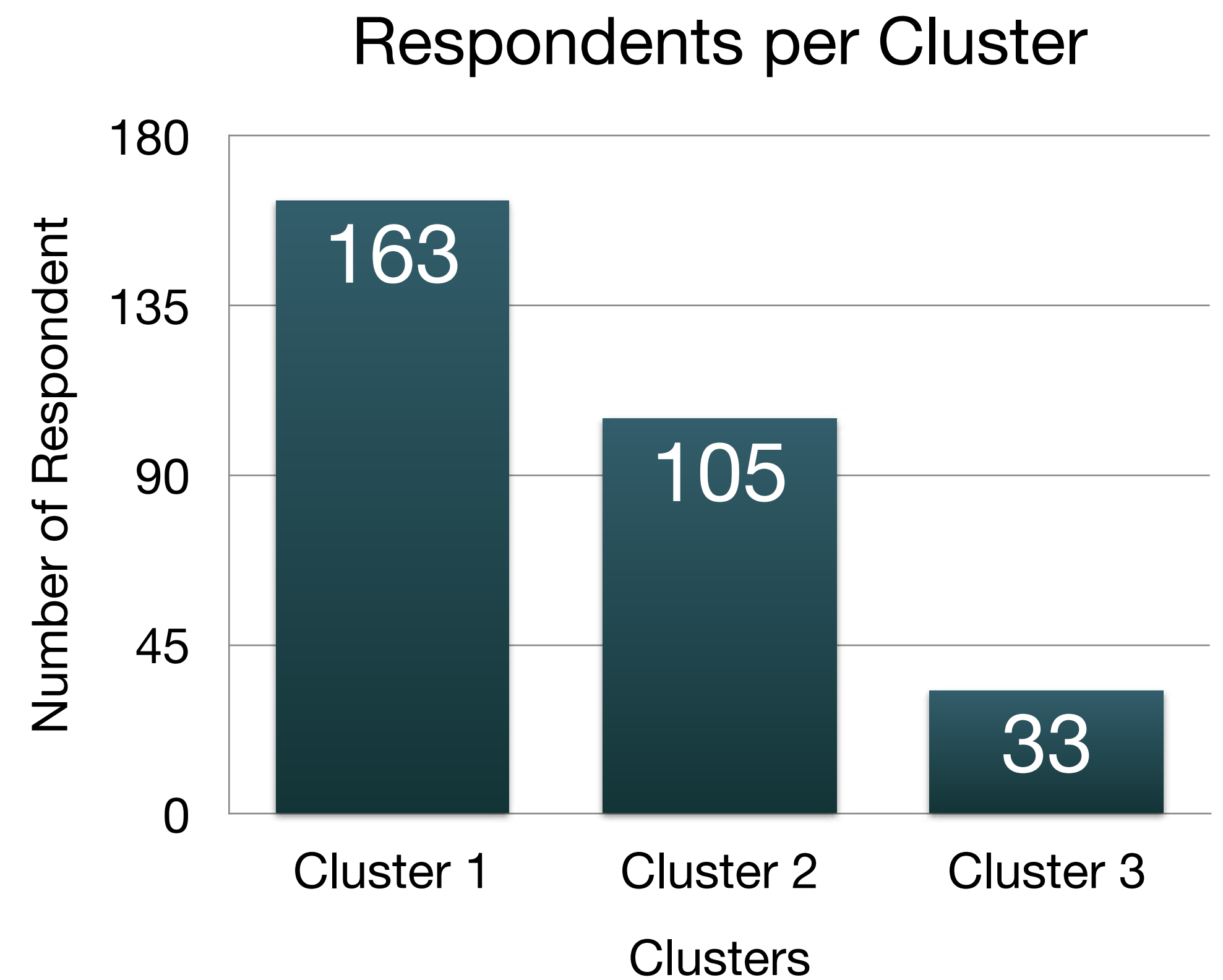
"Statement on Mobility Policy" VS "I noticed these changes during COVID"

- If the survey mirrors the London populace then and then by this clustering analysis:
- Having a total of 154 respondents (cluster 1+2) who are in favour of the mobility policy, it's almost a 51/49% chance that the populace might embrace an alternative or unorthodox mobility service or product.
- A company with an alternative or unorthodox mobility service or product must look for inventive ways to win some of the 147 respondents (49% of the population) who are against the policy. If they want to survive in this space.



Cluster Analysis 3

"Statement on Mobility Policy" VS "I support these policies"





Cluster Analysis 3



"Statement on Mobility Policy" VS "I support these policies"

- Cluster 1 — This is the group that do not support the actions being taken and wish they would cease.
- Cluster 2 — This is the group that strongly support London becoming more 'mobility-friendly' but do not believe progress is quick enough.
- Cluster 3 — This is the group that thinks the local government is on the right track and going at the right speed.



Cluster Analysis 3 — Comments



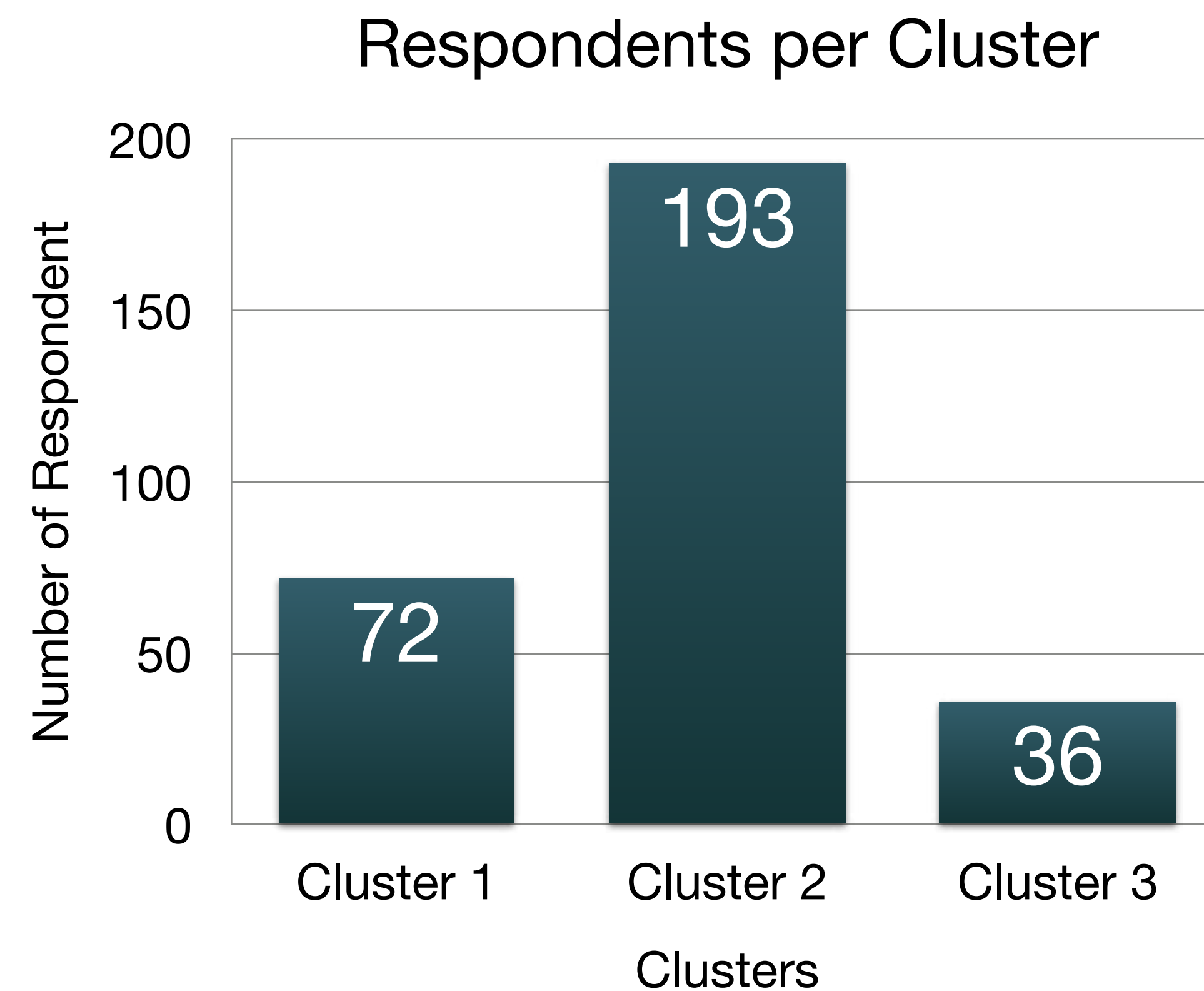
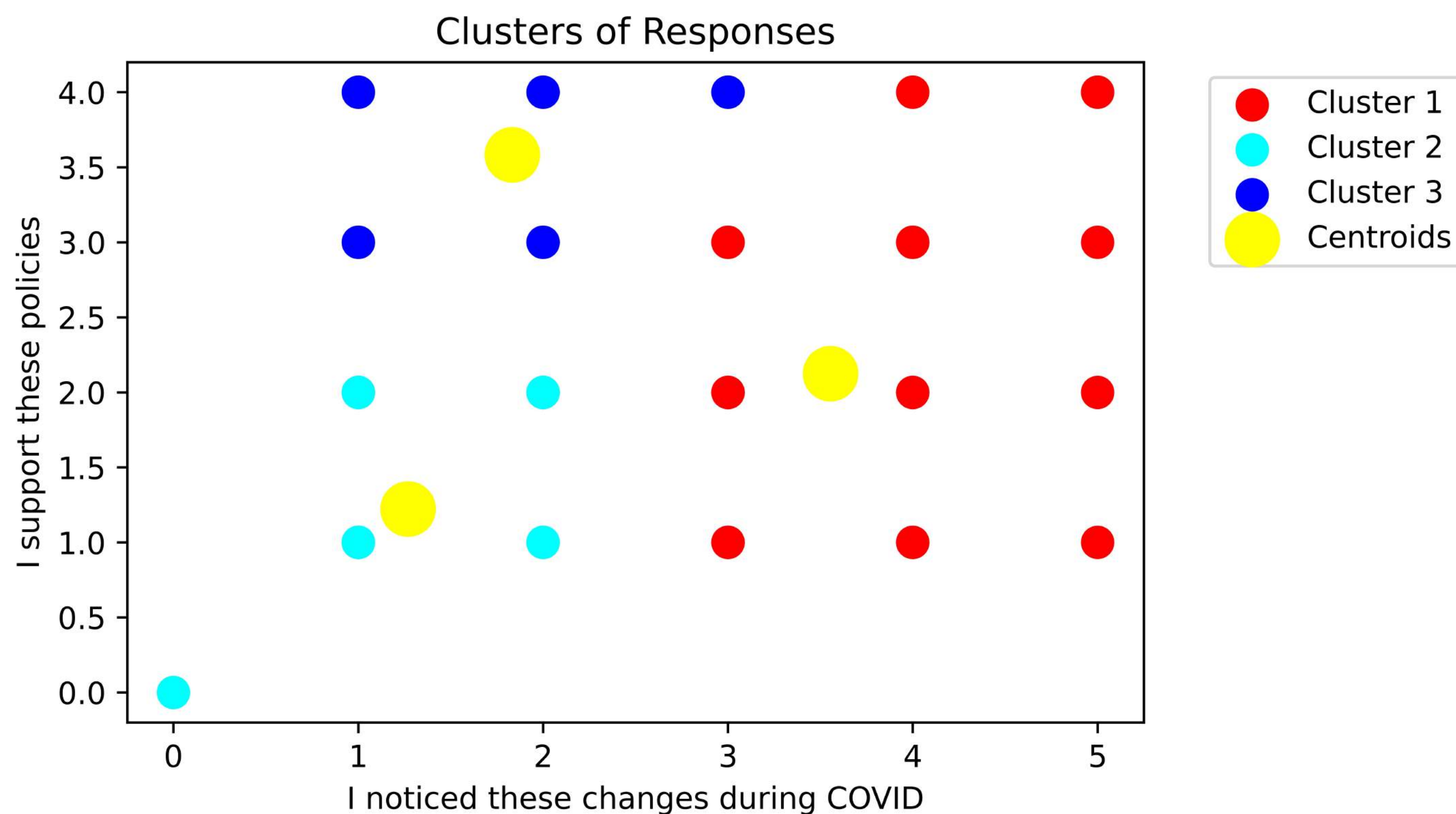
"Statement on Mobility Policy" VS "I support these policies"

- If this survey mirrors the London populace, then by this clustering analysis and then by this clustering analysis:
- Having a total of 163 respondents who do not support the actions being taken and wish it would cease, it's unlikely that the populace would embrace any alternative or unorthodox mobility service or product.
- A company with an alternative or unorthodox mobility service or product must look for inventive ways to win more than half of the populace who are against the policy. If they want to survive in this space they'd need to pour a lot of money into marketing and advertisement.



Cluster Analysis 4

"I noticed these changes during COVID" VS "I support these policies"





Cluster Analysis 4



"I noticed these changes during COVID" VS "I support these policies"

Cluster 1 — This is the group that strongly support London becoming more 'mobility-friendly' but do not believe progress is quick enough.

Cluster 2 — This is the group that do not support the actions being taken and wish they would cease.

Cluster 3 — This is the group that thinks the local government is on the right track and going at the right speed



Cluster Analysis 4 — Comments



"I noticed these changes during COVID" VS "I support these policies"

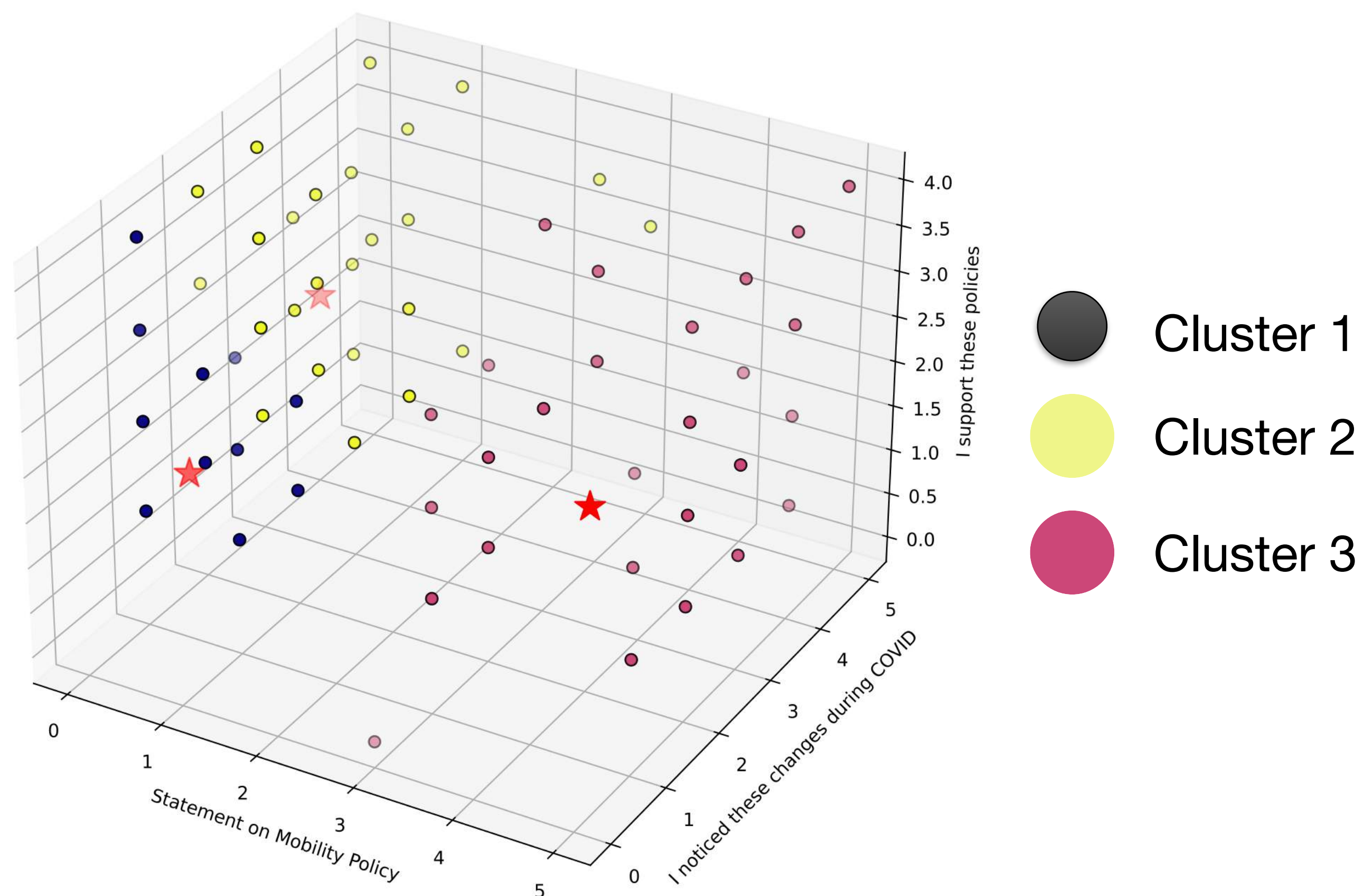
- If this survey mirrors the London populace and then by this clustering analysis and then by this clustering analysis:
- Having a total of 193 respondents who do not support the actions being taken and wish it would cease, it's unlikely that the populace would embrace any alternative or unorthodox mobility service or product.
- A company with an alternative or unorthodox mobility service or product must look for inventive ways to win more than half of the populace who are against the policy. If they want to survive in this space they'd need to pour a lot of money into marketing and advertisement.



Cluster Analysis 5

'Statement on Mobility Policy' VS 'I noticed these changes during COVID'
VS "I support these policies"

3D Clusters of Responses



Respondents per Cluster





Cluster Analysis 5



'Statement on Mobility Policy' VS 'I noticed these changes during COVID'
VS "I support these policies"

Cluster 1 — This is the group that do not support the actions being taken and wish they would cease.

Cluster 2 — This is the group that thinks the local government is on the right track and going at the right speed

Cluster 3 — This is the group that strongly support London becoming more 'mobility-friendly' but do not believe progress is quick enough.

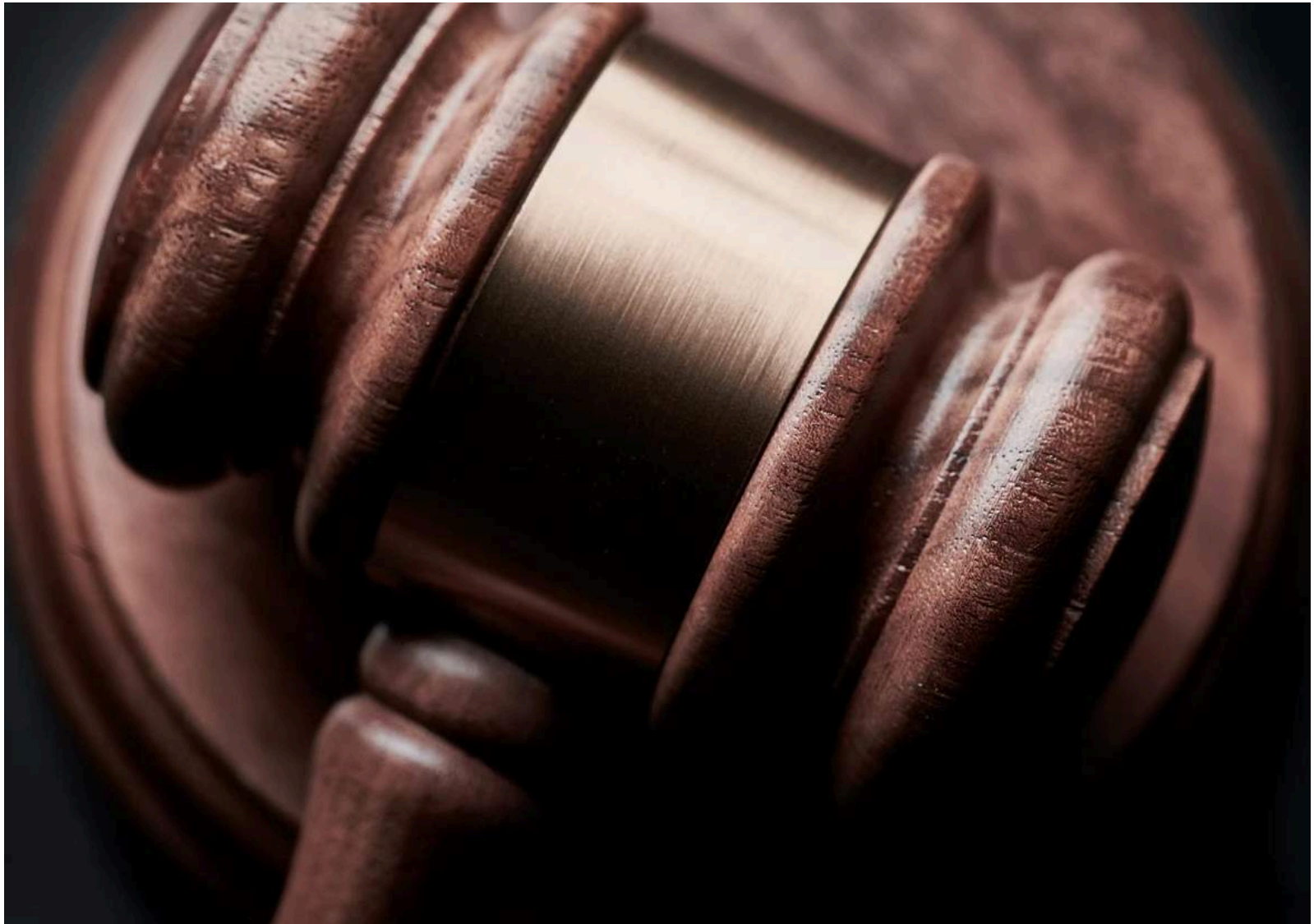


Cluster Analysis 5 — Comments



'Statement on Mobility Policy' VS 'I noticed these changes during COVID' VS
"I support these policies"

- If this survey mirrors the London populace, then by this clustering analysis:
- Having a total of 161 respondents (53% of the population) who do support the actions being taken by local government, it's likely that the populace would embrace any alternative or unorthodox mobility service or product.
- Even if the majority of the populace do embrace an unorthodox mobility service or product made available by a company, the company still has the task of trying to winning over some person in the remaining 47% populace.



Conclusion

Conclusion

- Created 3 archetypes/clusters and explored in 5 different cluster analyses.
- For all the scenarios explored in the 5 cluster analyses, approximately half of the respondent do not support policies and would rather have all actions being taken to cease.
- If these survey is a reflection of the whole London populace, then any company with unorthodox mobility service or product will have an hard time breaking even in London, although it is possible.
- For a mobility company (especially an unorthodox one), to succeed in London, it must look for inventive ways to appeal to the cluster of populace who do not support the mobility policies.

THANK
YOU





Cluster Analysis

Youpele Michael