Segmenting the potential market for the E-Systems® software development company

ALAA MAHJOUB

Thursday, 18 June 2020

AGENDA

- Project goal and objectives
- Data sources
- Data acquisition and cleaning
- Basis of E-Systems® market segmentation
- Using the K-Means clustering and agglomerative clustering algorithms
- Results
- Recommendations for E-Systems® management
- Conclusion

PROJECT GOAL AND OBJECTIVES

This project intends to support E-Systems® new market development strategy by segmenting the new world-wide market into three marketing segments

- **Objectives:**
  - Provide the management of E-Systems® with the information that helps them determining which capital cities will be managed by which marketing team
  - Provide the Marketing Department with the size of the new potential market in each market segment
  - Provide the Organization and Human Capital Planning Department with the information that helps them organize for the new market and uplift the capabilities of the company teams to ensure the success of the market expansion project
  - Support E-Systems® shareholders to better understand and monitor the market expansion project

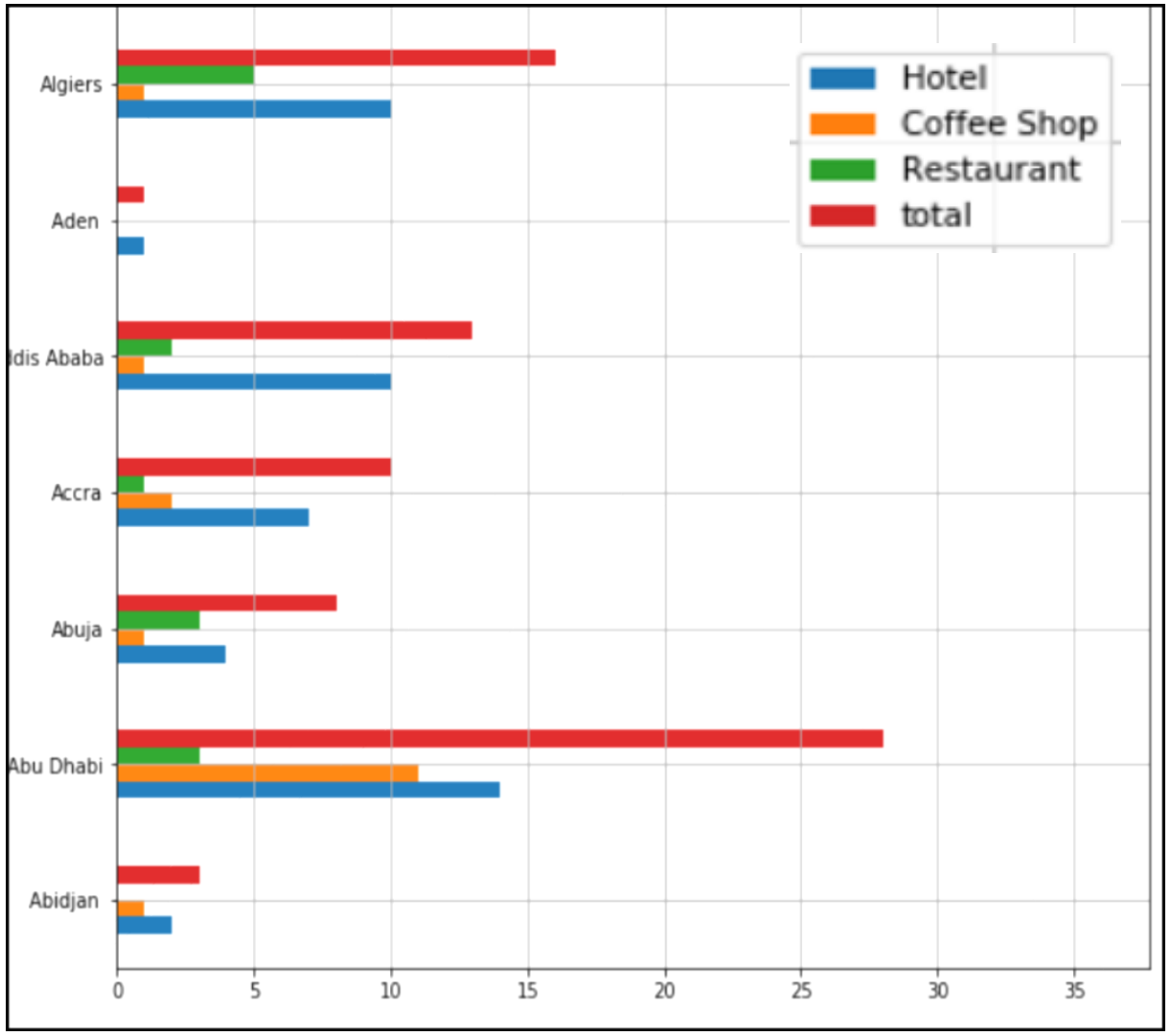
DATA SOURCES

- List of world-wide national capital cities is obtained from the Wikipedia website https://en.wikipedia.org/wiki/List_of_national_capitals
- Geo-Location data of each national capital city is obtained from the Internet using Python geocoding web services API
- World-wide potential-customers data is obtained by exploring the national capitals venues using the Foursquare API
- The world map GIS data is obtained from the Folium API

DATA ACQUISITION AND CLEANING

N0	Dataset	Dataset Type	Description	Number of records
1	National capital cities (raw)	Master data	Raw data of the national capital cities (extracted from the Wikipedia website)	260
2	National capital cities (used)	Master data	Data of national capital cities that have venue information in the Four-Square database	239
3	National capital cities venues	Features data	Data of the venues belonging to the 239 national capital cities	2520

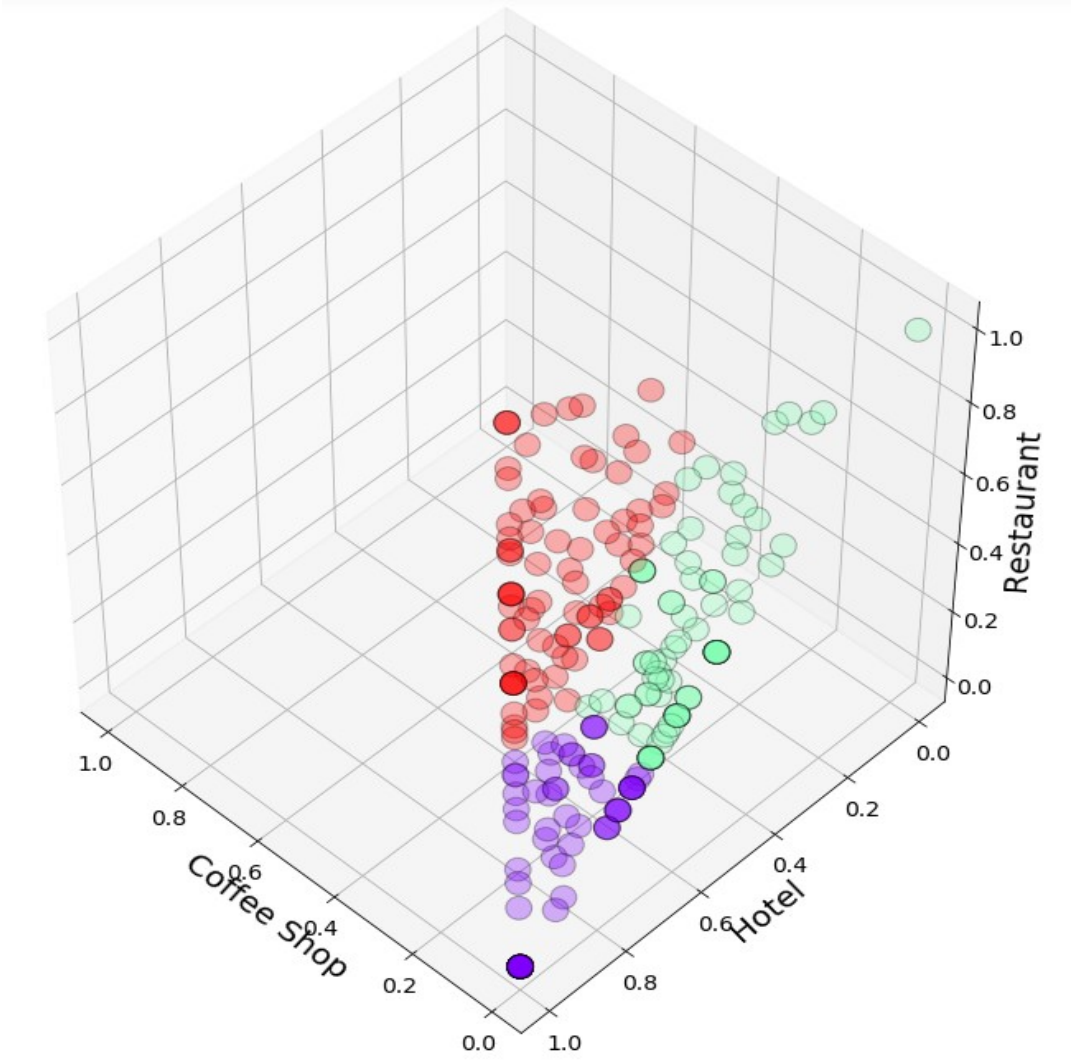
SEGMENTING E-SYSTEMS® WORLD-WIDE POTENTIAL MARKET



E-Systems® new world-wide market was segmented based on …

… the frequency of occurrence of coffeeshop, hotel and restaurant venues

USING THE K-MEANS CLUSTERING ALGORITHMS

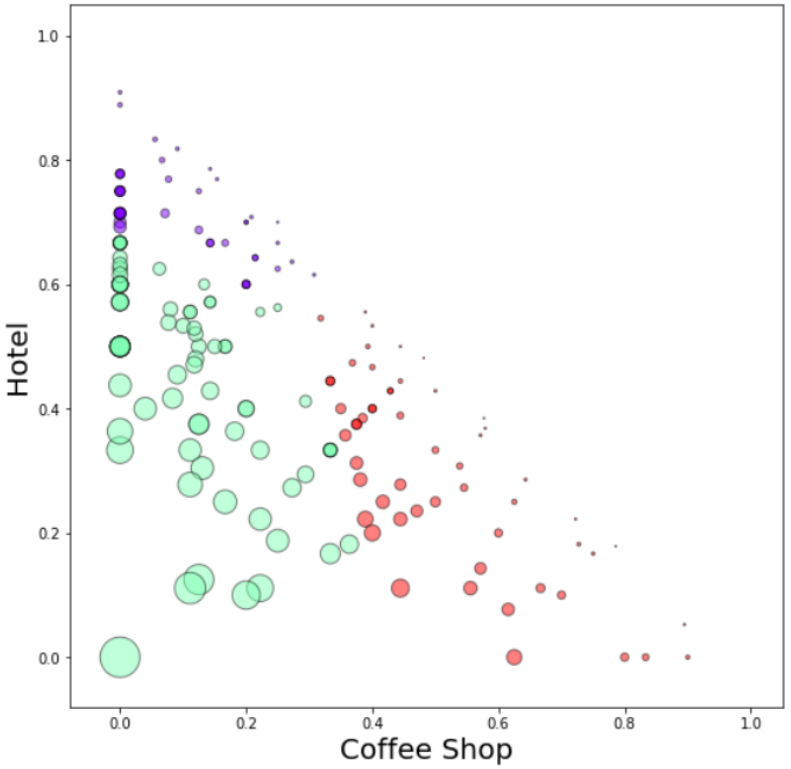


Distribution of cities based on the frequency of coffeeshop, hotel and restaurant venues

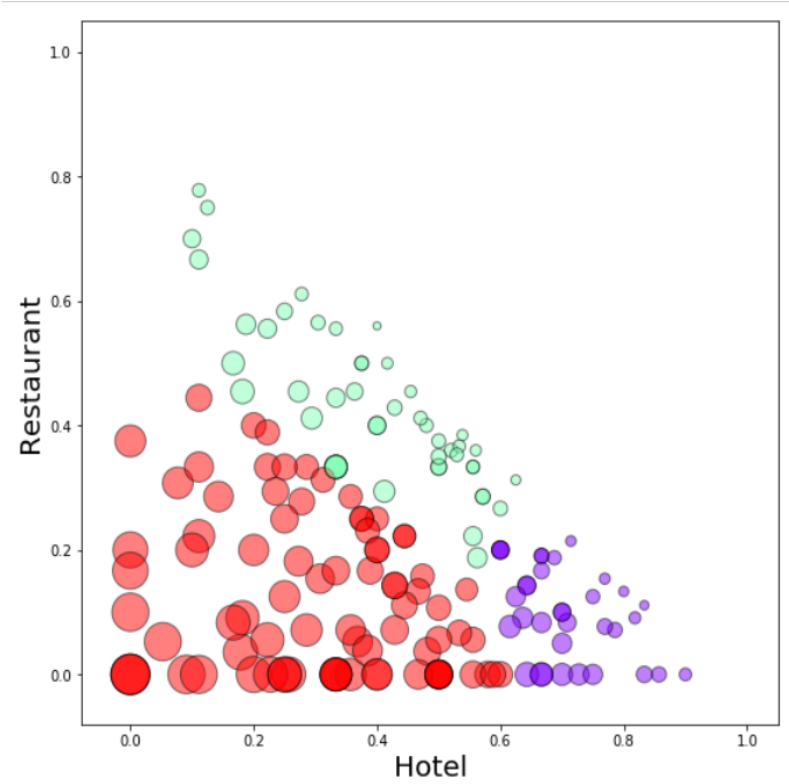
Both K-Means and Agglomerative Clustering algorithms were used to segment the E-Systems® market and …

… they produced the same results

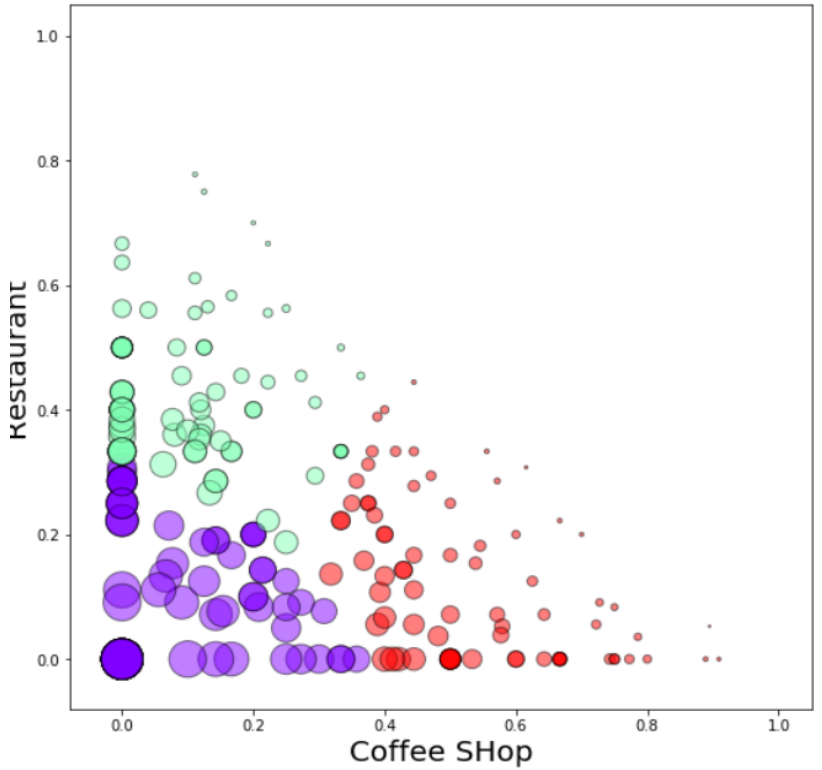
USING THE K-MEANS CLUSTERING AND AGGLOMERATIVE CLUSTERING ALGORITHMS



Distribution of cities based on the frequency of coffeeshop and hotel venues

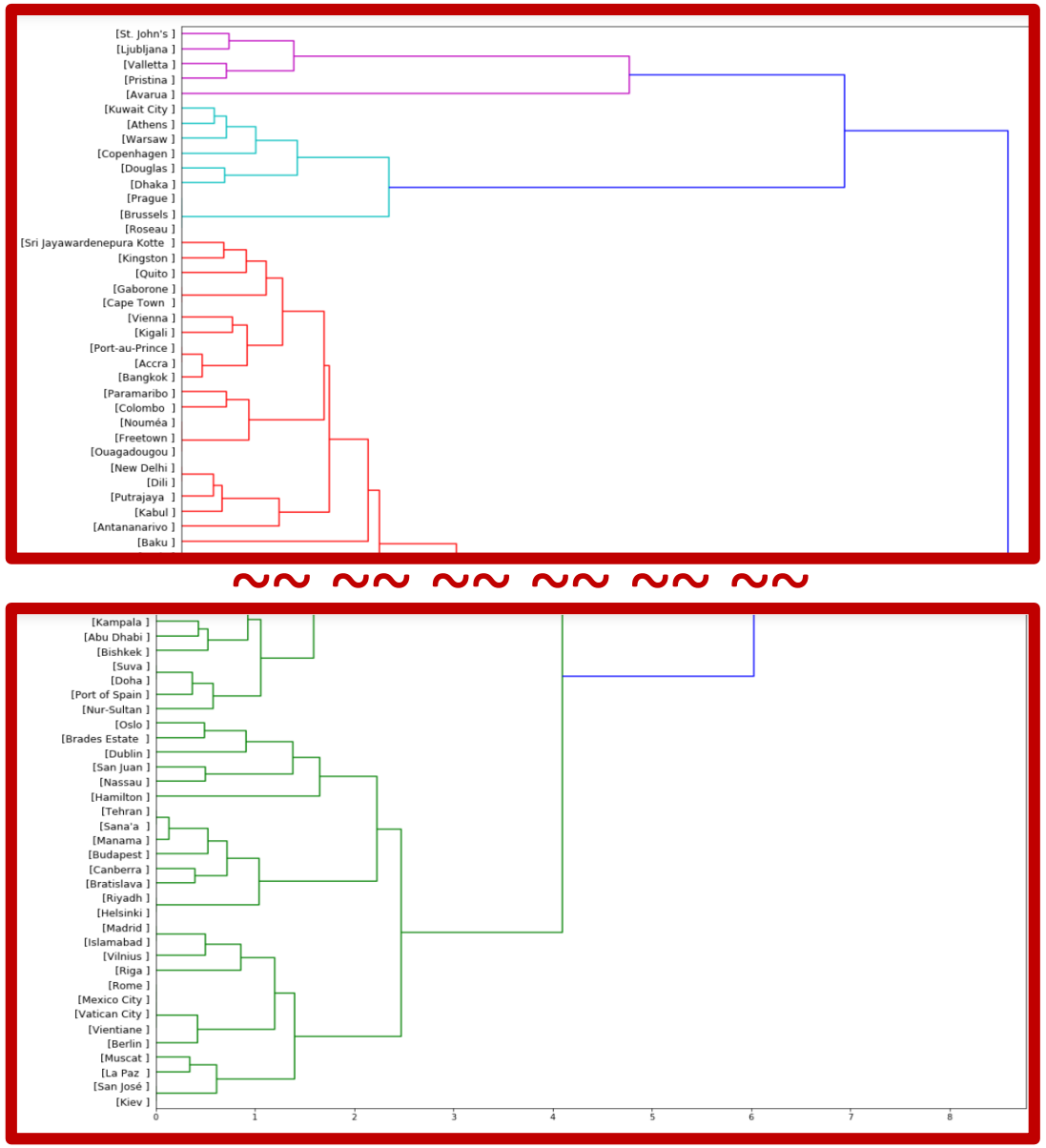


Distribution of cities based on the frequency of hotel and restaurant venues



Distribution of cities based on the frequency of coffeeshop and restaurant venues

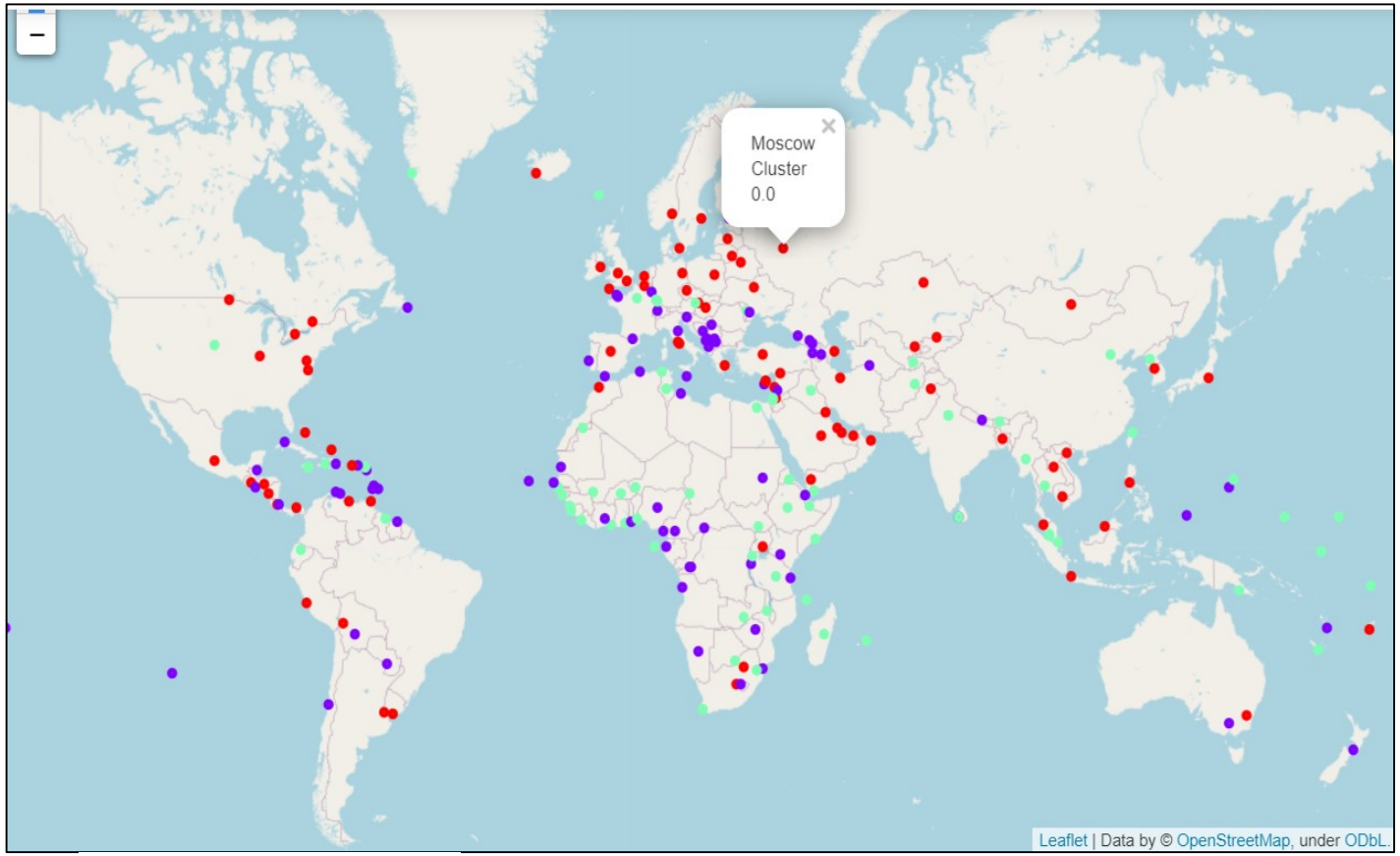
USING THE AGGLOMERATIVE CLUSTERING ALGORITHMS



- A Dendrogram generated by the hierarchical agglomerative clustering algorithm showing the distribution of national capital cities on the three market segments
- Applying the hierarchical agglomerative clustering algorithm gave the same clusters depicted by the diagrams in the two previous slides

RESULTS

Distribution of national capital cities on the three market segments



Market Segment 1 Market Segment 2 Market Segment 3

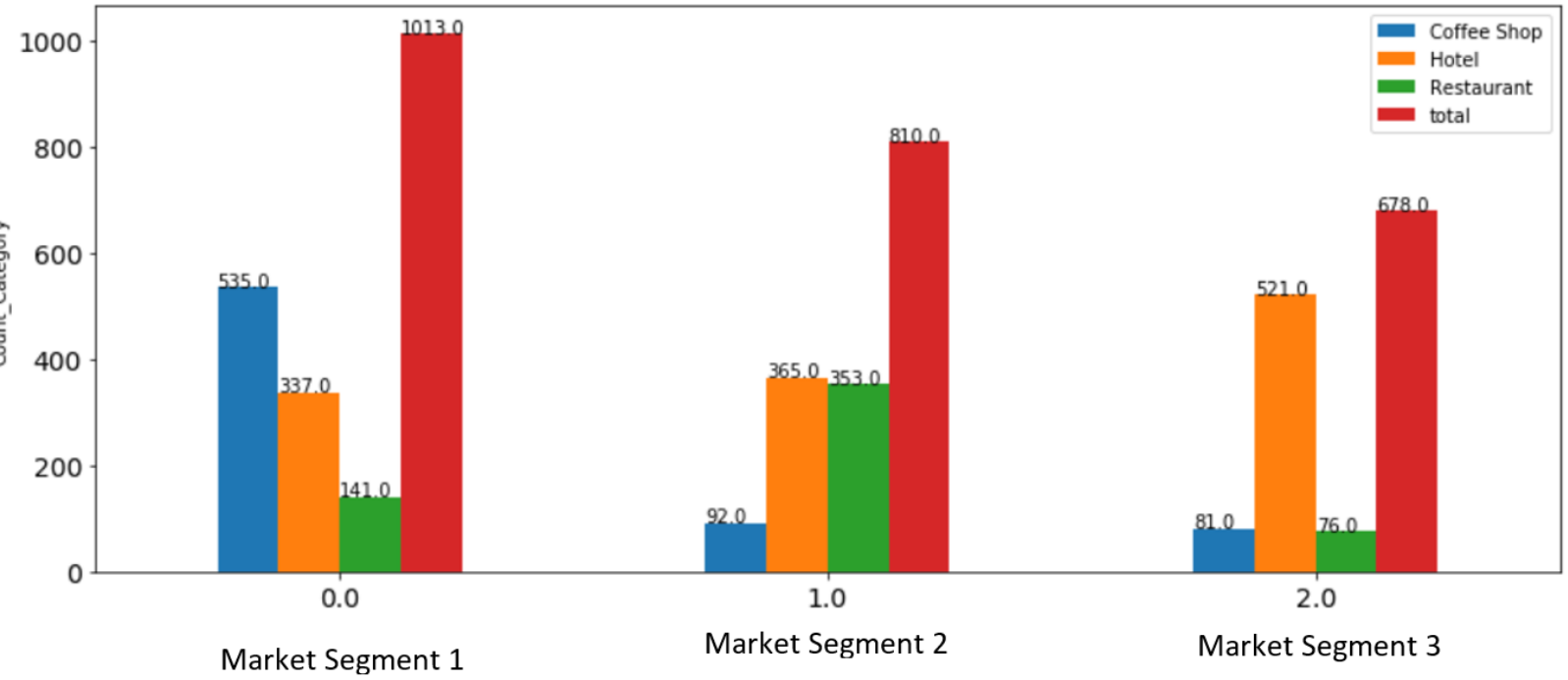
Number of coffeeshops, hotels and restaurants in each market segment

Cluster Labels	Coffee Shop	Hotel	Restaurant	total
0.0	535.0	337.0	141.0	1013.0
1.0	92.0	365.0	353.0	810.0
2.0	81.0	521.0	76.0	678.0

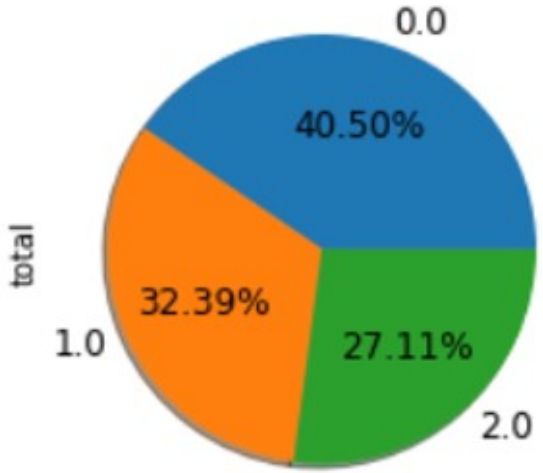
Number of cities in each market segment

Clus_km				
0	77	77	77	77
1	80	80	80	80
2	82	82	82	82

RESULTS



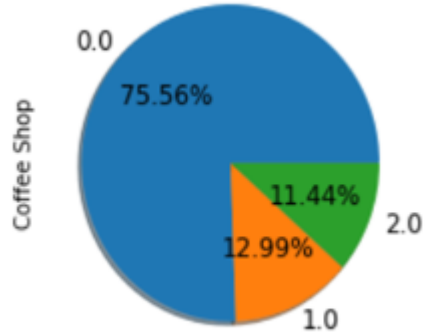
Distribution of coffeeshops, hotels and restaurants on market segments



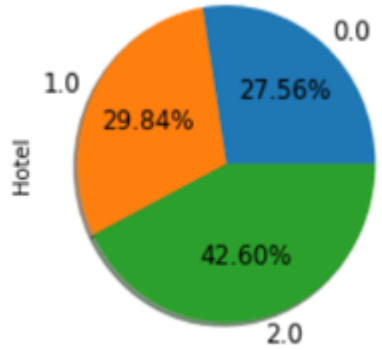
Distribution of all potential market customers on market segments

RESULTS

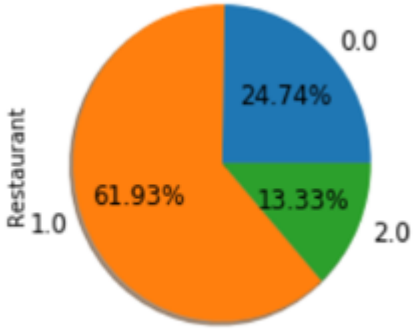
Density of coffeeshops, hotels and restaurants in each market segment



“Coffeeshop oriented” market segment



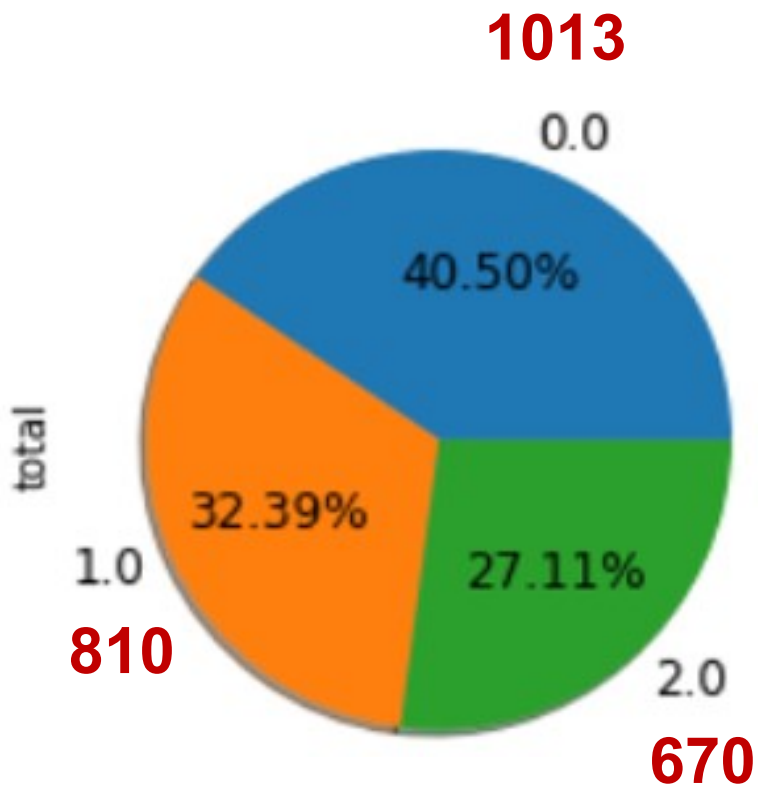
“Hotel oriented” market segment



“Restaurant oriented” market segment

RECOMMENDATIONS FOR E-SYSTEMS® MANAGEMENT

It is worth noting that the new market workload is not equally distributed among the three market segments. The number of potential customers in each market segment is somewhat different (1013, 810 and 670 for market segments 1,2, and 3 respectively). Based on this, the management of E-Systems® is recommended to consider adjusting its organization structure and providing the necessary human capital capabilities in order to successfully implement the new marketing strategy and cope with the new market requirements



Distribution of all potential customers on market segments

CONCLUSION

- This project used the K-Means and the Agglomerative clustering machine learning models to segment the new world-wide market of E-Systems®.
- The frequency of occurrence of hotel, restaurants and coffeeshops were identefied as important features that affect the segmentation of this market.
- These models can be very useful in helping E-Systems® management in several ways. For example, it could help develop a new organization chart, and plan the human capital and competencies necessary to implement the company's new marketing strategy

Thank You