

Capstone Project Report

The Battle of Neighborhoods in Berlin

Coursera Applied Data Science with Capstone



Source: <https://media.globalchampionstour.com/cache/750x429/assets/>

Agenda

Introduction

Data

Methodology

Results and Discussion

Conclusion

Introduction

- New café opening in Berlin, Germany
- Berlin is chosen because its known hipster community
- The new café is a bio vegan café
- Which place in Berlin is the most suitable for the new business venture?
- Important factor of choosing the place is the number of competition which offers similar services
- Recommendation based on venue locations and cluster analysis

Data

Data sources

- List of boroughs and their neighbourhoods in Berlin collected from Wikipedia
- Number of coffee places and location in every neighbourhood will be obtained using Foursquare API

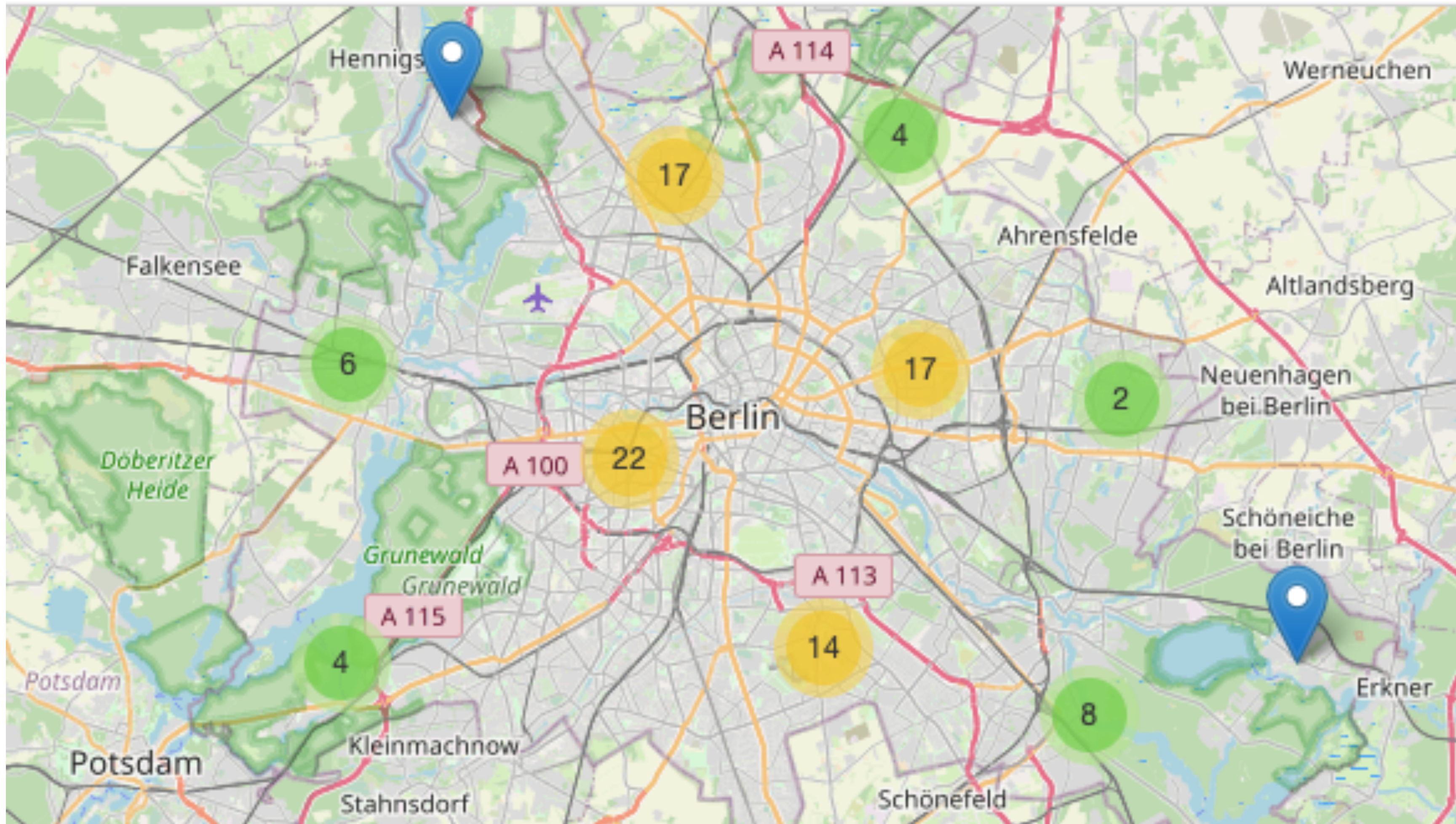
Data

Data extraction

- Data preparation
- Foursquare API usage
- Plot map of Berlin

Data

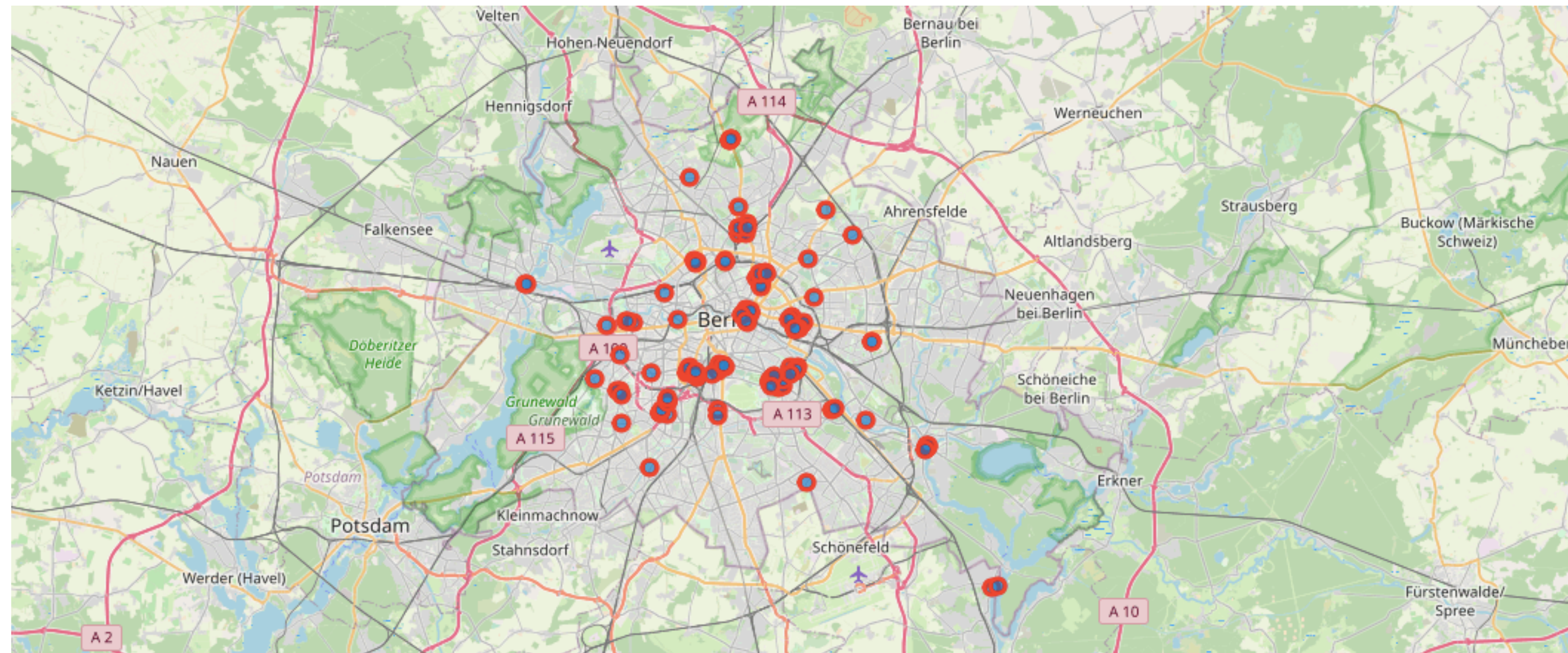
Map of Berlin



Methodology

Exploratory analysis

- Data filtering by “Café”, “Coffee Shop”, “Vegetarian/Vegan Restaurant”
- Plot map of the top venues based on the filtered data frame



Methodology

Exploratory analysis

- One hot encoding
- Grouped the samples by Neighborhood and by taking the each category's mean frequency
- For each one of the neighbourhood the frequency of the coffee places was extracted

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Alt-Treptow	Café	Vegetarian / Vegan Restaurant	Coffee Shop
1	Baumschulenweg	Café	Vegetarian / Vegan Restaurant	Coffee Shop
2	Blankenfelde	Café	Vegetarian / Vegan Restaurant	Coffee Shop
3	Charlottenburg	Café	Vegetarian / Vegan Restaurant	Coffee Shop
4	Dahlem	Café	Vegetarian / Vegan Restaurant	Coffee Shop
5	Falkenberg	Café	Vegetarian / Vegan Restaurant	Coffee Shop

Methodology

K-means clustering

- Number of clusters - 5
- Approach - popularity recommendation
- Combine the top venues with the cluster labels in one data frame

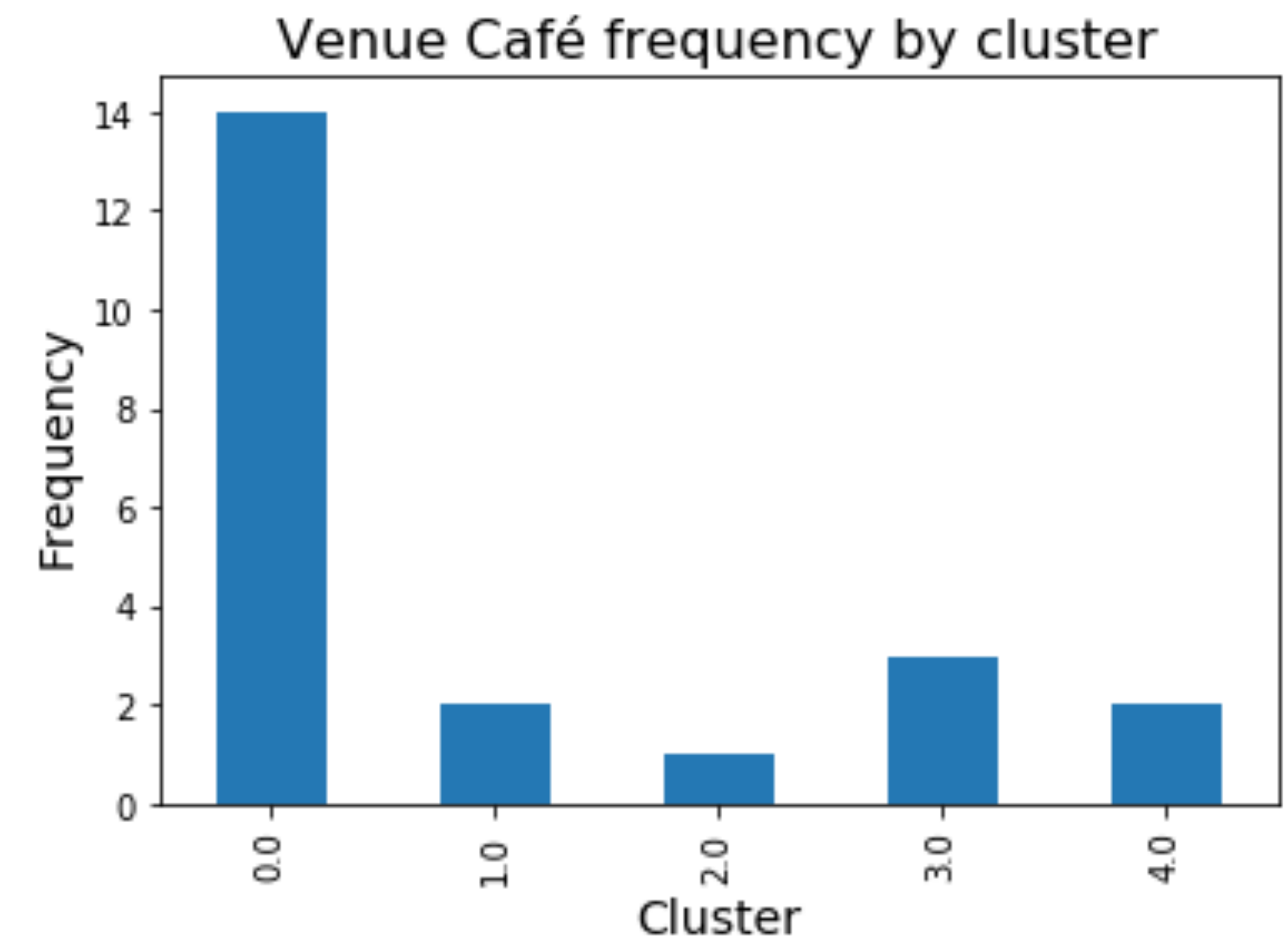
	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Charlottenburg-Wilmersdorf	Charlottenburg	52.516667	13.300000	0.0	Café	Vegetarian / Vegan Restaurant	Coffee Shop
1	Charlottenburg-Wilmersdorf	Wilmersdorf	52.483333	13.316667	2.0	Coffee Shop	Vegetarian / Vegan Restaurant	Café
2	Charlottenburg-Wilmersdorf	Schmargendorf	52.477222	13.288056	3.0	Café	Coffee Shop	Vegetarian / Vegan Restaurant
3	Charlottenburg-Wilmersdorf	Grunewald	52.483333	13.266667	0.0	Café	Vegetarian / Vegan Restaurant	Coffee Shop
4	Charlottenburg-Wilmersdorf	Westend	52.516667	13.283333	0.0	Café	Vegetarian / Vegan Restaurant	Coffee Shop

Results and Discussion

- The most popular place from the three categories was Café
- In some neighbourhoods such as Gesundbrunnen, Mitte, Prenzlauerberg, and Schmargendorf, the popularity was shared among Café and Coffee Shop
- Moreover, some neighbourhoods such as Friedrichshein and Neukölln, where the venue Vegan Restaurant was also a popular place

Conclusion

- The best area to open a new café is Cluster 2
- It is followed by Cluster 1 and Cluster 4
- There is no clear way to know, based only on the Foursquare information, which café is vegan or offers only organic products
- Further analysis should be carried out to create more precise recommendations for the new business venture



Thank you!