1. **Data**
   1. **Data sources**

The analysis is based on location data from Foursquare regarding venues around MPSes within radius of 1000 m in Bonn, Germany area. The list of MPSes has been searched based on Foursquare's category 'Mobile Phone Shop'.

* 1. **Data scrapping**

In this section I have identified the existing MPSes and presented them on the map. I have started from identification of central latitude and longitude of Bonn, Germany. Then, I have leveraged the FourSquare API to obtain URL that leads to the raw data in JSON form. I have searched for Mobile Phone Shops within radius of 10 km from the center of Bonn. I have found in total 35 MPSes.

I have realized that names of MPSes are not unique so I have added to the list of MPSes a unique identifier.

Next I have searched for venues around identified MPSes within radius of 1 km. Again I have used Foursqare API to obtain raw data in JSON form. I have scraped the raw data in this JSON file in order to retrieve the following attributes of the venues surrounding MPSes: name, category, latitude and longitude.

Final list of first five MPS is as follows.

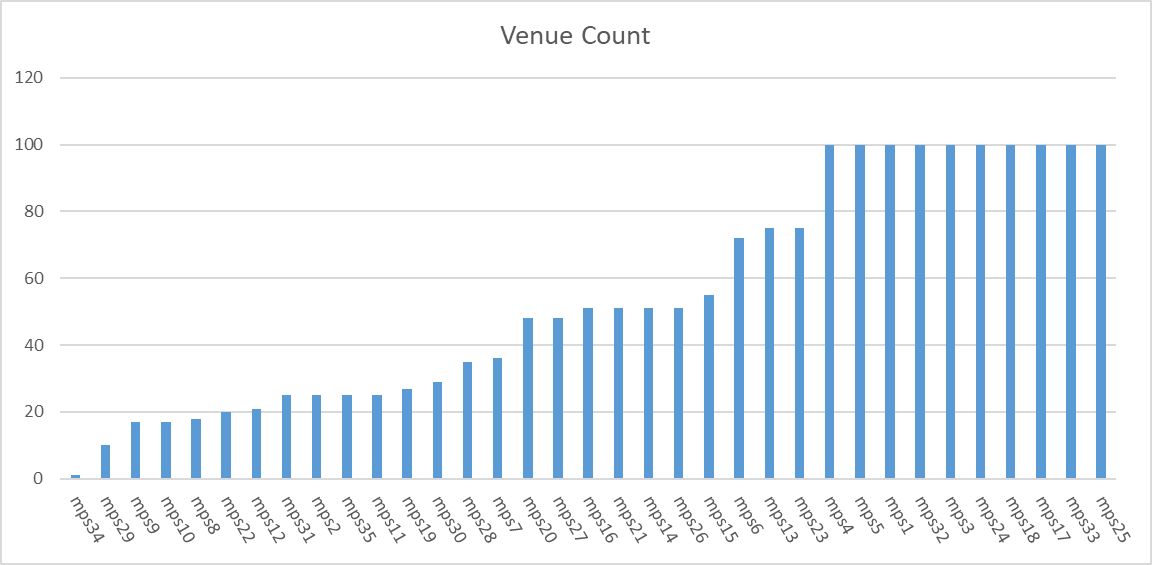
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | name | categories | lat | lng | id | mps |
| 0 | Telekom Shop | Mobile Phone Shop | 50.733911 | 7.100555 | 596d946a79f6c7178f5f5596 | mps1 |
| 1 | Vodafone Shop | Mobile Phone Shop | 50.814173 | 7.159805 | 536766bf498e423eae106986 | mps2 |
| 2 | o2 Shop Bonn | Mobile Phone Shop | 50.735623 | 7.098684 | 4c407f274a3e03bb03f56d0e | mps3 |
| 3 | Vodafone Shop | Mobile Phone Shop | 50.736130 | 7.098090 | 59a45dfd35d3fc3e2ecca93e | mps4 |
| 4 | Vodafone Shop | Mobile Phone Shop | 50.733718 | 7.099008 | 580e321238fa1f1ec7b8b010 | mps5 |

* 1. **Data cleaning**

First I have verified whether all results of search for MPSes are from category ‘Mobile Phone Shop’. I have found in the results some venues with different category although I have used in the search the identifier of this ‘MPS’ category provided by Foursquare. I have dropped those results.

Next I have verified the numbers of venues that have been searched around MPSes. For one of the MPSes there result was only ‘1’ so I have decided to exclude this MPS from clustering. The chart below presents the count of venues by MPSes.

Chart 1 – Number of venues by MPSes



In total I have found 1908 venues in 152 unique categories.

* 1. **Data preparation**

After cleaning the data still needed some more processing before it was suitable for clustering. First I have used one-hot encoding to get dummy variables.

Next I have grouped rows by MPS and calculated the mean of the frequency of occurrence of each venue category.

For further analysis I have used the top 10 venues for each MPS and put them into a new data frame.

The top 10 venues for a sample of MPS is presented below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| MPS | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
| mps1 | Café | Pub | Plaza | Italian Rest’nt | German Rest’nt | Bakery | Hotel | Bar | Coffee Shop | Park |
| mps10 | Clothing Store | Café | Tram Station | Italian Rest’nt | Big Box Store | Mobile Phone Shop | Electronics Store | Drugstore | Shopping Mall | Shopping Plaza |
| mps11 | Ice Cream Shop | Falafel Rest’nt | Castle | Chinese Rest’nt | Shopping Mall | Clothing Store | Department Store | Rest’nt | Drugstore | Optical Shop |
| mps12 | Ice Cream Shop | Greek Rest’nt | Train Station | Gym | Food | Chinese Rest’nt | Soccer Field | Bus Stop | Pizza Place | Gourmet Shop |
| mps13 | Italian Rest’nt | Hotel | Drugstore | Supermarket | Bakery | Middle Eastern Rest’nt | German Rest’nt | Rest’nt | Farmers Market | Metro Station |