

# Analysing the Neighbourhoods of Vienna (AT) to select a Restaurant Location

IBM Applied Data Science Capstone Project

Daniel Kunaver

20<sup>th</sup> February 2021



# Overview

- Introduction
- Data
- Methodology
- Analysis
- Discussion of Results
- Conclusion & Recommendation

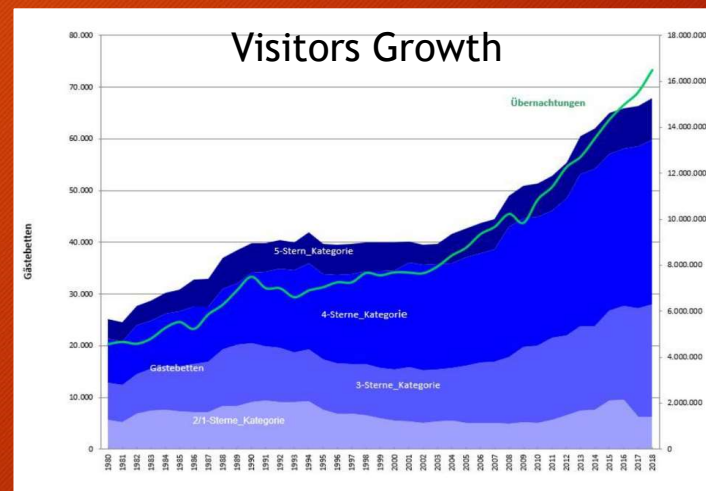
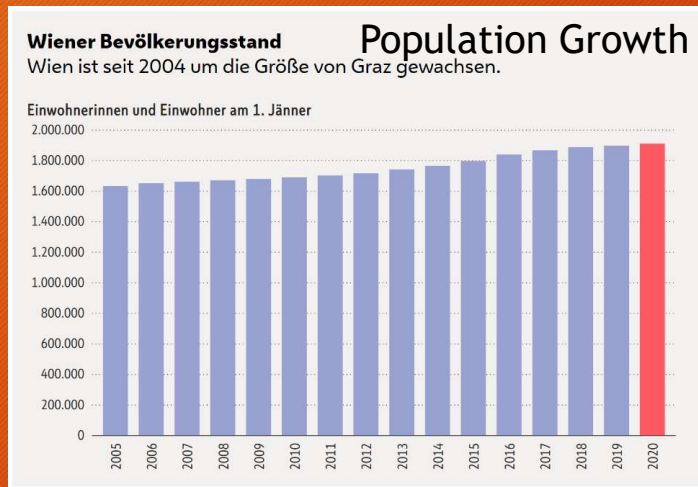




# Introduction



- City ranked top in quality of life
- City is growing in terms of population and tourist destination
- Opportunity to invest in restaurant to cater for locals and visitors



# Data



- The following data sources were used to get information about the city:
  - List of districts of Vienna (source: [https://de.wikipedia.org/wiki/Wiener\\_Gemeindebezirke](https://de.wikipedia.org/wiki/Wiener_Gemeindebezirke))
  - Income of people inside the districts (source: <https://www.wien.gv.at/statistik/arbeitsmarkt/tabellen/einkommen-gesamt-bez.html>)
  - Number of overnight stays by visitors in Vienna (source: <https://www.wien.gv.at/statistik/wirtschaft/tabellen/uebern-bezirk-zr.html>)
  - Foursquare data through the Foursquare developer API (see: <https://de.foursquare.com/developers/apps>)
  - Geolocation data through the geopy library (see: <https://geopy.readthedocs.io/en/stable/>)



# Methodology



- The following steps were taken for the analysis:
  - Data acquisition
  - Data exploration
  - Plot district data on map
  - Find venues in districts via Foursquare query
  - Do cluster analysis
  - Identify from clusters optimum solution
- The details are discussed in the next section

# Analysis - Data Exploration



- The data exploration found that some parameters of the city districts vary significantly and others factors are close together

	Area (hectare)	Population	Employees in District	Population Density (per km <sup>2</sup> )	Avg Annual Net Income (€)	Overnight Stays by Visitors
<b>count</b>	23	23	23	23	23	23
<b>mean</b>	1803.8	83095	42653	10753	23648	23648
<b>std</b>	2290.0	51538	25857	7835	2954	2954
<b>min</b>	109.0	16047	14364	1433	18528	18528
<b>25%</b>	291.8	46691	28474	4135	21387	21387
<b>50%</b>	810.3	76813	31901	8114	23971	23971
<b>75%</b>	2838.6	103776	54827	16918	25601	25601
<b>max</b>	10229.9	207193	108679	27397	29575	29575



# Analysis - Data Exploration

- The data in the graphs shows where there are large differences and where the data is only in a narrow distribution

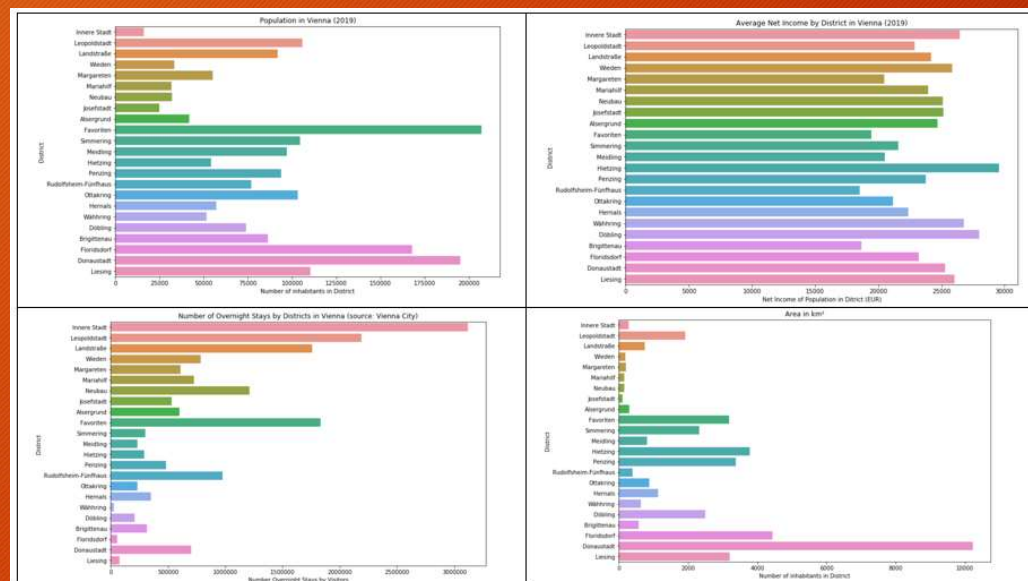


Figure 3: Bar plots on data by district - Top right: Population distribution, Top Left: Average net income (annual), Bottom Right Visitor overnight stays, Bottom Right: Area

# Analysis - Data from Foursquare

- Foursquare query resulted in the following list of restaurants by district
- The data does not line up with the chamber of commerce, which list over 2500 restaurants in the city compared to under 700 from Foursquare

Number of restaurants	
District	
Alsergrund	22
Brigittenau	33
Donaustadt	9
Döbling	28
Favoriten	26
Floridsdorf	16
Hernals	17
Hietzing	12
Innere Stadt	17
Josefstadt	20
Landstraße	22
Leopoldstadt	10
Liesing	19
Margareten	29
Mariahilf	26
Meidling	24
Neubau	20
Ottakring	36
Penzing	28
Rudolfsheim-Fünfhaus	31
Simmering	21
Wieden	25
Währing	33

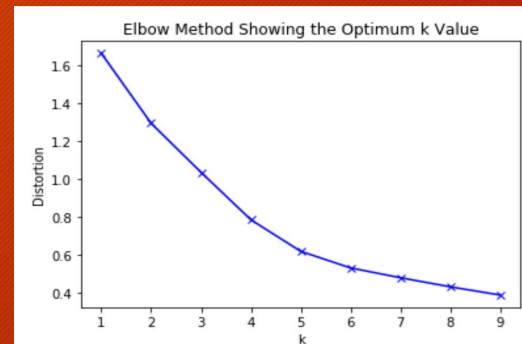




# Analysis - Clustering



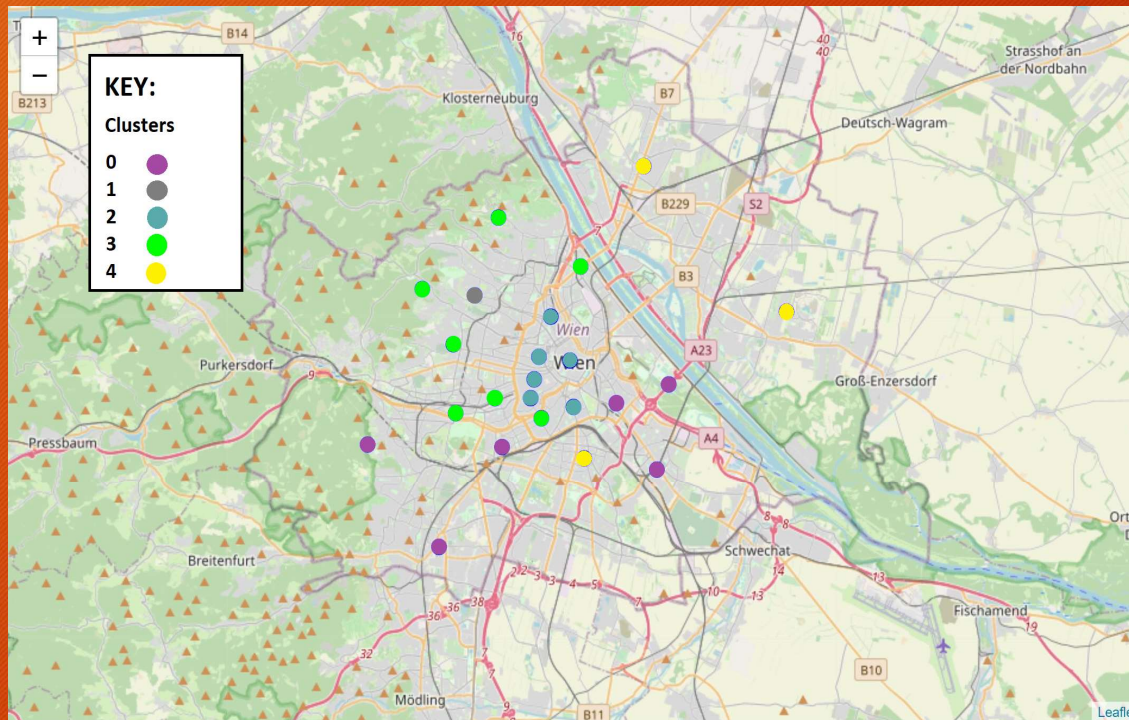
- Cluster algorithm used to find commonalities between districts based on:
  - The net purchasing power per district
  - Density of restaurants per local population
  - The number restaurants per overnight stay
  - The number of restaurants per number of employees in the district
- Carried out analysis (Elbow Method) first to find optimum number of clusters to use on the data set. The result of this work is seen right.
- Optimum k value is 5





# Analysis - Clustering Result of Districts

- The resulting clustering of the districts is shown on the map

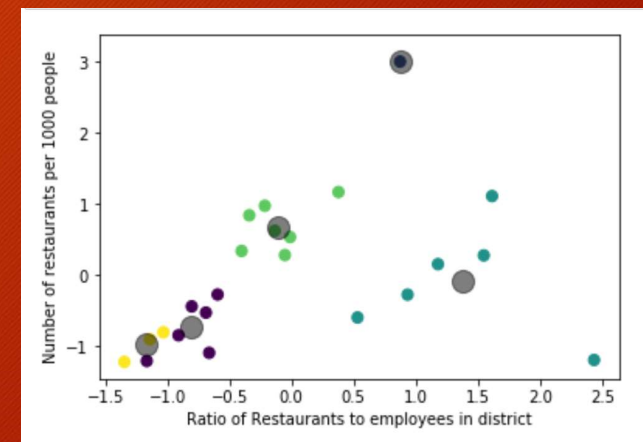
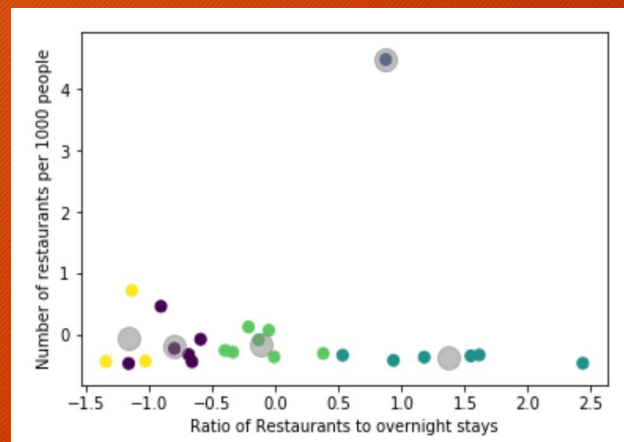
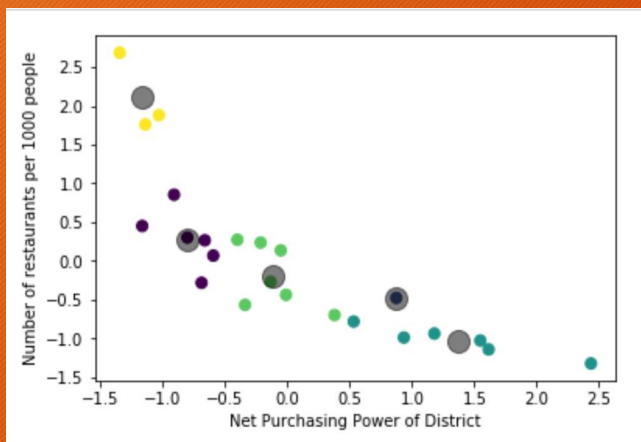




# Analysis - Clustering Results



- The results show that cluster 2 is the most favourable set of districts for opening a restaurant



- The plots show that in cluster 2 (turquoise) the economic parameters look most favourable for opening a restaurant.

# Discussion



- Results show almost predictable result of inner city districts being most favourable
- One orders about the clustering of districts in cluster 0 and 3, because here some high income areas are put together with low income areas
- As the data from Foursquare is not complete the results need to be taken with care



# Conclusion and Recommendation



- The analysis of districts in the city of Vienna used data from different sources and combined it with Foursquare location data to find a good location for opening a restaurant
- The result was that the inner city districts appear most favourable given that here a high density of people live combined with many people working in the city and the number of visitors
- One would need to check the analysis using Foursquare data with an alternative data source to verify the results prior to making an investment decision