**Michelin Star Restaurants Impact on Copenhagen**

Capstone project - Final report

IBM – Coursera

Data Science Specialization

Table of content:

[**I.** **Description Problem:** 2](#_Toc21588500)

[**II.** **Data description:** 3](#_Toc21588501)

1. **Description Problem:**

With a relatively small population compared to other capitals, Copenhagen is a city with many Michelin Star restaurants per capita. Noma restaurant in Copenhagen has been voted as the best Michelin Star Restaurant (MSR) in the world in 2012. The Michelin star experience is booming in Europe and in Copenhagen. People from all over the world, book top restaurants half a year in advance to be part of such an experience. Given that such experience is not low in price (sometimes starting at 350 Euros a person, not including wine pairing), and given that many people are okay with spending such amounts of money, how does an MSR impact its surroundings, neighborhoods, prices for housing and venues (shops, markets etc)?

From an investment perspective, is it worth it to open up a business around MSRs or not? Are housing prices near MSRs more expensive? Are there more food markets near MSRs compared to other areas? These are some of the questions we will try answer through our analysis.

Ultimately, the target audience for this data problem is:

* Investors that would like to open any type of business in the surrounding area
* Farmers, butchers, fishers that would provide MSRs with fresh produce
* Pupils that would like to buy or rent housing: would prices be higher in areas close to MSRs?

1. **Data description:**

To tackle the above problem we will use multiple internet sources to extract data with the help of Python packages such as beautifulsoup4 and pandas:

* For a list of all MSRs in Copenhagen, we will use FourSquare API to extract the list of names together with their geographical location. We can furthermore reconcile the complete list of MSRs against the https://guide.michelin.com website.
* For neighborhood names, location, postal codes we will use the Wikipedia page about Copenhagen. We will map all the neighborhoods and MSRs together on a folium map.
* For housing/rent prices by neighborhood, we will use the statistics Denmark website. This data, together with the zip code data, will be used to plot a folium map including a Chloropleth of prices by postal code or borough.
* For other venue, tips related information we will the FourSquare API. We will try and cluster all the venue data in the optimal number of clusters and map the results on top of the already created Folium map.
* For any other needed information we will scrape the internet with the help of BeautifulSoup4 and Pandas package