

Capstone Data science project

Starting my own pizza place



June 16, 2021

Aniket Khanna

Amsterdam Netherlands

Contents

[introduction 1](#_Toc74748209)

[Research problem & question 2](#_Toc74748210)

# 

# introduction

Many people enjoy good food. Especially fast food that is healthy and cheap. Eating pizza with your friends can be an enjoyment and an easy method to ease your hunger. Not only can we get pizza but we can also get various other consumables such as: soft drinks, cookies and deserts. Pizza stores are multi-purpose restaurants. It can be used for takeaways or eat your meal inside the store. The consumption of fast food has increased dramatically in the last few years. Franchises are expanding business and opening multiple new stores in a lot of location within Amsterdam The Netherlands. As a result of this, there are many different kinds of pizza places in Amsterdam. sole proprietorships as well as big company franchise stores. Opening new pizza stores comes with its own risks and gains. Ass there are many pizza places located in Amsterdam, what area is the best to open one?

This information is interesting for the audience that would like to open a pizza restaurant in Amsterdam or any other kind of business in the area of Amsterdam. as the API shows how many of a unique store is located in a cluster.

# Research problem & question

Research question is: in what area of Amsterdam are the least amount of pizza places and is it economically feasible to start a pizza place in a the area where there are the least amount of pizza places? Based on the outcome of the capstone data science project using data science methodology and machine learning techniques. By using these machine learning abilities, are we able to give concrete information about the best place to start a new pizza place.

The Foursquare API tells us how many of a store are in a certain location and how dens these businesses are among each other. (how close they are from each other). We can divide the density into separate clusters that can visualise the amount of stores in a specific neighbourhood in Amsterdam. with this information we can make an educated judgment on where there is a lot of competition and in which neighbourhoods, we have an opportunity to open a store. This code can be used for other countries and districts as well.

# Data

Based on what we want to know to solve this business case, there are some factors that will influence our outcome and predictions.

1. the amount of pizza places in Amsterdam and in their separate clusters.
2. The longitude and the latitude of the neighbourhoods where these pizza places are located in.
3. And the venue data that we will use to cluster the data.

We will use Wikipedia to extract data from their website. We will use their neighbourhoods in Amsterdam data. There are a total of 107 neighbourhoods in Amsterdam. we will be using webscraping to extract the data from their website. Furthermore, we will be using pandas to transform the data into dataframes that we can use for our clustering. We will use the foursquare API to gather the coordinates and venue data that will be relevant for our clustering.