IBM Data Science Professional Certificated

Capstone Project

Final Report

Zinan Lin

May 21st, 2019

Table of Contents

[Introduction 3](#_Toc9368073)

[Background 3](#_Toc9368074)

[Problem Scope 4](#_Toc9368075)

[Audience 4](#_Toc9368076)

[Data collection 4](#_Toc9368077)

[Data of Michelin Starred Restaurants 4](#_Toc9368078)

[Source 4](#_Toc9368079)

[Description 5](#_Toc9368080)

[Example 5](#_Toc9368081)

[Data of arrondissements in Paris 5](#_Toc9368082)

[Source 5](#_Toc9368083)

[Description 6](#_Toc9368084)

[Example 6](#_Toc9368085)

[Data of Restaurant Information 6](#_Toc9368086)

[Source 6](#_Toc9368087)

[Description 7](#_Toc9368088)

[Example 7](#_Toc9368089)

[Data from Foursquare API 7](#_Toc9368090)

[Source 7](#_Toc9368091)

[Description 7](#_Toc9368092)

[Example 7](#_Toc9368093)

# Introduction

## Background

Paris is the captain city of France, it is famous for its history, arts, and food. Paris has 20 arrondissements, which are regions that the city is separated into, like the picture [1] showed below. There are about 40,000 [2] restaurants in Paris and it is currently the city with the second most Michelin starred restaurants in the world [3]. I am a food lover and I am motivated to find out more about the fine-dining places in Paris.

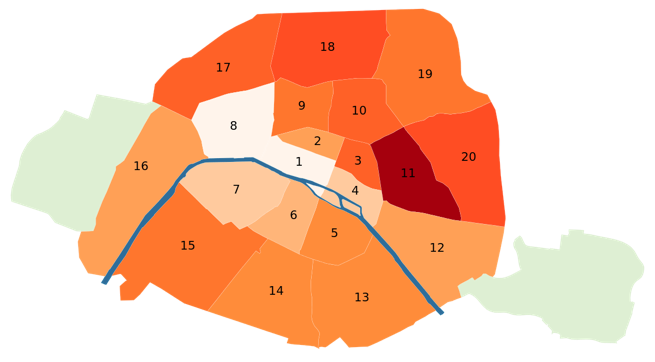


Fig. 1: The arrondissements of Paris

## Problem Scope

What are the similarities in the arrondissements of Paris in terms of restaurants?

Are the restaurants in the central (downtown) areas more favored by clients?

Optionally, how does the number of Michelin Starred restaurants affect the way that that restaurants are clustered?

We will need to find out about the above problems with the support of datasets.

## Audience

The audience of this problem are tourist, who planned to stay in Paris for vacation. Knowing which area/ areas are similar in terms of restaurants can help them deciding on where to live and where to eat.

This problem can also benefit restaurant owners/ businessmen who are interested in knowing the distribution of restaurants in Paris

# Data collection

## Data of Michelin Starred Restaurants

### Source

List of all Michelin starred restaurants:

<https://www.theupcoming.co.uk/2019/01/21/all-the-paris-michelin-star-restaurants-2019-on-a-map-and-full-list/>

### Description

Either using beautiful soup to scrape the web page or manually download the lines into an excel file. The data size is small (~100 rows, 2 columns) and the webpage is not very well organized, therefore it’s easy to manually create an excel file.

### Example

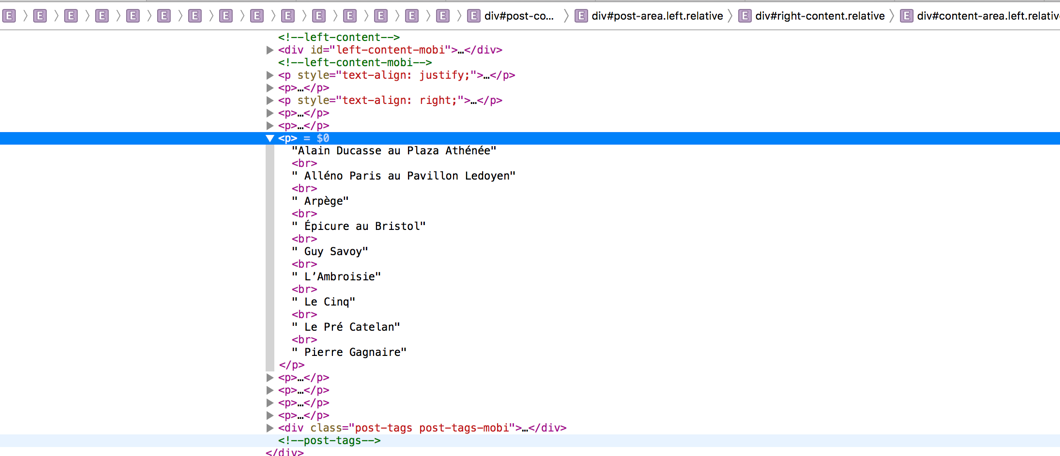


Fig 2. The website HTML, showing the labels of Michelin starred restaurants



Fig 3. Manually downloaded data

## Data of arrondissements in Paris

### Source

List of arrondissements in Paris

<https://opendata.paris.fr/explore/dataset/arrondissements/export/>

### Description

The data is nicely formatted from the official website of Paris open Data. It supports download, so I will import the data as a csv file using pandas.

### Example

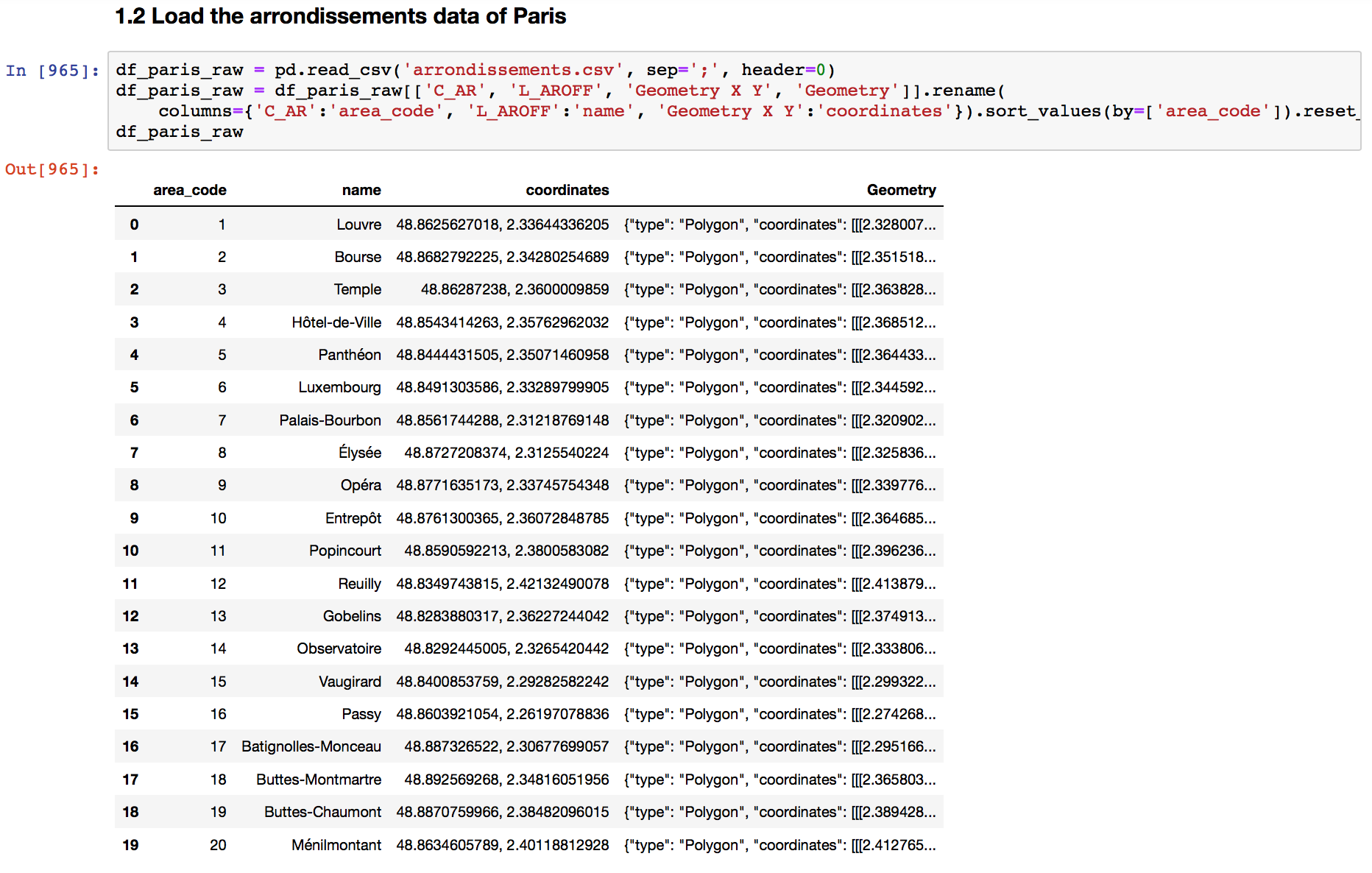


Fig 4. Code Snippet of importing arrondissement data as a data frame

## Data of Restaurant Information

### Source

Kaggle, TripAdvisor Restaurants Info for 31 Euro-Cities:

<https://www.kaggle.com/damienbeneschi/krakow-ta-restaurans-data-raw>

### Description

This is a dataset consisting Ratings and reviews for restaurants across 31 European cities, it’s 28.7MB. I will download it and import it as a CSV file. Later on, I will filter out only restaurants from Paris and use it to find out the details about restaurants I will be analyzing.

### Example

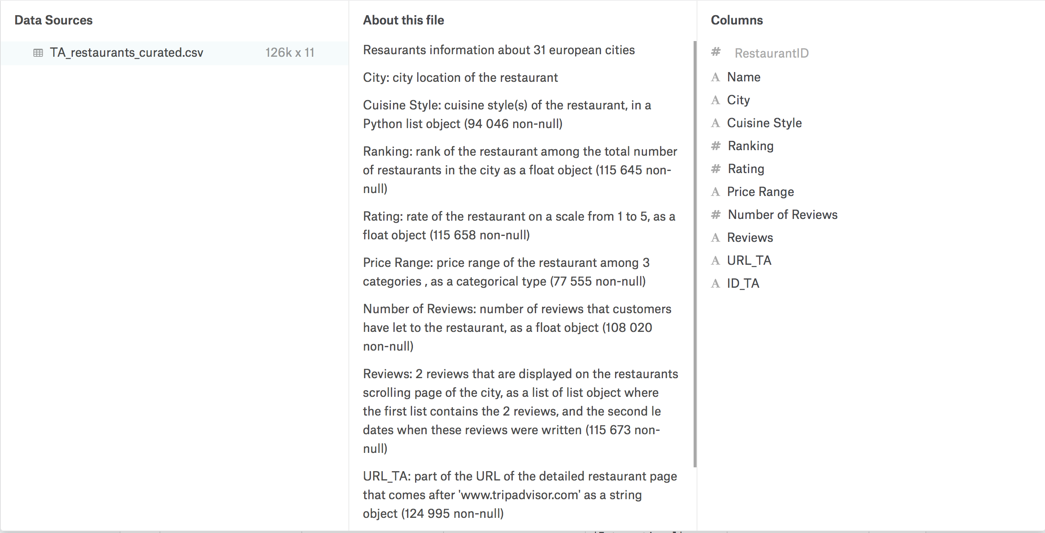


Fig 5, High level Overview of the data

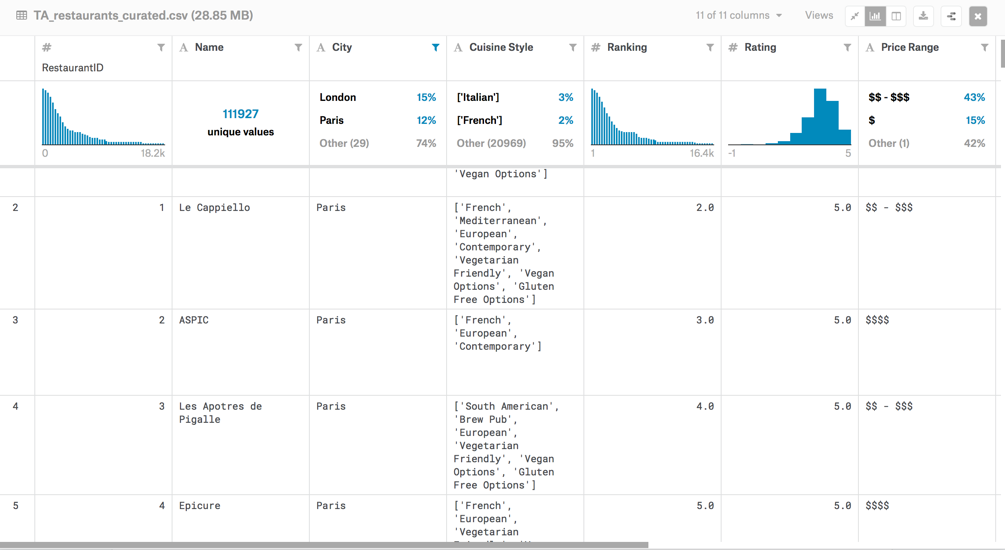


Fig 6, Example of the data

## Data from Foursquare API

### Source

Foursquare API:

<https://foursquare.com/developers/apps>

### Description

Using Foursquare API and it’s search endpoint, I will be able to get data on which arrondissement are restaurants from. As my current subscription, I can make 99,500 Regular Calls / Day, which should be more than enough for me as the total number of restaurants in Paris is about 40,000 [2].

### Example

Documentation for search endpoint: <https://developer.foursquare.com/docs/api/venues/search>



Fig 7, Code snippet of the function that makes call to the API

# Reference

[1] <https://en.wikipedia.org/wiki/Arrondissements_of_Paris>

[2] <https://www.quora.com/How-many-restaurants-are-in-Paris>

[3] <https://www.godsavethepoints.com/2019/02/15/13-most-michelin-starred-cities-in-the-world/>