

Exploring Paris to Make an Informed Choice about Relocating



WEEK 1 submission

IBM Coursera Data Science Capstone Project

Walter Belluco

May 24th 2020

Summary

This report presents the analysis of real estate, demographic, and location data of the city of Paris to help the reader make an informed decision about where to relocate.

The report is based on data in the public domain, retrieved from the opendata initiatives of the city of Paris and the French government. Since data are expressed with a different geographical segmentation, a preliminary work was necessary to aggregate a single dataset enabling the analysis. Location data concerning the most popular venues was obtained from Foursquare API.

In order to facilitate interpretation, results will be presented using a number of visual representations: choropleth maps, dot maps, scatter plots, line plots, heat maps. Content tables are available in the annex.

We shall look at different metrics to describe the 80 neighborhoods and 20 districts of Paris :

- rent value of real estate;
- selling value of real estate;
- location data.

The location data from Foursquare will be used to group neighborhoods into 4 different clusters. Combining main results into the same map will allow to identify at a glance the characteristics of a given neighborhood.

Table of Contents

Summary	2
Table of Contents.....	3
Introduction and business objectives	4
Data.....	4

Introduction and business objectives

Every year a number of people relocate to Paris, the French capital and global center of finance, fashion, arts, science, culture, and history.

Making an optimal choice when relocating can be overwhelming. Which neighborhood should be chosen? A number of factors including renting vs. purchasing price, as well as demographics and the characteristics of the neighborhood must be taken into account.

Several questions may arise :

- Is it better to buy or to rent ?
- Does population density affect real estate value ?
- What are the most common venues for each neighborhood?
- Where are similar neighborhoods located, and what do they have in common ?

This project aims to create a tool to help make informed decisions based on a comparison of available data : rent prices, sales prices, population density and location data using Foursquare.

Different users will be drawn to different conclusions, according to their preferences and financial disposition.

Data

A number of sources can be used to gather relevant information about Paris

- <https://parisdata.opendatasoft.com/> : rental data (80 neighborhoods)
- <https://www.data.gouv.fr/> : administrative districts (20 arrondissements).
- <https://cadastre.data.gouv.fr/> : property sales data in 2019, per district.
- <https://developer.foursquare.com/> : location data.
- <https://fr.wikidia.org/> : population data.

The data was gathered using web-scraping, making calls to the Foursquare API, and downloading geojson files.

Minor data cleansing was done including adding renaming columns, adding district column names, getting rid of duplicates, deleting unnecessary columns.