Capstone Project
Property Prices in Paris

February 2020 - Sebastien Serville
IBM Applied Data Science Course

Table of Content

Business Problem

Paris is one of the most expensive cities in the World when looking at properties. Our questions are:

1. What are the most expensive and cheapest districts or boroughs?
2. What is the average price per sqm for a property in Paris?
3. What is the volume of properties sold in Paris?
4. Do venues located close to a property have an impact on its price?

Data Sources

1. **Paris Property Sales data** (csv file): List of properties sold in Paris between January 18 and June 19 - https://www.data.gouv.fr/fr/datasets/5c4ae55a634f4117716d5656/#_
2. **Paris districts** (geojson file): List of Paris districts with geographical data - https://www.data.gouv.fr/fr/datasets/arrondissements-1/
3. **Paris boroughs** (geojson file): List of Paris boroughs with geographical data - https://www.data.gouv.fr/fr/datasets/quartiers-administratifs/
4. **Foursquare API:** To get list of venues based on a location - http://www.foursquare.com

Methodology

1. Identify data sources
   a. Property prices in Paris by districts and / or by boroughs
   b. Districts and boroughs of Paris with geospatial information
   c. Venues data with their categories
2. Prepare data
   a. Data cleansing and formatting
3. Process and analyse data
   a. Use standard formulas to calculate mean price per sqm and number of sales by district and boroughs
   b. Use data visualisation method to get insight from data
   c. Use clustering method (K-mean) to understand impact of venues on property prices
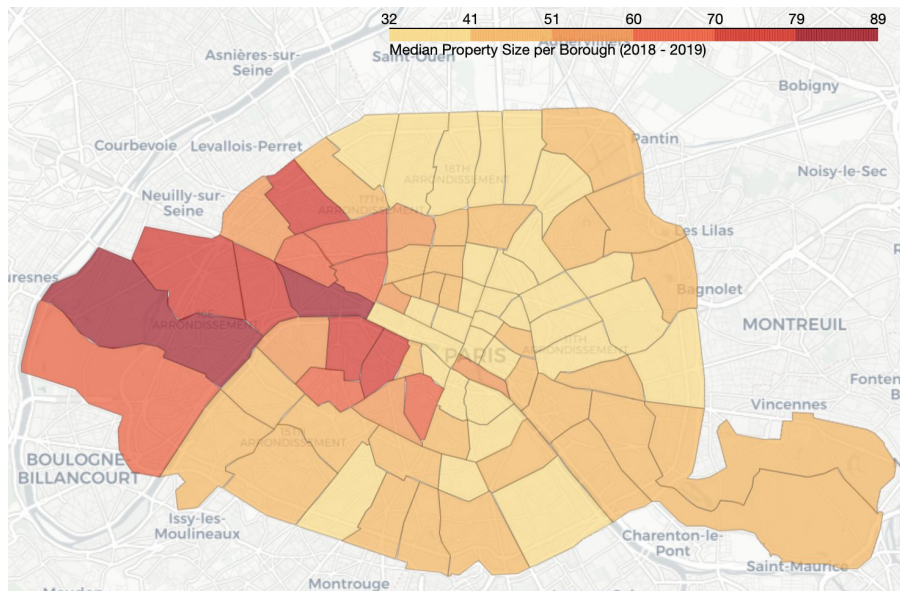4. Review results and conclude

Results - Property Sales

Number of Sales per Borough (2018 - 2019)
19 — 398 — 778 — 1,158 — 1,537 — 1,917 — 2,296
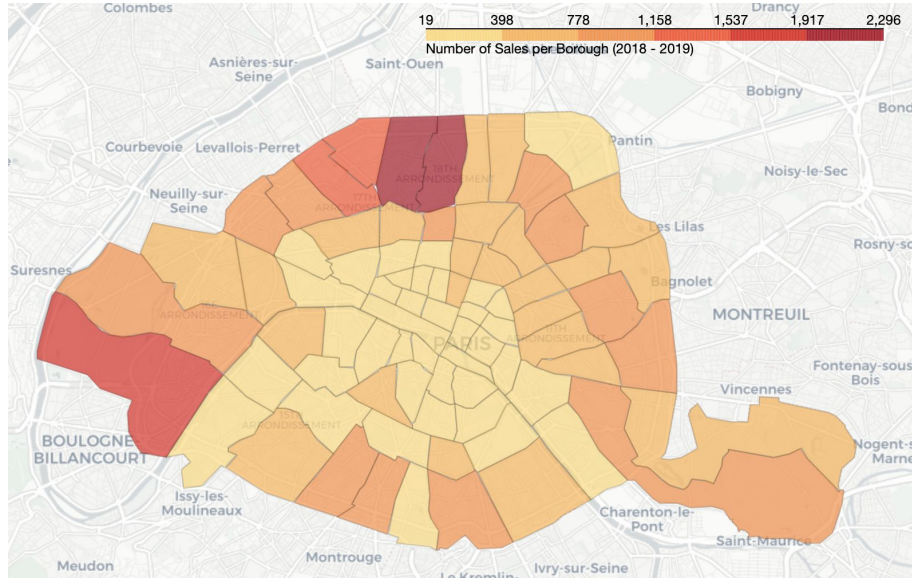
- 42,570 properties sold in Paris from January 2018 to June 2019
- Disparities between districts and boroughs: 563 sales in 1st district vs 5,509 sales in 18th district
- It will have an impact on property prices if less properties are available on the market
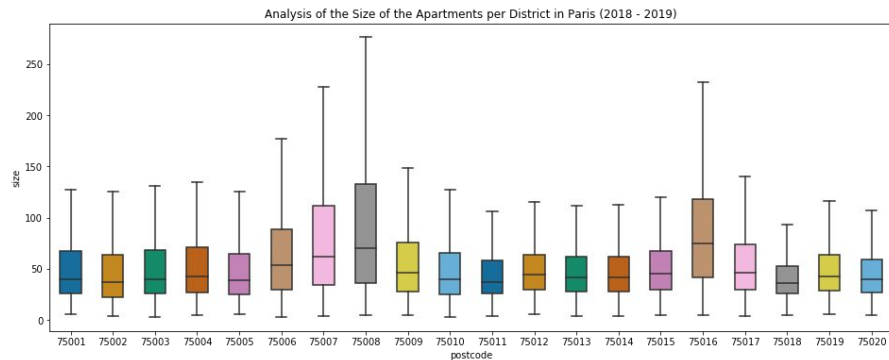

Number of House Sales per District in Paris (2018 - 2019)
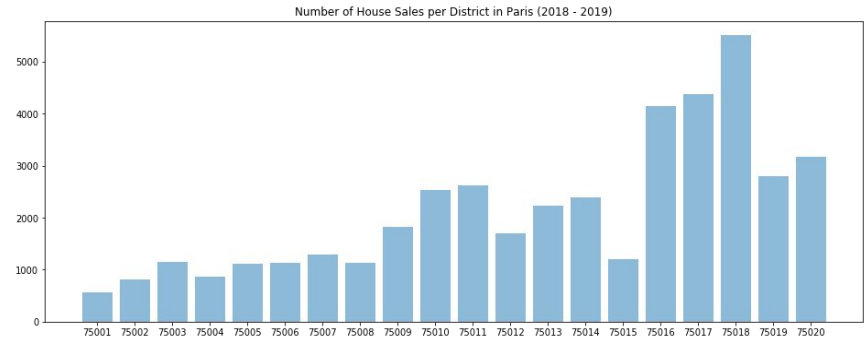
Results - Property Size
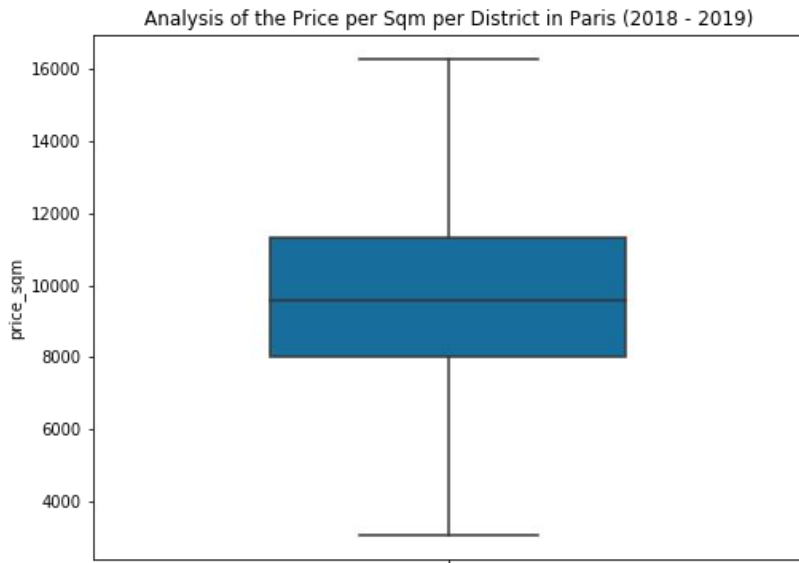

Median Property Size per Borough (2018 - 2019)

- Median size of apartments sold in Paris is 44 sqm
- Disparities between districts and boroughs:
  Largest apartments are in 8th and 16th districts
  and smallest ones are in 18th district
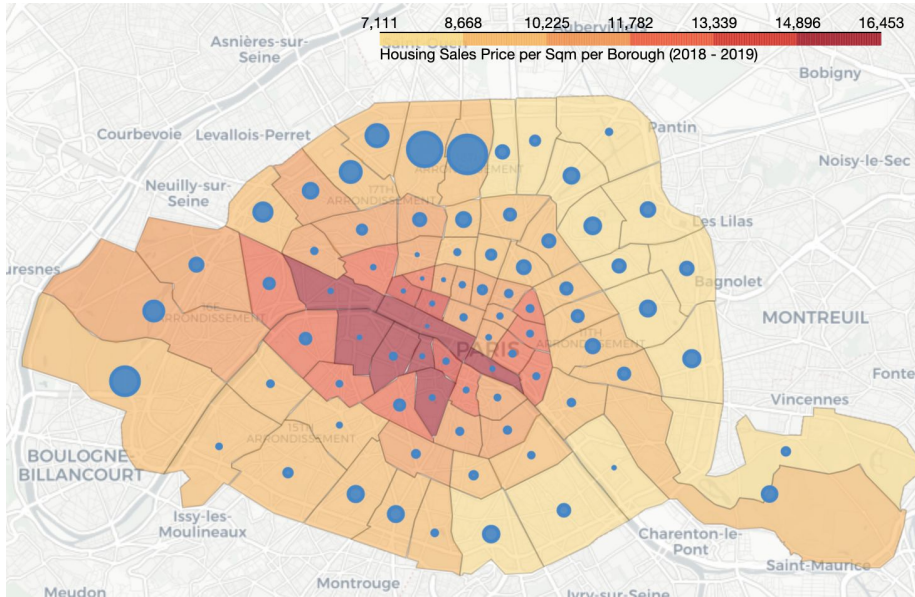
Outliers are not displayed on the chart below

Analysis of the Size of the Apartments per District in Paris (2018 - 2019)

Results - Property Prices 1/2

Analysis of the Price per Sqm per District in Paris (2018 - 2019)

- Mean price is 9,900 Euro per sqm

- Most properties are sold between 8,000 and 12,000 Euro per sqm

- Maximum price is 16,000 Euro per sqm

- Many factors can impact property prices: location, amenities, standing of the building, neighborhood, condition...
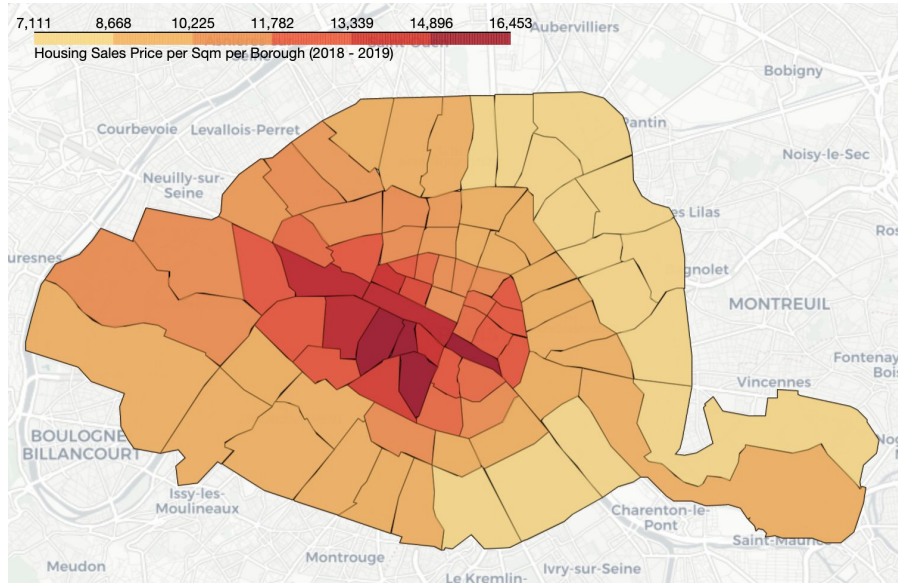
Outliers are not displayed on the chart

Results - Property Prices vs No of Sales

Housing Sales Price per Sqm per Borough (2018 - 2019)
7,111   8,668   10,225   11,782   13,339   14,896   16,453

- The number of sales usually has an impact on the price per sqm

- The lower the number of properties on the market, the more expensive the properties are

- There are exceptions on the East side of Paris so other factors impact the price of a property

Results - Property Prices 2/2

7,111   8,668   10,225   11,782   13,339   14,896   16,453

Housing Sales Price per Sqm per Borough (2018 - 2019)

- Mean price is 9,900 Euro per sqm

- Disparities between districts and boroughs: Most expensive areas are in Paris centre and cheapest areas are on the North-East, East and South-East of the city

Analysis of the Price per Sqm per District in Paris (2018 - 2019)
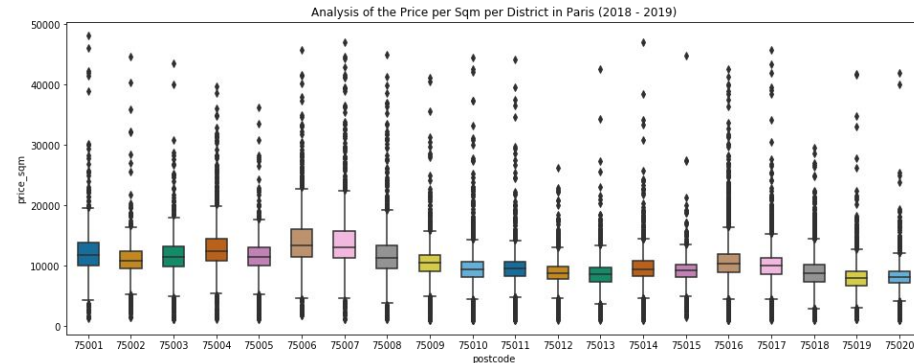
Results - Venues Analysis & Clustering

Venue Category	Venue
French Restaurant	704
Hotel	448
Italian Restaurant	207
Bar	169
Bakery	168
Café	159
Bistro	141
Coffee Shop	134
Japanese Restaurant	131
Plaza	117

- Use of Foursquare API to get top 100 venues per borough based on central point (except for 3 boroughs where central point was in parks)

- Top 10 venue categories in table

- 8 of the top 10 venues are related to food

- Use of K-mean clustering to categorise boroughs by type of venues

Results - Clusters Description

Best K to categorize the boroughs is 7

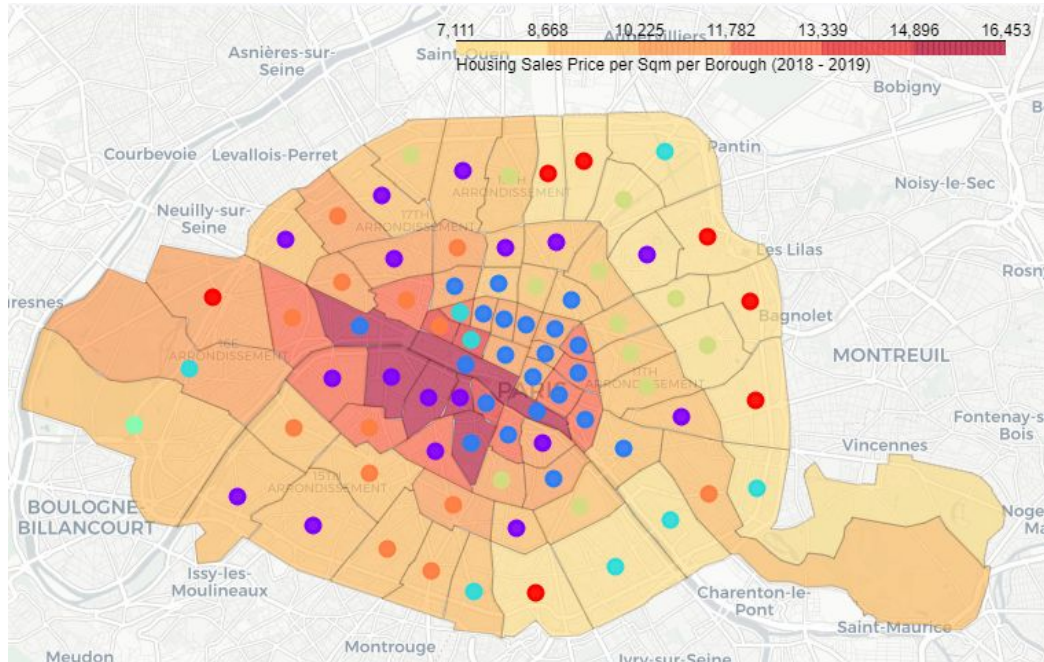| Cluster Id	Description	No of Borough	Color
0	Residential area with supermarkets, cafes, restaurants, fast foods and parks	7	🟥
1	Touristic area with many hotels, restaurants, places to go out, parks, museums and few shops	17	🟪
2	Shopping and touristic area with hotels, restaurants, many shops, museums and art galleries	22	🟦
3	Area with many hotels, international restaurants and theaters	8	🟦
4	Recreation area with venues for outdoor activities, sport and museums	1	🟩
5	Nightlife area with many bars, restaurants, bistros, cafes	12	🟨
6	Shopping areas with many shops, hotels, restaurants	13	🟧

Results - Clusters Mapping
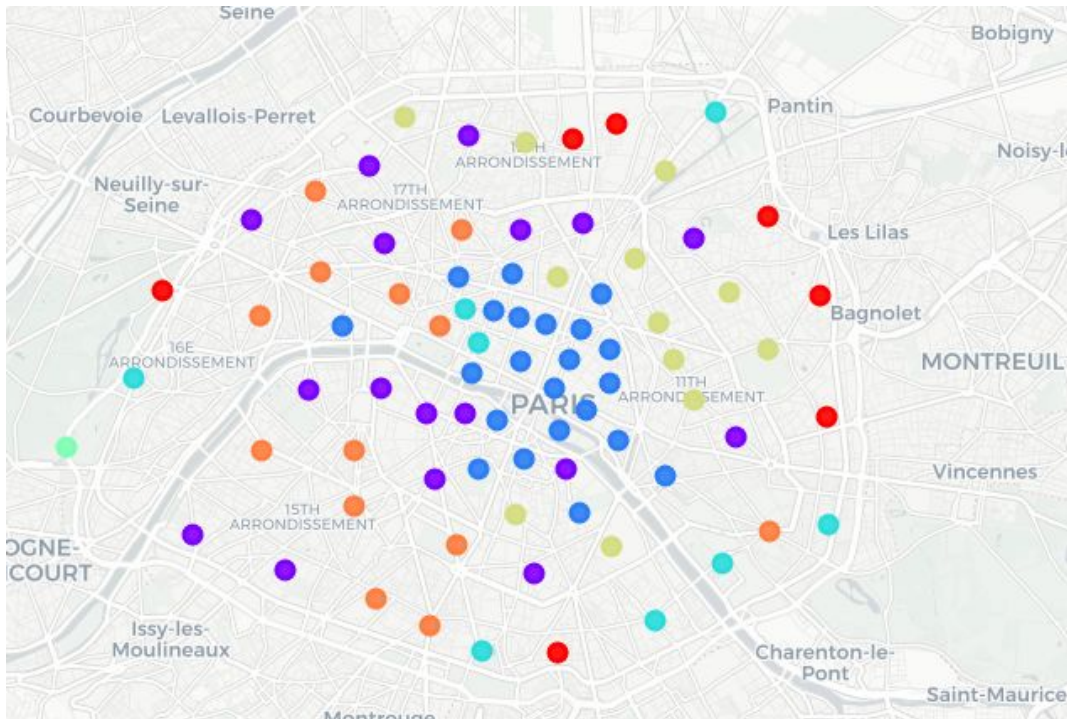
- Touristic areas are in the center of Paris

- Nightlife areas are on the East side

- Shopping areas are on the West side

- Residential areas are on the fringes

Cluster Id	Color
0	
1	
2	
3	
4	
5	
6	

Results - Venues vs Property Prices

Housing Sales Price per Sqm per Borough (2018 - 2019)
7,111   8,668   10,225   11,782   13,339   14,896   16,453

- Touristic areas are the most expensive ones

- Residential areas are the cheapest ones

Discussion & Conclusion

- The centre of the city is the most expensive because of touristic sites, famous restaurants and shops

- Neighborhoods in the North-East, East and South-East of Paris are the cheapest ones because they are more residential areas with many supermarkets, shops, parks...

- Average price per sqm is close to 10,000 Euro per sqm but it varies a lot (from 8,000 Euro to 12,000 Euro)

- Number of properties available on the market impacts the price per sqm

- Other factors could be analysed to understand how they impact the price of the properties like facilities, transportation, criminality, property features...

- Limitations due to data and hypothesis: incorrect records in dataset, central point of borough with 500m radius to look for venues, many restaurants in top 100 venues

Thanks

Refer to full report for more details

February 2020 - Sebastien Serville
IBM Applied Data Science Course