

Opening a new Luxury Wine Bar in London's ward.

The Battle of the Wards (Neighborhoods) in London

**Applied Data Science Capstone by IBM/Coursera -
Capstone Project**

◦

Sebastian Nastuta

1. Business Problem and Background

- The project will evaluate in which **wards** of London would be preferable to invest for opening a new **Wine Bar** or a chain of wine bars. The project will describe all analytical steps from data access, data analysis and visualisation to final recommendation of 5 to 10 wards in which is recommendable to open a new wine bar.
- **Business problem**
- Having a stable population of 8.6 million people and a daytime population of over 10 million people ([London Datastore](#)), London is the biggest city in Europe. London has a relatively young population, of people who mainly come here to study and work. London is extremely diverse from cultural and ethnic point of view has a strong hospitality sector consisting mainly in restaurant, hotels, pubs, bars and coffee shops. Especially pubs are extremely popular in London as they represent an important socialising place for Londoners and tourists. According to *2016 Pub Audit*, there were **3615** pubs in London a decline from 4835 pubs back in 2001 ([Estate agent London News](#) , 2018).

Socio-demographic profile of wine consumers

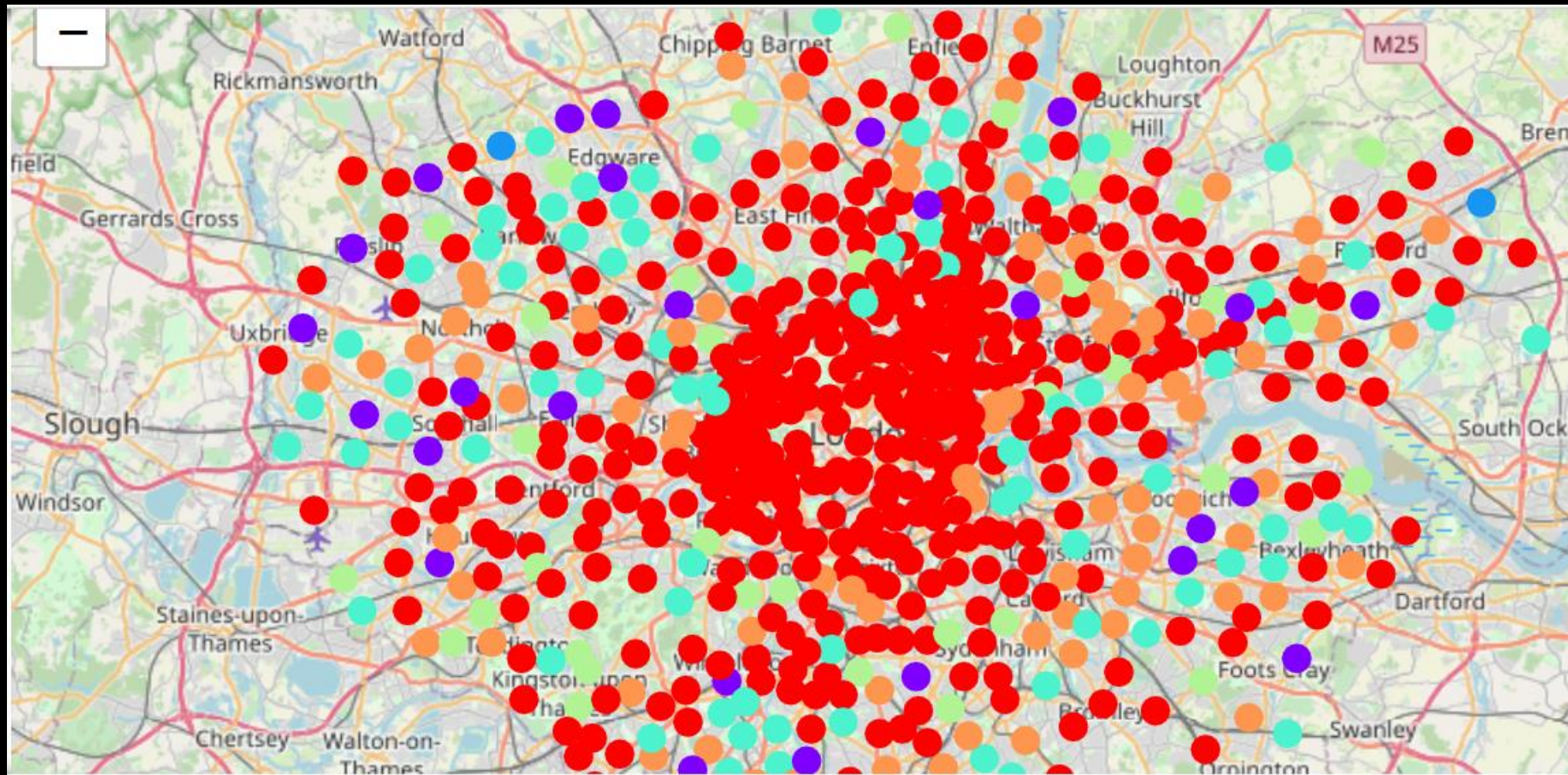
- According to a study proposed by [Duke University](#) (2014) wine consumers can overall be classified as **overwhelmed (23%)** by the available choices on store shelves, **Image Seekers (20%)** who use wine as a symbol of status, **traditionalists (16%)** who have clear established preferences, the **Savvy Shoppers (16%)** who "enjoy shopping for wine and discovering new varietal s on their own", **Satisfied Sippers (14%)** who simply like to drink and **wine enthusiasts (12%)** who "consider themselves knowledgeable about wine". Although this study was not performed in London, it is reasonable to assume the profiles can easily fit London wine market.

2. Methodology

- Several Python packages and statistical procedures will be involved in analysing the two datasets.
- Pandas **Crosstabulation** will be used to perform exploratory data analysis (**EDA**) on association between boroughs and cluster or between venues and clusters when the content of clusters will be described.
- Pandas package was used also for separating a string variable containing both Borough and Ward name into two separate columns, for excluding or keeping specific variables. Joining separate datasets was also performed using Pandas library.
- **K-means clustering** will be used to allocate each of London's wards into a cluster based on what venues they include,
- Matplotlib **histograms** and **scatter-plot graphs** will be used to plot the data and the association between demographic data,
- **Descriptive statistics** (means, etc) will be performed while grouping by boroughs to document the proposal we need to make regarding the wards where to open the wine bars. Following the 1st stage in which some candidate wards will be identified based on their componse in the 2nd stage of the analysis the socio-demographic profiles of the wards will be analysed.

4.2. Explore Wards in London through Venue Clustering

- The above created function was applied to repeat the same process to all the neighborhoods (**wards**) in London to identify all 12192 places of interest. While grouping the venues by **Venue Category** and sorting it descending by number it can be observed there are **841** Pubs, **710** Coffee shops, **630** Cafe and **438** Hotels. This indicate a strong demand for such leisure places which indicates potentially a solid investment opportunity. Interestingly, the histogram chart in this section indicates that out of 654 wards there are 400+ wards which have between 0 to 20 venues. This suggests these might be excluded from the list of potential investment places, reducing the number of neighborhoods of interest to around 200.



Map 2. Representation of all London Wards as Venue Clusters.

Description of the Cluster Segments

- **Cluster 0** can be labeled as **Gathering Places** as they include mainly pubs (18%), cafe-s (11%) and coffee shops (10%), hotels and indian restaurants.
 - Such wards are massively located in Central London in City of London and Westminster in which people come mainly for their work rather than living there.
 - Southwark, Hackney, Camden are strongly gentrified area in which younger people live and work or live closer to the corporate areas where they work. Hackney and Camden are strong touristic attractions also.
 - While Kensington and Chelsea is a more "posh" part of London having museums, restaurants and expensive residential area, Tower Hamlet is rather a place where people live and work.
- **Cluster 1** can be labelled as **Affluent Living Areas** cluster as it includes parks, yoga studios as 1st to 4th ranked venues. They are located at about 30-40 minutes from central London and they represent quiet areas where people mainly live.

Description of the Cluster Segments

- **Cluster 3** can be labelled as **Affordable Living Areas** cluster. This includes building sites probably for making new flats or office areas. These wards are in proximity with cluster 1.
 - These are in Croydon, Harrow, Hillingdon, Brent or Bexley.
- **Cluster 4** can be labelled as **Have a Pint on the Way to Home** as they include mainly pubs (60% as 1st venue). These are in zones 3-4 where people rather live than work and they presumably drink a pint of beer and meet some friends before going home.
 - These are spread in small numbers in most boroughs suggesting people gather in some wards before going home in the evening.
- **Cluster 5** can be labelled as **Shopping areas** as they include mainly grocery stores and supermarkets.

	C0	C1	C2	C3	C4	C5	All
3rd Most Common Venue							
All	408	28	2	82	45	84	649
Pub	31	3	0	4	6	7	51
Coffee Shop	36	1	0	1	2	3	43
Park	8	4	0	21	2	5	40
Café	29	0	0	3	2	1	35
Zoo Exhibit	14	8	0	1	3	3	29
Grocery Store	14	0	0	0	0	8	22
Hotel	13	1	0	1	1	2	18
Italian Restaurant	16	0	0	0	1	1	18
Indian Restaurant	11	0	0	1	0	5	17
Pizza Place	9	0	0	3	2	1	15

	C0	C1	C2	C3	C4	C5	All
2nd Most Common Venue							
All	408	28	2	82	45	84	649
Pub	58	0	0	1	11	7	77
Coffee Shop	38	0	0	0	1	2	41
Café	34	0	0	1	0	4	39
Grocery Store	17	2	0	2	1	15	37
Park	9	7	0	10	1	2	29
Zoo Exhibit	7	6	2	4	5	1	25
Indian Restaurant	18	1	0	3	0	2	24
Italian Restaurant	18	0	0	0	0	1	19
Supermarket	12	0	0	1	1	4	18
Hotel	16	0	0	0	0	2	18

	C0	C1	C2	C3	C4	C5	All
1st Most Common Venue							
All	408	28	2	82	45	84	649
Pub	71	1	0	1	27	0	100
Grocery Store	12	0	0	0	0	45	57
Café	51	0	0	0	0	0	51
Coffee Shop	38	0	0	0	0	1	39
Indian Restaurant	18	1	0	4	0	5	28
Hotel	26	0	0	1	0	0	27
Park	3	17	0	6	0	1	27
Construction & Landscaping	1	0	0	23	0	0	24
Convenience Store	15	0	0	4	2	3	24
Gym / Fitness Center	10	0	0	4	1	3	18

Fig 6. Average Percent Owners within Ward by Median Income

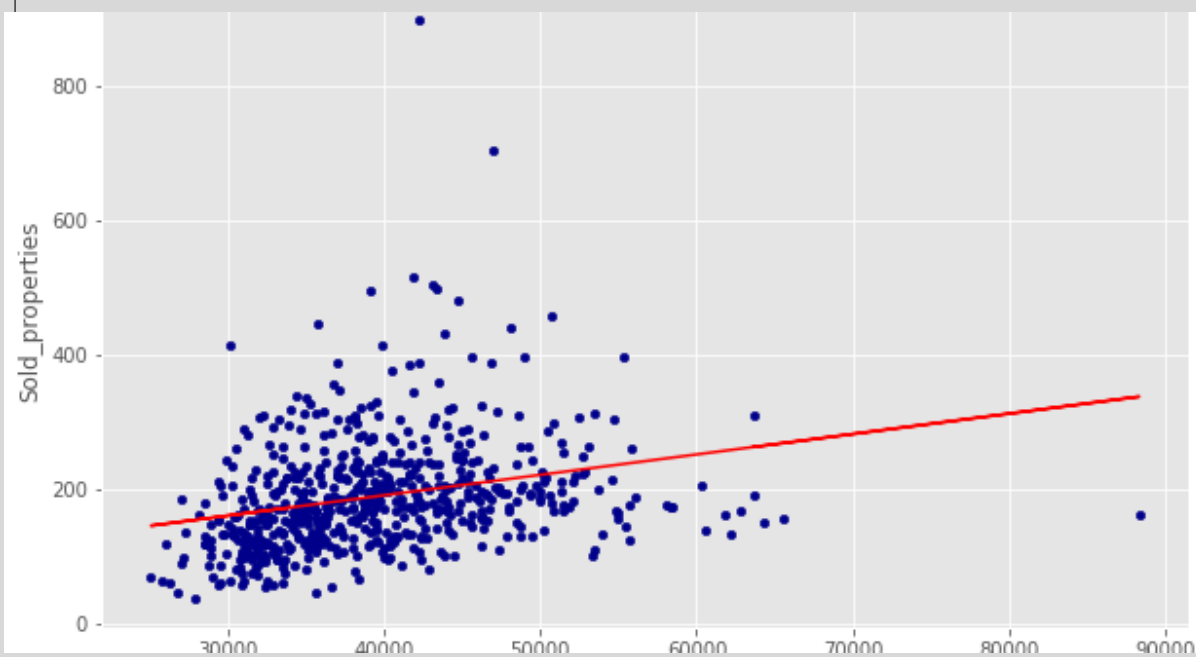
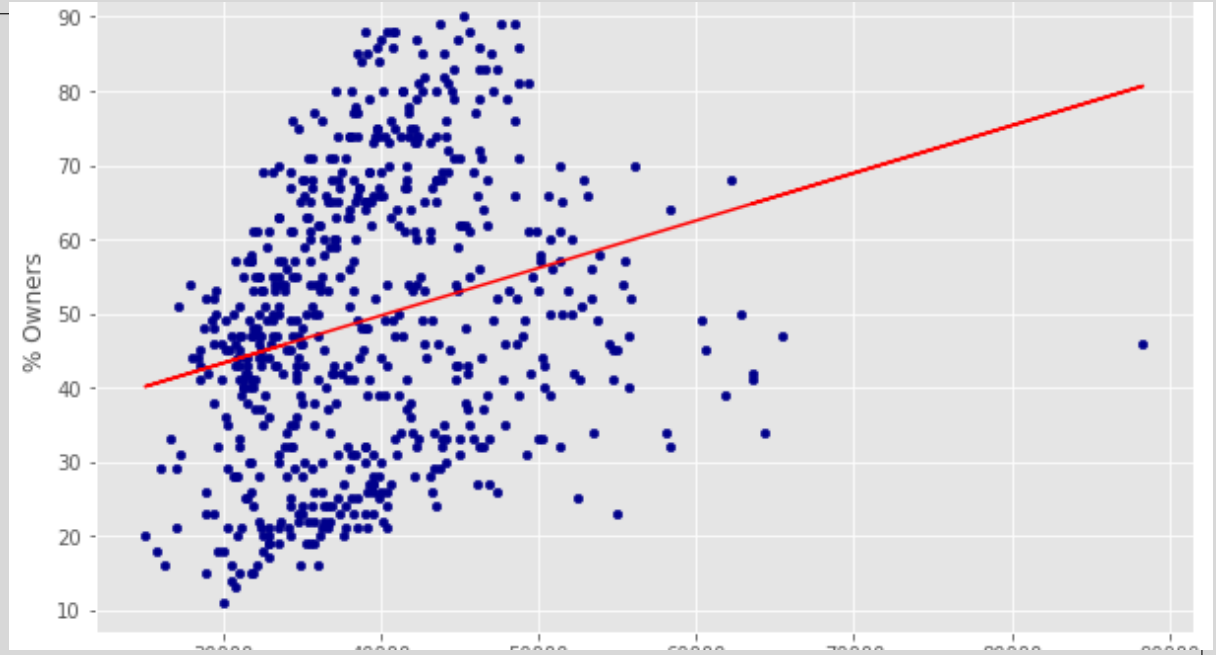
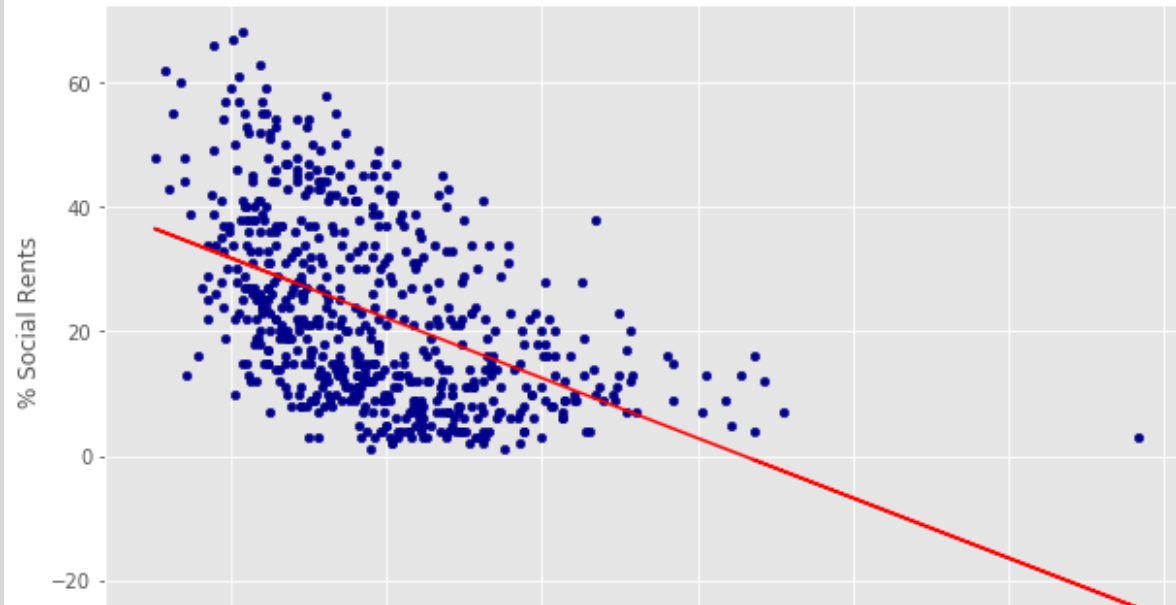


Fig 4. Population Size of Ward by Median Household Income

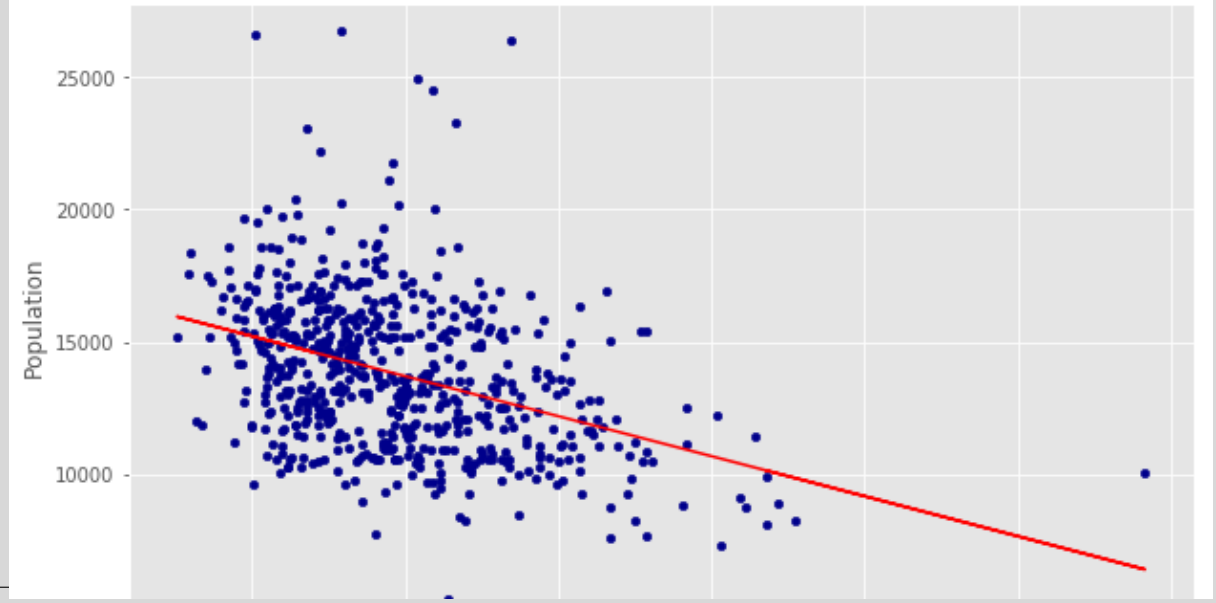


Figure 8. Average Percent Private Rents within Ward by Median Income



Fig 7. Average Percent Flats within Ward by Median Income

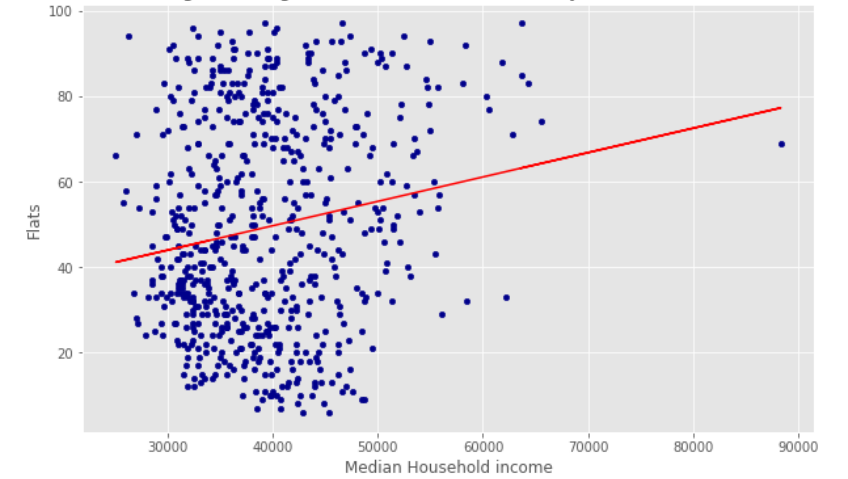
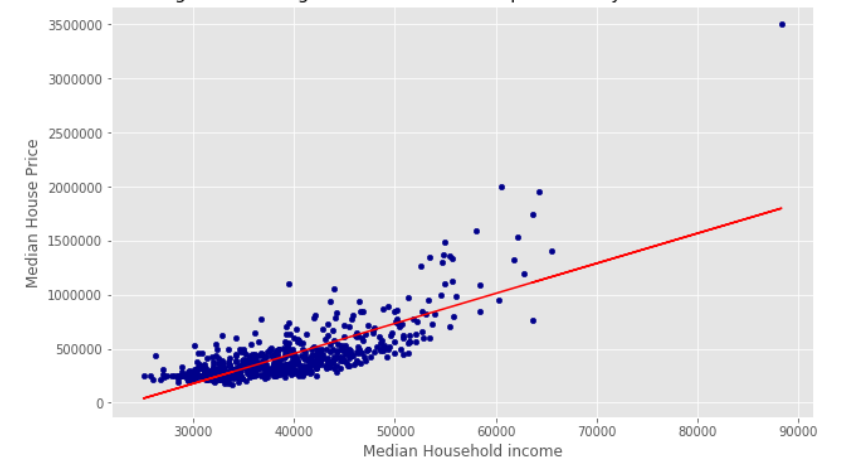
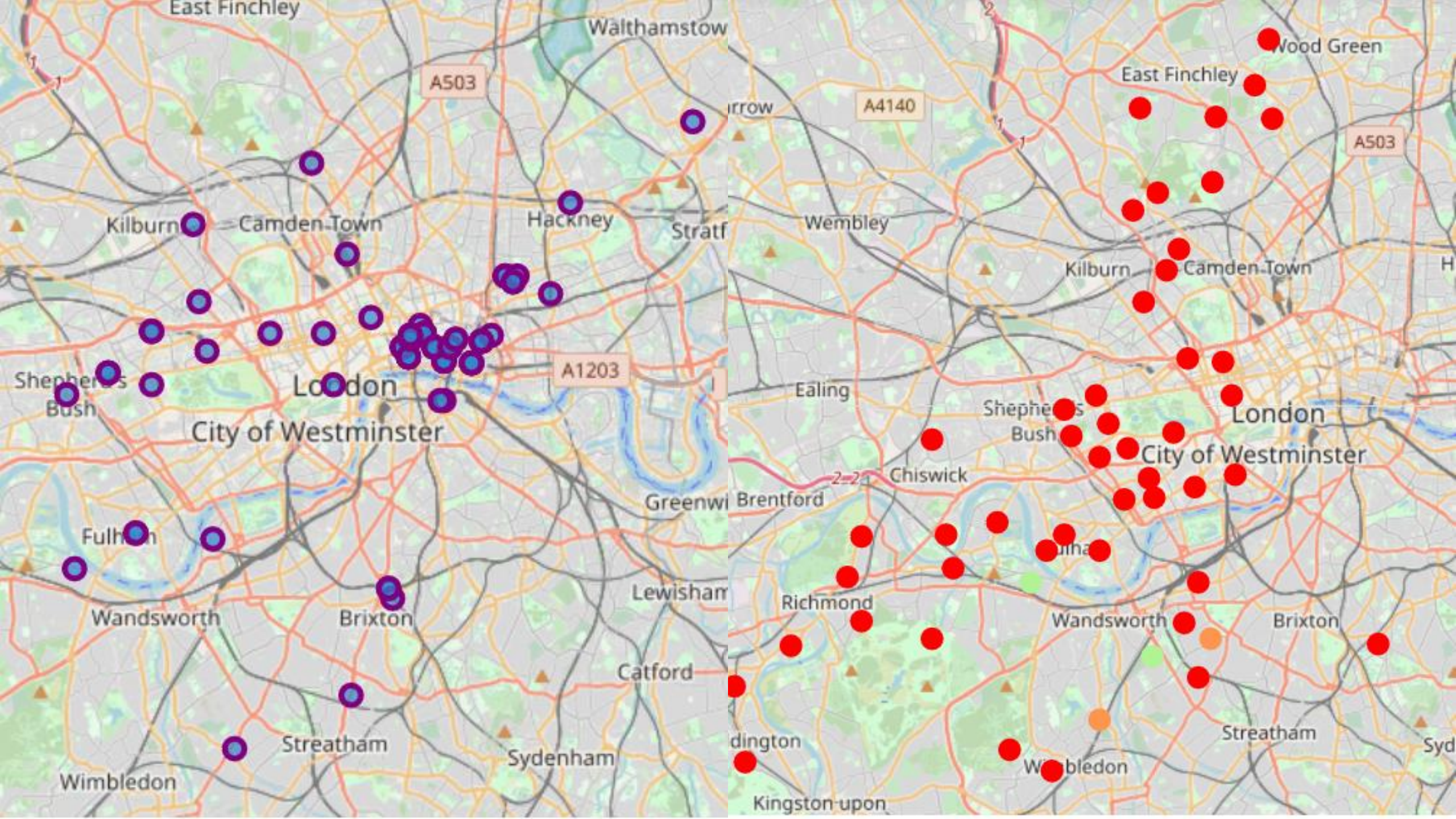


Figure 9. Average Median House Price per Ward by Median Income





Overall discussion of the results

- Once the 52 list of wards was selected, we created another map (**Map 4**) of the already existing Wine bars. Most of them are, interestingly, concentrated in City of London and otherwise scattered in different wards. Many of the wards we identified as potential location for luxury wine bars do not seem to have such a dedicated venue although they definitively have restaurants or other luxury places. The association table between Boroughs and clusters within the High-income dataset suggests the Boroughs in which the investment in some new luxury wine bars are Kensington and Chelsea, Richmond upon Thames, Wandsworth, Westminster, Camden, Haringey and Hammersmith and Fulham.

Recommendations

- Corroborating maps 3 and 4 the 1st venue where is feasible to open a luxury wine bar would be **Richmond upon Thames** which has 9 high income clusters but no dedicated wine bar.
- Although **Kensington and Chelsea** and **Westminster** have few wine pubs, the affluence of the area and the abundance of tourists allows the opening of another 1 or 2 wine bars. Wandsworth in the South part of London and Haringey in North are other two affluent areas which can support the opening of new luxury wine bars. Wards such as Alexandra (near Alexandra's Palace) or Hampstead Town in Haringey/Camden, West End in Westminster seem to be recommended for such investment as they attract not only tourists but they have affluent Londoners who might be targeted by such a new venue. Opening a wine bar in more isolated wards such as Teddington, Twickenham Riverside, Kew or Parsons Green and Walham would require to rely more on local people but seasonally on tourists also who might come in the area for tennis or rugby events.