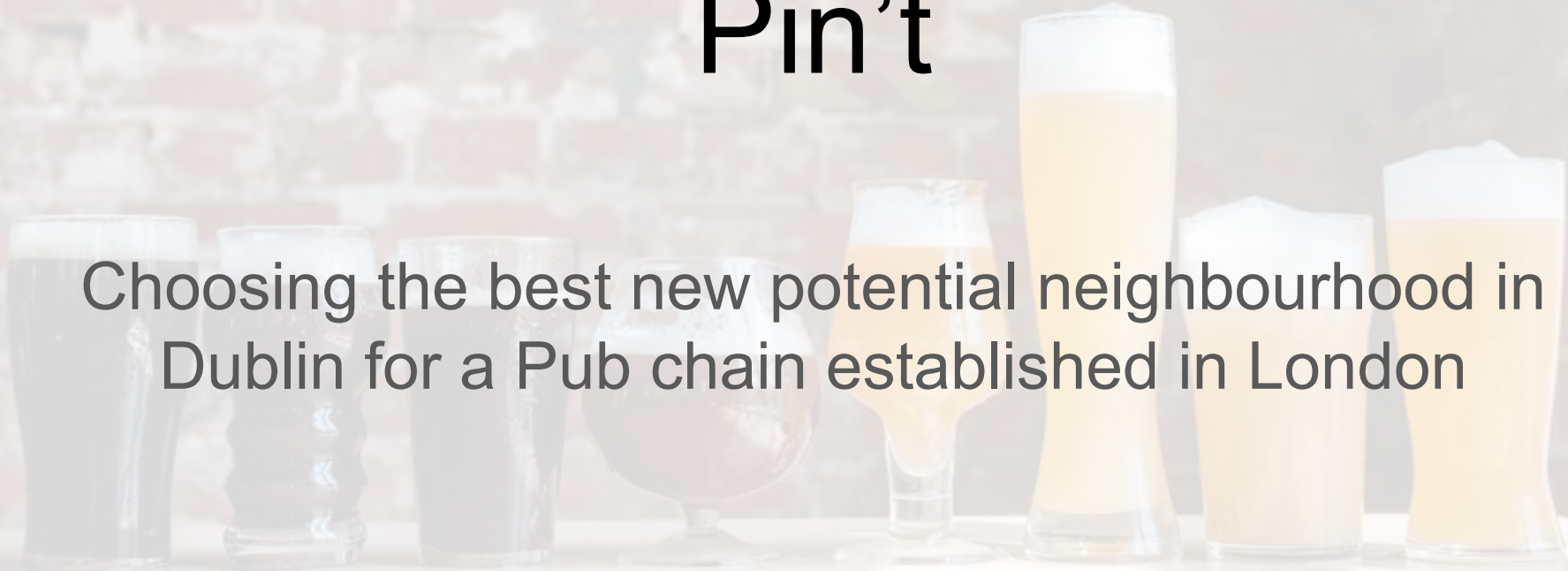


Pin't

Choosing the best new potential neighbourhood in Dublin for a Pub chain established in London



Predicting the right potential neighbourhood is important for the pub chain

- Taking the business to a new country for the first time is a huge step for the business. Making the right decisions will impact their success on the future openings in the rest of Ireland.
- The success of a pub is measured by the ROI (revenue on investment)
- Predicting the right neighbourhood based on established pubs ROI and locations will make other decisions easier. Also the pub chain will benefit from its past experiences in London

Data acquisition and cleaning

- Neighbourhoods' coordinates: I used python's geopy library to get the coordinates of each neighbourhood (existing and potential)
- Neighbourhoods' nearby venues: I used the Foursquare API to get maximum of 30 venues nearby each neighbourhood.
- The clustering of neighbourhoods was based on the frequency of categories of the venues nearby each neighbourhood.

Exploring distance from center of city as a variable

- The correlation between the existing pubs ROI was weak, so it was eliminated as a variable

```
2]:
```

	distance from central London	ROI
distance from central London	1.000000	0.045609
ROI	0.045609	1.000000

K-means clustering the pubs

```
2]: kclusters = 5

neighbourhoods_clustering = all_neighbourhoods.drop('Neighbourhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(neighbourhoods_clustering)
```

- Choosing the best cluster: I grouped the cluster by ROI mean

```
[218]: all_venues.groupby('Cluster Labels').mean()
```

```
[218]:
```

	ROI
Cluster Labels	
0	0.415000
1	0.475385
2	0.440000
3	0.610000
4	0.460000

K-means clustering the pubs

- Exploring the best cluster: I grouped the cluster by ROI mean

```
ROI    0.61
dtype: float64
```

24]:

	ROI	City	Neighbourhood
1	0.76	London	Barnet
10	0.78	London	Hackney
12	0.31	London	Haringey
25	0.59	London	Richmond upon Thames
32	NaN	Dublin	Ballsbridge and Donnybrook
33	NaN	Dublin	Christchurch
34	NaN	Dublin	Dalkey
35	NaN	Dublin	Drumcondra
36	NaN	Dublin	Malahide
37	NaN	Dublin	Ranelagh and Rathmines
38	NaN	Dublin	St. Stephen's Green

- The best cluster contains more than one neighbourhood from Dublin. I need to study the similarities between neighbourhoods in order to choose the closest one to Barnet's pub (highest ROI)

K-means clustering the pubs

- Exploring the best cluster:

	ROI	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
10	0.78	Hackney	Coffee Shop	Café	Pub	Park	Indie Movie Theater	Market	Canal Lock	Butcher	Flea Market
1	0.76	Barnet	Café	Park	Coffee Shop	Pub	Supermarket	Turkish Restaurant	Dessert Shop	Theater	Fish & Chips Shop
25	0.59	Richmond upon Thames	Park	Café	Garden	Coffee Shop	Hotel	Italian Restaurant	Bakery	Scenic Lookout	Pub
12	0.31	Haringey	Café	Mediterranean Restaurant	Turkish Restaurant	Park	Coffee Shop	Pizza Place	Gourmet Shop	Garden Center	Indie Movie Theater
32	NaN	Ballsbridge and Donnybrook	Café	Coffee Shop	Park	Hotel	Lounge	Concert Hall	Cocktail Bar	Outdoor Sculpture	Pizza Place
33	NaN	Christchurch	Café	Pub	Coffee Shop	Music Venue	Park	Cocktail Bar	Restaurant	Ice Cream Shop	Irish Pub
34	NaN	Dalkey	Beach	Coffee Shop	Scenic Lookout	Restaurant	Park	Pub	Café	Seafood Restaurant	Italian Restaurant
35	NaN	Drumcondra	Coffee Shop	Café	Pub	Clothing Store	Restaurant	Hotel	Donut Shop	Discount Store	Italian Restaurant
36	NaN	Malahide	Café	Italian Restaurant	Beach	American Restaurant	Pub	Gourmet Shop	Garden	Gym	Hotel
37	NaN	Ranelagh and Rathmines	Café	Park	Coffee Shop	Restaurant	Hotel	Burger Joint	Chinese Restaurant	Falafel Restaurant	Pub
38	NaN	St. Stephen's Green	Coffee Shop	Hotel	Park	Café	Burger Joint	Ice Cream Shop	Pub	Cheese Shop	Lounge

K-means clustering the pubs

- Making a decision:
-

notice that the 3 most common venues categories in the most successful pubs in London are: Coffee Shops, cafés ,pubs and parks

=> In order to make a decision I need to choose the neighbourhood which 3 most common venues categories are in the list

The potential neighbourhoods that check the criteria are:

- Ballsbridge and Donnybrook
- Christchurch
- Drumcondra
- Ranelagh and Rathmines

To refine my choices I noticed that **Ranelagh and Rathmines** is the most similar neighbourhood to Hackney where the most successful pub is established: with the same first 3 most common venues categories.

Conclusions

- I gathered useful data to help with the study
- I then built a model to predict the best potential neighbourhood in Dublin to open a new pub.
- Then I made a decision based on the mean ROI in clusters. And chose the most similar neighbourhood to the most successful neighbourhood in London.
- The study can include other features in the future: like proximity to other pubs (concurrent or from the same chain) and the age of inhabitants nearby pubs