

Capstone Project



Alexander Power

Introduction / Business Problem

There are **10,500** Fish & Chips shops in the UK, with an annual spend of **£1.2 billion**. **Bournemouth** was ranked as the **number one** most popular seaside resort of 2019.

We want to open our new Fish & Chips shop in Bournemouth to capitalise on its **popularity** with a food shop that is likely to be popular with the locals.

How someone can analyze the existing venues in a seaside town in England in order to decide on the **best place** in which to open a new Fish & Chips shop?



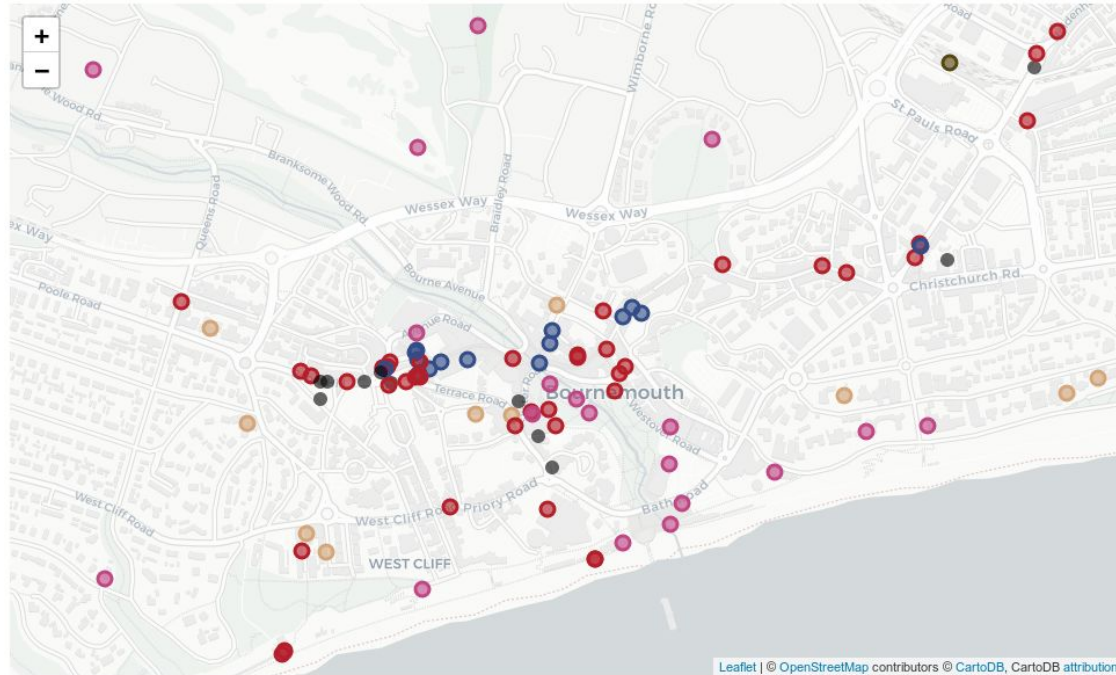
Data

We will use the **venue explore** API by **Foursquare** to gain insight into six types of venue in Bournemouth: Drink, Entertainment, Food, Hotel, Shopping, Transport.

We will use these groupings to understand how venues are dispersed by **business type**. Then, we will use **DBSCAN**, a data clustering algorithm, to cluster venues based on their **relative distance** from each other.

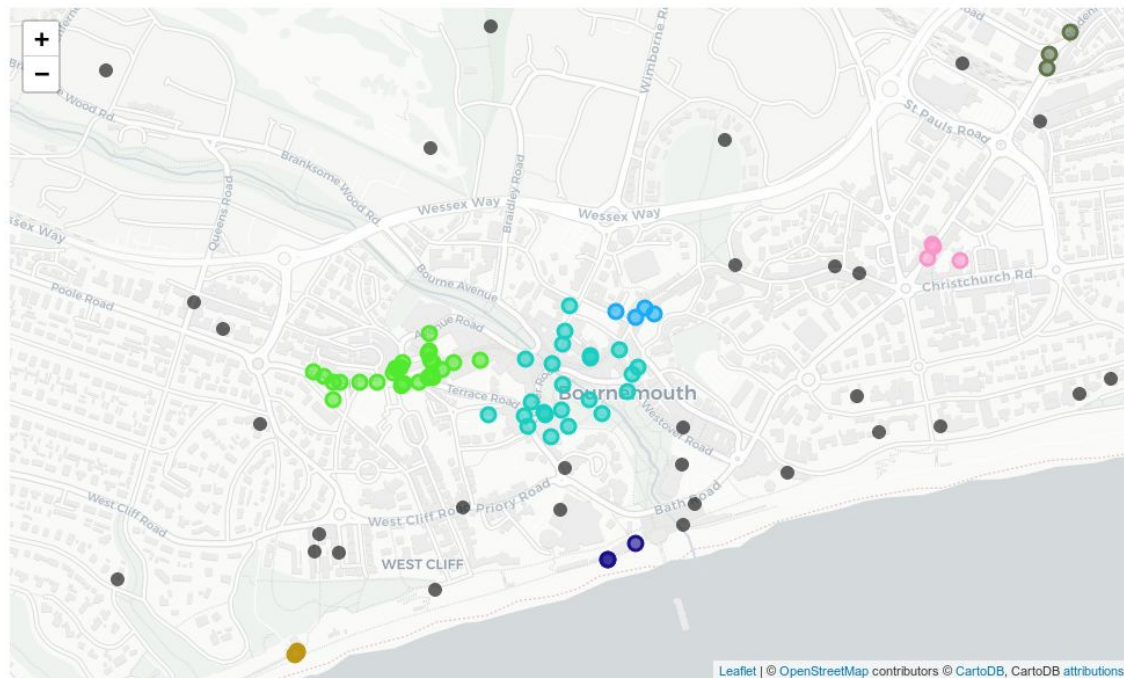
	Venue Name	Venue Category	Venue Latitude	Venue Longitude	Venue Group	Venue Cluster
0	South Coast Roast	Coffee Shop	50.720913	-1.879085	Shopping	0
1	DelMarco	Italian Restaurant	50.721370	-1.877221	Food	1
2	Lower Gardens	Park	50.719323	-1.878195	Entertainment	0
3	Bournemouth Gardens	Park	50.718990	-1.877733	Entertainment	0
4	Bournemouth Square	Plaza	50.720156	-1.879563	Shopping	0

Visualization: Groups



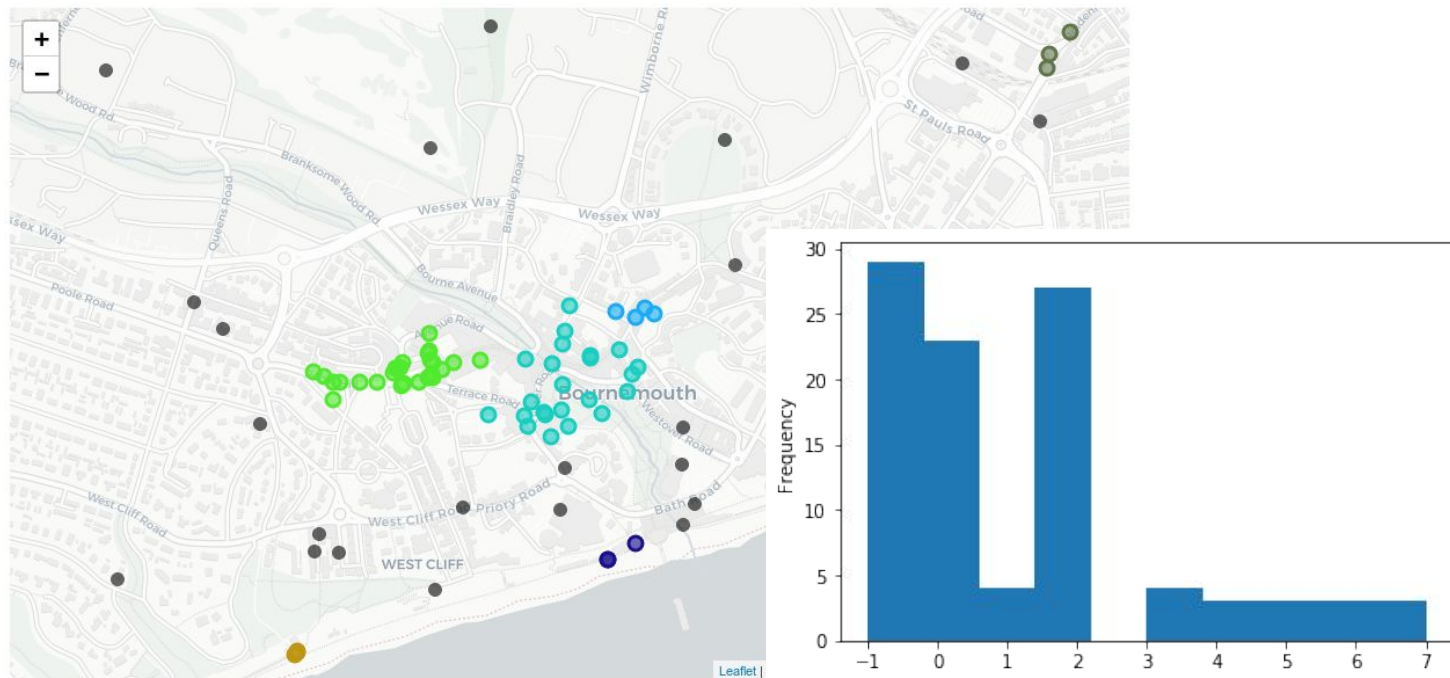
Venues, color-coded by group. **Entertainment** (pink) venues are closer to the beach, and most **Food** (red) and **Drink** (black) venues cluster in the center of town, along with all of the **Shopping** (dark blue) venues. **Hotels** (peach) are dispersed across town, and **Transport** (brown) is the furthest out of town, in the top-right of the map.

Visualization: Clusters



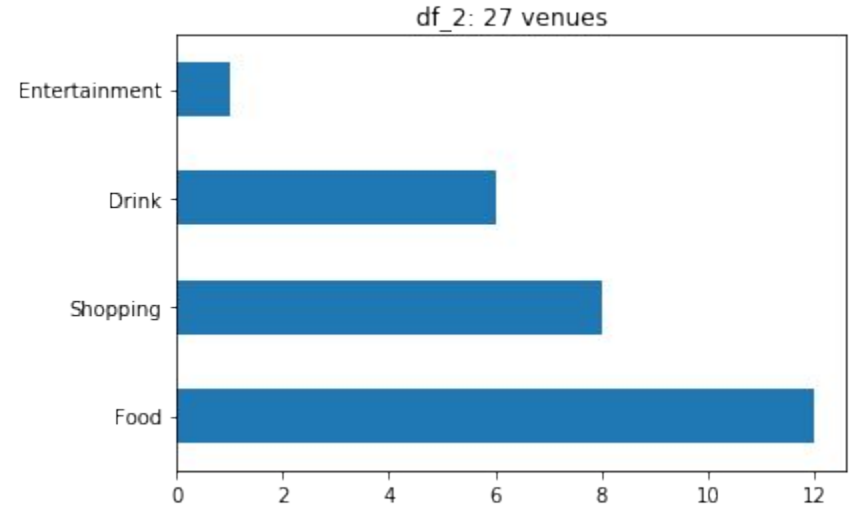
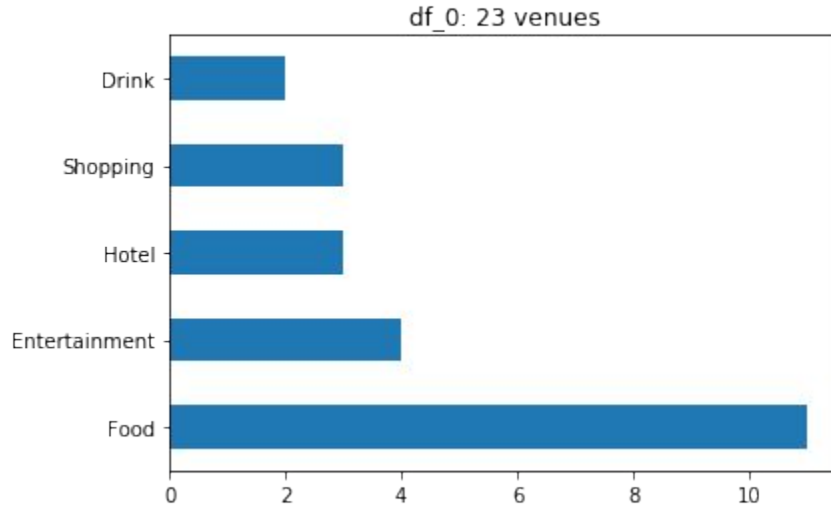
Venues, color-coded by their cluster category.

Visualization: Clusters



We see that, outside of the outlier class -1, the two most densely populated clusters are **0** (teal) and **2** (green), which are both in the center of town. We created data frames **df_0** and **df_2** containing data only from clusters 0 and 2, respectively.

Clusters



Cluster 0 has fewer **food** venues, which is desirable, but is only marginally lower than in cluster 2.

Cluster 2 has a more **non-food** venues than 0, which is more preferable. Areas with many places to **drink** are a hotspot for Fish & Chips food after a night of drinking, and people who do **Shopping** in the day time might like to eat Fish & Chips for lunch or dinner.

Conclusion

- We would want to open our Fish & Chips shop in **cluster 2**.
 - It is densely popular with venues that are not food related.
- Most other clusters were **too small**.
- Many venues were **outliers**.
 - Suboptimal locations because fewer customers would have been in these areas.

