

Data

We will be using foursquare API to obtain our data. We use the foursquare's explore API to find restaurant in the coordinates of Helsinki. We get the coordinates using geolocator. The data columns that we are interested are name, id, categories, latitude, longitude and postal code. Categories will help us to determine what type of restaurant we want to establish. Latitude, longitude and postal code tell us locations of the restaurants.

We will enrich the data with tip counts, likes, ratings and price tier. We will get this data using the foursquare's venue API. We query the API using ids of the venues that we obtained with the explore API. The price tier of the venues will help us to determine the what type of restaurant would be best suited to certain location. The other data we will use to categories the venues by their popularity.

[57]:

	name	categories	lat	lng	postalCode	Tip Count	Likes	Rating	Price tier
0	Café Trocadero	Mediterranean Restaurant	60.165103	24.942498	00120	33	91	9.0	2.0
1	Boulevard Social	Mediterranean Restaurant	60.166128	24.942034	00120	60	179	8.9	2.0
2	Gaijin	Asian Restaurant	60.165949	24.941759	00120	98	255	8.9	3.0
3	Patisserie Teemu & Markus	Bakery	60.167899	24.938190	00100	30	55	8.9	1.0
4	Passio	Scandinavian Restaurant	60.167088	24.937240	00100	27	92	8.9	2.0
5	Penny	American Restaurant	60.164708	24.941308	00120	22	89	8.9	2.0
6	Ravintola Vinkkeli	Scandinavian Restaurant	60.164107	24.946349	00130	13	43	9.1	2.0
7	Ragu	Restaurant	60.165831	24.944672	00130	31	111	8.7	2.0
8	Ravintola Muru	French Restaurant	60.165268	24.935957	00120	44	133	9.1	3.0
9	Zucchini	Vegetarian / Vegan Restaurant	60.166160	24.949551	00130	8	37	8.9	3.0

Picture 1. List of ten restaurants in our data

Picture 1 contains a snip of our dataframe. It includes ten first restaurants in our dataset. We have dropped all the Id -column. We only needed it to get more information about the venues. In addition, we have dropped all the restaurants that had NaN-values.