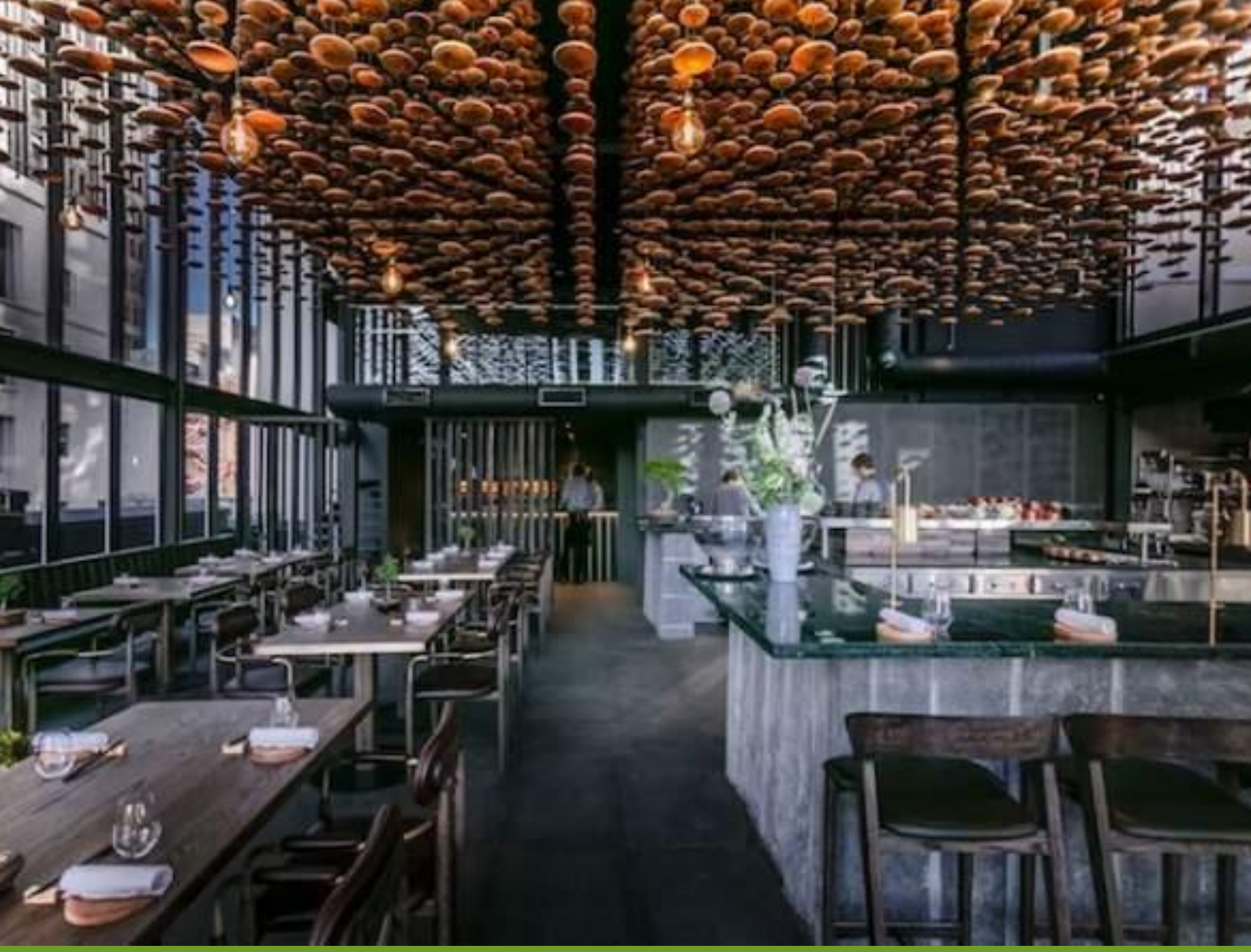

COURSERA
IBM APPLIED DATA SCIENCE
CAPSTONE PROJECT

The vegan/ vegetarian lifestyle

The data

Restaurants within a 50km radius from the center of Sandton and within the same radius from the center of Toronto will be downloaded from Foursquare.

The dataset will contain the names, addresses, city, postal code, state, country, latitude, longitude, distance from city center and restaurant category (e.g. Chinese, Indian, Italian, etc.). For a list of some of the categories of restaurants that can be obtained using Foursquare, please see Table 1.

Table 1 A list of some of the restaurants' categories and their ID codes

African Restaurant
4bf58dd8d48988d1c8941735
Ethiopian Restaurant
4bf58dd8d48988d10a941735
American Restaurant
4bf58dd8d48988d14e941735
New American Restaurant
4bf58dd8d48988d157941735
Asian Restaurant
4bf58dd8d48988d142941735
Chinese Restaurant
4bf58dd8d48988d145941735
Sushi Restaurant
4bf58dd8d48988d1d2941735
Thai Restaurant
4bf58dd8d48988d149941735
Fast Food Restaurant
4bf58dd8d48988d16e941735
Indian Restaurant
4bf58dd8d48988d10f941735
Italian Restaurant
4bf58dd8d48988d110941735
Mediterranean Restaurant
4bf58dd8d48988d1c0941735
Portuguese Restaurant
4def73e84765ae376e57713a
Vegetarian / Vegan Restaurant
4bf58dd8d48988d1d3941735

In Figure 1 an example of the data obtained from Foursquare is given.

	id	name	categories	referralId	hasPerk	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress
0	6050a765243d3a158f633049	Ocean Magic Seafood Restaurant	Asian Restaurant	v-1624434515	False	Cnr Rivonia Blvd, Shop GA3 Oriental City	9th Ave	-26.056591	28.060204	[{'label': 'display', 'lat': -26.056591, 'lng'...	1781	2191	ZA	Sandton	iGauteng	iNingizimu Afrika	[Cnr Rivonia Blvd, Shop GA3 Oriental City (9th...
1	4ea06d422c5b9f62fb3449dd	The Elm Tree Restaurant	Restaurant	v-1624434515	False	The Syrene Hotel	28 Wessel Rd, btwn 10th & 11th Ave	-26.052791	28.061692	[{'label': 'display', 'lat': -26.052790746599...	2199	2128	ZA	Rivonia	Johannesburg	iNingizimu Afrika	[The Syrene Hotel (28 Wessel Rd, btwn 10th & 1...
2	4cc95fa45c7e8cfa23ee4366	Thyme On Nicol Restaurant	Deli / Bodega	v-1624434515	False	William Nicol Dr	at Republic Rd	-26.089544	28.027193	[{'label': 'display', 'lat': -26.089544084013O...	3153	2196	ZA	Sandton	iGauteng	iNingizimu Afrika	[William Nicol Dr (at Republic Rd), Sandton, 2...
3	60b3b1efbc414b5b7f38bea7	Yuehua Restaurant	Chinese Restaurant	v-1624434515	False	Rivonia Central	NaN	-26.056688	28.060489	[{'label': 'display', 'lat': -26.056688, 'lng'...	1793	2191	ZA	Sandton	iGauteng	iNingizimu Afrika	[Rivonia Central, Sandton, 2191]
4	4bd9dac13904a59366e6439e	Maximillien Restaurant	French Restaurant	v-1624434515	False	Da Vinci Hotel & Suites, Maude St.	NaN	-26.107529	28.053211	[{'label': 'display', 'lat': -26.1075290072536...	4404	NaN	ZA	Sandton	iGauteng	iNingizimu Afrika	[Da Vinci Hotel & Suites, Maude St., Sandton]

Figure 1 Example of restaurant data obtained from Foursquare

Most of the columns contain text as values, with the exception of lat (latitude), long (longitude) and distance, that contain numerical values that calculations can be done on.

The latitudes and longitudes will be used to create maps of the two cities that indicate the different types of restaurants in each.

The data will be cleaned. Wherever missing values for latitude/ longitude or category appear, these observations will be removed. If venues “slip through” that are not restaurants, they will also be removed. It was for instance noted that a vegan pharmacy as well as a flower shop was downloaded. These are obviously not restaurants and will therefore be removed from the dataset.

4dc79925b0fb5556cd24803d	Veganview Dispensary	Pharmacy

Figure 2 Example of an item that has to be removed from the dataset

The categories will be used to group similar restaurants together so that a clear picture can be formed of how the two cities’ cuisines compare to one another.

	name	country	categories	lat	lng
0	Ocean Magic Seafood Restaurant	iNingizimu Afrika	Asian Restaurant	-26.056591	28.060204
1	The Elm Tree Restaurant	iNingizimu Afrika	Restaurant	-26.052791	28.061692
2	Thyme On Nicol Restaurant	iNingizimu Afrika	Deli / Bodega	-26.089544	28.027193
3	Yuehua Restaurant	iNingizimu Afrika	Chinese Restaurant	-26.056688	28.060489
4	Erawan Thai Restaurant and Bar	iNingizimu Afrika	Thai Restaurant	-26.058651	28.046673

Figure 3 Example of dataset where only the relevant columns were retained