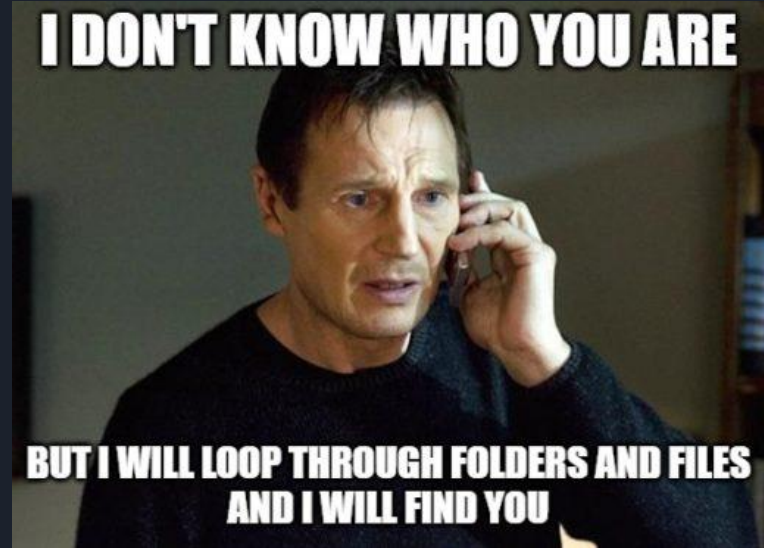# Mini-Project-2 (Calgary)

By: Sunny Bhandal

Introduction

Analyzing

YELP API
GOOGLE MAPS API

Structure

Purpose

Retrieve data and find insights by performing API calls

Methodology

Parse through and clean the data to obtain the necessary information

Results

The outcome of the analysis

Accuracy

How accurate are the results?

Purpose (YELP API)

- Build a database of bars, restaurants, and golf courses in the Calgary area (Pandas).
- Determine the most popular or highest rated venues on YELP.
- Apply the data in a meaningful way.

Methodology

- Set the correct parameters for the different criterias.

```python
# Define parameters for finding bars and look at top 50
params = {
    'term': 'bar',
    'limit': 50,
    'radius': 10000,
    'location': 'Calgary'
}
```

Methodology (Cont.)

- Setup a dataframe with the requested data.
- Variables chosen were name, rating, review count, and location.

Bar Data

	name	rating	review_count	location
0	Proof Cocktail Bar	4.5	61	1302 1st Street SW
1	Betty Lou's Library	4.5	76	908 17 Avenue SW
2	SHELTER	4.5	8	1210 1 Street SW
3	The Barn - Public House	5.0	5	1940 6th Avenue NW
4	One Night Stan's	4.0	9	520 17 Avenue SW

Results

- Combined the dataframes (bar (n=50), restaurant (n=50), and golf (n=15)) to obtain a final dataframe. The data may now be sorted based on rating or review_count.

	name	rating	review_count	location
3	The Barn - Public House	5.0	5	1940 6th Avenue NW
24	Brüch Bar	5.0	1	200 Barclay Parade SW
39	Yardarm	5.0	2	107-535 8th Ave SE
56	Alumni Sandwich & Liquor Bar	5.0	55	725 17 Ave SW
60	Nights & Weekends	5.0	7	821 1 Street SW
74	The Eden	5.0	21	1219 9 Avenue SE
100	Kyle Miller Golf	5.0	1	3434 48 Avenue NE
106	Screen Golf	5.0	1	122 5255 McCall Way NE
108	Metro Fairway Indoor Golfing	5.0	4	5220 4th Street NE
111	Willow Park Golf & Country Club	5.0	2	639 Willow Park Drive SE

Applications

Problem: A group of friends can't decide if they should go watch the game/get drinks, go try some high-end cuisine, or go for a round of golf.

Solution: Filter the data set by rating to obtain the top 10 highest rated venues. Run a random number generator from 1-10 to select the venue (it may be a bar, restaurant, or golf course.)

Accuracy

- Most companies that allow users to rate venues don't provide a specific criteria so ratings can be quite biased.
- Someone that had an average experience, may be less likely to leave a rating compared to someone who had a bad experience.
- Remove venues with a low review count (n < 5)
- Both points can **skew** the results left or right.

GOOGLE API

Travelling Salesman Problem

"Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?"

This is an example of an optimization problem. Algorithms are used to minimize the time of travel.

Methodology

- Modify the golf data to include a column for complete address (address, city, country)

- Sort the data by review count to find the most popular golf courses in the city.

- Use this more precise dataset to make API calls for accurate JSON output.

- Variables passed to the model are origins, destinations, and units (metric).

- Parsed the data to obtain a distance vector for all 10 venues.

- Functions: storing data, route modelling, returning node distance, and print solutions

Input

10 Most Popular Golf Courses (YELP)

	name	rating	review_count	address	city	country	location
1	Silverwing Golf Course	3.0	10	3434 48 Avenue NE	Calgary	CA	3434 48 Avenue NE,Calgary,CA
3	Confederation Park Golf Course	3.5	8	3204 Collingwood Drive NW	Calgary	CA	3204 Collingwood Drive NW,Calgary,CA
12	Inglewood Golf & Curling Club	2.5	7	19 Gosling Way SE	Calgary	CA	19 Gosling Way SE,Calgary,CA
2	McCall Lake Golf Course	4.5	5	1600 32 Ave NE	Calgary	CA	1600 32 Ave NE,Calgary,CA
7	National Golf Academy	4.0	4	50 50th Aveune SW	Calgary	CA	50 50th Aveune SW,Calgary,CA
8	Metro Fairway Indoor Golfing	5.0	4	5220 4th Street NE	Calgary	CA	5220 4th Street NE,Calgary,CA
4	Lakeview Golf Course	3.5	3	5840 19 Street SW	Calgary	CA	5840 19 Street SW,Calgary,CA
10	Golfuture YYC	4.5	3	7100 15 Street SE	Calgary	CA	7100 15 Street SE,Calgary,CA
5	Fox Hollow Golf Course	2.0	2	999 32 Ave NE	Calgary	CA	999 32 Ave NE,Calgary,CA
11	Willow Park Golf & Country Club	5.0	2	639 Willow Park Drive SE	Calgary	CA	639 Willow Park Drive SE,Calgary,CA

Results (Cont.)

Silverwing -> McCall -> Fox Hollow -> Inglewood -> Willow Park -> Golfuture -> National Golf -> Lakeview -> Confederation -> Metro Fairway -> Silverwing

Distance: 65 km

Time of Day: July 7th, 12:40 PM

```
Total Distance: 65 km
Route for the vehicle:
 0 -> 3 -> 8 -> 2 -> 9 -> 7 -> 4 -> 6 -> 1 -> 5 -> 0
```

Applications of TSP

Problem: How can a company save money on delivery costs (EX: gas or wages)?

Solution: Use a model of TSP to calculate the most efficient route each delivery driver would take.

- Potential Parameters: time of day, distance between houses, number of construction zones, number of traffic lights, and the number of LEFT TURNS.

Fun Fact: You save more gas by prioritizing right turns because of the time avoided idling during a typical left turn.

UPS found that by avoiding left turns on all the routes its trucks take, it was able to save 98 million minutes of idle time per year.

Model Accuracy

- Model doesn't take into account the time of day as the output is based on when the data was retrieved.
- Model can be modified to avoid tolls, change mode of travel, access traffic, etc.

Favorite API

- Google!!! Clean outputs and interesting functionality.
- Yelp was very user friendly.

End

References

- https://memegenerator.net/instance/69008866/crying-peter-parker-made-an-api-call-got-an-xml-response
- https://simpleexcelvba.com/programming-theme-memes-funny/
- https://en.wikipedia.org/wiki/Travelling_salesman_problem
- The History Behind the Google Logo I Express Writers
- https://developers.google.com/optimization/install
- https://developers.google.com/maps/documentation/directions/start
- https://www.yelp.com/developers/documentation/v3/business_search