SI 206 Final Project

By: Lauren Watsky, Katie Wolberg and Jamie Zuckerman

Original Goals:

———

To test whether the top 20 restaurants in 5 different college towns are rated similarly across three different rating platforms.

Achieved Goals:

We found that out of the 250+ restaurants that we collected, only 32 restaurants were in the top 20 on two different rating sites and only two restaurants were on all three of the rating sites.

APIs:

We used three APIs: Yelp, Google Places and Zomato. From each API we found the top 20 restaurants in Ann Arbor(University of Michigan), Austin (UT Austin), Bloomington (Indiana University), Gainesville (University of Florida) and Madison (University of Wisconsin).

Problems We Faced:

———

We originally wanted to use big cities, like New York and Chicago. However, since the cities are densely populated and filled with thousands of restaurants, there were barely any restaurants that were found in the top 20 list between the three APIs that we chose. We decided to switch to smaller cities where there are less restaurants in hopes for more crossover.

Limiting to 20 Responses

———

We used user-input to limit to 20 responses at a time. When a user would type in a city, the 20 top-rated restaurants for that specific city would be added to the database.

```python
city = input("Please enter the city you would like to add data for: ")
city = str(city)
```

```python
for i in range(20):
    _rest_name = rest_data[i]
    _rest_id = rest_id[i]
```

Tables Explanation:

We extracted data such as restaurant name, restaurant rating, and the city of each restaurant from each of the APIs.

We created a city id for each city, and used this as our shared key between tables.

We used the join method to create tables of the restaurants that crossed over between two of the APIs.

Table Pictures

___

	city_id	city_name
	Filter	Filter
1	1	Ann Arbor
2	2	Gainesville
3	3	Austin
4	4	Bloomington
5	5	Madison

Name

▼ ▦ Tables (7)
- ▶ ▦ CityIdConversion
- ▶ ▦ GoogleData
- ▶ ▦ GoogleZomato
- ▶ ▦ YelpData
- ▶ ▦ YelpGoogle
- ▶ ▦ YelpZomato
- ▶ ▦ ZomatoData

Table: GoogleData

	restaurant_name	restaurant_id	restaurant_rating	city_id
	Filter	Filter	Filter	Filter
11	Tomukun Korea...	3d549a4757f29...	4.5	1
12	Gandy Dancer	20063d1acd94...	4.5	1
13	Paesano Restaur...	34f2a6766b7a8...	4.5	1
14	Everest Sherpa ...	ae088c70276cb...	4.5	1
15	Aventura	786e2706a713...	4.5	1
16	Palm Palace	942a68390e39...	4.5	1
17	Black Pearl | An...	5c88860b929d...	4.5	1
18	Kang's Korean R...	abc8aea324095...	4.5	1
19	Gratzi Ann Arbor	b0dd95bde51fa...	4.2	1
20	Asian Legend	1460ef43ff0be...	3.9	1
21	Daybreak Pleasa...	e5373023dfb01...	4.8	2
22	WayWard Fox	f0d82e0bb97c6...	4.7	2
23	La Cocina de Ab...	089ec464cd52b...	4.7	2
24	Daily Green	fee4a0d6b4049...	4.7	2
25	Alpin Bistro	25c0ado34408f	4.7	2

Table: GoogleZomato

	restaurant_name	google_rating	zomato_rating	super_rating	city_id
	Filter	Filter	Filter	Filter	Filter
1	Pacific Rim By K...	4.5	4.4	8.9	1
2	Palm Palace	4.5	4.4	8.9	1
3	Satchel's Pizza	4.7	4.8	9.5	2
4	Uchi	4.7	4.6	9.3	3
5	Franklin Barbecue	4.7	4.7	9.4	3
6	Uptown Cafe	4.6	4.6	9.2	4
7	The Irish Lion R...	4.5	4.5	9.0	4
8	Mickies Dairy Bar	4.7	4.9	9.6	5
9	Weary Traveler ...	4.6	4.7	9.3	5
10	The Old Fashion...	4.5	4.8	9.3	5

Calculations:

—————

For the restaurants that were on two out the three sites, we created a "Super Rating." We added the ratings from both of the APIs to show an overall rating out of 10.

If a restaurant was on all three sites, we created a "Max Rating," out of 15.

```
for row in data:
    restname = row[0]
    googlerate = row[1]
    zomatorate = row[2]
    superrate = googlerate + zomatorate
    cityid = row[3]
    cur.execute("INSERT INTO GoogleZomato (restaurant_name, google_rating, zomato_rating, super_rating, city_id) VALUES (?, ?, ?, ?, ?)", (restname, googler
```

```
super_ratings = []
cityids = []
for row in data:
    names.append(row[0])
    super_ratings.append(row[1]+row[3])
    cityids.append(row[2])
```

** this max rating can
be seen in our
visualization

|	Table:	GoogleZomato				Ne
		restaurant_name	google_rating	zomato_rating	super_rating	city_id
		Filter	Filter	Filter	Filter	Filter
1		Pacific Rim By K...	4.5	4.4	8.9	1
2		Palm Palace	4.5	4.4	8.9	1
3		Satchel's Pizza	4.7	4.8	9.5	2
4		Uchi	4.7	4.6	9.3	3
5		Franklin Barbecue	4.7	4.7	9.4	3
6		Uptown Cafe	4.6	4.6	9.2	4
7		The Irish Lion R...	4.5	4.5	9.0	4
8		Mickies Dairy Bar	4.7	4.9	9.6	5
9		Weary Traveler ...	4.6	4.7	9.3	5
10		The Old Fashion...	4.5	4.8	9.3	5

Visualizations:

———

Yelp and Google
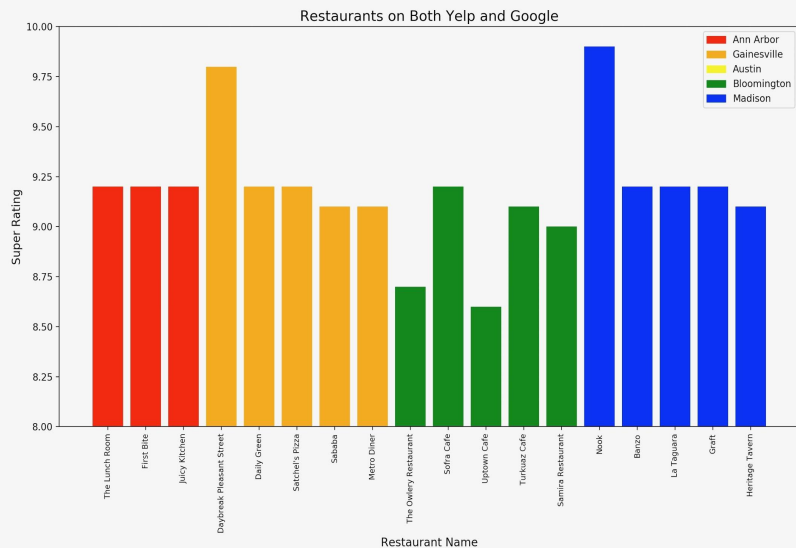
"Nook" in Madison has the highest rating out of all of the restaurants in this crossover.

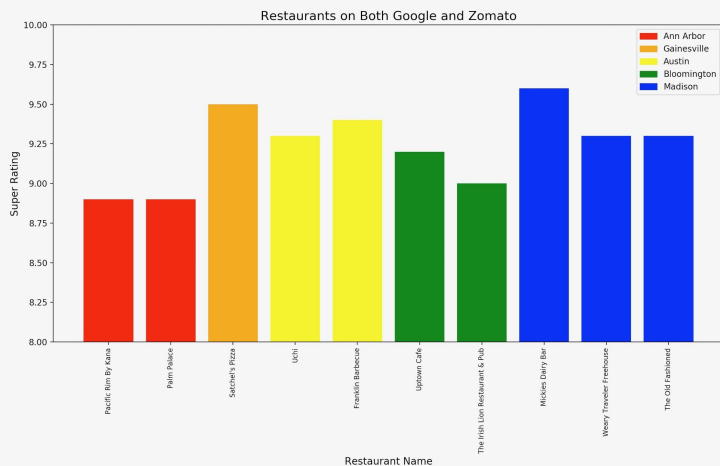## Yelp and Zomato
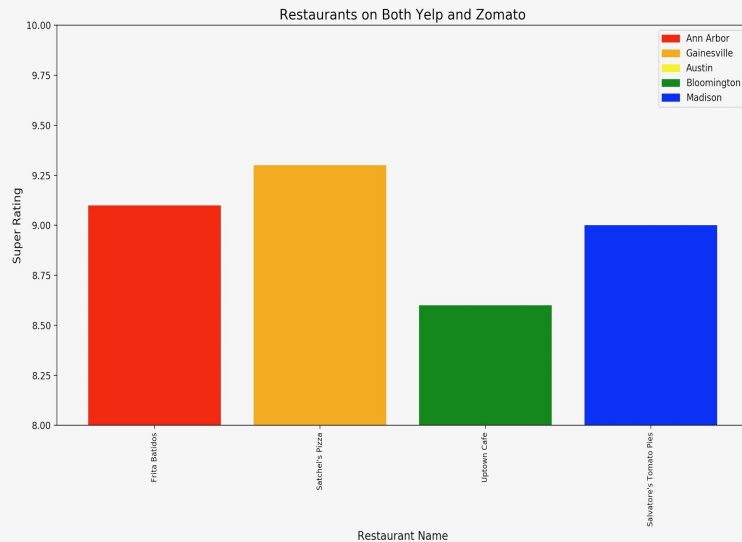
"Satchel's Pizza" in Gainesville has the highest rating out of all of the restaurants in this crossover.
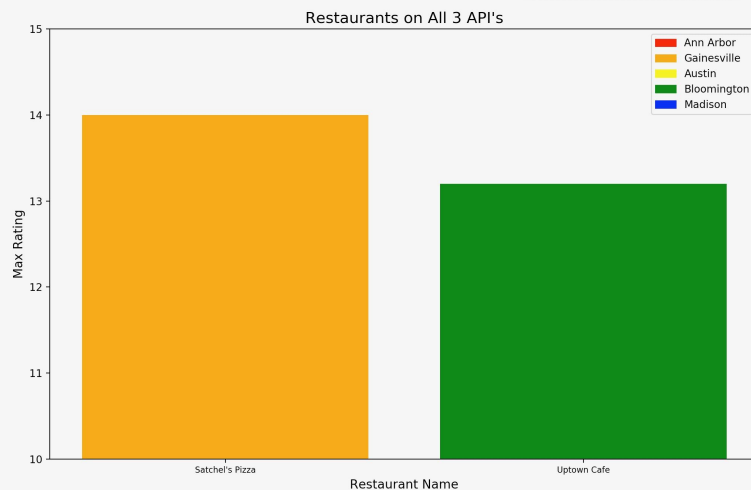
Visualizations Continued:

Google and Zomato

"Mickie's Dairy Bar" in Madison has the highest rating out of all of the restaurants in this crossover.

Restaurants on All Three API's

Both "Satchel's Pizza" and "Uptown Café" were on all three API's. "Satchel's Pizza" had the highest Max Rating.

Thank You!

Any Questions?

Github link:
https://github.com/lwatsky/final-project.git