



---

# YELP DATASET CHALLENGE

---

Group :18

Hanifah Othmal, Meghana Vijaykumar Rai, Monisha Thuppada Prakash



MAY 9, 2017

UNIVERSITY OF TEXAS, ARLINGTON

# Contents

Introduction .....	2
About the Dataset.....	3
Business Dataset .....	3
Review Dataset .....	4
User Dataset .....	4
Check-In Dataset.....	5
Tips Dataset .....	6
Data Processing and Filtering .....	7
Sqlite.....	7
TextBlob .....	7
Matplotlib .....	7
WordCloud.....	7
Our Analysis.....	8
What are the popular times and days to eat out in different countries? .....	8
What cuisines do Yelpers rave about in different countries? Does it differ? .....	12
Which countries have the harshest and/or nicest reviewers? .....	15
Top 25 Restaurants & Top 10 Cities .....	17
Limitations and Future Implications.....	22

## Introduction

Yelp.com is a website created and marketed by an American multinational corporation, Yelp, headquartered in San Francisco, California. Yelp publishes customer reviews of local businesses, and it also provides services such as Yelp Reservations and the online food delivery service known as Eat24. Additionally, it provides data about businesses, helps train small businesses, and hosts social events.

Our project uses a series of datasets which includes information from cities in select countries, and we will be using this data to analyze various cultural and seasonal trends. First off, according to Yelp, the dataset encompasses the following areas and includes the following information:

<ul style="list-style-type: none"><li>• 4.1M reviews and 947K tips by 1M users for 144K businesses</li><li>• 1.1M business attributes, e.g., hours, parking availability, ambience.</li><li>• Aggregated check-ins over time for each of the 125K businesses</li><li>• 200,000 pictures from the included businesses</li></ul>	<p>Countries</p> <ul style="list-style-type: none"><li>• U.K</li><li>• Germany</li><li>• Canada</li><li>• U.S.</li></ul>
--	--

Yelp provides, in a zipped file, multiple datasets which include general information on businesses, reviews received, information on users, check-ins logged, tips given, and user-uploaded images. For our project, we are working only on the data related to restaurants and food chains.

Using Yelp's dataset, we attempted to answer the following questions:

- What are the popular times and days to eat out in different countries?
- What are the most common cuisines and topics in different countries?
- Which countries have the harshest and/or nicest comments on the reviews?
- What are the top 25 restaurants based on number of reviews and best reviews? How about 25 restaurants with the least and worst reviews? And lastly, what are the 10 cities where Yelp is used the most?

### DATA SOURCE:

- [https://www.yelp.com/dataset\\_challenge](https://www.yelp.com/dataset_challenge)

## About the Dataset

Yelp provided the datasets as Json files. The following is detailed information on the different datasets included as well as the attributes contained in the Json file:

### Business Dataset

The business dataset provides general information about different businesses in different locations. Each business attribute includes a unique Business ID and other details such as latitude and longitude of the city it is situated in, Yelp stars received, and review counts.

```
{
  "business_id": "encrypted business id",
  "name": "business name",
  "neighborhood": "hood name",
  "address": "full address",
  "city": "city",
  "state": "state -- if applicable --",
  "postal_code": "postal code",
  "latitude": latitude,
  "longitude": longitude,
  "stars": star rating, rounded to half-stars,
  "review_count": number of reviews,
  "is_open": 0/1 (closed/open),
  "attributes": ["an array of strings: each array element is an attribute"],
  "categories": ["an array of strings of business categories"],
  "hours": ["an array of strings of business hours"],
  "type": "business"
}

{
  "business_id": "0DI8Dt2PJp07XkVvIEIcQ",
  "name": "Innovative Vapors",
  "neighborhood": "",
  "address": "227 E Baseline Rd, Ste J2",
  "city": "Tempe",
  "state": "AZ",
  "postal_code": "85283",
  "latitude": 33.3782141,
  "longitude": -111.936102,
  "stars": 4.5,
  "review_count": 17,
  "is_open": 0,
  "attributes": [
    "BikeParking: True",
    "BusinessAcceptsBitcoin: False",
    "BusinessAcceptsCreditCards: True",
    "BusinessParking: {'garage': False, 'street': False, 'validated': False, 'lot': True, 'valet': Fals",
    "DogsAllowed: False",
    "RestaurantsPriceRange2: 2",
    "WheelchairAccessible: True"
  ],
  "categories": [
    "Tobacco Shops",
    "Nightlife",
    "Vape Shops",
    "Shopping"
  ],
  "hours": [
    "Monday 11:0-21:0",
    "Tuesday 11:0-21:0",
    "Wednesday 11:0-21:0",
    "Thursday 11:0-21:0",
    "Friday 11:0-22:0",
    "Saturday 10:0-22:0",
    "Sunday 11:0-18:0"
  ],
  "type": "business"
}
```

## Review Dataset

The review dataset provides detail of reviews posted by Yelp users for different business. Each record contains attributes which have a unique Review ID, the User ID of the Yelp user, and the Business ID of the business reviewed by the Yelp user; it also includes other attributes like the number of stars posted by the user and date in which the review was posted.

```
{
  "review_id": "encrypted review id",
  "user_id": "encrypted user id",
  "business_id": "encrypted business id",
  "stars": star rating, rounded to half-stars,
  "date": "date formatted like 2009-12-19",
  "text": "review text",
  "useful": number of useful votes received,
  "funny": number of funny votes received,
  "cool": number of cool review votes received,
  "type": "review"
}

{
  "review_id": "_a7Zu2ZSEGO4b12gvu7OtQ",
  "user_id": "jhhHm3Vk9Z1P21WdY_5R0w",
  "business_id": "0cczfEgv9KAD4V1Ia7ANPNQ",
  "stars": 5,
  "date": "2009-04-10",
  "text": "I love Mint, and even though I'm a guy and there isn't much for me there, it's a great place for gifts fo",
  "useful": 2,
  "funny": 2,
  "cool": 1,
  "type": "review"
}
```

## User Dataset

The user dataset contains attributes that provide information on the different Yelp users registered on the site. Each record contains a unique User ID, the User IDs of other people the user is friends/followers with, the number of reviews posted by the user, how long has he/she has been a Yelp user, and other information pertaining to the reviews posted by the user.

```

{
  "user_id": "encrypted user id",
  "name": "first name",
  "review_count": number of reviews,
  "yelping_since": date formatted like "2009-12-19",
  "friends": ["an array of encrypted ids of friends"],
  "useful": "number of useful votes sent by the user",
  "funny": "number of funny votes sent by the user",
  "cool": "number of cool votes sent by the user",
  "fans": "number of fans the user has",
  "elite": ["an array of years the user was elite"],
  "average_stars": floating point average like 4.31,
  "compliment_hot": number of hot compliments received by the user,
  "compliment_more": number of more compliments received by the user,
  "compliment_profile": number of profile compliments received by the user,
  "compliment_cute": number of cute compliments received by the user,
  "compliment_list": number of list compliments received by the user,
  "compliment_note": number of note compliments received by the user,
  "compliment_plain": number of plain compliments received by the user,
  "compliment_cool": number of cool compliments received by the user,
  "compliment_funny": number of funny compliments received by the user,
  "compliment_writer": number of writer compliments received by the user,
  "compliment_photos": number of photo compliments received by the user,
  "type": "user"
}

{
  "user_id": "EZmocAborM6z66rTzeZxzQ",
  "name": "Rob",
  "review_count": 761,
  "yelping_since": "2009-09-12",
  "friends": [
    "iJg9ekPzF91kMuvjKYX6uA",
    "ctWAuzS04Xu01ke2Rop4lQ",
    "B8CqppjOne8X4RSJ5KY0vQ",
    "K9sK1A4fVkwI4hyGSPoPA",
    "Ec-epOsAWvjI6e90I1M8jw",
    "r2UUCzGxqI6WPsiWPgqG2A",
    "3ybkL7N63UdSn4wepINzUw",
    "d-1zusSagnkDuiyL1fF5pw",
    "Ydh2zA5wU1D-UbApp8toGA",
    "DeZhnC-RsNFmKS1I01Uksw",
    "NTuvVb-ZwQ_rFn6W9Krm7A",
    "PCdUS3L8LhQOereIyQ6_RA",
    "RYIhfNekKLDuqmwqkV13g",
    "i-mQUeJVKqXsclJ7XjT2w",
    "5T_2RWFzvf-nMZYD0RsvWg",
    "DTo6vHKzyx8RmMrVbTgowg",
    "kOUzW4porPxN0n54Z0eHMQ",
    "x-1u6ZBVTjyHFvoq3f0K4w",
    "8Gy1MJ4gCnOL8qiWd2ZGLQ",
    "8esbQFK-w5xRMnQZvnaADg",
    "YeEUeu7W3TM9PjBsGlyktg",
  ]
}

```

## Check-In Dataset

The check-In dataset includes a time attribute which provides information on the day, time, and hour the Yelp user checked in the restaurant, the Business ID of the restaurant, and the type of check-in.

```

{
  "time": ["an array of check ins with the format day-hour:number of check ins from hour to hour+1"],
  "business_id": "encrypted business id",
  "type": "checkin"
}

```

```
{
  "time": [
    "Fri-0:2",
    "Sat-0:1",
    "Sun-0:1",
    "Wed-0:2",
    "Sat-1:2",
    "Thu-1:1",
    "Wed-1:1",
    "Sat-2:1",
    "Sun-2:2",
    "Thu-2:1",
    "Wed-2:1",
    "Fri-3:1",
    "Sun-3:3",
    "Thu-4:1",
    "Tue-4:1",
    "Sun-6:1",
    "Wed-6:1",
    "Fri-10:1",
    "Sat-10:1",
    "Mon-11:1",
    "Wed-11:2",
    "Mon-12:1"
  ]
}
```

## Tips Dataset

The tip dataset is similar to review dataset in that both include attributes liked text, which contains reviews published by users on the restaurants (specifically geared towards advice to other users), a unique Business ID of those restaurants, and the likes a tip received from other users .

```
{
  "text": "text of the tip",
  "date": "date formatted like 2009-12-19",
  "likes": compliment count,
  "business_id": "encrypted business id",
  "user_id": "encrypted user id",
  "type": "tip"
}

{
  "text": "Pretty cool place. Good prices. Awesome finds.",
  "date": "2016-11-22",
  "likes": 0,
  "business_id": "ETJAKsH4DFL_G1dBtDL3FQ",
  "user_id": "_44C-qv9h_GBkqLJRTwOwg",
  "type": "tip"
}
```

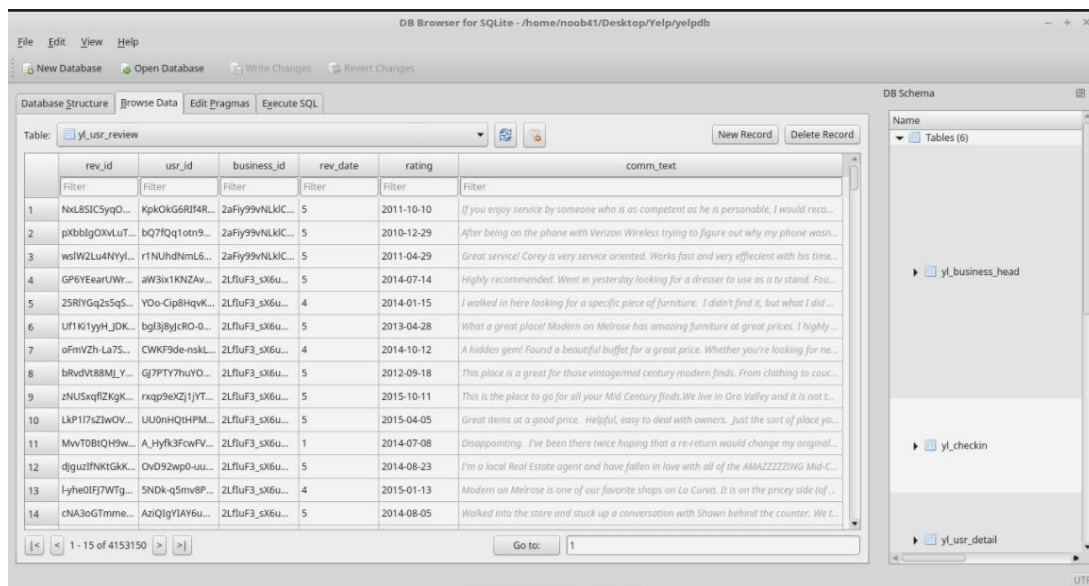
# Data Processing and Filtering

The following is a list of the different modules and tools used in the project:

1. SQLITE 3
2. TextBlob
3. Matplotlib
4. WordCloud
5. NLTK

## Sqlite

For our analysis, we were required to join attributes from different datasets. As our dataset was of size 3gb and more, we decided to use Sqlite. We created different tables for each dataset in Sqlite and in our code, we accessed the Sqlite database and joined tables based on required attributes; from there, we ran the analysis on the output.



## TextBlob

We used TextBlob to run polarity and subjectivity analysis on the desired combination of the dataset.

## Matplotlib

We used Matplotlib to provide graphical representation of different check-in time trends in different cities in our dataset.

## WordCloud

We use Wordcloud to find the most commonly used words in the reviews of the Yelp users for different cities in the countries.



## Our Analysis

### What are the popular times and days to eat out in different countries?

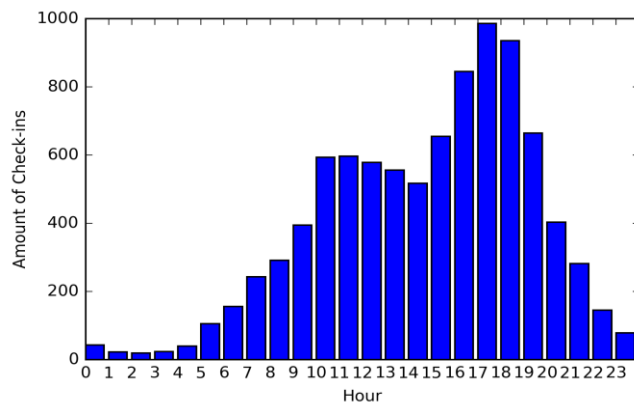
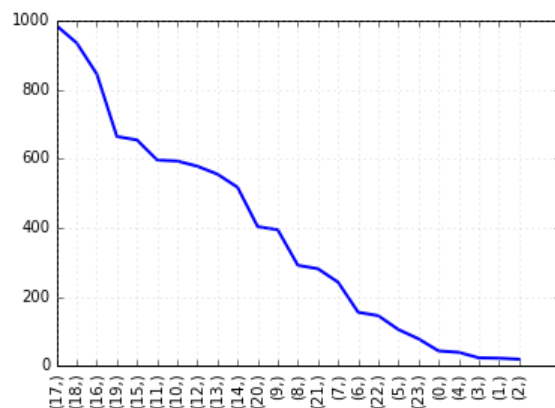
Using the business dataset and the check-in dataset, we analyzed how many times users checked in at a restaurant to find the peak time and day users from different countries liked to visit restaurants and food chains.

In the code labelled `import_json_checkin_peaktime.py`, we pulled data pertaining to the user check-in time and ran it through both a list and a counter dictionary. The list was used to calculate frequency and plot a frequency graph, while the dictionary was plotted using Matplotlib to create a bar graph. In the code labelled `import_json_checkin_weekend_trends.py`, we pulled data related to the day of the week the user checked in. In that code, a frequency graph was created.

We went through this process for different cities in all four countries available on our dataset.

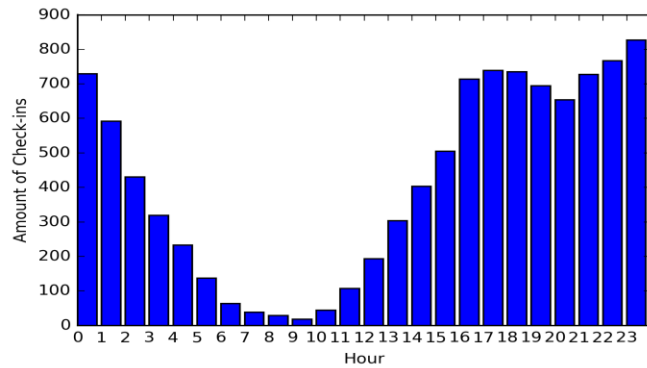
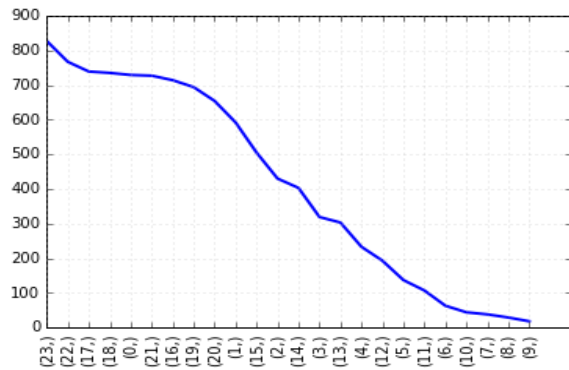
### Popular Check-In Times

#### Germany



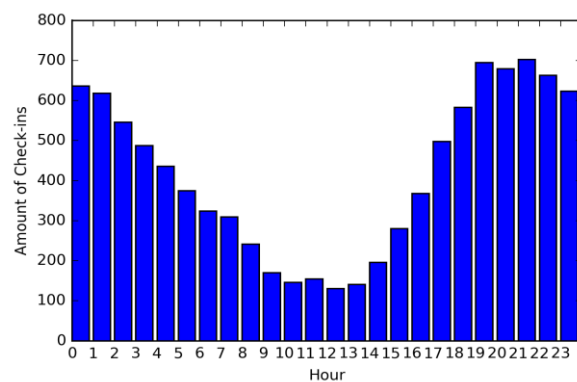
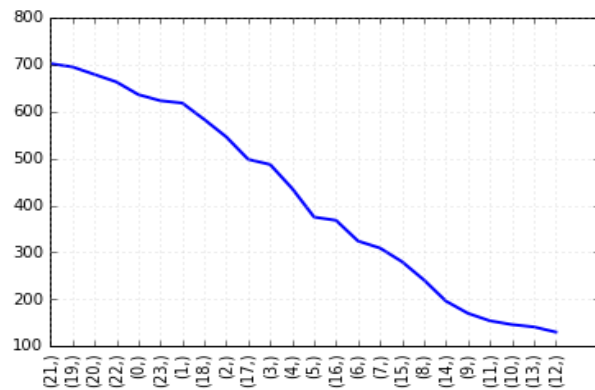
- **Peak check-in time:** 17:00-18:00, or 5:00-6:00 PM
- **Lowest check-in time:** 1:00-2:00
- Dinnertime seems to be the most popular time to eat out for German Yelpers, with lunch coming in second and brunch/breakfast coming in third. Early morning is expectedly the time the least people are eating out.

## Canada



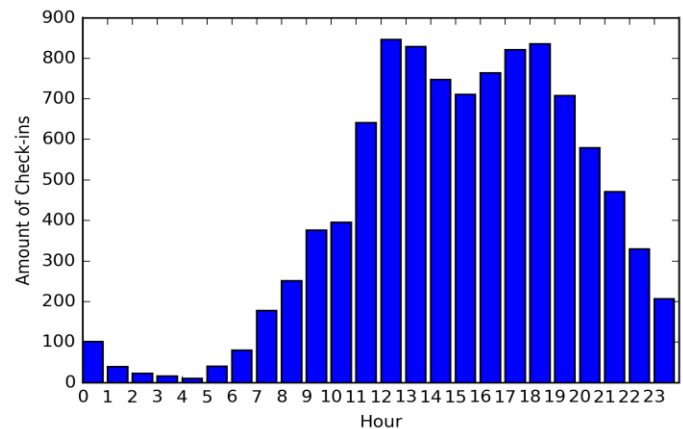
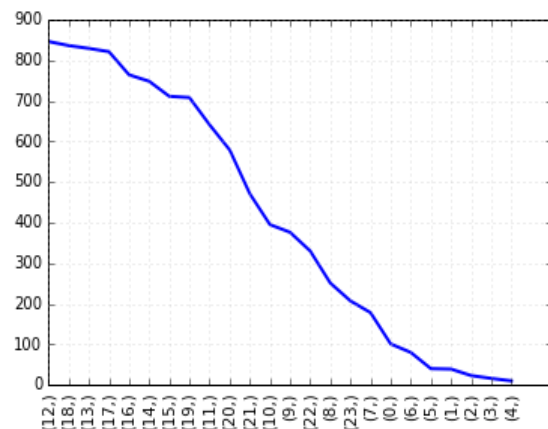
- **Peak check-in time:** 23:00-24:00, or 11:00 PM-12:00 AM
- **Lowest check-in time:** 9:00-10:00
- Once again dinnertime seems to be the most popular time to eat out, albeit more on the later side than was seen with Germany. Breakfast/brunch don't seem particularly popular times to eat here as well, with the number of check-ins increasing the later (and eventually, earlier) the day. A surprisingly amount of people checked in during the early mornings.

## USA



- **Peak check-in time:** 21:00-22:00, or 9:00-10:00 PM
- **Lowest check-in time:** 12:00-1:00 PM
- Dinnertime seems to be the most popular time to eat out for US users, though not as late as we saw with Canadian ones. The bar graph resembles to that of Canada's graph (with a lot of early morning check-ins), but a key difference here is that users checked in more during breakfast than they did during brunch and lunch.

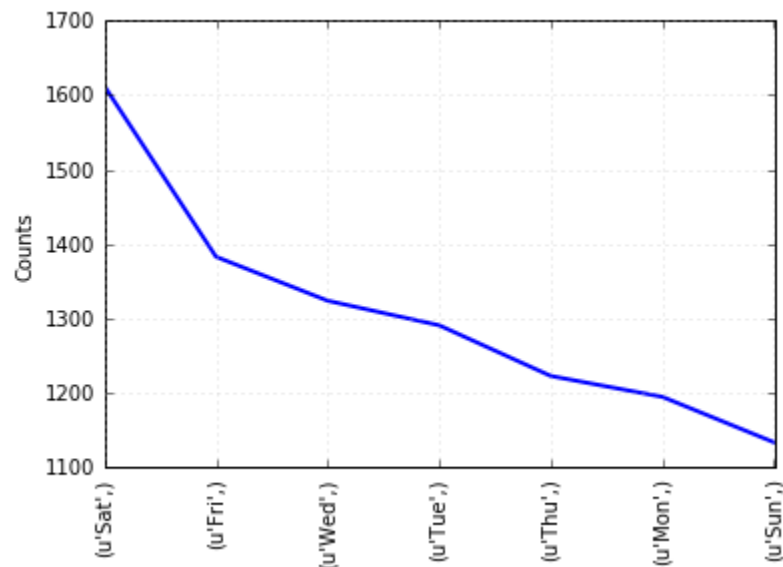
## U.K.



- **Peak check-in time:** 12:00-13:00, or 12:00-1:00 PM
- **Lowest check-in time:** 4:00-5:00 AM
- Lunch is the most popular time to eat out for users in the UK, though dinner and the time in between the two do not seem to be too far behind. The bar graph here resembles the one for Germany more so than it does the other two, with early morning and breakfast having the least amount of check-ins.

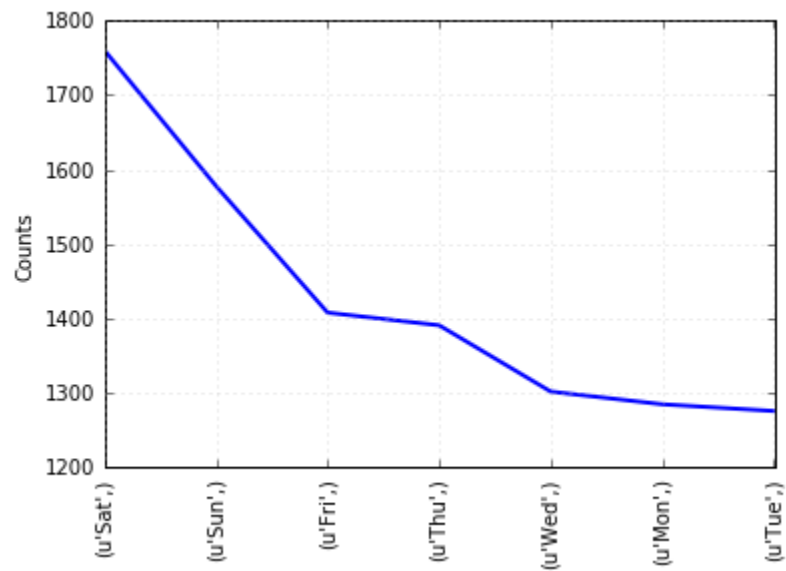
## Popular Check-In Days

### GERMANY



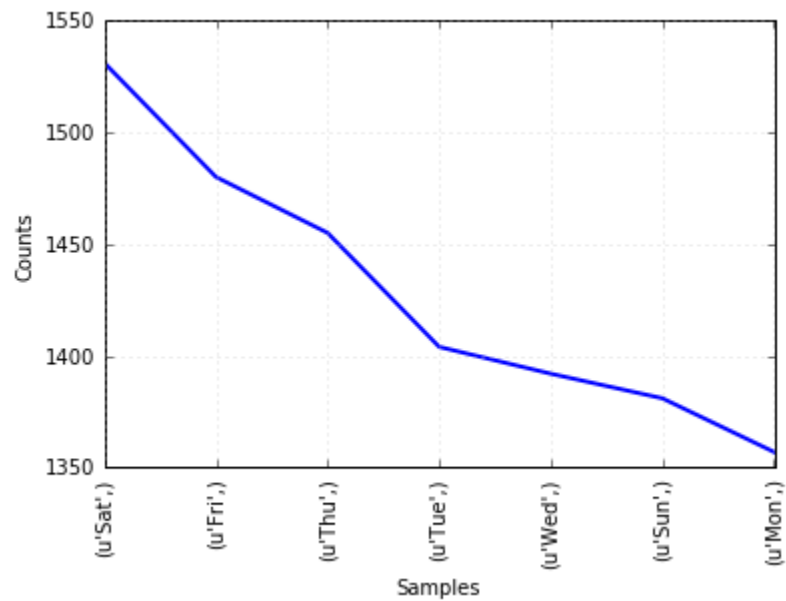
- **Most popular:** Saturday, followed by Friday
- **Least popular:** Sunday

## CANADA



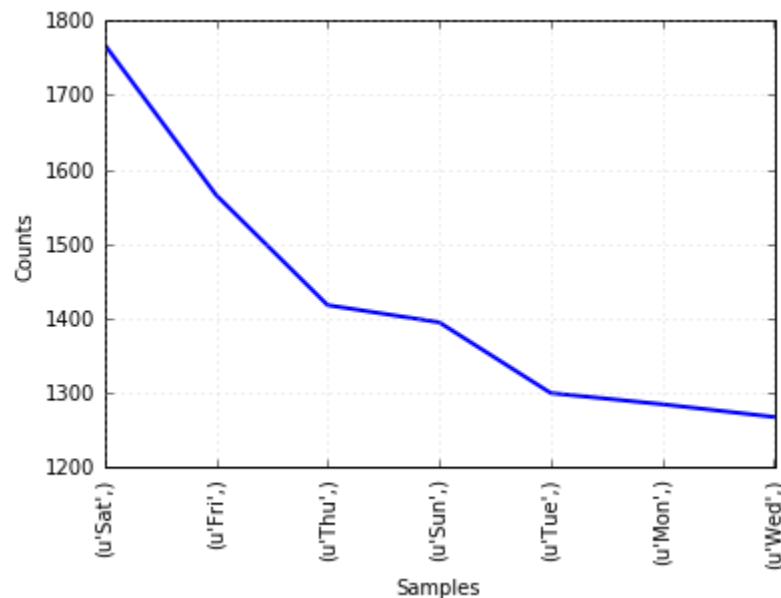
- **Most popular:** Saturday, followed by Sunday
- **Least popular:** Tuesday

## U.S.A



- **Most popular:** Saturday, followed by Friday
- **Least popular:** Monday

U.K.



- **Most popular:** Saturday, followed by Friday
- **Least popular:** Wednesday

### Analysis

- USA and Canada check-ins peaked during the early mornings/late evenings, while check-ins in the UK and Germany increased as the day went by and peaked around lunch/dinner.
  - All of the countries had peak-in check-in times in the evening (around dinnertime), except for the UK, which experienced its peak during noon/lunchtime.
- The shape of the bar graphs for check-in times for the USA resembled Canada's and the graph for the UK resembled Germany's.
  - The first pair are countries in North America while the second are countries in Europe, which may be the reason for the commonalities. These countries have different cultures and societies, but their (eating) habits seem to match their neighbors more than they do the countries across the Atlantic.
- All four of the countries experienced the most check-ins on Saturdays. Friday experienced the second greatest number of check-ins for every country except for Canada, where that spot belonged to Sunday.
- Except for Germany, where the least number of check-ins occurred on Sunday, users in all the other countries checked in the least on a regular weekday. This makes sense when you consider that most of these people probably work during the weekdays.

### What cuisines do Yelpers rave about in different countries? Does it differ?

For this question, we merged the review dataset with the business dataset to find the cuisines Yelpers rave about in different countries.

In `import_json_checkin_Wordcloud.py`, we used the output of the merged review dataset with the business dataset and generated a word cloud of the reviews. Stop words were set up so that irrelevant or meaningless words (like the term “restaurant” itself) would not be included.

The following is our analysis:

#### GERMANY



- **Notable words:** cafes, bars, Italian, German, Swabian, service, nightlife, health, tea, beer, hotels, art, garden, local, active, entertainment, coffee

#### CANADA



- **Notable words:** nightlife, active, tea, life, art, service, specialty, coffee, bar, Canadian, American, local, breakfast, brunch, sandwiches

U.S.A.



- **Notable words:** bar, art, active, nightlife, American, entertainment, local, life, service, traditional, sandwiches, hotel

U.K.



- **Notable words:** British, bakeries, delis, Chinese, Indian, bars, cafes, tea, nightlife breakfast, services, seafood, fast, sandwiches, Italian, brunch, pizza.

## Analysis

- A couple of words showed up in the WordClouds (in similar sizes) for all four countries; this indicates that some trends are popular irrespective of where the user was in terms of location.
  - Some of these words are: **bars**, **cafes**, and **nightlife**.
- All four countries except for the USA have **tea** and **coffee** in big font, indicating that tea and coffee are trends common and dear to users in the Germany, UK, and Canada. This may also mean that there are many establishments selling tea and coffee in those three countries.
- All four countries except for the UK have **art** and **service** in big font, showing that ambience and service is of great importance for them.
- Cuisines:
  - In **Germany**, some popular cuisines are **Italian**, **German**, and **Swabian**. The last one is notable (despite being smaller than the other two in terms of font size) because Swabia is a specific region in southwestern Germany.
  - In **Canada**, **American** and **Canadian** are the most popular cuisines reviewed.
  - In the **USA**, **American** seems to be the only notable cuisine listed in big font on the WordCloud.



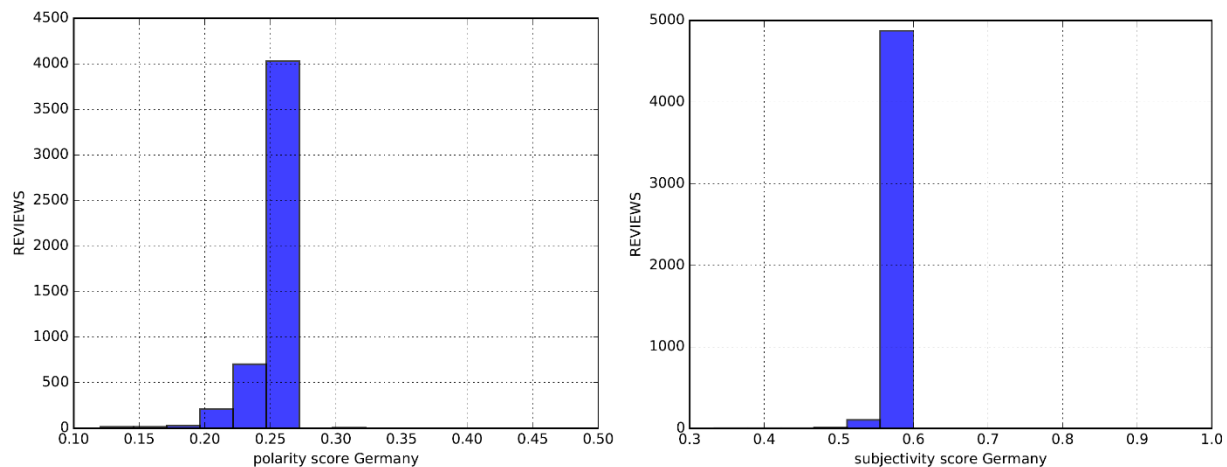
- In the **UK, British, Italian, Indian, Chinese, and American** show up as popular categories—the most out of all four countries reviewed.
- **American** cuisine is the one cuisine that shows up in the WordClouds for all the countries.
- The WordCloud for the UK seems to have the most variety in terms of popular food categories - **sandwiches, delis, fast, bakeries** and **seafood** are popular in this WordCloud but not in any of the others.
- **Breakfast** and **brunch** appear to be much more popular in Canada, UK, and USA, than in Germany.
- **Service** appears to be particularly important in Canada and Germany.

### Which countries have the harshest and/or nicest reviewers?

To answer this question, we ran sentimental analysis on the reviews contained in the review dataset. The code for this is contained in `import_json_business_sentimental.py`. Polarity measures how positive or negative a piece of text is, with negative numbers meaning negative sentiments, and positive numbers meaning positive sentiments. Subjectivity measures how personal the language used in the reviews were, and whether they contain strong emotions, feelings, opinions, etc. Those scores range from 0-1.0, with reviews closer to 0 being more objective, while reviews closer to 1.0 are more subjective.

The following is our analysis.

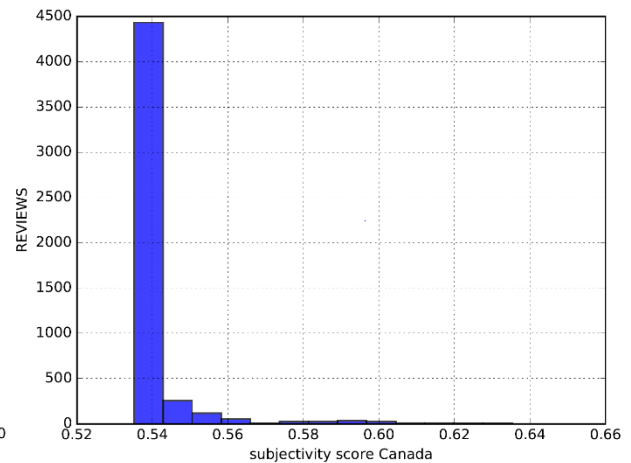
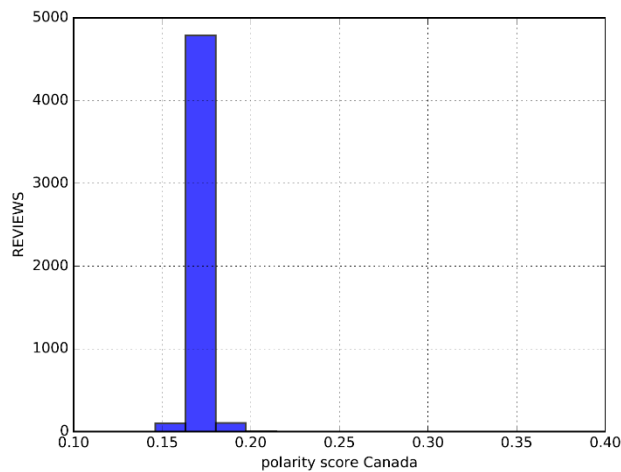
#### GERMANY



- In Germany, it appears that most reviews came in around 0.20-0.30 in polarity. Reviews seem to be very close to neutral, with only a slight positive bias.
- Subjectivity scores mostly came in between 0.5 and 0.6. We can interpret this as meaning that most reviews are subjective, but not overly so--fitting, given the nature of reviews.

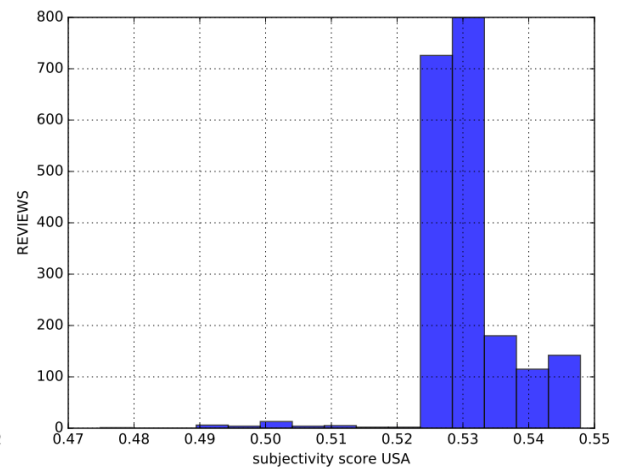
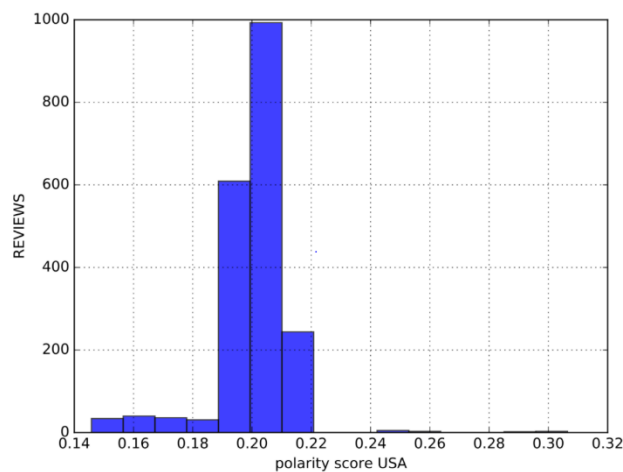


## CANADA



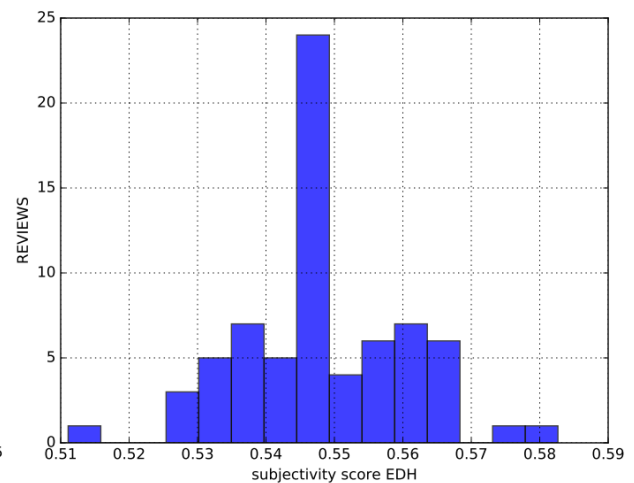
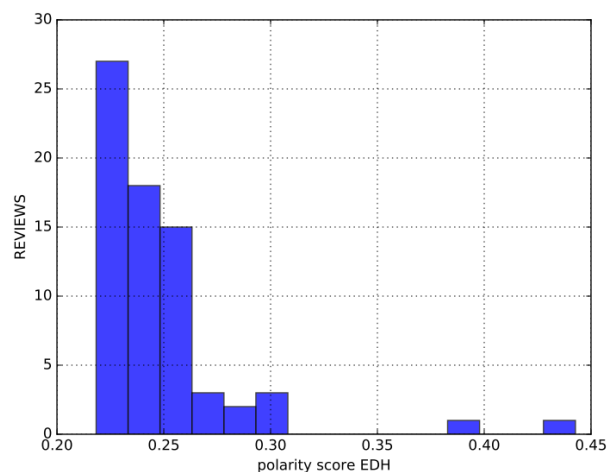
- Polarity scores for reviews in Canada fall between 0.15 and 0.20. For the most part they are neutral and not terribly positive.
- Subjectivity scores in Canada mostly fall into the range of 0.54 and 0.56--once again, subjective but not by a huge amount.

## U.S.



- Polarity scores for the US are spread across a longer range, but most of the reviews fall under 0.19-0.22. A little bit on the positive side, once again.
- Subjectivity scores here range from 0.52 to 0.55. Reviews seem personal for the most part, but not overwhelmingly so.

## U.K.



- Polarity scores here are spread across a range from about 0.22 to 0.31, though most reviews seemed to score around 0.22, indicating a slight positivity in reviews.
- Subjectivity is centered around 0.54-0.55, with a couple in the outskirts and none below 0.51; we can say from this that subjective language was used in reviews.

## Analysis

- There doesn't seem to be a noticeable variation across countries when it comes to polarity and subjectivity of reviews, therefore, we can't make a definitive conclusion on which countries give the harshest or nicest reviews.
  - Most polarity scores fell within the 0.15--0.30 range. This means most review held a bit of positivity in them, though not much.
  - Meanwhile, most subjectivity scores ranged from 0.50 to 0.60. This indicates the use of a notable amount of subjective/personal language in reviews. Reviews are intended to have some opinions in them, so this comes of no surprise.
  - If a country had to be chosen, however, **Germany** would be the country with the most "positive" sounding reviews (based on polarity scores), while **Canada** (and maybe the **US**) had the least.

## Top 25 Restaurants & Top 10 Cities

For this last analysis, we used information from the business dataset and the review dataset to capture the amount of reviews a given establishment had. From there, it was a matter of ordering the data in different ways:

- To find the top 25 restaurants with the most reviews, the list of restaurants was sorted in descending order based on review count and cut off at 25 restaurants. To find the 25 restaurants with the least amount of reviews, the list was simply sorted in ascending order.

- For restaurants with the greatest and worst ratings, the list was sorted in descending and ascending order of ratings respectively.
- To find the Top 10 where Yelp is used the most, cities, instead of restaurants, were sorted in descending order based on review count and only the top 10 results were taken.

The code for this portion is all included in **top25.py**. The following is the output that was generated:

#### Top 25 Restaurants with the Most Reviews:

Restaurant	Number of Reviews	Location (City)
Mon Ami Gabi	6414	Las Vegas
Bacchanal Buffet	5715	Las Vegas
Wicked Spoon	5216	Las Vegas
Gordon Ramsay BurGR	5116	Las Vegas
Earl of Sandwich	4655	Las Vegas
Gangnam Asian BBQ Dining	4120	Las Vegas
Serendipity 3	3911	Las Vegas
Hash House A Go Go	3881	Las Vegas
The Buffet	3676	Las Vegas
The Buffet at Bellagio	3481	Las Vegas
Lotus of Siam	3458	Las Vegas
Bouchon at the Venezia Tower	3292	Las Vegas
Secret Pizza	3258	Las Vegas
MGM Grand Hotel	3015	Las Vegas
Hash House A Go Go	2835	Las Vegas
Mesa Grill	2771	Las Vegas
Bachi Burger	2735	Las Vegas
Gordon Ramsay Steak	2700	Las Vegas
Holsteins Shakes and Buns	2474	Las Vegas
El Dorado Cantina	2316	Las Vegas
Burger Bar	2313	Las Vegas
Egg & I	2303	Las Vegas
The Peppermill Restaurant & Fireside Lounge	2260	Las Vegas
Pho Kim Long	2247	Las Vegas
Grand Lux Cafe	2218	Las Vegas

### Top 25 Restaurants with the Best Ratings:

Restaurant	Rating (1-5)	Location (City)
Feinkost Galeria	5.0	Ludwigsburg
Covered Wagon	5.0	Hawthorne
Lake City Tavern	5.0	Berea
Tropical Smoothie Cafe	5.0	Charlotte
Nordsee	5.0	Ludwigsburg
Cactus Cafe	5.0	Las Vegas
Kantina	5.0	Cleveland
Phyllo Bar Melina's	5.0	Montreal
Mi Pueblo Taco Shop	5.0	Las Vegas
Sincerely Yogurt	5.0	Pittsburgh
L'Express St Zotique	5.0	Montreal
Momo Sushi	5.0	Boulder City
Taqueria Castillo	5.0	Phoenix
Port of Subs	5.0	Las Vegas
El Ausente	5.0	Las Vegas
Vienna West	5.0	Cleveland
Thai Cuisine Experts	5.0	Mississauga
Mesopotamien	5.0	Ostfildern
Burrito Caliente Latin Grill	5.0	Toronto
Winter In July - A Pop-Up Lunch with Yonaka and Snow Shavery	5.0	Las Vegas
Chez Lavigne	5.0	Montreal
Oko Sushi	5.0	Kirkland
Pincho	5.0	Stuttgart
Murdock Meals	5.0	Las Vegas
BO'TECA di vino	5.0	Stuttgart

## 25 Restaurants with the Least Reviews:

Restaurant	Number of Reviews	Location (City)
La Prep	3	Montreal
Pho Viet	3	Phoenix
Covered Wagon	3	Hawthorne
McDonald's	3	Mississauga
El Paso Bar-B-Que Company	3	Surprise
Pizza Fresh	3	Fellbach
Da Toni Pizza und Eis	3	Toronto
New Spiceland Restaurant	3	Gilbert
Subway	3	Gilbert
Taste Cafe	3	Markham
Pizza Jarry & Restaurant	3	Montreal
Chen's Wok	3	Pittsburgh
Nordsee	3	Ludwigsburg
Casa De Angelina	3	Surprise
Ristorante Italiano	3	Las Vegas
Domino's Pizza	3	Chandler
Burger King Canada	3	Mississauga
Pizza Pinocchio	3	Hemmingen
Tandoor	3	Toronto
Crofters	3	Edinburgh
Nubo Sushi	3	Bois-Des-Filion
Burger King	3	Cleveland
Pizza Bellagio	3	Pittsburgh
Desi Hut Takeout & Catering	3	Toronto
The Cleveland Corned Beef Company	3	Goodyear

### 25 Restaurants with the Most Unfavorable Ratings:

Restaurant	Rating (1-5)	Location (City)
McDonald's	1.0	Maple Heights
Tandoor	1.0	Toronto
Burger King	1.0	Phoenix
Brooklyn Diner	1.0	Charlotte
Pizza Hut	1.0	Charlottes
Burger King	1.0	Champaign
Pizza Hut	1.0	Medina
Sonic Drive-In	1.0	Tolleson
Yummy Buffet	1.0	Madison
Wendy's	1.0	Urbana
Burger King	1.0	Las Vegas
Souper Salad	1.0	Mesa
Shanghai 1938	1.0	Urbana
Pizza Hut	1.0	Charlotte
Lady D's Pizzeria	1.0	Las Vegas
Wienerschnitzel	1.0	Las Vegas
Jacky's Seafood Restaurant	1.0	Toronto
Ho Lee Chow	1.0	York
Pizza Hut	1.0	Charlotte
Taco Bell	1.0	Cleveland
McDonald's	1.0	Charlotte
Pizza Hut	1.0	Solon
McDonald's	1.0	Pittsburgh
Tim Hortons	1.0	Scarborough
Subway	1.0	Castle Shannon

### Top 10 Cities Where Yelp is Used the Most:

City	Number of Reviews
Las Vegas	1295478
Phoenix	456673
Toronto	349695
Scottsdale	245105
Charlotte	183463
Pittsburgh	143116
Tempe	129786
Henderson	126750
Mesa	97842
Montreal	96510

## Limitations and Future Implications

While reviewing our analysis, it is important to keep in mind that A) Yelp did not provide all the information at their disposal and B) Yelpers may or may not be representative of a whole country's population. It is a good place to start, however.

With this in mind, the information and analysis gleaned from the results of this project could still be of much use to anyone from aspiring chefs to Yelp themselves in growing their business and/or assisting and guiding local restaurants on how to better fine-tune their marketing. Marketing strategies can be highly situational—what is effective in one country or city may not work in others due to differences in culture and lifestyle. This is especially prevalent in the graphs and illustrations regarding peak check-in times and the WordClouds generated for each of the four countries. Certain types of foods and restaurant environments may be appealing to the users of one location may not garner the same amount of attention in others (see: the UK and seafood or Germany and Swabian cuisine). Improving or capitalizing on items deemed important by certain users could eventually make its way back to Yelp, as more favorable ratings and reviews help restaurants climb higher in Yelp's internal system. This exposure could, in turn, potentially lead to an increase in ratings, reviews, and business.

All the data provided by Yelp were of countries in the West—in the future, it would be interesting to see the trends that appear amongst restaurant-goers in Asia, South America, Africa, or Australia, and whether there are any major differences between countries there.