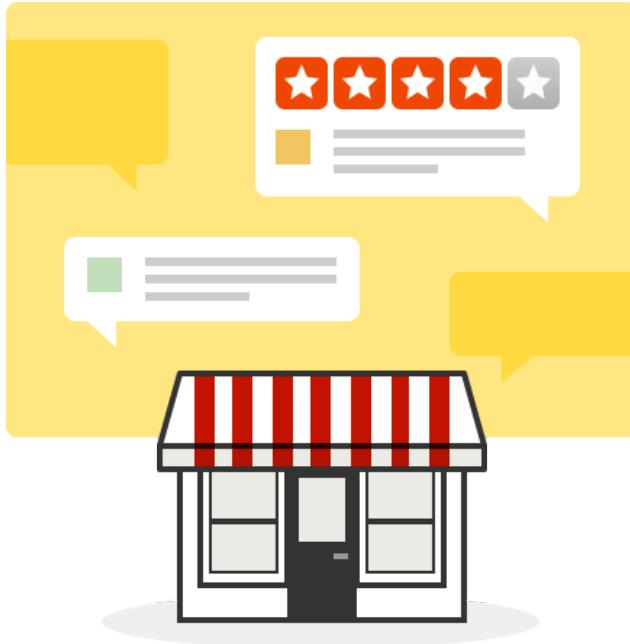




Seattle's Most
Engaging Businesses

Introduction



- 50 million businesses
- 27 countries
- Terabytes of log data every day

Approach

Search API

```
http://api.yelp.com/v2/search?  
parameters = term, location, category
```

Business API

```
http://api.yelp.com/v2/business/<business_id>
```

Approach

Building The Dataset

```
for hood in neighborhood_list:  
  
    for category in category_list:  
  
        call http://api.yelp.com/v2/search?hood&category  
  
##    response:  
  
##    728,304 businesses :-)
```

Approach

Dedupe The Dataset

```
no_dupes = {}

for row in dataset_file:

    if row['name'] in no_dupes:
        continue

## response:

## 26,088 businesses! :-(
```

Approach

In The End...

```
{  
    'categories': [],  
    'id': "",  
    'is_claimed': Boolean,  
    'location': {},  
    'name': "",  
    'rating': float,  
    'review_count': int,  
    'latitude': float,  
    'longitude': float  
}
```

Approach

Most Common & Most Reviewed Businesses

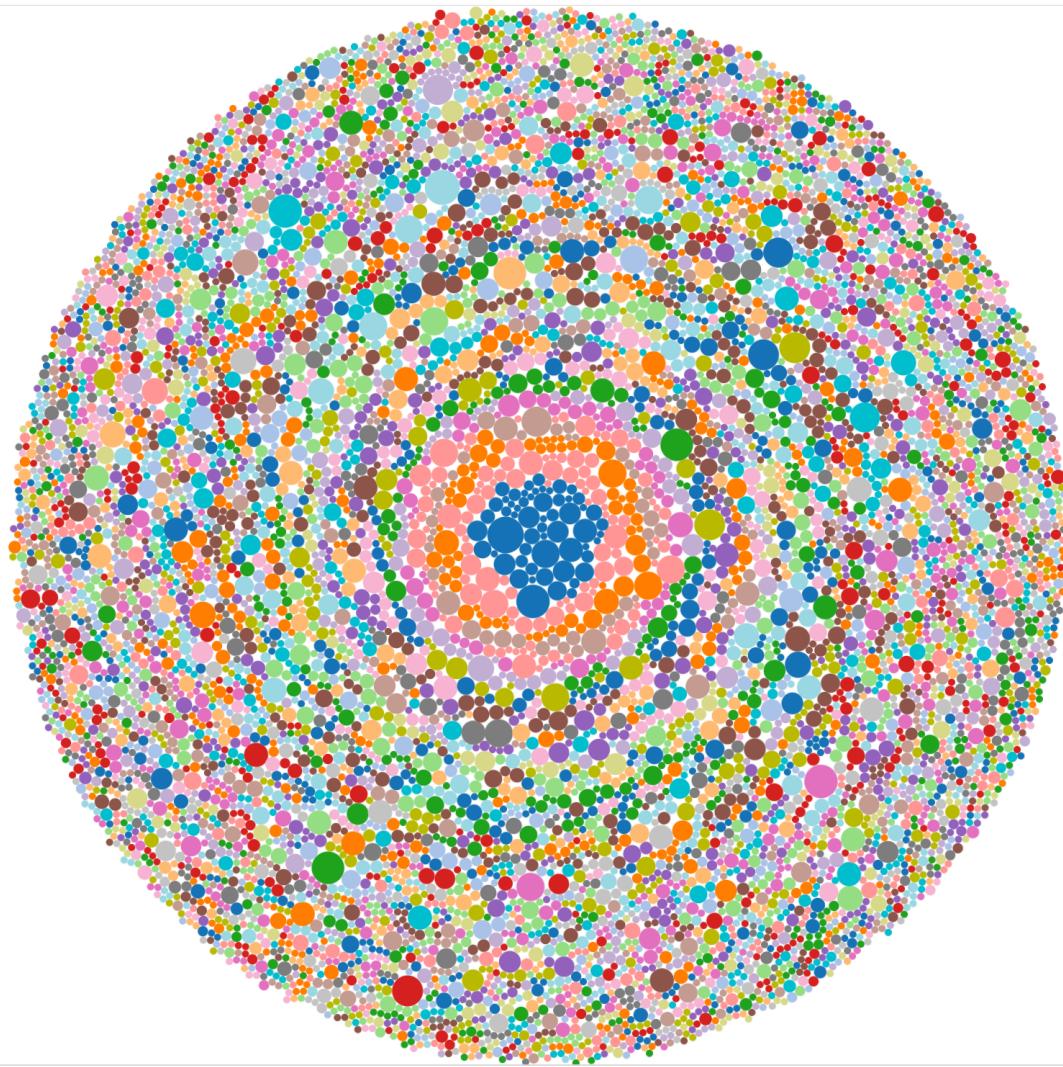
```
business_list = {}

for row in dataset_file:
    add row['neighborhood'] to the dict
    add row['category'] to the dict

most_common = count the number of times each
            category appears in each neighborhood

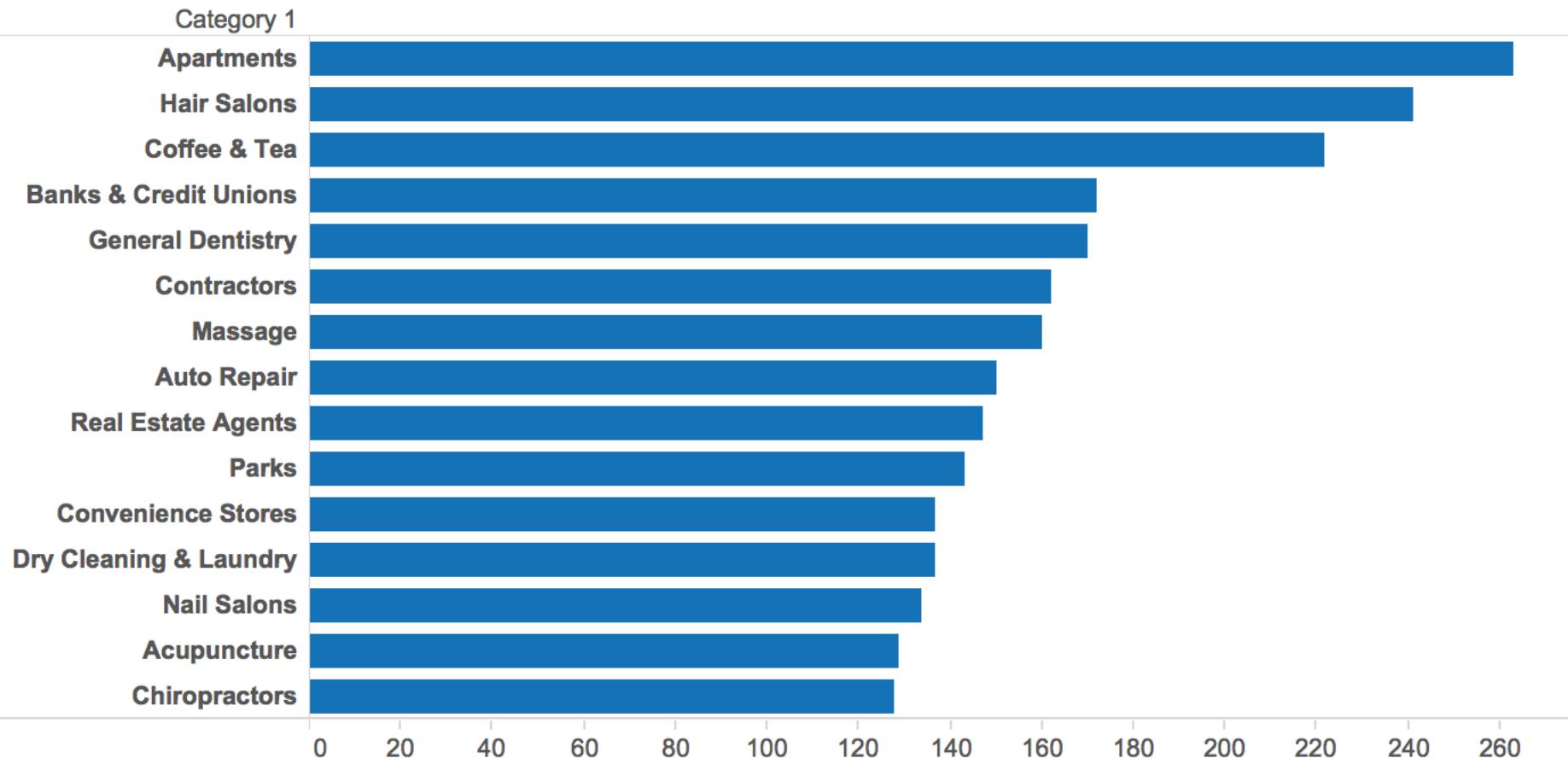
most_reviewed = total review counts for each
                category in each neighborhood
```

Analysis



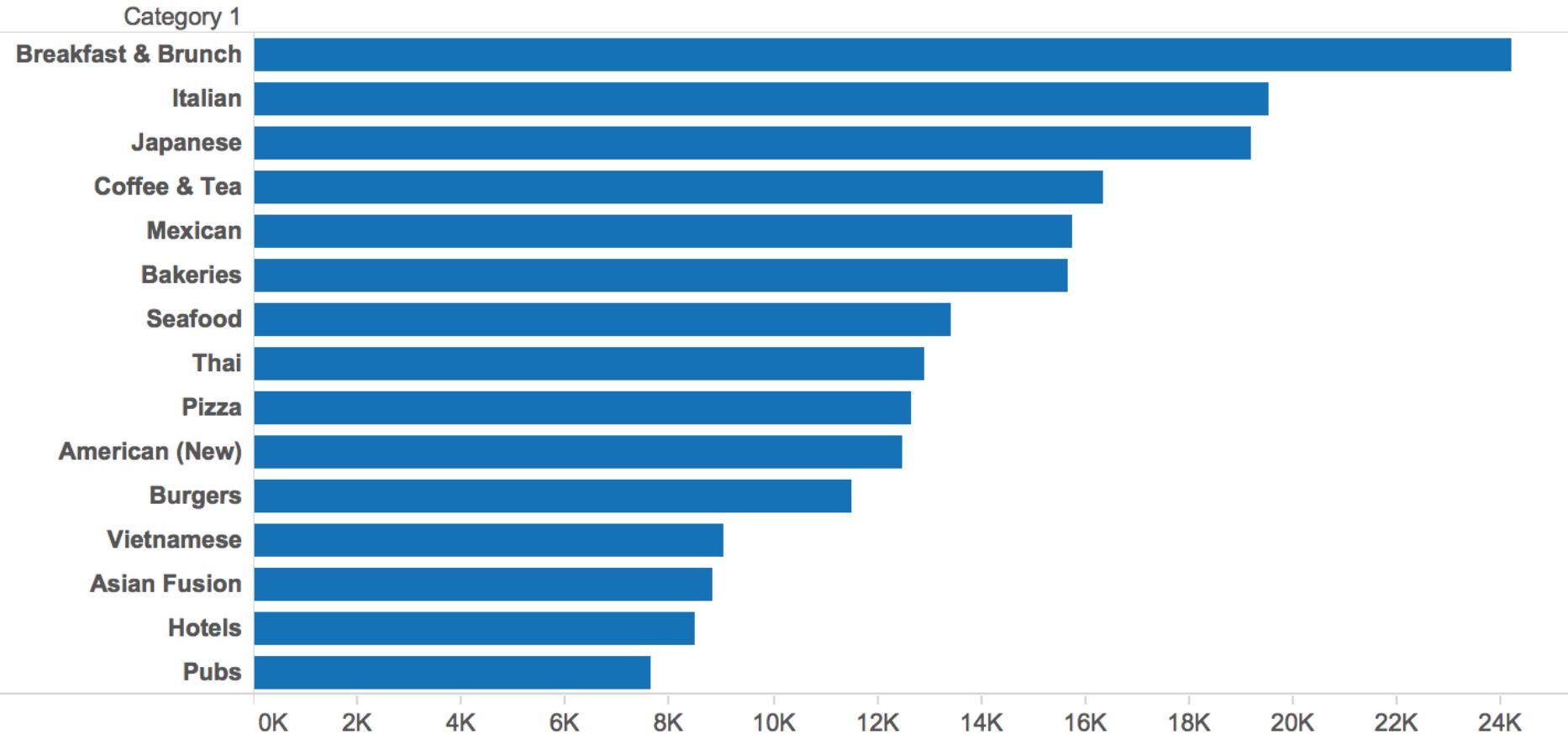
Analysis

What Are The Most Common Categories?



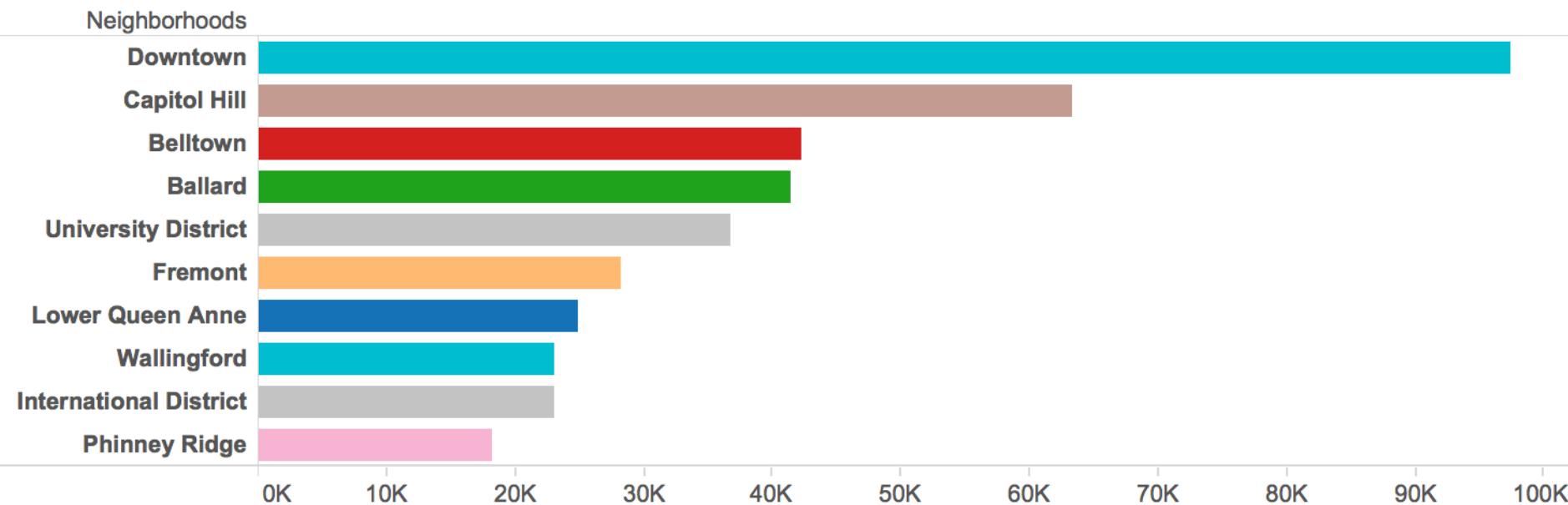
Analysis

What Are The Most Reviewed Categories?



Analysis

What Are The Neighborhoods With The Most Reviews?

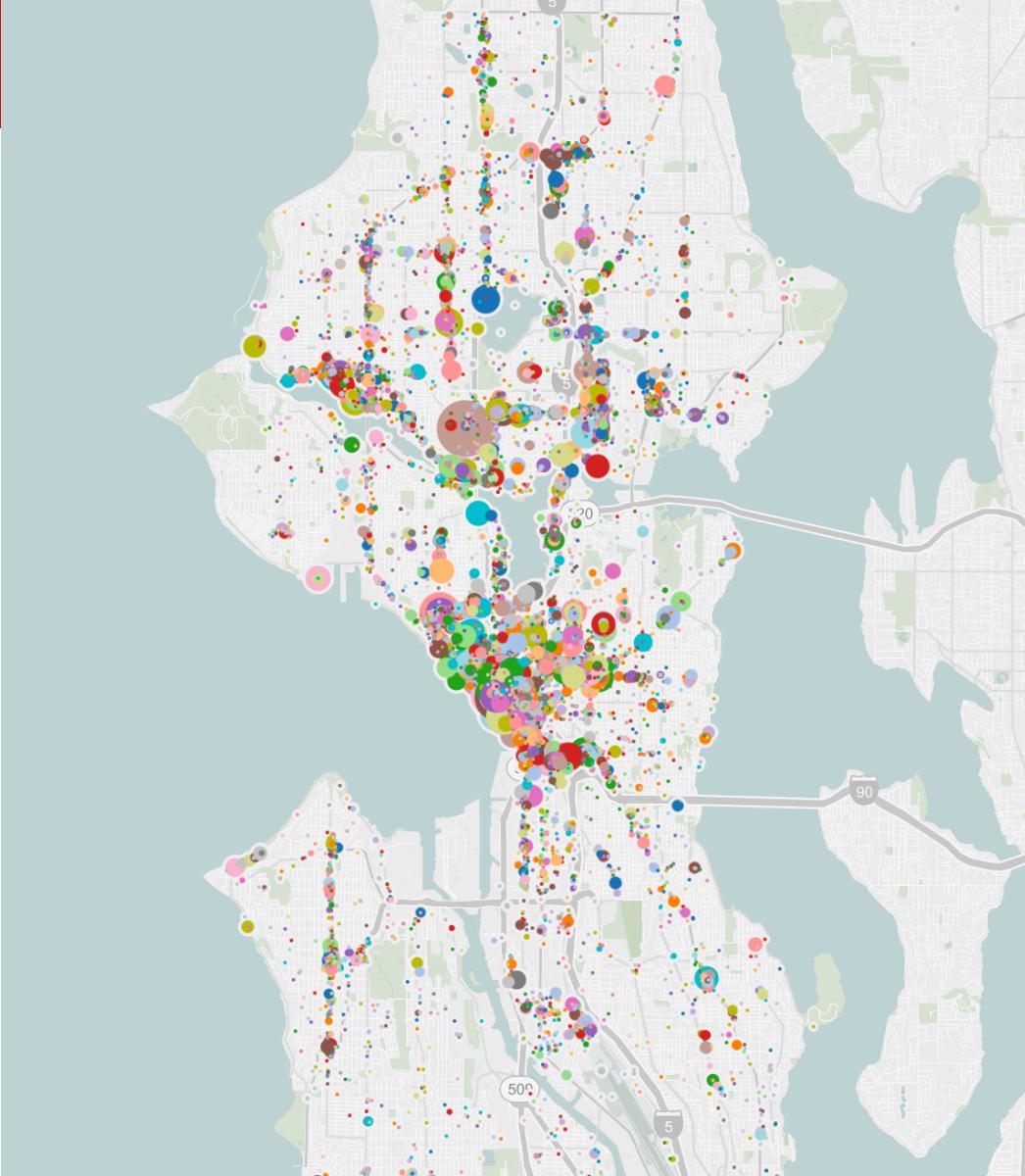


Analysis

What Are The Most Reviewed Categories In Each Neighborhood?

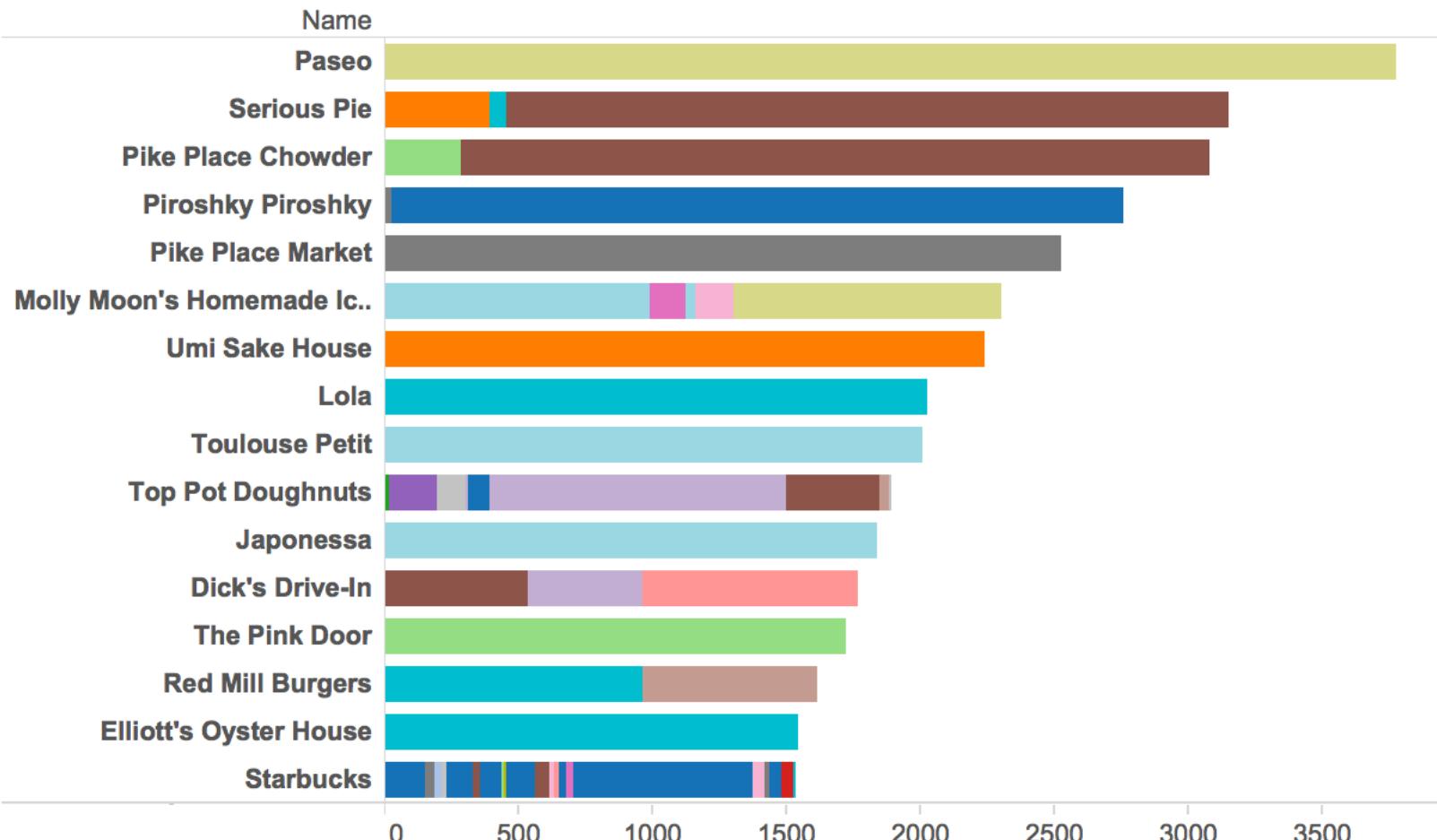
Downtown		Capitol Hill		Belltown		Ballard		University District	
Category	Reviews	Category	Reviews	Category	Reviews	Category	Reviews	Category	Reviews
Bakeries	6140	Mexican	4215	Japanese	5022	Breakfast & Brunch	1971	Thai	2361
Seafood	4056	Italian	2941	Breakfast & Brunch	3975	Mexican	1924	Breakfast & Brunch	1740
Italian	3593	Coffee & Tea	2502	Pizza	3597	Italian	1769	Indian	1496
Hotels	3577	Breakfast & Brunch	2195	Cocktail Bars	2139	Coffee & Tea	1606	Coffee & Tea	1412
Live/Raw Food	2620	Lounges	2010	Seafood	1910	Live/Raw Food	1131	Italian	939

Analysis



Analysis

What Are The Businesses With The Most Reviews?



Analysis

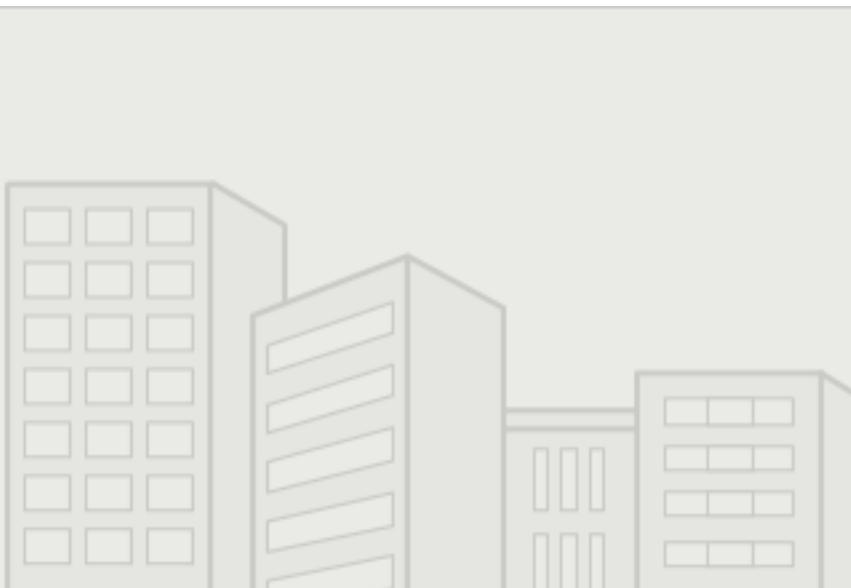
Claimed vs. Unclaimed

	Claimed	Unclaimed
Avg. Rating	4.16	3.81
Total Reviews	611802	207154

Top Neighborhoods With Unclaimed Businesses

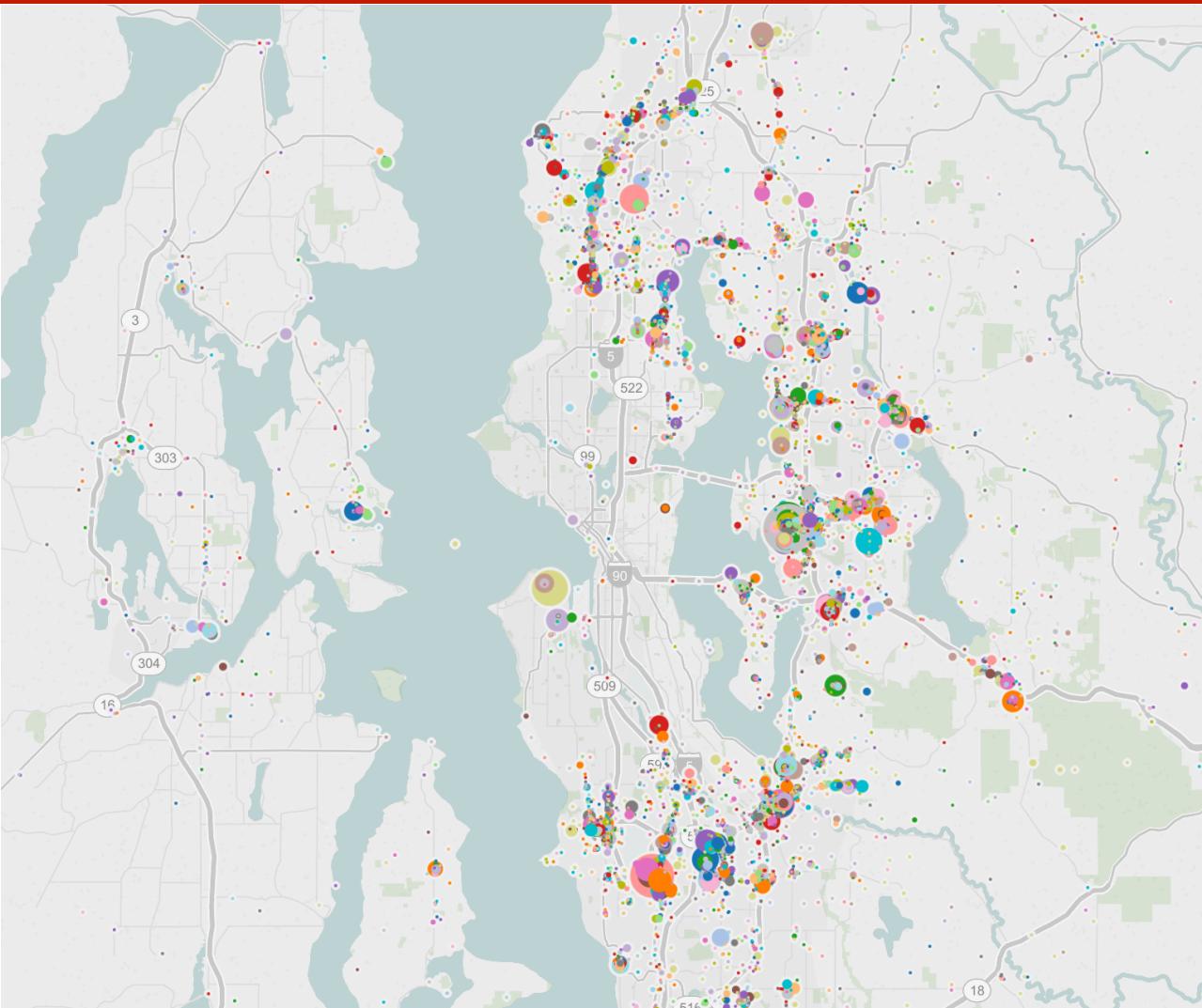
Downtown	823
U-District	375
Ballard	365
Capitol Hill	356
First Hill	300

Limitations



- Missing Businesses
- Neighborhood Tagging
- ‘none’ Neighborhood

Limitations



Thank You

