

4 Trending Diets On

Michelle Cascio
Joe Yon
Hayes Martens
Doris Pira

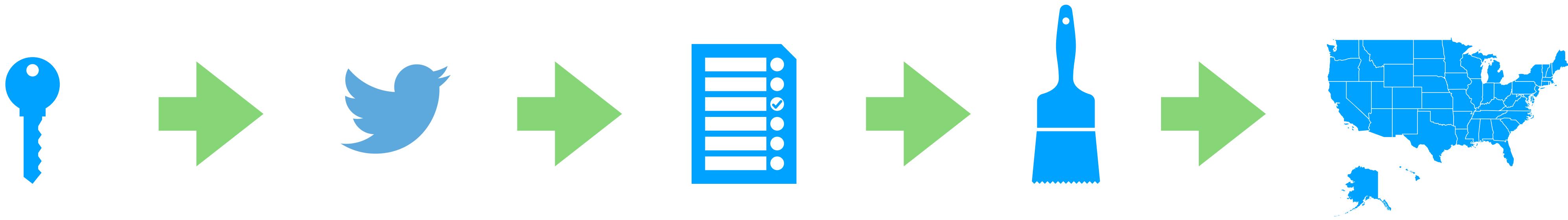
Objective

Where were Twitter hashtags for these diets trending in the US, within the last 30 days?



Keto
Paleo
Weight Watchers
Whole30

Process



Tools

- Twitter API
- Libraries: Pandas, NumPy, Matplotlib, Requests, JSON, pprint, Tweepy, Geocoder, gmaps
- Export/Read CSVs to prevent maxing out API requests
- Google maps to create the heat map
- Tableau was utilized to create bar graphs and overall activity map

Roadblocks

- Maxed out API requests for our API key
- Locations were inaccurate (ex: State,Country instead of City,State, and “from the sports grille”)
- Trying to grab the latitude longitude coordinates by using the Google Maps API, but then found the Geocoder library
- Number of locations and weights had to match to build heat map
- Only allowed us to pull 100 results, and only 70-75 results per diet had location values

Snippets

For loop to grab results from the Twitter API, and collect the correct location data. The geocoder library was extremely useful in pulling the lat/lng.

```
for tweet in r.json()["results"]:

    try:
        #Get and print location of tweet
        user_name = tweet["user"]["screen_name"]
        location = tweet["user"]["location"]
        tweet_location = tweet["place"]["full_name"]
        lat, lng = geocoder.arcgis(tweet_location).latlng
        print(f"The user, {user_name}, is located in {tweet_location}, {lat}, {lng}")

        #Build dictionary
        keto_location_dict = {
            "user": user_name,
            "tweet_location": tweet_location,
            "lat": lat,
            "lng": lng
        }

        # Add dictionary to list
        keto_location_list.append(keto_location_dict)

        # Increment count
        count = count + 1
    except:
        continue
```

Snippets

Aligning the length of weights to the length of geo coordinates

```
# Groupby state to get the average lat lng as the location
keto_map_locations = keto_data_US_extract.groupby("State").mean()
keto_map_locations
```

Snippets

Cleaning the location data, as not all users listed their location in the same format

```
# Update columns with city and state values as state and USA to blank columns
keto_data_US.iloc[9,5] = ""
keto_data_US.iloc[9,6] = ""
keto_data_US.iloc[15,5] = ""
keto_data_US.iloc[15,6] = ""
keto_data_US.iloc[20,5] = ""
keto_data_US.iloc[20,6] = ""
keto_data_US.iloc[29,5] = ""
keto_data_US.iloc[29,6] = ""
keto_data_US.iloc[35,5] = ""
keto_data_US.iloc[35,6] = ""
keto_data_US.iloc[39,5] = ""
keto_data_US.iloc[39,6] = ""
keto_data_US.iloc[42,5] = ""
keto_data_US.iloc[42,6] = ""
keto_data_US.iloc[43,5] = ""
keto_data_US.iloc[43,6] = ""
keto_data_US.iloc[48,5] = ""
keto_data_US.iloc[48,6] = ""
keto_data_US.iloc[49,5] = ""
keto_data_US.iloc[49,6] = ""
keto_data_US.iloc[59,5] = ""
keto_data_US.iloc[59,6] = ""
keto_data_US.iloc[70,5] = ""
keto_data_US.iloc[70,6] = ""

# Replace all blanks in the state column with correct state
keto_data_US.iloc[9,6] = " PA"
keto_data_US.iloc[15,6] = " AZ"
keto_data_US.iloc[20,6] = " AZ"
keto_data_US.iloc[29,6] = " AL"
keto_data_US.iloc[35,6] = " AL"
keto_data_US.iloc[39,6] = " OH"
keto_data_US.iloc[42,6] = " WV"
keto_data_US.iloc[43,6] = " AL"
keto_data_US.iloc[48,6] = " NY"
keto_data_US.iloc[49,6] = " WA"
keto_data_US.iloc[59,6] = " NY"
keto_data_US.iloc[70,6] = " TX"
```

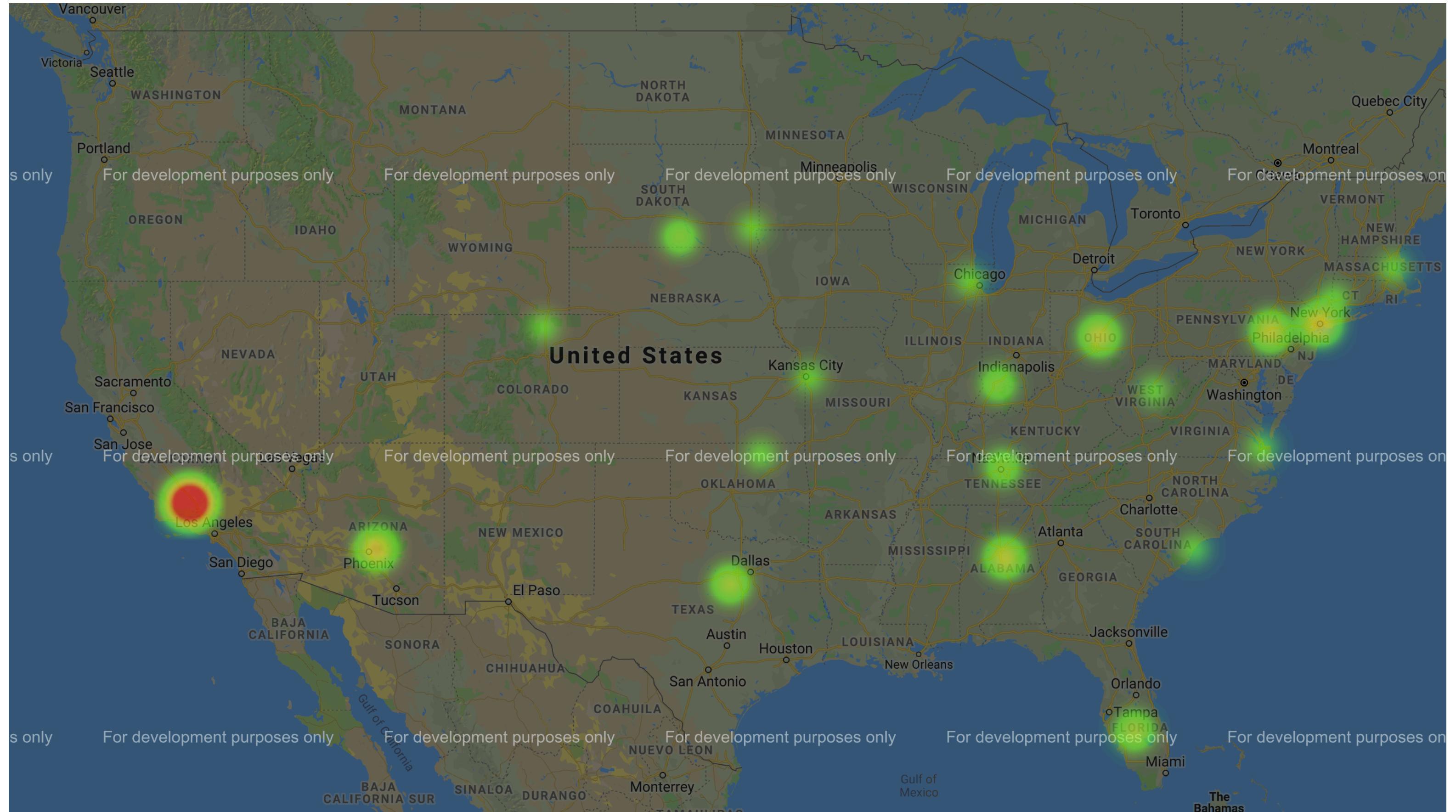
Snippets

Sorting the tweets per state and aligning with the geo coordinates for that state

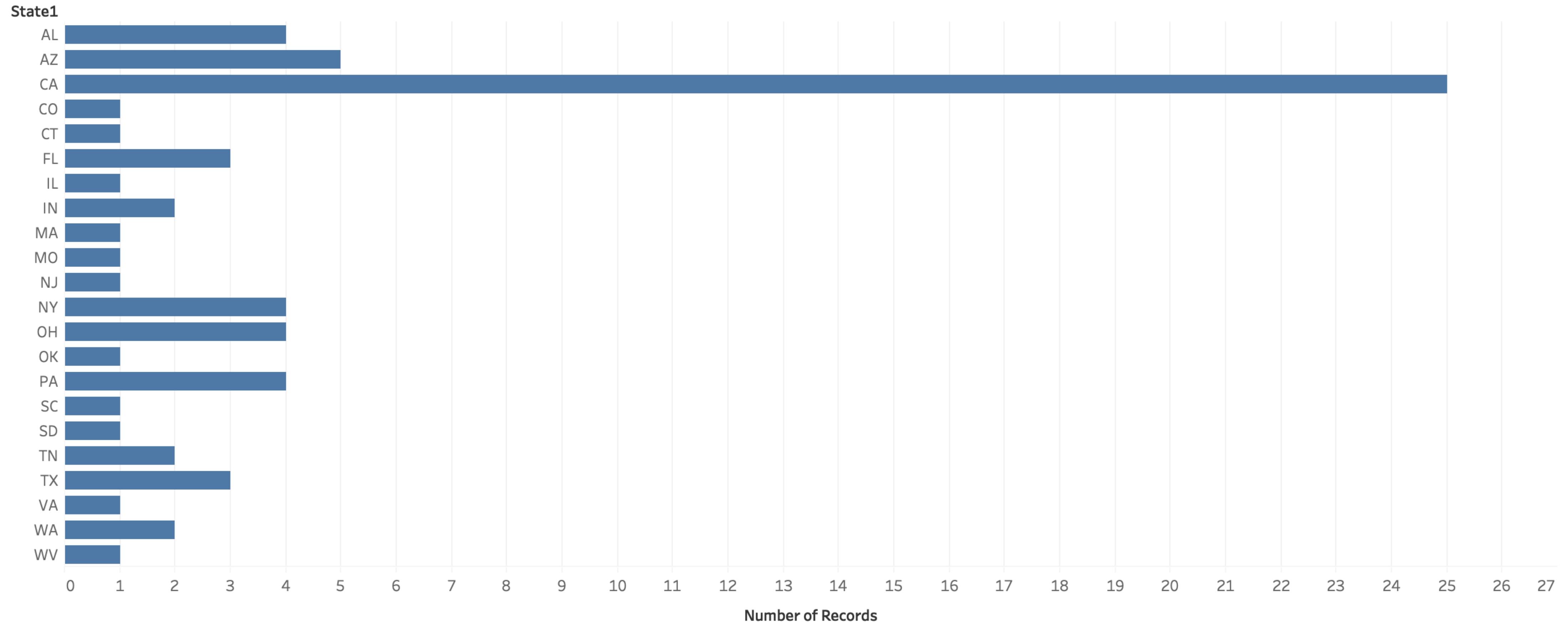
```
# Count the locations (number of states) and use as weights for heatmap
keto_map_locations_count = keto_data_US[ "State" ].value_counts()
reorg_keto_location_count = keto_map_locations_count.sort_index()
reorg_keto_location_count
```

Keto

Heat map of
tweets
with Keto hashtags

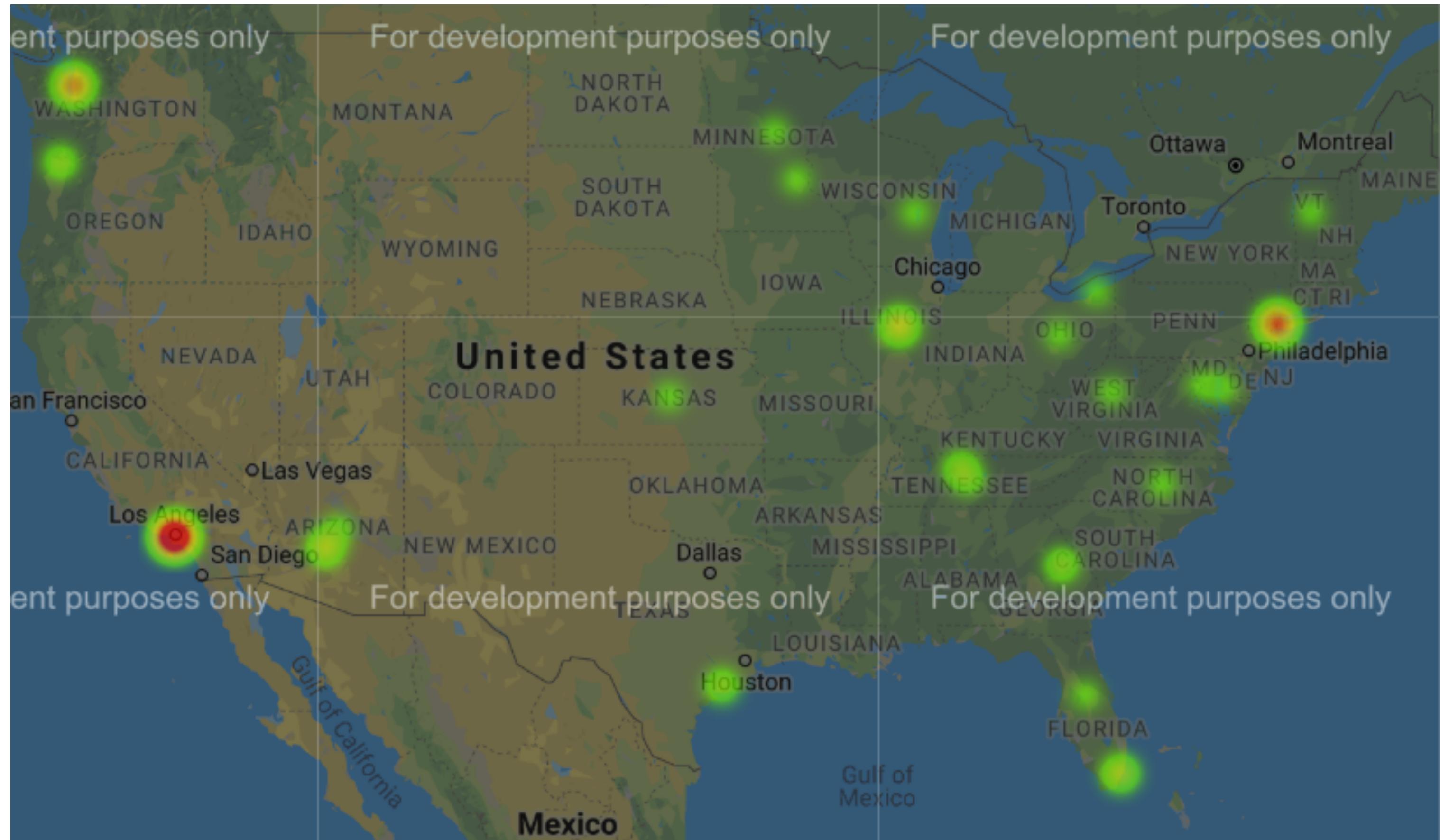


Keto

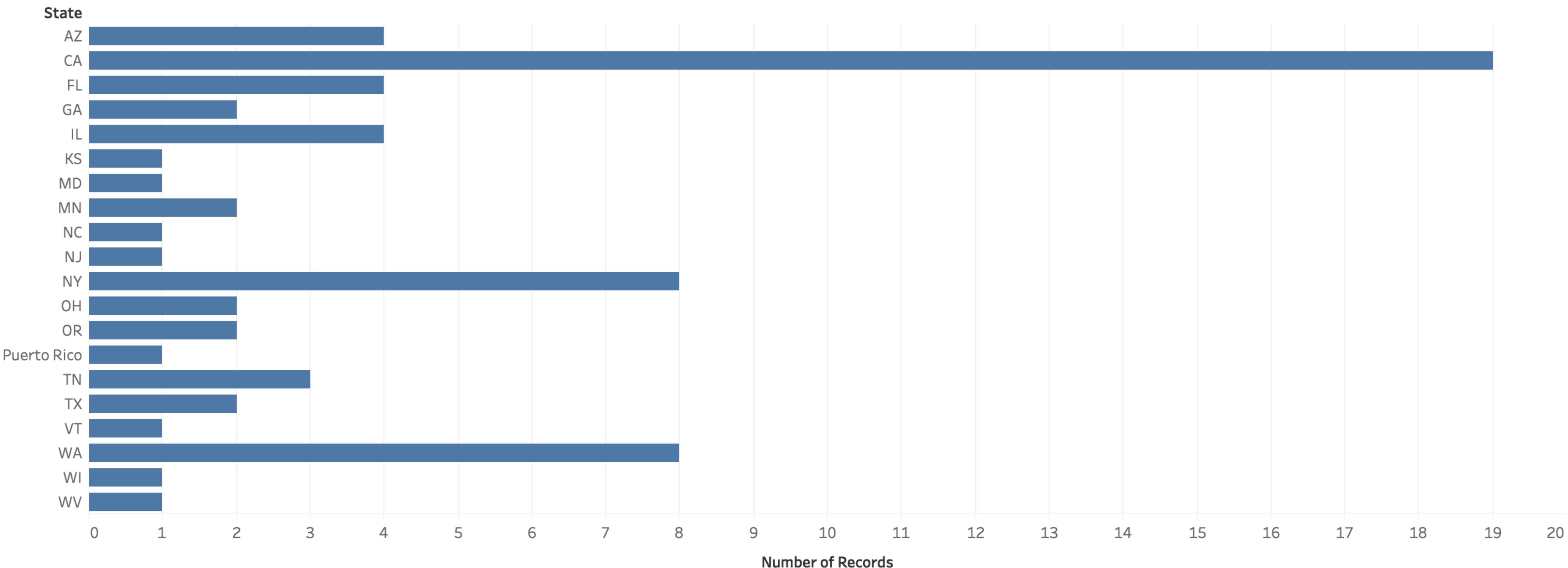


Paleo

Heat map of tweets
with Paleo
hashtags

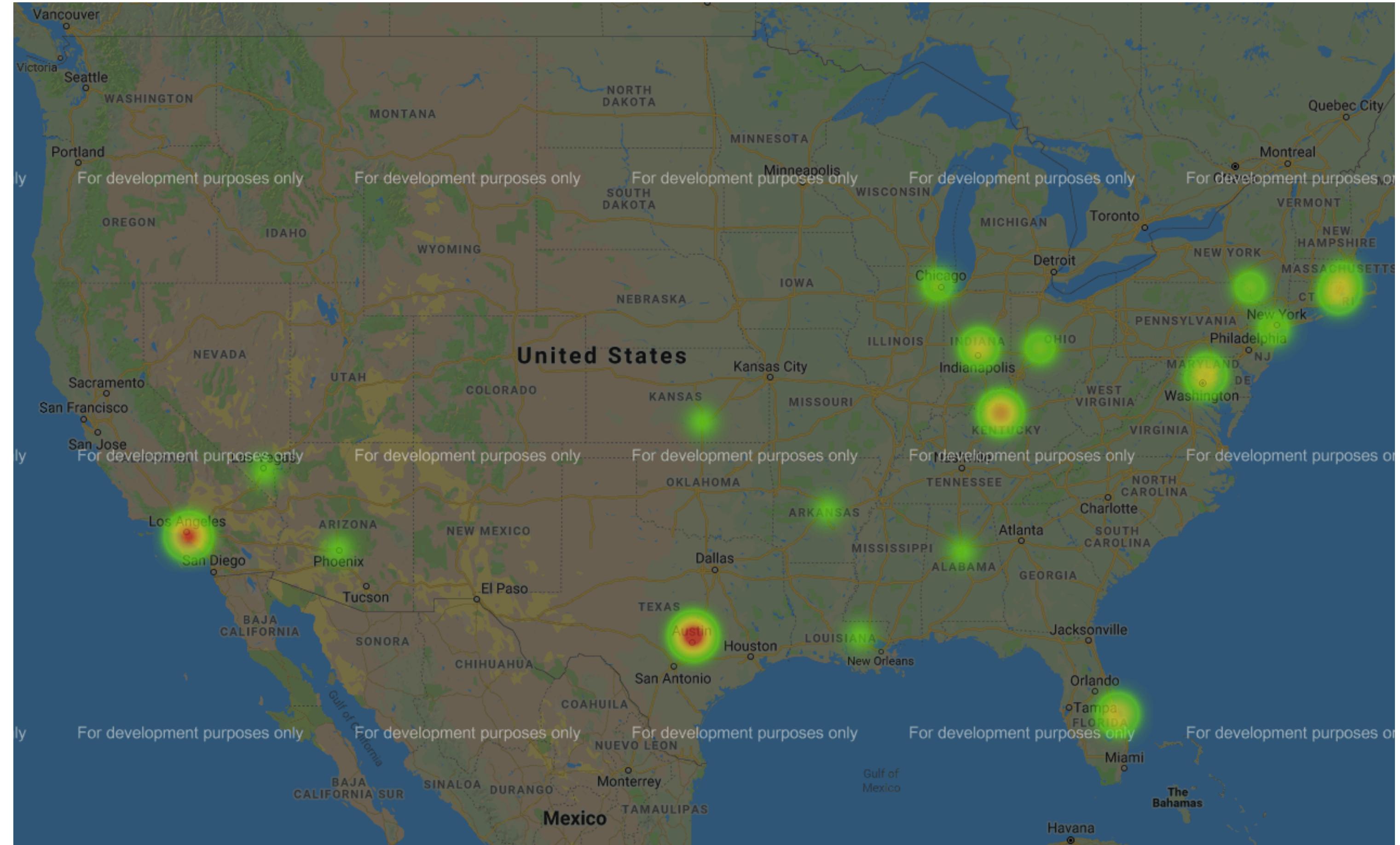


Paleo

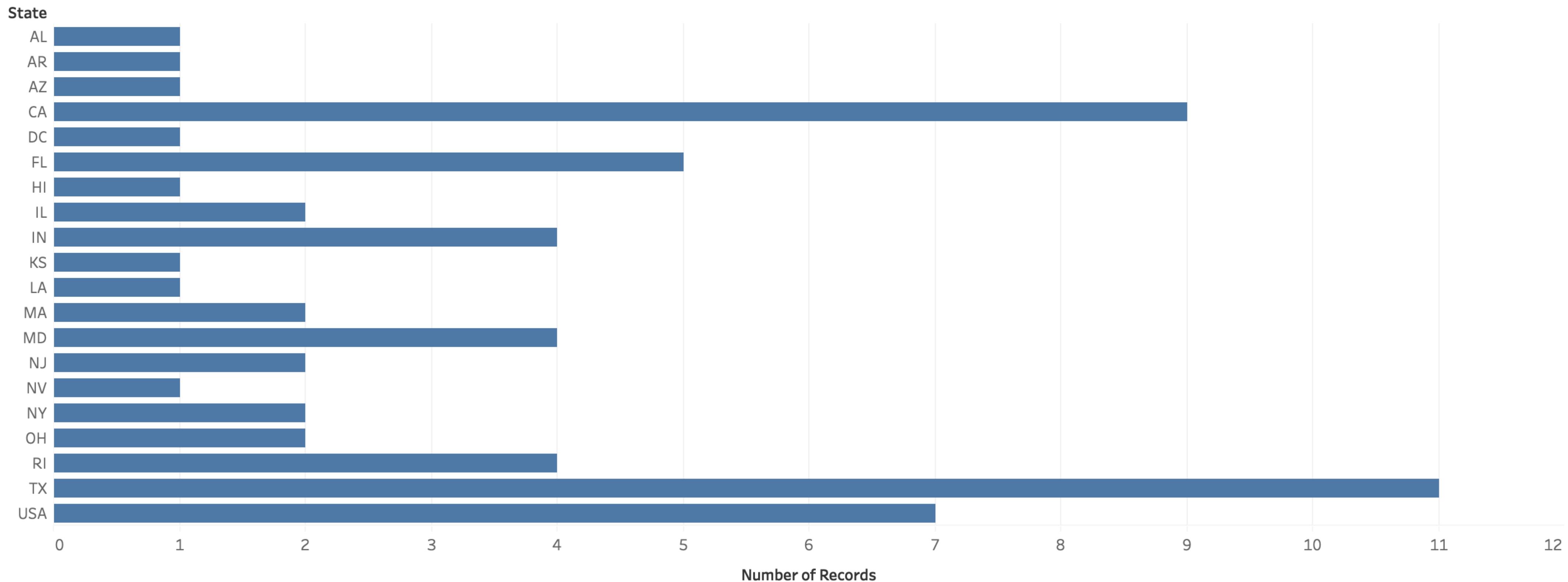


Weight Watchers

Heat map of tweets
with Weight
Watchers hashtags

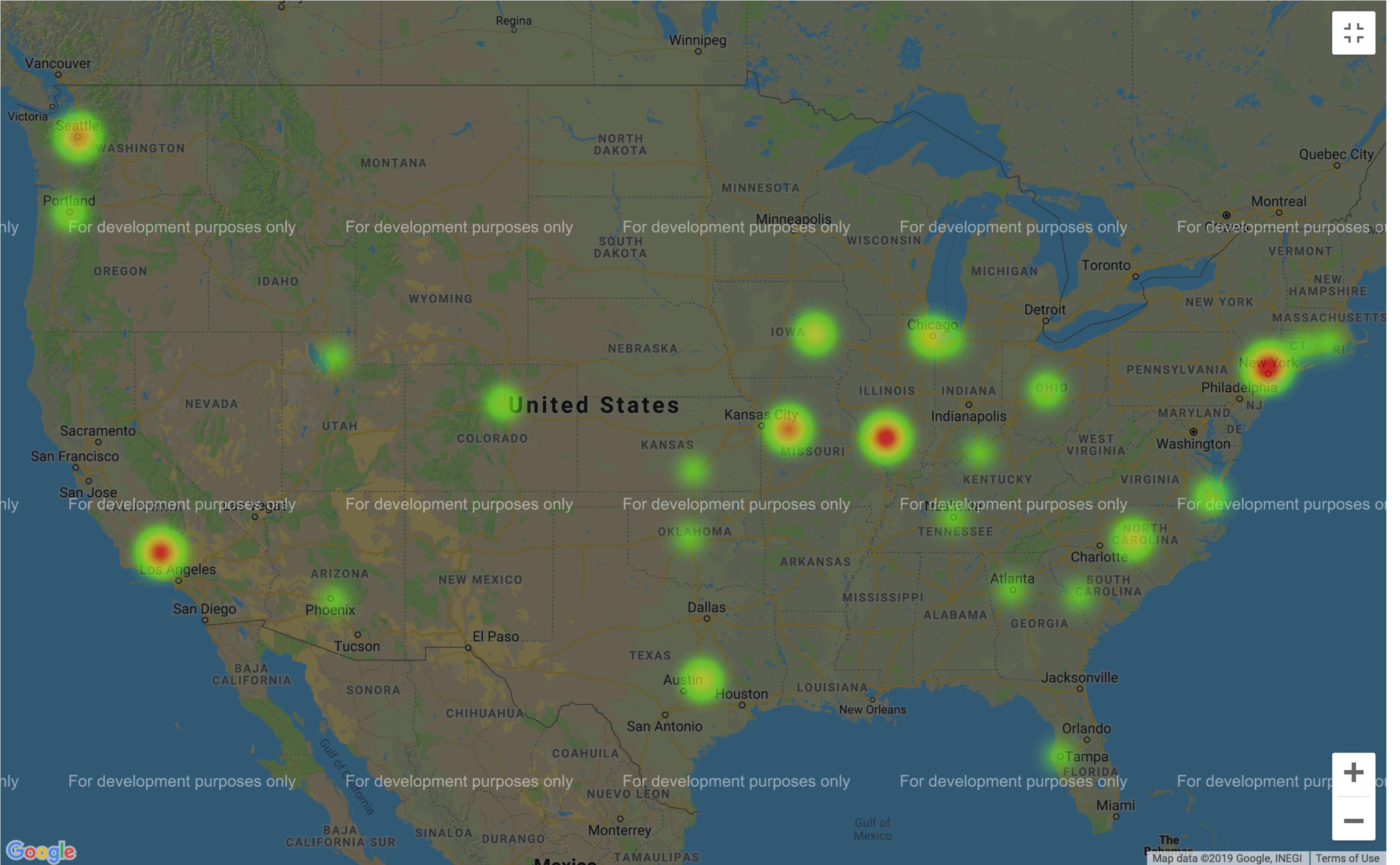


Weight Watchers

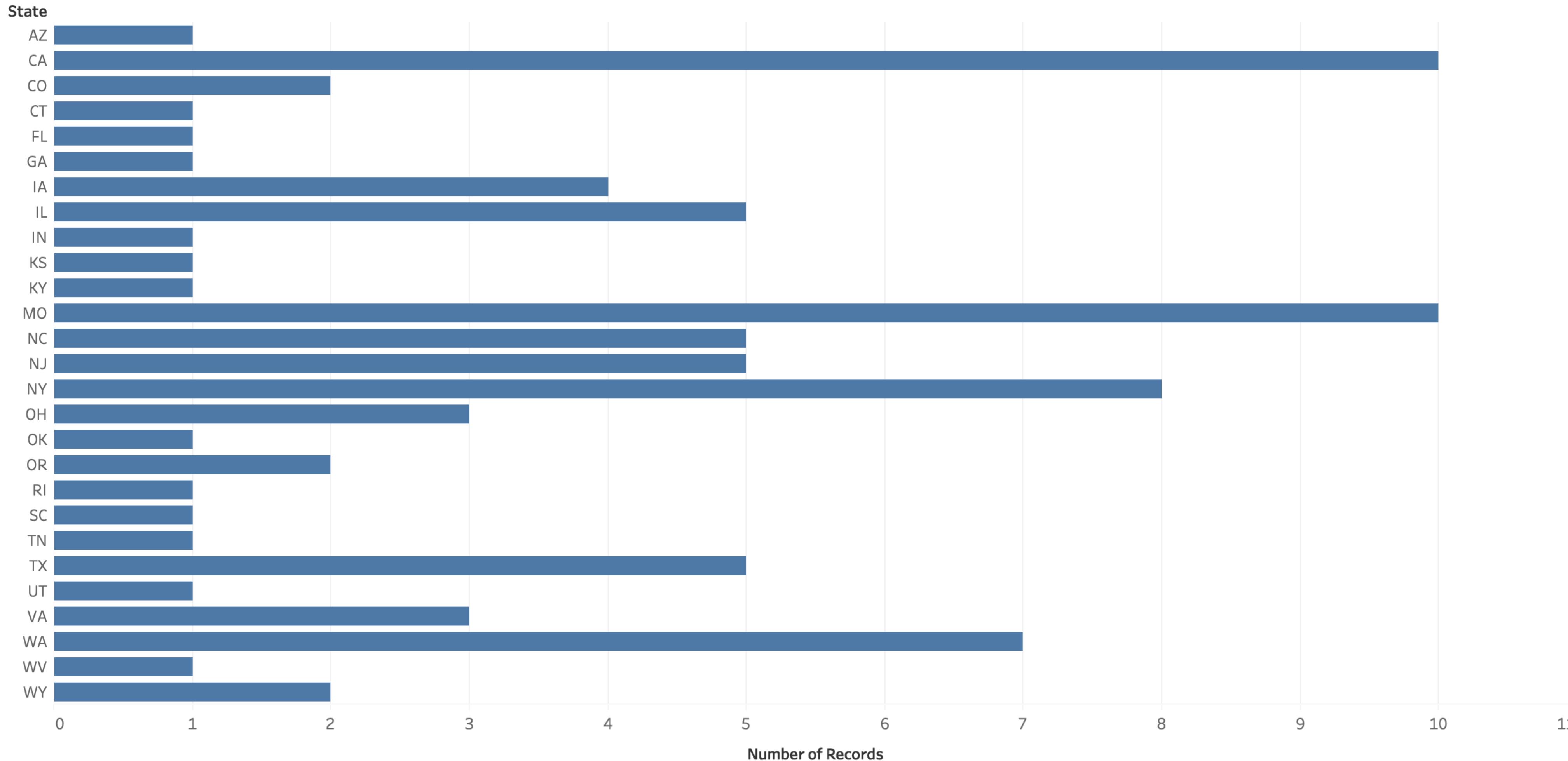


Whole30

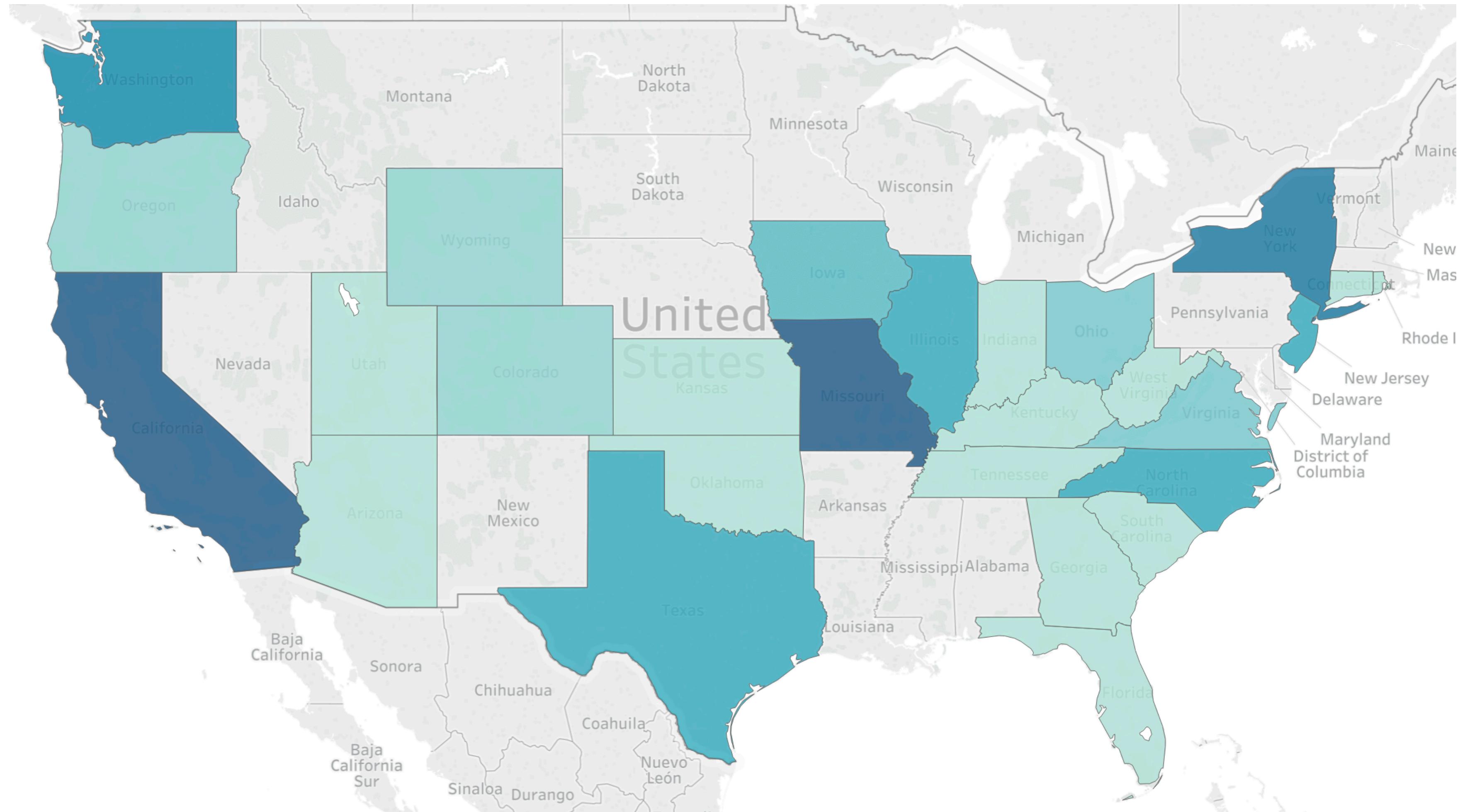
Heat map of
tweets
with Whole30
hashtags



Whole30



Overall Activity



Conclusion

Using Twitter to gather data on tweets based on location in regards to the 4 types of diets fads that were trending, we found:

- Weight Watchers was heavily trending in TX and CA. This could be because these two states are highly populated compared to the other 48 states.
- Whole30 was trending higher in CA and MO. It seems like MO is an outlier given its small population.
- Keto was trending highest in CA. When plotting by cities in CA, we noticed that SF and LA had a cluster of hits. These metros are more tech-savvy as compared to other cities in CA, which likely lead to more Twitter activity.
- Paleo was heavily trending in CA, NY, and WA. In this case, WA seems like an outlier against the large populations of CA and NY.

