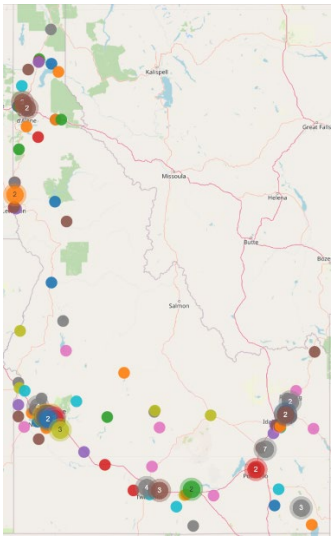
We are exploring the data at the Census Block Group level. This allows us to use as much of the data as possible, both from Safegraph and the Census data. We are helping with the decision of where a new upcoming ice cream shop would go, we can help them get an estimate of foot traffic from other ice cream businesses. So, we are predicting the foot traffic at a given location and using that as our target.



Data Wrangling:

We are focusing on ice cream shops in Idaho, so we all started with the following code to filter down the places table to get the following:

Using filter statements, we got to our starting table of just locations that sell ice cream at the CBG level.

Table +

	placekey	poi_cbg	location_name
1	zzy-229@5w9-cxy-snq	null	Canyon Creamery
2	228-222@5w9-jbw-pgk	160010001001	Fanci Freez
3	zzw-22j@5w9-jbw-rp9	160010001002	Stella's Ice Cream
4	225-222@5w9-jc4-6p9	160010001003	KIWI Shake & Bake
5	224-222@5w9-jc3-5s5	160010002024	Summer Sno
6	227-222@5w9-hvv-rff	160010003022	Baskin Robbins
7	zzw-224@5w9-hvx-rc5	160010003041	Yogurt Court

Landon Davis Feature: Average Distance from Home Traveled

Code Snippet: Creation of the Feature

```
# First, create a new dataframe that only includes the poi_cbg and distance_from_home columns
distance_df = frozen_patterns.select('poi_cbg', 'distance_from_home')

# Next, group the data by the poi_cbg column and compute the average distance from home
distance_features = distance_df.groupBy('poi_cbg').agg(F.avg('distance_from_home').alias('avg_distance_from_home'))

# Convert the avg_distance_from_home column from meters to miles
distance_features = distance_features.withColumn('avg_distance_from_home_miles', F.round(F.coalesce(distance_features['avg_distance_from_home'], F.lit(0)) / 1609.344, 2))
# Display new table
display(distance_features.select("poi_cbg", "avg_distance_from_home_miles"))
# Finally, join the distance_features dataframe with the original dataframe on the poi_cbg column
result = frozen_patterns.join(distance_features, on='poi_cbg', how='left')
```

Table: Average Distance Traveled to an Ice Cream Shop in a given CBG

	poi_cbg	avg_distance_from_home_miles
1	160550019002	24.31
2	160399603002	1.45
3	160010010002	5.28
4	160199706022	4.1
5	160010103311	2.94
6	160830007002	3.55
7	160550006022	3.8

Chart: Comparing each cbg to the amount traveled to go to an ice cream shop(grouped by cbg)