Establishing a Bicycle-based Bakery in Portland

Coursera Capstone Final Assignment by
Dustin Knudsen

Introduction

- Portland is home to many bicycle lovers
- The city itself is home to many routes, bicycle lanes, and bicycle shops
- By 2030, the city hopes to have over a quarter of all trips made by bike
- It’s also home to many health-conscious individuals who can be selective about what they eat.
- Menus at restaurants and other food and drink venues often cater to the healthy eating lifestyles of Portland’s inhabitants.

Problem

- Anne wants to start a gluten-free, lactose-free, organic bakery in Portland that uses a bicycle delivery system as part of its workforce
- Nearly 3,000 bicycle thefts are reported each year to the police, and this doesn’t include the larger number of unreported thefts.
- Anne wants her business in a neighborhood with low crime, close proximity to a bike shop, and fewer competitors

Data

- https://www.portlandoregon.gov/police/71978 - Portland Police Bureau 2018 Crime Data (CSV)
- https://gis-pdx.opendata.arcgis.com/datasets -PortlandMaps - Open Data for Neighborhoods and Bicycle Data Points (CSV, GeoJSON)

Methodology

Libraries

- NumPy - Handles Data in a Vectorized Manner
- Pandas - Library for Data Analysis
- Requests - Handles requests to API
- Json - Handles JSON files
- Geocoders - Uses Nominatim to convert address into latitude and longitude
- Folium - Map Rendering Library

Methodology (cont'd)

First used the Nominatim method, which is part of the Geocoder library to retrieve coordinates for the city of Portland

Retrieve geographical coordinates for the city of Portland

```
address = 'Portland'
geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude

print('The geograpical coordinate of Portland are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of Portland are 45.5202471, -122.6741949.
```

Methodology (cont'd)

Then created a dataframe based on the crime data. Afterwards, dropping unnecessary rows and making sure the columns were of type 'string.'

Create dataframe based on crime data. Drop unnecessary rows. Ensure columns are type 'string'

```
df = pd.read_csv('crimedata018.csv')
df = df.dropna()
df = df.dropna(axis=0, how='any')
df.columns=list(map(str, df.columns))
```

Methodology (cont'd)

Created a separate data frame which showed how many crimes were in each neighborhood

Use value_counts to get the number of crimes for each neighborhood

```
nList = df['Neighborhood'].value_counts().index
cList = df['Neighborhood'].value_counts().values
data = {'Neighborhood': nList, 'Count': cList}
zf = pd.DataFrame(data)
zf = zf.dropna()
zf
```

	Count	Neighborhood
0	3127	Hazelwood
1	2751	Downtown
2	1873	Lents
3	1786	Powellhurst-Gilbert

Methodology (cont'd)

Merged the new data frame with original data to create crime-count relationship

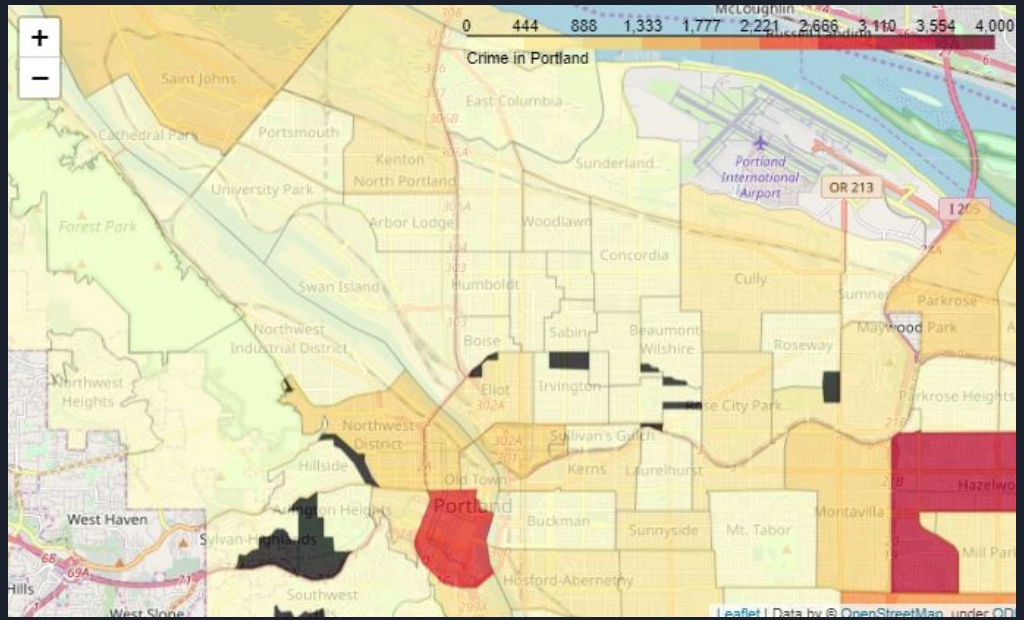
Merge original dataframe with new counts for each neighborhood

```
mergedDF = pd.merge(zf, df, how ='outer', on ='Neighborhood')
mergedDF.head()
```

	Count	Neighborhood	Address	CaseNumber	CrimeAgainst	OccurDate	OccurTime	OffenseCategory	OffenseTy
0	3127	Hazelwood	100 BLOCK OF NE 97TH AVE	18-223178	Property	7/2/2018	2030	Larceny Offenses	Theft From Motor Vehicle
1	3127	Hazelwood	100 BLOCK OF NE 99TH AVE	18-76681	Property	3/5/2018	1100	Fraud Offenses	Identity Th

Methodology (cont'd)

Create a choropleth map based on the amount of crime exists in each Portland neighborhood



Methodology (cont'd)

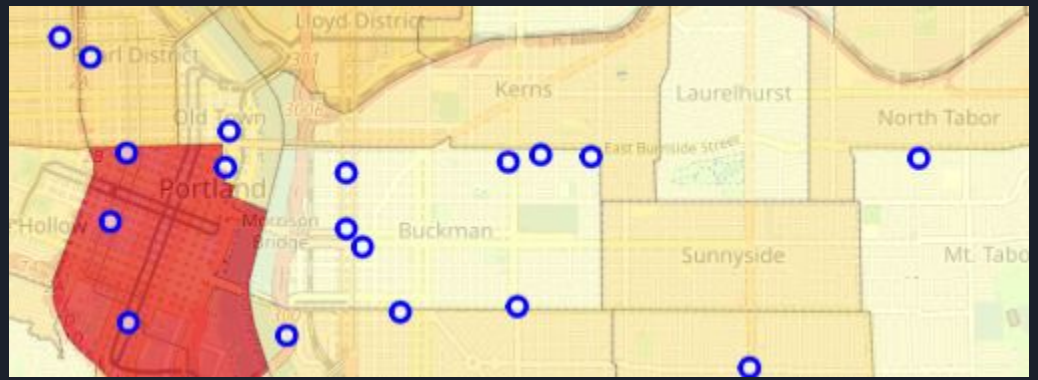
Acquired bike shop data from the csv file and made a data frame with it

```
bicycleData = pd.read_csv(r'Recommended_Bicycle_Route_Points.csv')

bikeDF = bicycleData.loc[bicycleData['Type']=='BIKESHOP']

bikeDF
```

152	-122.648927	45.473944	153	TP08-0000633	BIKESHOP	Gallery - Westmoreland	90.0	http://www.bikegallery.co pg6...
153	-122.573519	45.564961	154	TP08-0000634	BIKESHOP	Bike Tires Direct	90.0	http://www.biketiresdirect
				TP08-		Bikes For		

Methodology (cont'd)

Superimpose bike shop data points on to choropleth map to show where each bike shop is in relationship to the neighborhoods and their respective crime levels



Methodology (cont'd)

Use Foursquare API to generate data frame of venues near each bike shop

```
bike_venues = getNearbyVenues(names=bf['Neighborhood'],
                              latitudes=bf['Latitude'],
                              longitudes=bf['Longitude']
                             )
bike_venues
```

96	Brooklyn	45.500505	-122.654548	Genies Cafe	45.504973	-122.654345	Breakfast Spot
97	Brooklyn	45.500505	-122.654548	Brooklyn Park Pub	45.498094	-122.653834	Bar
98	Brooklyn	45.500505	-122.654548	Brooklyn Park	45.498512	-122.655367	Park
99	Brooklyn	45.500505	-122.654548	Taqueria Los Gorditos	45.504620	-122.653542	Mexican Restaurant

Methodology (cont'd)

Create a list of conditions to suggest which shops are competitors and make a dataframe out of that information

```
competitors = [

    'Food & Drink Shop',
    'Café',
    'Bakery',
    'Breakfast Spot',
    'Coffee Shop',
    'Donut Shop',
    'Dessert Shop',
    'Chocolate Shop'
]
```

```
competitorDF = pd.DataFrame(columns = bike_venues.columns)
for comp in competitors:
    x = bike_venues.loc[bike_venues['Venue Category']==comp]
    competitorDF = pd.concat([competitorDF, x])
```

Methodology (cont'd)

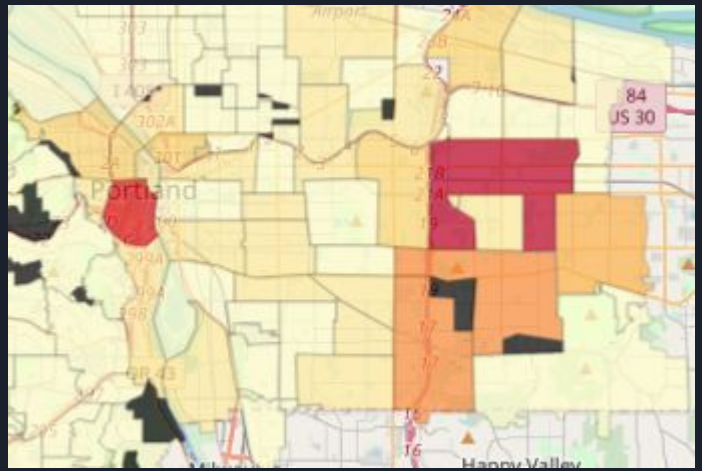
Using this information, Anne can decide which neighborhood is best to start her business!

```
competitorDF.sort_values(['Neighborhood'])
```

322	Buckman	45.521226	-122.661435	F & B Cafe	45.524005	-122.662990	Café
333	Buckman	45.521983	-122.646118	See See Motorcycle	45.524603	-122.648847	Coffee Shop
35	Burlingame	45.463662	-122.684853	Sol Station	45.462972	-122.684754	Coffee Shop
113	Mt Tabor	45.522232	-122.606919	Starbucks	45.522549	-122.606596	Coffee Shop
2	Multnomah	45.467571	-122.714317	John's Marketplace	45.467385	-122.713203	Food & Drink Shop
17	Multnomah	45.467571	-122.714317	Village Coffee	45.468267	-122.712522	Coffee Shop

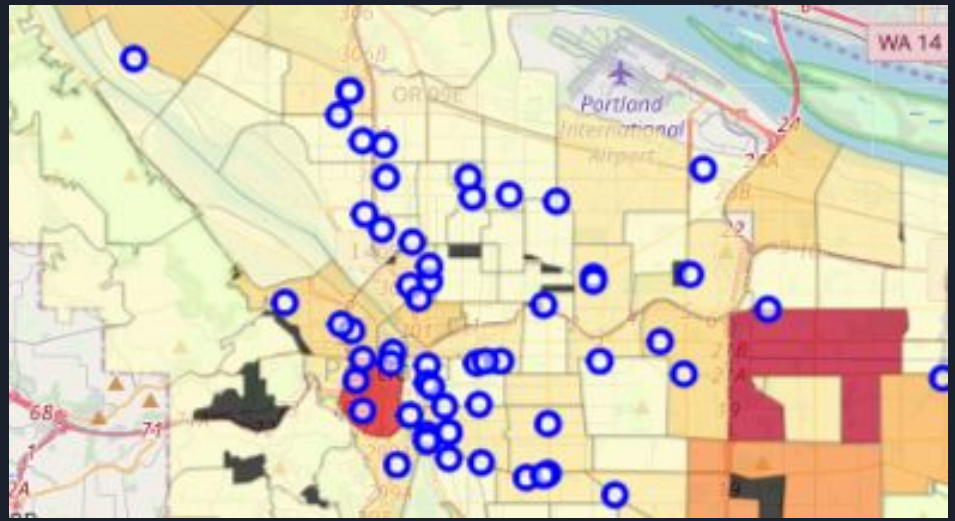
Results

- East Portland has more crime than the rest of Portland
- Downtown is especially heavy in crime



Results (cont'd)

- West Portland had very few bike shops
- Shops clustered around downtown and also up I-5 freeway



Discussion

- Gresham lies east of Portland. It’s known for high crime.
- Downtown has the highest concentration of bicycle shops
- Bicycle theft may be higher in downtown areas
- West Portland may have lower population and/or higher income suburbanites
- Shops along I-5 may indicate a main bike travel artery

Conclusion

- Anne chose Mount Tabor to build her business due to less crime, proximity to a bike shop, and fewer competitors.
- Buckman had a large number of bike shops, but also many competitors
- Downtown may still be more suitable for her business due to bicycling accomodations
- Population density in residential vs business should be considered