Aryan Chaudhary - BAC Insight-Advanced Team Application

September 28, 2023

1 BAC Insight/Advanced Team Application

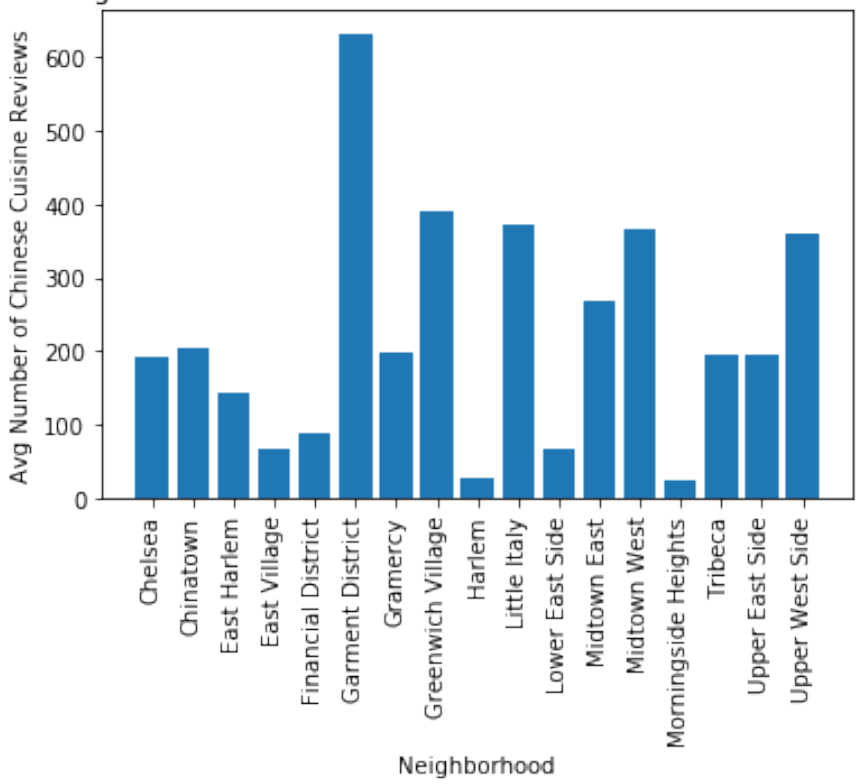
1.0.1 Question 4

One day, you decide to open a new Chinese family diner somewhere in New York City. You have access to Yelp review ratings and health inspection data for more than 3,000 restaurants (here) as well as supplementary visualizations (here). Based on the data, where would you open your restaurant? What potential difficulties might you face? (max 250 words)

Please upload your response and any supporting visualizations or model results.

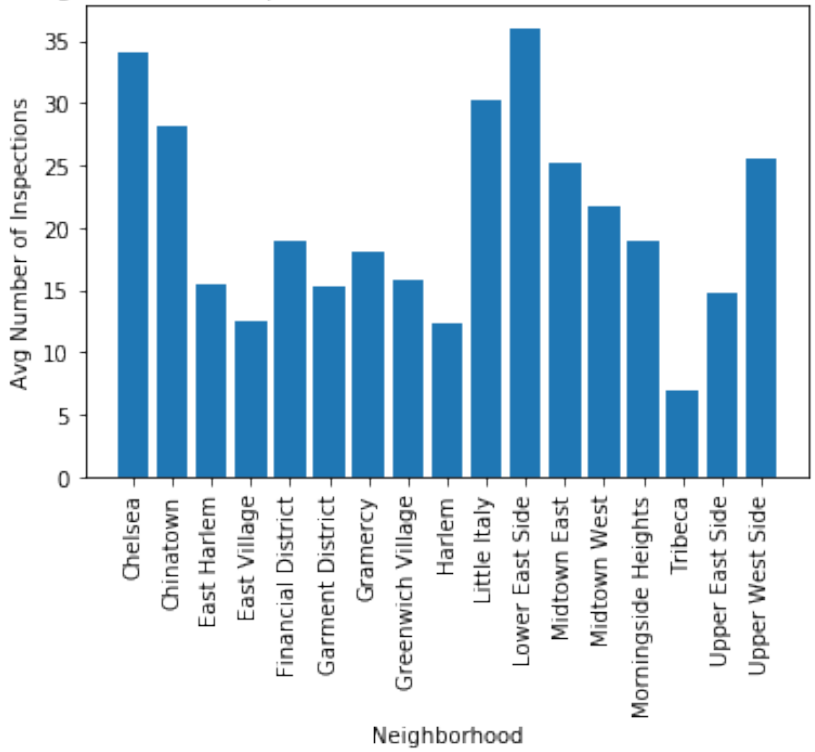
[29]:
```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('yelp.csv')
data_chinese_cuisine = data[data['CUISINE DESCRIPTION'] == 'Chinese']
average_review_count_by_neighborhood = data_chinese_cuisine.
 ↪groupby('neighborhood')['review_count'].mean().reset_index()
plt.figure(0)
plt.bar(average_review_count_by_neighborhood['neighborhood'],␣
 ↪average_review_count_by_neighborhood['review_count'])
plt.xticks(rotation = 90)
plt.xlabel('Neighborhood')
plt.ylabel('Avg Number of Chinese Cuisine Reviews')
plt.title('Figure 1: Avg Number of Reviews for Chinese Restaurants in Each␣
 ↪Neighborhood')
plt.show()
```

Figure 1: Avg Number of Reviews for Chinese Restaurants in Each Neighborhood

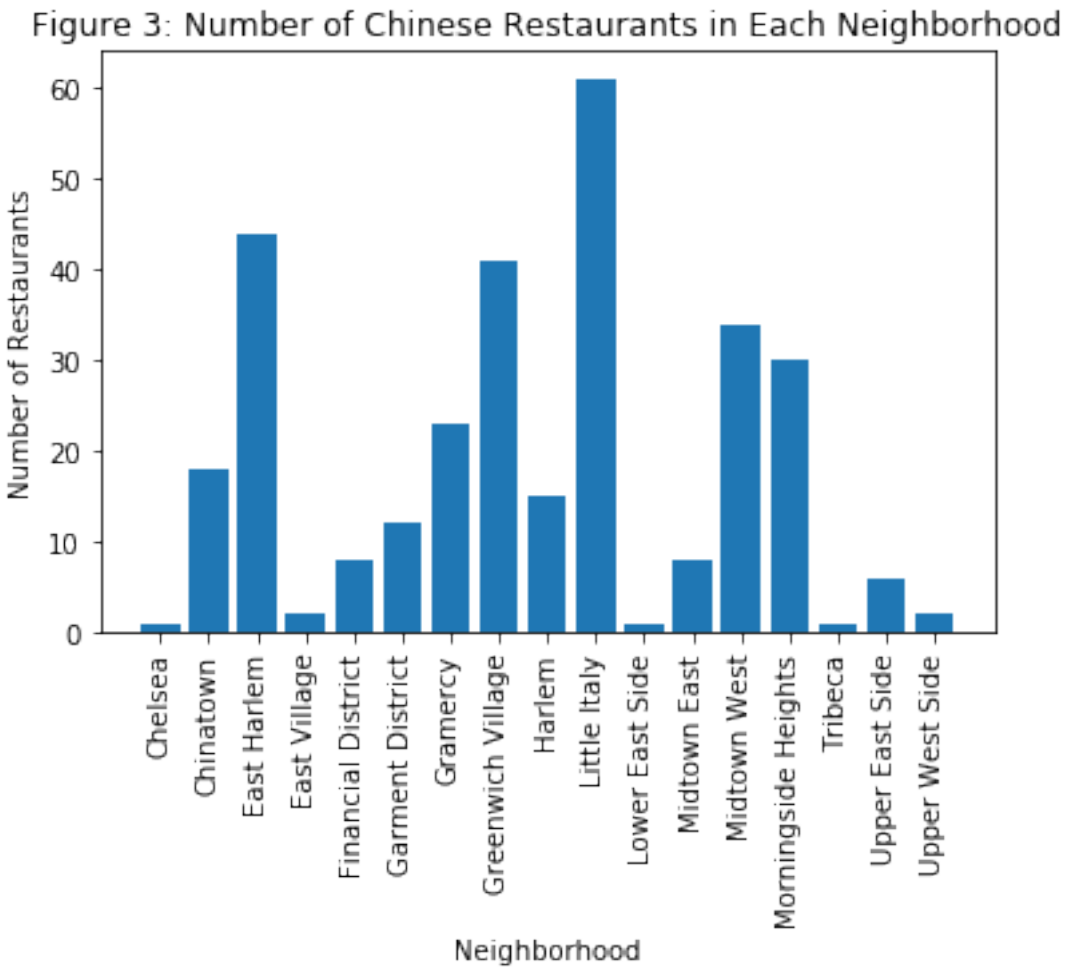


[28]:
```
average_number_of_inspections_by_neighborhood = data_chinese_cuisine.
  ↪groupby('neighborhood')['#_of_inspections'].mean().reset_index()
plt.figure(1)
plt.bar(average_number_of_inspections_by_neighborhood['neighborhood'],␣
  ↪average_number_of_inspections_by_neighborhood['#_of_inspections'])
plt.xticks(rotation = 90)
plt.xlabel('Neighborhood')
plt.ylabel('Avg Number of Inspections')
plt.title('Figure 2: Avg Number of Inspections for Chinese Restaurants in Each␣
  ↪Neighborhood')
plt.show()
```

Figure 2: Avg Number of Inspections for Chinese Restaurants in Each Neighborhood



[27]:
```
number_of_chinese_restaurants_by_neighborhood = data_chinese_cuisine.
  ↪groupby('neighborhood')['CUISINE DESCRIPTION'].count().reset_index()
plt.figure(2)
plt.bar(number_of_chinese_restaurants_by_neighborhood['neighborhood'],␣
  ↪number_of_chinese_restaurants_by_neighborhood['CUISINE DESCRIPTION'])
plt.xticks(rotation = 90)
plt.xlabel('Neighborhood')
plt.ylabel('Number of Restaurants')
plt.title('Figure 3: Number of Chinese Restaurants in Each Neighborhood')
plt.show()
```



Based on these figures and the data provided I would open my restaurant in the Garment District. Figure 1 shows that the Garment District has the largest average number of reviews between neighborhoods, indicating that Chinese restaurant-goers are the most active here. From Figure 2 I learned that the Garment District is not highly regulated although the industry as a whole is, and one usually needs to navigate a maze of permits, licenses, and health inspections while complying with labor laws. Opening in the Garment District would help avoid many of these time-consuming and expensive requirements, reducing operating costs while providing flexibility in practices. Finally, plotting the total Chinese restaurants in each neighborhood in Figure 3 gave me a glimpse into the market competition between them. A relatively low number of establishments in the Garment District makes it easier to gain market entry and establish a positive reputation in the community when opening a restaurant. Potential difficulties include rent because of real estate with prime location and commercial demand. Specialized infrastructure and amenities tailored to the needs of fashion businesses, such as specialized showrooms, sample-making facilities, and easy access to fabric suppliers justify higher rent costs for industry businesses. Seasonal variations in foot traffic due to the various fashion events the area hosts then also pose an issue. This industry specialization also creates staffing issues as recruiting skilled kitchen and front-of-house staff could be difficult in a competitive labor market diluted by the concentration of fashion-related businesses.