

## 1. Data Preparation and Filtering:

Using **pandas**, I carefully filtered the dataset to focus on specific locations and hotel chains of interest. I targeted five major cities using state and province keywords: **'NV'** (Las Vegas, Nevada), **'ON'** (Toronto, Ontario, Canada), **'AZ'** (Phoenix, Arizona), **'NC'** (Charlotte, North Carolina), and **'PA'** (Pittsburgh, Pennsylvania).

I also filtered hotel chains based on key brands such as **Marriott, Best Western, Hilton, Holiday Inn, Hyatt, Courtyard, Radisson, InterContinental, Sheraton, Fairfield Inn, Wyndham, Comfort Inn, La Quinta, Quality Inn, Econo Lodge, Days Inn, DoubleTree, Embassy Suites, Hampton Inn, and Extended Stay America.**

Additionally, I used a comprehensive list of keywords to isolate reviews that potentially indicated signs of **discrimination, bias, or unfair treatment**, including terms like *"racist"*, *"sexist"*, *"homophobic"*, *"discriminate"*, *"no wheelchair access"*, *"slur"*, *"rude"*, *"disrespectful"*, *"ignored"*, and more. (since those are the words commonly used in Yelp reviews that I have seen)

## 2. Data Output:

After cleaning and filtering, I generated structured CSV files to organize the focused dataset for further analysis.

## 3. Quality Control:

To ensure that the negative reviews truly reflected dissatisfaction, I excluded reviews with high star ratings (specifically, any reviews rated **4 stars or higher** were removed).

## 3. Categorization and Counting:

I categorized the filtered data into *positive* and *negative* sentiment based on review content and star ratings. I counted the **total number of reviews** in each category for both hotels and cities.

## 4. Negative Review Analysis:

I specifically calculated the **number and percentage of negative reviews** for each hotel chain and each city to understand where signs of discrimination or dissatisfaction were most prominent.

## 5. Visualization:

Finally, I created clear, visual summaries of the findings using tools like **Excel** and **Power BI**, making it easy to compare negative review rates across different hotels and cities.

