NYC Airbnb 2021 Data Insights and Analysis

**Visualization #1 (Choropleth Map of NYC Neighborhoods):** In this interactive choropleth map showcasing NYC's Airbnb landscape, we've utilized GeoPandas and Shapely points to pinpoint the precise locations of listings within the dataset. By organizing listings according to neighborhood, we offer insightful average prices and review scores for each area. Employing convex hull polygons, we outline the geographic boundaries of these neighborhoods, enhancing them with important details such as average price and review score. The color scheme intuitively represents the relative average price of each neighborhood, enabling users to swiftly identify areas within their budget. Hovering over the polygons reveals the names of neighborhoods, while clicking provides comprehensive averages for price and review scores. This interactive feature allows potential customers to identify neighborhoods that fit their budget and review score standards, facilitating informed decision-making.

**Visualization #2 (Bar Chart: Average Price by Room Type and Neighborhood):** The bar chart, created using Matplotlib, illustrates the average prices of various room types across neighborhoods in New York City, providing customers with valuable insights for their accommodation choices. It underscores significant price discrepancies among neighborhoods and offers clarity on the most popular room types in each area. Each bar represents the average price of a specific room type within a neighborhood, serving as visual cues that highlight price disparities among neighborhoods. The x-axis indicates different neighborhoods, organizing the data spatially, while the y-axis represents average prices, facilitating comparisons across room types and neighborhoods. This visual representation enables customers to easily pinpoint budget-friendly options and grasp market trends, empowering them to make well-informed decisions tailored to their preferences and budget constraints during their stay in New York City.

**Visualization #3 (Heatmap: Distribution of Room Types Across NYC Neighborhoods):** The heatmap offers guests valuable insights into the diverse range of room types available across New York City's neighborhoods. Each cell in the heatmap represents a combination of room type and neighborhood, with the presence of a mark indicating the count of listings with that specific room type in the corresponding neighborhood. The x-axis denotes different room types, while the y-axis represents neighborhood groups, offering spatial organization for identifying room type distribution. The color of each cell encodes the count of listings, with darker shades indicating higher counts and lighter shades representing lower counts, facilitating quick perception of room type frequencies. Numeric annotations within cells provide precise quantitative information, supplementing the visual encoding with exact counts of listings.

**Visualization #4 (Interactive Scatterplot: Price vs. Number of Reviews):** In the interactive scatter plot below, created using Altair, we illustrate the relationship between price and the number of reviews for a random sample of 5,000 listings. To enhance clarity, we capped the price at under $3,000, as most listings fell within the lower price range of $0 to $200. However, there were a few high-price outliers around $5,000, which could obscure the chart's readability. Each circle on the plot represents a listing, positioned based on its price and number of reviews. These circles are color-coded to denote different room types. This scatterplot enables customers to evaluate listing value and popularity by correlating price with review counts. Additionally, our interactive tool allows users to select their preferred room type, dynamically scaling the chart to display only listings of that type. This feature provides users with a clearer view of the general price and review distribution for a specific room type.

**Visualization #5 (Interactive Scatterplot: Various Review Score Attributes**): In our dataset, we found that Airbnb listings are evaluated on various sub attributes like location, cleanliness, and check-in process, which contribute to the overall review score. Our interactive scatter plot provides a comprehensive overview of overall review scores across NYC, allowing users to explore individual sub attributes. Positioned similarly to a geographical map, each circle represents a listing's location based on latitude and longitude. The color of each point corresponds to the displayed review score attribute, with darker shades indicating higher scores and lighter shades indicating lower scores on a scale of 1 to 5. This color scheme aids users in quickly identifying listings with lower scores amidst higher-scored listings.

Moreover, users can utilize the dropdown menu at the top of the plot to select their preferred sub attribute for analysis. This feature enables users to explore the distribution of review scores based on specific criteria, empowering them to make more informed decisions when choosing accommodations. Additionally, we've implemented a tooltip feature that allows users to hover over data points to access essential details about each listing. This includes the listing's name, price, room type, neighborhood, overall review score, and accommodation capacity.

**Link to Github Repository:** [**https://github.com/jonathanhng1117/DS4200Final**](https://github.com/jonathanhng1117/DS4200Final)

**Link to final webpage:** [**https://jonathanhng1117.github.io/DS4200Final/**](https://jonathanhng1117.github.io/DS4200Final/)