

M06 - Python for Data Analysis

● Live Lesson Challenge

The Airbnb NYC Adventure

A Journey through Data Exploration, Visualization, and Insights using Python

In this activity, you will apply the skills you've learned (data cleaning, exploration, grouping, and visualization) to analyze Airbnb listings in New York City. Use the dataset to uncover insights about prices, neighborhoods, room types, and reviews.

Database: [Airbnb NYC Listings Dataset from Kaggle](#)

Instructions:

1. Data Exploration

- Load the dataset into a Pandas DataFrame.
- Display the first 5 rows of the dataset.
- Check the number of rows and columns in the dataset.
- Identify missing values in the dataset.

2. Data Cleaning

- Drop rows with missing values in critical columns like **price** and **reviews_per_month**.
- Remove listings where the price is greater than 1000 to avoid outliers.

3. Analyze and Answer Key Questions Answer the following questions:

- What is the average price of listings in each **neighborhood group**?
- What are the **top 5 neighborhoods** with the most listings?
- What is the relationship between the number of **reviews per month** and **price**?
- What is the distribution of **room types**?

4. Data Visualization Create visualizations for the following:

- A bar plot showing the **average price by neighborhood group**.
- A bar plot showing the **top 5 neighborhoods** with the most listings.
- A count plot showing the **distribution of room types**.

- A scatterplot showing the relationship between **reviews per month** and **price**.
-