**Assignment 9**

**Puneet Kochar**

**University of Colorado Denver Business School**

* **Problem Statement**

New York City is one of the world's best-known cities. Each year it attracts millions of tourists which boosts our economy. Hence, NYC is one of Airbnb's hottest markets. New York City (NYC) has an active Airbnb market with over 48,000 rentals in the 2019 calendar year as of August. The management of the Airbnb wants to know the properties that are profitable in New York city and which properties of Airbnb needs to be focused on for the overall improvement. In order to answer these questions of management, I will focus on New York’s city Airbnb data and the goal of this project will be analyzing the data and finding answer for following questions:

* How are rental properties distributed among different neighborhood of New York?
* How do prices vary with respect to different neighborhood?
* What are different types of properties?
* Which properties are highly rated by the guest and which properties are poorly rated?
* Which areas have highly rated properties?
* Are the demand and prices of rental related?
* **Dataset**

The dataset for New York Airbnb listing is obtained from <https://www.kaggle.com/>. It is a detailed dataset with 16 columns which include host name, host id, Property name, Property type, Neighborhood, Latitude, Longitude location, Price of listing, number of reviews, average reviews per month etc. The data contains total of 48,895 listings.

In Order to preprocess data, we might need to perform a few imputations and transformations on our dataset. There is no significant inconsistencies or anomalies in the data, however most of the columns / features we were interested in do does not contain data in the correct format and therefore might need to be manipulated in a way that preserved their meanings. The data contains null values which needs to address after conducting detailed Exploratory Data Analysis.

* **Tools**

The tools that will be used in the projects will depend on different stages of project. For EDA analysis, I plan to use R and Excel. For data preprocessing and deep analysis, I plan to use Tableau and for creating dashboards and creating story for my analysis, I plan to use Tableau or PowerBI.