**IBM Data Science capstone project**

**Analyzing neighborhoods in Houston to start a private school**

**Abstract**

The purpose of this project is to analyze neighborhoods in Houston to find a best location to start a private school. Houston is a large metropolis in Texas with 2.3 million residents, is the fourth most populous city in the United States, trailing after New York, Los Angeles, and Chicago. The city is the largest in the South and the Southwest.

In this Machine Learning (ML) project I was able to present a clear picture by analyzing the number of population and private schools in the densely populated suburbs, and by visualizing the data points on Houston map to pin point a specific area to start a new private school with valuable reasoning.

This project will be highly valuable for new business owners and entrepreneurs. In this workflow, automated process for identifying best ML algorithm are used to find better recommendations for the stakeholders.

**Data overview**

* List of Houston Neighborhoods is obtained from https://en.wikipedia.org/wiki/List\_of\_Houston\_neighborhoods.
* The private school datasets used in this project is extracted from Kaggle which is originally obtained from the US Department of Homeland Security. It contains information about all public and private schools with attributes regarding their geographical distribution. <https://www.kaggle.com/andrewmvd/us-schools-dataset>.
* The geographical location required for this project is obtained from Foursquare API.