**CALIFORNIA HOUSING PRICES**

**GIYA BINOY**

**Student ID: 4354557**

**ABHISHEK KUTHUKALLINGAL SURENDRAN**

**Student ID: 4355412**

**Business Analytics – School of Business**

**ADMN5016-Applied Artificial Intelligence and Machine Learning**

**Prof.  Joseph Santarcangelo**

**July 03, 2023**

**DESCRIPTION OF THE PROJECT**

This is a dataset which contains the Median house prices for California districts derived from the 1990 census.

Based on the information from the 1990 census, the data is about the dwellings that may be in a certain California district and some basic statistics about them. The columns are as follows:

1. Median House Value: Median house value for households within a block
2. Median Income: Median income for households within a block of houses
3. Median age: Median age of a house within a block
4. Total rooms: Total number of rooms within a block
5. Total bedrooms: Total number of bedrooms within a block
6. Population: Total number of people residing within a block
7. Households: Total number of households, a group of people residing within a home unit, for a block
8. Latitude: A measure of how far north a house is; a higher value is farther north
9. Longitude: A measure of how far west a house is; a higher value is farther west
10. Distance to coast: Distance to the nearest coast point
11. Distance to Los Angeles: Distance to the centre of Los Angeles
12. Distance to San Diego: Distance to the centre of San Diego
13. Distance to San Jose: Distance to the centre of San Jose
14. Distance to San Francisco: Distance to the centre of San Francisco

**SUMMARY OF THE PROJECT**

Given the dataset containing information about houses in different California districts, including factors such as median house value, median income, median age, total rooms, total bedrooms, population, households, latitude, longitude, and distances to various cities, the problem is to develop a model that can accurately predict the median house value for a given block based on the available features.