PROJECT PROPOSAL

Student Name	DISSANAYAKE D.J.R H.E. SANDUNI SATHSARANI
Course	Data Science with Python
Supervisor Name	Pasindu Marasinghe
Title of Project	California Housing Price Prediction

Abstract of the project	**Problem Definition**
	
	The objective of this project is to develop a churn model for predicting housing price movements in California. The model will utilize various factors and characteristics related to properties and their surrounding areas to forecast whether the prices of houses are likely to increase or decrease. This prediction will assist potential buyers, sellers, and real estate investors in making informed decisions and understanding the market trends.
	
	Description of the dataset
	
	`#  Column  Non-Null Count  Dtype`
	`---  ------  --------------  -----`
	`0  longitude  20640 non-null  float64`
	`1  latitude  20640 non-null  float64`
	`2  housing_median_age  20640 non-null  int64`
	`3  total_rooms  20640 non-null  int64`
	`4  total_bedrooms  20433 non-null  float64`
	`5  population  20640 non-null  int64`
	`6  households  20640 non-null  int64`
	`7  median_income  20640 non-null  float64`
	`8  ocean_proximity  20640 non-null  object`
	`9  median_house_value  20640 non-null  int64`
	
	`dtypes: float64(4), int64(5), object(1)`
	`(Rows-20640, coloumns-10)`
	
	Proposed Solution
	
	The proposed solution for California Housing Price Prediction involves developing a machine learning model that utilizes data analytics and predictive modeling techniques to accurately predict housing prices. Collect comprehensive data on housing attributes and relevant features such as longitude, latitude, housing median age, total rooms, total bedrooms, population, households, median income, median house value, and ocean proximity.
	Cleanse and preprocess the data, handling missing values, outliers, and inconsistencies to ensure data quality. Develop a user-friendly interface that allows users (potential buyers, sellers, or real estate investors) to input relevant attributes and obtain predicted housing prices.
	Deploy the trained model into the system, making it accessible and interactive for users.