1. Local Home Builders Association(LHBA) is grappling with the challenge of accurately pricing homes in a dynamic real estate market, exacerbated by the recent economic crisis. Historically relying on county-assessed values, LHBA is now faced with the need for a more effective pricing model. The aftermath of the crisis and concerns about the impact of bubbles on home prices have underscored the limitations of their current approach. The lack of synchronization among executives regarding pricing methods and the demand for explainability from stakeholders further complicate decision-making.
    a) Model Development and Validation: How can LHBA build a pricing model that not only accurately predicts home selling prices but is also transparent and understandable to stakeholders?
    b) Impact of Crisis on Pricing: What are the significant differences between pre-crisis and post-crisis housing market dynamics, and how do these differences influence the factors affecting home prices?
    c) Stakeholder Acceptance: How can LHBA ensure that the new pricing model gains acceptance among stakeholders, particularly those who are accustomed to the traditional county-assessed values, and addresses their concerns about bubbles and market uncertainties?

2.
    a) What specific features or factors do stakeholders consider most crucial in determining the value of a home, and how have these considerations evolved in the post-crisis period?
    b) How do stakeholders currently perceive the influence of market bubbles on home prices, and what level of explainability do they expect from a new pricing model to gain their confidence?
    c) In what ways do LHBA executives envision utilizing the predictions of a new pricing model in their decision-making processes, and what key decisions hinge on the accuracy and reliability of these predictions?

¾.

Analytics Problem: Develop a predictive model for home selling prices that incorporates relevant features, taking into account the differences between pre-crisis and post-crisis periods.

Decision Impact: Improved accuracy in predicting home prices will empower LHBA executives to make informed decisions regarding construction, lending, and other operations, ultimately maximizing profitability and sustainability.

5.

Business Understanding: Gain a deep understanding of LHBA's objectives, challenges, and stakeholders' expectations.

Data Understanding: Explore and analyze Pre-Crisis and Post-Crisis datasets, identifying key features influencing home prices.
Data Preparation: Clean and preprocess data, handle missing values, and engineer relevant features for model development.
Modeling: Train predictive models, validating and fine-tuning to ensure accuracy and explainability.
Evaluation: Assess the model's performance on test data, interpreting errors and refining the model as needed.
Deployment: Implement the model into LHBA's decision-making processes, ensuring ongoing monitoring and updates.

6.

````
title: "Housing Prices"
output: html_document
date: "2023-10-25"
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

```{r}
library(ggplot2)
library(ggthemes)
library(scales)
library(dplyr)
library(data.table)
library(corrplot)
library(GGally)
library(e1071)
```

```{r}
train <-read.csv('../input/train.csv', stringsAsFactors = F)
test <-read.csv('../input/test.csv', stringsAsFactors = F)
dim(train)
str(train)
dim(test)
str(test)
```
````

````
```{r}
sum(sapply(train[,1:81], typeof) == "character")
sum(sapply(train[,1:81], typeof) == "integer")
```

```{r}
test$SalePrice<-rep(NA,1459)
house<-bind_rows(train,test)
```

```{r}
cat_var <- names(train)[which(sapply(train, is.character))]
cat_car <- c(cat_var, 'BedroomAbvGr', 'HalfBath', ' KitchenAbvGr','BsmtFullBath', 'BsmtHalfBath', 'MSSubClass')
numeric_var <- names(train)[which(sapply(train, is.numeric))]
```

```{r}
ggplot(train, aes(x = Neighborhood, y = SalePrice)) +
  geom_boxplot() +
  geom_hline(aes(yintercept=80),
             colour='red', linetype='dashed', lwd=2) +
  scale_y_continuous(labels=dollar_format()) +
  theme_few()
```

```{r}
doPlots(train1_num, fun = plotDen, ii = 2:6, ncol = 2)
doPlots(train1_num, fun = plotDen, ii = 7:12, ncol = 2)
doPlots(train1_num, fun = plotDen, ii = 13:17, ncol = 2)
```

```{r}
train <- house[house$isTrain==1,]
test <- house[house$isTrain==0,]
smp_size <- floor(0.75 * nrow(train))

set.seed(123)
train_ind <- sample(seq_len(nrow(train)), size = smp_size)
````
````

````

train_new <- train[train_ind, ]
validate <- train[-train_ind, ]
train_new <- subset(train_new,select=-c(Id,isTrain))
validate <- subset(validate,select=-c(Id,isTrain))
nrow(train_new)
nrow(validate)
str(validate)
```

```{r}
prediction <- predict(house_model,test)
````