

```
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error

train = pd.read_csv("/content/train.csv")
test = pd.read_csv("/content/test.csv")
sample = pd.read_csv("/content/sample_submission.csv")

features = ['GrLivArea', 'BedroomAbvGr', 'FullBath', 'HalfBath']


X = train[features].fillna(0)
y = train['SalePrice']
X_test = test[features].fillna(0)

X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)

model = LinearRegression()
model.fit(X_train, y_train)

val_pred = model.predict(X_val)
print("MSE:", mean_squared_error(y_val, val_pred))

test_pred = model.predict(X_test)
sample['SalePrice'] = test_pred
sample.to_csv("submission.csv", index=False)
```

 MSE: 2810942965.2180653