



Entrega final: House's price

Brian Said Montes Pacheco

Facultad de Ingeniería, Universidad de Antioquia

Inteligencia Artificial

Raul Ramon Pollan

18 de noviembre de 2022

Tabla de contenido

Introducción	3
Objetivo General	3
Objetivos específicos	3
Data set.....	3
Desarrollo del problema	4
Métricas usadas	4
Exploración y procesado de datos	5
Resultados	6
Conclusiones	11

Introducción

El mundo de la programación y los muchos lenguajes que en él existen deben encontrar una solución al problema, y este proyecto debe encontrar el precio más acertado y correcto de la vivienda según las características especificadas del comprador.

Objetivo General

Precisar un algoritmo que pueda clasificar diferentes tipos de inmuebles o viviendas según un precio que pueda pagar el comprador.

Objetivos específicos

Que en cada vivienda se le asigne el valor monetario adecuado a cada objeto o inmueble que este en la propiedad.

Según un precio que diga el comprador se pueda encontrar diferentes tipos de viviendas, que existan una muy buena variedad.

Que las viviendas que estén en las bases de datos tengan muchas localizaciones a nivel nacional e internacional.

Que el cliente sin importar su religión o etnia pueda encontrar una vivienda muy agradable la cual comprar según su capital.

Además de viviendas también se pueda agregar datos de inmuebles para negocios o microempresas.

Data set

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>

Desarrollo del problema

En la actualidad la adquisición o compra de una vivienda es una tarea que se considera desafiante en la evolución del ser humano actual, es considerada una actividad desafiante ya que se desarrolla en un ambiente competitivo. Un claro ejemplo de esta situación son las viviendas residenciales de Ames y Iowa, por el simple hecho de que no tienen precios de venta negociables con el cliente, a pesar de que se tiene en consideración muchos factores que influyen en este criterio ya que existen muchas consideraciones que catalogan a estos lugares como los más interesantes en cuestión de adquisición inmobiliaria por temas de ubicación, rutas de acceso entre otras cosas y aun así, el inmueble no es suficiente satisfactorio respecto a lo que puede aportar. Por lo tanto, se deberán tener en cuenta muchas más variables descritas en la competencia de Kaggle, y aun así los precios de venta de las viviendas residenciales continuarán siendo inexactos al momento de realizar la negociación con un comprador, pero la intención es buscar un porcentaje de acierto alto por medio del análisis de los datos encontrados en esta competencia y así definir costos de viviendas razonables, que motiven al cliente a tomar la decisión de adquirir la vivienda.

Métricas usadas

La evaluación de las predicciones obtenidas será muy importante para este proyecto, ya que mostrará qué tan preciso o alejado estuvo el resultado de sus objetivos. Por esta razón se utilizará la métrica RMSE (error cuadrático medio) entre el logaritmo del valor previsto y el logaritmo del precio de venta observado, es decir, RMSE medirá qué tanto coinciden los datos previstos con los datos observados, sin dejar a un lado que se busca un valor pequeño de RMSE ya que este significa que los valores considerados son cercanos .

Exploración y procesamiento de datos

En el desarrollo de este modelo simplificado surgió la idea de establecer un aplicativo capaz de responder a una problemática actual, esta se enmarca en el campo de saber cual es el precio más asequible frente a la adquisición de una propiedad. De primera instancia se optó por descargar la base de datos de kaggle, como nuestra data set principal

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>, se procesa la base y esto nos conlleva a los siguientes resultados.

Preprocesado de los datos

```

LotFrontage      259
Alley            1369
MasVnrType        8
MasVnrArea        8
BsmtQual         37
BsmtCond         37
BsmtExposure     38
BsmtFinType1     37
BsmtFinType2     38
Electrical        1
FireplaceQu      690
GarageType       81
GarageYrBlt      81
GarageFinish     81
GarageQual       81
GarageCond       81
PoolQC          1453
Fence           1179
MiscFeature     1406
dtype: int64

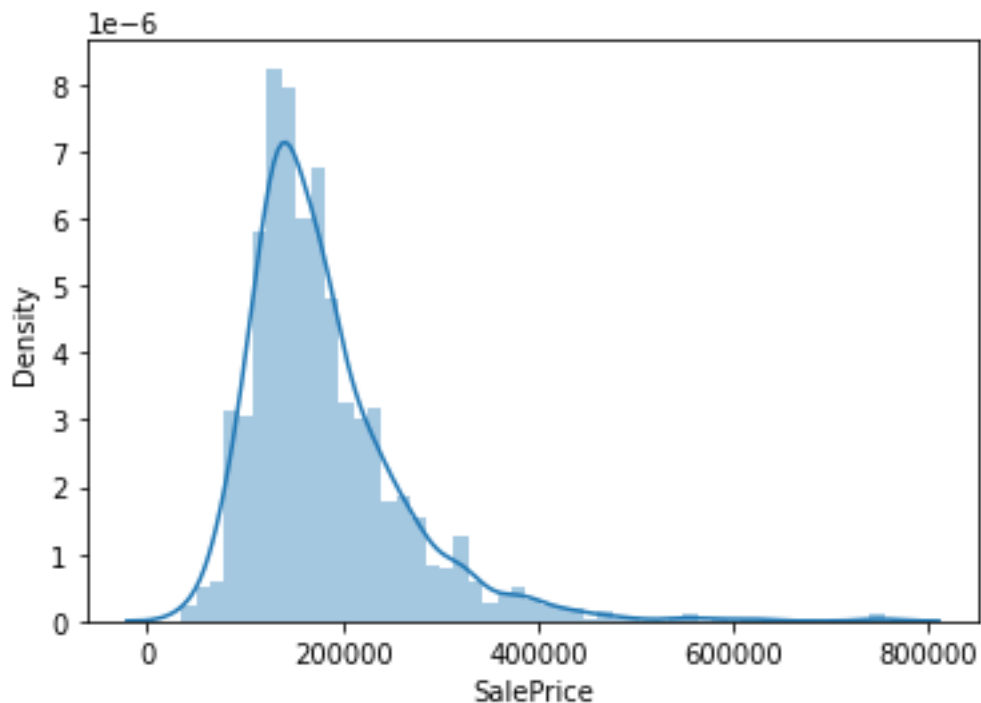
```

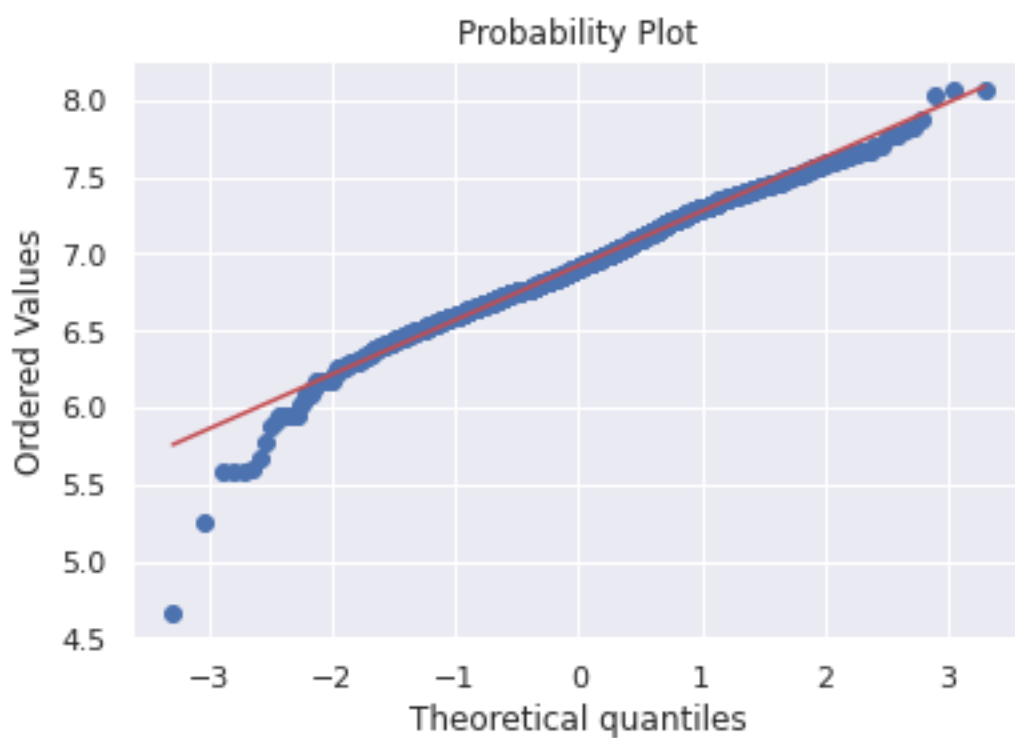
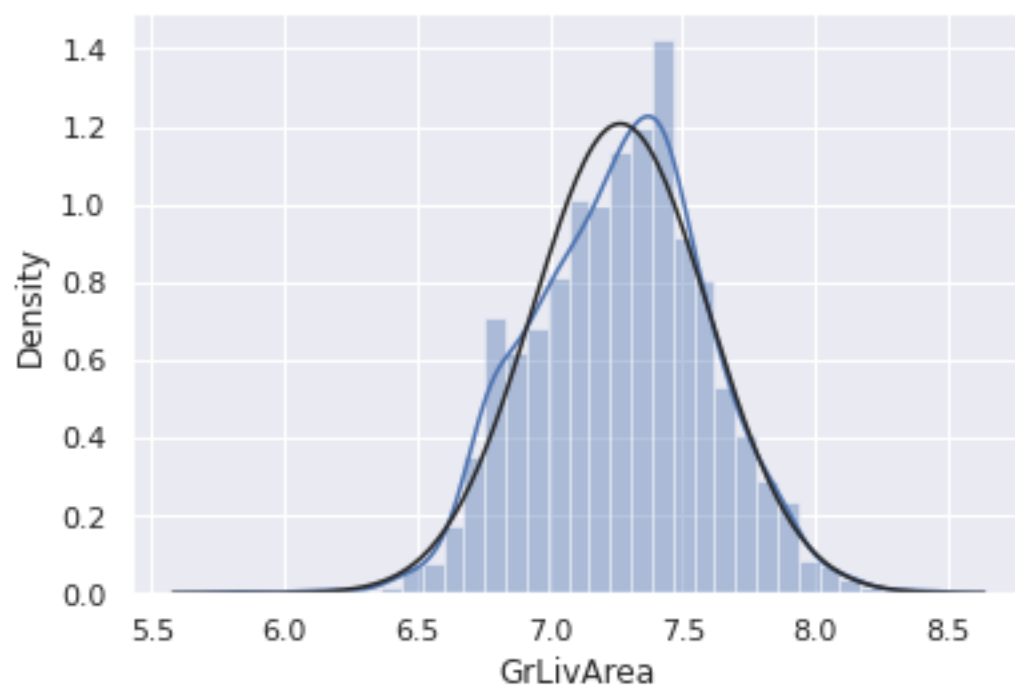
Se establece que la data set es una base muy completa que nos permite determinar un seriado de características, de las cuales podemos determinar un filtro de precios, equivalentes a estos. En principio tuvimos muchos errores respecto a la construcción del mismo aplicativo, pero en lo que respecta al proceso en el cual se analizaron los datos, se determinan muchos valores que nos dejan interpretar que si se desarrolla de buena medida.

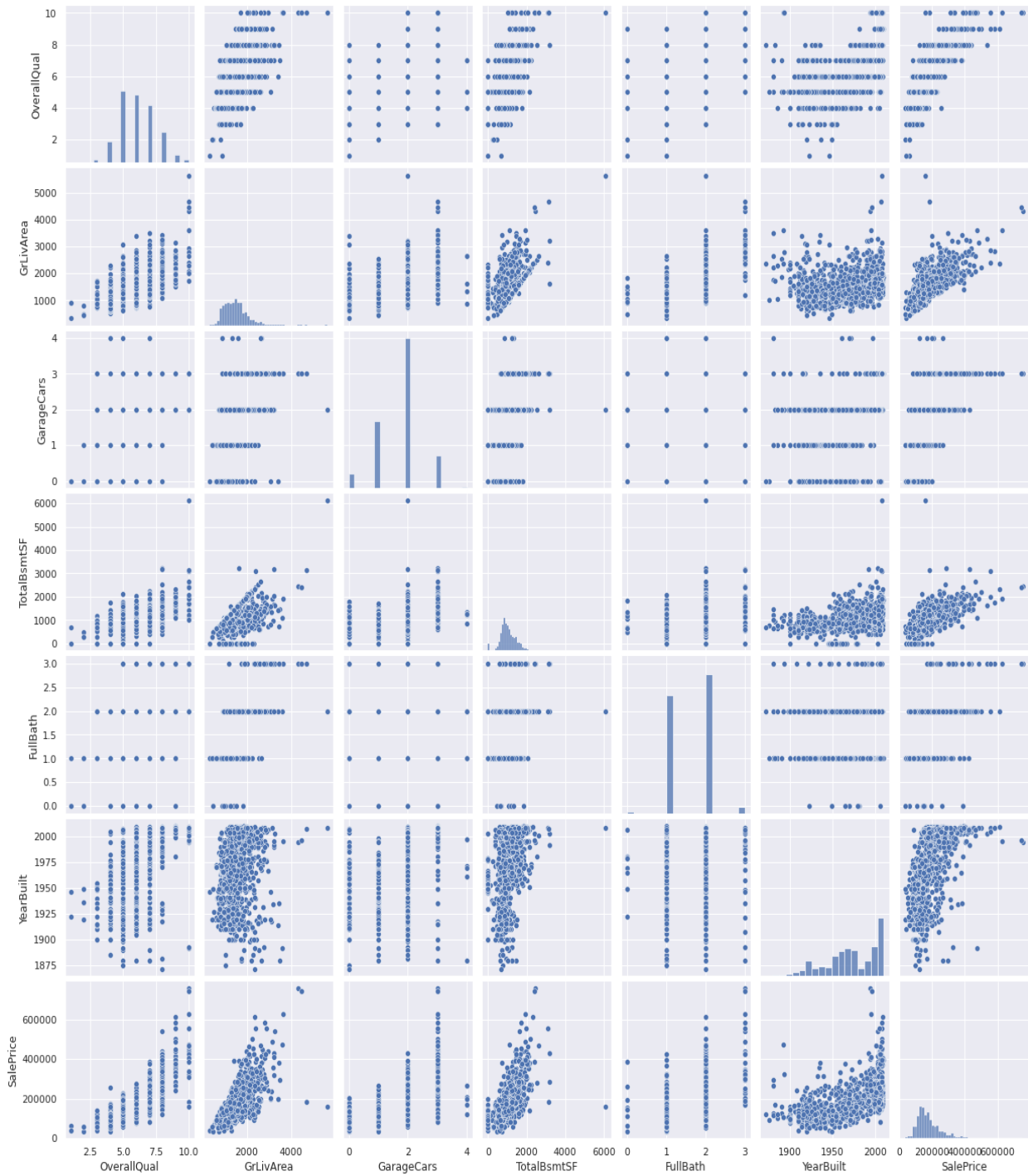
Resultados

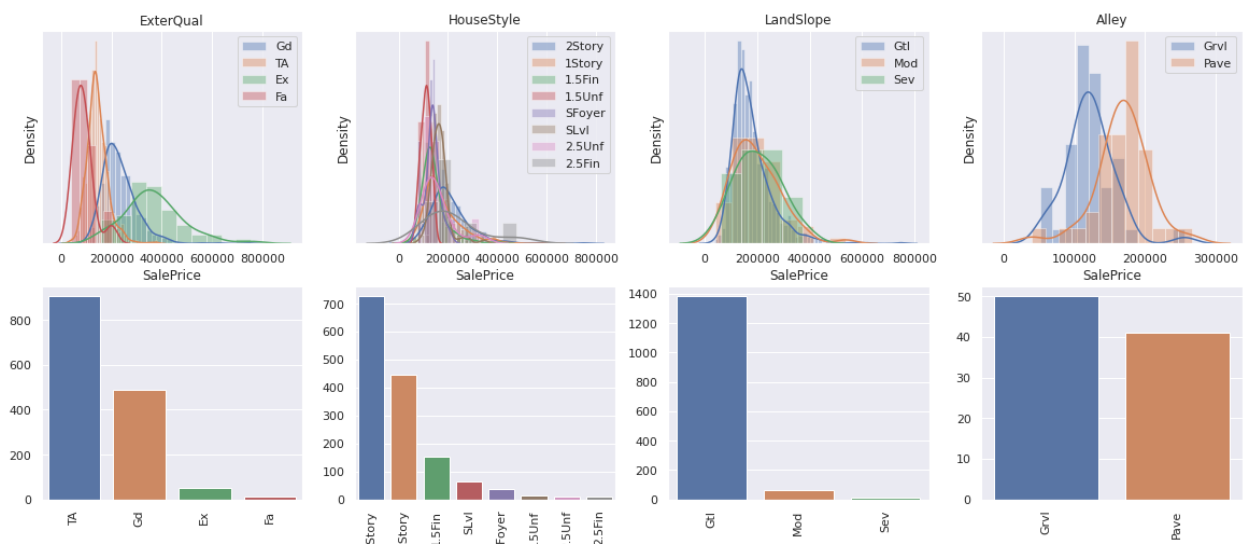
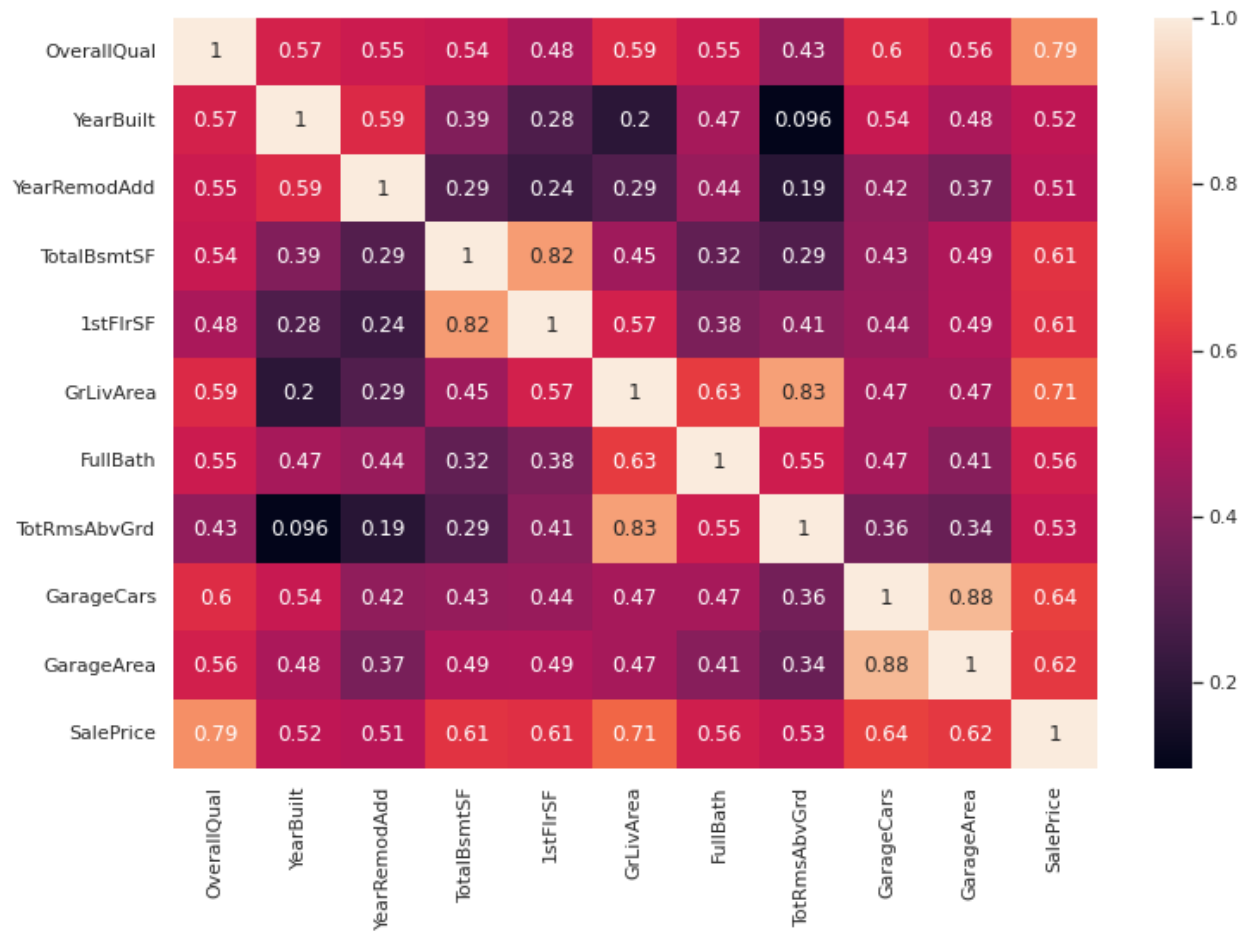
Luego de importar una gran variedad de librerías y datos al inicio, se creó un algoritmo para el problema ya descrito, el programa empezó a correr y según sus respectivos procesos para cada variable fue dando soluciones a cada especificación que se puso. Luego de tantos ensayos y errores dio muchas tablas y graficas con las cuales se puede dar solución al problema de comprar viviendas.

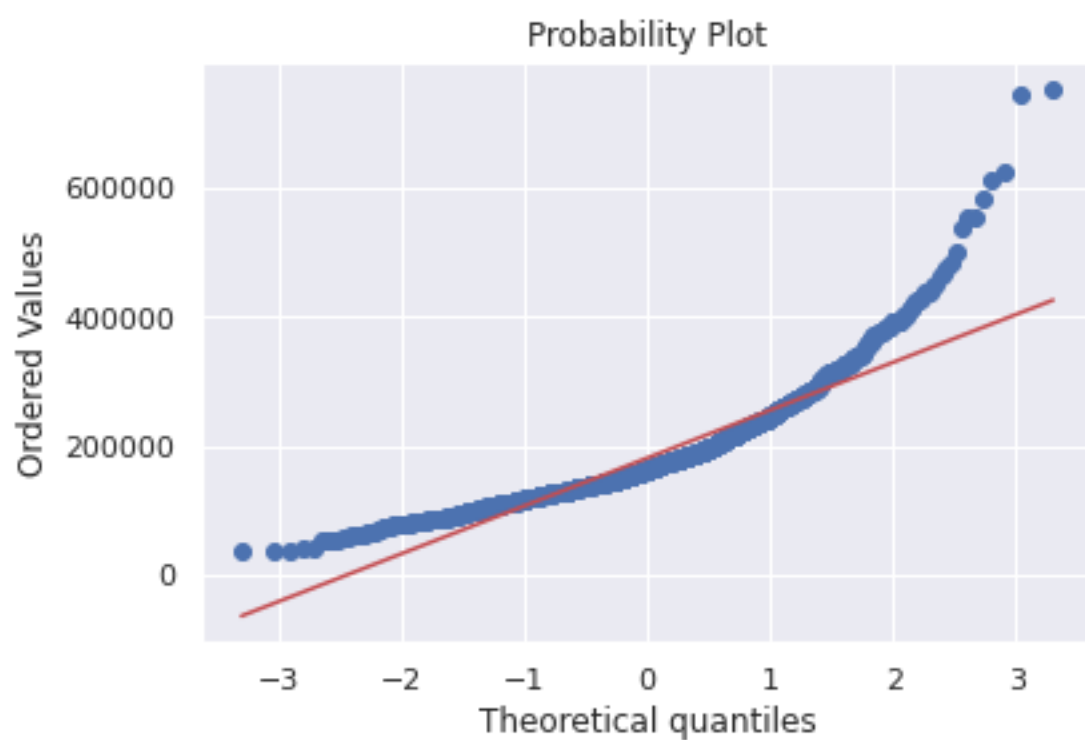
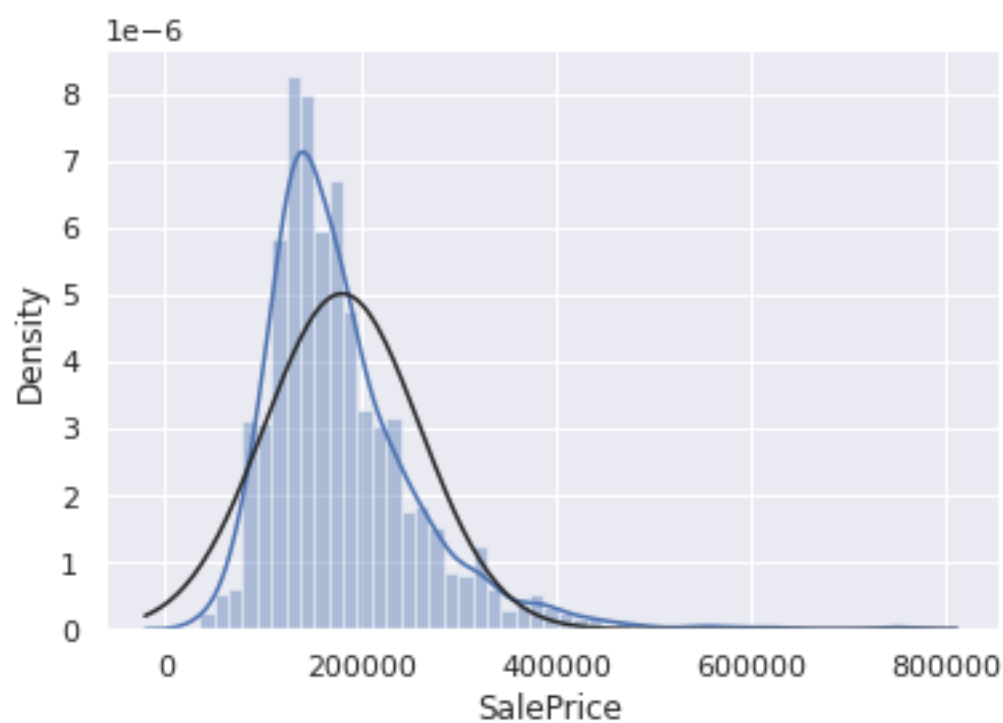
A continuación se muestran algunas graficas que dio como resultado el algoritmo :

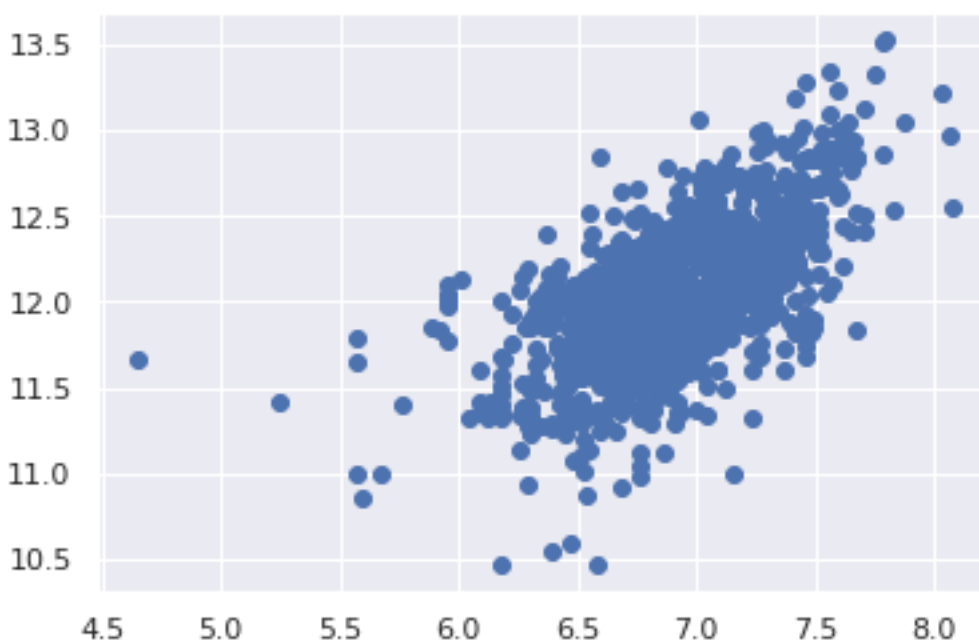












Una de las principales dificultades que se presentó en el transcurso de este proyecto se ve reflejado en la lectura de los datasets en los notebooks, ya que en determinadas ocasiones estas no eran leídas y nos aparecían muchos errores, en algunos casos se solucionaba eliminándolas y volviéndolas a cargar, pero luego dejaban de funcionar.

Conclusiones

En el mundo que vivimos actualmente hay innumerables problemas que la sociedad debe afrontar de la manera más eficiente y eficaz, gracias a los avances tecnológicos se pueden aplicar diversos lenguajes de programación para que de la manera adecuada se puedan solucionar dichos problemas, así como se hizo en este proyecto en el cual se manejó mucha información sustanciosa para dar una solución decente.