

# UNIVERSIDAD DEL VALLE DE GUATEMALA

## Minería de Datos



*Excelencia que trasciende*

**DEL VALLE**  
GRUPO EDUCATIVO

## **Avances Proyecto No. 2**

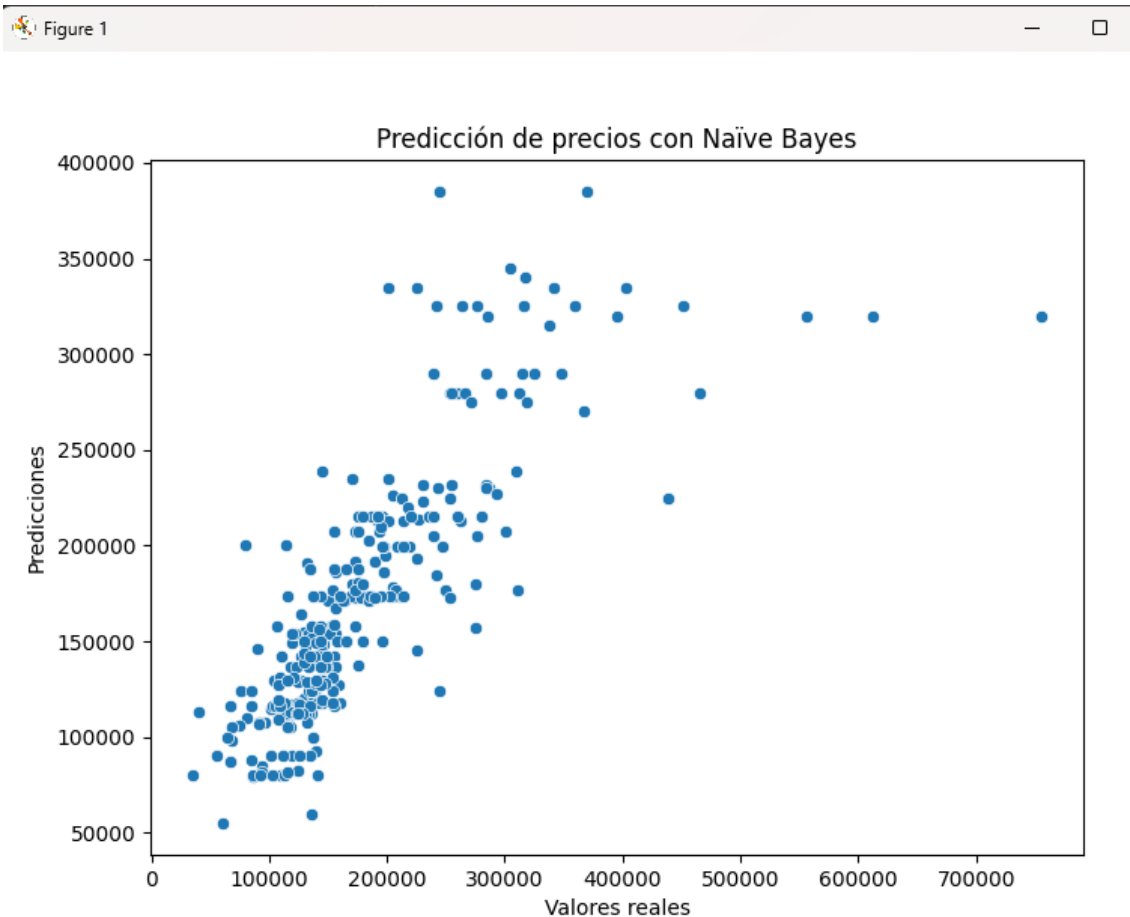
**Isabella Miralles #22293**

**Guatemala, 2025**

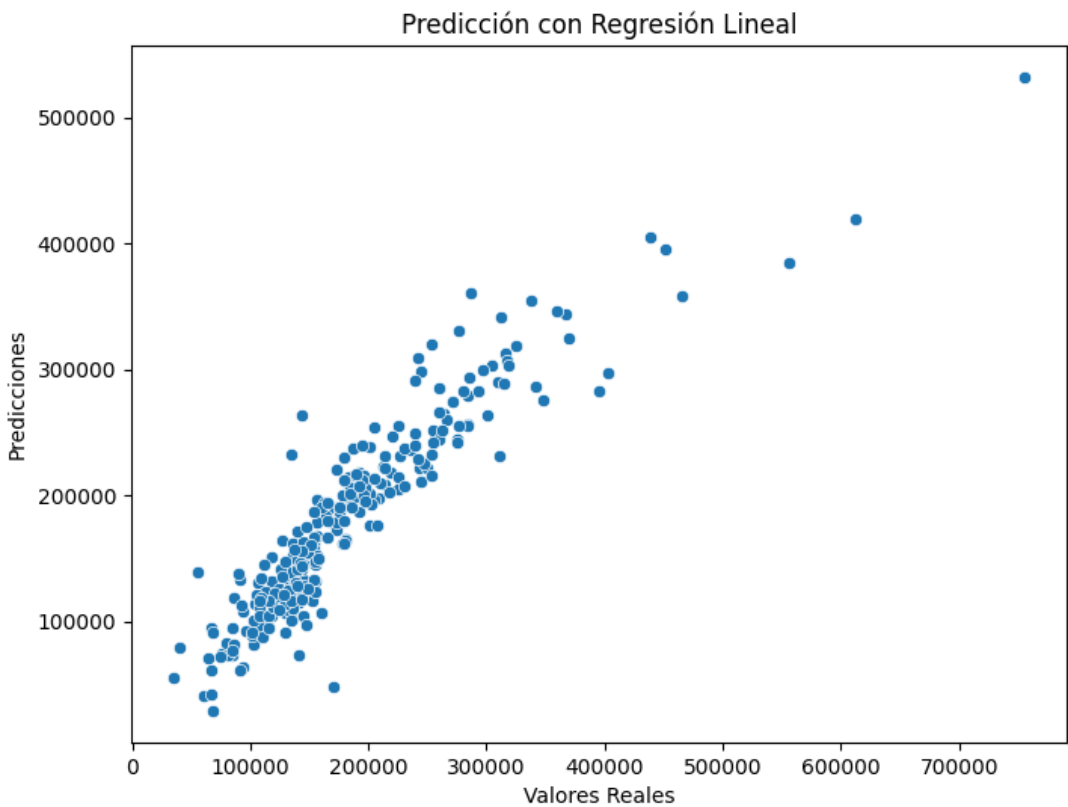
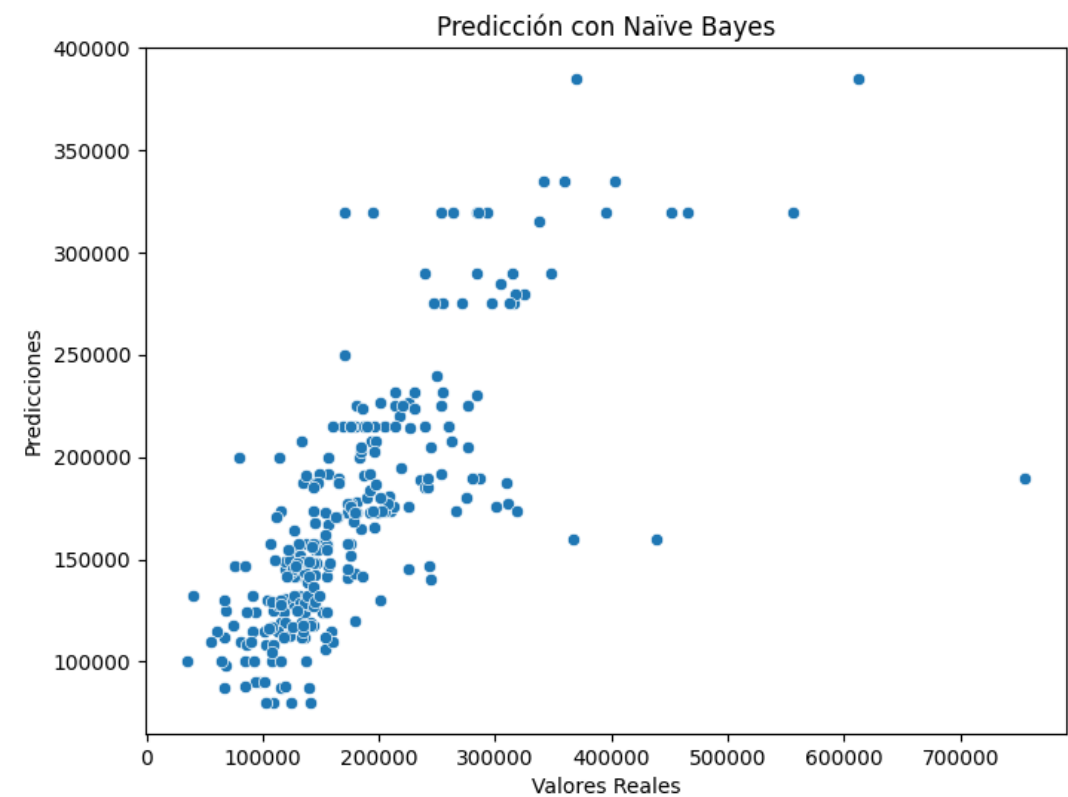
Elabore un modelo de regresión usando Bayes Ingenuo (Naïve Bayes), el conjunto de entrenamiento y la variable respuesta (meta) SalesPrice. Prediga con el modelo y explique los resultados que obtiene. Asegúrese que los conjuntos de entrenamiento y prueba sean los mismos para que los modelos sean comparables.

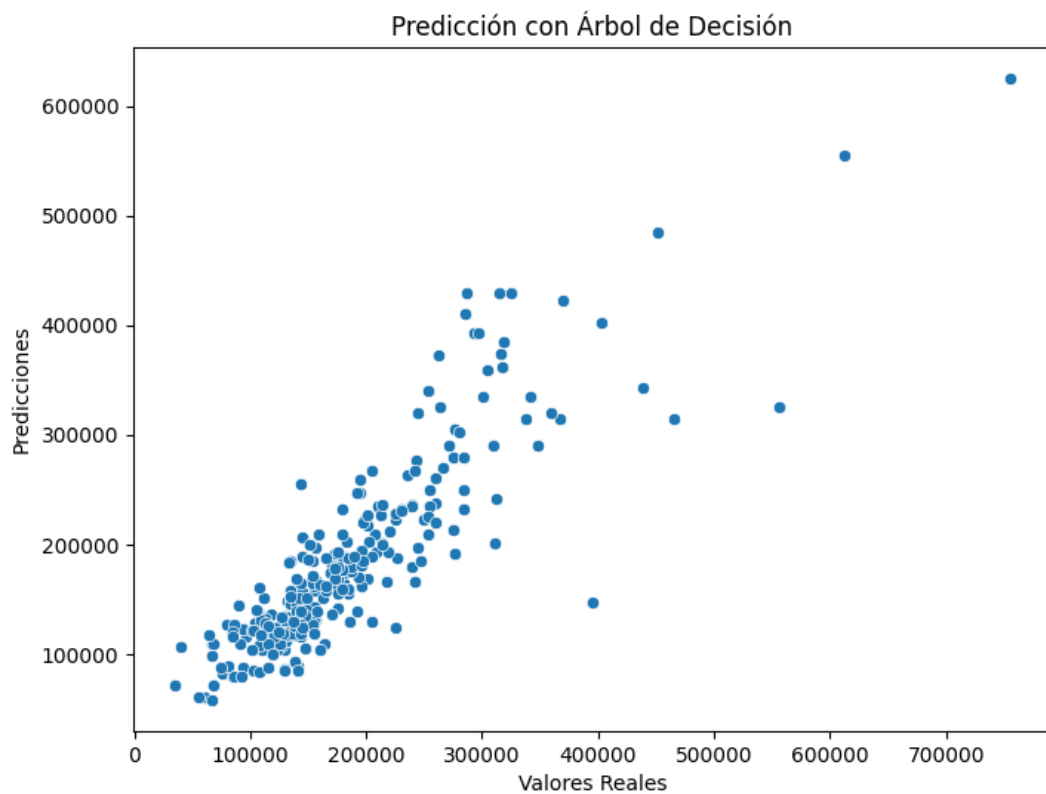
```
### Evaluación del Modelo de Regresión Naïve Bayes ###  
MSE: 2786752509.9623  
RMSE: 52789.7008  
MAE: 30966.6130  
R2: 0.6367
```

Analice los resultados del modelo de regresión usando Bayes Ingenuo. ¿Qué tan bien hizo las predicciones? Utilice las métricas correctas.



Compare los resultados con los de los modelos de Regresión Lineal y el Arbol de Decisión. ¿Cuál funcionó mejor?





```
PS D:\Documentos\Septimo semestre\Mineria de Datos\Proyecto1-MD> python comparacion_modelos.p
```

```
### Evaluación de Naïve Bayes ###
```

```
MSE: 3577906727.0856
```

```
RMSE: 59815.6060
```

```
MAE: 34851.6062
```

```
R2: 0.5335
```

```
### Evaluación de Regresión Lineal ###
```

```
MSE: 1196110750.1122
```

```
RMSE: 34584.8341
```

```
MAE: 21591.7391
```

```
R2: 0.8441
```

```
### Evaluación de Árbol de Decisión ###
```

```
MSE: 1798837350.8048
```

```
RMSE: 42412.7027
```

```
MAE: 27444.8185
```

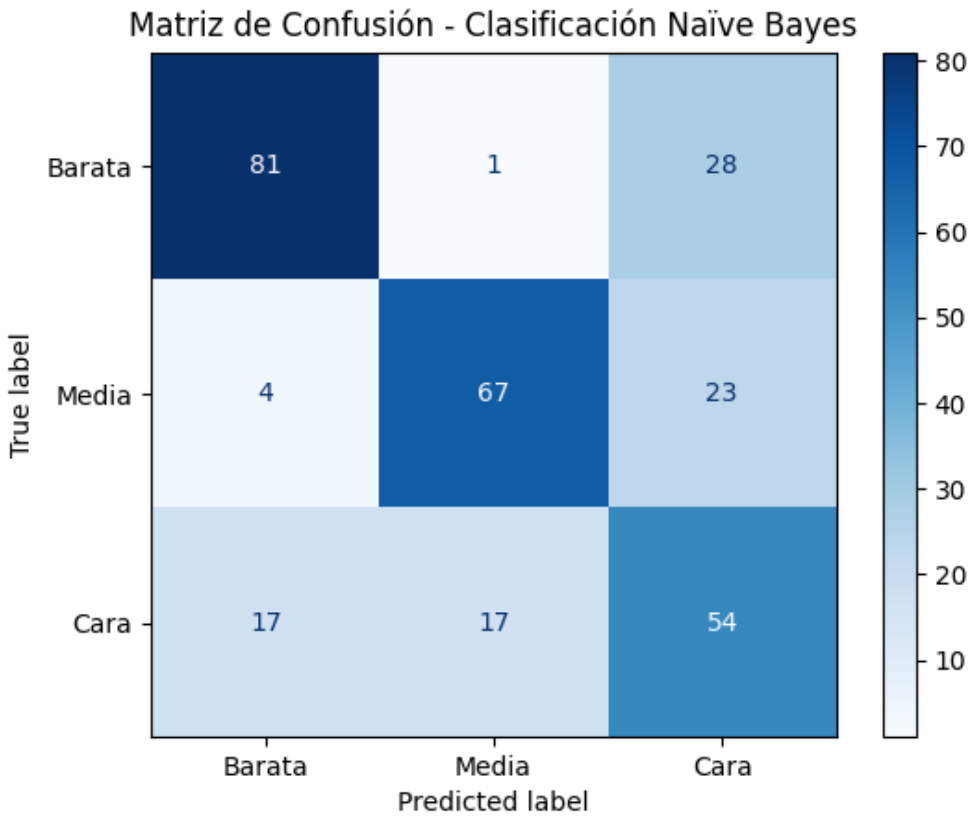
```
R2: 0.7655
```

Haga un modelo de clasificación con Naïve, use una variable categórica nueva en base al precio de las casas (barata, media y cara) y úsela como variable respuesta (meta)

```
### Evaluación del Modelo de Clasificación Naïve Bayes ###
Exactitud (Accuracy): 0.6917808219178082
```

	precision	recall	f1-score	support
Barata	0.79	0.74	0.76	110
Cara	0.79	0.71	0.75	94
Media	0.51	0.61	0.56	88
accuracy			0.69	292
macro avg	0.70	0.69	0.69	292
weighted avg	0.71	0.69	0.70	292

Utilice los modelos con el conjunto de prueba y determine la eficiencia de los algoritmos para predecir y clasificar



**Haga un análisis de la eficiencia del modelo de clasificación usando una matriz de confusión. Tenga en cuenta la efectividad, donde el algoritmo se equivocó más, donde se equivocó menos y la importancia que tienen los errores**