

# **Preço de vendas de imóveis**

The background of the slide features a series of overlapping, flowing waves in shades of orange and red. The waves originate from the bottom left and curve upwards towards the right, creating a sense of movement and depth. The colors transition from a bright orange on the left to a deeper red on the right, with various intermediate tones in between.

# Cenário

Para contextualizar os dados fiz uma breve pesquisa temporal no dataset, peguei o último ano de venda registrado que foi em 2010

Nesse período, os preços estavam relativamente baixos, o que impactava diretamente na quantidade de vendas devido a especulação do mercado.

Pessoas compravam imóveis como um investimento, pois tinha a esperança que o preço iria subir.

Boston também teve um número significativos de bairros que foram abalados pela gentrificação.

<https://blog.movingimoveis.com.br/dez-anos-apos-bolha-mercado-imobiliario-dos-eua-tem-recuperacao-gradual/>

<https://www.voltaaomundo.pt/2016/12/12/roteiro-de-boston-pelo-embaixador-dos-eua-e-fa-de-portugal/>

**Location,  location, location!**

## **As características avaliadas**

Um monte de nome estranho, mas a gente chega lá!

# Dentre as coisas que compõem o dado, estão:

- Zona em que o imóvel está.
- Condição da rua
- Se é perto ou não de grandes vias
- Tipo de energia instalado
- Comércio/centros comerciais próximos
- Material da construção
- Estado do edifício
- Entre outras muitas colunas

“

*81 colunas de muita confusão  
e NaNs...*

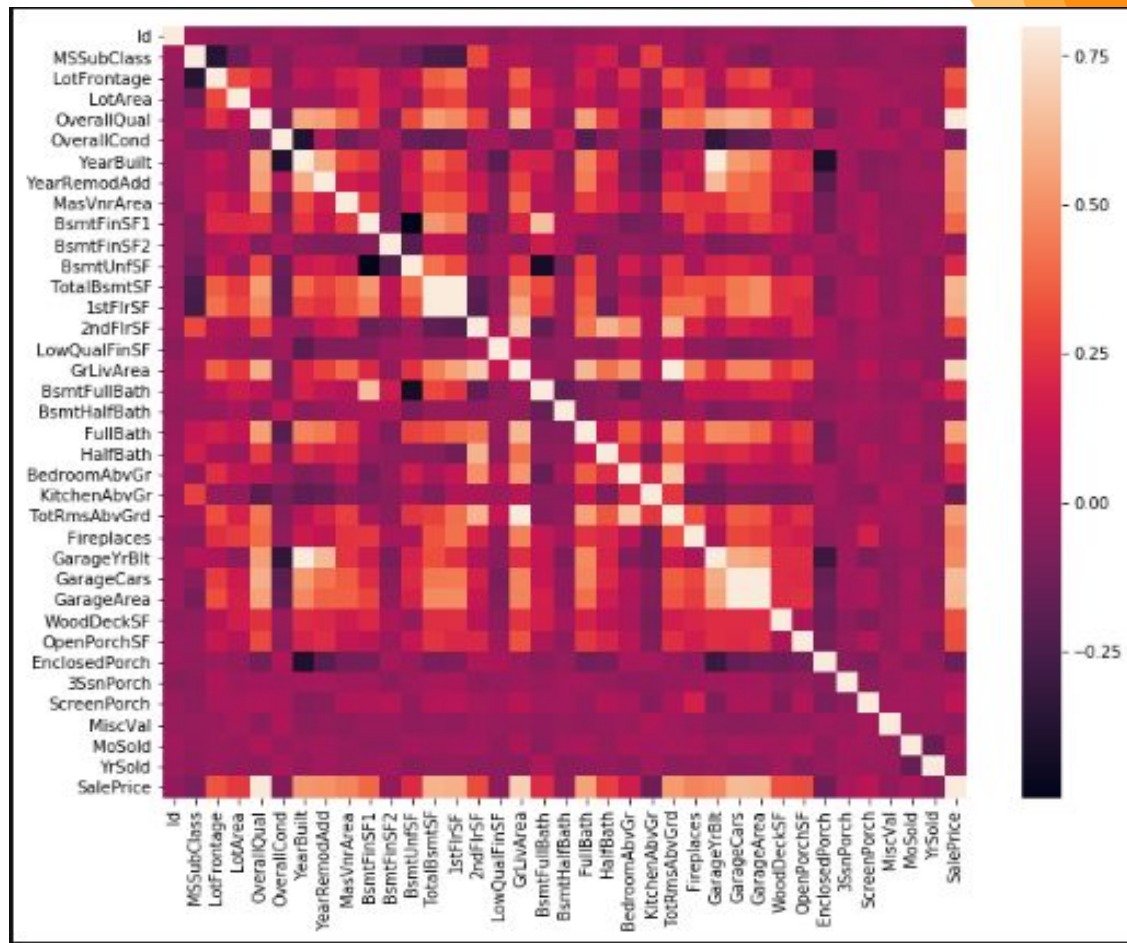
*Dropei 5 colunas e preenchi os  
nulos com valores, que em tese,  
não afetariam muito a análise  
final.*

**1.**

**Quero plotar tudooo'**

Mas querer não é poder. ;(

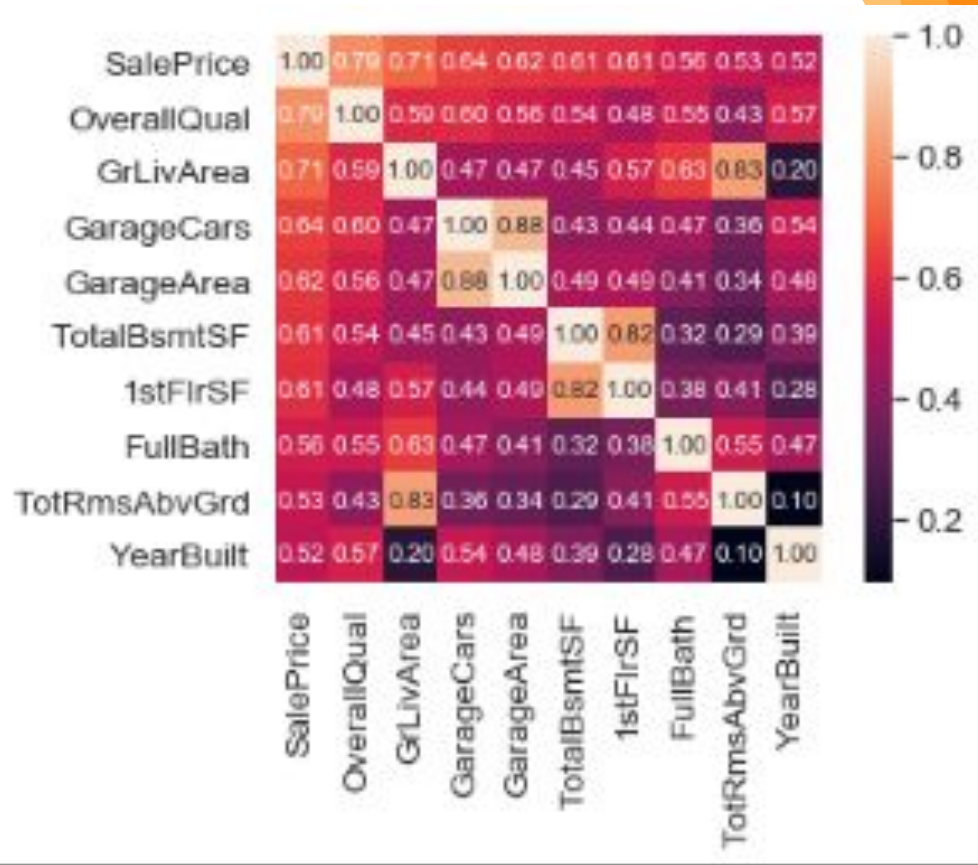
# Correlacionei tudo com tudo.



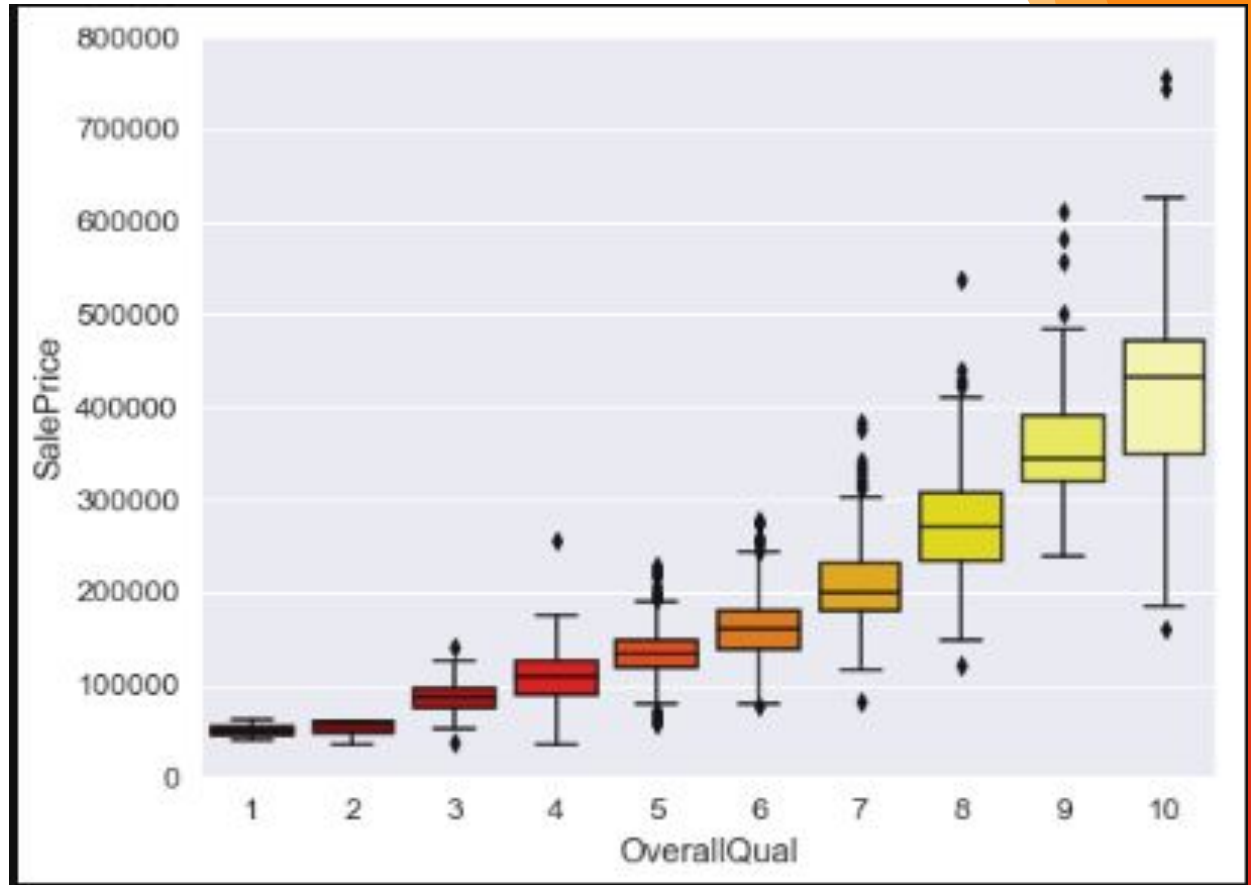
**Correlacionei  
tudo com  
tudo.**



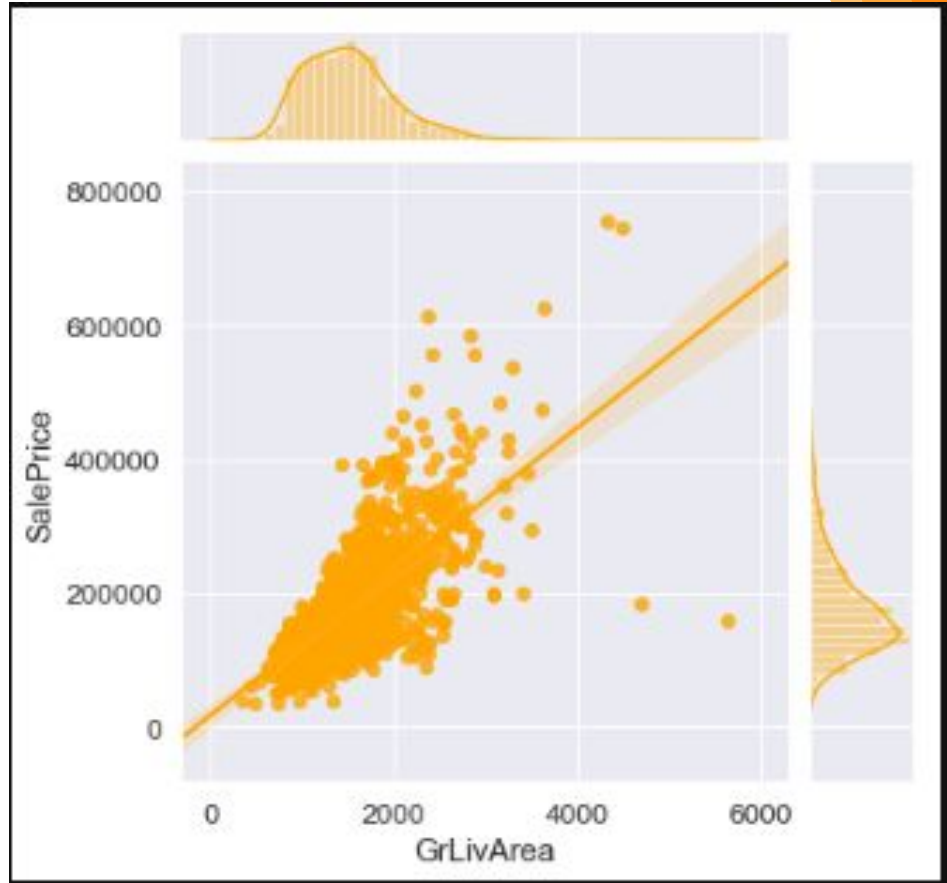
**Agora, são só  
as top  
categorias.**



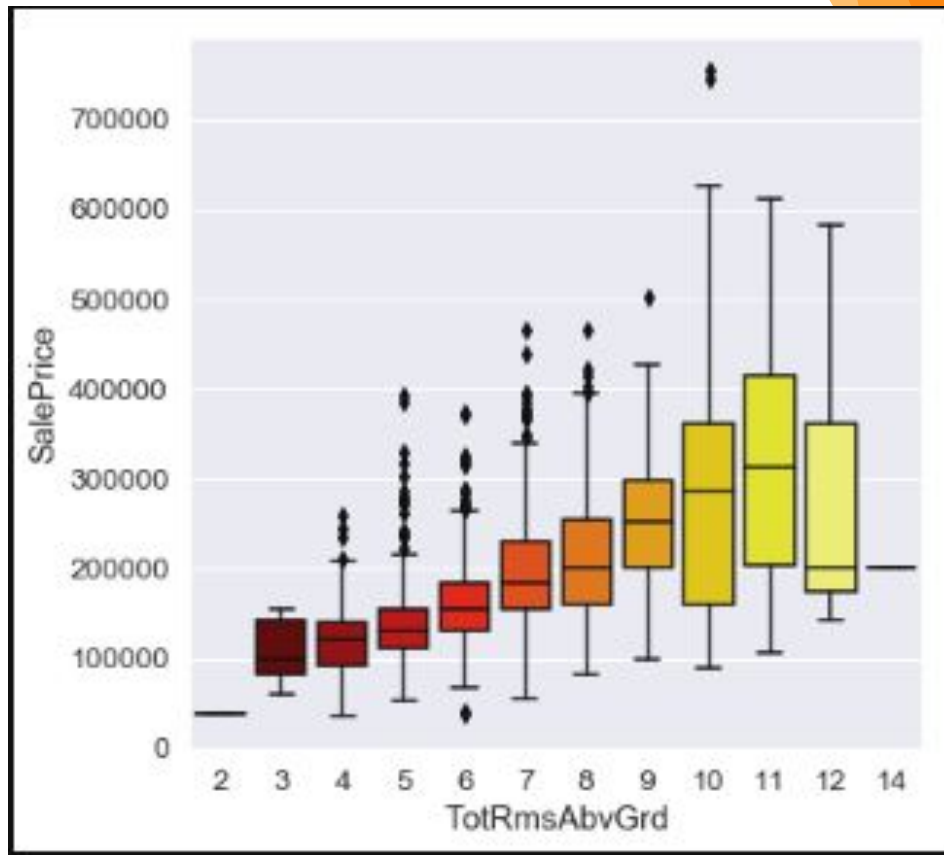
**Sim, pessoas  
pagam mais  
por qualidade.**



**Área verde  
também  
influencia no  
preço.**



**Qtd de  
comodos na  
casa. Detalhe  
nos acima de  
11.**



**Que preços são  
esses?!?**



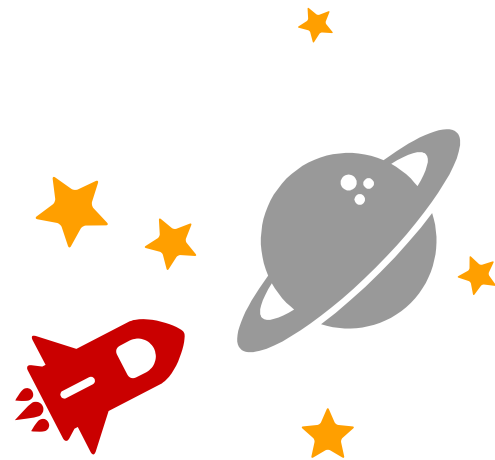


# Once upon a time in 2010...

|       |              |
|-------|--------------|
| count | 14600        |
| mean  | \$180.921,19 |
| std   | \$79.442,50  |
| min   | \$34.900,00  |
| 25%   | \$129.975,00 |
| 50%   | \$163.000,00 |
| 75%   | \$214.000,00 |
| max   | \$755.000,00 |

# Regressão Linear

Se foi, foi. Se não foi, ficou. O que ficou,  
está.



“

*No fim...deu isso.  
Usei um test de 40%*

```
R2 no set de treino: 0.93  
R2 no set de teste: 0.77
```



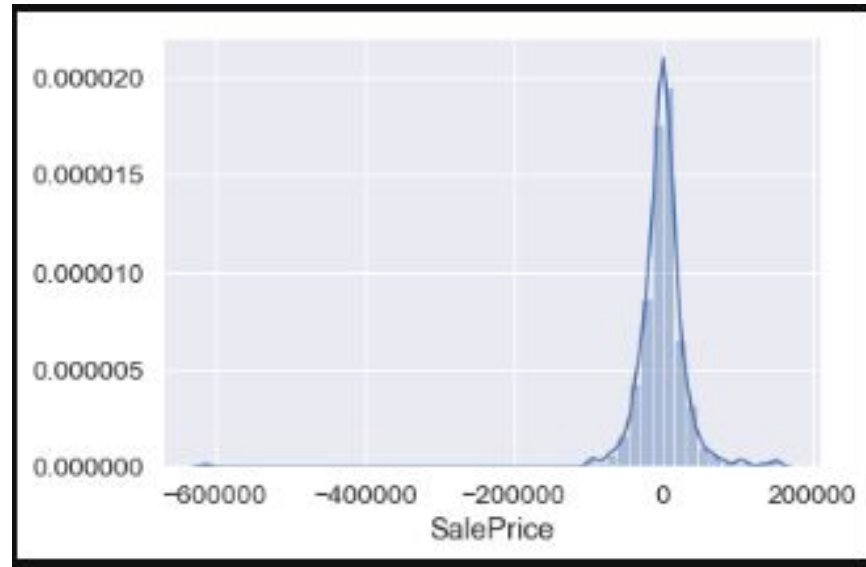
“

*Erros...*

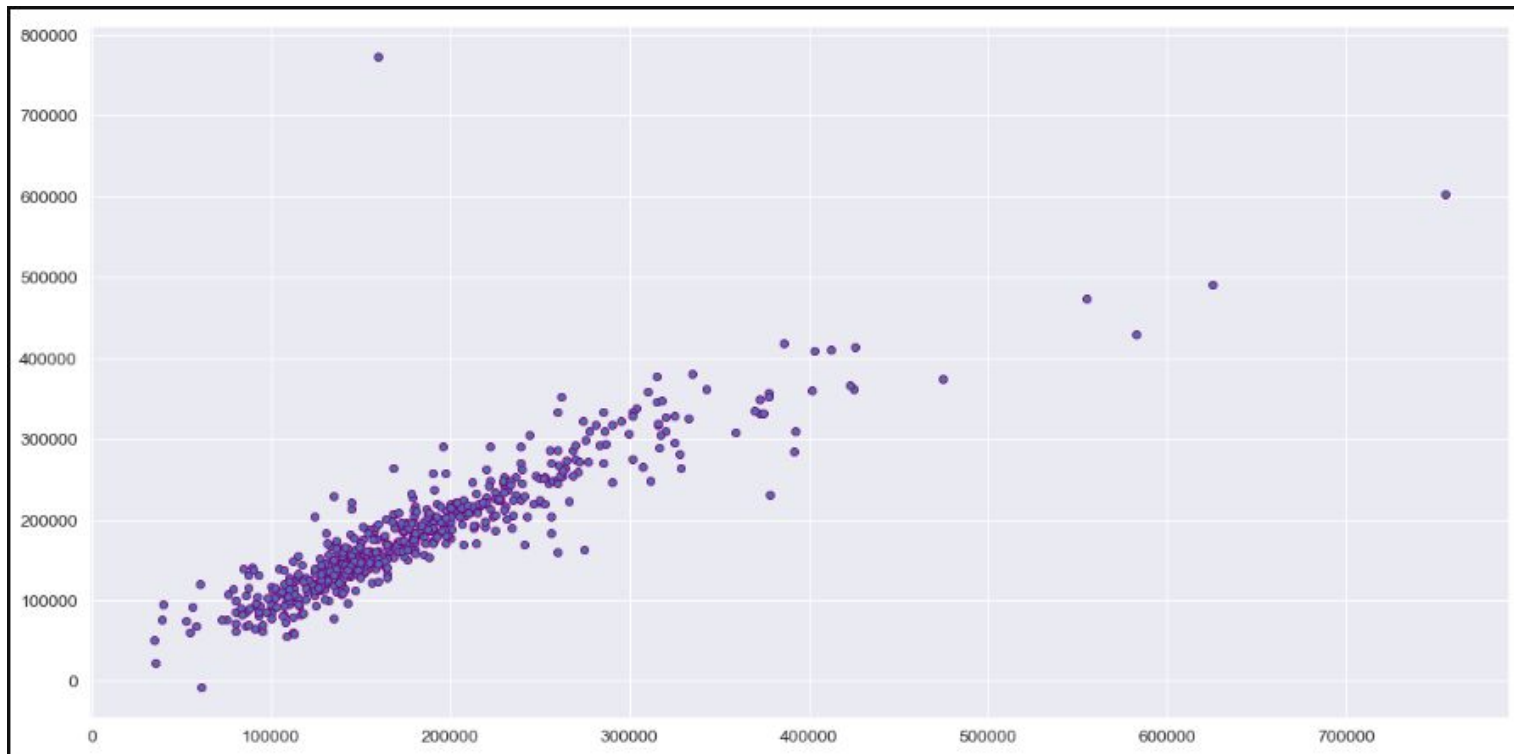
```
MAE: 20336.705168728018  
MSE: 1469805172.2677093  
RMSE: 38338.038190127954
```

“

*Erros...*



# E mais um plot fofo...



# A louca dos plots...

