



# ***HOUSING ANALYSIS***

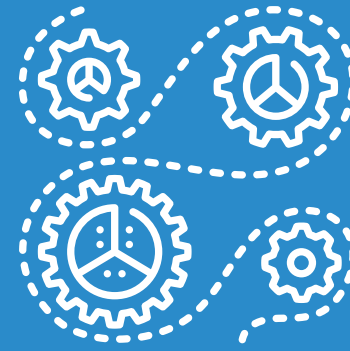
**Mid-term project | Regression**

Antonio Saleme Sastre & Marian Moreno Corral



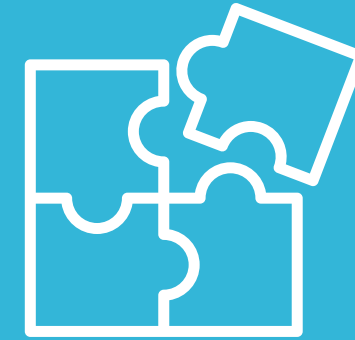
1

INTRO



2

PROCESSES



3

INSIGHTS



4

RESULTS

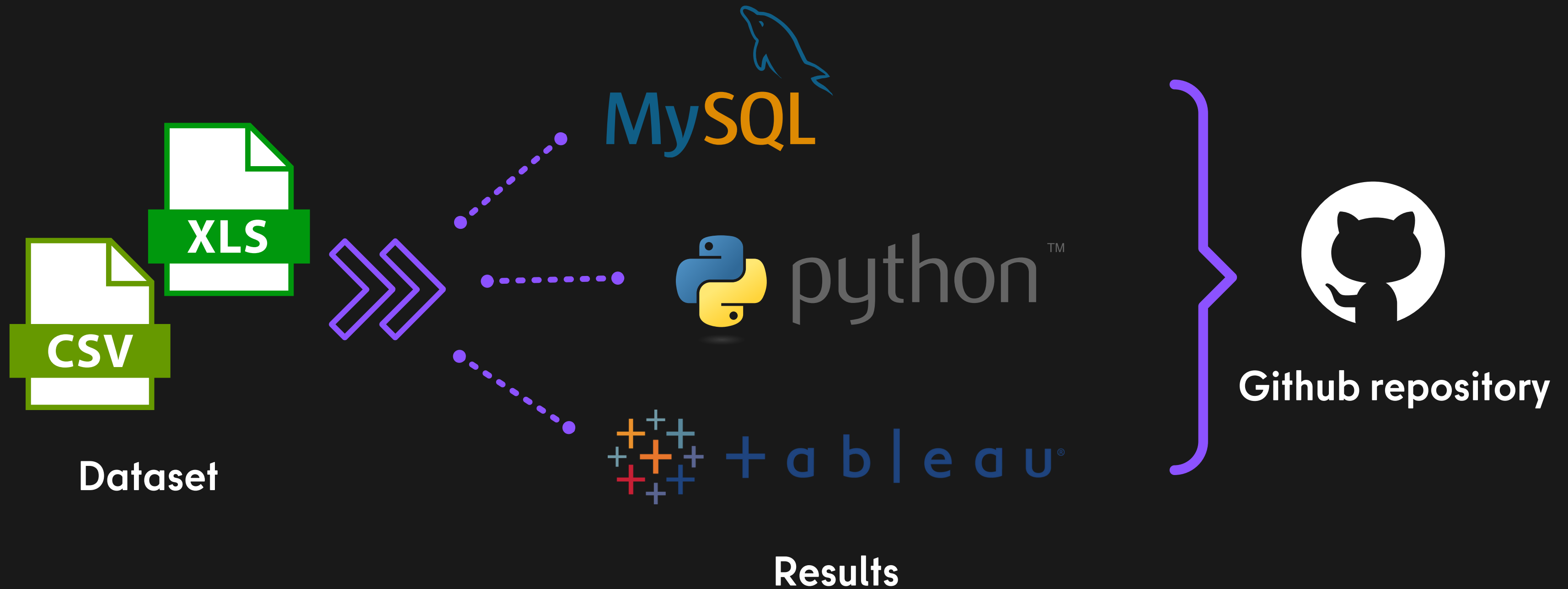
# INDEX OF CONTENTS

# INTRODUCTION

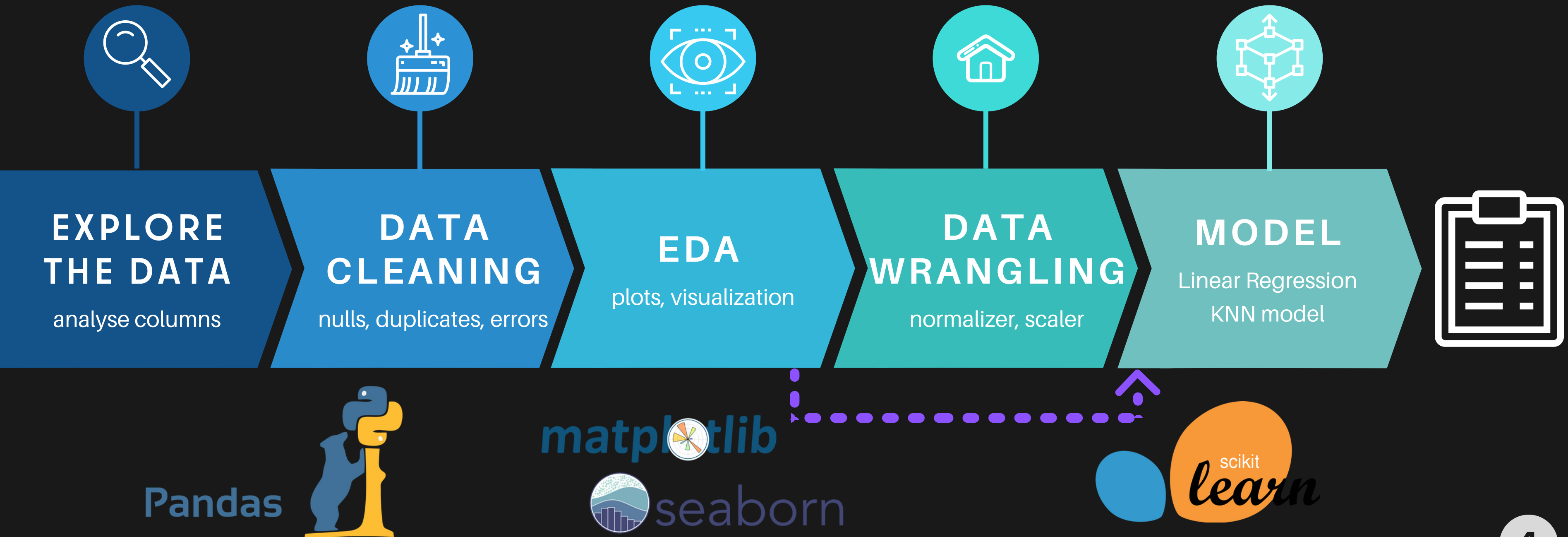
- The dataset consists of information on **22,000 properties sold** between 2014 to 2015.
- Explore and understand the **characteristics** of the houses.
- Build a Machine Learning model to **predict the price** of a house based on features provided in the data.



# PROCESSES



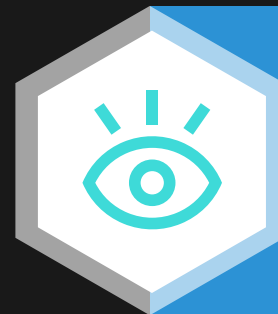
# PROCESSES



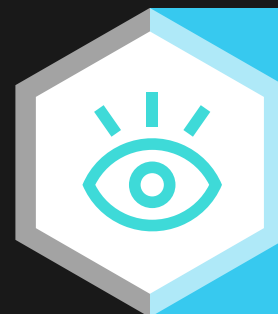
# INSIGHTS



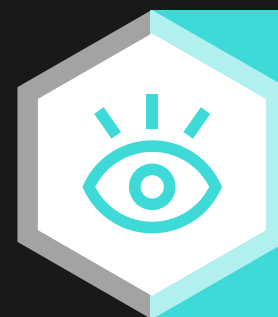
Newer houses, more expensive  
More sold houses built before 2000



Zipcode 98039 extremely  
expensive: Bill Gates, Jeff Bezos

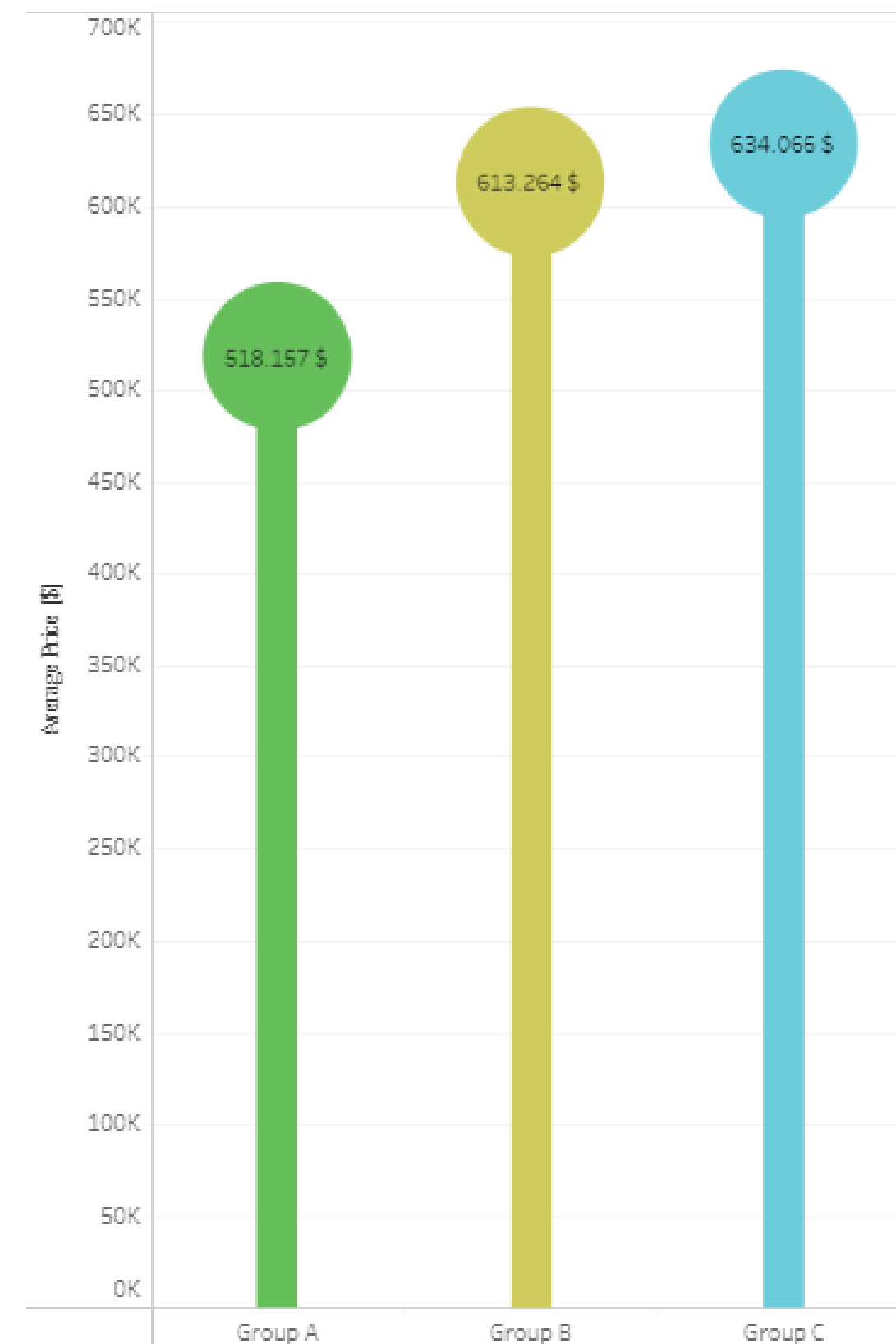


Price highly positively correlated with  
grade of the house and living area



Most popular houses: 3 bedrooms,  
followed closely by 4 bedrooms

Average selling price per category of year of construction







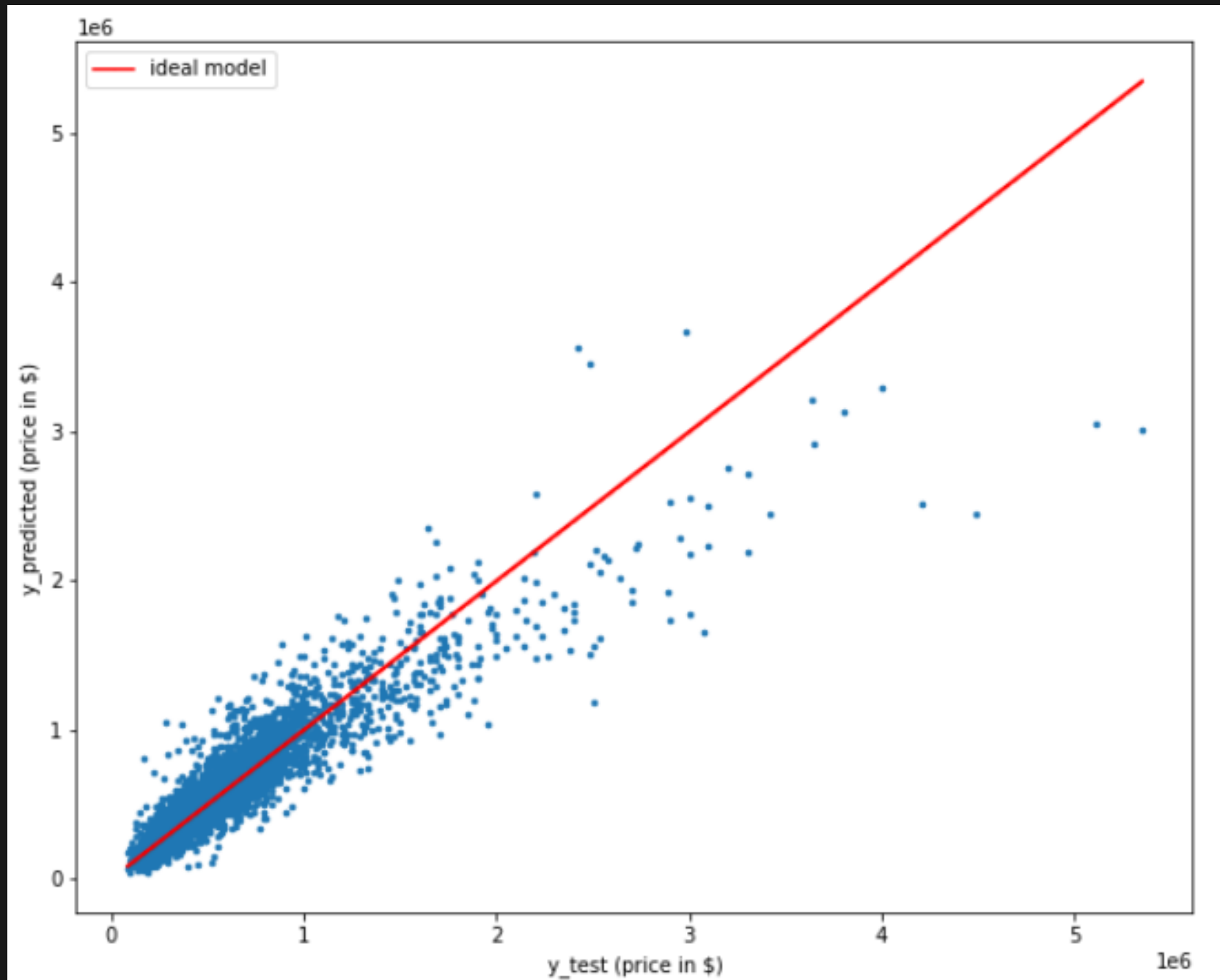
# RESULTS

Best model performance

**R2 : 84,45%**

**MAE: 86280\$**

- Standard Scaler
- One Hot Encoder





# THANKS FOR YOUR TIME!

We'll be happy to answer  
your questions

