

# Laptop Price Predictor

Hind Alamri  
Abrar Alsharidi

# CONTENTS

INTRODUCTION

DATA  
VISUALIZATION

AIM AND  
OBJECTIVES

TOOLS

WORKFLOW

CONCLUSION



Login/Register Now

Wishlist 0

Track Order

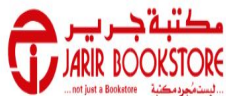
Service Orders

Help

Our Locations

Saudi Arabia | SR

عربي



Search

iPhone

Headphone

Apple Watch

iPad

Huawei

Samsung

TV



All Categories

Shop by Brand

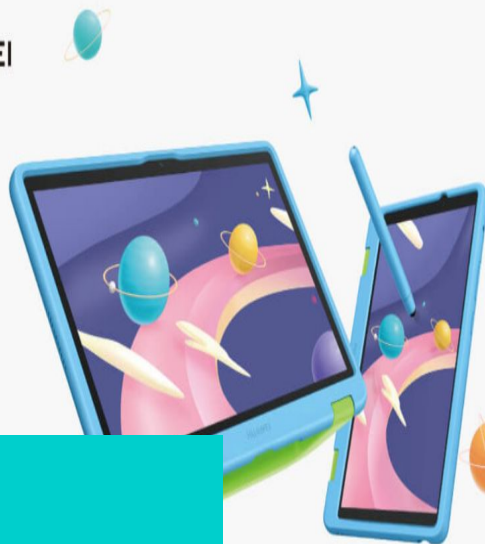
Back to School

Online Exclusive

Weekly Offers

New Arrivals

Renewed Products



## HUAWEI MatePad T Kids Edition

Kids Corner | Eye Protection | Parental Assistant

Starting from  
**SR549**

FREE Gifts worth SR501



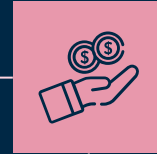
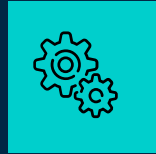
ePad T

\*Offer valid until 27 October 2021

# INTRODUCTION

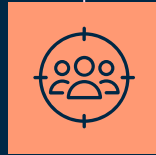
# AIMS AND OBJECTIVES

Perform web scraping on Jarir website



Create a dataset from the scraped data and perform EDA on it

See if our validation score close to the test and reach the goal of generalization



Choose the Best regression model for our data by comparing scores

# Tools

Web scraping with  
selenium

python

Pandas,  
matplotlib, numpy, statsmodels  
and sklearn libraries

# Workflow

Dataset and Methods

# Laptop Dataset

	Acer	Apple	Asus	Chuwi	Dell	Fujitsu	Google	HP	Huawei	LG	...	768.1	900	1080	1200	1440.1	1504	1600.1	1800	1824	2160.1
0	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0	0
1	0	1	0	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	0	0	...	0	0	1	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	1	0	0
4	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1003	0	0	0	0	0	0	0	1	0	0	...	1	0	0	0	0	0	0	0	0	0
1004	0	0	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	0	0	0	0
1005	0	0	0	0	0	0	0	1	0	0	...	1	0	0	0	0	0	0	0	0	0
1006	0	0	0	0	0	0	0	1	0	0	...	0	0	1	0	0	0	0	0	0	0
1007	0	0	0	0	0	0	0	1	0	0	...	0	0	1	0	0	0	0	0	0	0

1008 rows × 99 columns

- After cleaning the data we transform the categorical features into dummies
- We have 99 columns and 1008 rows
- Basic features: Company, TypeName, Inches, Screen Resolution, Cpu, Ram, Memory, Gpu, OpSys, Weight
- Target is price

# Method 1

We split our data into  
Train/Validation/Test

## First Model

Simple Linear Regression

## Second Model

Polynomial Regression  
Model

## Third Model

Ridge Regression

01

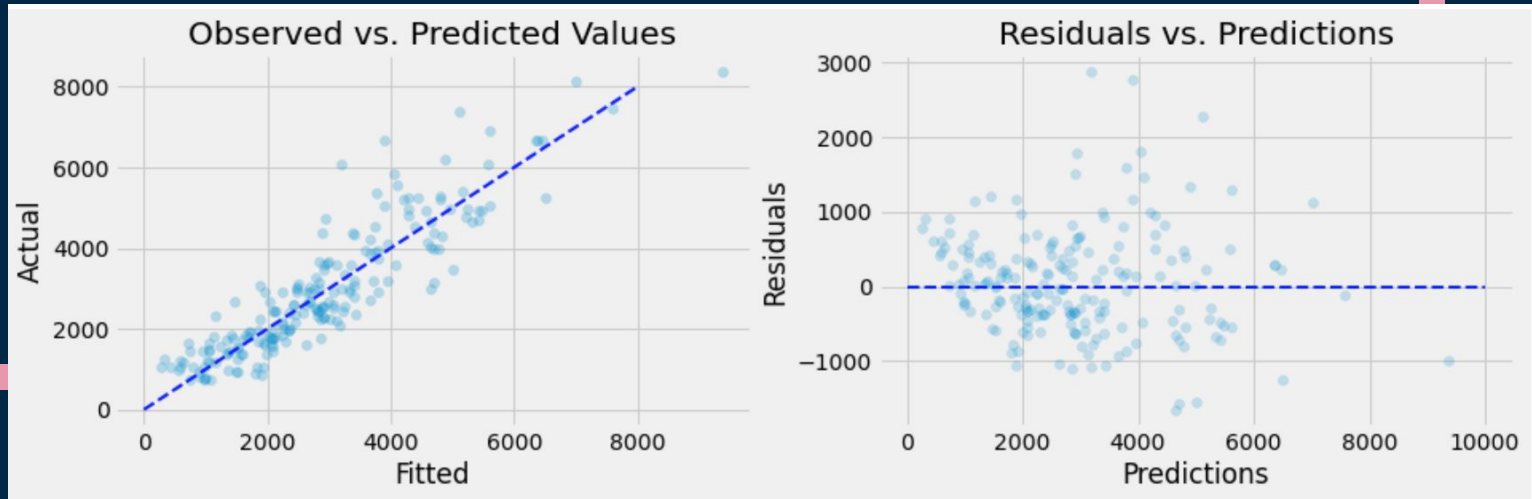
02

03



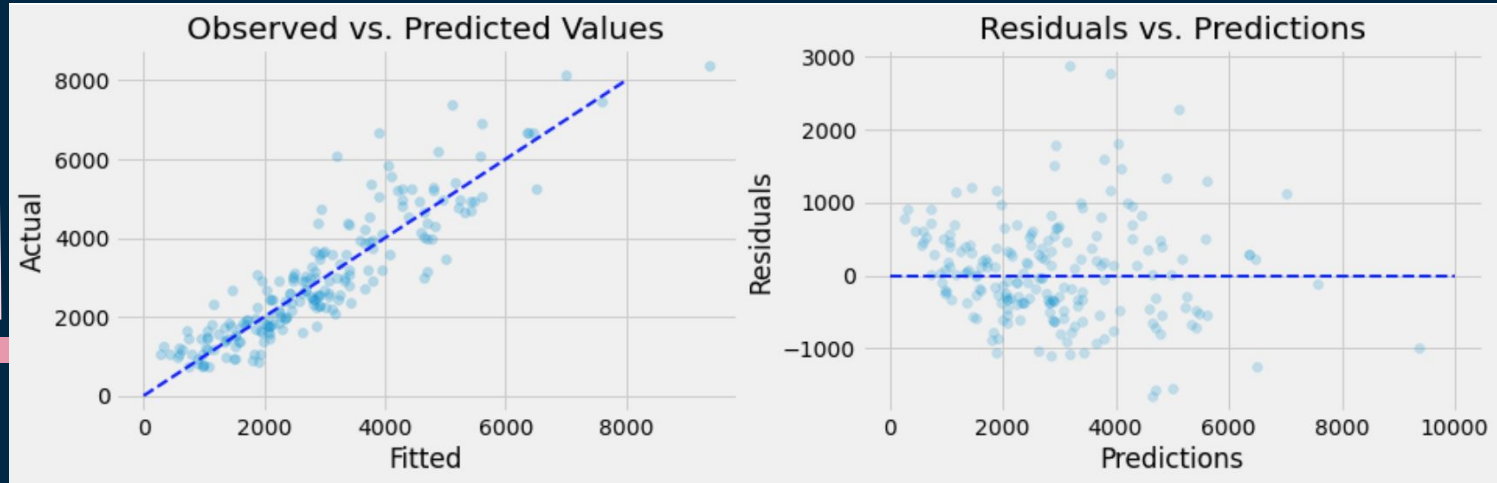


# First Model: Simple Linear Regression



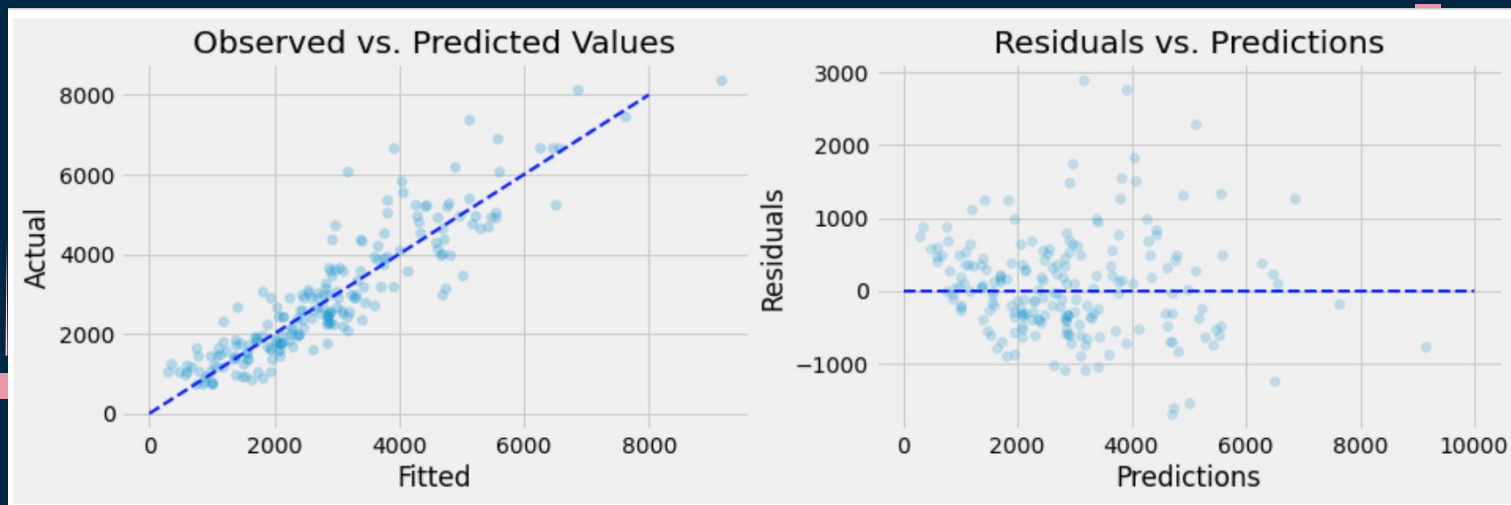
$$R^2 = 0.8098$$

## Second Model: polynomial Regression Model



$$R^2 = 0.8098$$

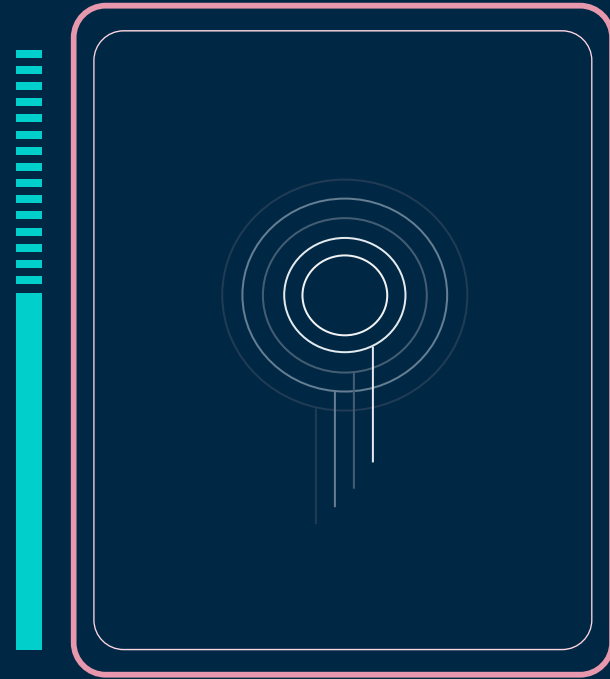
## Third Model: Ridge Regression



$$R^2 = 0.8137$$

# Best Model

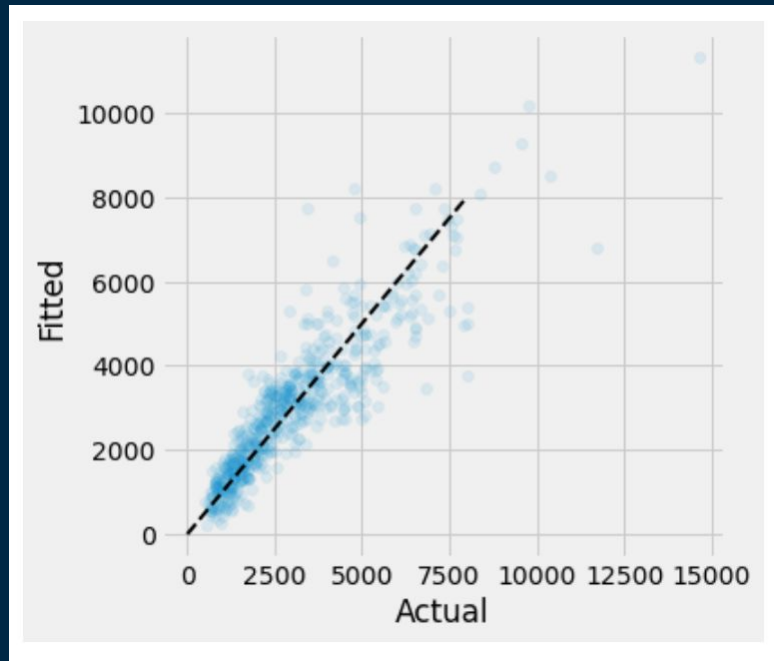
- We find the best model is Ridge Model , which is has validation score = 0.8137
- The test score is equal 0.8024
- So, there is no overfit



# Method 2

## Cross Validation/Test

We split our data into train validate and test, we choose to perform the cross validation on our data to make sure to have the best result



Linear Reg Mean Score: 0.709745

# CONCLUSION

- Ridge Regression Model is the best model to represent our data
- The R score of the test data is 0.80

The background is a dark blue field decorated with an abstract pattern of geometric elements. Scattered throughout are squares of various sizes and colors, including light blue, pink, orange, and teal. Some squares are solid, while others are outlined. Thin, light-colored vertical lines of varying lengths are also present, some extending from the top or bottom edges towards the center. The overall aesthetic is modern and minimalist.

# THANKS