

cimri.com

*SmartPhone  
Price Prediction  
Project*



FURKAN ŞÜKRÜ KÜTÜK  
UĞUR SELİM ÖZEN

# ***Problem***

“ We are data analytics consulting team that serve on marketplace domains.  
Our client requests **building a prediction model** for **actual smartphone prices**. ”

# ***Methodology***

- Data Collection with Web Scraping
- Data Cleaning and Transformation
- Exploratory Data Analysis (EDA)
- Feature Selection and Modeling
- Model Prediction
- Interpreting Results



## *Utilized Technologies*

## *Data Overview*

- Observation : **1282**
- Feature : **18**
- Target : **PRICE**
- Data Size : **318 KB**



# *Web Scraping Process*

1

- `getAndParseURL()`

2

- `getPageLinks()`

3

- `getProductLinks()`

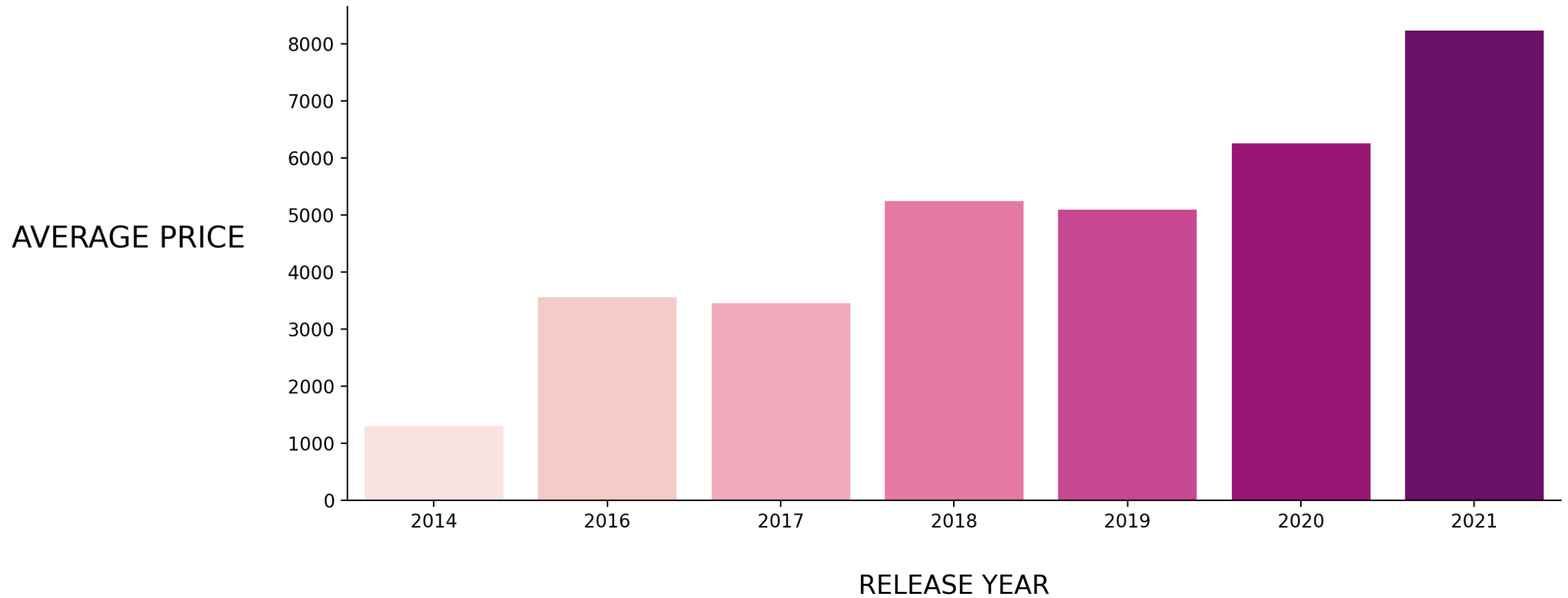
4

- `getDataFrame()`

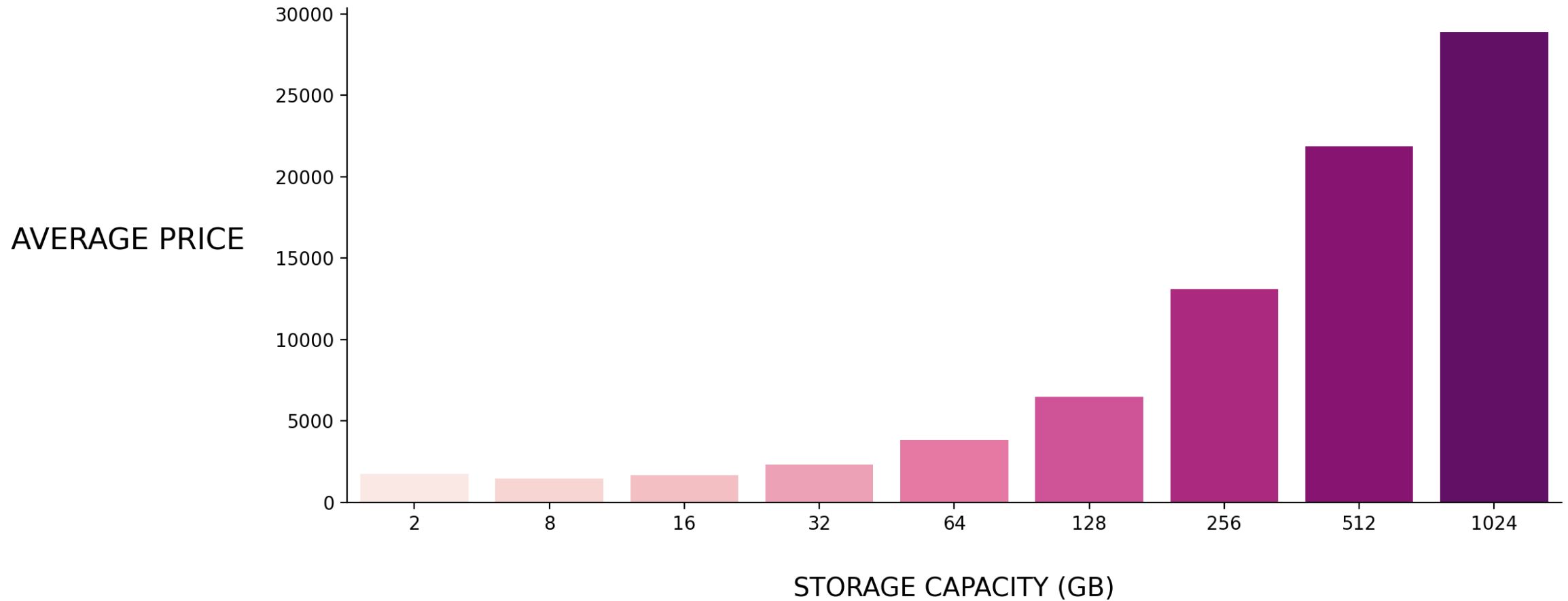
5

- `getResult()`

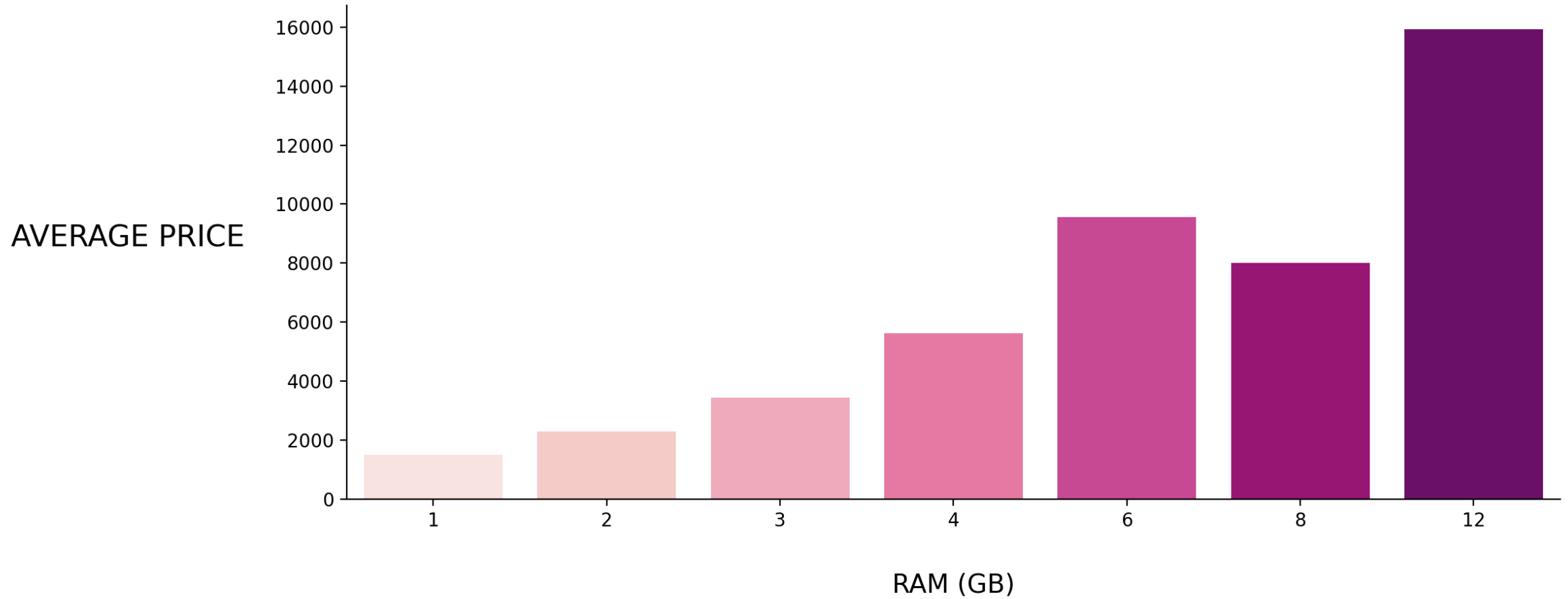
# ***EDA - I***



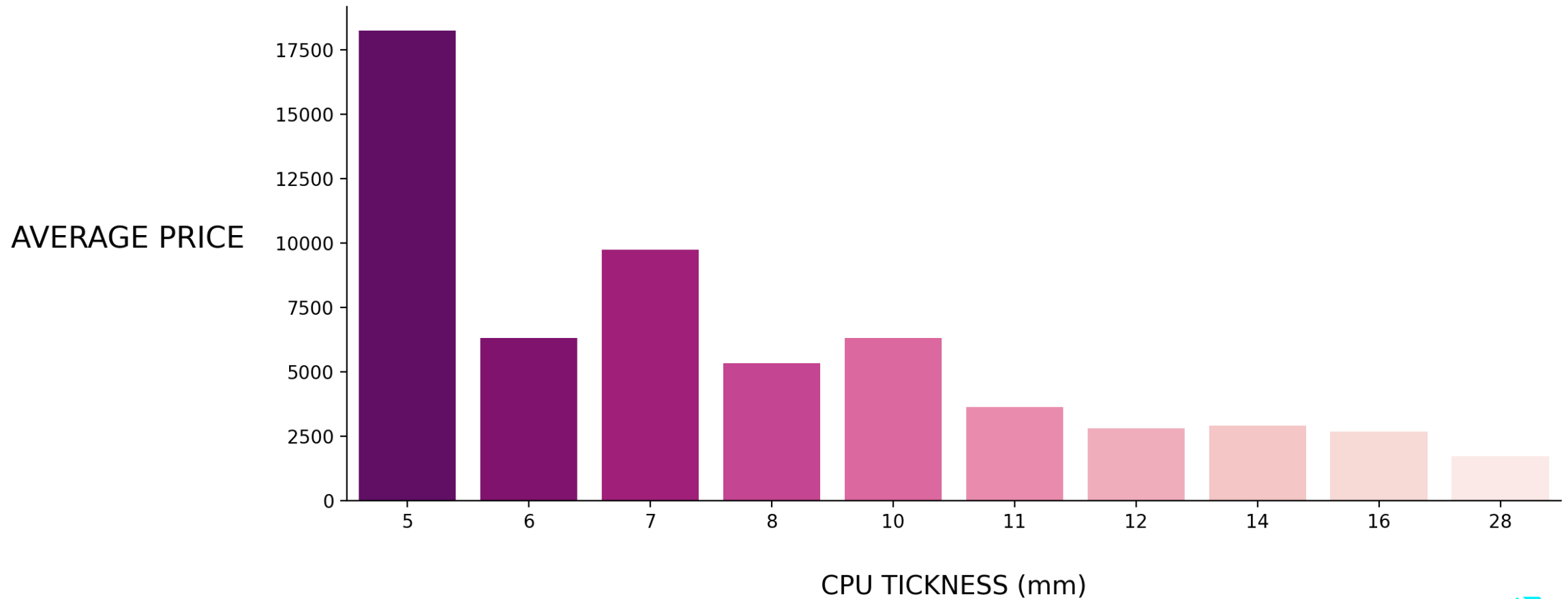
# ***EDA - 2***



# ***EDA - 3***

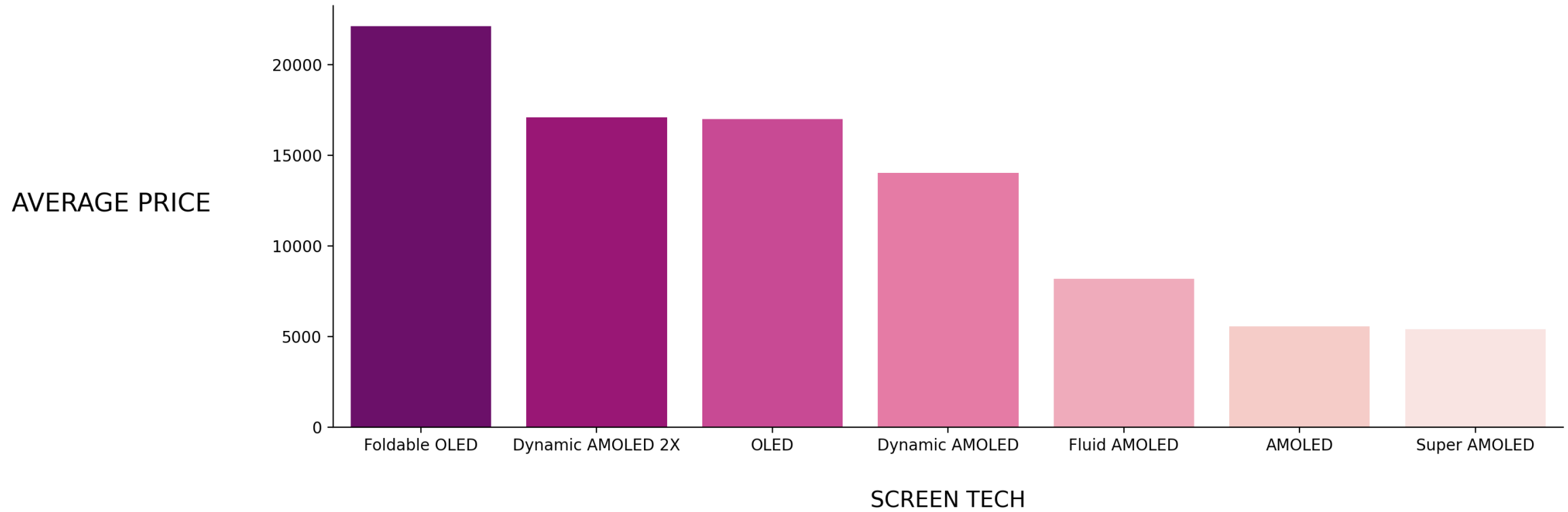


# ***EDA - 4***



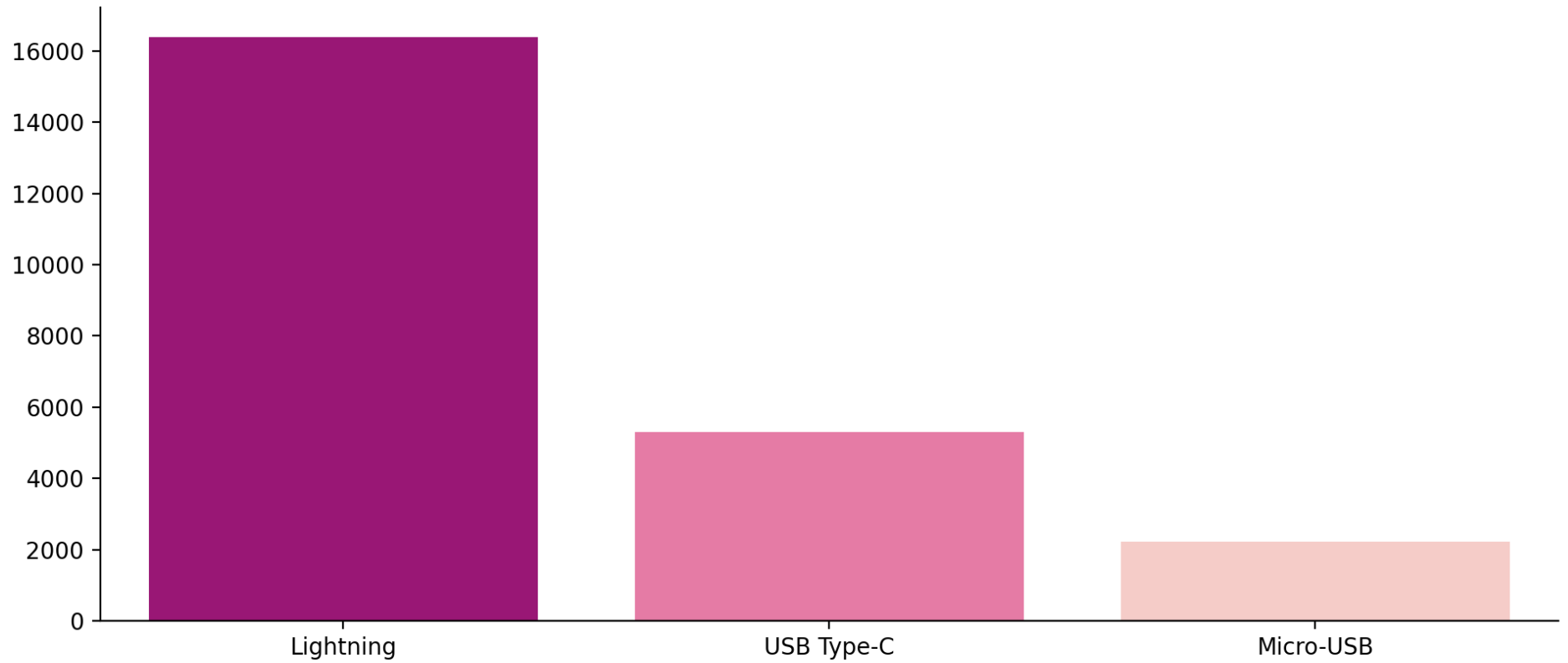


# ***EDA-5***



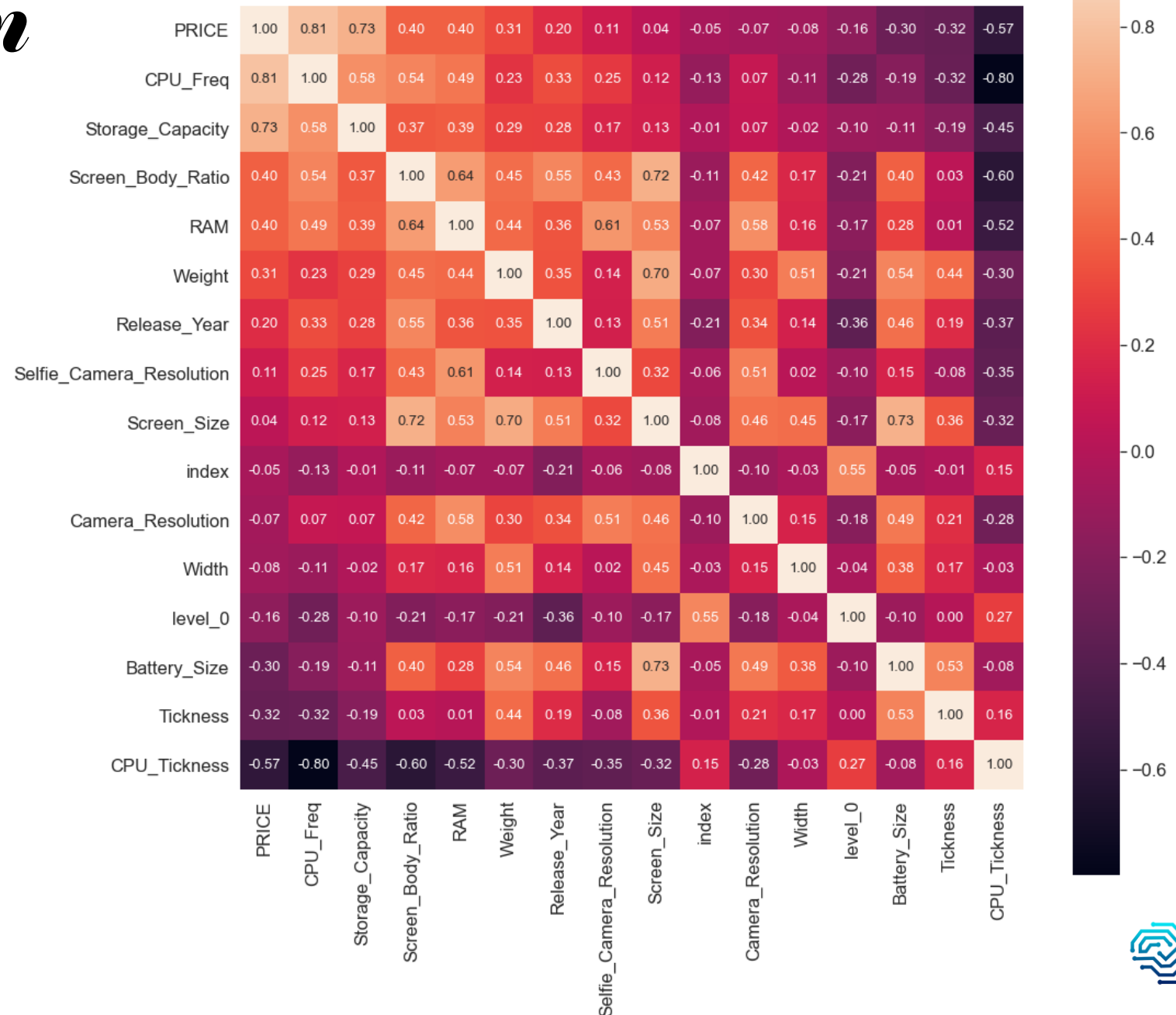
# ***EDA - 6***

AVERAGE PRICE



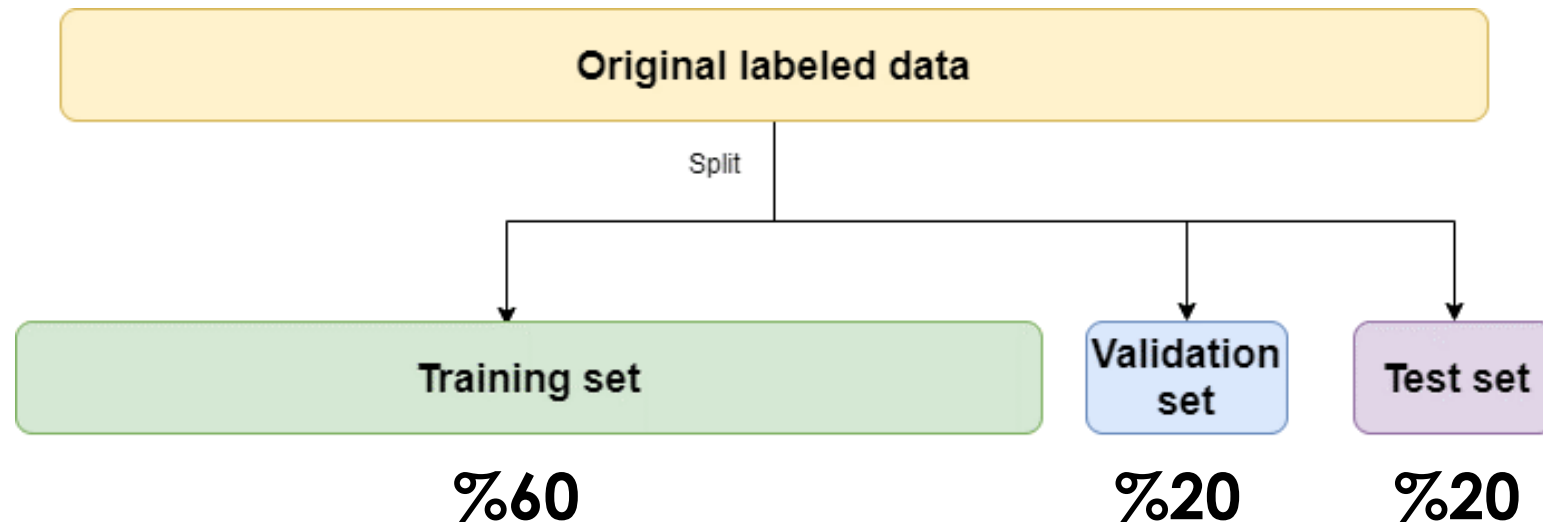
CHARGE TYPE

# Correlation Heatmap



# ***Data Modeling & Feature Selection***

- **Sorting** dataset **by product release year** for testing built model with the most up-to-date data.
- One-Hot Encoded Features
- **Screen\_Tech** , **CPU\_Model** , **GPU\_Model** , **Operating\_System** , **Charge\_Type**



# ***Model Prediction Test Results***

Model Algorithm	R2-Score (%60 Train)	R2-Score (%80 Train)	MSE (%60 Train)	MSE (%80 Train)	10-Fold Cross Validation
Linear Regression	0.923	0.863	5461971	9759771	0.917
<b>Ridge Regression</b>	<b>0.941</b>	<b>0.954</b>	<b>4206589</b>	<b>3261140</b>	<b>0.950</b>
Lasso Regression	0.942	0.950	4757696	3565501	0.948
Elasticnet Regression	0.825	0.849	12468972	10791699	0.821

*Thank you !*

*“ Ne kadar zor,  
işte bu kadar zor ! ”*

**Zekeriya Beşiroğlu**