cimri.com

SmartPhone Price Prediction Project

FURKAN ŞÜKRÜ KÜTÜK
UĞUR SELİM ÖZEN

ISTANBUL DATA SCIENCE ACADEMY

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 Tranformation
- Exploratory Data Analysis (EDA)
- Feature Selection and Modeling
- Model Prediction
- Interpreting Results

ISTANBUL DATA SCIENCE ACADEMY

Utilized Technologies

Data Overview

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

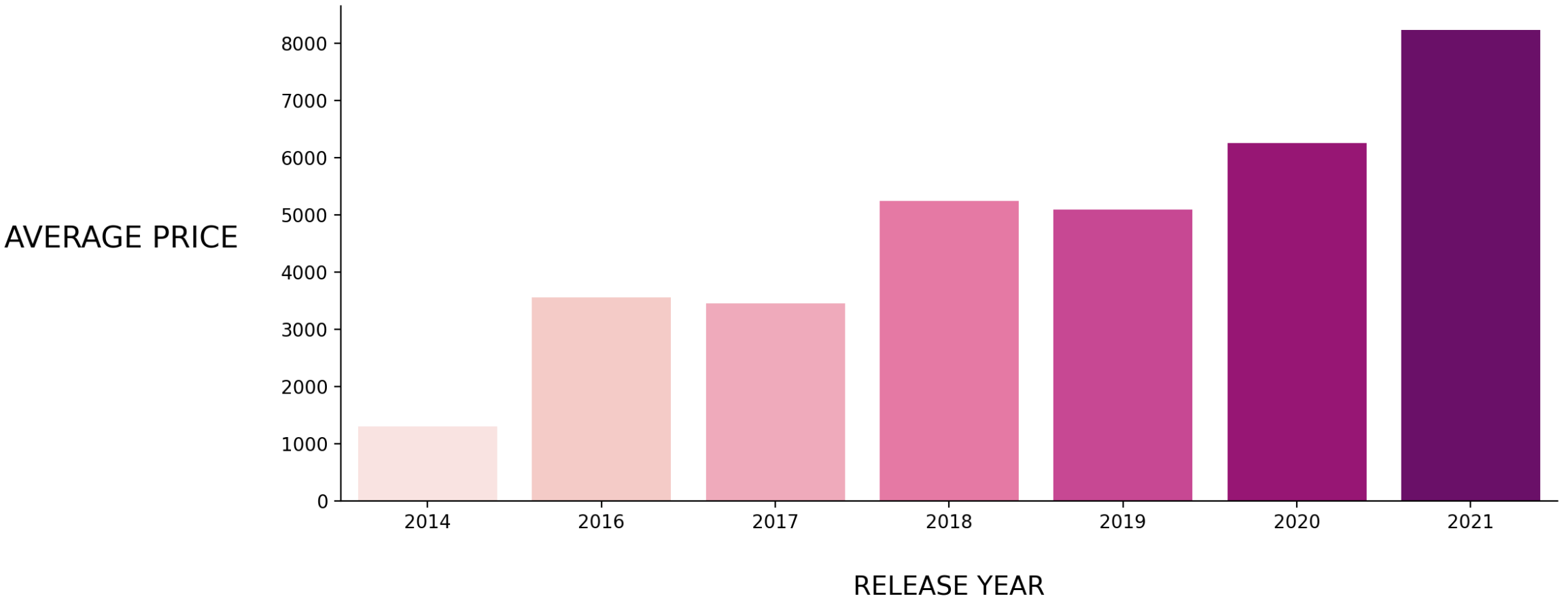
ISTANBUL DATA SCIENCE ACADEMY

Web Scraping Process

1	• getAndParseURL()
2	• getPageLinks()
3	• getProductLinks()
4	• getDataFrame()
5	• getResult()

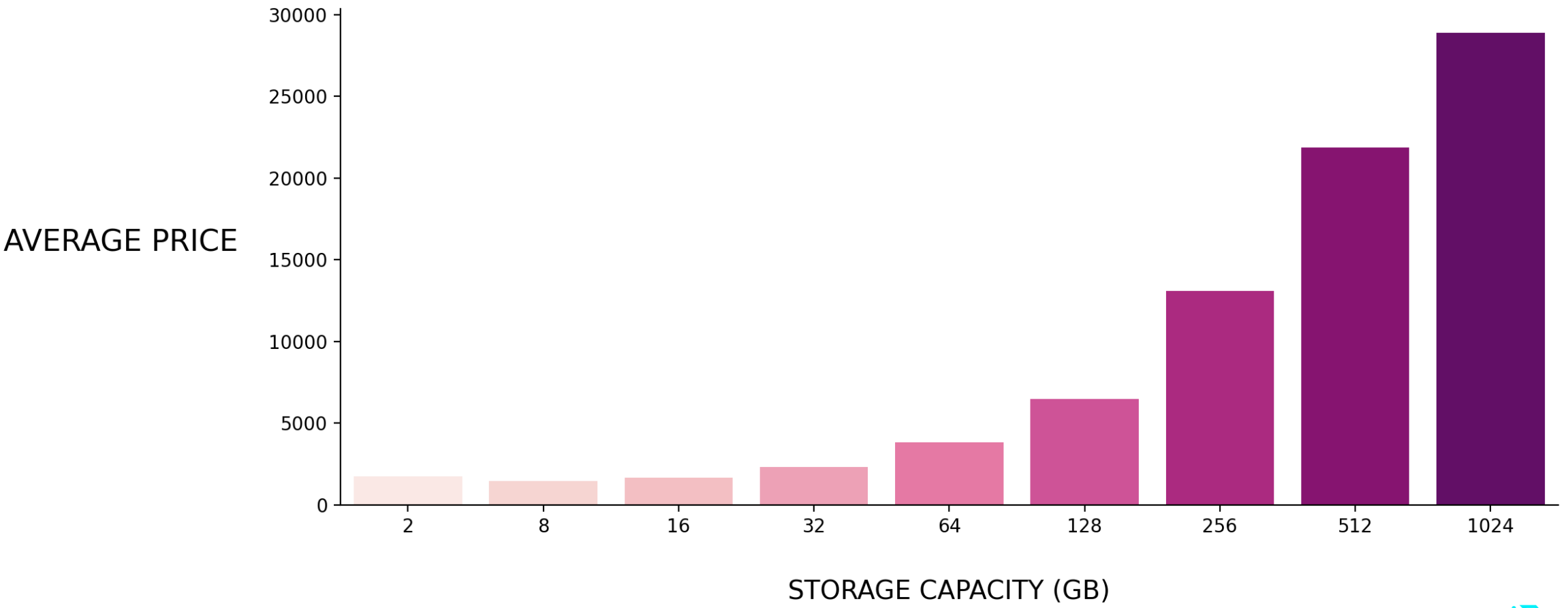
ISTANBUL DATA SCIENCE ACADEMY

EDA - 1



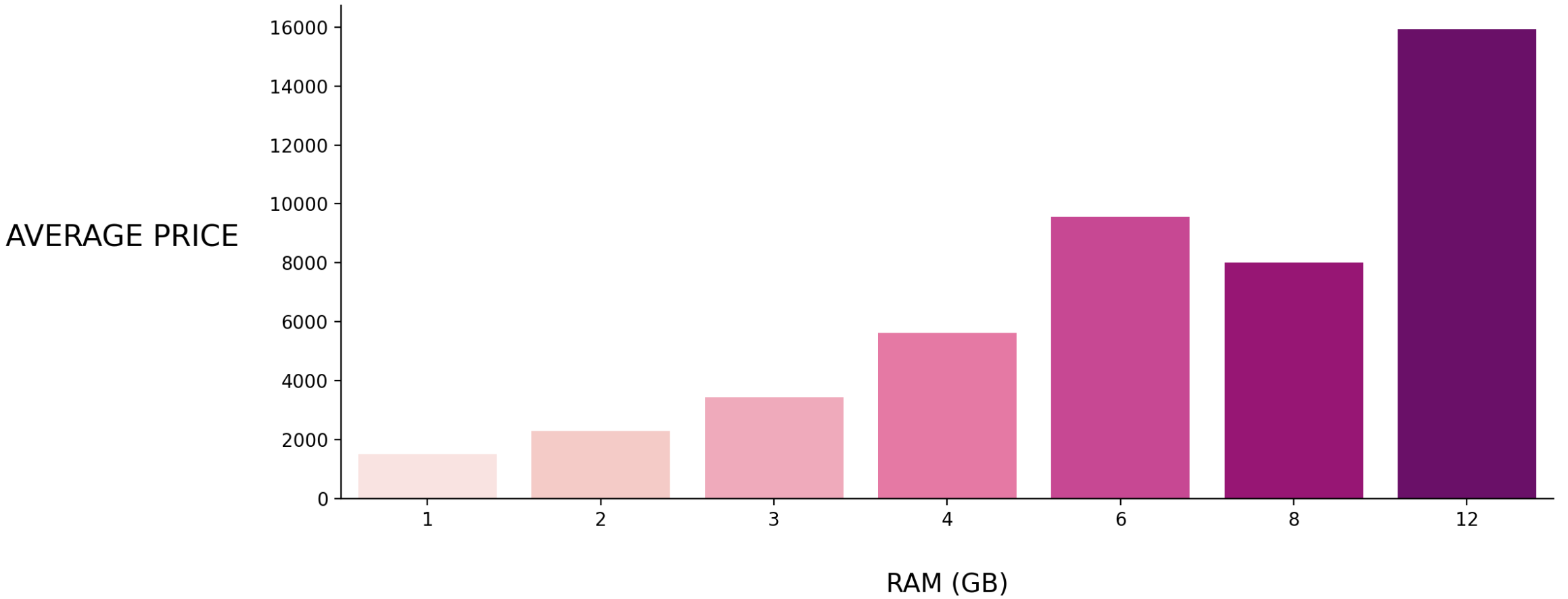
ISTANBUL DATA SCIENCE ACADEMY

EDA - 2



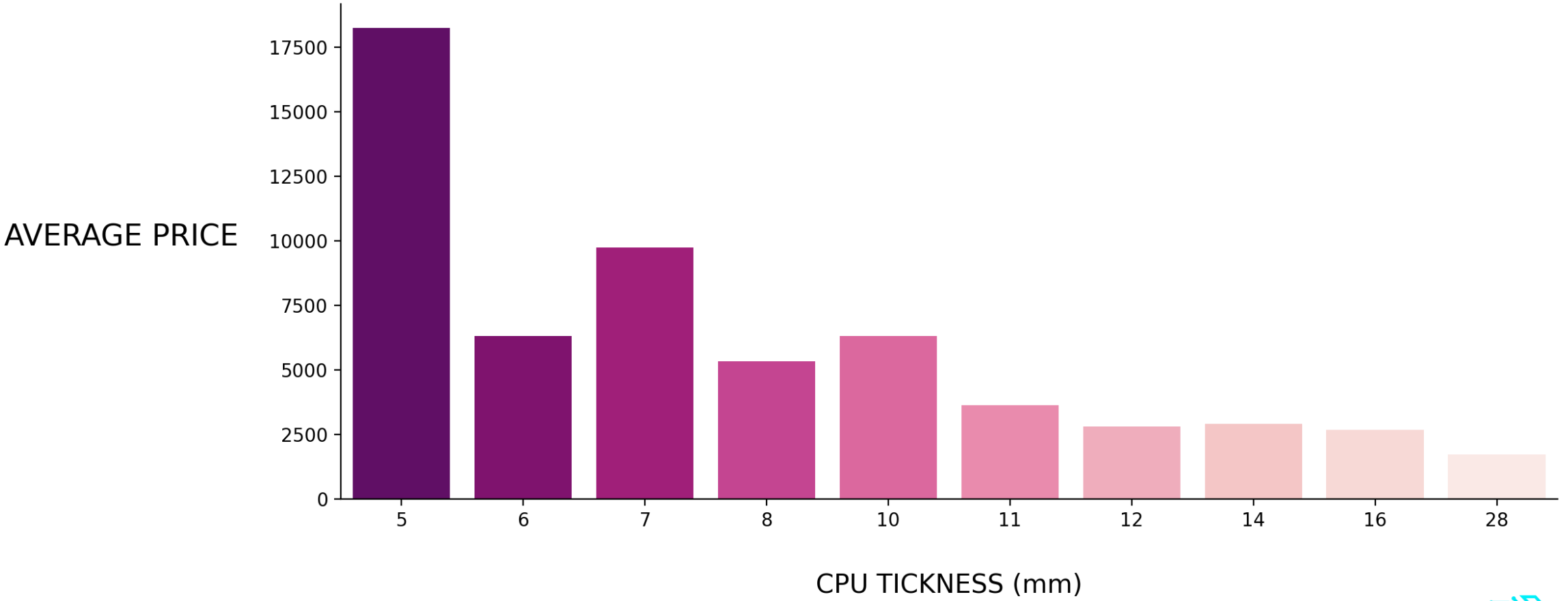
ISTANBUL DATA SCIENCE ACADEMY

EDA - 3



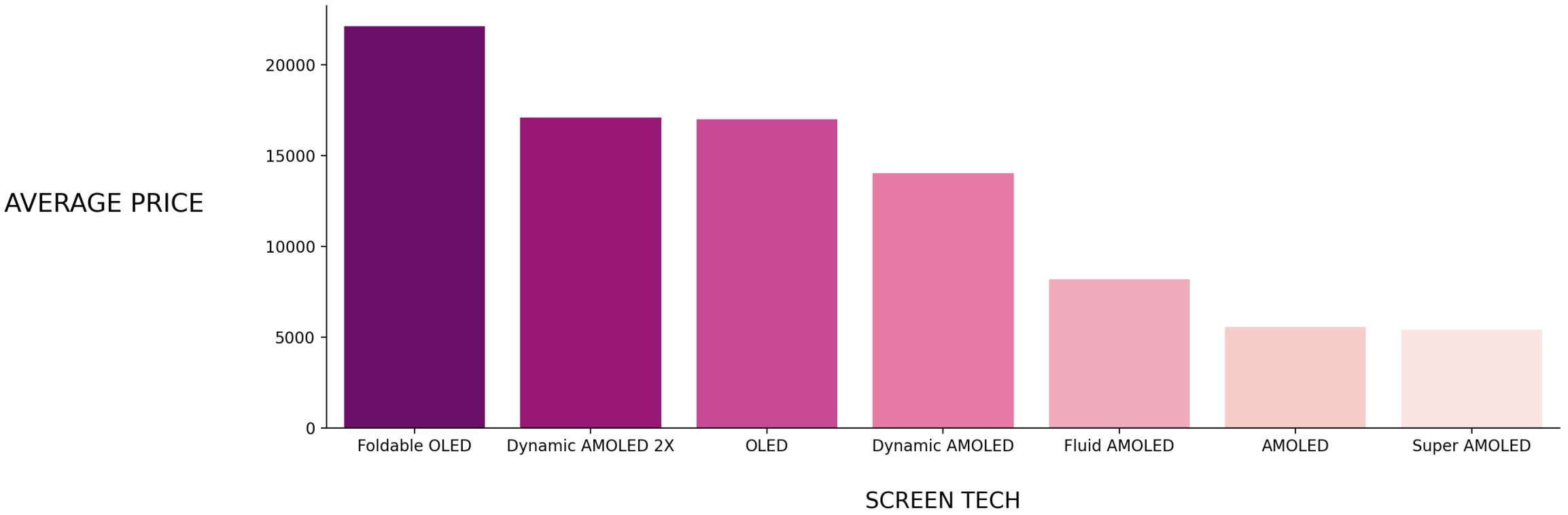
ISTANBUL DATA SCIENCE ACADEMY

EDA – 4



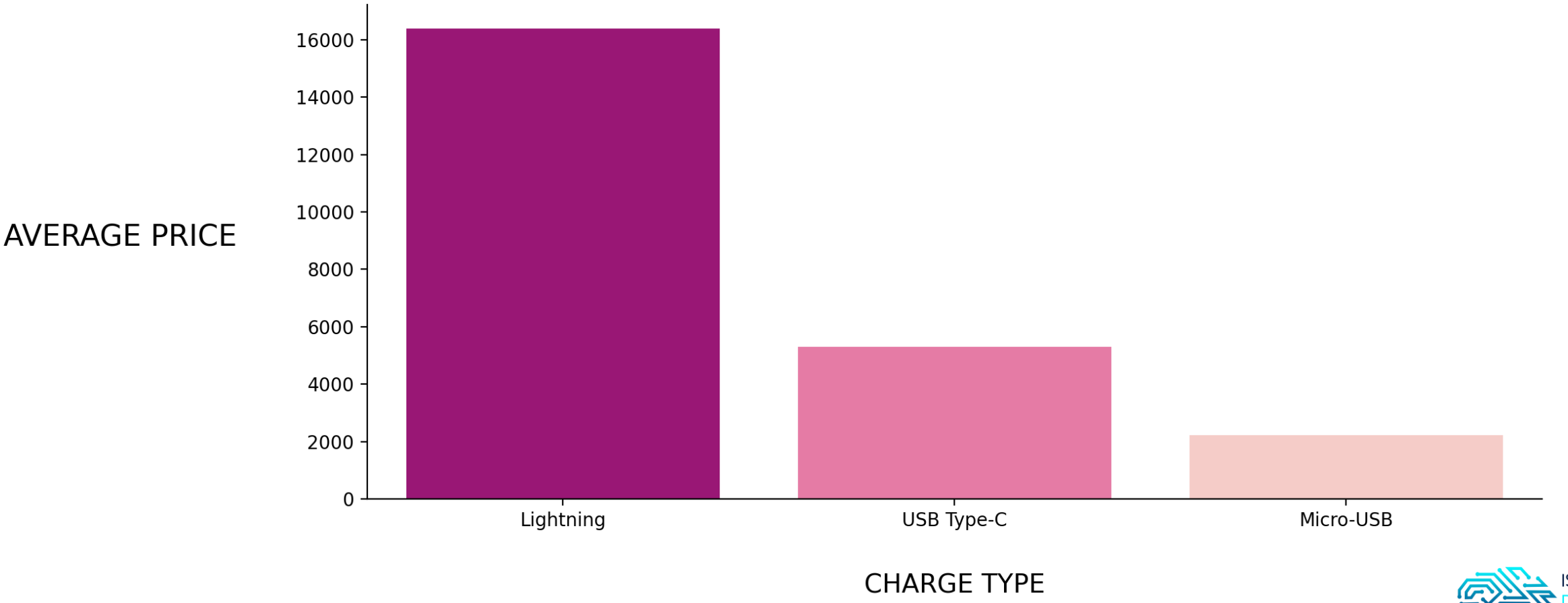
ISTANBUL DATA SCIENCE ACADEMY

EDA – 5



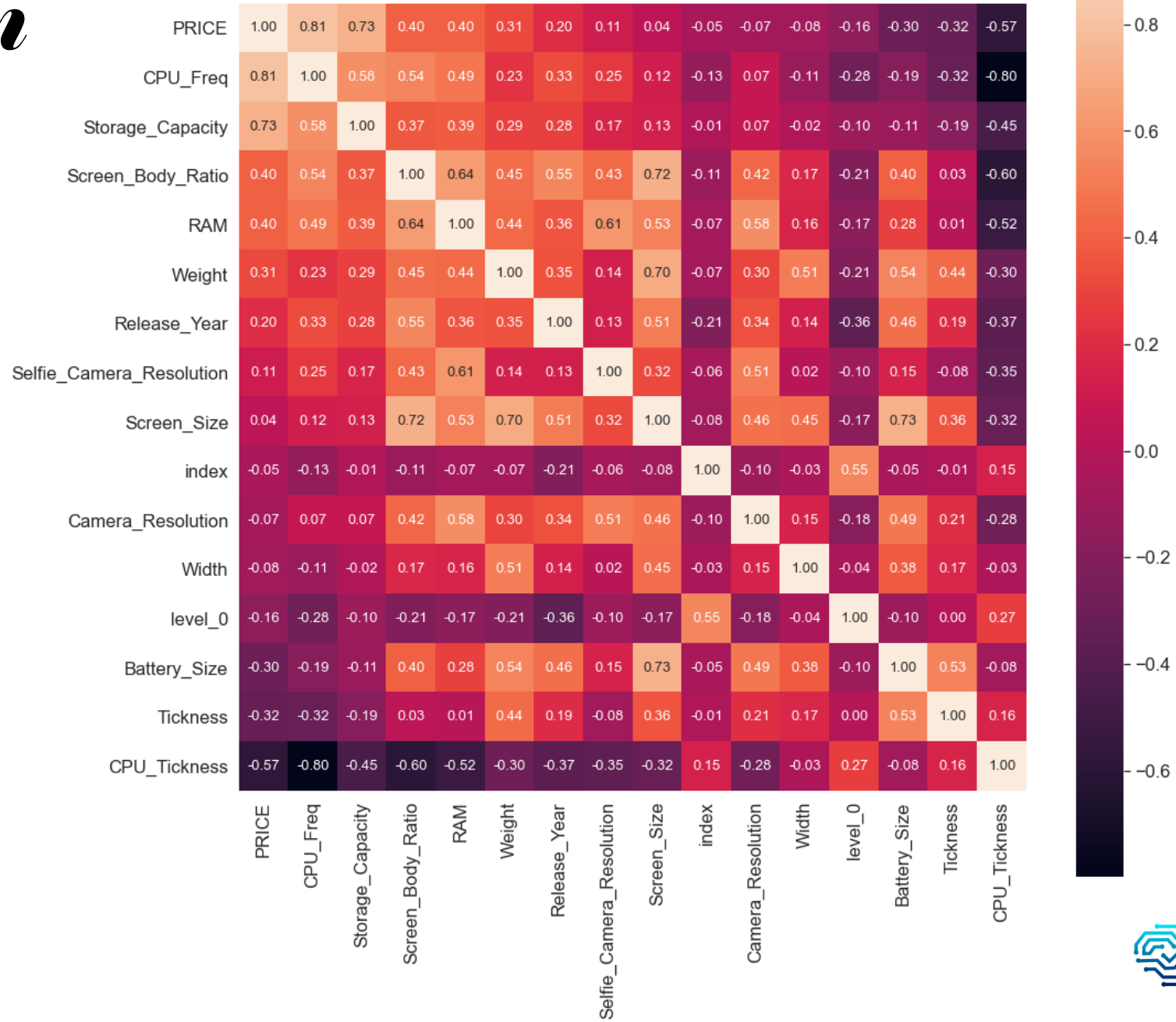
ISTANBUL DATA SCIENCE ACADEMY

EDA – 6



ISTANBUL DATA SCIENCE ACADEMY

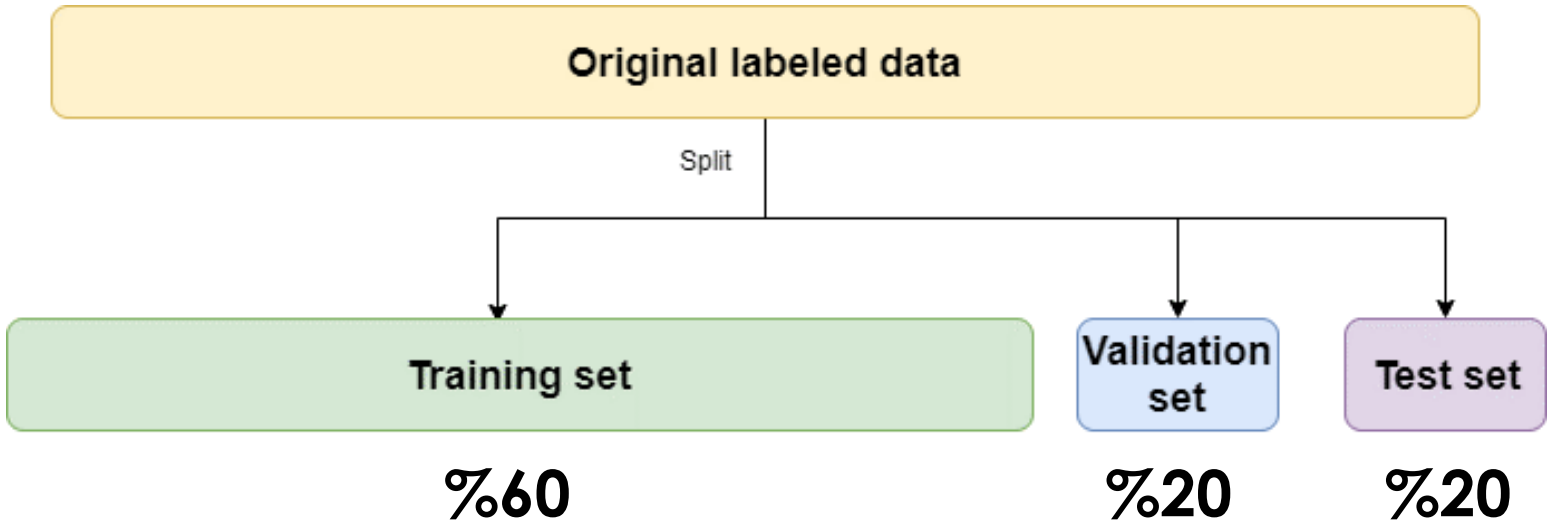
Correlation Heatmap



ISTANBUL DATA SCIENCE ACADEMY

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**



ISTANBUL DATA SCIENCE ACADEMY

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
Ridge Regression	**0.941**	**0.954**	**4206589**	**3261140**	**0.950**
Lasso Regression	0.942	0.950	4757696	3565501	0.948
Elasticnet Regression	0.825	0.849	12468972	10791699	0.821

ISTANBUL DATA SCIENCE ACADEMY

Thank you !

‘‘ Ne kadar zor,
işte bu kadar zor ! ’’

Zekeriya Beşiroğlu

ISTANBUL DATA SCIENCE ACADEMY