

# Web Scrapping ve Lineer Regresyon

**Hazırlayan: Nilüfer ÇATIKKAYA, Ümit EGBATAN, Tarık EMLİK**



# Amazon Sitesi Üzerinden Cep Telefonu Fiyat Tahminleme

- Apple
- Huawei
- Xiaomi
- Samsung
- Oppo
- General Mobile

Amozon sitesinde Web Scraping yaparak telefon özelliklerine fiyat tahminlemesi yapan Lineer Regresyon modeli oluşturduk.



https://www.amazon.com.tr/s?i=electronics&rh=n%3A12466496031%2Cn%3A13709880031%2Cn%3A13709907031%2Cp\_123%3A110955%7C32374%7C338933%7C46655%7C559198%7C584419&s=popularity-rank&dc&fs=true&ds=v1%3ABgDk1S0Q7ALazOmxyqPqm%2FCFLuixvXov5KBw09m%2BJoo&qid=1723377548&rnid=13709880031&ref=sr\_nr\_n\_3

amazon.com.tr/s?i=electronics&rh=n%3A12466496031%2Cn%3A13709880031%2Cn%3A13709907031%2Cp\_123%3A110955%7C32374%7C338933%7C46655%7C559198%7C584419&s=p... Gizli mod

Chrome otomatik test yazılımı tarafından kontrol ediliyor.

Tümü Çok Satanlar Günün Fırsatları Prime Yeni Çıkanlar Hediye Fikirleri Elektronik Moda Kitap Spor ve Açık Hava Aktiviteleri Amazon Ayrıcalıkları Kolay ve güvenli alışveriş keşfedin

1-24 / 153 sonuç Sıralama ölçüsü: Çok Satanlar

**Kargo Seçeneği**  
☐ 150 TL Üzeri Kargo BEDAVA  
Amazon tarafından gönderilen ürünlerde geçerlidir

**Teslimat Günü**  
☐ En Geç Yarın Teslimat

**Amazon Tarafından Gönderilir**  
☐ Amazon tarafından gönderilir

**Kategori**  
Elektronik  
Cep Telefonları ve Aksesuarlar  
Cep Telefonları

**Müşteri Yorumu**  
★★★★☆ ve Üzeri

**Markalar**  
Sil  
Apple  
HUAWEI  
Xiaomi  
SAMSUNG  
OPPO  
GENERAL MOBILE  
Reeder  
Daha fazlasını gör

**Kulaklık Girişi**  
☐ 3.5 mm

**Fiyat**  
4.800 TL – 68.500 TL ve üzeri

**Sonuçlar**  
Diğer satın alma seçenekleri için ürün sayfalarını kontrol edin.

**Amazon'un Seçimi**

**Apple iPhone 13 (128 GB) - Gece Yarısı**  
Seçenekler: 3 boyut  
★★★★☆ 2.284  
Geçen ay 200 adetten fazla satın alındı  
37.349<sup>00</sup> TL  
Sepete ekle

**Samsung Galaxy A35 5G Akıllı Cep Telefonu. 8 GB RAM. 256 GB Hafıza, Siyah (Samsung Türkiye Garantili)**  
★★★★☆ 103  
Geçen ay 500 adetten fazla satın alındı  
14.799<sup>00</sup> TL  
Samsung Akıllı Telefonlarda Takas Fırsatı!  
Sepete ekle

**Samsung Galaxy A24 Akıllı Cep Telefonu 6 GB RAM, 128 GB Depolama, Siyah (Samsung Türkiye Garantili)**  
★★★★☆ 339  
Geçen ay 1 B adetten fazla satın alındı  
8.599<sup>00</sup> TL  
Samsung Akıllı Telefonlarda Takas Fırsatı!  
Sepete ekle

**Apple iPhone 15 (128 GB) - Mavi**  
Seçenekler: 2 boyut  
★★★★☆ 417  
Geçen ay 100 adetten fazla satın alındı  
49.989<sup>00</sup> TL  
Stokta sadece 9 adet kaldı.  
Sepete ekle

**Samsung Galaxy S23 Ultra 256 GB | 8 GB RAM Cep Telefonu, Siyah (Samsung Türkiye Garantili)**  
★★★★☆ 78  
Geçen ay 100 adetten fazla satın alındı  
44.799<sup>00</sup> TL  
Samsung Akıllı Telefonlarda Takas Fırsatı!  
Sepete ekle

amazon.com.tr/s?i=electronics&rh=n%3A12466496031%2Cn%3A13709880031%2Cn%3A13709907031%2Cp\_123%3A110955%7C32374%7C338933%7C46655%7C559198%7C584419&s=p... Gizli mod

Chrome otomatik test yazılımı tarafından kontrol ediliyor.

**On Fotoğraf Sensörü Çözünürlüğü**  
☐ En Fazla 6,9 MP  
☐ 7,0-9,9 MP  
☐ 13,0 MP ve Üzeri

**GPS**  
☐ Doğru

**Yenileme Hızı**  
☐ En Fazla 88 Hz  
☐ 89-104 Hz  
☐ 105-120 Hz

**İnsan Arayüzü Girişi**  
☐ Kadran  
☐ Tuş Takımı  
☐ Mikrofon  
☐ Dokunmatik Ekran

**SIM Kart Yuvası Sayısı**  
☐ Tek SIM  
☐ Çift SIM

**Bağlantı Türü**  
☐ Lightning  
☐ Mikro USB  
☐ USB Type C

**Çözünürlük**  
☐ 1080 x 1920  
☐ 1080 x 2400  
☐ 1920 x 1080

**Pili Açıklaması**  
☐ Lityum Polimer  
☐ Lityum-İyon

**Kablosuz Ağ Teknolojisi**  
☐ CDMA  
☐ GSM  
☐ LTE  
☐ Wi-Fi

Diğer satın alma seçenekleri  
8.569,00 TL (4 yeni ürün)

**Samsung Galaxy A05S Akıllı Cep Telefonu 4 GB RAM, 128 GB Depolama, GÜMÜŞ (Samsung Türkiye Garantili)**  
★★★★☆ 63  
Geçen ay 100 adetten fazla satın alındı  
7.349<sup>00</sup> TL  
Stokta sadece 10 adet kaldı.  
Sepete ekle

**Samsung Galaxy S21 FE 5G 2.Nesil Akıllı Cep Telefonu, 8 GB RAM, 128 GB Hafıza, Snapdragon 888 İşlemcili, Yeşil (Samsung...)**  
★★★★☆ 123  
Geçen ay 50 adetten fazla satın alındı  
16.499<sup>00</sup> TL  
Samsung Akıllı Telefonlarda Takas Fırsatı!  
Sepete ekle

**Samsung Galaxy S23 256 GB | 8 GB RAM Cep Telefonu, Cream (Samsung Türkiye Garantili)**  
★★★★☆ 74  
Geçen ay 50 adetten fazla satın alındı  
32.197<sup>46</sup> TL  
Sepete ekle

**Samsung Galaxy A54 Akıllı Cep Telefonu 5G 8 GB RAM, 256 GB Depolama, Siyah (Samsung Türkiye Garantili)**  
★★★★☆ 480  
Geçen ay 50 adetten fazla satın alındı  
16.498<sup>99</sup> TL  
Samsung Akıllı Telefonlarda Takas Fırsatı!  
Sepete ekle

Önceki 1 2 3 ... 7 Sonraki

**Yardıma mı ihtiyacınız var?**  
Arama sonuçlarınız ile ilgili bilgi almak için yardım sayfamızı ziyaret edin veya bize ulaşın

# BeautifulSoup ve Selenium

[114] :

	product_title	price	marka	isletim_sistemi	bellek_depolama_kapasitesi	ekran_boyutu	model_adi	puan	degerlendirme_sayisi
0	Apple iPhone 13 (128 GB) - Gece Yarısı	37149.0	Apple	iOS 14	128.0	NaN	iPhone 13 128 GB Gece Yarısı	4.7	2275.0
1	Samsung Galaxy A35 5G Akıllı Cep Telefonu 8 GB...	14834.0	SAMSUNG	Android 14	8.0	66.0	Galaxy A35 256	4.5	100.0
2	Samsung Galaxy A24 Akıllı Cep Telefonu 6 GB RA...	8599.0	SAMSUNG	Android 130	128.0	NaN	NaN	4.3	336.0
3	Apple iPhone 15 (128 GB) - Mavi	50799.0	Apple	iOs	128.0	61.0	iPhone 15	4.6	414.0
4	Samsung Galaxy S23 FE 8 GB RAM. 256 GB Hafıza....	24999.0	SAMSUNG	Android 130	8.0	64.0	SM-S711BZACTUR	4.4	104.0
...	...	...	...	...	...	...	...	...	...
114	Apple iPhone 14 Plus (128 GB) - Gece yarısı	46499.0	Apple	iOs	128.0	67.0	iPhone 14 Plus	4.6	361.0
116	Samsung Galaxy A05S Akıllı Cep Telefonu 4 GB R...	6919.0	SAMSUNG	Android 130	4.0	67.0	1	4.4	62.0
117	Oppo Reno11 F 5G 8 GB RAM 256 GB Cep Telefonu ...	17149.0	OPPO	Android 130	256.0	NaN	Oppo Reno 11F 5G	5.0	2.0
118	Samsung Galaxy S21 FE 5G 2Nesil Akıllı Cep Tel...	16499.0	SAMSUNG	Android 130	8.0	64.0	Galaxy S21 FE 5G	4.3	122.0
119	Samsung Galaxy A54 Akıllı Cep Telefonu 5G 8 GB...	16199.0	SAMSUNG	Android 130	256.0	64.0	NaN	4.6	480.0

105 rows x 9 columns

# Corelasyonlarını Hesaplama

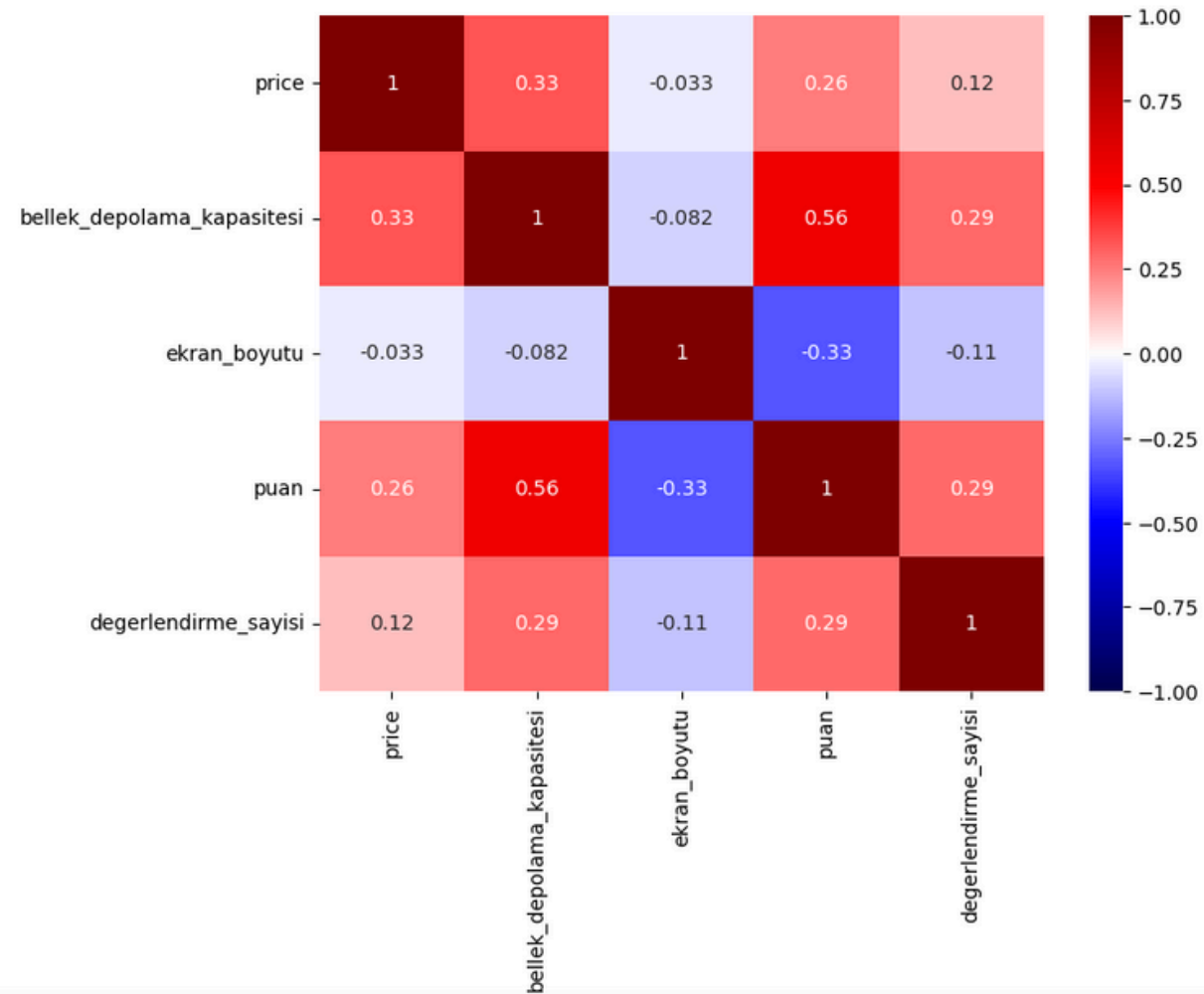
```
[134]: df2.corr()
```

```
[134]:
```

	price	bellek_depolama_kapasitesi	ekran_boyutu	puan	degerlendirme_sayisi
price	1.000000	0.330599	-0.032766	0.256734	0.118440
bellek_depolama_kapasitesi	0.330599	1.000000	-0.082199	0.555135	0.290518
ekran_boyutu	-0.032766	-0.082199	1.000000	-0.332296	-0.109941
puan	0.256734	0.555135	-0.332296	1.000000	0.286771
degerlendirme_sayisi	0.118440	0.290518	-0.109941	0.286771	1.000000

# Spearman ve Pearsona Heatmap

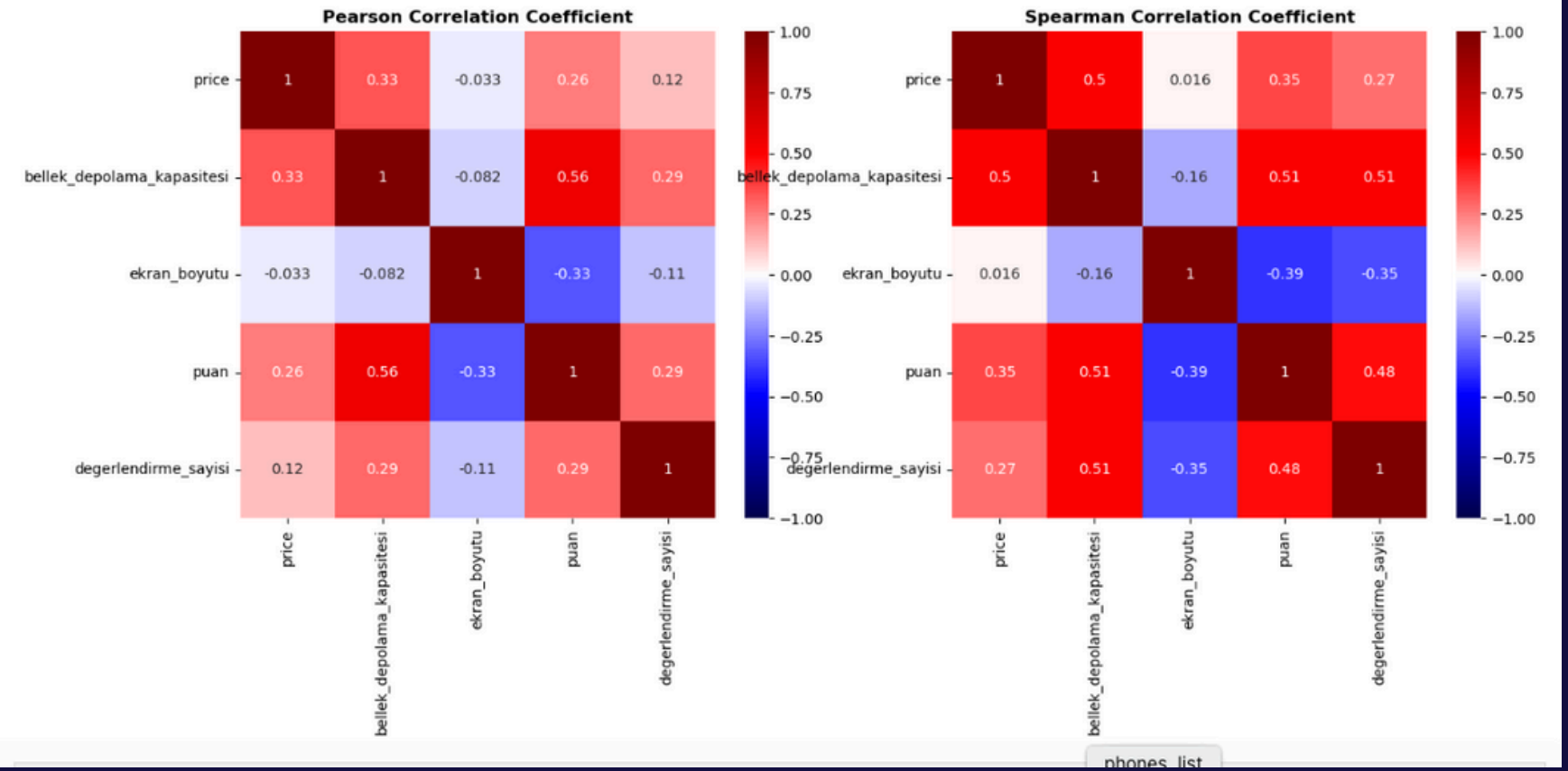
```
[137]: plt.figure(figsize=[8,6])
sns.heatmap(df2.corr(), cmap="seismic", annot=True, vmin=-1, vmax=1);
```



```
[139]: plt.figure(figsize=[16,6])

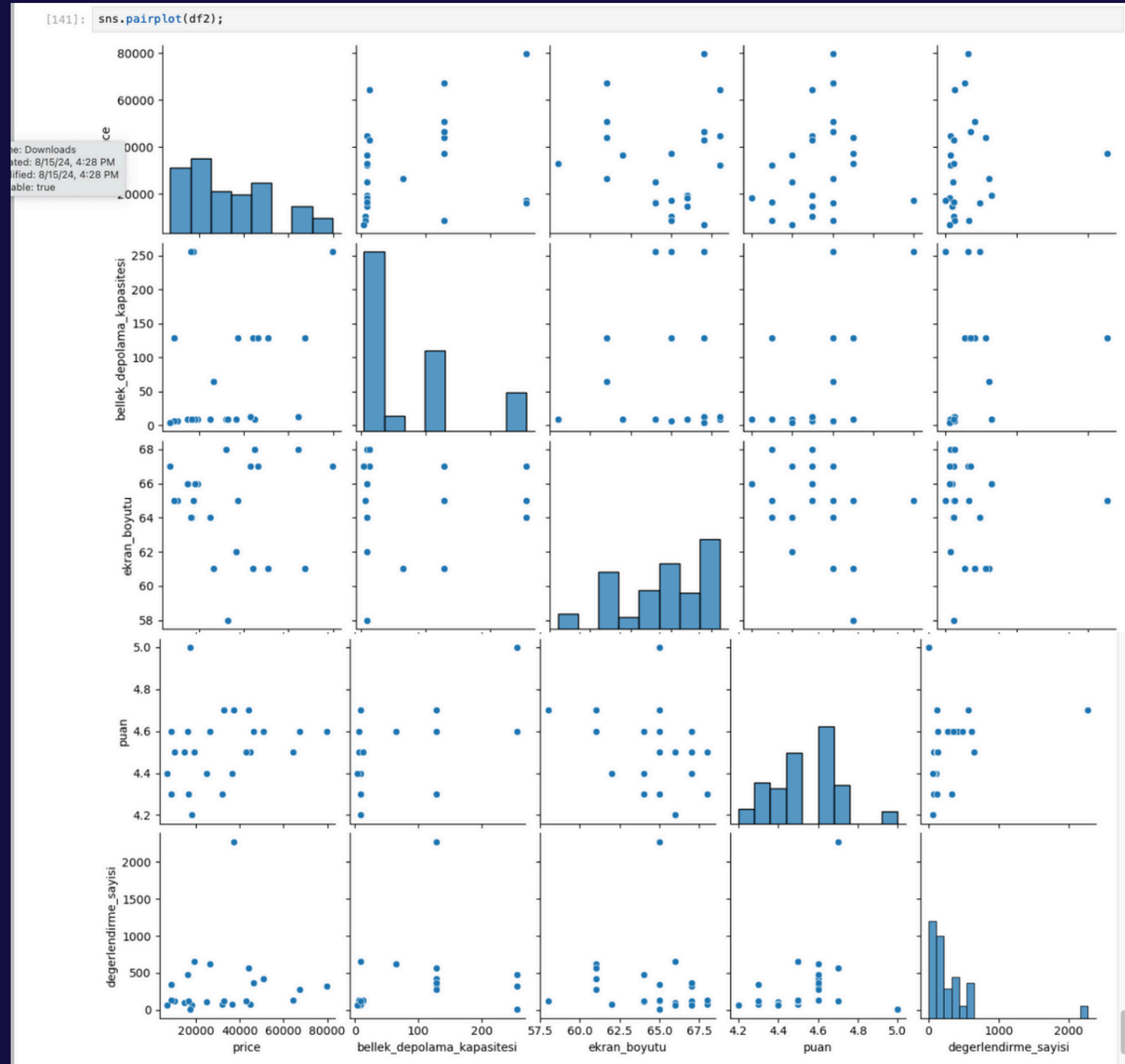
plt.subplot(1,2,1)
plt.title('Pearson Correlation Coefficient', weight='bold')
sns.heatmap(df2.corr(method='pearson'), cmap="seismic", annot=True, vmin=-1, vmax=1);

plt.subplot(1,2,2)
plt.title('Spearman Correlation Coefficient', weight='bold')
sns.heatmap(df2.corr(method='spearman'), cmap="seismic", annot=True, vmin=-1, vmax=1);
```





# Pair Plot



# Model Oluşturma Statsmodels

```
[225...] #Statsmodels Kullanarak Modelleme
# Değişkenlerin tanımlaması
y, X = patsy.dmatrices('price ~ bellek_depolama_kapasitesi + ekran_boyutu + puan + degerlendirme_sayisi', data=df2, return_type="dataframe")
# Modeli Oluşturma
model = sm.OLS(y, X)
# Modeli Eğiştirme
fit = model.fit()
fit.summary()
```

[225]:

## OLS Regression Results

Dep. Variable:	price	R-squared:	0.118			
Model:	OLS	Adj. R-squared:	0.083			
Method:	Least Squares	F-statistic:	3.340			
Date:	Fri, 16 Aug 2024	Prob (F-statistic):	0.0130			
Time:	16:03:59	Log-Likelihood:	-1182.7			
No. Observations:	105	AIC:	2375.			
Df Residuals:	100	BIC:	2389.			
Df Model:	4					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	-4.9e+04	9.31e+04	-0.526	0.600	-2.34e+05	1.36e+05
bellek_depolama_kapasitesi	62.2156	27.079	2.298	0.024	8.491	115.940
ekran_boyutu	213.8875	748.151	0.286	0.776	-1270.422	1698.197
puan	1.392e+04	1.46e+04	0.952	0.343	-1.51e+04	4.29e+04
degerlendirme_sayisi	0.5251	4.620	0.114	0.910	-8.640	9.691
Omnibus:	5.477	Durbin-Watson:	1.378			
Prob(Omnibus):	0.065	Jarque-Bera (JB):	3.208			
Skew:	0.223	Prob(JB):	0.201			
Kurtosis:	2.269	Cond. No.	2.67e+04			



# Etkili Kolonlar ile Model Eğitme ve Sonuçları

```
[219]: # Modeli Oluşturma
lm1 = smf.ols('price ~ bellek_depolama_kapasitesi + puan ', data=df2)

fit2 = lm1.fit()

fit2.summary()
```

```
[219]:
```

OLS Regression Results								
Dep. Variable:	price	R-squared:	0.117					
Model:	OLS	Adj. R-squared:	0.100					
Method:	Least Squares	F-statistic:	6.760					
Date:	Fri, 16 Aug 2024	Prob (F-statistic):	0.00175					
Time:	15:44:47	Log-Likelihood:	-1182.7					
No. Observations:	105	AIC:	2371.					
Df Residuals:	102	BIC:	2379.					
Df Model:	2							
Covariance Type:	nonrobust							
	coef	std err	t	P> t	[0.025	0.975]		
Intercept	-2.979e+04	5.99e+04	-0.498	0.620	-1.49e+05	8.9e+04		
bellek_depolama_kapasitesi	63.7141	26.216	2.430	0.017	11.715	115.713		
puan	1.273e+04	1.35e+04	0.946	0.346	-1.4e+04	3.94e+04		
Omnibus:	5.194	Durbin-Watson:	1.392					
Prob(Omnibus):	0.075	Jarque-Bera (JB):	3.202					
Skew:	0.238	Prob(JB):	0.202					
Kurtosis:	2.288	Cond. No.	3.68e+03					

Notes:

# Sklearn Modelleme Kullanarak Modelleme

```
[ ]: #Sklearn Kullanarak Modelleme
```

```
[217]: # Modeli Oluşturma
lr1 = LinearRegression()

# Featureları Oluşturma
X = df2.iloc[:, 1:]

# Targetı Oluşturma
y = df2.iloc[:, 0]

# Modeli Eğitme
lr1.fit(X, y)

# R2 Değeri
r_squared1 = lr1.score(X,y)
print("R2:", round(r_squared1,3))

# Adjusted R2 Değeri
adjusted_r_squared1 = 1 - (1-r_squared1)*(len(y)-1)/(len(y)-X.shape[1]-1)
print("Adjusted R2:", round(adjusted_r_squared1,3))
```

R2: 0.118

Adjusted R2: 0.083

# Regularization ile oluşturduğumuz veri seti ve modelleme sonucu

```
[ ]: #Regularization
```

```
[157]: from sklearn.preprocessing import StandardScaler, MinMaxScaler, RobustScaler
```

```
[159]: std_scale = StandardScaler()
```

```
[179]: X_ss = std_scale.fit_transform(X)  
df6 = pd.DataFrame(X_ss)  
df6
```

```
[179]:
```

	0	1	2	3	4
0	0.0	0.650684	0.183566	1.061316	4.552797
1	0.0	-0.749831	0.554227	-0.136944	-0.477392
2	0.0	0.650684	0.183566	-1.335204	0.068413
3	0.0	0.650684	-1.299081	0.462186	0.248806
4	0.0	-0.749831	-0.187096	-0.736074	-0.468141
...	...	...	...	...	...
100	0.0	0.650684	0.924889	0.462186	0.126231
101	0.0	-0.796515	0.924889	-0.736074	-0.565275
102	0.0	2.144566	0.183566	2.858706	-0.704039
103	0.0	-0.749831	-0.187096	-1.335204	-0.426512
104	0.0	2.144566	-0.187096	0.462186	0.401446

105 rows x 5 columns

# Regularization ile oluşturduğumuz veri seti ve modelleme sonucu

```
[215]: lr3 = LinearRegression()
lr3.fit(X_ss, y)
r_squared1 = lr3.score(X_ss, y)
print("R2:", round(r_squared1, 3))

# Adjusted R2 Değeri
adjusted_r_squared1 = 1 - (1 - r_squared1) * (len(y) - 1) / (len(y) - X_ss.shape[1] - 1)
print("Adjusted R2:", round(adjusted_r_squared1, 3))

R2: 0.118
Adjusted R2: 0.073
```



# Neleri daha iyi yapabildik?

Web Scraping yaparak Amazon sitesi üzerinde elde ettiğimiz veri seti oldukça az geldiğinden dolayı  $R^2$  değerimiz çok düşük çıktı. Veri setindeki satır sayımız yüksek olsaydı daha güvenilir sonuçlar elde edebilirdik.

Dinlediğiniz  
İçin  
**Teşekkürler !**