REFRIGERATOR PRICE FORECAST

ML Project

Prepared by:

Berrin Göçer

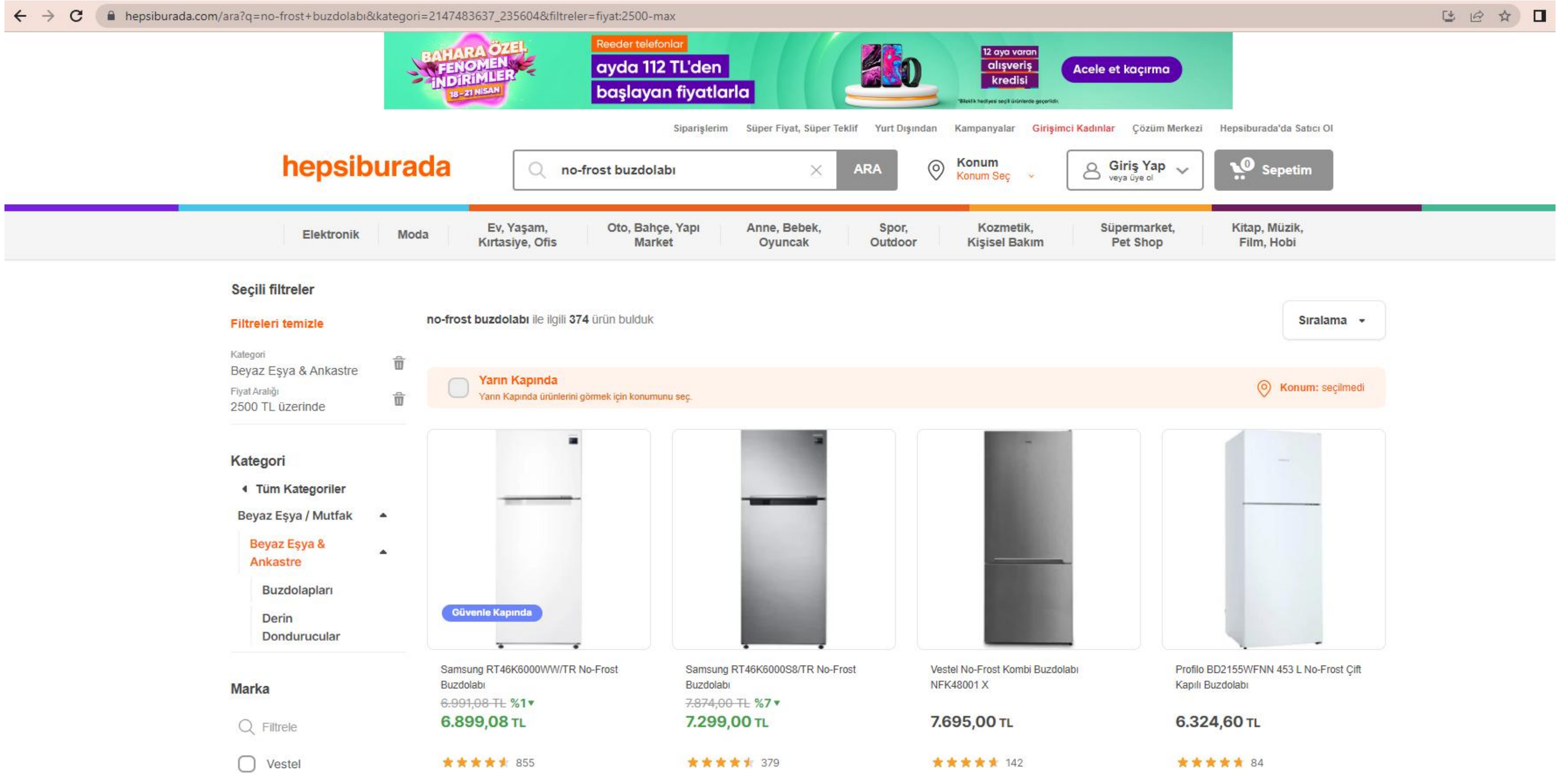
INTRODUCTION

PROBLEM	To make a price estimation according to product features for consumers who will buy refrigerators.
SOLUTION	Making the best estimation by scraping data from Hepsiburada website, which is one of the largest e-commerce sites in our country.
OBJECTIVE	Calculating the best R^2 score by establishing the best regression and making the best price estimation.

METHODOLOGY

DATA SOURCE	hepsiburada **URL:** ‘https://www.hepsiburada.com/ara?q=no-frost+buzdolab%C4%B1&kategori=2147483637_235604&filtreler=fiyat:2500-max’
TOOLS	jupyter, BeautifulSoup, NumPy, pandas, Scikit learn

ABOUT THE DATA



ABOUT THE DATA

	Urun_Adi	Uretici	Fiyat	Degerlendirme	Hacim	Enerji_Sinifi	Yillik_Enerji_Tuketimi	Renk
0	Samsung RT46K6000WW/TR No-Frost Buzdolabı	Samsung	6848	855	468.0	F	324	Beyaz
1	Samsung RT46K6000S8/TR No-Frost Buzdolabı	Samsung	7798	381	468.0	F	298	Inox
2	Vestel No-Frost Kombi Buzdolabı NFK48001 X	Vestel	7695	143	428.0	F	331	Gri
3	Profilo BD2155WFNN 453 L No-Frost Çift Kapılı ...	Profilo	6324	84	453.0	F	336	Beyaz
4	Vestel NF45001 No-Frost Buzdolabı	Vestel	6509	185	403.0	F	310	Beyaz
5	Bosch KGN86AIF0N 682 lt No-Frost Buzdolabı	Bosch	16885	13	619.0	E	378	Inox
6	Profilo BD3086WFDN 682 Lt No Frost Buzdolabı	Profilo	11025	54	682.0	F	384	Beyaz
7	Regal Nf 45010 402 Lt No-Frost Buzdolabı	Vestel	5298	55	402.0	F	310	Beyaz
8	Bosch KGN86AID1N 631 Lt No-Frost Buzdolabı	Bosch	18524	34	631.0	D	245	Inox
9	Vestel No-Frost Kombi Buzdolabı NFK48001	Vestel	7199	152	428.0	F	331	Beyaz

**First version of our dataset.

RESULTS

```
new_dff2["Renk"].replace({"Açık Gri": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Bej": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Gri": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Gümüş": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Inox": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Koyu Gri": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Siyah - Gri": "Gri"}, inplace=True)
new_dff2["Renk"].replace({"Bordo": "Renkli"}, inplace=True)
new_dff2["Renk"].replace({"Koyu Mavi": "Renkli"}, inplace=True)
new_dff2["Renk"].replace({"Kırmızı": "Renkli"}, inplace=True)
new_dff2["Renk"].replace({"Mavi": "Renkli"}, inplace=True)
new_dff2["Renk"].replace({"Turuncu": "Renkli"}, inplace=True)
new_dff2["Renk"].replace({"Açık Siyah": "Siyah"}, inplace=True)
new_dff2["Renk"].replace({"Siyah": "Siyah"}, inplace=True)
new_dff2["Renk"].replace({"Beyaz": "Beyaz"}, inplace=True)
```

```
new_dff34["Uretici"].replace({"Vestel": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Profilo": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Arçelik": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Beko": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Regal": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Altus": "Yerli"}, inplace=True)
new_dff34["Uretici"].replace({"Uğur": "Yerli"}, inplace=True)
```

```
scale_mapper = {"A":1, "D": 2,"E":3 ,'F':4}
new_dff55["Scale_Energy"] = new_dff55["Enerji_Sinifi"].replace(scale_mapper)
```

**We reorganized the values of some features on our dataset.

RESULTS

```
Data columns (total 19 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Yillik_Enerji_Tuketimi  356 non-null    int64  
 1   Degerlendirme           356 non-null    int64  
 2   Hacim                   356 non-null    float64
 3   Renk_Gri                356 non-null    uint8  
 4   Renk_Renkli             356 non-null    uint8  
 5   Renk_Siyah              356 non-null    uint8  
 6   Uretici_Electrolux      356 non-null    uint8  
 7   Uretici_Finlux          356 non-null    uint8  
 8   Uretici_Franke          356 non-null    uint8  
 9   Uretici_Hoover          356 non-null    uint8  
 10  Uretici_LG              356 non-null    uint8  
 11  Uretici_Liebherr        356 non-null    uint8  
 12  Uretici_Samsung         356 non-null    uint8  
 13  Uretici_Sharp           356 non-null    uint8  
 14  Uretici_Siemens         356 non-null    uint8  
 15  Uretici_Teka            356 non-null    uint8  
 16  Uretici_Yerli           356 non-null    uint8  
 17  Scale_Energy            356 non-null    int64  
 18  log_price               356 non-null    float64
```

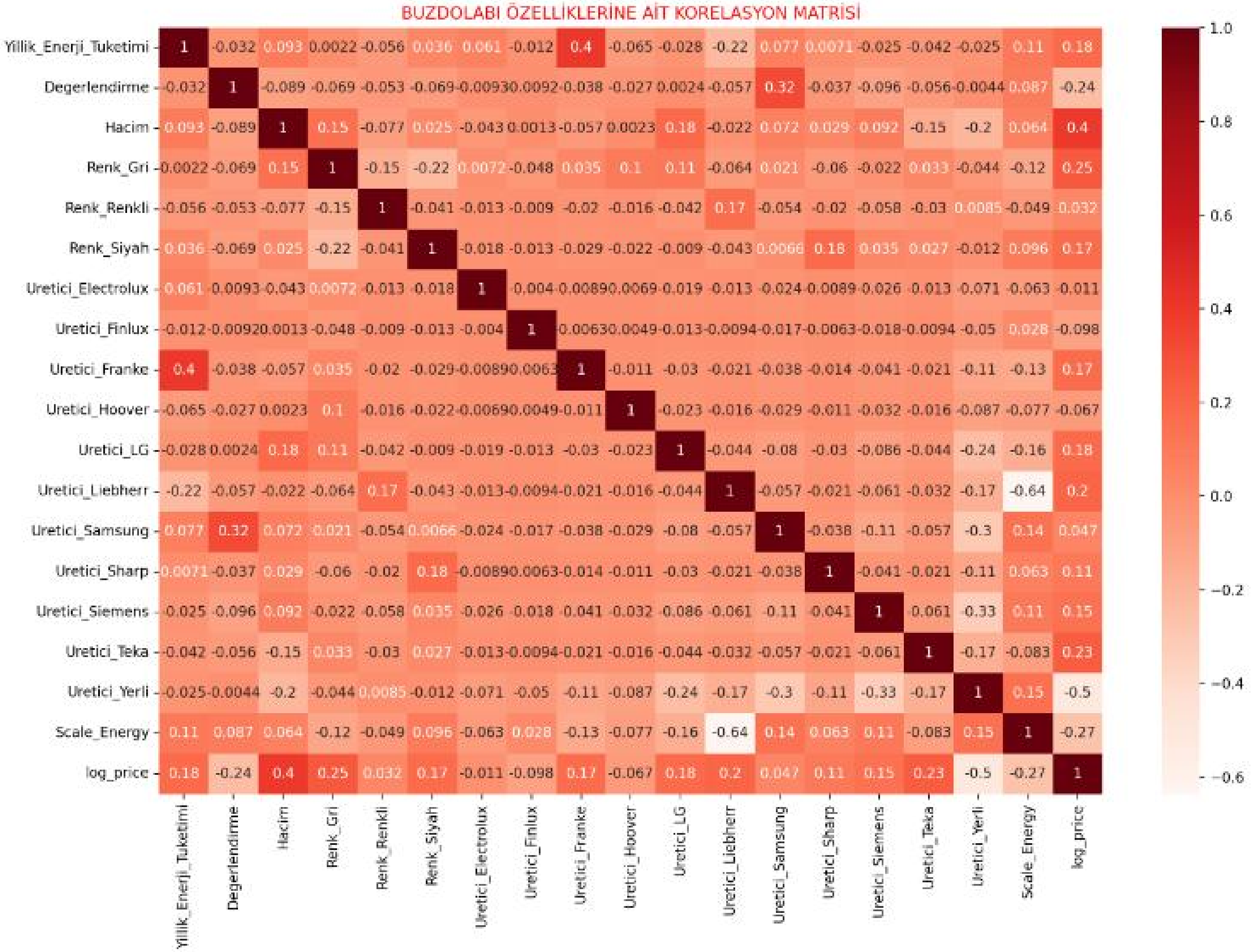
**We created dummies for Uretici and Renk feautures.

RESULTS

	Yillik_Enerji_Tuketimi	Degerlendirme	Hacim	Renk_Gri	Renk_Renkli	Renk_Siyah	Uretici_Electrolux	Uretici_Finlux	Uretici_Franke	Uretici_Hoover	Uretici_LG	Uretici_Liebherr	Uretici_Samsung	Uretici_Sharp	Uretici_Siemens	Uretici_Teka	Uretici_Yerli	Scale_Energy	log_price
Yillik_Enerji_Tuketimi	1.000000	-0.037940	0.273283	-0.035309	-0.059704	0.007863	0.068372	-0.161166	0.276188	-0.091408	-0.063476	-0.135202	0.005663	0.006055	-0.041747	-0.077055	0.107794	0.057422	0.045011
Degerlendirme	-0.037940	1.000000	-0.108961	-0.078711	-0.037432	-0.052893	-0.015017	-0.021517	-0.035111	-0.041662	0.011844	-0.040992	0.323527	-0.030476	-0.096069	-0.054944	-0.018738	0.063086	-0.257106
Hacim	0.273283	-0.108961	1.000000	0.194502	-0.056152	0.017200	-0.052939	-0.047863	-0.007461	0.002199	0.163532	-0.034642	0.059840	0.029386	0.128447	-0.208482	-0.125476	-0.063074	0.438830
Renk_Gri	-0.035309	-0.078711	0.194502	1.000000	-0.128742	-0.206228	0.055105	-0.090647	-0.021862	0.144031	0.043299	-0.058317	0.025340	0.055105	0.067686	0.003148	-0.112283	-0.078601	0.297834
Renk_Renkli	-0.059704	-0.037432	-0.056152	-0.128742	1.000000	-0.027461	-0.013957	-0.012070	-0.015627	-0.018543	-0.029320	0.452977	-0.039719	-0.013957	-0.047237	-0.024454	-0.044380	-0.205862	0.028640
Renk_Siyah	0.007863	-0.052893	0.017200	-0.206228	-0.027461	1.000000	-0.022358	-0.019335	-0.025032	-0.029703	0.084172	-0.029703	0.037043	-0.022358	-0.031866	0.115776	-0.043126	0.062562	0.197239
Uretici_Electrolux	0.068372	-0.015017	-0.052939	0.055105	-0.013957	-0.022358	1.000000	-0.009827	-0.012723	-0.015097	-0.023872	-0.015097	-0.032338	-0.011364	-0.038459	-0.019910	-0.107201	-0.052566	0.069377
Uretici_Finlux	-0.161166	-0.021517	-0.047863	-0.090647	-0.012070	-0.019335	-0.009827	1.000000	-0.011003	-0.013056	-0.020644	-0.013056	-0.027966	-0.009827	-0.033259	-0.017218	-0.092707	0.054029	-0.157913
Uretici_Franke	0.276188	-0.035111	-0.007461	-0.021862	-0.015627	-0.025032	-0.012723	-0.011003	1.000000	-0.016903	-0.026727	-0.016903	-0.036206	-0.012723	-0.043059	-0.022292	-0.120025	-0.136136	0.163018
Uretici_Hoover	-0.091408	-0.041662	0.002199	0.144031	-0.018543	-0.029703	-0.015097	-0.013056	-0.016903	1.000000	-0.031715	-0.020057	-0.042962	-0.015097	-0.051094	-0.026451	-0.142422	-0.120783	0.057597
Uretici_LG	-0.063476	0.011844	0.163532	0.043299	-0.029320	0.084172	-0.023872	-0.020644	-0.026727	-0.031715	1.000000	-0.031715	-0.067932	-0.023872	-0.080791	-0.041825	-0.225198	-0.153073	0.215561
Uretici_Liebherr	-0.135202	-0.040992	-0.034642	-0.058317	0.452977	-0.029703	-0.015097	-0.013056	-0.016903	-0.020057	-0.031715	1.000000	-0.042962	-0.015097	-0.051094	-0.026451	-0.142422	-0.528352	0.098784
Uretici_Samsung	0.005663	0.323527	0.059840	0.025340	-0.039719	0.037043	-0.032338	-0.027966	-0.036206	-0.042962	-0.067932	-0.042962	1.000000	-0.032338	-0.109443	-0.056658	-0.305065	0.075938	0.043277
Uretici_Sharp	0.006055	-0.030476	0.029386	0.055105	-0.013957	-0.022358	-0.011364	-0.009827	-0.012723	-0.015097	-0.023872	-0.015097	-0.032338	1.000000	-0.038459	-0.019910	-0.107201	0.062476	0.078020
Uretici_Siemens	-0.041747	-0.096069	0.128447	0.067686	-0.047237	-0.031866	-0.038459	-0.033259	-0.043059	-0.051094	-0.080791	-0.051094	-0.109443	-0.038459	1.000000	-0.067383	-0.362808	0.135472	0.206514
Uretici_Teka	-0.077055	-0.054944	-0.208482	0.003148	-0.024454	0.115776	-0.019910	-0.017218	-0.022292	-0.026451	-0.041825	-0.026451	-0.056658	-0.019910	-0.067383	1.000000	-0.187824	-0.114495	0.219555
Uretici_Yerli	0.107794	-0.018738	-0.125476	-0.112283	-0.044380	-0.043126	-0.107201	-0.092707	-0.120025	-0.142422	-0.225198	-0.142422	-0.305065	-0.107201	-0.362808	-0.187824	1.000000	0.160929	-0.470615
Scale_Energy	0.057422	0.063086	-0.063074	-0.078601	-0.205862	0.062562	-0.052566	0.054029	-0.136136	-0.120783	-0.153073	-0.528352	0.075938	0.062476	0.135472	-0.114495	0.160929	1.000000	-0.230208
log_price	0.045011	-0.257106	0.438830	0.297834	0.028640	0.197239	0.069377	-0.157913	0.163018	0.057597	0.215561	0.098784	0.043277	0.078020	0.206514	0.219555	-0.470615	-0.230208	1.000000

**We calculated the correlation values for our features.

BUZDOLABI ÖZELLİKLERİNE AİT KORELASYON MATRİSİ



CONCLUSION

```
from sklearn.linear_model import Ridge
from sklearn.datasets import make_blobs
model = Ridge(alpha=0.05, normalize=True)
model.fit(X, y)
new_input = [[256,105,524.0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,4]]
new_output = model.predict(new_input)
print(new_input, new_output)
```

```
[[256, 105, 524.0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 4]] [9.10725854]
```

Modeller ve Skor Değerleri:

LinearRegression_model_test: 47.66

LinearRegression2_model_test 50.36

ridgeReg_test 49.34

lassoReg_test -9.46

**The 3 features that are most correlated with log_price are ‘Uretici_Yerli’, ‘Hacim’, ‘Renk_Gri’

IF WE HAD MORE TIME

We could have analyzed all data in the refrigerator category, and not only no-frost refrigerators.

We could have analyzed other e-commerce websites with more products, and increased the number of our data’s.

The number of features can be increased on a per product basis.

THANK YOU FOR LISTENING