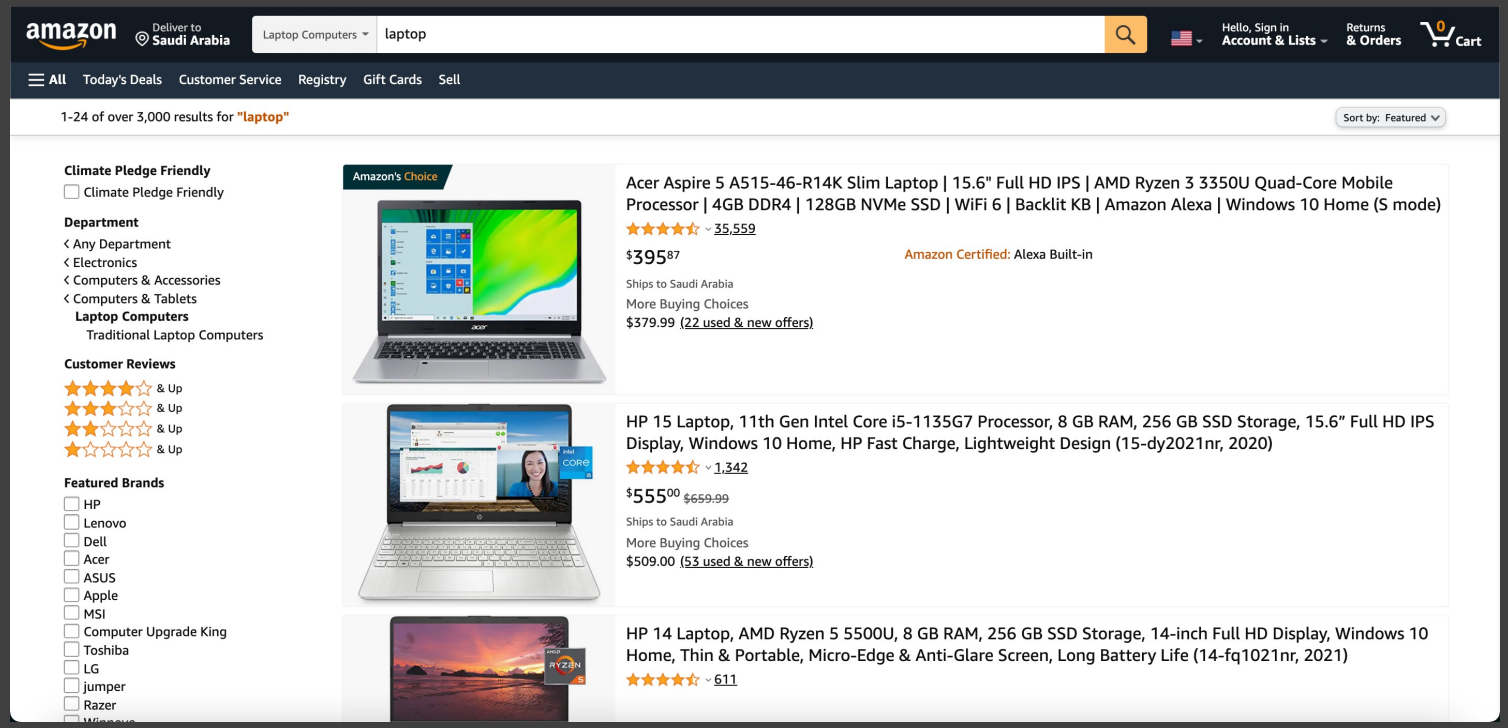
The Predicted Laptop price matches your needs By Amazon

Abdullah AlHuwaishel

Abdalla Alnujaidy

Tools

- Python
- Requests
- BeautifulSoup
- Webdriver
- Numpy & Pandans
- Sklearn
  - Linear Regression
- Matplotlib & Seaborn

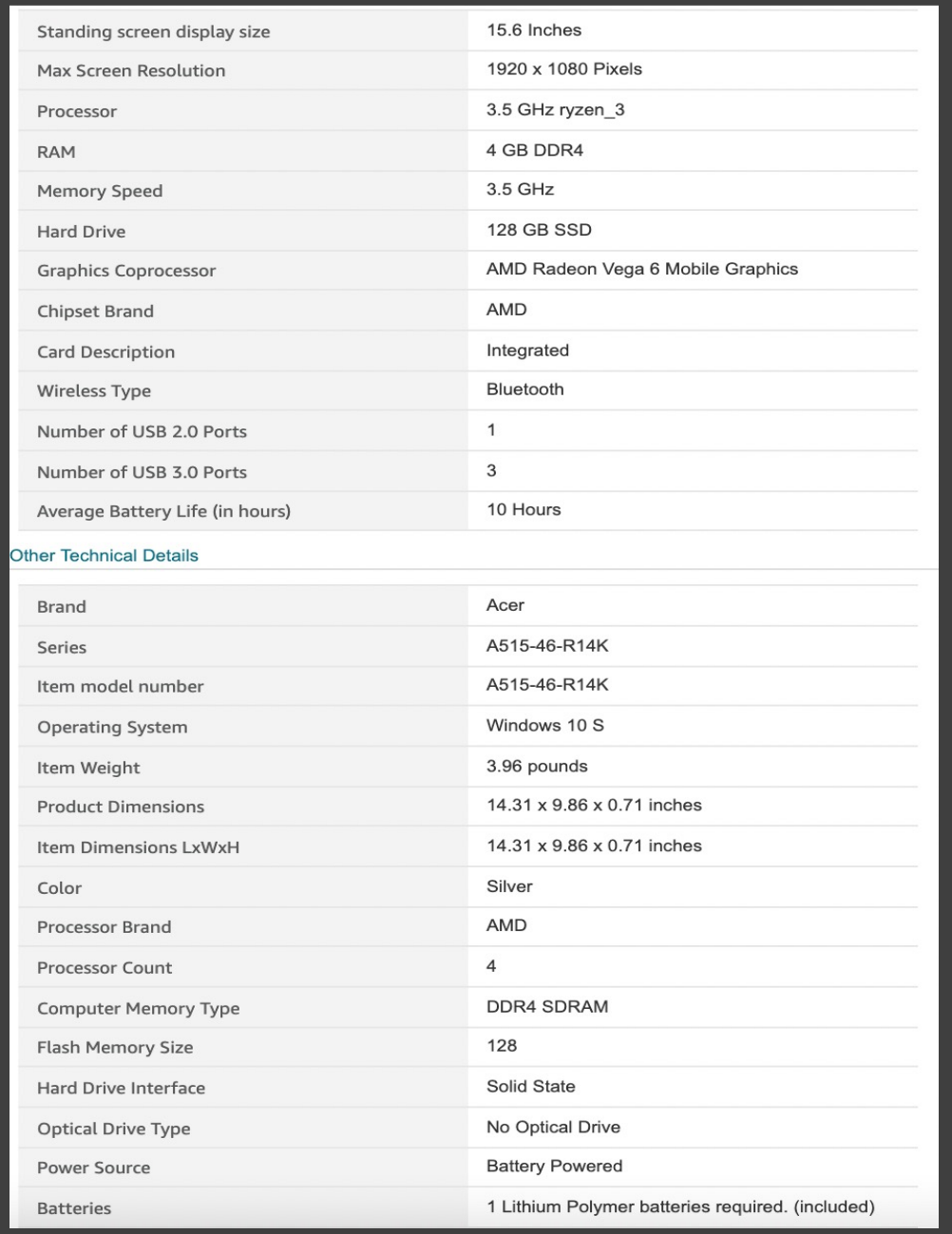


Features

- Price
- Brand
- Screen Size
- Screen Resolution
- Processor Speed
- RAM
- Hard drive
- Battery Life
- Weight
- Color
- Processor Brand
- Processor Cores

Observations

About 1758 Rows



Standing screen display size	15.6 Inches
Max Screen Resolution	1920 x 1080 Pixels
Processor	3.5 GHz ryzen_3
RAM	4 GB DDR4
Memory Speed	3.5 GHz
Hard Drive	128 GB SSD
Graphics Coprocessor	AMD Radeon Vega 6 Mobile Graphics
Chipset Brand	AMD
Card Description	Integrated
Wireless Type	Bluetooth
Number of USB 2.0 Ports	1
Number of USB 3.0 Ports	3
Average Battery Life (in hours)	10 Hours

Other Technical Details

Brand	Acer
Series	A515-46-R14K
Item model number	A515-46-R14K
Operating System	Windows 10 S
Item Weight	3.96 pounds
Product Dimensions	14.31 x 9.86 x 0.71 inches
Item Dimensions LxWxH	14.31 x 9.86 x 0.71 inches
Color	Silver
Processor Brand	AMD
Processor Count	4
Computer Memory Type	DDR4 SDRAM
Flash Memory Size	128
Hard Drive Interface	Solid State
Optical Drive Type	No Optical Drive
Power Source	Battery Powered
Batteries	1 Lithium Polymer batteries required. (included)

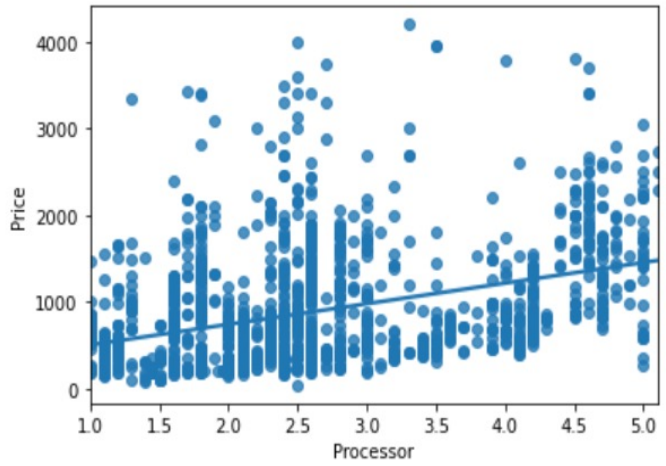
Linear Regression

```
In [282]: df_train, df_val = train_test_split(df2, test_size=0.25, random_state=3)

In [284]: m = LinearRegression()
          m.fit(df_train[['display_size','Processor','RAM','Hard_Drive','Battery_Life','Weight','Processor_Count_Cores','Resolutic
          m.score(df_train[['display_size','Processor','RAM','Hard_Drive','Battery_Life','Weight','Processor_Count_Cores','Resolut

Out[284]: 0.650107622529471

In [285]: m.score(df_val[['display_size','Processor','RAM','Hard_Drive','Battery_Life','Weight','Processor_Count_Cores','Resolutic

Out[285]: 0.6586609389637254

In [286]: m.coef_

Out[286]: array([ 1.35658795e+01,  6.05020599e+01,  2.88686391e+01,  1.87006976e-01,
                  2.62283629e+00,  2.99401938e+01,  3.54782662e+01,  1.44872464e+02,
                 -2.85589764e+01, -5.00592916e+01,  6.36170567e+02,  1.33698790e+02,
                  7.70621557e+02,  5.19951308e+02,  2.73377090e+02, -9.58486713e+01,
                  9.81788840e+01,  3.55833488e+02, -1.00464875e+02,  2.84576343e+02,
                 -1.20932544e+01,  1.65693666e+02, -1.80690880e+02,  6.97753598e+02,
                 -1.64062481e+02, -4.64088153e+02, -5.04997429e+02, -5.03747314e+02,
                 -5.26796799e+02, -4.27246113e+02, -5.09783995e+02, -5.13130604e+02,
                 -3.80179008e+01,  1.01868163e+02])

In [287]: m.intercept_

Out[287]: 120.4894821431759
```

Example

- Brand – Dell
- Processor speed – 2 GHz
- RAM – 8 GB
- Battery life – 10 Hr
- Hard drive – 1TB

```
In [291]: X = df2[['Brand_Dell','Processor','RAM','Battery_Life','Hard_Drive']]
          y = df2['Price']

In [292]: m.fit(X,y)

Out[292]: LinearRegression()

In [293]: ypred_P=m.predict(X)
          ypred_P

Out[293]: array([ 594.9639581 ,  903.70275212,  988.11612349, ...,  741.53504686,
                 1783.06720096,  729.34568891])

In [298]: sns.regplot(x="Processor", y="Price",ci=None, data=df2);
```




```
In [305]: (m.intercept_)+m.coef_[0]+m.coef_[1]*2+m.coef_[2]*8+m.coef_[3]*10+m.coef_[4]*1000

Out[305]: 1025.231863468368

In [306]: from sklearn.metrics import mean_squared_error, mean_absolute_error,r2_score

In [308]: r_sqr=r2_score(y,ypred_P)
          r_sqr

Out[308]: 0.5191871523578098
```

Conclusion