Predictive Analysis of Laptop Prices

1. **Chapter 1: Introduction**
    - **1.1 Background:** In the fast-paced technology market, laptops are essential and widely used devices. Consumers have a plethora of options, each with varying features and prices. Understanding the factors influencing laptop prices is crucial to make informed purchase decisions. Factors such as brand, specifications (CPU, RAM, resolution, etc.), and design contribute to the final price of a laptop.
    - **1.2 Objective:** The primary goal of this project is to build a predictive model that accurately estimates laptop prices based on their features. By analyzing historical data on laptops and their corresponding prices, we aim to develop a model that provides valuable insights for both consumers seeking the best value and manufacturers aiming to optimize pricing strategies.
2. **Chapter 2: Data Preprocessing**
    - **2.1 Data Acquisition:** The dataset used in this project was collected from reputable sources, including laptop specifications and their corresponding prices. The dataset includes information such as brand, specifications, weight, resolution, and more.
    - **2.2 Data Cleaning and Standardization:** Raw data is often messy and inconsistent. In this phase, we cleaned the dataset by handling missing values, removing duplicates, and addressing any data inconsistencies. Additionally, we standardized numerical features by scaling them to a similar range to avoid any bias during model training.
3. **Chapter 3: Outlier Detection and Handling**
    - **3.1 Outlier Identification:** Outliers, or unusual data points significantly different from the majority, can distort statistical analyses. We used various statistical methods, such as the Z-score or box plots, to identify outliers that might affect the accuracy of our predictive model.
    - **3.2 Outlier Removal:** After identifying outliers, we carefully removed them from the dataset to prevent them from unduly influencing the results of our analysis and subsequent predictive model.
4. **Chapter 4: Correlation Analysis**
    - **4.1 Correlation Calculation:** Correlation analysis helps us understand the relationships between different features and the target variable, price. We calculated the correlation matrix to quantify and evaluate the strength and direction of these relationships.
    - **4.2 Visualization:** The correlation matrix was then visualized using a correlation plot, making it easier to interpret and identify the most influential features in determining laptop prices.
5. **Chapter 5: Simple Linear Regression**
    - **5.1 Model Development:** Utilizing the preprocessed data, we developed simple linear regression models. These models predict laptop prices based on individual features such as resolution, CPU specifications, and more.
    - **5.2 Model Evaluation:** We evaluated the performance of the models using metrics like R-squared, mean absolute error, and mean squared error. Additionally, we compared multiple models using ANOVA to determine the most effective one for price prediction.
    - **5.3 Visualization:** Model evaluation metrics and residuals were visualized to gain insights into the accuracy and performance of the predictive models.
6. **Chapter 6: Making Predictions**

- **6.1 Predictions:** Using the selected optimal model, we made predictions on unseen data. This step demonstrates how the model can be utilized to estimate laptop prices based on their features.

7. Chapter 7: Conclusion

- **7.1 Summary:** This project provides valuable insights into the factors affecting laptop prices and presents a predictive model for estimating prices based on specifications. Consumers can utilize this model to evaluate the fairness of laptop prices in the market.
- **7.2 Contributions:** The predictive model developed in this project contributes to better decision-making for both consumers and manufacturers in the laptop market. Consumers can use the model to make informed purchasing decisions, while manufacturers can optimize pricing strategies for competitive advantage.
- **7.3 Future Work:** Future work could involve enhancing the model by incorporating more advanced machine learning techniques, considering a broader dataset for increased accuracy, and potentially extending the analysis to predict prices of other electronic devices.

8. Appendices

- **8.1 R Script:**

```
library(ggcorrplot)
library(car)

price<-read.csv("C:/Users/DELL/Desktop/price.csv")


# Look at the first 6 observations
head(price)
# Check the dimension
dim(price)
model1 <- lm(price$Price
~price$Sale+price$weight+price$resoloution+price$ppi+price$cpu.core
        +price$cpu.freq+price$internal.mem+price$ram+price$battery, data = price)

# Get the model residuals
model_residuals = model1$residuals

# Plot the result
hist(model_residuals)

# Plot the residuals
qqnorm(model_residuals)
# Plot the Q-Q line
qqline(model_residuals)


# Remove the sale column
reduced_data <- subset(price, select = -Sale)
```

```
# Compute correlation
corr_matrix = cor(reduced_data)

# Compute and show the  result
ggcorrplot(corr_matrix, hc.order = TRUE, type = "lower",lab = TRUE)

model2 <- lm(price$Price ~ price$Sale+price$resoloution+price$ppi+price$cpu.core
        +price$cpu.freq+price$internal.mem+price$battery, data = price[-c(33,48,75,77),])


price.col <- subset(price, select = -c(price$resoloution,price$weight))


pric <- price[, -which(names(price) %in% c("weight", "resoloution"))]
head(pric)
pric.rem<-pric[-c(33,48),]

model3<-lm(pric.rem$Price~.,data = pric.rem)

summary(model3)
avPlots(model3)

vif(model3)
influenceIndexPlot(model3,grid = T,id=list(n=10,cex=1.5,col="blue"))
influence.measures(model3)
qqPlot(model3)



# Get the model residuals
model_residuals = model2$residuals

# Plot the result
hist(model_residuals)

# Plot the residuals
qqnorm(model_residuals)
# Plot the Q-Q line
qqline(model_residuals)

# Anova test
anova(model1, model2)
```

```
# Print the result of the model
summary(model1)
summary(model2)
```

- **8.2 R Output:**

```
  Price Sale weight resoloution ppi cpu.core cpu.freq internal.mem   ram
1  2357   10  135.0        5.2 424        8     1.35           16 3.000
2  1749   10  125.0        4.0 233        2     1.30            4 1.000
3  1916   10  110.0        4.7 312        4     1.20            8 1.500
4  1315   11  118.5        4.0 233        2     1.30            4 0.512
5  1749   11  125.0        4.0 233        2     1.30            4 1.000
6  2137   12  150.0        5.5 401        4     2.30           16 2.000
  RearCam Front_Cam battery thickness
1   13.00         8    2610       7.4
2    3.15         0    1700       9.9
3   13.00         5    2000       7.6
4    3.15         0    1400      11.0
5    3.15         0    1700       9.9
6   16.00         8    2500       9.5

[1] 168  13

  Price Sale ppi cpu.core cpu.freq internal.mem   ram RearCam Front_Cam
1  2357   10 424        8     1.35           16 3.000   13.00         8
2  1749   10 233        2     1.30            4 1.000    3.15         0
3  1916   10 312        4     1.20            8 1.500   13.00         5
4  1315   11 233        2     1.30            4 0.512    3.15         0
5  1749   11 233        2     1.30            4 1.000    3.15         0
6  2137   12 401        4     2.30           16 2.000   16.00         8
  battery thickness
1    2610       7.4
2    1700       9.9
3    2000       7.6
4    1400      11.0
5    1700       9.9
6    2500       9.5

Call:
lm(formula = pric.rem$Price ~ ., data = pric.rem)

Residuals:
    Min      1Q  Median      3Q     Max 
-343.21 -106.23   -9.44  122.48  478.88 

Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1363.21346  133.13545  10.239  < 2e-16 ***
Sale         -0.02824    0.01152  -2.451   0.0154 *
ppi           1.17824    0.21605   5.454 1.91e-07 ***
cpu.core      54.09237    9.87183   5.479 1.69e-07 ***
cpu.freq      91.00261   40.14540   2.267   0.0248 *
internal.mem    7.12038    1.20818   5.893 2.28e-08 ***
ram          108.07416   25.65730   4.212 4.28e-05 ***
RearCam        4.16192    4.25660   0.978   0.3297
Front_Cam      10.39603    5.24779   1.981   0.0494 *
battery        0.02888    0.01405   2.055   0.0416 *
thickness     -57.60783    9.87456  -5.834 3.06e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 176 on 155 degrees of freedom
Multiple R-squared:  0.9497,	Adjusted R-squared:  0.9464
F-statistic: 292.4 on 10 and 155 DF,  p-value: < 2.2e-16

      Sale        ppi     cpu.core     cpu.freq internal.mem
  1.648396     4.548382     3.070938     3.078384     6.307796
       ram      RearCam    Front_Cam      battery    thickness
  8.991669     3.598920     2.711502     1.935813     2.466872

Influence measures of
	 lm(formula = pric.rem$Price ~ ., data = pric.rem) :

     dfb.1_  dfb.Sale   dfb.ppi  dfb.cp.c dfb.cp.f  dfb.int.   dfb.ram
1  -0.041314 -0.000301 -1.32e-01 -0.092495  0.18977  0.200552 -0.179699
2   0.094886  0.029854 -5.51e-05 -0.065444  0.03963 -0.056972  0.057227
3  -0.062926  0.015555 -3.59e-03  0.052375  0.03905  0.016535  0.000351
4   0.000329  0.000576  1.35e-03  0.001461 -0.02161 -0.008156  0.009468
5   0.094906  0.029956 -4.30e-05 -0.065442  0.03963 -0.057014  0.057249
6   0.068478  0.069310  2.80e-02  0.070507 -0.12605 -0.023799  0.071482
7   0.021999 -0.000795  2.40e-03 -0.012892 -0.04801  0.005138 -0.023644
8   0.068458  0.069242  2.80e-02  0.070486 -0.12602 -0.023775  0.071457
9   0.000323  0.000515  1.34e-03  0.001453 -0.02158 -0.008122  0.009443
10 -0.022776 -0.029757 -3.10e-02 -0.096861  0.18105  0.033208 -0.087320
11 -0.001729  0.069306  2.41e-01 -0.129774  0.06061  0.137956 -0.301601
12 -0.038478  0.037097 -7.45e-02  0.198776  0.05487  0.041112  0.006820
13  0.091318 -0.090910 -3.71e-02 -0.086470 -0.00654  0.115786 -0.149271
14 -0.010814 -0.008449  7.11e-03 -0.016908  0.02816  0.002413 -0.009284
15 -0.022771 -0.029669 -3.10e-02 -0.096864  0.18107  0.033179 -0.087315
16 -0.013661 -0.004712 -6.25e-03 -0.017858  0.02990  0.032783 -0.043674
```

17	-0.030439	-0.006343	4.24e-03	-0.000881	0.04314	0.004764	-0.012746
18	0.013166	-0.025327	1.58e-01	-0.141430	-0.07270	-0.165356	0.110668
19	-0.013669	-0.004839	-6.26e-03	-0.017865	0.02989	0.032818	-0.043681
20	0.021931	-0.001224	2.34e-03	-0.012917	-0.04794	0.005285	-0.023686
21	-0.001782	0.068690	2.40e-01	-0.129761	0.06059	0.138090	-0.301538
22	0.048710	0.224261	1.53e-01	-0.038648	-0.05595	-0.041710	0.024928
23	0.051029	-0.019590	-8.17e-02	-0.069620	-0.00831	0.065910	-0.068215
24	-0.010863	-0.008395	7.16e-03	-0.016991	0.02831	0.002391	-0.009318
25	0.182380	0.089205	-5.42e-02	0.062322	-0.17115	-0.032518	0.112637
26	-0.119141	0.043945	2.97e-02	0.038591	-0.04693	-0.015369	0.049015
27	-0.038507	0.037587	-7.46e-02	0.199172	0.05496	0.041034	0.006902
28	0.013215	-0.024979	1.59e-01	-0.141547	-0.07279	-0.165669	0.110849
29	-0.119143	0.043809	2.97e-02	0.038576	-0.04692	-0.015320	0.048989
30	0.138812	0.001994	-6.57e-02	-0.039731	-0.00411	0.015926	-0.004387
31	-0.030562	-0.006181	4.28e-03	-0.000867	0.04334	0.004718	-0.012774
32	0.050958	-0.019731	-8.16e-02	-0.069559	-0.00829	0.065896	-0.068166
34	-0.017784	0.013195	3.03e-03	-0.019545	0.01296	-0.034935	0.032555
35	0.011891	0.002357	5.55e-03	-0.004334	-0.00202	-0.000719	-0.000559
36	-0.017766	0.013142	3.03e-03	-0.019525	0.01294	-0.034880	0.032510
37	0.182082	0.087960	-5.43e-02	0.062148	-0.17095	-0.032074	0.112329
38	0.001791	-0.006999	6.54e-03	0.052610	-0.03579	0.021883	0.016092
39	0.053464	0.011459	-9.89e-02	0.054773	0.07386	0.072645	-0.054279
40	0.048478	0.222282	1.53e-01	-0.038732	-0.05583	-0.041071	0.024623
41	0.001785	-0.007045	6.53e-03	0.052556	-0.03576	0.021881	0.016068
42	-0.167684	0.035576	4.10e-02	0.079501	0.07643	0.070216	-0.146469
43	0.002969	0.011196	-9.23e-02	0.109415	0.08090	0.047883	0.015474
44	-0.167643	0.035300	4.10e-02	0.079445	0.07641	0.070283	-0.146455
45	-0.025470	0.058289	-7.13e-02	0.058681	0.03145	-0.037076	0.010990
46	-0.021613	0.017570	1.64e-02	0.026668	0.01370	0.006127	-0.007043
47	0.012032	0.002463	5.63e-03	-0.004375	-0.00205	-0.000755	-0.000552
49	0.002996	0.011475	-9.23e-02	0.109481	0.08092	0.047802	0.015525
50	0.010532	0.029899	5.93e-02	-0.040826	-0.00122	-0.034366	0.002399
51	-0.143640	-0.008403	-2.98e-02	0.132171	0.06187	0.040948	-0.011943
52	-0.048279	-0.008206	9.55e-03	0.040807	-0.01866	-0.001658	0.016851
53	-0.025467	0.058095	-7.13e-02	0.058627	0.03143	-0.036996	0.010957
54	-0.095823	-0.063595	-2.15e-01	0.203763	0.16372	-0.092332	0.365574
55	0.055262	-0.010131	5.36e-02	-0.057524	-0.05853	-0.014473	-0.033600
56	0.053599	0.012267	-9.89e-02	0.054910	0.07393	0.072443	-0.054209
57	-0.095835	-0.062944	-2.15e-01	0.203995	0.16384	-0.092658	0.365986
58	0.055501	-0.009729	5.39e-02	-0.057689	-0.05875	-0.014681	-0.033648
59	-0.021478	0.017038	1.62e-02	0.026419	0.01360	0.006219	-0.007052
60	-0.143790	-0.007235	-2.97e-02	0.132523	0.06196	0.040601	-0.011770
61	-0.082568	-0.162524	-8.82e-02	-0.048379	-0.06019	0.052039	-0.037415
62	-0.004647	0.053505	2.63e-02	0.031397	-0.00307	-0.043262	0.000791

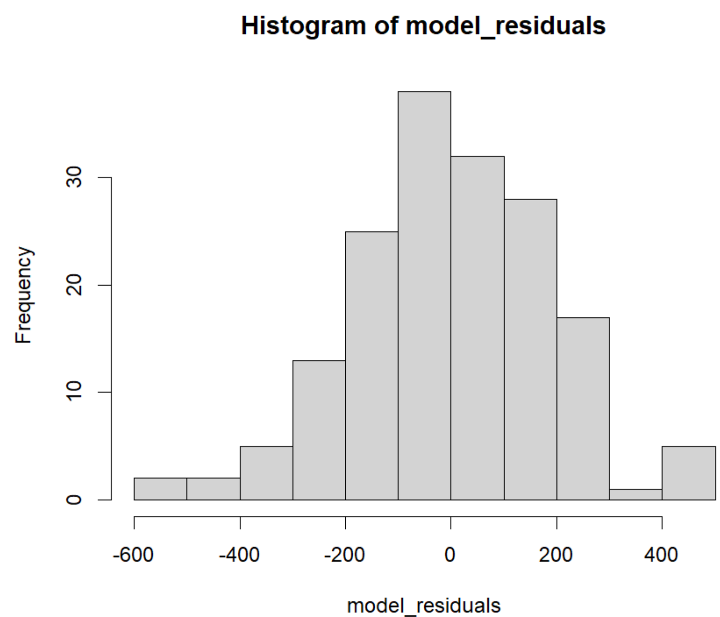
```
63 -0.048000 -0.008978  9.38e-03  0.040427 -0.01851 -0.001349  0.016589
64 -0.034980 -0.003898 -6.95e-04 -0.039715 -0.01961 -0.037835  0.044588
   dfb.RrCm  dfb.Fr_C  dfb.bttr  dfb.thck   dffit cov.r   cook.d    hat inf
1   0.11157 -2.25e-02  4.57e-02  0.048107 -0.3710 0.924 1.24e-02 0.0469
2  -0.08673 -2.95e-02 -5.86e-02 -0.059684  0.2299 0.941 4.77e-03 0.0235
3  -0.03813 -1.48e-02 -5.24e-03  0.055185 -0.0923 1.093 7.78e-04 0.0341
4   0.01310  3.80e-05  1.58e-02 -0.008229 -0.0477 1.094 2.08e-04 0.0244
5  -0.08677 -2.95e-02 -5.87e-02 -0.059706  0.2299 0.941 4.77e-03 0.0235
6  -0.06036 -1.21e-01 -4.28e-03 -0.066939 -0.2083 1.109 3.95e-03 0.0702
7   0.06573  1.13e-02  5.24e-02 -0.052523 -0.1540 1.038 2.16e-03 0.0268
8  -0.06034 -1.21e-01 -4.28e-03 -0.066916 -0.2082 1.109 3.95e-03 0.0702
9   0.01310  5.47e-05  1.57e-02 -0.008207 -0.0476 1.094 2.07e-04 0.0244
10  0.11658 -2.18e-02  1.02e-03 -0.000610  0.3120 0.991 8.80e-03 0.0500
11 -0.05219  4.95e-02  4.89e-02 -0.056020 -0.4528 0.933 1.84e-02 0.0653
12 -0.04371 -1.16e-01 -8.74e-02  0.054388  0.2306 1.120 4.85e-03 0.0817
13  0.14543  6.33e-02  4.09e-03 -0.067883  0.2528 0.998 5.79e-03 0.0388
14  0.00518  4.25e-03  1.24e-05  0.005047  0.0492 1.129 2.22e-04 0.0520
15  0.11658 -2.19e-02  1.02e-03 -0.000625  0.3120 0.991 8.80e-03 0.0500
16  0.00279  5.14e-02  2.86e-02 -0.007461 -0.1143 1.017 1.19e-03 0.0136
17  0.00482 -1.71e-03 -1.11e-02  0.027412  0.0655 1.112 3.92e-04 0.0411
18 -0.09764  1.03e-01  7.39e-02 -0.037636  0.2716 1.100 6.72e-03 0.0808
19  0.00282  5.14e-02  2.86e-02 -0.007438 -0.1143 1.017 1.19e-03 0.0136
20  0.06575  1.14e-02  5.23e-02 -0.052389 -0.1538 1.038 2.15e-03 0.0268
21 -0.05203  4.96e-02  4.88e-02 -0.055898 -0.4525 0.933 1.84e-02 0.0653
22 -0.00479 -4.00e-01  5.76e-02 -0.098354 -0.5353 0.977 2.58e-02 0.0946
23  0.10464  2.95e-02  4.01e-02 -0.028881 -0.1738 1.136 2.76e-03 0.0782
24  0.00518  4.25e-03  1.33e-05  0.005059  0.0495 1.129 2.24e-04 0.0519
25 -0.11184 -7.20e-02  1.41e-02 -0.168607 -0.3767 0.913 1.27e-02 0.0458
26  0.09596 -1.45e-01  6.44e-02  0.075010 -0.2943 0.830 7.73e-03 0.0214
27 -0.04388 -1.16e-01 -8.76e-02  0.054417  0.2311 1.120 4.87e-03 0.0817
28 -0.09783  1.03e-01  7.39e-02 -0.037737  0.2719 1.100 6.73e-03 0.0808
29  0.09598 -1.45e-01  6.44e-02  0.075025 -0.2943 0.830 7.72e-03 0.0214
30  0.03585 -3.25e-02 -3.46e-02 -0.119034  0.1652 1.088 2.49e-03 0.0495
31  0.00479 -1.77e-03 -1.11e-02  0.027508  0.0657 1.112 3.95e-04 0.0410
32  0.10456  2.95e-02  4.00e-02 -0.028823 -0.1737 1.136 2.75e-03 0.0782
34 -0.02536  1.65e-02  2.69e-03  0.019566 -0.0645 1.126 3.80e-04 0.0512
35 -0.00962 -1.91e-03 -6.15e-03 -0.008758  0.0222 1.103 4.49e-05 0.0280
36 -0.02532  1.65e-02  2.68e-03  0.019549 -0.0644 1.126 3.79e-04 0.0512
37 -0.11146 -7.16e-02  1.40e-02 -0.168244 -0.3760 0.913 1.27e-02 0.0458
38  0.01481 -2.11e-02 -1.15e-01  0.016392 -0.1505 1.241 2.07e-03 0.1430   *
39  0.01360 -7.40e-02 -4.79e-02 -0.019876  0.2134 0.971 4.12e-03 0.0251
40 -0.00441 -3.99e-01  5.75e-02 -0.097928 -0.5337 0.977 2.56e-02 0.0944
41  0.01481 -2.11e-02 -1.15e-01  0.016385 -0.1504 1.241 2.07e-03 0.1430   *
42  0.03754  6.85e-03  5.93e-02  0.128279 -0.2466 0.984 5.50e-03 0.0341
```

```
43 -0.02367 -1.01e-01 -1.32e-01  0.044894  0.2265 0.963 4.64e-03 0.0262
44  0.03759  6.92e-03  5.93e-02  0.128271 -0.2464 0.984 5.49e-03 0.0341
45  0.02329 -4.18e-03 -2.21e-02  0.034400 -0.1592 1.073 2.31e-03 0.0411
46 -0.03342 -7.46e-03 -3.69e-03  0.012404 -0.0535 1.089 2.61e-04 0.0226
47 -0.00974 -1.95e-03 -6.22e-03 -0.008869  0.0224 1.103 4.60e-05 0.0281
49 -0.02374 -1.01e-01 -1.33e-01  0.044866  0.2266 0.963 4.64e-03 0.0262
50 -0.08679  4.64e-02  2.47e-02 -0.019927 -0.1148 1.130 1.20e-03 0.0626
51 -0.02208 -6.26e-02 -2.91e-02  0.178540  0.2393 0.992 5.18e-03 0.0345
52  0.01934 -2.31e-03 -2.30e-02  0.041235 -0.0966 1.088 8.53e-04 0.0327
53  0.02331 -4.13e-03 -2.21e-02  0.034402 -0.1590 1.074 2.31e-03 0.0410
54 -0.14340 -9.22e-02 -2.72e-01  0.137935  0.5515 1.067 2.75e-02 0.1294
55 -0.00658  4.38e-02  6.31e-02 -0.059388  0.1138 1.097 1.18e-03 0.0419
56  0.01343 -7.43e-02 -4.80e-02 -0.020026  0.2137 0.971 4.13e-03 0.0251
57 -0.14368 -9.25e-02 -2.72e-01  0.137934  0.5519 1.066 2.75e-02 0.1293
58 -0.00671  4.38e-02  6.34e-02 -0.059671  0.1141 1.097 1.19e-03 0.0419
59 -0.03306 -7.29e-03 -3.67e-03  0.012369 -0.0529 1.089 2.56e-04 0.0225
60 -0.02239 -6.30e-02 -2.91e-02  0.178674  0.2397 0.992 5.20e-03 0.0345
61  0.28643  5.34e-02  7.96e-02  0.109517  0.3586 0.995 1.16e-02 0.0614
62 -0.03512  5.70e-03 -1.50e-02 -0.003620 -0.1029 1.094 9.67e-04 0.0372
63  0.01940 -2.06e-03 -2.28e-02  0.041064 -0.0960 1.089 8.41e-04 0.0328
64  0.02721  2.94e-03  3.22e-02  0.039784 -0.1004 1.149 9.22e-04 0.0741
 [ reached 'max' / getOption("max.print") -- omitted 104 rows ]

154 166
152 164

Analysis of Variance Table

Model 1: price$Price ~ price$Sale + price$weight + price$resoloution +
    price$ppi + price$cpu.core + price$cpu.freq + price$internal.mem +
    price$ram + price$battery
Model 2: price$Price ~ price$Sale + price$resoloution + price$ppi + price$cpu.core +
    price$cpu.freq + price$internal.mem + price$battery
  Res.Df     RSS Df Sum of Sq     F    Pr(>F)
1    158 5864419
2    160 7377262 -2  -1512843 20.38 1.337e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Call:
lm(formula = price$Price ~ price$Sale + price$weight + price$resoloution +
    price$ppi + price$cpu.core + price$cpu.freq + price$internal.mem +
    price$ram + price$battery, data = price)
```

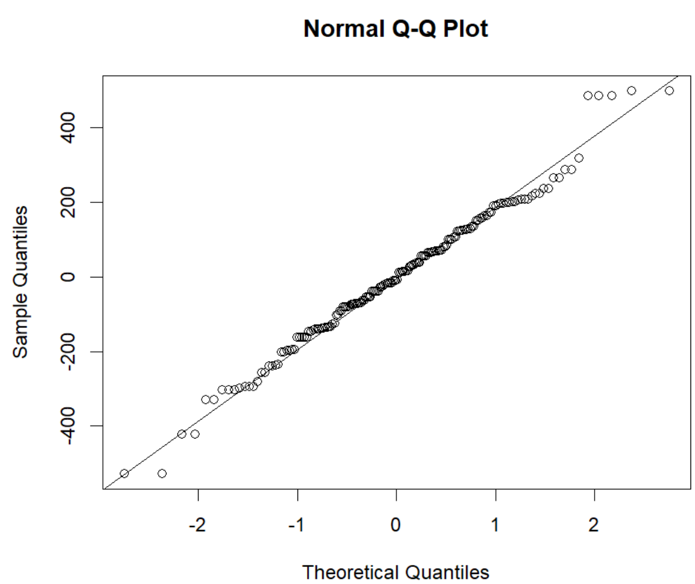
```
Residuals:
    Min      1Q  Median      3Q     Max 
-526.82 -132.17   -7.51  125.74  499.34 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      529.76591   80.66326   6.568 7.01e-10 ***
price$Sale        -0.01787    0.01166  -1.533 0.127347    
price$weight      -2.77889    0.63076  -4.406 1.94e-05 ***
price$resoloution  92.91155   35.23095   2.637 0.009194 ** 
price$ppi          1.04960    0.21197   4.952 1.87e-06 ***
price$cpu.core     79.97201    9.49071   8.426 2.07e-14 ***
price$cpu.freq    106.53689   48.57998   2.193 0.029768 *  
price$internal.mem   6.68644    1.30009   5.143 7.91e-07 ***
price$ram        115.99452   26.76665   4.334 2.60e-05 ***
price$battery      0.12969    0.03383   3.833 0.000182 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 192.7 on 158 degrees of freedom
Multiple R-squared:   0.94,	Adjusted R-squared:  0.9366 
F-statistic: 275.2 on 9 and 158 DF,  p-value: < 2.2e-16

Call:
lm(formula = price$Price ~ price$Sale + price$resoloution + price$ppi + 
    price$cpu.core + price$cpu.freq + price$internal.mem + price$battery, 
    data = price[-c(33, 48, 75, 77), ])

Residuals:
    Min      1Q  Median      3Q     Max 
-501.17 -142.61  -16.17  147.60  680.20 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      645.22809   80.82782   7.983 2.62e-13 ***
price$Sale        -0.02788    0.01280  -2.178   0.0308 *  
price$resoloution -38.36517   25.61486  -1.498   0.1362    
price$ppi          1.28479    0.22922   5.605 8.92e-08 ***
price$cpu.core    111.28701    8.82247  12.614  < 2e-16 ***
price$cpu.freq    238.69824   48.65666   4.906 2.27e-06 ***
price$internal.mem  12.08441    0.97653  12.375  < 2e-16 ***
price$battery      0.05722    0.02815   2.033   0.0437 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 214.7 on 160 degrees of freedom
Multiple R-squared: 0.9246,	Adjusted R-squared: 0.9213
F-statistic: 280.1 on 7 and 160 DF, p-value: < 2.2e-16

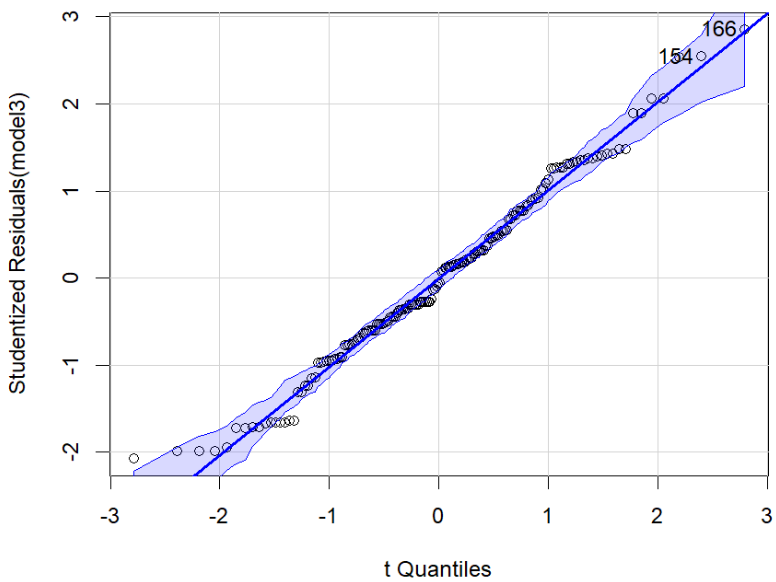
- **8.3 R Plots:**



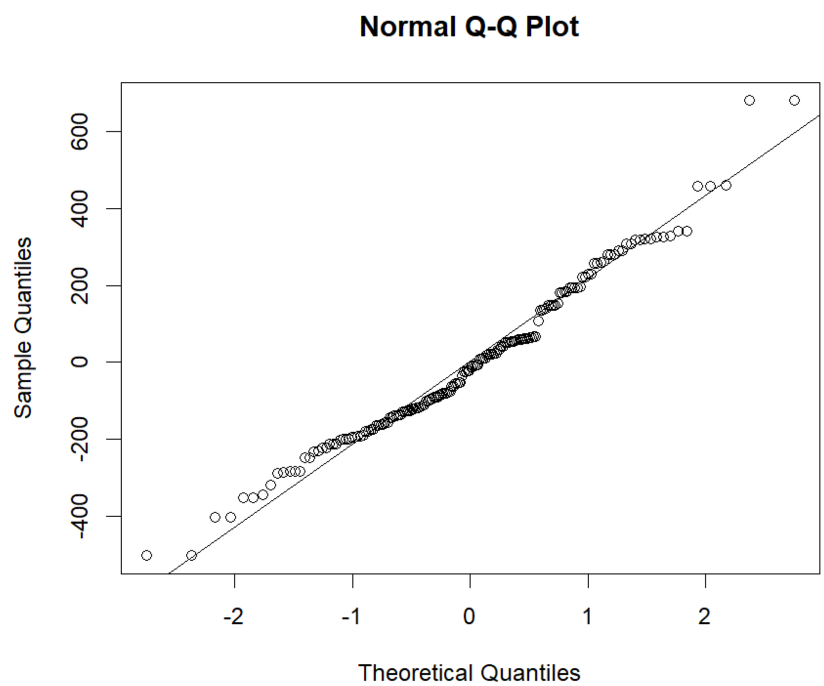
Central Tendency: The highest bar is at 0 on the x-axis, which suggests that the model's predictions are, on average, fairly accurate as most residuals are close to 0.
Skewness: The histogram appears to be slightly right-skewed. This means that there are some instances where the model's predictions are significantly lower than the actual values.
Outliers: The bar at -600 on the x-axis suggests that there are a few instances where the model's predictions are much higher than the actual values.



Normality: If the data were perfectly normal, the points would fall along the diagonal line. In this case, it seems like most of the points do fall along the line, suggesting that your data is approximately normally distributed.
Outliers: Points that deviate from the line can indicate outliers in your data. In this plot, there don’t appear to be any significant outliers.
Skewness and Kurtosis: The shape of the deviations from the line can provide information about the skewness and kurtosis of your data. In this case, there doesn’t appear to be any significant skewness or kurtosis.



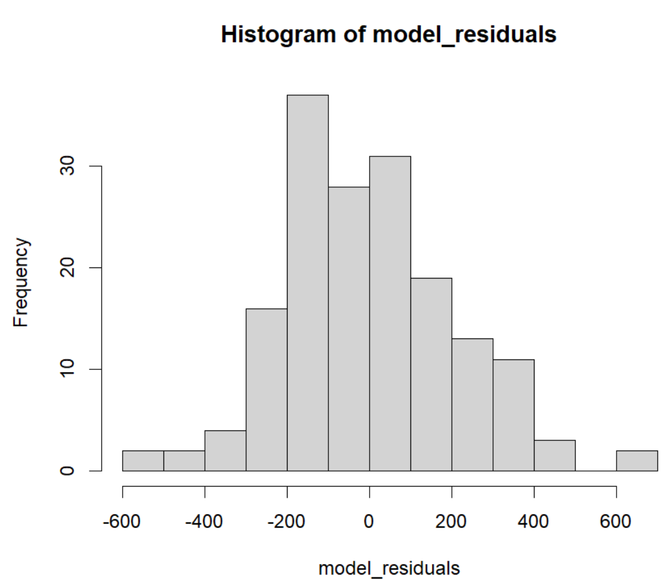
Linearity: The line of best fit has a positive slope, which suggests a positive linear relationship between the t quantiles and the studentized residuals. This means as the t quantiles increase, the studentized residuals also tend to increase.
Outliers: The points labeled “166” and “154” are above the line of best fit, which suggests that these observations have larger residuals than what would be expected given their t quantiles. These could potentially be outliers or influential points in your data.
Homoscedasticity: If the residuals are evenly distributed around the line of best fit, this would suggest homoscedasticity (equal variance) of the residuals. If they’re not, this could indicate heteroscedasticity (unequal variance), which is a violation of one of the assumptions of many statistical models.



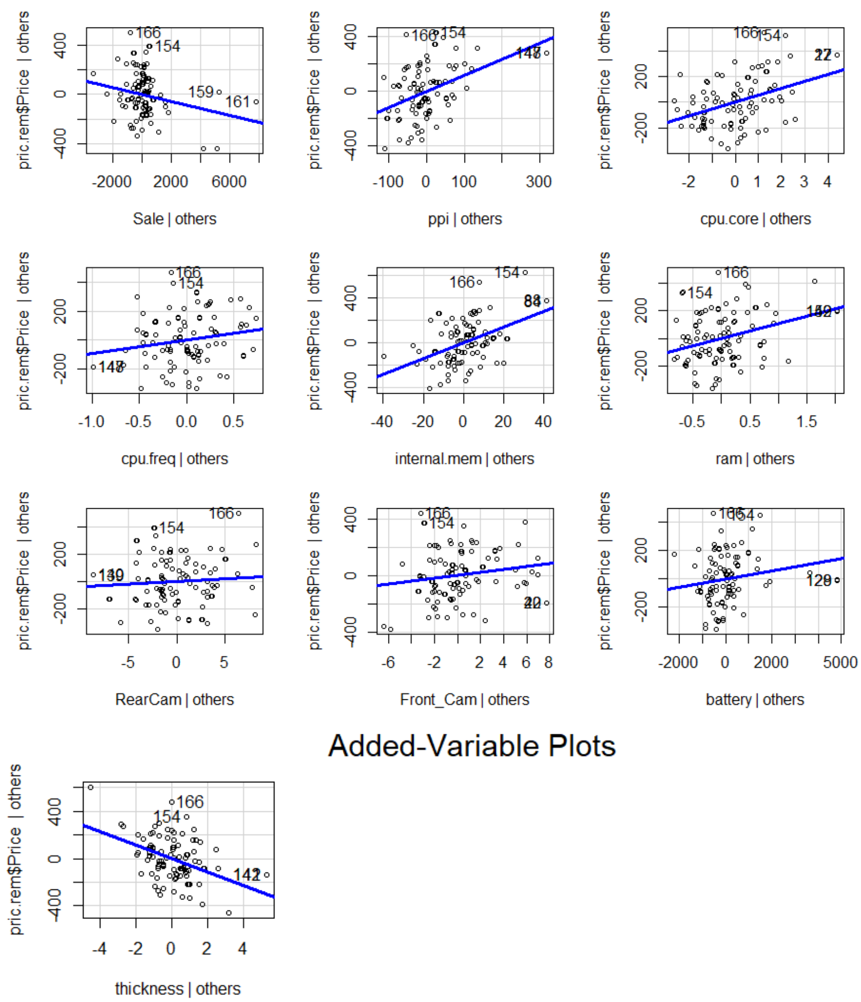
Normality: If the data were perfectly normal, the points would fall along the diagonal line. In this case, it seems like most of the points do fall along the line, suggesting that your data is approximately normally distributed.

Outliers: Points that deviate from the line can indicate outliers in your data. In this plot, there don't appear to be any significant outliers.

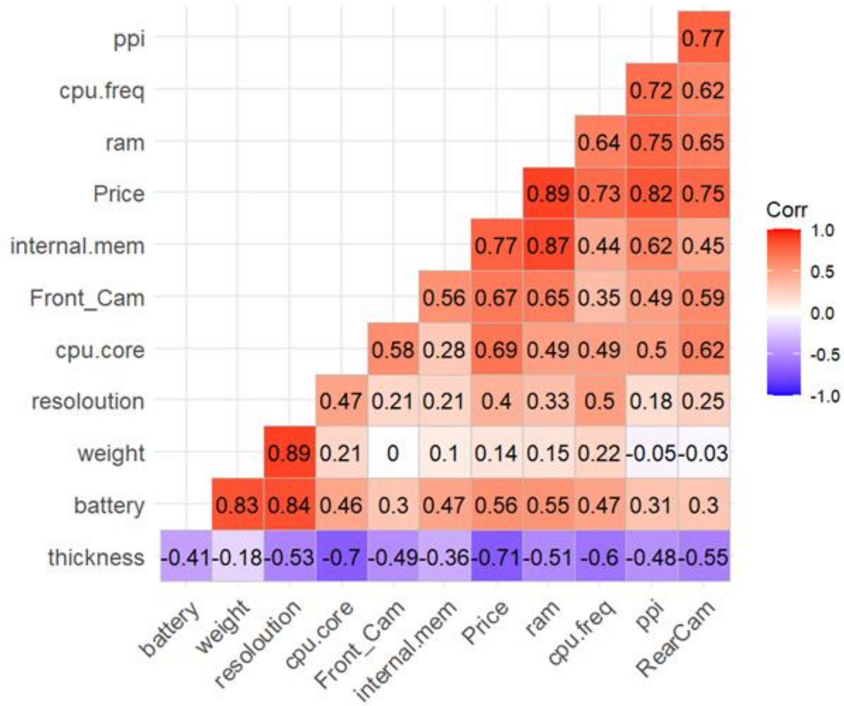
Skewness and Kurtosis: The shape of the deviations from the line can provide information about the skewness and kurtosis of your data. In this case, there doesn't appear to be any significant skewness or kurtosis.



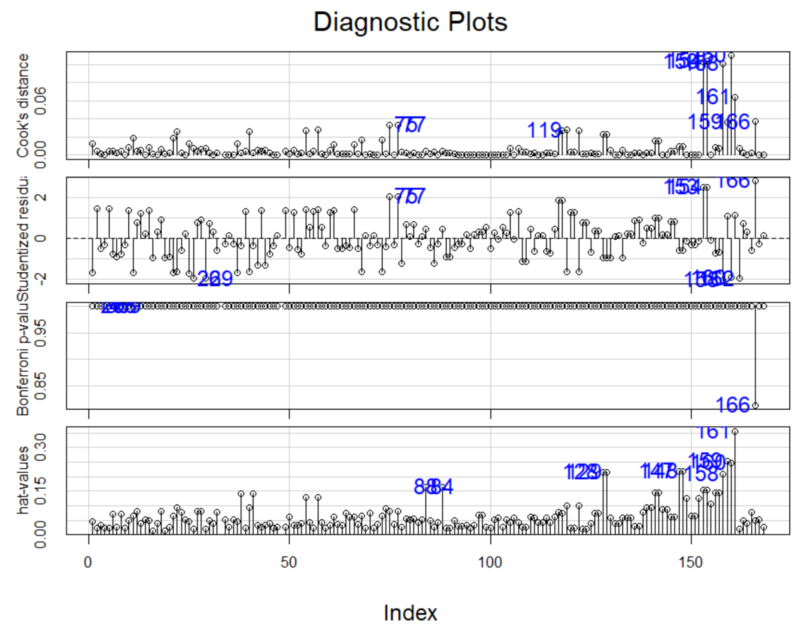
Central Tendency: The histogram is approximately bell-shaped, which suggests that the model's predictions are, on average, fairly accurate as most residuals are close to 0.
Symmetry: The histogram appears to be symmetric around 0, which indicates that the model's predictions are equally likely to be above or below the actual values.
Outliers: The bars at -600 and 600 on the x-axis suggest that there are a few instances where the model's predictions are much higher or lower than the actual values.



Correlation: The trend lines in each scatter plot show the correlation between the variables. A positive slope indicates a positive correlation (as one variable increases, so does the other), while a negative slope indicates a negative correlation (as one variable increases, the other decreases).
Strength of Relationship: The closer the data points are to the trend line, the stronger the relationship between the variables. If the data points are widely spread from the trend line, it indicates a weaker relationship.
Outliers: Any data points that deviate significantly from the trend line could be considered outliers. These could represent unusual observations that may need further investigation.



Correlation: The values range from -1 to 1, with -1 being a negative correlation (as one feature increases, the other decreases) and 1 being a positive correlation (as one feature increases, so does the other). The colors range from red to blue, with red being a negative correlation and blue being a positive correlation.
Diagonal Values: The diagonal values are all 1, as they represent the correlation of a feature with itself.
Positive Correlation: The highest positive correlation is between price and ram, with a value of 0.89. This suggests that as the RAM of a mobile phone increases, its price also tends to increase.
Negative Correlation: The lowest negative correlation is between battery and thickness, with a value of -0.81. This suggests that as the thickness of a mobile phone decreases (i.e., the phone becomes slimmer), its battery capacity also tends to decrease.



Residuals: Each blue dot represents an observation’s residual, which is the difference between the observed and predicted values. The residuals are scattered around the horizontal line at zero, which represents perfect prediction.
Outliers: The blue boxes with numbers such as “119”, “884”, and “166” likely represent outliers, which are observations that have large residuals. These could be due to unusual or extreme values in the data.
Model Fit: The black lines connecting some of the dots could represent a measure of model fit. If these lines closely follow the horizontal line at zero, it suggests that the model fits the data well.

9. Reference

- https://www.kaggle.com/datasets/muhammetvarl/laptop-price