Saajan Patel

PDF Bivariate Data Project

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	2375	2375.3	1.37	0.252
Top Speed	1	2375	2375.3	1.37	0.252
Error	29	50365	1736.7		
Lack-of-Fit	17	42782	2516.6	3.98	0.010
Pure Error	12	7583	631.9		
Total	30	52740			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
41.6740	4.50%	1.21%	0.00%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	154.3	22.1	6.99	0.000	
Top Speed	-0.391	0.334	-1.17	0.252	1.00

Regression Equation

Duration = 154.3 - 0.391 Top Speed

Fits and Diagnostics for Unusual Observations

Obs	Duration	Fit	Resid	Std Resid	
12	17.0	107.4	-90.4	-2.50	R X
17	50.6	104.3	-53.7	-1.55	X
23	213.0	123.0	90.0	2.22	R
27	204.0	117.2	86.8	2.20	R

Data:

Top Speed	Duration	Y	Residual
40	120.0	138.660	-18.660
40	170.0	138.660	31.340
48	120.0	135.532	-15.532
67	120.0	128.103	-8.103
60	160.0	130.840	29.160
40	150.0	138.660	11.340
72	165.0	126.148	38.852
70	150.0	126.930	23.070
93	120.0	117.937	2.063
57	135.0	132.013	2.987
60	150.0	130.840	19.160
120	17.0	107.380	-90.380
75	143.0	124.975	18.025
50	100.0	134.750	-34.750
61	195.0	130.449	64.551
70	125.0	126.930	-1.930
128	50.6	104.252	-53.652
60	145.0	130.840	14.160
30	120.0	142.570	-22.570
62	162.0	130.058	31.942
41	60.0	138.269	-78.269
50	120.0	134.750	-14.750
80	213.0	123.020	89.980
50	138.0	134.750	3.250
50	120.0	134.750	-14.750
51	120.0	134.359	-14.359
95	204.0	117.155	86.845
30	132.0	142.570	-10.570
50	110.0	134.750	-24.750
60	70.0	130.840	-60.840
67	126.0	128.103	-2.103

Descriptive Statistics: Top Speed, Duration

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Top Speed	31	0	62.16	4.09	22.77	30.00	50.00	60.00	70.00	128.00
Duration	31	0	130.02	7.53	41.93	17.00	120.00	126.00	150.00	213.00

The purpose of this study was that I was curious to see if there indeed was a relationship between the top speed of a roller coaster and the duration of the ride. I was curious because some roller coasters have a higher top speed, but last longer than the roller coasters who have a lower top speed. The data was collected by going to the websites of different roller coaster parks around the U.S. and seeing the average top speed and average duration of each ride. The average of each was better because if you were to measure it yourself, there could have been a mechanical issue with the ride, therefore leading to false data. So by taking the average, you eliminate the possibility of any wrong data.