Erica Christ

Bivariate Data Project Analysis

Purpose

The purpose of this study is to determine whether there is a linear relationship between the amount of calories and protein in a protein bar.

Description of Sample

I have a sample size of 30 different protein bars from Amazon.com, with the population of interest being protein bars. I chose the top 30 most popular bars on the website that weren't vegan or gluten free. I used the nutritional information section of the website to find each bar's protein and calories, which comes from the bars nutrition label. The information found on nutrition labels all comes from the FDA, so they are all calculated the same way.

Protein	Calories
20	250
10	190
15	150
16	150
12	190
21	250
20	230
20	250
10	230
15	170
8	170
12	170
11	260
15	260
11	250
10	250
9	190
12	210
10	200
12	240
8	80
8	150
10	200
14	250
8	220
18	270
12	250

7	170
7	90
20	360

Descriptive Statistics: Protein, Calories

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Protein	30	0	12.700	0.795	4.356	7.000	9.750	12.000	15.250	21.000
Calories	30	0	210.0	10.4	56.9	80.0	170.0	215.0	250.0	360.0

Regression Analysis: Calories versus Protein

Calories = 117.2 + 7.305 Protein

Model Summary

S	R-sq	R-sq(adj)
48.0452	31.24%	28.79%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	29366.5	29366.5	12.72	0.001
Error	28	64633.5	2308.3		
Total	29	94000.0			

Summarizing the Data

The correlation coefficient is 0.56, meaning there is a moderate linear correlation in this data. This means that the data has a decent positive correlation. The coefficient of determination is 31.24%. The minimum value for protein was 7, the maximum was 21, and the median value was 12. The minimum value for calories was 80, the maximum was 360, and the median value was 215. The median and mean for each data set were pretty similar, which shows that the data was symmetrical.