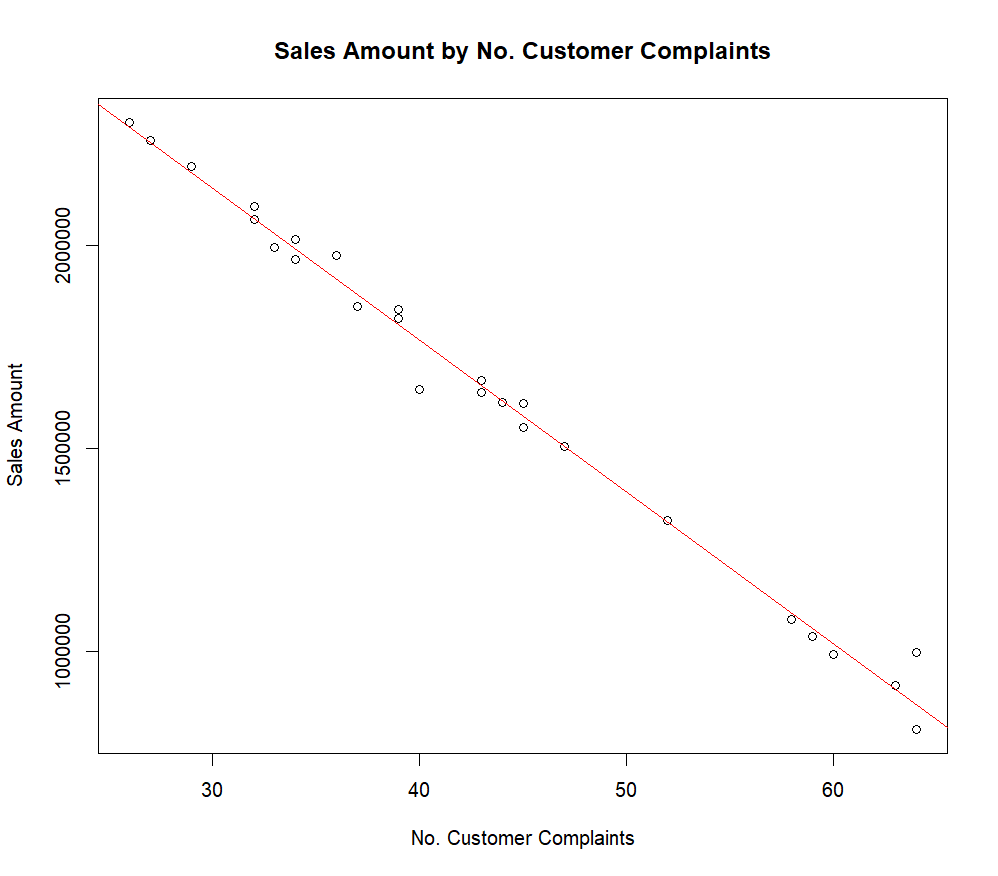


Starting this assignment like an EDA, we can use these descriptive statistics to find significant variables that affect the SalesAmt variable by finding their correlations. From looking at the corrgram graph, we can see that any correlation that has a graph with a line that looks straight is likely to be highly correlated, and the P value from rcorr() also corresponds to that. Some important variables with the best looking graphs that influence SalesAmt would be No.Employees, No.SalesCampaigns, and No.CustComplaints. One variable that had a very poorly related graph was the MaterialCost variable, as also seen through the 0.5046 P-value.

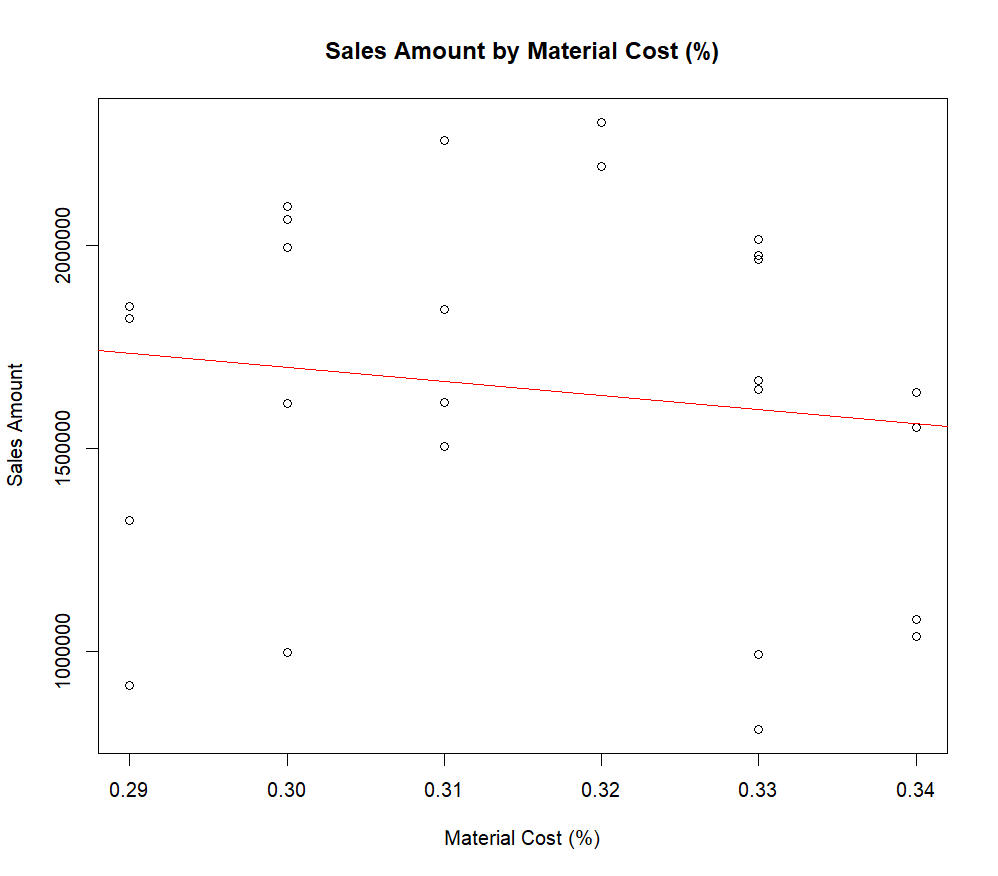


Standard Error: 44660

95% Confident that Slope is between: -38857 to -35489

R-Squared: 0.9904

With an R-squared value of 0.9904, we can tell that changes in the number of customer complaints accounts for 99.04% of the changes in sales for the weekly sales amount. This as well as low lowest standard error and spread in slope seen in the data, makes customer complaints the best determining factor in sales. This means that there is a very high probability of the sales amount decreasing when the number of customer complaints increase. This indicates that keeping customer complaints low is an important variable in maintaining a high sales amount.

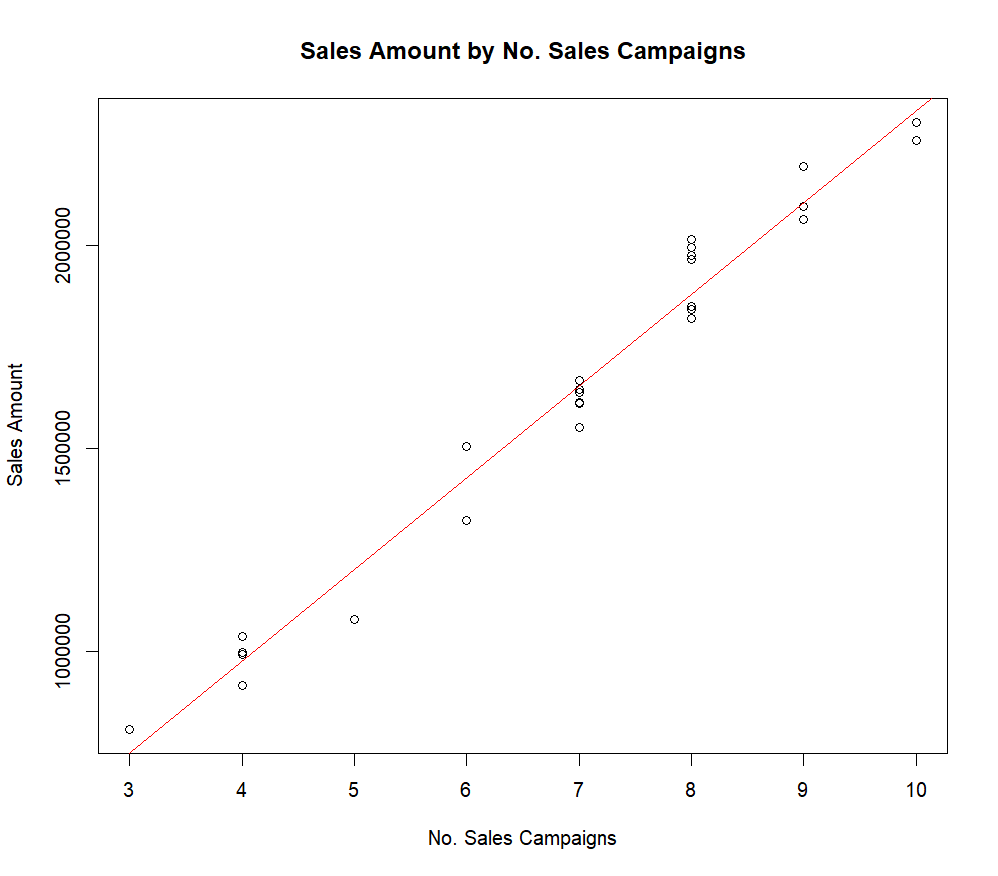


Standard Error: 450700

95% Confident that Slope is between: -13634014 to 6734926

R-Squared: 0.01876

With an R^2 of 0.01876, material cost is by far the worst determining factor of sales. Material cost is also seen to not be correlated to sales through our rcorr seen above.This means that only 1.8% of the data can be explained by our linear model with a high standard error and extremely high spread in our 95% confidence interval of the slope.. This could be due to many reasons but the main one could be that since material cost is a % and not a value amount, there is no correlation. This could also be because material cost across the board is very similar with only a 5% difference between the most material cost and the least.

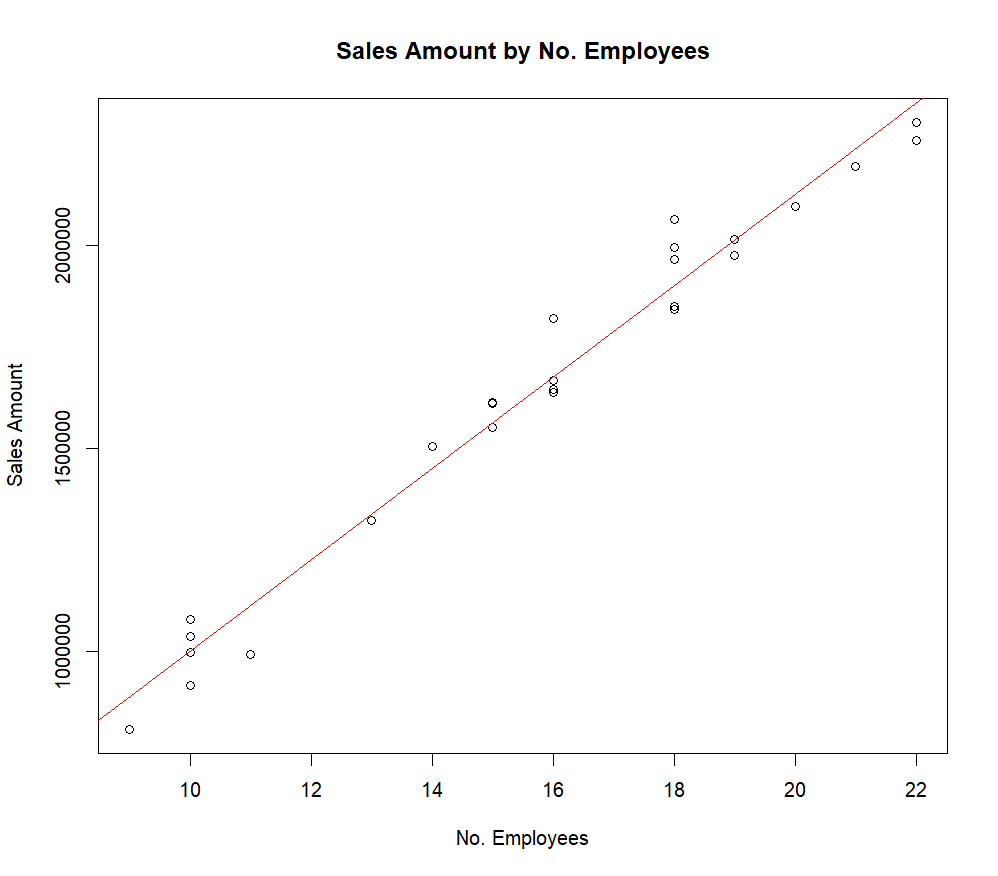


Standard Error: 73060

95% Confident that Slope is between: 210786 to 240778

R-Squared: 0.9742

With an R^2 of 0.9742 and a standard error of 73060, Number of sales campaigns is the 3rd best determining factor in sales. 97.42% of the data can be explained by our linear model which is quite high with a relatively low standard error and low spread in our 95% confidence interval of the slope.. This one is quite easy to explain as well as the more you advertise, the more likely people are to come to your business and buy the thing you advertised as well as more stuff while they are there.

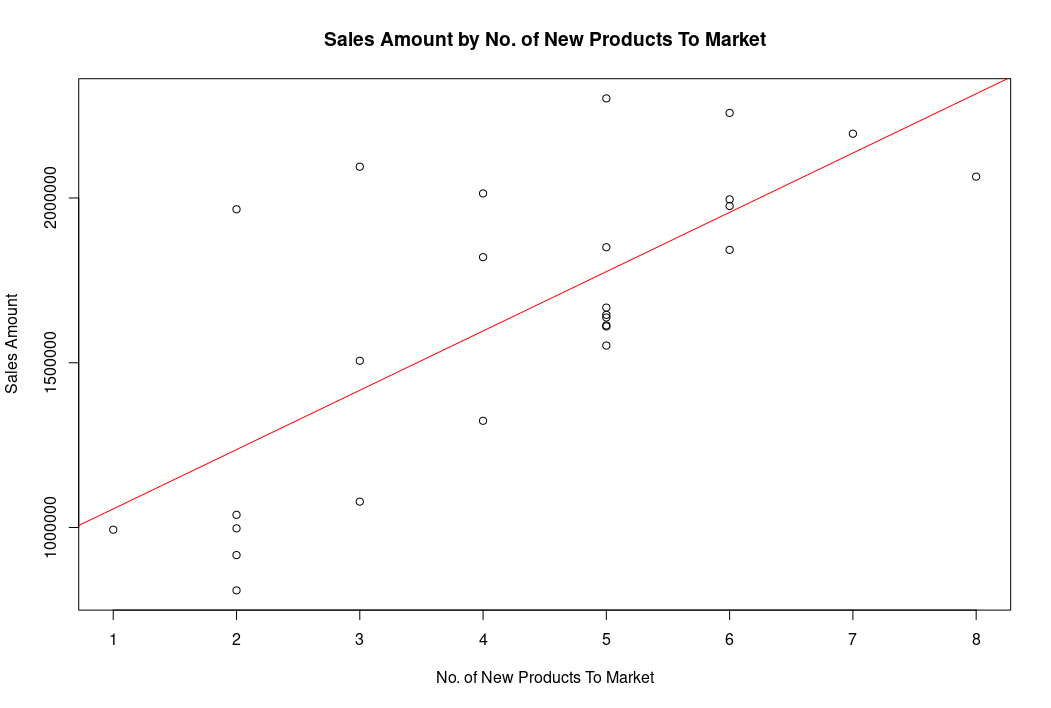


Standard Error: 73040

95% Confident that Slope is between: 104947 to 119871

R-Squared: 0.9742

With an R^2 of 0.9742 and a standard error of 73040, number of employees campaigns is the 2nd best determining factor in sales, only beating sales campaigns in standard error. 97.42% of the data can be explained by our linear model which again is quite high with a relatively low standard error and low spread in our 95% confidence interval of the slope. This one is also quite easy to explain as if you have more employees, you will also have a bigger store and quicker customer service.

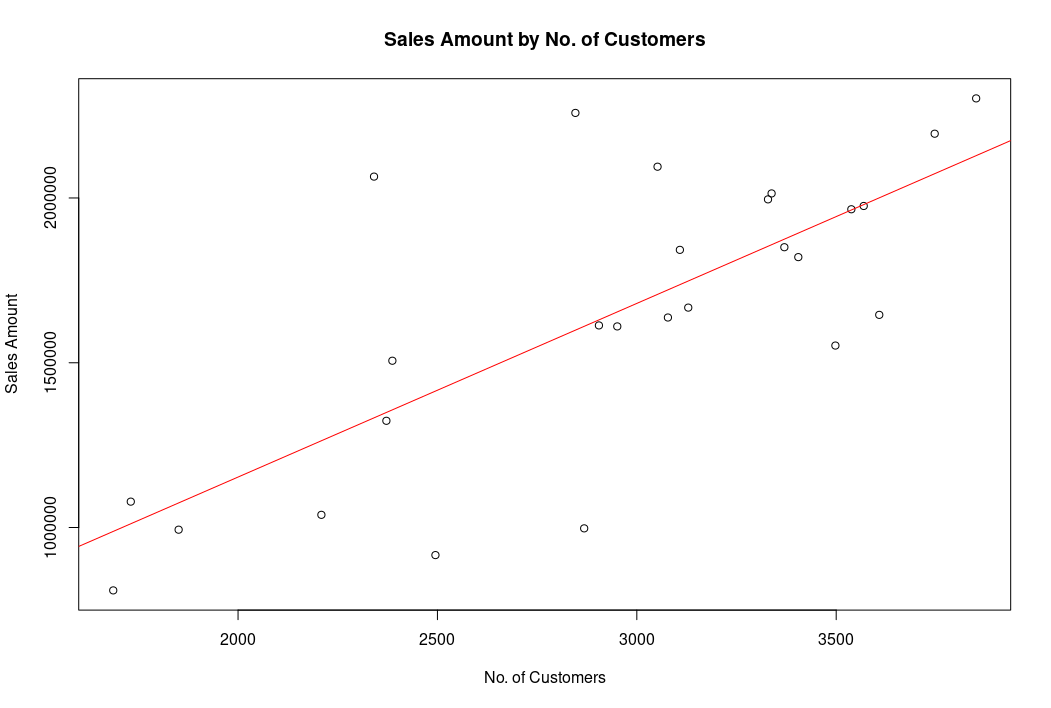


Standard Error: 316600

95% Confident that Slope is between: 108786 to 251178

R-Squared: 0.5158

At an R squared of 0.5158, number of new products on the market is the second worst determining factor of sales amount. Only 51.6% of the change in sales can be explained by the change in number of new products on the market. This paired with its high standard error and relatively high spread in slope makes it a worse determining factor that is actually correlated with sales. There could be a number of reasons for this from the stuff on the market not being worth buying sometimes to items being put out early, before a big season has actually started.

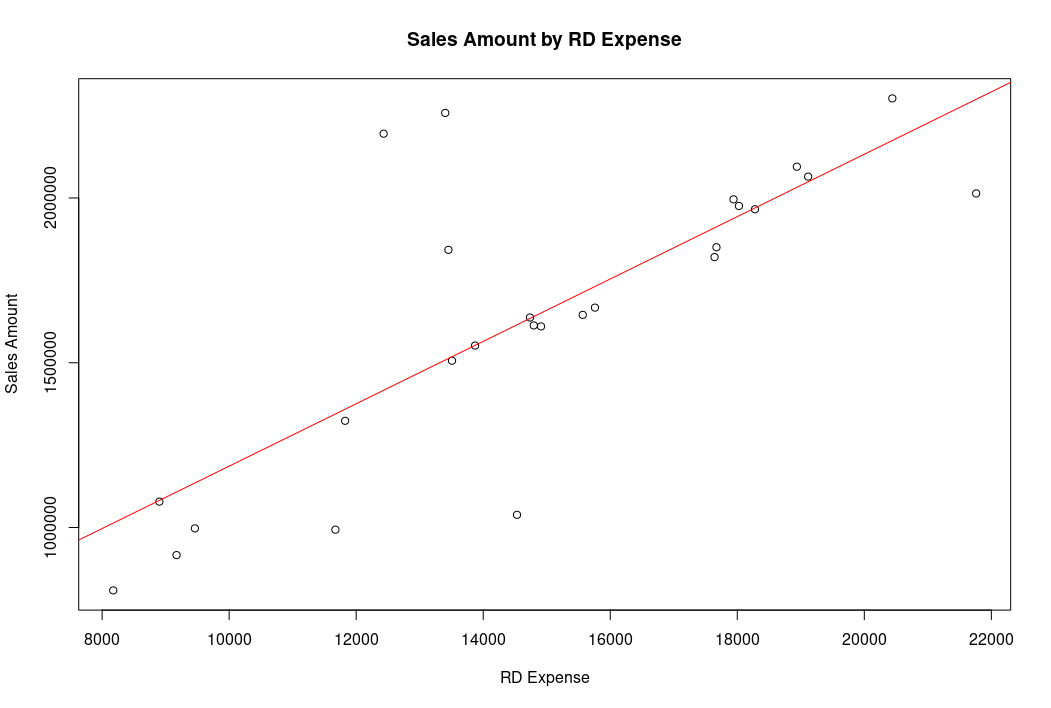


Standard Error: 307000

95% Confident that Slope is between: 330.35 to 723.59

R-Squared: 0.5449

With an R^2 value of 0.5449, number of customers is the 3rd worst determining factor of sales. Only 54.9% of data can be explained by our model as well as a high standard error and medium high relative spread in slope makes it a poor determining factor. This maybe could be explained by people coming into your store and not buying anything or people coming into your store and buying a lot of stuff. This may be caused by other factors such as customer complaints or advertising.

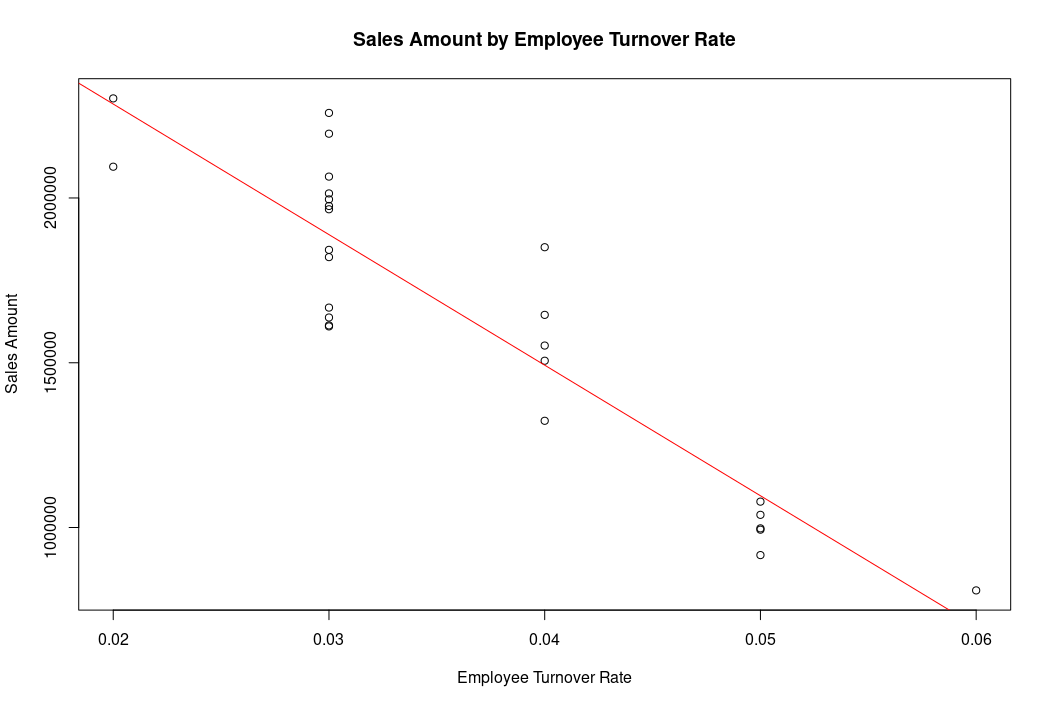


Standard Error: 285900

95% Confident that Slope is between: 63.46 to 125.86

R-Squared: 0.6053

With an R^2 of 0.6053, Research and Development expenses is the 4th worst determining factor of sales. It's only 60.5% of the change in data being able to be explained by rd expenses with a high standard error and relatively mid to high spread in slope makes it still a poor determiner of sales amount. This could be due to research going into the wrong things that aren’t helping the store or could also be that the standard company formula for getting sales works very well and needs very little improvement.



Standard Error: 189600

95% Confident that Slope is between: -47038525 to -32207561

R-Squared: 0.8263

At an r^2 of 0.8263, employee turnover rate is a pretty good determining factor of sales. 82.6% of the change in sales can be explained by the employee turnover rate of the store. This paired with a medium standard error and medium relative slope spread leads to employee turnover rate being a good factor to determine sales. The slope does loop very large because of the low small employee turnover rate but that is to be expected when the spread of turnover rate is 0.06%. It Is quite easy to explain as having a right turnover rate tends to be correlated with a more toxic and less caring work environment which leads to worse customer service and more customer complaints about poor treatment from workers.

List from best to worst determining factors of sales amount (with what they cause)

1. Customer complaints (Negative effect on sales)
2. Number of employees (Positive effect on sales)
3. Number of sales campaigns (Positive effect on sales)
4. Employee turnover rate (Negative effect on sales)
5. RD expenses (Positive effect on sales)
6. Number of customers (Positive effect on sales)
7. Number of new products on the market (Positive effect on sales)
8. Material cost (Not correlated)

