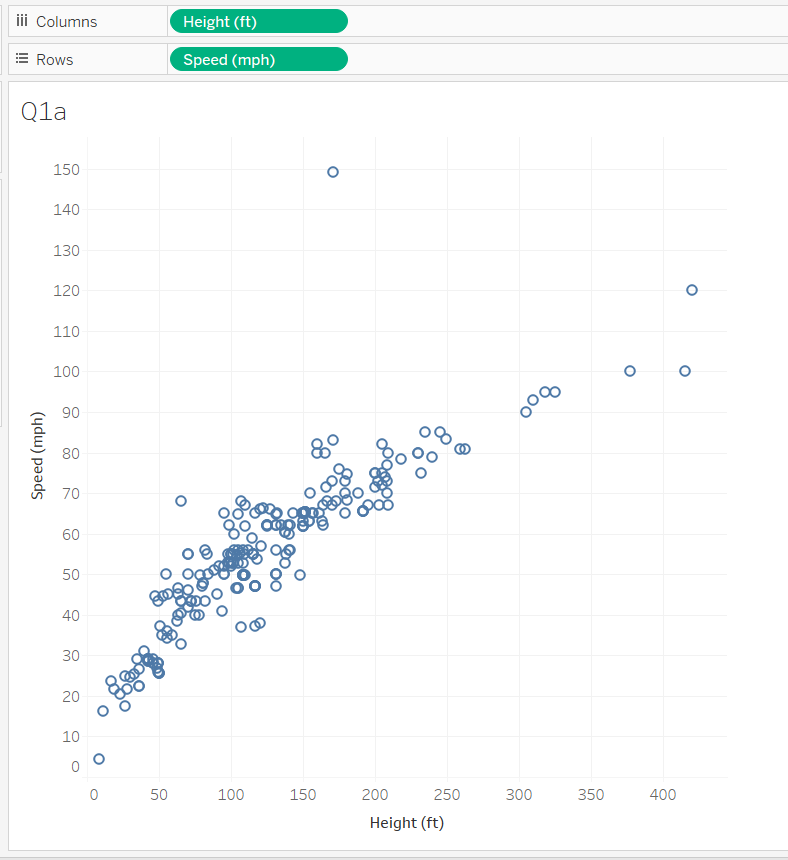
**DATA ANALYTICS AND VISUALIZATION WITH TABLEAU**

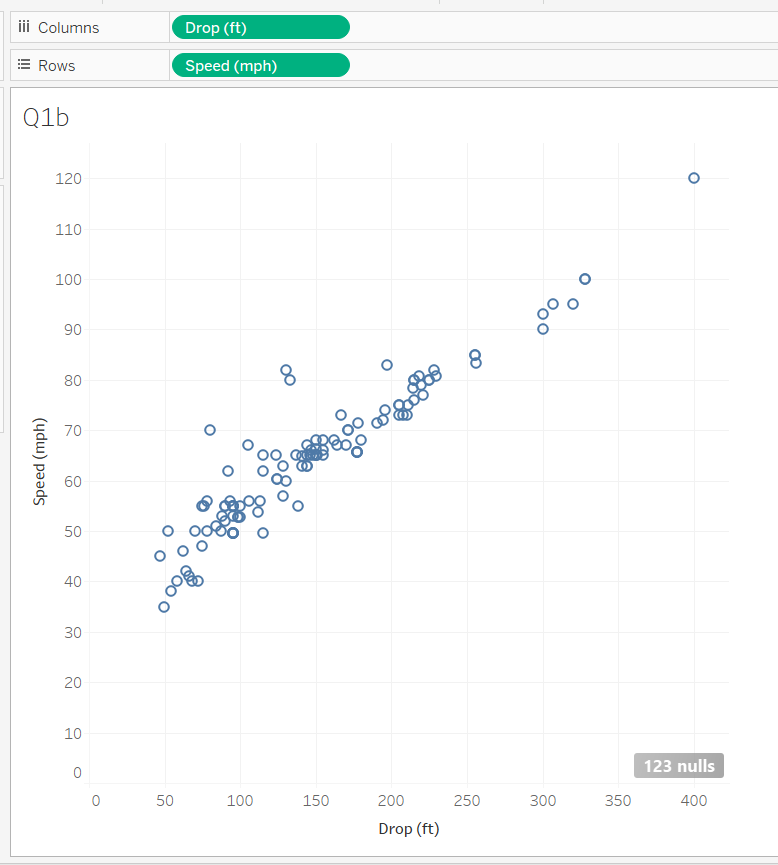
**Assignment: Day 4**

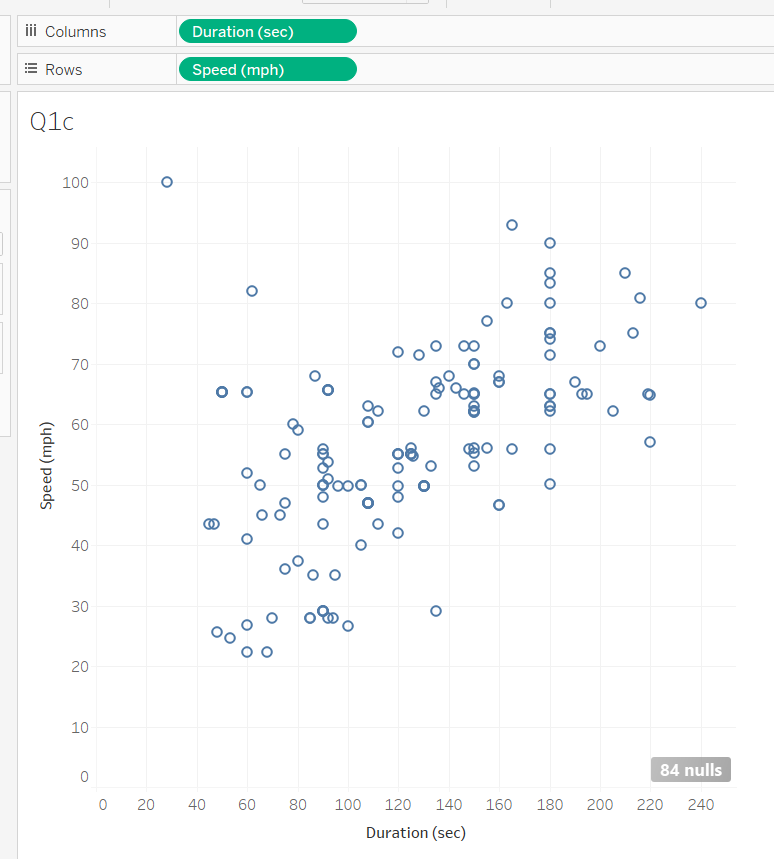
Rahul S Bhat

PES2UG19CS315

7th Semester

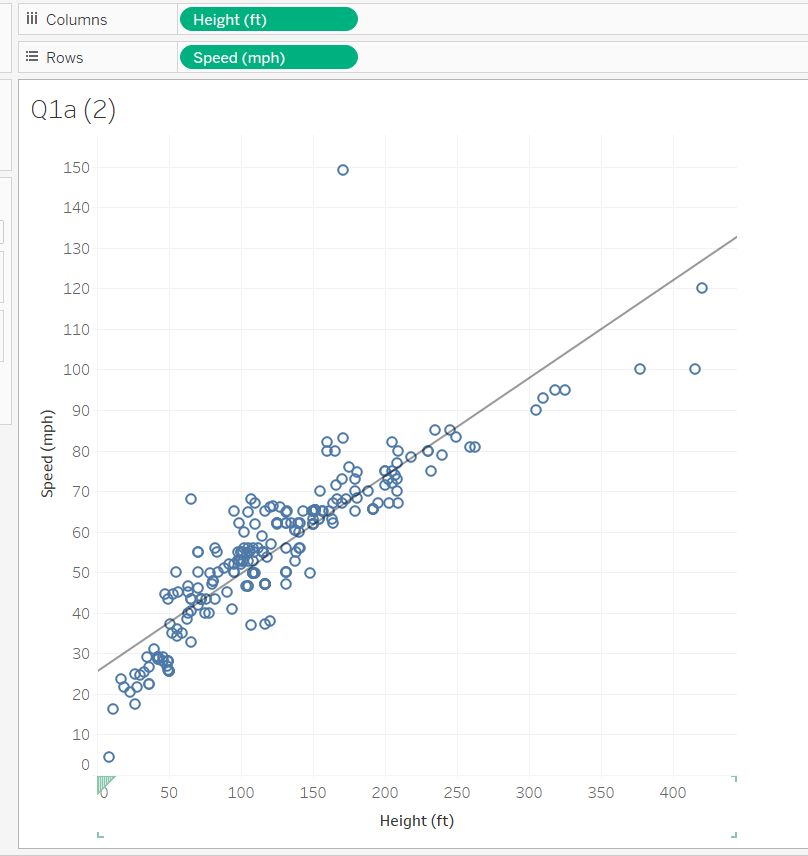


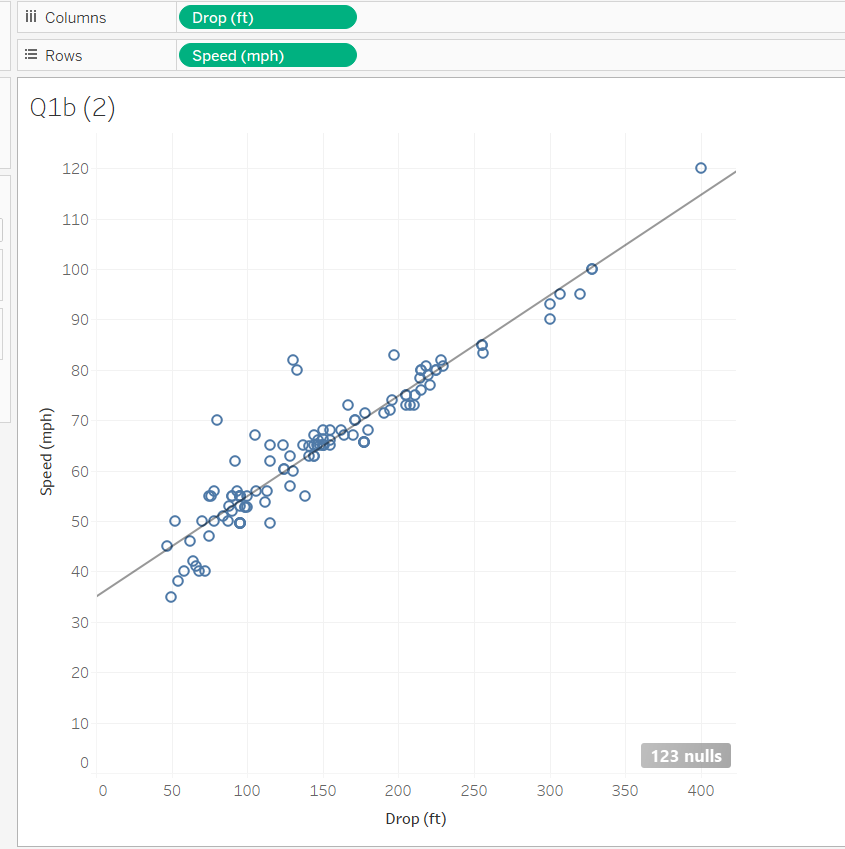




**Question One:** Evaluate the regression conditions for each plot. Explain why or why not it is appropriate to run a regression analysis on each plot. Please address all conditions covered in this module.

It is appropriate to run regression analysis on plots Q1a and Q1b as the points are not too scattered and can approximately fit a straight line. Q1c cannot be used for regression analysis as the data is all scattered and doesn’t seem to fit a line or a curve and hence it is not possible to predict the relation among the two plotted variables.



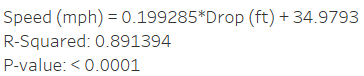


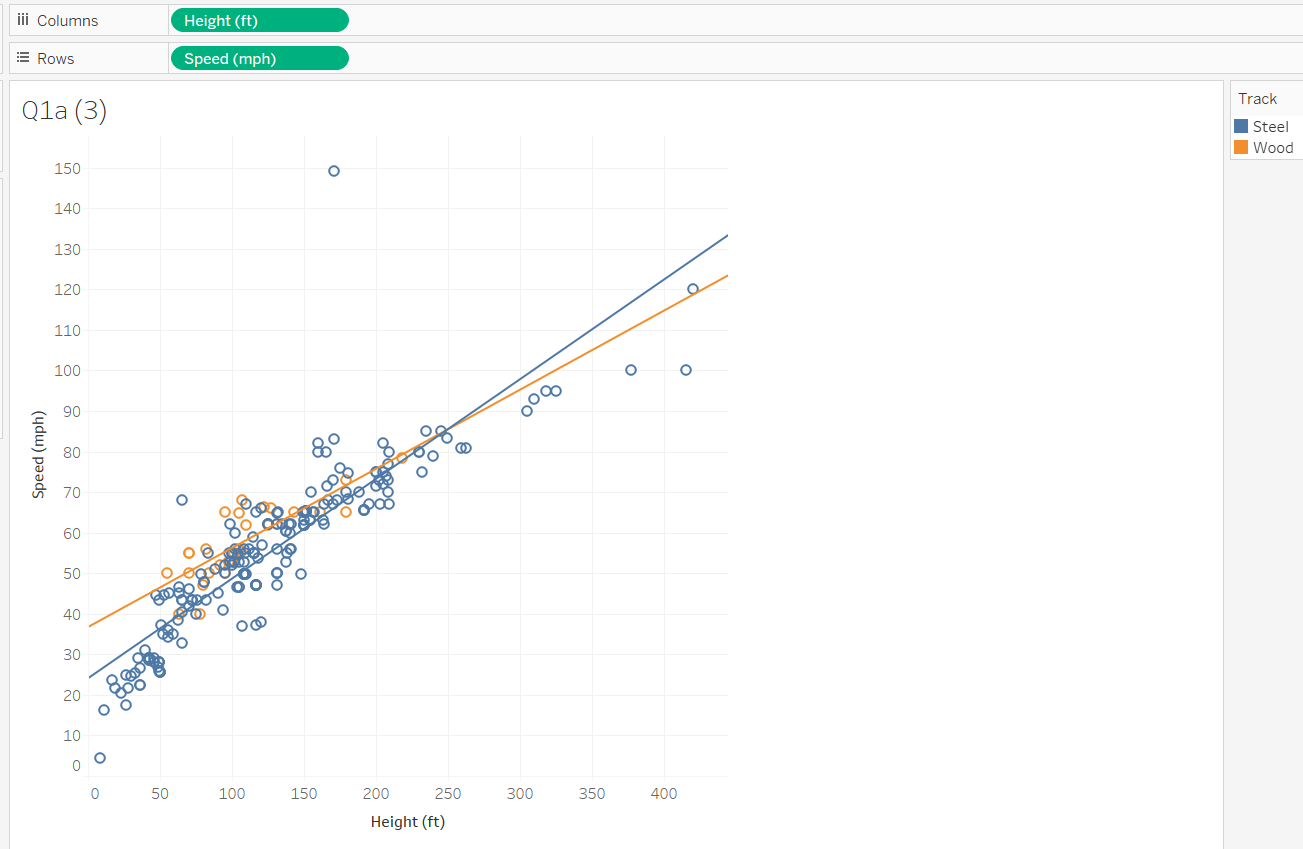
**Question Two:** Which variable(s) can you use to predict Speed?

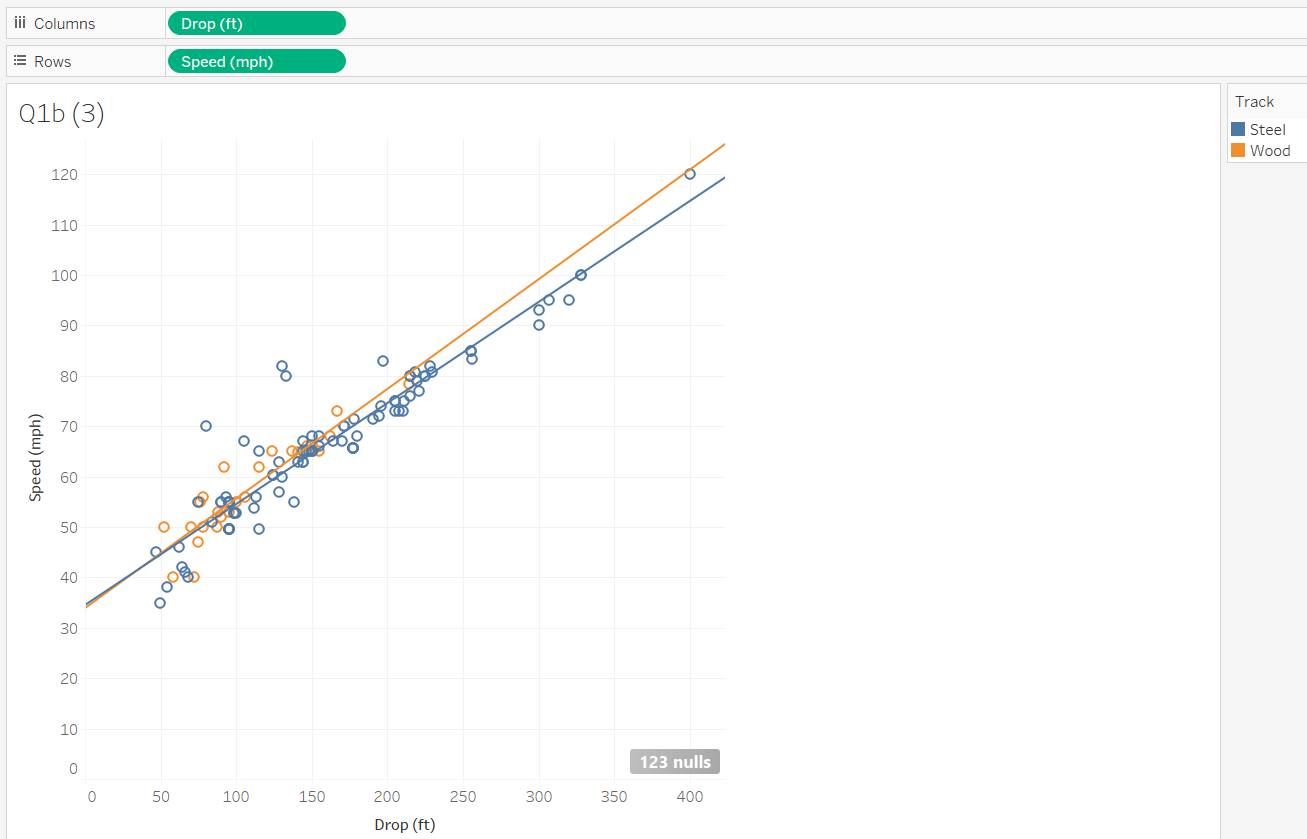
Drop can be used to predict speed as it fits the linear trend line better. (also higher R-Squared value)

**Question Three:** What is the regression equation for the trend line(s)?

For Q1a: 

For Q1b: 

****

****

**Question Four:** Are there any differences between steel and wooden coasters in your analyses? Explain in few sentences.

There are differences.

|  |  |
| --- | --- |
| Q1a(3) | Q1b(3) |
| Wood    Steel | Wood    Steel |

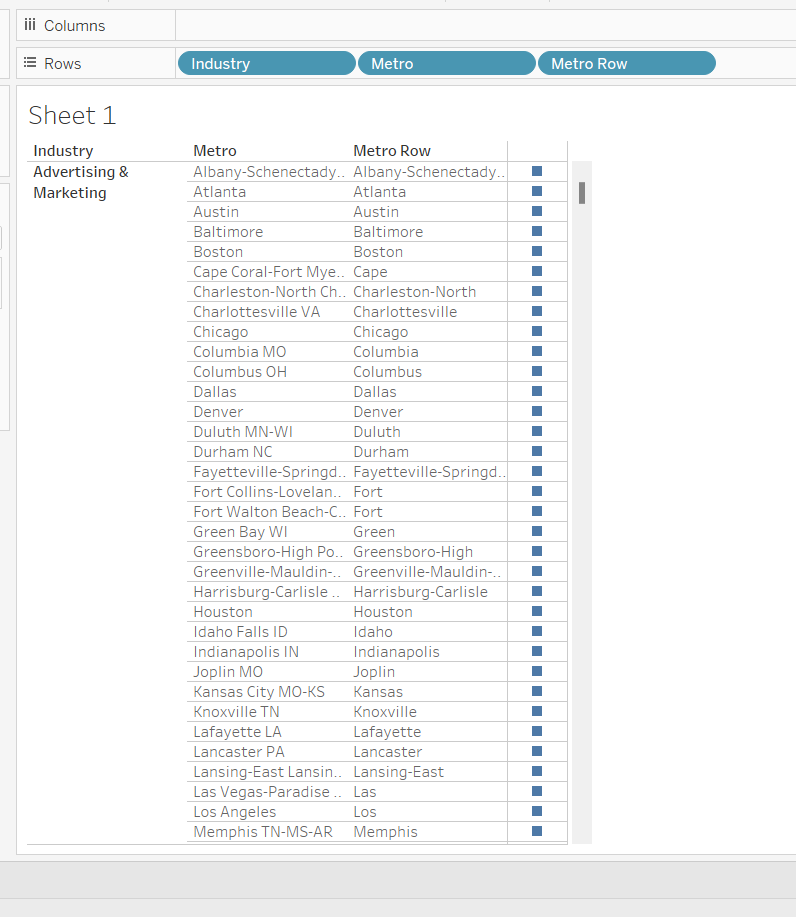
For Q1a(3), R-squared value for Steel is greater and hence we can say that the model fits better for steel category

Similarly in case of Q1b(3) R-squared value is greater for Steel and hence we can say that the model fits better for steel category and the more precisely, we can use drop values under steel track category to predict the speed

**Question Five**

*Demonstrate the following*

*row calculation*

**

calculate field

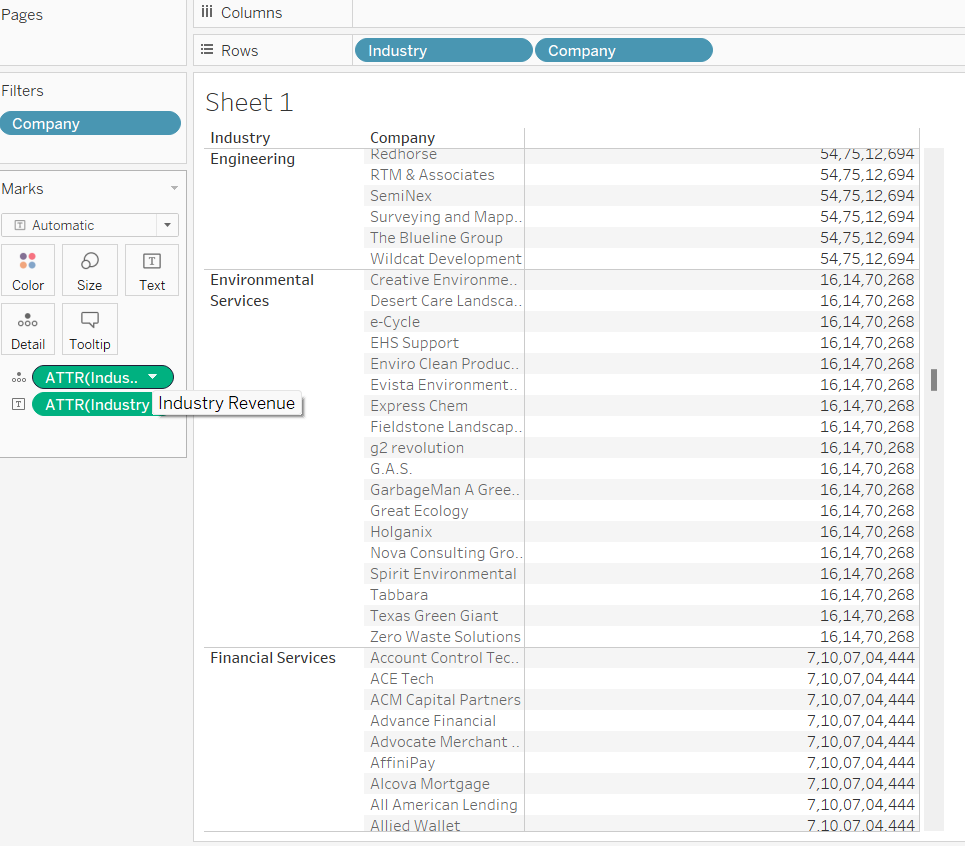
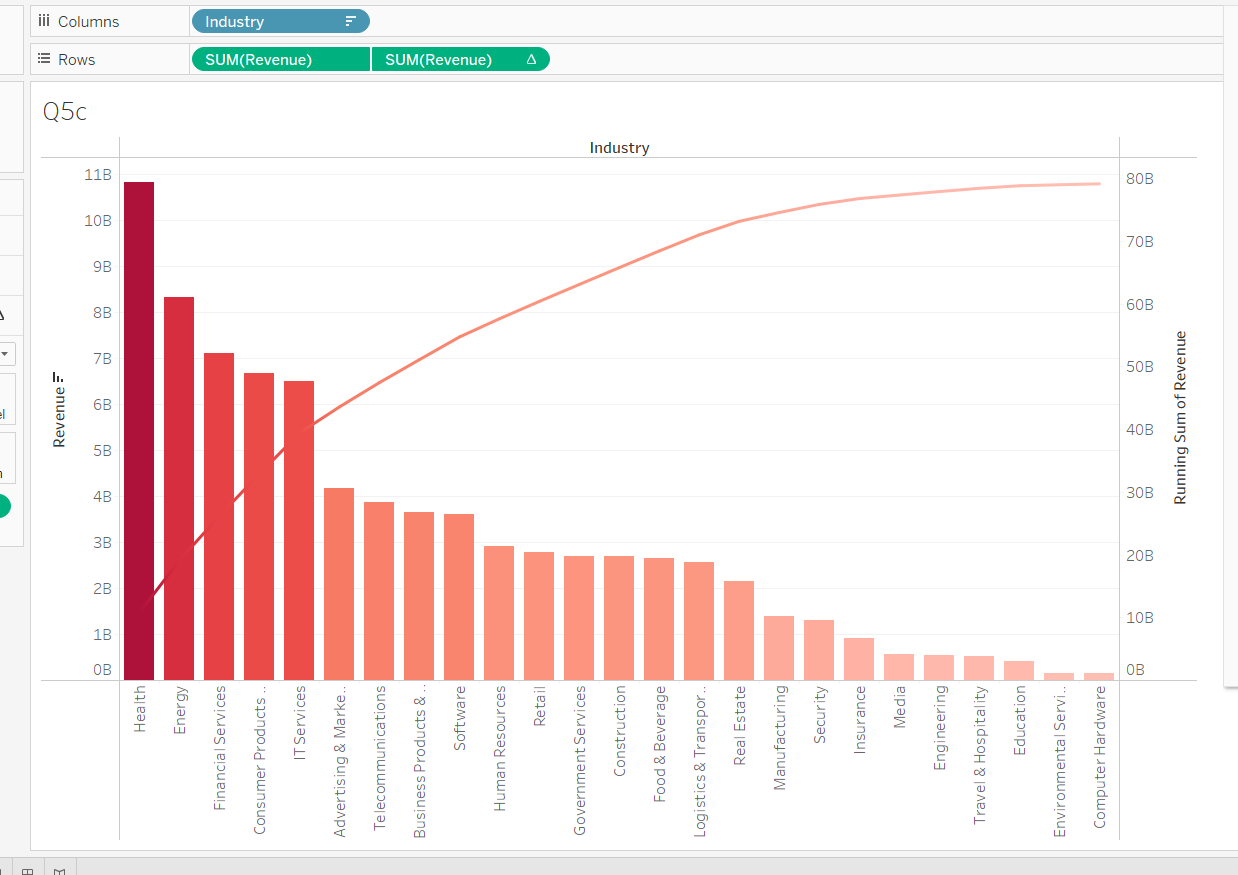


table calculation



ranking

