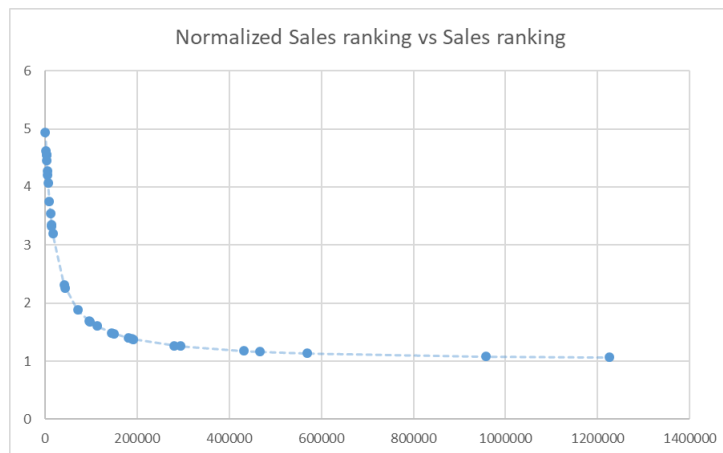


For this project, I have considered four metrics: **SalesRank**, **AvgRating**, **TotalReviews**, and **DegreeCentrality**.

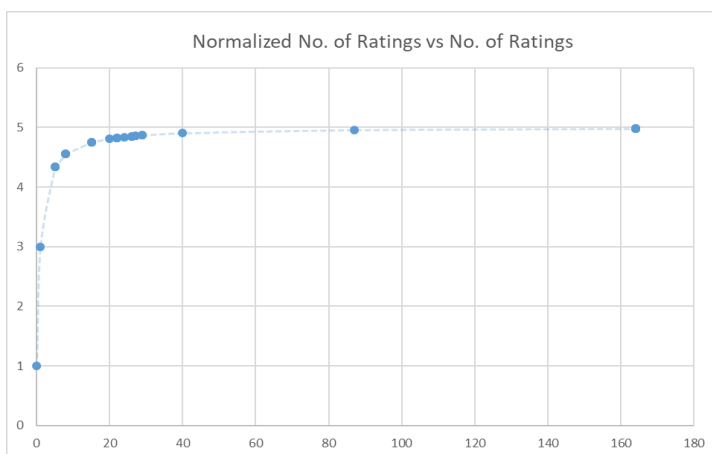
1. I first normalized **SalesRank** with scores in the range 1 to 5, where 1 indicated that an item had a high SalesRank number (and thus sold less) and 5 indicated a low SalesRank number. The formula I used to normalize is:

$$f(x) = \frac{ax+b}{cx+d} \quad \text{where } \frac{a}{c} = 5 \quad \text{and} \quad \frac{b}{d} = 1.$$



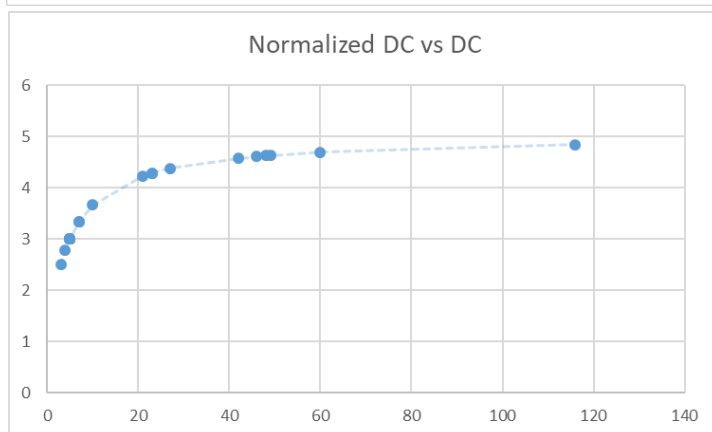
In this case, I chose $a = 5$, $c = 1$, and b & $d = 20,000$. And the subtracted this result from 6 to reflect the inverse nature of SalesRanking values. This gave me the curve on the left.

I chose these values so that the curve reflects the decreasing marginal importance of SalesRanking as the rank number increases. Thus, the numerically smallest SalesRanking have more variation in the normalized ratings.



2. I then normalized **TotalReviews** using the same formula, but with b & $d = 2$, which gave me the curve on the left.

These values reflect the high importance of the first few numbers of reviews, and the decreasing marginal utility of the numbers as the number of reviews increases.



3. I then normalized **DegreeCentrality** using the same formula, but with b & $d = 5$, which gave me the curve on the left.

The curve is less steep than the other two since a higher number for DegreeCentrality still merits a higher score to reflect its marginal utility.

Using these normalized values, and the **AvgRating** metric (which I left as-is, and is already in a range of 1 to 5), to come up with a composite metric. The weights I assigned each of these metrics can be seen in this equation:

$$\text{Composite measure} = \text{NormalizedSalesRank} * 0.3 + \text{NormalizedTotalReviews} * 0.2 + \text{NormalizedAvgRating} * 0.4 + \text{NormalizedDegreeCentrality} * 0.1$$

Highest importance was given to the average ratings of the book, and then the sales rank and number of total reviews since there is a similar pattern that appears for the original purchase.