Scholarship with Analysis Templates

Daniel Chen

Billboards dataset from raw to tidy to EDA. Data in its orginal form was in "long" format with a song on each row and a separate column for the ranking from week 1 to week 72. We "tidied" the data so it follows Hadley Wickham's tidy data principles.

Figure 1 shows the average rating of songs over the year with a different line drawn depending on which month the song was initially released.
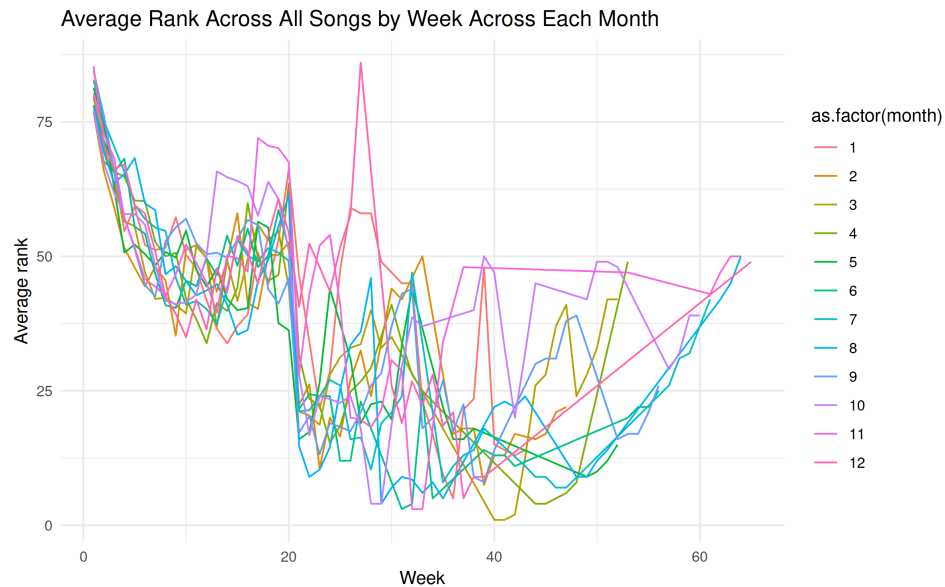

Figure 1: Average ratings over time by different starting months

Figure 2 shows the point estimate and 95% confidence interval looking at the rank of a song in the year 2000 as the weeks go by.
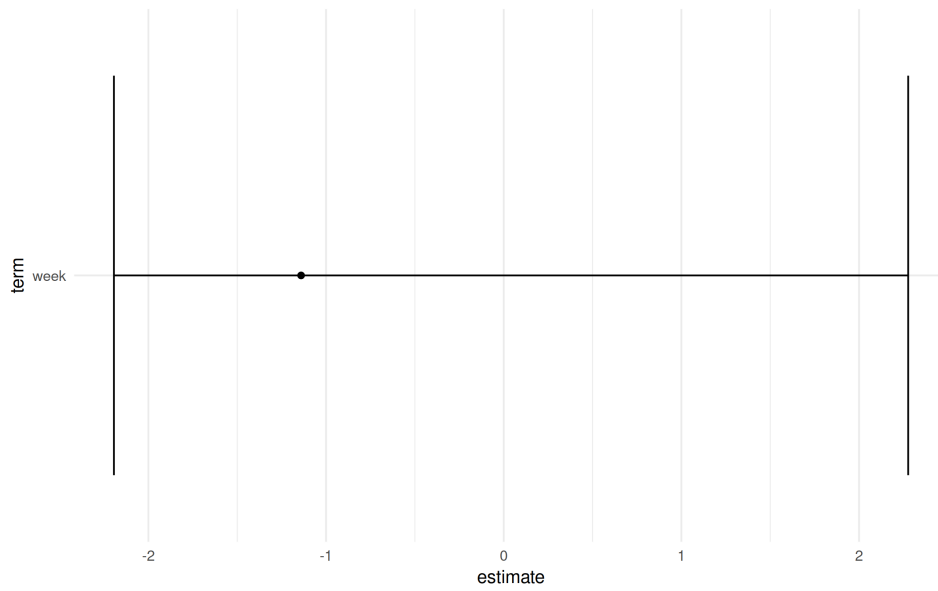
Figure 2: How the rank of a song changes weekly

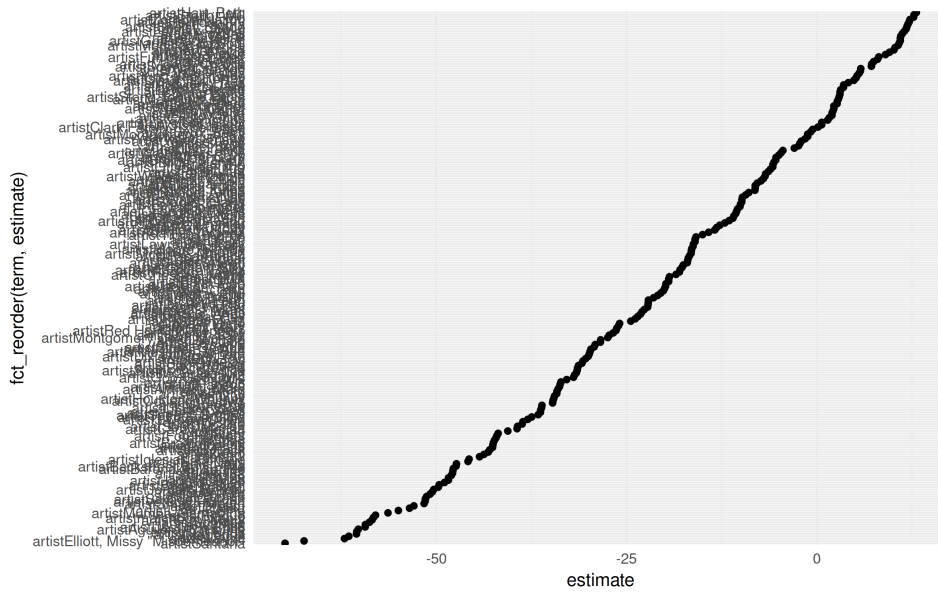In figure 3 we see that Missy Elliott isn't doing so great. Sad. But this analysis is a bit nonsensical anyway.

Figure 3: Ranks of songs by artist and week