New Music Friday

By Caitlin Nguyen

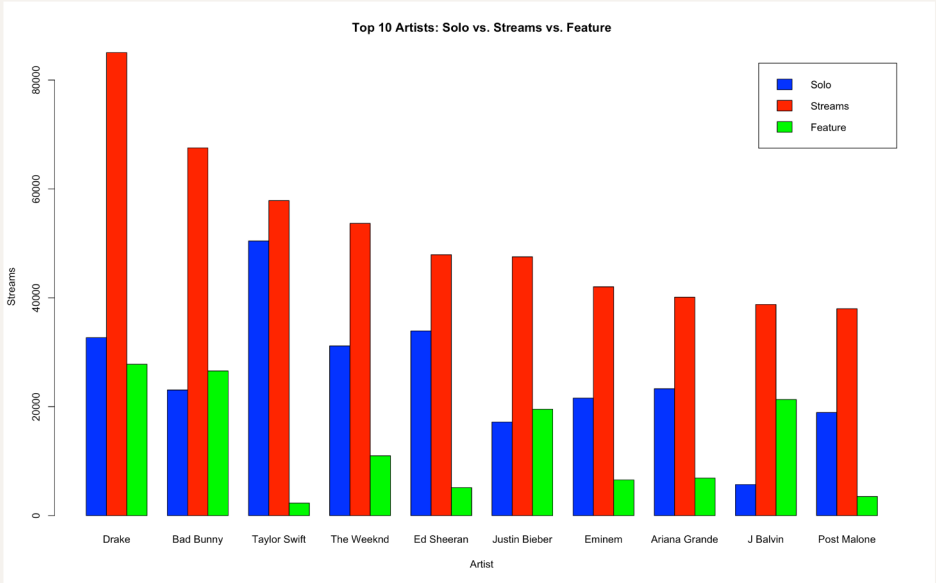
Problem

Universal Music Group wants to maximize the chances of getting one of their artist’s songs featured on popular Spotify playlists like "New Music Friday”

Hypothesis #1

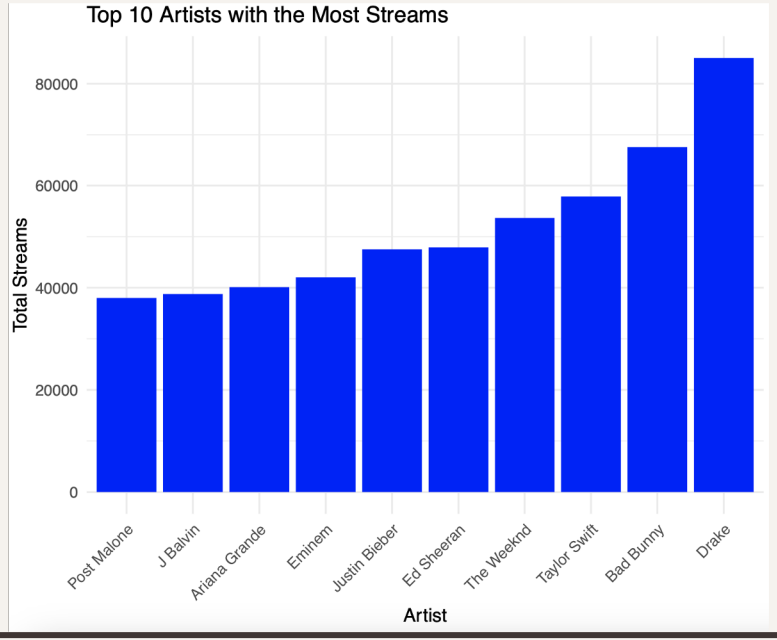
- Artists who have more solo streams tend to have higher total streams on Spotify

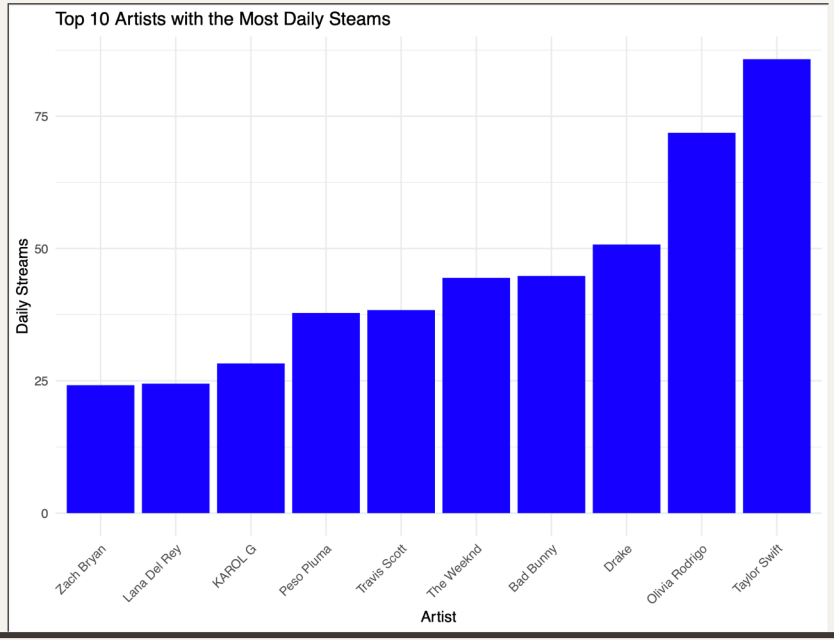
- This hypothesis finds support in the data, as it reveals that 7 out of the top 10 artists in our analysis have a notably higher number of streams for their solo work compared to their featured streams. This suggests that these artists, who already enjoy substantial streaming success, may not necessarily seek collaborations with other artists.
- The data implies that artists with already impressive streaming numbers may feel less compelled to engage in collaborations for exposure, as their solo endeavors are already attracting significant attention from listeners. This underscores the idea that high-performing artists often have the autonomy to focus on their individual projects, as they enjoy a dedicated and engaged fan base.



Hypothesis #2

Artists with a higher number of daily streams are more likely to get their songs featured on popular Spotify playlists like 'New Music Friday.





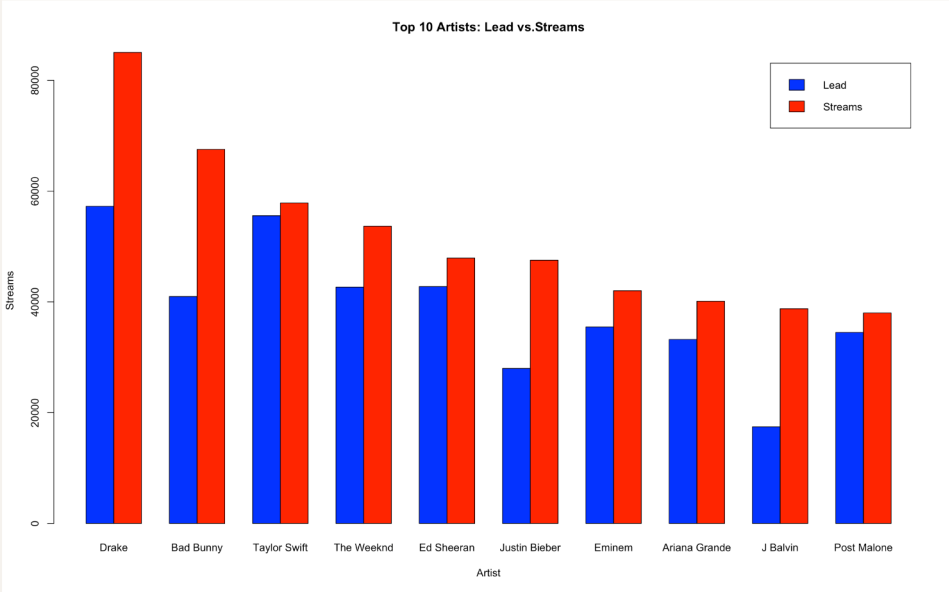
Hypothesis #2 research

- Being featured “New Music Friday” is based on not only one but both being top 10 Artists for Daily and Overall Streams
- 5 artists that are are on “Top 10 Artists with the Most Streams” are not on the “Top 10 Artists with the Most Daily Streams”
- Olivia Rodrigo and Peso Pluma are considered relatively new or emerging artists who have recently gained significant popularity. Despite not being among the top 10 most streamed artists overall, they have experienced a surge in daily streams due to their current success.
- These artists' current success and momentum make them strong candidates for playlist inclusion, emphasizing the importance of staying current and relevant in the music streaming landscape.
- Artists with high total streams have a proven track record of success, engaged fan bases, and industry support, all of which make them attractive candidates for playlist inclusion, including on high-profile playlists like "New Music Friday."

Hypothesis #3

Songs with a high 'lead' count, indicating the artist as the primary artist, are more likely to be featured on Spotify playlists like 'New Music Friday.

- This information can be valuable for strategizing song releases and collaborations to increase the likelihood of playlist features.
- The number of streams in the "lead" column reflects how many times songs where the artist is the lead artist have been streamed on Spotify.
- Artists who have a high number of streams in the "As lead" column are generally more popular when they release music on their own or when they are the primary focus of a track.



Solo Vs. Streams Vs. Feature

- “Artists with higher solo streams tend to have higher total streams" suggests that solo success is a significant factor in an artist's overall streaming performance. While feature collaborations can contribute to an artist's total streams, it's their ability to excel individually (solo) that often serves as a strong indicator of their potential for success on playlists like "New Music Friday." Artists with a substantial solo following may have a better chance of being featured on such playlists because they have already demonstrated their appeal to listeners on their terms.

Daily Streams Vs. Most Streams

- In summary, the relationship between the two lists indicates that while some artists with the most streams are established and maintain their popularity over time, there is also room for emerging artists to make an impact and gain recognition, particularly if they experience a surge in daily streams due to playlist placements or viral hits.
- This dynamic nature of the music industry keeps the balance between established and new talents.

Lead Vs. Most Streams

- The relationship between "Lead" and "Streams" suggests that an artist's success as the lead artist or their solo work can significantly contribute to their overall popularity and streaming numbers. It reflects the impact of their individual artistry on their streaming success, as opposed to collaborative efforts where they might be featured artists.

Final Recommendation

- By emphasizing the creation and promotion of solo work, artists and music labels can increase their chances of being featured on popular Spotify playlists like "New Music Friday."
- This approach leverages the strength of an artist's individual contributions and appeals to playlist curators seeking diverse and captivating music for their playlists.

Appendix: Data Set

	A	B	C	D	E	F
1	**Artist**	**Streams**	**Daily**	**Lead**	**Solo**	**Feature**
2	**Drake**	85,041.3	50.775	57,252.6	32,681.6	27,788.7
3	**Bad Bunny**	67,533.0	44.82	40,969.6	23,073.0	26,563.4
4	**Taylor Swift**	57,859.0	85.793	55,566.7	50,425.7	2,292.4
5	**The Weeknd**	53,665.2	44.437	42,673.3	31,164.2	10,991.9
6	**Ed Sheeran**	47,907.7	17.506	42,767.9	33,917.0	5,139.8
7	**Justin Bieber**	47,525.7	18.868	27,988.0	17,183.9	19,537.7
8	**Eminem**	42,029.0	20.175	35,475.8	21,576.7	6,553.2
9	**Ariana Grande**	40,111.0	17.158	33,219.8	23,307.3	6,891.2
10	**J Balvin**	38,774.8	11.784	17,450.7	5,699.8	21,324.2

Appendix

Solo-streams-feature.R

```r
# Read the Excel file into a data frame
library(readxl)
data_set <- read_excel("/Users/caitlin/Desktop/spotify.xlsx")

# Load the required libraries
library(dplyr)

# Ensure that the "Solo", "Streams", and "Feature" column is treated as a factor or character
data_set$Solo <- as.factor(data_set$Solo)
data_set$Streams <- as.factor(data_set$Streams)
data_set$Feature <- as.factor(data_set$Feature)

# Get the top 10 artists with the most streams
top_artists <- data_set %>%
  arrange(desc(Streams)) %>%
  head(10)

# Convert "Solo", "Streams", and "Feature" to factors if they are not already
top_artists$Solo <- as.factor(top_artists$Solo)
top_artists$Streams <- as.factor(top_artists$Streams)
top_artists$Feature <- as.factor(top_artists$Feature)

# Convert "Solo", "Streams", and "Feature" to character, replacing missing or zero values with "Missing"
top_artists$Solo <- as.character(top_artists$Solo)
top_artists$Solo[is.na(top_artists$Solo) | top_artists$Solo == "0"] <- "Missing"
top_artists$Streams <- as.character(top_artists$Streams)
top_artists$Streams[is.na(top_artists$Streams) | top_artists$Streams == "0"] <- "Missing"
top_artists$Feature <- as.character(top_artists$Feature)
top_artists$Feature[is.na(top_artists$Feature) | top_artists$Feature == "0"] <- "Missing"

# Create a matrix for the barplot data
barplot_data <- matrix(
  c(as.numeric(top_artists$Solo), as.numeric(top_artists$Streams), as.numeric(top_artists$Feature)),  # Convert to numeric
  nrow = 3,
  byrow = TRUE,
  dimnames = list(c("Solo", "Streams", "Feature"), top_artists$Artist)
)

# Create a grouped barplot
barplot(
  barplot_data,
  beside = TRUE,
  col = c("blue", "red", "green"),
  main = "Top 10 Artists: Solo vs. Streams vs. Feature",
  xlab = "Artist",
  ylab = "Streams",
  legend.text = TRUE
)
```

LeadVs.Streams.R

```r
# Read the Excel file into a data frame
library(readxl)
data_set <- read_excel("/Users/caitlin/Desktop/spotify.xlsx")

# Load the required libraries
library(dplyr)

# Ensure that the "Feature" column is treated as a factor or character
data_set$Feature <- as.factor(data_set$Feature)

# Get the top 10 artists with the most streams
top_artists <- data_set %>%
  arrange(desc(Streams)) %>%
  head(10)

# Convert "Solo" and "Feature" to factors if they are not already
top_artists$Solo <- as.factor(top_artists$Solo)
top_artists$Feature <- as.factor(top_artists$Feature)

# Convert "Solo" to character, replacing missing or zero values with "Missing"
top_artists$Solo <- as.character(top_artists$Solo)
top_artists$Solo[is.na(top_artists$Solo) | top_artists$Solo == "0"] <- "Missing"
top_artists$Feature <- as.character(top_artists$Feature)
top_artists$Feature[is.na(top_artists$Feature) | top_artists$Feature == "0"] <- "Missing"

# Create a matrix for the barplot data
barplot_data <- matrix(
  c(as.numeric(top_artists$Solo), as.numeric(top_artists$Feature)),  # Convert to numeric
  nrow = 2,
  byrow = TRUE,
  dimnames = list(c("Solo", "Feature"), top_artists$Artist)
)

# Create a grouped barplot
barplot(
  barplot_data,
  beside = TRUE,
  col = c("blue", "red"),
  main = "Top 10 Artists: Solo vs. Feature Streams",
  xlab = "Artist",
  ylab = "Streams",
  legend.text = TRUE
)
```