



## Super Bowl LVII

### 2023 Game Day Analytics Challenge

### *Kansas City Chiefs vs Philadelphia Eagles*



**Vamshikrishna Vuppunuthala**

## **What is the Game Day Analytics Challenge and how did it start?**

Dr. Chong Oh, Director of the Information Systems Undergraduate Program and Assistant Professor at the University of Utah, established the Annual Game Day Analytics Challenge thoroughly examine Twitter tweets pertaining to Super Bowl advertisements. Students get the chance to gain hands-on experience with data purification, visualization tools, and insights from the most effective commercials, thanks to the help of Firetoss, OptConnect, ThomasArts, Layton Construction, Pepsi, and Alteryx.

## **What do students get to do and what is expected?**

Teams were set out to use the tweets data, analyze it based on their perspective and gather insights. These insights are put in 3 deliverables i. Infographics, ii. White Paper, iii. Presentation. Some insights could be “What aspects of the commercial advertisements make them effective for the viewers” or “What is the reach of the commercial advertisements among the viewers?”. In Super Bowl LVII there were 140+ commercial ads with 30 seconds ad slot costing ~\$7 Million (the most expensive Super Bowl ad slot of all time).

## **Executive Summary of our analysis:**

- **Warner Bros** and **Kia Binky Dad** are two of the most popular brands on Twitter during the Super Bowl LVII.
- On the game day, the maximum number of tweets were tweeted between 8:00 PM – 8:15 PM, with 8:05 PM having the highest number of tweets (**11,326**).
- The hashtag **#TheFlashMovie** (Brand: Warner Bros) was the most used, with ~47k tweets containing this hashtag.
- **Kia Binky Dad | Kia Telluride X-Pro All-Terrain AWD SUV** tops the tweet count list during the game day period.
- **Limit Break** commercial has the highest number of re-tweets, **86%** more than **Paramount Plus: Stallone Face** commercial.
- **META** brand has the highest replies and **Warner Bros. - The Flash** commercial has the second highest.
- **MARVEL – GAURDIAN GALAXY** commercial has the highest number of impressions, 86% more than **Pepsi Zero Sugar: Ben Stiller** commercial.
- **MARVEL, DISNEY PLUS, WARNER BROS,** and **PARAMOUNT** lead in public engagement.

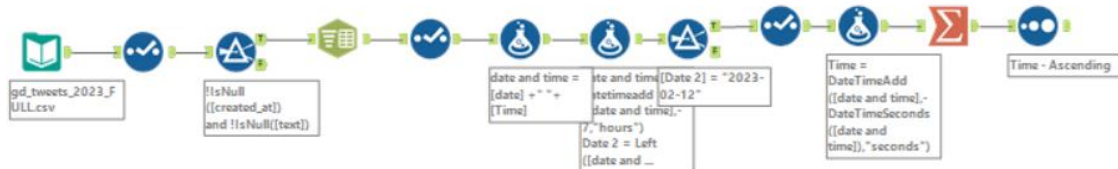
# Insights

*Note: We've used Alteryx as our base tool for the complete analysis*

## Tweets per minute on game day

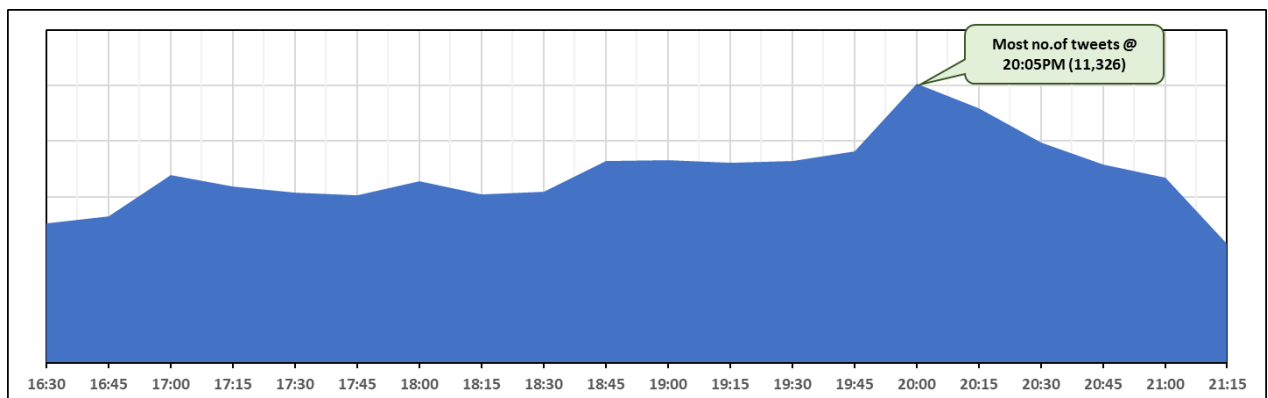
### Process Details:

- Input: "gd\_tweets\_2023\_FULL" file with all the tweets and time of the tweet
- Used the "created\_at" column which contained the time at which the tweet was created in UTC timezone format
- We then used the "Text to Columns" tool to parse the data from the column using the "T" delimiter, used the "datetimeadd" function to convert the data to MT timezone, rolled up the data (count of tweets) based on the time of creation to extract the final result
- Below is the workflow used:



### Observation:

- Tweets started to come in from 4:30 PM on the game day and we see that the maximum number of tweets were tweeted between 8:00 PM – 8:15 PM with 8:05 PM having the maximum number of tweets (11,326)



## Top Hashtags

People use the hashtag symbol (#) before a relevant keyword or phrase in their Tweet to categorize those Tweets and help them show more easily in Twitter searches and we've used those #'s as a part of the analysis to see which is the most used hashtag keyword during the Superbowl game.

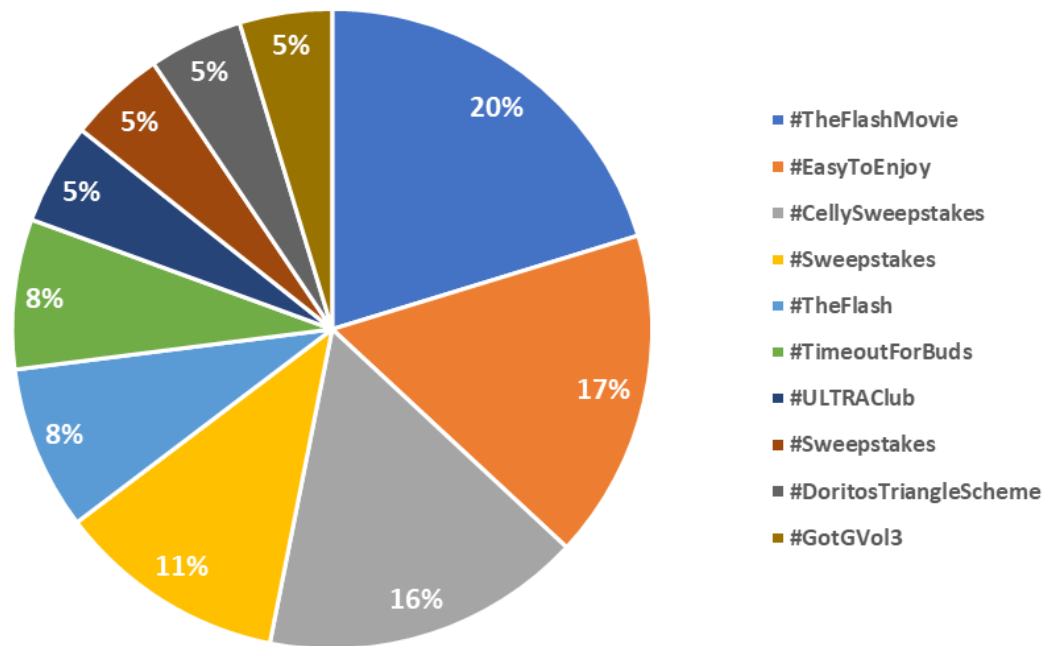
### Process Details:

- Input: "gd\_tweets\_2023\_FULL" file with all the tweets.
- Used "entities.hashtags" column which contained all the different hashtags used in that particular tweet.
- Data form in entities.hashtag column: [{"start": 37, 'end': 45, 'tag': '50years'}].
- We then used the "Text to Columns" tool to parse the data from the column using ":" & "\"" delimiters, used the "find and replace" tool to filter out the unwanted records after parsing and obtained the list of hashtags in each tweet in rows.
- We've then rolled up the count of hashtags using the "Summarize" tool to get the data cut for the count of tweets containing a particular hashtag.
- Below is the workflow used:

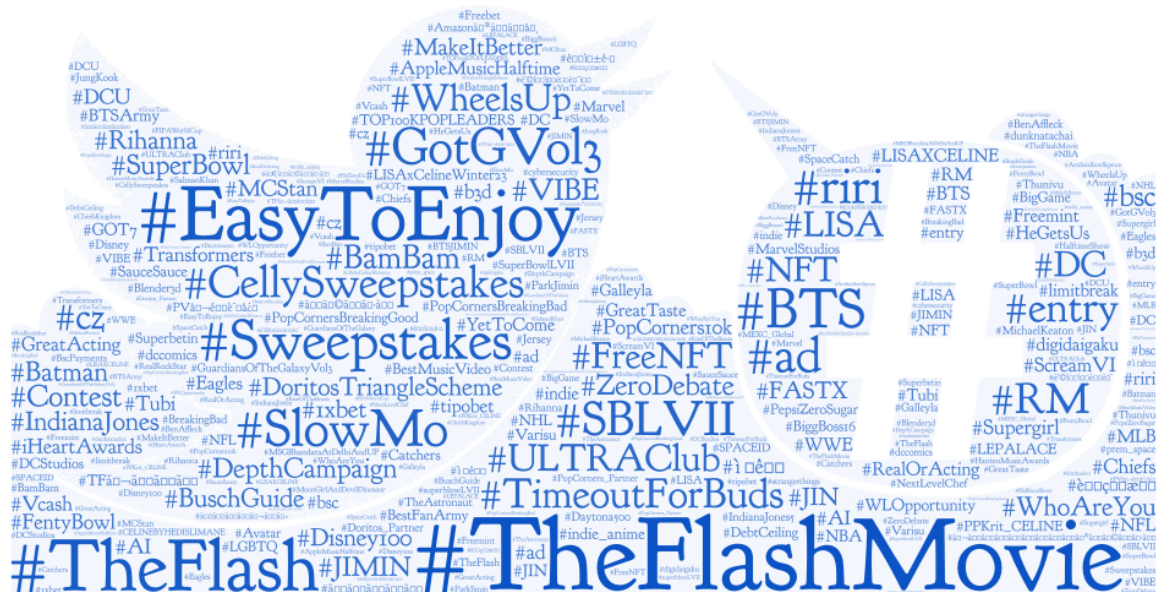


**Observation:**

- There was a total of ~**106k hashtags** used in tweets from the commercials of Super Bowl LVII.
- The hashtag **#TheFlashMovie** (*Brand: Warner Bros*) is used the most with ~**47k** tweets containing this hashtag. The brand related to this hashtag is **Warner Bros**



## Twitter Hashtags – Super Bowl LVII Word Cloud

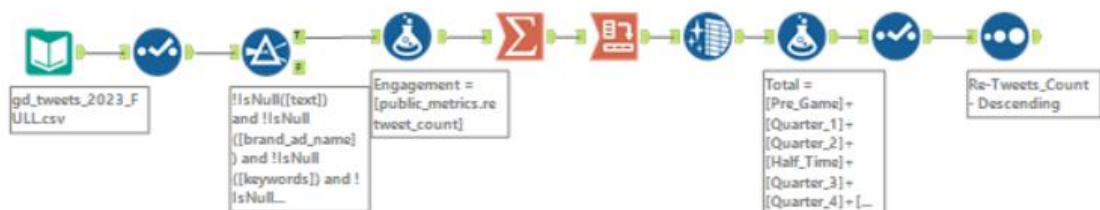


## Public Metrics:

- **Engagement:** The total number of times a user interacted with a Tweet. Clicks anywhere on the Tweet, including Retweets, replies, follows, likes, links, cards, hashtags, embedded media, usernames, profile photos, or Tweet expansion. In the current dataset,
- **Re-tweets:** Times a user retweeted the Tweet
- **Replies:** Times a user replied to the Tweet
- **Likes:** Times a user liked the Tweet
- **Impressions:** Times a user is served a Tweet in the timeline or search results
- **Engagement rate:** Number of engagements divided by impressions

## Process Details:

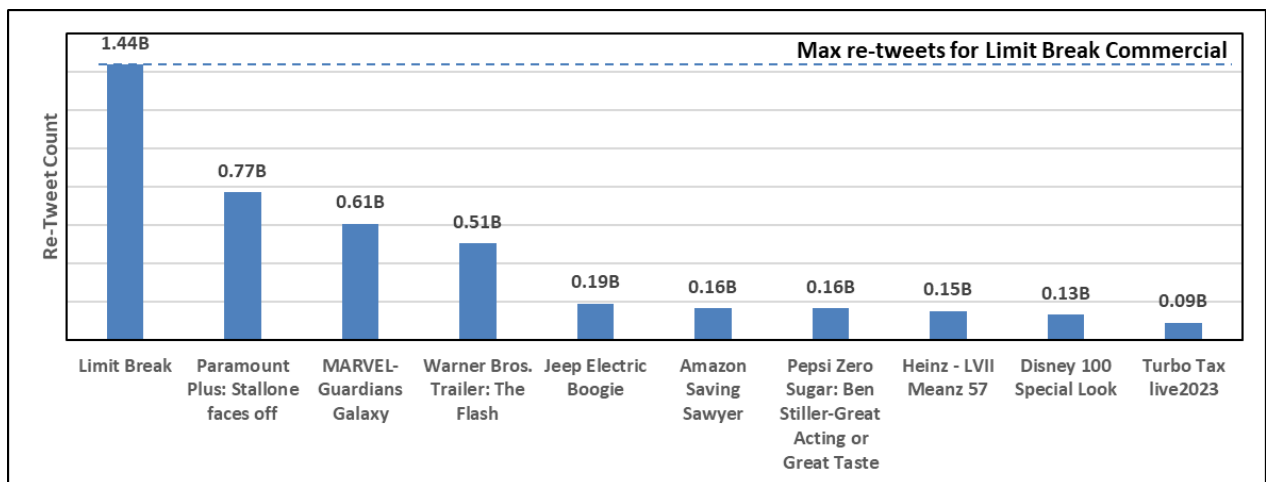
- Input: "gd\_tweets\_2023\_FULL" file with all the tweets
- Used "[public\_metrics.retweet\_count]", "[public\_metrics.likes\_count]", "[public\_metrics.reply\_count]", "[public\_metrics.impression\_count]" columns which give us each of the engagement metrics on each tweet.
- We converted the data type from "V\_String" to "Int64" to use and roll up the metrics for each brand.
- We then summed up all engagement metrics (re-tweets, replies, likes), named Tweet\_Engagement, and used this along with Impression\_count to calculate "Engagement\_Rate".
- Below is the workflow used for "[public\_metrics.retweet\_count]" and similar workflows are created for each of the other metrics:



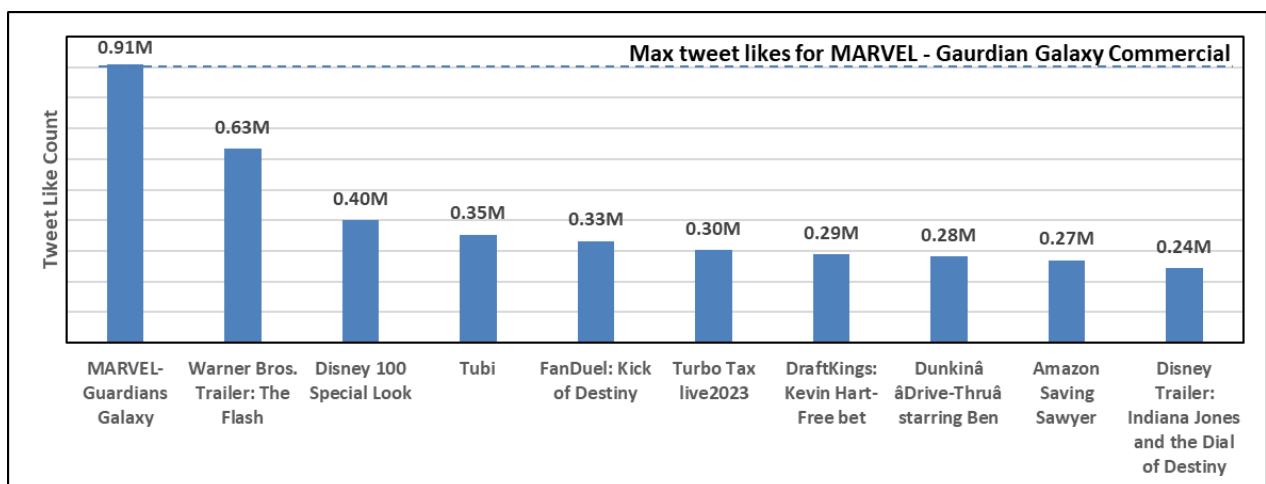
### Observation:

- **Limit Break** commercial has the highest number of re-tweets at **1.44B, (86% more)** than **Paramount Plus: Stallone Face** commercial which is #2.
- **MARVEL – GUARDIAN GALAXY** commercial has the highest number of likes at **0.91M**.
- **META** brand has the most number of replies at **~35k**, **Warner Bros – The Flash** commercial has the 2<sup>nd</sup> highest reply count at **~34k**
- **MARVEL – GUARDIAN GALAXY** commercial has the highest number of impressions at **0.87M, (86% more)** than **Pepsi Zero Sugar: Ben Stiller – Great Acting or Great Taste** commercial.
- **Limit break** has the highest engagement rate at **1,031 engagements per impression**
- Overall, most engagement is for genre movies with MARVEL, DISNEY PLUS, WARNER BROS, and PARAMOUNT leading the charts in all public metrics.

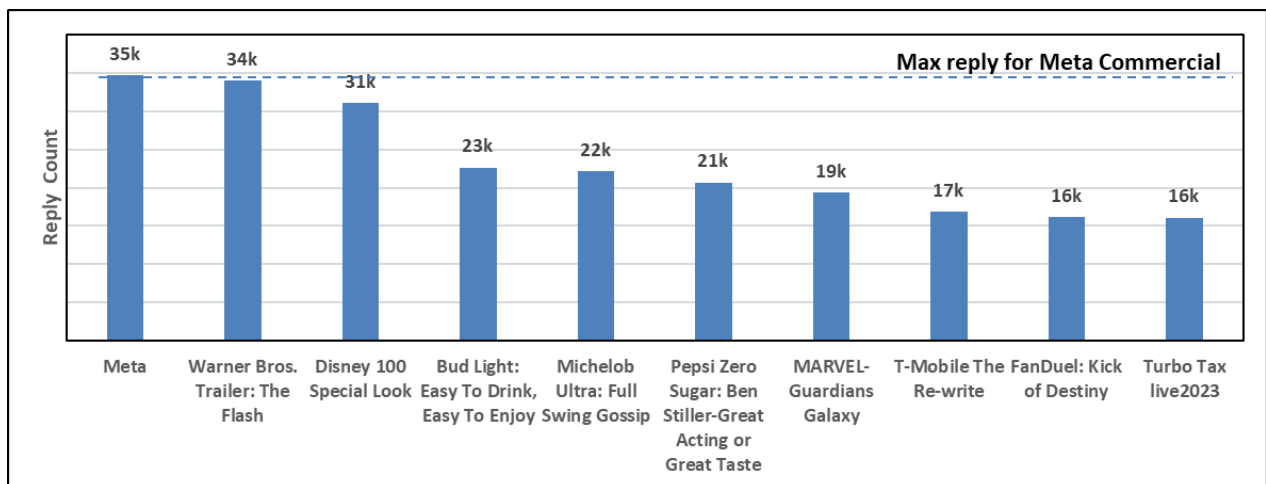
### Re-Tweet count:



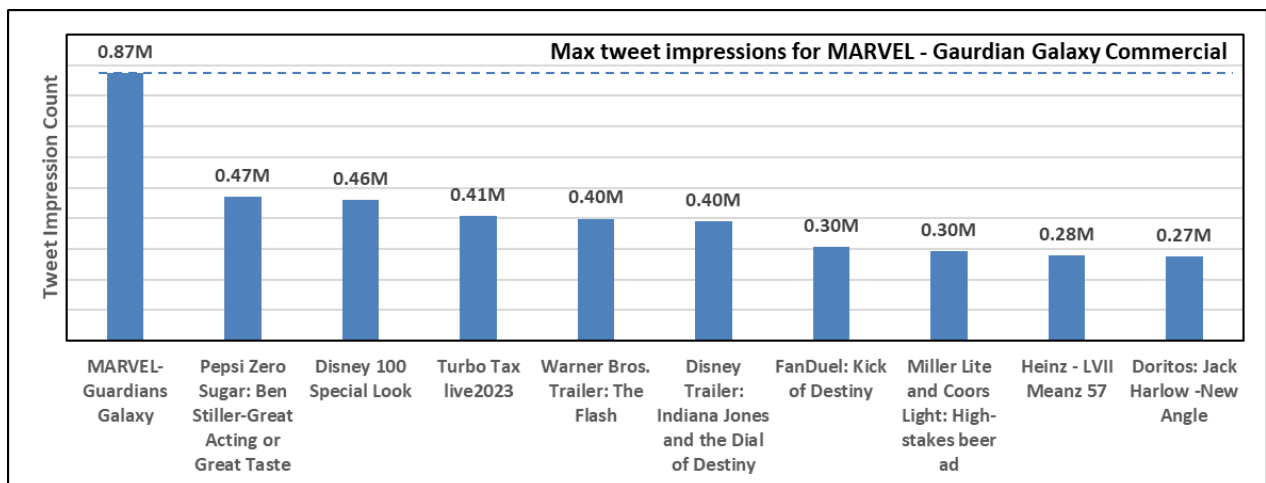
### Likes count:



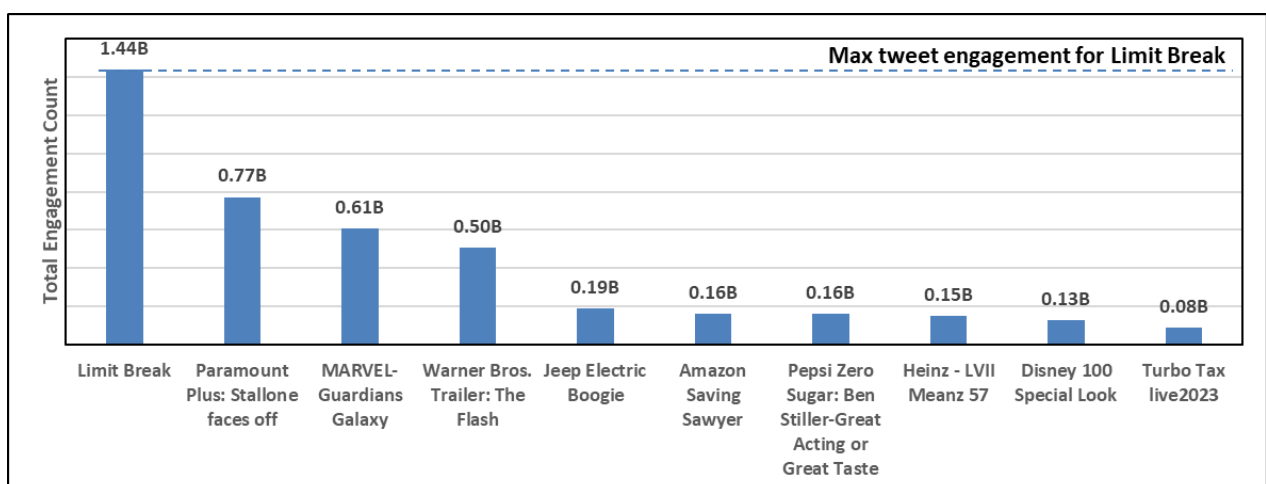
### Reply count:



### Impressions count:



### Total Engagement count:



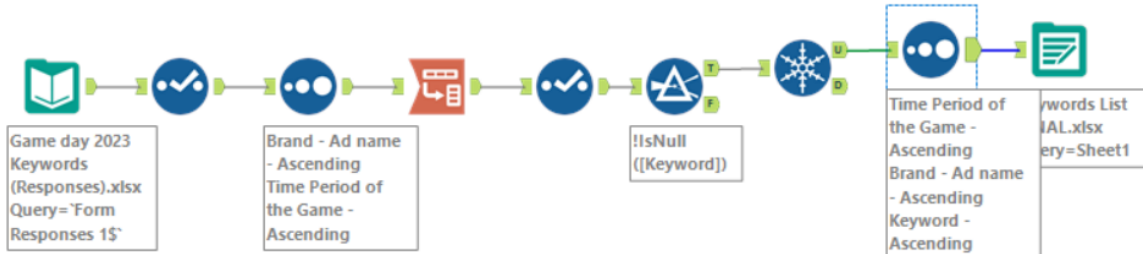


## Top Commercial advertisement for each day based on Tweet counts:

With the addition of a new dataset to the Game Day Analytics Challenge ("gdac\_2023\_keywords\_COUNTS"), which contains the number of tweets for each keyword during a specific timeframe, we tried to use this data to analyze the top tweets for each day.

### Process Details:

- Input 1: Keyword responses (Game Day 2023 Keywords (Responses))
- Since Input 1 is in wide format and we need a tall format for easy usage of this list, we used the "Transpose" tool to convert the format.
- Then the duplicates are removed at the Brand-Keyword level to extract the final keywords list.
- Input 2: Tweet count based on keywords ("gdac\_2023\_keywords\_COUNTS")
- The output of process 1 (Keywords List) is joined with Input 2 on keywords to map each record to a brand, to which the keywords are related.
- Tweet counts are rolled up on Brand-Ad Name and Start date to extract the Final Tweet\_Count for each brand on each day.
- The data is then sorted by Tweet\_Count (Desc) and then we choose the first record for each start date to get the highest Tweet\_Count brand for each day.
- Below are the workflows:
- To extract the final keywords list:

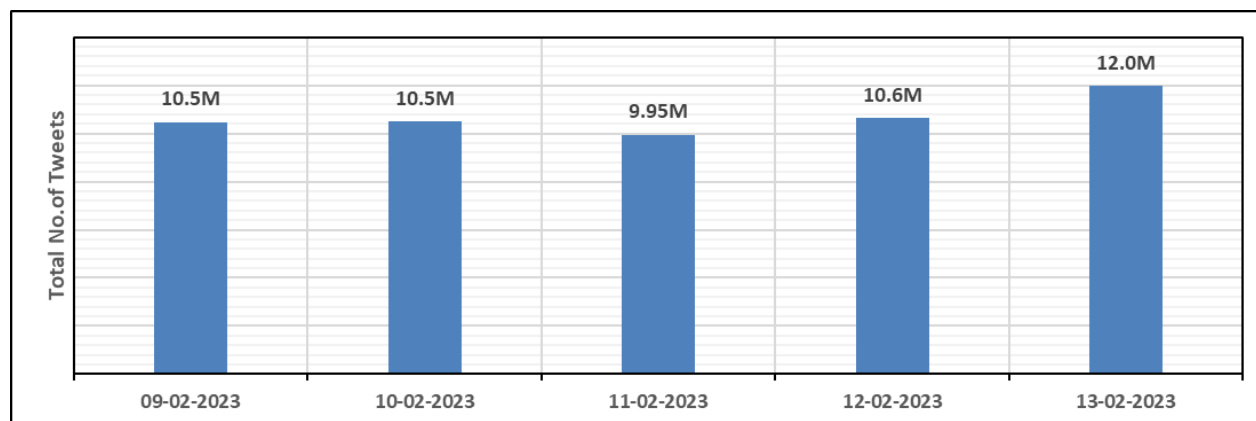


- To extract the max Tweet\_Count and the corresponding brand for each day:



**Observations:**

- Among all the brand and their commercial advertisements, **“Kia Binky Dad | The 2023 Kia Telluride X-Pro All-Terrain AWD SUV”** tops the list every day from 9<sup>th</sup> Feb, 2023 – 13<sup>th</sup> Feb 2023.



### **Brands – Commercial Ads. – Super Bowl LVII Word Cloud**

