

PART ONE: BRAINSTORMING A DATA SCIENCE QUESTION

For my project, I was curious about how effective Superbowl ads are. The Superbowl recently happened and I know the ads are always a cultural phenomenon. Many people watch just for the entertaining ads, and companies spend millions of dollars for 30-second spots. But is all of the fuss actually worth it?

This is a pretty broad question, so I used ChatGPT to try to narrow it down. I asked it for five ideas, and here's what it came up with:

Certainly! Here are five specific, measurable, and feasible data science questions related to Super Bowl advertisements:

- What is the immediate impact of Super Bowl ad viewership on brand sales in the weeks following the event?
- How does the effectiveness of Super Bowl ads vary across different demographic groups?
- What is the relationship between social media engagement during the Super Bowl and the subsequent virality of ads?
- Do brands that advertise during the Super Bowl experience a long-term increase in customer loyalty compared to those that do not?
- How does the timing of Super Bowl ad releases (e.g., pre-releases vs. game-day releases) affect overall ad performance and brand perception?

I thought many of these questions were pretty interesting, but it came down to which one I felt captured my original idea the most. I decided to go with the first suggestion, but also change the wording a bit to better suit my needs. My final question will be **“What was the immediate impact of Super Bowl 2024 ad viewership on brand interest and sales in the weeks following the event?”** This added response variable of interest can be measured many different ways, so I think that will give me a more robust answer.

My final question is specific, because it addresses specific aspects of what might make an ad “worth it” for a company. It's measurable and feasible as well. Off the top of my head, some ideas for data sources include google search trends, social media trends, viewership ratings on the ads, and perhaps sales data if the companies of interest release that.

PART TWO: IDENTIFYING POTENTIAL DATA SOURCES

Statista Data:

Statista has a couple datasets that could be of interest for my project, and I have linked them here.

[Super Bowl TV viewership 2025 | Statista](#)

[Super Bowl 30-second ad costs 2025 | Statista](#)

Statista is very easy to collect data from. As a student, I can directly download files in the form of spreadsheets.

Google Search Trends:

Google's data on search terms and their prevalence could help me see whether people became more interested than usual in the brands that advertised during the Superbowl. Here is the link to explore it.

<https://trends.google.com/trends/explore>

This data can also be downloaded in the form of CSVs.

X:

X is a social media site that I could use to better understand people's sentiments about brands during and immediately after the Superbowl. Here is the link to access their API, and another article I found giving some explanation of how to use it.

[How to get access to the X API - X](#)

[X API Guide for Non-Techies: How to Use It If You Don't Code](#)

This API will be a bit more difficult to access, but I should be able to get some usable data (up to 1500 posts) from X.

Youtube:

Youtube views on the ads could help me better assess their reach, because many people who may not watch the Superbowl on TV could be watching the ads on Youtube later on.

[Search: list | YouTube Data API | Google for Developers](#)

Youtube also has an API, and it includes view counts for videos.

The SEC:

The SEC reports data about public companies, which I could use to examine trends in sales.

[SEC.gov | Search Filings](#)

This data would be a little tricky because I would have to do some research to better understand what each report is telling me, but once I know what data I want to download, it can be in the form of CSV which makes it simple.

YouGov

YouGov gathers survey data about every topic imaginable. They have a report about superbowl ad effectiveness.

[Super Bowl 2024 advertising effectiveness report](#)

This report is in the form of a PDF which may be difficult to use

The Harris Poll

This site also has some very interesting reports about the ads. These will be very important because they talk about the *effect* of the ads on different perceptions about the brand before and after

[Brand-Bowl-2024-Final.pdf](#)

This also comes in the form of a pdf sadly

iSpot TV

This site has some data about the ads' likeability, attention index, etc.

[Top 2024 Super Bowl TV Commercials - iSpot.tv](#)

I can use web scraping to get data from the tables on this page

Kellogg School Super Bowl Ad Review:

Every Super Bowl, the Kellogg School of management conducts a poll among its MBA students about how well the ads landed, and grades them from A to D.

[Results | Kellogg Super Bowl Ad Review | Kellogg School of Management](#)

The results are displayed in tables, so I could use web scraping to get the data.

PART THREE: ITERATE

Based on my question "What is the immediate impact of Super Bowl 2024 ad viewership on brand interest and sales in the weeks following the event?" Does my available data directly contain the variable(s) I want to measure?

- Statista gives me general numbers for viewership and ad cost which I can use as context in my report.
- Google trends gives two columns, the date and the “interest” which shows the popularity of the term on that date, relative to its peak.
- X unfortunately might not work for this project, upon further investigation, because the free API doesn’t let you access tweets older than a week. Another potential social media source could be Reddit?
- Youtube should work for me to get viewership, likes/dislikes, and comment data for the different ads, but I definitely have more to learn to better understand APIs. Also, it will be tricky because the original ads aren’t always on the company’s channels anymore for some reason
- SEC.gov is honestly very confusing and not every company I’m interested in reports their data there.
- YouGov if I can use python to scrape a table from a pdf, this will work and will give me data about positive buzz surrounding brands after their ads
- Harris Polls would be a great resource, again if i can learn to scrape data from a pdf. The data talks about the effect of the ads on familiarity, momentum, and other important marketing metrics.
- The Kellogg Ad Review can give me an idea of what marketers and business experts think about the quality of an ad

Does my question need to be altered to match the data I have?

I don’t have data about sales or revenue, so I asked AI to help me adjust my question, and here’s what I came up with with its help:

How did Super Bowl 2024 ads influence short-term brand interest?