

FOR RELEASE FEBRUARY 6, 2024

Most Top-Ranked Podcasts Bring On Guests

Thousands of guests appeared on these podcasts in 2022, but a relatively small number accounted for a majority of appearances.

BY *Galen Stocking and Meltem Odabaş*

FOR MEDIA OR OTHER INQUIRIES:

Galen Stocking, Senior Computational Social Scientist

Aaron Smith, Director, Data Labs

Sogand Afkari, Communications Manager

202.419.4372

www.pewresearch.org

RECOMMENDED CITATION

Pew Research Center, February 2024,
"Most Top-Ranked Podcasts Bring On Guests"

About Pew Research Center

Pew Research Center is a nonpartisan fact tank that informs the public about the issues, attitudes and trends shaping the world. It does not take policy positions. The Center conducts public opinion polling, demographic research, computational social science research and other data driven research. It studies politics and policy; news habits and media; the internet and technology; religion; race and ethnicity; international affairs; social, demographic and economic trends; science; research methodology and data science; and immigration and migration. All of the Center's reports are available at www.pewresearch.org. Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder. This is the latest report in Pew Research Center's ongoing investigation of the state of news, information and journalism in the digital age, a research program funded by The Pew Charitable Trusts, with generous support from the John S. and James L. Knight Foundation.

© Pew Research Center 2024

How we did this

This study takes a close look at the guests hosted by the top-ranked podcasts in the United States in 2022. It is a continuation of work that Pew Research Center has done analyzing podcasts as a part of the news and information landscape in the U.S.

We identified these top-ranked podcasts by first analyzing daily lists of the top 200 podcasts on Apple Podcasts and Spotify from April 1 to Sept. 30, 2022 (chart data provided by [Podchaser Inc](#)). We calculated the average chart position of each podcast that appeared on either list to identify the top 300 podcasts from each site.

After combining daily Apple and Spotify top-300 lists and eliminating duplicates on both lists, we identified 451 top-ranked podcasts. Of these, 434 published at least one episode from Jan. 1 to Dec. 31, 2022. All 24,004 episodes from these podcasts that published in 2022 and were available from Apple and Spotify were analyzed in this report.

To identify the number of guests in each episode and their names, we used OpenAI's Generative Pre-Trained (GPT) model. Researchers gave the model specific instructions to extract the name of any guests in an episode from each of the episode descriptions in the dataset. We then ran accuracy checks for the model results, removed duplicate guests, and performed other standard data cleaning steps.

To learn more, [read the methodology](#) of the report and a [Q&A with the lead researchers](#).

Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder. This is the latest report in Pew Research Center's ongoing investigation of the state of news, information and journalism in the digital age, a research program funded by The Pew Charitable Trusts, with generous support from the John S. and James L. Knight Foundation.

Most Top-Ranked Podcasts Bring On Guests

Thousands of guests appeared on these podcasts in 2022, but a relatively small number accounted for a majority of appearances

By Galen Stocking and Meltem Odabaş

Podcasts are a unique, audio-first source of news and entertainment that has steadily [drawn more listeners since 2013](#), with about [half of Americans saying they have listened to a podcast in the past year](#).

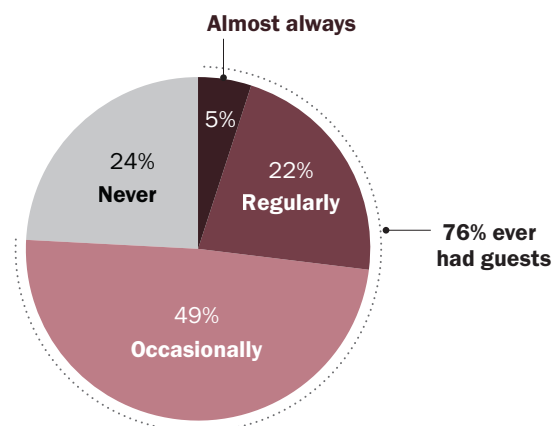
Podcasts are often based around the perspective of the host or hosts, but many podcasts bring on guests to provide additional viewpoints. A new study of 434 [top-ranked podcasts](#) shows just how common this practice is: **76% of top-ranked podcasts brought on at least one guest in 2022**. About one-in-five (22%) featured guests regularly, and 5% of the shows had guests in almost every episode (90% or more of the time).

Among the key findings in this report:

- **Thousands of individuals appear on top-ranked podcasts.** More than 7,000 people appeared on these podcasts in 2022. Because some individuals went on podcasts more than once, these 7,000 individuals made almost 11,000 guest appearances. All told, a guest appeared in about one-third of the approximately 24,000 episodes these podcasts released in 2022.
- **Podcasts on some topics are more likely to have guests, but true crime is a largely guest-free zone.** Sports, entertainment, politics and self-help podcasts are more likely to feature guests than other topics, with at least three-quarters of these podcasts hosting at least one guest in 2022. And over half of entertainment podcasts regularly had guests.

The majority of top-ranked podcasts brought on guests

% of top-ranked podcasts in the U.S. that ___ had guests in 2022



Note: Podcasts that “almost always” had guests hosted a guest in 90% or more of their episodes in 2022; those that “regularly” had guests did so in 50%-89% of their episodes; and those that “occasionally” had guests did so in less than 50% of their episodes. Read methodology for more details.

Source: Pew Research Center analysis of guests on all 2022 episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts.

“Most Top-Ranked Podcasts Bring On Guests”

PEW RESEARCH CENTER

By contrast, true crime is the most common genre among top-ranked podcasts. But about four-in-ten top-ranked true crime podcasts had no guests in 2022, and just 4% of true crime podcasts featured guests regularly.

- **Most guests make only one guest appearance and stick to a single show:** 78% of all guests on these top-ranked podcasts appeared just one time in 2022. And just 12% of guests appeared on multiple podcasts during the year.
- **But guests that made multiple appearances accounted for a significant share of all guest visits.** Just 22% of guests made more than one appearance on a top-ranked podcast in 2022, but these individuals made up half of all guest appearances.
- **Individuals affiliated with The Daily Wire are especially prolific podcast guests and often appear on each other's shows.** Five of the 18 people with the most guest appearances on top-ranked podcasts in 2022 were directly affiliated with the conservative news outlet. And the bulk of their appearances were on Daily Wire-produced shows.
- **Many of the people who appeared on the largest number of top-ranked podcasts went on The Joe Rogan Experience.** This podcast was a popular destination for people who frequently appeared on top-ranked podcasts in 2022: 44% of these guests went on Rogan's show that year.

Guests on these podcasts were identified using OpenAI's Generative Pre-Trained Transformer (GPT) 3.5 model. Researchers gave the model episode descriptions downloaded from Apple Podcasts and Spotify and instructed the model to identify all guests on each episode. For more information, [read the methodology](#) and a [Q&A with the lead researchers](#) on the report.

Keep reading or jump down to see:

- [Which podcasts bring on guests](#)
- [Which guests appear most often](#)
- [A network analysis of podcasts and their guests](#)

[BOX Rise of podcasts]

The rise of podcasts

Podcast listenership in the United States has [increased steadily since 2013](#). And a recent Pew Research Center survey found that [49% of U.S. adults have listened to a podcast in the past year](#).

This report is the latest in the Center's ongoing research into podcasts. Previous reports covered [how Americans engage with podcasts](#); [which podcasts they listen to most](#); how [differences in engagement with podcasts](#) varies based on factors like age and race; and an [in-depth analysis of top-ranked podcasts](#).

[/BOX]

Which podcasts bring on guests

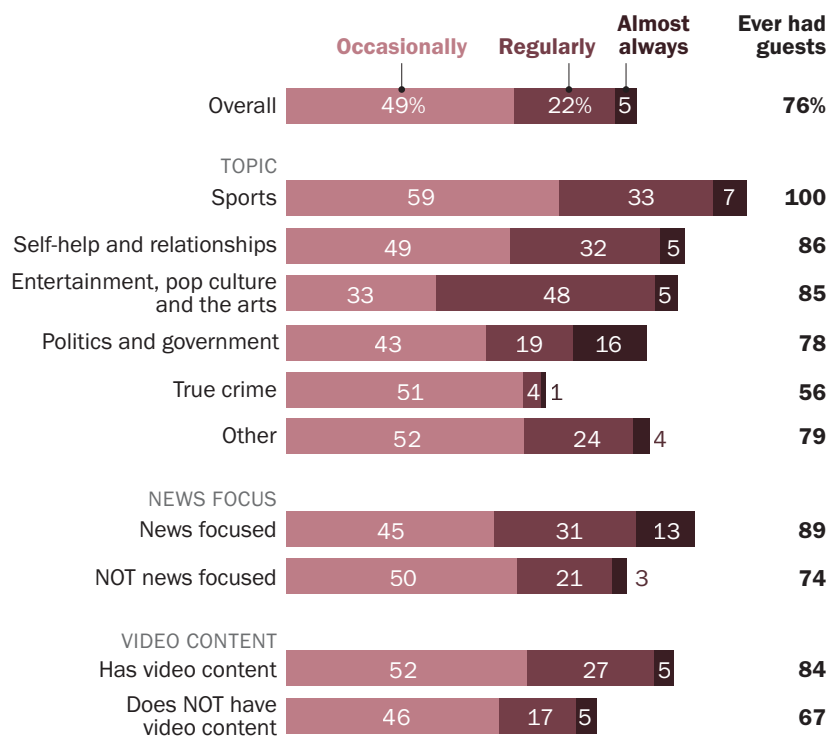
Top-ranked podcasts [cover a wide range of subjects](#). But podcasts about some topics were more likely to bring guests on than others in 2022. Notably, every top sports podcast had at least one episode with a guest. A majority of self-help (86%), entertainment (85%) and politics (78%) podcasts did so as well.

True crime podcasts were the least likely major genre to bring on guests: Slightly more than half (56%) brought a guest on in at least one episode.

Majorities of podcasts about each of these topics brought on at least one guest. But far fewer brought on a guest

16% of top-ranked podcasts about politics almost always had guests in 2022, more than any other topic

% of top-ranked podcasts in the U.S. that ___ had guests in 2022



Note: Podcasts that "almost always" had guests are those that hosted a guest in 90% or more of their episodes in 2022; podcasts that "regularly" had guests featured a guest on 50%-89% of their episodes; and podcasts that "occasionally" had a guest had at least one guest during the year but did not have a guest in at least 50% of their episodes. Read methodology for more details.

Source: Pew Research Center analysis of guests on all 2022 episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts.

"Most Top-Ranked Podcasts Bring On Guests"

PEW RESEARCH CENTER

regularly (in 50% to 89% of their episodes). Roughly half of entertainment podcasts (48%) hosted guests regularly, as did about one-in-three sports (33%) and self-help (32%) podcasts. In contrast, just 4% of true crime podcasts featured guests regularly.

A small share of podcasts (5%) almost always had a guest in their episodes in 2022 – that is, a guest appeared in 90% or more episodes. This was most common among politics and government podcasts, 16% of which had a guest on nearly every episode. No more than 10% of podcasts in all other categories had a guest this frequently.

News-focused podcasts

A modest share (15%) of top-ranked podcasts [focused on news of any topic](#) in 2022. But the vast majority (89%) of these news-focused podcasts hosted at least one guest during the year, a higher share than podcasts that were not news-focused.

Many news-focused podcasts frequently brought on guests. About a third (31%) did so regularly, and 13% *almost always* brought on guests – far higher than the share of podcasts that do not focus on news (3%).

Repeat guests

It was relatively common for top-ranked podcasts to bring on the same guests more than once in 2022. About half (49%) of all top-ranked podcasts had at least one of these *repeat guests* over the year. And around a quarter of these podcasts (23%) had more than five repeat guests that year.

At the same time, most podcasts didn't rely especially heavily on these repeat guests. Fewer than 1% of *all guest appearances* came from repeat guests on a typical podcast that ever had guests in 2022.

But podcasts that had guests more frequently tended to bring on the same guests more often. A typical podcast that regularly had guests brought on a repeat guest for 7% of all guest appearances. That share rose to 10% of guest appearances for podcasts that almost always had guests.

Which podcast guests appear most often

While podcasts [often brought on repeat guests in 2022](#), most guests did not appear very often: 78% of all guests appeared on just one episode of a top-ranked podcast during the year.

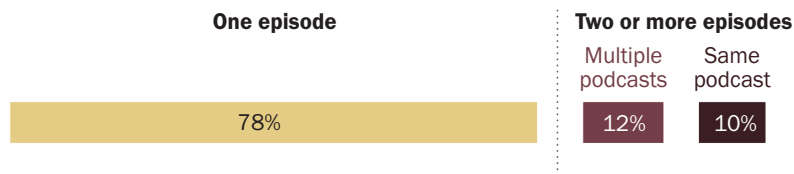
The 22% of guests who appeared on two or more episodes tended to be extremely prolific. Appearances from this minority of guests accounted for 49% of all guest appearances on top-ranked podcasts in 2022.

This 22% of guests who made multiple podcast appearances in 2022 is made up of:

- **12% of guests who appeared on *multiple podcasts*** – Together, this 12% of guests accounted for 28% of all guest appearances on top-ranked podcasts in 2022.
- **10% who appeared on a *single podcast multiple times*** – Together, this 10% of guests made up 21% of all guest appearances on top-ranked podcasts in 2022. These guests were often regulars for the podcasts they went on frequently: Many of the guests with the largest number of appearances on a single podcast were in 5% or more of that show’s episodes.

Most guests only made one appearance on any top-ranked podcast in 2022

% of guests who appeared on ____ of a top-ranked podcast in 2022



Note: N=10,996 guest appearances of N=7,212 distinct guests. Read methodology for more details.

Source: Pew Research Center analysis of guests on all 2022 episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts.

“Most Top-Ranked Podcasts Bring On Guests”

PEW RESEARCH CENTER

Guests who appeared on the most shows in 2022

The preceding “bird’s eye view” analysis looked at how often the same guests were appearing across podcasts or even within the same podcasts. In this section, we look at the *specific* guests who appeared on the highest number of podcasts and episodes.

As noted above, 12% of all guests appeared on more than one show in 2022. Sixteen of these individuals went on six or more top-ranked podcasts over the course of the year, accruing nearly 109 guest appearances between them.

The guests who went on the largest number of shows come from a diverse set of professional backgrounds, ranging from company founder and CEO Elon Musk to psychologist, author and commentator Jordan Peterson. Others in this list include:

- **Figures from the entertainment industry**, such as Andrew Schulz, Ari Shaffir, Bert Kreischer and Judd Apatow.
- **Scientists** like Neil DeGrasse Tyson or Anthony Fauci.
- **Journalists and political commentators** from across the media landscape, ranging from New York Times columnist Maggie Haberman to independent journalist Matt Taibbi.

Many of these guests made appearances on the same podcasts. Most notably, 14 of these 16 individuals appeared on The Joe Rogan Experience at one time or another during the year. Other

The most frequently appearing guests on top-ranked podcasts in 2022

Most frequently appearing guests on top-ranked podcasts in 2022 by the ...

Number of shows they appeared on		Number of episodes they appeared on	
Andrew Schultz	10	Ryen Rusillo	43
Ari Shaffir	9	Michael Knowles	34
Neil DeGrasse Tyson	8	Andrew Klavan	33
Elon Musk	7	Jeremy Boreing	31
Judd Apatow	7	Dane Brugler	28
Maggie Haberman	7	Matt Walsh	28
Matt Walsh	7	Robert Mays	26
Bert Kreischer	6	Peter Schragar	25
Dave Rubin	6	“Cousin Sal” Iacono	23
Anthony Fauci	6	Nate Tice	22
Jordan Peterson	6	Benjamin Solak	21
Josh Peck	6	Liz Loza	21
Matt Taibbi	6	Vincent Oshana	21
Susan Cain	6	Ben Shapiro	20
Yannis Pappis	6	Eric Moody	20
James O’Keefe	6	Mike Sando	20
		Rob Mahoney	20
		Kevin O’Connor	20

Note: N=434 top-ranked podcasts. These podcasts featured N=10,996 guest appearances of N=7,212 distinct guests. Table includes guests who appeared on six or more shows, or 20 or more individual episodes.

Source: Pew Research Center analysis of guests on all 2022 episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts.

“Most Top-Ranked Podcasts Bring On Guests”

PEW RESEARCH CENTER

shows that brought a relatively large number of these guests include The Megyn Kelly Show, Candace Owens, and the Are You Garbage? podcast.

Guests who appeared in the most episodes in 2022

Eighteen individuals made a total of 20 or more guest appearances on a top-ranked podcast in 2022, and these guests tended to come from one of two broad groupings.

Twelve of these 18 guests are sports journalists or commentators. Many are affiliated with The Ringer or The Athletic. And **five of these 18 guests are directly affiliated with the conservative outlet The Daily Wire.**

These five guests include Michael Knowles, CEO Jeremy Boreing, Matt Walsh, Andrew Klavan and site co-founder Ben Shapiro. All but Boreing host a podcast published by The Daily Wire. And all of them tend to appear on podcasts produced by that outlet. Some 90% of their nearly 150 guest appearances on top-ranked podcasts in 2022 were on podcasts produced by The Daily Wire.

Deep dive: Classifying podcasts around the guests they share


As noted above, 52% of top-ranked podcasts – 225 in total – shared at least one guest with another top-ranked podcast in 2022. We wanted to see if we could find any patterns in these guest appearances and identify clusters of podcasts that tend to share similar guests.


To do this, we visualized these podcasts and their guests as a network map. We then used a technique known as community detection to put them into groups based on the guests they bring on, how often they have those people as guests, and the other podcasts those guests make appearances on.


Here is an overview of how that process worked:

How we mapped guest appearances across multiple top-ranked podcasts in 2022

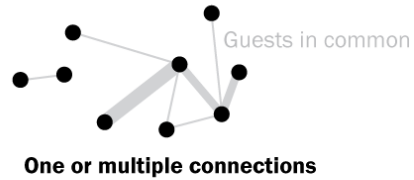
1 Our network diagram of podcast guests is made up of **circles** and **lines**.

 A single podcast is represented as a **circle**. If this was a map of just one podcast, with no shared guests, the map would be just a single circle on a blank canvas.

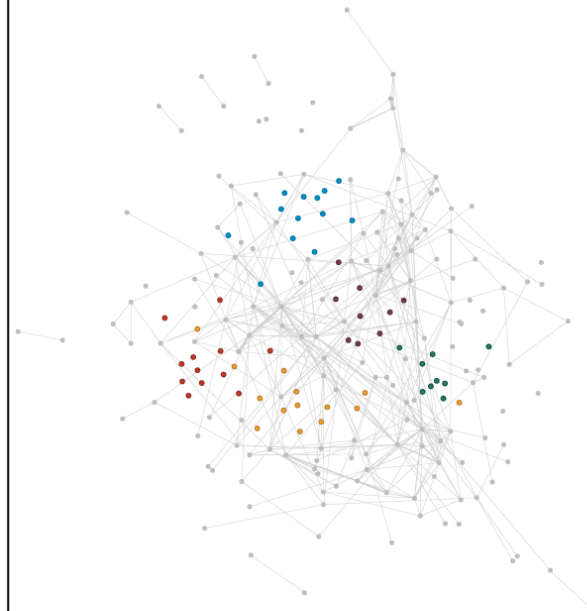
 When the map is expanded to add a second podcast, a second circle is added to that canvas. If those two podcasts had any guests in common in 2022, a **line** would connect the two circles.


 If two podcasts share more guests, we show this by increasing the **thickness** of the line between them. How close or far the podcasts appear on the map also reflects how many guests they share.

2 As more podcasts are added, the number of connections increases. And, the more guests the podcasts share, the thicker the lines.

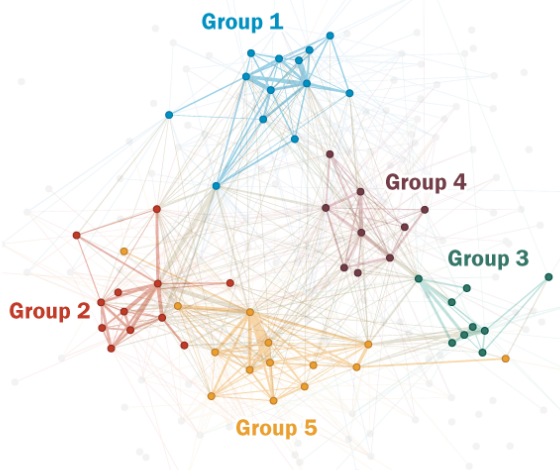


3 This is what the **full dataset of all 434 top-ranked podcasts** looks like. The 225 podcasts that shared guests with at least one other podcast have a large number of connections between them. The ones that didn't share any guests are isolated dots with no connections to the others.



 209 podcasts (48%) in our dataset had **no guests in common** with any other podcasts.

4 Here, we've **zoomed in** and faded out all but a few clusters of podcasts and their connections, so we can see them more clearly. **57** podcasts belong to one of these clusters, based on the number of guests they share.

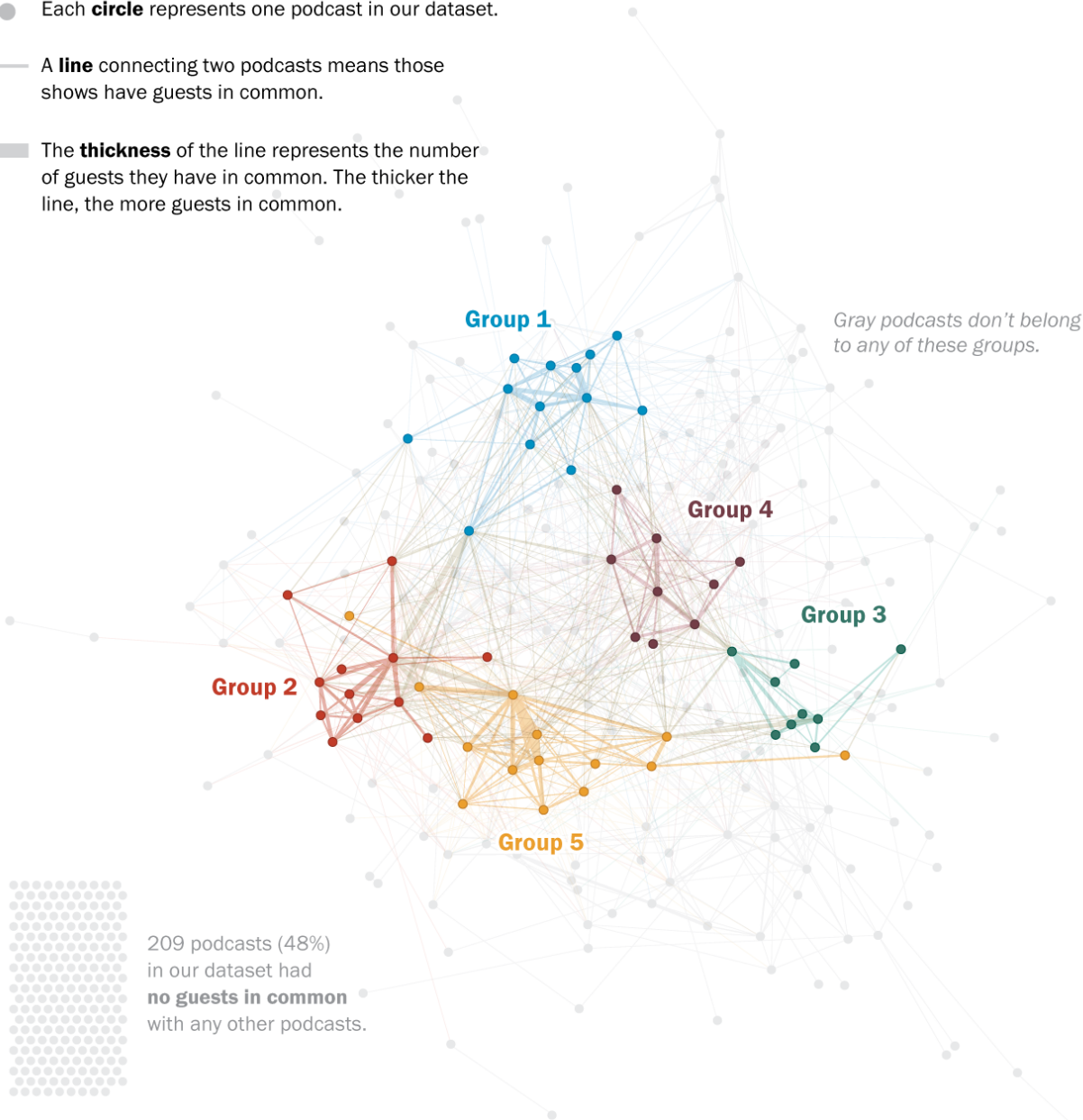


Source: Pew Research Center analysis of guests on all 2022 episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts. "Most Top-Ranked Podcasts Bring On Guests"

Examples of podcast ‘communities’ based on shared guests

Clusters of top-ranked podcasts that shared similar guests in 2022

- Each **circle** represents one podcast in our dataset.
- A **line** connecting two podcasts means those shows have guests in common.
- The **thickness** of the line represents the number of guests they have in common. The thicker the line, the more guests in common.



Note: Groups identified using community detection analysis.

Source: Pew Research Center analysis of guests on all episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts released in 2022. Network analysis includes N=225 podcasts that share at least one guest with another podcast.

"Most Top-Ranked Podcasts Bring On Guests"

PEW RESEARCH CENTER

When we ran our network of podcasts and guest appearances through the community detection algorithm, we found five distinct groups containing between nine and 14 shows each. Refer to the appendix for a table with [the full list of shows in each group](#).

All told, 57 podcasts were included in one of these five groups. The other podcasts on our map were not included in any groups. That is because they have fewer connections with other shows based on the guests they share.

Some of these groups – like Groups 3 and 4 – are close together on the map, which means that several guests made appearances on shows in each group. Others – like Groups 1 and 2 – are far apart, meaning there is less similarity in the guests they bring on, but still enough to be grouped together.

Beyond the guests they bring on, the podcasts that were assigned to each of the different groups were sometimes tied together by clear and consistent themes in the topics they cover or the people who host those shows. But in other cases, those similarities were not as obvious.

Here is a brief description of the shows in each group.

Group 1 (*news, political commentary and current events, often with a left-of-center perspective*)

All the podcasts in Group 1 are about news, current events and politics. These shows often – but not always – have a liberal or left-of-center viewpoint.

Although these shows tend to come from the political left, their guests can be drawn from other parts of the political spectrum. For instance, the community detection analysis placed the Real Time with Bill Maher podcast in this group: His show shared 17 guests with eight other podcasts in this group over the course of the year. But it also shared 23 guests with six podcasts in Group 2, many of whose shows tend to have a more conservative orientation.

Group 2 (*news, political commentary and current events, often with a right-of-center perspective*)

Group 2 also contains several shows about politics, news and current events. Many of these shows approach these topics from a conservative or right-of-center perspective. Several are affiliated with The Daily Wire.

The group also includes shows from journalists and media personalities who have been affiliated with Fox News, such as Megyn Kelly and Dan Bongino. Like Maher, Kelly's show is located toward

the midpoints of Groups 1 and 2 on our network map. Her show shared 60 guests with 10 podcasts in Group 2, but also 14 guests with five podcasts in Group 1 and 28 guests with eight podcasts in Group 5.

Group 2 also includes a few podcasts that touch on other topics like true crime or self-help and personal improvement.

Group 3 (*sports-centric commentary and analysis*)

The podcasts in this group tend to focus on sports news, commentary and analysis. Group 3 also contains two shows that focus on the broader world of pop culture but are affiliated with the Ringer Podcast Network.

Group 4 (*mostly entertainment, some sports*)

This group contains several shows hosted by actors or other figures from the world of entertainment and pop culture. It also includes interview podcasts hosted by sports personality Bill Simmons and NBA player Draymond Green, as well as NPR's Fresh Air podcast.

Group 5 (*a wide-ranging mash-up of comedy, sports, self-help, political commentary and more*)

This group is the most wide-ranging of the five identified here. Several of the shows in this group are hosted by figures from the world of stand-up comedy. But Group 5 also includes two sports shows on the Barstool Sports network and two shows that focus on topics related to self-help, life skills and financial advice.

It also includes several shows in which figures with ties to the world of comedy and entertainment discuss a wide-ranging set of issues, from pop culture to politics and current events. The most prominent of these is The Joe Rogan Experience, whose guests have a substantial amount of overlap with Group 2.

Appendix

Podcast groupings identified via network analysis

Group 1	Group 2	Group 3
<p>Group 1</p> <p>Fast Politics with Molly Jong-Fast</p> <p>In the Bubble with Andy Slavitt</p> <p>Kasich & Klepper</p> <p>On with Kara Swisher</p> <p>Pod Save America</p> <p>Real Time with Bill Maher</p> <p>Strict Scrutiny</p> <p>The Daily Beans</p> <p>The Ezra Klein Show</p> <p>The Journal</p> <p>The Problem with Jon Stewart</p> <p>The Wilderness</p> <p>V Interesting with V Spehar</p> <p>Group 4</p> <p>Armchair Expert</p> <p>Conan O'Brien Needs a Friend</p> <p>Fly on the Wall with Dana Carvey and David Spade</p> <p>Fresh Air</p> <p>Office Ladies</p> <p>Significant Others</p> <p>Smartless</p> <p>The Bill Simmons Podcast: The Interviews</p> <p>The Draymond Green Show</p>	<p>Group 2</p> <p>All-In Podcast</p> <p>Breaking Points with Krystal and Saagar</p> <p>Candace Owens</p> <p>Law&Crime Sidebar</p> <p>The Aubrey Marcus Podcast</p> <p>The Ben Shapiro Show</p> <p>The Dan Bongino Show</p> <p>The Glenn Beck Program</p> <p>The Jordan B. Peterson Podcast</p> <p>The Matt Walsh Show</p> <p>The Megyn Kelly Show</p> <p>The Michael Knowles Show</p> <p>Group 5</p> <p>2 Bears, 1 Cave with Tom Segura & Bert Kreischer</p> <p>Andrew Schulz's Flagrant with Akaash Singh</p> <p>Are You Garbage? Comedy Podcast</p> <p>Bussin' with the Boys</p> <p>Jocko Podcast</p> <p>Monday Morning Podcast</p> <p>Pardon My Take</p> <p>PBD Podcast</p> <p>Take Your Shoes off w/ Rick Glassman</p> <p>The Adam Carolla Show</p> <p>The Fantasy Footballers</p> <p>The Joe Rogan Experience</p> <p>The Tim Dillon Show</p> <p>Your Mom's House with Christina P. and Tom Segura</p>	<p>Group 3</p> <p>Fantasy Football Happy Hour with Matthew Berry Podcast</p> <p>Fantasy Football Podcast</p> <p>GOJO with Mike Golic Jr.</p> <p>Off the Pike with Brian Barrett</p> <p>Slow News Day with Kevin Clark</p> <p>The Athletic Football Show: A show about the NFL</p> <p>The Bill Simmons Podcast</p> <p>The Hottest Take</p> <p>The Ringer-Verse</p> <p>Group descriptions:</p> <p>Group 1: News, political commentary and current events, often with a left-of-center perspective</p> <p>Group 2: News, political commentary and current events, often with a right-of-center perspective</p> <p>Group 3: Sports-centric commentary and analysis</p> <p>Group 4: Mostly entertainment, some sports</p> <p>Group 5: Wide-ranging mash-up of comedy, sports, self-help, political commentary and more</p>

Note: Groups identified using community detection analysis.

Source: Pew Research Center analysis of guests on all episodes of 434 top-ranked podcasts on Spotify and Apple Podcasts released in 2022. "Most Top-Ranked Podcasts Bring On Guests"

PEW RESEARCH CENTER

Acknowledgments

Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder. This is the latest report in Pew Research Center’s ongoing investigation of the state of news, information and journalism in the digital age, a research program funded by The Pew Charitable Trusts, with generous support from the John S. and James L. Knight Foundation. This report is a collaborative effort based on the input and analysis of the Center’s data labs, news and information research team, communications, design, digital and editorial teams. Find related reports online at journalism.org.

Research Team

Galen Stocking, *Senior Computational Social Scientist*

Aaron Smith, *Director, Data Labs*

Meltem Odabaş, *Computational Social Scientist*

Samuel Bestvater, *Computational Social Scientist*

Katerina Eva Matsa, *Director, News & Information*

Jacob Liedke, *Research Analyst*

Emily Tomasik, *Research Assistant*

Christopher St. Aubin, *Research Assistant*

Communications and Web Publishing

Sogand Afkari, *Communications Manager*

Andrew Grant, *Communications Associate*

Sara Atske, *Digital Producer*

Editorial and Graphic Design

David Kent, *Senior Copy Editor*

Alissa Scheller, *Senior Information Graphics Designer*

Andrea Caumont, *Associate Director, Digital Outreach*

Methodology

This study takes a close look at the guests who appeared on top-ranked podcasts in the United States in 2022. The methodology for how the 451 top-ranked podcasts and their content are identified, as well as how their episode data were collected, is [available here](#). The vast majority of these podcasts (434 of 451) published at least one episode in 2022, with a total of 24,004 episodes. Roughly a third of these episodes (31%, or 7,423) included a guest, with a total of 7,212 guests appearing 10,996 times.

This is the latest report in Pew Research Center’s ongoing investigation of the state of news, information and journalism in the digital age, a research program funded by The Pew Charitable Trusts, with generous support from the John S. and James L. Knight Foundation.

Guest name extraction

All guests were identified through an analysis of published descriptions of podcast episodes. Those episode descriptions were collected from the Apple Podcasts and Spotify APIs in summer 2023.

There were 24,004 episodes collected from 434 podcasts included in the original study. The 17 podcasts that were not included either did not publish any episodes in 2022 or had removed episodes from those services when researchers collected episode descriptions.

Podcast episode descriptions are useful windows into the episode, as they often mention their guests by name. However, the language they use poses several challenges for an automated analysis.

First, podcast producers do not talk about their guests in consistent ways. For example, one description might introduce the guest with a phrase like “we are joined by [GUEST NAME]” while another might say “today [GUEST NAME] talks about...”

Second, even when the description mentions an individual by name, that person may not have been a guest. Instead, the podcast may have *talked about* this individual.

To overcome these difficulties, researchers used OpenAI’s Generative Pre-Trained Transformer model (GPT 3.5) to extract guest names from episode descriptions.

Using the three-part prompt structure below, we passed each episode description in turn to the gpt-3.5-turbo-0613 model endpoint:

Prompt for extracting guest names

Prompt description	Prompt
General instructions	You are an AI assistant that helps extract information from text. You will be given a set of podcast episode descriptions. Tell me if there are any guests featured on this episode. In a JSON titled 'guests', for each guest, tell me their name, a description of who they are, and the topic they are brought on to discuss. Return only guests, do not include the name of the regular podcast hosts or people who are mentioned, but do not appear in the episode. Do not summarize when performing this task, but draw only from the information included in the description, using the original language, and only return the JSON. If there are no guests, or there is not enough information to tell, just return '[]'
Specific context information	Episode description: <episode_description>
Notes and reminders	NOTE: Don't justify your answers. Don't give information not mentioned in the CONTEXT INFORMATION. Only return a valid JSON response.

PEW RESEARCH CENTER

While GPT processed the vast majority of these episode descriptions without difficulty, the built-in content moderation guardrails in the GPT endpoint refused to give us any data for episode descriptions it determined were about sex or violence. After some testing, researchers used a different endpoint (the commercial OpenAI endpoint) that was less sensitive to those issues. Overall, 966 episodes, or about 4%, could not be evaluated because of the content moderation guardrails.

Validation of GPT guest name results

Researchers then validated the output from GPT using a pair of trained human coders, who validated the output independently. There were two measures evaluated: *Has guests*, or whether an episode description included a guest(s) or not; and *Guest names*, or the list of guests. In other words, researchers were verifying that GPT didn't miss any guests and that the guests it identified were correct.

These human coders were trained to identify guests in episode descriptions. To confirm that coders were interpreting this task the same way, we measured intercoder agreement and Cohen's kappa across both variables:

- *Has guests*: 97% agreement; 0.91 kappa
- *Guest names*: 96% agreement; 0.76 kappa

After training, the human coders evaluated a subset of 9% of the GPT output:

GPT statistical performance scores

Measure	Measure description	Score
<i>Has guests</i>		
Accuracy rate	How often the model correctly identified the guest count	0.90
Precision	How often the model was correct when it predicted that the episode description had at least one guest name	0.83
Recall	How many of the actual episodes with at least one guest name the model was able to identify	0.88
<i>Guest names</i>		
Accuracy rate	How often the model was correct listing all guest names	0.90

Note: *Guest names* variable's statistical performance is measured by accuracy rate only, because the prediction results for this variable are not binary.

PEW RESEARCH CENTER

Data cleaning

As the metrics show, GPT did an acceptable job extracting the raw names of guest descriptions. But there were still several cleaning steps that needed to be performed:

- First, a small set (6%) of identified names only had a single name – which in many cases was not enough information to identify the guest.
- Second, the same identified name can refer to two different people – multiple people might be named John Smith, for instance.
- Finally, the same person may appear under different names – such as a person who goes by Elizabeth Smith and Lizzy Smith.

To identify any guests that needed this cleaning, researchers first removed any guests who were clearly a host or co-host on the podcast, and then hand-coded 966 guest names that fit those criteria. The final list included 7,212 guests.

Network analysis

To provide insight into the kinds of podcasts certain guests appeared on, researchers used network analysis methods that show guest sharing patterns across podcasts. About half (52%) of top-ranked podcasts (N=225) shared at least one guest with at least one other podcast.

These podcasts were analyzed using the walktrap community detection algorithm in [R's igraph package](#). To identify guest-sharing patterns for podcasts that commonly share guests, only podcasts that shared at least three guests with at least 12% overlap were included. This analysis uncovered five distinct groups (refer to the [Appendix](#) for a list of the podcasts in each group).

To illustrate the results of this network analysis, researchers used the network analysis tool [Gephi](#) to visualize a network map. In this map, each podcast is represented with a node, and the ties between podcasts indicate whether those podcasts share at least one guest. The thickness of the lines reflect the strength of these ties.