kaggle   **Competitions**   **Datasets**   **Scripts**   **Forums**   **Jobs**

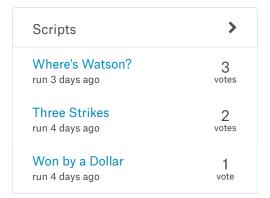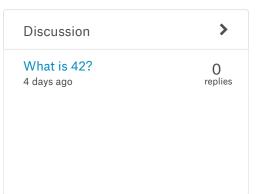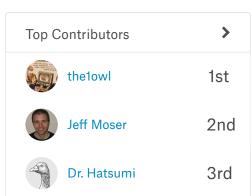# Jeopardy!

▲
9

Explore the popular game show's clues, responses, and contestants

by **Jerad Rose** · last updated 5 days ago

Scripts ❯		Discussion ❯		Top Contributors ❯	
Where's Watson? run 3 days ago	**3** votes	**What is 42?** 4 days ago	**0** replies	**the1owl**	**1st**
Three Strikes run 4 days ago	**2** votes			**Jeff Moser**	**2nd**
Won by a Dollar run 4 days ago	**1** vote			**Dr. Hatsumi**	**3rd**

Description

I'll take "Jeopardy Categories" for $200, Alex.

One of the more popular categories, particularly since it's introduction in 1997, it is known for combining two clues into one.

Find the answer to this clue, and discover many other interesting tidbits about the Jeopardy! game show by working your way through the data spanning the entire series*, starting with the first episode on September 10, 1984.

This data from j-archive.com includes data for each game, including contestant occupation & location, contestant scores at various checkpoints, correct and incorrect responses, and of course all clue and answer data.

Can you figure out:

- Who had the best comebacks?
- What categories are the easiest and hardest?
- Who ended the game with the biggest loss?
- Have contestants improved at categories over time?
- What locations breed the best and worst contestants?

- Do occupations play a factor in winnings?

Description   Scripts   Forum   Download Data (43.81 MB)   New Notebook   **New Script**

Note that this data was parsed from j-archive.com, which is a Jeopardy! fan site, and the data is crowdsourced by its members. This means that the data is not complete, particularly prior to 1997. If you'd like to help make this data more complete, break out your stash of VHS tapes (we know you recorded every episode!) and head over to j-archive.com to help complete their database of episodes. We will be periodically updating our dataset as new episodes are added to j-archive.com.

Unknown License

Files

categories.csv (936.54 KB)

Lookup table of all categories.

- **category_id** (integer) - PK
- **category** (text) - category name

clue_wrong_answers.csv.zip (528.72 KB)

Any wrong answers given for clues.

- **clue_id** (integer) - FK to *clues*
- **player_id** (integer) - FK to *players*
- **answer** (text) - the incorrect answer given by the player

clues.csv.zip (17.73 MB)

The clues and metadata for each game.

- **clue_id** (integer) - PK
- **game_id** (integer) - FK to *games*
- **round** (integer) - the round for the clue:
  - 1 = Jeopard!
  - 2 = Double Jeopardy!
  - 3 = Final Jeopardy!
- **column** (integer) - the column position (1-6) of the clue on the game board*
- **row** (integer) - the row position (1-5) of the clue on the game board*
- **value** (integer) - the dollar amount of the clue; for Daily Double clues, this is the wager amount*
- **is_daily_double** (integer) - 1 if the clue is a daily double; 0 if not
- **category_id** (integer) - FK to *categories*; the category of the clue
- **order_number** (integer) - the chronological order the clue appeared in its round*
- **clue** (text) - the clue
- **answer** (text) - the correct answer
- **answer_player_id** (integer) - FK to *players*; the player that gave the right answer (if any)*

*These fields will be *null* for Final Jeopardy! clues. Final Jeopardy! data is found in *final_jeopary_answers*.

database.sqlite.zip (24.46 MB)

SQLite database containing tables with identical data and structure to all the CSV files described in this data release

final_jeopardy_answers.csv (601.09 KB)

The responses to the Final Jeopardy! rounds.

- **clue_id** (integer) - FK to *clues*
- **player_id** (integer) - FK to *players*
- **answer** (text) - the answer given by the player
- **wager** (integer) - the amount wagered
- **is_correct** (integer) - 1 if the answer was correct; 0 if not

game_players.csv (760.45 KB)

The players and game player metadata for each game.

- **game_id** (integer) - FK to *games*
- **player_id** (integer) - FK to *players*
- **place** (integer) - where the player placed at the end of the game
- **first_break_score (integer)** - the amount the player had at the first commercial break
- **first_round_score** - the amount the player had at the end of the Jeopardy! round
- **second_round_score** - the amount the player had at the end of the Double Jeopardy! round
- **final_score** - the amount the player had at the end of the game
- **coryat_score** - the player's final score if all wagering is disregarded (more info)

games.csv (246.32 KB)

The list of games and metadata.

- **game_id** (integer) - PK (also corresponds to *game_id* on j-archive.com/showgame.php)
- **air_number** (integer) - the chronological position this game aired in the series
- **air_date** (text) - the date the game was originally aired
- **game_data_complete** (integer) - 1 if the game appears to have all player, clue, and score data; 0 if not
- **notes** (text) - any notes made about the game on j-archive.com; this may contain info about tournaments or other special game notes

players.csv (716.08 KB)

The list of all contestants and metadata.

- **player_id** (integer) - PK (also corresponds to *player_id* on j-archive.com/showplayer.php)
- **name** (text) - the full name of the player as mentioned on the airing
- **nickname** (text) - the short name as written by the player on their screen
- **occupation** (text) - the occupation of the player as mentioned on the airing
- **location** (text) - the location of the player as mentioned on the airing*
- **is_originally** (integer) - 1 if the mention of the player's location if preceded with "originally from"; 0 if just "from"

*In some special cases (i.e. celebrities, college students), *location* may not represent a geographical location, but rather a movie or college.

Versions

1,875 views · 90 downloads · 12 scripts · 1 topic

Explore More Datasets

Our Team   Careers   Terms   Privacy   Contact/Support