

Did James Holzhauer Change the Game of Jeopardy! Forever?

December 2025

Contents

Introduction	1
Results	2
Conclusion	8
Appendix	8

Introduction

For this final project I have analyzed a dataset from the TV game show Jeopardy! Jeopardy is a trivia game show where three contestants answer clues each with a dollar value assigned to them. Players buzz in and earn the corresponding dollar amount for correctly answering the clue, but the amount is deducted from their totals if they answer incorrectly. There are three rounds: Single Jeopardy has six categories each with five clues ranging from \$200 - \$1000; Double Jeopardy has the same number of categories and clues but values range from \$400 - \$2000; lastly there is Final Jeopardy where players all get the same clue and can wager up to their accumulated score. Importantly, there is also three daily doubles (one in Single Jeopardy and two in Double Jeopardy) where players can wager the maximum clue value on the board, or up to the amount they have accumulated already. Whoever has the highest score at the end of the three rounds gets to keep the money and is invited back to the next game as a returning champion. Read more about the show here: [<https://en.wikipedia.org/wiki/Jeopardy!>].

Anyone who is as much of a fan of the game as I am is familiar with the name James Holzhauer. When he entered the stage on April 4th, 2019, he was introduced simply as a professional gambler from Las Vegas. What would follow was one of the most astonishing feats the show has ever seen. During his 32 game winning streak, he accumulated \$2,462,216 in winnings, second only to Ken Jennings who took 74 games to win just over \$2,500,000. Being able to amass this amount of cash so fast was unprecedented, and James currently holds the top 12 spots for highest single game winnings, his top total being \$131,127 while second place (not including himself) is \$83,000 won by Matt Amodio, another super-champion. James owes his high winnings to his aggressive play style, using a technique known as the Forrest Bounce [https://en.wikipedia.org/wiki/Chuck_Forest], where he would hunt the high value clues, find daily doubles, and frequently bet all of his winnings. At the time this was uncommon, but since his appearance, this seems to be the normal style of play. Simply from my viewing perspective, it seems like James changed the game for good. So my question is simple, did James really change the game as much as it seems? To address this question statistically I will be comparing the average winning total before his appearance on the show to the average winning total after he appeared.

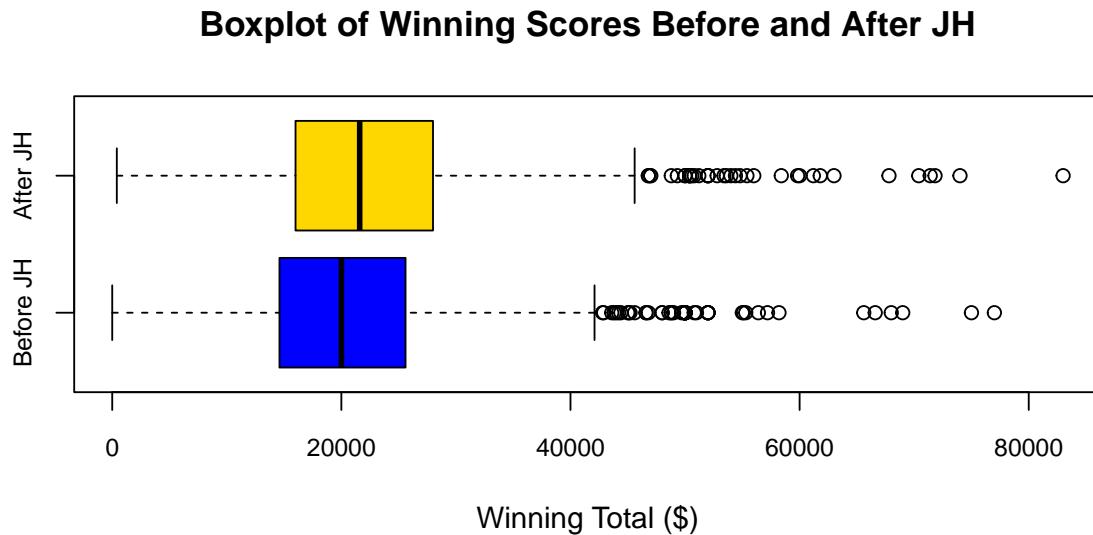
To analyze this data I used parametric and permutation methods to create null distributions and see if there was a significant difference between the mean winning total before and after James appeared on the show. I also created a bootstrap distribution and used parametric methods to create 95% confidence intervals. Finally, it is important to note that I am only considering games played on or after November 26, 2001 since that is when Jeopardy doubled the clue values and any games played before that would have lower winning totals due to the lower clue values.

The data from this analysis comes from a publicly available file on GitHub and contains data from every regular season game ever played up through season 41 (2024-2025) [https://github.com/jwolle1/jeopardy_clue_dataset/commit/2d6116506dfa74eb3bf750a9b039431bbf7b7484]. The data was originally scraped from a site called J! Archive that documents every game ever played [<https://j-archive.com/>]. It contains lots of variables including number of correct and incorrect answers, scores after each round, and most importantly for this project, final scores for each player. Lots of analyses have been done on the game, but I have not found another analysis looking at this exact question. The code for loading this data and assigning initial vectors can be found in the appendix.

Results

Data Visualization: Boxplot of Winning Totals Before and After JH Appearance

```
boxplot(winner_before, winner_after, horizontal = TRUE,
        names = c("Before JH", "After JH"), xlab = "Winning Total ($)",
        main = "Boxplot of Winning Scores Before and After JH",
        col = c("Blue", "Gold"), cex.axis = 0.8)
```



This boxplot shows the winning totals of games before and after James appeared on the show. This helps give a rough visualization of whether or not the median score has shifted. The boxplot is a good tool for visualizing this data because it helps separate out large outliers. Based on the boxplot

alone, it does seem that the average winning score has gone up slightly since James appeared. Both distributions both appear to be relatively normally distributed.

Hypothesis Tests: Permutation shuffle test and T-test for difference of means

Question: Have the average Jeopardy winnings increased since James Holzhauer appeared on the show?

Null Hypothesis: There is no difference in average winnings between before James appeared and after he appeared.

Alternative Hypothesis: The average winning total after James appeared is higher than the average winning total before he appeared.

$$H_0 : \mu_a = \mu_b$$

$$H_A : \mu_a > \mu_b$$

(a = after, b = before)

Alpha level: $\alpha = 0.01$

Permutation Hypothesis Test

```
# Calculate observed statistic
obs_stat <- mean(winner_after) - mean(winner_before)

# Create and visualize null distribution using permutation
# shuffling methods
comb_group <- c(winner_after, winner_before)

null_dist <- do_it(10000) * {
  shuff_group <- shuffle(comb_group)

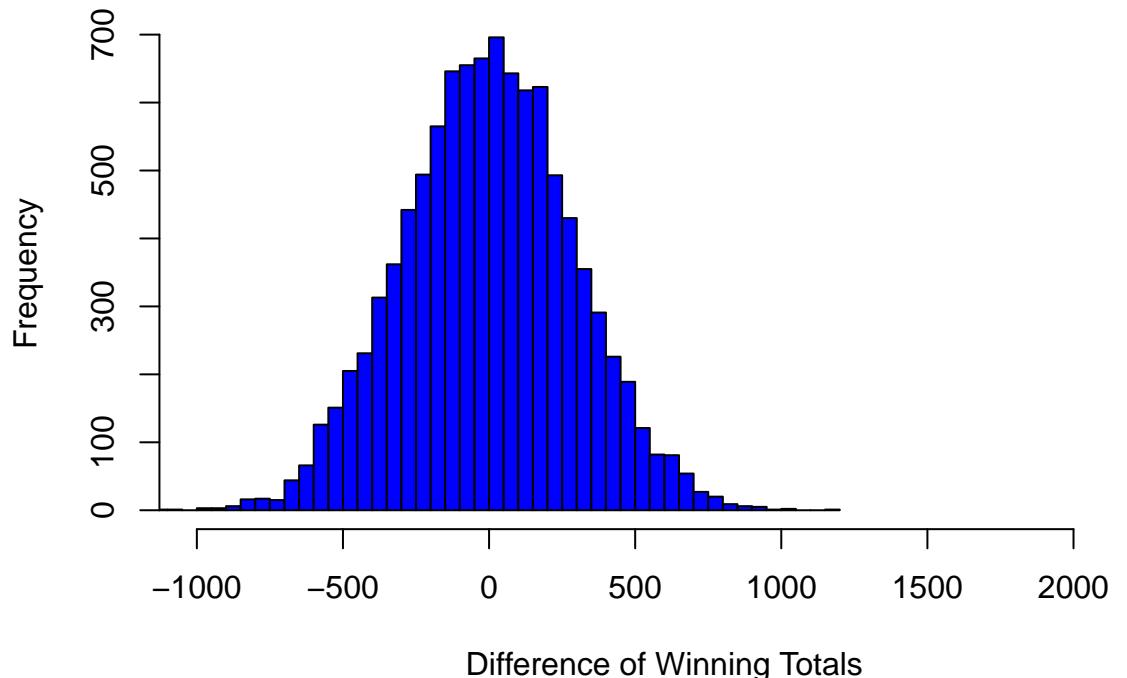
  shuff_after <- shuff_group[1:length(winner_after)]
  shuff_before <- shuff_group[(length(winner_after) + 1):length(comb_group)]

  mean(shuff_after, na.rm = TRUE) - mean(shuff_before, na.rm = TRUE)
}

hist(null_dist, breaks = 50, main = "Null Distribution of Difference of Winning Totals",
     xlab = "Difference of Winning Totals", col = "blue", xlim = c(-1000,
     2200))

abline(v = c(obs_stat), col = "red", lwd = 2)
```

Null Distribution of Difference of Winning Totals



```
# Calculate p-value
p_val <- pnull(obs_stat, null_dist, lower.tail = FALSE)

p_val

## [1] 0
```

Parametric Hypothesis Test

```
# Define variables and find degrees of freedom
xbar_after <- mean(winner_after)
xbar_before <- mean(winner_before)
n_after <- length(winner_after)
n_before <- length(winner_before)
s_after <- sd(winner_after)
s_before <- sd(winner_before)
df <- n_after - 1

# Calculate SE and t-statistic
SE <- sqrt((s_after^2/n_after) + (s_before^2/n_before))
```

```

t_stat <- (xbar_after - xbar_before)/SE

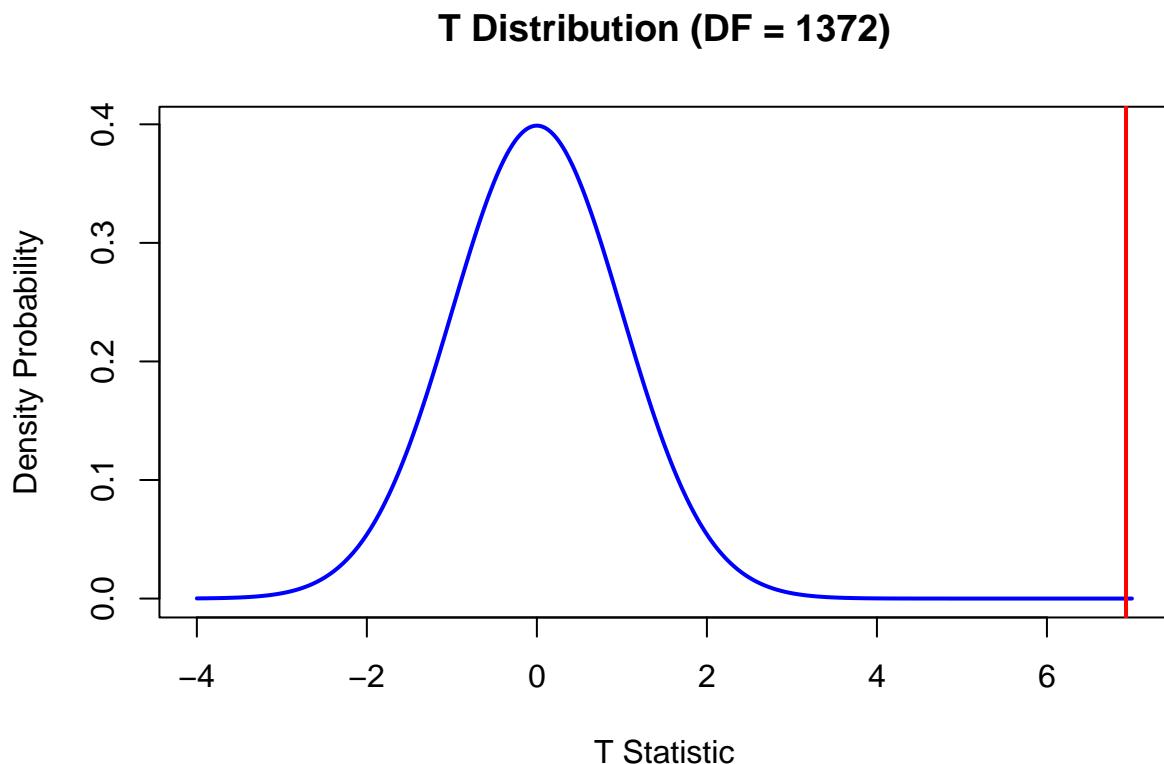
# Create and visualize null distribution using parametric
# methods
x_vals <- seq(-4, 7, length = 1000)

y_vals <- dt(x_vals, df)

plot(x_vals, y_vals, type = "l", main = "T Distribution (DF = 1372)",
      xlab = "T Statistic", ylab = "Density Probability", col = "blue",
      lwd = 2, xlim = c(-4, 7))

abline(v = t_stat, col = "red", lwd = 2)

```



```

# Calculate p-value
p_val <- pt(t_stat, df, lower.tail = FALSE)

p_val

```

[1] 3.223893e-12

Observed Statistic: 2193.6
T-Statistic: 6.93

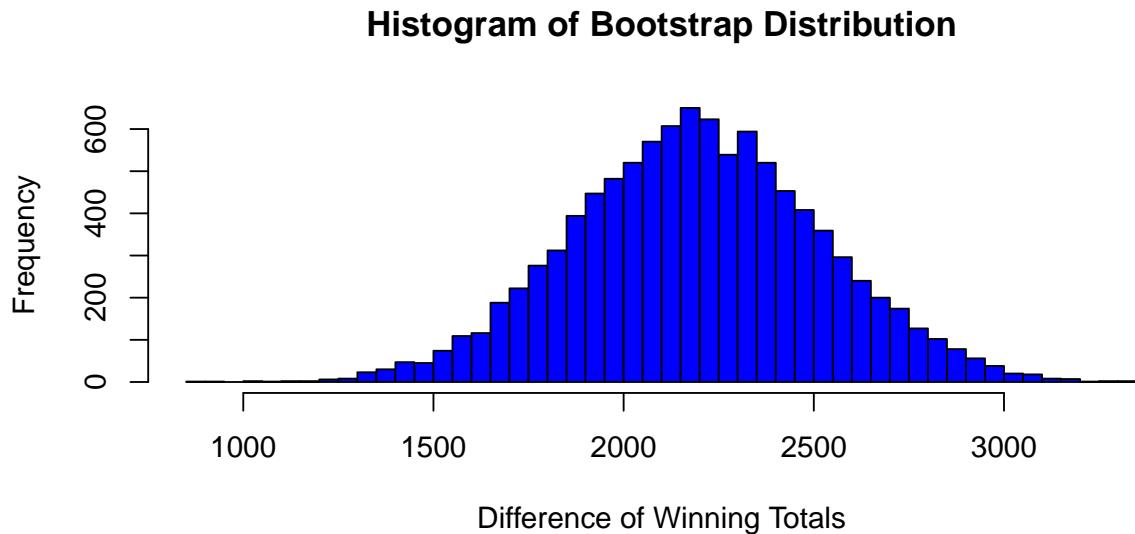
P-value for permutation test: 0
P-value for T-test: 3.2×10^{-12}

Conclusion: The p-values for each hypothesis test were very similar (0 and 0.000000000003). Using an alpha value of 0.01 they are both low enough to reject the null hypothesis and accept the alternative hypothesis, concluding that the average winning total after James appeared is significantly higher than the average winning total before he appeared. It is also important to note that because sample sizes exceed 30 in both groups and the distributions are approximately normal, the t-test is appropriate.

Confidence intervals: 95% CIs using bootstrap and parametric methods

Confidence interval using bootstrap method

```
# Create and visualize bootstrap distribution
boot_dist <- do_it(10000) * {
  boot_after <- sample(winner_after, 1298, replace = TRUE)
  boot_before <- sample(winner_before, 3985, replace = TRUE)
  mean(boot_after) - mean(boot_before)
}
hist(boot_dist, breaks = 50, main = "Histogram of Bootstrap Distribution",
  xlab = "Difference of Winning Totals", col = "blue")
```



```
# Determine SE from the bootstrap
SE <- sd(boot_dist)

# Create 95% CI using observed statistic and determined SE
obs_stat <- mean(winner_after) - mean(winner_before)
```

```

CI_upper <- obs_stat + 1.96 * SE
CI_lower <- obs_stat - 1.96 * SE

CI <- c(CI_lower, CI_upper)
CI

```

```
## [1] 1554.101 2833.174
```

Confidence interval using parametric method

```

# Assign variables and calculate degrees of freedom
n_after <- length(winner_after)
n_before <- length(winner_before)
s_after <- sd(winner_after)
s_before <- sd(winner_before)
df <- n_after - 1

# Calculate SE
SE <- sqrt((s_after^2/n_after) + (s_before^2/n_before))

# Find the critical values or t*
t_star <- qt(0.975, df)

# Create 95% CI using observed statistic and calculated SE
obs_stat <- mean(winner_after) - mean(winner_before)

CI_upper <- obs_stat + t_star * SE
CI_lower <- obs_stat - t_star * SE

CI <- c(CI_lower, CI_upper)

CI

```

```
## [1] 1572.693 2814.583
```

95% Confidence interval from bootstrap: [1554.101 2833.174]

95% Confidence interval using parametric methods: [1572.693 2814.58]

Both of these confidence intervals are relatively similar, and indicate that the average increase in total winnings after James' appearance is likely around \$1560 - \$2820. What these intervals really mean is that if we calculated them 100 times with different samples, then 95 of the intervals would contain the parameter of interest, or average winnings increase in this case.

Conclusion

Summary of Statistical Test Results

Test	P-value	95% CI
Permutation	0	[1554.10, 2833.17]
Parametric	3.2×10^{-12}	[1572.69, 2814.58]

The tests from this analysis ultimately support the idea that James Holzhauer did change the game of Jeopardy! forever. The results from both parametric and permutation tests were consistent with rejecting the null hypothesis and concluding that the average winnings have certainly statistically increased likely between \$1560 - \$2820 per game. While there could be confounding variables, and we cannot attribute causation directly to Holzhauer, the results strongly suggest a shift in contestant behavior consistent with his aggressive style.

Appendix

Here is the code used to load the data and create the data frame with necessary variables and vectors:

```
# This R chunk loads the data and creates a usable data
# frame
library( SDS1000 )
library( readr )
library( tidyverse )
library( dplyr )

# Load the TSV from GitHub
tsv <- read_tsv("https://raw.githubusercontent.com/jwolle1/jeopardy_clue_dataset/refs/heads/main/jeopardy_clues.tsv")

# Convert TSV to CSV and save as a data frame
write.csv(tsv, "scoring_season1-41.csv", row.names = FALSE)
jeopardy <- read.csv("scoring_season1-41.csv")
```

```

# Identify who won each game and create a variable with the
# winning total
jeopardy <- jeopardy %>%
  mutate(winner = case_when(final_left > final_middle & final_left >
    final_right ~ final_left, final_middle > final_left &
    final_middle > final_right ~ final_middle, final_right >
    final_left & final_right > final_middle ~ final_right,
    TRUE ~ NA_real_ # handle ties or missing data
  )) %>%
  filter(!is.na(winner)) # <-- drop NA winner games

# Add unique game id to each game
jeopardy <- jeopardy %>%
  mutate(game_id = row_number())

# Remove games before Novemeber 26, 2001
jeopardy <- jeopardy[jeopardy$game_id >= 3698, ]

# Create a vector for winning totals before JH appearance
winner_before <- jeopardy %>%
  filter(game_id < 7608) %>%
  pull(winner)

# Create a vector for winning totals after JH appearance
winner_after <- jeopardy %>%
  filter(game_id >= 7651) %>%
  pull(winner)

```