Data Dictionary: Contestants Dataset

Field Name	Data Type	Description	Example Value
Season	Integer	The season the contestant competed in.	1
Name	String	Full name of the contestant.	"John Doe"
Age	Integer	Contestant's age at time of competing.	30
Hometown	String	Contestant's hometown.	"Los Angeles, CA"
Profession	String	Contestants job at time of competing.	"Chef"
Finish	Integer	The spot the contestant finished in the competition.	16
Ethnicity	String	The contestant's ethnicity.	"African American"
Gender	String	The contestant's gender. Represented in a single character.	"F"
Lgbt	Boolean	Whether the contestant identifies as lgbt or not.	TRUE
Has Disability	Boolean	Whether the contestant had a disability during the competition or not.	TRUE
State	String	The State the contestant is from.	"CA"

Data Dictionary: Contestants Dataset

Field Name	Data Type	Description	Example Value
Country	String	The Country the contestant is from.	"USA"

Explanation of Fields:

- **Season**: A number identifying which season the contestants competed in. Some contestants competed in multiple seasons.
- **Name**: Self-explanatory; used for storing the contestant's name.
- **Age**: Stores the contestant's age when they competed, which is essential for age-related calculations.
- **Hometown**: Stores the City and State or Province the contestant is from. Typical formatting 'City, ST'
- **Profession**: The job title the contestant held at time of competing.
- **Finish**: The number in which the contestant placed in the competition. Should be a whole number.
- **Ethnicity**: Stores the contestant's ethnicity.
- **Gender**: Indicates the contestant's gender. Should be formatted as a single letter, either 'F', 'M', or 'N' (in the case of contestants who identify as non-binary)
- **Lgbt**: Boolean value indicating whether or not the contestant identifies as lgbt.
- **Has Disability**: Boolean value indicating whether or not the contestant had a disability during the competition. Does not indicate whether the contestant became disabled after.
- **State**: Stores the State or Province the contestant is from. Typically formatted as a state abbreviation such as 'KY'.
- **Country**: Stores the Country of origin of the contestant. Should only be 'USA' or 'Canada' at this time.

Data Dictionary: Stats Dataset

Field Name	Data Type	Description	Example Value
Season	Integer	The season the contestant competed in.	1

Data Dictionary: Stats Dataset

Field Name	Data Type	Description	Example Value
Contestant	String	Full name of the contestant.	"John Doe"
Survival Score	Float	The sum of Challenge Win %, Tribal Council %, and Jury Vote %	1.34
Survival Average	Float	Sum of fractional Challenge Wins, Tribal Council Ratio (Weighted), and a weighted Jury %	12.26
Challenge Wins	Float	Number of challenges won by the contestant. (A fraction is scored for team challenges)	5.87
Challenge Appearances	Float	Fractional number of challenges a contestant competed or could have competed in.	16.10
Challenge Win %	Float	Challenge Wins/Challenge Appearances	0.36
Sit Outs	Integer	Number of times contestant sat out of challenges.	2
Votes for Bootee	Integer	Number of times contestant voted for the person ultimately voted out at Tribal Council.	6
Votes Against (Total)	Integer	Total number of votes cast *against* the contestant.	0

Data Dictionary: Stats Dataset

Field Name	Data Type	Description	Example Value
Total Votes Cast	Integer	Total votes cast at Tribal councils the contestant attended	71
Tribal Council Appearances	Integer	Number of times contestant appeared and voted at Tribal.	11
Tribal Council %	Float	[VFB - (VA/TVC)] / TCA	0.55
Tribal Council Ratio (Weighted)	Float	TCR =2* [VFB / (4+VA)] x (14/TCA)	3.82
Jury Votes For	Integer	Number of Jury votes for the contestant.	3
Total Number Of Jurors	Integer	Total number of jurors in contestant's season finale.	7
Jury Votes %	Float	Jury Votes % = Juror Votes For/ Total Number of Jurors	0.43
Idols Found	Integer	Number of idols contestant found in the season	1
Idols Held	Integer	Number of idols held (given to contestant)	1
Idols Played	Integer	Number of idols played	0

Data Dictionary: Stats Dataset

Field Name	Data Type	Description	Example Value
Votes Voided	Integer	Total votes voided by idols played	6
Immunity Challenge Wins	Integer	Individual immunity challenges won	5
Immunity Challenge Appearances	Integer	Individual immunity challenges played in	6
Advantages Found	Integer	Total advantages found by contestant in season	2
Advantages Played	Integer	Total advantages played by contestant in season	1
Notes	String	Additional information on when the advantage was found, how it was used.	"50-50 coin; purchased, Day 22. Played at F7, voiding 2 votes (Day 31)."

Explanation of Fields:

- **Season**: A number identifying which season the contestants competed in. Some contestants competed in multiple seasons.
- **Contestant**: Used for storing the contestant's full name.
- **Survival Score**: The sum of Challenge Win% (ChW%) + Tribal Council% (TC%) + Jury Vote%, for a maximum possible score of 2 (3 after the finale).
- **Survival Average**: Sum of fractional Challenge Wins, Tribal Council Ratio (Weighted), and a weighted Jury%. The latter two max out at six points total, for a theoretical maximum score of around 18 or so.
- **Challenge Wins**: For tribal challenges, a contestant earns a fraction of 1 win, depending on if they participated (no points for sitting out). So in a five-person tribe's win, each participant gets (1/5) of a point, or 0.2 points. Individual challenge wins count as a full point.

- **Challenge Appearances**: Used to calculate Challenge Win %. Fractional for tribal challenges (same as Challenge Wins), except sit-outs get charged for an appearance, because they could have participated.
- **Challenge Win %**: Simply, Challenge Win % = Challenge Wins / Challenge Appearances.
- **Sit Outs**: The number of times a contestant sat out of a challenge.
- **Votes for Bootee**: The number of times the contestant has voted for the person who was ultimately voted out. Applies only to initial votes (no points for revotes in case of a tie). Special case: In a final three Tribal Council, where only one vote is cast (by the Final 3 Immunity Challenge winner), only that vote counts.
- **Votes Against (Total)**: The total number of tribal council votes cast against the contestant. Again, only initial votes count (no penalty for revotes), and here a hidden immunity idol (if played) erases the votes. In the special case of a final three tribal council above, only the Final 3 bootee receives a vote against.
- **Total Votes Cast**: Total votes cast during the tribal councils the player has attended (only initial votes count). Used to adjust for different vote totals as tribes shrink.
- **Tribal Council Appearance**: The number of times a contestant has attended tribal council (at which they voted).
- **Tribal Council %**: Attempts to reward voting for the bootee (which players controlling the vote almost always do), while punishing receiving votes yourself. The formula is: Tribal Council% = [Votes For Bootee - (Votes Against (Total)/Total Votes Cast)] / Tribal Council Appearances.
- **Tribal Council Ratio (Weighted)**: Very similar in intent to Tribal Council%, but calculated as a ratio of Votes For Bootee to Votes Against (Total), while also scaling to a uniform number of Tribal Council appearances. The formula is as follows: Tribal Council Ratio (Weighted) =2* [Votes For Bootee / (4+Votes Against)] x (14/Tribal Council Appearance). 14 used as the scaling factor for Tribal Council appearances because there are usually 14 episodes, then a final scaling factor of 2 to bring maximal scores up to roughly even with Challenge Wins and Juror Votes % high scores.
- **Jury Votes For**: Number of jury votes for the contestant to win. Maximum nine (as with Earl Cole, *Fiji*), theoretically.
- **Total Number of Jurors**: Total number of jurors. Necessary to not punish unanimous 7-juror winners (ie. JT Thomas, *Tocantins*).
- **Jury Votes %**: The percent of total jury votes cast for the contestant, or Jury Votes % = Juror Votes For/ Total Number of Jurors. This number is used, raw, in Survival Score, and is scaled in Survival Average (multiplied by six) to make it similar in size to Challenge Wins and Tribal Council Ratio (Weighted).
- **Idols Found**: Self-explanatory; In cases where multiple people find an idol together (*Nicaragua*, *Caramoan*, etc.), the person who keeps it gets credit for finding it.
- **Idols Held**: Generally only relevant if someone is given an idol by someone else (such as Russell from JT in *HvsV*, Sash from Marty in *Nicaragua*).
- **Idols Played**: The credit for playing (and voiding votes) goes to the person who first reveals the idol at Tribal Council, whether they hand it to Probst themself, or hand it off to someone else, who then plays it.

- **Votes Voided**: The number of votes eliminated against the contestant by the idol.
- **Immunity Challenge Wins**: The number of individual immunity challenges a contestant won.
- **Immunity Challenge Appearances**: The number of individual immunity challenges a contestant competed in.
- **Advantages Found**: Number of advantages found by the contestant.
- **Advantages Played**: Number of advantages played
- **Notes**: Additional information on when the advantage was found, how it was used.