DS 7330

Jeopardy: A Study of Clues and Contestants Through the Years

Proposal

1. **TEAM MEMBERS/INVESTIGATORS**

Megan Ball [ballm@mail.smu.edu](mailto:ballm@mail.smu.edu)

Matt Farrow [mfarrow@mail.smu.edu](mailto:mfarrow@mail.smu.edu)

Jake Harrison [harrisonjp@mail.smu.edu](mailto:harrisonjp@mail.smu.edu)

1. **PROPOSAL**

Our proposal involves using clues and metadata from the TV game show Jeopardy, with data coming from the fan-created archive at <http://j-archive.com>. With over 400,000 clues covering games back to 1985, the site contains several fields including contestant information, categories, comments, answers, and clue values.

The main objectives for this project are to

1. Determine insights and trends of Jeopardy clues, categories, and answers over 1985 to present
2. Complete a case study of notable contestant’s tenures on Jeopardy
3. Build a model to recommend wager for final jeopardy round

Our methodology includes the following:

1. Complete web scraping of <http://j-archive.com> by building upon existing jarchive script/package in R to collect additional data fields required for above objectives
2. Create schema and database of results from web scraping
3. Query from database
4. Perform analysis on data to meet above objectives