Some of the most important factors that contribute to finding a data-related job after graduation is building a portfolio and participating in data competitions. But, for some, this can be daunting! In order to ease you into these beneficial activities, Powered by Woz U will be hosting its first-ever data competition!

Data Science students and graduates will have from October 5 until November 1 at midnight to analyze data and create a presentation to showcase their work.

Prize Categories

There will be two prize categories:

Beginner: Any student in the first five modules

- Basic Statistics
- Statistical Programming in R
- Database Foundations
- Programming Foundations in Python
- Data Wrangling and Visualization

Advanced: Any graduate or student in the remaining five modules:

- Intermediate Statistics
- Machine Learning and Modeling
- Introduction to Big Data
- Metrics and Data Processing
- Final Project

Data

Who doesn't love watching Halloween movies in the month of October? And the quintessential Halloween genre is probably the horror movie.

Here is a dataset from IMDB on Horror Movies! You will use this data throughout the contest.

Evaluation Questions

You must answer at least one of the three below evaluation questions to be considered for the prize. Points will be awarded per question, so the more you are able to complete, the more likely you are to win.

But remember that this is all about practice! Even taking a crack at one of these will help your learning and skill-building, so don’t be afraid to try.

Evaluation Questions (answer at least 1)

- All of these movies are classified as horror. But some also fit into other genres as well. What information can you provide about these additional genres?
- What factors impact Review Rating? What about Movie Rating?

Presentation of Insights

Presentation of your data on November 2 at 4pm PST is desired, but not required. Public speaking and presentation of results is something you will do often on the job; it is thus recommended you begin practicing this valuable skill in a safe environment as soon as possible!

Rules

You may work on this project with any student(s) or graduate(s) in the Powered by Woz U Data Science program.

Help from outsiders, including family members, friends, and data professionals is NOT allowed. Students or graduates from other Powered by Woz U programs cannot participate.

Participants must answer at least one evaluation question using the data provided in order to be considered.

Entries turned in after November 1 at midnight local time will not be considered.

You must use the dataset provided to answer the evaluation questions.

Findings must be presented in presentation software, such as: Microsoft Powerpoint, Keynote, or Canva. The presentation MUST be saved as a PDF before submission.

All work must be included with submission. Work may include files such as: Python, R, or SQL files; links to Tableau Public workbooks; links to Trello boards guiding the project management process; data files that have been manipulated from the original.

All files for each team must be in one zipped file per evaluation question and must be slacked or emailed to Dr. Meredith Dodd by November 1 at midnight.

Entries without presentation or relevant work files will not be considered.

Scoring Rubric

Each evaluation question will be scored using the rubric below. For each evaluation question, participants can earn a total of three points for each of the five categories: completeness, accuracy, data interpretation, visualization, and analysis difficulty, for a total of 15 points per evaluation question. If participants complete both questions, they may earn a total of 30 points.

	Completeness	Accuracy	Data Interpretation	Visualization and Presentation	Analysis Difficulty
Exceeds Expectations	3	3	3	3	3
Meets Expectations	2	2	2	2	2
Does Not Meet Expectations	1	1	1	1	1

Completeness

The participant has fully answered the evaluation question and has used all of the appropriate data. Every data preparation or assumption step for the analysis performed has been completed.

Accuracy

Data was analyzed using the correct tests and the results from those tests were correct.

Data Interpretation

The results of the analysis were properly interpreted, and findings were discussed in an appropriate manner. All conclusions drawn from the data logically follow.

Visualization & Presentation

Data visuals are easily understood and interpreted by the layperson and are visually pleasing. Presentation is clean, easy to follow, and has no typos.

Analysis Difficulty

The participant will receive additional points for attempting more difficult analyses.