

This week's assignment (D7) pertains to nonparametric significance testing.

We would like you to write code that

- 1) Builds on the data we have been using since D1 (and in particular the framework from parametric significance testing in D6). So the same data should already be loaded. We'll reuse it here.
- 2) Uses nonparametric significance tests throughout this assignment. You can assume an alpha level of 0.05 throughout.
- 3) Identify the ratings from users who have seen the movies Indiana Jones and the Raiders of the lost Ark (1981), Indiana Jones and the last Crusade (1989), Indiana Jones and the Kingdom of the Crystal Skull (2008), Ghostbusters (2016), Wolf of Wall Street (2013), Interstellar (2014), and Finding Nemo (2003). Make sure you are not off by one due to Python indexing. When comparing the ratings of pairs of movies, you will sometimes want to use all ratings, and sometimes only the ratings of users that have jointly seen both movies. We recommend that you do both, and note when it matters. Check code sessions if this point or the previous one are unclear.
- 4) Tests whether the median ratings of Indiana Jones and the Raiders of the lost Ark (1981) and Indiana Jones and the last Crusade (1989) are different.
- 5) Tests whether the median ratings of Indiana Jones and the last Crusade (1989) and Indiana Jones and the Kingdom of the Crystal Skull (2008) are different.
- 6) Tests whether the median ratings of Indiana Jones and the Kingdom of the Crystal Skull (2008) and the Ghostbusters remake from 2016 are different.
- 7) Tests whether the ratings distribution of the Ghostbusters remake from 2016 and Finding Nemo (2003) are different.
- 8) Tests whether the ratings distribution of Finding Nemo (2003) and Interstellar (2014) are different.
- 9) Tests whether the ratings distribution of Interstellar (2014) and Wolf of Wall Street (2013) are different.
- 10) Tests whether the median ratings of Interstellar (2014) and Wolf of Wall Street (2013) are different.