

Data Wrangling importance

Poll 1: Before we start, how much time do you think a Data Scientist spends in munging the data?

1. ~30%
2. ~50%
3. ~70%
4. ~90%

Data Collection

Poll 2: Which html attribute indicates the hyperlink destination?

1. [] `div` tag
2. [] `a` tag
3. [] `href`
4. [] `ul` tag

Data Cleaning 1 *(multiple select)*

Poll 1: Pick the minimum number of columns that identify an unique row in `df_rating`:

1. `userId`
2. `timestamp`
3. `movieId`
4. `rating`

Data Cleaning 2 *(multiple select)*

Poll 2: Which column(s) should we use to join `df_rating` and `df_movies`?

1. `userId`
2. `timestamp`
3. `movieId`
4. `rating`
5. `title`
6. `genres`

Poll 5: Feedback

Q2 How did you like the session overall? (1 being the lowest and 5 highest.)

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5

Q3 How useful did you find the topics covered in the session?

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5

Q4 How confident do you feel in data wrangling now?

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5

Q5 How would you rate the instructor?

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5