

WS #4 - Verbs

Monday, September 14, 2026

DS 002R - Jo Hardin

Name: _____

Names of people you worked with: _____

Tell your group about one talk/performance/event/activity not related to your classes that you are looking forward to in the coming weeks.

Task: How would you wrangle each of the following data examples into a tidy format?¹

1. the emails in your inbox
2. social media posts
3. images
4. videos
5. audio files

¹Hint: what does tidy mean? What do you need to specify?

Solution:

The key is that for each type of data, you will need to sort it into rows and columns.

1. emails – each row is a single email. the columns could be things like: text of email, time stamp, who it is to, who it is from, subject line, whether or not there are attachments, whether or not it is a reply, whether or not it is a forward, etc.
2. social media – each row is a single post. the columns could be things like: text of the post, time of the post, platform, emojis from the post, whether it is a new post or a comment, etc.
3. images – each row is a single image. each column could represent the RGB coloring of a single pixel. there might be other columns like time stamp, name of the file / image, where the image was taken, etc.
4. videos – each row is a single frame of a single video. the columns could be pixel information and audio waveform information. alternatively, each row might be a single video, and the columns might be the name of the video, the duration of the video, the time stamp from when it was taken, the location of the video, etc.
5. audio – similar to images and video, the rows could either be a single audio file or a frame of the audio file. correspondingly, the columns will either be the meta data (time, name, etc.) or the audio waveforms.