w4

Irlanda Ayon-Moreno

4/7/2018

Organizing Data

1. Be consistent

2. Write dates as YYYY-MM-DD.

3. Fill in all of the cells.

4. Put just one thing in a cell.

5. Make it a rectangle.

6. Create a data dictionary.

7. No calculations in the raw data files.

8. Don't use font color or highlighting as data.

9. Choose good names for things.

10. Make backups.

11. Use data validation to avoid data entry mistakes.

12. Save the data in plain text files.

Data Frame

- data frames = special kinds of lists

- behave similar to matrices

- cells, rows, columns can be accessed using bracket notation with indices or logocal values

- useing a - in front of an index excludes that index

Storing Data Tables

- should not store data in excel bc they're enriched files

- excel files can't be opened with a text editor (the opposite is possible)

- 2 types of txt files: delimited & fixed-width formats

Importing Tables
- it gets imported as a data frame

- default: strings are converted to factors

- things to know before:

- field delimiter

- column/row names

- possible missing values & how they are codified