Lab 7: Tidying your dataset

Eric Simo

2023-07-09

Prune the dataset

Exercise 1

```
brauer2 <- brauer %>%
  subset(
    select = -c(GID, YORF, GWEIGHT)
    )
```

Enforcing one observation per row

Exercise 2

```
brauer3 <- brauer %>%
  pivot_longer(
    cols = G0.05:U0.3,
    names_to = "sample",
    values_to = "expression"
  )
```

Enforcing one column per variable and one value per cell

Exercise 3

```
brauer4 <- brauer3 %>%
  separate(
    col = sample,
    into = combine("nutriant","rate"),
    sep = "_",            #  Specifies the separator symbol(s) or position
    convert = TRUE         #  If TRUE, tries to set data type for new columns
  )
```

```
brauer4 <- brauer3 %>%
  separate(
    col = NAME,
    into = combine("gene_name", "biological_process", "molecular_function", "systematic_id", "n
    sep = "\\|\\|",            #  Specifies the separator symbol(s) or position
    convert = TRUE         #  If TRUE, tries to set data type for new columns
  )
```

Exercise 4

```
brauer5 <- brauer %>%
  separate(
    col = NAME,
    into = combine("gene_name", "biological_process", "molecular_function", "systematic_id", "n
    sep = "\\|\\|",            #  Specifies the separator symbol(s) or position
    convert = TRUE         #  If TRUE, tries to set data type for new columns
  )
```

Exercise 5

```
brauer_tidy <- brauer5 %>%
  mutate_at(vars(gene_name:systematic_id), str_trim)
```

Visualizations using the tidy dataset

Exercise 6