Not SEAtS!





Data Wrangling

PCHN63101

20 October 2025

SEAtS



182672



Assignment 1 Data Wrangling & Visualisation

50% of the unit mark

Due November 27th at 14:00 (2pm)
Or one week later with DASS extension

Make a project in RStudio!

Recap / Questions

head() / str() Inspect a dataframe filter() Choose certain rows select() Choose certain columns Create or edit a column mutate() recode() Renaming the contents of a column Split a column into multiple columns separate() pivot longer() Turn wide data into long data Turn long data into wide data pivot wider() Join two dataframes keeping rows from the left one left_join()

The pipe %>%

```
mpgdat <- group_by(mpg,manufacturer)
mpgdat <- summarise(mpgdat,mean_hwy = mean(hwy))
mpgdat <- filter(mpgdat,mean_hwy > 20)

OR

mpgdat <- group_by(mpg,manufacturer) %>%
  summarise(mean_hwy = mean(hwy)) %>%
  filter(mean_hwy > 20)
```

Demo

Section 3.18 from the course site

```
pivot_longer()
left_join()
```

Exercises

- 1. Have a look at the starwars data in the tidyverse. Then work out the mean height of humans
- 2. What else can you calculate other than the mean?
- 3. Can you edit the code so that it calculates the mean height of each species?
- 4. Type data() into the console to see all the available datasets. Choose one and find an interesting mean
- 5. In the Presidential data set use mutate() and recode() to change the party names to 'Dem' and 'Rep'
- 6. What happens when you run this code:
- df <- mtcars %>% mutate(car_info = paste(gear, carb, sep = "_")) %>% select(car_info)
- 7. Now use separate() to split the car_info column into two new columns
- 8. Practice using filter() and select() on a new dataset

SEAtS



182672

Not SEAtS!



