LCBC

LIFESPAN CHANGES
in brain and cognition

Introduction to Tidyverse

Athanasia Monika Mowinckel

Sept. 15th 2020

Timeline

- **Introduction to tidyverse concepts** (15 minutes)
- **Tidy data wrangling** - with translations to base-R (~ 2 hours)
    - plotting data with ggplot2 (~25 min)
    - sub-setting data with dplyr (~25 min)
    - chaining commands with the pipe `%>%` (~10 min)
    - adding and altering variables with dplyr (~25 min)
- **Lunch break** (30 min)
- **Tidy data reshaping & summaries** - avoiding loops (~ 1.5 hours)
    - pivoting data with tidyr (~25 min)
    - grouped summaries with dplyr (~25 min)
    - working with nested data using purrr (~25 min)

> All parts (except first intro) come with small exercises and breaks between.

What is the Tidyverse?

R packages for data science

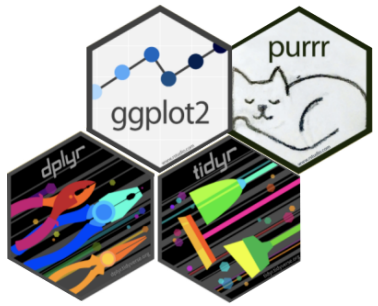
The tidyverse is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Install the complete tidyverse with:

```
install.packages("tidyverse")
```

The packages

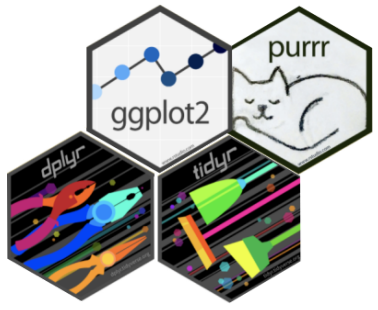
Covered here



Package	Functions
dplyr	data manipulation - altering and adding variables in a dataset
tidyr	data tidying - changing data shape and structure
ggplot2	data visualisation
purrr	Enhancements on functional programming

The packages

Covered here



Package	Functions
dplyr	data manipulation - altering and adding variables in a dataset
tidyr	data tidying - changing data shape and structure
ggplot2	data visualisation
purrr	Enhancements on functional programming

Not covered

Package	Functions
readr	easy and fast importing of data
tibble	variations on the R data.frame
forcats	working with factors/categorical data
stringr	working with strings/characters

What's so special about them?

- made to work with data sets (tibbles / data.frames)
  - geared towards data-scientists and folks working with rectangular data
- made to work together
  - naming and argument conventions
- use function names similar to spoken language
  - verbs and adverb function names
- arguments can be chained (piped) together
  - avoids saving intermediary objects
  - input always as first argument
- packages are actively developed and maintained by RStudio
  - functions will not abruptly or without warning and careful thought change behaviour