# Data analysis

Epidemiological data analysis with R

Lecture 3

MαMoδ Africa

AIMS

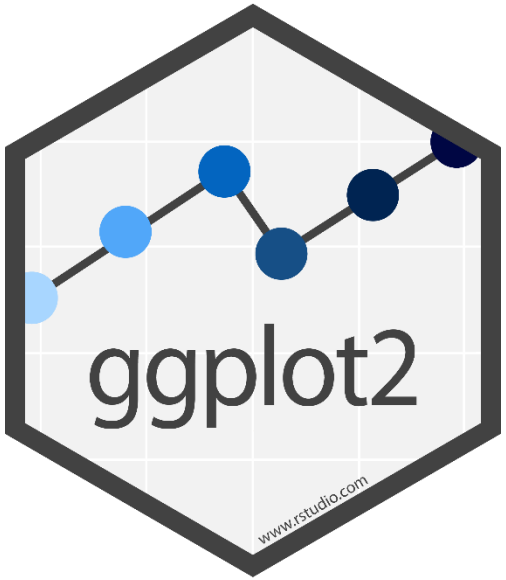
# Compute

- Confirmed cases in under and above 5 years by municipality in 2018
  - Same at national level
- Test positivity rate in all ages by municipality, month and age
- Incidence of confirmed cases in all ages in by municipality and month
- Incidence of confirmed cases in all ages by municipality in 2018

Hint: we are not interested by the number of health facilities – how can you have one row per Admin1 and month? How may rows are these?

# Basic plot with ggplot2

- ggplot2 is an R package for plots
- Modular: start with ggplot(), then add elements with +




data.frame

aesthetics: which column for x and y axis, colours,…

```
ggplot(TPR,
       aes(x=date_tested,y=TPR,color=age_group,group=age_group))+
  geom_line()+
  facet_wrap(Admin1~.)
```

line plot

one subplot for each value in this column

Cheatsheet

MαMoδ Africa

AIMS

# Feedback

https://docs.google.com/forms/d/e/1FAIpQLSemM01aWBsErD_RsLOBleRyQOpxai3J1vCZGRFpC9tZjV9osw/viewform?usp=sf_link

MαMoδ Africa

AIMS