

# Introduction to ggplot2: Creating simple publication quality figures

Matthew Castelo

June, 2021

scatter plot age by tumour size - color by sex - color by hospital - size by grade

boxplots

proportion as age increases

km curves

HR figures (catpillar plot)

Packages and working directory

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.3      v purrr  0.3.4
## v tibble  3.1.2      v dplyr  1.0.6
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(formatR)
```

```
db <- read.csv("C:/Users/matth/Documents/R/R code and education/Tutorials/Datasets/generic_cancer.csv")
```

```
head(db)
```

```
##      male pt_id age high_grade adverse_tumour_marker tumour_size diabetes
## 1   Male     1  62         No                No         0.41         No
## 2   Male     2  57         No                No         2.84         No
## 3 Female     3  66         No                No         3.35         No
## 4   Male     4  54         No                No         4.63         No
## 5   Male     5  49         No                No         1.29         No
## 6   Male     6  63         No                Yes         1.99         No
## heart_disease      hospital extended_resection postop_complication
## 1           No Community hospital                Yes                Yes
## 2           No Community hospital                No                 No
## 3           No Community hospital                No                 Yes
```

## 4	Yes Community hospital	No	Yes
## 5	No Community hospital	No	No
## 6	No Academic hospital	Yes	Yes
##	vital_status follow_up		
## 1	Alive 58		
## 2	Died 52		
## 3	Died 33		
## 4	Alive 58		
## 5	Alive 56		
## 6	Alive 41		