Learning Quarto

Tiago A. Marques

Table of contents

0.1 Quarto

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

0.2 Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.2     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.0
v ggplot2   3.4.3     v tibble    3.2.1
v lubridate 1.9.2     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to beco
```

```
1 + 1
```

```
[1] 2
```

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).

0.3 TAM's comments

Now if you need help with Quarto, you can find a lot here.

This github repository also includes extensive material of what others have done, videos, tutorials, templates, etc.

Look out for github actions which allows a document to be re-built every x minutes, updating e.g. a website.

```
5+6
```

```
[1] 11
```

1 Here I play with Saana's data

Stuff for this project, including data and code, is at this githubrepos.

```
#assumes file in same forder as .qmd file
load("1_tethys_docs_2023-08-13.Rd")
```

Now look at what got read in

```
dim(df_all)
```

```
[1] 8790   52
```

```
names(df_all)
```

```
 [1] "Title"                       "Link"
 [3] "Authors"                     "Month_year"
 [5] "Content_type"                "Author"
 [7] "Publication Date"            "Journal"
 [9] "Volume"                      "Pages"
[11] "Publisher"                   "Affiliation"
[13] "Technology"                  "Receptor"
[15] "Website"                     "Attachment"
[17] "Citation"                    "Abstract"
[19] "Stressor"                    "Issue"
[21] "Thesis Type"                 "Academic Department"
[23] "Place Published"             "Sponsoring Organization"
[25] "Document Number"             "Event Name"
[27] "Event Location"              "Book Title"
[29] "Magazine"                    "Edition"
[31] "Chapter"                     "Status"
[33] "Project Manager"             "Tech Developer"
[35] "Start Date"                  "Info Last Updated"
[37] "Contact"                     "Technology Subtype"
[39] "Support Structure"           "Project Scale"
[41] "Grid Connection"             "Installed Capacity"
[43] "Electrical Infrastructure"   "Country"
[45] "Physical Site"               "Water Depth"
[47] "Channel Width"               "Running Time"
[49] "End Date"                    "Researcher"
[51] "State"                       "Year"
```

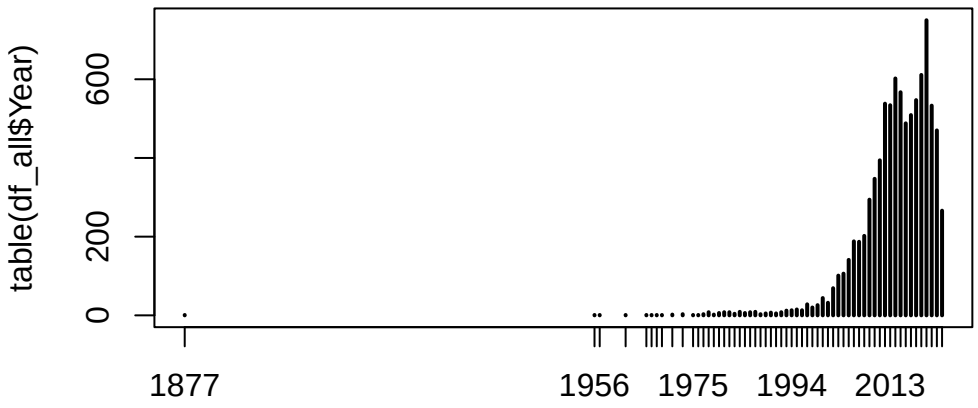
```
df_all %>%
    separate_rows(Technology, sep=",")  %>%
    mutate_at(vars(Technology), str_trim)
```

```
# A tibble: 16,128 x 52
   Title Link  Authors Month_year Content_type Author `Publication Date` Journal
   <chr> <chr> <chr>   <chr>      <chr>        <chr>  <chr>              <chr>
 1 "WIL~ /pub~ "Kuru,~ November ~ "Journal Ar~ Kuru,~ November 30, 2023  Expert~
 2 "Win~ /pub~ "Moust~ September~ "Journal Ar~ Moust~ September 20, 2023 Scienc~
 3 "Win~ /pub~ "Moust~ September~ "Journal Ar~ Moust~ September 20, 2023 Scienc~
 4 "A r~ /pub~ "Ma, C~ September~ "Journal Ar~ Ma, C~ September 15, 2023 Journa~
 5 "Int~ /pub~ "Gkeka~ September~ "Journal Ar~ Gkeka~ September 10, 2023 Journa~
 6 "Int~ /pub~ "Gkeka~ September~ "Journal Ar~ Gkeka~ September 10, 2023 Journa~
 7 "Co-~ /pub~ "Tursc~ September~ "Journal Ar~ Tursc~ September 1, 2023  Enviro~
 8 "Co-~ /pub~ "Tursc~ September~ "Journal Ar~ Tursc~ September 1, 2023  Enviro~
 9 "Tow~ /pub~ "Weiss~ September~ "Journal Ar~ Weiss~ September 1, 2023  Ocean ~
10 "Tow~ /pub~ "Weiss~ September~ "Journal Ar~ Weiss~ September 1, 2023  Ocean ~
```

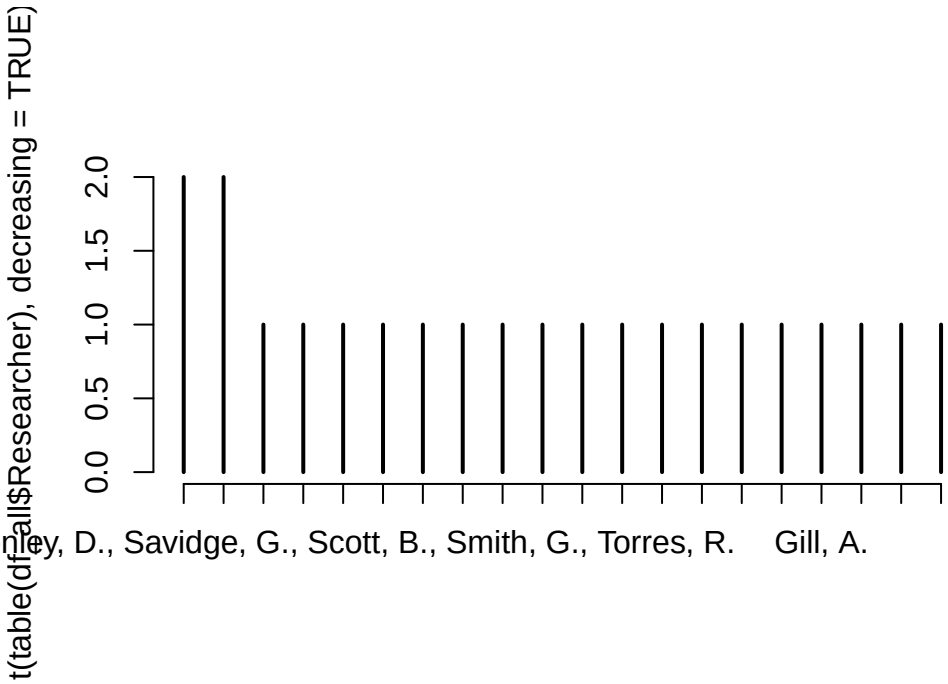
```
# i 16,118 more rows
# i 44 more variables: Volume <chr>, Pages <chr>, Publisher <chr>,
#   Affiliation <chr>, Technology <chr>, Receptor <chr>, Website <chr>,
#   Attachment <chr>, Citation <chr>, Abstract <chr>, Stressor <chr>,
#   Issue <chr>, `Thesis Type` <chr>, `Academic Department` <chr>,
#   `Place Published` <chr>, `Sponsoring Organization` <chr>,
#   `Document Number` <chr>, `Event Name` <chr>, `Event Location` <chr>, ...
```

Just taking a peak at the number of papers per year

```
plot(table(df_all$Year))
```




```
plot(sort(table(df_all$Researcher),decreasing=TRUE)[1:20])
```




```
# Pre-processing

  # Avoid NAs in filtering

  df_all$Technology[is.na(df_all$Technology)] <- ""
  df_all$Stressor[is.na(df_all$Stressor)] <- ""
  df_all$Receptor[is.na(df_all$Receptor)] <- ""

  # Identify based on institution

  df_all$USTAN <- str_detect(str_to_lower(df_all$Affiliation), "university of st andrews") |
    str_detect(str_to_lower(df_all$Affiliation), "sea mammal research unit") |
    str_detect(str_to_lower(df_all$Affiliation), "smru") |
    str_detect(str_to_lower(df_all$Affiliation), "creem")

  df_all$USTAN[is.na(df_all$Affiliation)] <- FALSE

  # View(df_all[df_all$USTAN,])

  names(df_all)[names(df_all)=="Sponsoring Organization"] <- "Sponsor"

  df_all$USTAN2 <- "Other"
  df_all$USTAN2[df_all$USTAN] <- "St Andrews"

  df_all$Marine <- str_detect(str_to_lower(df_all$Technology), "offshore") |
```

```
                  str_detect(str_to_lower(df_all$Technology), "wave") |
                  str_detect(str_to_lower(df_all$Technology), "tidal") |
                  str_detect(str_to_lower(df_all$Technology), "marine")


# Table 1, top-listed sponsors

  top_sponsors <-   df_all %>%
                      separate_rows(Sponsor, sep=",")  %>%
                      mutate_at(vars(Sponsor), str_trim) %>%
                      mutate(MarineTech = str_detect(str_to_lower(Technology), "offshore")
                               str_detect(str_to_lower(Technology), "wave") |
                               str_detect(str_to_lower(Technology), "tidal") |
                               str_detect(str_to_lower(Technology), "marine")) %>%
                      mutate(MarMammals = str_detect(str_to_lower(Receptor), "marine mammal
                      mutate(Birds = str_detect(str_to_lower(Receptor), "birds")) %>%
                      mutate(FishFisheries = str_detect(str_to_lower(Receptor), "fish")) %>
                      mutate(Inverteb = str_detect(str_to_lower(Receptor), "invertebrate"))
                      mutate(SocialHumanJustice = str_detect(str_to_lower(Receptor), "socia
                      group_by(Sponsor) %>%
                      drop_na(Sponsor) %>%
                      summarize(
                        N=length(Year),
                        USTAN=sum(USTAN),
                        MarineTech=mean(MarineTech, na.rm=T)*100,
                        MarMammals = mean(MarMammals, na.rm=T)*100,
                        Birds = mean(Birds, na.rm=T)*100,
                        Fish = mean(FishFisheries, na.rm=T)*100,
                        Inverteb = mean(Inverteb, na.rm=T)*100) %>%
                        arrange(desc(N),desc(USTAN)) %>%
                        filter(N>=20)

kable(top_sponsors,digits = 1)
```

Sponsor	N	USTAN	MarineTech	MarMammals	Birds	Fish	Inverteb
US Department of Energy (DOE)	153	2	64.7	14.4	28.1	20.3	7.2
Bureau of Ocean Energy Management (BOEM)	74	1	83.8	31.1	29.7	37.8	21.6
US Department of the Interior (DOI)	54	0	68.5	18.5	31.5	29.6	11.1
Scottish Government	53	20	88.7	43.4	26.4	15.1	1.9
Marine Scotland Science	47	10	87.2	29.8	40.4	12.8	4.3
Ocean Energy Systems (OES)	32	0	100.0	15.6	6.2	15.6	6.2

Sponsor	N	USTAN	MarineTem	MarMammals	Birds	Fish	Inverteb
National Renewable Energy Laboratory (NREL)	30	0	3.3	0.0	86.7	0.0	0.0
Collaborative Offshore Wind Research into the Environment (COWRIE)	28	2	89.3	21.4	60.7	21.4	3.6
The Crown Estate	28	2	92.9	10.7	28.6	7.1	0.0
Welsh Government	27	2	100.0	37.0	22.2	29.6	3.7
California Energy Commission	22	0	18.2	4.5	68.2	0.0	0.0
Scottish Natural Heritage	20	3	75.0	35.0	35.0	10.0	10.0
European Commission	20	0	85.0	10.0	10.0	5.0	0.0

Now looking at a plot

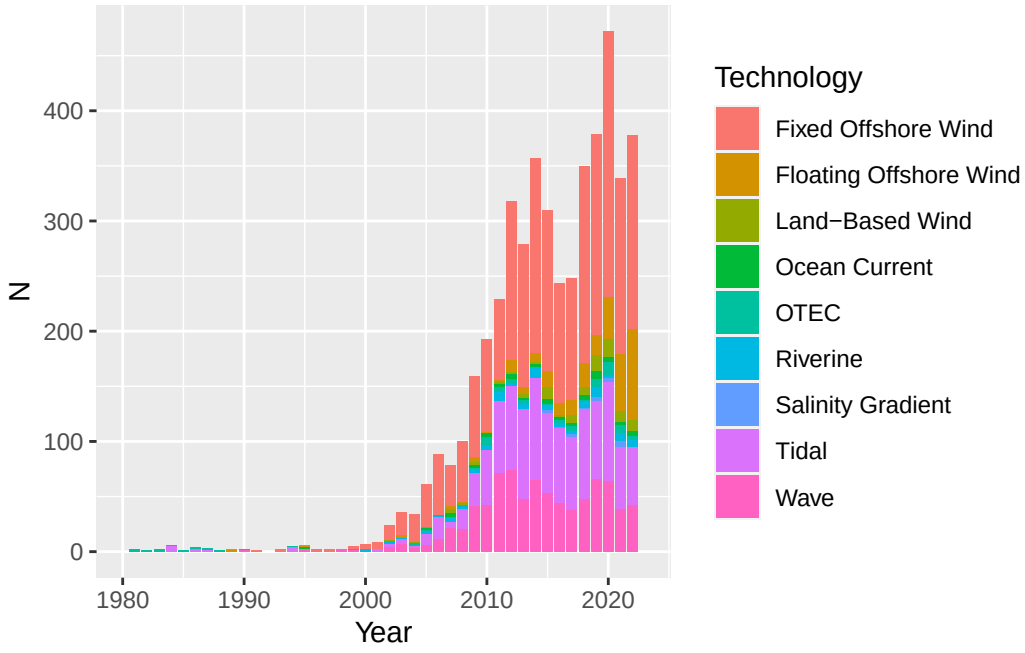
```
# Fig 2, Time series plot with technology type

df_all %>% filter(Marine) %>%
  separate_rows(Technology, sep=",")  %>%
  mutate_at(vars(Technology), str_trim) %>%
  filter(Technology!="Wind Energy") %>%
  filter(Technology!="Marine Energy") %>%
  group_by(Year, Technology) %>%
  summarize(N=length(Year))  %>%
  ggplot(aes(x=Year, y=N, group = Technology, fill = Technology)) +
  geom_bar(stat = "identity") +
  xlim(c(1980,2023))
```

```
`summarise()` has grouped output by 'Year'. You can override using the
`.groups` argument.
```

```
Warning: Removed 8 rows containing missing values (`position_stack()`).
```

```
Warning: Removed 10 rows containing missing values (`geom_bar()`).
```




```
# Fig 3, Database overview
p2 <- df_all %>%
  filter(Marine) %>%
  drop_na(Stressor) %>%
  filter(Stressor!="") %>%
  my_sankey(c("Technology","Stressor", "Receptor"), c(",",",",","), N_min = 0)
```

The Sankey plot is not working. No one knows why. Seems a bug, but only happens in my machine. Code works in both Enrico and Saana's PC's, and their code does not run on my PC… Must be my PC, but we have the same versions of `R`, `knitr` and `htmlwidgets`, so I am at a loss.

```
sankeyNetwork(Links = p2$Links,
    Nodes = p2$Nodes,
    Source = p2$Source,
    Target = p2$Target,
    Value = p2$Value,
    NodeID = p2$NodeID,
    #units = "TWh",
    fontSize = p2$fontSize,
    nodeWidth = p2$nodeWidth,
    iterations = p2$iterations,
    sinksRight = p2$sinksRight)
```

Looking at references per country.

```r
par(mar=c(8,4,0.2,0.2))
barplot(table(df_all$Country),las=2)
```

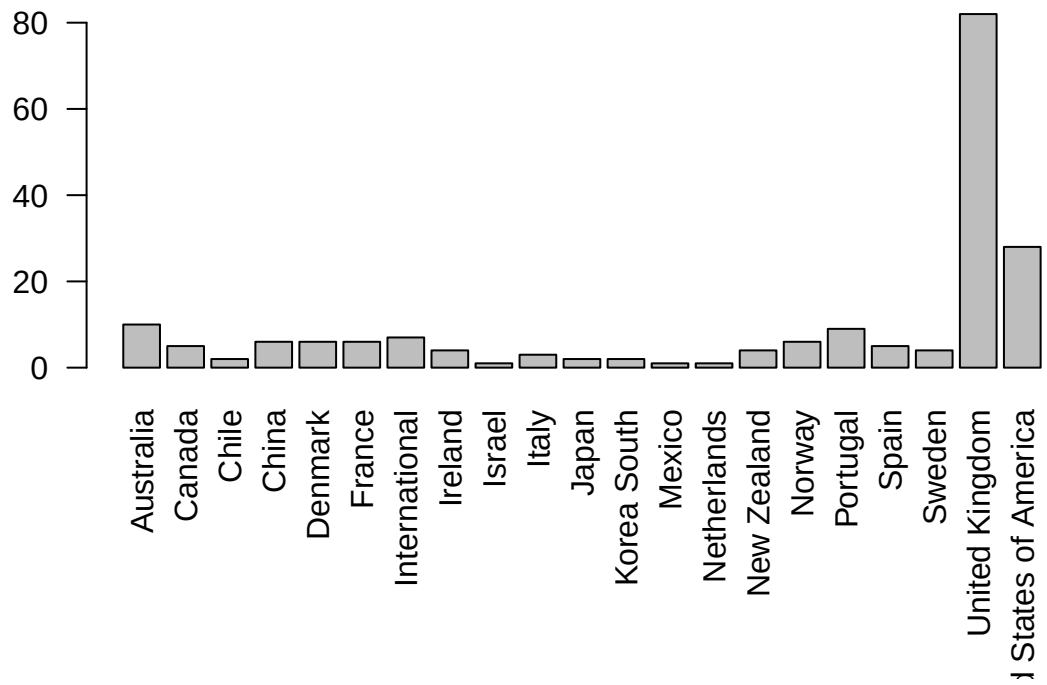


Figure 1: Number of papers per country.

We can see that in figure Figure 1.

Citing a book Buckland *et al.* (2001)

Adding math to text

$$\hat{N} = \frac{n}{\hat{P}}$$

References

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001). *Introduction to distance sampling: Estimating abundance of biological populations*. Oxford University Press, Oxford.