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 ──  
✔ dplyr 1.1.2 ✔ readr 2.1.4  
✔ forcats 1.0.0 ✔ stringr 1.5.0  
✔ ggplot2 3.4.3 ✔ tibble 3.2.1  
✔ lubridate 1.9.2 ✔ tidyr 1.3.0  
✔ purrr 1.0.2   
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

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](https://quarto.org/docs/guide/).

This [github](https://github.com/mcanouil/awesome-quarto) 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 [github](https://github.com/TiagoAMarques/LearningQuarto)repos.

#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)

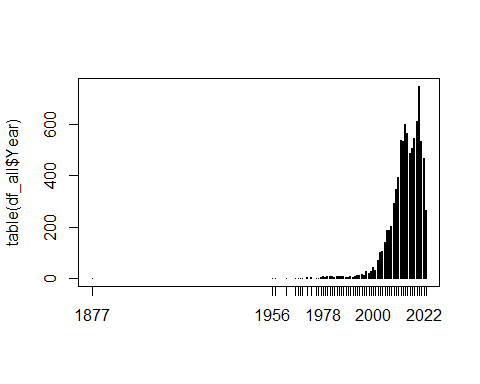
[1] "Title" "Link"   
 [3] "Authors" "Month\_year"   
 [5] "Content\_type" "Author"   
 [7] "Publication Date" "Journal"   
 [9] "Volume" "Pages"   
[11] "Publisher" "Affiliation"   
[13] "Technology" "Receptor"   
[15] "Website" "Attachment"   
[17] "Citation" "Abstract"   
[19] "Stressor" "Issue"   
[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)

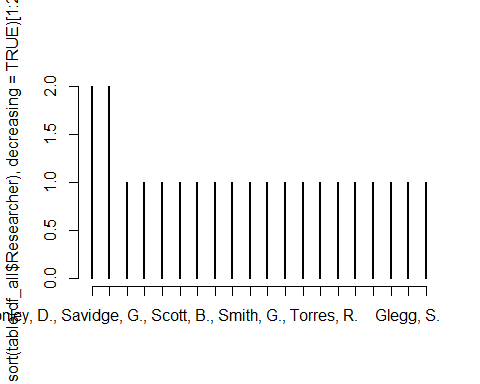
# A tibble: 16,128 × 52  
 Title Link Authors Month\_year Content\_type Author `Publication Date` Journal  
 <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>   
 1 "WIL… /pub… "Kuru,… November … "Journal Ar… Kuru,… November 30, 2023 Expert…  
 2 "Win… /pub… "Moust… September… "Journal Ar… Moust… September 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 mammals")) %>%  
 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), "social") | str\_detect(str\_to\_lower(Receptor), "human") | str\_detect(str\_to\_lower(Receptor), "justice")) %>%  
 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 |
| 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 |

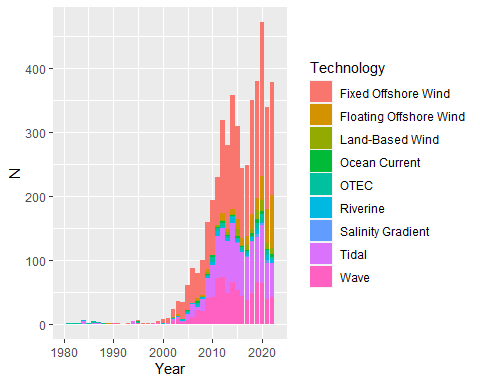
Now looking at a plot

# 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.

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](#fig-scatterplot).

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

Adding math to text

# 2. 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.