

My title*

My subtitle if needed

First author Another author

February 10, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

You can and should Cross-reference sections and sub-sections. We use R Core Team (2023) and Wickham et al. (2019).

The remainder of this paper is structured as follows. Section 2....

2 Data

Some of our data is of penguins (Figure 1), from Horst, Hill, and Gorman (2020).

Talk more about it.

And also planes (Figure 2). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

Talk way more about it.

```
#### Workspace setup ####  
library(tidyverse)  
library(knitr)  
library(dplyr)
```

*Code and data are available at: [LINK](#).



Figure 1: Bills of penguins



Figure 2: Relationship between wing length and width

```
#### Load data ####
```

```
raw_publication_data <- read_csv(here::here("data/raw_data/raw_data.csv"))
```

```
Rows: 684 Columns: 21
```

```
-- Column specification -----
```

```
Delimiter: ","
```

```
chr (21): Christopher Darnton, "The Provenance Problem: Research Methods and...
```

```
i Use `spec()` to retrieve the full column specification for this data.
```

```
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
#### Clean data ####
```

```
# Isolate publication data
```

```
publication_data <- raw_publication_data[c(119:685), c(0:21)]
```

```
# Convert top row to column names
```

```
# Uses stackoverflow query answer: https://stackoverflow.com/a/57531480
```

```
names(publication_data) <-
```

```
  publication_data |>
```

```
  slice(1) |>
```

```
  unlist()
```

```
publication_data <- publication_data |> slice(-1)
```

```
# Make journal name key
```

```
journal_names <- raw_publication_data[c(24:43), c(16,17)]
```

```
journal_names <-
```

```
  journal_names |>
```

```
  rename(name = ...16 , shortform = ...17)
```

```
# Isolate TRIP ratings (from table 1 data in data set)
```

```
rating <- raw_publication_data[c(24:43), c(20)]
```

```
rating <- rating |>
```

```
  rename(trip = "...20")
```

```
#### Build Figures ####
```

```
# Graph 1
```

```
# Isolate year, code, journal, and peer-reviewed status (0 for Foreign Policy
```

```

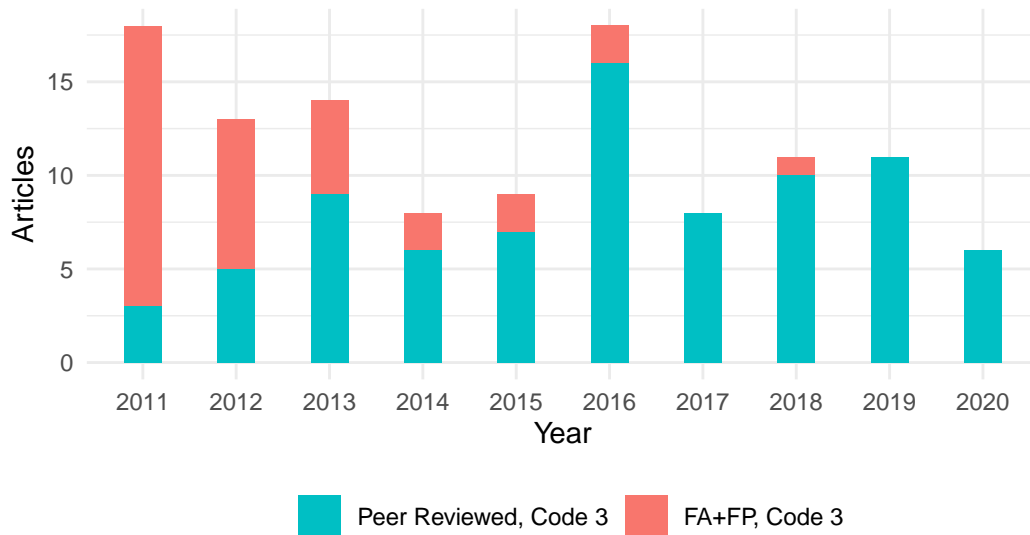
# and Foreign Affairs, 1 for all others.)
leak_references <-
  publication_data |>
  select(`Year`, `C`, `J`, `PEER`)

# Filter out all non code 3 rows (using leaked sources directly)
leak_references <-
  filter(leak_references, `C` == "3")

# Construct Graph
leak_references |>
  ggplot(aes(x= `Year`, fill = (`PEER` == 1))) +
  geom_bar(width = .4) +
  theme_minimal() +
  labs(x = "Year", y = "Articles") +
  theme(legend.title = element_blank()) +
  theme(legend.position = "bottom") +
  ggtitle("Articles Apparently Referencing Leaked
Material Directly, 2010-2020, in TRIP 2011-ranked
Journals (n=116)") +
  theme(plot.title = element_text(size=10)) +
  scale_fill_discrete(labels = c("FA+FP, Code 3", "Peer Reviewed, Code 3"))+
  guides(fill = guide_legend(reverse = TRUE))

```

Articles Apparently Referencing Leaked
Material Directly, 2010–2020, in TRIP 2011–ranked
Journals (n=116)



```
# Graph 2

# Select only number of cables cited
cables_cited <-
  publication_data |>
  select(`NUM CABLES`)

# Convert "??" to -1
cables_cited$`NUM CABLES`[cables_cited$`NUM CABLES` == "??"] <- "-1"

# Covert all entries to numbers
cables_cited$`NUM CABLES` <- as.numeric(cables_cited$`NUM CABLES`)

# Omit all N/A articles
cables_cited <-
  cables_cited |> na.omit(cables_cited)

# Make counts for each cable number
cables_cited<-
  cables_cited |> count(`NUM CABLES`, .drop = FALSE)

cables_cited <- cables_cited |>
  add_row("NUM CABLES" = 9, n = 0, .after = 10) |>
```

```

add_row("NUM CABLES" = 13, n = 0, .after = 14) |>
add_row("NUM CABLES" = 14, n = 0, .after = 15) |>
add_row("NUM CABLES" = 15, n = 0, .after = 16)

# Make column for combined cables cited
cables_cited <-
  cables_cited |>
  mutate(combined_cables = `NUM CABLES` * n)

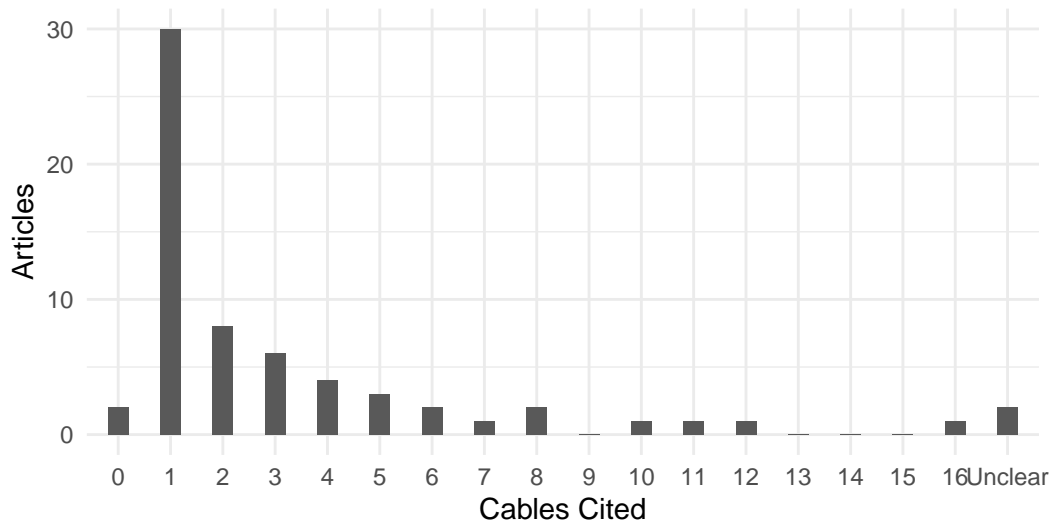
# Covert all entries to numbers and change -1 back to ?? or "Unclear"
cables_cited$`NUM CABLES` <- as.character(cables_cited$`NUM CABLES`)
cables_cited$`NUM CABLES`[cables_cited$`NUM CABLES` == "-1"] <- "Unclear"
cables_cited$`combined_cables`[cables_cited$`combined_cables` == -2] <- 2

# Reposition "unclear" to the bottom of the table
cables_cited <- cables_cited |>
  slice(2:18, 1)

# Construct Graph
cables_cited |>
  ggplot(aes(x= reorder(`NUM CABLES`, 1:18), y = n)) +
  geom_bar(stat = "identity", width = .4) +
  theme_minimal() +
  labs(x = "Cables Cited", y = "Articles") +
  ggtitle("Articles Apparently Referencing Leaked
          US Diplomatic Cables Directly (n=64), in
          Peer-reviewed, TRIP 2011-Ranked Journals,
          2010-2020, by Cables Cited") +
  theme(plot.title = element_text(size=10))

```

Articles Apparently Referencing Leaked
US Diplomatic Cables Directly (n=64), in
Peer-reviewed, TRIP 2011-Ranked Journals,
2010–2020, by Cables Cited



```
#### Build Table ####
```

```
table_data <- publication_data
```

```
# Isolate code, journal, and year
```

```
table_data <-  
  table_data |>  
  select(`C`, `J`, `Year`)
```

```
# sub table for c2
```

```
table_data_c2 <-  
  filter(table_data, `C` == 2)
```

```
# sub table for c3
```

```
table_data_c3 <-  
  filter(table_data, `C` == 3)
```

```
# counts for c2 articles
```

```
table_data_c2_counts <-  
  table_data_c2 |> count(`J`, .drop = FALSE, name = "c2_count")
```

```
# counts for c3 articles
```

```
table_data_c3_counts <-
```

```

table_data_c3 |> count(`J`, .drop = FALSE, name = "c3_count")

# add rows to table_data_c2 where count is 0. (FIX to omit)
table_data_c2_counts <- table_data_c2_counts|>
  add_row("J" = "FA", c2_count = 0, .after = 4) |>
  add_row("J" = "FP", c2_count = 0, .after = 5) |>
  add_row("J" = "IO", c2_count = 0, .after = 8) |>
  add_row("J" = "IR", c2_count = 0, .after = 9)

# merge counts
merged_c2_c3 <- cbind(table_data_c3_counts, table_data_c2_counts["c2_count"])

# sort journal names alphabetically
journal_table_names <- journal_names |> arrange(name)

# add long form journal names
c2_c3_data <- cbind(merged_c2_c3, journal_table_names["name"])

# move names to first column
c2_c3_data <- c2_c3_data |> relocate(name)

# make empty df for first year of code 3 article
first_pub <-
  tibble(years = rep(c(9980:9999)) )

# calculate first year published code 3 for each journal
publication_data

```

A tibble: 566 x 21

	C	Title	`Author(s)`	J	V	N	Year	Notes	CLASS	`NO SOURCE`
	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>
1	3	Democracy ~	Treisman, ~	APSR	114	3	2020	"p. ~	<NA>	<NA>
2	3	The juche ~	Howell, Ed~	IA	96	4	2020	"fn.~	<NA>	<NA>
3	3	The future~	Liao, Kai	IA	96	5	2020	"Bod~	<NA>	<NA>
4	3	Selective ~	Grigoryan,~	IS	44	4	2020	"fn.~ y		<NA>
5	3	All Hegemo~	Nicholls, ~	ISR	22	3	2020	"p. ~	<NA>	<NA>
6	3	Bucking th~	Roberts, D~	SS	29	2	2020	"Cit~ y		<NA>
7	3	Radicalism~	Marcusa, M~	CP	51	2	2019	"fn.~	<NA>	<NA>
8	3	The Respon~	Dunford, R~	EJIR	25	4	2019	"p. ~	<NA>	<NA>
9	3	South Afri~	Jordaan, E~	GG	25	1	2019	"p. ~	<NA>	<NA>
10	3	Who Does G~	Thrandardo~	GG	25	4	2019	"two~	<NA>	<NA>


```

# i 556 more rows
# i 11 more variables: `NO USG` <chr>, LEAK <chr>, EUPH <chr>, WLC <chr>,
#   `TO FROM FOUND` <chr>, U <chr>, CABLE <chr>, PEER <chr>, QUADRANT <chr>,
#   `BODY LEAK` <chr>, `NUM CABLES` <chr>

for (x in 1:20){
  for(y in 1:nrow(publication_data)){
    if(isTRUE(publication_data$J[y] == c2_c3_data$J[x]) == TRUE)
      if(isTRUE(publication_data$C[y] == 3) == TRUE)
        if(isTRUE(publication_data$Year[y] <= first_pub$years[x]) == TRUE)
          first_pub$years[x] = publication_data$Year[y]
        }
  }
}

# add first years to df
c2_c3_data <- cbind(c2_c3_data, first_pub["years"])

# order by c3 count
c2_c3_data <- c2_c3_data |> arrange(desc(c3_count))

# add trip rating
c2_c3_data <- cbind(c2_c3_data, rating["trip"])

# move years to last column
c2_c3_data <- c2_c3_data |> relocate(years, .after = trip)

# construct table
c2_c3_data |> kable(
  col.names = c("Title", " Short Form", "Code 3 Articles, 2010-2020",
                "Code 2 Articles, 2010-2020", "TRIP Rank, 2011", "First Code 3 Article"),
  booktabs = TRUE,
  caption = "Journals Publishing Work with Leaked Material"
)

```

Table 1: Journals Publishing Work with Leaked Material

Title	Short Form	Code 3 Articles, 2010-2020	Code 2 Articles, 2010-2020	TRIP Rank, 2011	First Code 3 Article
Foreign Policy	FP	25	0	9	2011
International Affairs	IA	14	15	13	2011

Title	Short Form	Code 3 Articles, 2010-2020	Code 2 Articles, 2010-2020	TRIP Rank, 2011	First Code 3 Article
Foreign Affairs	FA	10	0	3	2011
International Security	IS	10	6	4	2012
Review of International Studies	RIS	9	6	14	2013
Security Studies	SS	9	4	10	2013
International Studies Quarterly	ISQ	7	4	2	2012
Global Governance	GG	6	1	20	2014
International Studies Review	ISR	4	1	17	2013
European Journal of International Relations	EJIR	3	4	7	2012
International Organization	IO	3	0	1	2017
Review of International Political Economy	RIPE	3	1	15	2012
American Political Science Review	APSR	2	1	5	2016
Comparative Politics	CP	2	1	19	2017
International Relations	IR	2	0	18	2015
Journal of Peace Research	JPR	2	1	16	2017
World Politics	WP	2	2	6	2013
American Journal of Political Science	AJPS	1	1	12	2015
Journal of Conflict Resolution	JCR	1	2	8	2016
Millennium: Journal of International Studies	MIL	1	2	11	2011

3 Results

Our results are summarized in `?@tbl-modelresults`.

4 Discussion

4.1 First discussion point

If my paper were 10 pages, then should be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

4.2 Second discussion point

4.3 Third discussion point

4.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

.1 Diagnostics

References

- Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. <https://doi.org/10.5281/zenodo.3960218>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolmund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.