

Untitled

** Especially on Windows, run `quarto install tinytex` in the terminal. This may take a while. More info here: <https://quarto.org/docs/output-formats/pdf-engine.html>

MacOS and Unix may have different requirements.

ALSO: Use `{kableExtra}` to stylize and manipulate your table (headers, subtitles, totals, LaTeX features, etc.) Alternative to `{kableExtra}` is `{flextable}` See more at <https://davidgohel.github.io/flextable/reference/index.html> AND <https://ardata-fr.github.io/flextable-book/>

LAST bit of advice. I'm not a regular PDF document producer and hence I'm not familiar with all that `{kableExtra}` or `{flextable}` can do. But what they do is similar to `{gt}`. The difference is that `{gt}` output is not suitable for PDF. In any case, since I am familiar with `{gt}` I will show you some conceptual table editing. You can figure out how you do that with `{kableExtra}` or `{flextable}`. I suspect `{flextable}` has better documentation. See my example at the bottom of `basis_pivot_summarize_print.qmd`. In all cases, here and above, be sure to render the document to get the full effect.

```
library(tidyverse)
```

```
-- Attaching packages ----- tidyverse 1.3.2 --
v ggplot2 3.4.0      v purrr   1.0.0
v tibble  3.1.8      v dplyr   1.0.10
v tidyr   1.2.1      v stringr 1.5.0
v readr   2.1.3      v forcats 0.5.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

here's a summary

```
starwars |>
  select(name:hair_color) |>
  slice_head(n = 8)
```

name	height	mass	hair_color
Luke Skywalker	172	77	blond
C-3PO	167	75	NA
R2-D2	96	32	NA
Darth Vader	202	136	none
Leia Organa	150	49	brown
Owen Lars	178	120	brown, grey
Beru Whitesun lars	165	75	brown
R5-D4	97	32	NA

```
my_special_df <- starwars |>
  drop_na(mass, height) |>
  mutate(species = fct_lump_min(species, 3)) |>
  select(name:hair_color, species) |>
  group_by(species) |>
  summarise(mean(height), sd(mass))
my_special_df
```

species	mean(height)	sd(mass)
Droid	140.0000	51.03185
Human	179.5455	19.38334
Other	174.3438	229.48174
NA	178.0000	NA

```
economics |>
  sample_n(25)
```

date	pce	pop	psavert	uempmed	unemploy
1984-09-01	2533.8	236760.0	11.8	7.6	8367
1990-08-01	3848.3	250439.0	8.1	5.4	7188
1994-09-01	4775.0	264017.0	6.8	9.2	7734
1982-10-01	2125.8	232816.0	11.3	9.7	11529
1991-04-01	3907.1	252643.0	8.6	6.6	8439

date	pce	pop	psavert	uempmed	unemploy
1986-01-01	2827.1	239638.0	8.6	6.7	7795
2009-01-01	9783.8	306208.0	6.2	10.7	12058
2007-02-01	9546.8	300802.0	4.1	8.5	6927
1985-10-01	2755.8	239113.0	9.1	7.1	8298
1998-06-01	5871.7	275836.0	6.8	6.9	6212
1991-07-01	3966.0	253493.0	8.2	7.0	8586
1972-11-01	800.5	210656.0	13.6	5.7	4602
2001-12-01	7147.7	286570.0	4.5	8.2	8258
2015-01-01	12046.0	319928.6	7.7	13.2	8903
2004-08-01	8253.1	293719.0	5.2	9.2	7990
1967-12-01	525.1	199657.0	11.8	4.8	3018
2011-09-01	10738.1	312429.1	6.8	22.0	13948
1992-09-01	4255.3	257548.0	8.7	8.6	9781
2014-09-01	11957.4	319125.3	7.4	13.4	9262
1986-04-01	2835.2	240094.0	9.7	6.7	8364
1999-02-01	6101.8	277992.0	6.2	6.8	6111
1990-02-01	3728.2	248827.0	8.6	5.3	6651
1996-07-01	5251.9	269667.0	6.7	8.3	7337
1984-01-01	2419.4	235385.0	10.0	9.1	9008
1991-08-01	3969.1	253807.0	8.6	7.3	8666

pivot longer so that I can gather summary statistics

```
economics |>
  mutate(my_era = lubridate::year(date) < 1986) |>
  pivot_longer(cols = pce:unemploy, names_to = "my_variable", values_to = "value") |>
  sample_n(10)
```

date	my_era	my_variable	value
1993-08-01	FALSE	uempmed	8.2
1996-08-01	FALSE	uempmed	8.4
1997-02-01	FALSE	uempmed	8.1
1974-05-01	TRUE	pce	922.4
1999-02-01	FALSE	unemploy	6111.0
1970-10-01	TRUE	pce	658.3
1969-07-01	TRUE	pce	602.7
1996-07-01	FALSE	uempmed	8.3
2013-10-01	FALSE	pce	11419.8

date	my_era	my_variable	value
2003-12-01	FALSE	psavert	5.4

gather summary statistics

```
economics |>
  mutate(my_era = lubridate::year(date) < 1986) |>
  pivot_longer(cols = pce:unemploy, names_to = "my_variable", values_to = "value") |>
  group_by(my_era, my_variable) |>
  summarise(my_mean = mean(value))
```

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

my_era	my_variable	my_mean
FALSE	pce	7.007261e+03
FALSE	pop	2.812351e+05
FALSE	psavert	6.609659e+00
FALSE	uempmed	9.891761e+00
FALSE	unemploy	8.598108e+03
TRUE	pce	1.352150e+03
TRUE	pop	2.189859e+05
TRUE	psavert	1.167117e+01
TRUE	uempmed	6.574324e+00
TRUE	unemploy	6.460351e+03

pivot wider so that a data frame can be printed in a specific manner

```
economics |>
  mutate(my_era = lubridate::year(date) < 1986) |>
  pivot_longer(cols = pce:unemploy, names_to = "my_variable", values_to = "value") |>
  group_by(my_era, my_variable) |>
  summarise(my_mean = mean(value)) |>
  pivot_wider(names_from = my_era, values_from = my_mean)
```

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

my_variable	FALSE	TRUE
pce	7.007261e+03	1.352150e+03
pop	2.812351e+05	2.189859e+05
psavert	6.609659e+00	1.167117e+01
uempmed	9.891761e+00	6.574324e+00
unemploy	8.598108e+03	6.460351e+03

print data frame in a specific manner

```
economics |>
  mutate(my_era = lubridate::year(date) < 1986) |>
  pivot_longer(cols = pce:unemploy, names_to = "my_variable", values_to = "value") |>
  group_by(my_era, my_variable) |>
  summarise(my_mean = mean(value)) |>
  pivot_wider(names_from = my_era, values_from = my_mean)
```

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

my_variable	FALSE	TRUE
pce	7.007261e+03	1.352150e+03
pop	2.812351e+05	2.189859e+05
psavert	6.609659e+00	1.167117e+01
uempmed	9.891761e+00	6.574324e+00
unemploy	8.598108e+03	6.460351e+03