

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
1971-06-01	699.8	207462.0	14.7	5.7	4949
1985-12-01	2811.3	239477.0	8.6	6.8	8138
2013-03-01	11227.1	315662.2	5.9	17.6	11689
2011-06-01	10636.8	311791.2	7.2	22.4	13962
1975-08-01	1047.0	216195.0	13.0	9.2	7928

date	pce	pop	psavert	uempmed	unemploy
1988-08-01	3368.0	245240.0	8.4	5.9	6843
1998-03-01	5750.3	275047.0	7.5	6.8	6422
1968-05-01	549.8	200361.0	12.0	4.4	2740
1974-11-01	956.2	214625.0	13.8	5.2	6140
1967-11-01	517.4	199498.0	12.8	4.7	3066
1987-11-01	3144.2	243639.0	8.5	6.2	7035
2012-02-01	10953.5	313339.0	8.0	19.7	12813
1974-01-01	884.5	212932.0	14.3	5.0	4644
1986-12-01	2997.1	241620.0	7.0	7.1	7883
1986-03-01	2823.6	239928.0	9.9	6.8	8383
1975-01-01	975.6	214931.0	13.2	6.3	7501
2005-04-01	8645.6	295490.0	3.1	9.0	7672
1984-04-01	2457.5	235839.0	11.5	8.2	8762
1981-02-01	1884.2	229071.0	10.8	7.1	8051
2010-08-01	10228.2	309957.8	6.9	21.0	14648
1997-02-01	5434.0	271585.0	6.2	8.1	7102
1977-12-01	1336.0	221303.0	11.4	6.8	6386
1983-03-01	2202.8	233613.0	10.6	10.4	11408
1999-02-01	6101.8	277992.0	6.2	6.8	6111
2009-07-01	9841.7	307439.0	6.0	16.0	14601

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
1976-08-01	TRUE	pce	1158.0
2004-09-01	FALSE	pce	8321.1
1978-07-01	TRUE	pop	222585.0
2014-11-01	FALSE	psavert	7.3
2004-11-01	FALSE	pop	294466.0
1984-08-01	TRUE	pce	2512.2
1991-12-01	FALSE	unemploy	9198.0
1993-11-01	FALSE	pce	4554.1
2005-08-01	FALSE	psavert	2.7

date	my_era	my_variable	value
2004-03-01	FALSE	uempmed	10.2

gather summary statistics

```
economics |>
  sample_n(25) |>
  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	7.619817e+03
FALSE	pop	2.877397e+05
FALSE	psavert	5.716667e+00
FALSE	uempmed	1.080833e+01
FALSE	unemploy	9.557500e+03
TRUE	pce	1.178954e+03
TRUE	pop	2.156889e+05
TRUE	psavert	1.139231e+01
TRUE	uempmed	5.838462e+00
TRUE	unemploy	5.517615e+03

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

```
economics |>
  sample_n(25) |>
  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	7.495782e+03	1295.4875
pop	2.850568e+05	216577.8750
psavert	6.741176e+00	11.7125
uempmed	1.070588e+01	5.9000
unemploy	8.765706e+03	5531.8750

print data frame in a specific manner

```
economics |>
  sample_n(25) |>
  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	7941.89286	1.474409e+03
pop	288525.27143	2.199335e+05
psavert	5.95000	1.109091e+01
uempmed	10.58571	6.863636e+00
unemploy	8942.92857	6.619818e+03