

# 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
2008-02-01	9913.4	303711.0	4.1	8.7	7497
2005-07-01	8829.5	296186.0	2.2	8.8	7406
1986-01-01	2827.1	239638.0	8.6	6.7	7795
1986-10-01	2932.9	241274.0	8.4	7.0	8243
1985-02-01	2640.8	237602.0	9.1	7.1	8321

date	pce	pop	psavert	uempmed	unemploy
1979-02-01	1517.8	224053.0	11.1	5.9	6173
2000-04-01	6671.1	281653.0	5.0	6.1	5481
1981-05-01	1913.8	229575.0	11.0	6.9	8174
1996-09-01	5296.6	270284.0	6.7	8.5	6979
1982-11-01	2149.3	232993.0	10.9	10.0	11938
1969-01-01	583.5	201760.0	10.3	4.4	2718
2014-07-01	11860.5	318662.4	7.5	13.1	9608
2004-12-01	8481.5	294694.0	6.9	9.5	7934
1992-06-01	4177.1	256589.0	10.1	8.7	10040
2007-09-01	9797.9	302546.0	3.5	8.7	7170
1976-09-01	1168.8	218440.0	11.4	7.8	7380
1969-09-01	613.2	203090.0	11.6	4.7	3040
2013-12-01	11517.9	317411.6	6.4	17.3	10404
2004-09-01	8321.1	293971.0	4.6	9.6	7927
2008-04-01	9996.8	304117.0	3.4	9.4	7637
1991-05-01	3933.2	252913.0	8.4	6.4	8736
1990-11-01	3871.9	251346.0	7.9	5.7	7764
1982-02-01	2021.2	231313.0	12.1	7.5	9705
1983-10-01	2357.6	234907.0	9.7	9.4	9887
2001-02-01	6995.8	284137.0	4.9	6.1	6089

pivot longer so that I can gather summary statistics

```
economics |>
  mutate(my_era = lubridate::year(date) < 1980) |>
  pivot_longer(cols = pce:unemploy, names_to = "my_variable", values_to = "value") |>
  sample_n(10)
```

date	my_era	my_variable	value
2005-08-01	FALSE	unemploy	7345.0
1993-11-01	FALSE	pce	4554.1
1970-01-01	TRUE	pce	628.7
1983-01-01	FALSE	unemploy	11534.0
1987-05-01	FALSE	pce	3048.4
1985-08-01	FALSE	pop	238679.0
1979-08-01	TRUE	pce	1615.6
1978-08-01	TRUE	psavert	10.5
2012-01-01	FALSE	psavert	8.0

date	my_era	my_variable	value
1974-02-01	TRUE	pop	213074.0

gather summary statistics

```
economics |>
  mutate(my_era = lubridate::year(date) < 1980) |>
  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	6.192141e+03
FALSE	pop	2.730622e+05
FALSE	psavert	7.337028e+00
FALSE	uempmed	9.553538e+00
FALSE	unemploy	8.671962e+03
TRUE	pce	9.417707e+02
TRUE	pop	2.122084e+05
TRUE	psavert	1.204467e+01
TRUE	uempmed	5.938000e+00
TRUE	unemploy	5.225467e+03

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

```
economics |>
  mutate(my_era = lubridate::year(date) < 1980) |>
  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	6.192141e+03	941.77067
pop	2.730622e+05	212208.40000
psavert	7.337028e+00	12.04467
uempmed	9.553538e+00	5.93800
unemploy	8.671962e+03	5225.46667

print data frame in a specific manner

```
economics |>
  mutate(my_era = lubridate::year(date) < 1980) |>
  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	6.192141e+03	941.77067
pop	2.730622e+05	212208.40000
psavert	7.337028e+00	12.04467
uempmed	9.553538e+00	5.93800
unemploy	8.671962e+03	5225.46667