

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 big 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. Inanycase, since i am familiar with `{gt}` I will show you some conceptual table editing. Youc an 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_sumarize_print.qmd`. In all cases, here and above, besure 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
1977-11-01	1327.0	221109.0	11.2	7.0	6815
1974-11-01	956.2	214625.0	13.8	5.2	6140
1972-10-01	794.9	210479.0	13.0	5.7	4875
2003-08-01	7834.5	291072.0	6.0	10.1	8896
2012-09-01	11061.5	314646.7	8.2	18.8	12115

date	pce	pop	psavert	uempmed	unemploy
1985-08-01	2752.1	238679.0	8.2	7.1	8196
1976-09-01	1168.8	218440.0	11.4	7.8	7380
2006-02-01	9090.1	297854.0	4.2	9.1	7184
2014-11-01	12051.4	319564.2	7.3	13.0	9090
2009-12-01	9998.9	308633.0	5.9	20.1	15098
1997-08-01	5587.0	273237.0	6.0	7.8	6608
1981-02-01	1884.2	229071.0	10.8	7.1	8051
1982-12-01	2161.6	233160.0	10.9	10.2	12051
1998-09-01	5965.6	276714.0	6.4	6.7	6300
1984-05-01	2474.5	235993.0	11.1	9.1	8456
2003-01-01	7533.1	289518.0	5.5	9.6	8520
1997-02-01	5434.0	271585.0	6.2	8.1	7102
2003-06-01	7678.6	290584.0	5.6	11.5	9266
2007-05-01	9651.3	301483.0	4.0	8.2	6766
1991-11-01	4003.6	254718.0	9.0	7.5	8931
1994-11-01	4825.6	264559.0	7.0	9.0	7375
2015-04-01	12193.8	320402.3	7.6	11.5	8526
1990-03-01	3754.9	249012.0	8.3	5.1	6598
2009-01-01	9783.8	306208.0	6.2	10.7	12058
2015-01-01	12046.0	319928.6	7.7	13.2	8903

pivot longer so that I can 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")
```

date	my_era	my_variable	value
1987-05-01	FALSE	pce	3048.4
1987-05-01	FALSE	pop	242423.0
1987-05-01	FALSE	psavert	8.2
1987-05-01	FALSE	uempmed	6.6
1987-05-01	FALSE	unemploy	7574.0
2014-02-01	FALSE	pce	11566.2
2014-02-01	FALSE	pop	317753.9
2014-02-01	FALSE	psavert	7.3
2014-02-01	FALSE	uempmed	15.9
2014-02-01	FALSE	unemploy	10349.0

date	my_era	my_variable	value
1993-02-01	FALSE	pce	4355.3
1993-02-01	FALSE	pop	258919.0
1993-02-01	FALSE	psavert	8.9
1993-02-01	FALSE	uempmed	8.5
1993-02-01	FALSE	unemploy	9183.0
1984-06-01	TRUE	pce	2495.6
1984-06-01	TRUE	pop	236160.0
1984-06-01	TRUE	psavert	11.1
1984-06-01	TRUE	uempmed	7.5
1984-06-01	TRUE	unemploy	8226.0
1970-09-01	TRUE	pce	659.1
1970-09-01	TRUE	pop	205540.0
1970-09-01	TRUE	psavert	12.9
1970-09-01	TRUE	uempmed	5.2
1970-09-01	TRUE	unemploy	4456.0
1989-08-01	FALSE	pce	3620.6
1989-08-01	FALSE	pop	247573.0
1989-08-01	FALSE	psavert	7.6
1989-08-01	FALSE	uempmed	5.0
1989-08-01	FALSE	unemploy	6511.0
1978-09-01	TRUE	pce	1452.9
1978-09-01	TRUE	pop	223053.0
1978-09-01	TRUE	psavert	10.6
1978-09-01	TRUE	uempmed	5.6
1978-09-01	TRUE	unemploy	6125.0
1974-02-01	TRUE	pce	889.7
1974-02-01	TRUE	pop	213074.0
1974-02-01	TRUE	psavert	14.2
1974-02-01	TRUE	uempmed	5.1
1974-02-01	TRUE	unemploy	4731.0
1983-03-01	TRUE	pce	2202.8
1983-03-01	TRUE	pop	233613.0
1983-03-01	TRUE	psavert	10.6
1983-03-01	TRUE	uempmed	10.4
1983-03-01	TRUE	unemploy	11408.0
1996-12-01	FALSE	pce	5378.6
1996-12-01	FALSE	pop	271125.0
1996-12-01	FALSE	psavert	6.4
1996-12-01	FALSE	uempmed	7.8
1996-12-01	FALSE	unemploy	7253.0
2009-02-01	FALSE	pce	9766.0

date	my_era	my_variable	value
2009-02-01	FALSE	pop	306402.0
2009-02-01	FALSE	psavert	5.5
2009-02-01	FALSE	uempmed	11.7
2009-02-01	FALSE	unemploy	12898.0
1972-12-01	TRUE	pce	806.1
1972-12-01	TRUE	pop	210821.0
1972-12-01	TRUE	psavert	13.7
1972-12-01	TRUE	uempmed	6.1
1972-12-01	TRUE	unemploy	4543.0
1970-05-01	TRUE	pce	642.4
1970-05-01	TRUE	pop	204607.0
1970-05-01	TRUE	psavert	12.4
1970-05-01	TRUE	uempmed	4.7
1970-05-01	TRUE	unemploy	3919.0
1978-06-01	TRUE	pce	1425.8
1978-06-01	TRUE	pop	222379.0
1978-06-01	TRUE	psavert	10.0
1978-06-01	TRUE	uempmed	6.0
1978-06-01	TRUE	unemploy	6028.0
1995-08-01	FALSE	pce	5005.3
1995-08-01	FALSE	pop	266843.0
1995-08-01	FALSE	psavert	6.7
1995-08-01	FALSE	uempmed	8.3
1995-08-01	FALSE	unemploy	7484.0
1998-12-01	FALSE	pce	6070.5
1998-12-01	FALSE	pop	277526.0
1998-12-01	FALSE	psavert	5.8
1998-12-01	FALSE	uempmed	6.8
1998-12-01	FALSE	unemploy	6032.0
1990-08-01	FALSE	pce	3848.3
1990-08-01	FALSE	pop	250439.0
1990-08-01	FALSE	psavert	8.1
1990-08-01	FALSE	uempmed	5.4
1990-08-01	FALSE	unemploy	7188.0
1975-01-01	TRUE	pce	975.6
1975-01-01	TRUE	pop	214931.0
1975-01-01	TRUE	psavert	13.2
1975-01-01	TRUE	uempmed	6.3
1975-01-01	TRUE	unemploy	7501.0
2005-11-01	FALSE	pce	8938.5
2005-11-01	FALSE	pop	297207.0

date	my_era	my_variable	value
2005-11-01	FALSE	psavert	3.5
2005-11-01	FALSE	uempmed	8.5
2005-11-01	FALSE	unemploy	7566.0
2005-08-01	FALSE	pce	8832.4
2005-08-01	FALSE	pop	296440.0
2005-08-01	FALSE	psavert	2.7
2005-08-01	FALSE	uempmed	9.2
2005-08-01	FALSE	unemploy	7345.0
1998-01-01	FALSE	pce	5689.9
1998-01-01	FALSE	pop	274626.0
1998-01-01	FALSE	psavert	7.4
1998-01-01	FALSE	uempmed	7.4
1998-01-01	FALSE	unemploy	6368.0
1972-09-01	TRUE	pce	781.1
1972-09-01	TRUE	pop	210278.0
1972-09-01	TRUE	psavert	12.2
1972-09-01	TRUE	uempmed	5.6
1972-09-01	TRUE	unemploy	4849.0
1992-02-01	FALSE	pce	4099.5
1992-02-01	FALSE	pop	255448.0
1992-02-01	FALSE	psavert	9.8
1992-02-01	FALSE	uempmed	8.2
1992-02-01	FALSE	unemploy	9454.0
2010-04-01	FALSE	pce	10112.9
2010-04-01	FALSE	pop	309191.2
2010-04-01	FALSE	psavert	6.4
2010-04-01	FALSE	uempmed	22.1
2010-04-01	FALSE	unemploy	15325.0
2006-07-01	FALSE	pce	9322.6
2006-07-01	FALSE	pop	298996.0
2006-07-01	FALSE	psavert	3.4
2006-07-01	FALSE	uempmed	8.0
2006-07-01	FALSE	unemploy	7175.0

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	7.618344e+03
FALSE	pop	2.864290e+05
FALSE	psavert	6.994444e+00
FALSE	uempmed	1.160000e+01
FALSE	unemploy	9.419167e+03
TRUE	pce	1.536900e+03
TRUE	pop	2.225169e+05
TRUE	psavert	1.142857e+01
TRUE	uempmed	6.871429e+00
TRUE	unemploy	7.220143e+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	7.881439e+03	1.394986e+03
pop	2.896725e+05	2.183624e+05
psavert	6.422222e+00	1.155714e+01
uempmed	1.187778e+01	6.257143e+00

my_variable	FALSE	TRUE
unemploy	9.561833e+03	6.156857e+03

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	7.426306e+03	1384.9625
pop	2.850551e+05	218630.5000
psavert	6.558823e+00	11.0750
uempmed	1.067059e+01	6.0625
unemploy	8.495118e+03	6243.1250