

# Writing reproducible manuscripts in R Markdown and Pagedown

Camille Landesvatter

Mannheim Centre for European Research (MZES), University of Mannheim

✉ [camille.landesvatter@uni-mannheim.de](mailto:camille.landesvatter@uni-mannheim.de)

🌐 [MZES Website](#)

👤 Camille Landesvatter

Dr. Paul C. Bauer

Mannheim Centre for European Research (MZES), University of Mannheim

✉ [mail@paulcbauer.de](mailto:mail@paulcbauer.de)

🌐 [MZES Website](#)

👤 Paul C. Bauer

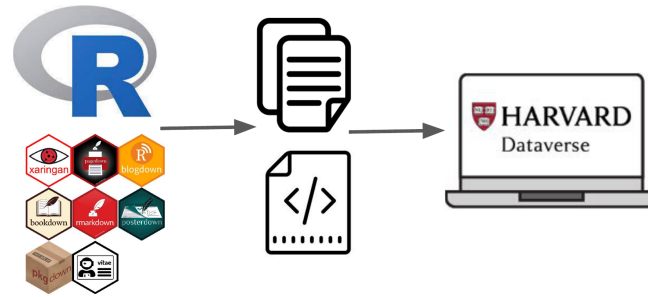
Mannheim Open Science Day Reprohack

Oktober 11, 2022



# Why reproducible research?

- **Reproducibility:** Everyone (including yourself) can trace the steps that led to a particular result or statement. This increases transparency (and credibility).
- **Access:** Open access is given when research is freely available to everyone without financial, legal, or technical barriers (UNESCO 2022), for example commercial software. E.g.,



- **Errors:** Everyone (including yourself) can spot errors. Plus, automation (e.g., code that produces findings is tied to the document itself) avoids a slow and error-prone cut-and-paste approach.
- **Revisions:** Revision of manuscripts and other pieces of research gets easier when steps are reproducible and automated. Plus, easy version control with git, etc.

# Tools for reproducible manuscripts in R

- **rmarkdown**

- combines R code and markdown-formatted text to produce HTML, PDF or Word documents

- **pagedown** (*Xie et al. 2022*)

- R Markdown based
- creates paginated HTML reports and prints to PDF
- this includes headers, footers, etc.
- highly customizable due to HTML and CSS styling sheets (YAML Metadata)
- many available templates, e.g. CV, thesis, letter, journal templates

- Other (R Markdown based): posterdown, bookdown, xaringan, etc.



Source: Indrajeet Patil on Twitter

# Two templates (on OSF, Gdrive and Github)

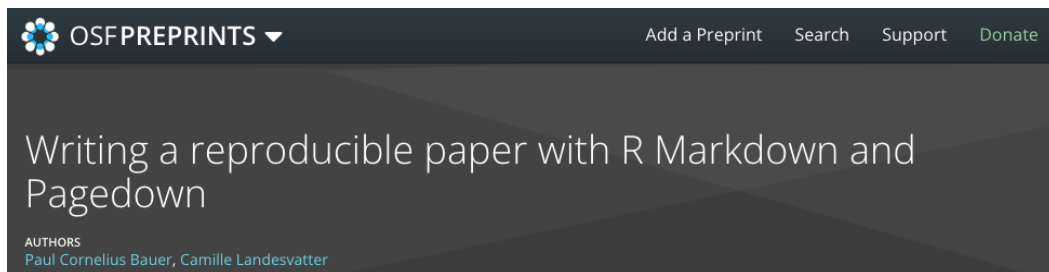
Bauer, P. 2021:



: [/paulcbauer/Writing\\_a\\_reproducible\\_paper\\_in\\_rmarkdown](https://github.com/paulcbauer/Writing_a_reproducible_paper_in_rmarkdown)

: [/2021\\_Writing\\_a\\_reproducible\\_paper\\_in\\_rmarkdown](https://drive.google.com/file/d/12021_Writing_a_reproducible_paper_in_rmarkdown/view)

Bauer, P.; Landesvatter, C. 2022:



: [/paulcbauer/Writing\\_a\\_reproducible\\_paper\\_in\\_pagedown](https://github.com/paulcbauer/Writing_a_reproducible_paper_in_pagedown)

: [/2021\\_Writing\\_a\\_reproducible\\_paper\\_in\\_pagedown](https://drive.google.com/file/d/12021_Writing_a_reproducible_paper_in_pagedown/view)

How can you use these templates?

- download or clone all available files:
  - 'paper.rmd'
  - 'references.bib'
  - 'data.csv'
  - 'american-sociological-association.csl'
  - styling files for pagedown:  
'wp\_paged.html', 'wp.css', 'wp-fonts.css',  
'wp-pages.css'
- open .Rproj
- start creating your first manuscript in rmd, use our working paper (wp) style or modify the styling sheets

# Pagedown: Working Paper (wp) Style

## Writing a reproducible paper with R Markdown and Pagedown<sup>2</sup>

Paul C. Bauer  
Mannheim Centre for European Social  
Research  
`mail@paulcbauer.eu`

Camille Landesvatter  
Mannheim Centre for European Social  
Research  
`camille.landesvatter@mzes.uni-  
mannheim.de`

First version: 20 June, 2021  
This version: 04 Oktober, 2022

Download: `https://osf.io/k8jhx`

---

### Abstract

The present paper provides a template to write a reproducible scientific paper with R Markdown and Pagedown.<sup>1</sup> Below we outline some of the “tricks”/code (e.g., referencing tables, sections etc.) we had to figure out to produce this document. The underlying files which produce this document can be `downloaded here`. Importantly, we also provide different CSS and HTML files that can be used to achieve a pdf output with the look of a “working paper”. We are convinced that in the future there will be many improvements and developments with regards to RStudio, R markdown and Pagedown. We intend to update this file when we discover more convenient code. You can follow any updates on the `github repository`.

*Keywords:* open science, transparency, replication, reproducible research, reproducibility, R, markdown, pagedown.

---

# Pre-Requisites for R Markdown and Pagedown

Both require: installation of R + recommended: Rstudio IDE

- R Markdown:
  - rmarkdown package + other dependencies, e.g. knitr package
  - for working without Rstudio: Pandoc Installation
  - for PDF output: a LaTeX distribution, e.g. tinytex (a lightweight, cross-platform LaTeX distribution)

```
install.packages(c('tinytex', 'rmarkdown'))  
tinytex::install_tinytex()
```

- Pagedown:
  - rmarkdown package + other dependencies, e.g. knitr
  - pagedown package
  - no LaTeX!

```
remotes::install_github('rstudio/pagedown')
```

# Three Examples

---

1. Inline code & results

2. Figures

3. Tables

# 1. Inline code & results

## Example 1:

Writing ``r 3 + 7`` (use back ticks!) will calculate and insert 10 into your text.

## Example 2:

Imagine you want to automate the display of regression coefficients within your manuscript.

```
data ← read.csv("data.csv")
```

```
M1 ← lm(Fertility ~ Education + Agriculture, data = data)
```

Now we can write some plain text with inline code: The coefficient of `Education` of the model we ran above is -0.96. -0.96 was printed by writing: ``r round(coef(M1)[2], 2)`` (use back ticks!).

*What advantages does such a workflow bring?*



## 2. Figures

Example: You want to create, edit and insert a ggplot2 graph alongside its code and the text.

```
data$Education_cat <- cut(data$Education,  
                           breaks=c(-Inf, 6, 12, Inf), labels=c("low", "middle", "high"))
```

```
```{r fig-1, fig.cap="Fig 1. Fertility and Agriculture by Education", fig.align="center",  
fig.width=6, fig.height=3, eval=T, echo=T}  
ggplot(data, aes(x=Catholic, y=Fertility, shape=Education_cat)) + geom_point() +  
  labs(x="Agriculture", y = "Fertility", shape="Education") + theme_classic()  
```
```

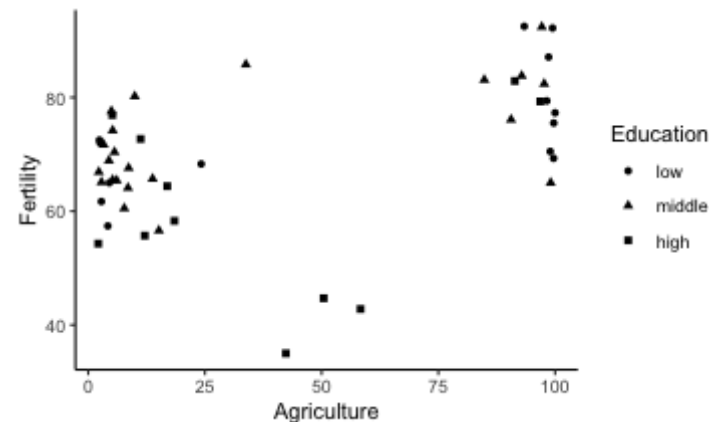


Fig 1. Fertility and Agriculture by Education

## 2. Figures

- For cross-referencing within your manuscript you can refer to Figure 1 by using `\@ref(fig:fig-1)`.
- But pay attention to your code chunk options. In the previous example:

```
{r fig-1, fig.cap="Fig 1. Fertility and Agriculture by Education",  
fig.align="center", fig.width=6, fig.height=3, eval=T, echo=T}
```

*What do the different code chunk options mean?*

# 3. Tables

Example 1: You want to show summary statistics of your data.

```
library(stargazer)
stargazer(data,
  title = "Summary table with stargaz",
  label="tab1",
  header=FALSE,
  table.placement = "H",
  type="html",
  font.size = "small")
```

**Summary table with stargazer**

| Statistic        | N  | Mean   | St. Dev. | Min    | Max     |
|------------------|----|--------|----------|--------|---------|
| Fertility        | 47 | 70.143 | 12.492   | 35.000 | 92.500  |
| Agriculture      | 47 | 50.660 | 22.711   | 1.200  | 89.700  |
| Examination      | 47 | 16.489 | 7.978    | 3      | 37      |
| Education        | 47 | 10.979 | 9.615    | 1      | 53      |
| Catholic         | 47 | 41.144 | 41.705   | 2.150  | 100.000 |
| Infant.Mortality | 47 | 19.943 | 2.913    | 10.800 | 26.600  |

### 3. Tables

Example 1b: You want to show summary statistics of a subset of your data.

```
library(stargazer)
stargazer(data %>% filter(Education ≥ 15),
  title = "Summary table with stargaz",
  label="tab1",
  header=FALSE,
  table.placement = "H",
  type="html",
  font.size = "small")
```

| Summary table with stargazer |         |        |          |        |     |
|------------------------------|---------|--------|----------|--------|-----|
| Statistic                    | N       | Mean   | St. Dev. | Min    | Max |
| Fertility                    | 854.013 | 13.239 | 35.000   | 76.900 |     |
| Agriculture                  | 824.750 | 14.964 | 1.200    | 46.600 |     |
| Examination                  | 826.125 | 7.791  | 16       | 37     |     |
| Education                    | 828.125 | 11.692 | 15       | 53     |     |
| Catholic                     | 825.738 | 21.527 | 2.150    | 58.330 |     |
| Infant.Mortality             | 818.875 | 3.634  | 10.800   | 23.000 |     |

# 3. Tables

```
```{r table-1, eval=T, echo=T, results='hide'}
M1 <- lm(Fertility ~ Education + Agriculture, data = data)
M2 <- lm(Fertility ~ Education + Catholic, data = data)
models <- list("M1" = M1, "M2" = M2)

library(gt)
library(modelsummary)

modelsummary(models,
  title = 'Table 1. Linear regression',
  output = 'gt',
  notes = "Notes: some notes ..." ) %>%
  tab_spanner(label = 'Dependent variable: Fertility', columns = 2:3) %>%
  tab_options(
    table.font.size = 16, data_row.padding = px(1),
    table.border.top.color = "white", heading.border.bottom.color = "black",
    row_group.border.top.color = "black",
    row_group.border.bottom.color = "white", table.border.bottom.color = "white",
    column_labels.border.top.color = "black", column_labels.border.bottom.color = "black",
    table_body.hlines.color = "white"
  )
```
```

- For cross-referencing within your manuscript you can refer to Table 1 by using `\@ref(tab:table-1)` (use table-1 as the chunk name!).

| Table 1. Linear regression |                               |         |
|----------------------------|-------------------------------|---------|
|                            | Dependent variable: Fertility |         |
|                            | M1                            | M2      |
| (Intercept)                | 84.080                        | 74.234  |
|                            | (5.782)                       | (2.352) |
| Education                  | -0.963                        | -0.788  |
|                            | (0.189)                       | (0.129) |
| Agriculture                | -0.066                        |         |
|                            | (0.080)                       |         |
| Catholic                   |                               | 0.111   |
|                            |                               | (0.030) |
| Num.Obs.                   | 47                            | 47      |
| R2                         | 0.449                         | 0.575   |
| R2 Adj.                    | 0.424                         | 0.555   |
| AIC                        | 349.7                         | 337.6   |
| BIC                        | 357.1                         | 345.0   |
| RMSE                       | 9.17                          | 8.06    |

Notes: some notes...

# Xaringans' Infinite Moon Reader

- xaringan package (Xie 2021) includes R Studio add-in Infinite Moon Reader
- preview live changes while editing your files or even manuscript
- avoids manually knitting your document each time

```
xaringan::inf_mr()
```

- launches a local web server via the servr package (Xie 2021) and displays your pages in the RStudio viewer
- each time you save your document (Ctrl/Cmd+S) xaringan updates your pages in the viewer.

# Xaringans' Infinite Moon Reader

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for setting up the Infinite Moon Reader. The code includes comments about the package's features and the `xaringan::inf_mr()` function call.
- Console:** Shows the output of the `serve` command, indicating that a local web server is running at `http://127.0.0.1:4321/mzes-xaringan-metropolis.html`.
- Environment:** Lists the loaded packages: `magick` and `pkgs`.
- Preview:** Displays the generated HTML slide titled "Pagedowns' Infinite Moon Reader". The slide content includes a list of features and the `xaringan::inf_mr()` function call.

```
320
321
322 ---
323 # Pagedowns' Infinite Moon Reader
324
325 * xaringan package (Xie 2021) includes R Studio add-in Infinite Moon Reader
326 * preview live changes while editing your manuscript
327 * avoids manually knitting your document each time
328
329 ```{r xaringan-inf-mr, eval=F, echo=T}
330 xaringan::inf_mr()
331 ```
332
333 * launches a local web server via the servr package (Xie 2021) and displays your pages in the RStudio viewer
334 * Each time you save your document (Ctrl/Cmd+S) xaringan updates your pages in the viewer.
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

```
R 4.2.0 ~ /Documents/GitHub/reproducible-manuscripts-with-RMD/
r /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/library/rmarkdown/rmarkdown/latex-div.lua --wrap pre
serve -V 'mathjax-url=https://mathjax.rstudio.com/latest/MathJax.js?config=TeX-MML-AM_HTML' -V 'title-slide-class=cent
er, middle, inverse, title-slide' --standalone --section-divs --template /Library/Frameworks/R.framework/Versions/4.2-a
rm64/Resources/library/xaringan/rmarkdown/templates/xaringan/resources/default.html --no-highlight --css jobtalk_ks_mth
eme.css --css jobtalk_ks_fonts_mtheme.css --include-in-header /var/folders/zm/d2gy0g_x16z6d1jbkjpjpmf880000gn/T//RtmpgM4Lk6/rmarkdown-str1223815e465cc.html --include-before-body /var/folders/zm/d2gy0g_x16z6d1jbkjpjpmf880000gn/T//RtmpgM4Lk6/
xaringan1223815e465cc.md --include-after-body /var/folders/zm/d2gy0g_x16z6d1jbkjpjpmf880000gn/T//RtmpgM4Lk6/xaringan1223
876b8ae5d.js --include-after-body insert-logo.html --variable title-slide=true --variable math=true

Output created: mzes-xaringan-metropolis.html
To stop the server, run servr::daemon_stop(1) or restart your R session
Serving the directory /Users/camillelandesvatter/Documents/GitHub/reproducible-manuscripts-with-RMD at http://127.0.0.
1:4321/mzes-xaringan-metropolis.html
```

**Pagedowns' Infinite Moon Reader**

- xaringan package (Xie 2021) includes R Studio add-in Infinite Moon Reader
- preview live changes while editing your manuscript
- avoids manually knitting your document each time

```
xaringan::inf_mr()
```

- launches a local web server via the `servr` package (Xie 2021) and displays your pages in the RStudio viewer
- Each time you save your document (Ctrl/Cmd+S) xaringan updates your pages in the viewer.

Slide 12 of 13

# R trackdown package for improving collaborative writing

 [ClaudioZandonella/trackdown](https://github.com/ClaudioZandonella/trackdown)

- collaborative workflow for the writing and editing process with R Markdown
- upload local .Rmd (or .Rnw) file as a plain-text file to Google Drive
- use Markdown (or LaTeX) syntax and Google Docs options (e.g. suggest edits, add comments)
- review and integrate all contributions
- the final document can be downloaded and rendered locally



# R trackdown package for improving collaborative writing

```
remotes::install_github("claudiozandonella/trackdown",  
                        build_vignettes = TRUE)
```

```
upload_file(file = "trackdown_example.Rmd",  
            gpath = "trackdown_example",  
            hide_code = TRUE)
```

- `download_file()`
- `update_file()`
- `render_file()`

# Thank you!

---

: `clandesv/reproducible-manuscripts-with-RMD`

Xaringan Presentation on GitHub Pages: <https://clandesv.github.io/reproducible-manuscripts-with-RMD/mzes-xaringan-metropolis.html>

R Markdown Template on OSF: <https://osf.io/q395s>

R Markdown and Pagedown Template on OSF: <https://osf.io/k8jhx/>