Your Substantive Title Here

Your Name1* Your Name2†

2025-01-02

Abstract

This is our informative abstract of fewer than 250 words. It describes what we investigate, how we investigate it, and what we find.

1 Introduction

In this section, we introduce the reader to the phenomenon we investigate. We describe the way in which our analysis contributes to an important intellectual debate, or how it answers a pressing political or social question. We introduce our research question, hypotheses, data, and results. We signpost for the reader what's coming in the rest of the paper.

We remember that our paper is not a mystery novel. We note our core results early and often.

Throughout our paper, we use active, first-person language and avoid the passive voice. For example, we write "we examine the relationship between X and Y"; we do not write "the relationship between X and Y was examined." Where we do the analysis, we speak about it transparently. We use the present tense; for example, "In this paper, we argue …" and "Paper XYZ demonstrates the relationship between …".

2 [Our Substance and Context Section Title Here]

Here we go deeper into the intellectual debate, the political and social context of our investigation. To give the reader a clear sense of why we are writing this paper, we describe the relevant scholarly, technical, or popular literature. We give this section a meaningful *substantive* title; it is not entitled "Literature Review", for example. We cite at least five published, peer-reviewed scholarly works, as long as they are relevant. For example, we could cite Moore and Reeves (2020) or Moore and Ravishankar (2012)[1], which we discussed in class.[2] We only cite others' work in our paper when it enhances the reader's understanding of what we, the authors of this paper, are doing. We connect everything we cite to *our* investigation. This is our original research, not a book report or an annotated bibliography.

We do not cite paper titles or journal names, unless our paper is about someone else's paper or about the set of articles in a journal. We do not cite authors' first, given names. We do not cite the

*American University
†American University

[1] There should always be a space before the "(" in a citation date.

[2] To cite a paper within parentheses, use, e.g., (Moore 2012).

universities or institutes with which authors are affiliated. We can refer to either what an author does, or what a paper does, but we should be consistent. For example, "Moore and Ravishankar (2012) argue that we should …" refers to what the authors do; "Moore and Ravishankar (2012) argues that we should …" refers to what the paper – Moore and Ravishankar (2012) – does.

In order to integrate citations into the References section below, we add entries into our file `main.bib`. This is a plain-text file that we edit in RStudio (or BibDesk, or similar). We store `main.bib` in the same folder as our paper's `.qmd` and `.pdf` files. Its entries are formatted so that they can be knit to `.pdf`; see https://j.mp/2UzTXEZ for example entries for articles, books, and miscellaneous. We can get these entries automatically from Google Scholar by turning on BibTeX in the Google Scholar Settings - Bibliography Manager. Perhaps we use a tool like free, open-source BibDesk to help us manage the `.bib` file.

3 Data and Methods

This section describes the data we analyze. We describe the source of the data, and its primary features. We cite our data. We describe the methods we use to answer our question and to test our hypotheses.

If our data were `cars`, loaded in a chunk above, we note that our data have 50 observations. Our unit of analysis is the cars; each row represents a different car that was measured.

We refer to concepts and label them appropriately. We state our outcome and how it is measured: "Our outcome is stopping distance, measured in feet." We state our key predictor and how it is measured: "Our key predictor is the car's speed, measured in miles per hour."

We almost never refer to specific variable, object, function, or data frame names (such as `var_x`, `ourdata`, `this_useful_func`, or `df`). These particular names are almost never of interest or use to the reader.

We explain important decisions and codings in this section. "We collect our data from source X. We code the outcome as 1 if the registrant turned out in the 2022 election, 0 if they did not." A table can serve as an efficient way to detail several such decisions.

Where there are less-critical details that we implement to improve our analysis or presentation, we do not explain them in the paper. Our paper does not say, for example, "we reorder the levels of the factor variable from alphabetical 'high, low, medium' to the sensible 'low, medium, high.' " Of course, we implement that reordering, but it should not appear in our paper.

We cite the software we use. For example, we conduct our analysis using R version 4.4.2 (R Core Team 2024). We rely on several elements of the `tidyverse` (Wickham et al. 2019). Of course, we do not cite software that we do not use.

4 [Our Results Section Title Here]

Here, we explain and interpret our results. We try to learn as much as we can about our question as possible, given the data and analysis. We present our results clearly. We interpret them for the reader with precision and circumspection. We avoid making claims that are not substantiated by our data. We are careful about causality. When we describe associations, we avoid language like "effects" and "increases"; we only describe "effects" or "impacts" when we have a causally well-identified research design.

Note that this section may be integrated into Section 3, if joining the two improves the overall presentation.

4.1 Predicting Distance with Speed

Our results for the `cars` data include estimating the linear model

$$\text{Distance}_i = \beta_0 + \beta_1(\text{Speed}_i) + \epsilon_i.$$

Perhaps we start by plotting the data, as in Figure 1.

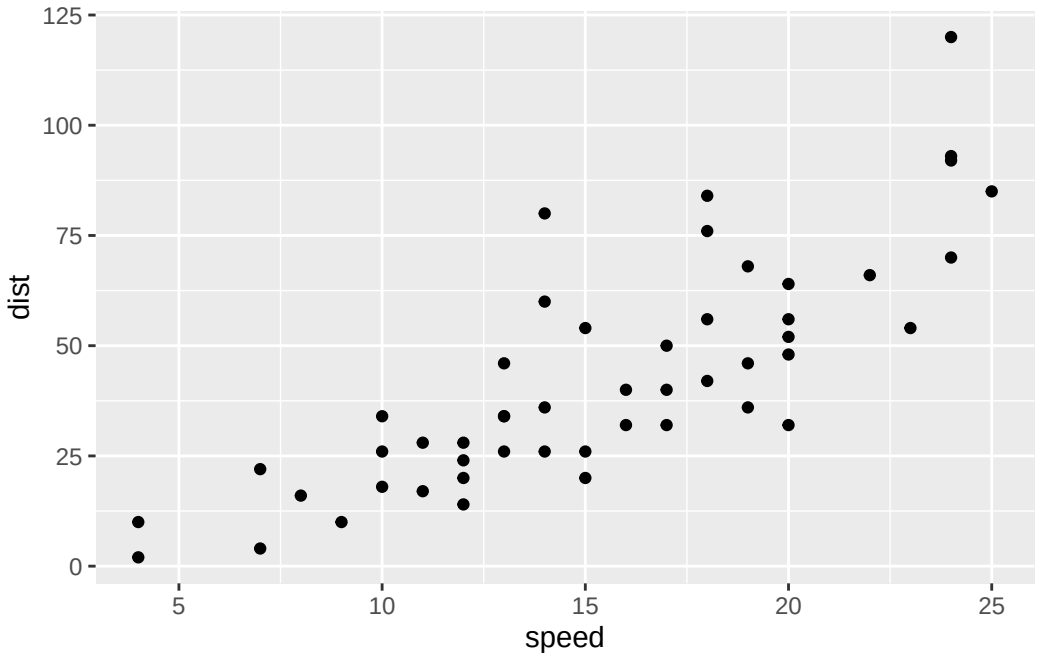

Figure 1: Distance on Speed

The data may be roughly linear, though there may be some non-linearity we should incorporate.

Below we show the model estimates. The first table uses `xtable()`, the second uses `stargazer()` (Hlavac 2018).

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-17.58	6.76	-2.60	0.01
speed	3.93	0.42	9.46	0.00

Table 1: Our Informative Caption

Using the `cars` data, we find that each unit of speed is associated with 3.9 more units of distance. We draw out what this really means, and what it implies. For example, if a typical difference among our observations is 7 units of speed, then our model estimates that a typical difference in distance among our observations is $7 \times 3.9 = 27.3$ units of distance. We describe the substantive relevance of this number.

Table 2: Our Informative Title

	Outcome
	dist
speed	3.93***
	(0.42)
Constant	−17.58**
	(6.76)
Observations	50
R^2	0.65
Adjusted R^2	0.64
Residual Std. Error	15.38 (df = 48)
F Statistic	89.57*** (df = 1; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

We do not report estimates like `p = 3.242e-15`, since these are computational zeros. Instead, we write $p < 0.001$ or $p \approx 0$, as appropriate.

We do not report quantities to unhelpful degrees of precision. Although there were 112,030,874 votes cast from voting-eligible population of 242,690,810 in the U.S. in 2022, it is not helpful to report turnout as 46.1619762%; writing 46.2% suffices.

4.2 Comparing Distances between High- and Low-Speed Cars

To report the results of a t-test, we do so in text, and perhaps in a well-formatted table as well, such as Table 3. Here, as above, we report the important details in text. For example, when we define "high-speed" cars as those traveling above the mean speed, the difference between the high-speed and low-speed group means is 32.25, with a 95% confidence interval that covers $(20.6, 43.9)$.

Table 3: Distance by Speed Group

Parameter	Group	hs = 0	hs = 1	Difference	95% CI	t(42.87)	p
dist	hs	27.50	59.75	-32.25	(-43.90, -20.60)	-5.58	< .001

If I have tests of two outcomes from the same data, I can bind them together, as in Table 4:

Table 4: Distance and Square-root Distance by Speed Group

Parameter	Group	hs = 0	hs = 1	Difference	95% CI	t	df	p
dist	hs	27.50	59.75	-32.25	(-43.90, -20.60)	-5.58	42.87	< .001
dist_sqrt	hs	4.99	7.60	-2.61	(-3.49, -1.74)	-5.99	47.82	< .001

If we have trouble formatting using `{parameters}`, we can use `kable()` for one test, as in Table 5, or two tests as in Table 6:

Table 5: Distance by Speed Group using `kable`

Diff in Means	Group 0	Group 1	t-statistic	p-value	df	CI Lower	CI Upper
-32.25	27.5	59.75	-5.58	0	42.87	-43.9	-20.6

Table 6: Distance and Sqrt Distance by Speed Group using `kable`

Outcome	Diff Means	Slow	Fast	t-stat	p-value	df	CI Lower	CI Upper
Dist	-32.25	27.50	59.75	-5.58	0	42.87	-43.90	-20.60
Dist Sqrt	-2.61	4.99	7.60	-5.99	0	47.82	-3.49	-1.74

5 Discussion

We remind the reader what this paper was about, why it was important, and what we found. We reflect on limitations of the data or methods. If we have specific advice for someone picking up where we leave off, we provide that guidance. We avoid making trite statements like "more research should be done".

References

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.

Moore, Ryan T. 2012. "Multivariate Continuous Blocking to Improve Political Science Experiments." *Political Analysis* 20 (4): 460–79. https://doi.org/10.1093/pan/mps025.

Moore, Ryan T., and Nirmala Ravishankar. 2012. "Who Loses in Direct Democracy?" *Social Science Research* 41 (3): 646–56. https://doi.org/10.1016/j.ssresearch.2011.10.003.

Moore, Ryan T., and Andrew Reeves. 2020. "Defining Racial and Ethnic Context with Geolocation Data." *Political Science Research and Methods* 8 (4): 780–94. https://doi.org/https://doi.org/10.1017/psrm.2020.10.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing* (version 4.4.2). Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.