

# Ideas - Updated to Control for Post Date

Sally Cochrane

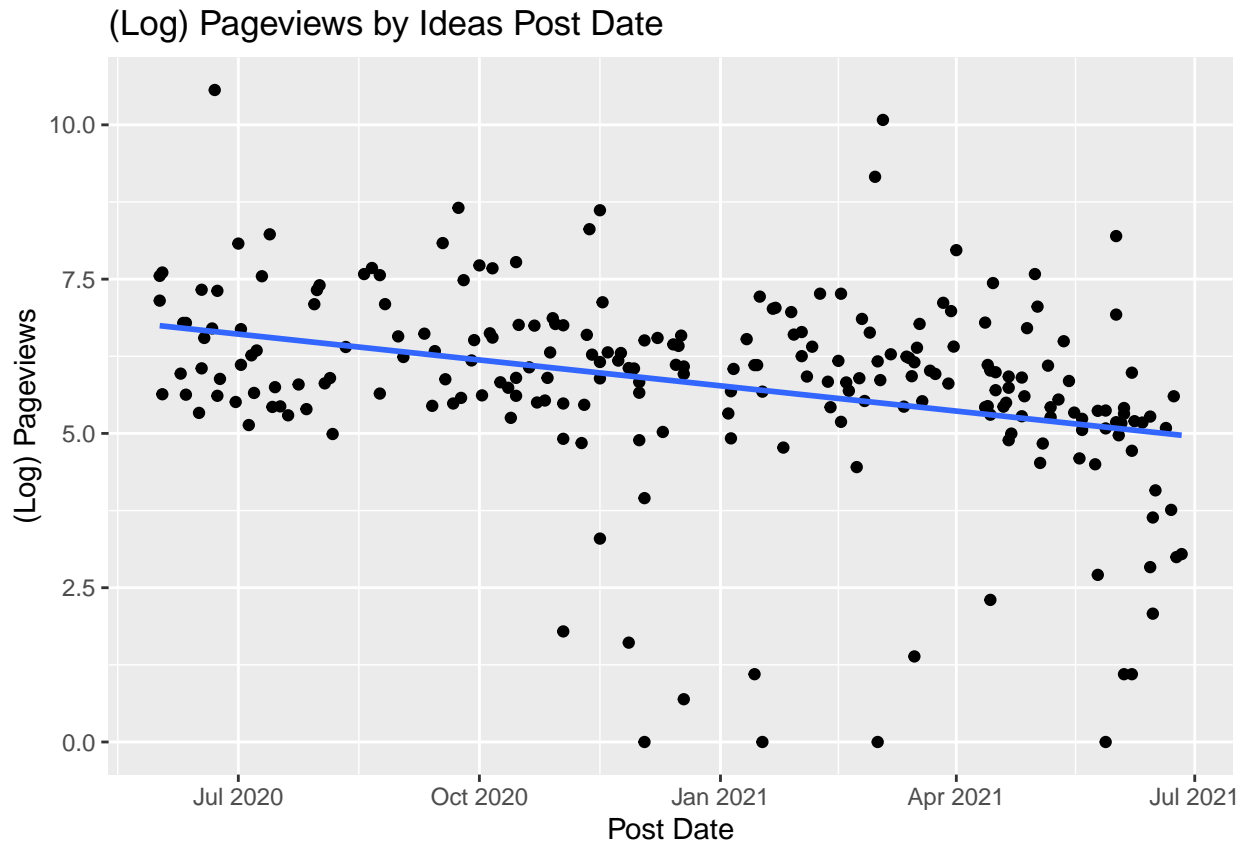
8/22/2021

## Procedure:

- Data came from Google analytics: Date range June 1 2020 - June 28 2021. Site content > content drilldown > Ideas. (Note: for some reason, this procedure returned 3 Ideas posts that were published online after June 28, 2021 (podcast-we are not born submissive; on self interest; and who was euclid. I'm not sure why this happened).
- In google analytics, there are multiple rows of data for what appear to be the same Ideas post. I combined the data from these rows as if it was one row of data, because I think it was a mistake that they were separated on google analytics. If this is wrong, let me know and I can change it.
- I made URL's from all the google analytics' "Ideas" pages rows, and used the URL's to scrape the PUP website for the content tag.
- To distinguish interviews, excerpts, and essays: I scraped the PUP website for alternating bold-face text to tag interviews, and I scraped for the word "excerpt" to tag excerpts.
- I removed rows that didn't lead to an Ideas URL (mostly strings of letters and numbers), or that went to the more general search landing pages ("podcasts", "videos", etc.).
- I scraped for the publication date of the post on the Ideas page, and limited the analysis to Ideas posts that were published between June 1 2020 and June 28 2020 (same as the google analytics data). While some older posts got views during this time, they only got a small number of views, probably due to their age. It seemed to be a fairer comparison to look at posts published within the last year.
- I took into account the time the post had been "live" using that time as a fixed effect in a model for the pageviews, bounce rate, etc.

## How Ideas Post date relates to Pageviews:

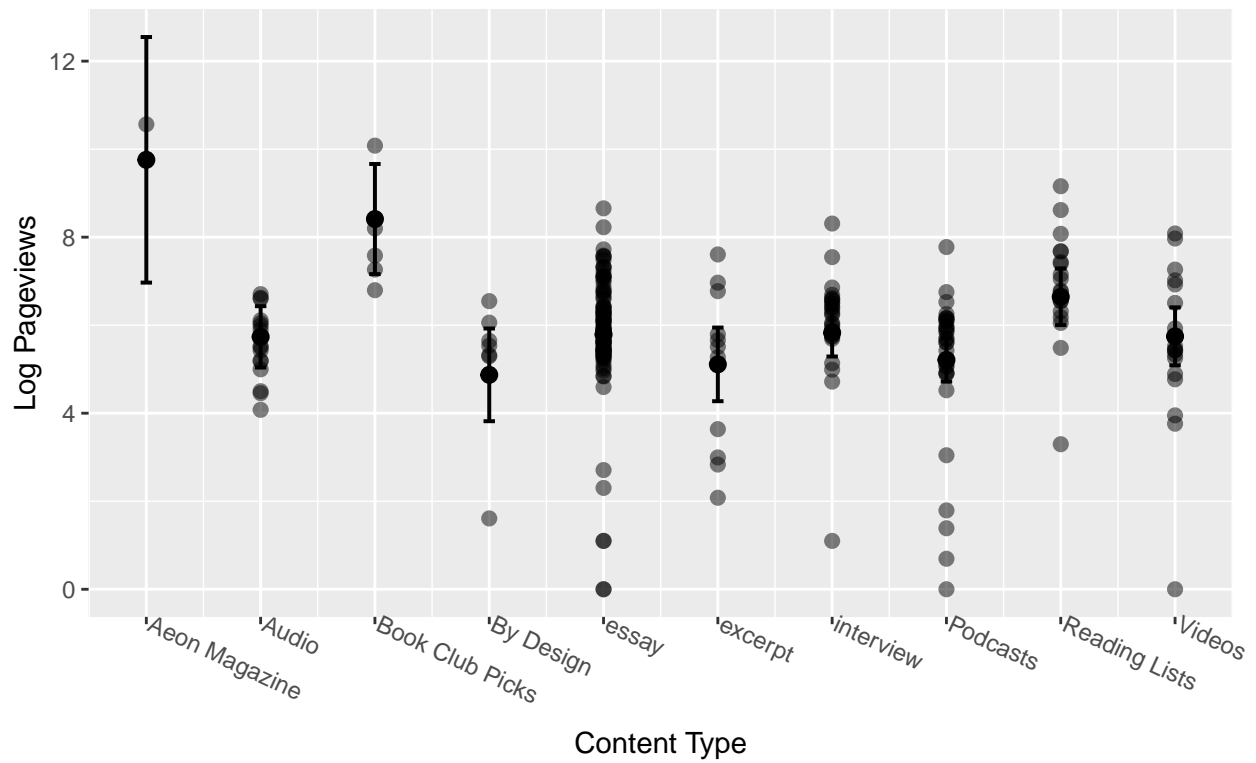
This graph shows that more recently published Ideas posts have fewer pageviews, probably because they haven't been up for as long, so we need to adjust for the "time live" to compare pageviews, etc., between posts.



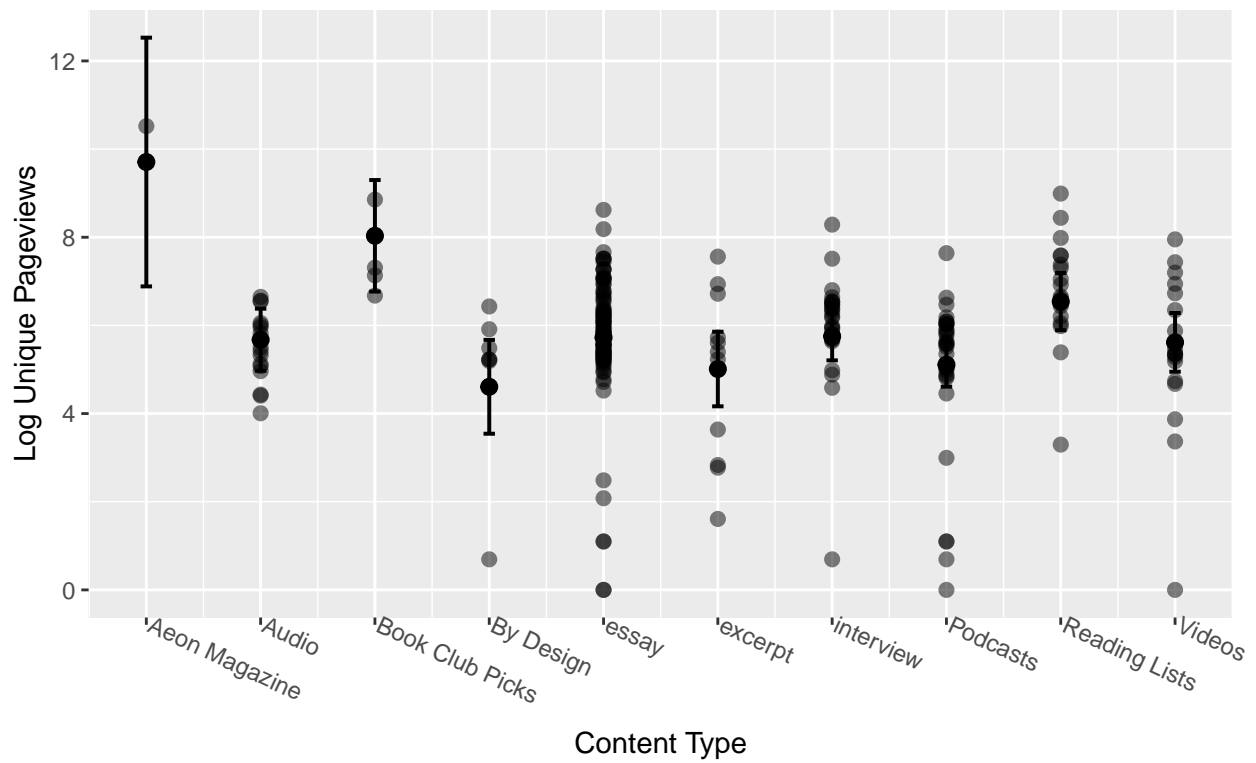
#### Plots of Performance by Content Type:

- Controlling for the amount of time the Ideas post has been “live” (published), Aeon Magazine posts and Book Club Picks posts have a higher number of pageviews and unique pageviews than the other types of content. Reading lists have a slightly higher number of pageviews than the other types of content minus Aeon and Book Club Picks (i.e., essays, interviews, excerpts, podcasts, etc). Podcasts have a slightly lower number of pageviews than the other types of content.
- However, there is only one Aeon Magazine post in this dataset, so take this with a grain of salt. There are also a lot of outliers for essays, so I think that group might warrant further exploration (i.e., what *kinds* of essays get more traffic?).

Log Pageviews by Content Type,  
with bars for prediction interval

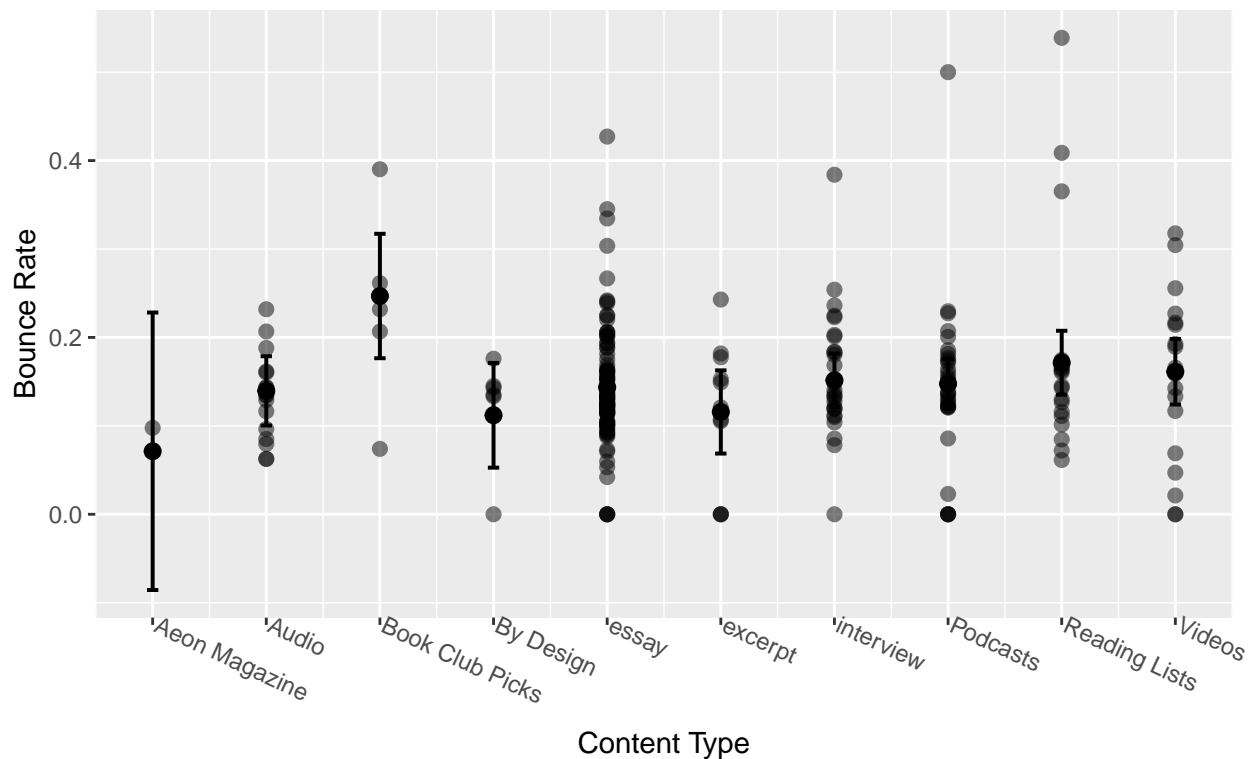


Log Unique Pageviews by Content Type,  
with bars for prediction interval



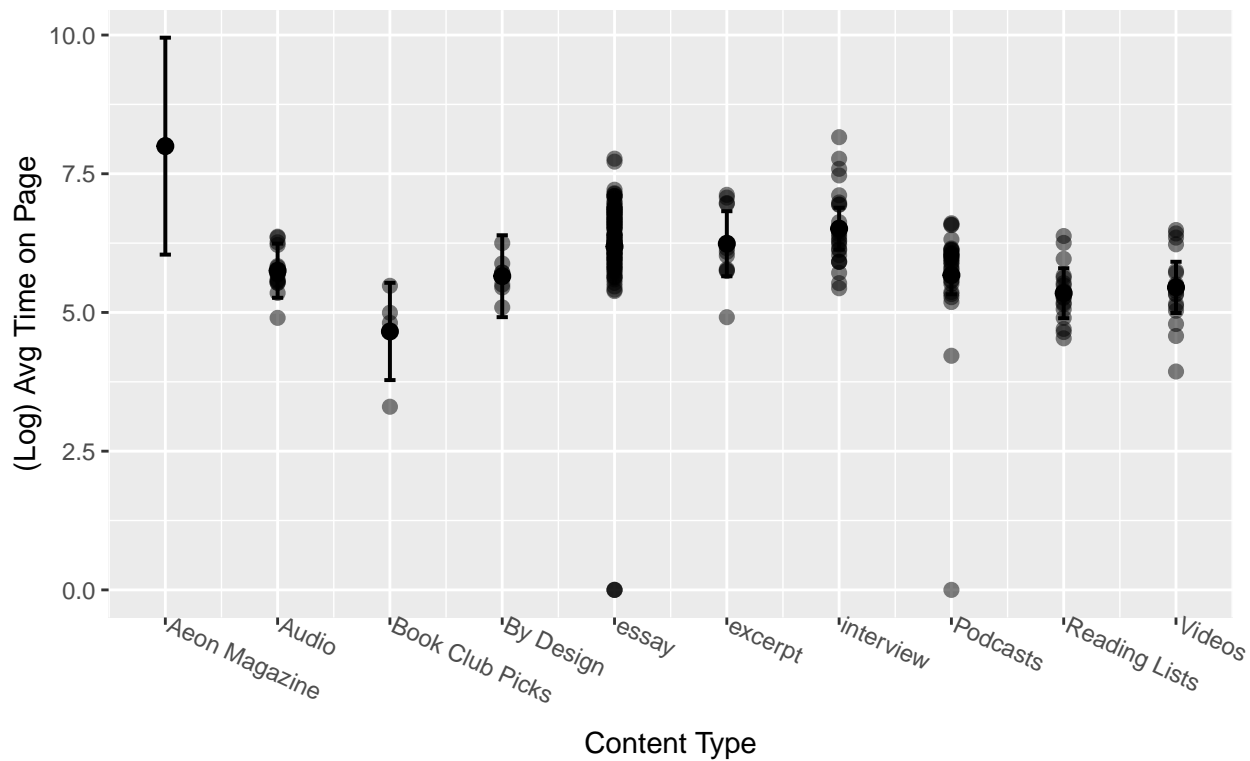
- Controlling for the amount of time the Ideas post has been live, all content types have approximately the same bounce rates, except for Book Club Picks which have slightly higher bounce rates.

Bounce Rate by Content Type,  
with bars for prediction interval



- Book Club Picks had the lowest average time on page. Reading Lists, Videos, and By Design had similarly low average time on page to Book Club Picks. Aeon Magazine, essays, excerpts, and interviews had the most time on the page, followed by Podcasts and Audio.

(Log) Average Time on Page by Content Type,  
with bars for prediction interval



- Aeon Magazine posts, essays, excerpts, and interviews also have the highest percent exit rates, meaning people leave the PUP website after visiting those pages. Book Club Picks, Reading Lists, Podcasts, and By Design have the lowest percent exit rate, meaning people stay on the PUP website after viewing them more often:

