Analyzing Book Checkout Trends at the Seattle Public Library: A 5-Year Data Study (2018-2022)

Sarvesh Fotedar, Ekrem Kizilkaya, Abbas Shaikh, Cody VanZandt, Andy Wang

February 22, 2023

Contents

1   Introduction

We present an preliminary examination of 11,749,255 checkouts from the Seattle Public Library system over a five-year period from 2018 to 2022. We intend to use this data – in concert with book review data from Goodreads – to examine Seattle's literary tastes and reading habits. Animating this analysis are more general questions about the contours of contemporary American readership. Why do some books enjoy widespread acclaim while others fizzle? Do popular books explode onto the scene or accumulate readers more gradually? How do popular success and critical acclaim differ? How has recent publisher conglomeration of publishers changed the literary marketplace? And, ultimately, to what degree can the success or failure of a book be predicted? The answers, we hope, will be illuminating for authors, publishers, critics, tastemakers, and readers alike.

2   First Plots

2.1   Plot 1: Checkouts by Usage Class

Figure 1 shows the number of physical and digital checkouts from the Seattle Public Library from 2018 to 2022. While the number of digital checkouts has shown consistent and steady growth within this time period, the number of physical checkouts per year has been more sporadic. Most notably, there was significant decrease of physical checkouts in 2020, most likely due to the onset of the COVID-19 pandemic, and the number of physical checkouts per year after 2020 has still not recovered to the levels seen in 2018 and 2019. Because of this downturn in physical checkouts, there were more digital than physical checkouts in years 2020 and 2021.
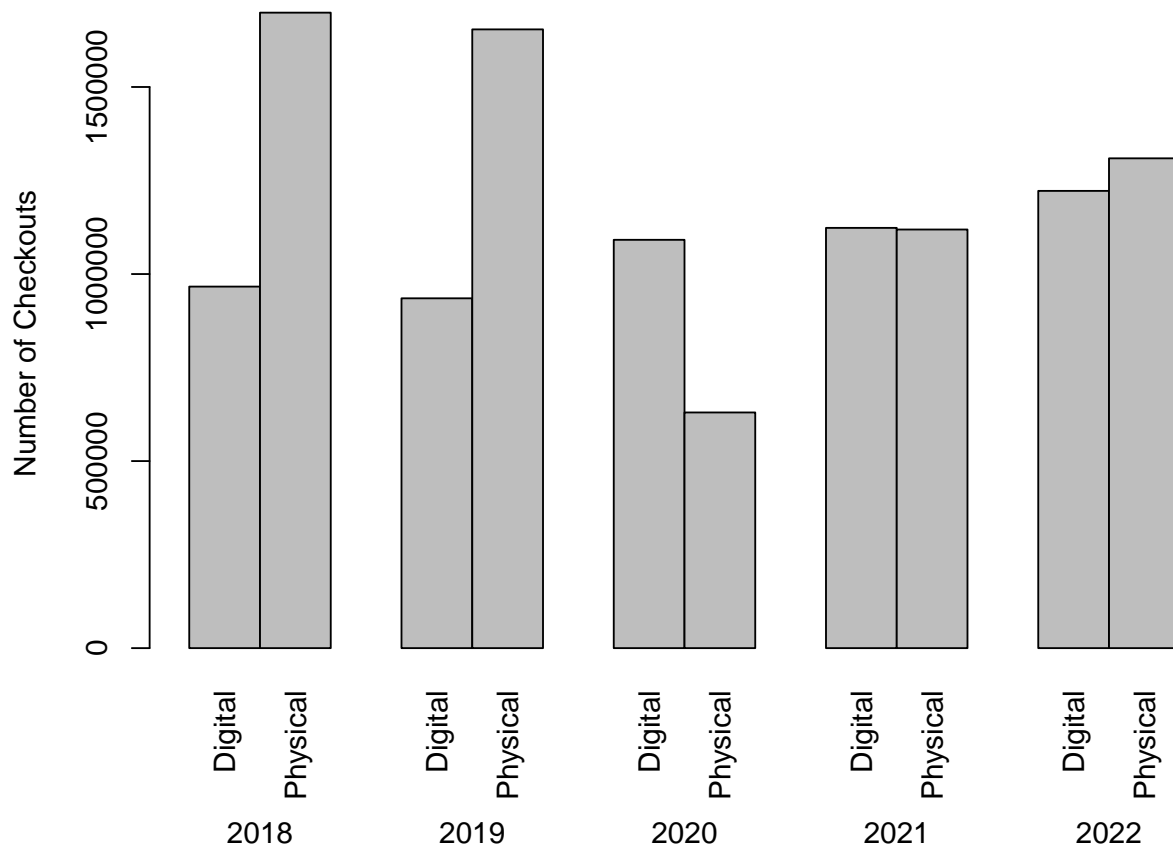
Figure 1: Number of Checkouts by Usage Class (Physical or Digital) per Year from 2018 to 2022.

2.2 Plot 2: Most Popular Subjects

Figure 2 shows the seven most popular book subjects of all books checked out between 2018 and 2022 in descending order of frequency. Books can have multiple subjects, such as both "Fiction" and "Literature." Fiction is the most popular topic with a little over 2.5 million occurrences, and Nonfiction had around 1.5 million. This is unsurprising since fiction and nonfiction are broad identifiers. In addition, in this plot, subjects are re-counted when their corresponding book is checked out again.

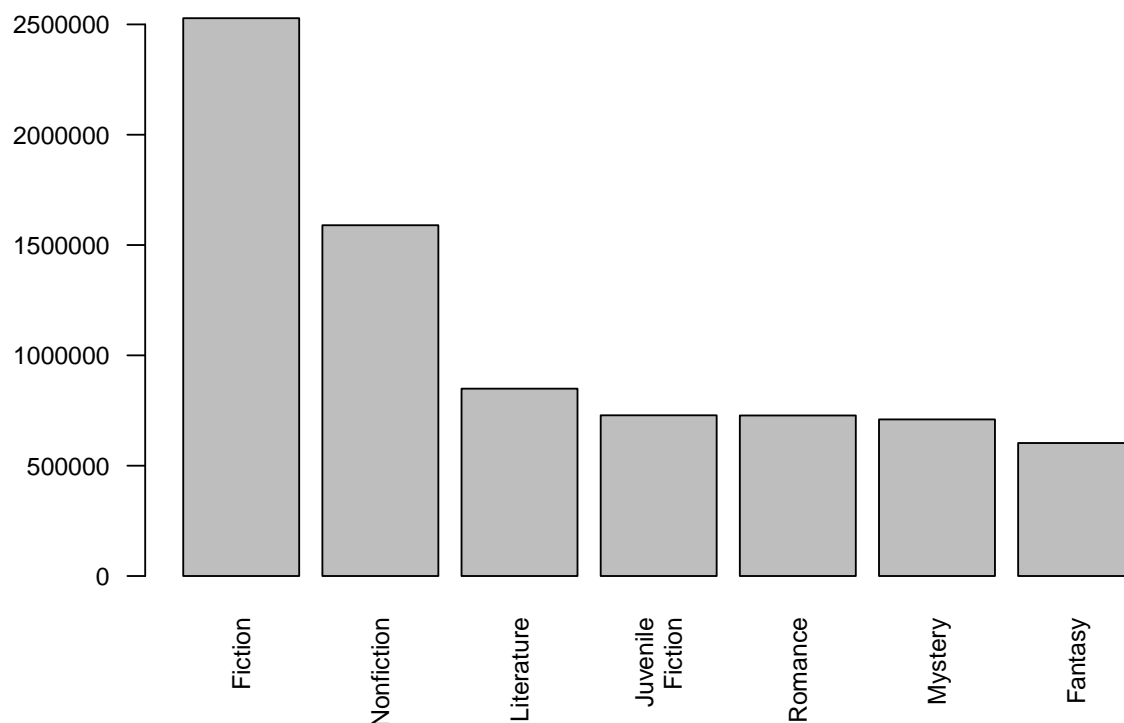**Top 7 Checked–Out Book Subjects, 2018–2022**

Figure 2: Top 7 subjects of all books checked-out from 2018 to 2022.

2.3   Plot 3: Most Popular Physical Books

Figure 3 shows the relative popularity of the seven most popular physical books with checkouts in each year between 2018 and 2022. We use "relative" popularity to signify that these are not necessarily the books that were most popular in each year. Rather, these are the most popular books whose popularity endured over the period in question. Novels and memoirs dominate, although their popularity curves exhibit a marked difference: popular memoirs (Educated; Becoming) appear to lack the staying power of similarly popular novels (Where the Crawdads Sing; There There; The Overstory). This figure could serve as ground zero for an investigation into the popularity curves and how they vary both within and across generic boundaries. We also note the curious prevalence of reading lines (A Memoir; a Novel). Reading lines have a long and complex history, beginning with 17th century publishers who sought to delineate new, experimentally-realistic work from older romances and epics. Early 20th century writers, by contrast, often troubled this binary by affixing "a Novel" to works that few readers would deem "realistic." Finally, post-45 American authors and publishers used (and continue to use!) reading lines to assert, aspirationally, that their book belongs among that most "serious" and "literary" in-group: prizewinners. An examination linking reading lines to popularity certainly bears further investigation.
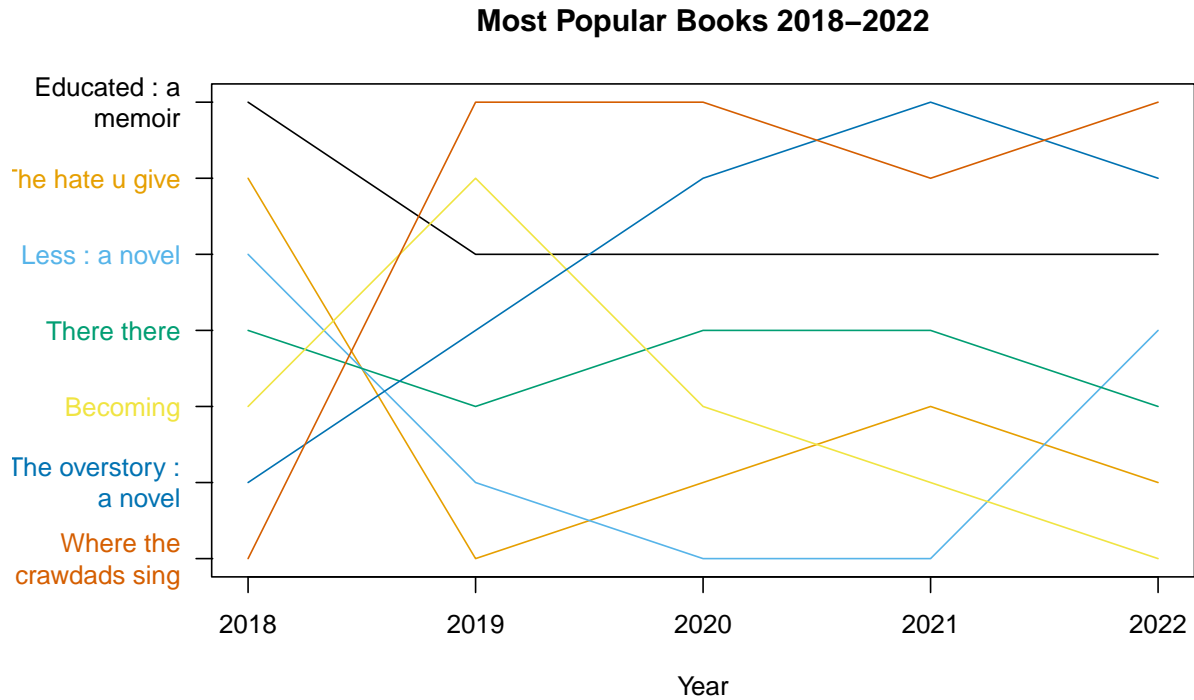
Most Popular Books 2018–2022

Figure 3: Relative popularity of the most popular physical books 2018-2022

2.4 Plot 4: Prior Year vs. Current Year Popularity

Figure 4 demonstrates the degree to which current-year popularity is predicted by prior-year popularity for years 2019-2022. We present a log-log plot of current vs. prior popularity and overlay a line of best fit. Prior popularity seems a fair – though by no means perfect – predictor of future popularity. Noticeably, there two groups straddle the line of best fit: one with over-predicted popularity and one under-predicted. These groups might be a consequence of the nature of popularity; namely, it's either rising or falling. Or this early analysis could be indentifying two categorically different groups of books: successes and flops.
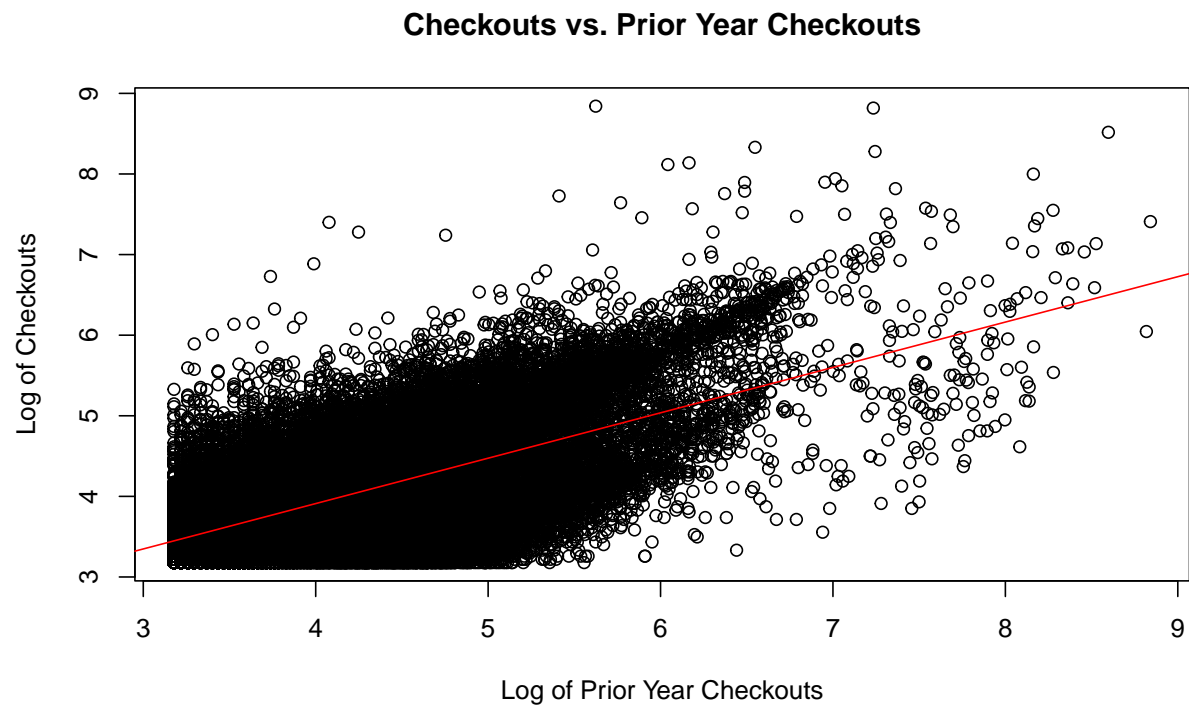
Figure 4: Current year vs. prior year checkouts, log-log scale