

# Forbush Collection & Circulation Analysis Report

(as of April 2025)

C.R. Gette, Ph.D.

May 16, 2025

## Abstract

This report analyzes the Forbush Memorial Library's collection and circulation data (from an export dated 4/15/25) to identify patterns and trends. The primary focus is on circulation performance in the previous full year (PY 2024), differentiating between items existing before 2024 and those newly added during 2024. Key findings highlight the strong performance of new acquisitions, Juvenile materials, and Large Print (Kendig Room), while Adult Nonfiction shows lower circulation relative to its size. Recommendations include promoting new items, reviewing the Adult Nonfiction collection, and leveraging insights from high-performing sections.

## Executive Summary: Key Findings

This report analyzes the Forbush Memorial Library's collection and circulation data (from an export dated 4/15/25) to identify patterns and trends. The primary focus is on circulation performance in the previous full year (PY 2024), differentiating between items existing before 2024 and those newly added during 2024.

## Key Circulation Insights

1. **"Newness" Drives Circulation:** Perhaps unsurprisingly, across all shelving locations and material types, items added in 2024 had a significantly higher average circulations than older items. For each library location, the "PY Circs" ranged from 1.6 (Young Adult) to 3.7 (Audio/Visual) average circulations per item for the "new" items, while the existing collection ranged from 0.75 (Adult Nonfiction) to 1.6 (Juvenile) average circulations per item. Additionally, between 60-85% of the "new" items circulated within 2024, while existing items ranged from about 35-60%. Patrons tend to like the new stuff! (See. Figs 1, 2)
2. **Top Performing Sections (PY 2024 Circs/Item):**
  - **Juvenile:** Consistently high average circulation per item for both existing and new materials. Additionally, had a higher proportion of total number of circulations: 56% (Juvenile Circulations/Total Circulations) than the current library proportion: 45% (Juvenile Item Count/Total Collection Count). (Fig. 3)
  - **Kendig Room (Large Print):** Shows very strong average circulation per item, especially for new additions. This also had a higher proportion of total circulations to library proportion (3.5% vs 3.2%), although the effect is much less given the smaller size of the Kendig Room collection. (Fig. 3)
3. **Underperforming Section (PY 2024 Circs/Item):**

- **Adult Nonfiction:** Has the lowest circulation per item average, indicating a potential mismatch between collection size/scope and patron demand or discoverability issues. This section also had the greatest mismatch between proportion of the library collection (15%) and proportion of circulations (9%). (Fig. 3)

#### 4. High Engagement Sections (Proportion of Collection Circulating in PY 2024):

- **Juvenile & Kendig Room:** A large proportion of items in these existing collections circulated at least once (55% and 57% respectively for items that existed prior to 2024). The Kendig Room also had the highest proportion of new items (added in 2024) that circulated at least once (87%) (Fig. 2)
- **Audio-Visual:** Newly added Audio-Visual items received very high circulations per item (3.6).

#### 5. Fiction Genre Insights (PY 2024):

- **Thrillers:** Dominate both the fiction collection size and total circulations. Newly added thrillers have a very high circulation per item.
- **Literary Fiction & Romance:** Also form significant parts of the collection and see strong circulation.
- **High Circulation for New Additions:** Newly added books in Historical Fiction, Christian Fiction, and Thriller showed particularly high average circulations (number of circulations/items in category) with  $> 4$  circulations per-item in 2024. (Fig. 5)
- **Proportion of Genre Circulating:** Almost all of the genres noted have very high proportion of the newly added items circulate at least once (generally  $> 70\%$ ) while the existing items lagged behind with no genre circulating more than 40 %. (Fig. 6)

### Key Recommendations

1. **Promote New Acquisitions:** Continue to capitalize on the high interest in new items through targeted displays, newsletters, and social media.
2. **Review Adult Nonfiction:** Investigate the low circulation. Consider weeding, targeted acquisition based on demand, or improved promotion.
3. **Leverage Strengths:** Analyze the success factors of Juvenile and Kendig Room (e.g., collection development, programming, layout) for potential adoption elsewhere.
4. **Monitor Genre Performance:** Continue tracking fiction genre circulation. Ensure acquisition in high-demand genres like Thrillers is maintained, and explore growth opportunities in other popular or newly circulating genres.

# 1 Introduction

This report provides an Exploratory Data Analysis (EDA) of the Forbush Memorial Library's collection inventory, focusing on circulation patterns. The data, exported on April 15, 2025, comprise approximately 50,500 records. The primary goal is to understand how different segments of the collection are utilized, with a particular emphasis on circulation performance during the previous full calendar year (PY 2024). This analysis aims to provide actionable information for collection development, patron engagement, and resource allocation.

## 2 Data Overview & Preparation

The dataset includes detailed information for each item, such as:

- Bibliographic details (Title, Author, Publisher, Call Number, etc.)
- Collection metadata (Owning Library, Shelving Location, CWMARS Format, ARIS Format, etc.)
- Circulation statistics (Total Circs, Current Year (CY) Circs, Previous Year (PY) Circs)
- Key dates (Create Date/Time, Active Date/Time, Last Circ in EG, etc.)

Initial data preparation involved standardizing date formats and creating a flag (`added_2024`) to distinguish items added to the collection during 2024 from those existing prior. A thorough data quality check was performed; specific anomalies and potential data entry issues are detailed in Appendix A for review by cataloging staff. For circulation analysis, items in non-circulating categories (e.g., "Library Use Only," "Noncirculating," "DO NOT USE," "Adult Reference," "Special Collection") were filtered out to focus on materials available for checkout.

## 3 Circulation Analysis

The core of this analysis examines circulation data, primarily focusing on "PY Circs" (circulations in 2024) to assess recent performance. Key metrics include:

- **Circulation Average:** Average circulations per item (PY Circs / Number of Items).
- **Proportion of Collection Circulating:** Percentage of items in a category that circulated at least once in PY 2024.
- **Collection Makeup vs. Circulation Share:** Comparing a category's share of the total collection to its share of total circulations. (Number of items in Category/Total number of items vs. Category Circs/Total Circs)

### 3.1 Circulation by Shelving Location

- **Circulation Average (Circs/Item) (Fig. 1):**
  - Items added in 2024 ("Newly Added") consistently outperformed "Existing" items across all shelving locations.
  - **Juvenile** and **Kendig Room** (identified as primarily Large Print) demonstrated the highest circulation average for existing items.
  - **Audio-Visual**, **Kendig Room**, and **General** saw the most significant circulation average for newly added items.

- **Adult Nonfiction** had the lowest circulation average for existing items and second lowest for newly added items (Young Adult was lowest).
- **Young Adult** showed a smaller difference in circulation average between existing and newly added items compared to other categories, and lowest average circulations per item for newly added items.

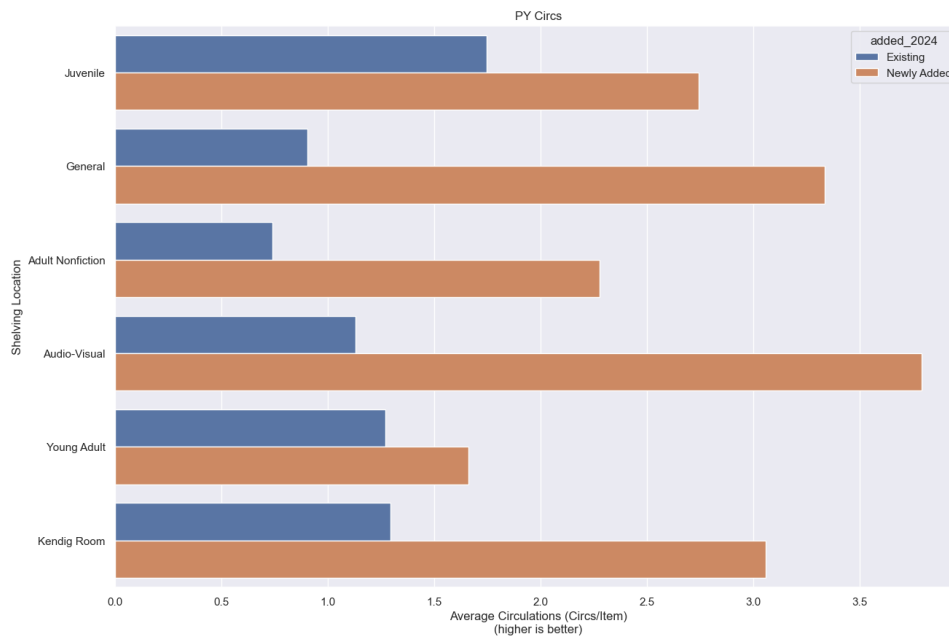


Figure 1: PY 2024 Circulation Average (Circs/Item) by Shelving Location: Existing vs. Newly Added.

- **Proportion of Collection Circulating (Fig.2):**

- A high proportion of newly added items circulated across all categories.
- For existing items, **Juvenile** and **Kendig Room** had the largest proportion of their collections circulate.
- **Adult Nonfiction** and **General** had the lowest proportion of their existing collections circulate.

- **Library Makeup vs. Circulation Share (PY 2024) (Fig. 3):**

- **Juvenile:** Comprises the largest segment of the circulating collection (approx. 45%) and accounts for the largest share of circulations (approx. 56%).
- **General:** The second largest collection segment (approx. 22%), contributing about 18% of circulations.
- **Adult Nonfiction:** Represents about 15% of the collection but only 9% of circulations, indicating lower usage relative to its size.
- **Audio-Visual:** Makes up about 12% of the collection and 10% of circulations.
- **Kendig Room & Young Adult:** Smaller collection segments with circulation shares roughly proportional to their size.

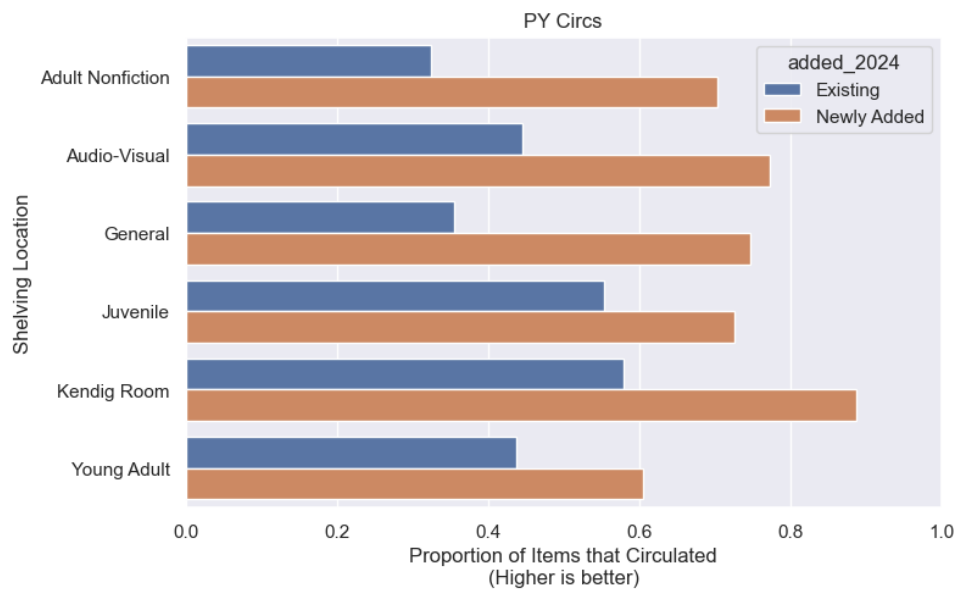


Figure 2: Proportion of Collection Circulating in PY 2024 by Shelving Location: Existing vs. Newly Added.

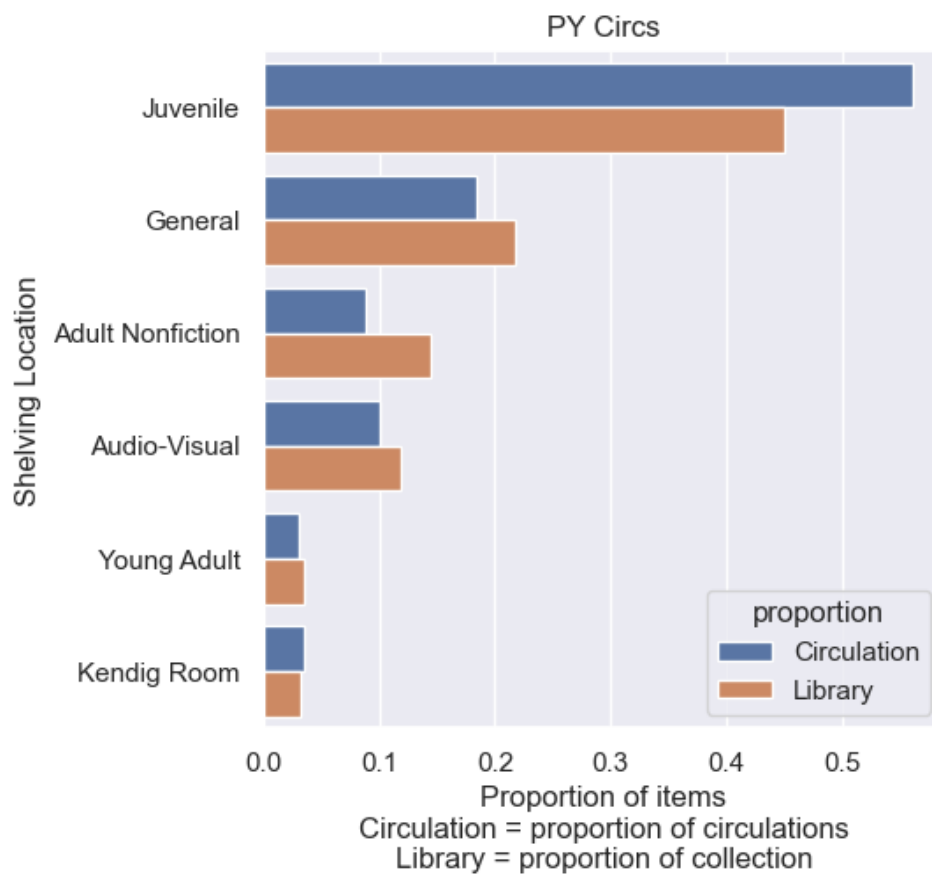


Figure 3: PY 2024 Library Collection Makeup vs. Circulation Share by Shelving Location.

### 3.2 Circulation by CWMARS Format (within Shelving Locations) (Fig. 4)

Breaking down shelving locations by CWMARS Format provides more granular insights:

- **Kendig Room:** Confirmed to be almost exclusively "Large Type," which circulates well.
- **Young Adult:** Interestingly showing a higher average circulation of existing "YA Non-fiction" rather than newly added.
- **General:** A mix of "Fiction," "Mystery," "Romance," and other formats. "Equipment" (though small in number) showed a very high circulation average.
- **Juvenile:** Highly diverse, with "Juv Picture Hardcover," "Juv Fiction Hardcover," "Beginner Reader," and "Juv Graphic" being prominent. Many newly added Juvenile formats (e.g., "Juv Graphic," "Juv Board Book," "Talking Book," "Game") had a very high proportion of items circulate.
- **Adult Nonfiction:** Newly added items within this section showed better circulation average than existing items.

### 3.3 Fiction Genre Analysis (Items with "FIC" Call Number)

Fiction items (identified by call numbers starting with "FIC") were further classified into common genres using an AI tool (Gemini) to understand sub-genre performance. To get each genre, the title and author were reviewed by Google's Gemini "thinking model" (Gemini-2.5-pro). The instructions were to assign one of the specified Genres and only if those genres did not apply, to then assign Literary Fiction, Unknown, or Other. The specifics of this implementation are not described here, but can be shared upon request. The genre labels are additionally shared as a csv accompanying this report.

- **Circulation average (Circs/Item) by Genre (PY 2024) (Fig. 5):**
  - Newly added items consistently outperformed existing items across all genres (generally >3 circs/item for newly added and <1 for existing)
  - **Thriller, Historical Fiction, and Christian Fiction** showed the highest circulation average for newly added books.
  - For existing books, **Romance, Historical Fiction, Fantasy, and Western** had an average above 1.0 (circulating more than once per item on average).
- **Proportion of Collection Circulating by Genre (PY 2024) (Fig. 6):**
  - A very high percentage of newly added **Christian Fiction, Historical Fiction, and Thriller** titles circulated. Sci Fi as well, however with only 5 newly added items, it is a negligible number.
  - For existing collections, genres had at most around 40% of the collection circulate at least once.
  - **Sci Fi** had the lowest proportion of the existing collection circulate.
- **Library Fiction Makeup vs. Circulation Share (PY 2024) (Fig. 7):**
  - **Thriller:** Largest genre by collection size (approx. 34%) and circulation share (approx. 32%).
  - **Literary Fiction:** Second largest by collection size (approx. 22%), contributing about 20% of circulations.

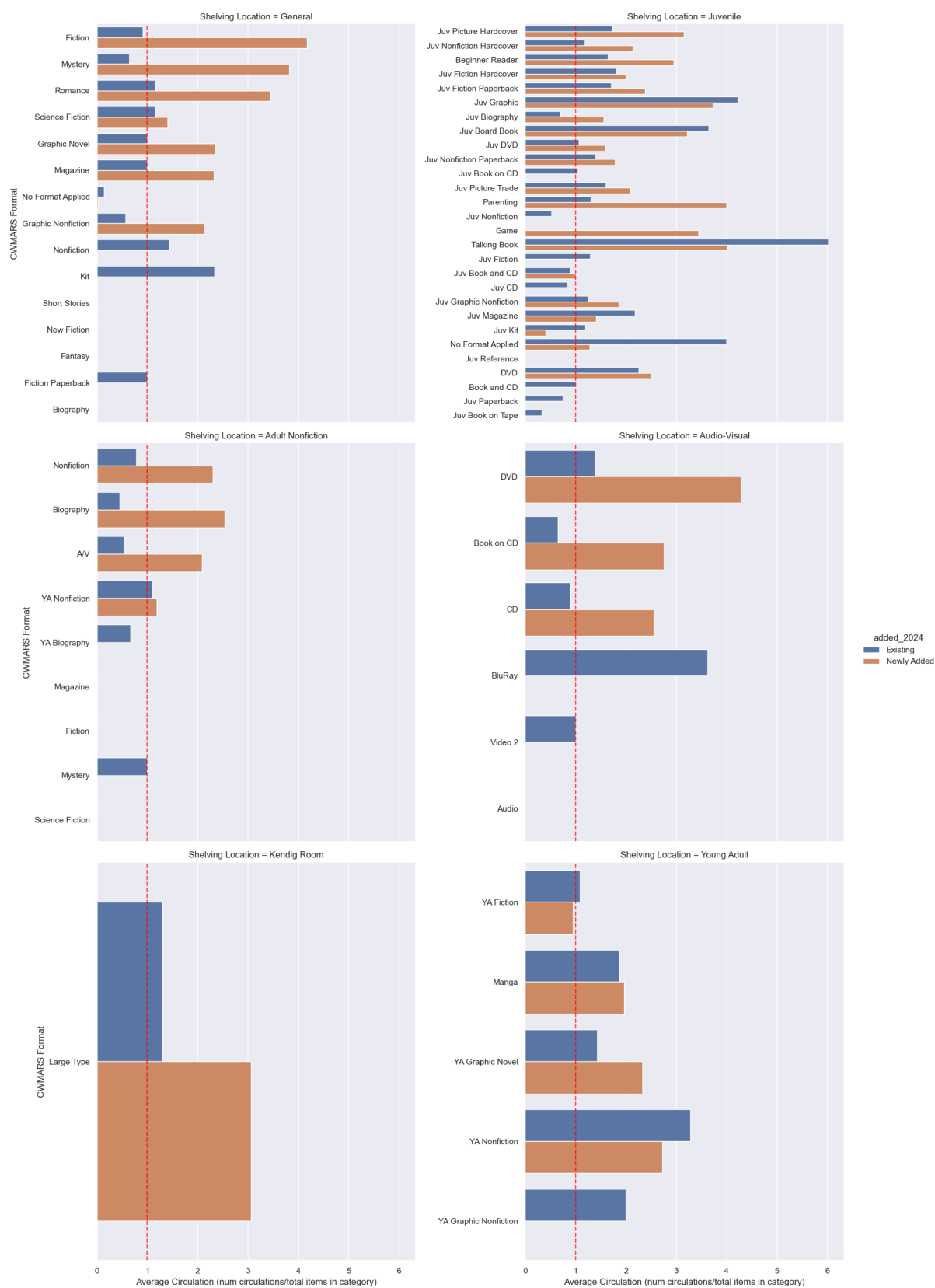


Figure 4: PY 2024 Circulation average by CWMARS Format for Selected Shelving Locations.

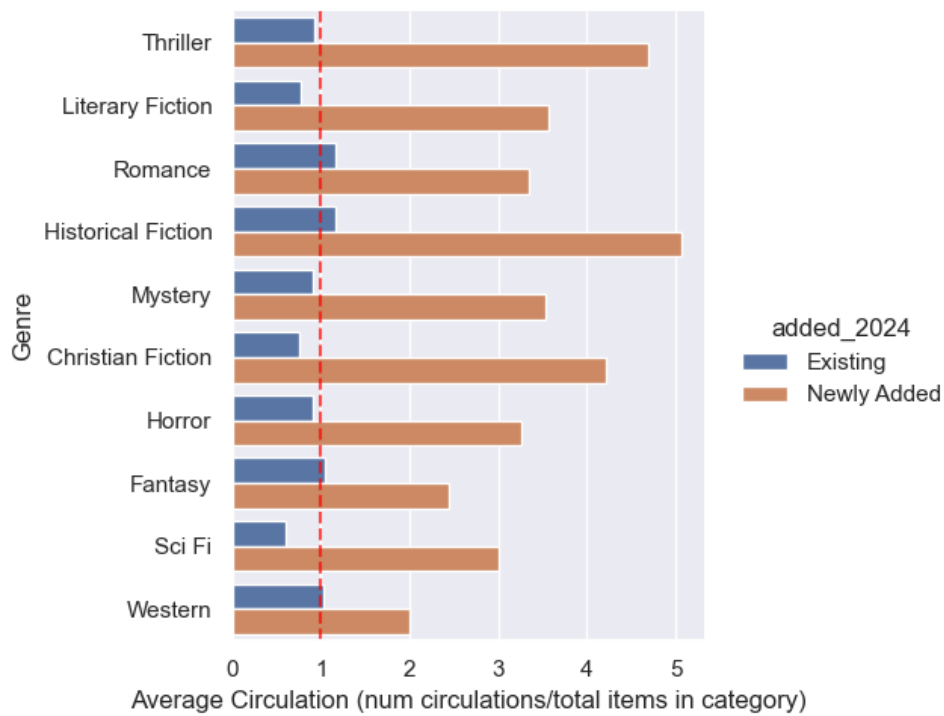


Figure 5: PY 2024 Circulation average (Circs/Item) by Fiction Genre: Existing vs. Newly Added.

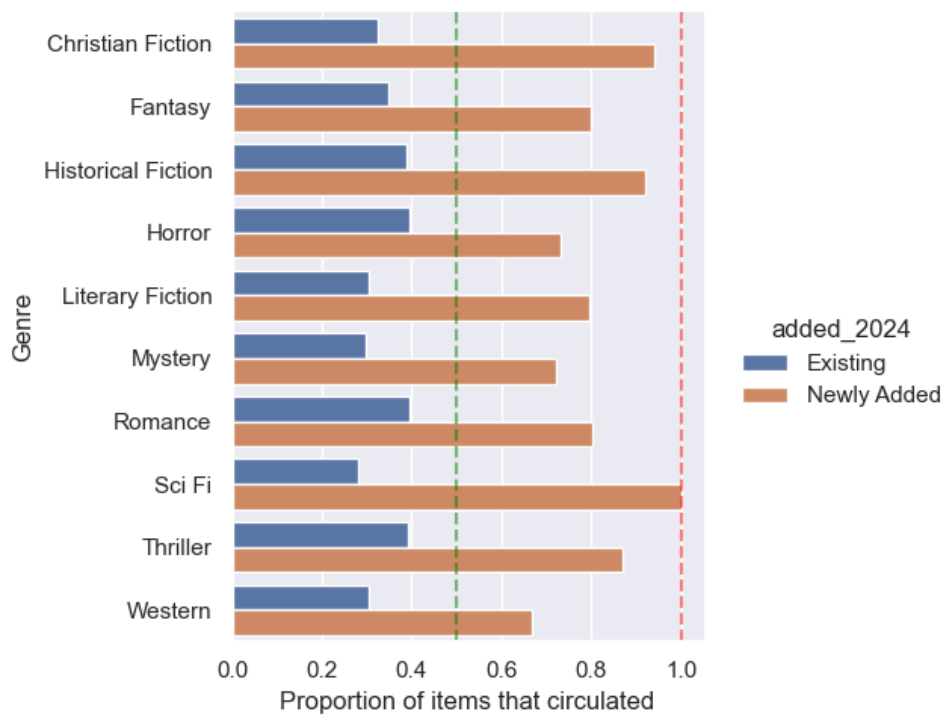


Figure 6: Proportion of Collection Circulating in PY 2024 by Fiction Genre: Existing vs. Newly Added.



- **Romance:** About 18% of the collection, accounting for a higher share of circulations (approx. 21%).
- **Historical Fiction:** Approx. 13% of the collection, 15% of circulations.
- Other genres (Mystery, Christian Fiction, Horror, Fantasy, Sci-Fi, Western) make up smaller portions of the collection and circulations.

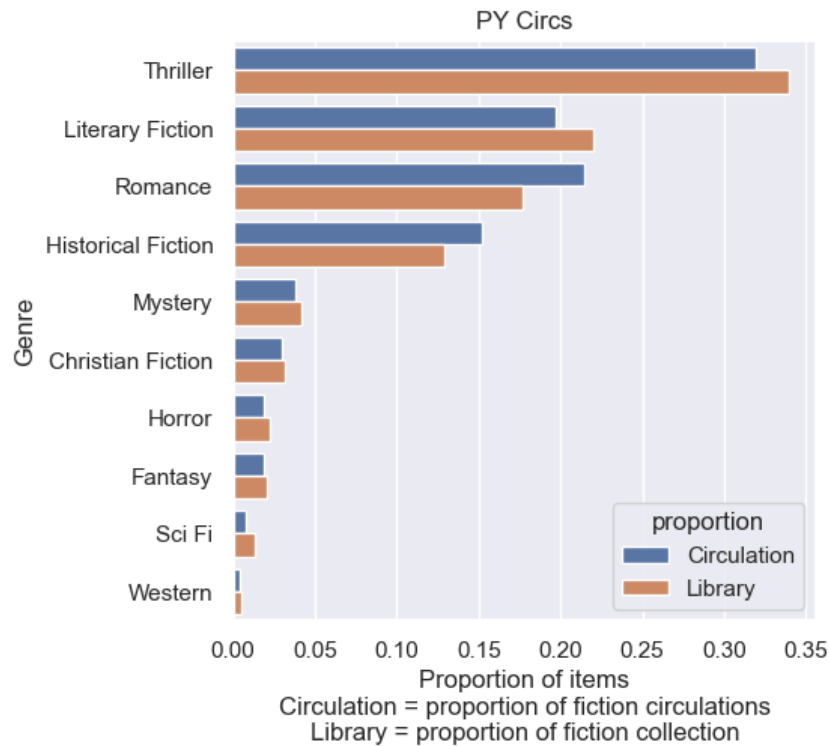


Figure 7: PY 2024 Library Fiction Collection Makeup vs. Circulation Share by Genre.

## 4 Conclusion

The analysis shows strong patron interest in new materials across all categories. Juvenile and Kendig Room (Large Print) are standout locations. Adult Nonfiction presents an area for potential review and strategic adjustment. The detailed breakdown by CWMARS format and fiction genre offers deeper insights for targeted collection development and promotion. Continued monitoring of these trends, especially for newly added items and specific genres, will be beneficial for optimizing the library's collection to meet patron needs.

## A Data Quality & Anomaly Notes for Cataloger Review

During the EDA, several data characteristics and potential anomalies were observed:

- **CN Prefix Column:** Appears to be entirely empty (NaN) and can likely be disregarded or dropped.
- **Price Column:**
  - Many missing values, often corresponding to items with "On order" status, which is logical.
  - A few items had unusually high prices (e.g., >\$500). These might be valid special items or data entry errors. (Barcode: 37365000074256, 37365003097296)
  - Only 5 "Available" items had a missing price; these should be reviewed. (Barcode: 37365003133448, 37365002919276, 37365003064320, 37365002705550, 37365002855181)
- **CN Suffix Column:** Rarely used (234 instances). Predominantly associated with "Library Use Only" or "Noncirculating" items, often reference materials or specific discs in a set.
- **Part Column:** Used in ~10% of records. Often indicates volumes, disc numbers, or sometimes years (e.g., "1971," "1952"). More prevalent in "Special Collection."
- **Circ Modifier Column:** Only one record was found with a missing Circ Modifier. This should be an easy fix. (Barcode: 230527833)
- **Title Column:** Two records were found with missing titles. (Barcode: 37365004346577, 37365004345884)
- **Author Column:** Approximately 6,343 records have missing authors. This is expected for many Audio-Visual items (DVDs, CDs) but also occurs for some non-fiction and juvenile books.
- **Published Year Column:**
  - Contains non-numeric entries (e.g., "[1984], c1965", "19uu", "c1961"). These prevent direct numeric conversion. ~380 such items were identified, mostly in "Special Collection."
  - This field is currently object/string type.
- **Barcode Column:**
  - Contains "On Order [number]" for items not yet received, which is logical but prevents numeric conversion. ~600 such items, all correctly marked "On order" under **Item Status**.
  - One unique instance of "Lumberjack" for an ILL item in "DO NOT USE" shelving location. This appears to be a placeholder or special case. (Barcode: Lumberjack)
- **Shelving Location - "DO NOT USE":** Used for 54 items, apparently for Inter Library Loan (ILL) items. This seems to be a functional, though perhaps unconventional, use.
- **Shelving Location - "Juvenile Reference":** Contains only 1 item. Informed this category is being phased out. (Barcode: 37365002815672)
- **"FIC" Call Numbers with "Nonfiction" CWMARS Format:** A few items (4) with call numbers starting "FIC" were categorized as "Nonfiction" in CWMARS Format. This is a minor discrepancy.

These observations are provided to aid in ongoing data maintenance and ensure the highest quality data for future analyses and system operations.

## B Additional Figures

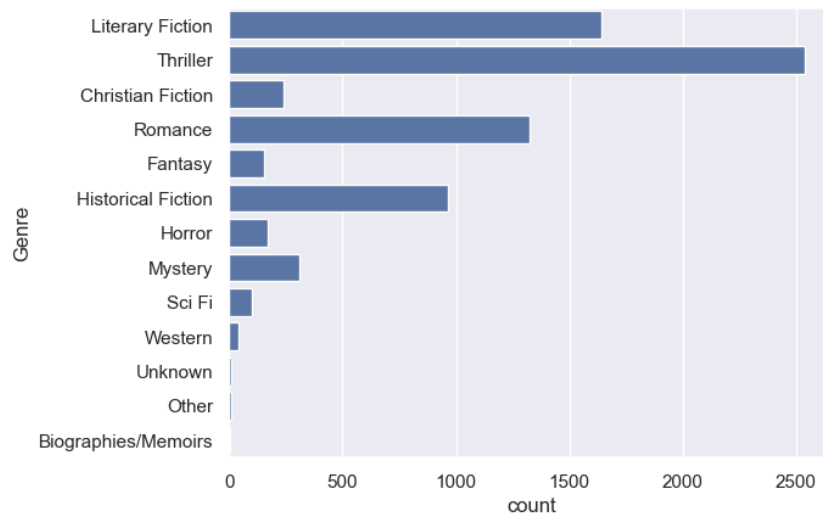


Figure 8: PY 2024 Library Fiction Collection Counts by Genre.

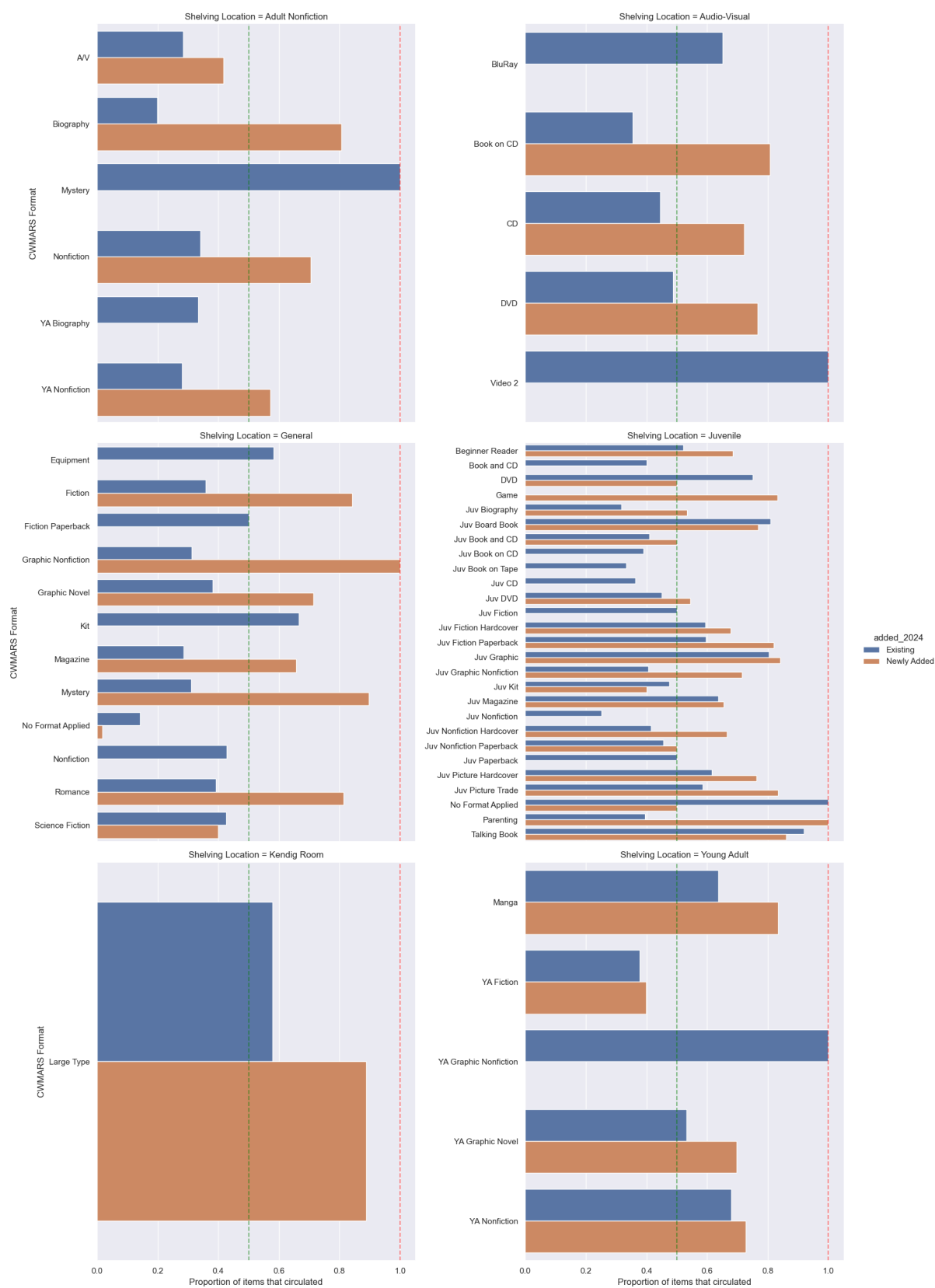


Figure 9: PY 2024 Proportion Circulating by CWMARS Format for Selected Shelving Locations.