Week1 Practice

Jonathan Walls

5/30/2021

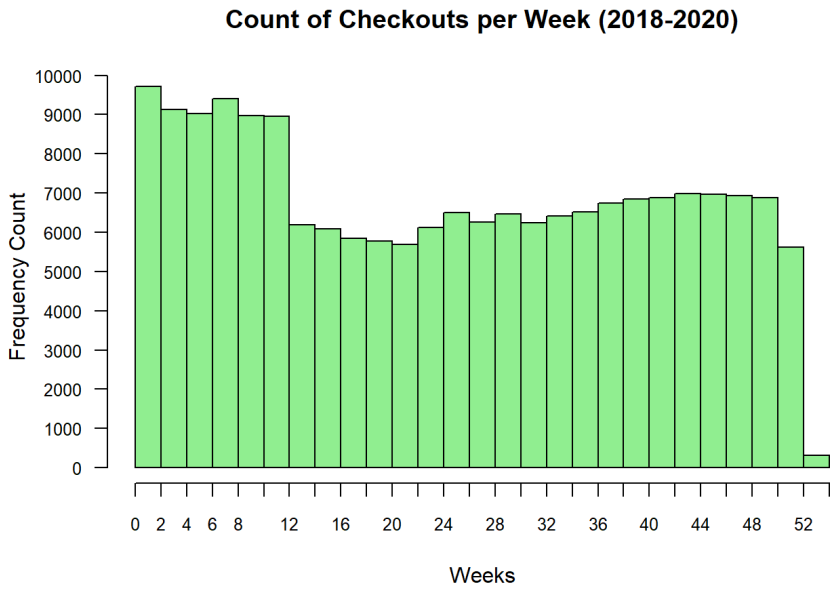
Introduction

In this report, I am going to look at some checkout data from the Seattle Library and answer a couple key questions which I have outlined below:

1. During which weeks throughout the year, does the library experience the most checkouts?
2. Throughout the week, which day is the busiest at the Seattle library, according to our sample?
3. Is the busiest day influenced by the age of the person checking out the book (Adults vs Juveniles)?
4. Throughout any given day, what are the peak hours at the library?
5. Of all fiction books checked out, what is the most popular sub-category and do different age groups prefer different fiction books?

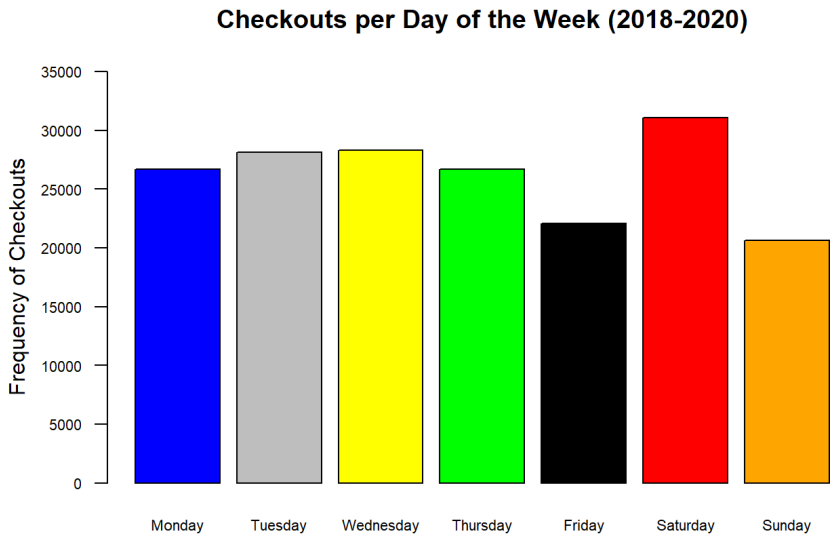
Prior to examining these questions, I will note some data manipulation which needed to occur in order to analyze the above questions. First, I needed to extract the following for all data records: the name of the day of the week, the number of the week within each year, the hour a book was checked out (In a 24 hour format), as well as the number of the day within each week (1-7). Next, I needed to create a unified age group as some of the data contained the indicator “teen” which did not add value to my analysis. Finally, I needed to create multiple frequency tables, which I used to create barplots for my data.

Examining Weekly Checkout Data Throughout the Year



In the above histogram, one will notice how the data appears to be positively skewed with the library receiving the most checkout of books in the first 12 weeks of the year. However, I am not positive that it is positvely skewed, as the mean and median are both approximately 25. I will propose that the reason for the spike of checkouts in week 1-12 could be due to everyone’s New Years’ Resolutions to read more, and then this spike drops off in week 13. One will then notice a couple weeks of decline followed by upticks through the end of the year until December. I would think as people settle in for the holidays, the library then sees a decline in the number of books checked out.The final note I will make is that week 53, does not occur every year, which is why the number of checkouts is so much lower.

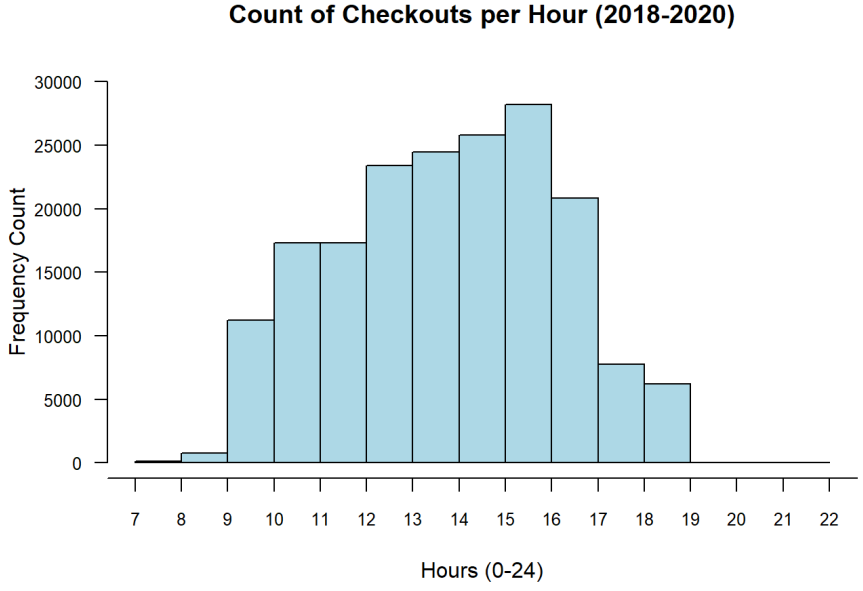
Examining Checkout Data by Days of the Week & Age Grouping




```
##
##              Adult Juvenile
##   Monday    24889     1796
##   Tuesday   26375     1736
##   Wednesday 26451     1829
##   Thursday  24977     1697
##   Friday    20443     1628
##   Saturday  28692     2366
##   Sunday    18826     1787
```

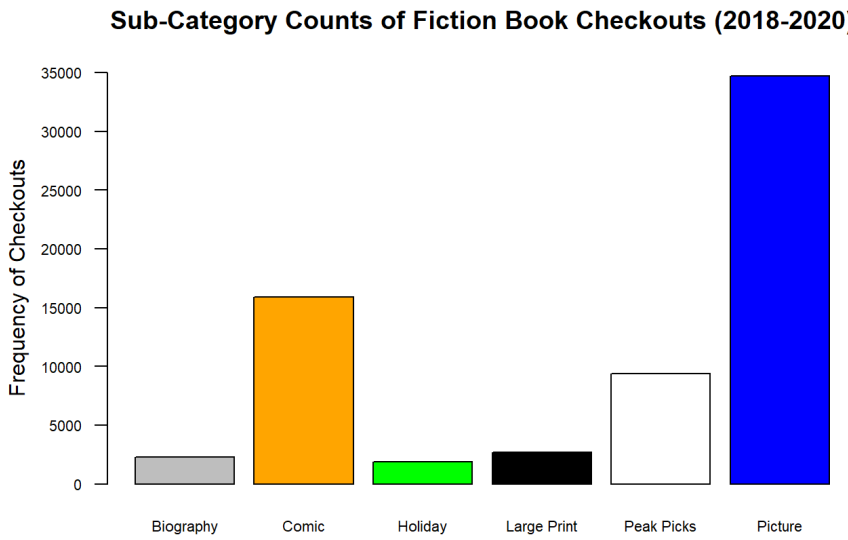
Above are two graphics, in the first, one will see a bar plot which highlights the busiest day of the week (Saturday) for checkouts of books at the Seattle library.I have also further broke down this data by age grouping (Adult/Juvenile) and created a cross-tabulation as the second image. In the second table, one will notice that Saturday is the busiest day for both adults and juveniles, but that juveniles visit the library proportionately less all other days of the week than adults. It appears as though adults check out books on Tuesdays and Wednesdays almost as much as they do on Saturdays. This makes sense since juveniles most likely need their parents to get to the library and therefore cannot go as much during the week days when parents are busier.

Examining the Distribution of Checkouts by Hour



In this graphic, one will notice the distribution of checkout times appear to be slightly positively skewed, and that the peak time of checkout is between 3:00 PM and 4:00 PM, and there are far fewer checkouts early in the morning when the library first opens and later in the evening just before the library closes.

Examining Fiction Book Checkouts, and the Relationship to Age Group




```
##
##                Adult Juvenile
##   Biography     2216       96
##   Comic        11412     4490
##   Holiday       1855       50
##   Large Print   2591       81
##   Peak Picks    9168      232
##   Picture      33871      811
```

Above are my final two images, in the first, one will see a graph of all the fiction books with a sub-classification. The most popular fiction books are picture books, with the second most popular book being the comic book. I have also included a cross-tabulation of the adult/juvenile classification. One will notice that adults check out picture books the most, where juveniles check out comic books the most.

Bibliography

1. Kabacoff, Robert. R In Action: Data Analysis and Graphics with R. Manning, 2015
2. Bluman, Allan G. Elementary Statistics: a Step by Step Approach. McGraw-Hill Education, 2018
3. *Get Week Number of Date in R (2 Examples)* (n.d.). Statistics Globe https://statisticsglobe.com/get-week-number-of-date-in-r (https://statisticsglobe.com/get-week-number-of-date-in-r)
4. *Dates and Times in R* (n.d.). Berkeley Statistics Department https://www.stat.berkeley.edu/~s133/dates.html (https://www.stat.berkeley.edu/~s133/dates.html)
5. *R: scaling axis for different limits in hist* (2018, December). BioStars https://www.biostars.org/p/352966/ (https://www.biostars.org/p/352966/)
6. *How to suppress the code but have the plots displayed in R markdown?* (2017, September, 25). stackoverflow https://stackoverflow.com/questions/46414779/how-to-suppress-the-code-but-have-the-plots-displayed-in-r-markdown (https://stackoverflow.com/questions/46414779/how-to-suppress-the-code-but-have-the-plots-displayed-in-r-markdown)
7. *R Markdown Basics* (2016, October, 24). RPubs by RStudio https://rpubs.com/Rtudes/RMarkdownBasics# (https://rpubs.com/Rtudes/RMarkdownBasics#):~:text=To%20write%20text%20in%20italic,before%20and%20after%20th