

ProjectMilestone1

Elizabeth Edminster & Sara Evans

2024-03-15

Project: Analysis of the trends that lead to the survival & downfall of American public libraries

Problem: Figuring out what makes certain libraries sustainable, trends in libraries from different states, & other assorted trends & survival data from 2006-2021

Unusually named Columns: BKVOL - PRINT MATERIALS EBOOK - ELECTRONIC BOOKS AUDIO_PH - AUDIO PHYSICAL UNITS AUDIO_DL - AUDIO DOWNLOADABLE UNITS VIDEO_PH - VIDEO PHYSICAL UNITS VIDEO_DL - VIDEO DOWNLOADABLE UNITS LOANTO - INTER-LIBRARY LOANS OUT LOANFM - INTER-LIBRARY LOANS RECEIVED PITUSR - PUBLIC INTERNET COMPUTER USES PER YEAR KIDCIRCL - CIRCULATION OF KIDS MATERIALS TOTCIR - TOTAL CIRCULATION VISITS - LIBRARY VISITS C_FSCS - (Public library system meets all criteria in the definition)

Guiding Question: Our guiding questions are “what features make some libraries successful, some libraries popular, and some libraries disappear? How do different libraries react to changes throughout the years from 2006-2021? How do different locations affect these features & responses?”

Domain: This review is under the domain of public resources, education, & to some extent, sociology. The implications of the results would fall under the fields of education, public policy, library science, & business.

Data: <https://www.ims.gov/research-evaluation/data-collection/public-libraries-survey> This data is from the Institute of Museum & Library Services. They have data on these libraries nationally from 2005 to 2021. The selected columns for our preliminary exploration are listed below. Further columns & information about the data can be found at this link.

Following questions: We would also like to look at the affects of national events such as the 2008 financial crisis, covid, & more on library enrollment. We may also look into other data on things like literacy rates, childhood technology use, & public perspectives on public libraries to compare the data on a matched timeline.

```
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
## Warning: package 'dplyr' was built under R version 4.3.2
```

```
## Warning: package 'stringr' was built under R version 4.3.2
```

```
## Warning: package 'lubridate' was built under R version 4.3.2
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.4
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2     3.4.4      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.0
## v purrr       1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
data <- read.csv("C:/Users/lized/OneDrive/Documents/GitHub/R-Analysis-Project/pls_fy2021_csv/PLS_FY2021.csv")
#This data specifically is the 2021 data
df <- data %>% select(LIBID, LIBNAME, STABR, LOANTO, LOANFM, PITUSR, TOTCIR, KIDCIRCL, BKVOL, EBOOK, AUDIO_DL, VIDEO_PH, VISITS, C_FSCS)
df$C_FSCS <- as.factor(df$C_FSCS)
summary(df)
```

```
##      LIBID      LIBNAME      STABR      LOANTO
## Length:9215      Length:9215      Length:9215      Min.   :   -3.0
## Class :character      Class :character      Class :character      1st Qu.:   31.5
## Mode  :character      Mode  :character      Mode  :character      Median :   548.0
##                                         Mean   :  6969.2
##                                         3rd Qu.:  5436.5
##                                         Max.   :706516.0
##      LOANFM      PITUSR      TOTCIR      KIDCIRCL
## Min.   :   -3      Min.   :   -3      Min.   :   -3      Min.   :   -3
## 1st Qu.:   50      1st Qu.:   306      1st Qu.:   7546      1st Qu.:   1900
## Median :   593      Median :   1090      Median :   25986      Median :   7403
## Mean   :   7153      Mean   :   5751      Mean   :  168778      Mean   :   55391
## 3rd Qu.:  4832      3rd Qu.:   3610      3rd Qu.:   90500      3rd Qu.:   29406
## Max.   :919237      Max.   :1484987      Max.   :17408320      Max.   :5778440
##      BKVOL      EBOOK      AUDIO_PH      AUDIO_DL
## Min.   :   -3      Min.   :   -3      Min.   :   -3      Min.   :   -3
## 1st Qu.:  13698      1st Qu.:  14555      1st Qu.:   316      1st Qu.:   5003
## Median :  26873      Median :  41154      Median :   1041      Median :   16032
## Mean   :  71831      Mean   :  113585      Mean   :   4072      Mean   :   55286
## 3rd Qu.:  60241      3rd Qu.: 105022      3rd Qu.:   3314      3rd Qu.:   34278
## Max.   :22168629      Max.   :2189199      Max.   :511890      Max.   :16457956
##      VIDEO_PH      VISITS      C_FSCS
## Min.   :   -3      Min.   :   -3      N: 190
## 1st Qu.:  1080      1st Qu.:   3250      Y:9025
## Median :   2490      Median :   11000
## Mean   :   7221      Mean   :   45306
## 3rd Qu.:   6006      3rd Qu.:   36216
## Max.   :525571      Max.   :6722578
```