

FBN Data Report #2

9/14/2023

Salome “Sam” Grasland
samgrasland@gmail.com





What We've Got So Far & How We Got Here



1, 836 Rows of Data

- In total we have spent **1,201 days (3.2 years)**
- On **average** a person can index **17 pages** in a day
- The **most pages** indexed in a day were **144**

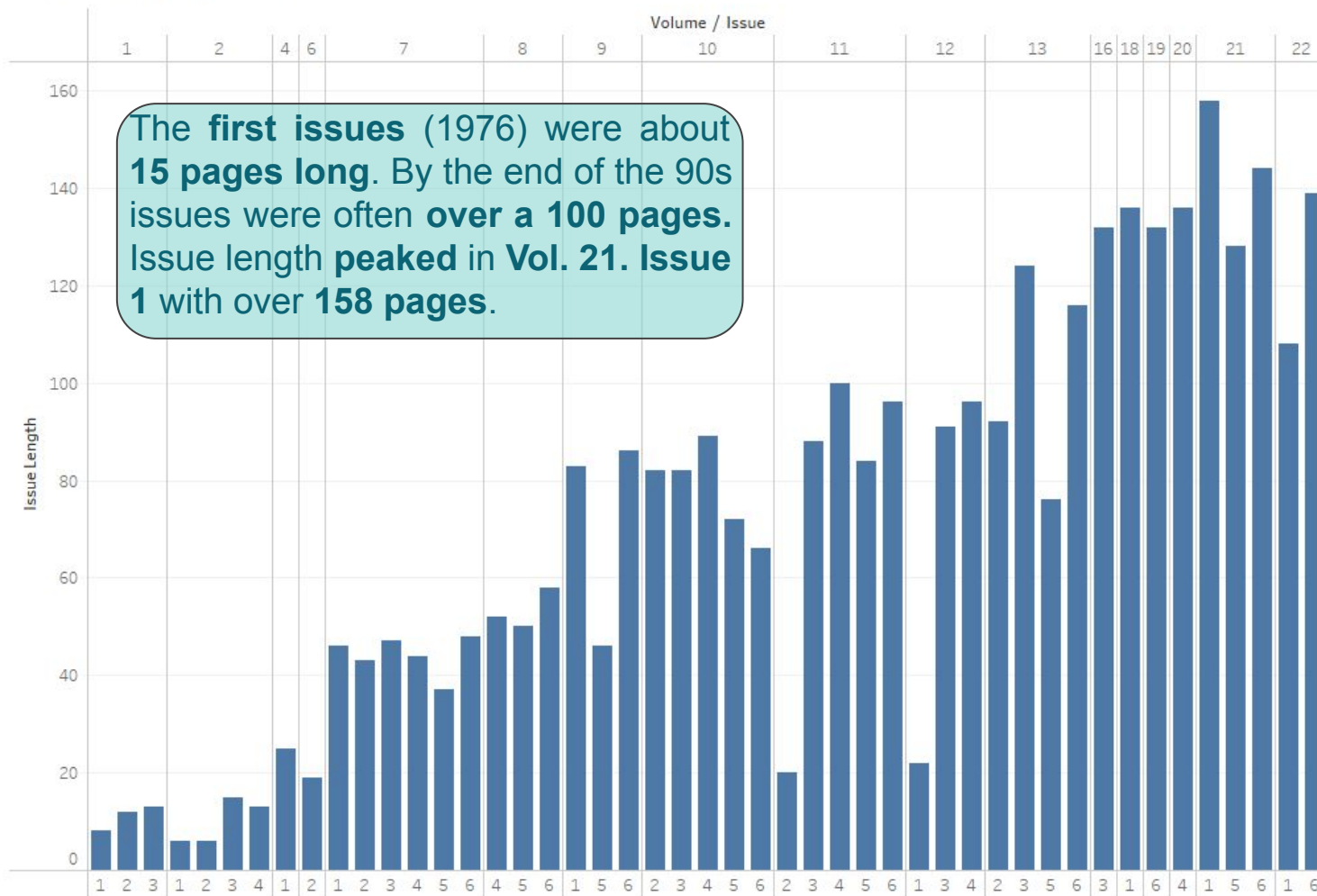
19 Different Contributors

- **Completed Issues**
Li-Anne Wright : 11
Andréanne Wahlman: 7
Rose Norman: 6
- **Completed Rows**
Li-Anne Wright: 273
Elizabeth Lanyon: 215
Andréanne Wahlman: 203

17/22 of Volumes
partially indexed
and 20/25 years
partially indexed

- **Volumes Missing:**
3, 5, 14, 15, 17
- **Years Missing**
1979, 1981, 1992,
1994, 1996

Issue Length Over Time





Storing and Managing Data



Data Entry and Cleaning

- FBN data is being **stored and managed** on a google sheet
- **Pros:** easily accessible to a wide user base
- **Cons:** google sheet will result in unorganized data
- **For example:** to the right I have pasted an image currently showing **all iterations of month**
 - Contains **input errors** like spelling errors, abbreviations, or inconsistent formatting.
 - While this can be cleaned it **adds processing time**

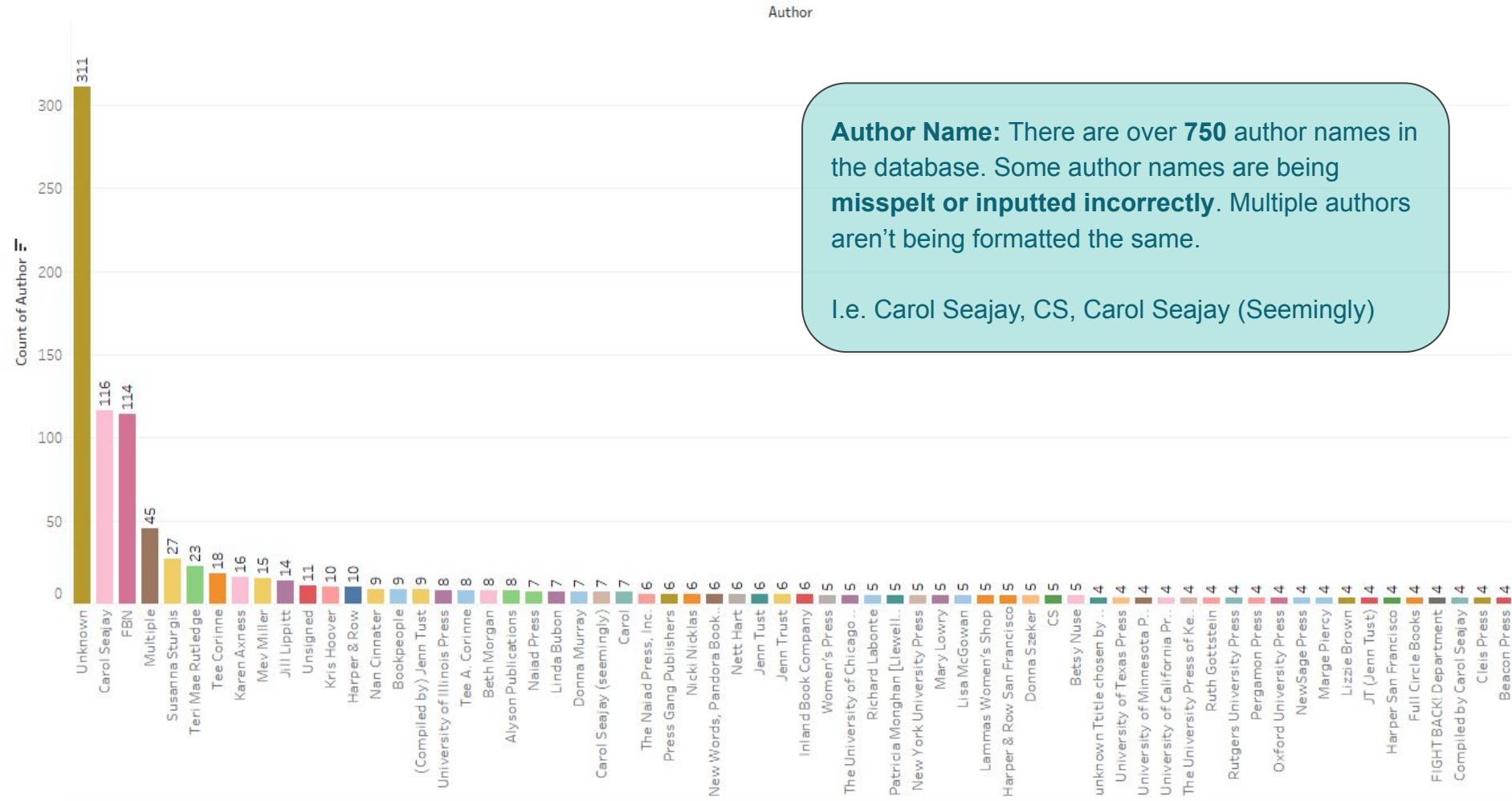
Month
Null
3
?
(Sep/Oct?)
April
August
August
Dec
December
Feb
Jan/Feb
January
June
Mar/Apr
March/April
May-June
May/June
Nov
Nov/Dec.
November
November/December
October
Sept / Oct
Sept/Oct
September

Data That Need to Be Standardized

	Article Length	Issue Length
count	1836	1836
mean	3	93
std	6	40
min	1	6
25%	1	72
50%	1	92
75%	2	132
max	88	158

- **Issue number:** Issue numbers that are **slashed** (i.e. Issue %) are being **formatted as a date**, we will have to make sure it gets formatted as a number. Do these issues **have to be slashed** can they each have their own row?
- **Month:** Use **full name of month** selected from a drop down menu
- **Continued Articles:** The last page of a continued article is being put in as a page end, this is giving us **inaccurate article length**. One article was recorded being **88 pages long!**

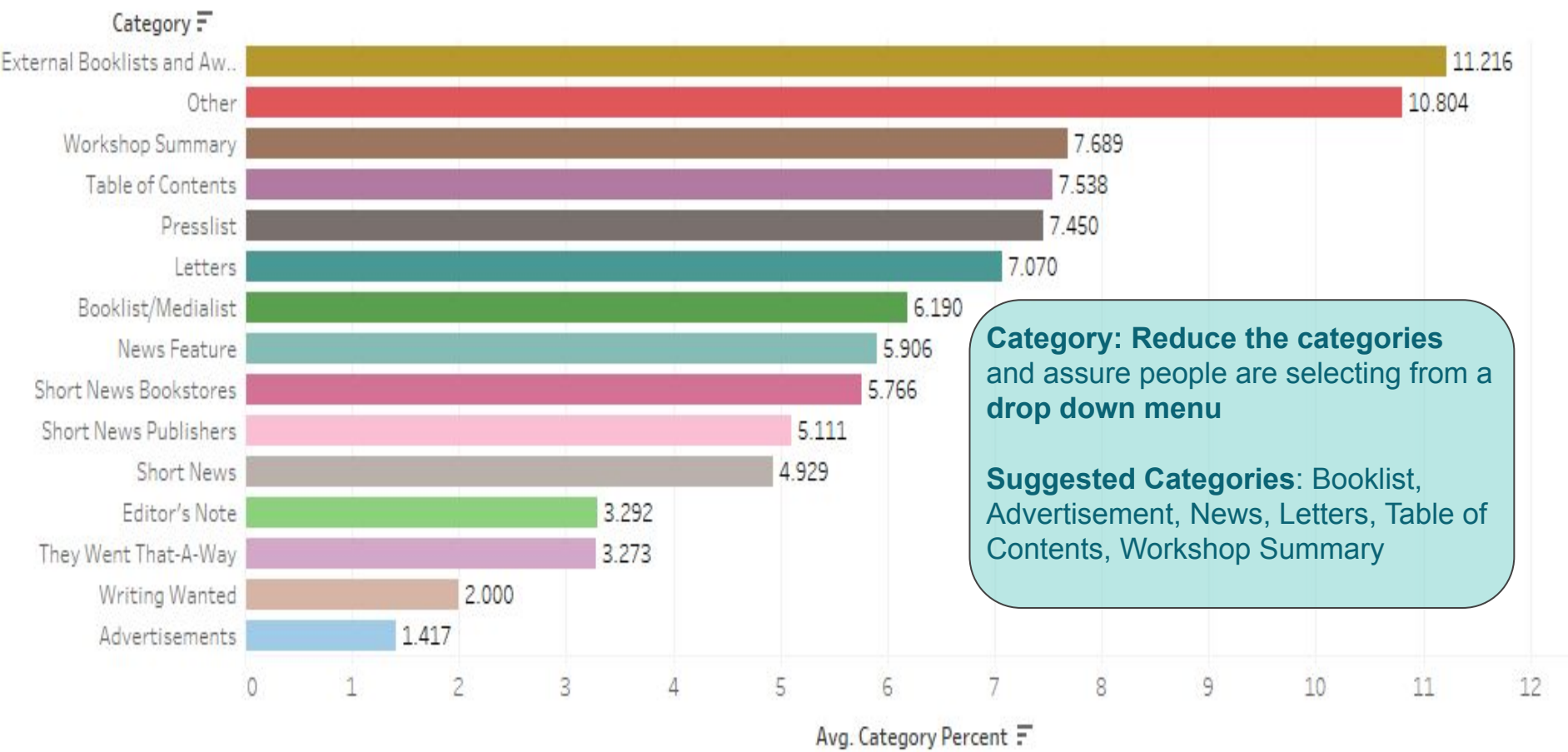
Author Count (1970s)



Author Name: There are over **750** author names in the database. Some author names are being **misspelt or inputted incorrectly**. Multiple authors aren't being formatted the same.

I.e. Carol Seajay, CS, Carol Seajay (Seemingly)

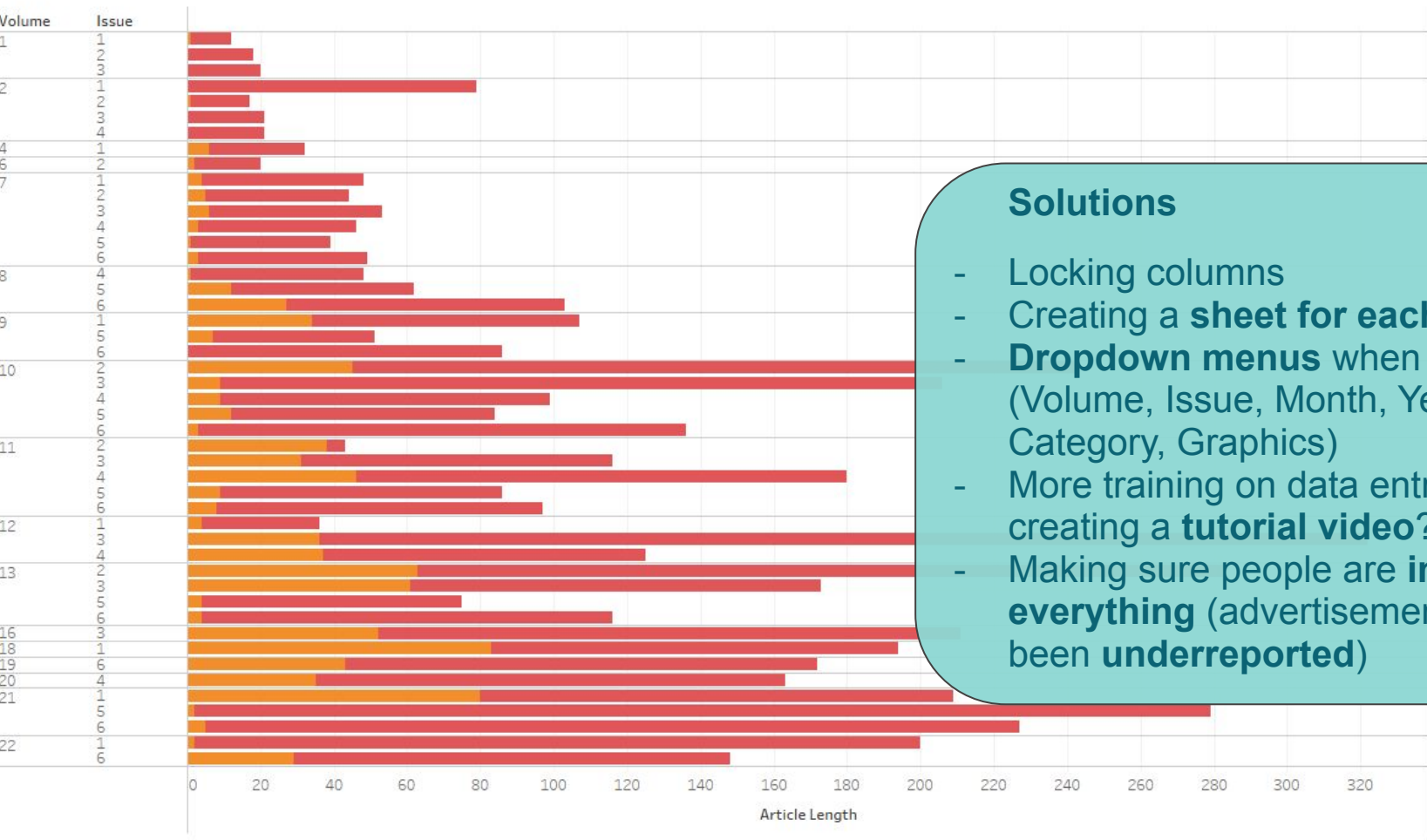
Category Percent in Issues



Category: Reduce the categories and assure people are selecting from a drop down menu

Suggested Categories: Booklist, Advertisement, News, Letters, Table of Contents, Workshop Summary

Advertisements vs Not Advertisements

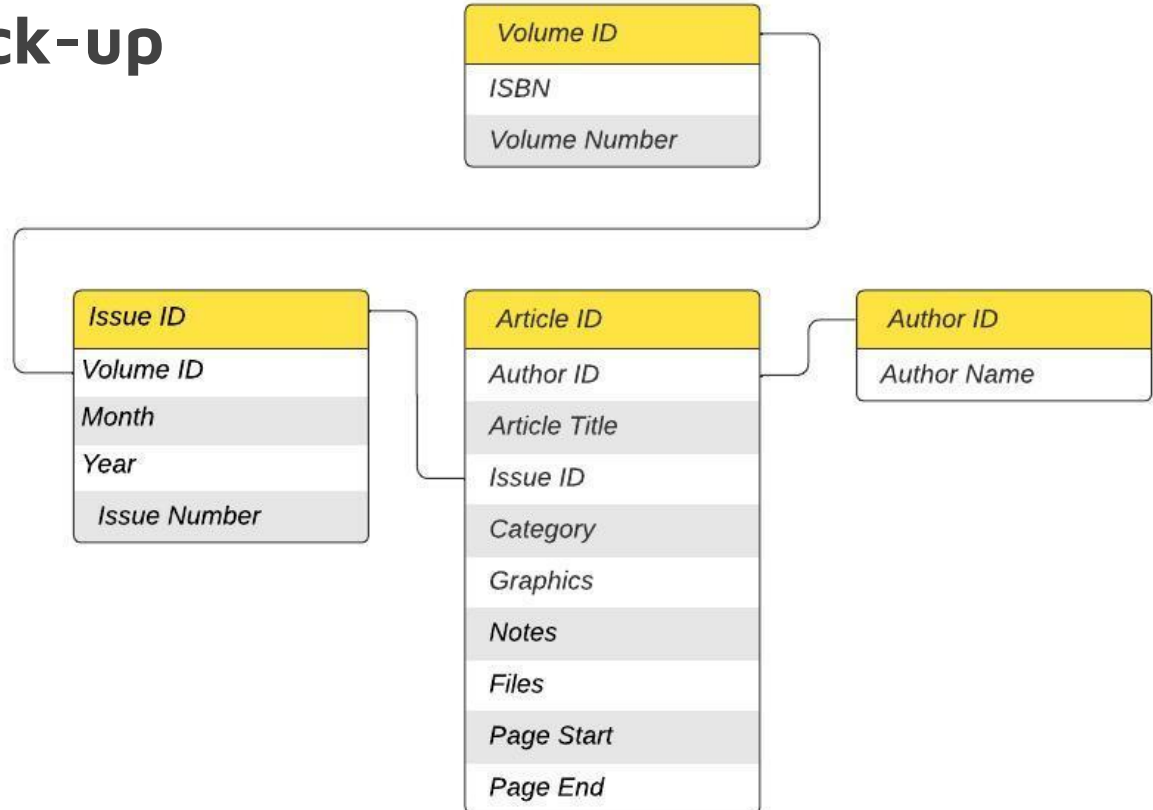


Solutions

- Locking columns
- Creating a **sheet for each volume**
- **Dropdown menus** when possible (Volume, Issue, Month, Year, Category, Graphics)
- More training on data entry, perhaps creating a **tutorial video**?
- Making sure people are **inputting everything** (advertisements have been **underreported**)

Entity Relationship Diagram Mock-up

Currently, FBN data is being kept in a flat excel file. Eventually, we would like to keep this information in a relational database using tables. This will allow the information to avoid repetition be more easily managed, modified, and searched. To the right is a mock-up of a potential table layout we could use in a relational database.



Closing Ideas





What comes next?

- Continue considering trends we can visualize in the data
- Collecting more data by completing issues
- Creating a new google sheet and saving the current one
- Scraping Issues
- Moving data to a relational database



Resources

The Manga Guide to Databases is an Excellent resource to learn more about database management.

Takahashi M, Azuma S. Databases
Databases Databases Th E M a N G a G U
I de TM To. Accessed August 23, 2023.
<https://oberstar.eu.org/share/Documents/The-Manga-guide-to-databases.pdf>

Data Access

Below are links to my Tableau Public Access and Github where I am storing data related to the project

<https://github.com/SSGrasland/FBN->

<https://public.tableau.com/app/profile/salome.grasland>