“The more that you read, the more things you will know. The more that you learn, the more places you’ll go.” – Dr. Seuss

Project Gutenberg is a digital library with over 75,000 books. Launched in 1971, it was the first eBook project on the Internet and remains one of the oldest online content providers still operating today.

This week’s #TidyTuesday dataset involved a lot of text data, so I expected to do some string manipulation. As I explored, I found the usual classics like Moby Dick, Romeo and Juliet, Pride and Prejudice, and The Bible. There were also a surprising number of historical texts from countries around the world.

I’ve been eager to try mapping again, so I decided to visualize how many books in the library are about countries in Latin America. I used library(sf) to build a simple map and overlaid the counts of how many books had the country’s name in the subject list.

The dark background was loosely inspired by a stunning visualization X did for the DnD project last week.

Check out my full notebook on GitHub: https://github.com/stevenvillalon/portfolio/tree/main/TIDYTUESDAY/2025-06-03-GUTENBERG

#DataScience #RStats #TidyTuesday #LatAm #BookNerd #DataViz