Best selling Books: Data Analysis

Weekly Update

Data collected from:

GOODREADS

https://github.com/andre-st/goodreads-toolbox/blob/master/help/savreviews.md

Collected review data from Goodreads as I was having a lot of issues with collecting Amazon review data.

In a .txt file.

TODO: Collect for all 51 books on Amazon.

```
2020/10/01 #3575175883

As with Kristen Hannah&quot;s previous books, this started with so much promise....until the female protagonist had sex and predictably became pregna

------------------------------------------------------------
2020/09/08 #3501576000

Thank you to Net Galley and St. Martin's Press for a DARC of this title in exchange for my honest review. Kristin Hannah won my heart with this origi

------------------------------------------------------------
2020/09/05 #3532836691

Â <a href="http://readmore-sleepless.blogspot.com/" rel="nofollow noopener">\n  <br>\n  <img src="https://i.gr-assets.com/images/S/compressed.photo.g
rc="https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/hostedimages/1472313906i/20179658.png" class="gr-hostedUserImg"></a><a href="mail

------------------------------------------------------------
2020/10/04 #3551012373

Reading

------------------------------------------------------------
2020/08/15 #3478962678

Boy howdy, this one&apos;s pretty bleak, especially considering the current economic climate. That said, it was a strong story about an indomitable wo

------------------------------------------------------------
```
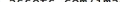
Amazon Data Scraping Attempts

Not eligible for Amazon API/still couldn't figure out BeautifulSoup (headings).

Tried Octoparse, ParseHero, two different python libraries to no avail.

Any free suggestions?

Plans for Next Meeting:

- Goodreads Reviews.
- Figure out Amazon Reviews and Product Details.
- Potentially start clustering.