

Concept

Looking at the dataset, one interesting characteristic that Seattle Public Library has over others is the cataloging of checkouts, and their timespan. To push this idea further, one could say that checkout and checkin time could be effectively seen as an absence in the collection, a missing piece of the whole that is the libraries existence. Taking this into mind, I found it interesting to have the ability to mark, and eventually visualize such an abstract concept as absence.

Approach

My first attempt at this was to look for all the database entries where the checkin was empty or NULL, however this only yielded entirely NULL entries. I realize that a database entry is completed with a checkout *and* checkin.

I then changed my query to look at time differences between checkouts and checkins that would be considered greater than normal. Initially I looked at differences in terms of days, then eventually moved into years. But still, because the dataset was so large, I had to further narrow my search to the dewey 700 and 800 sections of the library. I was mainly interested in looking a leisure books, which would add a more human element to their absence from the catalogue.

Query

To gather a larger set of data, I did two separate queries for each hundred section of the dewey classes.

```
SELECT
  title,
  cout,
  cin,
  deweyClass,
  (TIMESTAMPDIFF(YEAR, cout, cin))
FROM
  spl2.inraw
WHERE
  (itemtype = 'acbk'
  AND (deweyClass > 700 AND deweyClass < 800)
  AND (TIMESTAMPDIFF(YEAR, cout, cin) > 1)
  AND (YEAR(cout) > 2005 AND YEAR(cout) < 2014))
```

```
SELECT
  title,
  cout,
  cin,
  deweyClass,
  (TIMESTAMPDIFF(YEAR, cout, cin))
FROM
  spl2.inraw
WHERE
  (itemtype = 'acbk'
  AND (deweyClass > 800 AND deweyClass < 900)
  AND (TIMESTAMPDIFF(YEAR, cout, cin) > 1)
  AND (YEAR(cout) > 2005 AND YEAR(cout) < 2014))
```

The first query took about 26.935 sec and resulted in of 6312 rows data

The second query took 21.370 sec and resulted in 4792 rows

Analysis

Since my concept was mainly about visualizing a concept within the collection, there was not too much that I could do to analyse the data. Usually books were replaced or returned after two to three years, then it gradually died down.

