1 Data Query

I chose to explore the relationship between banned books and their popularity before, during, and after being "banned". The American Library Association puts out a list every year of the top ten most frequently challenged books and their reasons for being challenged. My goal is to glean any kind of connection between the popularity of a particular title and it's addition or placement on the list. The list of the top most frequently challenged books can be found at http://www.ala.org/bbooks/frequentlychallengedbooks/top10.

As a note on how these lists are generated, the ALA says "The Top Ten Most Frequently Challenged books list is compiled by the Office for Intellectual Freedom (OIF) and is based on anecdotal data derived from media stories and voluntary reports sent to OIF about book challenges in communities across the United States. Challenges are documented requests to remove materials from schools or libraries, thus restricting access to them by others. In some cases OIF may get numerous details about who challenged a book, why they are complaining about the book, what happened during the challenge, and the current status of the book. In other cases, few details are supplied beyond the fact of the challenge and the reasons for the challenge. Sometimes OIF receives information as the challenge is happening, and sometimes OIF receives an online report years later. All of this can affect the total number of challenges reported in any given year and how we inform the public about those challenges. Thus the Top Ten Most Frequently Challenged Books list should be seen as a snapshot of the reports OIF receives and not an exhaustive report."

To generate this particular query I first pulled the list of the Top 10 most frequently challenged books from each year during 2006-2014 (the 2015 list is not out yet), and the library checkout data only goes back to 2006. I put these into a spreadsheet and made a pivot grid so I could easily view duplicates (see BannedBooks.xlsx). From there, I modified the data in excel to extrapolate the year by year breakdown of the checkout frequency for each title.

1.1 Syntax

```
1  SELECT title, checkouts.checkOut, checkouts.bibNumber, checkouts.itemType
   FROM spl3.`_rawXmlDataCheckOuts` AS checkouts
2  WHERE ((checkouts.title LIKE "%bad␣boy␣can␣be␣good%")
3  OR (checkouts.title LIKE "stolen␣life␣a␣memoir%")
4  OR (checkouts.title LIKE "starting␣with␣alice")
5  OR (checkouts.title LIKE "%tango␣makes␣three%")
6  OR (checkouts.title LIKE "%athletic␣shorts%")
7  OR (checkouts.title LIKE "%bless%ultima%")
8  OR (checkouts.title LIKE "bone␣vol␣1%")
9  OR (checkouts.title LIKE "brave␣new␣world")
10 OR (checkouts.title LIKE "adventures␣of␣captain␣underpants%")
11 OR (checkouts.title LIKE "crank")
```

```
12  OR (checkouts.title LIKE "drama")
13  OR (checkouts.title LIKE "fifty shades of grey")
14  OR (checkouts.title LIKE "flashcards%life%")
15  OR (checkouts.title LIKE "gossip girl")
16  OR (checkouts.title LIKE "golden compass")
17  OR (checkouts.title LIKE "i know why the caged bird%")
18  OR (checkouts.title LIKE "its perfectly normal%")
19  OR (checkouts.title LIKE "looking for alaska")
20  OR (checkouts.title LIKE "lush")
21  OR (checkouts.title LIKE "my moms having a baby%")
22  OR ( checkouts.title LIKE "my sisters keeper a novel")
23  OR ( checkouts.title LIKE "nickel and dimed on%")
24  OR ( checkouts.title LIKE "olives ocean")
25  OR ( checkouts.title LIKE "persepolis")
26  OR ( checkouts.title LIKE "revolutionary voices%")
27  OR ( checkouts.title LIKE "saga volume one")
28  OR ( checkouts.title LIKE "scary stories to tell%")
29  OR ( checkouts.title LIKE "absolutely true diary%")
30  OR (checkouts.title LIKE "adventures of huck%")
31  OR (checkouts.title LIKE "bluest eye")
32  OR (checkouts.title LIKE "catcher in the rye")
33  OR (checkouts.title LIKE "chocolate war")
34  OR (checkouts.title LIKE "color of earth")
35  OR (checkouts.title LIKE "color purple")
36  OR (checkouts.title LIKE "earth my butt%")
37  OR (checkouts.title LIKE "glass castle")
38  OR (checkouts.title LIKE "hunger games")
39  OR (checkouts.title LIKE "kite runner")
40  OR (checkouts.title LIKE "perks of being%")
41  OR (checkouts.title LIKE "thirteen reasons why%")
42  OR (checkouts.title LIKE "to kill a mock%")
43  OR (checkouts.title LIKE "ttyl%" AND callNumber LIKE "%myracle%")
44  OR (checkouts.title LIKE "twilight")
45  OR (checkouts.title LIKE "uncle bobby%")
46  OR (checkouts.title LIKE "what my mother does%"))
47  AND ((checkouts.checkout BETWEEN "2005-01-01" AND "2014-12-31")
48  AND (checkouts.`itemType` LIKE  "%bk%"))
49  ORDER BY title, checkout
```

1.2 Query Details

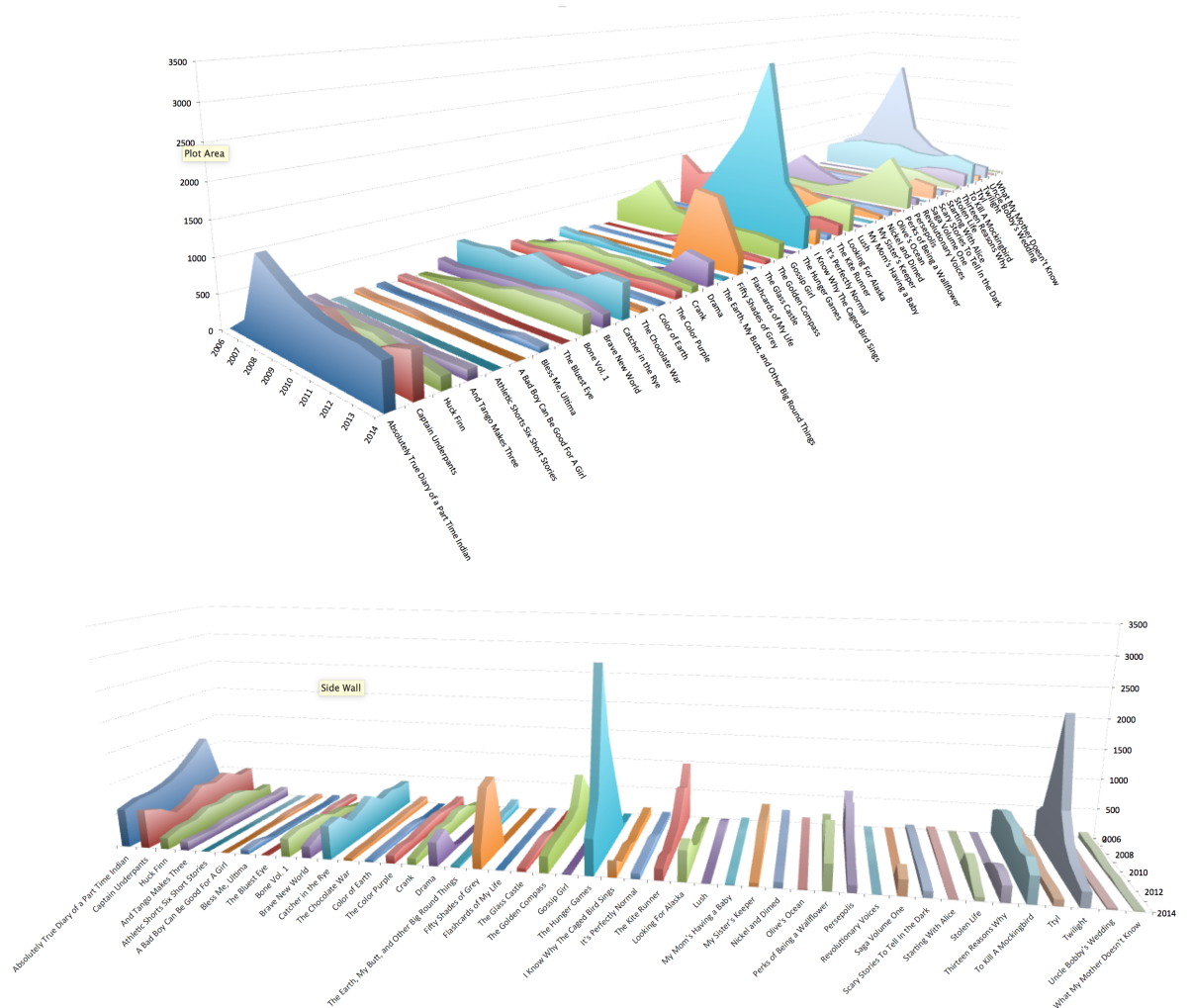The query took approximately 513.74 seconds to run when fully completed. It was built line by line to include each book title on the list, and the filtered for those titles. The query essentially pulls each checkout for each title for the time period between 2006 and 2014. This returned 80,875 checkouts for all of these titles over that time period. From there, I created a data table that broke the

titles down by year and did a frequency count of each title and checkout year.

1.3   Final Results

	Totals	2006	2007	2008	2009	2010	2011	2012	2013	2014	Totals
absolutely true diary of a part time Indian	6054	0	240	1229	1027	846	737	681	682	612	6054
adventures of Captain Underpants an epic novel	4641	491	474	614	538	596	425	351	529	623	4641
adventures of Huckleberry Finn	2342	195	191	302	349	305	335	240	243	182	2342
And Tango makes three	1304	160	139	159	156	148	153	115	130	144	1304
Athletic shorts six short stories	64	16	14	11	10	8	2	3	0	0	64
bad boy can be good for a girl	260	64	40	55	58	21	16	3	0	3	260
Bless me Ultima	440	56	23	28	39	41	43	36	99	75	440
bluest eye	446	95	95	111	75	38	18	4	5	5	446
Bone Vol 1 Out from Boneville	1898	29	67	124	206	218	314	314	333	293	1898
Brave new world	1946	189	149	183	201	224	227	301	286	186	1946
catcher in the rye	4128	396	395	442	399	562	455	375	565	539	4128
chocolate war	709	119	93	105	86	79	76	54	58	39	709
color of Earth	320	0	0	0	55	74	55	57	50	29	320
color purple	1390	156	131	163	191	177	160	127	157	128	1390
Crank	1528	66	123	236	260	195	223	179	143	103	1528
Drama	919	0	0	0	0	0	0	74	468	377	919
earth my butt and other big round things	707	154	115	126	79	64	63	47	29	30	707
Fifty shades of Grey	2682	0	0	0	0	0	0	1228	1159	295	2682
Flashcards of my life a novel	123	32	21	19	18	10	13	8	2	0	123
glass castle	507	0	0	0	16	115	130	104	82	60	507
golden compass	3900	362	581	886	443	356	346	333	327	266	3900
Gossip girl	205	51	39	62	37	0	0	0	0	16	205
hunger games	8540	0	0	70	658	1247	1858	3046	1079	582	8540
I know why the caged bird sings	1108	132	104	102	98	56	128	106	111	271	1108
Its perfectly normal a book about changing bodies growing up sex and sexual health	819	67	67	76	73	131	117	99	97	92	819
kite runner	4040	960	652	747	489	307	265	183	251	186	4040
Looking for Alaska	1860	9	33	105	120	87	172	302	541	491	1860
Lush	314	0	79	82	58	42	25	14	10	4	314
My moms having a baby	333	41	38	54	41	41	35	30	35	18	333
My sisters keeper a novel	2169	305	299	338	368	325	201	146	126	61	2169
Nickel and dimed on not getting by in America	1311	218	149	178	171	135	114	119	131	96	1311
Olives ocean	475	102	73	72	61	51	45	26	26	19	475
perks of being a wallflower	3085	184	127	211	179	179	297	549	941	418	3085
Persepolis	2945	611	347	639	399	240	248	169	171	121	2945
Revolutionary voices a multicultural queer youth anthology	7	2	1	2	1	1	0	0	0	0	7
Saga Volume one	615	0	0	0	0	0	0	22	339	254	615
Scary stories to tell in the dark	899	70	93	102	89	87	112	124	120	102	899
Starting with Alice	359	57	66	55	56	53	33	18	11	10	359
stolen life a memoir	762	0	0	0	0	0	295	262	140	65	762
Thirteen reasons why a novel	1035	0	0	0	0	0	169	290	318	258	1035
To kill a mockingbird	4768	398	534	619	606	625	507	491	563	425	4768
Ttyl	1095	126	106	175	144	127	93	106	115	103	1095
Twilight	7089	431	627	1379	2301	907	537	371	298	238	7089
Uncle Bobbys wedding	273	0	0	62	43	35	37	35	31	30	273
What my mother doesnt know	461	100	82	84	56	33	49	22	19	16	461

1.4 Query Analysis

The numbers in red in the results table correspond to the years that the book was on the most frequently challenged list. It is quite easy to see from just glancing through the table rows that for many of the titles, the year or years that the title made the list corresponded to a year of a very high checkout rate. There seems to be at least some positive correlation between the book being on the list and the rate at which it is checked out. Of the 45 titles that were in my final set of data, 30 of the them were placed on the list during one of the top three years they were most checked out. 15 were on the list the exact same year as the year they were most checked out.

It is difficult to say whether a title receives more checkouts because of its

placement on the frequently challenged list, or whether being placed on the frequently challenged list increases the number of checkouts for a particular title. It seems logical to conclude that some mixture of both of these is at play. There are several examples where it seems clear that the current popularity of a title is what landed them on the frequently challenged list. The Golden Compass fits this model. The book was originally published in 1995, but a movie was released in 2007. 2007 was the only year the book made the frequently challenged list and there were 581 checkouts that year, second only to the 886 checkouts in 2008. It seems reasonable to conclude that the popularity of the movie increased the popularity of the book, at which point people decided to start filing complaints about it. Both the number of checkouts and the number of challenges toward the book have subsequently decreased.