

Question:

My question involves an interest in how both time of day, and day of week affect peoples lending habits. What is the average number of books borrowed in a day and what is the standard deviation that each day of week has from that average. Is the distribution different for different genres (i.e. comic books being borrowed more in the afternoon, while home goods books being borrowed during the day)

Query:

To gather my data, I ran this query for each day of the week (if we were able to store tables, I might have been able to do this all in one query, but since we are not I decided to take the engineering route of “well... it works”).

```
SELECT count(case when deweyclass >= 740 or deweyclass<750 then 1 else 0 end)
FROM spl2.inraw
where date(cout)>="2013-03-01" and date(cout)<"2013-04-01" and dayofweek(date(cout))=7
group by hour(cout)
order by hour(cout)
```

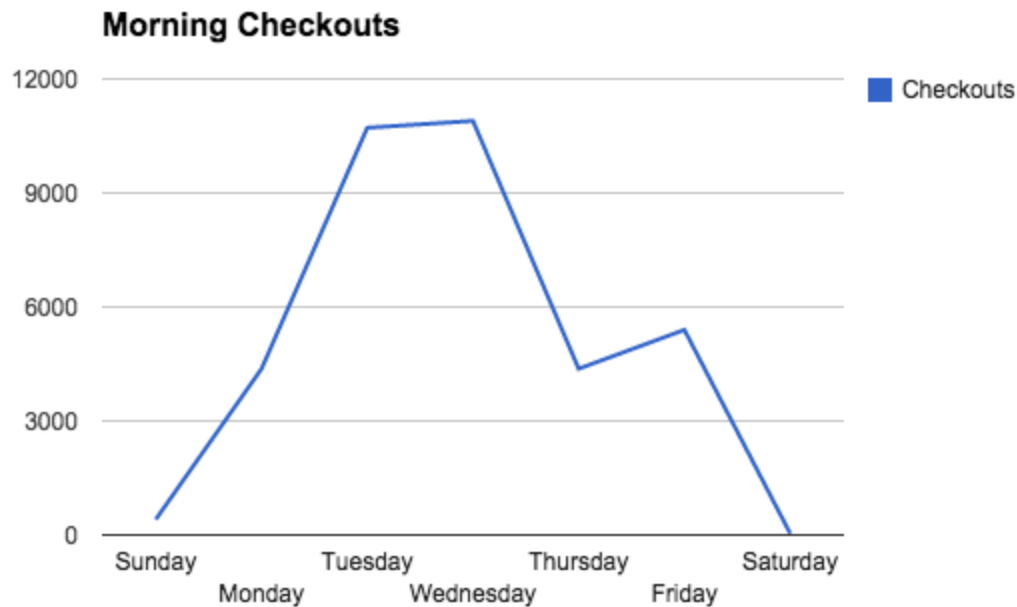
Processing Time

On average, each query took between 57 and 59 seconds.

Analysis

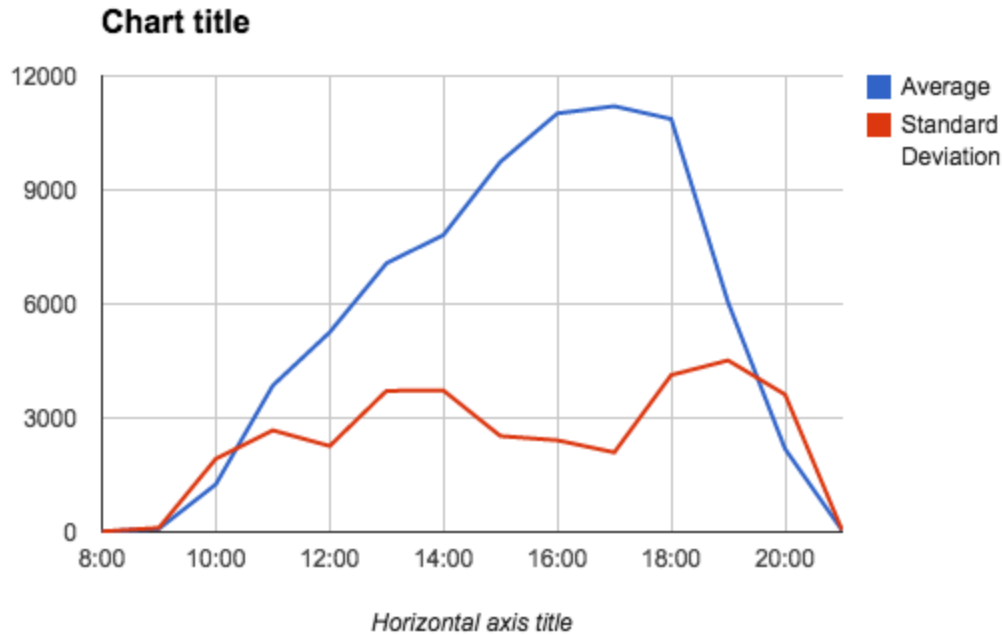
	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
8:00		1	14	34	1	1	
9:00		15	273	83	15	40	
10:00	16	195	4243	3885	195	240	1
11:00	396	4163	6186	6895	4163	5117	42
12:00	214	5771	5575	6699	5771	6557	6195
1:00	36	7093	7796	7536	7093	7179	12782
2:00	985	8115	7942	7705	8115	8025	13827
3:00	12676	7899	8223	8828	7899	8633	14040
4:00	14065	9352	10176	10092	9352	9162	14902
5:00	13794	10637	8860	10166	10637	9876	14464
6:00	16705	10635	6598	5098	10635	11513	14895
7:00	1613	7792	6109	4951	7792	161	13839
8:00	21	7482	83	59	7482	5	177
9:00		90		3	90	1	

So right off the bat, the first thing that I noticed was that there was a defined curve of when people check out books in the morning. It would seem that for some reason, in the middle of the week people are significantly more likely to check out books than on a Monday or a Friday (and of course extremely unlikely on a weekend since there are fewer morning hours for people to check out books).



Unfortunately, I was unable to find any interesting information from looking into the daily number of evening checkouts, except for the fact that almost all checkouts on Saturdays are done in the evening.

Another very interesting correlation is noticed here when I compare the average with the standard deviation. Notice that between 3:00PM and 5:00 PM there is an extreme dip in the standard deviation, which directly linked to an extreme growth in the average. Unlike 6:00 to 7:00 where the average is still high, but mostly being held up by Friday, Saturday, and Sunday, 3:00 to 5:00 are high every single day of the week. This seems to point towards an interest in drawing books among students, since these are the hours that students would just be getting out of school and adults would still be in work.



In the future, I intend to run these same queries on different genres of books to see if other interesting correlations could be found. Would science books have the same peak hours as home and living books? Are there significant changes in standard deviations during the summer when students are no longer in school? Questions like these will be addressed in future projects, and will hopefully draw some interesting correlations.