MAT259 Proj 1: Data Query

Song Gao

Geography Department (Email: sgao@geog.ucsb.edu)

Question: Does the rainy weather in Seattle influence on the check-in and check-out behavior in the Seattle Public Library?

Seattle is famous for its rainy weather and the average annual rainfall is approximately 38 inches. According to the *Weatherdb* website[1] Seattle has a relatively high amount of days per year with precipitation (158), compared to such rainy places as New York (119), Boston (127), and Nashville (119). All cities that get an average of about 16% more rain per year than Seattle, but also average between them about 36 less days a year of precipitation. Because of the entertainment industry producing things like Sleepless in Seattle, Frasier, and the works which portray it as such, most people think to go outside in Seattle without an umbrella is tantamount to committing troublesome. It inspired me to investigate whether the rainy weather did affect the users' check-in and check-out behavior in the Seattle Public Library. The intuition is that the probability of check behavior will decrease if it does rain heavily outside in Seattle, which has influenced on the urban transportation and daily commuting behavior of citizens. To this end, I need to compare the number of total check-outs and check-ins with the seasonal, monthly and daily patterns of rainfalls in Seattle. The weather data can be queried and downloaded from the U.S. National Weather Service Forecast Office website[2], but it only limits on latest five years, i.e. starting from 1 January, 2009. Therefore, I will also query the library data from then on. Note that the Winter season usually starts from last December, so we also need to include the data during that time period.

MySQL Query and Explanation

(1) Monthly check-outs

SELECT year(cout), month(cout), count(*)   /*count the number of transactions by month*/

FROM inraw   /*select the data from the target table inraw*/

WHERE date(cout)>"2008-12-1" AND date(cout)<"2009-12-31"   /*set the time period*/

---

[1] http://average-rainfall.weatherdb.com/
[2] http://www.nws.noaa.gov/climate/index_nonjs.php?wfo=sew

GROUP BY year(cout), month(cout)   /*Group the results by different year, month*/

ORDER BY year(cout), month(cout); /*Ranking the results by year, month*/

(2) Monthly check-ins

SELECT year(cin), month(cin), count(*)

FROM inraw

WHERE date(cin)>"2008-12-1" AND date(cin)<"2009-12-31"

GROUP BY year(cin), month(cin)

ORDER BY year(cin), month(cin);

(3) Daily check-outs

SELECT dayofyear(cout), count(*)   /*count the number of transactions by day*/

FROM inraw   /*select the data from the target table inraw*/

WHERE date(cout)>"2009-01-01" AND date(cout)<"2009-12-31"   /*set the time period*/

GROUP BY dayofyear(cout)   /*Group the results by different days*/

ORDER BY dayofyear(cout);   /*Ranking the results by day*/

(4) Daily check-ins

SELECT dayofyear(cin), count(*)

FROM inraw

WHERE date(cin)>"2009-01-01" AND date(cout)<"2009-12-31"

GROUP BY dayofyear(cin)

ORDER BY dayofyear(cin);

Processing Time:

413.824 seconds for monthly check-outs in year 2009

418.442 seconds for monthly check-ins in year 2009

497.097 seconds for daily check-outs in year 2009

511.693 seconds for daily check-ins in year 2009

Results:

These exported SQL query results can be accessed at:

http://www.geog.ucsb.edu/~sgao/mat259/2009monthly_checkouts.csv

http://www.geog.ucsb.edu/~sgao/mat259/2009monthly_checkins.csv

http://www.geog.ucsb.edu/~sgao/mat259/2009daily_checkouts.csv

http://www.geog.ucsb.edu/~sgao/mat259/2009daily_checkins.csv

http://www.geog.ucsb.edu/~sgao/mat259/2009daily_precipitation.csv

Here, it shows the post-processing tables by combining results together:

Year2009	Precipitation (Inch)	Checkouts	Checkins
Jan.	5.4	999538	964974
Feb.	1.51	883648	914368
Mar.	4.16	949770	1004110
Apr.	3.36	924968	908917
May	3.61	905096	939789
June	0.18	951739	966204
July	0.06	957653	977699
Aug	1.16	971625	935868
Sep.	1.75	704144	768425
Oct	5.54	926774	970585
Nov.	8.96	848099	878695
Dec.	2.75	803490	847211

The seasonal patterns can be got by summarize the number of transactions by the months in each season.

Season (months)	Precipitation (Inch)	Checkouts	Checkins
Winter (12, 1, 2)	12.34	2645328	2694984
Spring (3, 4, 5)	11.13	2779834	2852816

Summer (6, 7, 8)	1.4	2881017	2814910
Fall (9, 10,11)	16.25	2479017	2617705

Comment & Analysis:

Based on the comparison of monthly precipitation with the library transactions in Fig.1 & 2, I didn't find any necessary linkage between them. Both of the check-out and check-in numbers have small fluctuation randomly around the monthly average value (Avg-cout: 902212, Avg-cin: 923070).

But for the seasonal patterns, I can clearly see the reverse correlation between the amount of rain and the library transactions, i.e., the rank of season by the numbers of check-out transactions from the largest to smallest is *Summer, Spring, Winter and Fall,* while it will got the reverse order by the amount of precipitation (in Fig. 3 & 4). It usually rains more in the fall and winter than during the summer in Seattle.

For the daily patterns in Fig. 5 & 6, a large amount of rainfall (such as large than 1 inch, about 25.4 millimeters) might be troublesome for outdoor behaviors such as going library to check-out or check-in if someone's location is far away from the library. Zooming to November 2009, the continuous rains on days 6, 7 reduce the number of check-outs and even affect the behavior on 8 (Fig. 7 & 8). Although we know that there are many factors influencing on the library transactions behaviors, we still believe the local weather could be one of them.

To sum up, I did a quantitative analysis on the patterns of check-outs and check-ins compared with the seasonal, monthly and daily amounts of rainfalls in Seattle and found the seasonal and daily influence on the transaction behaviors.
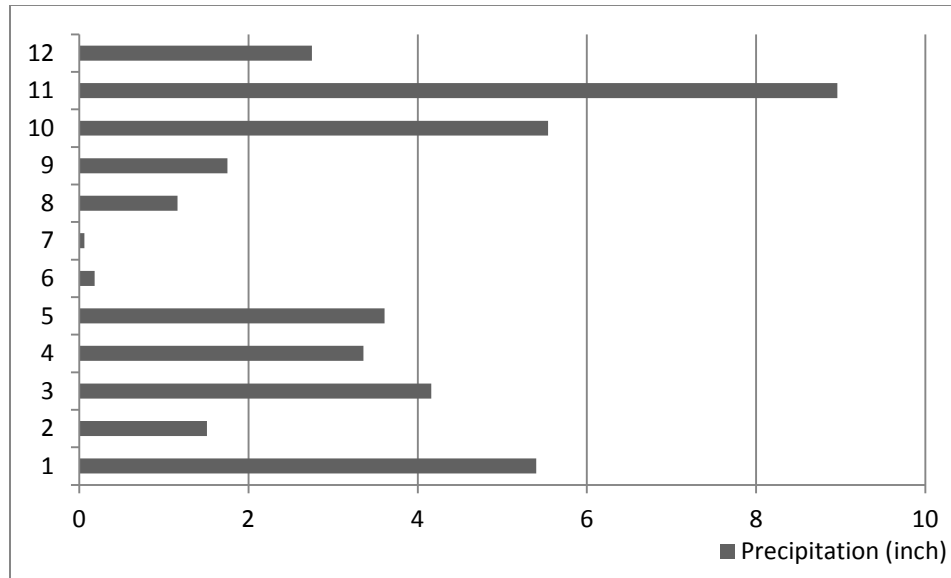
Figures

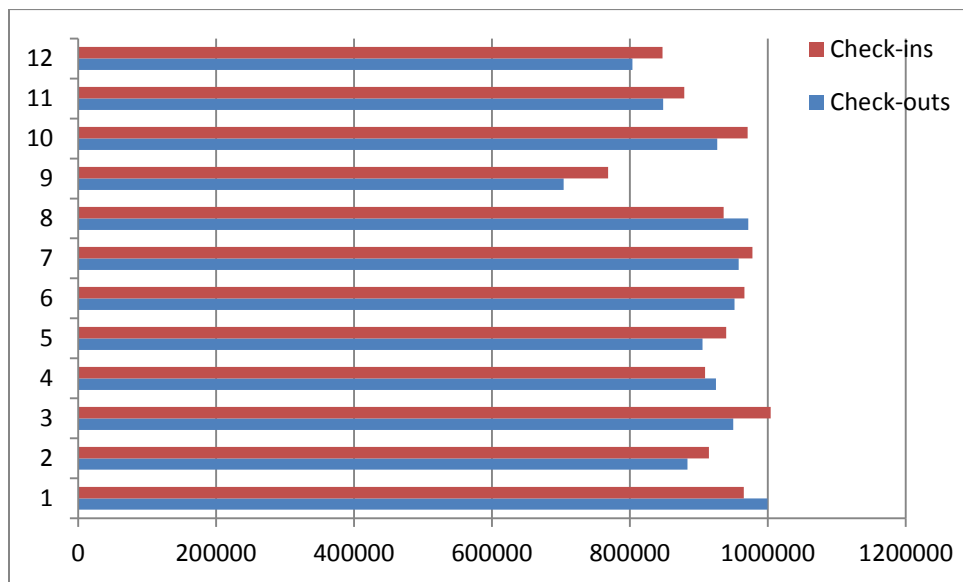Fig. 1 Seattle Monthly Precipitation in the year 2009

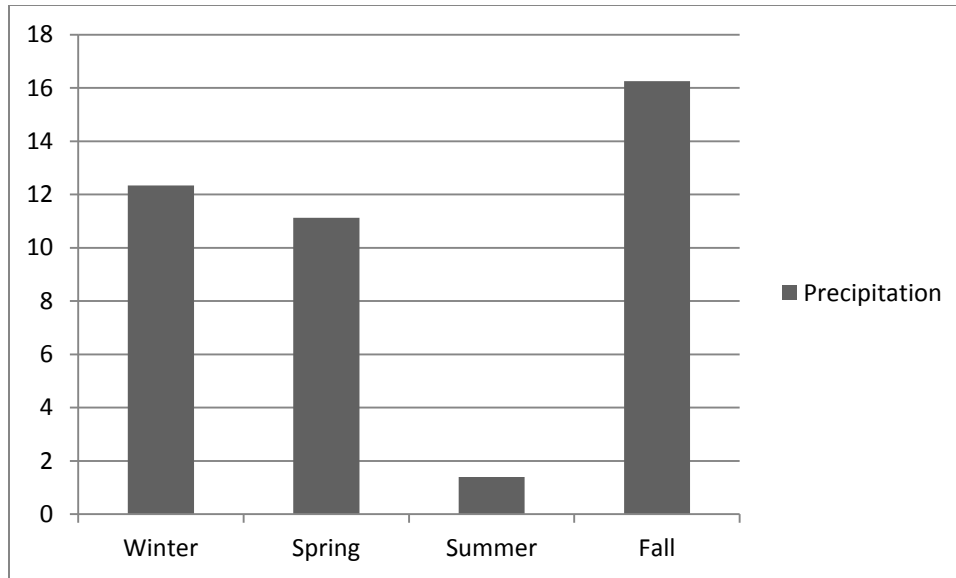Fig. 2 Library Monthly Transactions in the year 2009

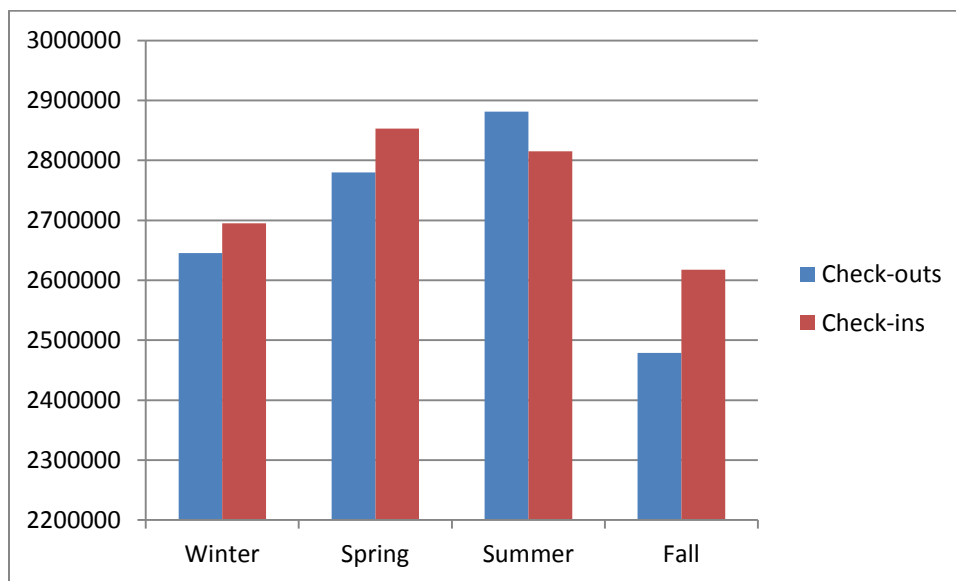Fig. 3 Seattle Seasonal Precipitation in the year 2009

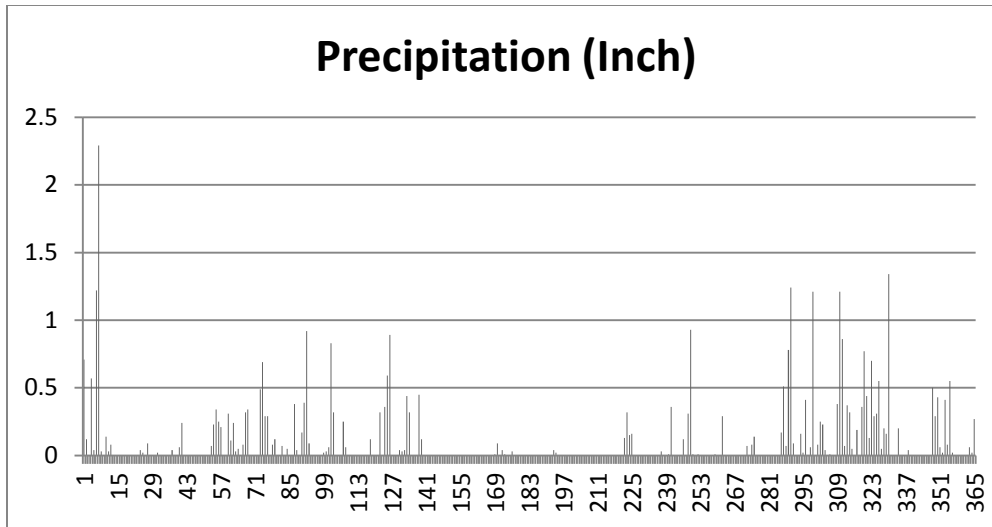Fig. 4 Library Seasonal Transactions in the year 2009

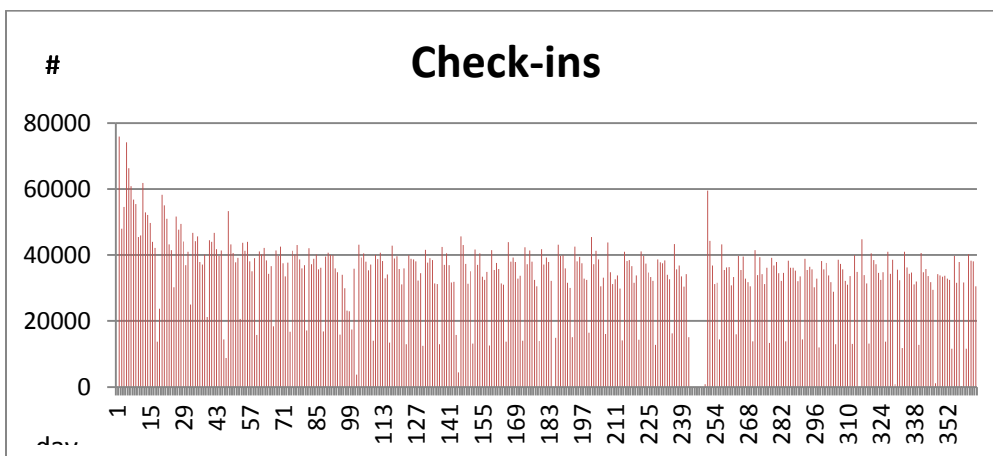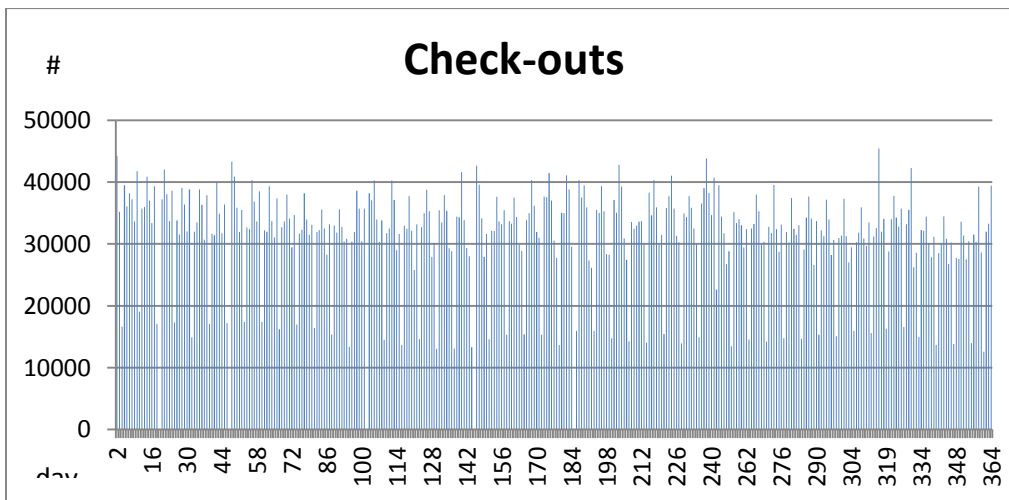Fig. 5 Seattle Daily Precipitation in the year 2009

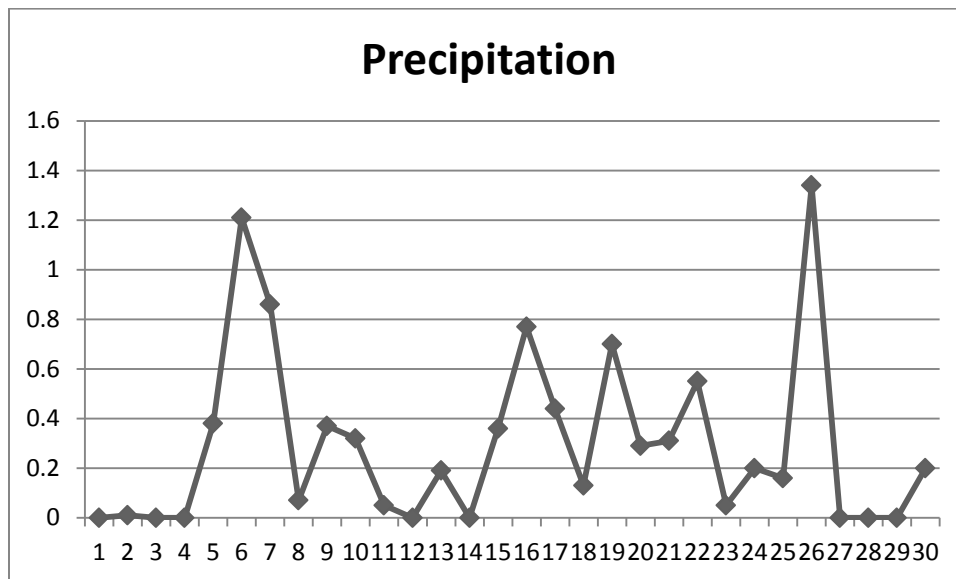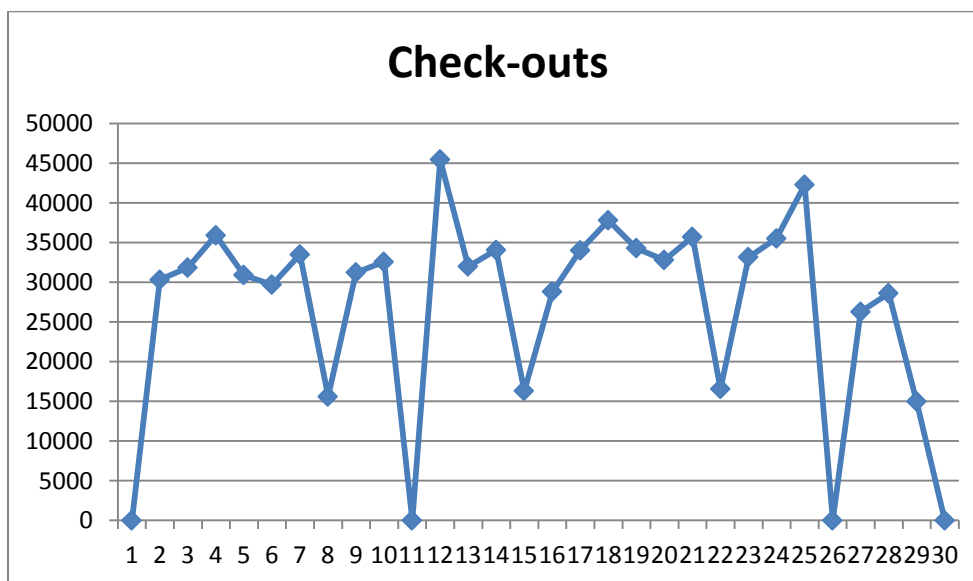Fig. 6 Library Seasonal Transactions in the year 2009

Fig. 7 Seattle Daily Precipitation in November 2009

Fig. 8 Library Daily Check-outs in November 2009