HIGH RATE OF CAR BREAK-INS DURING LATE MORNING HOURS OF SEATTLE SUMMERS

Seattle's police department has released their crime statistics for the summer months of 2014 for public access. Looking at this information we can find some interesting conclusions. I focused on the most common crime and wanted to answer the following questions

- What specific crime is most common in Seattle during this period?
- At what time of day is this particular crime most likely to happen?
- Does the time of day pattern change from month to month over the summer?

I will be using R to do my data investigations.

INTRODUCTORY INVESTIGATIONS

First of all, we'll start with the surface level data. The first thing I did was figure out which crimes were common, and if there were any standout crimes that were more common than any other. The figure below shows the top five crimes, which was easily found by a group_by table:
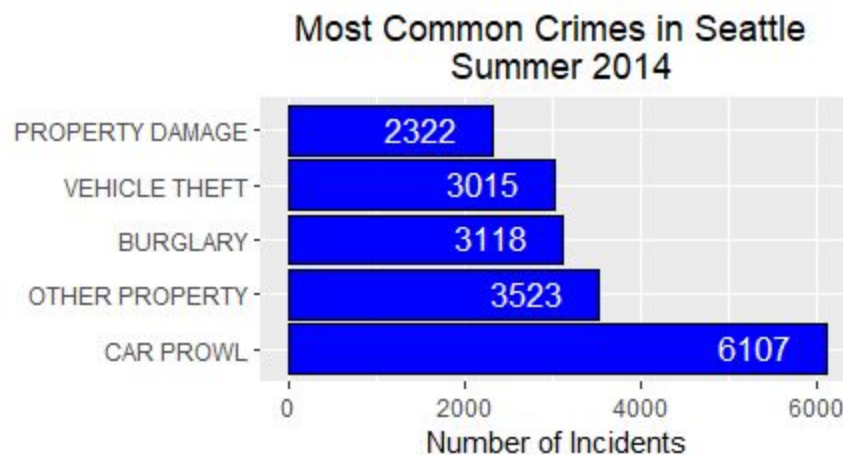
As you can see, the crime known as car prowl happens way more commonly than any other crime, close to doubling any other individual report. Car prowl, by the way, is the Seattle police team for when items are stolen from a person's car, usually via breaking into the vehicle.

CAR PROWL INCIDENTS VERSUS TIME OF DAY

So, knowing this, I went to look at if there were any particular observations that I could pull from the car prowl data. I was most interested in the times in which the crime was most likely to happen so I broke it down by hour. To do this I converted the time into a proper time(POSIXlt) variable and then extracted the hours column, thus putting the data in one hour intervals. I used a step plot to indicate that the time distribution is not truly continuous, since every hour data point is a collection of incidents that happened over the 60 minutes.
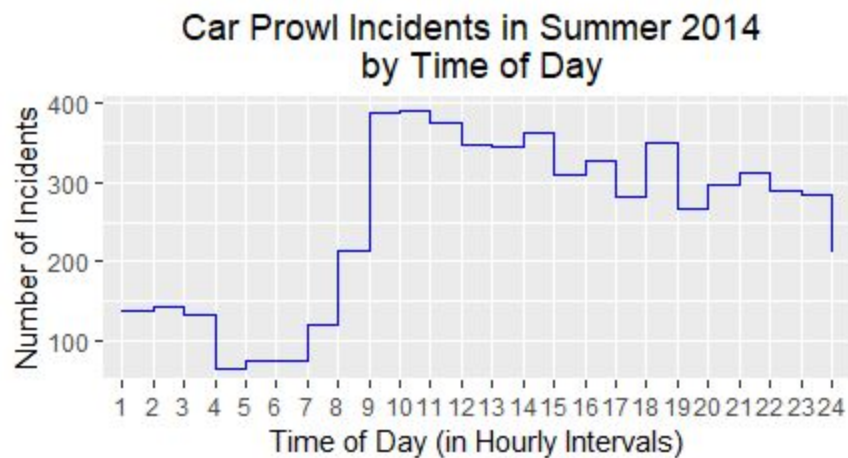

The immediate interesting thing to notice is the high spike of incidents that occurs between 7-10 am. I imagine this might come from the fact that people have commuted to work and parked over this period, resulting in a lot of unattended cars in parking lots that can be broken into. In comparison, the earlier morning hours are where people are likely to be preparing for the workday or commuting, so either in their car or in a position where they would enter them soon. The rate gently descends from there until the steep cliff that happens between late night and early morning, where I assume most of the cars are gone from the open parking lots and are instead parked in home driveways and apartment lots (which are harder to access and safely steal from).

## TIME OF DAY TRENDS CHANGING OVER THE SUMMER

What remains is to look at the way this trend changes over the months. To do this I subdivided the data by month of incident, removed the outlier months, and then created new time distribution step charts. The first thing of note is that there were slightly more incidents in June(2188), while July and August were roughly equal (1965 and 1954, respectively). The resultant plots are as follows

Car Prowl Incidents in June 2014 by Time of Day
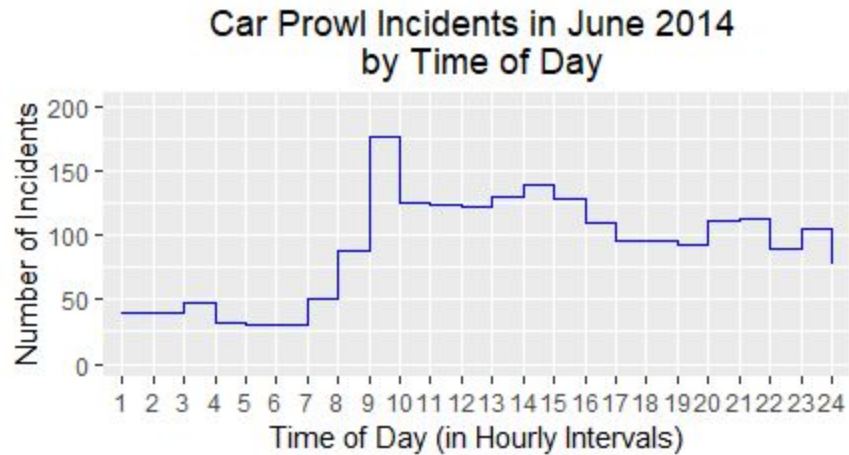
The June data roughly mirrors the structure of the overall incident rate over the summer, but with a more marked spike in the 9-10 interval range.

Car Prowl Incidents in July 2014 by Time of Day
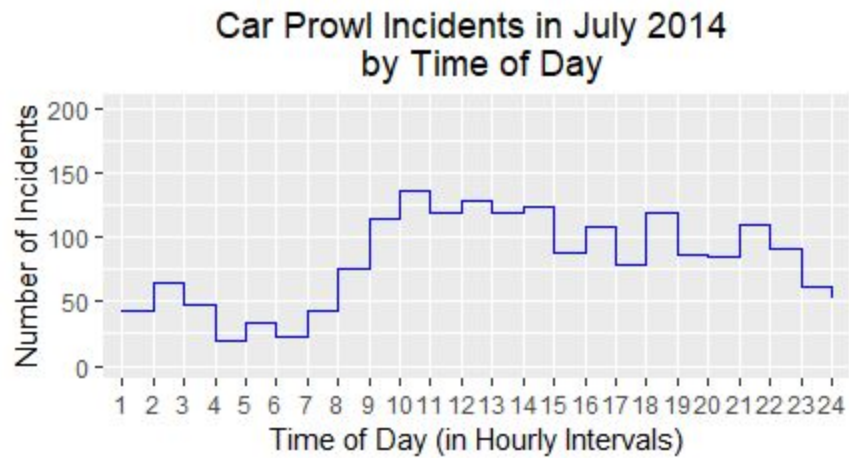
By contrast, the morning spike seen in July's data is much less pronounced. Interestingly there's a small bump at about 2am.

Car Prowl Incidents in August 2014 by Time of Day
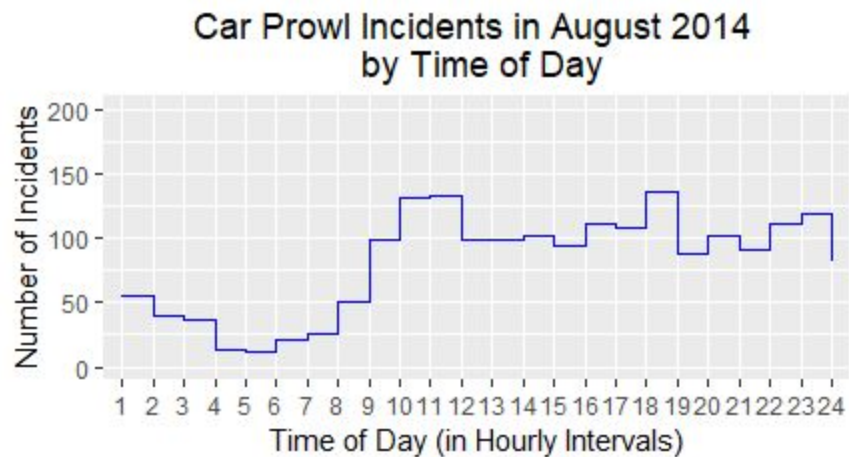
The August graph mirrors the overall pattern with a bit of shifting - the late morning peak is not particularly high and is in fact matched by a small spike around dinnertime.

TENTATIVE CONCLUSIONS AND THOUGHTS

Taking all three of these graphs in total makes me wonder if the crime statistics are influenced by summer vacations. People absent from work due to time off are not parking their cars in lots, and thus their cars cannot be broken into. July and August are the most common vacation months in the year, so are most likely to have people not commuting to work. I hypothesize that if people were parking their cars for non work reasons (errands, shopping, eating) it would be more evenly distributed throughout the midday and afternoon, which might create a result like the july and august plots.

More data would help substantiate this hypothesis. If I were to look into more data, I would be interested to see if September marked a return to the June pattern of car prowl break ins, or if it was something that was unique to the summer months