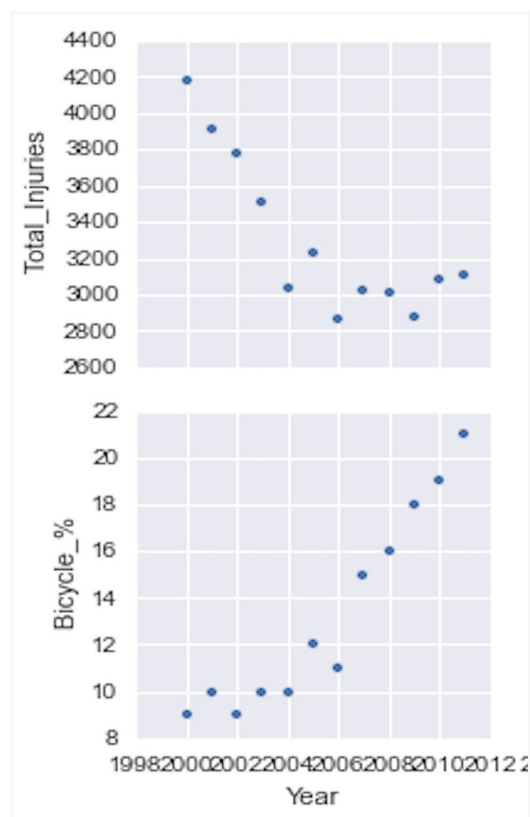


SF_DAT_15 Course Project

San Francisco Cycling Safety -- Michael Phillips

Cycling is an increasingly popular sport and mode of transit in America; and particularly in dense cities that don't require car ownership, it can be the fastest and most cost effective means of travel to work and many other destinations. In part because of the recent increase in cycling popularity, there has been a staggering increase in the percentage of all injury causing collisions that bicycles are involved in, while the total injury volume across all vehicle collisions has significantly decreased

(<http://j.mp/1lf0zJ2>):

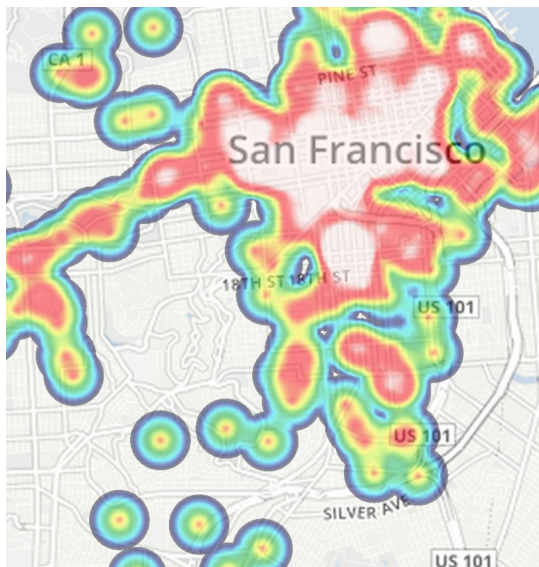


San Francisco is beginning to take steps towards making the city a more bike friendly place, but so far, many of these efforts have been controversially focused on cracking down on cyclists that violate traffic laws (<http://j.mp/1f0lfw>).

In contrast, significant attention has been given to studying particular areas where motor vehicle collisions occur, and that has greatly impacted how the city plans construction and times traffic lights (<http://j.mp/1f0zJ2>).

In this project, I analyzed reports of cycling injuries throughout the city, along with location data, in order to plot the most dangerous locations.

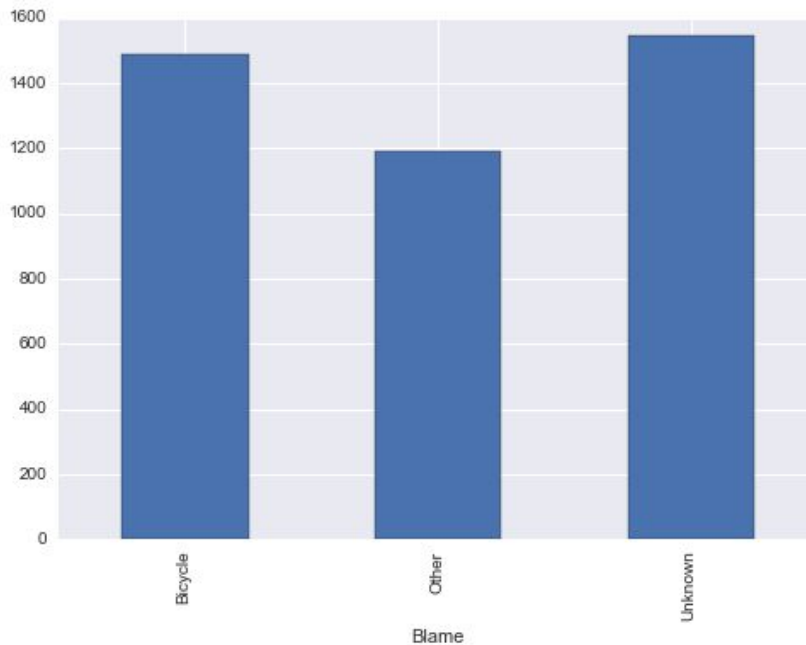
Sample JSON location data:



```
{  
  "lat":37.7724134,  
  "severity":"u'INJURY",  
  "street1":"11th St",  
  "street2":"Market",  
  "image_id":1,  
  "year":"2012",  
  "date":"u'20120419",  
  "lng":-122.4147158  
},
```

Extending beyond the location data, I also recognized interesting trends in the blame assigned in bicycle collisions. While significant efforts have been made to prevent injuries via ticketing traffic law breaking cyclists, only slightly more than half of collisions where blame is assigned are attributable to cyclists (other half being motorists and

pedestrians). Even more concerning is the fact that the majority of accidents do not have any recorded fault.

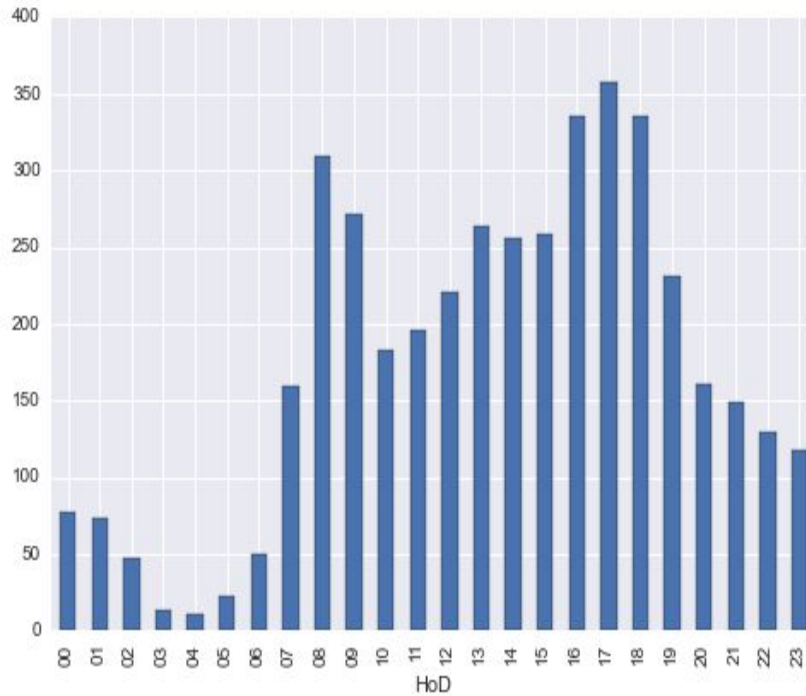


```
# Who's at fault??
accidents2.chp_vehtype_at_fault.value_counts()

def Blame(n):
    if ( n == 'Bicycle' ): return "Bicycle"
    elif ( n == '(null)' or n == 'Not Stated or Unknown (Hit and Run)' ): return "Unknown"
    elif ( n != 'Bicycle' ): return "Other"

accidents2['Blame'] = accidents2.chp_vehtype_at_fault.apply(Blame)
accidents2.groupby('Blame').chp_vehtype_at_fault.count().plot(kind='bar')
```

Additionally, there is a clear trend towards collisions during rush hour. This has significant implications for where the city should focus prevention efforts, as well as emergency response staffing plans. Further research is required to know if engineering efforts around alleviating some of the most heavily impacted rush hour roadways in San Francisco could reduce cycling accidents during those hours.



Next steps for this project are to gather more complete location and traffic data and plot time over time comparisons. More features should also be included such as road conditions and route features (bike lanes, police ticketing rates), and statistical analyses can be run to quantify the impact of each variable.