Team Activity 05.1

Names: Blair Lee, Aryan Sunkersett, Harry Tong, Cole Hannum, Adam Goldstein

In [ ]:
```
from IPython.display import Image
Image("RenderedImage Medium.jpeg")
```

Out[ ]:

In [ ]:
```
import numpy as np
import pandas as pd
from datetime import datetime

traffic = pd.read_csv("tc.csv")
traffic.head()
```

Out[ ]:

	DR Number	Date Reported	Date Occurred	Time Occurred	Area ID	Area Name	Reporting District	Crime Code	De
0	200704291	01/07/2020	01/01/2019	1033	7	Wilshire	793	997	C
1	192004028	01/01/2019	01/01/2019	820	20	Olympic	2088	997	C
2	191418727	08/25/2019	01/01/2019	2250	14	Pacific	1494	997	C
3	190104049	01/01/2019	01/01/2019	1630	1	Central	111	997	C
4	190104026	01/01/2019	01/01/2019	330	1	Central	129	997	C

5 rows × 24 columns

In [ ]:
```
traffic["Date Occurred"] = pd.to_datetime(traffic["Date Occurred"])
traffic["Date Reported"] = pd.to_datetime(traffic["Date Reported"])
traffic["Time Occurred"] = pd.to_datetime(traffic["Time Occurred"], format =
traffic.head()
```

Out[ ]:

	DR Number	Date Reported	Date Occurred	Time Occurred	Area ID	Area Name	Reporting District	Crime Code	Desc
0	200704291	2020-01-07	2019-01-01	10:33:00	7	Wilshire	793	997	TI COL
1	192004028	2019-01-01	2019-01-01	08:20:00	20	Olympic	2088	997	TI COL
2	191418727	2019-08-25	2019-01-01	22:50:00	14	Pacific	1494	997	TI COL
3	190104049	2019-01-01	2019-01-01	16:30:00	1	Central	111	997	TI COL
4	190104026	2019-01-01	2019-01-01	03:30:00	1	Central	129	997	TI COL

5 rows × 24 columns

In [ ]:
```
DR_number = traffic["DR Number"]
dr_number = traffic.iloc[:, 0]

print(DR_number)
print(dr_number)
```

```
# It is likely much more helpful to use column names
# since large dataframes will likely have many indices.
```

```
0         200704291
1         192004028
2         191418727
3         190104049
4         190104026
            ...
134168    230204688
134169    230904001
134170    222019939
134171    230106871
134172    231808015
Name: DR Number, Length: 134173, dtype: int64
0         200704291
1         192004028
2         191418727
3         190104049
4         190104026
            ...
134168    230204688
134169    230904001
134170    222019939
134171    230106871
134172    231808015
Name: DR Number, Length: 134173, dtype: int64
```

In [ ]:
```
print(traffic["Victim Age"].describe())
print(traffic["Time Occurred"].describe())
```

```
count    123750.000000
mean         40.279394
std          17.365037
min          10.000000
25%          27.000000
50%          37.000000
75%          51.000000
max          99.000000
Name: Victim Age, dtype: float64
count       133243
unique        1292
top       20:00:00
freq          2071
Name: Time Occurred, dtype: object
```

In [ ]:
```
n_by_area = traffic.groupby("Area Name")["Area ID"].count()
n_by_area.head()
```

Out[ ]:
```
Area Name
77th Street    10866
Central         6049
Devonshire      6304
Foothill        5046
Harbor          5219
Name: Area ID, dtype: int64
```

In [ ]:
```
fatality = np.where(traffic["MO Codes"].str.contains("3027").fillna(False))
print(fatality)
```

```
(array([   282,    702,    749, ..., 134018, 134048, 134058]),)
```

In [ ]:
```
traffic.iloc[fatality]["Area Name"].describe()

# It appears that the most dangerous intersection would be 77th Street as
# they have the most accidents and the most number of fatalities.
```

Out[ ]:
```
count            1097
unique             21
top       77th Street
freq              109
Name: Area Name, dtype: object
```