**Project\_2**

**Team**: Samuel Omosuyi, Arun Surendranath

**Github Repo:** https://github.com/UC-Berkeley-I-School/Project2\_Sam\_Arun.git

**Dataset**: https://data.boston.gov/dataset/employee-earnings-report, https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system

Boston crime data, Boston city spending data (employee earnings report)

Checkbook Explorer, Employee Earnings report, Crime incident reports ( data.boston.gov)

**Questions:**

1. How does the crime in city of Boston look like over the last few years?
2. What is the city spending money on?
3. How does the city employee pay over the years look like?
   1. How many instances of overtime over last few years?
   2. Which department is taking the most overtime?
   3. If the overtime is more with the police department, does that reflect with the crime data (is there a reduction in the overall crime)
4. Any other significant finding

**Data Cleaning:**

1. Cleaning of the data and combining data from various years to make it possible into a master data sheet.
2. Identify Nan values and finding methods to replace them with appropriate values based on column, row data.

**Report Format:**

1. Key findings from employee earning report, and highlighting which department are taking the highest amount of overtime and spending
2. If police department has high spending, does it tie with any change in crime data, linking crime data to city police spending
3. Trends over last 5 years, report on the trends on overtime trend and crime
4. Any other significant findings to be reported too.

**Possible Charts to use:**

1. Bar, line, pie chart to showcase the spending by the city
2. Scatter plot possible to look at any correlation between city spending and crime data.
3. Other visualization of data based on data analysis.