Data analysis and mining.

After performing the data cleaning process we tried to analyze the data.

With the use of R/Rattle we have tried to analyze the distribution and modelling of the available data which helped us extract accurate information about the different crime occurred.

Decision tree to classify the violence or crime rate with respect to each feature like poverty, unemployment and population among others.

R Graphics: Device 2 (ACTIVE)

Decision Tree CommViolPred (1).csv $ rapesPerbinary

```
                            1
                         .50  .50
                          100%
              ┌────────────┴────────────┐
            yes      pctPoverty < 6.7      no
                                          3
                                       .31  .69
                                         62%
                              ┌───────────┴───────────┐
                        pctFemDivorc < 13
                              6
                           .51  .49
                             23%
                   ┌──────────┴──────────┐
             houseVacant < 780
                                        13
                                     .34  .66
                                       7%
                              ┌─────────┴─────────┐
                        ownHousUperQ >= 201e+3
```

2	12	26	27	7
0	0	0	1	1
.81 .19	.60 .40	.92 .08	.27 .73	.19 .81
38%	16%	1%	7%	39%

Rattle 2015-Mar-13 16:19:27 Apurwa

R Graphics: Device 2 (ACTIVE)

Decision Tree CommViolPred (1).csv $ murder.binary

```
                            0
                         .51  .49
                          100%
              ┌────────────┴────────────┐
            yes   kidsBornNevrMarr < 615   no
              2
           .70  .30
             64%
     persPoverty < 948
                              6
                           .61  .39
                             40%
                    pctPopDenseHous < 2.1
                                        11
                                     .50  .50
                                       20%
                                 pctHousOccup >= 89
                              22
                           .55  .45
                             17%
                    houseVacant < 270
                                        46
                                     .51  .49
                                       14%
                                 pctCollGrad < 36
```

4	19	44	90	91	23	3
0	0	0	0	1	1	1
.84 .16	.72 .28	.79 .21	.56 .44	.23 .77	.25 .75	.17 .83
25%	19%	3%	12%	2%	4%	36%

Rattle 2015-Mar-13 16:08:43 Apurwa

Decision Tree CommViolPred (1).csv $ larc.binary

Decision Tree CommViolPred (1).csv $ burg.binary

Rattle 2015-Mar-13 16:36:08 Apurwa

Decision Tree CommViolPred (1).csv $ assault.binary

```
                                    0
                                 .50 .50
                                  100%
```

yes — pctKidsBornNevrMarr < 3 — no

```
        2                                              3
        0                                              1
     .69 .31                                        .13 .87
      66%                                            34%
```

pctAllDivorc < 9.1   pctPoverty < 7

```
                    5
                    0
                 .55 .45
                  38%
```

pctPopDenseHous < 2

```
                        11
                         1
                      .41 .59
                       21%
```

houseVacant < 551

```
                            23
                             1
                          .29 .71
                           13%
```

pctUnemploy < 4.1

```
  4          10          22          46          47          6           7
  0           0           0           0           1           0           1
.89 .11    .73 .27    .59 .41    .64 .36    .20 .80    .63 .37    .10 .90
 28%        17%         8%         3%        10%         2%        32%
```

Rattle 2015-Mar-13 16:34:34 Apurwa

Decision Tree CommViolPred (1).csv $ assault.binary

```
                                    0
                                 .50 .50
                                  100%
```

yes — pctKidsBornNevrMarr < 3 — no

```
        2                                              3
        0                                              1
     .69 .31                                        .13 .87
      66%                                            34%
```

pctAllDivorc < 9.1   pctPoverty < 7

```
                    5
                    0
                 .55 .45
                  38%
```

pctPopDenseHous < 2

```
        10                                      11
         0                                       1
      .73 .27                                 .41 .59
       17%                                     21%
```

pct2Par >= 71   houseVacant < 551

```
            21                                      23
             1                                       1
          .43 .57                                 .29 .71
           4%                                      13%
```

landArea < 9.6   pctUnemploy < 4.1

```
  4       20        42        43        22        46        47        6        7
  0        0         0         1         0         0         1        0        1
.89 .11 .81 .19   .69 .31   .15 .85   .59 .41   .64 .36   .20 .80  .63 .37  .10 .90
 28%     14%        2%        2%        8%        3%       10%       2%      32%
```

Rattle 2015-Mar-13 16:25:02 Apurwa

R Graphics: Device 2 (ACTIVE)

Decision Tree CommViolPred (1).csv $ arson.binary

```
                        1
                      .50 .50
                      100%
        yes    pctFemDivorc < 12    no
                                              3
                                              1
                                            .34 .66
                                            57%
                                    persUrban < 86e+3
                              6
                              1
                            40  60
                            48%
                    kidsBornNevrMarr < 362
                                          13
                                          1
                                        .32 .68
                                        31%
                                pctFgnImmig.5 >= 44

    2          12              26          27          7
    0          0               0           1           1
  .71 .29    .53 .47         .83 .17      .30 .70     .08 .92
   43%        17%             1%          29%          9%
```

Rattle 2015-Mar-13 16:43:42 Apurwa

R Graphics: Device 2 (ACTIVE)

Decision Tree CommViolPredupdated.csv $ violent.binary

```
                                0
                              .51 .49
                              100%
                        kidsBornNevrMarr < 456
           yes                                      no

    0                                                      1
  .74 .26                                                .20 .80
   57%                                                    43%
  pctAllDivorc < 11                                      pctWdiv >= 46

              0                                3
            .52 .46                          0
            19%                            .52 .48
      pctBornStateResid >= 30              7%
                                       pct12.17w2Par >= 76
         0
       .55 .45
       18%
    pctPopDenseHous < 2.9

       0
     .56 .34
     12%
  pctLowEdu < 12

 4        48       47       27       11       12       13       7
 0        0        1        1        1        0        1        1
.86 .14  .73 .27  .39 .61  .36 .64  .00 1.00 .75 .25  .22 .78  .14 .86
 37%      9%       3%       6%       1%       4%       3%       37%
```

Rattle 2015-Mar-13 17:08:41 Apurwa

Distribution of Community.Area (sample)

Rattle 2015-Mar-13 13:46:24 Apurwa