

```
In [1]: import csv
```

This notebook explores the Crime head-wise persons arrested under crime against children during 2001-2012.

- dataset : [link \(https://data.world/bhavnachawla/crime-rate-against-children-india-2001-2012\)](https://data.world/bhavnachawla/crime-rate-against-children-india-2001-2012) ##### Following lines just load the dataset and explores it. ##### There are two ways to read the dataset - as dataset is a csv file, either csv library can be used or the famous pandas. ##### This notebook uses csv library and read dataset in a list data structure where each element is a dictionary.

```
In [2]: path = 'Crime head-wise persons arrested under crime against children during 2001-2012.csv'
f = open(path)
```

```
In [3]: reader = csv.reader(f)
```

```
In [4]: header = next(reader)
```

```
In [5]: header
```

```
Out[5]: ['STATE/UT',
        'CRIME HEAD',
        '2001',
        '2002',
        '2003',
        '2004',
        '2005',
        '2006',
        '2007',
        '2008',
        '2009',
        '2010',
        '2011',
        '2012']
```

```
In [6]: dataset = []
        for line in reader:
            d = dict(zip(header, line))
            d['2001'] = int(d['2001'])
            d['2002'] = int(d['2002'])
            d['2003'] = int(d['2003'])
            d['2004'] = int(d['2004'])
            d['2005'] = int(d['2005'])
            d['2006'] = int(d['2006'])
            d['2007'] = int(d['2007'])
            d['2008'] = int(d['2008'])
            d['2009'] = int(d['2009'])
            d['2010'] = int(d['2010'])
            d['2011'] = int(d['2011'])
            d['2012'] = int(d['2012'])
            dataset.append(d)
```

```
In [7]: dataset[0]
```

```
Out[7]: {'STATE/UT': 'ANDHRA PRADESH',
         'CRIME HEAD': 'INFANTICIDE',
         '2001': 1,
         '2002': 1,
         '2003': 3,
         '2004': 0,
         '2005': 0,
         '2006': 0,
         '2007': 1,
         '2008': 0,
         '2009': 5,
         '2010': 6,
         '2011': 1,
         '2012': 6}
```

```
In [8]: print("DATSET SIZE(ROWS IN CSV FILES): ", len(dataset))
```

DATSET SIZE(ROWS IN CSV FILES): 494

Dataset Analysis

```
In [9]: CRIME_HEAD = []
        STATE = []
        for d in dataset:
            if not d['CRIME HEAD'] in CRIME_HEAD:
                CRIME_HEAD.append(d['CRIME HEAD'])
            if not d['STATE/UT'] in STATE:
                STATE.append(d['STATE/UT'])
```

```
In [10]: CRIME_HEAD
```

```
Out[10]: ['INFANTICIDE',  
          'MURDER OF CHILDREN',  
          'RAPE OF CHILDREN',  
          'KIDNAPPING and ABDUCTION OF CHILDREN',  
          'FOETICIDE',  
          'ABETMENT OF SUICIDE',  
          'EXPOSURE AND ABANDONMENT',  
          'PROCURATION OF MINOR GIRLS',  
          'BUYING OF GIRLS FOR PROSTITUTION',  
          'SELLING OF GIRLS FOR PROSTITUTION',  
          'PROHIBITION OF CHILD MARRIAGE ACT',  
          'OTHER CRIMES AGAINST CHILDREN',  
          ' TOTAL CRIMES AGAINST CHILDREN']
```

```
In [11]: print("Categories of crime: ",len(CRIME_HEAD))
```

```
Categories of crime:  13
```

```
In [12]: print("List of STATE/UT : ", STATE)
```

```
List of STATE/UT :  ['ANDHRA PRADESH', 'ARUNACHAL PRADESH', 'ASSAM', 'BIHAR',  
                    'CHHATTISGARH', 'GOA', 'GUJARAT', 'HARYANA', 'HIMACHAL PRADESH', 'JAMMU & KAS  
HMIR', 'JHARKHAND', 'KARNATAKA', 'KERALA', 'MADHYA PRADESH', 'MAHARASHTRA',  
                    'MANIPUR', 'MEGHALAYA', 'MIZORAM', 'NAGALAND', 'ODISHA', 'PUNJAB', 'RAJASTHA  
N', 'SIKKIM', 'TAMIL NADU', 'TRIPURA', 'UTTAR PRADESH', 'UTTARAKHAND', 'WEST  
BENGAL', 'TOTAL (STATES)', 'A & N ISLANDS', 'CHANDIGARH', 'D & N HAVELI', 'DA  
MAN & DIU', 'DELHI', 'LAKSHADWEEP', 'PUDUCHERRY', 'TOTAL (UTs)', 'TOTAL (ALL-  
INDIA)']
```

Q 1 How many crimes are reported each year in India from 2001 to 2012?

```
In [13]: dataset[0]
```

```
Out[13]: {'STATE/UT': 'ANDHRA PRADESH',  
          'CRIME HEAD': 'INFANTICIDE',  
          '2001': 1,  
          '2002': 1,  
          '2003': 3,  
          '2004': 0,  
          '2005': 0,  
          '2006': 0,  
          '2007': 1,  
          '2008': 0,  
          '2009': 5,  
          '2010': 6,  
          '2011': 1,  
          '2012': 6}
```

```
In [14]: Total = {}  
for d in dataset:  
    if d['STATE/UT'] == 'TOTAL (ALL-INDIA)' and d['CRIME HEAD'] == 'TOTAL CRIMES AGAINST CHILDREN':  
        Total = d  
Total
```

```
Out[14]: {'STATE/UT': 'TOTAL (ALL-INDIA)',  
          'CRIME HEAD': 'TOTAL CRIMES AGAINST CHILDREN',  
          '2001': 13401,  
          '2002': 12507,  
          '2003': 13524,  
          '2004': 16663,  
          '2005': 17353,  
          '2006': 20870,  
          '2007': 22432,  
          '2008': 25766,  
          '2009': 26012,  
          '2010': 27403,  
          '2011': 35427,  
          '2012': 42117}
```

```
In [15]: import matplotlib.pyplot as plt
```

```
In [16]: X = list(Total.keys())
```

```
In [17]: Y = list(Total.values())
```

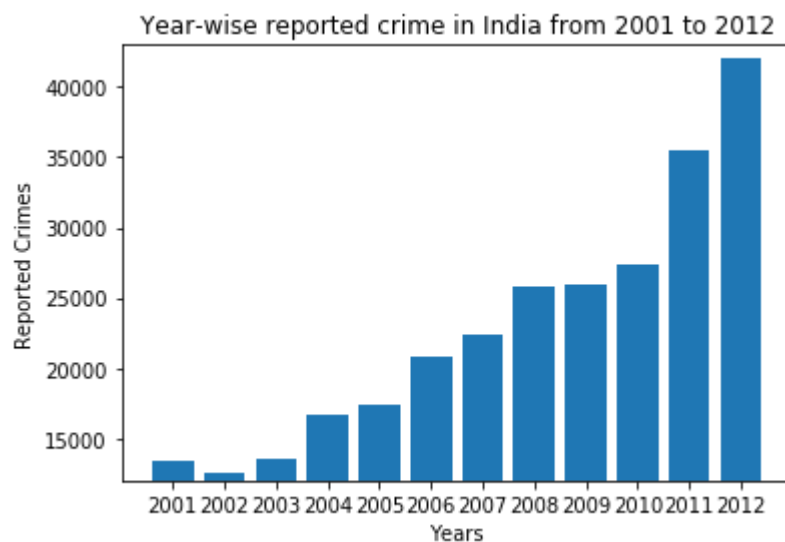
```
In [18]: X.pop(0)  
X.pop(0)  
X
```

```
Out[18]: ['2001',  
          '2002',  
          '2003',  
          '2004',  
          '2005',  
          '2006',  
          '2007',  
          '2008',  
          '2009',  
          '2010',  
          '2011',  
          '2012']
```

```
In [19]: Y.pop(0)
Y.pop(0)
Y, len(Y)
```

```
Out[19]: ([13401,
12507,
13524,
16663,
17353,
20870,
22432,
25766,
26012,
27403,
35427,
42117],
12)
```

```
In [20]: plt.bar(list(range(0,12)),Y)
plt.xticks(list(range(0,12)),X)
plt.xlabel("Years")
plt.ylim(12000,43000)
plt.ylabel("Reported Crimes")
plt.title("Year-wise reported crime in India from 2001 to 2012")
plt.show()
```



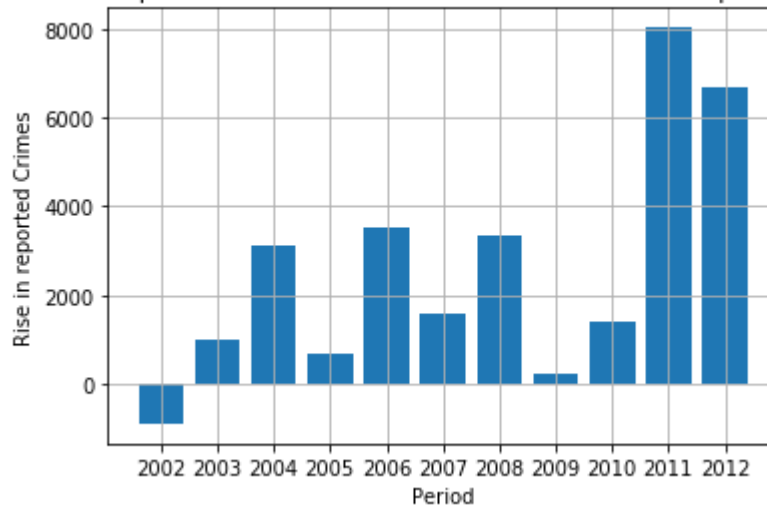
Q-2 Rise in reported child crime from 2001 to 2012

```
In [21]: Y_diff = []
for i in range(len(Y)):
    if not i is 0:
        Y_diff.append(Y[i]-Y[i-1])
Y_diff
```

```
Out[21]: [-894, 1017, 3139, 690, 3517, 1562, 3334, 246, 1391, 8024, 6690]
```

```
In [22]: plt.bar(list(range(0,11)),Y_diff)
plt.xticks(list(range(0,11)),list(range(2002,2013,1)))
plt.xlabel("Period")
plt.ylabel("Rise in reported Crimes")
plt.grid()
plt.title("Year-wise rise in reported crime in India from 2002 to 2012 as compare to previous year")
plt.show()
```

Year-wise rise in reported crime in India from 2002 to 2012 as compare to previous year



Q 3 Which Category of Crime has the highest rate of reporting?

```
In [23]: CRIME_HEAD
```

```
Out[23]: ['INFANTICIDE',
'MURDER OF CHILDREN',
'RAPE OF CHILDREN',
'KIDNAPPING and ABDUCTION OF CHILDREN',
'FOETICIDE',
'ABETMENT OF SUICIDE',
'EXPOSURE AND ABANDONMENT',
'PROCURATION OF MINOR GIRLS',
'BUYING OF GIRLS FOR PROSTITUTION',
'SELLING OF GIRLS FOR PROSTITUTION',
'PROHIBITION OF CHILD MARRIAGE ACT',
'OTHER CRIMES AGAINST CHILDREN',
' TOTAL CRIMES AGAINST CHILDREN']
```

```
In [24]: len(CRIME_HEAD)*len(STATE) == len(dataset)
```

```
Out[24]: True
```

```
In [25]: dataset_India = []
for d in dataset:
    if d['STATE/UT'] == 'TOTAL (ALL-INDIA)' and not d['CRIME HEAD'] == ' TOTAL
CRIMES AGAINST CHILDREN':
        dataset_India.append(d)
dataset_India[0]
```

```
Out[25]: {'STATE/UT': 'TOTAL (ALL-INDIA)',
'CRIME HEAD': 'INFANTICIDE',
'2001': 117,
'2002': 124,
'2003': 84,
'2004': 63,
'2005': 80,
'2006': 87,
'2007': 147,
'2008': 118,
'2009': 40,
'2010': 82,
'2011': 67,
'2012': 68}
```

```
In [26]: len(dataset_India)
```

```
Out[26]: 12
```

```
In [27]: from operator import itemgetter
```

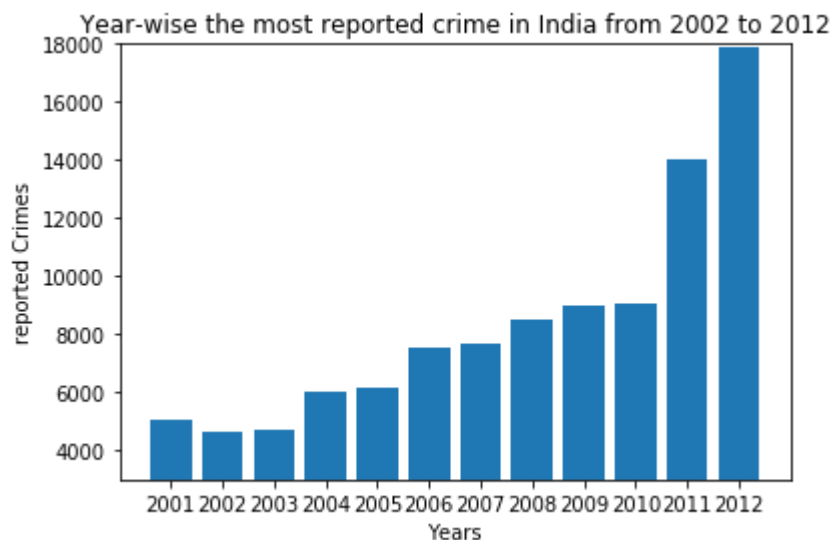
```
In [28]: def sort_list_of_dictionary_key(crime_India,keys):
    sortd_crime_India = sorted(crime_India,key=itemgetter(keys),reverse=True)
    return sortd_crime_India[0]['CRIME HEAD'],sortd_crime_India[0][keys]
```

```
In [29]: CRIME_CATEGORIES = []
YEARS = ['2001','2002','2003','2004','2005','2006','2007','2008','2009','2010',
'2011','2012']
NO_OF_CRIMES = []
for year in YEARS:
    x,y = sort_list_of_dictionary_key(dataset_India,year)
    CRIME_CATEGORIES.append(x)
    NO_OF_CRIMES.append(y)
```

In [30]: CRIME_CATEGORIES

Out[30]: ['OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'OTHER CRIMES AGAINST CHILDREN',
'KIDNAPPING and ABDUCTION OF CHILDREN',
'KIDNAPPING and ABDUCTION OF CHILDREN',
'KIDNAPPING and ABDUCTION OF CHILDREN']

In [31]: plt.bar(list(range(0,12)),NO_OF_CRIMES)
plt.xticks(list(range(0,12)),YEARS)
plt.xlabel("Years")
plt.ylim(3000,18000)
plt.ylabel("reported Crimes")
plt.title("Year-wise the most reported crime in India from 2002 to 2012")
plt.show()



Conclusion :Yearwise The Most repoted crime category is 'KIDNAPPING and ABDUCTION OF CHILDREN' from the year 2010 to 2012 and earlier the category is 'OTHER CRIMES AGAINST CHILDREN'

Q-4 Which States have the highest reported crime category?

In [32]:

```
def sort_list_of_dictionary_key2(crime_India,keys):
    soretd_crime_India = sorted(crime_India,key=itemgetter(keys),reverse=True)
    for d in soretd_crime_India:
        if not (d['STATE/UT'] == 'TOTAL (ALL-INDIA)' or d['STATE/UT'] == 'TOTAL (UTs)' or d['STATE/UT'] == 'TOTAL (STATES)'):
            return d['STATE/UT'],d[keys]
```



```
In [33]: STATE_OR_UT = []
YEARS = ['2001', '2002', '2003', '2004', '2005', '2006', '2007', '2008', '2009', '2010',
          '2011', '2012']
NO_OF_CRIMES = []
for year in YEARS:
    x,y = sort_list_of_dictionary_key2(dataset,year)
    STATE_OR_UT.append(x)
    NO_OF_CRIMES.append(y)
```

```
In [34]: len(STATE_OR_UT),len(NO_OF_CRIMES)
```

```
Out[34]: (12, 12)
```

```
In [35]: from collections import defaultdict
```

```
In [36]: RatingOfSTATE = defaultdict(list)
```

```
In [37]: j=0
k=0
for i in STATE_OR_UT:
    RatingOfSTATE[i].append((YEARS[j],NO_OF_CRIMES[k]))
    j = j + 1
    k = k + 1
RatingOfSTATE
```

```
Out[37]: defaultdict(list,
                        {'UTTAR PRADESH': [('2001', 6319),
                                             ('2008', 5760),
                                             ('2011', 8560),
                                             ('2012', 11470)],
                        'MADHYA PRADESH': [('2002', 3178),
                                             ('2003', 3668),
                                             ('2004', 4530),
                                             ('2005', 4949),
                                             ('2006', 5062),
                                             ('2007', 5305),
                                             ('2009', 5829),
                                             ('2010', 5846)]})
```

Conclusion : There are only two states which reports to the most of the child crimes : UTTAR PRADESH AND MADHYA PRADESH from the span of 2001 to 2012.

Q-5 Which States have the lowest reported child crime category?

```
In [38]: def sort_list_of_dictionary_key2(crime_India,keys):
soretd_crime_India = sorted(crime_India,key=itemgetter(keys))
for d in soretd_crime_India:
    if not (d['STATE/UT'] == 'TOTAL (ALL-INDIA)' or d['STATE/UT'] == 'TOTAL (UTs)' or d['STATE/UT'] == 'TOTAL (STATES)'):
        return d['STATE/UT'],d[keys]
```

```
In [39]: STATE_OR_UT_GOOD = []
YEARS = ['2001', '2002', '2003', '2004', '2005', '2006', '2007', '2008', '2009', '2010',
          '2011', '2012']
NO_OF_CRIMES_LESS = []
for year in YEARS:
    x,y = sort_list_of_dictionary_key2(dataset,year)
    STATE_OR_UT_GOOD.append(x)
    NO_OF_CRIMES_LESS.append(y)
```

```
In [40]: len(STATE_OR_UT_GOOD),len(NO_OF_CRIMES_LESS)
```

```
Out[40]: (12, 12)
```

```
In [41]: from collections import defaultdict
```

```
In [42]: RatingOfSTATE2 = defaultdict(list)
```

```
In [43]: j=0
k=0
for i in STATE_OR_UT_GOOD:
    RatingOfSTATE2[i].append((YEARS[j],NO_OF_CRIMES_LESS[k]))
    j = j + 1
    k = k + 1
RatingOfSTATE2
```

```
Out[43]: defaultdict(list,
                        {'ARUNACHAL PRADESH': [('2001', 0),
                                                ('2002', 0),
                                                ('2003', 0),
                                                ('2007', 0),
                                                ('2009', 0),
                                                ('2010', 0),
                                                ('2011', 0),
                                                ('2012', 0)],
                        'ANDHRA PRADESH': [('2004', 0),
                                              ('2005', 0),
                                              ('2006', 0),
                                              ('2008', 0)]})
```

Conclusion: ANDHRA PRADESH and ARUNACHAL PRADESH has the lowest reported crime in atleast one category from the span of 2001 to 2012.