

# EP219: Data Analysis and Interpretation

## Assignment Report 1



By Team: *Significantly Different*

August/06/2018 - August/13/2018

# Contents

<b>Problem Statement</b>	<b>3</b>
<b>Code</b>	<b>4</b>
<b>Histograms</b>	<b>6</b>
<b>Conclusion</b>	<b>9</b>
<b>Team Contribution</b>	<b>10</b>

# Problem Statement

Our aim is to extract data about Pre-Primary school and number of teachers in a Numpy Array and then making a Histogram of number of states versus certain number of teachers per pre-primary school.

X-axis of the histogram should be the number of teachers per school and Y-axis is the number of states.

There are 3 seperate Histogram:

- For Urban
- For Rural
- For combined Data (i.e Urban and Rural)

# Code

```
#Importing libraries numpy for data handling and matplotlib for plotting histogram
import numpy as np
import matplotlib.pyplot as plt

#Reading data from appropriate columns of csv file , Typecasting data from
#String to Float and Storing it in list
array = []
for i,line in enumerate( open('pre-primary-school-data.csv')):
    if i > 0:
        full_row = line.split(',')
        req_row = []
        for j in range(3):
            req_row.append(float( full_row[12 + j] ))
        array.append(req_row)

#List converted to numpy main-matrix
main_matrix = np.array(array)

#Getting the size of the numpy main-matrix
(a, b) = main_matrix.shape

#Declaring and initializing array to store total number of teachers
total_teachers_array = np.zeros(a)
for i in range(a):
    total_teachers_array[i] = (main_matrix[i][1]+main_matrix[i][2])

#Converting the array to a matrix for concatenation to main-matrix
total_teachers_matrix = total_teachers_array.reshape(a, 1)
main_matrix = np.concatenate( (main_matrix, total_teachers_matrix), axis=1)

#Declaring and initializing an array to store number of teachers per school
teachers_per_school_array = np.zeros(a)
for i in range(a):
    if main_matrix[i][0] != 0:
        teachers_per_school_array[i] = main_matrix[i][3] / main_matrix[i][0]

#Converting the array to a matrix for concatenation to main-matrix
teachers_per_school_matrix = teachers_per_school_array.reshape(a, 1)
main_matrix = np.concatenate( (main_matrix, teachers_per_school_matrix), axis=1)

#Declaring and initializing a dictionary for teachers per school
teachers_per_school={}
teachers_per_school[ "rural" ] = np.zeros( int(a / 3) - 1)
teachers_per_school[ "urban" ] = np.zeros( int(a / 3) - 1 )
teachers_per_school[ "total" ] = np.zeros( int(a / 3) - 1)
for i in range(a - 3):
    current=teachers_per_school_array[i]
    if i % 3 == 0:
        teachers_per_school[ 'rural' ][ int(i / 3) ] = current
    elif i % 3 == 1:
        teachers_per_school[ 'urban' ][ int(i / 3) ] = current
    else:
        teachers_per_school[ 'total' ][ int(i / 3) ] = current
```

```

#Plotting and showing required histograms with optimum bins
plt.hist( teachers_per_school[ 'rural' ], bins = 7, color = 'b')
plt.xlabel('Teachers per school rural')
plt.ylabel('No of states')
plt.title('Pre primary teachers per school distribution - rural')
plt.show()

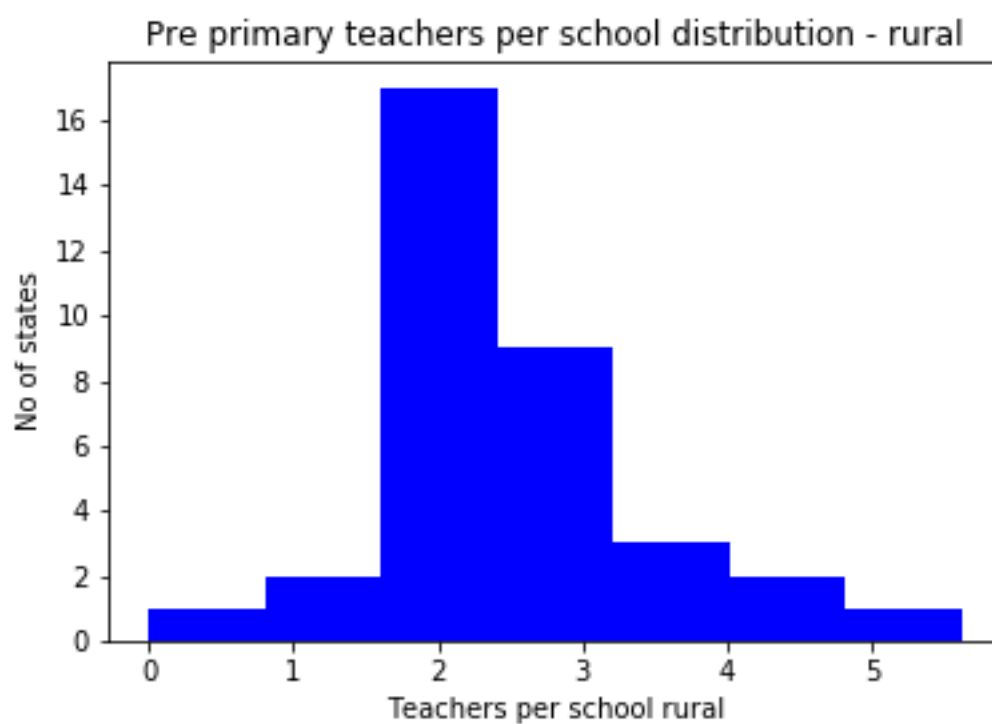
plt.hist( teachers_per_school[ 'urban' ], bins = 5, color = 'b')
plt.xlabel('Teachers per school urban')
plt.ylabel('No of states')
plt.title('Pre primary teachers per school distribution - urban')
plt.show()

plt.hist( teachers_per_school[ 'total' ], bins = 7, color = 'b')
plt.xlabel('Teachers per school total')
plt.ylabel('No of states')
plt.title('Pre primary teachers per school distribution - total')
plt.show()

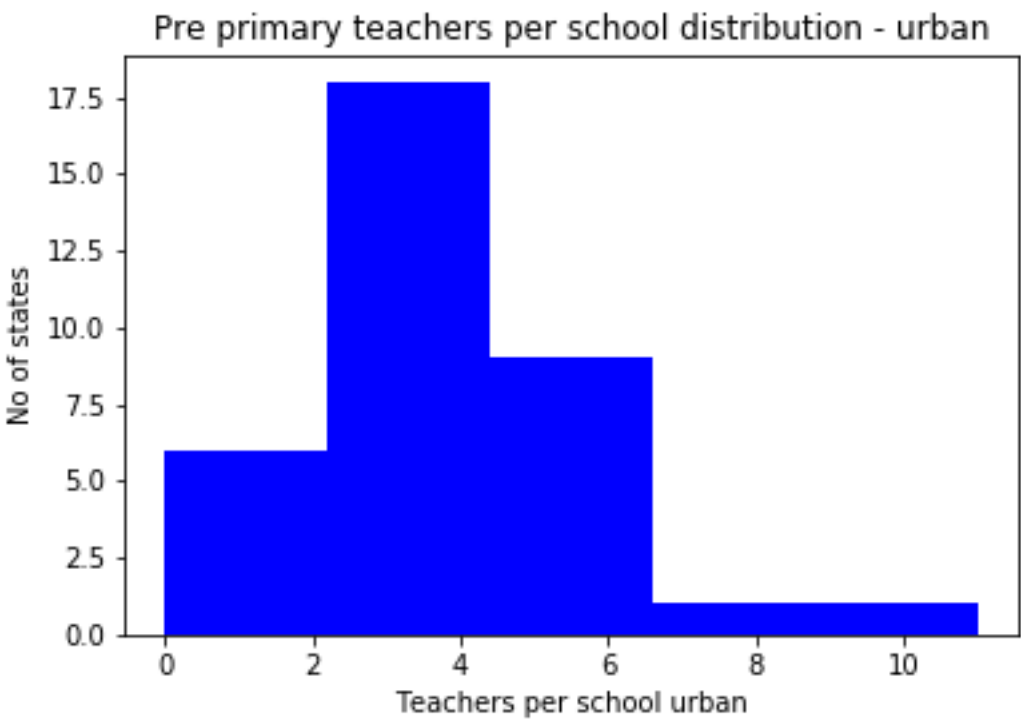
```

# Histograms

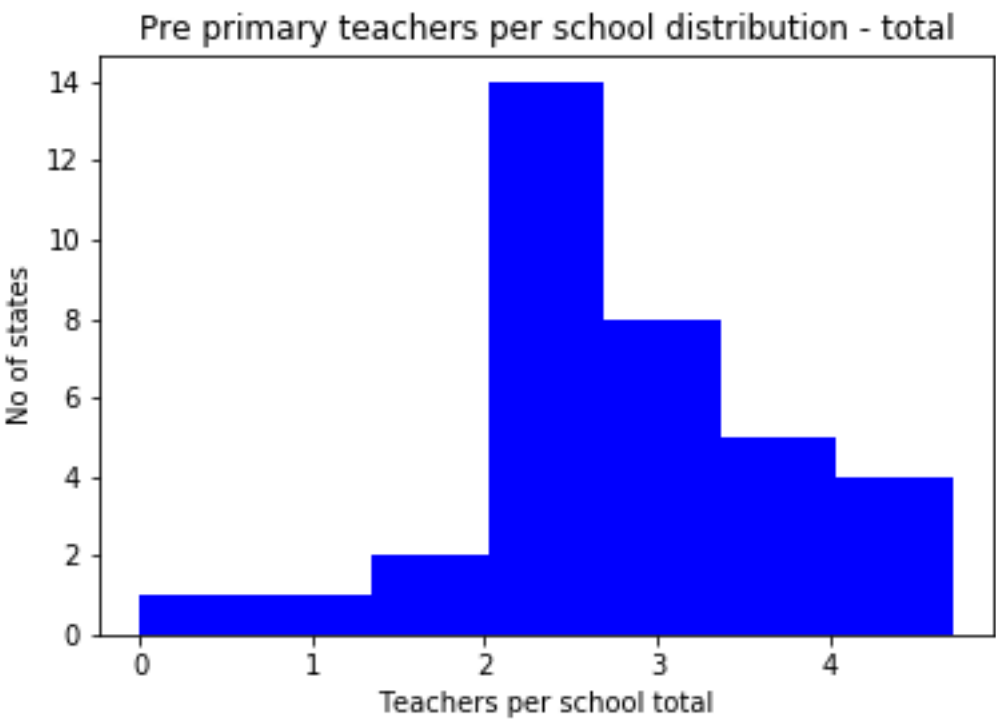
A) Histogram of number of states vs number of teachers per pre-primary school in **Rural Areas**.



B) Histogram of number of states vs number of teachers per pre-primary school in **Urban Areas**.



C) Histogram of number of states vs number of teachers per pre-primary school combined in **Rural and Urban Areas**.





# Conclusion

By plotting this data we conclude that:

- a) In Rural Areas of most of the states there are around **2 - 2.5** teachers per pre-primary school.
- b) In Urban Areas of most of the states there are around **3.5 - 4** teachers per pre-primary school.
- c) And in the combined data, most of the states have around **2.5 - 3** teachers per pre-primary school.

## Team Contribution

- a) **Vashishtha Kochar** - Web Developer ..... 25%
- b) **Nihal Barde** - Team Leader ..... 25%
- c) **Adeem Jassani** - Programmer ..... 25%
- d) **Ram** - Report writer ..... 25%