mean, median, mode, variance and standard deviation

In [41]:
```
import collections
import math

class Statistics:

    def __init__(self, data_set):
        self.data_set = data_set
        self.data_set.sort() #sorted list
        self.num_value = len(self.data_set)    # number of observations
        self.sum_value = sum(self.data_set)    # average of data

    def mean(self):
        self.avg_value = (self.sum_value/self.num_value)
        return self.avg_value

#     def mean(self):          # sum of obser/num of obser
#         value = 0
#         num_value = len(self.data_set)
#         for i in range(num_value):
#             value = value + self.data_set[i]
#             avg_value = value/num_value
#             return avg_value

#     def median_odd(self):
#         '''if len(data) is odd, the median is the middle element'''
#         mid_value = (self.num_value + 1)/2
#         return mid_value

    def median_even(self):
        '''if len(data) is even, the median is the middle element'''
        mid_value = (self.num_value/2) + ((self.num_value/2) + 1)
        return mid_value

    def mode(self):
        data = collections.Counter(self.data_set)  #calculate the frequency of each item
        most_value = data.most_common(1)
        return most_value

    def variance(self):
        deviations = [(x - self.avg_value) ** 2 for x in self.data_set]   #statistics.pvariance(data_set)
        variance_val = sum(deviations)/self.num_value   #(n-1)
        return variance_val

    def std_deviation(self):
        variance = self.variance()
        std_dev = math.sqrt(variance)
        return std_dev

data_set = [4, 8, 6, 5, 3, 2, 8, 9, 2, 5]
stats = Statistics(data_set)

mean = stats.mean()
print('mean value: ', mean)

median_even = stats.median_even()
print('median value: ', median_even)

mode = stats.mode()
print('mode value: ', mode)

variance = stats.variance()
print('variance value: ', variance)

std_deviation = stats.std_deviation()
print('standard deviation value: ', std_deviation)
```

```
mean value:  5.2
median value:  11.0
mode value:  [(2, 2)]
variance value:  5.76
standard deviation value:  2.4
```

In [ ]:
```
import statistics

data_set = [4, 8, 6, 5, 3, 2, 8, 9, 2, 5]

var =  statistics.pvariance(data_set)  #returns the variance of the population   #N
var1 = statistics.variance(data_set)   #returns the variance of the sample   #n-1

std_dev = statistics.pstdev(data_set)  #returns the dev of the population
std_dev1 = statistics.stdev(data_set)  #returns the dev of the sample
```

In [ ]:
```

```