datascience_assignment_16.2

July 13, 2018

0.1 Find the variance for the following set of data representing trees in California (heights in feet):

3, 21, 98, 203, 17, 9

0.2 Solution:

- Write a function variance which takes list/tuple which returns variance with Bessel's correction. Note I have assumed the data reprenting samples of population of trees in California, that is why Bessel's correction is needed
- The function variance first calculates mean
- From each element subtract mean and square it and sum the whole resultant elements, which is divided by (No of elements - 1) to get the variance with Bessel's correction

```
In [5]: import functools
        import math
        def variance(element_list):
            # Input element_list must be either tuple or list else ValueError is raised
            if ( not isinstance(element_list, tuple)) and ( not isinstance(element_list, list))
                raise ValueError("element_list must be either list or tuple")

             # Input element_list should not be None or empty else ValueError is raised
            if element_list is None or len(element_list) == 0:
                raise ValueError("element_list can not be empty")

            # If number of elements is 1 return 0
            if len(element_list) == 1:
                return 0

            # Caculate mean by using reduce function with lambda expresssion as sum of two num
            # divide by number of elements
            mean = functools.reduce(lambda x,y: x+y, element_list)/len(element_list)

            # From each element subtract mean and square it and sum the whole resultant elemen
            #  to calculate variance with Bessel's correction
            result = sum([(x - mean) ** 2 for x in element_list]) / (len(element_list) - 1)
            return result
```

```
In [6]: element_list  = [3, 21, 98, 203, 17, 9]
        res = variance(element_list)
        print("standard deviation = " + str(res) + " square feet")
```

```
standard deviation = 6219.9 square feet
```