Stat 2:

stat session 2:

MCT ---> measure of central tendency.
   --> to explain about variable performance with a single value.

components of mct.
   mean, mode, median

Mean--> average --> most of values are close to mean
Mode --> highly repeated. --> most probably the next event would be "mode"
Median --> central value.--> 50% of data are above to median
                                     50% of data is below to median

Mean --->
-----
    is an average.

   3 types of means.

   1. Arithmetic mean
   2. Geometric mean
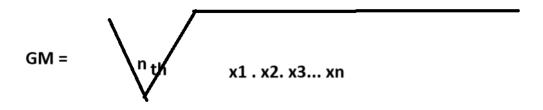   3. Harmonic mean (scientific mean)

1. AM -->   sum/n

```
x = [10,20,30,40,50,60,70]
len(x)
```

```
tot = sum(x)
cnt = len(x)
am = tot/cnt
print("Arithmetic mean ", am)
```

```
def amean(x):
    return sum(x)/len(x)
```

```
amean(x)
```

2. Geometric mean:

$$GM = \sqrt[n^{th}]{x1 . x2. x3... xn}$$

```python
def gmean(x):
    n = len(x)
    m = 1
    for v in x:
        m *= v
    gm = m**(1/n)
    return gm
```

```
gmean(x)
```

3. Harmonic mean:

$$hm = \frac{n}{1/x1 + 1/X2 + 1/X3 .... + 1/Xn}$$

```python
def hmean(x):
    n = len(x)
```

```
  ix = [1/v for v in x]
  s = sum(ix)
  return n/s


hmean(x)
```

```
am = amean(x)
gm = gmean(x)
hm = hmean(x)

mean = (am, gm, hm)
print(mean)
```

```
(40.0, 33.800151591412956, 26.99724517906336)
```

Relationship between am, gm, hm

Am > gm > hm

When to use, what type of mean ?????????

----

case1
    -- > comparision of two person weights.

                You   Me

                67.9kg 67.3kg
                ----
        who is fatty ?
                -- > both at same level

    ---> fractions are neglectable

---> am

case2

    ---> comparision of two Gold arnment weights

       g1   g2

       10.32   10.58gr

  ---> fractions are not neglectable. (important)
  ---> gm
case 3:
    air temperature

      day 1 --> 34.56 degree
      day 2 --> 34.42 degree
    --> which day more hotty (sunny)

      ==> both days are hotty

  ---> fraction is neglectable.
  ---> am

case 4:
    rain fall in cente meters

      day --> 4.1 cm
      day2 --> 4.6 cm

      which day is more rained ?
      --> day2

  ---> fraction is important.
  ---> gm

when ever fractions are neglectable --> use "Arithmetic mean"
if fractions are important ---> use "Geometric Mean".

Harmonic mean: (scientific mean)
----
two situations for hm:

  1. minute (sub) fractions are also important.

[where values can not be roundable]

2. if outliers existed in data (variable values).

situation 1>
minute (sub) fractions are also important.
        [where values can not be roundable]

ex: 1. anti missile launching time
     2. dimond carrot calculations (light reflection degrees).
     3. chemical compound weights

                        weights

person ---->   67.83   67.35   (fraction neglectable)--> am

gold   ---->   30.83   30.833   (fraction important)
                                        (sub fraction neglectable)-- gm

chemical -->   0.0003245934568   0.000324589123 --> sub fractions also imp-- hm
            0.0003246   0.0003246

situation 2:
----

2. if outliers existed in data (variable values).

case1:
        Sachin

          60
          70
          80
        ----
    tot = 210
    cnt = 3
   am = 210/3 = 70 (mean)

interpretation --> most of his scores close to 70

lets test it

Sachin
        ------
          60 --> close to mean
          70 --> equal to mean
          80 --> close to mean

```
        -------
above interpretation is correct.

case2:
        Sachin
        ------
          60
          70
          80
       290 --> irregular behaviour ---> outlier
        -------
   tot = 500
   cnt = 4
   am = 500/4 = 125

interpretation --> most of his scores are close to 125.

lets test:

        Sachin
        ------
          60  ---> not close to mean
          70 -->  not close to mean
          80 -->  not close to mean
       290 -->  not close to mean

here, your interpretation is wrong.


case3:
        Sachin
        ------
          60
          70
          80
         1 --> irregular behaviour ---> outlier
        -------
   tot = 210
   cnt = 4
    am = 210/4 --> 52


we need to neglect outliers.

--------
Below is "Google Colab" notebook of above code.
```

https://colab.research.google.com/drive/14BeCfMBx_j7JlxlvvY6HD_NJnjyu6tHn?usp=sharing