

# basic\_statistics

January 13, 2022

## 1 Participant

Title= "Mr" Name= "Khizar Jahangir" email = "khizerjahangir786@gmail.com" whatsapp = "00923212878267"

## 2 Basic Statistics

### 2.1 Data Types

- 1 (i)- Cross sectional (Data collected at single point)
- 1 (i)- Time series (Data collected at different time points)
- 2 (i)- Univariate (Data contains a single variable to measure entity)
- 2 (ii)- Multi-variate (Data contains >2 variables to measure)

### 2.2 Variable types

- Categorical
  - 1 (i)- Binomial (etc: True/False)
  - 1 (ii)- Multinomial (etc: more than two choices)
  - 1 (iii)- Nominal (Groups with no rank or order between them etc: colors)
  - 1 (iv)- Ordinal variables (Groups that are ranked in a specific order etc: ranking)
- Quantitative
  - 1 (i)- Ratio Data (Measurement in units and ratios are continuous variables)
  - 1 (ii)- Interval data (ordered and characterized data)

```
[ ]: import pandas as pd
df= pd.read_csv("bangla.csv")
df.describe()
```

```
[ ]:      Area Code  Element Code  Item Code      Year Code      Value  Note
count      19.0         19.0       19.0  1.900000e+01  19.000000   0.0
mean       16.0        6132.0    210011.0  2.009201e+07  20.568421  NaN
std         0.0          0.0         0.0  5.627877e+04   1.926728  NaN
```

min	16.0	6132.0	210011.0	2.000200e+07	15.900000	NaN
25%	16.0	6132.0	210011.0	2.004701e+07	19.700000	NaN
50%	16.0	6132.0	210011.0	2.009201e+07	20.000000	NaN
75%	16.0	6132.0	210011.0	2.013702e+07	21.850000	NaN
max	16.0	6132.0	210011.0	2.018202e+07	23.500000	NaN

```
[ ]: df.mean()
```

C:\Users\hp\AppData\Local\Temp\ipykernel\_9428\3698961737.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.mean()
```

```
[ ]: Area Code      1.600000e+01
      Element Code   6.132000e+03
      Item Code      2.100110e+05
      Year Code      2.009201e+07
      Value          2.056842e+01
      Note           NaN
      dtype: float64
```

```
[ ]: df.median()
```

C:\Users\hp\AppData\Local\Temp\ipykernel\_9428\530051474.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.median()
```

```
[ ]: Area Code      16.0
      Element Code   6132.0
      Item Code      210011.0
      Year Code      20092011.0
      Value          20.0
      Note           NaN
      dtype: float64
```

```
[ ]: df.mode()
```

```
[ ]:   Domain Code      Domain Area Code      Area \
0      FS Suite of Food Security Indicators      16.0 Bangladesh
1      NaN      NaN      NaN      NaN
2      NaN      NaN      NaN      NaN
3      NaN      NaN      NaN      NaN
4      NaN      NaN      NaN      NaN
5      NaN      NaN      NaN      NaN
6      NaN      NaN      NaN      NaN
```

7	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN

	Element Code	Element	Item Code \
0	6132.0	Value	210011.0
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	NaN	NaN	NaN
6	NaN	NaN	NaN
7	NaN	NaN	NaN
8	NaN	NaN	NaN
9	NaN	NaN	NaN
10	NaN	NaN	NaN
11	NaN	NaN	NaN
12	NaN	NaN	NaN
13	NaN	NaN	NaN
14	NaN	NaN	NaN
15	NaN	NaN	NaN
16	NaN	NaN	NaN
17	NaN	NaN	NaN
18	NaN	NaN	NaN

	Item	Year Code	Year \
0	Number of people undernourished (million) (3-y...	20002002	2000-2002
1	NaN	20012003	2001-2003
2	NaN	20022004	2002-2004
3	NaN	20032005	2003-2005
4	NaN	20042006	2004-2006
5	NaN	20052007	2005-2007
6	NaN	20062008	2006-2008
7	NaN	20072009	2007-2009
8	NaN	20082010	2008-2010
9	NaN	20092011	2009-2011
10	NaN	20102012	2010-2012
11	NaN	20112013	2011-2013

12		NaN	20122014	2012-2014
13		NaN	20132015	2013-2015
14		NaN	20142016	2014-2016
15		NaN	20152017	2015-2017
16		NaN	20162018	2016-2018
17		NaN	20172019	2017-2019
18		NaN	20182020	2018-2020

	Unit	Value	Flag	Flag Description	Note
0	millions	20.0	F	FAO estimate	NaN
1	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN

```
[ ]: import pandas as pd
df2 = pd.Series([30,30,35,45,10,61,70,90,115])
df2.mean()
```

```
[ ]: 54.0
```

```
[ ]: df2.median()
```

```
[ ]: 45.0
```

```
[ ]: df2.mode()
```

```
[ ]: 0    30
dtype: int64
```

```
[ ]: df2.std()
```

```
[ ]: 33.26409475695979
```

```
[ ]: df2.var()
```

```
[ ]: 1106.5
```

```
[ ]: df2.min()
```

```
[ ]: 10
```

```
[ ]: df2.max()
```

```
[ ]: 115
```