

## LAPORAN STATISTIKA INFERENSIAL

### (Angka Kelahiran Pada Perempuan Usia 15-19 Tahun Menurut Provinsi tahun 2017)

Nama : Dian Ayu Fauziah

NIM : 22031554011

Kelas : Sains Data 2022B

Data yang digunakan adalah data Angka Kelahiran Pada Perempuan Usia 15-19 Tahun Menurut Provinsi pada Tahun 2017, dengan jumlah total data keseluruhan adalah 34 data. Data acuan diskusi kelompok kami dapat diakses pada link berikut ini : [https://www.bps.go.id/indikator/indikator/view\\_data/0000/data/1397/sdgs\\_3/1](https://www.bps.go.id/indikator/indikator/view_data/0000/data/1397/sdgs_3/1)

1. Langkah pertama yaitu, mengimport data ke Rstudio

**Gambar 1.1**

	A	B
1	Provinsi	Angka Kelahiran Pada Perempuan Usia 15-19 Tahun Menurut Provinsi 2017
2	ACEH	21
3	SUMATERA UTARA	29
4	SUMATERA BARAT	16
5	RIAU	54
6	JAMBI	46
7	SUMATERA SELATAN	46
8	BENGKULU	49
9	LAMPUNG	32
10	KEP. BANGKA BELITUNG	50
11	KEP. RIAU	21
12	DKI JAKARTA	20
13	JAWA BARAT	35
14	JAWA TENGAH	38
15	DI YOGYAKARTA	15
16	JAWA TIMUR	29
17	BANTEN	27
18	BALI	22
19	NUSA TENGGARA BARAT	42
20	NUSA TENGGARA TIMUR	49
21	KALIMANTAN BARAT	63
22	KALIMANTAN TENGAH	83
23	KALIMANTAN SELATAN	56
24	KALIMANTAN UTARA	53
25	KALIMANTAN TIMUR	33
26	SULAWESI UTARA	36
27	SULAWESI TENGAH	60
28	SULAWESI SELATAN	45
29	SULAWESI TENGGARA	46
30	GORONTALO	44
31	SULAWESI BARAT	46
32	MALUKU	51
33	MALUKU UTARA	57
34	PAPUA BARAT	39
35	PAPUA BARAT	65


2. Lakukan pengambilan random sampel dari setiap populasi menggunakan Rstudio

**Gambar 1.2**

```
> data <- data.frame(ID = 1:34)
> sampling <- data[sample(nrow(data), 30), ]
> print(sampling)
[1] 8 21 19 11 26 16 18 7 24 30 10 20 9 3 5 14 33 12 31 13 28 15 32 25 2 34
[27] 6 23 4 27
```

Gambar 1.2 melakukan pengambilan sampel dari populasi secara random dengan cara melabeli setiap populasi dan mengambil 30 data untuk dijadikan sampel dari data populasi keseluruhan

3. Langkah selanjutnya yaitu membuat file excel dari random sampel populasi kemudian file tersebut diimport kembali ke Rstudio untuk dicari nilai mean, modus, median, dll. Lalu, didapatkan hasilnya pada gambar di bawah ini:

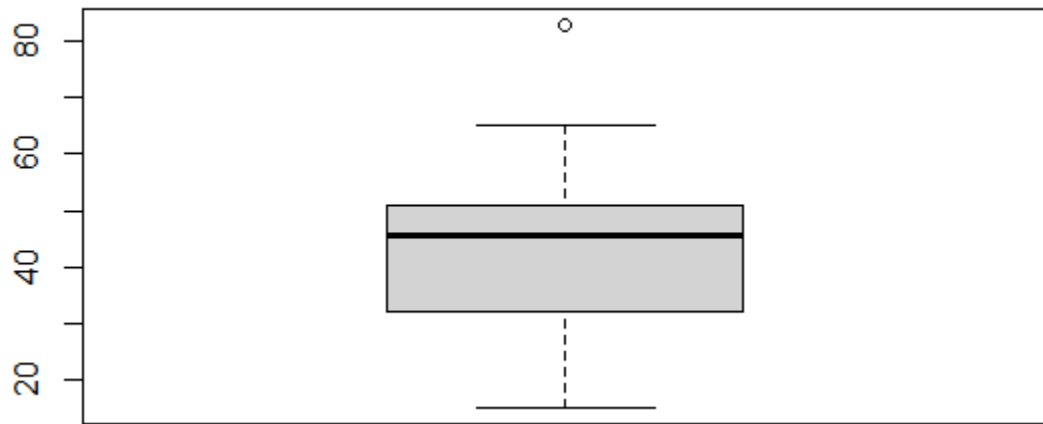


```
R 4.3.1 ~ /
> mydata<-read_excel("D:/User/Downloads/sampel random populasi.xlsx")
> mean(sampel_random_populasi$`nilai random sample`)
[1] 42.5
> median(sampel_random_populasi$`nilai random sample`)
[1] 45.5
> var(sampel_random_populasi$`nilai random sample`)
[1] 239.9138
> sd(sampel_random_populasi$`nilai random sample`)
[1] 15.48915

> getmode <- function(v) {
+   uniqv <- unique(v)
+   uniqv[which.max(tabulate(match(v, uniqv)))]
+ }
> v <- sampel_random_populasi$`nilai random sample`
> hasil <- getmode(v)
> print(hasil)
[1] 46
> summary(sampel_random_populasi$`nilai random sample`)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  15.00  32.25   45.50   42.50   50.75   83.00
```

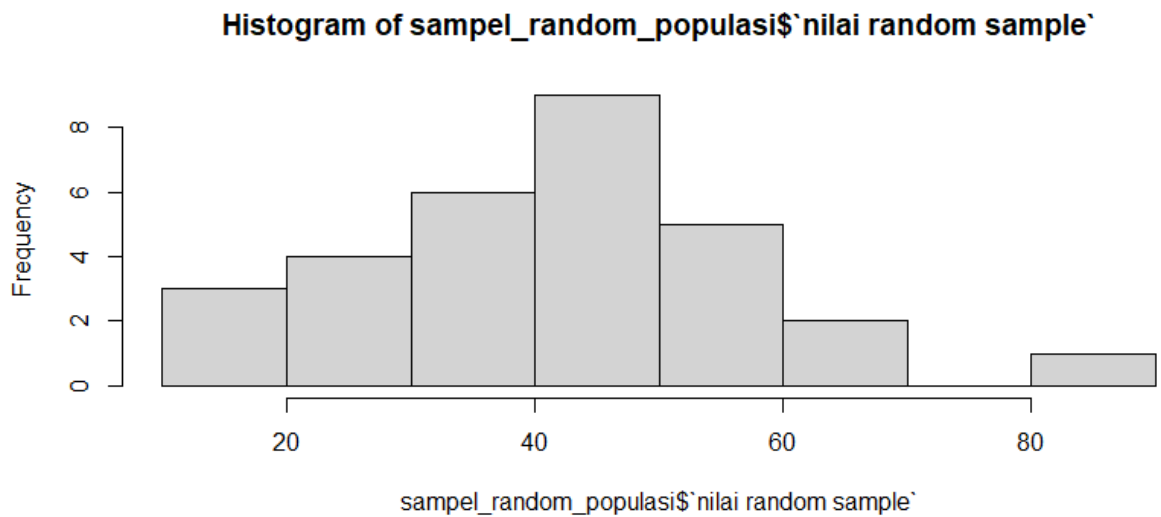
4. Setelah mencari nilai mean, modus, median, dll. Selanjutnya adalah membuat boxplot dan juga histogram dari random sampel populasi tersebut

**Gambar 1.3**



Gambar 1.3 merupakan boxplot dari random sampel populasi

**Gambar 1.4**



Gambar 1.4 merupakan Histogram dari random sampel populasi

5. Terakhir yaitu menghitung secara manual di excel untuk membandingkan hasil perhitungan dari Rstudio apakah valid atau tidak

**Gambar 1.5**

1	Provinsi	Nilai random sampel	Mean	X-X_bar	^2			
2	Lampung	32	42.5	-10.5	110.25			
3	Kalimantan Tengah	83	42.5	40.5	1640.3			
4	Nusa Tenggara Timur	49	42.5	6.5	42.25	sd	15.489	
5	DKI Jakarta	20	42.5	-22.5	506.25	modus	146	
6	Sulawesi Tengah	60	42.5	17.5	306.25	median	45.5	
7	Banten	27	42.5	-15.5	240.25	var	239.91	
8	Nusa Tenggara Barat	42	42.5	-0.5	0.25			
9	Bengkulu	49	42.5	6.5	42.25			
10	Kalimantan Timur	33	42.5	-9.5	90.25			
11	Sulawesi Barat	46	42.5	3.5	12.25			
12	Kep. Riau	21	42.5	-21.5	462.25			
13	Kalimantan Barat	63	42.5	20.5	420.25			
14	Kep. Bangka Belitung	50	42.5	7.5	56.25			
15	Sumatera Barat	16	42.5	-26.5	702.25			
16	Jambi	46	42.5	3.5	12.25			
17	Jogja	15	42.5	-27.5	756.25			
18	papua Barat	39	42.5	-3.5	12.25			
19	Jawa Barat	35	42.5	-7.5	56.25			
20	Maluku	51	42.5	8.5	72.25			
21	Jawa Tengah	38	42.5	-4.5	20.25			
22	Sulawesi Tenggara	46	42.5	3.5	12.25			
23	Jawa Timur	29	42.5	-13.5	182.25			
24	Maluku Utara	57	42.5	14.5	210.25			
25	Sulawesi Utara	36	42.5	-6.5	42.25			
26	Sumatera Utara	29	42.5	-13.5	182.25			
27	Papua	65	42.5	22.5	506.25			
28	Sumatera Selatan	46	42.5	3.5	12.25			
29	Kalimantan Utara	53	42.5	10.5	110.25			
30	Riau	54	42.5	11.5	132.25			
31	Sulawesi Selatan	45	42.5	2.5	6.25			
32					6957.5			

Berikut adalah hasil perhitungan manual statistika descriptive, berupa Mean, Median, dan Modus dan juga Dispersion berupa Range, IQR, Standard Deviasi, dan juga Varians

**Gambar 1.6**

Random sampling :									
32	83	49	20	60	27	42	49	33	46
21	63	50	16	46	15	39	35	51	38
46	29	57	36	29	65	46	53	59	45
Random sampling setelah diurutkan dari yang terkecil :									
15	16	20	21	27	29	29	32	33	35
36	38	39	42	45	46	46	46	46	49
49	50	51	53	54	57	60	63	65	83
J. Statistika Deskriptif									
Mean									
$\text{Mean} = \frac{\sum_{i=1}^n x_i}{n}$									
$= \frac{15 + 16 + 20 + 21 + 27 + 29 + 29 + 32 + 33 + 35 + 36 + 38 + 39 + 42 + 45 + 46 + 46 + 46 + 46 + 49 + 49 + 50 + 51 + 53 + 54 + 57 + 60 + 63 + 65 + 83}{30}$									
= 42,5									
Median									
$15, 16, 20, 21, 27, 29, 29, 32, 33, 35, 36, 38, 39, 42, 45, 46, 46, 46, 46, 49, 49, 50, 51, 53, 54, 57, 60, 63, 65, 83$									
$45 + 46 = \frac{91}{2} = 45,5$									
Modus (data yang paling banyak muncul)									
∴ modus = 46									
nilai 46 muncul sebanyak 4 kali pada random sampel populasi tsb.									

Gambar 1.6 merupakan hasil perhitungan manual statistika descriptive dari data random sampel populasi

Gambar 1.7

II. Dispersion

a.) Range  
 $\text{range} = \text{nilai maks} - \text{nilai min}$   
 $= 83 - 15$   
 $= 68$

b.) IQR  
 $\text{IQR} = Q_3 - Q_1$

Data = 15 16 20 21 27 29 29 32 33 35 36 38 39 42 45  
 46 46 46 46 49 49 50 51 53 54 57 60 63 65 83

lower ( $Q_1$ )  
 upper ( $Q_3$ )

$\text{IQR} = Q_3 - Q_1$   
 $= \left( \frac{51}{2} \right) - \left( \frac{32}{2} \right)$   
 $= (25,5) - (16)$   
 $= 9,5$

Standart Deviasi  
 $\bar{x} = 42,5$   
 $\text{sd} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$

$= \sqrt{\frac{(15-42,5)^2 + (16-42,5)^2 + (20-42,5)^2 + (21-42,5)^2 + (27-42,5)^2 + (29-42,5)^2 + (29-42,5)^2 + (32-42,5)^2 + (33-42,5)^2 + (35-42,5)^2 + (36-42,5)^2 + (38-42,5)^2 + (39-42,5)^2 + (42-42,5)^2 + (45-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (49-42,5)^2 + (49-42,5)^2 + (50-42,5)^2 + (51-42,5)^2 + (53-42,5)^2 + (54-42,5)^2 + (57-42,5)^2 + (60-42,5)^2 + (63-42,5)^2 + (65-42,5)^2 + (83-42,5)^2}{30-1}}$

$\text{sd} = \sqrt{\frac{756,25 + 702,25 + 506,25 + 462,25 + 240,25 + 182,25 + 182,25 + 110,25 + 90,25 + 56,25 + 42,25 + 20,25 + 12,25 + 0,25 + 6,25 + 12,25 + 12,25 + 12,25 + 12,25 + 42,25 + 42,25 + 132,25 + 210,25 + 306,25 + 420,25 + 506,25 + 1640,25}{29}}$

$\text{sd} = \sqrt{\frac{6057,5}{29}} = 15,489$

Gambar 1.8

Varian

$\bar{x} = 42,5$   
 $\sigma^2 = \frac{\sum (x - \bar{x})^2}{n-1}$

$= \frac{(15-42,5)^2 + (16-42,5)^2 + (20-42,5)^2 + (21-42,5)^2 + (27-42,5)^2 + (29-42,5)^2 + (29-42,5)^2 + (32-42,5)^2 + (33-42,5)^2 + (35-42,5)^2 + (36-42,5)^2 + (38-42,5)^2 + (39-42,5)^2 + (42-42,5)^2 + (45-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (46-42,5)^2 + (49-42,5)^2 + (49-42,5)^2 + (50-42,5)^2 + (51-42,5)^2 + (53-42,5)^2 + (54-42,5)^2 + (57-42,5)^2 + (60-42,5)^2 + (63-42,5)^2 + (65-42,5)^2 + (83-42,5)^2}{30-1}$

$= \frac{6057,5}{29} = 209,91$

Gambar 1.7 dan 1.8 merupakan hasil perhitungan manual Dispersion dari data random sampel populasi

## Kesimpulan

Dari data Angka Kelahiran Pada Perempuan Usia 15-19 Tahun Menurut Provinsi tahun 2017 sebanyak 34 data setelah dilakukan perhitungan *Central Tendency*, *Dispersion*, dan juga Varians dengan menggunakan tiga cara yaitu, manual, excel, dan Rstudio menunjukkan hasil yang sama

- *Central Tendency*  
Mean = 42,5  
Modus = 46  
Median = 45,5
- *Dispersion*  
Range = 68  
IQR = 9,5  
Standard Deviasi = 15,489
- Varians  
Var = 339,91