

LAPORAN PRAKTIKUM
STATISTIKA INFORMATIKA
“Pertemuan ke-2: LANGKAH PRAKTIKUM
PEMBENTUKAN DISTRIBUSI FREKUENSI
KUMULATIF”

Diajukan untuk memenuhi salah satu praktikum Mata Kuliah Statistika Informatika yang di
ampu oleh:

Ir., Sri Winiarti, S.T., M.Cs.



Disusun Oleh:

Mohammad Farid Hendianto 2200018401

A / Rabu 10.30 – 13.30 Lab. Jaringan

PROGRAM STUDI INFORMATIKA
UNIVERSITAS AHMAD DAHLAN
FAKULTAS TEKNOLOGI INDUSTRI

TAHUN 2023

KASUS 1:

Kasus 1 berikut ini:

Diberikan data data nilai mahasiswa sebagai berikut yang akan dibuat tabel distribusi frekuensi :.

35 38 43 48 49 51 56 59 60 60

61 63 63 63 65 66 67 67 68 70

70 70 70 71 71 71 72 72 72 73

73 74 74 74 74 75 75 76 76 77

78 79 79 80 80 80 80 81 81 81



82 82 83 83 83 84 85 86 86 87

88 88 88 88 89 90 90 90 91 91

91 92 92 93 93 93 95 97 98 99

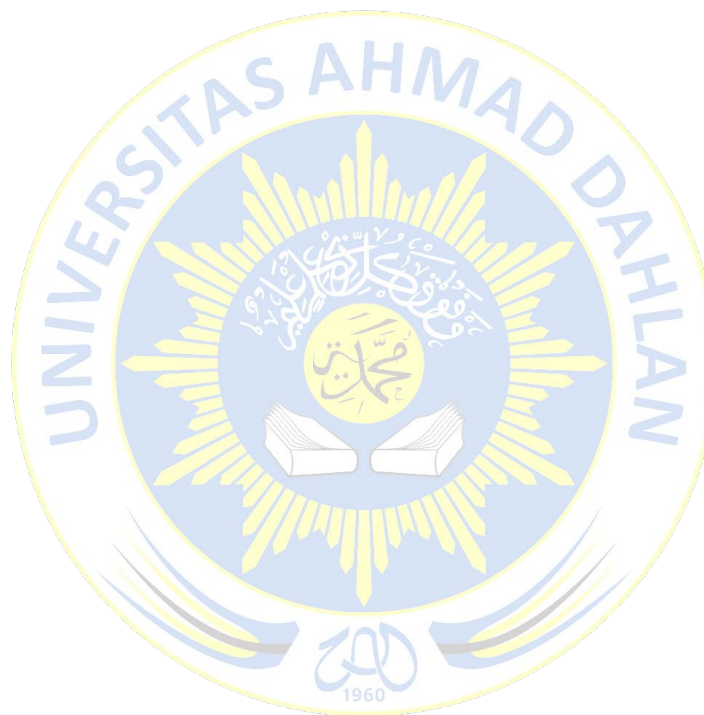
1. Berdasarkan kasus 1 , buatlah table distribusi frekuensi kumaltifnya sesuai langkah 1-5

Berikut adalah tabel yang sudah dibuat di SPSS

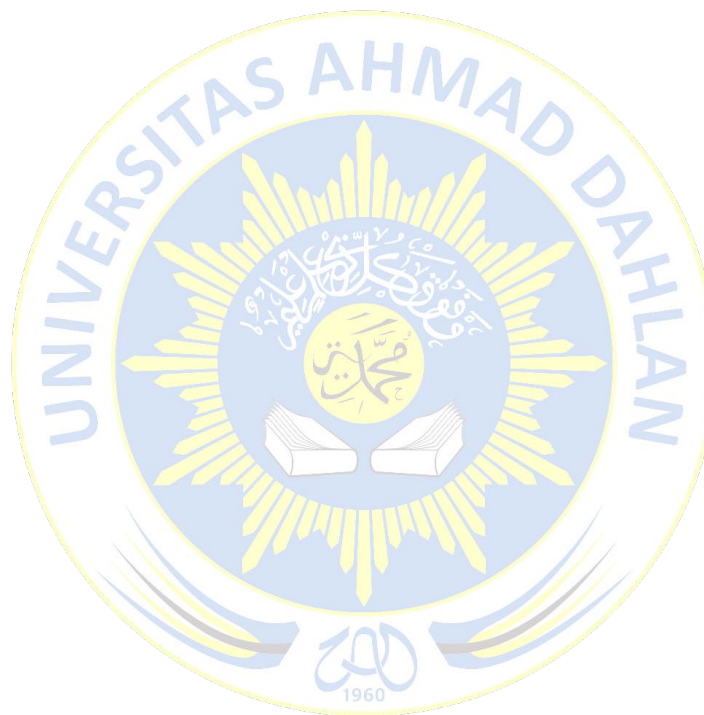
	 valueStudent	 intervalClass	
1	35	1	
2	38	1	
3	43	1	
4	48	2	
5	49	2	
6	51	2	
7	56	3	
8	59	3	
9	60	4	
10	60	4	
11	61	4	
12	63	4	
13	63	4	
14	63	4	
15	65	4	
16	66	4	
17	67	4	
18	67	4	
19	68	5	
20	70	5	
21	70	5	
22	70	5	
23	70	5	
24	71	5	
25	71	5	
26	71	5	
27	72	5	
28	72	5	
29	72	5	
30	73	5	
31	73	5	
32	74	5	
33	74	5	
34	74	5	
35	74	5	
36	75	5	
37	75	5	
38	76	6	

ValueStudent | Interval Class

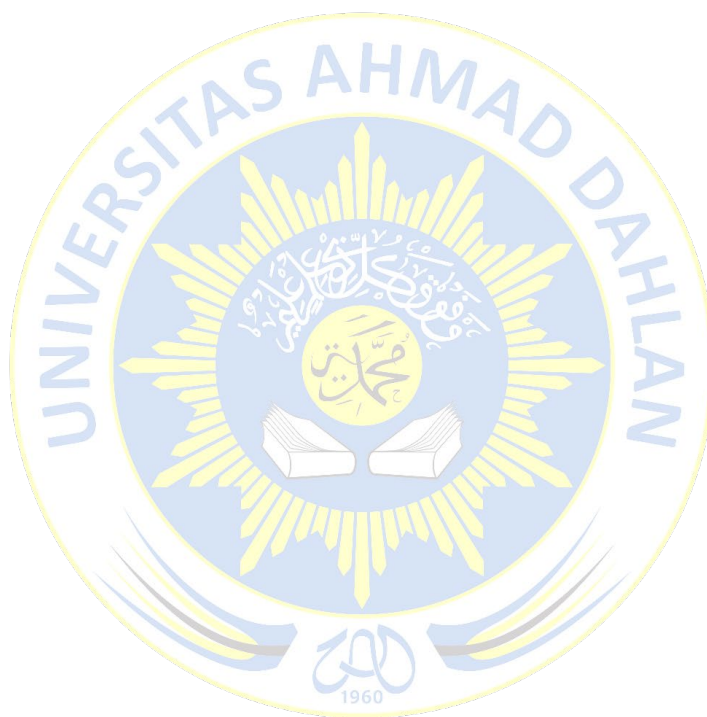
35	1
38	1
43	1
48	2
49	2
51	2
56	3
59	3
60	4
60	4
61	4
63	4
63	4
63	4
65	4
66	4
67	4
67	4
68	5
70	5
70	5
70	5
70	5
71	5
71	5
71	5



72	5
72	5
72	5
73	5
73	5
74	5
74	5
74	5
74	5
75	5
75	5
76	6
76	6
77	6
78	6
79	6
79	6
80	6
80	6
80	6
80	6
81	6
81	6
81	6
82	6
82	6
83	6



83	6
83	6
84	7
85	7
86	7
86	7
87	7
88	7
88	7
88	7
88	7
89	7
90	7
90	7
90	7
91	7
91	7
91	7
92	8
92	8
93	8
93	8
93	8
95	8
97	8
98	8
99	8



Cara mudah untuk membuat kelas interval menggunakan visual binning. Sebelum itu hitunglah R,k dan C nya.

$$R = 99 - 35 = 64$$

$$k = \text{ceil}(1 + 3.332 \log 80) = 8$$

$$C = \text{ceil}(64 / 8) = 8$$

Untuk cutpoints (bagian tengah dipotong) jadinya itu adalah bagian tengah antara kelas atas data 1 dan kelas bawah data ke dua dan seterusnya

Misalnya

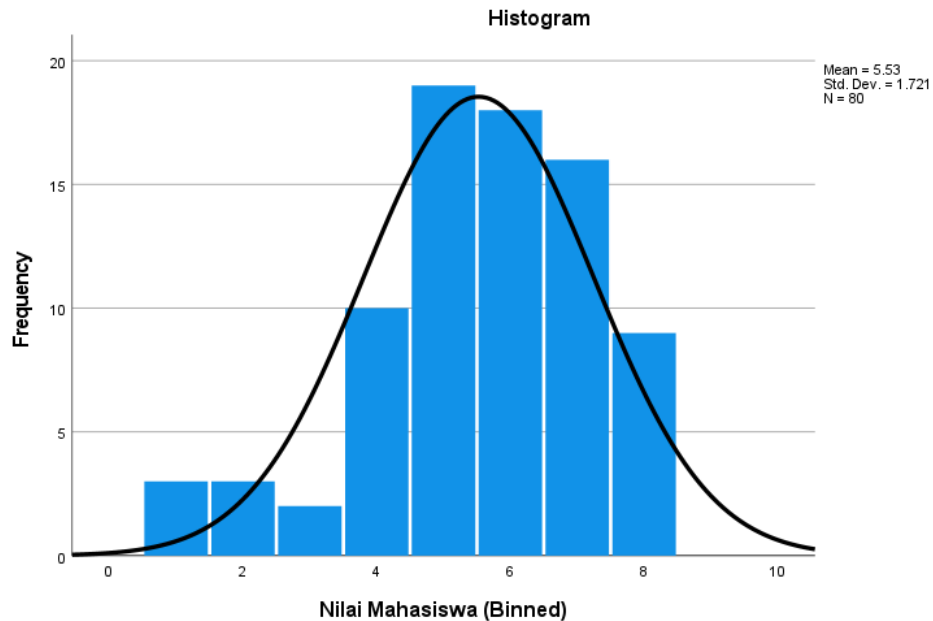
The screenshot shows the SPSS Visual Binning dialog box and the resulting grid table. The dialog box is set to 'Visual Binning' with 'Nilai Mahasiswa [valueStudent]' as the current variable. The binned variable is 'Nilai Mahasiswa (Binned)'. The minimum value is 35 and the maximum is 99. The grid table shows 8 bins with labels '35-43', '43-51', '51-59', '59-67', '67-75', '75-83', '83-91', and '91-99'. The 'Make Cutpoints' button is highlighted.

Grid	Value	Label
1	35.0	35-43
2	43.0	43-51
3	51.0	51-59
4	59.0	59-67
5	67.0	67-75
6	75.0	75-83
7	83.0	83-91
8	91.0	91-99

Akan otomatis membuat label ke interval

Visual binning

2. Buatlah Grafik Frekuensi kumulatifnya



3. Tujukkan hasil hasil olahan data dari kasus 1 dari aplikasi SPSS yang dibuat yang berupa Tabel Distribusi Frekuensi Kumulatif

Statistics

Nilai mahasiswa

N	Valid	Missing
	80	0
Mean	5.53	
Median	6.00	
Mode	5	
Minimum	1	
Maximum	8	
Sum	442	

Nilai Mahasiswa					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	35-43	3	3.8	3.8	3.8
	43-51	3	3.8	3.8	7.5
	51-59	2	2.5	2.5	10.0
	59-67	10	12.5	12.5	22.5
	67-75	19	23.8	23.8	46.3
	75-83	18	22.5	22.5	68.8
	83-91	16	20.0	20.0	88.8
	91-99	9	11.3	11.3	100.0
	Total	80	100.0	100.0	

Berdasarkan data yang diberikan, berikut adalah analisis yang dapat dilakukan:

Data ini menunjukkan statistik nilai mahasiswa. Ada total 80 pengamatan (N=80) dan tidak ada data yang hilang (Missing=0).

Deskripsi Statistik:

Rata-rata (Mean) nilai mahasiswa adalah 5.53.

Nilai tengah atau Median dari kumpulan nilai tersebut adalah 6.

Nilai yang paling sering muncul atau Mode adalah 5.

Nilai terendah atau Minimum dalam kumpulan data ini adalah 1.

Nilai tertinggi atau Maximum dalam kumpulan data ini adalah 8.

Jumlah total semua nilai atau Sum dari seluruh pengamatan adalah 442.

Distribusi Frekuensi:

Distribusi frekuensi menunjukkan sebaran nilai mahasiswa dalam rentang tertentu:

Ada sekitar 3,8% mahasiswa dengan nilai antara 35-43 dan juga antara 43-51

Sekitar 2,5% mahasiswa mendapatkan skor antara 51-59

Sekitar 12,5% skor antara range skor diantara 59-67

Sekitar 23,8% mendapatkan skor diantara 67-75

Sekitar 22,5% mendapatkan skor diantara 75-83

Sekitar 20% mendapatkan skor diantara 83-91

Dan 11,3 % mendapat skor diantara nilai 91-99

Dari distribusi frekuensi tersebut kita bisa melihat bahwa mayoritas siswa (46.3%) memperoleh skor antara "67" hingga "75", dengan persentase kumulatif mencapai hingga "68.8%" untuk siswa yang mencetak kurang dari "83".

4. Bagaimana hasil analisa dari kasus 1...? Jelaskan dengan menentukan:

a. Jelaskan nilai frekuensi kumulatif tertinggi

Nilai frekuensi kumulatif tertinggi adalah 100%, yang mencakup semua data. Nilai ini menunjukkan bahwa semua nilai mahasiswa (100% dari total) berada dalam rentang 35-99.

b. Dari grafik yang dihasilkan tuliskan analisa sebaran tertinggi ada dikelompok data berapa..?

Sebaran nilai tertinggi ada pada kelompok 67-75 dengan persentase sebesar 23.8%. Ini berarti bahwa hampir seperempat dari seluruh mahasiswa memiliki nilai dalam rentang ini.

KASUS 2:

Dilakukan survey terhadap 60 pasien di daerah XYZ yang diduga mengalami positif Covid 19. Usia rata pasien terkena covid antara 20 -50 tahun. Hasil pendataan usia pasien tersebut disajikan sebagai

berikut:

20 22 25 30 34 35 40 45 40 50

47 45 50 55 45 37 30 38 35 55

22 25 33 30 45 40 45 30 50 52

44 22 45 47 30 55 50 45 30 23

25 25 28 30 33 35 40 46 50 57

57 55 50 48 48 30 25 27 30 30

56 50 45 45 40 30 33 36 25 27

1. Berdasarkan kasus 2 , buatlah table distribusi frekuensi kumulatifnya sesuai langkah 1-5

Berikut tabel yang sudah dibuat

patientAge

20

22

25

30

34

35

40

45

40

50

47

45

50

55

45

37

30

38

35

55

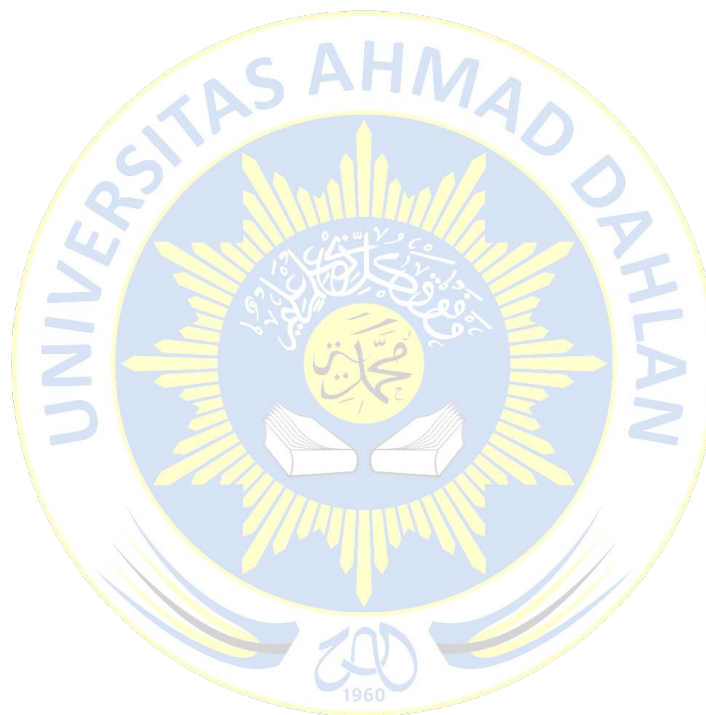
22

25

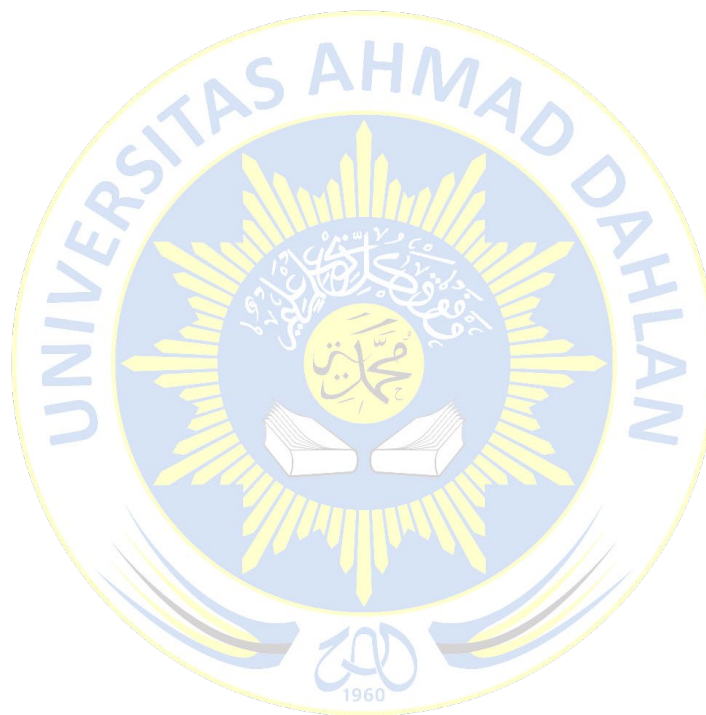
33

30

45



40
45
30
50
52
44
22
45
47
30
55
50
45
30
23
25
25
28
30
33
35
40
46
50
57
57
55



50

48

48

30

25

27

30

30

56

50

45

45

40

30

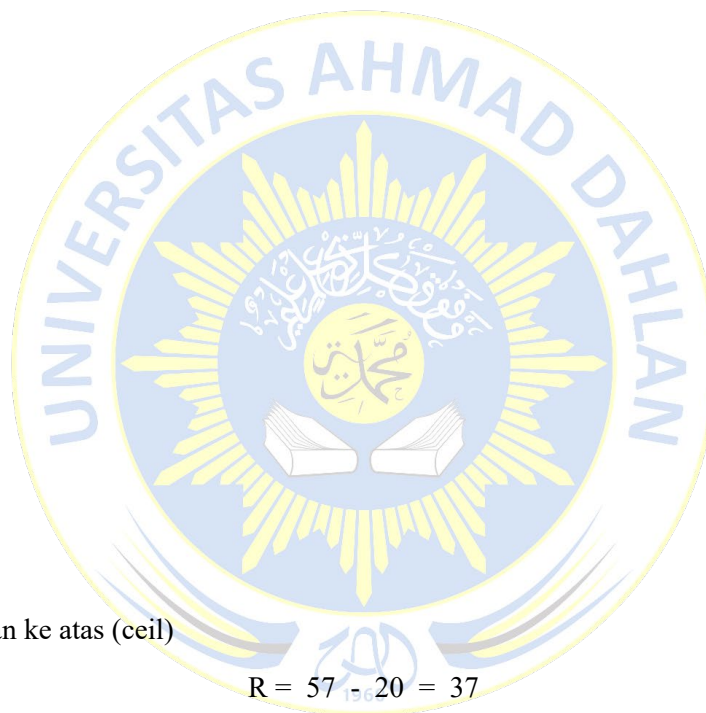
33

36

25

27

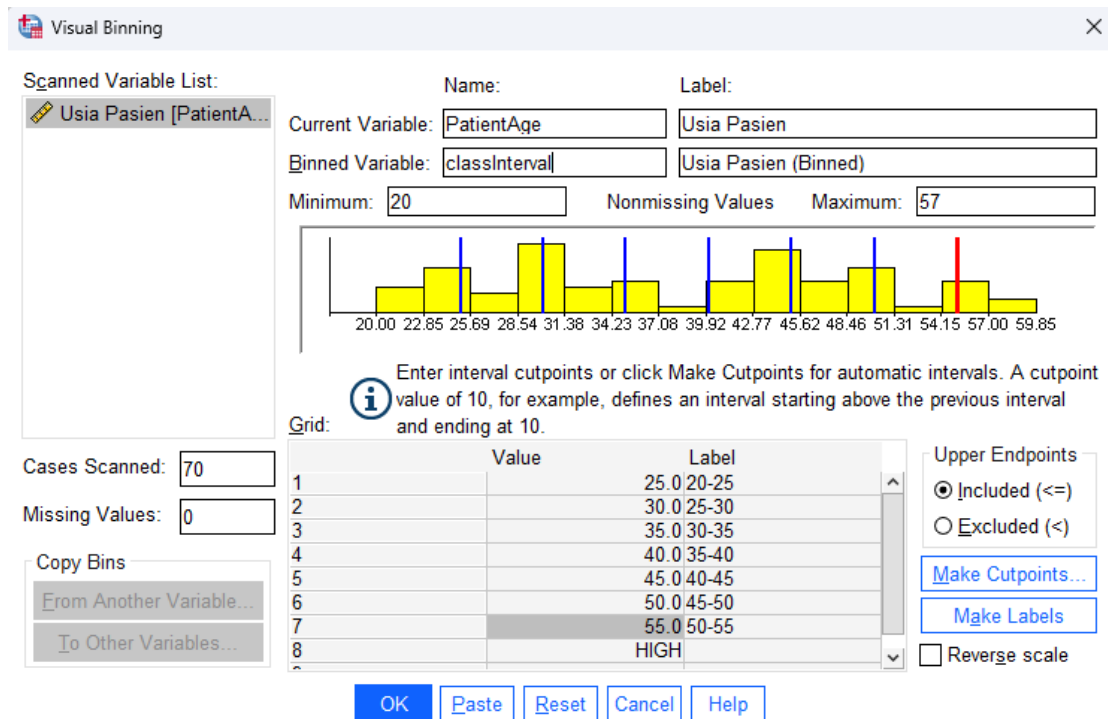
Ingat pembulatan ke atas (ceil)



$$R = 57 - 20 = 37$$

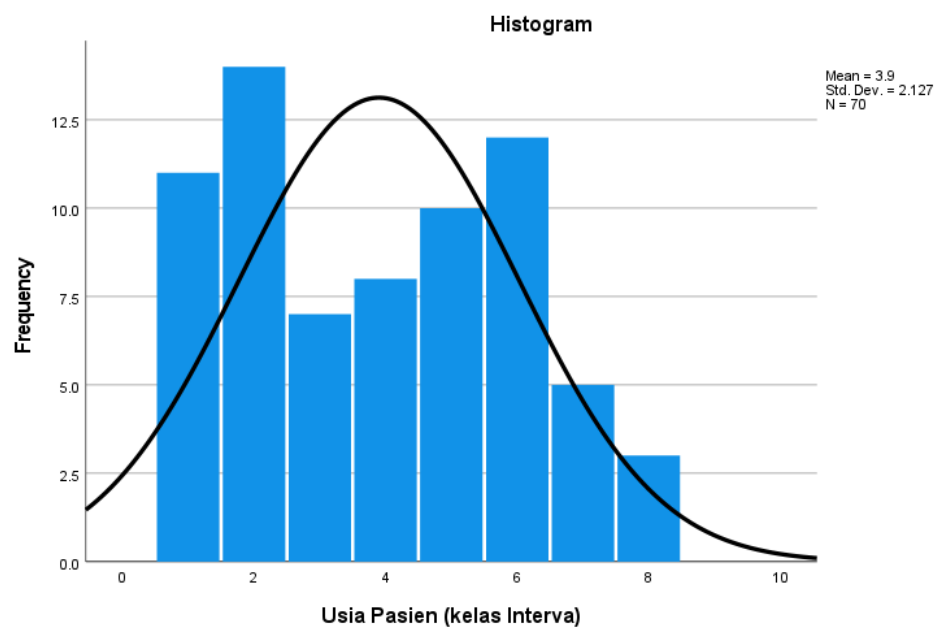
$$k = \text{ceil}(1 + 3.332 \log 70) = 8$$

$$C = \text{ceil}(37 / 8) = 5$$



Tabel yang sudah terbuat

2. Buatlah Grafik Frekuensi kumulatifnya



3. Tujukkan hasil hasil olahan data dari kasus 1 dari aplikasi SPSS yang dibuat yang berupa Tabel Distribusi Frekuensi Kumulatif

Statistics		
Usia Pasien (kelas Interval)		
N	Valid	70
	Missing	0
Mean		3.90
Std. Error of Mean		.254
Median		4.00
Mode		2
Std. Deviation		2.127
Variance		4.526
Range		7
Minimum		1
Maximum		8
Sum		273

Usia Pasien (kelas Interval)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-25	11	15.7	15.7	15.7
	25-30	14	20.0	20.0	35.7
	30-35	7	10.0	10.0	45.7
	35-40	8	11.4	11.4	57.1
	40-45	10	14.3	14.3	71.4
	45-50	12	17.1	17.1	88.6
	50-55	5	7.1	7.1	95.7
	55-60	3	4.3	4.3	100.0
	Total	70	100.0	100.0	

Data ini menggambarkan usia pasien dalam kelas interval, yaitu rentang usia tertentu. Berikut adalah analisis statistik yang dapat kita lakukan berdasarkan data yang diberikan:

Total Pasien: Ada total 70 pasien dalam data ini dan tidak ada data yang hilang.

Usia Rata-rata: Rata-rata usia pasien adalah 3.90, tetapi karena ini dalam bentuk kelas interval, nilai rata-ratanya mungkin tidak secara langsung mewakili usia sebenarnya. Misalnya, jika kita mengasumsikan bahwa angka-angka dari 1 sampai 8 mewakili kelas interval (mis., 1 = 20-25 tahun, 2 = 25-30 tahun, dst.), maka rata-rata "3.90" akan lebih dekat ke kelas interval keempat (35-40 tahun).

Median Usia: Median adalah nilai tengah dari set data ketika diurutkan dari terendah ke tertinggi. Dalam hal ini, median adalah 4 (kelas interval keempat atau rentang usia 35-40).

Modus Usia: Modus merujuk pada nilai yang paling sering muncul dalam set data dan dalam kasus ini modusnya adalah '2' yang menunjukkan bahwa kelompok umur antara '25-30' memiliki jumlah pasien terbanyak.

Standar Deviasi & Varians Usia: Standar deviasi dan varians merupakan ukuran sebaran atau variasi dalam set data tersebut. Standar deviasi dari usia pasien adalah sekitar '2', dan variannya sekitar '4'. Ini menunjukkan bahwa ada cukup variasi di antara usia pasien.

Rentang Usia: Rentang merujuk pada perbedaan antara nilai maksimum dan minimum dalam suatu set data; disini itu mencapai angka '7', mencerminkan perbedaan antara kelompok umur termuda ('20-25') dan termuda ('55-60').

Distribusi Frekuensi Usia Pasien: Dari tabel frekuensi kita bisa melihat distribusi persentase pasien berdasarkan kelompok umur:

Kelompok umur dengan jumlah terbanyak adalah '45-50' dengan persentase ~17%.

Diikuti oleh kelompok umur '25-30' (~20%), kemudian '40-45' (~14%), lalu '20-25' (~16%).

Kelompok umur lainnya memiliki persentase lebih rendah.

Dari analisis di atas dapat disimpulkan bahwa mayoritas pasien berada pada rentang umur antara '45-50', namun ada juga signifikan jumlah individu pada rentang umur '25-30' dan '40-

45'. Variasi usia yang cukup besar menunjukkan bahwa kondisi kesehatan atau penyakit yang sedang diteliti mungkin mempengaruhi berbagai kelompok umur.

Dari 70 pasien, sebanyak 11 orang atau 15.7% berada dalam rentang usia 20-25 tahun. Pasien berusia antara 25-30 tahun merupakan kelompok terbesar kedua dengan 14 orang atau sekitar 20% dari total populasi sampel. Jumlah ini membuat total persentase kumulatif menjadi sekitar 35.7%.

Selanjutnya, ada penurunan jumlah pada kelompok umur selanjutnya, yaitu pada rentang usia antara 30-35 tahun dengan hanya ada sebanyak tujuh orang atau sekitar sepuluh persen dari populasi sampel. Ini membawa jumlah kumulatif menjadi sekitar 45.7%.

Rentang umur selanjutnya adalah antara 35-40 tahun yang mencakup delapan individu atau kurang lebih memuat satu per empat belas bagian (11.4%) dari populasi sampel dan menjadikan total persentase kumulatif menjadi lebih dari setengah, tepatnya menjadi sekitar lima per tujuh bagian (57.1%).

Kelompok umur selanjutnya adalah antara umur '40-45' yang memiliki sepertiga lebih banyak dibandingkan dengan kelompok umur '35-40', yaitu terdiri atas sepuluh individu atau membawahi dua belas bagian ratusan (14%) dari populasi sampel sehingga membuat total persentase kumulatif melebihi tiga per empat bagian (71%).

Kelompok umur '45-50' memiliki dua belas individu atau memuat dua belas bagian seratusan (17%) dari populasi sampel dan merupakan kelompok dengan jumlah individu terbesar diantara semua kelompok yang lain.

Pasca itu, ada penurunan jumlah signifikan pada kelompok umur '50-55' yang hanya memuat lima individu saja atau kurang lebih memuat satu per lima belas bagian (~7%) dan menjadikan total persentase kumulatif mendekati sempurna yaitu ~96%.

Terakhir, pada rentang usia tertua diantara semua rentangan yaitu '55-60', terdapat tiga individu saja atau hanya mengisi empat ratusan bagian (~4%). Meskipun begitu, keberadaan mereka membuat kesimpulan bahwa semua individu telah masuk kedalam analisis ini.

Dengan demikian dapat disimpulkan bahwa rentang usia pasien sangat beragam, mulai dari 20 hingga 60 tahun, dengan jumlah terbanyak berada dalam rentang usia 25-30 tahun dan jumlah terkecil pada rentang usia 55-60 tahun.

4. Bagaimana hasil analisa dari kasus 1...? Jelaskan dengan menentukan:

a. Jelaskan nilai frekuensi kumulatif tertingginya

Nilai frekuensi kumulatif tertinggi dalam data ini adalah 100%, yang mencakup semua pasien dalam sampel. Ini tercapai pada rentang usia terakhir, yaitu 55-60 tahun. Frekuensi kumulatif ini berarti bahwa 100% dari total pasien (70 orang) berada dalam rentang usia antara 20 hingga 60 tahun.

b. Dari garfik yang dihasilkan tuliskan analisa sebaran tertinggi ada dikelompok data berapa..?

berdasarkan tabel frekuensi yang diberikan, sebaran tertinggi ada di kelompok umur '25-30' dan '45-50'. Kelompok umur '25-30' memiliki frekuensi sebesar 20% dan kelompok umur '45-50' memiliki frekuensi sebesar 17.1%. Jadi, jika kita membuat grafik berdasarkan tabel ini, puncaknya kemungkinan akan ada di kelompok umur '25-30' atau bisa juga memiliki dua puncak di kelompok umur '25-30' dan '45-50'.

