**Probability and Statistics**

**Assignment 1**

**Total Mark:100**

**Question No.1: (5)**

Listed are the weights of the NBA’s top 50 players. Construct a grouped frequency distribution and a cumulative frequency distribution with 8 classes. Analyze the results in terms of peaks, extreme values, etc.

240 210 220 260 250 195 230 270 325 225

165 295 205 230 250 210 220 210 230 202

250 265 230 210 240 245 225 180 175 215

215 235 245 250 215 210 195 240 240 225

260 210 190 260 230 190 210 230 185 260

**Solution:**

Min = 165

Max = 325

K = 8

Range = Max – Min

= 325 – 165 = 160

Width = Range / K = 160 / 8 = 20

|  |  |  |  |
| --- | --- | --- | --- |
| **Class Limit** | **Class Boundaries** | **Frequency** | **Cumulative Frequency** |
| 165 – 185 | 164.5 – 185.5 | 4 | 4 |
| 186 – 206 | 185.5 – 206.5 | 6 | 10 |
| 207 – 227 | 206.5 – 227.5 | 15 | 25 |
| 228 – 248 | 227.5 – 248.5 | 13 | 38 |
| 249 – 269 | 248.5 – 269.5 | 9 | 47 |
| 270 – 290 | 269.5 – 290.5 | 1 | 48 |
| 291 – 311 | 290.5 – 311.5 | 1 | 49 |
| 312 – 332 | 311.5 – 332.5 | 1 | 50 |

A peak occurs in class 207–227 (206.5–227.5). There are no gaps in the distribution, and there is one value in each of the three highest classes.

**Question No.2: (5)**

The ages of the signers of the Declaration of Independence are shown. (Age is approximate since only the birth year appeared in the source, and one has been omitted since his birth year is unknown.) Construct a grouped frequency distribution and a cumulative frequency distribution for the data using 7 classes.

41 54 47 40 39 35 50 37 49 42 70 32

44 52 39 50 40 30 34 69 39 45 33 42

44 63 60 27 42 34 50 42 52 38 36 45

35 43 48 46 31 27 55 63 46 33 60 62

35 46 45 34 53 50 50

**Solution:**

Min = 27

Max = 70

K = 7

Range = Max – Min

= 70 – 27 = 43

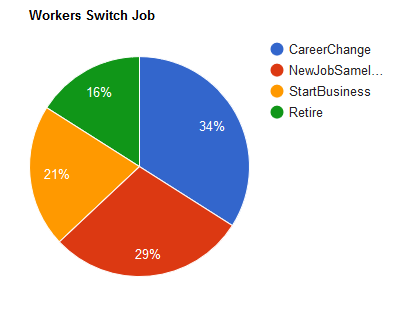
Width = Range / K = 43 / 7 = 6

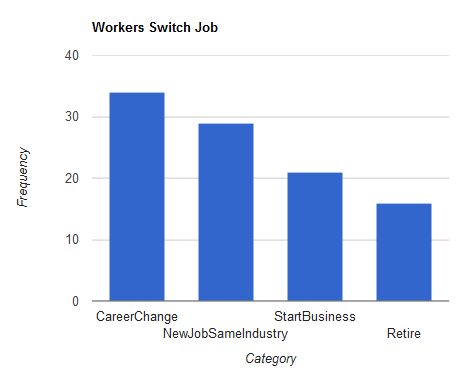
|  |  |  |  |
| --- | --- | --- | --- |
| **Class Limit** | **Class Boundaries** | **Frequency** | **Cumulative Frequency** |
| 27 – 33 | 26.5 -33.5 | 6 | 6 |
| 34 – 40 | 33.5 – 40.5 | 15 | 21 |
| 41 – 47 | 40.5 – 47.5 | 15 | 36 |
| 48 – 54 | 47.5 – 54.5 | 11 | 47 |
| 55 – 61 | 54.5 – 61.5 | 3 | 50 |
| 62 – 68 | 61.5 – 68.5 | 3 | 53 |
| 69 – 75 | 68.5 -75.5 | 2 | 55 |

**Question No.3: (10)**

In a recent survey, 3 in 10 people indicated that they are likely to leave their jobs when the economy improves. Of those surveyed, 34% indicated that they would make a career change, 29% want a new job in the same industry, 21% are going to start a business, and 16% are going to retire. Make a pie chart and a Pareto chart for the data. Which chart do you think better represents the data?

**Solution:**

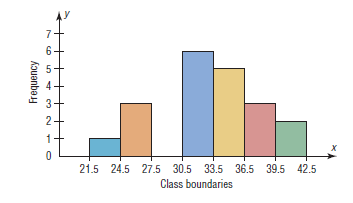




The Pareto chart better represent the data, because the differences among the categories are more obvious in the Pareto chart than the pie graph.

**Question No.4: (10)**

Using the histogram shown here, do the following.



a. Construct a frequency distribution; include class limits, class frequencies, midpoints, and cumulative frequencies.

b. Construct a frequency polygon.

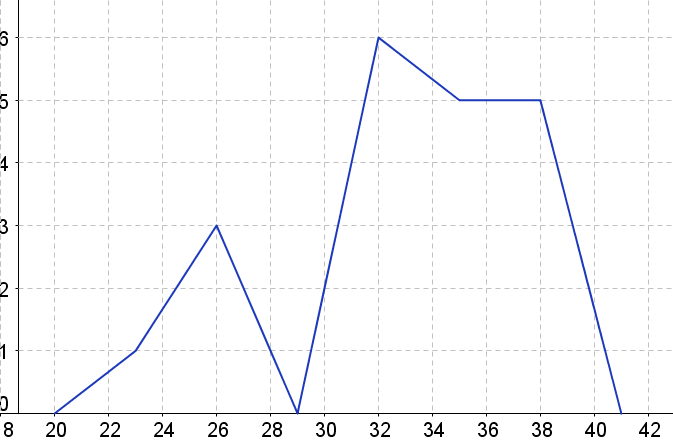
c. Construct an ogive.

**Solution:**

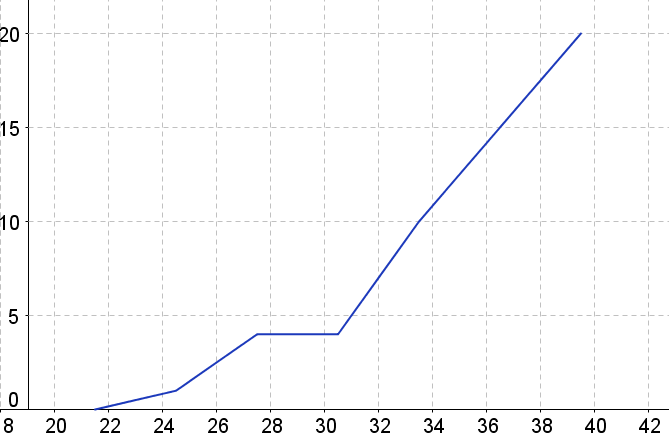
**a.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Class Boundaries** | **Midpoints** | **Frequency** | **Cumulative Frequency** |
| 21.5 – 24.5 | 23 | 1 | 1 |
| 24.5 - 27.5 | 26 | 3 | 4 |
| 27.5 – 30.5 | 29 | 0 | 4 |
| 30.5 – 33.5 | 32 | 6 | 10 |
| 33.5 – 36.5 | 35 | 5 | 15 |
| 36.5 – 39.5 | 38 | 3 | 18 |
| 39.5 – 42.5 | 41 | 2 | 20 |

**b.**



**c.**



**Question No.5: (10)**

The amount of protein (in grams) for a variety of fast-food sandwiches is reported here. Construct a frequency distribution using 6 classes. Draw a histogram, a frequency polygon, and an ogive for the data, using relative frequencies. Describe the shape of the histogram.

23 30 20 27 44 26 35 20 29 29

25 15 18 27 19 22 12 26 34 15

27 35 26 43 35 14 24 12 23 31

40 35 38 57 22 42 24 21 27 33

**Solution:**

Min = 12

Max = 57

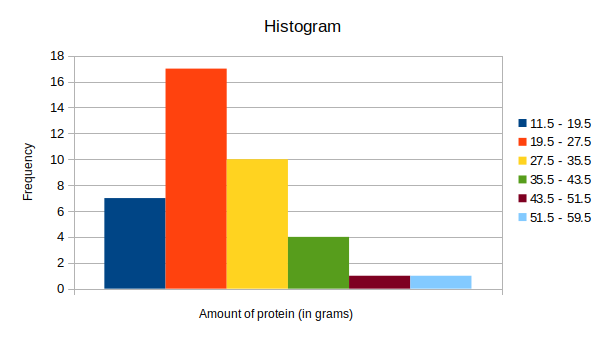
K = 6

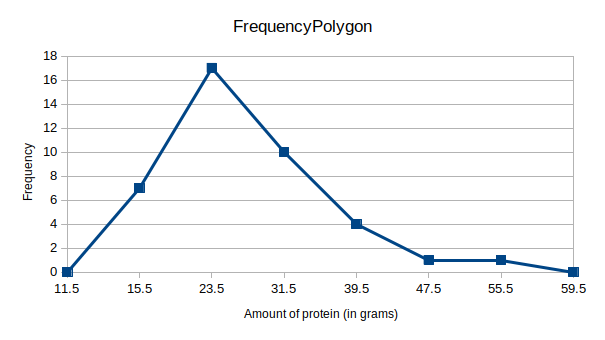
Range = Max – Min

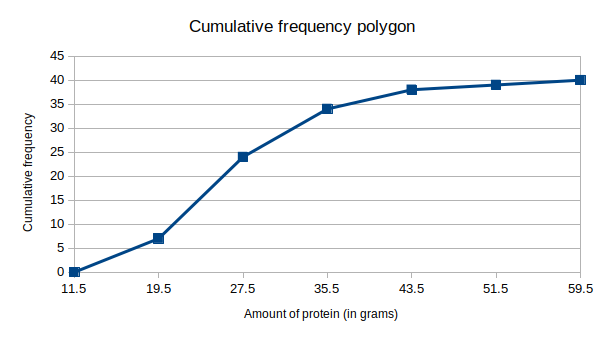
= 57 – 12 = 45

Width = Range / K = 45 / 6 = 7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class Limit** | **Class Boundaries** | **Midpoints** | **Frequency** | **Cumulative Frequency** |
| 12 – 19 | 11.5 - 19.5 | 15.5 | 7 | 7 |
| 20 – 27 | 19.5 – 27.5 | 23.5 | 17 | 24 |
| 28 – 35 | 27.5 – 35.5 | 31.5 | 10 | 34 |
| 36 – 43 | 35.5 – 43.5 | 39.5 | 4 | 38 |
| 44 – 51 | 43.5 – 51.5 | 47.5 | 1 | 39 |
| 52 – 59 | 51.5 – 59.5 | 55.5 | 1 | 40 |







**Question No.6: (5)**

A group of 20 people played a PC game. The table below shows the frequency distribution of their scores.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Score | 1 | 2 | 4 | x |
| Number of people | 2 | 5 | 7 | 6 |

Given that mean score is 5. Find the value of x.

**Solution:**

|  |  |  |
| --- | --- | --- |
| **Score(y)** | **Number of people (f)** | **F \* y** |
| 1 | 2 | 2 |
| 2 | 5 | 10 |
| 4 | 7 | 28 |
| X | 6 | 6x |

sum (f \* y) = 40 + 6x

sum (f) = 20

Mean = 5

Mean= sum (f \* y) / sum (f)

5 = (40 + 6x) / 20

5 \* 20= 40 + 6x

100 = 40 + 6x

x = (100 – 40) / 6

x = 10

**Question No.7: (10)**

A recent survey of a new diet cola reported the following percentages of people who liked the taste. Find the weighted mean of the percentages.

|  |  |  |
| --- | --- | --- |
| **Area** | **% Favored** | **Number surveyed** |
| 1 | 40 | 1000 |
| 2 | 30 | 3000 |
| 3 | 50 | 800 |

**Solution:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Area** | **% Favored(w)** | **Number surveyed(x)** | **w \* x** |
| 1 | 40 | 1000 | 40000 |
| 2 | 30 | 3000 | 90000 |
| 3 | 50 | 800 | 40000 |

Sum (w \* x) = 40000 + 90000 + 40000 = 170000

Sum (x) = 1000 + 3000 + 800 = 4800

Weighted mean of the percentages = Sum (w \* x) / Sum (x)

= 170000/4800 = 35.4 %

**Question No.8: (5)**

Consider the following data which represent lives of batteries.

|  |  |
| --- | --- |
| **Class Boundary** | **Frequency** |
| 1.45-1.95 | 2 |
| 1.95-2.45 | 1 |
| 2.45-2.95 | 4 |
| 2.95-3.45 | 15 |
| 3.45-3.95 | 10 |
| 3.95-4.45 | 5 |
| 4.45-4.95 | 3 |

Find the Mean.

**Solution:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Class Boundary** | **Frequency(f)** | **Midpoints(x)** | **F \* x** |
| 1.45-1.95 | 2 | 1.7 | 3.4 |
| 1.95-2.45 | 1 | 2.2 | 2.2 |
| 2.45-2.95 | 4 | 2.7 | 10.8 |
| 2.95-3.45 | 15 | 3.2 | 48 |
| 3.45-3.95 | 10 | 3.7 | 37 |
| 3.95-4.45 | 5 | 4.2 | 21 |
| 4.45-4.95 | 3 | 4.7 | 14.1 |

sum (f \* y) = 136.5

sum (f) = 40

Mean= sum (f \* y) / sum (f) = 136.5 / 40 = 3.4125

**Question No.9: (10)**

The data shown are the total compensation (in millions of dollars) for the 50 top-paid CEOs for a recent year.

17.5 18.0 36.8 31.7 31.7

17.3 24.3 47.7 38.5 17.0

23.7 16.5 25.1 17.4 18.0

37.6 19.7 21.4 28.6 21.6

19.3 20.0 16.9 25.2 19.8

25.0 17.2 20.4 20.1 29.1

19.1 25.2 23.2 25.9 24.0

41.7 24.0 16.8 26.8 31.4

16.9 17.2 24.1 35.2 19.1

22.9 18.2 25.4 35.4 25.5

Find the Median.

**Solution:**

Arrange the value in ascending order:

16.5, 16.8, 16.9, 16.9, 17.0, 17.2, 17.2, 17.3, 17.4, 17.5, 18.0, 18.0, 18.2, 19.1, 19.1, 19.3, 19.7, 19.8, 20.0, 20.1, 20.4, 21.4, 21.6, 22.9, 23.2, 23.7, 24.0, 24.0, 24.1, 24.3, 25.0, 25.1, 25.2, 25.2, 25.4, 25.5, 25.9, 26.8, 28.6, 29.1, 31.4, 31.7, 31.7, 35.2, 35.4, 36.8, 37.6, 38.5, 41.7, 47.7

n = 50

Median = ½ [the value of (n/2) th + the value of ((n/2) + 1) th]

= ½ [the value of (50/2) th + the value of ((50/2) + 1) th]

= ½ [the value of (25) th + the value of (26) th]

= (23.2 + 23.7) / 2

Median = 23.45

**Question No.10: (10)**

Thirty automobiles were tested for fuel efficiency (in miles per gallon). This frequency distribution was obtained. Find the Median.

|  |  |
| --- | --- |
| **Class boundaries** | **Frequency** |
| 7.5–12.5 | 3 |
| 12.5–17.5 | 5 |
| 17.5–22.5 | 15 |
| 22.5–27.5 | 5 |
| 27.5–32.5 | 2 |

**Solution:**

|  |  |  |
| --- | --- | --- |
| **Class boundaries** | **Frequency** | **CF** |
| 7.5–12.5 | 3 | 3 |
| 12.5–17.5 | 5 | 8 |
| 17.5–22.5 | 15 | 23 |
| 22.5–27.5 | 5 | 28 |
| 27.5–32.5 | 2 | 30 |

N=sum(f)/2=30/2=15

So median class is 17.5-22.5

H=22.5-17.5=5

F=15

C=8

L=17.5

Median=l+(h/f) ((n/2) -c)

M= 17.5+(5/15) (15 -8)

M=17.5+(0.33) (7)

M=19.81

**Question No.11: (10)**

**Earnings of Nonliving Celebrities** *Forbes* magazine prints an annual Top-Earning Nonliving Celebrities list (based on royalties and estate earnings). Find the Median and Mode. Figures represent millions of dollars.

**Kurt Cobain**  50 **Ray Charles** 10

**Elvis Presley** 42 **Marilyn Monroe** 8

**Charles M. Schulz** 35 **Johnny Cash** 8

**John Lennon** 24  **J.R.R. Tolkien** 7

**Albert Einstein** 20 **George Harrison** 7

**Andy Warhol** 19 **Bob Marley** 7

**Theodore Geisel** 10

**(exercise 3.1 question 6 page 118)**

**Solution:**

Arrange the value in ascending order:

7, 7, 7, 8, 8, 10, 10, 19, 20, 24, 35, 42, 50

n = 13

Median = [the value of (n+1)/2 th]

= [the value of (14/2) th]

= [the value of (7) th]

Median = 10

Most occurring value is 7 so

Mode = 7

**Question No.12: (10)**

Consider the following data.

|  |  |  |
| --- | --- | --- |
| **Class Limit** | **Class Boundary** | **Frequency** |
| 7-18 | 6.5-18.5 | 6 |
| 19-30 | 18.5-30.5 | 10 |
| 31-42 | 30.5-42.5 | 13 |
| 43-54 | 42.5-54.5 | 8 |
| 55-66 | 54.5-66.5 | 5 |
| 67-78 | 66.5-78.5 | 6 |
| 79-90 | 78.5-90.5 | 2 |

Find the mode.

**Solution:**

Modal Class = 30.5 – 42.5

F1 = 10

Fm = 13

F2 = 8

H = 42.5 – 30.5 = 12

L = 30.5

Mode = L + ((Fm - F1) / ((Fm - F1) + (Fm – F2))) \*H

= 30.5 + ((13- 10) / ((13- 10) + (13– 8))) \*12

Mode = 35