CENTRAL UNIVERSITY

CENTRAL BUSINESS SCHOOL

DEPARTMENT OF BANKING AND FINANCE

CBFW 102 (3 CREDITS)

BUSINESS STATISTICS

LEVEL 100

AUGUST, 2018

DURATION: 3 Hours

STUDENT ID No..

INSTRUCTIONS

ANSWER ALL QUESTONS IN SECTION A AND ANY TWO IN SECTION B

DO NOT TURN OVER THIS PAGE UNTIL YOU HAVE BEEN TOLD TO DO SO BY THE INVIGILATOR

LECTURER: JUSTICE KUDJO AMPIAH

1. Statistics can be defined as,

    A. methods of organizing, summarizing and presenting data in an informative way.
    B. the science of collecting, organizing, presenting, analyzing, and interpreting data to assist making more decisions.
    C. the methods used of find out something about a population, based on a sample.
    D. a collection of all possible individuals, objects, or measurements of interest.

2. Data can be classified according to levels of measurement. Which of the following is not a level of measurement of data?

    A. Nominal Level
    B. Ration Level
    C. Ordinal Level
    D. Interval Level

3. A violin student records the number of hours she spends practicing during each of ten consecutive weeks:-

    6.0   5.0   4.0   7.0   5.0   7.0   8.0   1.0   6.0   3.0
    What is the median number of hours spent practicing per week during this period?

    A. 5.5 hours
    B. 8.0 hours
    C. 6.0 hours
    D. 5.0 hours

4. A Type II Error is committed when;

    A. A incorrect H_o is accepted
    B. A correct H_o is rejected
    C. An H_1 is accepted
    D. An H_1 is rejected

2

5. The ages of eighteen (18) faculty members in a Business Department of the Novoya University are represented in a stemplot as follows:

$$
\begin{array}{r|l}
3 & 2\ 4\ 8\ 9\ 9 \\
4 & 0\ 3\ 5\ 6\ 9 \\
5 & 3\ 4\ 7\ 8\ 9\ 9 \\
6 & 3\ 8
\end{array}
$$

What are the actual values in the fourth row data?
   A. 63 68
   B. 3 8
   C. 3 4 7 8 9 9
   D. 53 54 57 58 59 59

6. A value between zero and one, inclusive, describing the relative possibility (chance or likelihood) an event will occur is called

   A. Experiment
   B. Outcome
   C. Sample space
   D. Probability

7. A survey

   A.   is a study of a portion or part of the population
   B.   is a study of the entire population
   C.   is a body of facts that are in a format suitable for decision making.
   D.   is a portion, or part, of the population of interest.

8. To select a sample of undergraduate students in Ghana, a simple random sample of four regions was selected. From each of these four regions, a simple random sample of two polytechnics was selected. Finally, from each of these eight polytechnics, a simple random sample of 20 undergraduates was selected. What is the sample size?

   A. 20
   B. 160
   C. 80
   D. 40

3

9. Population can be defined as,

    A. the science of collecting, organizing, presenting , analyzing, and interpreting data to
       assist in making more effective decisions
    B. the methods of organizing, summarizing, and presenting data in an informative way.
    C. an item of interest that can take on many different numerical values.
    D. the collection of all possible observations of a specified characteristic of interest.

10. An equation that determines the relationship between two variables is

    A. Functional Equation
    B. Complex Equation
    C. Correlation Equation
    D. Regression Equation

11. Which of the following is **not** a measure of dispersion?

    A. Mean
    B. Quartile Deviation
    C. Mean deviation
    D. Standard deviation

12. The variance is:

    A. the square of the standard deviation
    B. the square root of the standard
    C. square of the mean deviation
    D. square of the standard error.

13. Stratified sampling scheme can only be used when the population is

    A. Heterogeneous
    B. Homogeneous
    C. Large
    D. Infinite

14. Which of these is an advantage of personal interview?

    A. Saves the researcher time and money.
    B. High response rate
    C. Respondents answer questions the way the understand them.

4

D. No interviewer-bias

15. The coefficient of correlation is a measure of

    A. the strength of the linear relationship between two variables.
    B. how close the actual values are to the regression line.
    C. the difference between a sample statistic and its corresponding population parameter.
    D. the percent of the variation in Y that is explained by the variation in X

16. A large mass of data can best be summarized pictorially by means of:

    A. the range
    B. a histogram
    C. the mean
    D. the frequency table

17. For a symmetric distribution, the mean, median and mode are

    A. insufficient information
    B. the same
    C. always different
    D. possibly the same, possibly different

18. Marital Status and Gender are examples of:

    A.   Quantitative Variable
    B.   Qualitative Variable
    C.   Continuous Variable
    D   Discrete Variable

19. The following is the percent change in net income from 1997 to 1998 for a sample of 12 construction companies in Donvar.

    5   1   -10   -6   5   12   7   8   2   5   -1   11.

What is the mode?

    A. 12
    B. 11
    C. 5
    D. 3

20. If the mean, median and mode of a distribution are 5, 6, 7 respectively, then the distribution is:
   A. skewed negatively
   B. not skewed
   C. skewed positively
   D. symmetrical

21. Consider these five values of a population: 10, 3, 7, 3, and 4. Determine the range.

   A. 4
   B. 10.0
   C. 5.0
   D. 7.0
   E. 3.0

22. A Hypothesis is

   A. a statement about the sample developed for the purpose of testing
   B. a statement about the value of the population parameter
   C. a statement about a population developed for the purpose of testing
   D. a statement about the study.

23. In Linear Regression and Correlation an independent variable is

   A.   the variable that is being predicted or estimated.
   B.   plotted on Vertical axis
   C.   a variable that provides the basis for estimation
   D.   the Output Variable.

24. Pearson's Product Moment Correlation Coefficient is used when data is in

   A.   Nominal Scale and Ordinal Scale
   B   Interval Scale and Ratio Scale
   C.   Ordinal Scale Interval Scale
   D.   Ratio Scale and Nominal Scale

25. Which of the following is a source of primary data?

   A. Personal investigation

B. Non Official Statistics
C. Official Statistics
D. Published statistics.

26. The act of making an estimate of response within the observed range of the independent variable or inside the observed region of a set of independent variables is known as:

    A.   Intrapolation
    B.   Exterpolation
    C.   Extrapolation
    D.   Interpolation

27. Which of the following is a source of secondary data?

    A. Published statistics
    B. Teams of investigators
    C. Postal questionnaires
    D. Personal investigation

28. Which of the following is a true statement?

    A.   The mean is necessarily one of the original values.
    B.   A set of data has only one mean
    C.   The mean is not affected by extreme values.
    D..   In a grouped frequency distribution, the mean cannot be estimated.

29. Which of the following is an advantage of the postal questionnaire?

    A.   Absence of Interviewer Bias
    B.   Time consuming and expensive.
    C.   Unrepresentativeness
    D.   High response rate.

30. The value of the coefficient of correlation lies in the range:
    A. $0 \leq r \leq +1$
    B. $0 < r < +1$
    C. $-1 \leq r \leq +1$
    D. $-1 < r < +1$

31. The study of statistics is usually divided into two categories. Which of the following are the two categories?.

    A. Qualitative and Quantitative

B. Parameters and Statistics
C. Descriptive and Inferential
D. Census and Survey

32. When we are choosing a random sample and we place chosen units back into the population, we are:

A. Sampling without Replacement
B. Sampling with Replacement
C. Using a Systematic Sample
D. Using a Voluntary Response Sample

33. Which of the following is an example of a continuous variable?

A. Number of children in a family.
B. Gender
C. Race
D. Age

34. In finding the standard deviation we must first determine

A. the median
B. the mean
C. kurtosis
D. skewness

35. Ms. Bergen is a loan officer at Coast Bank and Trust. Based on her years of experience, she estimates that the probability is .025 that an applicant will not be able to repay his or her installment loan. Last month she made 40 loans. What is the probability that 3 loans will be defaulted?

A. 0.0000156
B. 0.0613
C. 0.3678
D. 0.075

36. In frequency distribution, the height and sharpness of the peak relative to the rest of the data are measured by a number called

A. kurtosis.
B. coefficient of variation
C. coefficient of skewness
D. coefficient determination

37. Sample size is

    A. the number of items/persons/objects that must be in the sample.

    B. the number of items/persons/objects that must be in the sampling frame

    C. the number of items/persons/objects that must be in the population.

    D. the number of items/persons/objects that must be in the census.

38. The annual dividends, in percent, of four oil shares are: 4.91, 5.75, 8.12, and 21.6. What is the geometric mean dividend?

    A. 10.095

    B. 6.935

    C. 21.60

    D. 8.389

39. In constructing a good questionnaire which of the following must be avoided?

    A.   The questionnaire should be as brief as possible.
    B.   Questions should be short and ambiguous.
    C.   Questions should be in logical sequence
    D.   The questions should be adequate to obtain the information required.

40. Methods of organizing, summarizing and presenting data in an informative way is known as

    A. Descriptive Statistics
    B. Informative Statistics
    C. Narrative Statistics
    D. Inferential Statistics

41. Spearman's rank coefficient of correlation is used when data is in the

    A.   Nominal Scale

B.   Ratio Scale
C.   Ordinal Scale
D.   Interval Scale

42. A portion, or part, of the population of interest is known as

    A. Survey
    B. Statistics
    C. Parameter
    D. Sample

43. A sampling technique where the samples are gathered in a process that does **not** give all the individuals in the population equal chances of being selected is known as

    A.   Random Sampling
    B.   Non Random Sampling
    C.   Probability Sampling
    D.   Fair Sampling

44. Consider the data set 1 3 2 3 5. What is the sample average?

    A. 3
    B. 2
    C. 2.5
    D. 2.8

45. A table has six parts. Which of the following is not a part of a table?

    A. Table Number
    B. Page Number
    C. Table Title
    D. Source Note

46. If we wish to draw a sample of size 100 from a population of 2,000 using the stratified random sampling method, what would be the sampling fraction?

    A. 100
    B. 10
    C. 20
    D. 0.05

47. A sample of 10 observations is selected from a normal population for which the population's standard deviation is known to be 5. The sample mean is known to be 20. The 95% confidence for the population mean is,

A. $20 \pm 1.96 \frac{5}{\sqrt{10}}$

B. $20 \pm 2.58 \frac{5}{\sqrt{10}}$

C. $20 \pm 1.645 \frac{5}{\sqrt{10}}$

D. $20 \pm 0.95 \frac{5}{\sqrt{10}}$.

48. The set of all possible sample points (experimental outcomes) is called,

    A. the sample space.
    B. the sampling frame.
    C. sampling interval.
    D. sample statistics.

49. A population characteristic, such as population mean, is called

    A. a parameter.
    B. a statistic.
    C. descriptive statistic.
    D. Inferential statistic.

50. When the characteristic or variable being studied is numeric it is called:

    A.   Discrete variable
    B.   Continuous variable
    C.   Quantitative variable
    D.   Attribute

51. Of 180 calculators sampled, 5 were defective. The proportion of defects is found by:

    a. $\dfrac{\text{Number of defects in sample}}{\text{Number sampled}}$

    b. $\dfrac{Number\ sampled}{Number\ of\ defects}$

    c. $n = \left(\dfrac{z.s}{E}\right)^2$

    d. $\dfrac{X_1 + X_2}{n_1 + n_2}$

52. Which of the following is **not** a potential source of error and bias in using secondary data:

    A. Sampling error

    B. Response error

    C. Non-response error

    D. Standard error

53. Which of the following would constitute a set of discrete data?

    A. Time taken to travel to work each day over one year.

    B. Weights of a consignment of tins of plum tomatoes.

    C. Number of cars passing a census point each minute over a 3-month period.

    D. Age of applicants applying for catering jobs over a 3-month period in a large hotel chain

54. A card was drawn from a pack of 52 playing cards. What is the probability that it is an ace?

    a. $\dfrac{1}{13}$

    b. $\dfrac{1}{54}$

    c. $\dfrac{4}{13}$

    d. $\dfrac{1}{4}$

55. If the lower and upper quartiles of a frequency distribution were 10 and 15, the quartile deviation of the distribution would be

    A 25

    B 3.5

    C 2.5

    D 12.5

56. Which of the following is not a characteristic of the normal probability distribution and its accompanying normal curve?

    A. The normal curve is bell-shaped

    B. The normal probability distribution is symmetrical about its mean.

    C. The normal curve falls off smoothly in either direcrion from the central value.

    D. The random variable is the result of counting the number of successes in a fixed number of trials.

57. An investment analyst judges that a particular investment has a probability of 0.6 of returning a profit within two years. This is an example of

    A. classical probability

    B. subjective probability

    C. empirical probability

    D. relative frequencies

58. A brokerage survey reports that 30% of individual investors have used a discount broker; that is, one which does not charge the full ommission. In a random sample of 3 individuals, what is the probability: none of them have used a discount broker?

    A. 0.343
    B. 1.0
    C. 0.0
    D. 0.3

59. Two coins are tossed. What is the probability that both will land tail up?

    A. 0.5
    B. 0.25
    C. 1.0
    D. 0.2

60. The Board of Directors of a company consists of 3 accountants, 3 managers, and 2 engineers. A planning committee of 3 is chosen at random from the Board. What is the probability that all 3 members of the sub-committee are accountants?

    A. $\dfrac{3}{8}$

    B. $\dfrac{3}{56}$

    C. $\dfrac{1}{8}$

    D. $\dfrac{1}{56}$

(60 Marks)

Section B.   Answer two questions

1. A cola-dispensing machine is set to dispense on average 7.0 ounces of cola per cup..
    The standard deviation is 0.10 ounces. Calculate the probability that a machine will
    dispense:
    a. Between 7.10 and 7.25 ounces of cola?   (8 Marks)
    b. Between 6.8 and 7.25 ounces of cola?   (7 Marks)
    d. How much cola is dispensed in the largest 1 percent of the drinks?   (5 Marks)

                                                                    (Total 20 Marks)

2. The following sample observations were randomly selected.

X:	5	3	6	3	4	4	6	8
Y:	13	15	7	12	13	11	9	5

    i)   Draw the scatter diagram of the sample observations .   (4 Marks)
    ii) Calculate the coefficient of correlation   (10 Marks)
    iii) Determine the coefficient of determination.   (3 Marks)
    iv) Interpret your result in (ii) and (iii)   (3 Marks)

                                                                    (Total   20 Marks)

3. A sample of 50 antique dealers in Ghana revealed the following sales last year:

Sales ($ thousands)	Number of Firms
100 up to 120	5
120 up to 140	7
140 up to 160	9
160 up to 180	16
180 up to 200	10
200 up to 220	3

    a) Draw a histogram of the distribution.   (3 Marks)
    b) Calculate:
        (i)   the mean;   (4 Marks)
        (ii)   the mode;   (3 Marks)
        (iii)   the standard deviation;   (6 Marks)
        (iv)   the coefficient of skewness;   (2 Marks)

14

of the distribution.

c) Comment on the mean, the standard deviation and the skewness of the distribution.

(2 Marks)
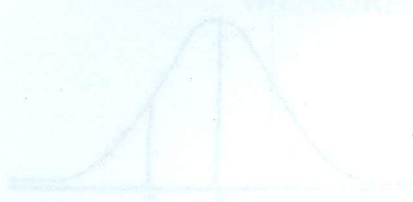
(Total 20 Marks)

4. The Jampkuoz Pizza Chain claims that the mean waiting time of customers for service is normally distributed, with a mean of 3 minutes and a standard deviation of 1 minute. The quality assurance department found in a sample of 50 customers at the Dansoman Branch conclude that the mean waiting time was 2.75 minutes. Can the quality assurance department conclude that the mean waiting time is less than 3 minutes? Conduct the test of the hypothesis at the 0.05 significance level.

(i)	State the null and alternate hypotheses		(4 Marks)
(ii)	State the level of significance		(1 Marks)
(iii)	Compute the test statistic		(8 marks)
(iv)	State the critical value		(3 Marks)
(v)	Make a conclusion		(4 Marks)

(Total 20 Marks)

120

TABLE FOR AREAS UNDER THE STANDARD NORMAL CURVE

z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

MEASURES OF CENTRAL TENDENCY

$$\bar{X} = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$Mode = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) \times c$$

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$GM = \sqrt[n]{(x_1)(x_2)..(x_n)}$$

$$Median = L_1 + \left(\frac{\frac{N}{2} - (\sum f)_1}{f_{median}}\right) \times c$$

Hamonic Mean = $\dfrac{1}{H} = \dfrac{\frac{1}{r1} + \frac{1}{r2} +}{N}$

MEASURES OF DISPERSION

$$Range = X_{max} - X_{min}$$

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

Mean Deviation = $\dfrac{\sum f|x - \bar{x}|}{\sum f}$

$$\sigma = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

$$s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{\sum f}}{n-1}$$

$$c.v = \frac{\sigma}{\bar{x}} \times 100$$

$$Sk = \frac{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^3}{(N-1)s^3}$$

$$Sk. = \frac{mean - mode}{S..D} \qquad\qquad Sk. = \frac{3(mean - median)}{S.D}$$

$$Kurtosis = \frac{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^4}{(N-1)S^4}$$

REGRESSION AND CORRELATION

$$a = \frac{\sum Y}{n} = m, \quad b = slope = \frac{n\left(\sum xy\right) - \left(\sum x \sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$$

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \qquad\qquad r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

PROBABILITY

P (A or B) = P(A) + P(B)

P(A or B or both) = P(A) + P(B) - P(A and B).

P(A and B) = P(A)P(B)

P(A and B) = P(A) P(B|A)

$$P(A_1|B) = \frac{P(A1)P(B|A1)}{P(A1)P(B|A1) + P(A2)P(B|A2)}$$

n! = n(n-1)(n-2) . . .3.2.1

$$^nP_r = \frac{n!}{(n-r)!}$$

$$^nC_r = \frac{^nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

Expected Value

$$\mu = E(x) = \sum[xP(x)],$$

Variance

$$\sigma^2 = \sum[(x - \mu)^2 P(x)]$$

$$\sigma^2 = \sum x^2 P(x) - \mu^2$$

Weighted Indexes

$$P = \frac{P_t}{P_0} \times 100$$

$$P = \frac{\sum P_i}{n}$$

$$P = \frac{\sum P_t}{\sum P_0} \times 100$$

$$P = \frac{\sum P_t q_0}{\sum P_0 q_0} \times 100$$

$$P = \frac{\sum p_t q_t}{\sum p_0 q_t} \times 100$$

Fisher's Ideal Index $= \sqrt{(Laspeyres'\,index)(Paasche's\,index)}$

Binomial Distribution

$$P(x) = \frac{n!}{x!(n-x)!}\pi^x(1-\pi)^{n-x}$$

Poisson Distribution

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

Normal Distribution

$$z = \frac{X - \mu}{\sigma} \qquad \sigma \neq 0$$

Normal Approximation to the Binomial Distribution

$$\mu = np$$

$$\sigma = \sqrt{npq}$$

Continuity Correction

$$x \pm .5$$

$$z = \frac{x + .5 - np}{\sqrt{npq}}$$

Confidence Intervals

$$\bar{x} \pm \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} \pm z \frac{s}{\sqrt{n}}$$

$$p \pm z \sqrt{\frac{p(1-p)}{n}}$$

Test Statistics

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$z = \frac{p - \pi}{\sqrt{(p)(1-p)/\sqrt{n}}}$$

Index Numbers

Simple Indexes