CHAPTER-1

INTRODUCTION TO STATISTICS

1.1 Meaning of statistics:-

Statistics is derived from the Latin word 'status' or the Italian word 'statista' which means 'a political state'

The word statistics is used in two different ways: plural and singular. As a plural noun, it refers to numerical data or quantitative data relating to finished data. We often says that "the statistics of production in agriculture in Maharashtra so and so" such statements gives certain quantitative data about agriculture production in Maharashtra.

As a singular noun, the word refers to statistical methods or methodology of statistics, i.e. collection, presentation, analysis and interpretation of quantitative data.

It is believed that Statistics is in use from the time when man began to count and measure. In ancient days kings used to maintain records of land, agricultural yield, wealth, taxes, livestock, soldiers, weapons, deaths and births etc.

1.1.2 Definitions of statistics:

Croxton and Cowden:-

Statistics or statistical methods are treated as a branch of science which deals with Collection, tabulation, presentations, analysis and interpretation of data.

Sir R.A.Fisher:-

Statistics is a branch of applied Mathematics which specializes in data

Dr. M. H. Lohgaonkar:-

Quantitative things which can understand layman is called statistics.

OR

Pattern of anything which can understand layman is called statistics.

1.2 Importance of Statistics OR Utility of Statistics

- [i] Statistical methods enable to condense the data. It facilitates several functions apart from summarization.
- [ii] Statistical methods give tools of comparison.
- [iii] Estimation, prediction is also possible using statistical tools.
- [iv] We can get idea about the shape, spread, symmetry of the data.
- [v] Inter-relation between two or more variables can be measured using statistical techniques.
- [vi] Statistical methods help in planning, controlling, decision-making.
- [vii] The use of statistical methods is important because considerable amount of timely money, manpower can be saved.
- [viii] Uncertainties can be reduced to get reliable results.
- [ix] Statistical methods give systematic methods of data collection and investigation.
- H.G. Wells expresses the importance and need of statistics "Statistical thinking will one day be necessary for effective citizenship as the ability to read and write"

From the above definitions, we can highlight the major characteristics of statistics as follows:

- [i] Statistics are the aggregates of facts. It means a single figure is not statistics. For example, national income of a country for a single year is not statistics but the same for two or more years is statistics.
- **[ii] Statistics are affected by a number of factors**. For example, sale of a product depends on a number of factors such as its price, quality, competition, the income of the consumers, and so on.
- [iii] **Statistics must be reasonably accurate.** Wrong figures, if analysed, will lead to erroneous conclusions. Hence, it is necessary that conclusions must be based on accurate figures.

- [iv] Statistics must be collected in a systematic manner. If data are collected in a haphazard manner, they will not be reliable and will lead to misleading conclusions.
- [v] Collected in a systematic manner for a pre-determined purpose
- [vi] Lastly, Statistics should be placed in relation to each other. If one collects data unrelated to each other, then such data will be confusing and will not lead to any logical conclusions. Data should be comparable over time and over space.

1.3 Scope of Statistics

The tools and techniques of statistical methods are used in following fields at several phases.

1.3.1 Statistics in industry:-

- [i] Industry makes use of statistics at several places such as administration, planning, production, growth and development.
- [ii] In many industries 'statistical quality control' division is separately operating, mainly, whether manufactured goods possess a desirable standard or not is examined using various control charts. These inspections are done at the time of production.
- [iii] On-line process capability goods or semifinished goods are inspected using acceptance sampling plans of various types, ISO 9000 is used and designs of experiment is also used.
- [iv] Newly installed machinery is tested for its performance using statistical methods.
- [v] Sampling technique is used, Multiple regression planes are used for forecasting, when several factors are interlinked.

1.3.2 Statistics in Biological Sciences:-

[i] Scientist Sir Francis Galton (1822-1911) used Statistics in studying heredity. He pioneered the use of statistical methods in biological sciences. The regression analysis was used by him in the field of genetics.

- [ii] Scientist Mendel performed number of experiments in the field of genetics.
- [iii] Demography uses, statistical methods for forecasting population measuring death rates, birth rates, growth rates etc. Average longevity of human can be estimated using statistical tools.
- [iv] Infant mortality rates can also be measured using statistics. It serves as in indicator of health condition of a country.
- [v] Estimation of sex ration is possible using statistical tools.

1.3.3 Statistics in Agriculture:

- [i] Designs of experiment technique is used to test the interaction effects. For example, the interaction effect of fertilizer and irrigation . it saves amount of time due to statistical methods. Otherwise separate experiments are to be performed to verify the various factors, several times. These experiments help the scientists in developing various hybrid species having specific characteristics such as disease resistance, high yield, requirement of less irrigation facility etc.
- [ii] Various methods of irrigation such as drip irrigation pot, irrigation, sprinkler etc. can be compared using statistical tests.
- [iii] The effect of alkalinity of water, hardness, impurities in the water on growth of plants which receive such a water, can be tested using statistical tests.
- [iv] Estimation of number of trees in a jungle, forest density, number of animals in a jungle, fish in a lake etc. can be done using various statistical techniques.

1.3.4 Statistics in Medical Sciences:-

In the field of medical sciences statistical methods are used to test various claims, such as (i) Does smoking increase the possibility of proneness of cancer? (ii) Is there any correlation between age and blood pressure? (ii) Is a

particular vaccine useful in controlling a particular disease? (iv) Whether growth of a baby is normal? The conjectures(guess) made by experts can be supported by statistical data. Such conjectures can also be tested. Effectiveness of remedial medicines can be tested statistically. Thus statistics helps every respect in this field.

1.3.5 Statistics and Economics:

- [i] In the field of Economics, huge amount of data are needed to be processed and interpreted. Statistics is very much helpful in this field. In order to collect data, various statistical methods of investigations are used.
- [ii] Estimation of national income, per capita income, poverty line, industrial production etc. is done using statistical techniques. Probability distribution of income can be useful in various economic activities.
- [iii] A tool known as index number is used to measures average increase in prices, production income, volume of import, export etc. Index numbers are called as economic barometers. Index numbers are used in determining real income, deflation, cost of living index numbers. To measure the changes in prices of shares in stock market index number provides the best tool.
- [iv] Several interlinked activities in economics can be studied. For example,
- (a) the relation between prices and supply (b) the relation between demand and prices (c) the relation between sales and profit.

Demand analysis, time series analysis techniques are mainly developed to study Economics. They are the gifts given by statistics.

1.3.6 Statistics in Social Science :-

The social sciences we need to test association between two variables such as

- [i] Education and criminality
- [ii] Education and marriage adjustment score
- [ii] Sex and education

[iv] Richness and criminality etc.

1.3.7 Statistics and Management Sciences:-

- [i] For efficient working various sections of management such as sales, production, marketing statistical method are used.
- [ii] Different statistical tools such as forecasting, tests of significance, index numbers, time series analysis, statistical quality control, estimation play vital role in management activities.
- [iii] Apart from this, various optimization techniques known as linear programming, transportation techniques, job assignment problems, sequencing, CPM and PERT, replacement problems, inventory control are also useful.

1.3.8 Statistics in Insurance:

[i] Life table or mortality rates play key role in life insurance policies. In order to decide the premium, insurance company has to use mortality rates which are determined using statistical method. Mortality rates depend upon age, sex, occupation, residential area, heredity.

1.3.9 Statistics in Psychology and Education:-

- [i] In the field of psychology human traits are interrelated. The powerful techniques of measuring such dependence is correlation. We can study relation between productivity and intelligence, productivity and emotional quotient etc. multivariate analysis, correlation and regression are the most applicable tools of statistics which are used in the field of psychology.
- [ii] Analysis of variance is developed to test the most effective methods of teaching and communication. The relation between interrelated variables is studied using correlation, regression. Factor analysis and multivariate techniques are also used to great extent. Forecasting and predictions are required every now and then.

[iii] In order to study equivalence between two examinations for example, SSC board score and CBSE broad score or score of arts and that science students can be compared using percentile ranks.

1.3.10 Statistics in Computer Science:-

[i] Now a days several statistical software packages like MINITAB, MATLAB, STATPACK, SAS, SPSS, SYSSTAT etc. are used for data analysis. Forecasting, prediction, estimation curve fittings etc. are the commonly used statistical techniques. The use of software packages provides the unusual opportunities to get the data summarized in appropriate way. The suitability of modal used for analysis can be quickly determined by means of software, package, otherwise it is a time consuming and tedious procedure. Although software packages are useful to great extent, it cannot replace totally the necessity of statistician. In order to interpret the output or to decide the suitability of statistical model for analysis, to design the questionnaire, to design the experiment of statisticians help is essential.

[ii] Computer is an assembly of several components. The life of each component is a variable having some probability distribution. The average life of each component as well the assembled product can be determined using statistical methods. Reliability of component and system may help the manufacturer to decide the guarantee period of computer as well to user to decide the policy of replacement of spare parts.

1.4 Statistical Organizations in India:-

Statistical organization in operation today is gradually developed over a considerable time period. Mughal emperors, Bristishers set some statistical system in India. After independence, according to Constitution of India, the Central Government modified the systems. The Central Statistical Organization (CSO) associated with Cabinet Secretariat at the centre, looks after various activities such as collection and compilation of data. The

concerned regular publications are taken care of by CSO. It is the main coordinating agency of various statistical organizationks in cetre as well as states of our country. We study some important statistical organizations in India.

1.4.1 Central Statistical Organization (C.S.O.):-

Central Government has established central statistical organization for coordination of statistical organization within the states. This was set in May 1951. CSO works under the Cabinet Secretariat. CSO was established with a view of (i) laying down the standards in the statistical fields (ii) Costatistical activities at the cenre and in the states (iii) ordination maintaining liaison in statistical matters with international agencies (iv) playing an advisory role in statistical matters. The National Income Unit was transferred from Ministry of Finance to the CSO in the year 1954. In 1957 industrial statistics was transferred to CSO. In 1961 Department of statistics was set up in the Cabinet Secretariat and CSO become part of it. Hence the additional duties CSO had stated performing are as follows: (i) preparation of national accounts, (ii) processing and publications of industrial statistics (iii) conduct of economic census and surveys, (iv) maintenance of consumer price index number, (v) organizing various training programmes in official statistics (vi) organizing various conferences on a regular basis.

1.4.2 National Sample Survey Organization (N.S.S.O.).

NSS was set up in **1950** under the guidance of **P.C. Mahalanobis** and it was **reorganized in 1970** under the name **National Sample Survey Organization (NSSO)**. All aspects of survey work were considered together to form NSSO. In functions under the Department of Statistics. It is headed by a Chief Executive Officer. It functions through four divisions for conducting large scale sample surveys (i) Survey Design and Research (ii) Field Operations (iii) Data Processing (iv) Co-ordination and Publications. It

has about 170 offices throughout the country. The activities of NSSO are guided by Governing Council. Reconstitution of Governing Council took place in 1993.

The main functions of the Directorate of NSSO are as follows:

Data collection for Estimation of National Income, for the activities of Planning Commission, for the activities of various Ministries. Collection of Socio economic and demographic data.

Reliability of Surveys:-

Every attempt is made to give correct and most reliable results. It is the primary concern of NSSO. The following are the main three steps taken in this regard.

Internal checks:- Expert statisticians directly collect the primary data and compare the results with those obtained by investigators.

External checks: The results of surveys are compared with those obtained by other entirely independent sources.

Comparison of estimates with the actual values.

1.4.3 Indian Statistical Institute (ISI):-

The Indian Statistical Institute at Calcutta was set up on 28th April 1932 by P.C. Mahalanobis. In the year 1959, the Indian Parliament passed an act and accordingly ISI is treated as an Institute of national importance. Moreover it is authorized to confer degrees and diplomas like any university.

The main functions of ISI are as follows:

- [i] To carry out research qualifying high standards.
- [ii] To provide training and to conduct statistical projects.
- [iii] To provide technical and computational to NSSO.
- [iv] ISI runs B. Stat., M. Stat courses, Ph.D. and research activities.
- [v] ISI is world renowned institute for its high standards research in Statistics and training.

[vi] The 'Sankhya' is regular publication for its high standards research in Statistics and training.

1.4.4 International Institute for Population Sciences (IIPS) :-

The international Institute for Population Sciences (IIPS), formerly known as Demographic Training and Research Centre, was established at Mumbai in July 1956 to serve as the regional centre for training and research in Population Studies for the countries of Asia and Pacific region. It was renamed as International Institute for Population Studies in April 1971 and was redssignated to its present title in March 1984 to facilitate expansion of its academic activities. The institute was declared as Deemed University in August 1985 by the Ministry of Human Resource Development, Government of India, New Delhi. It is an autonomous institution under the administrative control of Department of Family welfare, Ministry of Health and Family Welfare, Government of India. This is the only Institute of Its kind in the world completely devoted to teaching and research in the population related areas.

Over the years, the institute has helped in building a nucleus of professional in the field of population in various countries in the ESCAP region. Many who are trained at the Institute now occupy key positions in the field of population in reputed national and international organizations. During the past forty-five years, the institute has trained 1,956 students i.e. 1,321 from India and 635 form 39 different countries.

The institute offers following regular teaching programmes:

- [i] Diploma in Population Studies (S.P.S.)
- [ii] Master of Pollution Studies (M.P.S) .
- [iii] Master of Philosophy (M.Phil) in Population Studies.
- [iv] Master of Population Studies. (M.P.S) (Correspondence Course)
- [v] Doctor of Philosophy (Ph.D.) in Population Studies.

1.4.5 Statistical Organization in the Maharashtra

Bureau of Economics and Statistics

The present statistical system in India does not specify the jurisdiction of center and state clearly. Hence statistical units exist in Central Ministries as well as in State Government according to the subjects.

Statistical system in states varies from state to state. In Maharashtra State Statistical Bureau is functioning for various activities such as (i) Statistical co-ordination (ii) State Income (iii) Socio economic survey. Apart from this different departments have their own statistical units in respective fields.

The specific functions of Bureau of Economics and Statistics, Bombay are described below:

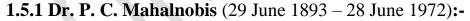
Co-ordinate statistics collected by various departments of State Government.

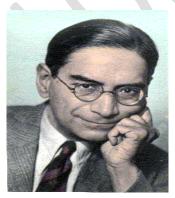
Provide guidance regarding statistics to various departments.

Collect statistical information, conduct statistical enquires and statistical surveys. Provide liaison between state and CSO. Conduct economic and statistical research. Provide statistical assistance to state planning agencies.

Compile economic indicators and give State Income estimates. Publish, Annual State Statistical Abstract and Quarterly District Statistical Abstracts, which include all the essential statistics of the state.

1.5 Statistical Heritage (Indian Perspective):-





Prasanta Chandra Mahalanobis Born in 29 June 1893. Mahalanobis received his early schooling at the Brahmo Boys School in Calcutta,

graduating in 1908. He joined the Presidency College, then affiliated with the University of Calcutta, where he was taught by teachers who included Jagadish Chandra Bose, and Prafulla Chandra Ray. Others attending were Meghnad Saha, a year junior, and Subhas Chandra Bose, two years his junior at college, Mahalanobis received a Bachelor of Science degree with honours in physics in 1912. He left for England in 1913 to join the University of London.

After missing a train, he stayed with a friend at King's college Cambridge. He was impressed by King's College Chapel and his host's friend M. A. Candeth suggested that he could try joining there, which he did. He did well in his studies at King's, but also took an interest in cross-country walking and punting on the river. He interacted with the mathematical genius Shrinivasa Ramanujan during the latter's time at Cambridge. After his Tripos in physics, Mahalanobis worked with C.T.R. at the Cavendish Laboratory. He took a short break and went to India, where he was introduced to the Principal of Presidency College and was invited to take classes in physics.

After returning to England, Mahalanobis was introduced to the journal **Biometrika**. This interested him so much that he bought a complete set and took them to India. He discovered the utility of statistics to problems in meterology and anthology, beginning to work on problems on his journey back to India.

He is best remembered for the **Mahalanobis distance**, a statistical measure, and for being one of the members of the first **Planning Commission of India**. He made pioneering studies in anthropometry in India. He founded the **Indian Statistical Institute** on **28 April 1932**. He contributed to the design of large-scale sample surveys. For his contributions, awarded The **Padma Vibhushan** which is the second-highest civilian award

of the Republic of India. Mahalanobis has been considered the father of modern statistics in India.

1.5.2 Dr. V. S. Huzurbazar (15 September 1919 – 15 November 1991):-



Prof.V.S.Huzurbazar was born on 15th September,1919 at Kolhapur in Maharashtra State, INDIA. As a student he had a brilliant career in University of Bombay, Banaras Hindu University and University of Cambridge, England where he obtained his Ph.D. degree for his thesis entitled "Properties of Sufficient Statistics" written under the supervision of Sir Harold Jeffreys.

Professor Huzurbazar is internationally well-known for his work on maximum likelihood estimation, invariants for probability distributions and sufficient statistics.

During his long career, Professor Huzurbazar has received many recognitions and Honours. He played an active role in the development of Statistics and Mathematics in India. The Government of India recognized his contributions to education by awarding him the "Padma Bhushan" in 1974.

After his return from Cambridge in 1949, Dr. Huzurbazar worked in the University of Gauhati, Lucknow and also in the Bureau of Economics and Statistics Of Government of Bombay. He Joined the University of Pune as Professor and Head of the Department of Mathematics and Statistics in 1953

1.5.3 Dr. P. V. Sukhatme (27 July, 1911 – 28 January, 1997):-



Pandurang Vasudeo Sukhatme was an award-winning Indian statistician. He is known for his pioneering work of applying random sampling methods in agricultural statistics and in biometry, in the 1940s. He was also influential in the establishment of the Indian Agricultural Statistics Research Institute. As a part of his work at the Food and Agriculture Organization in Rome, he developed statistical models for assessing the dimensions of hunger and future food supplies for the world. He also developed methods for measuring the size and nature of the protein gap.

His other major contributions included applying statistical techniques to the study of <u>human nutrition</u>. One of his ideas, the Sukhatme–Margen hypothesis, suggested that at low calorie intake levels, stored energy in the body is used with greater metabolic efficiency and that the metabolic efficiency decreases as the intake increases above the homeostatic range. This involved paying attention to intra-individual variability that was found to be more than the inter-individual variability in protein or calorie intake. He gave a genetic interpretation of the intra-individual variation jointly with P. Narain.

He was awarded the **Padma Bhushan** by the <u>Government of India</u> in 1971

1.5.4 Dr. C. R. Rao (Born 10 September 1920):-

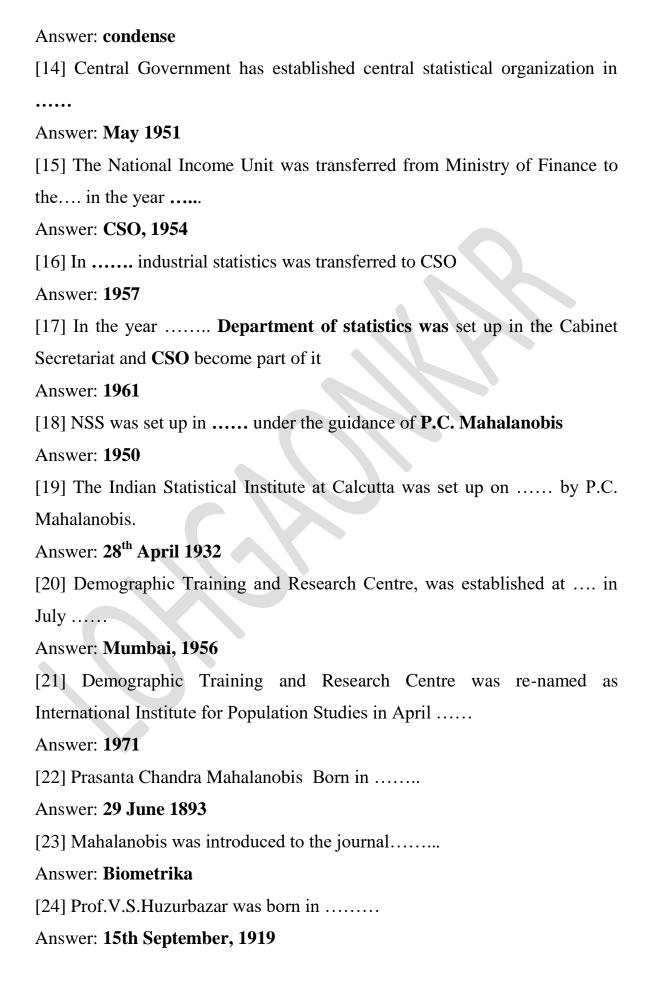


Calyampudi Radhakrishna Rao, Currently serving as the Professor Emeritus at Penn State University &Research Professor at the University at Buffalo, he has various honours and accolades under his belt. Being described as a living legend by the American Statistical Association, his work has influenced not just statistics but various other fields like economics, genetics, demography, biometry, medicine and more.

After having completed M.A in mathematics at the Andhra University, he did his M.A in statistics from Calcutta University. He worked at the Indian Statistical Institute for 40 years after which he moved to USA to take teaching at University of Pittsburgh and Pennsylvania State University.

Having worked in the area of multivariate analysis, combinatorial design, statistical genetics, statistical inference and linear models, he is a big inspiration for those planning to dive into the field of statistics. Among his best-known discoveries are the Cramer-Rao bound and the Rao-Blackwell theorem both related to the quality of estimators. For his contributions, awarded The **Padma Vibhushan**

Objective type questions:
(A) Fill in the blanks
[1] Statistics is both, aand an
Answer: Science, Art
[2] Statistics can prove
Answer: Anything
[3] Statistics deals with only
Answer: Quantitative data
[4] Statistical analysis helps in theof results.
Answer: Interpretation
[5] Statistics is not applicable toobservation
Answer: Single
[6] Statistics are numericalof facts, but all numerical statements are
not
Answer: Statements, Statistics
[7] Statistics does not study
Answer: Individuals
[8] Statistics is the arithmetic of human
Answer: Welfare
[9] Statistics is liable to be
Answer: misused
[10] Statistics is derived from the Latin word
Answer: 'status'
[11] Statistics is derived from the Italian word
Answer: 'statista'
[12] The word statistics is used in two different waysand
Answer: plural and singular
[13] Statistical methods enable tothe data



[25] Pandurang Vasudeo Sukhatme was b	oorn in							
Answer: 27 July, 1911								
[26] Pandurang Vasudeo Sukhatme was	[26] Pandurang Vasudeo Sukhatme was awarded the Pdma Bhushan by the							
Government of India in								
Answer: 1971								
(B) Multiple choice questions: Choose	e the correct alternative from the							
following:								
[1] Statistics is not applicable toobservation								
(a) Single	(b) Double							
(c) More than two	(d) NOTA							
Answer: Single								
[2] Statistics is derived from the Latin word								
(a) Status	(b) Statista							
(c) Statistik	(d) NOTA							
Answer: 'status'								
[3] Statistics is derived from the Italian w	vord							
(a) Status (b) Statista								
(c) Statistik	(d) NOTA							
Answer: 'statista'								
[4] The word statistics is used in two diff	erent ways							
(a) Plural	(b) Singular							
(c) Plural and Singular	(d) NOTA							
Answer: plural and singular								
[5] Statistical methods enable tothe	data							
(a) Condense	(b) Expand							
(b) Both (a) & (b)	(c) NOTA							
Answer: condense								
[6] Central Government has established	d central statistical organization in							
•••••								

	(a) May 1951	(b) April 1951					
	(c) May 1952	(d) April 1952					
	Answer: May 1951						
	[7] The National Income Unit was transferred from Ministry of Finance to the						
	CSO in the year						
	(a) 1951	(b) 1952					
	(c) 1953	(d) 1954					
	Answer: 1954						
	[8] In industrial statistics was transferred to CSO						
	(a) 1955	(b) 1956					
	(c) 1957	(d) 1954					
	Answer: 1957						
[9] In the year Department of statistics was set up in the Cabinet							
Secretariat and CSO become part of it							
	(a) 1961 (b) 1952						
	(c) 1963	(d) 1954					
	Answer: 1961						
[10] NSS was set up in under the guidance of P.C. Mahalanobis							
	(a) 1951	(b) 1950					
	(c) 1953	(d) 1954					
	Answer: 1950						
	[11] The Indian Statistical Institut	te at Calcutta was set up on by P.C.					
	Mahalanobis.						
	(a) 28 th April 1951	(b) 28 th April 1952					
	(c) 28 th April 1932	(d) 28 th May 1954					
	Answer: 28 th April 1932						
	[12] Demographic Training and Research Centre, was established at Mumbai						
	in July						
(a) 1951 (b) 1952							

(c) 1955	(d) 1956							
Answer: 1956								
[13] Demographic Training	and	Research	Centre	was	re-named	as		
International Institute for Population Studies in April								
(a) 1971	(b) 1952							
(c) 1973	(d) 1961							
Answer: 1971								
[14] Prasanta Chandra Mahalanobis Born in								
(a) 28 th April 1851	h April 1851 (b) 28 th April 1881							
(c) 29 th June 1893	(d) 29 th June 1892							
Answer: 29 June 1893								
[15]founded the India	ın Sta	tistical Inst	itute on 2	28 Ap	ril, 1932			
(a) Prof. P.C. Mahalano	(a) Prof. P.C. Mahalanobis (b) Prof. V.S. Huzurbazar							
(c) Prof. P. V. Sukhatme	V. Sukhatme (d) NOTA							
Answer: P.C. Mahalanobis								
[16] Prof.V.S.Huzurbazar was	born i	n						
(a) 28 th April 1893			(b) 28 th A	April	1881			
(c) 15 th September 1919			(d) 29 th J	June 1	892			
Answer: 15th September, 191	Answer: 15th September, 1919							
[17] Pandurang Vasudeo Sukhatme was born in								
(a) 28 th April 1893								
(c) 15 th September 1919			(d) 27 th J	July 19	911			
Answer: 27 July, 1911								

EXERCISE:

- [1] Give definition of statistics and explain its meaning.
- [2] Explain scope of statistics in Management science.
- [3] Explain scope of statistics in Industry.
- [4] Explain scope of statistics in Economics.

- [5] Explain the concept of population and sample.
- [6] Explain SRSWR method.
- [7] Define the term statistics. Explain its scope of statistics in management sciences.
- [8] Define the term statistics. Explain its scope in Economics.
- [9] Explain in details utility of Statistics.
- [10] Write a short note on
- (i) Dr. P. C. Mahalnobis
- (ii) Dr. V. S. Huzurbazar
- (iii) Dr. P. V. Sukhatme
- (ii) Dr. C. R. Rao