Emerging trends in social science research

Presentation at the Interactive Session with UG and PG students held at PES University, Bangalore South Campus, Electronic City, Bangalore 560100, India, on December 21, 2015

Stephen ARO-GORDON, Ph.D.
Baze University Abuja, Nigeria
Department of Financial Mathematics
Faculty of Computing and Applied Sciences
Email: stephen.aro-gordon@bazeuniversity.edu.ng /; getyouthsworking@yahoo.com

BAZE UNIVERSITY ABUJA

December 21, 2015 @ PES University, Bangalore, India

Objectives…

BAZE UNIVERSITY ABUJA

1. To revisit some general areas of concerns in contemporary social science research.
2. To review some of the emerging shifts in social science research especially in emerging market economies including some of the noted implications for the benefit of future researchers in the field.

Resources…

BAZE UNIVERSITY ABUJA

1. *Research Methodology: Methods and Techniques. Third Edition.* By Kothari, C. R., & Garg, G (2014). New Delhi: New Age International (P) Limited, Publishers.
2. *Research Methodology in the Social and Management Sciences* by U. G. Moti, Ed. (2005). Kaduna: Joyce Graphic Printers & Publishers Company.
3. *Discovering Facts Using SPSS* by Andy Field
4. *Building Financial Models with Microsoft Excel: A Guide for Business Professionals,* 2nd edn by S. Proctor (2009). Hoboken, NJ: Wiley.
5. *Introduction to Linear Regression Analysis*, 3rd edition, by Montgomery, Peck, & Vining (2003). Wiley India.
6. *Introductory Econometrics for Finance, 2nd edn.* by Chris Brooks (2008). Cambridge University Press.

Context…

BAZE UNIVERSITY ABUJA

What really is ‘science’?

Understanding what 'science' is all about….(Encarta Dictionary, 2009)

BAZE UNIVERSITY ABUJA

1. ***Study of physical world*:** The study of the physical and natural world and phenomena, especially by using systematic observation and experiment. ***Branch of science*:** a particular area of study or knowledge of the physical world.
2. ***Systematic body of knowledge*:** A systematically organized body of knowledge about a particular subject. ***Something studied or performed methodically*:** An activity that is the object of careful study or that is carried out according to a developed method.

Science as a body of knowledge . . .

BAZE UNIVERSITY ABUJA

1. *Social sciences*

2. Physical sciences or natural sciences

3. Life sciences

4. Management sciences, etc.

Exploits of physical science research on CNN this morning! (21/12/2015)

BAZE UNIVERSITY ABUJA

December 21, 2015 @ PES University, Bangalore, India

Social science: Science of society…

BAZE UNIVERSITY ABUJA

- ***Social science*** is a key discipline dealing with aspects of the society – relationships among groups of people within a community.
- Established thinkers, notably, Karl Marx, Weber, Georg Simmel, and Emile Durkheim
- Contemporary influencers – Karl Popper and Talcott Parsons

- ***Typically*** – Economics, Political Science, Public Administration, International Relations, Geography, Demography, Sociology,
- ***In wider context*** – Anthropology, Archeology, Psychology, History, Law, Criminology, Linguistics, Education, Communication Arts, Business / Management Studies, Development Studies, Environmental Studies, Library Science, Marketing
- ***Boundary is blurring*** – Sociobiology, Neuropsychology, Behavioural Finance, Bioeconomics, Social Medicine

What 'research' involves…

BAZE UNIVERSITY ABUJA

- ***Research*** = *"to go about seeking*" or 'search' (French "*recherché*")
- Research is basically a creative work conducted in a systematic manner to add to stock of knowledge and help in devising new applications that would make the society a better place.
- Research has also been defined variously as:
- ✓ Data / information / facts collection for advancement of knowledge
- ✓ A process…
- ✓ A studious inquiry or examination…
- ***Empirical research*** is not derived from application of logic; rather, it is based on:
- ✓ Observation (sensory)
- ✓ Experiment
- ✓ Practical experiment

Revisiting the purpose of research - The BIG PICTURE

BAZE UNIVERSITY ABUJA

- Well-being of the society…
- The current massive data availability – the Internet age / smart phones
- Rapid and monumental changes and implications on the future of the professions – e.g. technology ‘disruptions’ rendering traditional approaches antiquated, opaque, and unaffordable (Susskind &Susskind, 2015)

The question of 'methodology' and 'methods'

BAZE UNIVERSITY ABUJA

?

The recurring question of 'methodology' and 'methods'

BAZE UNIVERSITY ABUJA

- ***Methodology*** is the general research strategy – outlining the way in which you want to undertake the research, the methods you will use, + the theoretical justification for using the methods
- 'Methodology' is <u>not</u> synonym for method…
- ***Methods*** captures the means or modes for collecting data (publication research, interviews, surveys, etc.) and in some cases, how you want to analyze, compute or infer from the result.

Research methods

BAZE UNIVERSITY ABUJA

Exploratory research

-Helping to identify and define research problems

Constructive research

- Testing theories and proposing solutions

Empirical Research

-Testing the feasibility of solution using real-world data

Two major types of empirical social research design (Gorard, 2013)

BAZE UNIVERSITY ABUJA

Qualitative research

-Understanding human behaviour

Philosophy: Social constructionism

Quantitative research

- Uses numerical data and statistical methods

Philosophy: Positivism

Formal social research process…

BAZE UNIVERSITY ABUJA

1. Identification of research problem – the research gap that you intend to narrow
2. Literature review
3. Specify purpose of your research
4. Determine the specific research questions
5. Formulate your hypotheses – a testable prediction which designates the relationship between two or more variables.

6. Select a suitable methodology
7. Data collection
8. Analyze and interpret the data

*9. Report and assess the research results (was the Null hypothesis validated or rejected?)

*10. Communicate the research findings, recommendations and *scope for further research*.

Wider variety of data sets…

BAZE UNIVERSITY ABUJA

- Fundamental economics
- Real estate
- Human resources management
- Accounting
- Advertising
- Agriculture
- Banking & Finance
- Business

- Finance & investments
- Marketing
- Opinion polls
- Transportation
- Sports
- Life sciences
- Physics and engineering, etc.

BAZE UNIVERSITY ABUJA

Errors in data analytics

Areas to watch out for

December 21, 2015 @ PES University, Bangalore, India

Statistics revisited… means or end?

BAZE UNIVERSITY ABUJA

- Statistics is the scientific method of collection, classification, representation, analysis and interpretation of numerical data with a view to making meaningful inferences on the objects on which the data were collected.
- Statistics answers questions using data – not numbers only; we also use pictures, graphs, tables, etc.
- Statistical models are used to simplify reality and help us to answer questions.

- Capacity to make desirable, informed, result-oriented decision.
- one thing is clear: we use statistical processes to serve as guide towards making qualitative decisions.

Are 'numbers' the end?

BAZE UNIVERSITY ABUJA

- "Far too many reviews are dominated by dry discussions of numbers… The review should be a creative exercise, not a drill where people regurgitate data."

- L. Bossidy & R. Charan (2002) in ***Execution: The discipline of getting things done.*** **New York: Crown Business, p. 208**

Increased awareness of statistical errors Continuation

BAZE UNIVERSITY ABUJA

- ***Wrong methodology*** – A good number of unintentional mistakes come from using the wrong method to address research questions, thereby leading to misinterpretation of results.
  - Choose and use correct statistical methods for every problem

- ***Data organization***: Avoid error here by clearly defining your variables (observations) before you record your data – what does *X1, X2, X3*, … mean?
  - The name of a variable should describe its attributes.

Why exercising good judgement is imperative…

BAZE UNIVERSITY ABUJA

- Analytical tools are as important to the modern executive as pliers and screwdrivers are necessary to the auto mechanic.
- Like a mechanic, the analyst must know his business well enough to choose the proper statistical tool to solve the problem at hand.

Resolving several social questions in organizational settings …

BAZE UNIVERSITY ABUJA

- What are your competitors doing to increase market share?
- What are your competitors doing differently in the area of customer service?
- How are your best-in-class competitors handling cost, quality, technology and HR?
- Does your organization have the cost structure (or capital structure) that will allow it to compete profitably?
- Assume your organization needs a new organogram; what new sales management skills will be required?

How sensible is this?
One research approach to every problem?

BAZE UNIVERSITY ABUJA

"I suppose it is tempting, if the only tool you have is a hammer, to treat every problem as if it were a nail."

- Abraham Maslow

Minimizing statistical errors
Further tips…

BAZE UNIVERSITY ABUJA

- Understand the underlying business theory / issues / questions first before performing statistical analysis – this is pivotal to drafting meaningful questionnaires.
- The central point of statistics is problem-solving – how are your analyses helping your organization or country to make better decisions or policies?
- Don't carelessly round up data – check to ensure that your data add up – data credibility

- The need to produce better decisions and insights from the massive data amount generated in today's world of business and science.
- Technology – Computers now perform most of the calculations that once dominated statistics and related courses
- Use Statistical packages / software carefully – interpretation of results must make sense to you first before it can make sense to your audience

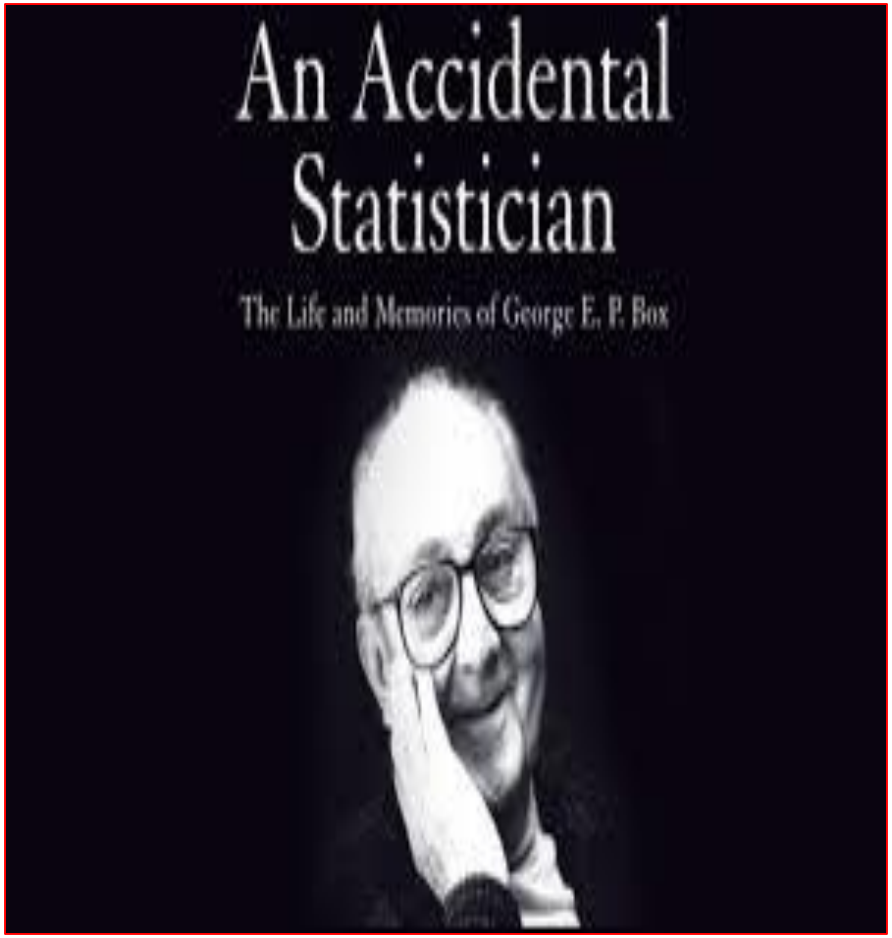
Minimizing statistical errors

Emerging concerns…

BAZE UNIVERSITY ABUJA

"All models are wrong, but some are useful."

- George E. P. Best



Data normalization and standardization: now a major issue…

BAZE UNIVERSITY ABUJA

- Basically, to normalize data, traditionally this means to fit the data within unity (1), so all data values will take on a value of 0 to 1 (***Ben Etzkom, 2012***). Where applicable, data should be normalized or standardized to bring all of the variables into proportion with one another.
- This is important so that the coefficients associated with each variable will scale appropriately to <u>adjust for the disparity in the variable sizes,</u> thereby reflecting meaningful relative activity between each variable, i.e., a positive coefficient will mean that the variable acts positively towards the objective function, and vice versa.

On data normalization and standardization

Emergence of Post-Modern Portfolio Theory (PMPT)

BAZE UNIVERSITY ABUJA

- Associated with software entrepreneurs Brian M. Rom & Kathleen Ferguson, PMPT is an expanded risk-return paradigm designed to address the major practical limitations of CAPM/MPT – the assumption that of a discrete, normal (mean-variance) distribution that may not accurately reflect investment reality.
- Thus, the lognormal distribution was introduced as a more robust model for the pattern of investment returns.

Avoiding data confusion / mishandling…

BAZE UNIVERSITY ABUJA

- ***Qualitative data*** – Categorical - Discrete - Nominal – Ordinal (Likert scale) – we cannot easily measure or count; e.g. gender, behaviour, quality…
- Performing purely quantitative techniques such multiplication and division on categorical data will yield meaningless results.
- Don't put ordinal data in a pie chart!
- Don't carelessly round off data particularly in pie charts.
- Be careful with elaborate graphs
- Clarity in knowledge of Mean, Mode, & Median…

- ***Quantitative data*** – data that we can easily measure and count; e.g. age, weight, height, sales, production output, prices…Numerical – Continuous – Interval – Ratio
- ***Time series (trend analysis)*** – changing values of a variable over time / at different times.
- ***Cross-sectional data*** – data that measure attributes of different objects at the same time – one-shot data.
- ***Panel Data*** – Data collected on various objects (individuals, countries, etc.) for sequential periods – a combination of time-series and cross-sectional data.

Self-study…

BAZE UNIVERSITY ABUJA

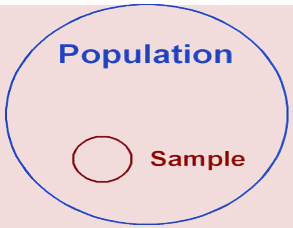
Parametric

versus

non-parametric

- In what situations should you use chi-square?

Ensuring robust sampling



BAZE UNIVERSITY ABUJA

- The essence of sampling is to obtain maximum, accurate, and reliable information about the universe with the minimum sacrifice of money, time, and energy.
- Statistical regularity - randomness
- Inertia of large numbers – more inert (constant) than small ones
- Persistence – same attributes as the universe
- Optimization – cost-effective & efficiency
- Validity – selected at random, scientifically done

On Sample Size…

BAZE UNIVERSITY ABUJA

- The issue is about the accuracy of your data, perhaps, not necessarily your sample size.
- The point is: How representative is your sample? - (the young girl thought it was ice cream, but a 'little' taste was enough to disappoint her!!)
- It is better to get a representative sample than a very large sample [See Vijayalaskshmi & Sivapragasam (2008) on *Research Methods: Tips and Techniques*]

Some tips on sample size…

(*V*ijayalaskshmi & Sivapragasam, 2008)

BAZE UNIVERSITY ABUJA

- The more varied the data, the smaller the sample size needed to attain the same level of accuracy – if we have many variables in the study, the sample size should be relatively small.
- On the other hand, if we have only a few variables, we should have a larger sample.
- In cross tabulations, the desirable sample size depends on the number of intended cells – a rough guide is to have at least 20-30 study units per cell.
- In the final analysis, the eventual sample size depends on a compromise between what is desirable and what is feasible.

Dealing with sampling errors…

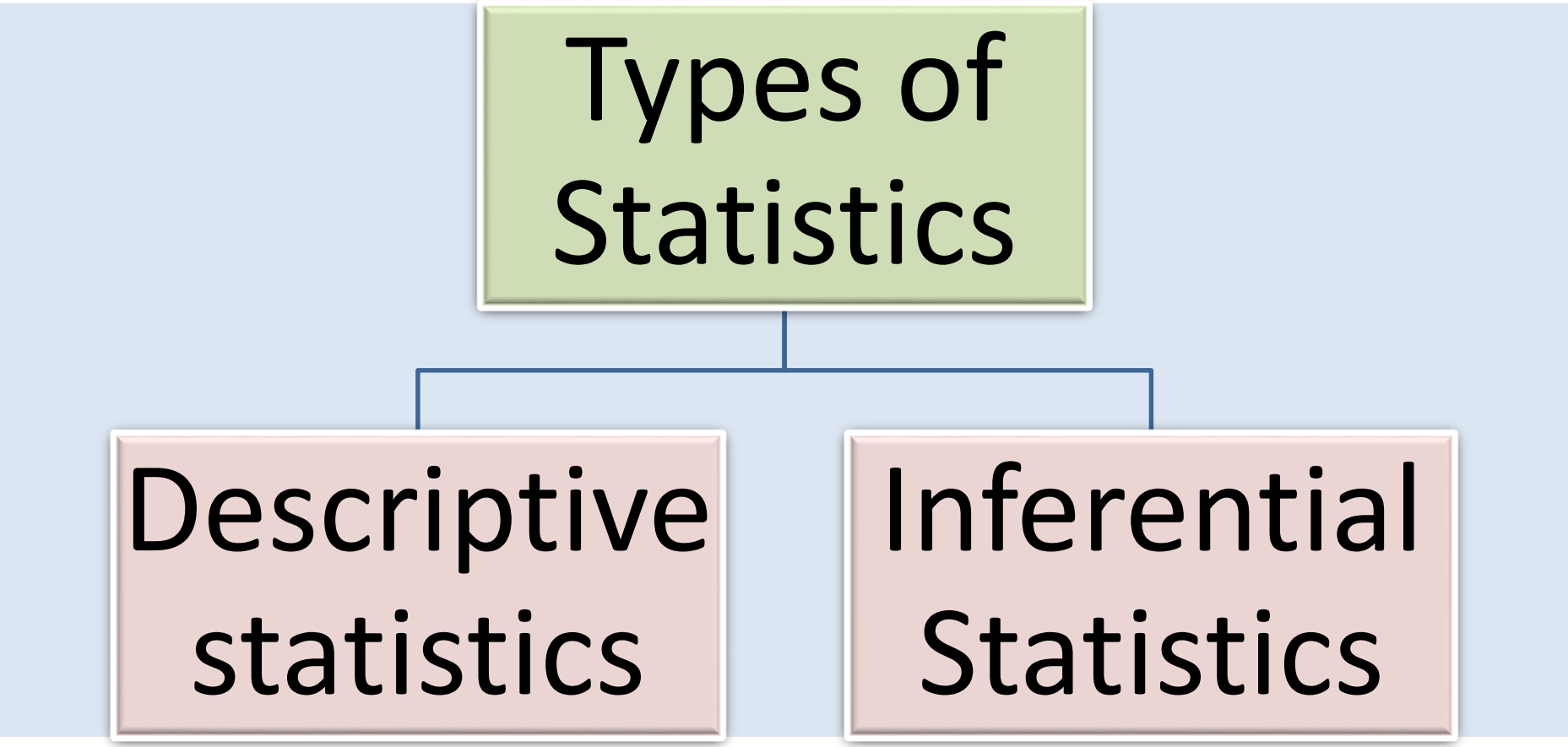
BAZE UNIVERSITY ABUJA

- ***Biased errors*** – prejudice of the analyst / investigator
- ***Note:*** Increasing sample size will not cure biased errors.

- ***Unbiased errors*** – accidental or arising in the course of events or survey
- Also called random sampling error – only a part of the universe is ultimately observed.
- ***Solution:*** Ensure that the sample size is reasonably large to neutralize this type of error

Clarity of statistical typology

BAZE UNIVERSITY ABUJA



Inferential statistics…

BAZE UNIVERSITY ABUJA

- What is ‘inferential’?

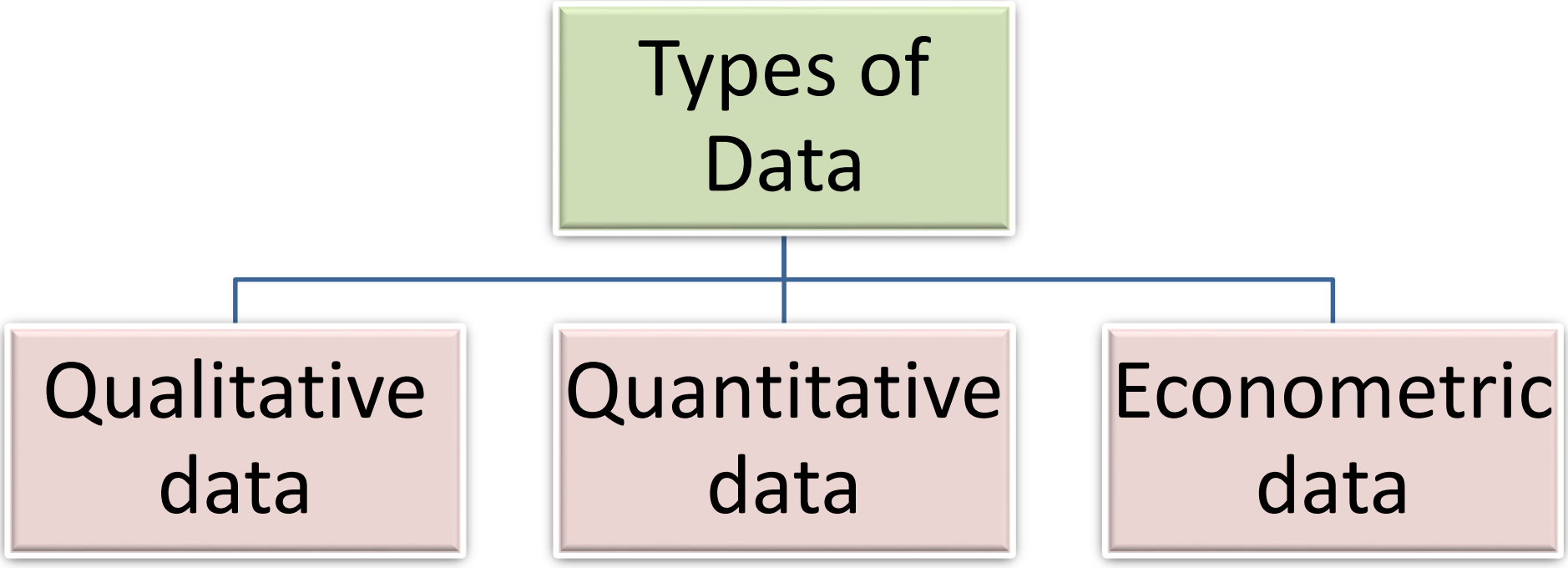
What is 'inferential statistics'?

BAZE UNIVERSITY ABUJA

- ***Inferential statistics*** – statistical approach that helps you to draw clear or evident implications from analytical results

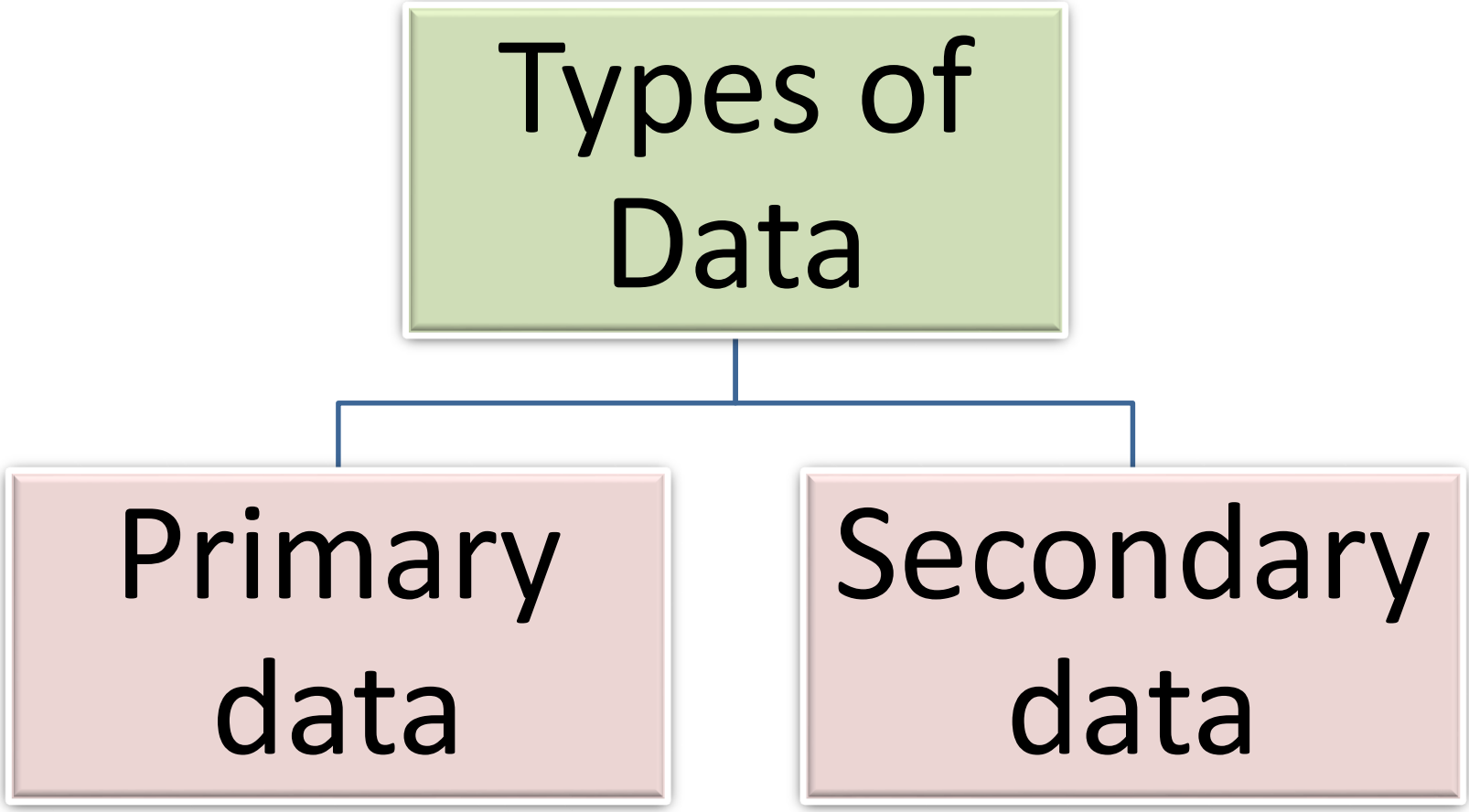
Understanding data typology…

BAZE UNIVERSITY ABUJA



Still on data typology…

BAZE UNIVERSITY ABUJA



Choosing between primary and secondary data… things to consider:

BAZE UNIVERSITY ABUJA

- Nature and scope of your research – more data could be needed to reinforce or modify current knowledge.
- Availability of resources – quality primary survey could take a lot of time and finance.
- You can never have enough of data any way – secondary data can provide supplementary evidence.
- How credible and reliable is the source – data-collection agency
- Why insist on primary data when the data needed already exist

Era of massive data…

BAZE UNIVERSITY ABUJA

S/NO	SOURCES	Examples of data
1	ARCHIVES	General Information
2	PRINT MEDIA (Newspapers, daily, weekly or monthly magazine)	Election result, Academic Result, admission list of Schools, and other news items
3	ELECTRONIC MEDIA	Census result, election result, Radio, television, internet/websites, News items and general information.
4	MUSEUM (a cultural archive)	Information on antiquities, cultural information.
5	NATIONAL BUREAU OF STATISTICS PLANNING COUNCILS, CENTRAL BANKS, WORLD BANK, IMF, ETC.	Wages and salaries of workers, facts on national economy.

Secondary data:

Have they passed the test of scrutiny?

BAZE UNIVERSITY ABUJA

- A major feature of secondary data is that it may be fraught with bias, and may not be so accurate or adequate, depending on the purpose of the investigation.
- Consequently, since the researcher is typically not the original compiler of the data, there is a need for proper editing and scrutiny of the secondary data in order to make it appropriate for usage.
- To do this, answers must be provided to the pertinent questions listed here – (right side of this slide).
- When answers are satisfactorily provided for each of these afore mentioned questions satisfactorily; the secondary data are said to have been *scrutinized* and edited, thus made reliable for use by researchers.

1. What is the type and purpose of the institution which the data emanated?
2. Are the data accurate and adequate?
3. Are the data biased?
4. In what types of units are the data expressed?
5. Are the data related to the problem under study?

'Data' vs. 'information' – any real difference?

BAZE UNIVERSITY ABUJA

- Not all pieces of information gathered about a particular object can be referred to as statistical data. Why?
- *****<u>**Data is factual information**</u> often in the form of ***facts or figures obtained from experiments or surveys***, and used as a basis for making calculations or drawing conclusions. (Encarta Dictionary, 2009)

Revisiting…
Methods of data collection

BAZE UNIVERSITY ABUJA

?

On methods of data collection… integrity & finance issues…

BAZE UNIVERSITY ABUJA

1. Direct Personal Observation - FGDs
2. Personal interviewing – directly or indirectly (due to language barrier) - robust dialogue among business leaders, executives, policy makers, customers, etc.

3. Data collected through experimentation…

4. Data collected through questionnaires… reliability issue – how truthful are the respondents – how truthful are the enumerators, etc.

Data collection:

Problems and type of errors that may occur …

- **BIAS:** The enumerator and respondents may colour their information with ethnic or tribal sentiments.
- **SAMPLING:** Sample may be an incorrect representation of the population.
- **ENUMERATOR:** The questionnaire may not be properly administered, thus required data may not be collected as it ought to be

- Deficiency of Questionnaires
- Reluctance on the part of informant

Variables – revisited …

BAZE UNIVERSITY ABUJA

Types of Variable

- Discrete variables
- Continuous variables

On discrete & continuous variables

BAZE UNIVERSITY ABUJA

❑ *Discrete Variable*

- These are quantitative properties of objects that can be expressed in terms of whole numbers only.
- For example, the number of a particular book sold by a bookseller can only be counted whole numbers: 10,12, 14 etc.

❑ *Continuous Variable*

- These are quantitative properties of objects that are measured on a continuous scale rather than counting.
- For example, the heights of school pupils are measured in feet to the nearest centimeters 1.89 metres

Why you need to classify data properly

BAZE UNIVERSITY ABUJA

- To ensure that data so collected by any of the data collection methods be well-organized thereafter.
- To improve our understanding of problems and helps us to take wise decisions.
- To eliminate unnecessary details.
- To bring out clearly points of similarity and dissimilarity.
- To enable us to form clear mental picture of objects.
- To enable us to make comparison and draw meaningful inferences.

How to make meaningful data presentation…

BAZE UNIVERSITY ABUJA

?

December 21, 2015 @ PES University, Bangalore, India

Presenting statistical data

BAZE UNIVERSITY ABUJA

- The simplest and perhaps the most important way of presenting data is by means of diagrammatic form, design, or chart, not only to catch the eyes, but also to convey the information easily.

- ✓ Pie chart
- ✓ Bar chart
  - ❖ Simple bar chart
  - ❖ Component bar chart
  - ❖ Multiple bar chart
- ✓ Histogram
- ✓ Frequency polygon

Handling 'normal distribution'…

BAZE UNIVERSITY ABUJA

?

Handling correlation analysis

BAZE UNIVERSITY ABUJA

?

Beta - a key application area in social sciences, particularly the business world

BAZE UNIVERSITY ABUJA

- *Betas* vary significantly across industries.
- Hence there are differences in business risks across industries (Chandra, 2008)

Interpretation of correlation coefficient: Summary

BAZE UNIVERSITY ABUJA

S/No	***r***	***Interpretation***
1	= 1	Perfect positive relationship
2	= -1	Perfect negative relationship
3	= 0	No relationship
4	= 0.1 to 0.4	Positive and weak relationship
5	= 0.5 to 0.9	Positive and strong relationship
6	= -0.4 to -0.1	Negative and weak relationship
7	= -0.9 to 0.5	Negative and strong relationship

Dealing with regression analytics…

BAZE UNIVERSITY ABUJA

?

Note the essence of regression analytics (Montgomery et al, 2003)…

BAZE UNIVERSITY ABUJA

- Modeling a complex system
- Data description
- Prediction & estimation
- Parameter estimation & control
- Model validation

○ Note: There is no 'best' regression estimation (Montgomery *et al,* 2003, p. 291)

Recap:
Interpretation of goodness of goodness of fit

BAZE UNIVERSITY ABUJA

S/No	*R^2*	*Interpretation*
1	= 1	Perfect fit
2	= 0	No fit
3	= 0.1 to 0.4	Poor fit
4	= 0.5 to 0.9	Good fit

Recall interpretation of correlation coefficient:

BAZE UNIVERSITY ABUJA

S/No	*r*	*Interpretation*
1	= 1	Perfect positive relationship
2	= -1	Perfect negative relationship
3	= 0	No relationship
4	= 0.1 to 0.4	Positive and weak relationship
5	**= 0.5 to 0.9**	**Positive and strong relationship**
6	= -0.4 to -0.1	Negative and weak relationship
7	= -0.9 to 0.5	Negative and strong relationship

Recall:
Interpretation of 'goodness of fit':

BAZE UNIVERSITY ABUJA

S/No	*R^2*	*Interpretation*
1	= 1	Perfect fit
2	= 0	No fit
3	= 0.1 to 0.4	Poor fit
4	*= 0.5 to 0.9*	*Good fit*

Revisiting the R-squared metric…

BAZE UNIVERSITY ABUJA

- R-Squared is seen as a non-decreasing measure meaning that even as more explanatory variables enter the model, the R-squared value will keep increasing thereby failing to account for the loss of degree of freedom arising from additional independent variables in the model.
- This deficiency is addressed by ***adjusted R-squared*** which gives a truer picture of the goodness of fit of a statistical model.
- R^2 or adjusted $R^2 > 0.5$ is thought to imply that the regression model has a good fit and it is therefore okay for forecasting – ***Really?!!***

How reliable are social science research results these days?

‘Hypothesis testing’ revisited

BAZE UNIVERSITY ABUJA

December 21, 2015 @ PES University, Bangalore, India

Revisiting…
What "statistical significance" really means

BAZE UNIVERSITY ABUJA

- When a statistic is ‘significant’, what this really means in statistics is that you are very sure that the statistic is **reliable**, not that the finding is important or that it has any decision-making utility to your client or audience.
- Significance is a statistical term that indicates how sure or confident you are that a difference or relationship exists, and how strong or weak it is.
- Significant differences can be large or small depending on your sample size.
- In effect, what we are talking about is not the ordinary sense of significance but ‘statistical significance’.

What 'hypothesis' is all about…

BAZE UNIVERSITY ABUJA

- Recall: ***Hypothesis*** is simply a theory needing investigation: a tentative explanation for a phenomenon, used as a basis for further investigation
- Hypothesis is therefore an assumption or a statement that is assumed to be true for the sake of argument.
- As analysts, we usually want to start our investigations with basic assumptions or *intelligent* guess concerning the behaviour of variables or issue under study.

That market 'assumption' – test it before launching that product!

BAZE UNIVERSITY ABUJA

- ***Hypothesis testing*** – testing the validity of your results / claims, scientific verification or validation to confirm whether the assumptions made about the subject variables are true or false.
- Types of hypothesis testing - what do you understand by
  - One-tailed test?
  - Two-tailed test?

On whether to use a one-tailed or two-tailed test of significance?

BAZE UNIVERSITY ABUJA

?

Whether to use a one-tailed or two-tailed test of significance?

BAZE UNIVERSITY ABUJA

- The answer largely depends on your hypothesis.
- It is generally safest to use two-tailed tests

[https://www.statpac.com/surveys/statistical-significance.htm – 29/11/2015]

Hypothesis Testing (HT): Revisiting its essence…

BAZE UNIVERSITY ABUJA

- HT is perhaps an area in which statistics finds its greatest applications – <u>testing validity of claims</u>.
- Testing is needed in practically every field of human activity, particularly business and management research because <u>new ideas and products should be properly evaluated before</u> they are launched into the market place.

Hypothesis testing: continuation…

BAZE UNIVERSITY ABUJA

- Hypothesis tentatively explains an observation that can be tested (i.e. proved or disproved) by further investigation.
- Figuring out the solution to the problem, i.e. "hypothesizing", before you start will help build a roadmap for approaching the problem
- You can express hypothesis as possible root causes of the problem
- Breaking down the problem into key drivers (root causes) or aspects can help formulate hypothesis.

Four key attributes of a good hypothesis

BAZE UNIVERSITY ABUJA

1. Simplicity, clarity, and precision – not vague
2. Verifiable – testable
3. It should have clear nexus with the research problem, the research questions, and the researcher's specified objectives.
4. It should have clear nexus with the relevant theory – it shouldn't be a wild guess.

Formulating your hypothesis…

BAZE UNIVERSITY ABUJA

❑ *Null Hypothesis: H_o*

- Ho is set up by the investigator with the intention of being rejected based on the available statistical evidence.

❑ *Alternative Hypothesis: H_1*

- H_1 is the hypothesis accepted by the investigator after the Ho has been rejected.

Type I and Type II decision errors

BAZE UNIVERSITY ABUJA

- ***Type I error – alpha (α)error:***
  - True H_o is rejected
  - False H_1 is accepted

- ***Type II error – beta (β)error:***
  - True H_1 is rejected
  - False H_o is accepted
  - Why is Type II error usually emphasized in statistical literature?

Traditional testing statistics…

BAZE UNIVERSITY ABUJA

- *t*-test statistics (the student *t*-test) – small sample test, i.e. $n < 30$
- *Z*-test statistics (normal test) – large sample test i.e. $n \geq 30$
- *F*-test statistics (joint test)
- Chi-square distribution X^2

‘Sacrosanct’ confidence levels…

BAZE UNIVERSITY ABUJA

- Confidence level – the extent of confidence or certainty that the investigator has; conventionally (2-tailed):

- $Z_T = Z\alpha = 1\% = 2.58$
- $Z_T = Z\alpha = 5\% = 1.96$
- $Z_T = Z\alpha = 10\% = 1.65$

Decision Rule:

- Calculated test statistic > tabular (critical) value: Accept H_1 and reject H_o

i.e. the parameter testes is statistically significant

- Calculated test statistic < tabular value: Accept H_o and reject H_1

i.e. the parameter testes is statistically insignificant

Statistical testing:

On simpler (less confusing!) approaches…

BAZE UNIVERSITY ABUJA

- The *p*-value is the probability that the test statistic (z^*, t^*, x^*, F^*) will be exceed, and thus *p* is called <u>*the observed level of significance*</u>, in contrast to the α-value which is a priori-level of significance.
- The default value of α = 0.05, and the relationship between *P* and α is as stated below:

i. If $p \geq \alpha$, do not reject Ho

ii. If $p < \alpha$, reject Ho

Statistical testing:

On simpler (less confusing!) approaches…

BAZE UNIVERSITY ABUJA

- In essence, using a table is not necessary when you have the exact probability for a statistic.
- Your suitable computer applications can calculate exact probabilities for most test statistics. If you have an exact probability output from computer software, you simply compare it to your critical alpha level.
- If the exact probability is less than the critical alpha level, your finding is significant.
- if the exact probability is greater than your critical alpha level, your finding is not significant.

Type I and Type II decision errors…

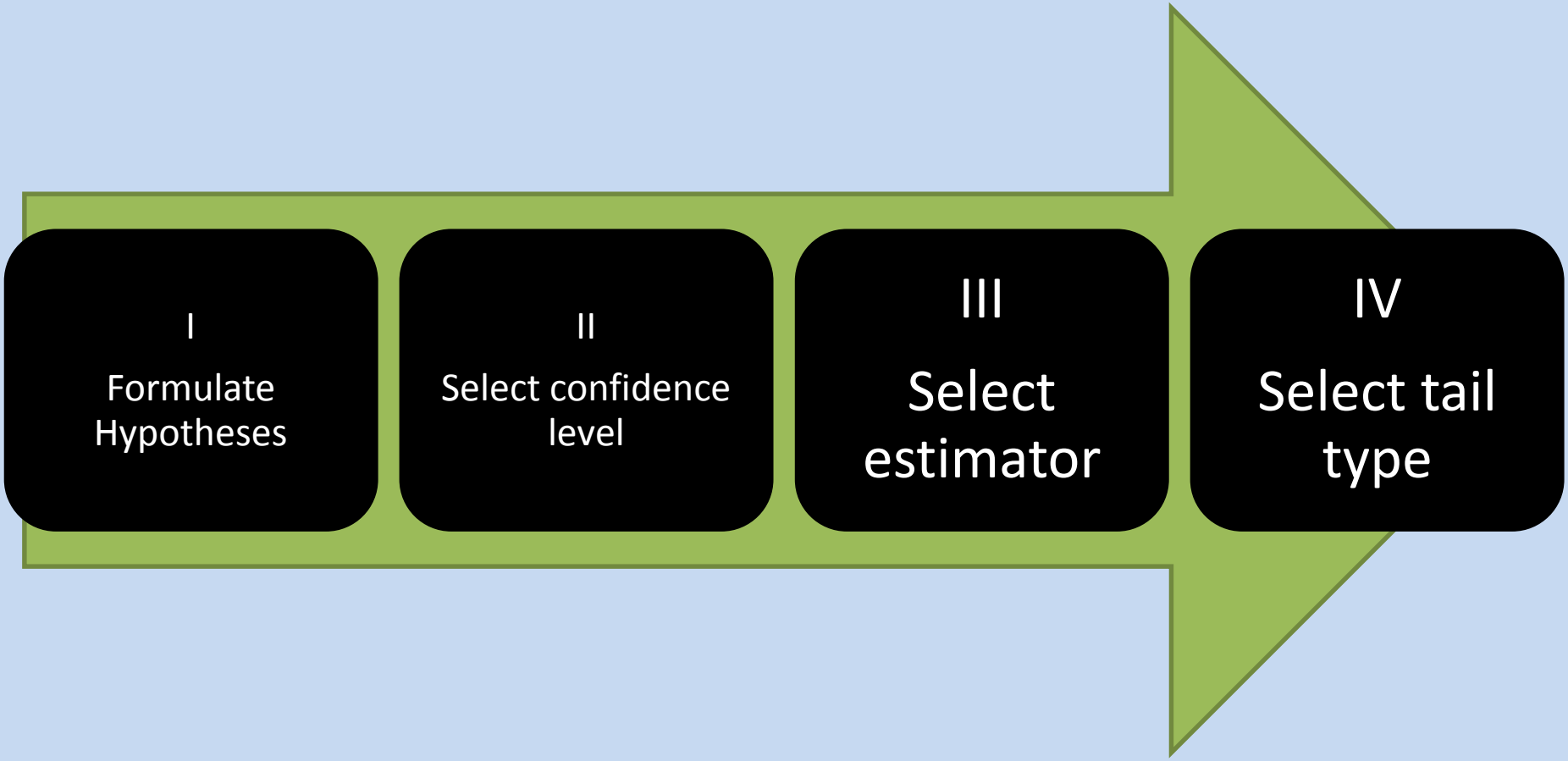
BAZE UNIVERSITY ABUJA

- Pick hypothesis before looking at the data to avoid bias – *apriori* expectation
- Don't confuse statistical importance with substantive importance.
- The *p*-value is not necessarily the probability that the null hypothesis (*Ho*) is true; the *p*-value already assumes that *Ho* is true.
- Rather, it is the probability of rejecting Ho incorrectly on the basis of your results that is displayed in sample assessment.

- You have learnt about <u>how to interpret the adequacy of statistical models</u>, such as beta coefficients, R-squared, adjusted R-squared, finding the critical value of *F* statistic and verifying the *F* calculated value, but ensure that you able to summarize your results in everyday language – less jargons, please!

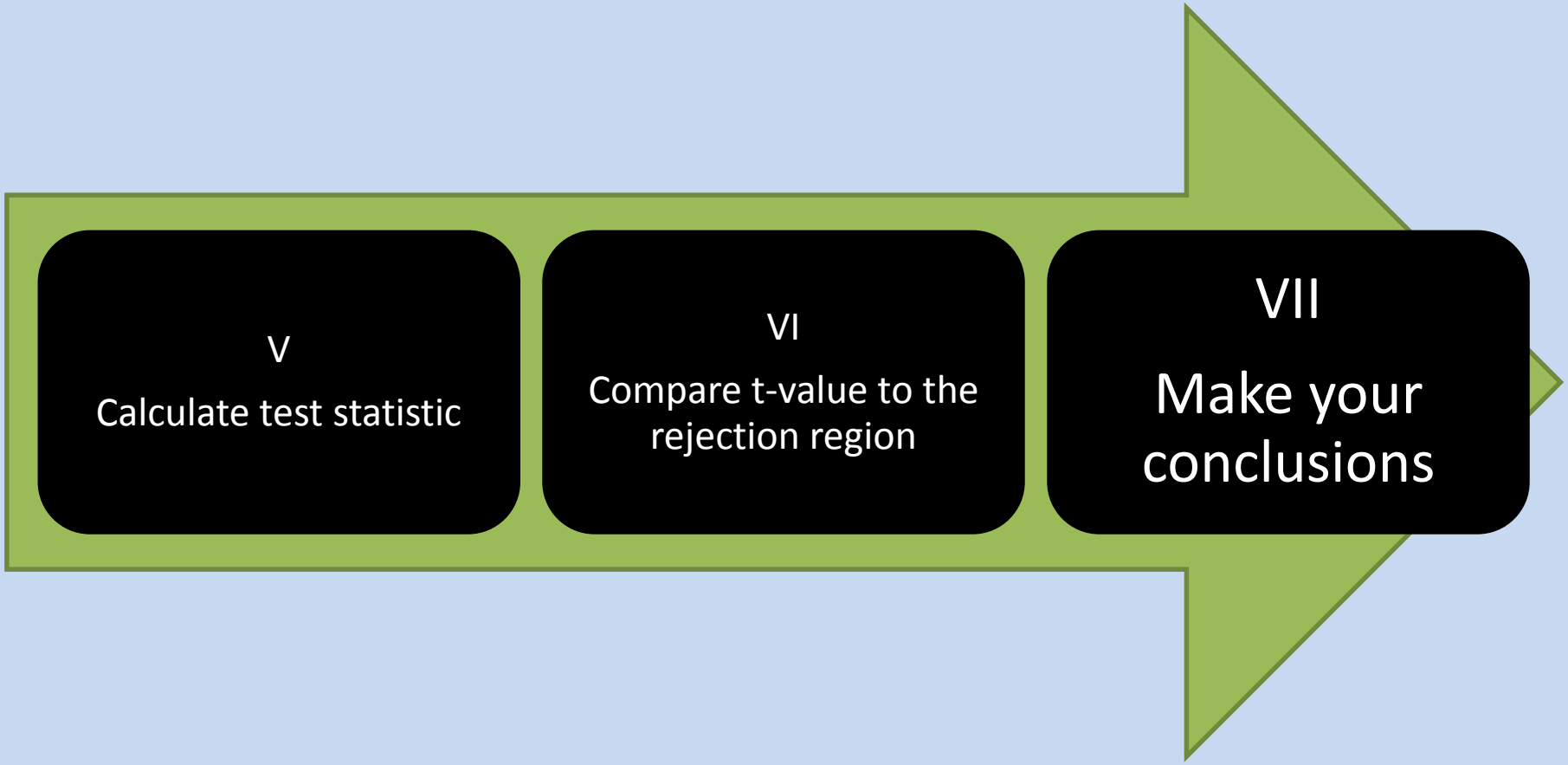
Hypothesis testing
The standard procedure revisited…

BAZE UNIVERSITY ABUJA



Hypothesis testing
The standard procedure revisited…

BAZE UNIVERSITY ABUJA



Statistical testing: Technology is trending, but…

BAZE UNIVERSITY ABUJA

- The manual procedure for hypothesis testing can be tedious.
- Happily, there are statistical tools available these days to solve the problems with relative ease, once they have been carefully / properly defined.
- All you need to do is enter the given data and then issue the <u>appropriate</u> commands.
- Software supporting hypothesis tests: SPSS, MINITAB, EViews, MS Excel, and so on.
- When the results of manual solutions are compared to the computer solutions, we see that they are identical.
- But…

some provisos…

BAZE UNIVERSITY ABUJA

- ✓ Understand the methodology – underlying principles and procedure.
- ✓ Be familiar with the output generated by the computer (you will see some examples of these later in this presentation).
- ✓ Be knowledgeable enough to interpret the computer results meaningfully (back to the recurring imperative of proper understanding of the related theory, the business and the question at hand).

Statistical computing
The case of SPSS…

BAZE UNIVERSITY ABUJA

- ***IBM SPSS*** – Statistical Package for Social Sciences, is widely used computer-based analysis.
- **What it does / can do:**
  - Enables you to input your raw data
  - You can modify your inputted data

The case of SPSS continuation …

BAZE UNIVERSITY ABUJA

- ***With SPSS, you can carry out a wide range of simple, statistical and multivariate analyses such as:***

✓ Calculating means

✓ Cross-tabulation

✓ Correlation analysis

✓ Regression analysis

✓ ANOVA

✓ Discriminant analysis, factor analysis, principal component analysis, etc.

Checking the effectiveness of your research tool:

The 4 Big Questions…

BAZE UNIVERSITY ABUJA

- ✓ ***Reliability*** – how reproducible or repeatable is your diagnosis /results by another (independent) researcher using the same tool / data?
- ✓ ***Validity*** – Is the measure really measuring what it claims to measure – i.e. really serving the purpose for which it was constructed?
- ✓ ***Practicability*** – How feasible, in terms of ease of administration, scoring, interpretation, economy of time, cost and energy?
- ✓ ***Objectivity*** – Will there be disagreement as to what the correct answer / solution is, or, it 'depends'…?! [If it's objective, then, there should be only one correct answer]

To sum up - Common errors in data analytics: Some areas to watch out for

BAZE UNIVERSITY ABUJA

1. Fixation on numbers at the expense of intelligent decision-making - deficient knowledge of what data analysis or statistics is all about.
2. Wrong statistical tools/methods – wrong judgment
3. Foggy definition of variables – incomprehensible questionnaires
4. Un-standardized disparity in variable sizes.
5. Assumption of normal distribution viz-a-viz non-normal reality.
6. Using un-scrutinized data, especially secondary data (what Prof Toyin Falola, 2015 called "propagandistic data", in *The Punch*, 11/6/2015, p. 26)

7. Deficient knowledge of data and various typologies and how each system should be handled differently in data analytics.

8. Bias and sampling errors.

9. Fixation on computer outputs - wrong interpretation of analytical results, e.g. hypothesis test results

10. Absence or inadequate training of enumerators - absence of pilot tests.

BAZE UNIVERSITY ABUJA

Next Session…

Broad areas of emerging shifts in social science research & some implications for researchers

December 21, 2015 @ PES University, Bangalore, India

The four global forces

breaking all the trends (Dobbs et al, 2015)…

BAZE UNIVERSITY ABUJA

1. *Economic shift towards the emerging markets* – Asia, Latin America, Middle East, & Africa.
2. *Technology* – boundless information.
3. *Demographics* – falling fertility / migration / urbanization.
4. *Trade and finance – increased global inter-connectedness* – Asia becoming the world's largest trading region / rising volume of trade between China and Africa (from $9bn in 2000 to $211bn in 2012).

Emerging shifts in social research priorities (www.norrag.org)…

BAZE UNIVERSITY ABUJA

1. Qualitative research vs. quantitative research: The flight to massive analytics / digitization.
2. Increasing recognition / value attached to research for knowledge advancement and robust policy-making – [e.g. ICSSR (India), NISER (Nigeria)]
3. Moving from basic research to applied/empirical research
4. Does blind application of new methodological techniques to observational dataset represent progress in applied research?
5. Increasing involvement in research management, no longer just research production.
6. Mixed-method research getting common – qualitative + quantitative / primary data + secondary data
7. Phenomena of 'big data'.

8. Policy-relevant research as opposed to 'abstract' research / 'academic' versus 'practice-based' research.

9. Specificity: Shift to location/region or sector-specific – case studies

10. Duration: Shift to shorter duration; 3-6 months if not shorter

11. Increased interdisciplinary research.

12. Research professionalization: Private-sector involvement – research-consultants – data collection & compilation, not just data analysis.

13. Increased emphasis on quality of reports (with beautiful graphics, printing, binding and all!)

14. Terminology change – 'research study' or 'research report'?

Other related issues…

[see Kara, 2012 in *Research and evaluation for busy practitioners: A time saving guide.* Bristol: The Policy Press)]

BAZE UNIVERSITY ABUJA

1. ***Anti-methodology / anti-methodological hegemony /imperialism*** – saying 'No' to one-cap-fits-all scientific methodology, or imposition of a particular methodology (see also point 5 below).
2. ***Anti-secondary data*** – fixation on so-called 'primary data', but you can use secondary data whenever justified

3. ***Quality of teaching vs. Quality of teaching***

4. The issue of '***original research '***- research that produces completely new knowledge; i.e. not based on existing knowledge in a new form.

5. Domination (via funding + publications + paper selection) of research 'prowess' / gatekeeping by ***rich countries.***

6. Academic integrity – Honesty, trust, fairness, respect & responsibility *[Dr. (Mrs.) A. T. Alabi, 2005 , in her 'Ensuring academic integrity in research and publications'].* – role of project supervisors in ensuring this is being stressed.

7. Accounting properly for stochastic processes in **time-series** data.

Developments in analytical tools and dimensions…

BAZE UNIVERSITY ABUJA

1. SPSS, MS Excel, EViews, etc.
2. Model building - Correlation - Regression
3. Factor Analysis – Multivariate Data Analysis – Multiple Responses – Principal Component Analysis – Dimension Reduction
4. Defining variables
5. Correspondence Analysis
6. Discriminant Analysis
7. Cluster Analysis, etc.

A word about research funding…

BAZE UNIVERSITY ABUJA

- Corporate Sources
- Private Foundations
- Government research councils

BAZE UNIVERSITY ABUJA

Re-emergence of time series data!

In brief

December 21, 2015 @ PES University, Bangalore, India

The issue of '*spurious correlation*'…

BAZE UNIVERSITY ABUJA

"First impressions can often lead to wrong conclusions."

-*Our Daily Bread,* 2015, p. 346.

'Spurious Correlation'…

BAZE UNIVERSITY ABUJA

- Correlation does not mean causation
- Consequently regression may be also spurious and interpretation may not be valid.

Understanding time-based data…

BAZE UNIVERSITY ABUJA

- In time-based data, the change in a variable is an important concept.
- The change in a variable y_t, also known as its first difference, is given by $\Delta y_t = y_t - y_{t-1}$.
- Δy_t is the change in the value of the variable y from period t - 1 to period t

Understanding time series analysis…

BAZE UNIVERSITY ABUJA

- Time-based data – showing the dynamic movement of a phenomenon over a period of time, usually at equal intervals.
- Used in any domain of applied sciences and engineering that involve temporal measurements.
- Sequence of data points – successive movements over time, e.g. ocean tides, closing stock prices on Indian stock markets, etc.

Beyond analytics: KIV important workplace skills…

BAZE UNIVERSITY ABUJA

- ✓Written & oral communication
- ✓Problem-solving abilities – analytical reasoning
- ✓Creative thinking
- ✓Interpersonal skills (team work)
- ✓Time management
- ✓General professionalism, personal effectiveness, and work ethics

A word about research 'recommendations'…

BAZE UNIVERSITY ABUJA

- It's fine to include 'recommendations' at the end of your paper in applied research to show that you understand the importance / implications of your research.
- However, let you're the recommendations have clear nexus with the text, specifically the evidence / results that you have presented in the paper.
- Don't add 'recommendations' just for the sake of it!

Keywords

BAZE UNIVERSITY ABUJA

Descriptive Statistics, Inferential Statistics, Time Series, Econometric Data, Bias, Sampling Errors, Spurious Correlation & Regression, Time Series Data, Stationarity, Data Collection, Data Analysis, Correlation, Regression, Hypothesis Testing, Enumerators, Pilot Tests, Technology. SPSS, MS Excel, EViews, Model Building, Factor Analysis, Multivariate Data Analysis, Multiple Responses, Principal Component Analysis, Dimension Reduction, Defining Variables, Correspondence Analysis, Discriminant Analysis, Cluster Analysis.

Concluding remarks

BAZE UNIVERSITY ABUJA

- This presentation tried to revisit some of the evolving areas of concerns in contemporary social science research. Deficient knowledge of data and various typologies and how each system should be handled differently in data analytics, bias and sampling errors, fixation on computer outputs - wrong interpretation of analytical results, e.g. hypothesis test results, and inadequate training of enumerators or pilot studies are among areas of increasing concern in research practice.
- Increasing emphasis on interdisciplinary / collaborated research, technology-aided empirical research, relatively shorter duration of research in a 'big data' era, as well as quality and integrity of research reports, among others, have attracted increased attention in contemporary social science research. All these dynamics should drive the experienced practitioner and the budding researcher to sustained production of more focused, quality research for shared well-being, prosperity, and progress.

Stephen ARO-GORDON, Ph.D.
Department of Financial Mathematics
Faculty of Computing & Applied Sciences
Baze University Abuja, NIGERIA
Email: stephen.aro-gordon@bazeuniversity.edu.ng / getyouthsworking@yahoo.com

BAZE UNIVERSITY ABUJA

December 21, 2015 @ PES University, Bangalore, India

Caveat / feedback …

BAZE UNIVERSITY ABUJA

- All the materials used in this presentation are solely for educational purpose, advancement of knowledge and improvement of educational practices to move the society forward. The author will be pleased to make good any omissions or rectify any mistakes brought to his attention at the earliest possible opportunity.