Introduction to Business Problem Solving

LEARNING OBJECTIVES

Upon completing this session, you should be able to do the following:

- define analytics and understand its role in business problem solving
- know major phases of the business problem solving / research process
- understand problem discovery and definition
- understand research design methods
- describe main data types

WHAT IS BUSINESS ANALYTICS?

- Business Analytics is

  "*a process of transforming data into actions through analysis and insight in the context of organisational decision making and problem solving*"

  (Liberatore & Luo 2010).

- Analytics is the use of data, information technology, statistical analysis, and mathematical or computer-based models to help decision makers gain improved insight about their business operations and make better, fact-based decisions.

BUSINESS ANALYTICS APPLICATIONS

- Management of customer relationships
- Financial modelling
- Supply chain management
- Pricing decisions
- Market research
- Social media analysis
- New Product Development
- Product design
- Manufacturing
- Product placement
- Predicting customer churn
- Allocation of resources
- Investments
- Web analytics

AN ILLUSTRATIVE EXAMPLE

- How a bank can increase its customer base for its credit card offerings?

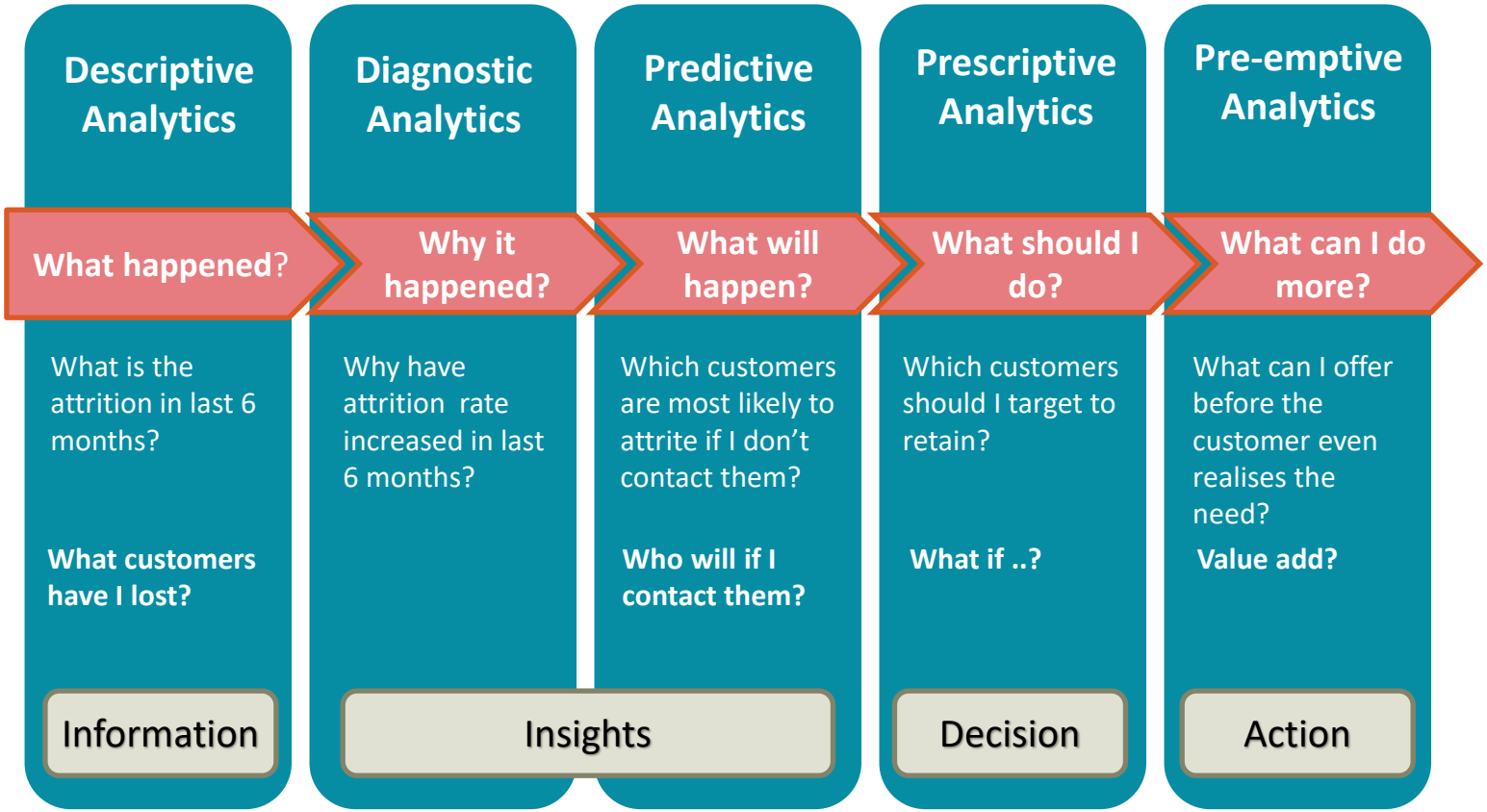
1. Build a model based on historical data to assign credit score to new costumers.
2. What variables the manager should take into account in order to predict a potential client credit score?
3. Decide who should be approached with an offer based on the calculated credit score.

SCOPE OF BUSINESS ANALYTICS

- BA begins with the collection, organisation and manipulation of data and is supported by three major components:
  - Descriptive Analytics (including diagnostic)
  - Predictive Analytics
  - Prescriptive Analytics (including pre-emptive)

- Although the tools used in descriptive, predictive and prescriptive analytics can be different, many applications involve all three.

ANALYTICS TECHNIQUES OVERVIEW

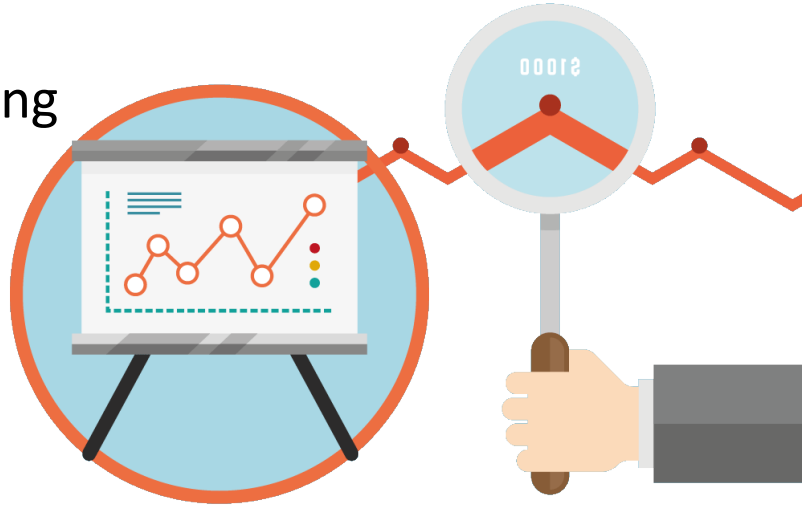


DESCRIPTIVE ANALYTICS

- Most commonly used and well understood type of analytics.
- Uses data to understand past and present performance and make informed decisions.
- Uses fundamental tools and methods of data analysis and statistics, focusing on:
  - Descriptive statistical measures and data visualisation
  - Probability distributions, sampling and estimation
  - Statistical inference

PREDICTIVE ANALYTICS

- Analyses past performance in an effort to predict the future by examining historical data, detecting patterns or relationships in these data, and then extrapolating these relationships forward in time.
- Techniques include:
  - Regression and Forecasting
  - Data mining and machine learning



PRESCRIPTIVE ANALYTICS

- Also referred to as Decision Analytics*
- Uses optimization to identify the best alternative to minimize or maximize some objective.
- Techniques include the use of advanced models.

* Please note these models will not be covered in MIS771

EXAMPLE
RETAIL MARKDOWN

- Most department stores clear seasonal inventory by reducing prices.
- The question is:

  "When to reduce the price and by how much?"

Descriptive analytics:

Examine historical data for similar products (prices, units sold, advertising, …)

Predictive analytics:

Predict sales based on price

Prescriptive analytics:

Find the best sets of pricing and advertising to maximize sales revenue

YOUR TURN…
DESCRIPTIVE, PREDICTIVE, PRESCRIPTIVE?

- RED© is planning to introduce its first mobile handset to the market in 2023. Question is, which feature of a newly engineered handset will be received well in the marketplace?
- The analysis should also help RED determine the price points of the new model in various markets across the globe…

HYDROGEN™

YOUR TURN...

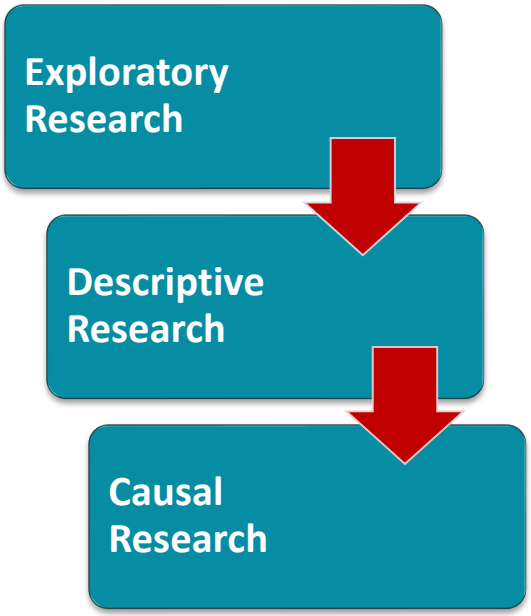
DESCRIPTIVE, PREDICTIVE, PRESCRIPTIVE?

- airbnb© - an online marketplace and hospitality service – is planning to offer a limited number of vouchers which costs $200/customer to its customers.
- They have two customer base: Those with Platinum M Cards and those with Gold M Cards.
- The offer breakseven, if a customer makes a purchase of minimum $5,000 in his entire lifetime.
- airbnb wants to target customers who are likely to make a purchase of $5,000 as early as possible.
- Question is, which customer base should you target for this offer?

WHAT THEN IS STATISTICS?

- Statistics relates to the collection, analysis, interpretation, and presentation of data.
- Statistical methods underpin many of the techniques used in business analytics. The methods can be used to:
  - Summarise a collection of data
  - Draw inferences about an entire population
  - Make predictions or forecasts

TYPES OF BUSINESS RESEARCH



- Exploratory Research:

  Conducted to clarify ambiguous situations or discover ideas that may be potential business opportunities.

- Descriptive Research:

  Describe characteristics of objects, people, groups, organizations, or environments.

- Causal Research

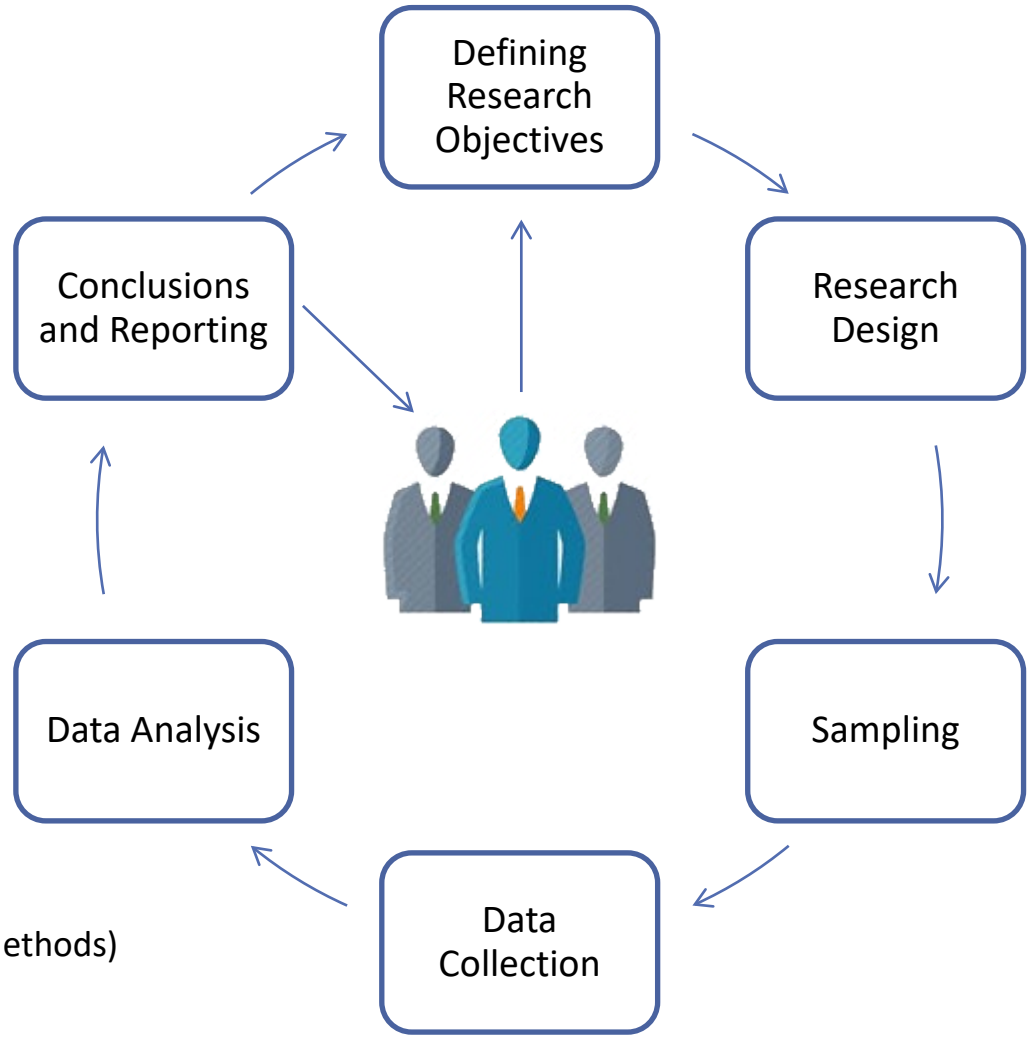
  Allows causal references to be made; seeks to identify cause-and-effect relationships.

CHARACTERISTICS OF DIFFERENT TYPES OF BUSINESS RESEARCH

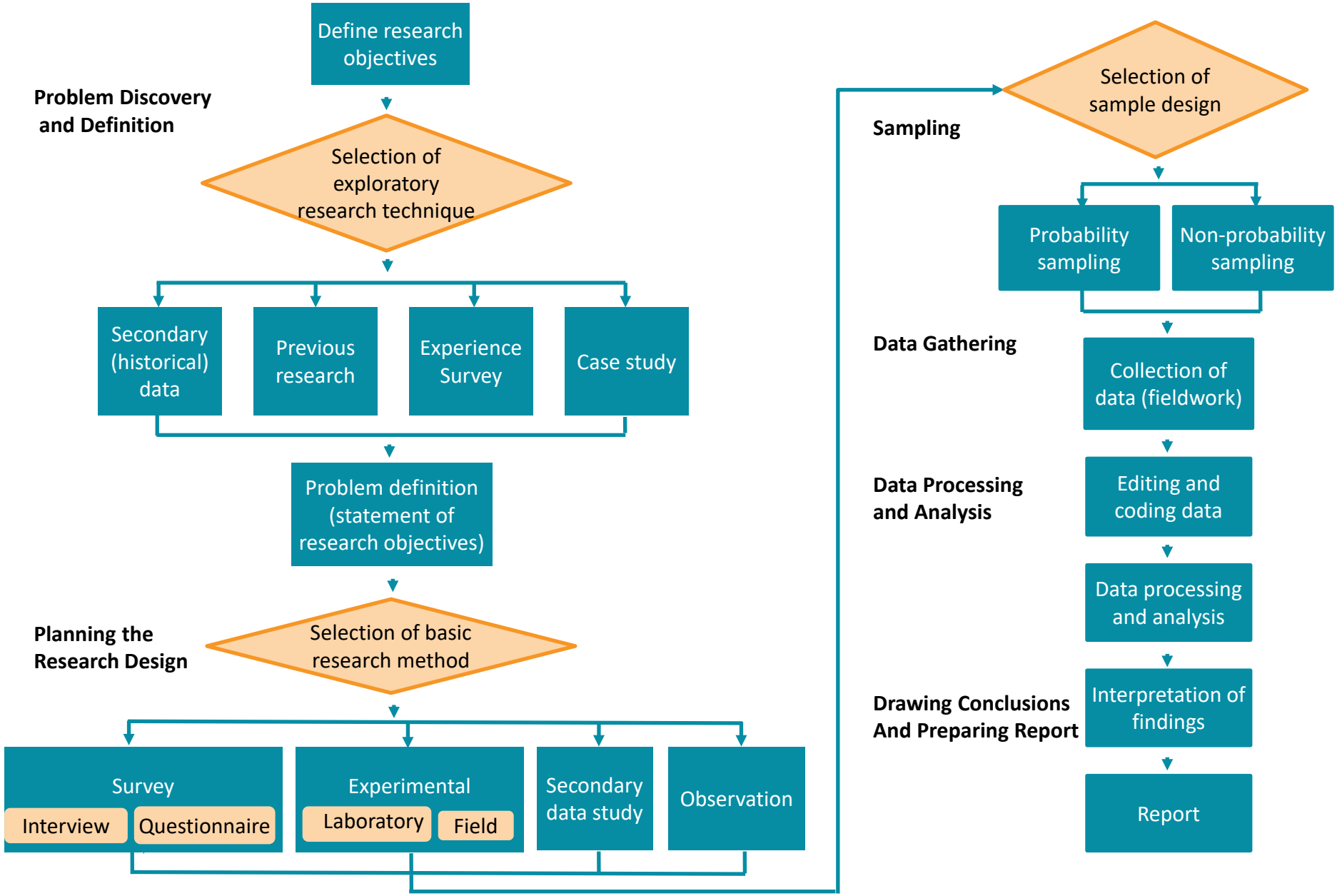
	Exploratory Research	Descriptive Research	Causal Research
Amount of Uncertainty Characterising Decision Situation	Highly ambiguous	Partially defined	Clearly defined
Key Research Statement	Research question	Research question	Research hypothesis
When conducted?	Early stages of decision making	Later stages of decision making	Later stages of decision making
Usual Research Approach	Unstructured	Structured	Highly structured
Discovery of Results	Discovery oriented, productive, but still speculative. Often in need of further research.	Can be confirmatory although more research is sometimes still needed. Results can be managerially actionable.	Confirmatory oriented. Fairly conclusive with managerially actionable results often obtained.

STAGES OF BUSINESS RESEARCH PROCESS

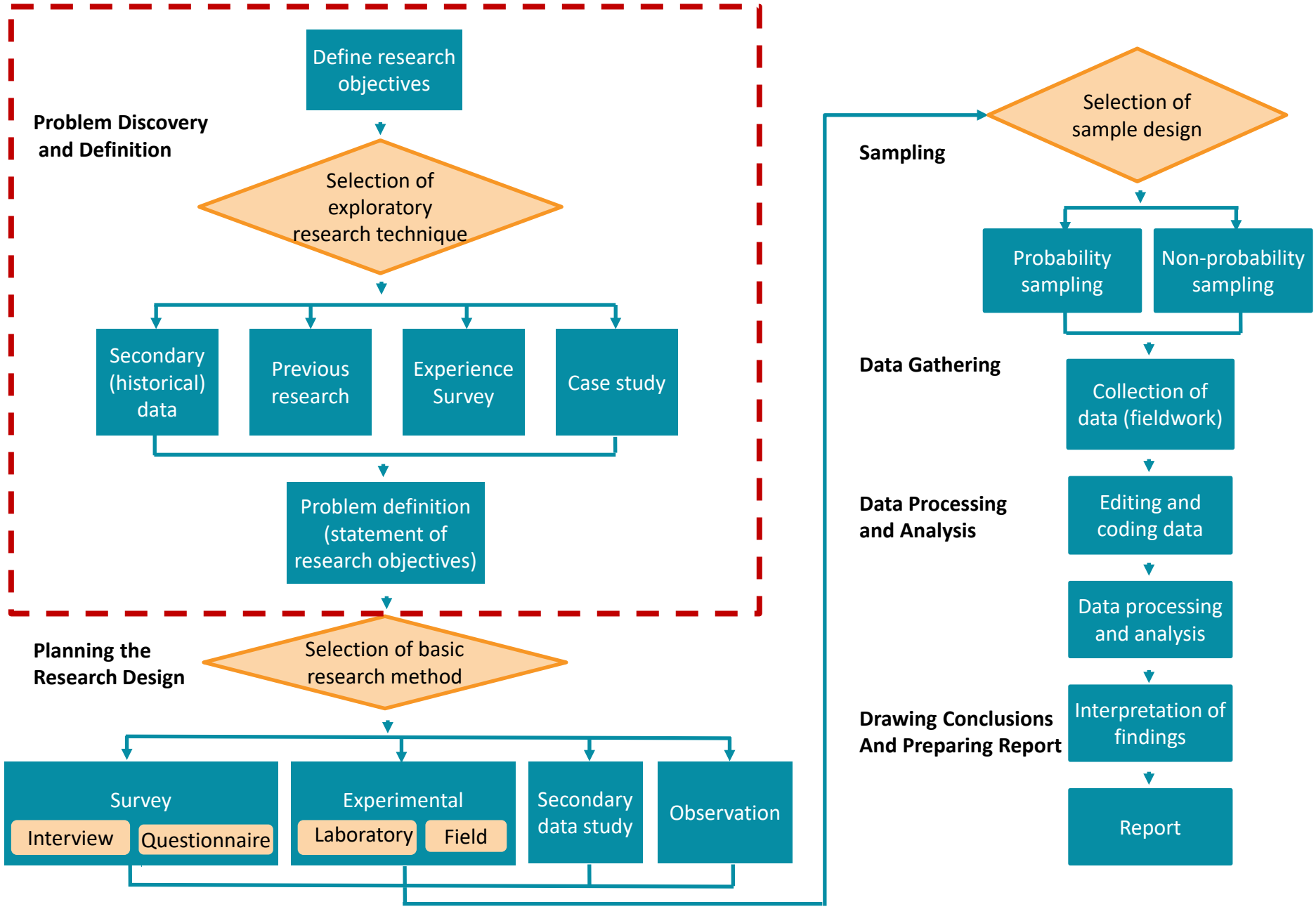
1. Defining Research Objectives
2. Research Design
3. Sampling
4. Data Collection
5. Data Analysis
6. Conclusions and Reporting



(Source: Zikmund et al, Business Research Methods)



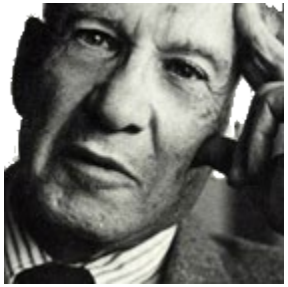
BUSINESS RESEARCH PROCESS FLOW CHART



BUSINESS RESEARCH PROCESS FLOW CHART

PROBLEM DISCOVERY AND DEFINITION

The most serious mistakes are not being made as a result of answers. The true dangerous wrong thing is asking the wrong question.



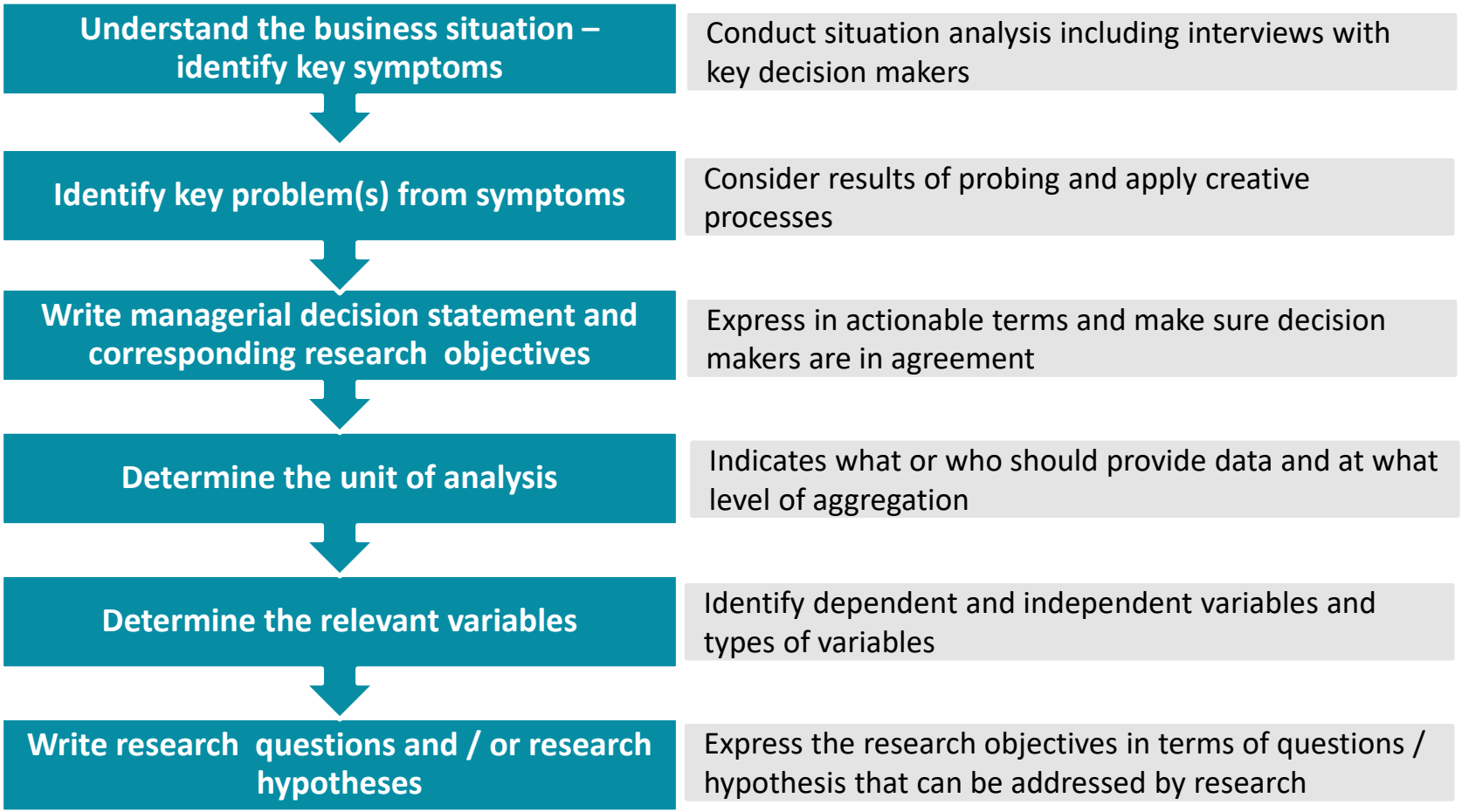
Peter Drucker

PROBLEM DISCOVERY AND DEFINITION

- Problems and symptoms can be confusing and may only be symptoms of a deeper problem.

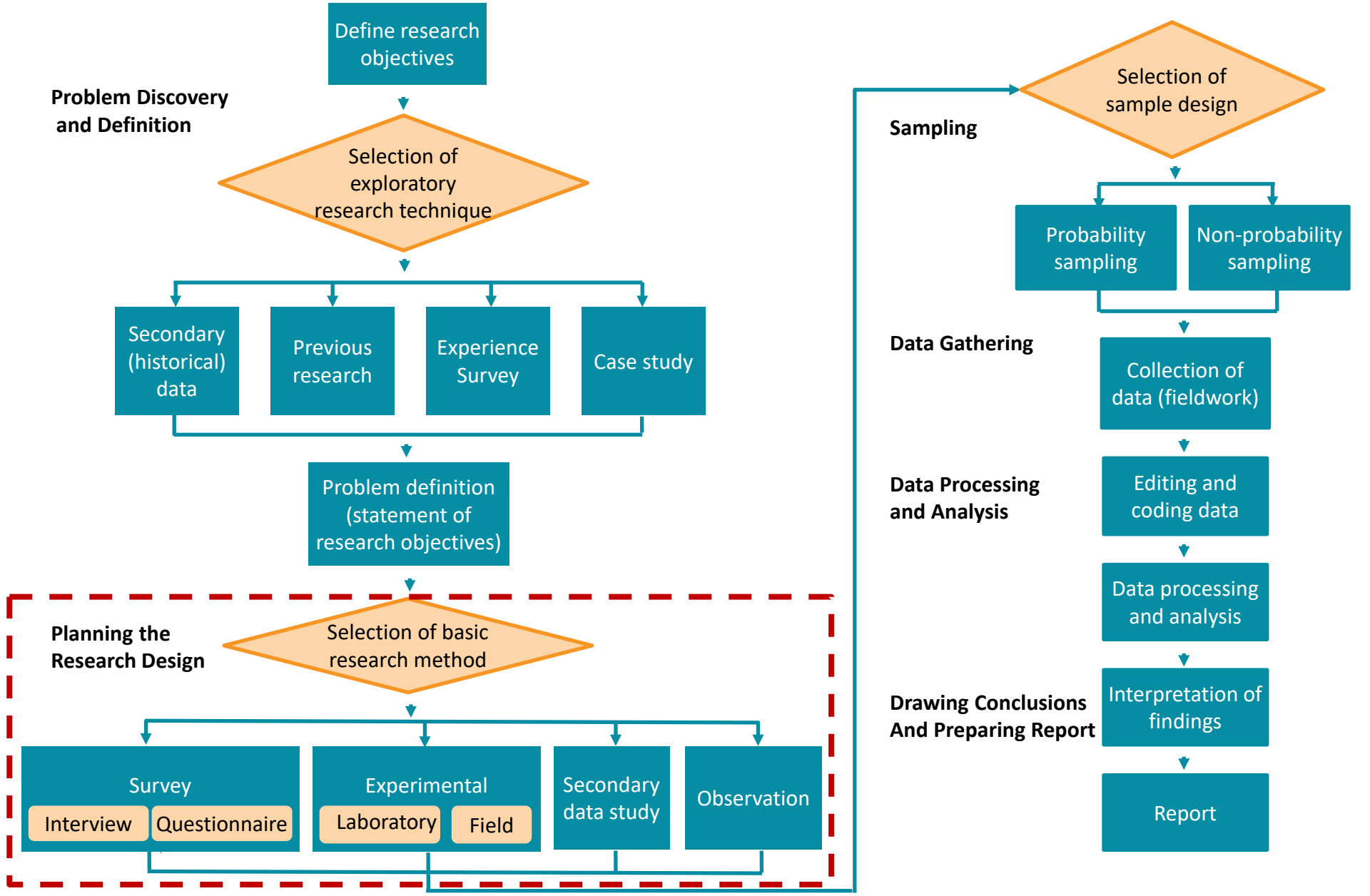
Firm	Symptoms	Based on Symptoms	Underlying Causes
Manufacturer of orange soft drinks	Consumers say the sugar content is too high	Determine consumer preferences for alternative levels of sugar content	

THE PROBLEM-DEFINITION PROCESS



EXAMPLE
THE PROBLEM-DEFINITION

Situation	Manufacturer of palm-sized computer with wireless internet access believes B2B sales are too low.
Symptoms	• Distributors complain prices are too high. • Business users still use larger computers for displaying information to customers or smart phones for other purposes.
Likely Problem (s)	• Business users do not see the advantages of smaller units. • The advantages are not outweighed by costs. • The transition costs may be a drawback for B2B customers more than for B2C customers.
Decision Statement	What product features can be improved and emphasized to improve competitive positioning of the new product in B2B markets?
Research Objectives	List actions that may overcome the objections (switching costs) of B2B customers toward adoption of the new product.
Research Questions	What are the factors that most lead to perceptions of high switching costs?
Research Hypothesis	• ***Perceived difficulty*** in learning how to use the new device is related to ***switching costs***. • ***Price*** is positively related to ***switching costs***. • ***Knowledge*** of new product is positively related to ***switching costs***.



BUSINESS RESEARCH PROCESS FLOW CHART

RESEARCH DESIGN

- After the research problem is formulated, the research design must be developed.
- A research design is a master plan that specifies the methods and procedures for collecting and analysing the needed information.
- A research design provides a framework or plan of action for the research.
- The objectives of the study, the available data sources, the urgency of the decision, and the cost of obtaining the data will determine which method should be chosen.

EXPERIMENTAL RESEARCH

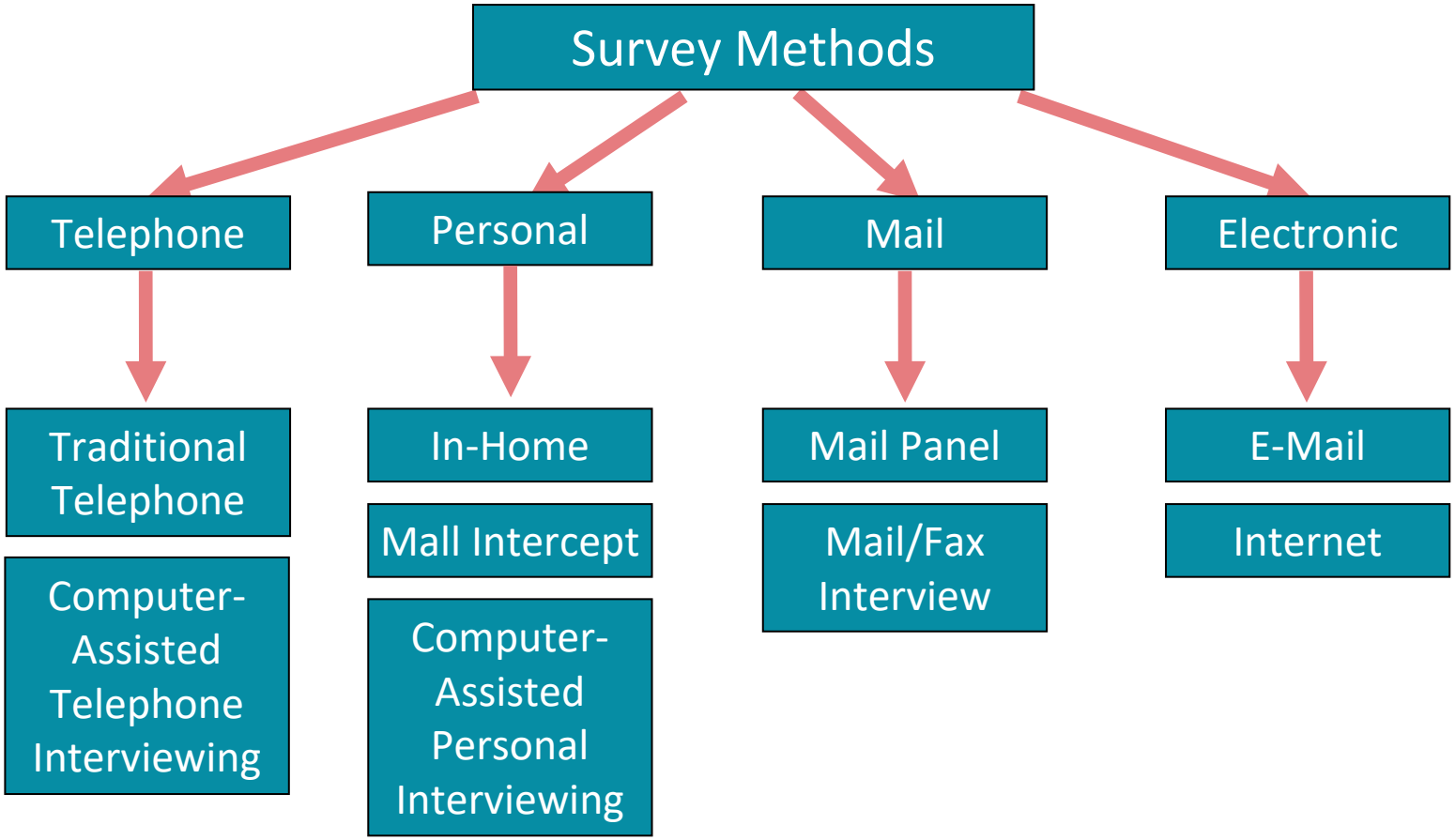
- With experimental design, variables/items of interest are first identified. Then one or more factors are controlled so that data can be obtained about how the factors influence the variables

OBSERVATION

- Observation is the systematic process of recording the behavioural patterns of people, objects and occurrences as they are witnessed.
- No questioning or communicating with people
- The researcher witnesses and records information as events occur, or compiles evidence from records of past events.

Phenomenon	Example
Physical action	A shopper's movement pattern in a store
Verbal behaviour	Statements made by airline travellers while waiting in line
Expressive behaviour	Facial expressions, tones of voices and other forms of body language
Spatial relations and locations	How close visitors at an art museum stand to paintings
Temporal patterns	How long fast-food customers wait for their orders to be served
Physical objects	What brand-name items are stored in consumers' pantries
Verbal and pictorial records	Bar codes on product packages

SURVEY


Survey Methods
Telephone
Personal
Mail
Electronic
Traditional Telephone
Computer-Assisted Telephone Interviewing
In-Home
Mall Intercept
Computer-Assisted Personal Interviewing
Mail Panel
Mail/Fax Interview
E-Mail
Internet

SECONDARY DATA

- Most companies today have access to numerous sources of external data as well as internal sources from enterprise-wide systems.
- The data deluge is the result of the prevalence of:
  - Automatic data collection
  - Electronic instrumentation
  - Online transactional processing
- There is a growing recognition of the untapped value in these databases.
- This recognition is driving the development of business analytics.

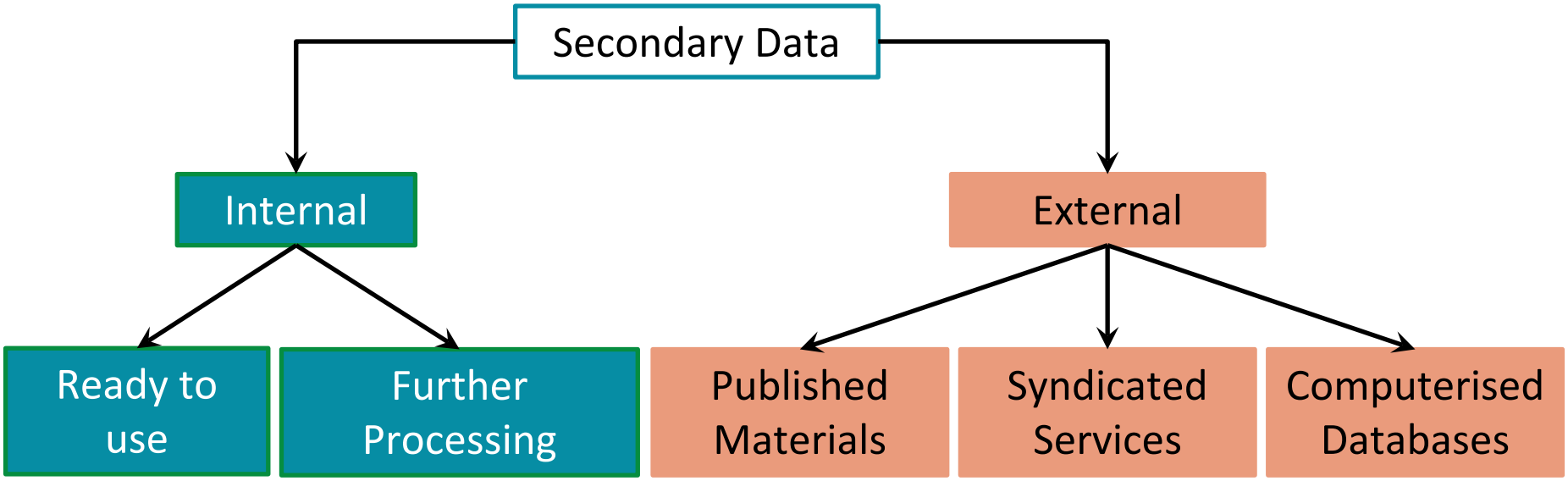
EXAMPLE

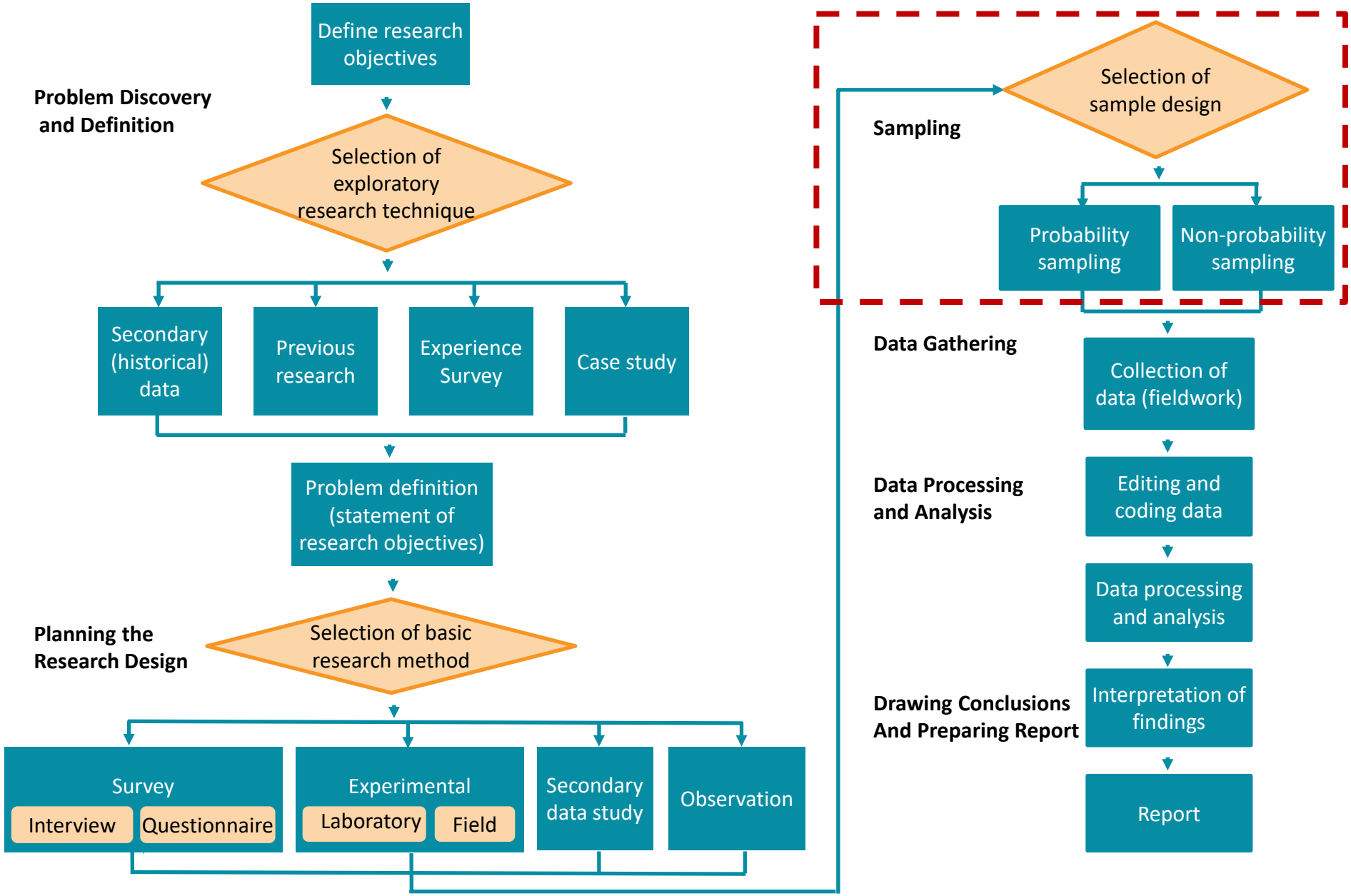
SECONDARY DATA

- Bounce.io (Founded by Scott Brown – 2013) is an online tool for users to learn more about email bounces.

- Bounce.io takes bounced emails and does two things:
  1. It adds advertising to the bounced emails that actually came from people who, for whatever reason, were sending an email to the parked domain.
  2. The email from spam bots gets sold to data security companies that view it as a fresh source never tapped before.

SECONDARY DATA





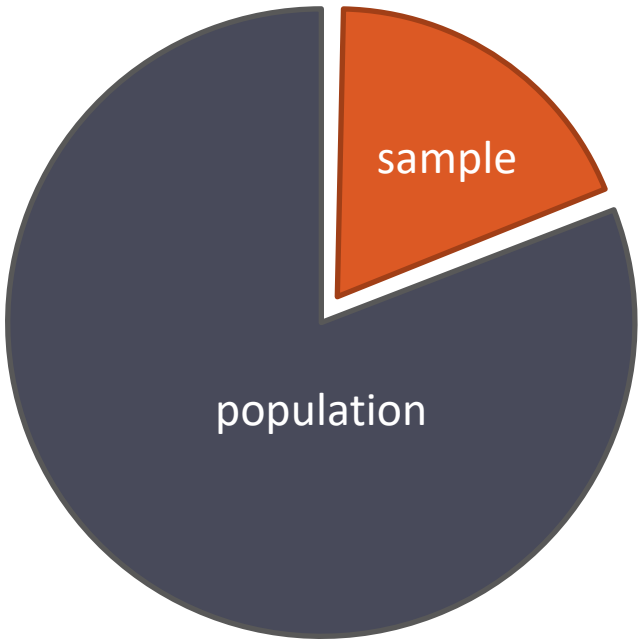
BUSINESS RESEARCH PROCESS FLOW CHART

SAMPLING

POPULATION
Any complete group of entities that share some common set of characteristics

SAMPLE
A subset, or some part, of a larger population

CENSUS
An investigation of all the individual elements that make up a population

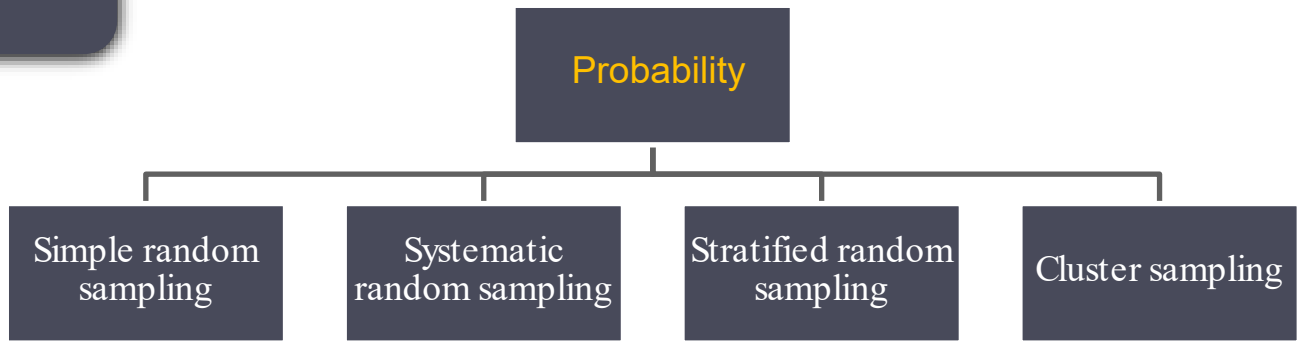


WHY SAMPLE?

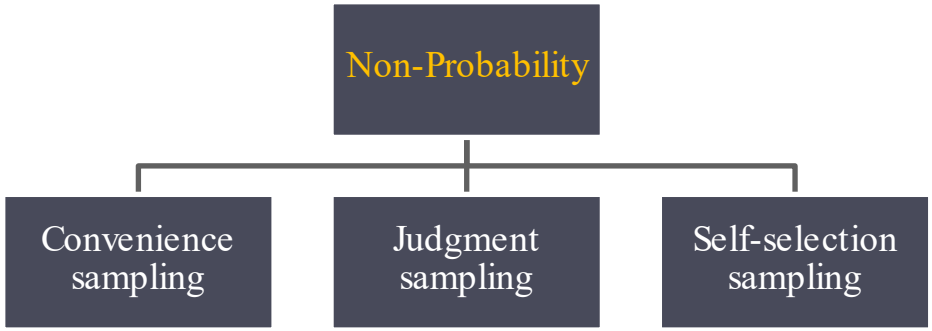
- Pragmatic Reasons
  - Budget and time constraints
  - Some of the population is inaccessible or impossible to contact
- Accurate and Reliable results
  - Most properly selected samples give results that are quite accurate.
- Destruction of Test Units
  - e.g. If a manufacturer of firecrackers wished to find out whether each unit met a specific production standard, no product would be left after the testing.
  - e.g. If an experiment sale presentation is presented to every potential customer, no prospects would remain to be contacted after the experiment.

SAMPLING TECHNIQUES

PROBABILITY SAMPLING
Known, nonzero probability for every element



NON-PROBABILITY SAMPLING
Probability of selecting any particular member is unknown



NON-PROBABILITY (NON-RANDOM) SAMPLING

- Where the probability of any particular element of the population entering the sample is unknown, and
- Some individuals or items in the population have a greater chance of selection than others.

This means that statistically valid statements or inferences cannot be made about the precision of the estimates.

NON-PROBABILITY (NON-RANDOM) SAMPLING

- There are advantages to non-random sampling:
  - The sampling costs are lower and the implementation is easier.
  - Sometimes there is no alternative, as a random sample cannot be taken.

- However, almost always, this means that:
  - Inferential statistical techniques cannot be applied.
  - In particular, no statistical statement should be made about the precision of the result.

SELF-SELECTION - CONVENIENCE SAMPLING (NON-RANDOM)

- Self-selection:

  People are invited to submit a questionnaire.

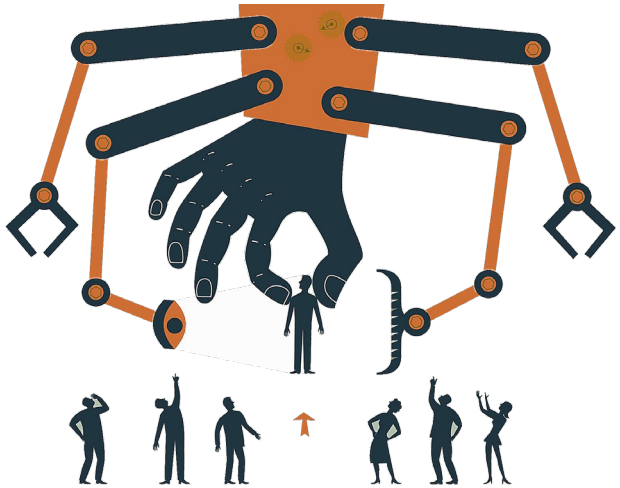
  For example, through a magazine or on-line.

- Convenience:

  Where the elements that are included in the sample are chosen because of their accessibility or willingness.

  For example, students are not randomly sampled, but interviewed as they enter the library.

JUDGMENT SAMPLING
(NON-RANDOM)



- Judgment:

  A knowledgeable person selects sampling units that he/she feels are the most representative of the population.

  – The quality of the result is dependent on the judgment of the person selecting the sample.

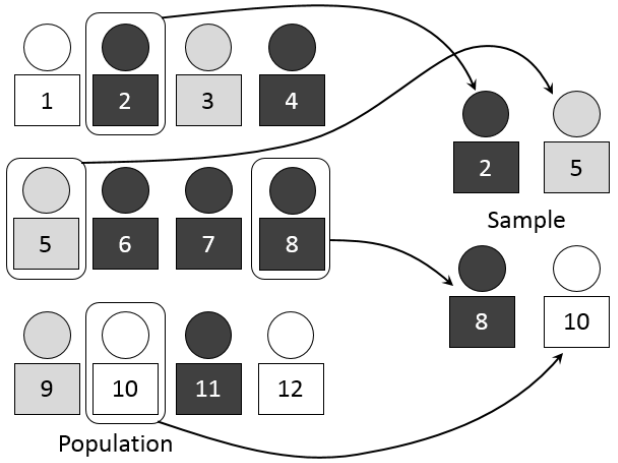
  – The person may be well-qualified, and their judgment may be highly regarded, but it does not give the results statistical validity.

PROBABILITY SAMPLING
(RANDOM)

- Each element of the population has a known (non-zero) chance of being included in the sample chosen.
- Should be used where possible.
- Inferential statistics requires random samples.
- For the rest of this course we assume that all samples are random samples.

SIMPLE RANDOM SAMPLING

- Every individual or item from the sampling frame (list of eligible elements from the population) has an equal chance of being selected.
- In addition, every sample of a fixed size has the same chance of selection as every other sample of that size.
- Most elementary random sampling technique.
- Samples normally obtained from computer random number generators.

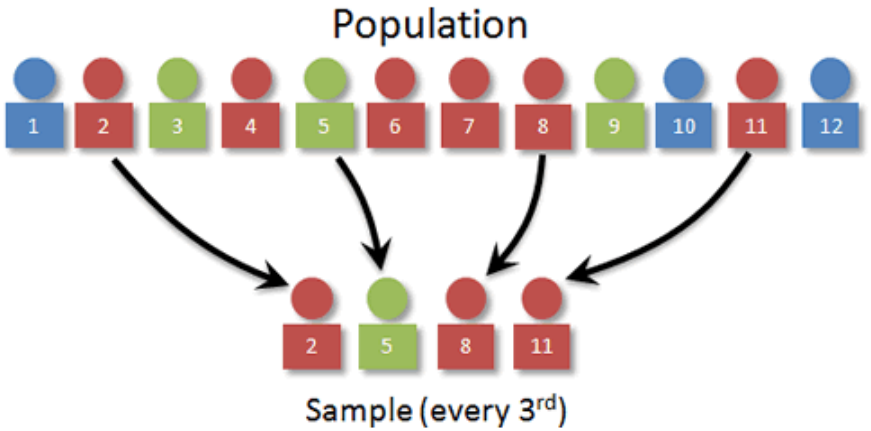


SIMPLE RANDOM SAMPLING
(AN EXAMPLE)

- A components manufacturer makes approximately 2000 of a particular type of part per week. 60 parts are to be selected in a simple random sample to check for defects.
- If each part has it's own unique part number, then a list can be generated with all 2000 part numbers.
- Using a random number generator, 60 part numbers would be chosen at random in this range.
- The 60 parts with the corresponding numbers would be those included in the sample.

SYSTEMATIC SAMPLING

- Decide on sample size.
- Divide frame of N individuals into groups of k individuals: $k = N/n$.
- Randomly select one individual from the 1^{st} group.
- Select every k^{th} individual thereafter.
- Some situations it is more convenient/faster than a simple random sample.

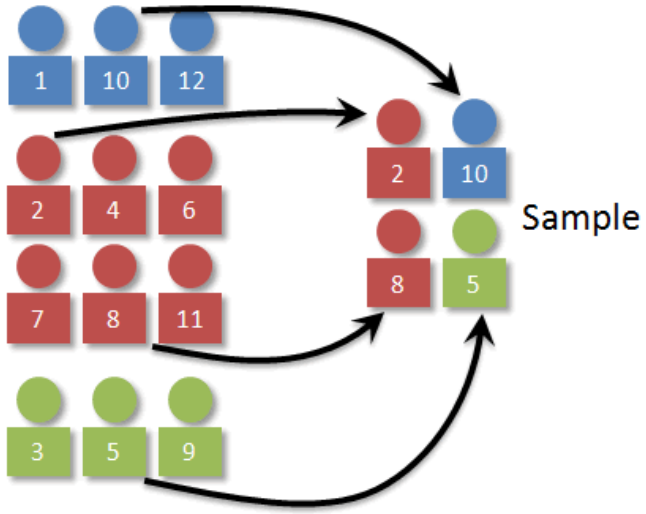


SYSTEMATIC SAMPLING
(AN EXAMPLE)

- With the manufacturing example, 60 parts are to be randomly selected from 2,000.
- This means k = 2000/60 = 33.3 or one in every 33 is to be selected.
- Begin by randomly selecting a number from 1 to 33.
- Say the number chosen was 17.
- We would then choose every 33rd part made from 17, that is 17, 50, 83, 116, 149, …

STRATIFIED SAMPLING

- Divide population into two or more subgroups (called strata) according to some common characteristic.
- Examples of strata – gender, age groups, states etc.
- A simple random sample is selected from each subgroup (stratum), with sample sizes proportional to strata sizes.
- Samples from each stratum are combined into one.

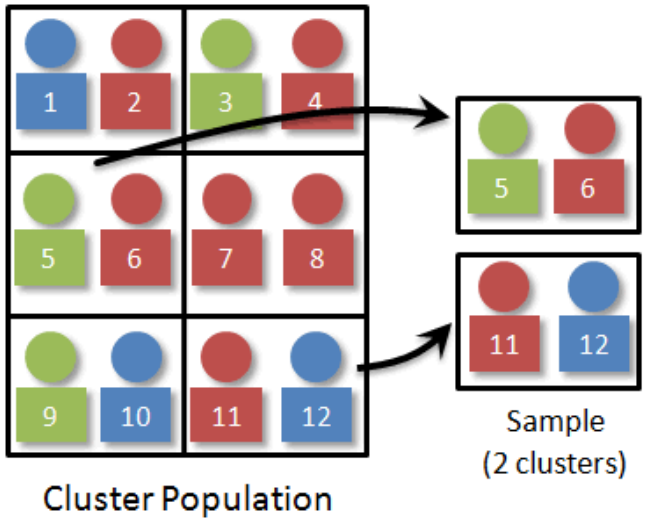


STRATIFIED SAMPLING
(AN EXAMPLE)

- The manufacturing company runs a morning, afternoon and evening shift. Approximately half of the components are made during the morning shift, 30% during the afternoon shift, and the remaining 20% during the evening shift.
- To take a stratified random sample of 60 parts across the three shifts will mean:
  - $0.50 \times 60 = 30$ parts from the morning shift
  - $0.30 \times 60 = 18$ parts from the afternoon shift
  - $0.20 \times 60 = 12$ parts from the evening shift

CLUSTER SAMPLING

- Population is divided into several ‘clusters’, each representative of the population.
- A simple random sample of clusters is selected
  - All items in the selected clusters can be used, or items can be chosen from a cluster using another probability sampling technique.
- These clusters are often based on geographical groupings, such as postcode areas, electorates, city blocks or streets, where the cost of accessing individual households can be greatly reduced.

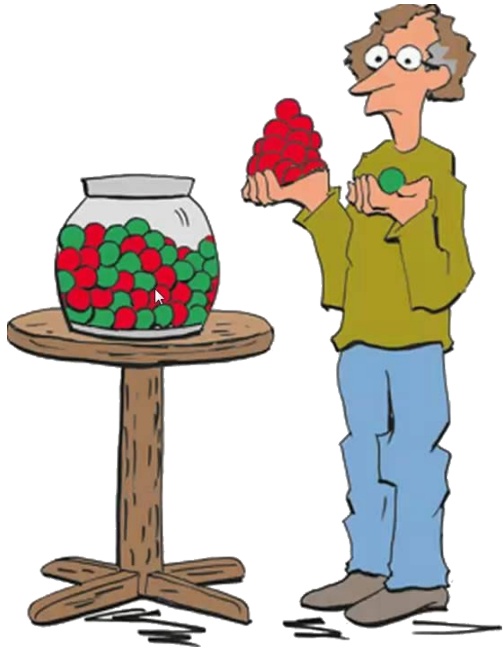


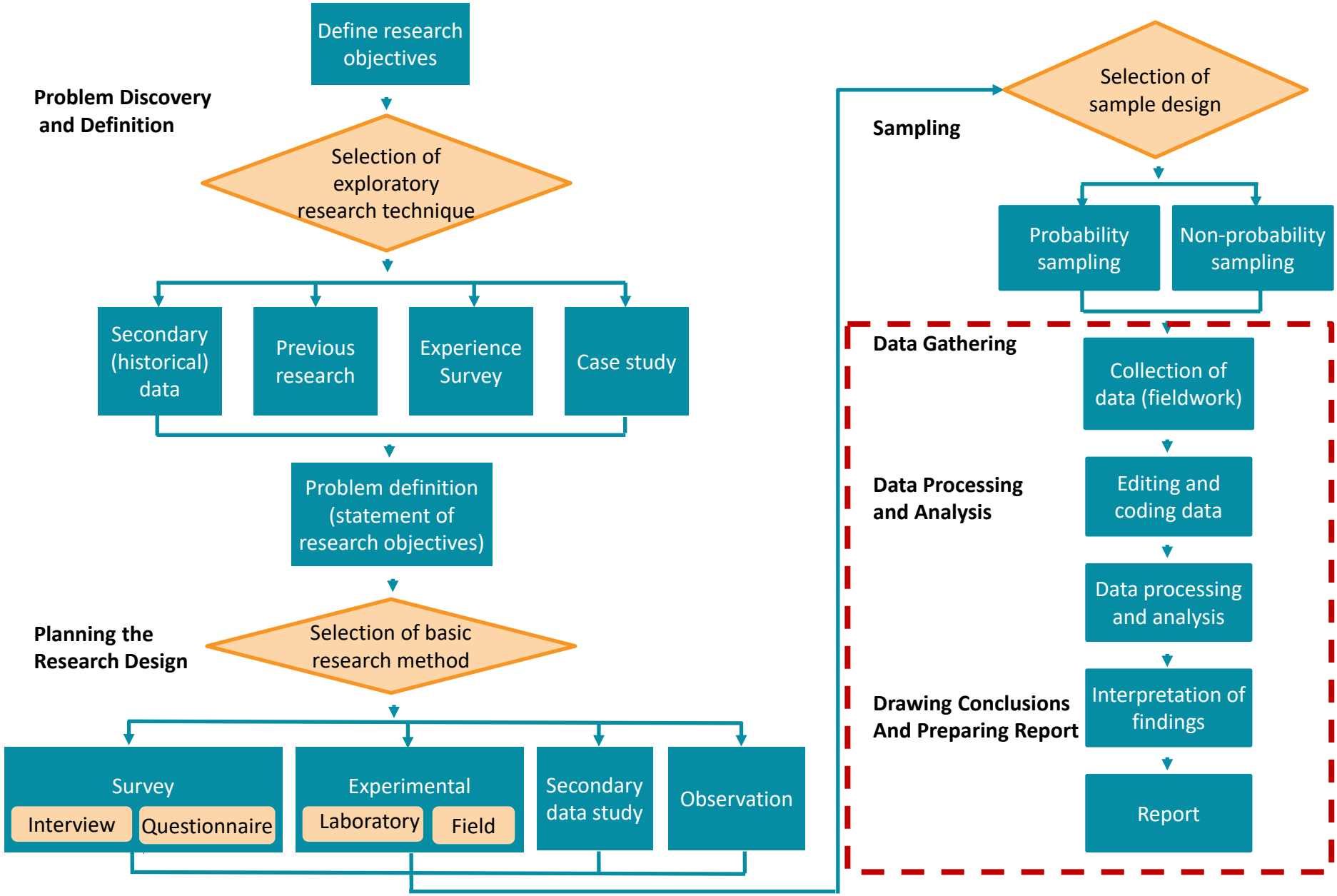
CLUSTER SAMPLING
(AN EXAMPLE)

- We randomly select 30 statistical units / streets from the 400.
- In this example, all elements in a cluster are included as part of the sample.
- That is, we would interview every household in any street chosen.
- Thus, it is much less expensive for interviewers to visit every house in a street than to visit households scattered all over a city.
- But the sample is still a random one.

SAMPLING ERRORS

- Random sampling errors
  - Difference between the sample result and the result of a census conducted using identical procedures.
  - Function of sample size. As sample size increases, random sampling error decreases.
- Systematic (non-sampling) errors
  - Result from primarily the nature of a study's design and the correctness of execution.
  - Errors due to sample selection problems (sampling frame error, non response error)





BUSINESS RESEARCH PROCESS FLOW CHART

GATHERING, PROCESSING AND ANALYSING DATA

- Gathering Data
  - Process of gathering or collecting information.
  - Data may be gathered by human observers or interviewers, or they may be recorded by machines.
- Processing Data
  - Gathered data must be converted and ‘cleaned’ into a suitable format.
  - Generally begins with editing and coding data.
- Data Analysis
  - The application of reasoning to understand the data that have been gathered.
  - May involve determining consistent patterns and summarizing the relevant details revealed in the investigation.

CATEGORICAL DATA

TYPES OF DATA

- Labels or names used to identify an attribute of each entity.
- Often referred to as qualitative data.
- Can be recorded in either numeric or nonnumeric format.
- Appropriate statistical analyses are rather limited.

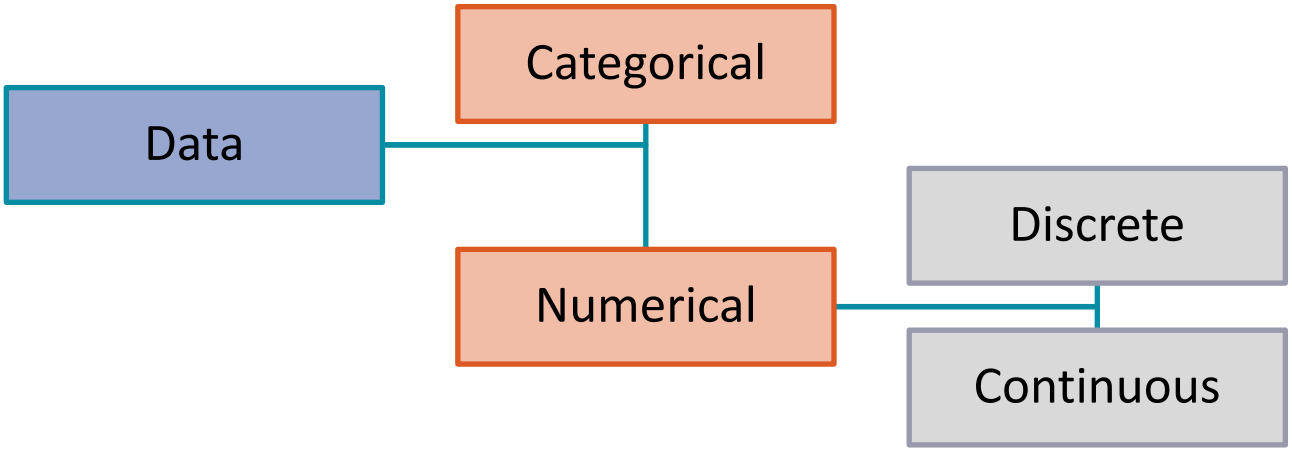
NUMERICAL DATA
TYPES OF DATA

- Numerical data indicate how many or how much:
  - Discrete: if measuring how many
  - Continuous: if measuring how much

- Often referred to as quantitative data.
- Ordinary arithmetic operations are meaningful for quantitative data.

NOTE

- Numerical data can be converted to categorical data.
- A common mistake is to treat data collected using a rating/ranking (e.g. Likert) scale as numerical data (this can only be justified in certain circumstances).



SCALE OF MEASUREMENT

Measurement Scale	Examples	Numerical Operations	Descriptive Statistics
Nominal	Yes – No Male – Female	**Counting**	• Frequencies • Mode
Ordinal	Student class rank (Dissatisfied, Satisfied, Very Satisfied, Delighted) Level of Education (Some high school, High school diploma, Some college, College degree, Graduate degree)	**Counting** **Ordering**	• Frequencies • Mode • Median • Range
Interval	Student grade point average (GPA) Temperature (Celsius and Fahrenheit)	**Common arithmetic operations**	• Frequencies • Mode • Median • Range • Mean • Variance • Standard deviation
Ratio	Amount spent on last purchase Salesperson sales volume Number of stores visited on a shopping trip	**All arithmetic operations**	

QUESTIONS?