

# **Data Collection**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes.







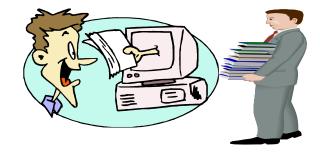
Objective Method



Subjective Method



Use of Existing Records







# Objective Method

- The data are collected through measurement, counting or by observation.
- This method requires the use of a measuring or counting instrument.









# Subjective Method

- The information is provided by identified respondents.
- The instrument used to gather data may take the form of a questionnaire.
- The researcher collects data by:
  - conducting personal interviews either face-to-face, telephone call, mailing or online







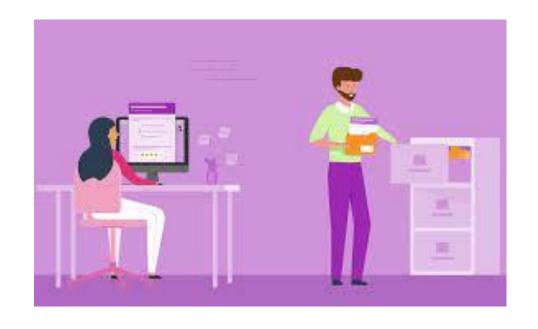






# Use of Existing Records

This method uses data which have been previously collected by another person or institution for some other purposes.







# Quantitative Data collection Tools

Survey

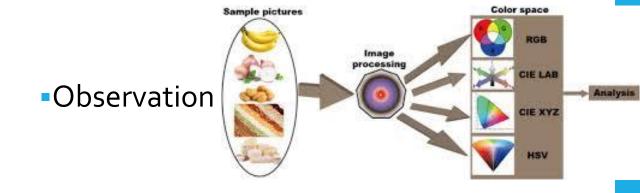


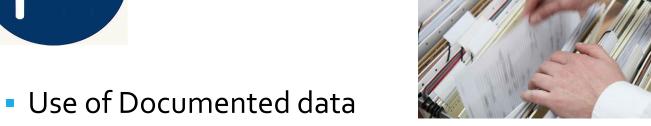
Experimentation



Registration









# Qualitative Data Collection Tools

Key Informant Interview



Participant Observation



Focus Group Interview







Using Documents as Bibliography/ Autobiography and Diaries

# Basic Concepts in Sampling

#### Survey

- = uses **survey sample**
- = inquiries: more in-depth
- = reasons:

lesser cost, greater speed, greater scope, greater accuracy, necessity, ethical consideration

#### **Features:**

- Identifiable
- Measurable
- Describable
- Sampling
- Estimable
- Randomness







# Basic Concepts in Sampling



#### **CENSUS**

- = complete enumeration of the units in the universe
- = inquiries: straightforward counts

#### **EXAMPLES:**

- Census of Population and Housing
- Census of Philippine Business and Industry
- Census of Agriculture and Fisheries
- Barangay Management Information System





# Think about these...

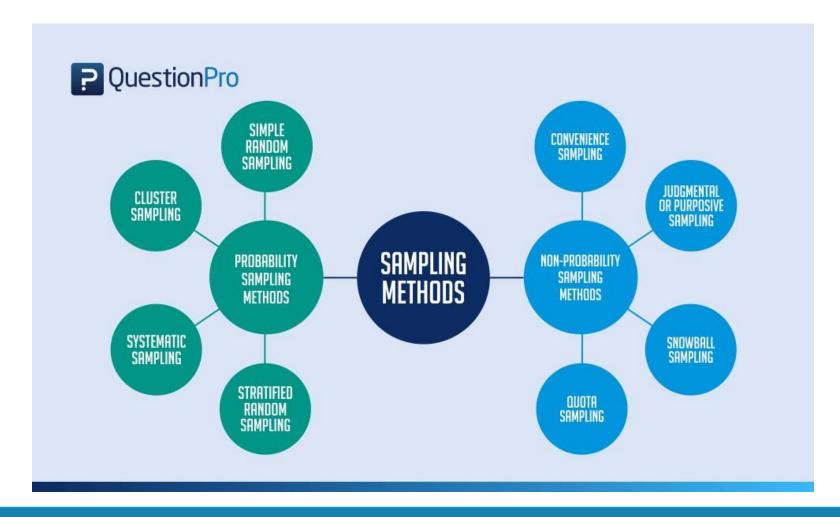
- Demand Analysis for laboratory equipment in hospitals and clinics in Bulacan
- 2. Monitoring and Control of infectious diseases in Metro Manila cities
- Forecast of volume of production of distilled water in Region III
- 4. Birth control: Health effects of birth control pills and menstruation control pills
- Levels of physical and mental health, and personal hygiene of CS students in BULSU

Which method(s) and tool(s) of data collection is(are) most appropriate?





# SAMPLING







# Sampling Design

- It depicts the approachable plans of an investigation that the researcher employs in arriving at a population sub-group to be investigated rather than the entire population elements (Kabir, 2016).
- It entails selecting elements of the investigation which the researcher asks questions from and in which such information is used for generalization on the whole population of the study.





# Basic Concepts in Sampling

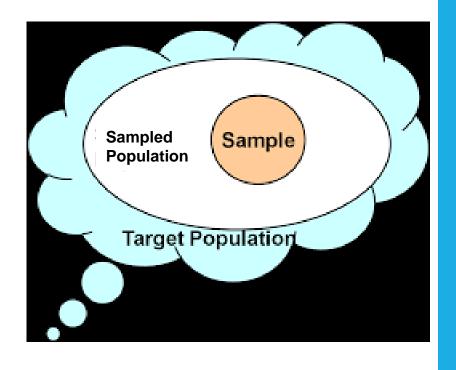
#### **Target versus Sampled Population**

#### Target Population

 the population from which information is desired for study

#### Sampled Population

• the population from which sample is actually taken based on the sampling frame.



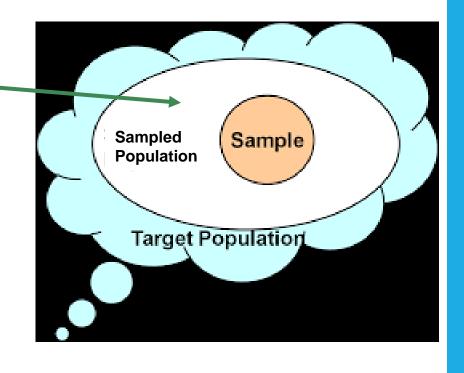




# Basic Concepts in Sampling

#### **Sampling Frame**

- list or mapping of all the elements or sampling units in the sampled population (items, persons, households, etc.)
- □ The frame may not include all the elements in the population because of missing elements due to problem of updating, clustering and blanks. It is also possible that there is duplicate listing







# Example: List

Excerptonly

3,982 Passers in August 2023 **Medical Technologist Licensure Exam** 

#### Complete list of passers in Food Technologist Licensure Exam (August 2023)







5 AGCAOILI, DAVE ANTONIO 6 AGUINALDO, JANET MUÑIZ

7 AGUIRRE, MARIA ALEXANDRA SEGARRA

8 AGUSTIN, PHOEBE MAUREEN TAMAYO

9 AGUSTIN, RENZO REY MENDOZA

10 ALAMBAN, TIMMY ANGELINA LLACUNA

11 ALARCON, JOSHUA CASONA

12 ALBETIA, JEAN MARGARET MATULLANO

13 ALCANTARA, LEVI-JUSTINE BAFUL

14 ALEGRE, MICKEY DANIELLE QUESADA

15 ALMONTERO, MARCO JAY MAGLALANG

16 ALMOSARA, TONI ABEGAIL ENRIQUEZ

17 ALTERADO, DRIEZEL BALINGIT

18 ALTURAS, CHRISTINE LOUANNE BUAN

19 ALVARADO, ROXANNE MARIE BUAYABAN

20 ANAPI, GERIEKA RAMOS

21 ANDALION, RENIEL REYES

22 ANDARZA, MARY ANTONETTE GARRINO

23 ANDRES, ELLAIZA RAE DIZON







# Example: List



49 manufacturing companies in Bulacan

#### Moscel Metal Trading

Trading company involved in small scale mining for various metal and non-metal ores in the Philippines.

Address: Blk 38 Lot 4 Heritage Homes Phase 1, Loma de Gato Marilao, Bulacan Business type: Trading Company

#### Chateau Valenzuela, Condo-Ville

The first class "Condo-Ville" Project of Pag-IBIG FundPag-IBIG Fund and Globe Asiatique Realty has once again come together to give life the idea "luxurious living within our each". We present you...CHATEAU VALENZUELA, ...

Address: 16P. Gregorio St., Brgy Lingunan Valenzuela, Bulacan Business type: Others

#### Hope Mart International Trading Co., Ltd.

We are Hope Mart International Trading Company Limited, with offices in Bulacan, Kalaw and China (Shanghai). We export cavendish banana class-a.

Address: 1888 Balantok St, Bo., Pantoc Meycauyan Bulacan 3020 Meycauyan, Bulacan Business type: Trading Company

#### Agate Enterprises

We collect flint, amber and green cullets. Reprocess/clean them according to glass plants specifications of allowable contaminants.

Address: 181 Gov. F. Halili ave. Turo Bocaue, Bulacan Business type: Trading Company

#### AMB Cyber Cafe

PLEASE GIVE ME A RESETER ...

Address: MEYCAUAYAN BULACAN BULACAN, BULACAN Business type: Others

#### KY plastic trading

We are Ky plastic trading, manufacturer of pellets for 4 years now. specializes in PE, PP, OPP, HD. also buy and sell scraps, pellets. produces 15tons a day, trading locally wanted to be more expose internationally, need long term ...

Address: B4 Greenmiles compound, sta. rosa 1, marilao bulacan, bulacan Business type: Manufacturer

#### San Mig's Sweets

supplier of low-price, high quality pastillas and other sweets including sitcharon( cracked pork skin)from san miguel, bulacan philippines.





# Example: Map









### **EXERCISE**

San Miguel Foods, a division of San Miguel Food and Beverage, Inc., is comprised of leading food companies in the Philippines. Its products and services span across the entire value chain—from animal feeds, fresh chicken, and fresh meats to processed meats, ready-to-eat food, flour, dairy, spreads, and coffee—making it undoubtedly the most diversified food player in the country today. As an annual activity, they need to know the satisfaction level of consumers in Bulacan, Batangas and Misamis Oriental. A sample survey was implemented using PSA masterlist of households in Bulacan for the year 2024. Results estimated that about 90% of the respondents are satisfied with the safety and quality of foods produced. Specify the following:

- a) Target population
- b) Sampled population
- c) Sampling frame
- d) Estimate/Statistic





# Basic Concepts in Sampling

# Methods of Sampling:

Characteristics	Probability Sampling	Non-Probability Sampling
Statistical inference	Valid	Not Valid
Probabilities of selection	Known	Unknown
Obtained through	Objective chance mechanisms	purposively, voluntarily
Need frame	Yes	No
Research design	Usually quantitative or mixed method	Often qualitative or exploratory





# Basic Concepts in Sampling

# Cases Wherein Nonprobability Sampling is Useful

- Only few are willing to be interviewed
- □Extreme difficulties or more expensive in locating or identifying subjects
- □You may want to gain the views of only a niche or targeted set of people, based on their location or characteristics.
- If there is a target market that you want to enter, it may be worthwhile doing a small pilot or exploratory research to see if new products and services are feasible to launch.





# **Purposive Sampling**

 sets out to make the sample agree with the population in regard to certain characteristics

## **Quota Sampling**

a specific number of particular type of elements are selected

# Convenience Sampling

chooses units which come to hand or are convenient to choose.
 Snowballing or friendship pyramiding are parts of convenient sampling.

## **Judgment Sampling**

selects sample in accordance with an expert's judgment





# Simple Random Sampling



- basic method of drawing a probability sample
- a scheme of selecting a sample of size n from a population of size N such that each of the distinct possible samples has equal chance of being drawn
- appropriate when the sampling units are quite homogeneous
- This is through Draw-lots method, random number generation and Table of random numbers
- Inclusion Probability:  $\frac{n}{N}$



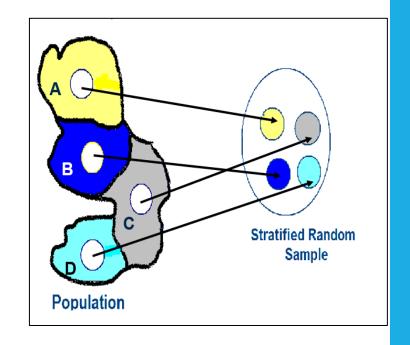


# Stratified Random Sampling

- » The target universe is grouped into so-called strata.
- » For each stratum, independent simple random samples are selected.
  - Allocation can be equal or proportional.
- » Different sampling method can be used for each stratum
- **»** Inclusion Probability :  $\frac{n_h}{N_h}$

#### **Reasons for stratifying:**

- 1. Better precision of estimates
- 2. Administrative convenience
- 3. Objective of the survey
- 4. Nature of the target population







# Stratified Random Sampling Notes:

- Stratified random sampling is stratified sampling wherein SRS is used in the selections of sample from each stratum
- 2. Inclusion probability:  $\underline{n_h}$





# Systematic Random Sampling

adopts skipping pattern in the selection of the sample units from the sampled population or stratum

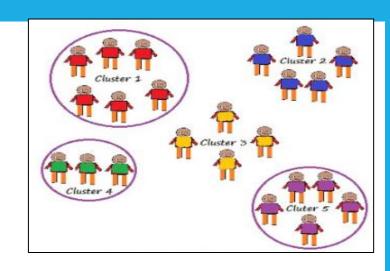
$$k = \frac{N}{n}$$

- first unit in the sample is selected at random and the rest are selected at regular interval k.
- Random start is any number from 1 to k-1





# Cluster Sampling



- Not an element sampling but sampling unit is a cluster of elements which is heterogenous in nature
- all elements within each selected cluster are enumerated
- Advantage: simpler frame and simplifies survey operation
- Disadvantage: less efficient than element sampling





# **Multi-Stage Sampling**

- involves sample selection done in stages
- use for large-scale surveys
- more complicated form of estimators

Primary sampling units PSU's:

sampling units at the first stage

Secondary sampling units SSU's:

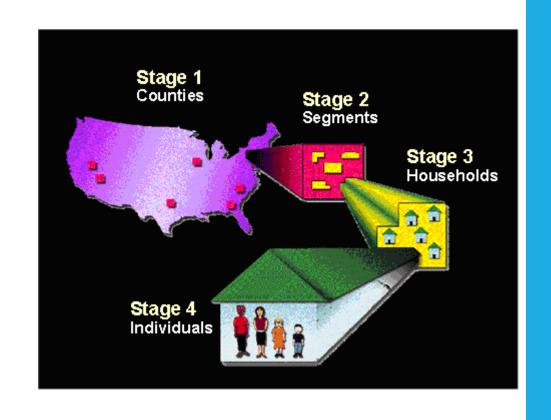
sampling units at the second stage

Tertiary sampling units TSU's:

sampling units at the third stage

Ultimate sampling units USU's:

sampling units at the last stage







# SAMPLE SIZE DETERMINATION

#### Sample Size Considerations:

- 1. The larger the population, the more samples must be taken.
- 2. The more variable or heterogenous the population, the more samples must be taken.
- 3. Cost
- 4. **Complexity** of planned analysis

#### **Approaches in Sample Size Determination:**

- 1. Sampling fraction/rates: Example: Drawing a 10% or 30% sample
- 2. Subjective approach: Example: Drawing a sample of at least 25 or 30
- 3. Precision point of view (margin of error).
- 4. Consider effect size and power





#### For Simple Random Sampling:

Using the formula,

$$n = \frac{\frac{Z^2 \sigma^2}{\sigma^2}}{1 + \frac{1}{N} \frac{Z^2 \sigma^2}{\sigma^2}}$$

where: N= population size d=margin of error  $\sigma$  = standard deviation Z = standard normal variate

#### **Notes:**

- ightharpoonup 99% confidence → Z=2.575
- ightharpoonup 95% confidence → Z=1.96
- ightharpoonup 90% confidence ightharpoonup Z=1.645





#### For Simple Random Sampling:

Using this formula,

$$n = \frac{\frac{Z^2 PQ}{d^2}}{1 + \frac{1}{N} \left( \frac{Z^2 PQ}{d^2} - 1 \right)}$$

Where: N = population size

d = margin of error

P = perceived value of the proportion

Q = 1-P

#### **Notes:**

- ightharpoonup 99% confidence  $\rightarrow$  Z=2.575
- ightharpoonup 95% confidence  $\rightarrow$  Z=1.96
- $\triangleright$  90% confidence  $\rightarrow$  Z=1.645





If again, the percentage is not given, use the highest value of the product of equivalent proportion of P and Q, that is:

Р			
Percentage	Proportion	Q= 1-P	PxQ
10%	.10	.90	0.09
20%	.20	.80	0.16
30%	.30	.70	0.21
40%	.40	.60	0.24
50%	.50	.50	0.25
60%	.60	.40	0.24
70%	.70	.30	0.21
80%	.80	.20	0.16
90%	.90	.10	0.09

Highest product of all possible combination of P and Q is 0.25





#### Let's now compute:

• Formula:

$$n = \frac{\frac{Z^2 PQ}{d^2}}{1 + \frac{1}{N} \left( \frac{Z^2 PQ}{d^2} - 1 \right)}$$



- Let's compute first  $\frac{Z^2PQ}{d^2}$
- - Compute the Z<sup>2</sup> then multiplied by the product of P and Q
- - The product of Z<sup>2</sup>PQ is divided by d<sup>2</sup>
- Then compute for  $\frac{Z^2PQ}{d^2}$  1
- Next multiply this to (1 divided by your population size). Then add to 1
- Finally, divide  $\frac{Z^2PQ}{d^2}$  by the denominator  $1 + \frac{1}{N} \left( \frac{Z^2PQ}{d^2} 1 \right)$





Sample Size Determination in Stratified Random Sampling:

**Approach 1: Separate sample size** determination for **each stratum** and **then combine** them to get the sample size.

- » via design effect
- measure of the **precision gained or lost** of the use of **more complex design instead an SRS**.
- 100% gain in efficiency when stratified sampling is used instead of SRS when deff is set at 0.5

 $n = deff \times sample size using SRS$ 

**Approach 2:** Aggregate sample size *n* and then **allocate** this to different strata





Sample Size Determination in Systematic Random Sampling:

This also applies the design effect, hence:

$$n_{sys} = n_{srs} \times .50$$





# Sample Size From G\*Power

- G\*power is a free statistical software that allows the user to determine statistical power based on a wide variety of tests.
- Here, the user can specify the type of test being run, their desired level of power, and alpha level to determine the sample size needed.
- In statistics, power (1- $\beta$ ) or the ability to detect an effect if there is one present, is an important tool used to reduce the chances of a type II error ( $\beta$  or the error when you fail to reject  $H_o$  when in fact it is false).
- Power can be thought of as sensitivity, meaning that the more power that is present, the more likely we are to be able to detect an effect if it is there.
- The common power level is 0.80







# Sample Size From G\*Power

- Effect size how meaningful the relationship between variables or the difference between groups is. It indicates the practical significance of a research outcome.
- A large effect size means that a research finding has practical significance, while a small effect size indicates limited practical applications.

Effect Size	Analysis
Cohen's d	Independent samples t test
Odds ratio (OR)	Logistic regression
Pearson correlation (r)	Correlation analysis
R squared (R2)	Multiple regression
Omega squared (ω²)	ANOVA

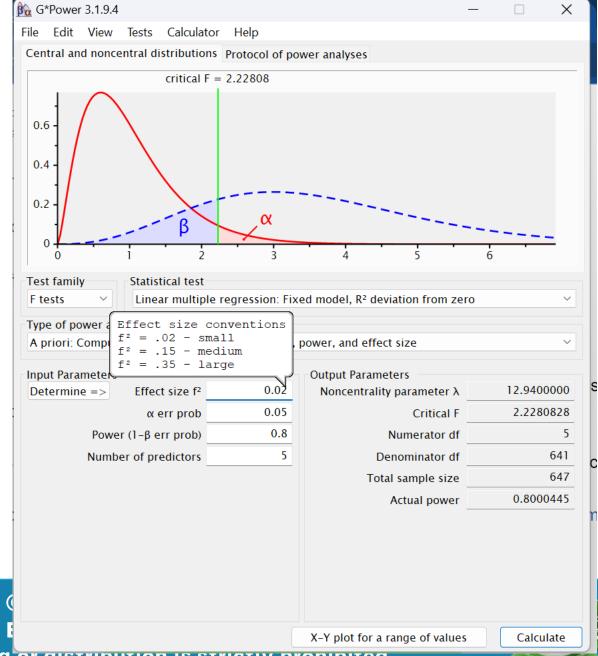
TYPE OF STATISTIC	EFFECT SIZE
t-test	cohen's d
ANOVA	eta <sup>2</sup>
chi square	phi coefficient
correlations	pearson's r
multiple regression	R <sup>2</sup>
logistic regression	analogous to R <sup>2</sup> (pseudo R <sup>2</sup> )







# Sample Size From G\*Power





**BULACAN STATE UNIVERSITY** 

For the exclusive use of Officially I

Unauthorized use, reproduction, sharing or distribution is strictly pronibited.

# Sample Selection

# Simple Random Sampling



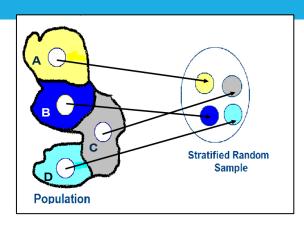
- After computing for the sample size, perform simplest method, that is the draw-lots
- Write down names of sampling units into a piece of paper.
- Place in a box and draw piece of paper based on the number of samples.
- Alternatively, number is written on the piece of paper. The set of numbers corresponds to the numbered names in the sampling frame.





# Sample Selection

# **Stratified Random Sampling**



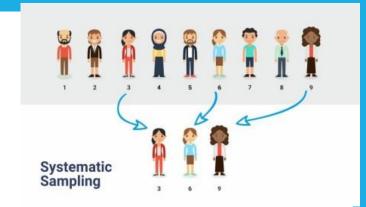
- Identify the grouping of the sampled population.
- Within each stratum, perform either simple random sampling or systematic random sampling selection





# Sample Selection

# Systematic Random Sampling



- Compute for the skipping pattern which is the ratio of population and sample sizes or k = N/n
- Determine the random start from 1 to k-1
- Using the sampling frame, start the counting from the random start until k. First selection is the first sampling unit. Count again using k, where second selection is the second sampling unit, and so on until the sample size is reached.





# References

- Philippine Statistical Research and Training Institute (PSRTI) Training Module on Research and Basic Statistics
- 2. Albacea and Reyes (2011): STAT 173 Survey Operations: A Handbook of Slide Presentation
- 3. Albacea and Reaño (2010) Elementary Statistics: A Handbook of Slide Presentation, Institute of Statistics, University of the Philippines Los Baños
- 4. https://www.niu.edu/rcrportal/datamanagement/dctopic.html
- 5. Lohr (1999) Sampling: Design and Analysis
- 6. Lucero, Jeffrey A. (Published on Aug 26, 2018). Education. Retrieved from: https://www.slideshare.net/mayamangjeff/quantitative-research-design-111
- Menoy, Jesus, Millano, Karen and Tuason, Juan Martin (2018). <u>Practical Research 1 for Senior High School</u> (Focus on Qualitative Research)
- 8. Prieto, Nelia, Naval, Victoria, and Carey, Teresita (2017). <u>Practical Research 1 for Senior High School Qualitative</u>
- 9. https://www.statisticshowto.com/sampling-design/
- 10. https://www.qualtrics.com/experience-management/research/non-probability-sampling/#:~:text=Non%2Dprobability%20sampling%20is%20typically,on%20their%20location%20or%20ch aracteristics.
- 11. https://www.statisticssolutions.com/how-to-determine-sample-size-from-gpower/



# Thank you

Ma'am Meann Magtulis
Assistant Professor III
Bulacan State University (Main)
maryann.magtulis@bulsu.edu.ph



