BUSINESS STRATEGY
SURVEY PLAN IN HONG KONG & MACAU

SURVEY OBJECTIVE

The survey aims at collecting information regarding the financial and banking habits of undergraduate students in Hong Kong and Macau. The data collected will be applied to assess the students' satisfaction levels of their current banking services and discover new banking products that have enormous potential for debut in Hong Kong and Macau marketplace.

SURVEY DESIGN

The sampling method chosen is Two Stages with Primary Units Selected by Probability Proportional to Size and Secondary Units Selected with S.R.S. based on the survey objective. In our survey, universities in Hong Kong and Macau are set for the primary units and the undergraduate students studying in the institutions are viewed as the secondary units.

Firstly, a probability sample of universities will be randomly chosen. institutions are to be selected, therefore random numbers between 1 and the population 118,803 will be chosen from the random number table. By locating the drawn numbers in the cumulative range column, the selection process of primary units completes. If there is primary unit selected more than once, then an independent drawing of same amount of secondary units will be made without replacement from the complete primary unit instead each time. Invitations for conducting the survey will be sent to the research units of the chosen universities and then take simple random samples within each selected university. It is planned that the research units will assist in the sample selection of secondary units and survey conduction. After a simple random sampling of same amount of people in each selected primary unit, the survey will take place at Google Form with a variety of question types. A sample of questionnaire and its output are attached in the Appendix. Interviewers should approach the respondents in a rigorous and ethical manner. Meanwhile, certain pieces of important information must be clearly recorded. For example, the details of contacting potential respondents, the final number of respondents, the characteristics of people refuse to participate in the research, the administration of survey and the response rate. (Kelley et al., 2003) Data collected will be sent to us for further data analysis.

Justification

At first, Probability sampling is introduced to the sampling process. The classical formulation of parameter estimations requires randomness built into the sampling design. Therefore, the property of the estimators such as sampling error and bias can be assessed probabilistically in a systematic way. (Scheaffer, R., 2012)

It is essential to provide definition involved in the sampling process. The population is clearly defined as the undergraduate students in Hong Kong and Macau. In two-layer design, clusters or primary units are the universities in Hong Kong and Macau while secondary unit is the undergraduate students within their primary unit. Sampling unit is an individual undergraduate student. The frame is the undergraduate student lists of the universities. It is assumed that a student cannot study in two university in Hong Kong and Macau at the same time.

The unique attribute combinations such as remaining years of study and financial knowledge are observed among the elements, it is expected that the measurements between universities is small, but the measurements between students is large. It means that the clusters are similar to another and it is representative to sample a few clusters for reflection of the whole population. According to the data of student headcounts by

Universities, the clusters are large and differ considerably in their size. It is suggested to apply probability proportion to size which is a sampling procedure under which every primary unit of the population has chance proportional to its own size of the ultimate unit (population). (Singh, Chaudhary, & Chaudhary, F. S., 1986) So as to ensure all elements have same probability to be selected, ignore the size of their primary unit. (Lohr, S.,1999)

In general, one-stage cluster sampling, all elements in the selected clusters are supposed to be survey. According to the figures of student headcount in Hong Kong and Macau, there are more than 118800 undergraduate students in academic year 2018/19. It is barely able to afford a census in Hong Kong and Macau and the costs of making observation and travel are expected to increase as the distance between elements increase. Therefore, randomly selecting elements in the cluster is applied as alternatives to solve the problem. It is also possible to conduct analysis at the two levels of data aggregation, including the analysis between students by using data describing individual students and institutions by using data based on the sample parameters. (Ross, K., 2005)

Access to the student lists from the tertiary institutions which is the frame of this sampling is restrictive to the unauthorized party. The suggested sampling process requires participation of universities' research unit because of the restrictive use of the student lists. The researchers can also share the roles of coordinators and data managers. This arrangement is used out of necessity, for reasons of economy. (Ross, K., 2005)

The reason of choosing Google Form as the survey platform is the compromise between costs and special features. From a cost standpoint, Google Form is significantly cheaper than Qualtrics which is a professional survey software launched by SAP. The base package for Qualtrics is around $5,000, while Google provides enterprise plan starting at $25/month. Qualtrics provides high-quality and actionable data than Google Form with not only a variety of question types, reporting graphs, but also security services and cloud for data storage. Meanwhile, some of the graphs are also available in Google Form. Basically, the survey aims at collecting information from the students, a series of statistical analysis will be conducted by us with controlled sampling error. It is more economic to choose Google Form as a data collection platform.

QUESTIONNAIRE

The full version of the questionnaire is attached to the Appendix. The draft is also attached to reveal the changes done in pre-test. Below I will investigate the flow of the questionnaire and the intention behind the questions and explain how the collected data will be analyzed.

- **Identity confirmation:** It aims at confirming that the respondent is eligible for the survey and getting information of the location of his/her institution.

- **Personal Particular (Part 1):** Students studying in Hong Kong and students studying in Macau will be directed to different sites for same questions. This section consists of questions about the institution he/she currently studying at and his/her experience in banking services. If the respondent does not have any experience in banking services, the following sections will be skipped, and the respondent will be directed to the last section to enquire the his/her interest in banking products.

- **Experience in Banking Service (Part 1):** This section asks a question about the bank that the respondent is currently using. Students studying in Hong Kong and students studying in Macau will have different response sets in accordance with the list of licensed banks in Hong Kong and Macau.

- **Experience in Banking Service (Part 2):** This section aims at getting information of respondents' banking habit and time length of using bank services digitally and physically.

- **Satisfaction level of banking services:** This section gathers respondents' attitudes towards the general experience in banking services and towards specific products. Likert scales are applied to seek the underlying relationship between the satisfaction level of the products and the intention of switching bank to another and reveal the essential factors of achieving satisfaction in banking services.

- **Opinion on banking products and financial services:** This section attempts to collect respondents' opinion on the service form of banking service (Digital or physical) and their preference on sharing customer data with the bank and third parties for personalized services. It also asks the respondents about the trust in values of banking technology and products that they look forward to discovering in the marketplace. This section ends with an open question, it intends to collect opinion on banking services and suggestion for future improvement.

- **Personal Particular (Part 2):** This section collects the basic information of the respondent including year of study and sex.

DATA ANALYSIS

Supplemental statistics

Statistic which is collected from the government officials can serve as auxiliary purpose. For example, the University Grants Committee provides the figures of Student Enrolment of UGC-funded Programmes by University, Level of Study and Sex. It can be treated as a close estimation of the population size because the enrolment number barely changes every year. The summary table of the figures is attached to the Appendix.

Descriptive statistic

Generally, the questionnaire consists of 5 types of questions including dichotomous, multiple choices, open questions and 5-point Likert scale. Example charts can be viewed in the Appendix section.

Quantitative questions (Applicable to Q8 – Q11)

Notation

- N : number of primary units in the population
- M_i: number of secondary units in the ith primary unit
- $y_i = \sum_{j=1}^{M_i} y_{ij}$
- $Population\ total: \tau = \sum_{i=1}^{N} \sum_{j=1}^{M_i} y_{ij}$

- $\mu = \frac{\tau}{M}$, where $M = \sum_{i=1}^{N} M_i$
- n : number of primary units selected in the first stage
- m_i: number of secondary units selected in the second stage
- $\alpha = 5\%$ for convention

These questions require the respondent to answer the frequency and time spent in banking services. The format of answers are integers. Therefore, with the Hansen-Hurwitz estimator, estimation of population mean is first to obtain the sample mean of each selected primary unit and then taking average of the sample means, $\bar{y} = \frac{\sum_{i=1}^{n} \bar{y}_i}{n} = \frac{\sum_{j=1}^{\bar{m}} y_{i,j}}{n\bar{m}}$. The estimated variance of $\bar{y}$ is $Va\hat{r}(\bar{y}) = \frac{\sum_{i=1}^{n}(\bar{y}_i - \bar{y})^2}{n(n-1)}$, the confidence interval can be obtained by the formula: $\bar{y} \pm z_{1-\frac{\alpha}{2}}\sqrt{Va\hat{r}(\bar{y})}$. (Frerichs, 2004)

Dichotomous (Applicable to Q1, Q3, Q5, Q12, Q14, Q17- Q23, Q25, Q30)

Notation

- n: number of primary units selected in the first stage
- a: the number of persons with the attribute of interest
- m: number of elements in population
- i,j: assign variables at the two levels (Primary, Secondary) of the sampling process.

- $a = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \sum_{j=1}^{\bar{m}} a_{i,j}$
- $p_i = \frac{\sum_{j=1}^{\bar{m}} a_{i,j}}{\bar{m}}$
- $\alpha = 5\%$ for convention

These questions require the respondent to answer either/ or questions. Respondents with the characteristics interested is coded as 1, otherwise 0. The estimation of population mean is first to obtain the sample mean of each selected primary unit and then taking average of the sample means, $\hat{p} = \frac{\sum_{i=1}^{n} \hat{p}_i}{n}$. The estimated variance of $\hat{p}$ is $Va\hat{r}(\hat{p}) = \frac{\sum_{i=1}^{n} (p_i - \hat{p})^2}{n(n-1)}$, the confidence interval can be obtained by the formula: $\hat{p} \pm z_{0.95}\sqrt{Va\hat{r}(\hat{p})}$. (Frerichs, 2004)

5-point Likert Scale (Applicable to Q13, Q15, Q16)

Notation

- $\bar{r}$: is the average of the $\frac{K(K-1)}{2}$ correlations between K components.

The application of Likert Scale is to realize the general opinion about the subjects of the question. The three questions attempt to discover respondents' attitude towards banking services and what factors of banking services do the respondents value. 5-point Likert Scale reveals 3 pieces of information including direction, intensity and a middle-ground position. Considering the internal consistency, Cronbach's α is used to check for consistency. $\alpha = \frac{K\bar{r}}{1+(K-1)\bar{r}}, \alpha \leq 1$.

Open Question

The questionnaire introduces examples of products have not appear or have not been prevalent in Hong Kong. The open question aims at receiving feedback on banking services in this two cities after the questionnaire. For example, with guidance of the questions, the respondents want to share their view on the banking services on different aspects and ideas of products that have not been provided or prevalent in Hong Kong in Macau. Qualitative analysis will be undergone in this part.

REQUIRED SAMPLE SIZE

Notation

fsu: first stage sampling units
n: number of primary units
m_i: number of secondary units to be drawn from primary unit i.
p_i: the probability of selecting the ith primary unit at each draw, where $\sum p_i = 1$ and $\bar{y}_i = \frac{\sum_{j}^{m} y_{ij}}{m_i}$

$S_{w_i}^2$: sample variance within primary units
C_0: the overhead cost
c_1: cost of selection of a primary unit and prepartion of a frame
c_2: cost of selection of a secondary unit and analyses

m: number of clusters selected
n_0: number of secondary selected in each selected primary cluster

The required sample size can is derived from the formula of the sampling error. Sampling error equals to the product of z-value and estimated standard deviation divided by the square root of the sample size: $z_{1-\frac{\alpha}{2}} \times \frac{Va\hat{r}(\hat{y}_{pps})}{\sqrt{n}}$. Set the desired bound of error to within 10% and $\alpha = 5\%$ for convention, by rearranging the terms in the formula, the required sample size of each question is $\left(\frac{0.1}{1.96}\right)^{-2} \Sigma_{i=1}^{n}(p_i - \hat{p})^2 = n(n-1)$. Survey in format of questionnaire depends more on the number of selected schools than on the number of students surveyed in each selected school. (Ross, K., 2005) Hedayat and Sinha (1991) provides a solution to the optimization problem, taking account of the discrete nature of the sample sizes n and m. (Thompson, S., & Ebrary, Inc., 2012) In Singh and Chaudhary (1986), there are formulas for optimal subsample sizes are available for the sampling design of this survey. Firstly, the unbiased estimator of population mean is given by the formula $\hat{y}_{pps} = \frac{1}{n}\Sigma_i^n \frac{M_i y_i}{p_i}$. It can be further discovered that the sampling variance of the estimator is

$$Var(\hat{y}_{pps}) = \frac{1}{n}\Sigma_i^N p_i \left(\frac{Y_i}{p_i} - Y\right)^2 + \frac{1}{n}\Sigma_i^N \frac{\frac{M_i^2}{p_i}(1-f_{2i})S_{wi}^2}{m_i}.$$ The unbiased estimator of $Var(\hat{y}_{pps})$ is $Va\hat{r}(\hat{y}_{pps}) =$

$\Sigma_i^n \frac{\left(\frac{M_i \bar{y}_i}{p_i} - \hat{y}\right)^2}{n(n-1)}$. It is a good estimation because the estimator is simply the combination of the variance between primary units and the variance within the primary units. With the formula of the within component $Va\hat{r}_w(\hat{y}_{pps}) = \Sigma_i^n \frac{M_i^2(1-f_{2i})}{n^2 p_i^2}\frac{S_{wi}^2}{m_i}$, the variance between primary units can be computed be subtraction from the variance estimator.

The optimal choice of sample sizes in this two-stage cluster sampling design is decided by the combination of the cost function $C_0 + c_1 m + c_2 m n_0 \leq C^*$ with constraints $2 \leq m \leq M$ and $2 \leq n_0 \leq N_0$ and the variance formula. By assumption of variance within clusters greater than variance between clusters, Hedayat and Sinha (1991) suggested that the variance can be rearranged in the form of $Var(\hat{y}_{pps}) = S_w^2 \left[\frac{1}{mn_0} + \frac{\frac{S_B^2}{S_W^2}-1}{mN_0}\right] - \frac{S_B^2}{MN_0}$. It decreases with mn_0 and increases with m. The ultimate aim is to find the smallest possible m and the largest possible n_0 for the cost function close to C^*. The search for optimum values should initiate with the least feasible value of m for which $2 \leq \frac{C^*-C_0-c_1 m}{c_2 m} \leq N_0$. (Hedayat, Sinha, & Sinha, Bikas Kumar., 1991)

It is interesting to note that for variance within clusters less than variance between clusters, the optimum choice of (m, n_0) depends on the ratio $\frac{S_B^2}{S_W^2} = \alpha$. (Hedayat, Sinha, & Sinha, Bikas Kumar., 1991) An algorithm is established for selecting the best combination of (m, n_0). For $\alpha > 1$, if (m', n_0') and (m'', n_0'') both satisfy the cost function, while $m' \geq m''$ and $n_0' \geq n_0''$, (m', n_0') is more preferred than (m'', n_0''). (Hedayat, Sinha, & Sinha, Bikas Kumar., 1991)

MINIMIZATION OF NON-SAMPLING ERRORS

Launch of the pre-test

A pre-test of the questionnaire was launched on 29th April and 30th April, successfully collected 20 replies. It aims at:

1. Estimating parameters required for determining sample size
2. Correct the typos of questions
3. Improve the questionnaire design including the ordering, descriptions of questions

The responses provided by closed questions may not be able to represent the true view of the respondents. Therefore, three open questions are added to determine possible fixed alternatives for the closed questions.

Result of the pre-test

Errors were corrected and descriptions were added to increase the understandings of the questions and responses. Some examples are presented below:

- Options edited (Q15, 16), (Q12, 19)

As two of the pre-test respondents indicated that fee is one of the significant factors while considering using service from a bank, corresponding option 'Fee (administrative fee, service charge)' is added to the question 15 & 16. The neutral positions are removed from the questionnaire as the survey focuses on the direction of the respondent. No middle ground is offer to prevent the respondents taking the option too often as an easy way out. (Scheaffer, R., 2012)

- The phrasing of options changed (Q19, 20, 21)

Question: Do you favour or oppose combining technology with banking services?
Old responses: 'Yes' & 'No'
New responses: 'Favour' & 'Oppose'.

- Question changed (Q8, Q10)

Old question: How often do you usually use Branch service?
New question: How often do you use Branch services once?

- Description of Question added (Q8, Q10, Q23)

To prevent a type of nonresponse error, the inability of responding to come up with the answer to the question, in some of the questions, supplementary information is provided to ensure that the respondents understand the questions and all the options. (Scheaffer, R., 2012)
Question: How often do you use Branch services once?
Description: Please answer the frequency of using Branch services in terms of days, example: Usually use Branch services every 5 days, write down "5".

Question Design

It is hoped that a careful planning on question design can reduce the common non-sampling errors such as the non-response rate and refusal rate.

- Choice on point of Likert Scale

Likert Scale is to reveal the attitude towards a certain subject. There are three important pieces of information including direction, intensity and a middle-ground position. In this context, 5-point scale performs way better than 3-point scale in term of indication of these information. The studies from Preston and Colman (2000) signified that 5-point scale is easier to use compared to 3-point scale, except for the quickness. A study concluded that 5 is one of the optimum number of alternatives on the reliability and validity of rating scales. (Lozano, L., García-Cueto, E., & Muñiz, J., 2008) In a compromise between quickness (prevent boredom due to a large number of questions) and reliability, some questions asking for a general view asks in either/or format, when a subject interested relates to many concepts/ values/ items, for example, attitude towards banking services, 5-point Likert Scale is applied.

- Order of questions

Sometimes the respondents would be annoyed by answering a large number of questions consecutively. In order to prevent bias introduced by the length of the question, breaks are given during doing the questionnaire. Each section consists of less than eight questions and each grid-question consists of less than fifteen items.

- Order of responses

Sometimes the respondents would be annoyed by reading a long list of responses. In order to prevent bias introduced by the length of the responses, each opinion-based question with multiple choices consist of less than four choices. Only factual questions, for example, the question asking the respondents about the school, provides a list of the schools in alphabetical order and common abbreviations used by the institutions.

- The phrasing of questions

The wordings of questions are attempted to be balanced by listing both supportive and negative phrases in the opinion-based questions and keeping the phrases in a formal and neutral tone. Moreover, the question should avoid bringing argument or counterargument as less as possible so as to prevent leading the respondents to agree with the stance of the interviewer/ questionnaire.

- Description of Question

In order to make sure that the respondents understand the context and the meaning of both the questions and answers, the necessary description is attached to some of the questions. For example, the questions 'Do you favour or oppose the following banking products or services in the market?' involves banking jargon, additional information is attached to the questions to explain the function of some products rarely be seen in the marketplace.

- Screening question and response

Generally, human tends to choose easier options, experiments prove that people tend to choose a middle-ground option if the option is included in the list of possible responses. (Scheaffer, R., 2012) The response of 'Don't know' and 'Neutral' reduces the sample size and increases the sample error. Such error has to be eliminated or minimized. Therefore, a screening question of the understanding of banking technology is asked and middle-ground options were not included in some questions in order to reduce the possibility of answering 'Don't know' or 'Neutral'.

Incentives

Lilywhite as a globally prestigious bank, if the survey explains the benefits of completing the survey, it may improve the response rate. To the respondents, the interviewer should approach them in a friendly manner to build trust between the interviewer and the respondent. Moreover, it is essential to let the respondents know how can they potentially benefit from the survey and arouse respondents' interest in the survey. (Scheaffer, R., 2012) For example, the completion of survey helps improve the future product design, they may be able to enjoy the products specialised to what they stated in the survey. If the budgets allows, cash payment can also be given to the respondents so as to improve the response rate and error rate from frivolous responses. (Scheaffer, R., 2012)

Lilywhite can avoid cumbersome administrative work and difficulty of getting a more updated frame if having assistance from the selected universities. For the university research units, other than offering budget for their work, the data collection may also benefit their research such as business strategy related to department of marketing and product design related to the department of finance. Unless the edit would make the questionnaire unable to meet the objective of the survey, it is possible to edit the questionnaire as they require for their research. Therefore, both parties can be benefited from the survey.

POTENTIAL WEAKNESS

Due to limitations from multiple aspects, the survey plan may suffer weakness. Below are the description of the potential weakness and suggestion for minimization of the errors or bias.

- **Errors of Coverage:** Student list of the universities is confidential. The problem of sampling frame not matching up with the population exists, it introduced errors which is difficult to measure into the sampling process. It is only possible to obtain a list of the universities and their classes. It is suggested to seek assistance of the frame and survey conduction from the department of universities. Therefore, the gap between the population and frame can be reduced to permit inferences to be made about the population. (Scheaffer, R., 2012).

- **Insufficient fund:** In this proposal, we are considering a two-stage cluster sampling. It requires cooperation of the tertiary institutions. Partnership with them in a project may require funding and budget, while this cost is unclear. Further confirmation of cost should be undergone. If the total cost exceeds the budget, an alternative sampling method should be considered, for example systematic sampling which requires less information of the population.

- **Insufficient power of pre-test:** The pre-test surveyed 20 respondents while 17 of them are from CUHK which is one of the clusters. The insufficient data may lead to bad estimates of the parameters and unsuccessful estimation of required sample size. Also, the respondents of pilot samples were invited on

online platform, it may harm the randomness of sampling and thus lead to an unrepresentative result. It is suggested that a pre-test in a properly larger scale and from more than 2 clusters should be conducted to increase the accuracy and precision of the estimation.

- **Too many response options:** Regarding to the questions about the factors of banking service, the satisfaction level of the products, there are more than 10 sub-questions in each question. Risk of dropping out lies ahead if the survey becomes excessively long or irrelevant to the respondents. (Scheaffer, R., 2012). In consideration of this problem, some uninterested items such as products that Lilywhite does not plan to develop can be eliminated from the response option. Current solution is to arrange breaks between the exhaustive questions and reduce some general questions from Likert Scale to either/or questions so as to reduce stress of doing the questionnaire.

- **Sensitive question:** The questionnaire concerns about the financial habit of the respondents, it may raise fear related to leakage of sensitive data. The respondents may refuse to answer survey questions or give a dishonest answer and thus the response error deteriorates. The questionnaire has a preface statement to assure them a strict confidentiality of personal data management and clearly stated the objective of the survey so as to alert them in advance. As Scheaffer (2012) wrote, the high refusal rates occur overall in ethnic and salary groups, the questionnaire avoids questions about sensitive data such as ID no., name, the income and the amount of investment in the bank.

FINANCIAL PLAN

Creating a financial plan is where all of the survey planning comes together. Up to this point you have identified the target market and target population, along with pricing. These items along with the assumptions, will help estimate what expenses expect to incur. This is important on an ongoing basis to see when you are efficient and statistically sufficient.

At a minimum, this section should include the estimated start-up costs, along with a summary of the assumptions you are making with these projections.

Here is how to set up a financial plan:

1. **Start-Up Costs:** Until the selection of primary units completes, no expenses is expected to incur. To partner with research units, it is necessary to reserve a lump sum for administrative cost and cash incentives for the respondents. Moreover, for the sake of internet security and technical support, it is recommended to subscribe the Enterprise plan provided by Google, depend on the project duration.

2. **Actual Budget:** What are you able to spend? If you have a little extra in your budget, this plan can include equipment or staffing that you would like to have.

3. **Choose experts, financial and administrative setup:** There are three people who will be invaluable in guiding through survey set up: a practice start-up consultant, a project coordinator and an interviewer. These experts will help:

   a. Set up, plan and help execute the survey

   b. Hire and train interviewers

The financial plan consists of the necessary items, name and budget in a budget plan, the other three columns for tracking the actual expenses.

COST OF SURVEY

COST ITEMS	DATE DUE	BUDGET	ACTUAL	REMARK
ADMINISTRATIVE/GENERAL				
Software (General)		$300		Subscription fee of Google Form
Miscellaneous		$500		Reserve
LABOR EXPENSES				
Payroll		$47,000		Budget for the survey conduction and project consultation in the selected institutions
Training		$500		
Travel		$500		Travel between the institutions in Hong Kong and Macau
Miscellaneous (Contingency reserve/launch party/etc.)		$500		Reserve for
ESTIMATED BUDGET		**$49,300**		

Should you have enquiries, please do not hesitate to contact me via email 1155077960@link.cuhk.edu.hk. Thank you for reading the proposal.

END

Supplemental Statistic The model below shows necessary figures in statistical analysis. It can be utilized as population of specific groups. The top portion of the table shows projected the student headcounts by broad academic programme category. The next section itemizes the headcounts by schools in Hong Kong and Macau. These should be consistent with the figures in the prior section.

Student Enrolment in Academic Year 2018/19			
BROAD ACADEMIC PROGRAMME CATEGORY	HONG KONG	MACAU	TOTAL
Medicine, Dentistry and Health	11168	2295	13463
Sciences	17078	2040	19118
Engineering and Technology	19747	1108	20855
Business and Management	18063	14952	33015
Social Sciences	13498	6732	20230
Arts and Humanities	12623	5078	17701
Education	5903	2006	7909
Total	98079	24995	123074

SCHOOL IN HONG KONG			
	MALE	FEMALE	TOTAL
City University of Hong Kong (CityU)	6421	6343	12764
Hong Kong Baptist University (BU)	2491	4495	6986
Lingnan University (LU)	899	1636	2535
The Chinese University of Hong Kong (CUHK)	7983	9298	17281
The Education University of Hong Kong (EduHK)	1187	2539	3726
The Hong Kong Polytechnic University (PolyU)	7211	7552	14763
The Hong Kong University of Science and Technology (HKUST)	5844	3959	9803
The University of Hong Kong (HKU)	7962	8974	16666
Total	39728	44796	84524

SCHOOL IN MACAU			
	MALE	FEMALE	TOTAL
Academy of Public Security Forces	43	3	46
City University of Macau	2824	2733	5557
Kiang Wu Nursing College of Macau	64	363	427
Macau Institute of Management	568	1067	255
Macao Institute for Tourism Studies	568	1067	1635
Macau Millennium College	76	67	143
Macau Polytechnic Institute	1382	1887	3269
Macau University of Science and Technology	4926	6379	11305
University of Macau	4526	5888	10414
University of Saint Joseph	438	790	1228
Total	14928	19351	34279

REFERENCE

1. Frerichs, R. R. (2004). Cluster sampling. Retrieved May 3, 2020, from
   http://www.ph.ucla.edu/epi/41808/418cmat/chap5rapid_1.pdf

2. Kate, K., Belinda, C., Vivienne B., & John, S. (2003). International Journal for Quality in Health Care, Volume 15, Issue 3, May 2003, Pages 261–266. Singh, Chaudhary, & Chaudhary, F. S. (1986). Theory and analysis of sample survey designs. New York: Wiley.

3. Hedayat, Sinha, & Sinha, Bikas Kumar. (1991). Design and inference in finite population sampling (Wiley series in probability and mathematical statistics. Probability and mathematical statistics). New York: Wiley.

4. Lohr, S. (1999). Sampling : Design and analysis. Pacific Grove, CA: Duxbury Press.

5. Lozano, L., García-Cueto, E., & Muñiz, J. (2008). Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 4(2), 73-79.

6. Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. Acta Psychologica, 104(1), 1-15.

7. Ross, K. (2005). Sample design for educational survey research: Module 3. Retrieved May 3, 2020, from https://unesdoc.unesco.org/ark:/48223/pf0000214550

8. Scheaffer, R. (2012). Survey sampling (7th ed., int'l ed.). Australia: Brooks/Cole, Cengage Learning.

9. Singh, Chaudhary, & Chaudhary, F. S. (1986). Theory and analysis of sample survey designs. New York: Wiley.

10. Thompson, S., & Ebrary, Inc. (2012). Sampling (3rd ed., Wiley series in probability and statistics). Hoboken, N.J.: Wiley.