

```
# project owner: https://www.linkedin.com/in/mrtnz/
# → data driven research fundamental capstone course (requirement)
# → under dap döst project sparta data scientist pathway
```

CASE OVERVIEW

Smarta, a local telecommunications company, is considering a marketing agency for a door-to-door sampling campaign to distribute free prepaid SIM cards. This initiative aims to expand its subscriber base and generate future revenue as more SIM cards are activated. Ultimately, the project seeks to boost subscriber numbers and achieve net profit.

SAMPLING STRATEGY

The marketing agency proposes a door-to-door sampling system to distribute free prepaid SIM cards. The initiative will be executed by a team of ten samplers each day in a designated barangay. Key details include:

- The samplers will be paid minimum wage by the agency and will work from 08:00 AM to 05:00 PM, with a one-hour lunch break from 12:00 PM to 01:00 PM. The team assembles in BGC at 08:00 AM and returns by 05:00 PM.
- Upon arrival, samplers will knock on doors, covering various streets for maximum reach. If interested, they will provide one SIM card and request a signature. Each household may receive only one SIM card.
- The interaction involves an average wait of 20 seconds for the door to be answered, followed by the spiel and questions, and about 10 seconds for signing. If there's no response after 30 seconds or if the household is not interested, the sampler moves on. Samplers walk at an average speed of 4 km/h, with an assumption that 10% of houses will not answer.
- The agency reports a 75% success rate, meaning about 7.5 out of every 10 responding households will accept the SIM card. Of the SIM cards distributed, 25% is activated, generating an average annual revenue of PHP 200.00 per activated card. Each SIM card costs PHP 40.00 to manufacture, which includes the initial PHP 50.00 credit.
- The agency proposes a PHP 5,000,000.00 contract with Smarta for 100 days of sampling, which covers the salaries of the samplers and van expenses estimated at PHP 10,000.00 per day.
- It's important to note that the contract excludes SIM card costs. Throughout the sampling project, it is assumed that there will be enough SIM cards available for all 10 samplers each day.

SECTION 0. CAPSTONE INSTRUCTIONS

Using the concepts you've learned from Data-Driven Research, develop a research paper of no more than five pages to present your analytics solution for assessing whether a door-to-door caravan project should be contracted to the Agency or managed in-house.

Ensure that your research paper has a strong business focus and a coherent narrative, with all decisions supported by published literature on analytics, statistics, arithmetic computations, visualizations, the analytics implementation framework, and logical reasoning. You may make additional assumptions as necessary and incorporate relevant datasets from the Philippine Statistics Authority (PSA).

SECTION 1. VARIABLE NAMES AND VALUES

Table 1 shows the list of variables that will be considered in the study. Moreover, Table 2 presents the baseline data provided by the contracted marketing agency. The data has been converted to align with the units in Table 1.

Table 1. List of Variables

Variable	Variable Name	Unit
t_s	Sampling time	min
t_T	One Way Travel time	min
R_U	Ratio of unresponsive houses	—
t_I	Interaction time	min house
t_{Wa}	Waiting time for an unresponsive house	min house
W	Average house width	m
R_{SM}	Sampler movement rate	m min
N_n, N_T	Sampled houses in (n) samplers, (T) total	house/s
R_H	Ratio of houses who accepted the SIM card	—
N_{Sg}	Number of SIM card given	SIM card
R_C	Ratio of houses who used the SIM card	—
N_{Sa}	Number of SIM card activated	SIM card
PM	Profit Margin	%

Table 2. Baseline Data

Variable	Value	Unit
R_U	0.10	—
t_I	0.67	min house
t_{Wa}	0.50	min house
R_{SM}	66.67	m min
R_H	0.75	—
R_C	0.25	—

SECTION 2. CRITERIA FOR SELECTING BARANGAYS

Equation 1 shows the number of houses sampled in a barangay per workday. Using seven meters as the average width of houses located in urban areas (Teoalida, 2012), the maximum number of houses that can be sampled in a barangay per workday is 6,332 houses.

With the average household size of 4.3 people per house in the National Capital Region (Census of Population and Housing, 2010), the targeted barangay should have a minimum 27,228 people to maximize the sampling time.

$$N_n = \frac{n(t_s)}{(1-R_U)t_l + R_U t_{wa} + \frac{w}{R_{SM}}} \quad \text{Equation 1}$$

$$t_s = 480 - 2 t_T \quad \text{Equation 2}$$

SECTION 3. PROFIT MARGIN AND PROJECT FEASIBILITY

Assuming that the five million investment is fixed, regardless of whether the door-to-door caravan sampling is conducted in-house or with the help of a marketing agency, the profit margin of the project will be expressed as Equation 3.

$$PM = 1 - \frac{0.2}{R_C} - \frac{25000}{(N_T)(1-R_U)(R_H)(R_C)} \quad \text{Equation 3}$$

$$N_T = \sum_{i=1}^{100} N_{ni} \quad \text{Equation 4}$$

Using values from Table 2 and Section 2, the profit margin of the project will be -3.40% at best. Based on the baseline data provided, the project will generate a loss rather than a profit.

In addition, as shown in Table 3, if the project is conducted with the help of a marketing agency, Smarta Telecommunications will be unable to utilize PHP 3.46 million, or 69.3% of its total investment.

Therefore, under the baseline conditions, the project will not be profitable, and it would be better to carry it out in-house to maximize the investment.

Table 3. Breakdown of 5M PHP Investment

Investment	5,000,000.00 PHP
(-) Van Expense (x1 at 10,000.00 PHP/day, 100 days)	1,000,000.00 PHP
(-) Salary (x10 at 537.00 PHP/day, 100 days)	537,000.00 PHP
Remainder	3,463,000.00 PHP

SECTION 4. KEY VARIABLES TO CONSIDER

Table 4 shows different trends of Equation 3 as it approaches ideal conditions. Among them, Condition 6 shows the most favorable trend, characterized by the fewest variables and the highest profit margin.

Thus, to increase the profit margin, Smarta Telecommunications needs to increase the total number of sampled houses (N_T) and the ratio of houses who used the SIM card (R_C).

Table 4. Profit margin trend as Eq. 3 approaches ideal conditions

No.	Condition	$PM = 1 - \frac{0.2}{R_C} - \frac{25000}{(N_T)(1 - R_U)(R_H)(R_C)}$
1	$R_C \rightarrow 1$	-3.40% → 74.15%
2	$R_H \rightarrow 1$	-3.40% → 2.45%
3	$N_T \rightarrow \infty$	-3.40% → 20.00%
4	$R_U \rightarrow 0$	-3.40% → -1.06%
5	$R_C \rightarrow 1 \& R_H \rightarrow 1$	-3.40% → 75.61%
6	$R_C \rightarrow 1 \& N_T \rightarrow \infty$	-3.40% → 80.00%
7	$R_C \rightarrow 1 \& R_U \rightarrow 0$	-3.40% → 74.74%
8	$R_H \rightarrow 1 \& N_T \rightarrow \infty$	-3.40% → 20.00%
9	$R_H \rightarrow 1 \& R_U \rightarrow 0$	-3.40% → 4.21%
10	$N_T \rightarrow \infty \& R_U \rightarrow 0$	-3.40% → 20.00%
11	$R_C \rightarrow 1 \& R_H \rightarrow 1 \& N_T \rightarrow \infty$	-3.40% → 80.00%
12	$R_C \rightarrow 1 \& R_H \rightarrow 1 \& R_U \rightarrow 0$	-3.40% → 76.05%
13	$R_C \rightarrow 1 \& N_T \rightarrow \infty \& R_U \rightarrow 0$	-3.40% → 80.00%
14	$R_H \rightarrow 1 \& N_T \rightarrow \infty \& R_U \rightarrow 0$	-3.40% → 20.00%
15	$R_C \rightarrow 1 \& R_H \rightarrow 1 \& N_T \rightarrow \infty \& R_U \rightarrow 0$	-3.40% → 80.00%

SECTION 5. LIMITS TO CONSIDER

As shown in Equation 1 and Equation 4, the total number of sampled houses $N_T(n, t_s)$ is a function dependent on the number of samplers and the sampling time.

Additionally, the number of samplers (n) is limited by the allocated budget for the project, which is 5 million PHP.

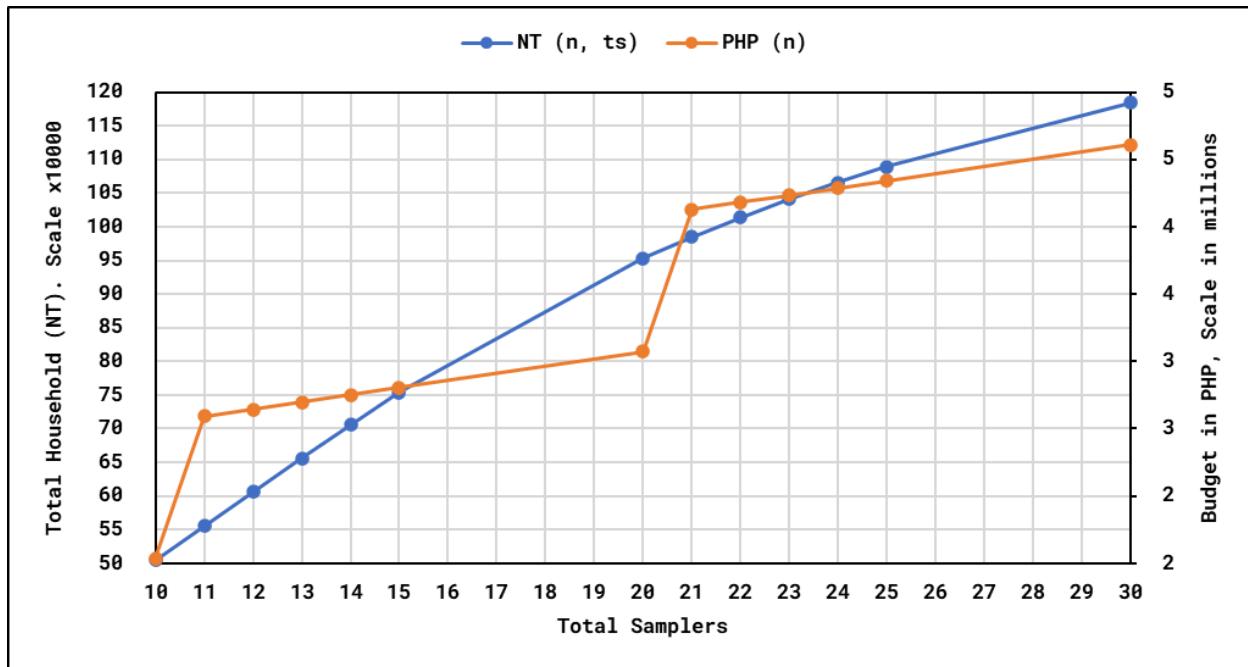


Figure 1. NT, n, and Budget Relationship Line Diagram

Figure 1 shows the relationship between the total number of households sampled (N_T), the number of samplers (n), and the budget needed for the sampling.

In the budget calculation for Figure 1, it is assumed that each van can carry 10 samplers. Moreover, Equation 5 was used to determine the maximum possible number of samplers within the 5 million PHP budget. Using Equation 5, the project can utilize a maximum of 30 samplers and 3 vans.

$$5,000,000 \geq 1,000,000 \left(\frac{n}{10} \right) + 53,700(n) \quad \text{Equation 5}$$

The total number of households sampled (N_T) was calculated using travel time data from Google Maps, Equation 1, Equation 4, the average household size (Census of Population and Housing, 2010), and data on the top 100 most populated barangays in NCR (Philippine Statistics Authority, 2015).

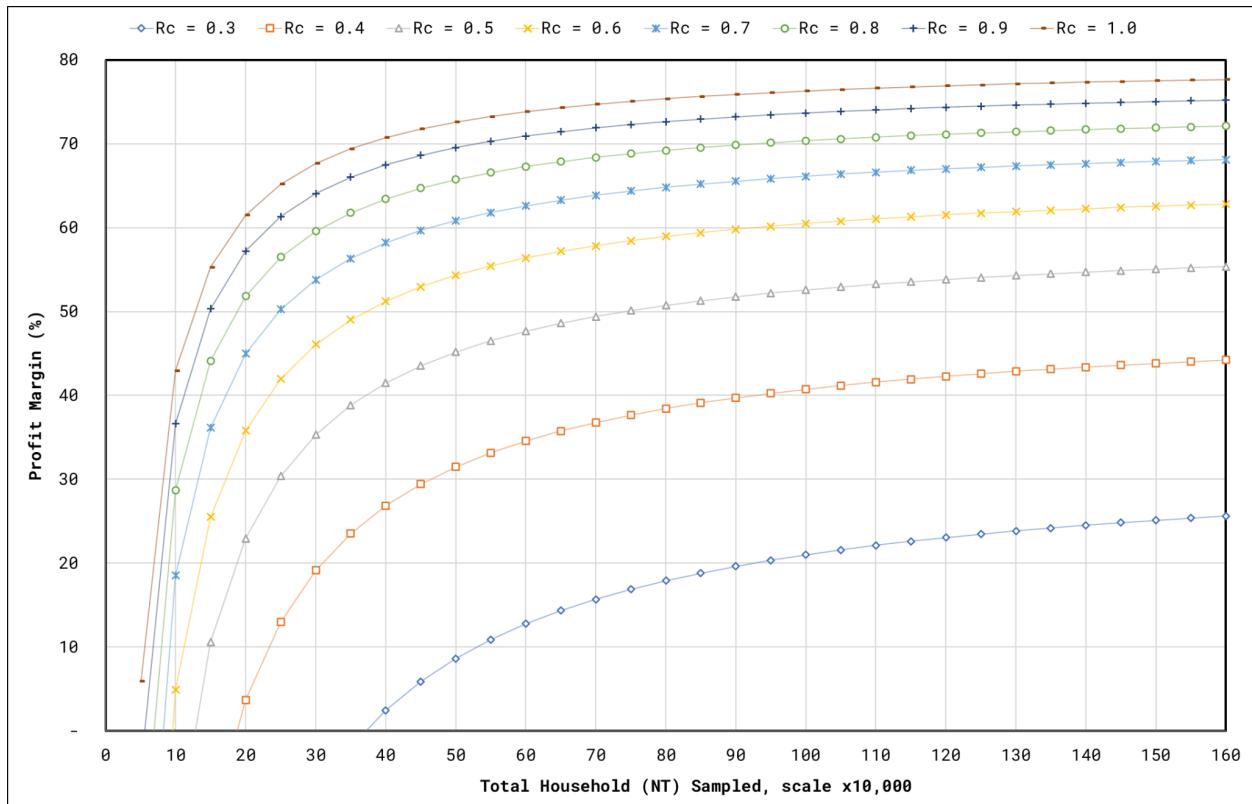


Figure 2. N_T, R_c, and PM Relationship Line Diagram

In addition, Figure 2 is the representation of Condition 4 in Table 4, which shows the relationship between the total number of sampled houses (N_T), the ratio of houses who used the SIM card (R_c), and profit margin.

Figures 1 and 2 are graphical guides essential for estimating and balancing the key variables for the project's success. For example, Figure 1 indicates that a budget of approximately 4.6 million PHP is needed to sample around 1,180,000 households with 30 samplers and three vans. Figure 2 shows that this can result in a profit margin between 23% and 77%, depending on the ratio of houses using the SIM card, which ranges from 0.3 to 1.0.

SECTION 6. RECOMMENDATIONS

An increase in both the total number of sampled houses (N_T) and the ratio of houses who used the SIM card (R_c) is essential for achieving a higher profit margin. Considering the details of the sampling strategy, between the two, it is safer to maximize the total number of sampled houses because there is not enough information provided on how the team can effectively increase the ratio of houses who used the SIM card.

With this approach, Smarta Telecommunications needs to spend 4,611,000 PHP to hire 30 samplers and three vans to sample 1,184,268 households, resulting in a profit margin of 7.49%. Additionally, when the ratio of houses using the SIM card increases above 0.25, a higher profit margin is expected.

SECTION 7. APPENDIX

Equation 1:

A. Sampling Time = Responsive Interaction + Unresponsive Interaction + Movement

$$B. t_s = (1 - R_U) (N) (t_I) + (R_U)(N) (t_{Wa}) + (N) \left(\frac{W}{R_{SM}} \right)$$

$$C. t_s = N[(1 - R_U) (t_I) + (R_U)(t_{Wa}) + \left(\frac{W}{R_{SM}} \right)]$$

$$D. N = \frac{(t_s)}{(1-R_U) t_I + R_U t_{Wa} + \frac{W}{R_{SM}}} \text{ where sampler} = 1$$

$$E. N_n = \frac{n(t_s)}{(1-R_U) t_I + R_U t_{Wa} + \frac{W}{R_{SM}}} \text{ where sampler} > 1$$

Equation 2:

A. Sampling Time = Working Time – Total Travel Time

$$B. t_s = 480 - 2 t_T$$

Equation 3:

$$A. \text{Profit Margin} = \frac{\text{Revenue} - \text{Cost}}{\text{Revenue}}$$

- Revenue = (Revenue per SIM Card)(SIM Card Activated)

- Revenue = $(200)(N_n)(1 - R_U)(R_H)(R_C)$

- Cost = (Manufacture Cost per SIM Card)(Sim Card Given) + 5000000

- Cost = $(40)(N_n)(1 - R_U)(R_H)$

$$B. PM = \frac{(200)(N_n)(1-R_U)(R_H)(R_C) - [(40)(N_n)(1-R_U)(R_H) + 5000000]}{(200)(N_n)(1-R_U)(R_H)(R_C)}$$

$$C. PM = 1 - \frac{(40)(N_n)(1-R_U)(R_H)}{(200)(N_n)(1-R_U)(R_H)(R_C)} - \frac{5000000}{(200)(N_n)(1-R_U)(R_H)(R_C)}$$

$$D. PM = 1 - \frac{0.2}{R_C} - \frac{25000}{(N_n)(1-R_U)(R_H)(R_C)}$$

$$E. PM = 1 - \frac{0.2}{R_C} - \frac{25000}{(N_T)(1-R_U)(R_H)(R_C)} \text{ where } N_T = \sum_{i=1}^{100} N_{ni}$$