



**UNIVERSITY OF THE PHILIPPINES
OPEN UNIVERSITY**

Master of Information Systems

SAMUEL PATRICK R. PASAMONTE

**SURVEY MANAGEMENT INFORMATION SYSTEM
FOR THE NOVALICHES DISTRICT HOSPITAL**

Thesis/Dissertation Adviser: ✓
Prof. Concepcion Khan,
Faculty of Information and Communication Studies

Date of Submission
12 May 2018

Permission is given for the Following people to have access to this thesis/dissertation:

Available to the general public	Yes
Available only after consultation with author/thesis/dissertation adviser	Yes
Available only to those bound by confidentiality agreement	No

Student's Signature: _____
Signature of Thesis/ Dissertation Adviser: _____

University Permission Page

"I hereby grant the University of the Philippines a non-exclusive, worldwide, royalty-free license to reproduce, publish and publicly distribute copies of this thesis or dissertation in whatever form subject to the provisions of applicable laws, the provisions of the UP IRR policy and any contractual obligations, as well as more specific permission marking on the Title Page."

"Specifically, I grant the following rights to the University:

- a) To upload a copy of the work in the theses database of the college/school/institute/ department and in any other databases available on the public internet;*
- b) To publish the work in the college/school/institute /department journal, both in print and electronic or digital format and online; and*
- c) To give open access to above-mentioned work, thus allowing "fair use" of the work in accordance with the provisions of the Intellectual Property Code of the Philippines (Republic Act No. 8293), especially for teaching, scholarly and research purposes."*

Student Name over Signature and Date

This Special Project titled
SURVEY MANAGEMENT INFORMATION SYSTEM
FOR THE NOVALICHES DISTRICT HOSPITAL

is hereby accepted by the Faculty of Information and Communication Studies in
partial fulfillment of the requirements for the Master of Information Systems.

Adviser

Date

Program Chair

Date

Dean

Faculty of Information and Communication Studies

Date

ACKNOWLEDGEMENTS

To **GOD**, who through His Infinite Wisdom and Grace gave me the ability to finish this thesis while juggling important projects at work. Thank You for blessing me and allowing me to finish this capstone project. You have blessed me more than I could ask or think. Thank You from the bottom of my heart.

To my thesis adviser, **Prof. Concepcion Khan, PhD**, who was very understanding of my thesis, from proposal to inception. Thank you. God bless your heart.

To all of my professors in the **Master of Information Systems** program, thank you so much for sharing your knowledge in the field. Thank you, for your invaluable lessons enabled me to make this program work. I would especially like to thank **Asst. Prof. Mari Anjeli Lubrica-Crisanto**, for her lessons in Web Information Systems proved essential to the success of this project.

To my professors in the UP Diliman **BS Statistics** program, I would also want to extend my gratitude, since the lessons I learned in my undergraduate statistics classes were put to great use in this thesis. I would especially like to thank **Prof. Kevin Carl Santos, PhD** for his work and his lessons in sampling were highly utilized in this project.

To **my family**, I would not have been able to do this work without you. If not for your support, I would have a harder time implementing and finalizing this work. Thank you for bearing with me when I needed to decompress; thank you also for cheering me when I got stressed and burnt out due to the combined rigors of work and study.

To **my parents**, thank you for inspiring me to continue with this. To Papa, who himself has a Master of Statistics (MoS) degree, and who inspired me to take up statistics as my bachelor's degree; and to Mama, who never ceased to encourage me to finish this thesis; thank you.

David, thanks ever so much for being a great brother, for praying for me, and for encouraging me to finish this work.

To my officemates in the **Decision Management** department of **Citibank, NA, Philippines**, thank you for inspiring me to work harder, yet understanding me when I need to lay low at work due to the rigors of my studies. I would especially like to thank my manager, **Rona Bautista**, for allowing me to do this alongside my work.

To the **Novaliches District Hospital**, specially **William Christian V. Reboton, MD**, thank you for allowing me to help you, through the implementation of this project. May this system be of great use to your institution.

To my churchmates at the **Cathedral of Praise**, especially to my connect group: Ate Aime, Ate Verna, Ate Sheng, Andhie, Lois, Kenneth, Ate Juli, Ken, Arvin, Ate Leng, Ate Mye, to my ministry pastors, **Ptr. Rafael Manalo** and **Ps. Lani Pagadora** and to our North Campus pastors: **Ps. Bheng Peralta** and **Ps. Mel Mercano**, thank you for the prayers.

To my friends from the **UP Campus Crusade for Christ**: **Razel**, Shai, Olga, Micah, Carlo Bantola, and my ministry partner, Ptr. Hoshea Felix, I really appreciate the support you've been giving me. Thank you so much. God bless your hearts.

ABSTRACT

The Survey Statistics Information System is intended for use by the research and administrative staff of the Novaliches District Hospital to streamline their patient satisfaction survey process. Since the current process is, depending on department, either to get all patients, or to use convenience sampling, there would either be too much work (if all patients were surveyed), or a possibility of bias (since not all patients have an equal chance of selection). The objective of this information system is to move surveys towards a probability sampling, which saves time and effort on the staff implementing the survey. The information system also provides automated statistical reporting with visualizations, which translates into actionable insights on the part of hospital administration, without needing to hire trained statisticians to interpret statistical reports.

Keywords: survey, sampling, statistics, hospital, patient survey, Novaliches District Hospital

TABLE OF CONTENTS

I. THE PROBLEM DOMAIN	1
A. Statement of the Problem	1
B. Background and Objectives of the Project	1
C. Significance and Scope of the Project	3
D. Documentation of Existence and Seriousness of the Problem	4
II. REVIEW OF EXISTING ALTERNATIVES	5
A. Surveying software	5
B. Statistical analysis	6
C. Statistical reporting	7
III. METHODOLOGY	9
A. Conceptual Framework	9
B. Rationale for the Framework	10
C. Technologies Used	10
IV. IMPLEMENTATION	12
A. Concept	12
B. Implementation Methods	13
C. User Testing and Project Assessment	16
V. RESULTS AND DISCUSSIONS	17
VI. CONCLUSIONS	23
VII. RECOMMENDATIONS	25
V. REFERENCES	27
VI. APPENDICES	30
A. Deliverables and Milestones	30
B. Qualifications	30
C. Resource Requirements	31
D. Software Requirements Specifications	32

E. Setup Guide	55
F. User Manual	73
G. Sample Survey Forms	95

TABLES

Table 1. Use Cases.	14
Table 2. Features.	15

FIGURES

Figure 1 – Process Diagram for the Survey Statistics Information System	9
Figure 2 - Entity-Relationship Diagram	15
Figure 3 - The Survey Statistics Information System Login	17
Figure 4 - Sample Creation in the Survey Statistics Information System	18
Figure 5 - Sample Analyses in the Survey Statistics Information System	18
Figure 6 – Descriptive Statistics of the Patient Survey	19
Figure 7 - Encoder's Landing Page	20
Figure 8 - Encoder's Manual Input page	20
Figure 9 - Encoder's Survey Input	20
Figure 10 - Add Patient Record	21
Figure 11 - Remove Patient Record	21
Figure 12 – Administration Screen	22
Figure 13 – User Profile Creation	22
Figure 14 – User Profile Deletion	22

Chapter 1

THE PROBLEM DOMAIN

The Novaliches District Hospital (NDH) was established in 2001 to provide quality healthcare services to residents of the Greater Novaliches Area, as well as other neighboring communities. It envisions itself as a general hospital that makes the 4As of health possible: readily Available, Accessible, Affordable and Attainable healthcare for the public (NDH, 2017). This hospital serves, on the average, around 2,000 outpatient patients, 700 in-patient admissions, and 500 emergency-room cases every month. (Reboton, personal interview, April 2018)

A. Statement of the Problem

This information system project was designed to facilitate research conducted by the Novaliches District Hospital, such as its patient satisfaction surveys. The system did this by providing a means of selecting those to be sampled, as well as a way for survey data to be entered, and statistical analysis to be conducted, with minimal manual intervention, rather than getting information from more than 3,000 patients every month.

B. Background and Objectives of the Project

The NDH is required to perform patient satisfaction surveys, both as a Department of Health requirement, and for its quality management system (QMS) to maintain and facilitate its ISO quality standards. These surveys are manually implemented, tabulated manually into Microsoft Excel, and presented as-is without statistical analysis.

There are two different patient satisfaction surveys being implemented. One survey is given to in-patients, who stay in the hospital's wards, while the other survey

is given to patients who seek out-patient treatment. These surveys are given to selected patients in the four departments of the hospital: Pediatrics, Internal Medicine, Surgery, and Obstetrics/Gynecology. These are manually implemented by the nurses of the NDH. For departments with many patients (i.e. Pediatrics), convenience sampling is performed – in this way, only some patients are surveyed (as chosen by the nurses). For other departments, surveys cover all patients as they are being discharged. Samples of the in-patient and out-patient surveys are available in the Appendix.

An issue raised with this is that the nurses, having other tasks to attend to, tended to deprioritize survey implementation. This is because they had to get responses from as many patients as possible. Statistical sampling will be of help to them, since if they do not have to get all samples for the survey, their workload is lessened and they could focus on other, more critical tasks.

Another issue that was raised was that there was little, if any, statistical training received by hospital staff, including administrators, physicians, nurses, and medical records staff. Due to budget constraints, the hospital could not afford to hire staff with statistical backgrounds to serve as statisticians for the survey. This is why the hospital does not perform statistical analysis on the collected data.

Statistical analysis that can be performed includes hypothesis testing (On average, is the data collected better than the defined null hypothesis?), as well as correlation and relationship tests (Is there a relationship between time spent in line and patient satisfaction?). Visualizations could also be added to the statistical analysis for the researchers to more easily understand this analysis.

The objectives of this project were to:

- Provide a sampling frame from which patients may be selected for the patient satisfaction survey
- Efficientize the survey process by only selecting samples, instead of collecting data from all patients. This is done based on currently accepted statistical practice
- Provide means for data entry for the sample that is secure, and keeps the patient's confidentiality; and that lends itself to statistical analysis
- Provide means for seamless, appropriate, and user-friendly statistical analysis based on the data entered. Reports output as a result of statistical analysis must be standardized to align with best practices. There must also be a visual component to these analyses for the appreciation of the hospital staff.
- Provide these tools in a cost-effective manner, using free/open-source tools.

C. Significance and Scope of the Project

This project is intended to facilitate patient surveys conducted within the Novaliches District Hospital. Because the Department of Health requires hospitals to conduct patient surveys, with the objective of providing better service to their patients, a way to simplify the survey process for them was developed.

This project made the survey process for the hospital more efficient. By automating the survey sampling, data collection, and statistical analysis of the collected data, the project has lessened the workload of nurses, of administration staff, and has enabled the hospital administration to obtain insights from the data. These insights will have led to better patient service, which was intended to result in greater positive feedback.

In addition, as it is the vision of the hospital to become a Level 3 (teaching) hospital, which trains residents to perform clinical research. It is hoped that once this has occurred, the project could be extended to facilitate the clinical research performed by its residents.

D. Documentation of Existence and Seriousness of the Problem

Before this system, there was currently no information system handling survey sampling, as the patient satisfaction survey was manually implemented. This resulted in less manpower available to do other tasks. On the other hand, survey work load was on top of the other work performed by implementors.

The current implementors of this survey are the nursing staff, who also have to take care of the patients. As such, survey implementation usually took a lower priority compared to other duties (Reboton, personal interview, October 2017).

In addition, no statistical analysis was done on this survey, due to a lack of trained manpower – the administrative staff have little, if any, statistical training. Only manual calculations of totals, and graphs of the totals, were performed on this, until the information system was developed.

Chapter 2

REVIEW OF EXISTING ALTERNATIVES

Existing alternatives were sought for each stage of the process – whether sample survey selection, statistical analysis, or statistical report generation. These technologies were compared, and approaches from these technologies which can be leveraged, were leveraged.

A. Surveying Software

R has the survey package, which is free, open-source, and allows users to take a statistical sample. (Lumley, 2017)

SAS has SAS/STAT – where there is a command, proc surveyselect, that creates statistical samples. This creates samples, but the cost of SAS licensing is US\$8,700 for the first year, and 25-30% of that amount thereafter (Dinsmore,2014). For the purposes of this project, then, SAS would be too cost-prohibitive for analysis delivery.

Open Epi is a web software that calculates optimal sample sizes based on the type of study to be conducted. According to the documentation, “OpenEpi provides statistics for counts and measurements in descriptive and analytic studies, stratified analysis with exact confidence limits, matched pair and person-time analysis, sample size and power calculations, random numbers, sensitivity, specificity and other evaluation statistics, R x C tables, chi-square for dose-response, and links to other useful sites.” (Dean et al, 2013). However, since this is publicly available on the Internet, it could not directly interface with current patient records due to security and confidentiality concerns.

B. Statistical Analysis

The following software can perform the necessary statistical analysis:

SPSS Statistics is a widely used program for statistical analytics, which contains a comprehensive set of statistical tools. Its drag and drop interface makes it simple and easy to perform statistical analysis across multiple data sources (SPSS, 2017). However, it is priced at USD 99 per month for a single user subscription.

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. (R Project, 2017) As R is free, and open-source, it provides the best fit for this system. Moreover, it could be integrated into web programs / servers through extensions such as RApache (Horner, 2013)

SAS (once known as the Statistical Analysis System) (SAS, 2017) is a software suite that can mine, alter, manage and retrieve data from a variety of sources and perform statistical analysis on it. SAS provides a graphical point-and-click user interface, called the SAS Enterprise Guide, for non-technical users and more advanced options through the SAS language. (Encyclopedia of Research Design, 2010)

Excel could also perform limited statistical analysis via the Analysis ToolPak.

However, this may not include nonparametric statistics, which are also going to be used for this project.

C. Statistical Reporting

Statistical analysis can be done via software. However, reporting is done manually by statisticians. One could automate this, using solutions similar to those offered by Narrative Science (Quill) – which has pioneered in a field called Natural Language Generation (NLG).

Leveraging Natural Language Generation for statistical report automation will greatly help increase the understanding of statistical reports, without need for a statistical background. (Narrative Science, 2017)

In addition, for any free-text items, word cloud analysis could be used. This is a visual representation of the frequency of words that were submitted during the survey. (Better Evaluation, 2017)

Tools for visualization of the statistical analysis are also considered. This includes such software as R (R Project, 2017), and Weave (Weave Technology, 2017). These platforms help transform statistical data and analysis into visual insights.

Approaches for survey results generation were also considered. For the purpose of this research, surveys will be divided into large and small. Large means there were more observations than the optimal survey size, while small means that there are less observations than the optimal survey size. Large surveys can be generalized, hence lend themselves well to parametric statistical analysis, while smaller surveys lend themselves to nonparametric statistical analysis.

Smaller sample sizes generally do not approximate the normal distribution, as per the Central Limit Theorem. As such, the most popular statistical tests, such as the z-test for hypothesis testing, the chi-square test of independence, and the Pearson's r

correlation cannot be reliably used for analysis. (Mood et al, 1978) However, a certain class of tests, called nonparametric statistical tests, can be used with smaller samples. These tests make no assumptions about the statistical distributions of variables; hence, these are useful for smaller sample sizes (Hollander et al, 2003). Examples of nonparametric statistical tests include the Wilcoxon signed rank test (an analogue to t-tests for hypothesis testing), as well as the Kendall's tau test of correlation (an analogue to the Pearson's r correlation test).

Larger sample sizes make for more random variables, which would approximate the normal distribution, by virtue of the Central Limit Theorem (Mood et al, 1978). With a sufficiently large sample, it is more likely that the averages of the (random) observations would be close to the expected value (or mean), by virtue of the Weak Law of Large Numbers (Mood et al, 1978). As such, assumptions about their distribution can be made (invariably, this would be the Normal distribution) (Mood et al, 1978). This allows more popular statistical tests, such as the z-test for hypothesis testing, the chi-square test of independence, and the Pearson's r correlation to be used, as long as the sample is of sufficient size. (Mood et al, 1978)

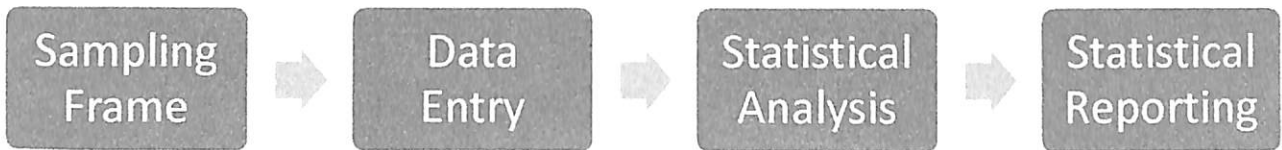
Chapter 3

METHODOLOGY

A. Conceptual Framework

The framework that was implemented for this project can be summarized in the process diagram below:

Figure 1. Process diagram for the Survey Statistics Information System:



The sampling frame is the source of the sample to be taken (Sarndal, et. al., 2003). As such, this interfaced with the existing patient database of the Novaliches District Hospital.

To save on costs, the survey proper was still implemented manually by the nursing staff, who provided all sample sheets to an encoder for data entry.

Once the data is entered into a database of sample responses, statistical analysis could be performed on the data. Statistical analysis included creation of descriptive statistics (to further summarize the data, as compared to previously only graphing totals), hypothesis testing (against a null hypothesis set by the researchers), and correlation/independence testing (between two variables / questionnaire responses).

Statistical reporting was thus automated. The web page itself provided directions as to how to create reports, with sample values provided. The output of this statistical reporting, both tabular and graphical in format, was readable to laymen, although a dictionary was also provided for the hospital staff to understand outputs.

The information system that was implemented is a Web information system, that is accessible only within the intranet of the hospital (to maintain the system's security), and that can access application programming interfaces (APIs) as necessary.

B. Rationale for the Framework

The rationale for the framework is that it aligns to currently accepted statistical practice – survey sampling requires a sampling frame, after which data from that statistical survey will be tabulated; and only after that can statistical reports be generated.

C. Technologies Used

To implement the information system, the following technologies were used:

In order to create the sampling frame, to sample patients, and to add their inputs to the system, a database management system (DBMS) was required. For this implementation, MySQL was used, as this is the most popular open-source database solution (MySQL, 2017).

For statistical analysis, graphing, and visualization, the open-source language R was used. In order to implement this properly within the confines of the Web

information system, this was used alongside Apache in a server-side system called RApache.

The R language supports sampling, statistical analysis, and graphical output, making this the backbone for all statistical calculation and sample generation (Lumley, 2017) (R Project, 2017). Since R was free to install, and open-source, it was cost-effective, yet extensible using packages, which made this system powerful. This was why R was used in the system.

This web application was created using HTML5 standards, enhanced by Cascading Style Sheets (CSS) files ported from Bootstrap, in order to create an aesthetically pleasing and user-responsive design. This was implemented by the free (for Windows users) software Microsoft Expression Web 4 as an HTML page designer and PHP script generator.

The server implementation for this web application was done in Ubuntu, with a Linux, Apache, MySQL, and PHP (LAMP) stack; in order to enhance the server, RApache was necessary (Horner, 2013). Results processing was done via a combination of JavaScript Object Notation (JSON) outputs from RApache's R script handling process, and PHP scripting to output a workable set of statistics outputs.

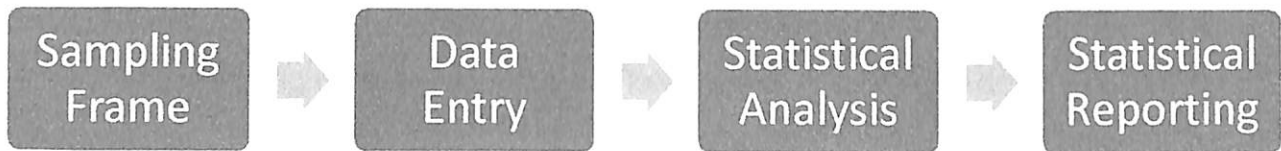
Chapter 4

IMPLEMENTATION

A. Concept

Figure 1 is repeated below, as this shows the concept of the Survey Statistics Information System:

Figure 1. Process diagram for the Survey Statistics Information System:



The design of this project is a system that closely matches the existing currently accepted statistical practice of:

1. Sampling Frame – source of observations to be sampled – this was the patient database, which was also the source of survey sampling.
2. Data Entry and storage – once the survey / trial has been performed, data entry was performed for each result. There is a choice for the encoders to either enter each entry into the database, or upload one CSV file containing the results of all surveys that were performed. This was then stored in a database for later statistical analysis.
3. Statistical Analysis Module – once data has been input into a database, this module would run, and create reports with all appropriate statistical tests being done:

Hypothesis Testing (if selected, the hypothesized null value must be provided)

Correlation Testing (a correlation table will be output)

Tests for Independence (this would be done pairwise)

The report would also contain descriptive statistics, to make it easier to identify demographic trends within the data. It will also contain graphs and charts, in order to provide visualizations of the data, and of the statistical tests being done to the data.

Statistical tests in the report can be done in a parametric, or non-parametric approach, depending on the sample size. These were discussed in detail in Chapter 2 – Review of Alternatives.

B. Methods

The following methods were used to create, prototype, test, and complete the project:

Review of Existing Systems: Existing systems were documented based on interviews with the client. It was found out that surveys were done manually, so this information system automated the process due to client need.

Literature review and Assessment of Existing Alternatives: For this project, both statistical literature and systems development literature are intended to be consulted and mined for best practices. In addition, existing systems such as SAS, SPSS, R, and Excel's Analysis Toolpak were compared. Weak points were identified, which were addressed within this project. The results of this literature review could be found in Chapter 2 – Review of Alternatives.

Production of prototypes – The plan of action was to produce a minimum viable product by January 2018, for the hospital to try out its latest survey data on. There were changes in the use cases and functional requirements due to identified changes in the ability of the program to be implemented, given time constraints. For instance, instead of granting the administrator all powers, his role was reduced to profile management. He could create dummy users for the other roles.

This is shown in Tables 1 and 2 (Use Cases and Functional Requirements, respectively)

Table 1. Use Cases

User	Use Case
System Administrator	UC-01: Application Access
System Administrator	UC-02: Login
System Administrator	UC-03: Logout
System Administrator	UC-04: Create user
System Administrator	UC-05: Modify user
System Administrator	UC-06: Delete user
Medical Records Officer	UC-07: Modify own profile
Medical Records Officer	UC-08: Create additional patient record
Medical Records Officer	UC-09: Modify Patient Record
Medical Records Officer	UC-10: Delete patient record (for erroneous patient record)
Researcher	UC-11: Modify own profile
Researcher	UC-12: Sample Size Determination
Researcher	UC-13: Creation of the Actual Sample for Implementation
Researcher	UC-14: Statistical Analysis Selection
Researcher	UC-15: Statistical Report Generation
Encoder	UC-16: Modify own profile
Encoder	UC-17: Import from Excel / CSV
Encoder	UC-18: Data Entry

Table 2. Functional Requirements

Functional Requirements
F-00: Main Menu
F-01: Login/Logout
F-02: Profile Management
F-03: Sampling Frame
F-04: Data Entry Module
F-05: Statistical Analysis and Report Generation Module

Part of the system design involved database creation. The following databases were created for the system: Logins (a listing of all authorized users in the system), Patient Information (pertinent patient details for the sampling frame), In-Patient Survey Data (survey information taken from the in-patient survey), and Out-Patient Survey Data (survey information taken from the out-patient survey). Figure 2 contains the Entity-Relationship Diagram of these databases:

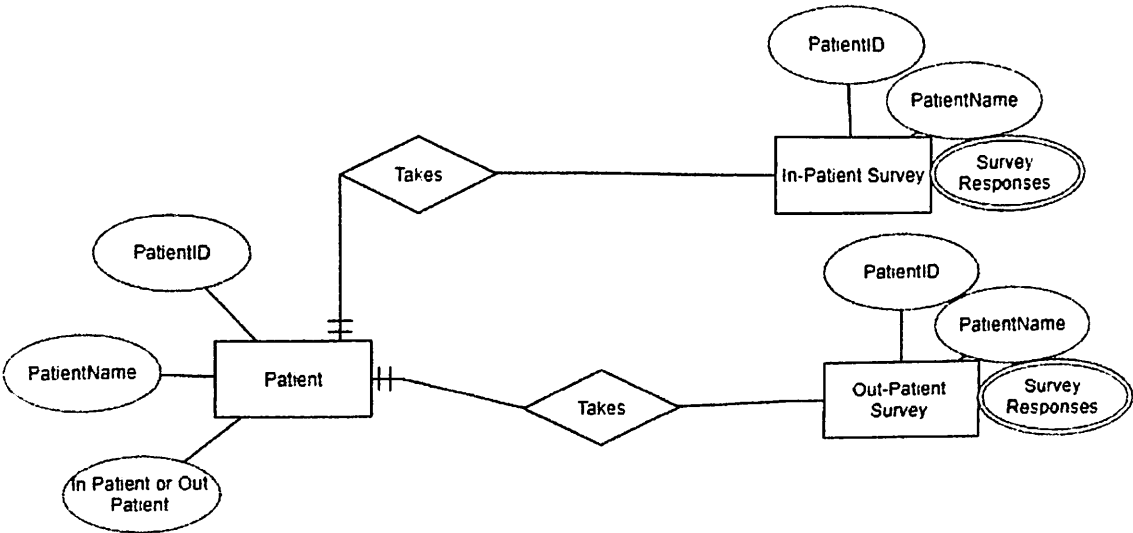


Figure 2. Entity-Relationship Diagram

Prototype testing was performed with the main proponent of this system. The system was demonstrated to him, and he performed User Acceptance Testing (UAT) on this system, as well as answered the software usability survey for this.

Preparation of final project documentation/report was performed after this testing.

C. User Testing and Project Assessment

As soon as the system was completed, a meeting was held in order to test the prototype. For this purpose, a dummy database was constructed with the same fields as the patient database to be created. Samples were obtained from the patient database, which were given a mock survey. Data entry was simulated, and the survey databases for in-patient and out-patient surveys were checked for issues.

Chapter 5

RESULTS AND DISCUSSIONS

In order to fulfill the objectives of this project, which are to provide a sampling frame, to efficienize the survey process by selecting samples, to provide means for data entry and simplified statistical analysis, and to provide these in a cost-effective manner, a web application was designed for the Novaliches District Hospital's patient survey.

The web application involved logins to ensure security, which was done for four classes of user: Administrators, responsible for user profile creation; Researchers, who will select the sample and create the statistical analysis; Medical Records staff, responsible for the patient database; and Encoders, who are responsible for data entry of survey responses. Figure 3 shows the login page.

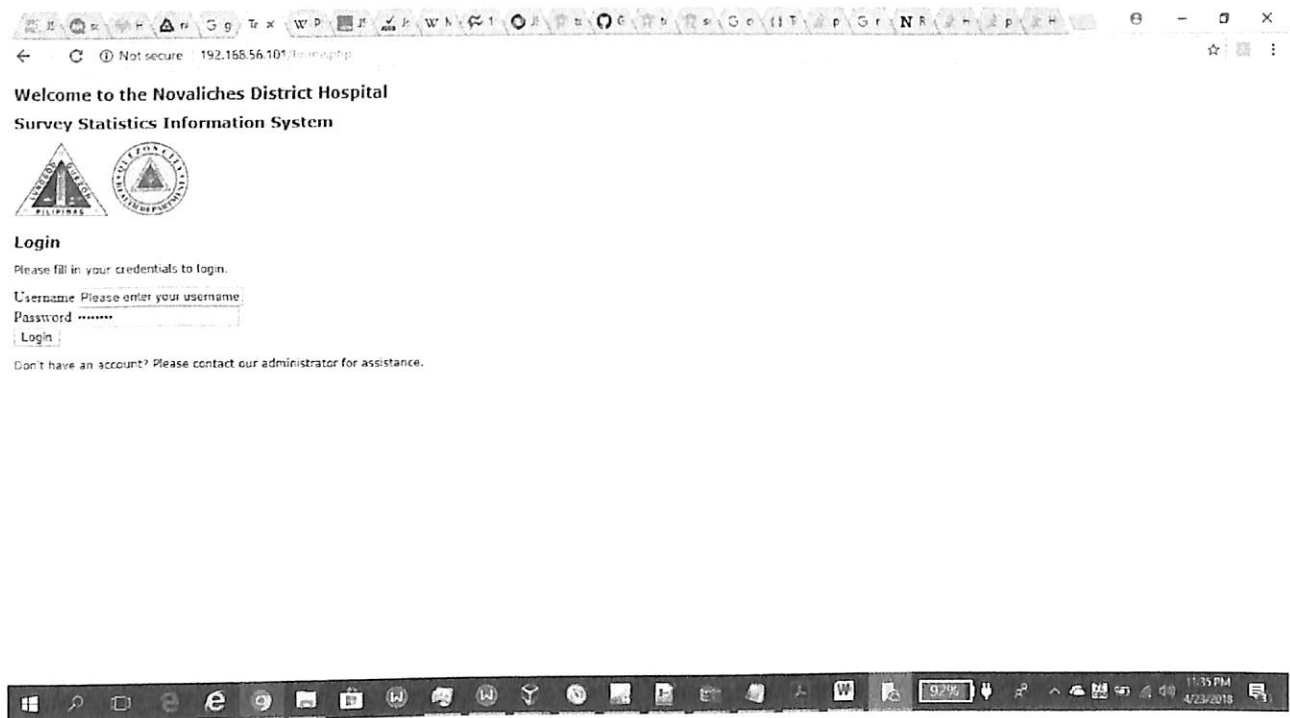


Figure 3. The Survey Statistics Information System Login.

Researchers have access to sample creation, wherein they will input the statistical tests to be done, as well as statistical requirements for these tests. These will then use the power calculation methodology of R to create a sample survey, which would then be manually implemented, and then encoded separately. Figure 4 below shows how a researcher could create a sample:

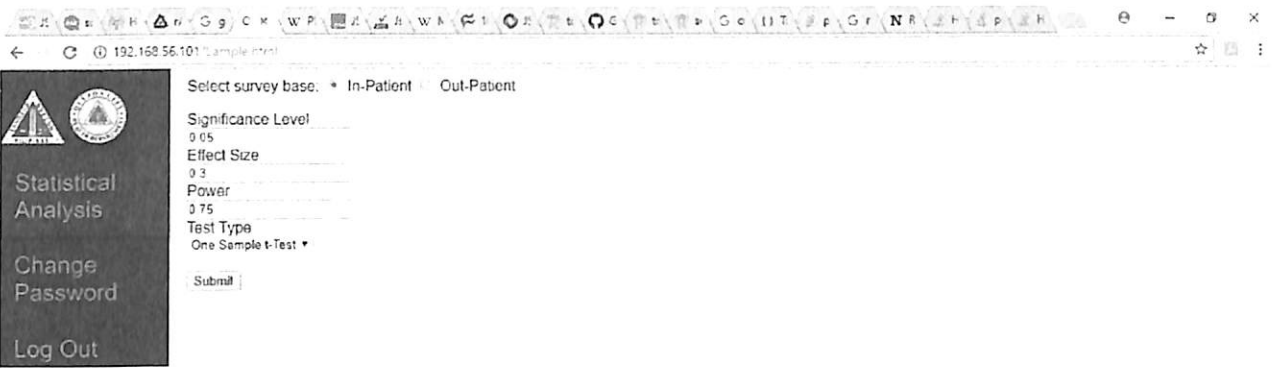


Figure 4. Sample Creation in the Survey Statistics Information System.

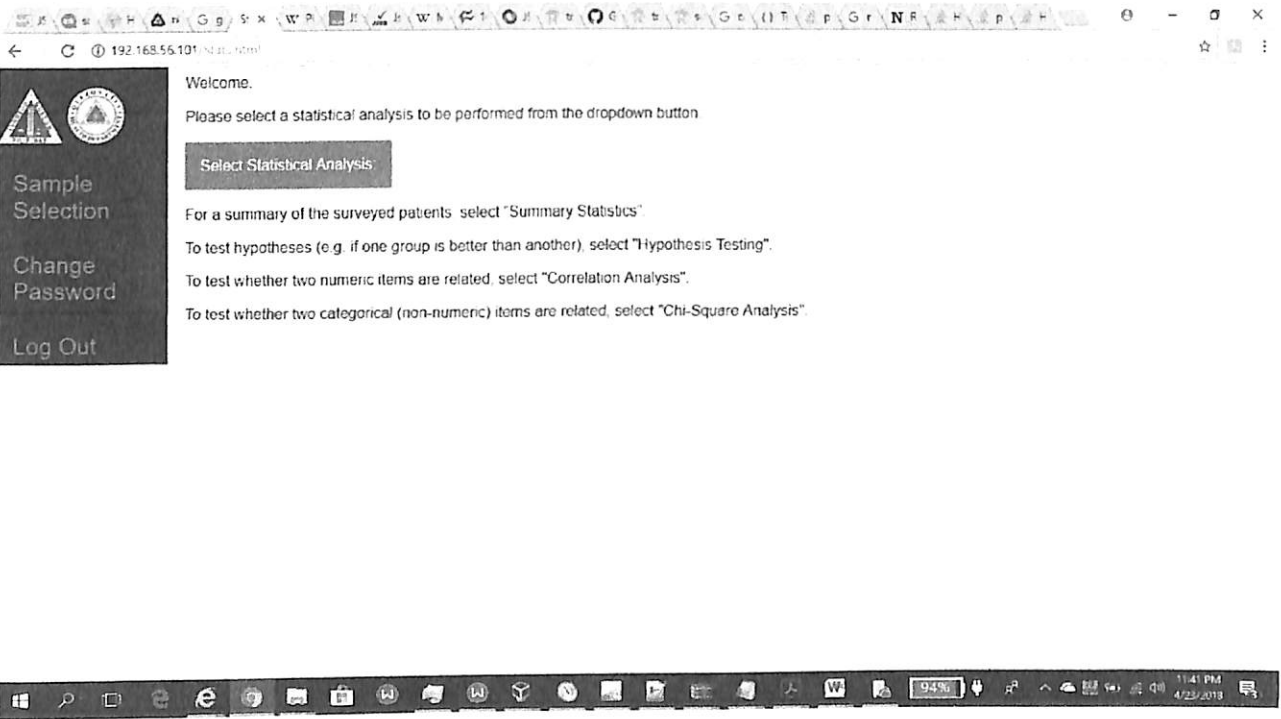


Figure 5: Statistical Analyses in the Survey Statistics Information System.

There are a number of statistical analyses that can be performed on the survey data. For instance, one can obtain summary statistics (with corresponding graphs comparing average scores for each employee/resource type in the hospital). These are shown in Figure 6 below.

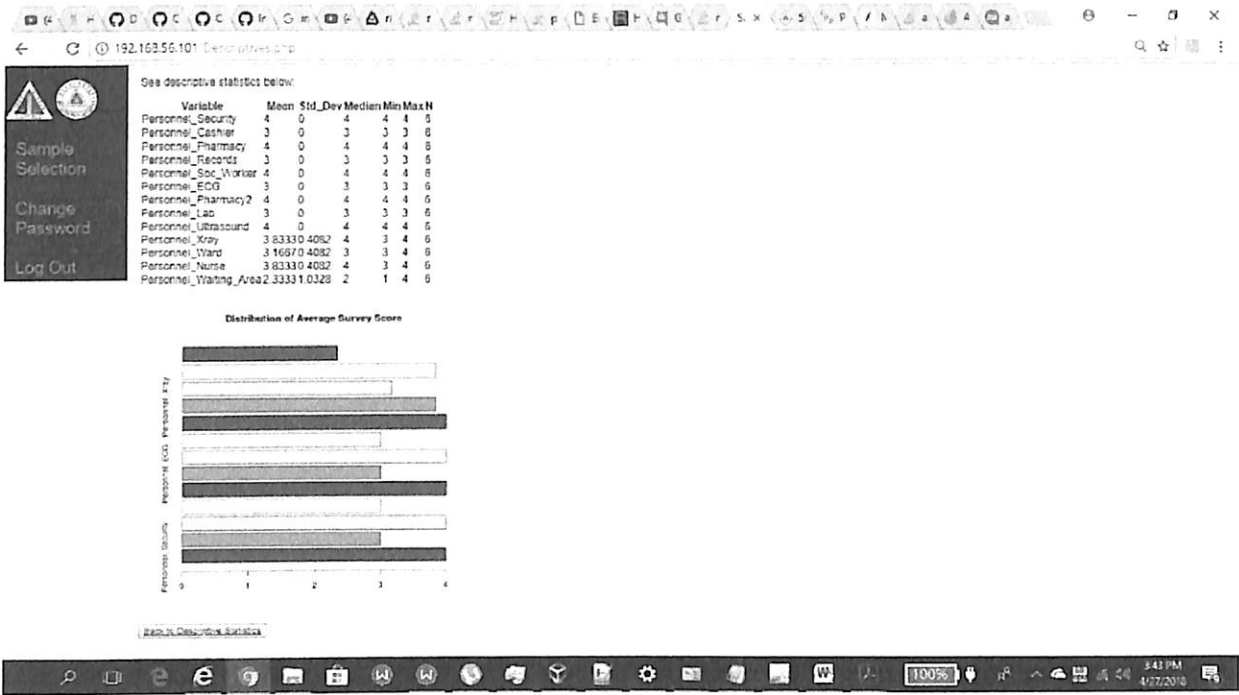


Figure 6. Descriptive Statistics of the Patient Survey.

Encoders have access to the data entry module of survey responses. They could enter survey responses, which were already designed in the system, to make it easier for them to enter data. Figure 7, 8, and 9 shows how the encoder can input data into the data entry module. Figure 8, in particular, shows how the encoder can input either into the In-Patient survey or the Out-Patient survey:

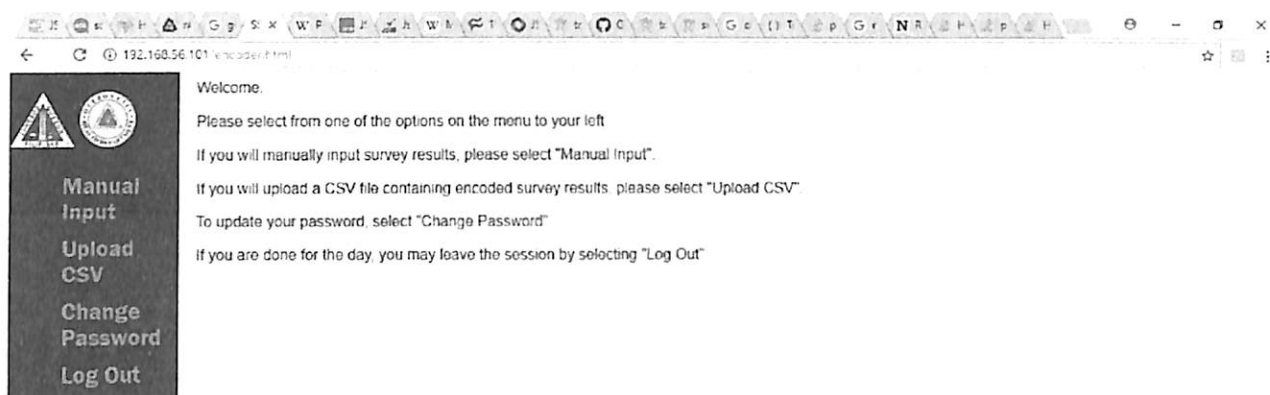


Figure 7: Encoder's Landing Page



Figure 8: Encoder's Manual Input page.

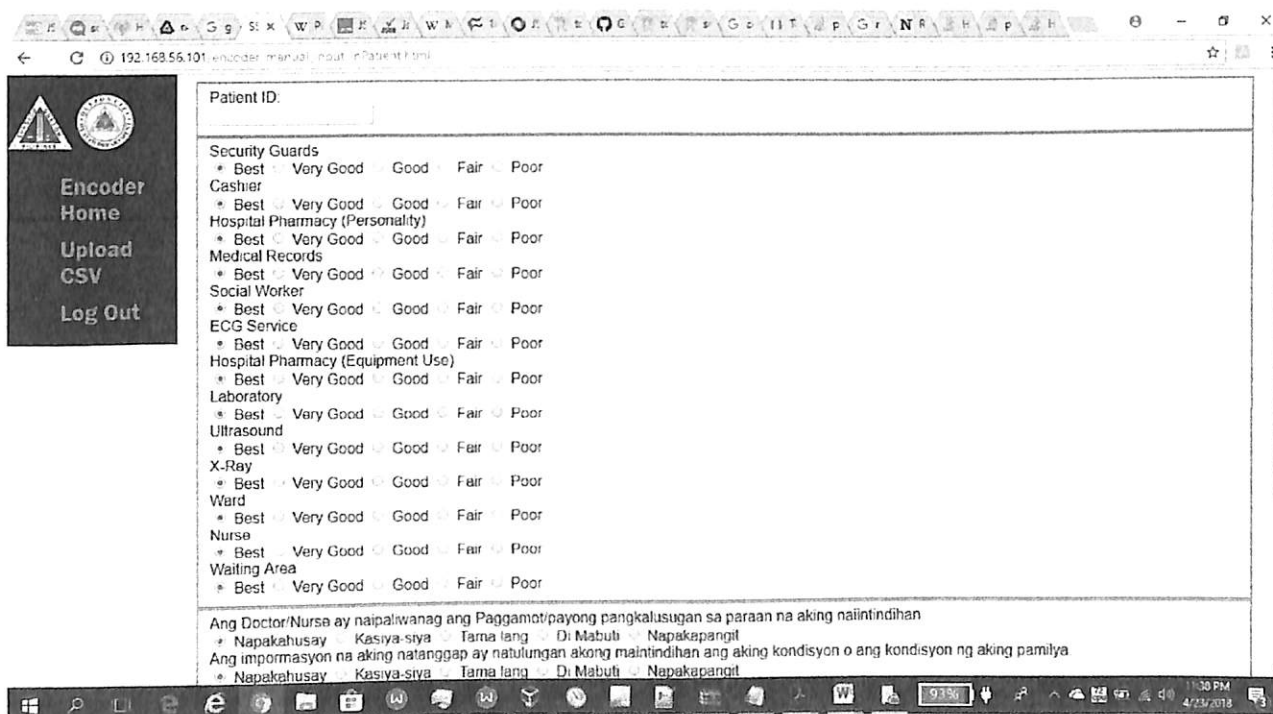


Figure 9. Encoder's Survey Input.

The Medical Records staff has access to the entire patient database. They could add or remove patients as necessary from the entire sampling frame. This is shown in Figures 10 and 11:

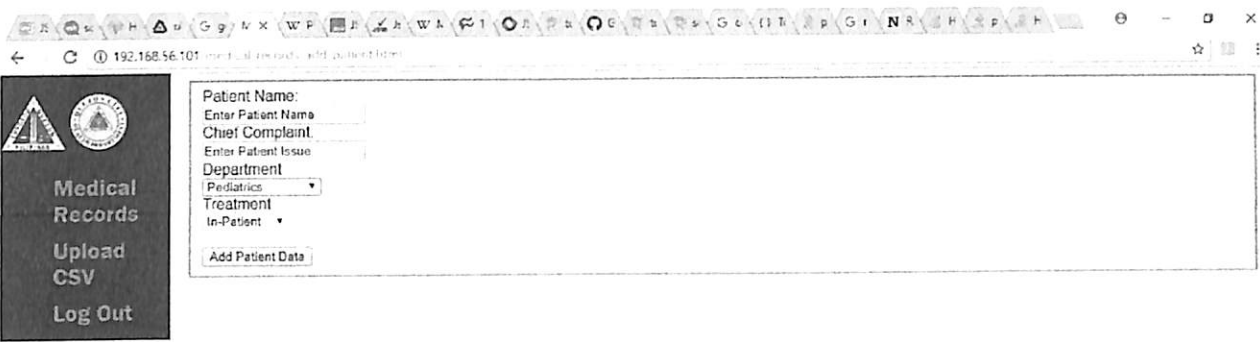


Figure 10: Add Patient Record

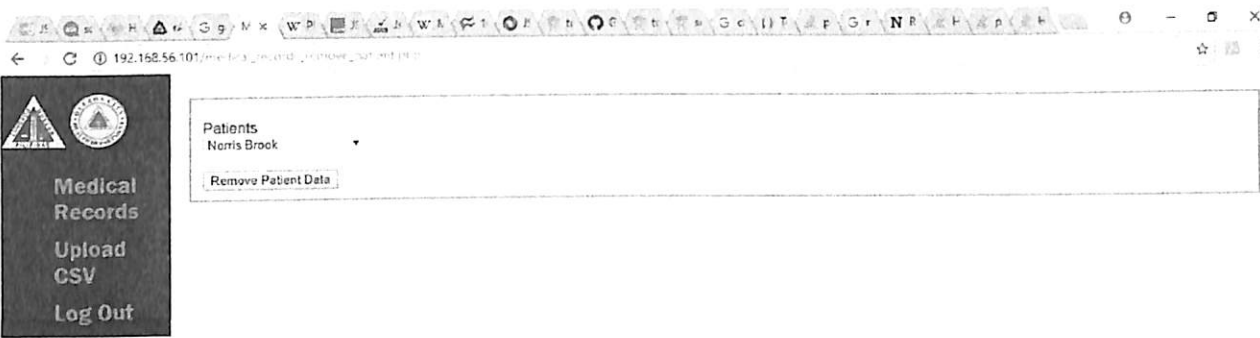


Figure 11: Remove Patient Record

Administrators have access to the profile creation area, where they could create or delete user profiles for the other users. This is shown in Figure 12 – the Administrator screen. Figure 13 and 14 deal with profile creation and deletion, respectively.

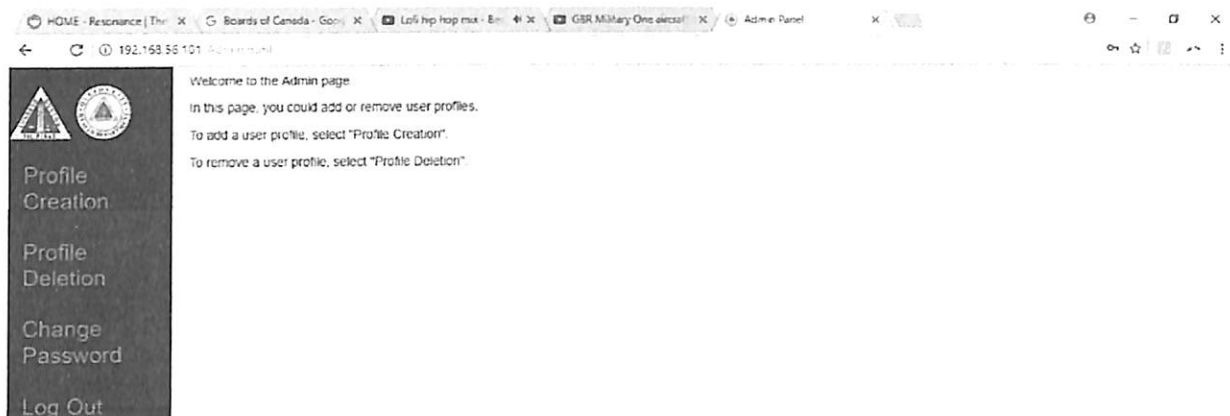


Figure 12. The Administrator screen



Figure 13: Profile Creation

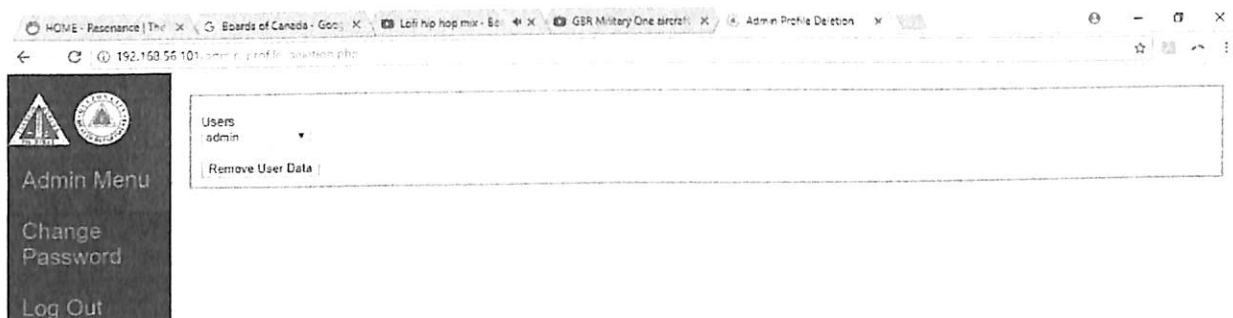


Figure 14: Profile Deletion

Chapter 6

CONCLUSIONS

The patient satisfaction survey, as a requirement of the Department of Health and a quality management tool, is important to gauge how patients see the hospital's services, thereby providing a way for the hospital to improve its services.

The Novaliches District Hospital, as a secondary hospital, is required to implement this survey. In order to make the process more efficient for them, the project aimed to implement a web application that could serve as sampling frame, source of survey gathering information, and hub for simplified statistical analysis.

The project's objectives, which are reiterated here, were successfully attained by this system.

Objective 1 – to create a sampling frame – was accomplished by the sampling module of this system, which took the patient database and used it to create random samples.

Objective 2 – to efficienize the survey process by selecting samples – was also accomplished within the sampling module of this system, since instead of creating a census, the hospital only has to extract information from fewer subjects.

Objective 3 – to provide data entry for the survey's results – was accomplished by the data entry module, which provided encoders two means of inputting data: manually for each survey, or via uploading an existing CSV file.

Objective 4 – seamless, appropriate, user-friendly, and visual statistical analysis – was accomplished by the researchers'/statistical analysis module, which calculated statistics from the patient data obtained using the data entry module

Objective 5 – cost-efficiency- was provided by the fact that all of the tools utilized in the creation of this survey were free and open-source, from the R program to the server side programs being used.

The technologies and tools used for this are appropriate, such that for as long as there is a working intranet, and server, this program could work in the hospital's network.

The researcher learned more about web development, about coding in PHP, parsing JSON outputs between R and PHP, HTML forms, and the creation of a web information system. Database knowledge was also learned, through the setup of databases for patients, survey results, and logins.

Chapter 7

RECOMMENDATIONS

Although the Survey Statistics Information System was completed, the researcher has found a number of possible improvements which could be done in the future. These include adding the date the survey was taken, and extracting statistics from survey date (for trend and seasonality analysis – if the average scores went up or down for one personnel type, as an example). One could also create plots over time to visualize these trends. One could also extract the actual responses by patients via a database export feature, in the future.

As data privacy is a pressing topic, especially considering hospital data, it is also recommended to enhance the security of this system, with regards to its handling of patient data.

The researcher also has recommendations to policy – the Department of Health's policy requiring hospitals to perform patient surveys must be revisited. If this policy is to continue, there must be informed consent, where the patient knows what his data is being used for. This is in line with provisions in the Philippines' Data Privacy Act of 2012.

There were alternatives that were explored, but were not implemented in this research for security reasons – these required outside application program interfaces (APIs) which may take the data provided to it and use it in a manner inconsistent with the patients' rights to data privacy. These included Natural Language Generation, as well as word cloud analysis. These could be used in systems that do not require as much confidentiality as hospital information systems.

For further research, it is suggested to create similar programs for other hospitals, subject to their survey needs. Other applications could also be explored for the creation of survey data, such as in market research and others.

The future researcher may wish to improve this system by adding other statistical tests such as, but not limited to, analysis of variance (ANOVA), regression analysis, and exploratory data analysis. Furthermore, since the system runs on R, it could leverage other web applications that use R, such as Shiny (R Studio, 2018). The researcher also suggests exploring the use of these web applications to develop statistical information systems. Also, for the medical records section, there may be an easier way to remove patient records than doing a dropdown, especially when there are over 3,000 records in the dropdown.

Chapter 8

REFERENCES

BetterEvaluation.(2017) *Word Cloud* <http://www.betterevaluation.org/en/evaluation-options/wordcloud> Accessed Oct 21, 2017

ComputerWorld (2017). *8 Cool Tools for Data Analysis, Visualization, and Presentation*. <https://www.computerworld.com/article/2502891/business-intelligence/business-intelligence-8-cool-tools-for-data-analysis-visualization-and-presentation.html?page=4> Accessed Nov 25, 2017.

CommVault (2017). *Hardware specifications for Web Server*.
http://documentation.commvault.com/commvault/v10/article?p=system_requirements/commcell_sizing/webserver.htm Accessed Nov 25, 2017

Dean, A.G.; Sullivan, K.M.; Soe, M.M..*OpenEpi: Open Source Epidemiologic Statistics for Public Health*, www.OpenEpi.com, Accessed Oct 20, 2017

Dinsmore, T.W. (2014). *SAS vs R Part Two*.
<https://thomaswdinsmore.com/2014/12/15/sas-versus-r-part-two/> Accessed Oct 20, 2017

Horner, J. (2013). *rApache: Web application development with R and Apache*.
<http://www.rapache.net/>

Lumley, T. (2017) *Package "Survey"*. R (CRAN) <https://cran.r-project.org/web/packages/survey/survey.pdf> Accessed Oct 20, 2017

Microsoft (2017). *Excel Analysis ToolPak* <https://support.office.com/en-us/article/Use-the-Analysis-ToolPak-to-perform-complex-data-analysis-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6> Accessed Oct 21, 2017

Mood, F.M; Graybill. F.A.; and Boes, D.C. (1974) *Introduction to the Theory of Statistics, Third Edition*.

Narrative Science (2017). *The Quill Platform* <https://narrativescience.com/Platform>. Accessed Oct 21, 2017

Nonparametric Statistics (Hollander M., Wolfe D.A., Chicken E. (2014). *Nonparametric Statistical Methods*

Novaliches District Hospital. <http://quezoncity.gov.ph/index.php/qc-department/243-ndh> Accessed Oct 15, 2017

StatSoft, Inc. (2013). *Electronic Statistics Textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/>.

QuickR. (2017). *Graphs*. <https://www.statmethods.net/graphs/index.html> Accessed Nov 24, 2017

R Project (2017). *The R Project for Statistical Computing*. <https://www.r-project.org/> Accessed Oct 20, 2017

R Studio (2018). *Shiny, from RStudio*. <https://shiny.rstudio.com/> Accessed 23 April 2018.

Salkind, N.J. (2010). *Encyclopedia of Research Design*. Accessed Oct 20, 2017

Särndal, C; Swensson, B; Wretman, J (2003). *Model assisted survey sampling*. Accessed Oct 20, 2017

SAS Institute (2017) *About SAS* https://www.sas.com/en_us/company-information.html#history Accessed Oct 20, 2017

SPSS (2017). *SPSS Statistics* <https://www.ibm.com/products/spss-statistics>
Accessed Oct 20, 2017

Weave Technology (2017). *About Weave* <http://www.iweave.com/about2.html>
Accessed Nov 25, 2017

APPENDICES

Annex A. Deliverables and Milestones

1. Proposal - submitted October 21, 2017
2. Revisions – submitted before November 30, 2017
3. Data Gathering (patient surveys, test patient database) – February-March 2018
3. Prototype – was constructed April 7, 2018
4. User Testing and Assessment Plans – finalized April 14, 2018
4. Report of results from user testing – finalized April 18, 2018

Annex B. Qualifications

Samuel Patrick R. Pasamonte

BS Statistics, University of the Philippines – Diliman, class of 2012

Decision Management / Data Capability Analyst, Citibank NA, 2012 – 2017

As data capability analyst, is proficient in creating, maintaining, and reporting on databases using SQL.

As statistician, has assisted in multiple thesis statistical analyses, both at the undergraduate level and at the graduate level.

Proficient in SAS, SQL, and R as a result of undergraduate training.

Learning agenda: The researcher intends to learn web development, and how to integrate APIs to an existing program, in order to complete this project.

Annex C. Resource Requirements

This project required the following resources:

1. Hardware: Two computers are required:

a. One desktop/server computer that will run 24/7 as a server for the information system, with the following specifications:

- 100 GB of disk space for cache
- 8 CPU cores
- 16 GB RAM
- Server operating system (Ubuntu Server 16.04)

b. One desktop/laptop to be used for development, with the following specifications:

- Windows 7/10 (with an Ubuntu virtual machine) or Ubuntu
- At least 2 CPU Cores (rated at least 2GHz)
- 8 GB RAM

2. Patient Database Access – or, at least, fields that could be used in creating a dummy patient database for proof-of-concept testing

3. Manpower support – of, at least, an encoder.

4. Access to the project site – Novaliches District Hospital – and its IT spaces

Annex D. Software Requirement Specifications

INTRODUCTION

The Survey Statistics Information System is intended for use by the research and administrative staff of the Novaliches District Hospital to streamline their patient satisfaction survey process. Since the current process is to get all patients, or to use convenience sampling, depending on department, not all patients have an equal chance of selection, leading to the possibility of bias. The objective of this information system is to move surveys towards a probability sampling, which saves time and effort on the staff implementing the survey. The information system also provides automated statistical reporting, which translates into actionable insights on the part of hospital administration without needing to hire trained statisticians to interpret statistical reports.

USE CASES

The Use Case Suite is arranged based on the four main classes of user:

- System Administrator
- Medical Records Staff
- Researcher
- Encoder

USE CASE SUITE

Release Information	
Project	Survey Statistics Information System
Internal Release Number	1.0
Related Documents	Use Cases
	Use Case Diagram
	Feature Set

Use Cases by All users:

UC-01: Application Access

UC-02: Login

UC-03: Logout

Use Cases by System Administrator:

UC-04: Create user

UC-05: Modify user

UC-06: Delete user

UC-07: Sample Size Determination

UC-08: Creation of the Actual Sample for Implementation

UC-09: Import from Excel / CSV

UC-10: Setup Number and Type of Questions

UC-11: Data Entry

UC-12: Statistical Analysis Selection

UC-13: Statistical Report Generation

For medical records officer (access to the Patient Records Database)

UC-14: Modify own profile

UC-15: Create additional patient record

UC-16: Modify Patient Record

UC-17: Delete patient record (for erroneous patient record)

For researcher/admin staff

UC-18: Modify own profile

UC-19: Sample Size Determination

UC-20: Creation of the Actual Sample for Implementation

UC-21: Setup Number and Type of Questions

UC-22: Statistical Analysis Selection

UC-23: Statistical Report Generation

For encoder

UC-24: Modify own profile

UC-25: Import from Excel / CSV

UC-26: Data Entry

USE CASES

Release Information	
Project	Survey Statistics Information System
Internal Release Number	1.0
Related Documents	Use Case Suite Use Case Diagrams Feature Set

Default Aspects of All Use Cases

Direct Actors:

Researcher – one who obtains results from the survey, and receives statistical analysis on data obtained via the survey

Medical Records staff – one who has access to the patient database

System Administrator – user with the highest level of access to the application

Data Encoder – user responsible for inputting data obtained via survey; has the least access to the application.

Stakeholders: Project owners, as well as researchers from the Administration Department of the Novaliches District Hospital

Prerequisite: Application must already have been installed in the machine.

Use Frequency: This depends on the user of the application:

Data Encoder – every day, as survey results come in

Medical Records staff – every other day, to ensure that the patient database is still accessible

Researcher – every day, to obtain fresh samples

Administrators – every week, to keep the system operating efficiently.

Use Cases by All Users:

UC-01: Application Access

Summary: This is expected to be a web application, which should only be accessible within the intranet of the Novaliches District Hospital. Application must be accessible, online, and the main menu/homepage should display a login screen.

Use Frequency: This depends on the user of the application:

Data Encoder – every day, as survey results come in

Medical Records staff – every other day, to ensure that the patient database is still accessible

Researcher – every day, to obtain fresh samples

Administrators – every week, to keep the system operating efficiently.

UC-02: Login

Summary: As a user, I have to be able to log into the system. Once logged in, I should only have access to those features that my role has access to.

Use Frequency: This depends on the user of the application:

Data Encoder – every day, as survey results come in

Medical Records staff – every other day, to ensure that the patient database is still accessible

Researcher – every day, to obtain fresh samples

Administrators – every week, to keep the system operating efficiently.

UC-03: Login

Summary: As a user, at the end of the working day, I have to be able to log out of the system.

Use Frequency: This depends on the user of the application:

Data Encoder – every day, as survey results come in

Medical Records staff – every other day, to ensure that the patient database is still accessible

Researcher – every day, to obtain fresh samples

Administrators – every week, to keep the system operating efficiently.

Use Cases by Administrator:

UC-04: Create user

Summary: As a system administrator, I need to be able to create user profiles.

Priority: Essential

Use Frequency: Once per user to be created

Prerequisites: Login as an administrator

Scenario:

- Enter "Profile Management" page
- Select "Create New User Profile"
- Input information on the new user profile
- Input the user's role (medical records staff, data encoder, researcher, or system administrator)
- Create username and password (to be provided to the user, who will be required to change password)

UC-05: Modify user

Summary: As a system administrator, I need to be able to modify user profiles.

Priority: Essential

Use Frequency: Once per user to be modified

Prerequisites: Login as an administrator

Scenario:

- Enter "Profile Management" page
- Select "Modify Profile"
- Select profile to be modified
- Modify information in the user profile

UC-06: Delete user

Summary: As a system administrator, I need to be able to delete user profiles.

Priority: Essential

Use Frequency: Once per user to be modified

Prerequisites: Login as an administrator

Scenario:

- Enter "Profile Management" page
- Select "Delete Profile"
- Select profile to be deleted
- A confirmation box should appear ("Delete this user?")
- Confirm
- User should be deleted. Profile should no longer be visible

Use Cases by Medical Records Staff:

UC-07: Edit Own Profile

Summary: As a user, I should be able to edit my own profile information

Priority: Expected

Use Frequency: As there are changes in my information (seldom)

Prerequisites: Login as medical records staff

Scenario:

- Select "Profile Management"
- Select "Edit Profile"
- Change any of the fields available to the user, such as Address, Phone, Email, and Password.
- For password changes, place the current password, and the new password twice.

UC-08: Create additional patient record

Summary: As a medical records staff member, I should be able to create new patient entries, modifying the patient database through this tool, in addition to other tools where I can modify the patient database.

Priority: Desired

Use Frequency: Everyday

Prerequisites: Login as medical records staff

Scenario:

- Select "Patient Database"
- Select "Create New Record"
- Input patient information
- Click OK
- A dialog box will appear, asking you to confirm. Click OK.

UC-09: Modify Patient Record

Summary: As a medical records staff member, I should be able to modify patient entries through this tool, in addition to other tools where I can modify the patient database.

Priority: Desired

Use Frequency: Everyday

Prerequisites: Login as medical records staff

Scenario:

- Select "Patient Database"
- Select "Modify Record"
- Select a patient to modify. You may search by name, by patient ID number, or select from a drop-down
- Modify patient information as needed
- Click OK
- A dialog box will appear, asking you to confirm. Click OK.

UC-10: Delete patient record (for erroneous patient record)

Summary: As a medical records staff member, I should be able to remove erroneous patient entries through this tool, in addition to other tools where I can modify the patient database.

Priority: Desired

Use Frequency: Everyday

Prerequisites: Login as medical records staff

Scenario:

- Select "Patient Database"
- Select "Delete Record"
- Select a patient whose record needs to be removed. You may search by name, by patient ID number, or select from a drop-down
- Click Delete.
- A dialog box will appear, asking you to confirm. Click OK.

Use Cases by Researchers:

UC-11: Modify own profile

Summary: As a user, I should be able to edit my own profile information

Priority: Expected

Use Frequency: As there are changes in my information (seldom)

Prerequisites: Login as researcher

Scenario:

- Select "Profile Management"

- Select "Edit Profile"
- Change any of the fields available to the user, such as Address, Phone, Email, and Password.
- For password changes, place the current password, and the new password twice.

UC-12: Sample Size Determination

Summary: As a researcher, I need to be able to determine how many samples I should survey, to optimize the survey's power.

Priority: Essential

Use Frequency: Daily

Prerequisites: Login as researcher

Scenario:

- Enter "Sampling Frame" page
- Click "Select Survey Sample".
- Two options appear: Enter Sample Size and Use Best Sample Size
- If "Use Best Sample Size" is selected, the sample size in "Enter Sample Size" becomes the optimum sample size (as determined based on statistical best practices)

UC-13: Creation of the Actual Sample for Implementation

Summary: As a researcher, I need to be able to create a sample to implement the survey.

Priority: Essential

Use Frequency: Daily

Prerequisites: Login as researcher

UC-19 must be completed.

Scenario:

- Once one option is selected, click on "Generate Sample".
- A random sample, consisting of the sample size input (if entered), or of the optimum sample size (based on statistical best practice) is created.
- The random sample is exported via CSV or Excel and saved to the user's computer

UC-14: Statistical Analysis Selection

Summary: As a researcher, I need to be able to request certain statistical analyses to be created depending on the survey requirements.

Priority: Essential

Use Frequency: Daily

Prerequisites: Login as researcher

Scenario:

- Select the Statistical Analysis Module
- Select "Select Statistical Tests"
- Select any of the statistical tests: Hypothesis Testing, Correlation, Relationship.
- Depending on the statistical test, level of significance may be asked. There will be a default, in case the user does not know what level of significance is to be used.

UC-15: Statistical Report Generation

Summary: As a researcher, I need to be able to generate reports with actionable insights, which are automatically generated from UC-22

Priority: Essential

Use Frequency: Weekly

Prerequisites: Login as an administrator

Data must be entered (UC-09, UC-11, UC-25, or UC-26)

Scenario:

- Select the Statistical Analysis Module
- Select "Generate Report"
- Wait for report generation to complete
- File to be exported is either PDF or DOC. This file will be saved to the computer of the one running the report.

Use Cases by Encoders:

UC-16: Modify own profile

Summary: As a user, I should be able to edit my own profile information

Priority: Expected

Use Frequency: As there are changes in my information (seldom)

Prerequisites: Login as encoder

Scenario:

- Select "Profile Management"
- Select "Edit Profile"
- Change any of the fields available to the user, such as Address, Phone, Email, and Password.
- For password changes, place the current password, and the new password twice.

UC-17: Import from Excel / CSV

Summary: As an encoder, I need to be able to import encoded data from Excel/CSV.

Priority: Expected

Use Frequency: Every other day

Prerequisites: Login as encoder

Scenario:

- Select the Data Entry module
- Select "Import from File"
- Select the file containing tabulated surveys. This file will follow a specific format.
- There will be a confirmation that data is OK.
- Data will be viewable within the Data Entry Module.

UC-18: Data Entry

Summary: As an encoder, I need to be able to manually input questionnaire results.

Priority: Essential

Use Frequency: Every other day

Prerequisites: Login as encoder

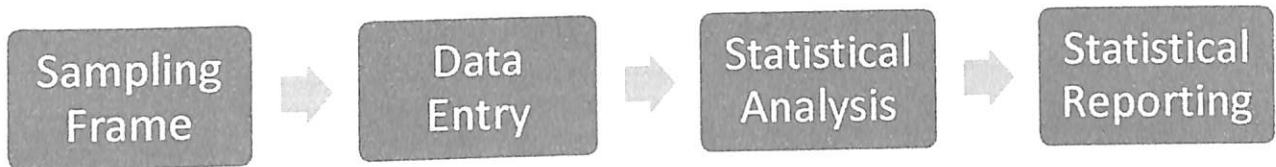
Scenario:

- Select the Data Entry module
- Select "Manual Data Entry"
- Enter answers to the survey questions. Click "OK" when done.

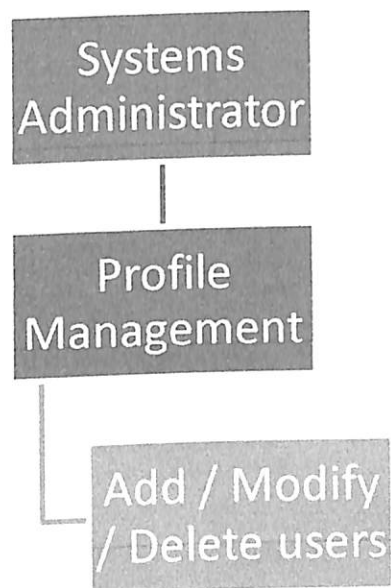
- Confirm that the answers are OK.
- There will be a confirmation that the answers to one questionnaire have been saved.

USE CASE DIAGRAMS

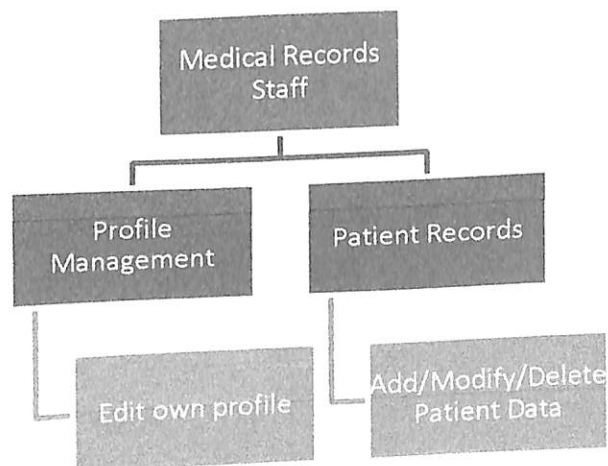
Process Diagram for the Information System



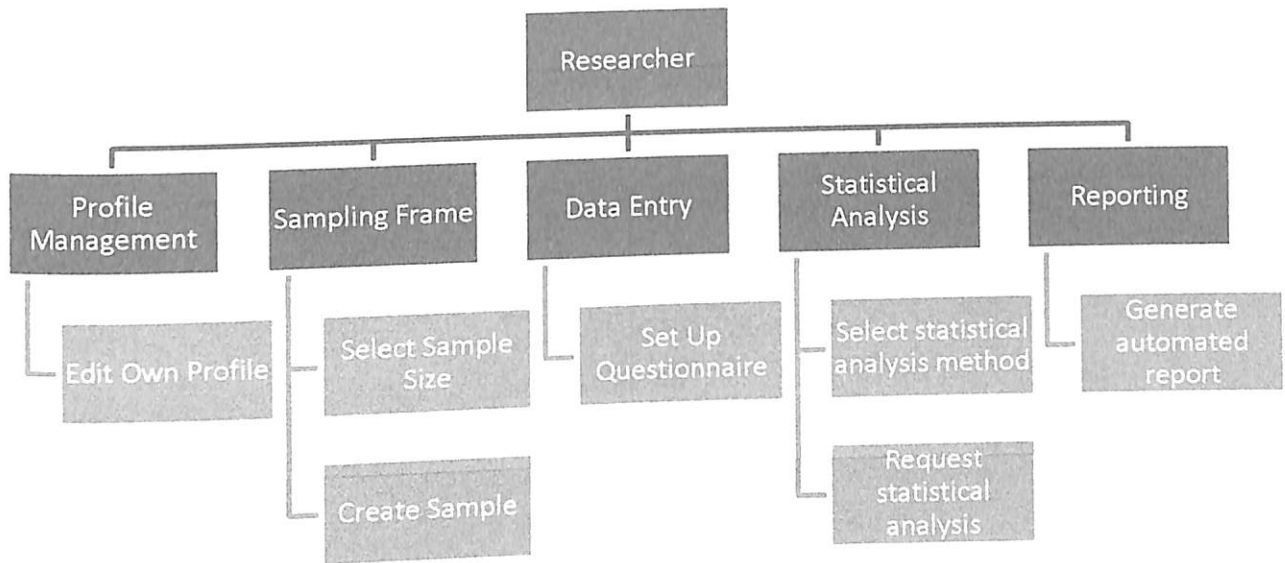
Systems Administrator Use Case Diagram



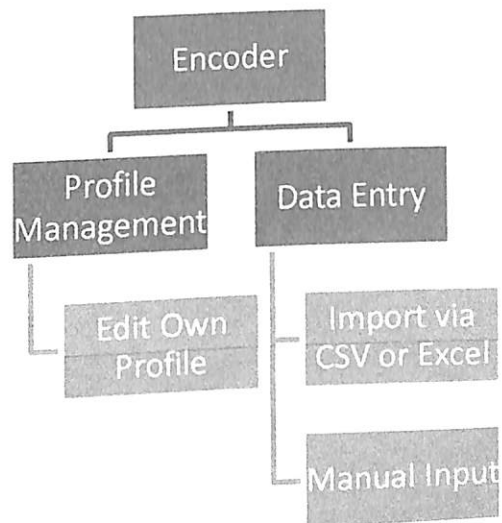
Medical Records Staff Use Case Diagram



Researcher Use Case Diagram



Encoder use case diagram



FUNCTIONAL REQUIREMENTS

The Survey Statistics Information System aims to provide statistical sampling capability, as well as automated statistical analysis of surveyed patients, to the Novaliches District Hospital. Here are some of its functional requirements:

Release Information	
Project	Survey Statistics Information System
Internal Release Number	1.0
Related Documents	Use Case Suite Use Cases Use Case Diagram Features

FEATURE SET

Features by Priority:

Essential

- F-00: Main Menu
- F-01: Login/Logout
- F-02: Profile Management
- F-03: Sampling Frame
- F-04: Data Entry Module
- F-05: Statistical Analysis Module
- F-06: Report Generation Module

Features by Risk:

1-Risk:

- F-00: Main Menu
- F-01: Login/Logout
- F-02: Profile Management

F-03: Sampling Frame

F-04: Data Entry Module

Safe:

F-05: Statistical Analysis Module

F-06: Report Generation Module

Features by Functional Area:

Main Menu

F-00: Main Menu

Profile Management

F-02: Profile Management

Sampling System

F-03: Sampling Frame

Data Entry

F-04: Data Entry Module

Data Analysis

F-05: Statistical Analysis Module

F-06: Report Generation Module

FEATURES

F-00 Main Menu

Priority – Essential

Effort – Hours

Risk – 1-Risk – as this enables access, any delays in this will impact the project

Functional Areas – Main Menu

Use Cases <application access>

Description:

The main menu is the initial screen that will be presented to anyone who opens the program. It will contain a welcome note and an area by which the user is required to login before accessing more advanced tasks.

Once the user has logged in, the main menu will contain links to the other modules, depending on the user's access level.

- Profile Management – for all

- Sampling Frame – for system administrators, medical records staff, and researchers

- Data Entry Module – for system administrators, researchers, and encoders

- Statistical Analysis Module – for system administrators and researchers

It should be noted that system administrators have access to all modules within the project.

Notes and Questions: Access levels to be determined by system administrators via profile management

F-01 Login/Logout

Priority – Essential

Effort – (coding) – Hours

Risk – 1-Risk: If not coded correctly, this can cause user lockouts on the database.

Functional Areas – Main Menu

Use Cases <Login>

Description:

From the main menu, a user should be able to log in to access other modules, depending on the user's access level.

There should be capability for a logged-in user to log out after they are done editing. Once logged out, the system goes back to the main menu.

Notes and Questions: Access levels to be determined by system administrators via profile management

F-02 Profile Management

Priority – Essential

Effort – Hours

Risk – 1-Risk – as this determines user access, any delays in this will impact the project

Functional Areas – Profile Management

Use Cases <Profile Creation>

Description:

There is an area where user profiles are created and/or deleted by system administrators. This area should also be accessible to users, who can edit their own profiles, updating fields such as:

- Name
- Address
- Contact Number
- Email
- Additional Info
- Password

For profile creation/deletion, there would be another view, with the same fields, as well as:

- Role (whether Medical Records staff, Encoder, Researcher, or System Administrator)

There would be buttons for “Save Profile”, “Delete Profile” (only visible to administrators”, and “Create Profile (only visible to administrators)

Notes and Questions: Access levels to be determined by system administrators via profile management

F-03 Sampling Frame

Priority – Essential

Effort – Days

Risk – 1-Risk – Issues in this will affect statistical analysis

Functional Areas – Sampling System

Use Cases <Sampling>

Description:

The sampling frame contains the entire universe, or list of all patients to be surveyed. This is done via interface to the existing patient database of the Novaliches District Hospital. This module also contains a way to create random samples, based on either the optimum sample size (which is based on current statistical best practice), or on a user-defined sample size).

Output from this module would consist of a list of sampled patients. They could be identified by name or by patient ID number. The same should be exportable via CSV or Excel, to facilitate identification and implementation of the survey.

Notes and Questions: Different sampling methods may need to be introduced here

F-04 Data Entry

Priority – Essential

Effort – Days

Risk – 1-Risk – Issues in this will affect statistical analysis

Functional Areas – Data Entry

Use Cases <Profile Creation>

Description:

The Data Entry Module consists of three submodules:

1. Questionnaire Setup
2. Manual Data Entry
3. Import via Excel or CSV

Questionnaire setup is a one-time process for each survey – the number of questions, as well as the type of questions (whether the answers are numeric, scale (1-10), free text, or multiple choice) is input.

Manual Data Entry depends on the setup of the questionnaire – it can be simply inputting numbers, or a drop-down for the scale, typing free text, or picking choices

The data could also be imported via Excel or CSV, if tabulation was done in a separate process.

Once imported, the data should take the form of an SQL database containing the following:

- Patient ID Number (names need not be entered, to maintain confidentiality)
- Answers to each question

Notes and Questions:

SQL database should be readable by the statistical analysis module.

F-05 Statistical Analysis Module

Priority – Essential

Effort – Days

Risk – Safe

Functional Areas – Data Analysis

Use Cases <Profile Creation>

Description:

There should be a way for the user to select what statistical analyses will be done within the system: whether hypothesis testing (where the null hypothesis needs to be selected), correlation testing (between two variables), and relationship testing (between two variables).

If the analysis selected is hypothesis testing, then there must be a way to input the null hypothesis, as well as the significance level (which is, by default, 0.05, as per generally accepted statistical practice).

If the analysis selected is correlation or relationship testing, then the two variables for comparison are selected.

At the backend, the sample size is evaluated, as the statistical analysis approach will depend on the sample size.

Notes and Questions:

SQL database from feature F-04 should be readable by the statistical analysis module.

F-06 Statistical Report Generation

Priority – Essential

Effort – Days

Risk – Safe

Functional Areas – Data Analysis

Use Cases <Profile Creation>

Description:

This should be able to read from feature F-05 and create statistical reports, which contain:

1. Descriptive statistics for each of the questions entered
2. Visualizations for descriptive statistics in (1)
3. Word cloud analysis for any open-ended, free text questions
4. Statistical tests requested in feature F-05
5. Visualizations for statistical tests in (4), as appropriate
4. Simple text writeup of the statistical tests and descriptive statistics using statistical natural language generation APIs

Notes and Questions:

SQL database from feature F-04 should be readable by the statistical analysis module.

NON FUNCTIONAL REQUIREMENTS

Usability Requirements

This system is designed to provide easy access to statistical analysis and reporting to users. The main targets of this project, the staff of the Novaliches District Hospital, have minimal if any statistical training; hence, one of the objectives of this project is to provide a user-friendly interface to the entire program.

Ease of access and usability is provided via a GUI which can be logged in to. There will be clickable links (which, to enhance user experience, will take the form of icons), which will lead to each module. Each module may request for data, but pre-defined defaults will be placed; in case the user does not know what to place, he may use the default value, which is in line with statistical best practices.

Reliability and Up-Time requirements

User requirement is for a continuous survey. As such, the system must be reliable, and capable of continuous uptime. This is done through the assistance of the system administrators, who should be able to respond and fix any bugs that occur throughout the lifetime of the application. This project shall be monitored by the administrators, who would have the power to effect small changes in the program without contacting the developer.

Security requirements

As the data being accessed by the application includes the hospital's confidential patient records, there should be a high level of security built in to this application. Logins should be instituted for all users, consisting of a username and password. The only users that can change patient records are the medical records staff and the administrators. Other users cannot access patient records via this program, only having access to the samples (for researchers).

Data encoders would neither have access to the patient database nor the statistical analysis, since they do not have a need to access either module.

Performance and Scalability Requirements

The system is expected to show good performance and scalability, as it is a light program, which should be able to handle both small samples (from a small patient database), and large samples (from a large patient database).

Maintainability and Upgradability Requirements

These requirements are inextricably tied with reliability, performance, and scalability. Proper maintenance by the system administrators will promote easy maintainability. The system is expected to be upgradeable, depending on whether further statistical analyses are needed (as the hospital transitions to a Level 3 teaching hospital with more advanced clinical trials).

ENVIRONMENTAL REQUIREMENTS

System Hardware Requirements

The application will be installed in the existing hardware systems of the Novaliches District Hospital. As this hospital is a government-run entity, outdated systems are expected. As such, the minimum specifications of this information system are expected to closely align with the minimum specifications of Windows Vista:

- 800-megahertz (MHz) processor (either 32 or 64-bit)
- 512 megabytes (MB) of system memory.
- Minimum HDD space of 500MB

System Software Requirements

The application is expected to run on a Windows platform (32-bit) PC. It is not expected to require an excessive amount of computer resources, since the modules are expected to be thin and light. The backend will be via R (via RApache), Apache server, PHP, and MySQL; the frontend is expected to be Web – a web browser such as Firefox or Microsoft Internet Explorer is required.

Application Program Interfaces

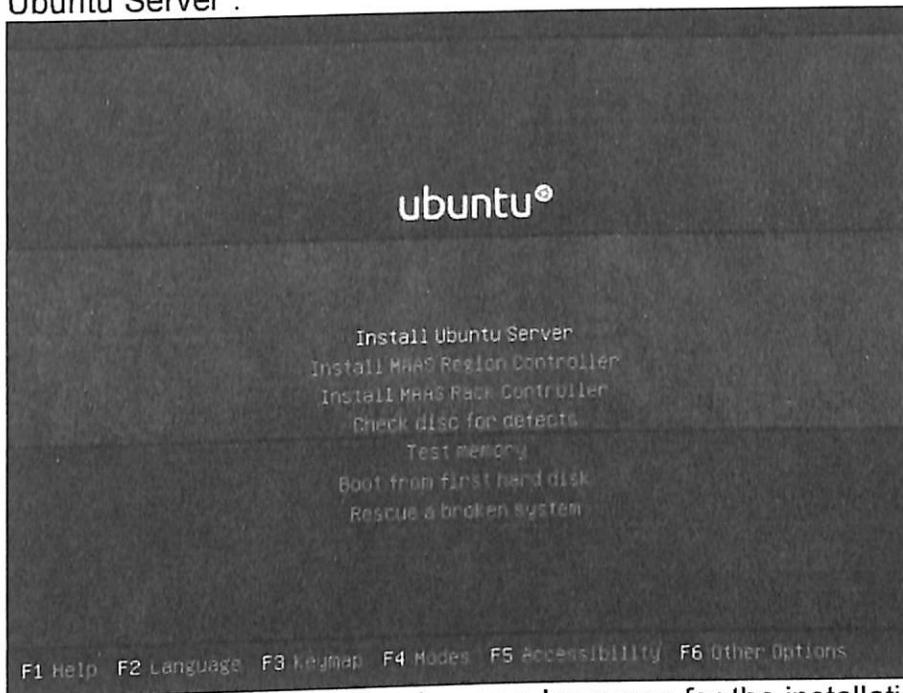
The development of this program did not require any APIs to be called.

Data Import and Export Requirements

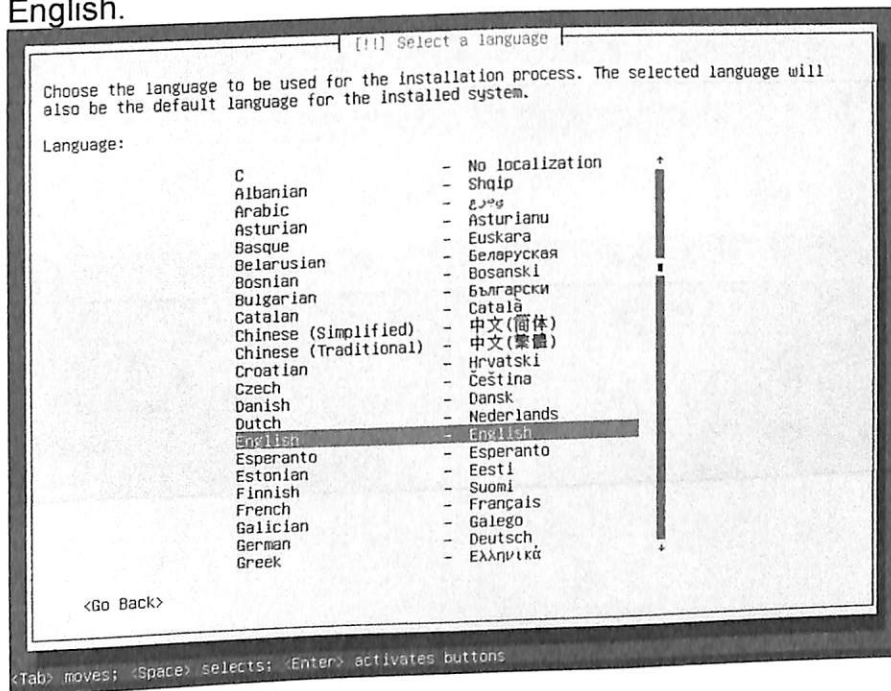
Survey samples (as culled from the patient database) are expected to be exported to CSV. This is part of feature F-03. In addition, data entry may be facilitated by a CSV input file, which is part of feature F-04.

Annex E. Setup Guide

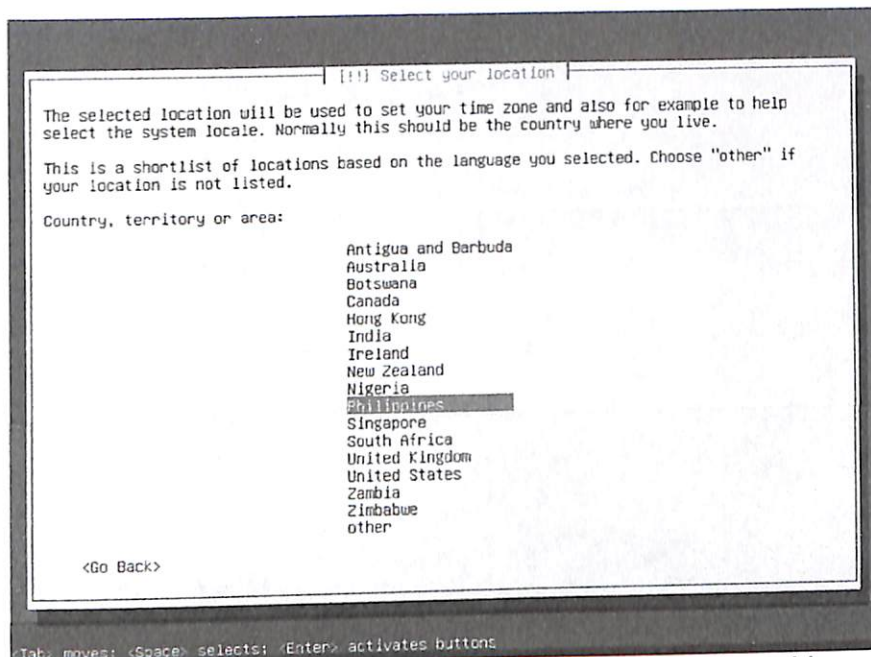
1. To install Ubuntu, insert a CD containing the installer. Then, Select "Install Ubuntu Server":



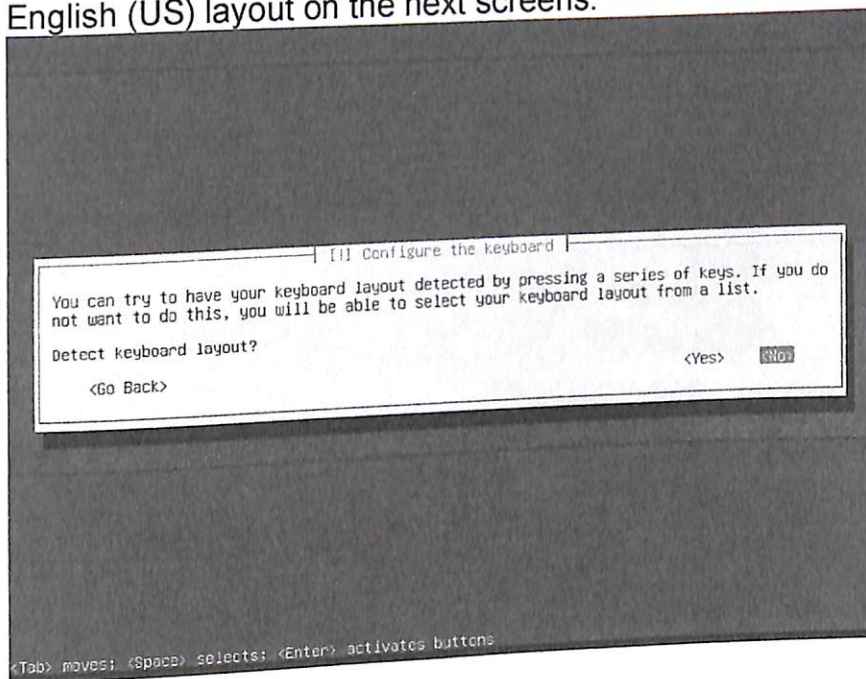
2. You will again be asked to choose a language for the installation itself. Select English.

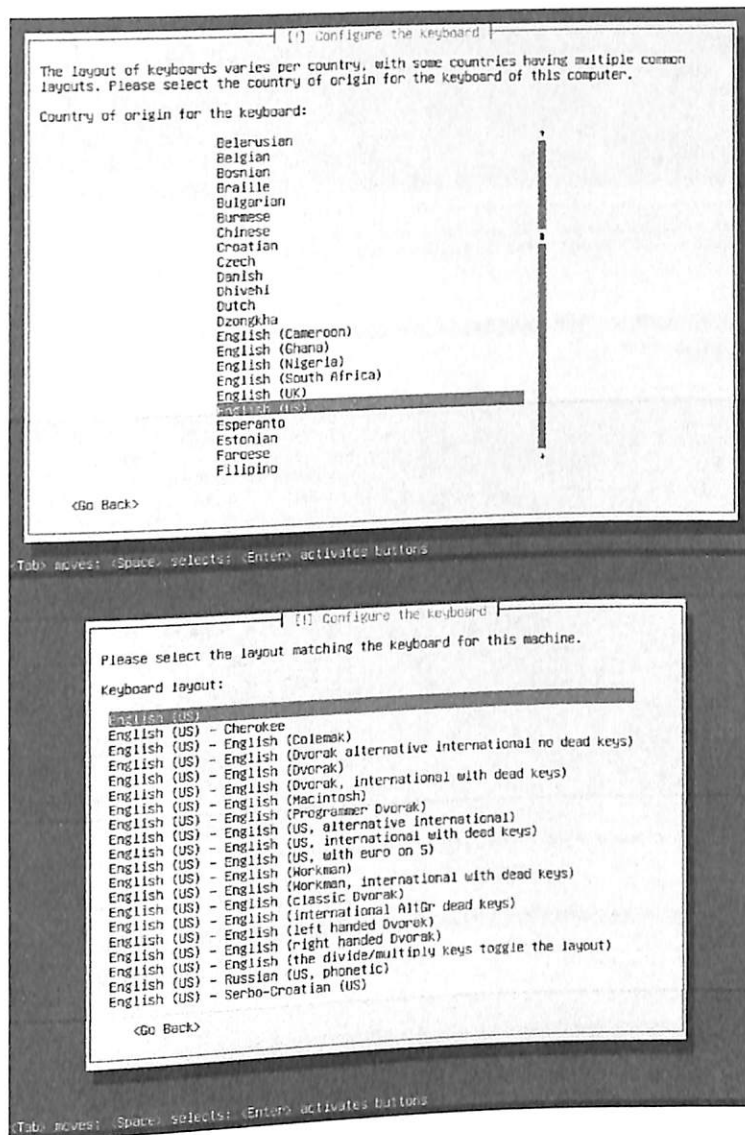


3. You will be asked for your location. Since we are in the Philippines, we select the same from the menu.

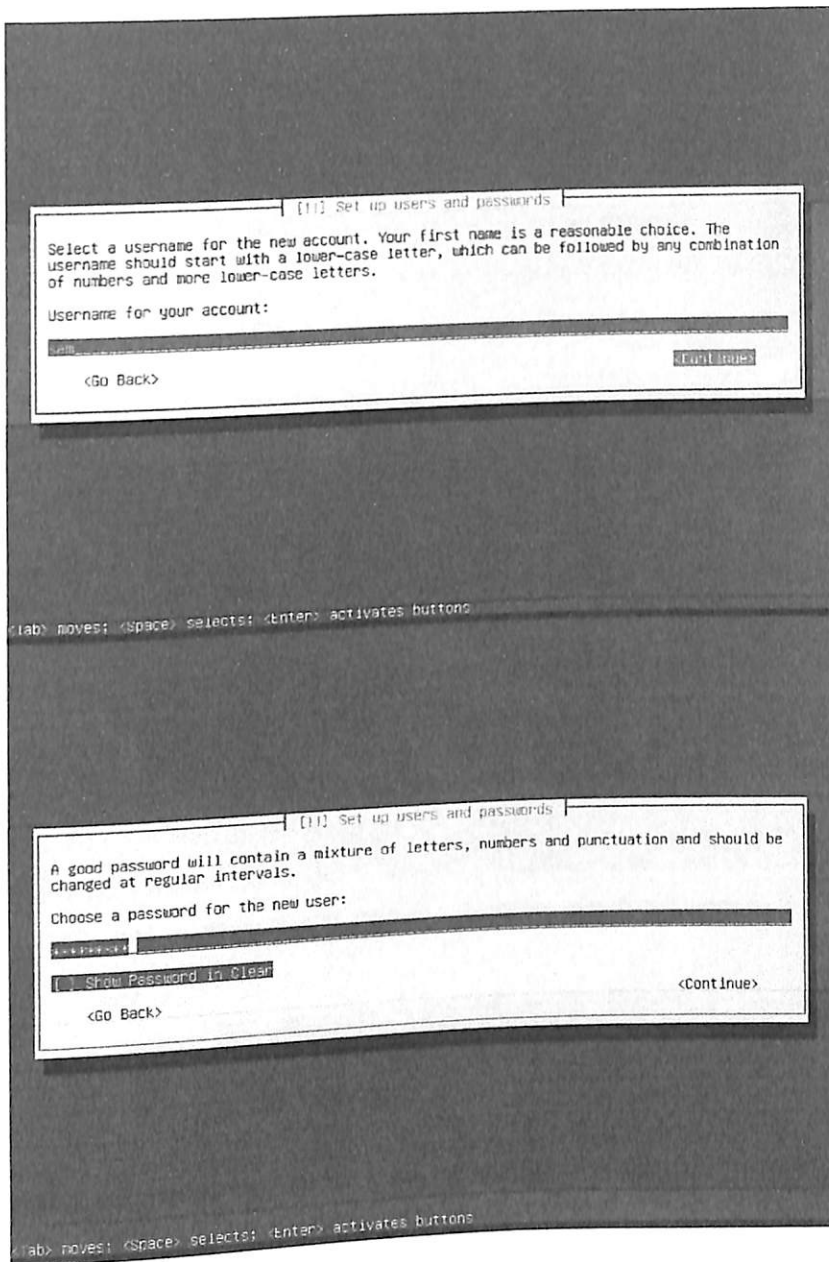


4. The program will ask to automatically detect keyboard layout. For a faster install (assuming a US English keyboard), we select "No" here and select the English (US) layout on the next screens.

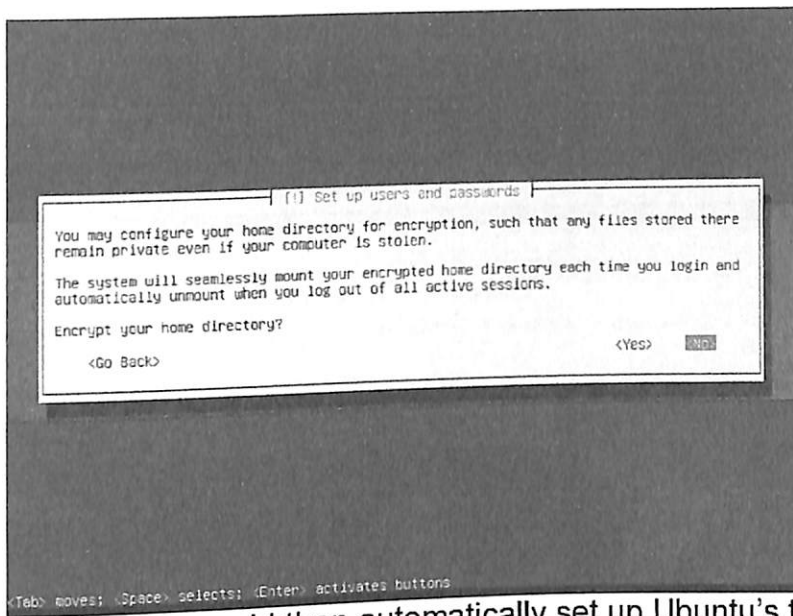




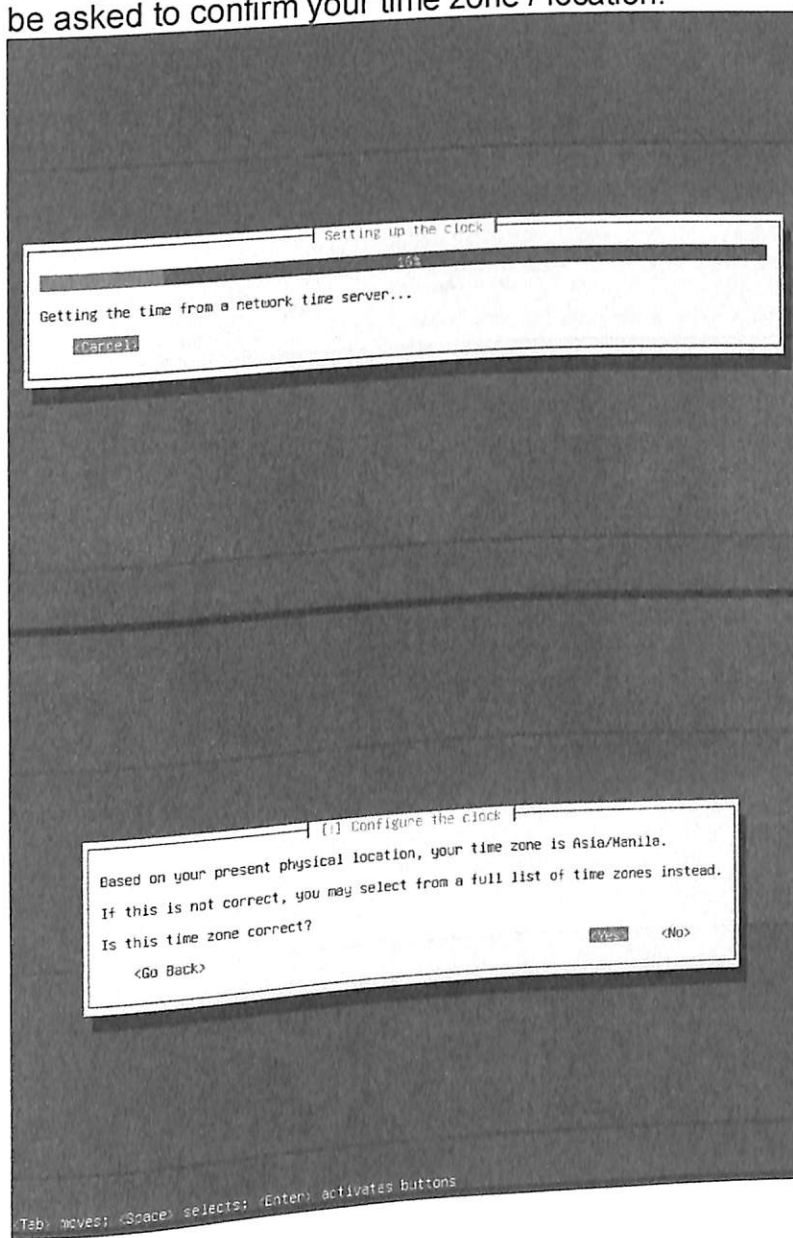
- After this, the next few screens will ask for you to input your hostname (this would be your computer name), user account names, and user password. Choose a hostname and password you can easily remember. Remember to choose strong passwords.



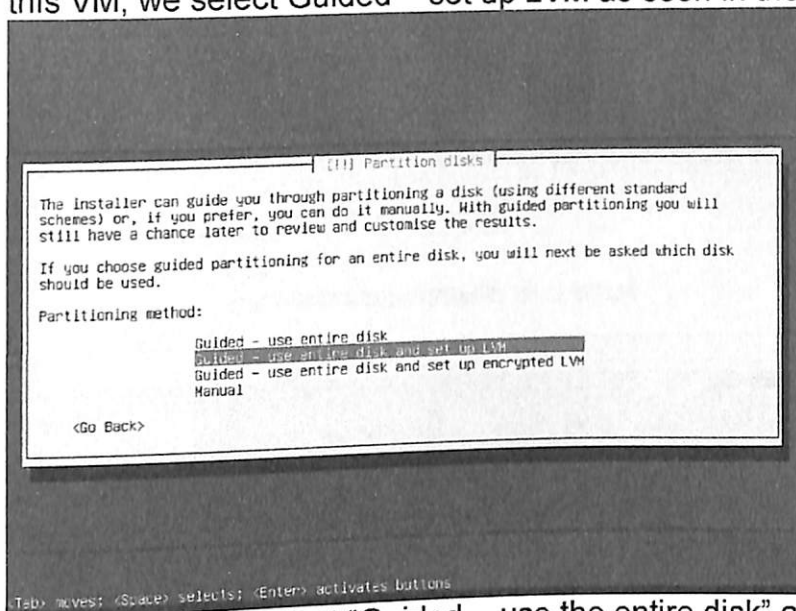
6. Setup may include steps to encrypt the home directory. You may select "Yes" or "No" depending on whether you prioritize efficiency (No) or file security (Yes).



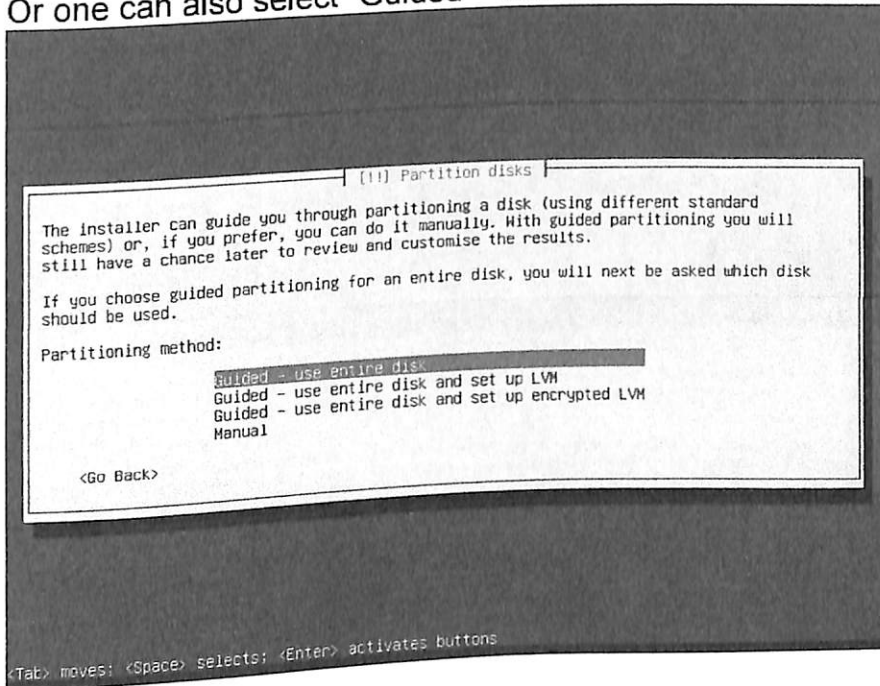
7. The system would then automatically set up Ubuntu's time server. You may be asked to confirm your time zone / location:



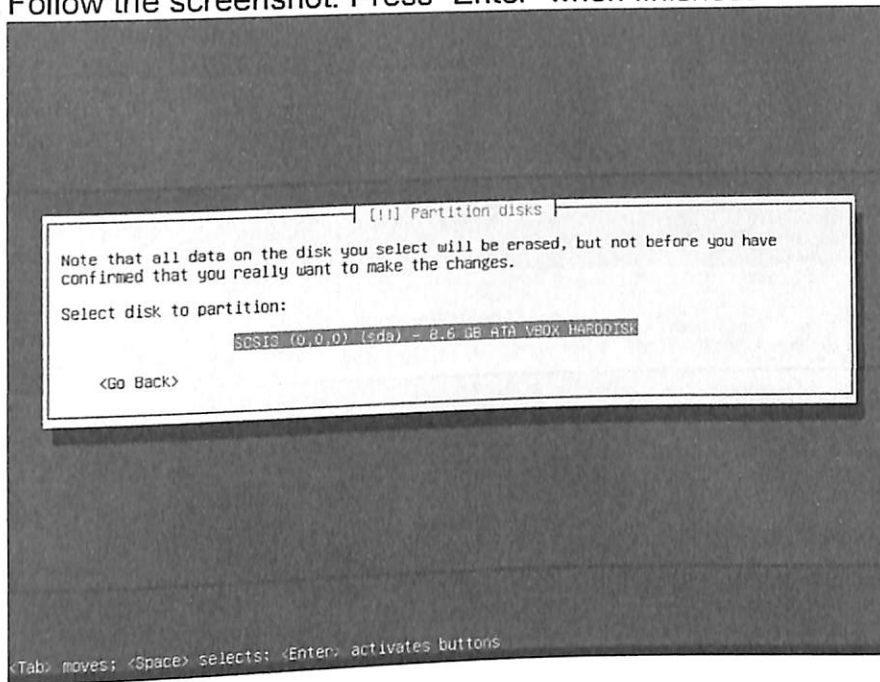
8. The Ubuntu server installation can automatically partition the disk used. For this VM, we select Guided – set up LVM as seen in the screen below:



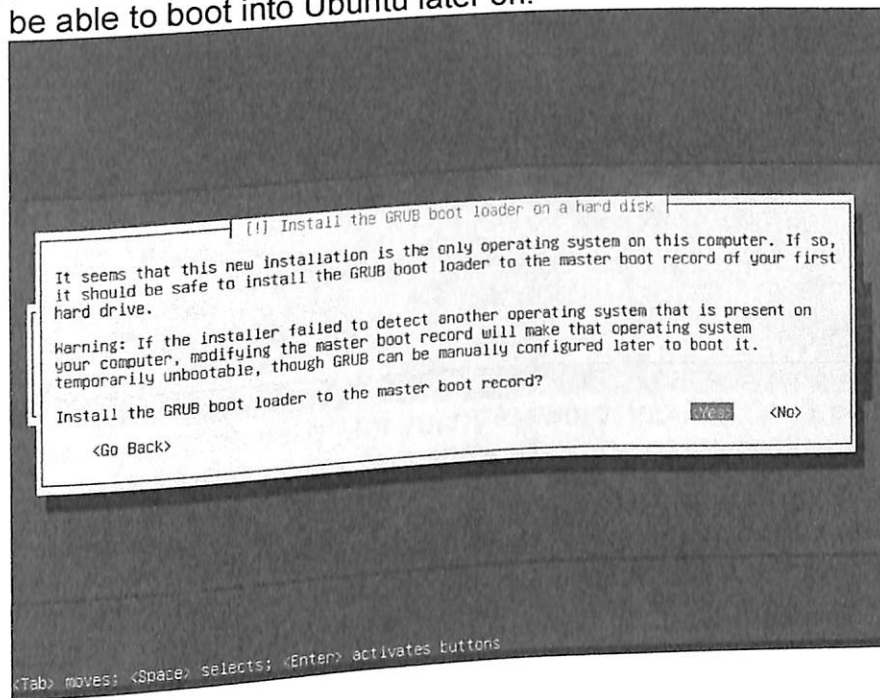
9. Or one can also select "Guided – use the entire disk" as below:



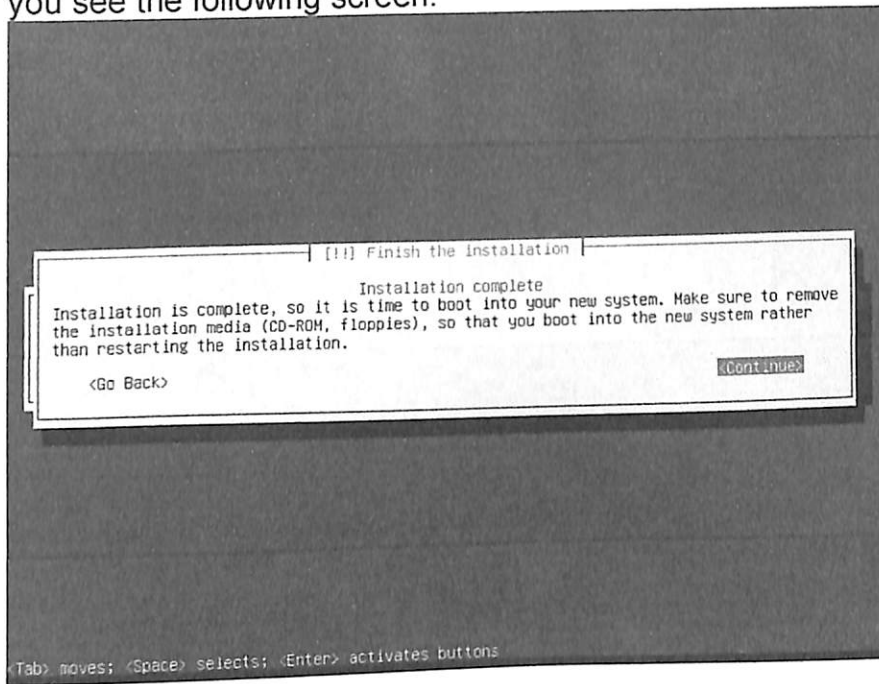
10. Follow the screenshot. Press "Enter" when finished.



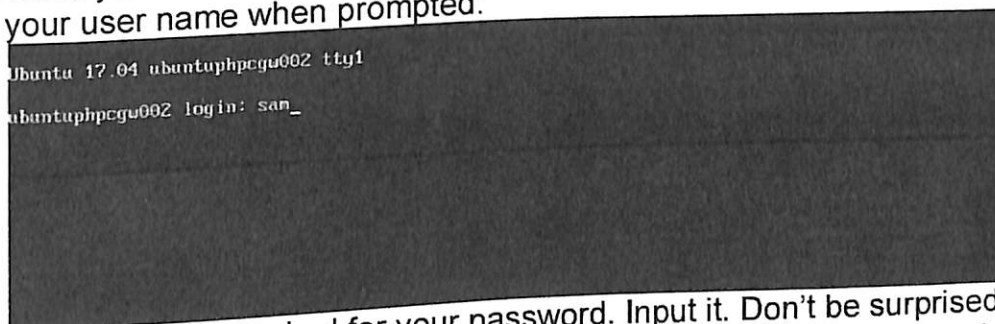
11. If asked to install the GRUB boot loader, select "Yes". Otherwise, you may not be able to boot into Ubuntu later on:



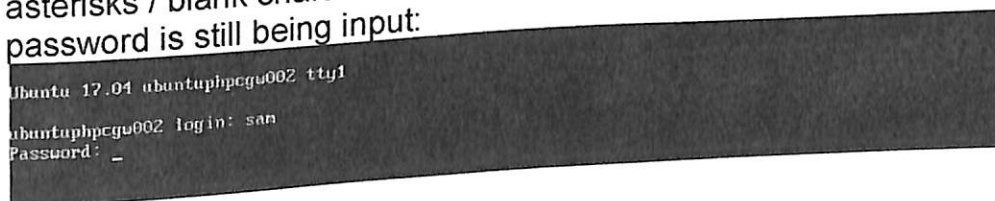
12. Wait for the screens to finish loading. Installation will have completed when you see the following screen:



13. Once you select "Continue", you will be presented with your first login. Input your user name when prompted:



14. You will also be asked for your password. Input it. Don't be surprised if no asterisks / blank characters show while the password is being typed – the password is still being input:



15. You should see a screen like this:

```
Ubuntu 17.04 ubuntu@pcgu002 tty1
ubuntu@pcgu002 login: sam
Password:
Welcome to Ubuntu 17.04 (GNU/Linux 4.10.0-19-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

Ubuntu 12.04 LTS end-of-life is April 25, 2017 -- Upgrade your Precise systems!
$ sudo do-release-upgrade -n server

21 packages can be updated.
14 updates are security updates.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

sam@ubuntu@pcgu002:~$
```

16. Among the things that need to be installed next is vim (to edit text files). This is done by typing the command "sudo apt-get install vim".

```
Ubuntu 17.04 ubuntu@pcgu002 tty1
ubuntu@pcgu002 login: sam
Password:
Welcome to Ubuntu 17.04 (GNU/Linux 4.10.0-19-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

Ubuntu 12.04 LTS end-of-life is April 25, 2017 -- Upgrade your Precise systems!
$ sudo do-release-upgrade -n server

21 packages can be updated.
14 updates are security updates.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

sam@ubuntu@pcgu002:~$ sudo apt-get install vim
```

In this command, sudo forces this user to become root, that is, to run the command using elevated privileges. The apt-get command is used to get packages for installation, and install them. Vim is a package used to install text files.

/Tip: If you don't want to always use sudo, input the command *sudo su*. This forces all commands to be run under the root user.

17. Since the command requires elevated privileges, it will request the user password again. If one makes a mistake in typing the password, he could always try again.

```
Ubuntu 17.04 ubuntu@pcgu002 tty1
ubuntu@pcgu002 login: san
Password:
Welcome to Ubuntu 17.04 (GNU/Linux 4.10.0-19-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

Ubuntu 12.04 LTS end-of-life is April 25, 2017 -- Upgrade your Precise system!
$ sudo do-release-upgrade -n server

21 packages can be updated.
14 updates are security updates.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

san@ubuntu@pcgu002:~$ sudo apt-get install vim
[sudo] password for san:
Sorry, try again.
[sudo] password for san: _
```

18. Once installation is complete, the screen should look like this:

```
ubuntu@pcgu002 login: san
Password:
Welcome to Ubuntu 17.04 (GNU/Linux 4.10.0-19-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

Ubuntu 12.04 LTS end-of-life is April 25, 2017 -- Upgrade your Precise system!
$ sudo do-release-upgrade -n server

21 packages can be updated.
14 updates are security updates.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

san@ubuntu@pcgu002:~$ sudo apt-get install vim
[sudo] password for san:
Sorry, try again.
[sudo] password for san:
Reading package lists... Done
Building dependency tree
Reading state information... Done
vim is already the newest version (2:8.0.0095-1ubuntu3).
0 upgraded, 0 newly installed, 0 to remove and 21 not upgraded.
san@ubuntu@pcgu002:~$ _
```

19. Next, enter the command "sudo apt-get update". This updates the list of what files can be downloaded. This is needed to get the latest version of other applications necessary to run Moodle.

The programs included with the Ubuntu system are free software; the exact distribution terms for each program are described in the individual files in /usr/share/doc/*/*copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by applicable law.

To run a command as administrator (user "root"), use "sudo <command>". See "man sudo_root" for details.

```
sam@buntuphpcgu002:~$ sudo apt-get install vim
[sudo] password for sam:
Sorry, try again.
[sudo] password for sam:
Reading package lists... Done
Building dependency tree
Reading state information... Done
vim is already the newest version (2:8.0.0095-1ubuntu3).
0 upgraded, 0 newly installed, 0 to remove and 21 not upgraded.
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$ sudo
```

```
No command 'sudox' found, did you mean:
Command 'sudo' from package 'sudo' (main)
Command 'sudo' from package 'sudo-ldap' (universe)
sudox: command not found
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$ sudo apt-get update
```

Installing the LAMP Stack (Linux/Apache/PHP/MySQL)

1. After this, we install necessary files such as apache2 (the server program), MySQL (both mysql-client and mysql-server), to serve as the database, and php7.0 and related libraries (for a server-side programming language). This comprises the LAMP stack (Linux, Apache, MySQL, and PHP). Type in below command to install all of these programs in one go:

```
sudo apt-get install apache2 mysql-client mysql-server php7.0
```

libapache2-mod-php7.0

See "man sudo_root" for details.

```
sam@buntuphpcgu002:~$ sudo apt-get install vim
[sudo] password for sam:
Sorry, try again.
[sudo] password for sam:
Reading package lists... Done
Building dependency tree
Reading state information... Done
vim is already the newest version (2:8.0.0095-1ubuntu3).
0 upgraded, 0 newly installed, 0 to remove and 21 not upgraded.
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
apt-get install sudo vim
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$ sudo
```

```
No command 'sudox' found, did you mean:
Command 'sudo' from package 'sudo' (main)
Command 'sudo' from package 'sudo-ldap' (universe)
sudox: command not found
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$
sam@buntuphpcgu002:~$ sudo apt-get update
```

```
Hit:1 http://ph.archive.ubuntu.com/ubuntu zesty InRelease
Get:2 http://ph.archive.ubuntu.com/ubuntu zesty-updates InRelease [89.2 kB]
Hit:3 http://ph.archive.ubuntu.com/ubuntu zesty-backports InRelease
Get:4 http://security.ubuntu.com/ubuntu zesty-security InRelease [89.2 kB]
Fetched 178 kB in 4s (40.3 kB/s)
Reading package lists... Done
sam@buntuphpcgu002:~$ sudo apt-get install apache2 mysql-client mysql-server php7.0 libapache2-mod-
php7.0
```

2. If, in the middle of an apt-get command, you get "After this operation, XX MB of additional disk space will be used. Do you want to continue?" – always answer with Y. Otherwise the install will not continue.


```

san@ubuntu:~$ sudo apt-get update
Hit:1 http://ph.archive.ubuntu.com/ubuntu zesty InRelease
Hit:2 http://ph.archive.ubuntu.com/ubuntu zesty-updates InRelease [89.2 kB]
Hit:3 http://ph.archive.ubuntu.com/ubuntu zesty-backports InRelease
Get:4 http://security.ubuntu.com/ubuntu zesty-security InRelease [89.2 kB]
Fetched 178 kB in 4s (40.3 kB/s)
Reading package lists... Done
san@ubuntu:~$ sudo apt-get install apache2 mysql-client mysql-server php7.0 libapache2-mod-
php7.0
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  apache2-bin apache2-data apache2-utils libaio1 libapr1 libaprutil1 libaprutil1-dbd-sqlite3
  libaprutil1-ldap libbrotli-fast-perl libbrotli-perl libcode-locale-perl libevent-core-2.0-5
  libfcgi-perl libhtml-parser-perl libhtml-tagset-perl libhtml-template-perl libhttp-date-perl
  libhttp-message-perl libio-html-perl liblua5.2-0 liblwp-mediatypes-perl libtimedate-perl
  liburi-perl libwrap0 mysql-client-5.7 mysql-client-core-5.7 mysql-common mysql-server-5.7
  mysql-server-core-5.7 php-common php7.0-cli php7.0-common php7.0-json php7.0-opcache
  php7.0-readline ssl-cert tcpd
Suggested packages:
  www-browser apache2-doc apache2-suexec-pristine | apache2-suexec-custom php-pear
  libdata-dump-perl libipc-sharedcache-perl libwww-perl mailx tinycss openssl-blacklist
The following NEW packages will be installed:
  apache2 apache2-bin apache2-data apache2-utils libaio1 libapache2-mod-php7.0 libapr1 libaprutil1
  libaprutil1-dbd-sqlite3 libaprutil1-ldap libbrotli-fast-perl libbrotli-perl libcode-locale-perl
  libevent-core-2.0-5 libfcgi-perl libhtml-parser-perl libhtml-tagset-perl libhtml-template-perl
  libhttp-date-perl libhttp-message-perl libio-html-perl liblua5.2-0 liblwp-mediatypes-perl
  libtimedate-perl liburi-perl libwrap0 mysql-client mysql-client-core-5.7 mysql-client-core-5.7
  mysql-common mysql-server mysql-server-5.7 mysql-server-core-5.7 php-common php7.0 php7.0-cli
  php7.0-common php7.0-json php7.0-opcache php7.0-readline ssl-cert tcpd
0 upgraded, 42 newly installed, 0 to remove and 21 not upgraded.
Need to get 25.4 MB of archives.
After this operation, 193 MB of additional disk space will be used.
Do you want to continue? [Y/n]

```

3. You will be asked to nominate a password for your MySQL root user. Do so, for security purposes. I use the same password as my Ubuntu login to make this easier to remember. Type it twice, then select "OK" and press Enter.

Package configuration

Configuring mysql-server-5.7

While not mandatory, it is highly recommended that you set a password for the MySQL administrative "root" user.

If this field is left blank, the password will not be changed.

New password for the MySQL "root" user:

<OK>

Package configuration

Configuring mysql-server-5.7

Repeat password for the MySQL "root" user:

<Ok>

Configuring mysql-server-5.7

Repeat password for the MySQL "root" user:

<Ok>

Selecting previously unselected package libapril:amd64.
Reading database ... 59659 files and directories currently installed.
Preparing to unpack .../00-libapril_1.5.2-5_amd64.deb ...
Unpacking libapril:amd64 (1.5.2-5) ...

4. The rest of the applications will be set up on the system.

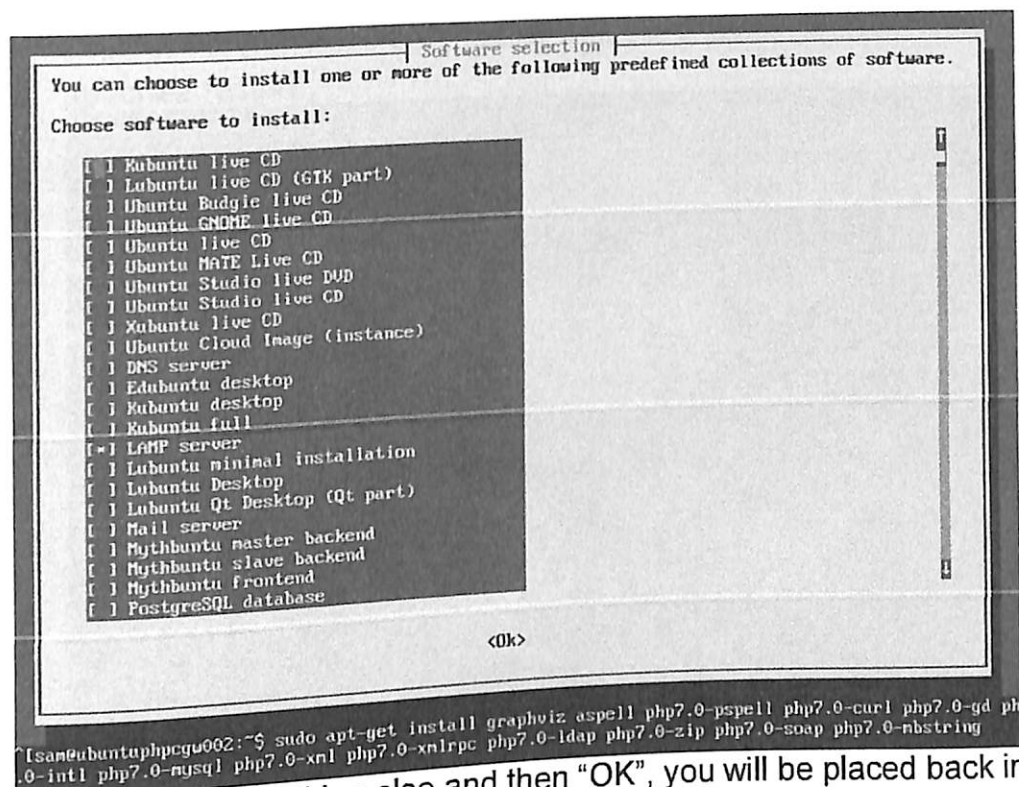
```
(Reading database ... 59659 files and directories currently installed.)
Preparing to unpack .../00-libapr1_1.5.2-5_and64.deb ...
Unpacking libapr1:amd64 (1.5.2-5) ...
Selecting previously unselected package libaprutil1:amd64.
Preparing to unpack .../01-libaprutil1_1.5.4-3_and64.deb ...
Unpacking libaprutil1:amd64 (1.5.4-3) ...
Selecting previously unselected package libaprutil1-dbd-sqlite3:amd64.
Preparing to unpack .../02-libaprutil1-dbd-sqlite3_1.5.4-3_and64.deb ...
Unpacking libaprutil1-dbd-sqlite3:amd64 (1.5.4-3) ...
Selecting previously unselected package libaprutil1-ldap:amd64.
Preparing to unpack .../03-libaprutil1-ldap_1.5.4-3_and64.deb ...
Unpacking libaprutil1-ldap:amd64 (1.5.4-3) ...
Selecting previously unselected package liblua5.2-0:amd64.
Preparing to unpack .../04-liblua5.2_5.2.4-1.1build1_and64.deb ...
Unpacking liblua5.2-0:amd64 (5.2.4-1.1build1) ...
Selecting previously unselected package apache2-bin.
Preparing to unpack .../05-apache2-bin_2.4.25-3ubuntu2_and64.deb ...
Unpacking apache2-bin (2.4.25-3ubuntu2) ...
Selecting previously unselected package apache2-utils.
Preparing to unpack .../06-apache2-utils_2.4.25-3ubuntu2_and64.deb ...
Unpacking apache2-utils (2.4.25-3ubuntu2) ...
Selecting previously unselected package apache2-data.
Preparing to unpack .../07-apache2-data_2.4.25-3ubuntu2_all.deb ...
Unpacking apache2-data (2.4.25-3ubuntu2) ...
Selecting previously unselected package apache2.
Preparing to unpack .../08-apache2_2.4.25-3ubuntu2_and64.deb ...
Unpacking apache2 (2.4.25-3ubuntu2) ...
Selecting previously unselected package mysql-common.
Preparing to unpack .../09-mysql-common_5.8+1.0.2ubuntu1_all.deb ...
Unpacking mysql-common (5.8+1.0.2ubuntu1) ...
Selecting previously unselected package libaio1:amd64.
Preparing to unpack .../10-libaio1_0.3.110-3_and64.deb ...
Unpacking libaio1:amd64 (0.3.110-3) ...
Selecting previously unselected package mysql-client-core-5.7.
Preparing to unpack .../11-mysql-client-core-5.7_5.7.17-0ubuntu1_and64.deb ...
Unpacking mysql-client-core-5.7 (5.7.17-0ubuntu1) ...
```

5. One can install the desktop onto the server:

```
sudo taskset

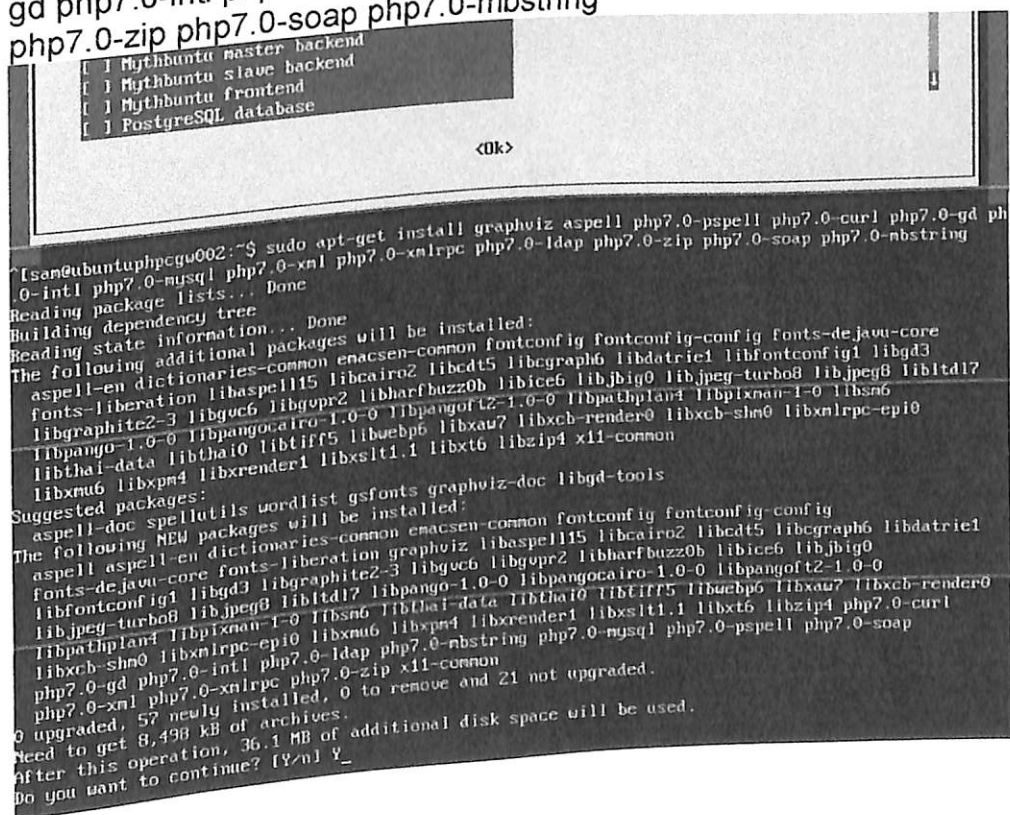
Enabling conf charset.
Enabling conf localized-error-pages.
Enabling conf other-vhosts-access-log.
Enabling conf security.
Enabling conf serve-cgi-bin.
Enabling site 000-default.
Created symlink /etc/systemd/system/multi-user.target.wants/apache2.service → /lib/systemd/system/apache2.service.
Created symlink /etc/systemd/system/multi-user.target.wants/apache-htcacheclean.service → /lib/systemd/system/apache-htcacheclean.service.
Setting up mysql-client (5.7.17-0ubuntu1) ...
Setting up php7.0-cli (7.0.15-1ubuntu4) ...
update-alternatives: using /usr/bin/php7.0 to provide /usr/bin/php (php) in auto mode
update-alternatives: using /usr/bin/phar7.0 to provide /usr/bin/phar (phar) in auto mode
update-alternatives: using /usr/bin/phar.phar7.0 to provide /usr/bin/phar.phar (phar.phar) in auto mode
Creating config file /etc/php/7.0/cli/php.ini with new version
Setting up mysql-server-5.7 (5.7.17-0ubuntu1) ...
update-alternatives: using /etc/mysql/mysql.cnf to provide /etc/mysql/my.cnf (my.cnf) in auto mode
Renaming removed key_buffer and myisam-recover options (if present)
Created symlink /etc/systemd/system/multi-user.target.wants/mysql.service → /lib/systemd/system/mysql.service.
Setting up mysql-server (5.7.17-0ubuntu1) ...
Setting up libapache2-mod-php7.0 (7.0.15-1ubuntu4) ...
Creating config file /etc/php/7.0/apache2/php.ini with new version
Module mpm_event disabled.
Enabling module mpm_prefork.
apache2_switch_mpm Switch to prefork
apache2_invoke: Enable module php7.0
Setting up php7.0 (7.0.15-1ubuntu4) ...
Processing triggers for libc-bin (2.24-0ubuntu2) ...
Processing triggers for ureadahead (0.100.0-19) ...
Processing triggers for systemd (232-21ubuntu2) ...
Processing triggers for ufw (0.35-4) ...
sam@ubuntu:~$ sudo taskset
```

Select Ubuntu Desktop:



6. If you select nothing else and then "OK", you will be placed back into the command line. You will need to install additional applications for the Moodle to function well. Type the below command. When asked, always reply with "Y" to continue the install.

```
sudo apt-get install graphviz aspell php7.0-pspell php7.0-curl php7.0-gd php7.0-intl php7.0-mysql php7.0-xml php7.0-xmlrpc php7.0-ldap php7.0-zip php7.0-soap php7.0-mbstring
```



7. Once done with the install, restart important services. Type "sudo service apache2 restart"

```
Setting up graphviz (2.38.0-1ubuntu1) ...
Processing triggers for libapache2-mod-php7.0 (7.0.15-1ubuntu4) ...
Processing triggers for libc-bin (2.24-9ubuntu2) ...
Processing triggers for ureadahead (0.100.0-19) ...
Processing triggers for systemd (232-21ubuntu2) ...
Processing triggers for dictionaries-common (1.27.2) ...
aspell-autobuildhash: processing: en [en-common].
aspell-autobuildhash: processing: en [en-variant_01].
aspell-autobuildhash: processing: en [en-variant_11].
aspell-autobuildhash: processing: en [en-variant_21].
aspell-autobuildhash: processing: en [en-u_accents-only].
aspell-autobuildhash: processing: en [en-wo_accents-only].
aspell-autobuildhash: processing: en [en_AU-variant_01].
aspell-autobuildhash: processing: en [en_AU-variant_11].
aspell-autobuildhash: processing: en [en_AU-u_accents-only].
aspell-autobuildhash: processing: en [en_AU-wo_accents-only].
aspell-autobuildhash: processing: en [en_CA-variant_01].
aspell-autobuildhash: processing: en [en_CA-variant_11].
aspell-autobuildhash: processing: en [en_CA-u_accents-only].
aspell-autobuildhash: processing: en [en_CA-wo_accents-only].
aspell-autobuildhash: processing: en [en_GB-ise-u_accents-only].
aspell-autobuildhash: processing: en [en_GB-ise-wo_accents-only].
aspell-autobuildhash: processing: en [en_GB-ize-u_accents-only].
aspell-autobuildhash: processing: en [en_GB-ize-wo_accents-only].
aspell-autobuildhash: processing: en [en_GB-variant_01].
aspell-autobuildhash: processing: en [en_GB-variant_11].
aspell-autobuildhash: processing: en [en_US-u_accents-only].
aspell-autobuildhash: processing: en [en_US-wo_accents-only].
san@buntuphpcgu002:~$
mysql-server      php7.0-mysql      sudo
apache2           php7.0            taskset
apt-get           php7.0-curl       update
aspell            php7.0-gd         vin
graphviz          php7.0-intl       php7.0-xmllrpc
install           php7.0-ldap       php7.0-zip
libapache2-mod-php7.0  php7.0-mbstring  sudo
mysql-client      php7.0-mbstring
san@buntuphpcgu002:~$ sudo service apache2 restart_
```

Once done with initial LAMP stack setup, the following programs need to be installed:

R:

To install R, the following needs to be done.

1. Set a CRAN mirror. We use the PH server since it's the nearest. Type the

command below:

```
sudo echo "deb http://cran.stat.upd.edu.ph/bin/linux/ubuntu xenial/" | sudo tee
-a /etc/apt/sources.list
```

2. Add R to Ubuntu Keyring:

Type these two commands in sequence:

```
gpg --keyserver keyserver.ubuntu.com --recv-key E084DAB9
gpg -a --export E084DAB9 | sudo apt-key add -
```

3. Install R. Type below:

```
sudo apt-get update
```

```
sudo apt-get install r-base r-base-dev
```

Rapache

Type the commands below in sequence:

```
sudo add-apt-repository ppa:opencpu/rapache
```

```
sudo apt-get update
```

```
sudo apt-get install libapache2-mod-r-base
```

OpenSSL.

Type the command below:

```
sudo apt-get update
```

```
sudo apt-get install libapache2-mod-r-base
```

Once in R, the following packages also need to be installed. In the R environment, type the following in sequence:

```
## Default repo
```

```
local({r <- getOption("repos")
```

```
  r["CRAN"] <- "https://cran.stat.upd.edu.ph"
```

```
  options(repos=r)
```

```
})
```

```
#Add library path. This is the path where R saves its other libraries
```

```
.libPaths("/home/sam/R/i686-pc-linux-gnu-library/3.2")
```

```
.Library.site <- file.path("/home/sam/R/i686-pc-linux-gnu-library/3.2")
```

```
#Install libraries
```

```
install.packages(pwr)
```

```
install.packages(MASS)
```

```
install.packages(jsonlite)
```

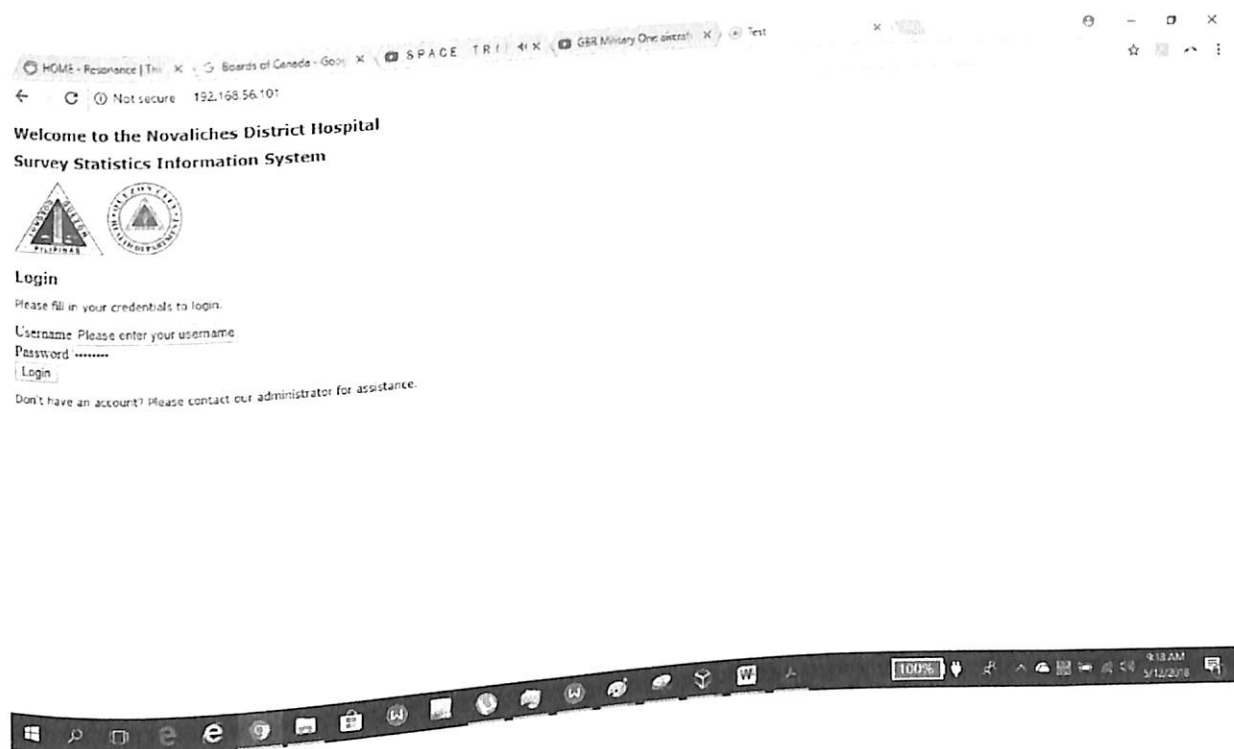
```
install.packages(splitstackshape)
```

```
install.packages(qwraps2)
```

Annex F. User Guide

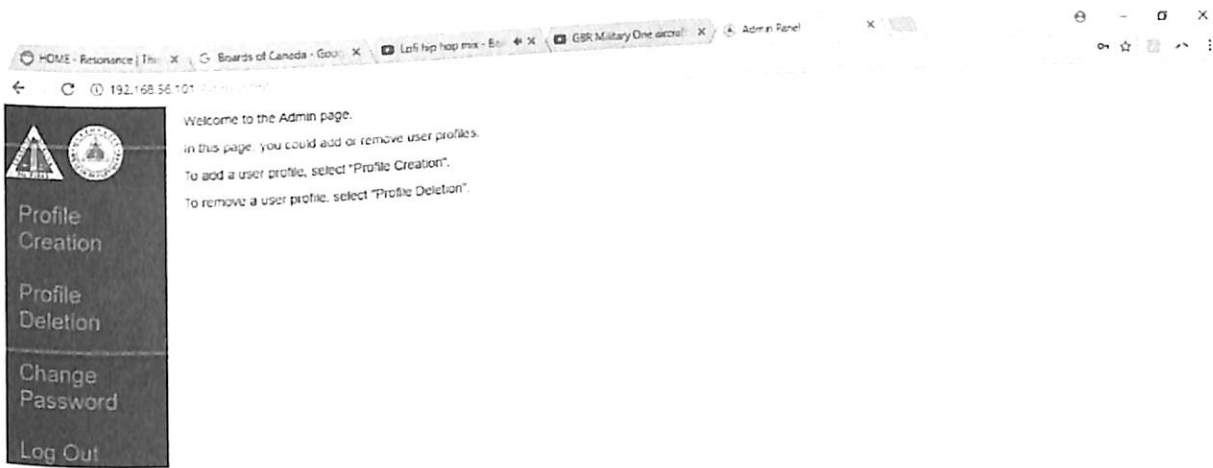
Login

Go to the current IP address of the system (192.168.56.101). You will be greeted by a login screen.



Enter your login details. If you have valid login details, you will automatically be redirected to your view of the system.

Administrator View



In the Administrator view, you will find four menu options: Profile Creation, Profile Deletion, Change Password, and Log Out.

Profile Management

Profile Management involves two menu options: Profile Creation and Profile Deletion.

Profile Creation



As an administrator, you can create new user profiles. To do this, fill out the form as follows:

Username: The User ID of the profile

Password: An initial password to access the profile. The profile's owner needs to change this once he receives his profile.

Confirm Password: Repeat the initial password in this box.

Name: The name of the profile's owner

Email: Email Address of the profile's owner

Role: Select a role for the profile: Admin, Researcher, Encoder, or Medical

Records.

The screenshot shows a web browser window with the address bar displaying '192.168.56.101/admin/profile_creation.php'. The page has a dark sidebar on the left with the following menu items: 'Admin Menu', 'Change Password', and 'Log Out'. The main content area is titled 'Create Profile' and contains the instruction 'Please fill this form to create an account.' Below this, there are input fields for 'Username' (containing 'TestEncoder'), 'Password' (masked with dots), 'Confirm Password' (masked with dots), 'Name' (containing 'Test User'), and 'Email' (containing 'test@testtest'). There is a dropdown menu for 'Role' with 'Encoder' selected. At the bottom of the form are 'Submit' and 'Reset' buttons.

Once the form is completed, click "Submit". A pop-up will show saying "Congratulations. User created successfully."

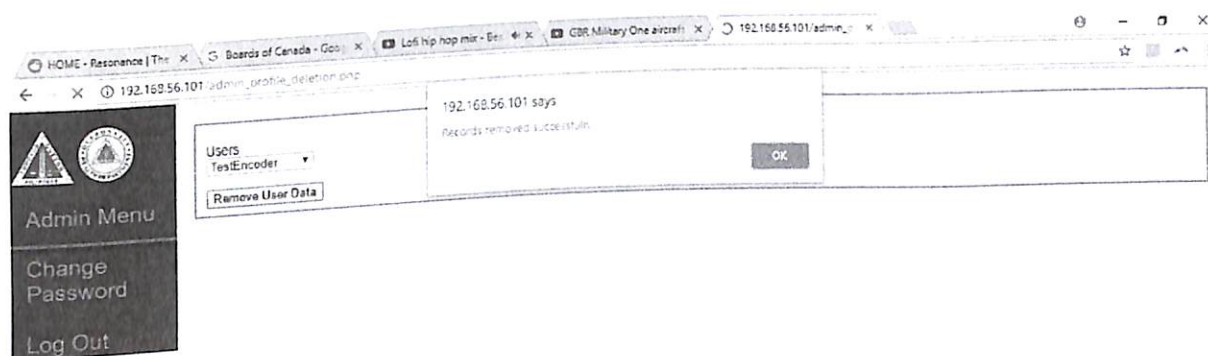
This screenshot shows the same 'Create Profile' form as the previous one, but with a success message pop-up overlaid. The pop-up is a white box with a grey border, containing the text '192.168.56.101 says' and 'Congratulations. User created successfully.' with an 'OK' button in the bottom right corner. The form fields and sidebar are still visible in the background.

Profile Deletion

If a profile no longer needs to be used (for example, if the user resigned or was reassigned to another department), it could be deleted by the administrator.



All the administrator needs to do is select the user from the dropdown, and click "Remove User Data". A pop-up will show saying "Records removed successfully."



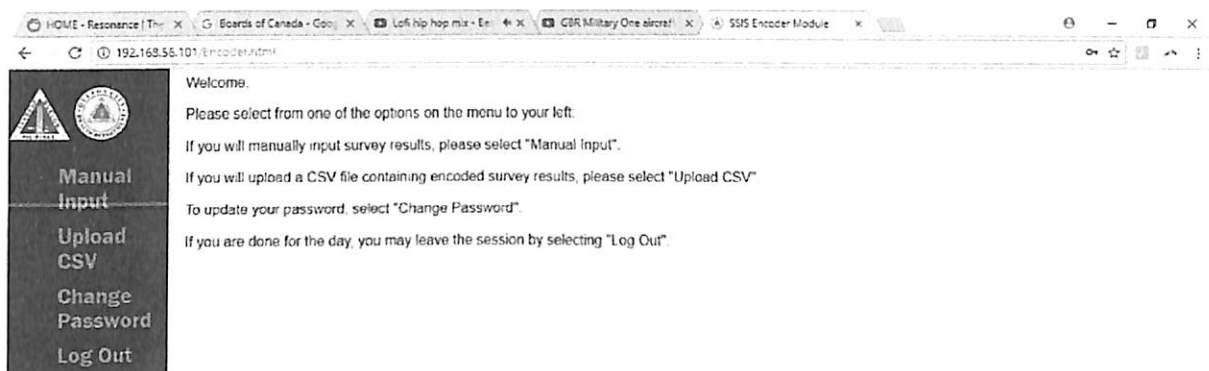
Log Out

Once done with the session, user can log out. He will be given an opportunity to log in again as another user by clicking "Log In".



Data Encoder

Upon login, the data encoder will be able to see the following view:



In the Data Encoder view, you will find four menu options: Manual Input, Upload CSV, Change Password, and Log Out.

Manual Input

For manual input, you will be asked to select a survey type (In-Patient or Out-Patient). The questions are largely similar, but there is a question specific to Out-Patient surveys which is not implemented in the In-Patient survey.



Once a survey is selected, the responses can be recorded. Enter the patient ID and the responses recorded for that patient.

The screenshot shows a web browser window with the URL 192.168.56.101/encoder_manual/about/Patient.html. The page has a sidebar on the left with the following menu items: Encoder Home, Upload CSV, and Log Out. The main content area is titled "Patient ID: 15" and contains a list of services with radio button options for rating: Best, Very Good, Good, Fair, and Poor.

Services and their corresponding radio button options:

- Security Guards: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Cashier: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Hospital Pharmacy (Personality): ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Medical Records: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Social Worker: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- ECG Service: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Hospital Pharmacy (Equipment Use): ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Laboratory: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Ultrasound: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- X-Ray: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Ward: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Nurse: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor
- Waiting Area: ☐ Best ☐ Very Good ☐ Good ☐ Fair ☐ Poor

Below the list of services, there is a section for "Ang Doctor/Nurse ay napalitan ang Paggamot/payong pangkalusugan sa paraan na aking naintindihan" with radio button options: ☐ Napakalusay ☐ Kasiya-siya ☐ Tama lang ☐ Di Mabuti ☐ Napakapangit.

Once done, click "Submit Survey".

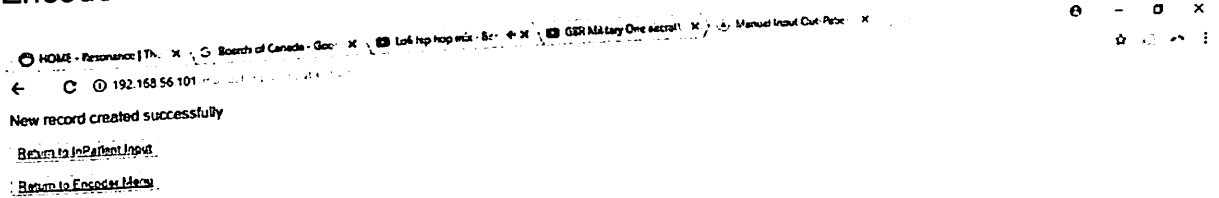
The screenshot shows the same web browser window as the previous one, but with additional questions at the bottom of the form. The "Submit Survey" button is visible at the bottom left of the form area.

Additional questions and options:

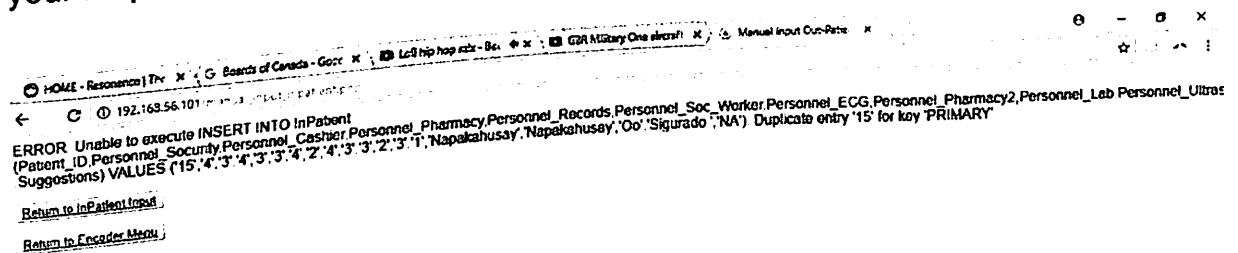
- Ang Doctor/Nurse ay napalitan ang Paggamot/payong pangkalusugan sa paraan na aking naintindihan: ☐ Napakalusay ☐ Kasiya-siya ☐ Tama lang ☐ Di Mabuti ☐ Napakapangit
- Ang impormasyon na aking natanggap ay natulungan akong maintindihan ang aking kondisyon o ang kondisyon ng aking pamilya: ☐ Napakalusay ☐ Kasiya-siya ☐ Tama lang ☐ Di Mabuti ☐ Napakapangit
- Aking irekomenda sa aking mga kamag-anak at kaibigan ang serbisyo ng ospital na ito: ☐ Oo ☐ Hindi
- Ano ang posibilidad na sasadya kayong muli sa aming ospital sa mga sumusunod na panahon upang magpatingin/gumamit ng serbisyo? ☐ Sigurado ☐ Malamang ☐ May posibilidad ☐ Baka hindi na ☐ Hinding hindi ☐ Hindi alam
- Mga Mungkahi (suggestions/recommendations):

The "Submit Survey" button is located at the bottom left of the form area.

A page will show, displaying the text "New Record Created Successfully" and giving you two buttons to either return to InPatient Input or return to the Encoder menu.



Other pages may show if there were errors. Make sure that your Patient ID number is not reused; otherwise, an error similar to below will show up and your response will not be recorded:



Data Entry using CSV

For the encoder who already has an Excel/CSV file, he could simply upload this CSV file with responses.

The dictionary in the next page must be strictly followed for the responses to be properly uploaded to the system:

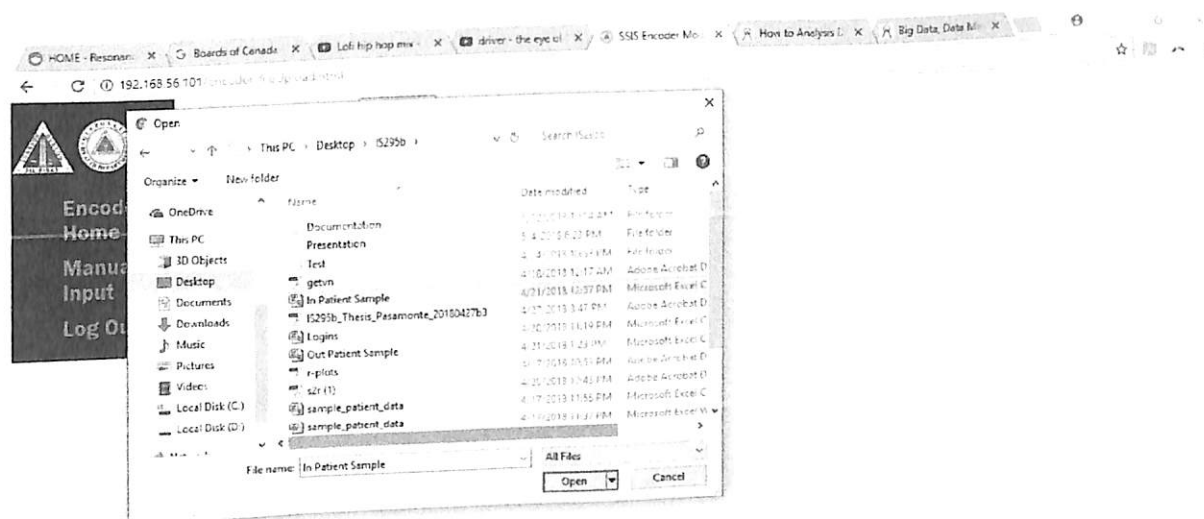
Header	Value
Patient_ID	Unique Patient ID - there CANNOT be any duplicates here, or an error will occur
Personnel_Security	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Cashier	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Pharmacy	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Records	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Soc_Worker	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_ECG	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Pharmacy2	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Lab	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Ultrasound	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Xray	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Ward	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Nurse	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Personnel_Waiting_Area	Rating scale from 1-5 (5 is "Best", while 1 is "Poor")
Ease_of_Understanding_Treatment	Actual text: Napakahusay Kasiya-siya Tama lang Di Mabuti Napakapangit
Ease_of_Understanding_Condition	Actual text: Napakahusay Kasiya-siya Tama lang Di Mabuti Napakapangit
Recommendation	Yes/No (Oo/Hindi)

Header	Value
Will_Return	Actual text: Sigurado Malamang May posibilidad Baka hindi na Hinding-hindi Hindi alam
Suggestions	Free text

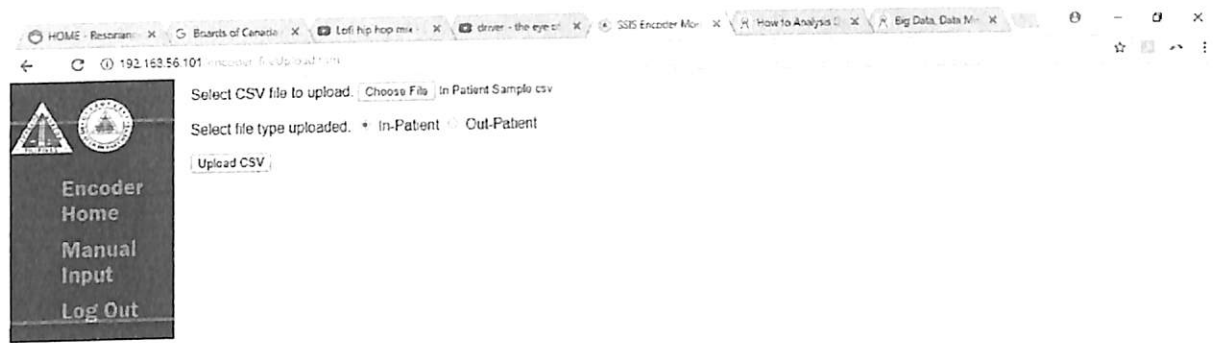
Once the encoder has a CSV file following the dictionary, he could upload it via this page:



Click "Choose File", and proceed to the location of the encoded file:

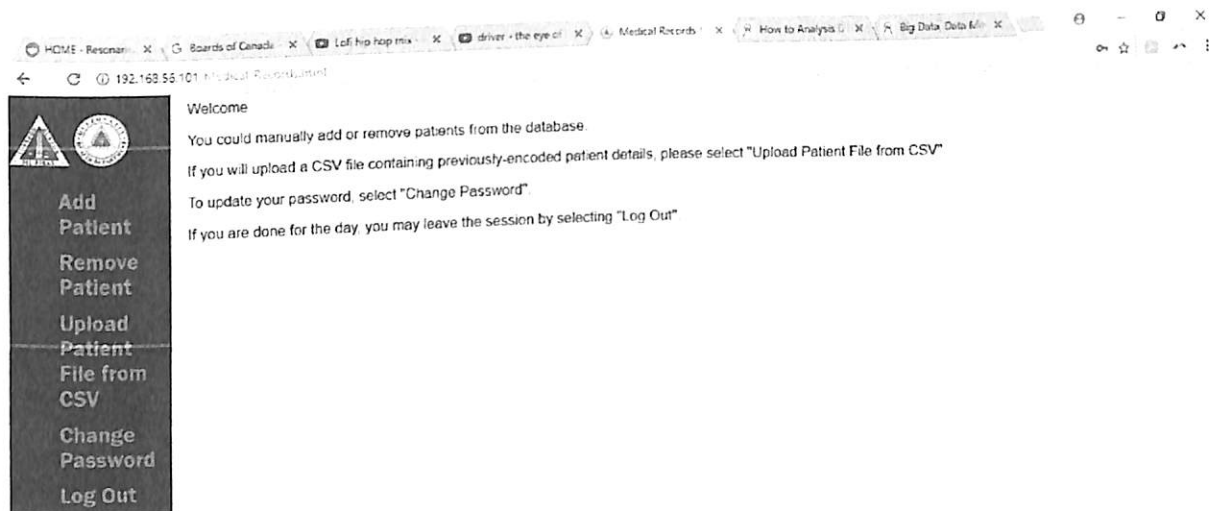


Then click "Upload CSV".



For the Medical Records Staff

Medical Records staff would have this initial landing page:

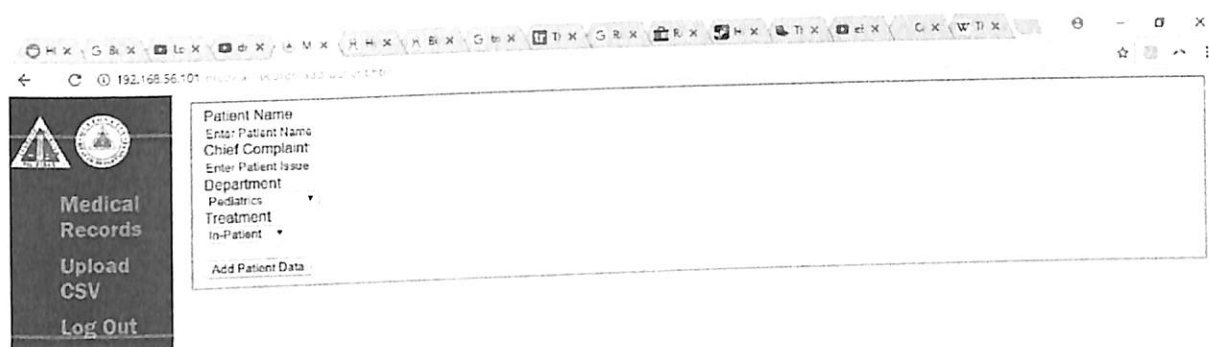


Medical Records staff could add patient records, remove patient records (in case of erroneous addition to patient records), upload a patient file from CSV, change passwords, and log out.

Manual Data Entry

Medical Records staff could add and remove patients manually.

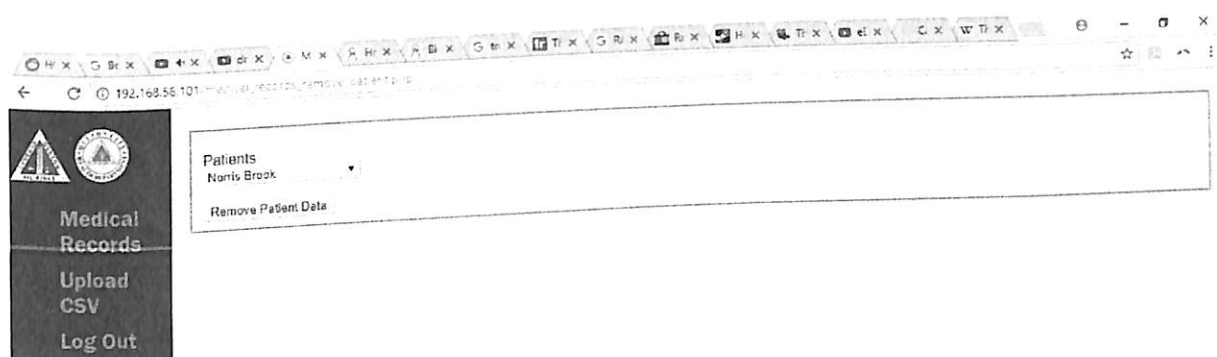
Add Patient



The screenshot shows a web browser window with the URL 192.168.56.101. The page has a dark sidebar on the left with the following menu items: Medical Records, Upload CSV, and Log Out. The main content area contains a form with the following fields: Patient Name, Enter Patient Name, Chief Complaint, Enter Patient Issue, Department (a dropdown menu with 'Pediatrics' selected), Treatment (a dropdown menu with 'In-Patient' selected), and an Add Patient Data button.

Adding patients involves filling up the form in the "Add Patients" page. This involves filling out the patient's name, chief complaint, department, and whether the patient's treatment is in-patient or out-patient.

Remove Patient

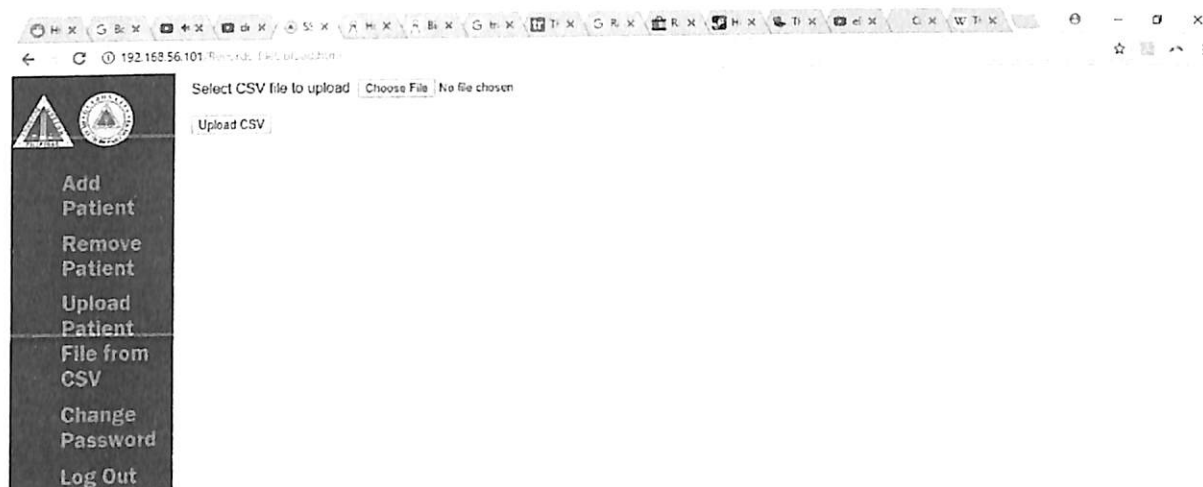


The screenshot shows a web browser window with the URL 192.168.56.101. The page has a dark sidebar on the left with the following menu items: Medical Records, Upload CSV, and Log Out. The main content area contains a form with the following fields: Patients (a dropdown menu with 'Naris Brook' selected), and a Remove Patient Data button.

Removing patients involves selecting the patient from the dropdown, and removing him manually by clicking "Remove Patient Data".

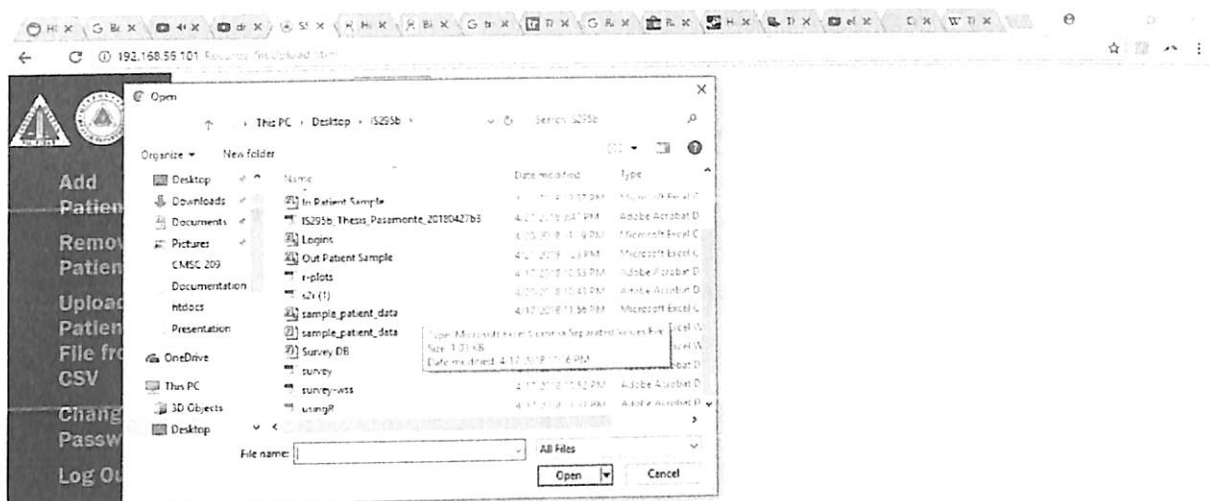
Upload CSV

Patient records could be encoded in a CSV file and uploaded to the database, too. This could be done via the CSV upload module:



The medical staff selects a CSV file. This file must follow the dictionary below:

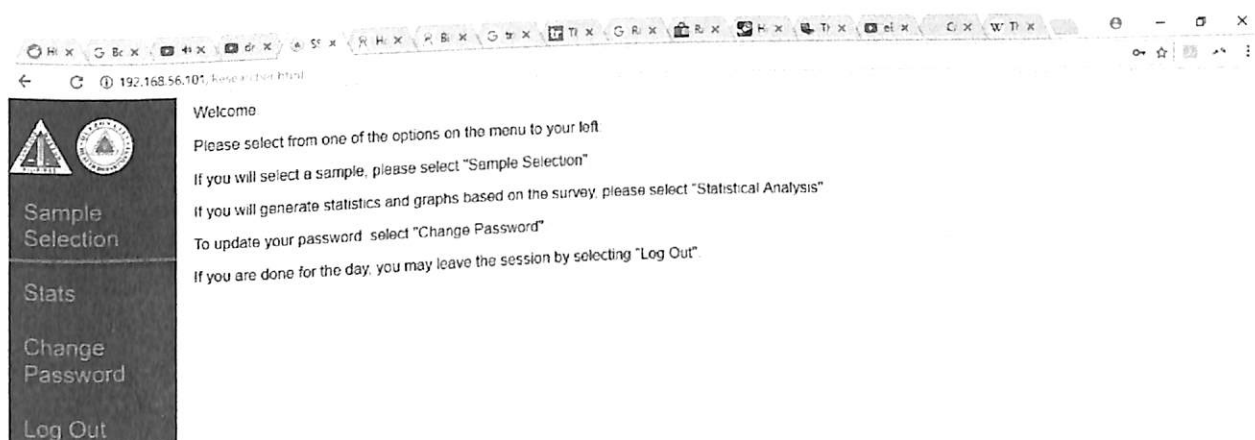
Header	Value
patientID	Unique patient ID - this will also be used in the survey by the encoders. There CANNOT be any duplicates here, or an error will occur
patientName	Patient Name
Chief_Complaint	Issue which the patient came in for a consultation
Department	Hospital Department where patient will consult
In_Out_Patient_Tag	Treatment Type, whether In Patient or Out Patient



Once a CSV is selected, click "Upload CSV". A page should show, saying the file has been successfully uploaded, or if there are any errors.

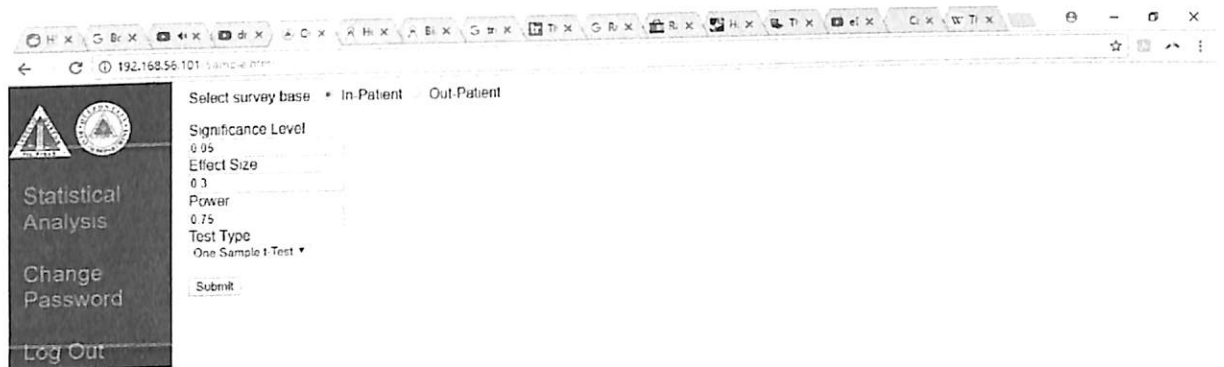
For the Researcher

The researcher view has the following menu options: Sample Selection, Stats, Change Password, and Log Out.



Sample Selection

The researcher can select who among the listed patients could be sampled for the patient survey.



The screenshot shows a web browser window with a URL bar displaying '192.168.56.101:5000/sample.htm'. The page features a sidebar on the left with a logo and three menu items: 'Statistical Analysis', 'Change Password', and 'Log Out'. The main content area contains a form with the following fields and options:

- Select survey base:** Radio buttons for 'In-Patient' (selected) and 'Out-Patient'.
- Significance Level:** A text input field containing '0.05'.
- Effect Size:** A text input field containing '0.3'.
- Power:** A text input field containing '0.75'.
- Test Type:** A dropdown menu showing 'One Sample t-Test'.
- Submit:** A button at the bottom of the form.

The important item here is to select which sample should be taken – whether In-Patient or Out-Patient.

The other items in the form are as follows.

Significance Level – Usually this is set to 0.05, as is commonly-accepted statistical practice. More restrictive statistics set this to 0.01.

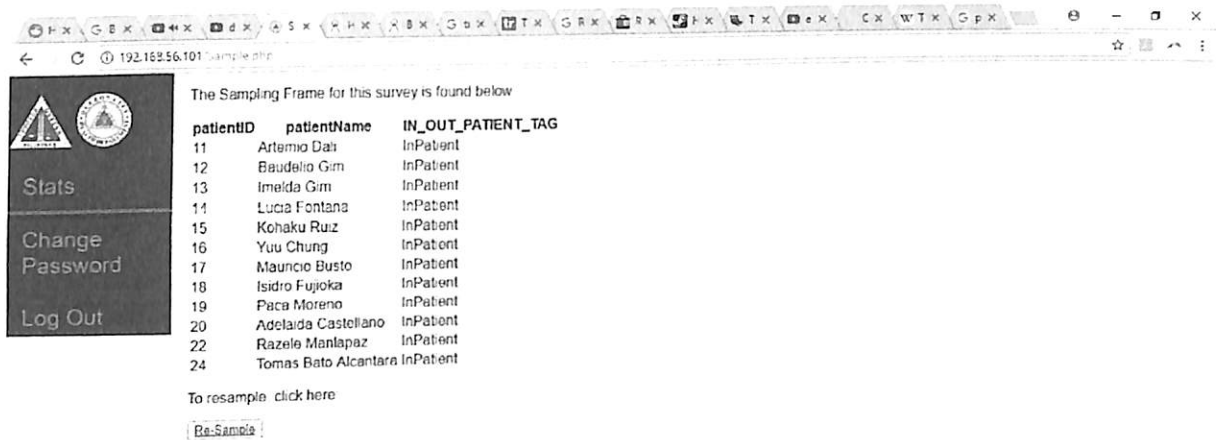
Effect Size – if one wants to test whether there are significant differences between groups, you could increase this. Otherwise, keep as is.

Power – This is the probability of rejecting a false null hypothesis. We would want to keep this high (around 0.70 to 0.75), but not too high, since this would lead to an excessively high sample size requirement.

Test Type – If the researcher has some statistical test in mind, such as a one-sample t-test or a paired t-test, he could select those here.

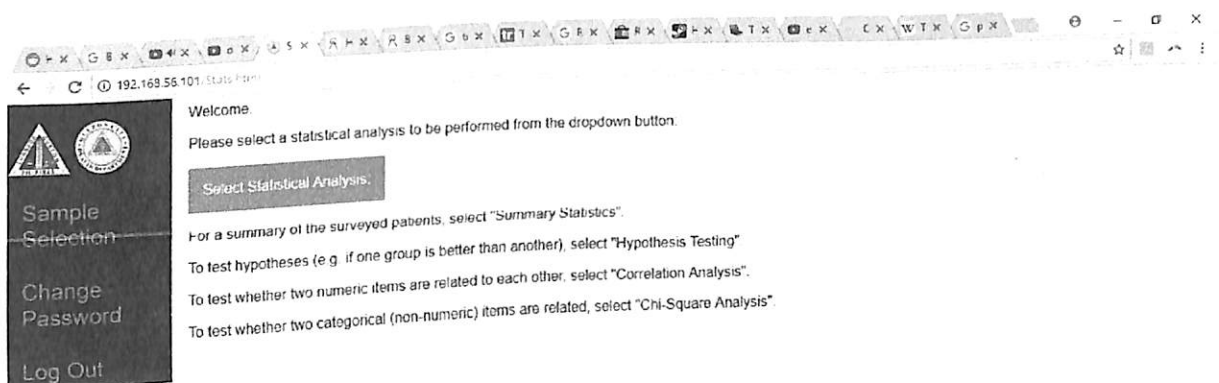
If one is not sure, he could leave these as is, and a workable sample could still be obtained. Once done with these values, the researcher clicks "Submit".

A list of sampled patients is output. The survey would then be implemented on these patients, and they would be encoded by the data encoders in a separate step.

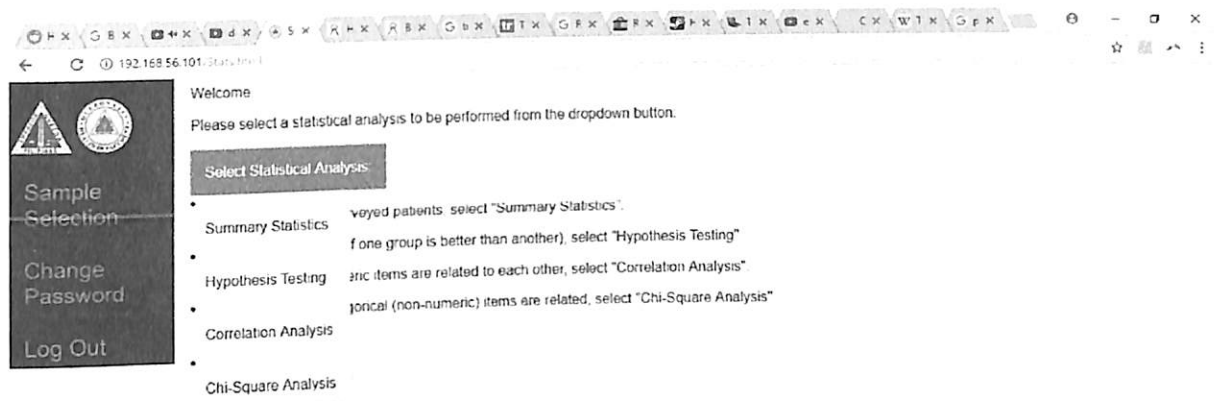


Statistical Analysis

Another feature of this system is the ability to perform some statistical analyses on the sampled data, once data encoders have encoded this data. The researcher clicks the "Stats" link. It will take the researcher to this page:



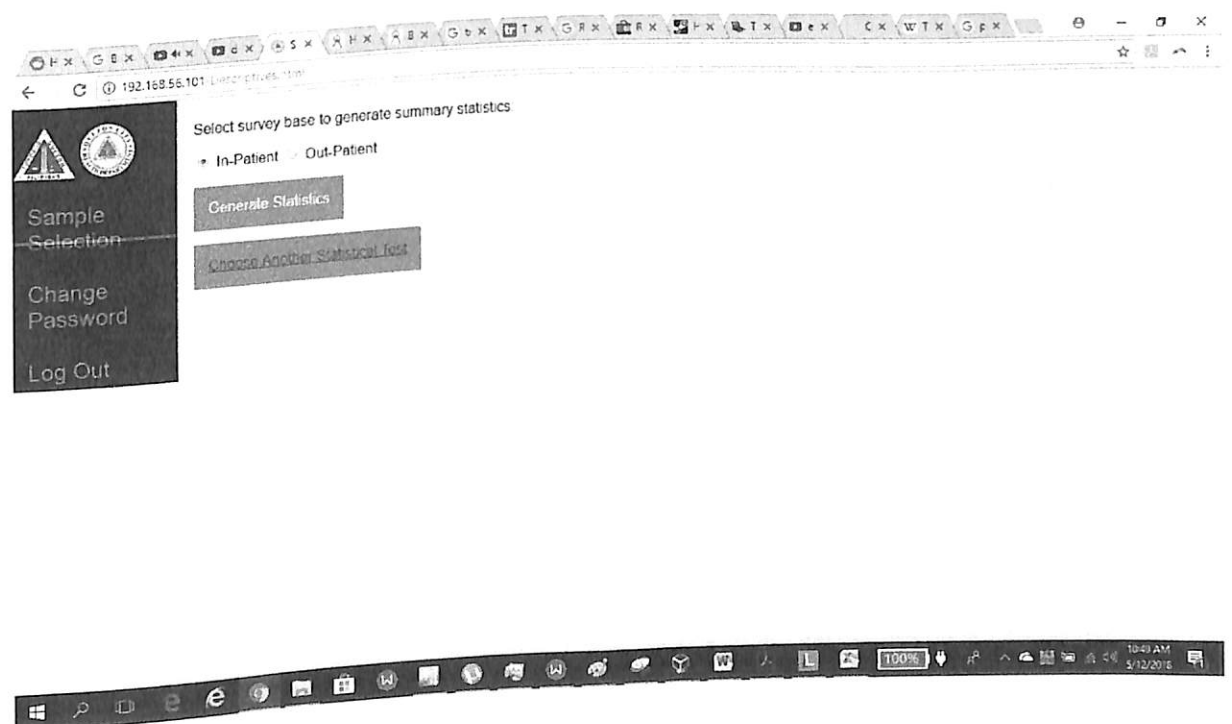
Once the researcher clicks the blue "Select Statistical Analysis" button, a drop-down menu will show, listing the various statistical analyses he could perform.



Currently, the researcher could obtain summary statistics and perform three tests: Hypothesis Testing (between numerical scores, such as any personnel scores), Correlation Analysis (also between numerical scores), and Chi-Square Analysis.

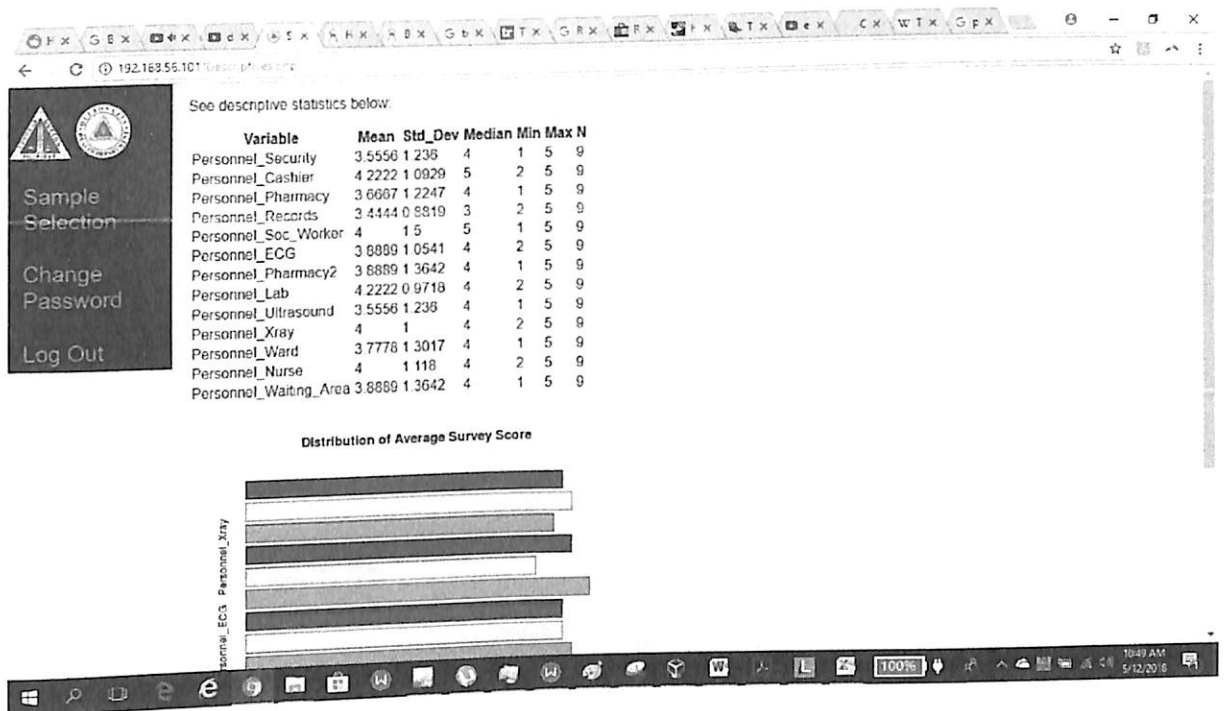
Summary Statistics

To obtain summary statistics, click on "Summary Statistics".

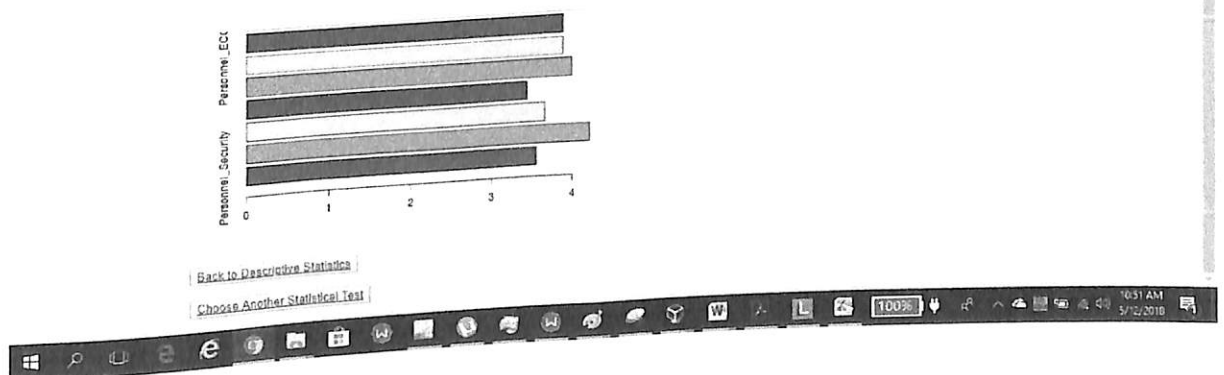


You will need to select survey base to get the statistics from: whether In-Patient or Out-Patient. Once you select this, click "Generate Statistics"

This set of descriptive statistics will be output, consisting of the variable name, mean, standard deviation, median, minimum, maximum, and number of observations:

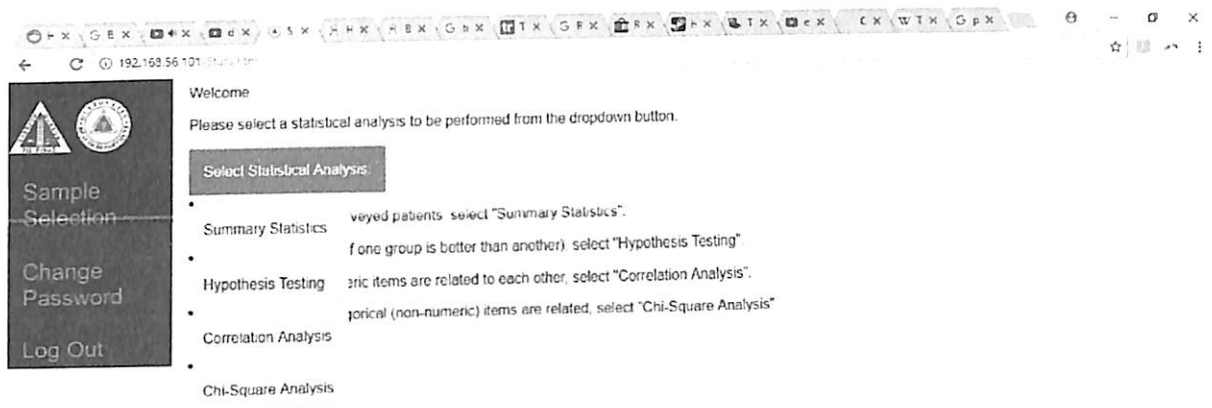


Once done with the descriptive statistics, you could either get the descriptive statistics for the other set, or select another statistical test to be done on the data.

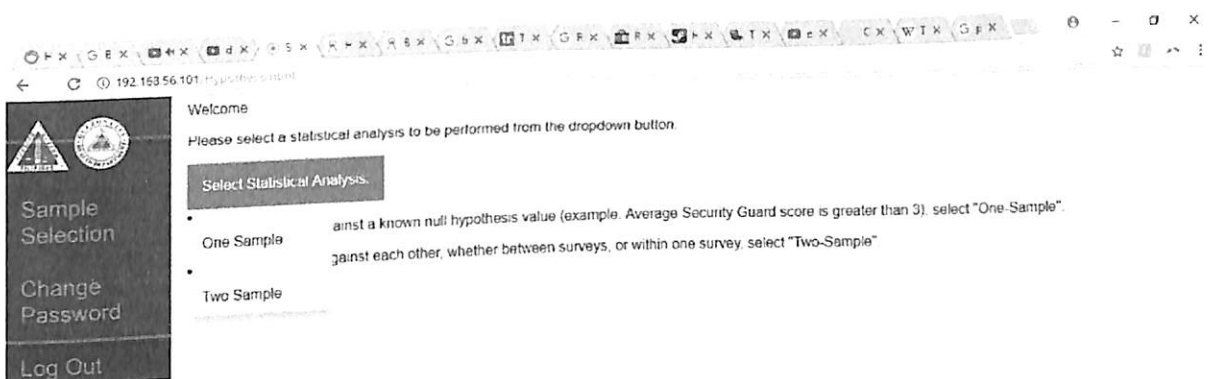


Hypothesis Testing

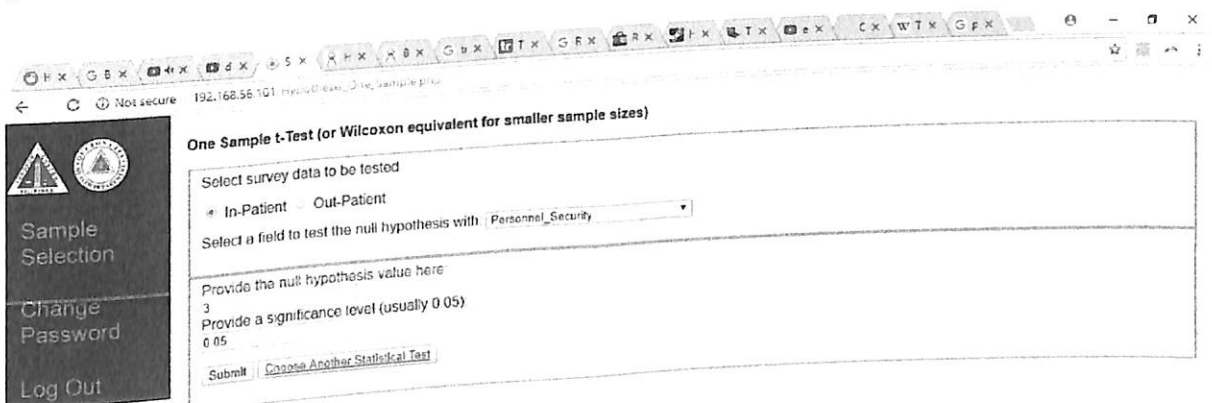
Select "Hypothesis Testing" from the dropdown menu:



You will be given two options, depending on what test you would want to do.



One Sample tests are comparisons against a known null hypothesis (for example, if the average scores for the effectivity of the security guard are greater than 3 – these would at least not be in the "bad" buckets)



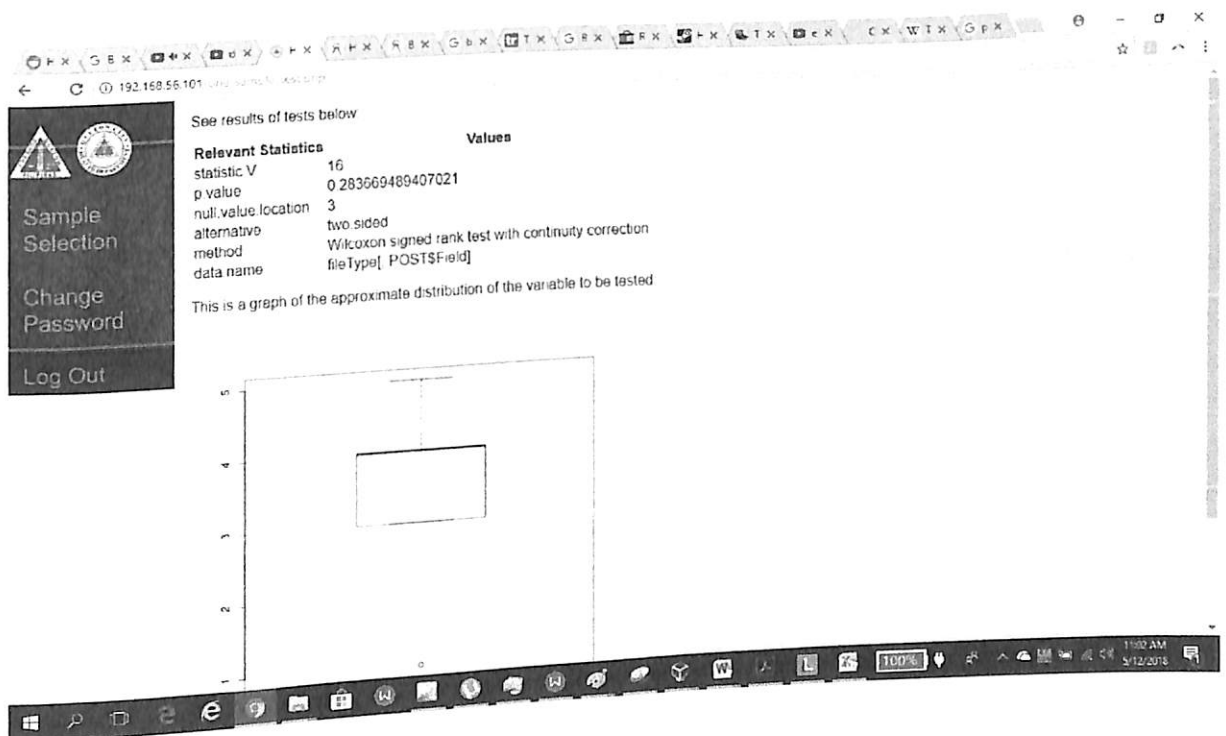
It should be noted here that for smaller sample sizes, the nonparametric analogue to the t-test, the Wilcoxon Signed Rank Test, is used.

One selects the survey data to be tested (in-patient vs out-patient), the field for testing of the null hypothesis, then fills out the following details:

Null Hypothesis Value (Are values greater than this? Less than this?)

Significance Level – usually 0.05 as per commonly accepted statistical practice, but can be lowered to 0.01 for strict statistics, and raised to 0.10 for more relaxed statistics.

Once these are input, click "Submit". The output would look similar to below:



Note the relevant statistics:

Statistic.V refers to the output of the t-test (or the corresponding Wilcoxon test, as the case may be).

The p-value shows the probability of the null hypothesis being true, given the data. This simply means the probability that the hypothesis alternative is false. For example, we wish to test that the average score for security guards is greater than 3. The null hypothesis is that it is less than/equal to 3.

If the p-value is less than our level of significance, we can reject the null hypothesis, and conclude that our test is successful. Otherwise, we cannot reject the null hypothesis.

The null value location is what we set in the earlier step.

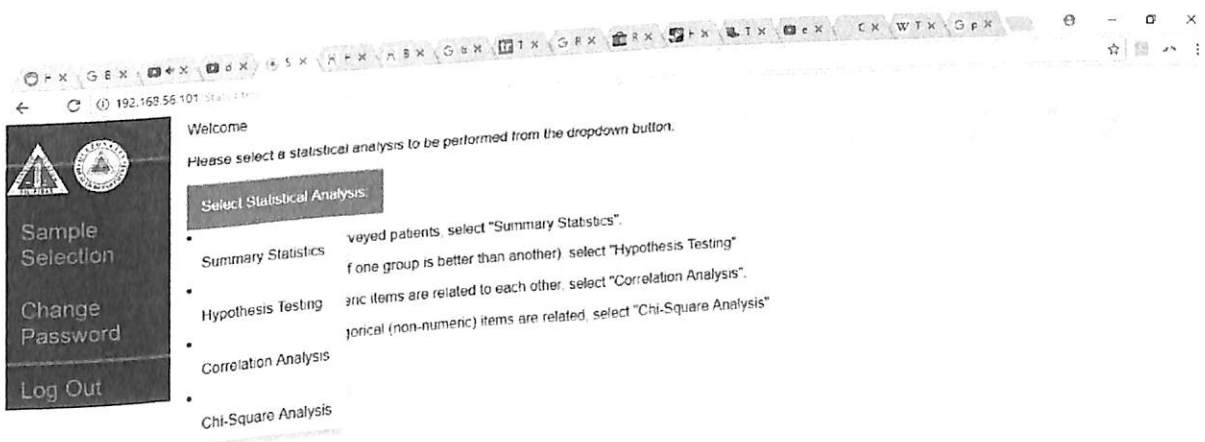
The method is either t-test or Wilcoxon Signed Rank Test, depending on the sample size.

Data name will be improved in a further revision. This should be the name of the variable being tested.

There is also a box plot of the test being performed, for the user to see the distribution of the variable.

Correlation Analysis

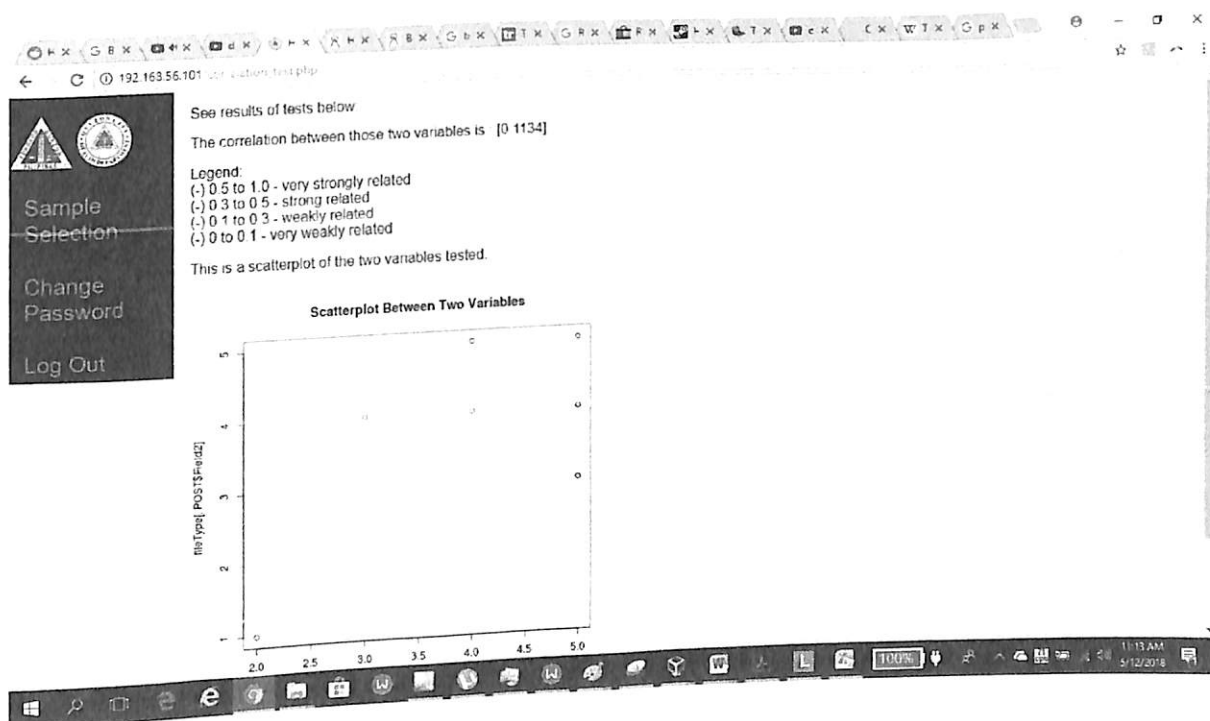
To do correlation analysis, select correlation analysis in the menu:



You will be greeted by this correlation analysis menu.

The screenshot shows a web browser window with the address bar displaying '192.168.56.101: Correlation.php'. The page title is 'Pearson's r-correlation (or Kendall's tau equivalent for smaller sample sizes)'. On the left is a dark sidebar with a logo and four menu items: 'Sample Selection', 'Change Password', and 'Log Out'. The main content area has a form with the following fields: 'Select survey data to be tested' with a dropdown menu showing 'In-Patient' and 'Out-Patient'; 'Select two fields to compare' with a note 'Both must be numeric (i.e. patient scores beginning with Personnel_)' and a dropdown menu showing 'Patient_ID'; and a 'Patient_ID' input field. At the bottom of the form are two buttons: 'Submit' and 'Choose Another Statistical Test'.

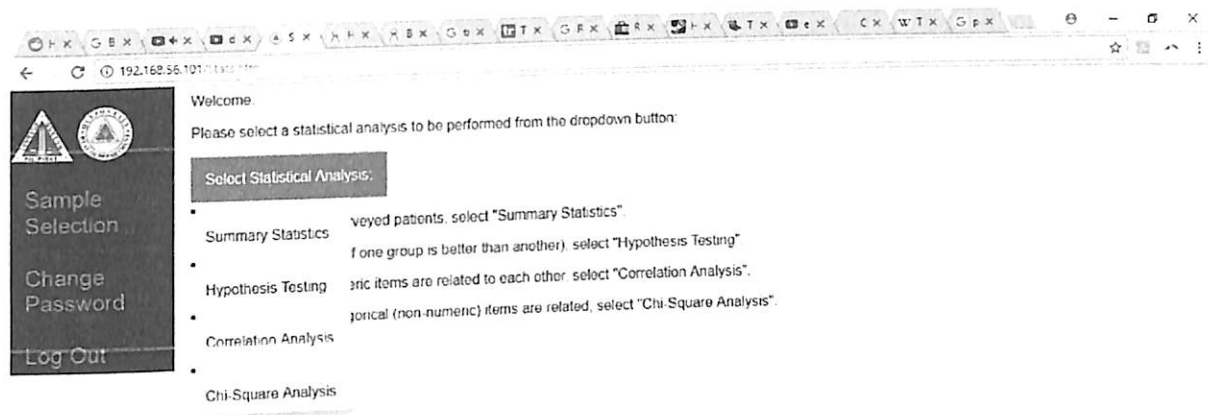
Again, we need to select survey data to be tested (whether in-patient or out-patient). We also need to select two numeric variables for which we are looking for their correlation. Then click "Submit". The following output will be shown:



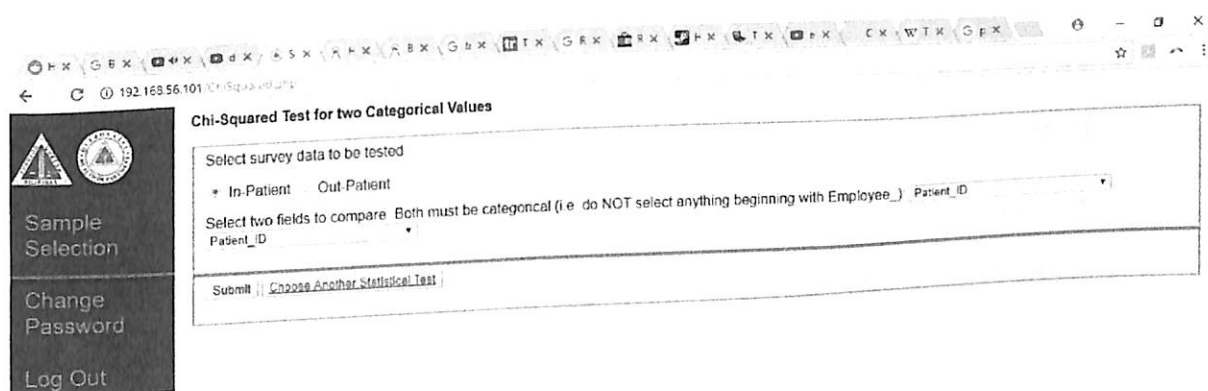
The correlation score will be displayed, as well as a legend showing the meaning of each correlation score. A scatterplot is also shown for visual appreciation of the relationship between the two variables.

Chi-Square Analysis

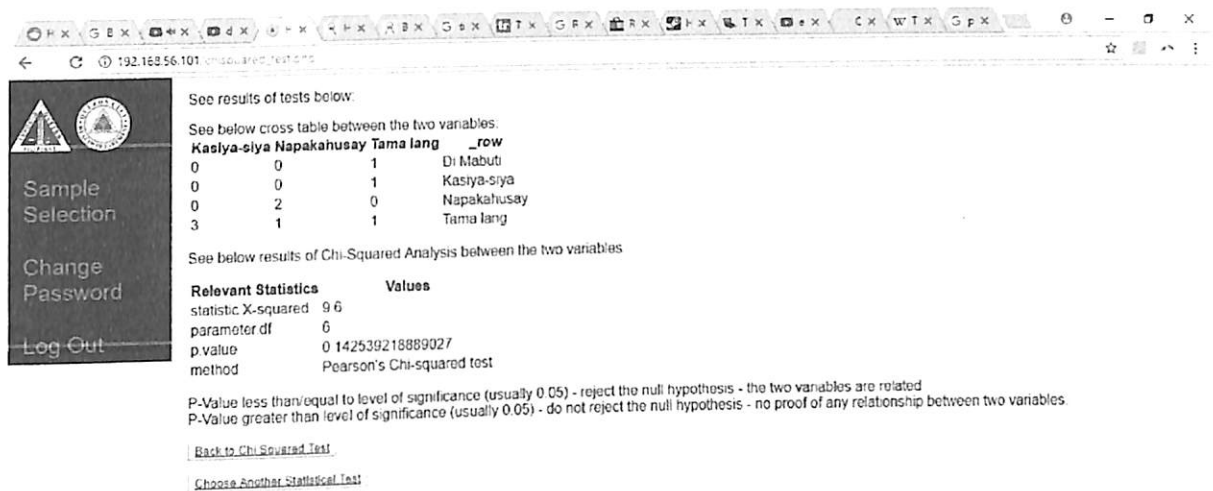
To do chi-square analysis, select Chi-Square Analysis in the menu. This is to determine the relationship of any categorical data:



You will be taken to this menu. Select the survey data, then select two categorical fields. For the surveys, do NOT select anything beginning with "Employee" as these are considered numeric by the system.



You will see this output:



See results of tests below:

See below cross table between the two variables:

Kasiya-siya	Napakahusay	Tama lang	_row
0	0	1	Di Mabuti
0	0	1	Kasiya-siya
0	2	0	Napakahusay
3	1	1	Tama lang

See below results of Chi-Squared Analysis between the two variables

Relevant Statistics	Values
statistic X-squared	9.6
parameter df	6
p value	0.142539218889027
method	Pearson's Chi-squared test

P-Value less than/equal to level of significance (usually 0.05) - reject the null hypothesis - the two variables are related
P-Value greater than level of significance (usually 0.05) - do not reject the null hypothesis - no proof of any relationship between two variables

[Back to Chi-Squared Test](#)

[Choose Another Statistical Test](#)

The way to interpret these statistics are:

Statistic.X-squared – the actual chi-squared statistic calculated in R.

Parameter.df – degrees of freedom in the chi-squared analysis – number of observations, less 1.

p.value – Since the chi-squared test evaluates the null hypothesis that there is no relationship between these two variables, there is a calculated p-value. If this is less than the level of significance (usually 0.05), we could say that there is some association between the two variables; otherwise, these two variables are said to be independent of each other.

Annex G. Patient Survey Forms

(Kindly answer the following)

Serbisyo ng ospital (Hospital Services)

OUT-PATIENT DEPARTMENT

Pakikopyan po lamang ng check ang mga kahon na nangkop sa inyong na obserbahan o naranasan. Mangyari po lamang na iwasan ang blangko. (Please put check in the boxes provided based on what you have observed or experienced. Please avoid leaving any blanks)

Personnel/Staff	ATAT (S)	VERIFIED (A)	GOOD (A)	Tali (2)	NOT GOOD (1)
Security Guards					
Cashier					
Hospital Pharmacy					
Medical Records					
Social Worker					
ECG service					
Hospital Pharmacy					
Laboratory					
Ultrasound					
X-Ray					
WARD					
Nurses/Nurse aides					
Waiting Area					

Gaano katagal ka naghintay para sa konsultasyon?

☐ 0-30 minuto ☐ 31-59 minuto ☐ 1-2oras ☐ 2 oras o higit pa

Ang Doctor/Nurse ay naipaliwanag ang Paggamot/payong pangkalusugan sa paraan na aking naiintindihan.

☐ Napakahusay ☐ Kasiya-siya ☐ Tama lang ☐ Di-Mabuti ☐ Napakapangit

Ang impormasyon na aking natanggap ay natulungan akong maintindahan ang aking kondisyon/kondisyon ng aking pamilya.

☐ Napakahusay ☐ Kasiya-siya ☐ Tama lang ☐ Di-Mabuti ☐ Napakapangit

Aking irerekomenda sa aking mga kamag-anak at kaibigan ang serbisyo ng ospital na ito.

☐ Oo ☐ Hindi

Ano ang posibilidad na sasadya kayong muli sa aming ospital sa mga sumusunod na panahon upang magpatingin/gumamit ng serbisyo?

☐ Sigurado ☐ Baka hindi na
☐ Malamang ☐ Hinding hindi
☐ May Posibilidad ☐ Hindi alam

Mungkahi (Suggestion/Recommendation)

Paksa ng pagtugon sa survey na ito _____

Pangalan ng pasyente/bantay _____

Ward _____

Oras _____

MARAMING SALAMAT PO!!!

Serbisyo ng ospital (Hospital Services)

IN-PATIENT DEPARTMENT

Pakiusap po lamang ng check ang mga kahon na nangkop sa inyong na obserbahan o naranasan.

Mangyari po lamang na iwasang mag-iwan ng blanko. (Please put check in the boxes provided based on what you have observed or experienced. Please avoid leaving any blank)

Personnel/Staff	BEST (5)	VERY GOOD (4)	GOOD (3)	FAIR (2)	NOT GOOD (1)
Security Guards					
Cashier					
Hospital Pharmacy					
Medical Records					
Social Worker					
ECG service					
Hospital Pharmacy					
Laboratory					
Ultrasound					
X-Ray					
WARD					
Nurses/Midwives					
Waiting Area					

Ang Doctor/Nurse ay naipaliwanag ang Paggamot/payong pangkalusugan sa paraan na aking naiintindihan.

☐ Napakahusay ☐ Kasiya-siya ☐ Tama lang ☐ Di-Mabuti ☐ Napakapangit

Ang impormasyon na aking natanggap ay natulungan akong maintindahan ang aking kondisyon/kondisyon ng aking pamilya.

☐ Napakahusay ☐ Kasiya-siya ☐ Tama lang ☐ Di-Mabuti ☐ Napakapangit

Aking irerekomenda sa aking mga kamag-anak at kaibigan ang serbisyo ng ospital na ito.

☐ Oo ☐ Hindi

Ano ang posibilidad na sasadya kayong muli sa aming ospital sa mga sumusunod na panahon upang magpatingin/gumamit ng serbisyo?

☐ Sigurado ☐ Baka hindi na
☐ Malamang ☐ Hinding hindi
☐ May Posibilidad ☐ Hindi alam

Mungkahi (Suggestions/Recommendation).

Petsa ng pagtugon sa survey na ito _____

Pangalan ng pasyente/bantay: _____

Ward: _____

Oras: _____

MARAMING SALAMAT PO!!!