

**SPATIAL DATA VISUALIZATION FOR SMOKING  
CESSATION THERAPY PROGRAM**




**A THEMATIC PAPER SUBMITTED IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF MASTER OF SCIENCE  
(INFORMATION TECHNOLOGY MANAGEMENT)  
FACULTY OF GRADUATE STUDIES  
MAHIDOL UNIVERSITY**

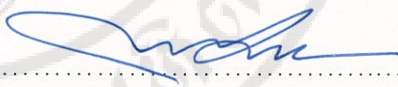
Copyright by Mahidol University 2018


**COPYRIGHT OF MAHIDOL UNIVERSITY**


Thesis  
entitled  
**SPATIAL DATA VISUALIZATION FOR SMOKING  
CESSATION THERAPY PROGRAM**

  
.....  
Mr. Thanachot Techawattanasirikul  
Candidate

  
.....  
Asst. Prof. Sotarath Thammaboosadee,  
Ph.D. (Information Technology)  
Major advisor

  
.....  
Asst. Prof. Taweesak Samanchuen,  
Ph.D. (Electrical Engineering)  
Co-advisor

  
.....  
Assoc. Prof. Varaporn Akkarapatumwong,  
Ph.D. (Science)  
Acting Dean  
Faculty of Graduate Studies  
Mahidol University

  
.....  
Asst. Prof. Supaporn Kiattisin,  
Ph.D. (Electrical and Computer  
Engineering)  
Program Director  
Master of Science Program in  
Information Technology Management  
Faculty of Engineering  
Mahidol University


Thesis  
entitled  
**SPATIAL DATA VISUALIZATION FOR SMOKING  
CESSATION THERAPY PROGRAM**

was submitted to the Faculty of Graduate Studies, Mahidol University  
for the degree of Master of Science (Information Technology Management)

on  
December 16, 2018

ธนาโชต เทชวาทันศิริกุล

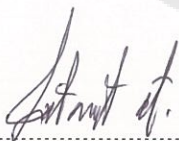
Mr. Thanachoat Techawattanasirikul  
Candidate



Prof. Somkiat Wattanasirichaigon,  
M.D., FRTSC (General Surgery)  
Chair



Asst. Prof. Taweesak Samanchuen,  
Ph.D. (Electrical Engineering)  
Member



Asst. Prof. Sotarath Thammaboosadee,  
Ph.D. (Information Technology)  
Member

Orasa Patsadu

Orasa Patsadu  
Ph.D. (Computer Science)  
Member



Assoc. Prof. Varaporn Akkarapatumwong,  
Ph.D. (Science)  
Acting Dean  
Faculty of Graduate Studies  
Mahidol University



Asst. Prof. Jackrit Suthakorn,  
Ph.D. (Robotics)  
Dean  
Faculty of Engineering  
Mahidol University

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Asst. Prof. Sotarath Thammaboosadee, Ph.D. for the support throughout the course of this project. From the beginning of the project when I figure out the research topic and throughout research and writing of this thematic paper. I would like to express the appreciation to Thai Physician Alliance Against Tobacco (TPAAT) who provided the data in this research. I would like to express the appreciation Prof. Dr. Somkiat Wattanasirichaigoon, Dr.Taweesak Samanchuen and Dr.Orasa Patsadu to give the suggestion to greatly improve my thesis.

Finally, I would like to thank IT Management division, Faculty of engineering, Mahidol University for their help and support. The last one, I also would like to thank my friends and family for the encouragement, and endless support.

Thanachot Techawattanasirikul

**SPATIAL DATA VISUALIZATION FOR SMOKING CESSATION THERAPY PROGRAM**

**THANACHOAT TECHAWATTANASIRIKUL 5937839 EGIT/M**

**M.Sc. (INFORMATION TECHNOLOGY MANAGEMENT)**

**THEMATIC PAPER ADVISORY COMMITTEE: SOTARAT THAMMABOOSADEE, Ph.D., TAWEESAK SAMANCHUEN Ph.D.**

**ABSTRACT**

This research aims to develop a BI system for data display formats, for data views, easy data analysis and easy-to-understand overview. The purpose of the project is to help smoking treatment. For this project, the research used data of Fah Sai Clinic Project from the year of 2015 to 2017, in which total number of 34,851 patients participated. The data contained information about patients, hospital and follow-up treatment separated by a period of 3 months and 6 months. The attribute data were selected from the dataset; data cleansing was carried out and then the coordinate location of each hospital was collected for plotting location on map. After that, the data were entered for the development of BI. Dashboards were designed, which consisted of 8 dimensions, and all 8 dimensions were considered to create a storyboard which could help the project manager see and easily understand information, assist in the planning and help to decide project planning fast and effectively. The assessment was carried out to evaluate the project, and the result showed the overall rating ranged from 3.75 to 4.75 out of total assessment score 5. The results of this research can be used to build several aspects; for example, if information on drug prices is available in the future, another dimension “Budget” can be added to the developed dashboard.

**KEY WORDS: SMOKING CESSATION / DATA ANALYZE /  
DATA VISUALIZATION / BI TOOL / CRIPS-DM**

36 pages

การทำทัศนภาพข้อมูลเชิงภูมิศาสตร์สำหรับ โปรแกรมการรักษาโรคติดยาสูบ

SPATIAL DATA VISUALIZATION FOR SMOKING CESSATION THERAPY PROGRAM

ธน โชนิต์ เศษวัฒน์สิริกุล 5937839 EGIT/M

วท.ม. (การจัดการเทคโนโลยีสารสนเทศ)

คณะกรรมการที่ปรึกษาสารนิพนธ์: โยทศร์รัต ธรรมบุญดี, Ph.D., ทวีศักดิ์ สมานชื่น, Ph.D.

#### บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบ BI สำหรับ การนำข้อมูลมาแสดงผลในรูปแบบทัศนภาพข้อมูล เพื่อให้สามารถเห็นวิเคราะห์ข้อมูลได้ง่าย และเข้าใจภาพรวมข้อมูลได้ง่าย จุดประสงค์ของโครงการคือ ช่วยบำบัดผู้ติดยาสูบ ขอบเขตของงานวิจัยได้ข้อมูลจากทางโครงการ โดยข้อมูลที่ได้นำมาเป็นข้อมูลระหว่างปี 2558 ถึง 2560 รวมจำนวนผู้ป่วยที่เข้ารับการรักษาทั้งหมด 34,851 คน โดยข้อมูลประกอบด้วย ข้อมูลผู้ป่วยที่เข้ารับการรักษ ข้อมูลโรงพยาบาลที่เข้าร่วม และ ข้อมูลติดตามผลการรักษาโดยแยกออกเป็นช่วง 3 เดือน และ 6 เดือน หลักจากนั้น ได้นำข้อมูลมาคัดเลือกแอททริบิวต์ที่สำคัญ การทำความสะอาดข้อมูล และการหาข้อมูลพิกัดโรงพยาบาลเพิ่ม ได้นำข้อมูลมาเข้าสู่กระบวนการ การพัฒนา BI มี 8 มิติ โดยแต่ละมิติมีความสัมพันธ์กันของข้อมูล และแสดงข้อมูลเฉพาะในแต่ละมิติ อีกทั้งยังสามารถแสดงข้อมูลเฉพาะบางส่วนที่ต้องการดูเป็นพิเศษ หลังจากได้นำทั้ง 8 มิติมาทำเป็น Storyboard จะทำให้ผู้บริหาร โครงการสามารถเห็น และ เข้าใจข้อมูลได้ง่าย และช่วยในการตัดสินใจวางแผนบริหาร โครงการต่อได้อย่างรวดเร็ว และมี ประสิทธิภาพ โดยได้ทำแบบประเมินให้ผู้ควบคุมโครงการทำการประเมิน ผลประเมิน โดยรวมอยู่ในช่วง 3.75 ถึง 4.75 จากผลประเมินเต็ม 5 ซึ่งได้ ในอนาคตงานวิจัยนี้สามารถต่อยอดได้หลายด้าน เช่น ถ้าหากมีข้อมูลรายกายก็จะสามารถเพิ่มมิติในด้านงบประมาณที่ใช้ได้

36 หน้า

## CONTENTS

	<b>Page</b>
<b>ACKNOWLEDGEMENTS</b>	<b>iii</b>
<b>ABSTRACT (ENGLISH)</b>	<b>iv</b>
<b>ABSTRACT (THAI)</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>viii</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>CHAPTER I INTRODUCTION</b>	<b>1</b>
1.1 Background and Problem Statement	1
1.2 Objective	2
1.3 Scope of Work	2
1.4 Expected Result	2
<b>CHAPTER II LITERATURE REVIEW</b>	<b>3</b>
2.1 General information about tobacco and treatment of tobacco dependence	3
2.2 Data analysis	6
2.3 Business Intelligence	7
2.4 Business Intelligence Architecture	9
2.5 Data Visualization	10
2.6 Relevant Research	11
<b>CHAPTER III RESEARCH METHODOLOGY</b>	<b>13</b>
3.1 Study the inventory models	13
3.2 Select the suitable method	13
3.3 Data Collection	14
3.4 Data Understanding	14
3.5 Data Preparation	20
3.6 BI Development	21
3.7 Dashboard Design	21

**CONTENTS (cont.)**

	<b>Page</b>
3.8 Evaluation	21
3.9 Research Timeline	22
<b>CHAPTER IV RESULTS AND DISCUSSION</b>	<b>23</b>
4.1 Results of problem analyze from collection data	23
4.2 Analyze and design dashboard	23
4.3 Evaluation	29
<b>CHAPTER V CONCLUSION</b>	<b>31</b>
5.1 Conclusions	31
5.2 Limitations and future works	31
<b>REFERENCES</b>	<b>33</b>
<b>BIOGRAPHY</b>	<b>34</b>

**LIST OF TABLES**

<b>Table</b>	<b>Page</b>
3.1 Data structure from smoking cessation clinic.	15
3.2 Data structure from smoking cessation clinic, Follow up 3 months	16
3.3 Data structure from smoking cessation clinic, Follow up 6 months	18
3.4 Excluded Attributes	20
3.5 Research Timeline	22
4.1 Evaluation Result	30

## LIST OF FIGURES

<b>Figure</b>	<b>Page</b>
2.1 Business Intelligence Workflow	8
2.2 Business Intelligence Architecture	9
2.3 Business Intelligence	10
3.1 Methodology Workflow	14
3.2 Example of data in smoking cessation clinic before data preparation	19
3.3 Example of follow up 3 months	19
3.4 The picture is use the web for find out latitude and longitude	21
4.1 Flow display sequence storyboard	24
4.2 Introduction of Storyboard	24
4.3 Map show total of patient by region	25
4.4 Map show total of hospital by region	25
4.5 Pie chart display relationship of patient can quit smoking in each region	26
4.6 Relationship between pie chart NCDs information and SEX information	27
4.7 Bar chart compare between patients has NCDs and not have.	28
4.8 Bar graph display top 10 highest hospital that have the most people quit smoking	28
4.9 Bar graph display top 10 highest hospital that have the most people quit smoking	29

## **CHAPTER I**

### **INTRODUCTION**

#### **1.1 Background and Problem Statement**

Smoking in Thailand has tended to rise. Thailand's National Statistical Office (NSO) reported the results from the survey on smoking and drinking behavior of population in 2014 that a total of 11.4 million smokers were recorded, including 10 million regular smokers. The number increased from 19.9 percent in 2011 to 20.7 percent (National Statistical Office, 2014). The 2009 study of the diseases burden from risk behaviors in Thailand revealed that the mortality from smoking was with 11.6 percent of total deaths, including Oral cancer, Lung cancer, Chronic obstructive pulmonary disease, Ischemic Stroke, and other relevant diseases (International Health Policy Program, Thailand, 2009). Supported by the research on smoking behavior smokers' life expectancy is 10 years (U.S. Department of Health and Human Services, 2014) less than non-smokers while those succeeding in smoking cessation before 40 years of age results have 90 percent (U.S. Department of Health and Human Services, 2014) (Jha P et al., 2013) of reduction of smoking-induced death.

Successful smoking cessation is considered to be difficult. Nearly all Smokers know and realize about how harmful smoking affects to their health, and desire to quit smoking. Annually, 30.4 percent represents the smokers trying to quit annually. Only 5.4 percent tries 5 times and over, without success ( American Psychiatric Association, 1996). Thailand at present implements the 20-year strategic plan aimed at people's longer life, 80 years old at least ( Office of the National Economics and Social Development Board, 2017) The life expectancy of Thai males stays at 72 years and females at 78 years (Institute for Population and Social Research, 2017). They experience the shorter years of living and lower quality of life. The major causes of death are premature death from non-communicable diseases and risk factors, for instance smoking (International Health Policy Program. Thailand, 2009). Helping smokers to quit smoking is an important duty of medical personnel whose roles should

focus on proactive services through health promotion and disease prevention rather than treatment. An explicit example of health promotion is the provision of guidance on smoking cessation, which later resulting in less caregiving burdens for patients with chronic diseases from smoking.

Therefore, the data was run on the process of Data Visualization which passed ETL and Transformation by the information came from interviews the executives and personnel for making BI the overview of the smoker's information in each region that has density of smokers. The project will predict budget in curing smokers' addiction in each region and country. And it also shows the regional drug information for each type that is most effective. And can also be used in other areas.

## **1.2 Objectives**

Preparing presentation by Data Visualization using BI in overview of the number of patients in each region of the country, the status of the treatment, the type of drug, and the overall budget spent.

## **1.3 Scope of Work**

Study on the therapy, drug dispensing information and duration of smoking cessation from patient history data with track 6 months within the smoking cessation clinic. The participating Fah Sai Clinics of a medical professional to control tobacco consumption between 2558 and 2560 and using data prepared by BI tools to make a report and presentation to see the results or looking the information easily and can be implemented through the BI tool utilize for the future.

## **1.4 Expected Result**

Personnel can bring overall information to diagnose or study to help treat smoking addiction personnel and related agencies plan strategic and budget to be allocated to each clinic or hospital to help treat smoking addiction.

## **CHAPTER II**

### **LITERATURE REVIEW**

In this study, Spatial Visualization for Smoking Cessation Therapy Program: Success and Choices of Medication Approaches, the following relevant concepts, theories and studies guide the research framework are following.

#### **2.1 General information about tobacco and treatment of tobacco dependence**

##### **2.1.1 Tobacco information**

Tobacco is a product extracted from leaves of the tobacco plant, called *Nicotiana Tabacum*. It is consumed by smoking, sucking, chewing, eating, blowing and inhaling through mouth or nose with either electric or nonelectric instrument for the similar result (Tobacco Control Act, 1992).

##### **2.1.2 Smoking cessation assistance**

The treatment in the present day has different options that reinforce smoking cessation. These options come with various processes and durations (Somkiat Wattanasirichaigoon, 2007). The two treatment options are pharmacological treatment and non-pharmacological treatment.

###### **1. Pharmacological treatment**

This treatment reduces the suffering from nicotine withdrawal symptoms. However, it contributes only 10-30 percent of cessation success for one-year duration (U.S. Department of Health and Human Service, 2000). In the longer term, as stated in medical evidences, the pharmacological method provides more chances of cessation. Consequently, most national and international organizations strongly encourage patients in utilizing this method. The exception goes to women

with pregnancy or breastfeeding, people smoking less than 10 cigarettes a day and those in restricted medication. The pharmacological treatment has two types of drug, namely the first-line drug and the second-line drug (Suthat Rungruanghiranya, 2010).

### 1.1 Nicotine replacement therapy (NRT)

The nicotine replacement therapy is widely known. Nicotine patches and gums are mostly used in Thailand. The patch type activates slowly through skin. In contrary, nicotine in the gum is absorbed and activates much faster. Both drugs are prohibited in pregnant or breastfeeding patients. Smokers having the severe premature ventricular contraction are suggested to not avoid this therapy, not

#### - Nicotine replacement therapy (NRT)

The nicotine replacement therapy is widely known. Nicotine

patches and gums are mostly used in Thailand. The patch type activates slowly through skin. In contrary, nicotine in the gum is absorbed and activates much faster. Both drugs are prohibited in pregnant or breastfeeding patients. Smokers having the severe premature ventricular contraction are suggested to not avoid this therapy, not until the medical proof allows.

#### - Nicotine Gum

The dose varies by the number of cigarettes smoked per day. The suggestion for those smoking more than 20 cigarettes is the 21-milligram patch per day and 14-milligram patch per day for less than 20 cigarettes. General instructions are that, one patch is required per day for consecutive days for four weeks. Then a dose reduction to 14 and 7 milligrams per day every two to four weeks. As soon as starting the patch, the patient must quit smoking until the medical proof allows.

### 1.2 Drug for nicotine withdrawal and craving

Bupropion is a 150-milligram pill reducing craving for tobacco. Application of this nicotine-free drug is different from the above nicotine replacement therapy. The doctor requires the patient to set the target quit date. The start of taking the pill is scheduled one week before the set date. The instruction includes taking one pill daily in the morning and in the first three days, increasing the

dose to one pill twice daily for three to six months and quitting completely from the eighth day after starting using this pill for best results. Regularly found side effects are difficult sleeping (35-40 percent), dry mouth and throat (10-70 percent), constipation (9 percent) nausea (6 percent) and seizure (0.1 percent). To cure difficult sleeping, the patient should take the second pill in the evening, not afternoon, or eight hours after the first pill in the morning. This pill is prohibited in patients having records of epilepsy, eating disorders and taking monoamine oxidase inhibitors in the past two weeks. Pregnant or breastfeeding women should avoid taking this pill unless necessary. Varenicline is both a nicotine partial agonist and an antagonist against nicotinic acetylcholine receptor. Activation of the partial agonist of alpha4 beta2 nicotine receptors in the ventral tegmental area causes dopamine release, making smokers feel nicotine. Varenicline interacts with nicotine receptors instead of nicotine and creates nicotinic actions without nicotine intake, leading to pleasurable effects of smoking and decreasing tobacco craving. In addition, while stimulating nicotine receptors, this pill blocks them and prevents nicotine from attaching to them. Studies at present find the varenicline successful in smoking cessation at 44 percent. The appropriate dose is one milligram twice (morning and evening) daily. Nortriptyline is a tricyclic antidepressant blocking norepinephrine and 5-hydroxytryptamine in the peripheral nervous system and reducing nicotine withdrawal symptoms. Many studies reveal the higher rate of tobacco cessation as compared with placebo. The suggestion is the 25-milligram dose daily for the beginning and then a slight increase of dose is feasible up to the 75 milligram daily. The start of drug taking, however, should be 10-28 days before the target cessation date. Normally, the treatment takes around 12 weeks. Common side effects are blurring vision, dry mouth, drowsiness, urinary retention, dizziness, shaky hands and cardiac arrhythmia. Patients with cardiovascular disease should avoid taking the nortriptyline.

## 2. Non-pharmacological treatment

Advice and behavioral change are popular options among health personnel. Group or individual advising and advice via telephone are effective.

### **2.1.3 Factors influencing smoking cessation**

According to the medical practice for nicotine addiction treatment among smokers, (Waraporn Bhumisawat et al, 2003) introduces two variables that influence the success rate of smoking cessation.

□ Variables influencing the success rate of tobacco cessation are: Patient having high incentive to quit smoking; Patient having preparedness for changes (ready to quit within one month); Patient having confidence, at moderate to high level, in their ability (believing that they can quit smoking) and; Patient having social support (having no smoking family member or smoking friend during smoking cessation).

□ Variables influencing the low success rate of tobacco cessation are: Patient having high nicotine addiction, smoking the first cigarette within 30 minutes after waking up in the morning and smoking more than 20 cigarettes per day (with severe withdrawal symptoms during smoking cessation); Patient having psychiatric disorder (depression, alcohol or substance induced psychotic disorder) and; Patient with high level of stress (experiencing, at present or in the future, life transition such as divorce and employment change).

## **2.2 Data analysis**

Data analysis is the information obtained. Understanding Data analysis and make new Information, such as adding relevant information. Or the existing data has been analyzed. Get data ready to implement Data Visualization. By the way, all three methods are as follows one part is based on the CRISP-DM process.

### **2.2.1 Business Understanding**

It is a step in understanding the problem or business opportunity then convert the problem into a suitable format for analysis.

### 2.2.2 Data Understanding

It is a step in collecting relevant data for analysis using data mining techniques. It should also be considered that the information obtained from reliable sources is sufficient for the analysis.

### 2.2.3 Data Preparation

It is the process of preparing the data and takes the longest time since the data mining model will get the right results depending on the quality of data used. If the information used is incorrect. The results will be wrong as well. By preparing the information. Can be divided into 3 sub-stages as below

**1. Data Selection** - it is a decision making to select important data or factors for analysis in order to create data mining.

**2. Data Cleansing** - it is a process to allow data prompt to application such as making data accurate, addressing noising data and deleting incorrect data.

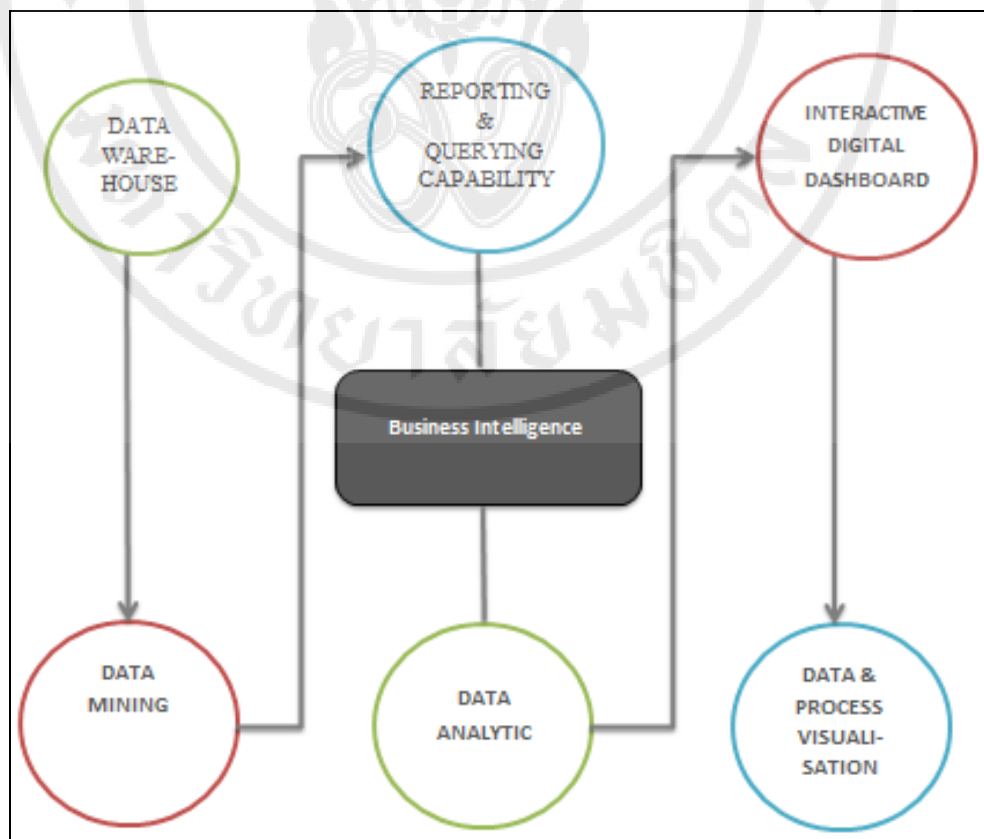
**3. Data transformation** - it is to change data to be the appropriate model for analysis application according to the algorithm and the model in creating data mining.

## 2.3 Business Intelligence

It is a framework designed to support decision making processes. This structure includes database architecture, analytic tools and applications. Business Intelligence is an important part of business intelligence. Business Intelligence includes data warehousing, business analytics, performance, strategy, and user interface. The principle is to get information from a variety of sources. The data warehouse is then managed and used. Authorized users can access this information and work with this information to get the results they need. This result will be shared with decision makers. These results can be published via a dashboard or share point. (Wang, Alexander, 2015)

### Benefits of Business Intelligence

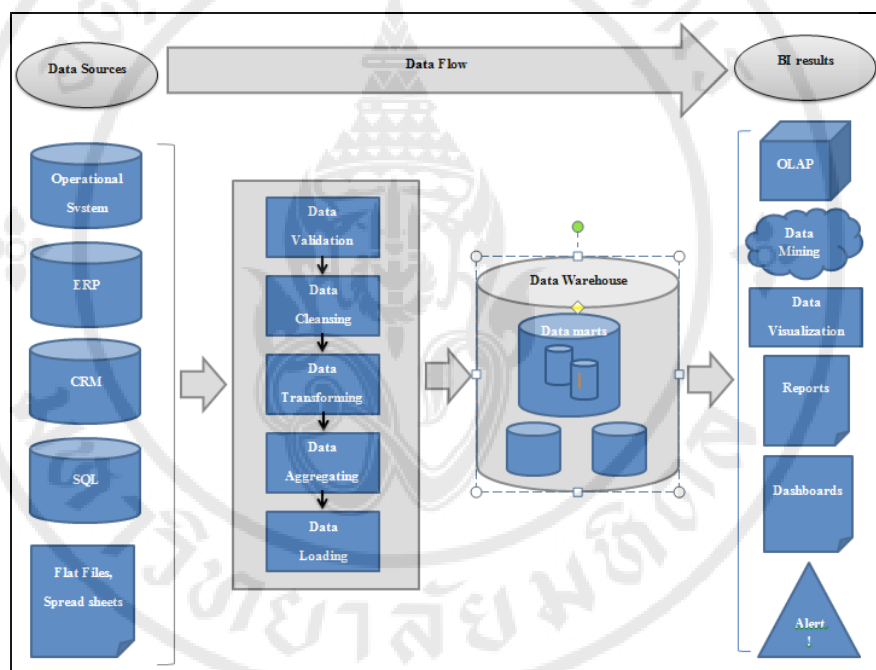
- Business intelligence is faster more accurate process of reporting critical information.
- Business intelligence facilitates better and efficient decision-making process.
- Business intelligence provides timely information for better customer relationship management.
- Business intelligence improves profitability of the company
- Business intelligence provides a facility of assessing organization' s readiness in meeting new business challenges.
- Business intelligence supports usage of best practices and identifies every hidden cost.



**Figure 2.1** Business Intelligence Workflow

## 2.4 Business Intelligence Architecture

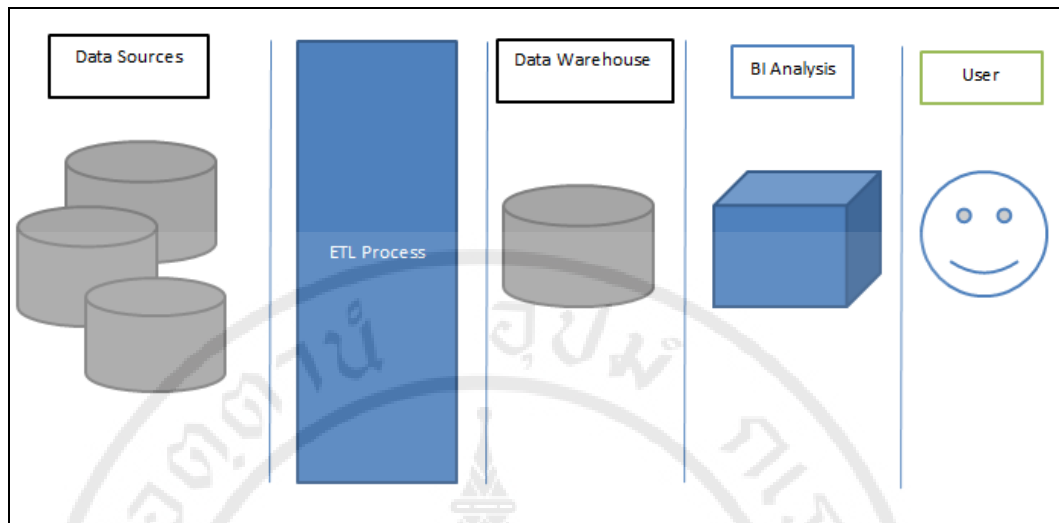
A business intelligence architect (BI architect) is a top-level sort of business intelligence analyst who deals with specific aspects of business intelligence, a discipline that uses data in certain ways and builds specific architectures to benefit a business or organization. The business intelligence architect will generally be responsible for creating or working with these architectures, which serve the specific purpose of maximizing the potential of data assets. (Sharan, 2016)



**Figure 2.2** Business Intelligence Architecture

The key elements of business intelligence architecture are:

- Source systems
- ETL process
- Data modeling
- Data warehouse
- Enterprise information management (EIM)
- Appliance systems
- Tools and technologies



**Figure 2.3** Business Intelligence.

## 2.5 Data Visualization

It is viewed by many disciplines as a modern equivalent of visual communication. It involves the creation and study of the visual representation of data, meaning "information that has been abstracted in some schematic form, including attributes or variables for the units of information". (DataViz, Inc., 2015)

Data visualization refers to the techniques use to communicate data or information by encoding it as visual objects (e.g., points, lines or bars) contained in graphics. The goal is to communicate information clearly and efficiently to users. It is one of the steps in data analysis or data science. According to Friedman (2008) the "main goal of data visualization is to communicate information clearly and effectively through graphical means. It doesn't mean that data visualization needs to look boring to be functional or extremely sophisticated to look beautiful. To convey ideas effectively, both aesthetic form and functionality need to go hand in hand, providing insights into a rather sparse and complex data set by communicating its key-aspects in a more intuitive way. Yet designers often fail to achieve a balance between form and function, creating gorgeous data visualizations which fail to serve their main purpose — to communicate information".

Data visualization is closely related to information graphics, information visualization, scientific visualization, exploratory data analysis and statistical graphics. In the new millennium, data visualization has become an active area of research, teaching and development. According to Post et al. (2002), it has united

Scientific and information visualization. ([https://en.wikipedia.org/wiki/Data\\_visualization](https://en.wikipedia.org/wiki/Data_visualization))

## 2.6 Relevant Research

Exploring the current related researches on Classification related with quit smoking, it was found that data mining has been applied with many works, such as:

- Kelvin KF Tsoi et al. (2016) proposed the research “Data Visualization Tools for Investigating Health Services Utilization among Cancer Patients” They will have to develop tools for visualizing information; generating insight; and creating actionable, on-demand knowledge for clinical decision making. This chapter has three objectives: 1) to review the data visualization tools that are currently available and their use in oncology; 2) to discuss implications for research, practice, and decision making in oncology; and 3) to illustrate the possibilities for generating insight and actionable evidence using targeted case studies. A few innovative applications of data visualization are available from the clinical and research settings. We highlight some of these applications and discuss the implications for evidence generation and clinical practice. In addition, we develop two case studies to illustrate the possibilities for generating insight from the strategic application of data visualization tools where the interoperability problem is solved. Using linked cancer registry and Medicare claims data available from the National Cancer Institute, we illustrate how data visualization tools unlock insights from temporal event sequences represented in large, population-based datasets. We show that the information gained from the application of visualization tools such as Event Flow can define questions, refine measures, and formulate testable hypotheses for the investigation of cancer-related clinical and process outcomes.

- Eberechukwu Onukwugha et al. (2017) proposed the research “ Data Visualization on Global Trends on Cancer Incidence an Application of IBM Watson

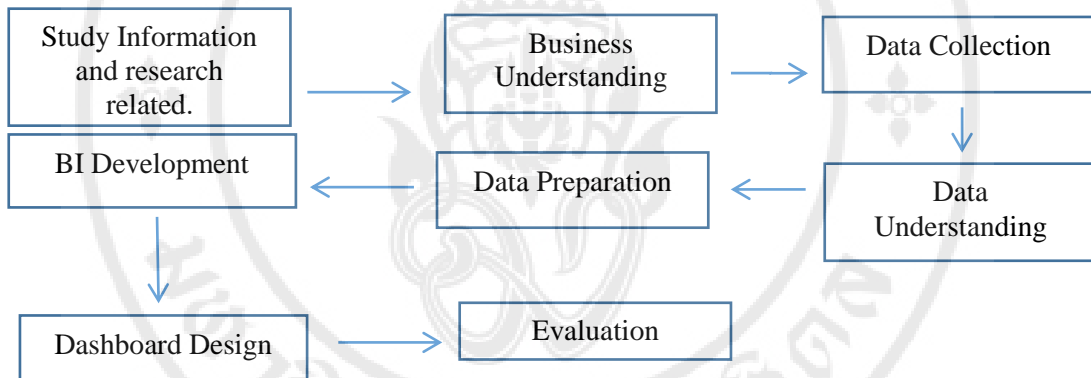
Analytics” Visual analytics is widely used to explore data patterns and trends. This work leverages cancer data collected by World Health Organization (WHO) across over a hundred of cancer registries worldwide. In this study, we present a visual analytics platform, IBM Watson Analytics, to explore the patterns of global cancer incidence. We included 26 cancers from different geographic regions. An interactive interface was applied to plot a choropleth map to show global cancer distribution, and line charts to demonstrate historical cancer trends over 29 years. Subgroup analyses were conducted for different age groups. With real-time interactive features, we can easily explore the data with a selection of any cancer type, gender, age group, or geographical region.

- Oluwakemi Ola et al. (2016) proposed the research “Beyond simple charts: Design of visualizations for big health data” Health data is often big data due to its high volume, low veracity, great variety, and high velocity. Big health data has the potential to improve productivity, eliminate waste, and support a broad range of tasks related to disease surveillance, patient care, research, and population health management. Interactive visualizations have the potential to amplify big data’s utilization. Visualizations can be used to support a variety of tasks, such as tracking the geographic distribution of diseases, analyzing the prevalence of disease, triaging medical records, predicting outbreaks, and discovering at-risk populations. Currently, many health visualization tools use simple charts, such as bar charts and scatter plots, that only represent few facets of data. These tools, while beneficial for simple perceptual and cognitive tasks, are ineffective when dealing with more complex sense making tasks that involve exploration of various facets and elements of big data simultaneously. There is need for sophisticated and elaborate visualizations that encode many facets of data and support human-data interaction with big data and more complex tasks. When not approached systematically, design of such visualizations is labor-intensive, and the resulting designs may not facilitate big-data-driven tasks. Conceptual frameworks that guide the design of visualizations for big data can make the design process more manageable and result in more effective visualizations. In this paper, we demonstrate how a framework-based approach can help designers create novel, elaborate, non-trivial visualizations for big health data. We present four visualizations that are components of a larger tool for making sense of large-scale public health data.

## CHAPTER III

### RESEARCH METHODOLOGY

Data from the smoking cessation clinic are increase from such information can benefits of data visualization. To be provide an overview of who is admitted in each region. And the budget spent on treatment from the patient. To be able to take that information to analyze, plan to continue the budget of the drug in the treatment.



**Figure 3.1** Methodology workflow

### 3.1 Study the inventory models

From the study of information and the related research found Data Visualization that is very useful for large data processing. And display in a diagram format for help decision makers and to the organization in the planning.

### 3.2 Select the suitable method

Study report of smoking from Treatment of tobacco dependence by utilizing the 5A's strategy (Fiore MC, et al. 2000) (Medical Guideline for Smoking treatment in Thailand for Medical Doctors and Health Professionals, 2009), Smoking cessation assistance from the two treatment options are pharmacological treatment and

non-pharmacological treatment ( Somkiat Wattanasirichaigoon, 2007) , Stepwise approach to pharmacological treatment ( Suthat Rungruanghiranya, 2010) Factors influencing smoking cessation (Waraporn Bhumisawat, 2003) and other factors.

### **3.3 Data Collection**

It is a process of collecting data from the source. This research received and study more to get the data needed to use Data Visualization. The information is as follows.

3.3.1 Data from the National Bureau of Statistics is a survey of smoking and alcohol drinking behavior in the population in 2014.

3.3.2 Drug Information for Smoking Cessation

- Bupropion price of 25 baht per tablet should be taken 2 tablets per day.

- Varenicline

- Nortriptyline

3.3.3 Information on the location of hospitals and clinics.

### **3.4 Data Understanding**

This study makes an exploration of data concerning smoking treatment, dispensary and duration of their smoking cessation. As provided by clinics participating in the smoking cessation program under the Thai Physicians Alliance Against Tobacco (TPAAT), the mentioned data are gained from smoking treatment records of smokers or patients having doctors' follow-up within six months. Data was collected in 2015 to 2017.

**Table 3.1** Data structure from smoking cessation clinic.

Attributes	Definition
ID	Numeric, Auto increment
Hospital ID	Hospital ID
Patient ID	Patient identification
Register Date First treatment date	Register Date First treatment date
Age (year)	Numeric
Gender	Gender code 1= male, 2 = female
Marital status	,single, 1=married, 2=widowed=0 divorced=3
History of Alcohol Disease	yes, 2=sometimes, 3=no=1 =1none, =2diabetes, =3hypertension, =4coronary artery disease, =5stroke, =6emphysema, =7cancer, =8liver, =9nephropathy, ,asthma=10 ,allergy =11 others=12
Tobacco Type	,filter =1 ,hand rolling =2 others =3
Smoking duration (years)	Total of years smoking
Number of cigarettes on days (total)	Average number of cigarettes

**Table 3.1** Data structure from smoking cessation clinic. (cont.)

<b>Attributes</b>	<b>Definition</b>
Time to first cigarette	Minutes
Carbon monoxides (ppm)	The measurement of carbon monoxide test (ppm)
Counseling	Person code to counseling
Medication	Used / not used
First line medication	1=Nicotine gum 2=Nicotine patch 3=Nortriptyline 4=Bupropion 5=Varenicline 6=Wash mouth 7= Vernonia 8=Others
Second line medication	1= Nicotine gum 2= Nicotine patch 3= Nortriptyline 4= Bupropion 5= Varenicline 6= Wash mouth 7= Vernonia 8= Others

**Table 3.2** Data structure from smoking cessation clinic, Follow up 3 months.

<b>Attributes</b>	<b>Definition</b>
Treatment Date	Date
CO	The measurement of carbon monoxide test (ppm)
Weight	Weight of patients

**Table 3.2** Data structure from smoking cessation clinic, Follow up 3 months. (cont.)

<b>Attributes</b>	<b>Definition</b>
Status of quit smoking	Status code = 1 quit smoking = 2 smoking
Withdrawal severity	Scale 1-10
Number of cigarettes on days (total)	Average number of cigarettes consumed on days
Medication	Used / not used
First line medication	1=Nicotine gum 2=Nicotine patch 3=Nortriptyline 4=Bupropion 5=Varenicline 6=Wash mouth 7= Vernonia 8=Others
Second line medication	1= Nicotine gum 2= Nicotine patch 3= Nortriptyline 4= Bupropion 5= Varenicline 6= Wash mouth 7= Vernonia 8= Others

**Table 3.3** Data structure from smoking cessation clinic, Follow up 6 months.

<b>Attributes</b>	<b>Definition</b>
Treatment Date	Date
CO	The measurement of carbon monoxide test (ppm)
Weight	Weight of patients
Status of quit smoking	Status code = 1 quit smoking = 2 smoking
Withdrawal severity	Scale 1-10
Number of cigarettes on days (total)	Average number of cigarettes consumed on days
Medication	Used / not used
First line medication	1=Nicotine gum 2=Nicotine patch 3=Nortriptyline 4=Bupropion 5=Varenicline 6=Wash mouth 7= Vernonia 8=Others
Second line medication	1= Nicotine gum 2= Nicotine patch 3= Nortriptyline 4= Bupropion 5= Varenicline 6= Wash mouth 7= Vernonia 8= Others

1. ประวัติผู้ป่วย (สำหรับ visit 1 / แรกเริ่ม)								
id	รหัสโรงพยาบาล	รหัสผู้ป่วย	เดือนวัน ลงทะเบียน	อายุ	เพศ	สถานะภาพ	ดื่มสุรา	รวมโรคประจำตัว
1			140807	33	1	0	1	35
2			140807	30	1	0	2	1
3			140807	29	2	0	1	1
5			140807	38	2	1	1	1
6			140807	47	2	1	2	3
9			140807	37	1	1	2	1
11			140807	46	2	2	2	3
12			140807	37	2	0	1	1
13			140807	27	2	0	1	234
14			140807	28	2	0	2	1
15			140807	30	2	0	1	1
16			140807	33	1	0	3	1
17			140807	32	1	0	3	1
18			140807	32	1	0	3	1
19			140813	41	1	1	2	34
20			140813	40	2	1	1	1

Figure 3.2 Example of data in smoking cessation clinic before data preparation.

4. ประวัติการติดตามผล 3 เดือน						
ปีที่เข้ารับการรักษา	ค่า CO	น้ำหนัก (กก.)	หรีได้ตามวันที่	ครบรุนแรงอาการ	ยังสูบบุหรี่	ยังสูบบุหรี่วันละกี่มวน (ระบุจำนวนที่สูบ)
140807	20	ไม่มีข้อมูล	2	5	1	20
140807	5	ไม่มีข้อมูล	1	0	2	0
-	-	-	-	-	-	-
-	-	-	-	-	-	-
150407	0	ไม่มีข้อมูล	1	0	2	0
-	-	-	-	-	-	-
-	-	-	-	-	-	-
140807	6	ไม่มีข้อมูล	1	0	2	0
140807	20	ไม่มีข้อมูล	1	0	2	0
150407	0	ไม่มีข้อมูล	1	0	2	0
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
150217	0	ไม่มีข้อมูล	2	2	1	5
-	-	-	-	-	-	-
-	-	-	-	-	-	-

Figure 3.3 Example of follow up 3 months.

### 3.5 Data Preparation

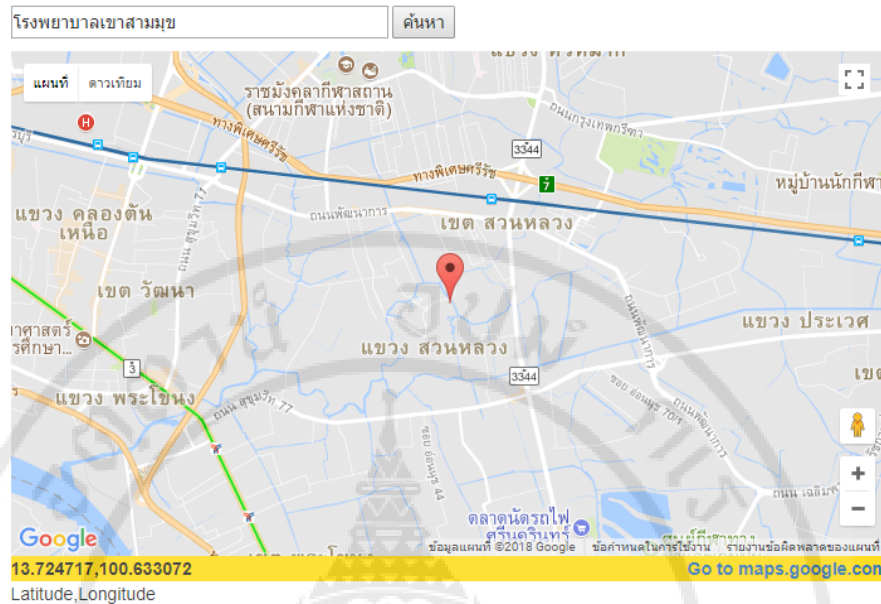
This step summarizes the important data that needs to be implemented in Data Visualization to obtain a Dashboard. The process of importing data to be considered, including the data format. The selection of unwanted information. To reduce the amount of data to consider and to increase the efficiency of data visualization.

3.5.1 The first step leave out the unrelated data and non-affecting data in order to reduce the number of factors needed consideration and to increase their validity to the results of the data mining as below.

**Table 3.4** Excluded Attributes.

N	Attributes	Note
1	Patient ID	Data Privacy
2	Register Date	
4	Carbon monoxide value	The data in the system is very low.
5	Weight	The data in the system is very low.
6	Counseling Code	No import data because of unethical .research

3.5.2 Researching and preparing additional information Use hospital and clinical names to find latitude and longitude (<http://map.rimnam.com/>)



**Figure 3.4** The picture is use the web for find out latitude and longitude.

### 3.6 BI Development

In our research we designed Information in the following dimensions:

- Treatment Information
- Location of hospital
- Drug price information
- Budget estimate data

### 3.7 Dashboard Design

In this process, we will design components of the dashboard. It consists of many sheet we have designed sheet as follows below.

### 3.8 Evaluation

Dispatch to the agency and the people who want to use it, see and evaluate or feedback.

### 3.9 Research Timeline

Table 3.5 Research Timeline

Activity	Dec (2017)	Jan (2018)	Feb (2018)	Mar (2018)	Apr (2018)	May (2018)	Jun (2018)	Jul (2018)	Aug (2018)	Sep (2018)	Oct (2018)	Nov (2018)	Dec (2018)
Study Information and research related													
Business Understanding													
Data Collection													
Data Understanding													
Data Preparation													
BI Development													
Design													

## **CHAPTER IV**

### **RESULTS AND DISCUSSION**

In this chapter are presented the result of design of story and dashboard of Information about treatment smoker. Moreover, to describe the result on questionnaire in evaluation by Project Administrator of CLINIC FHASAI project.

#### **4.1 Result of problem analyze from collection data**

The source of the data comes from 2 parts. The first data from Project Administrator. The other parts was collection of information from other source in internet that of all are as follows:

##### **4.1.1 Project Administrator interview**

The interview form Project Administrator is about design dashboard of visualization which he needs to see it.

Project Administrator need to show patient can stop smoking in each 3 month and 6 month.

Project Administrator need to show summary of project in total number of patients and total number of hospital.

Project Administrator need to show relationship between data patient and NCDs (Non-Communicable Diseases)

#### **4.2 Analyze and design dashboard**

After analyzing the data from chapter 3 that has been prepared for Tableau Program by present related data to show that information of CLINIC FHASAI Project in dashboard for usefully decision making and support the Project administrator. The development dashboard was implementing on data visualization follow:

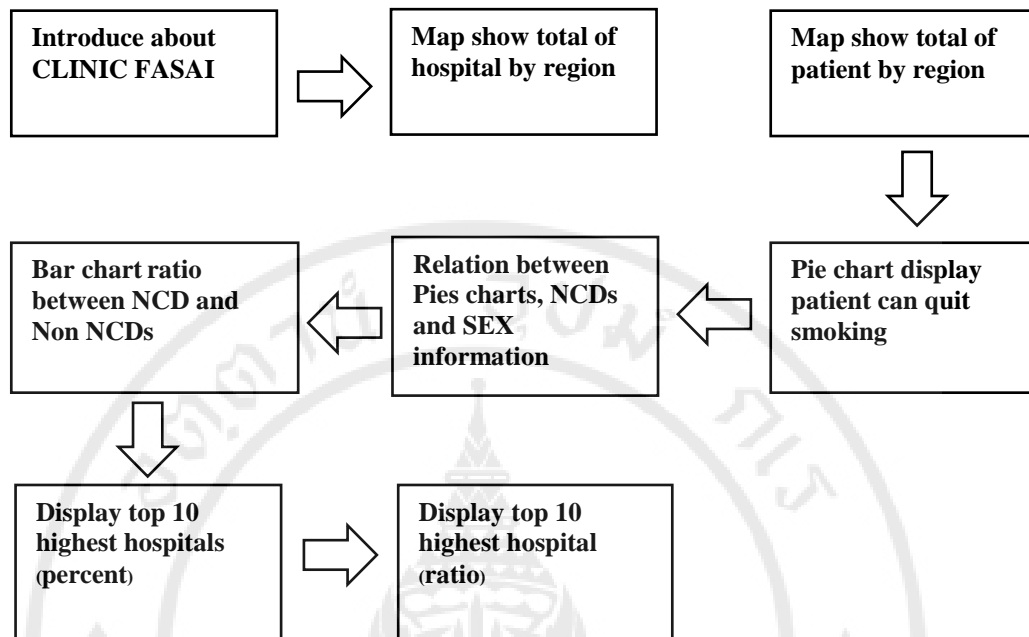


Figure 4.1 Flow display sequence storyboard

โครงการเครือข่ายคลินิกฟ้าใส  
อยู่ภายใต้การกำกับดูแลของ  
เครือข่ายวิชาชีพแพทย์



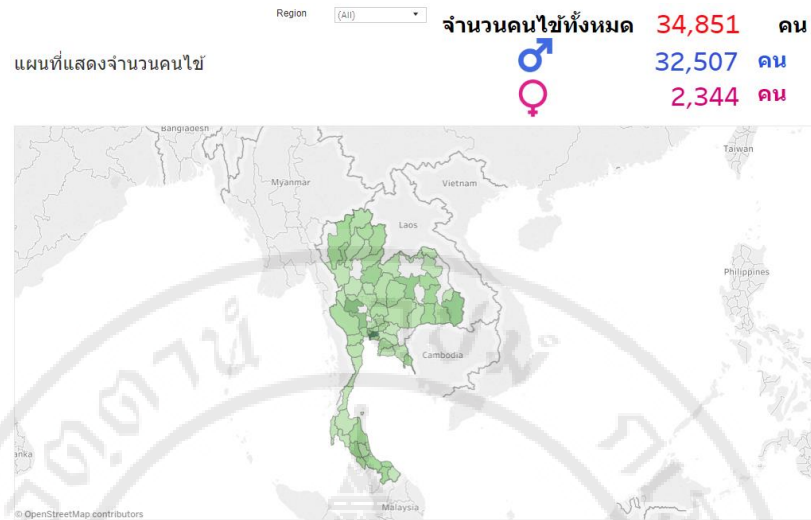
วัตถุประสงค์เพื่อบำบัดรักษา  
โรคเรื้อรัง จัดตั้งขึ้นเมื่อวันที่  
3 มิถุนายน 2553

ปัจจุบันมีโรงพยาบาลเข้าร่วม  
โครงการ 552 แห่งทั่วประเทศ



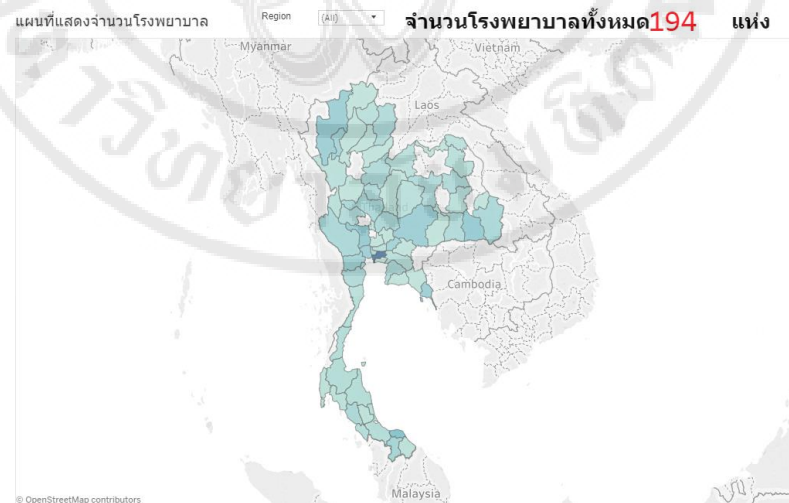
Figure 4.2 Introduction of Storyboard

Map show detail about total of patient in CLINIC FHASAI Project in Each region that how many people was treatment.



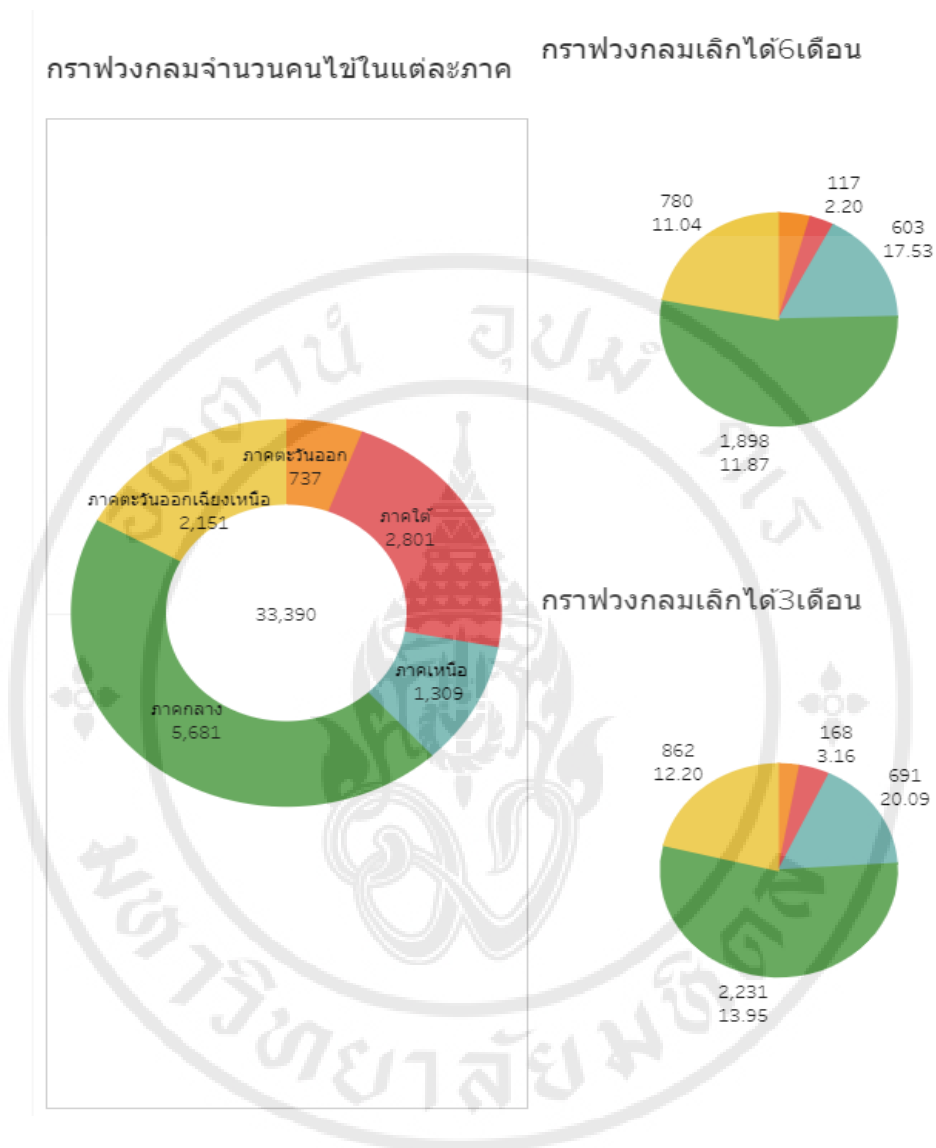
**Figure 4.3** Map show total of patient by region

Map show detail about total of patient in CLINIC FHASAI Project in each region that how many people was treatment.



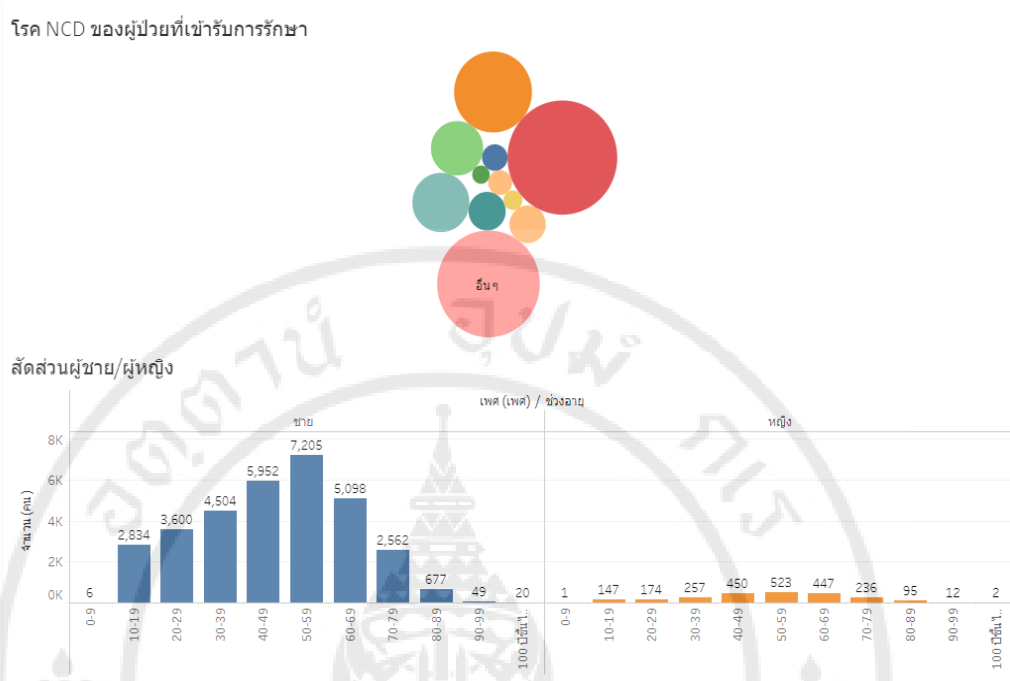
**Figure 4.4** Map show total of hospital by region

Map show detail about total of hospital in CLINIC FHASAI Project in each region



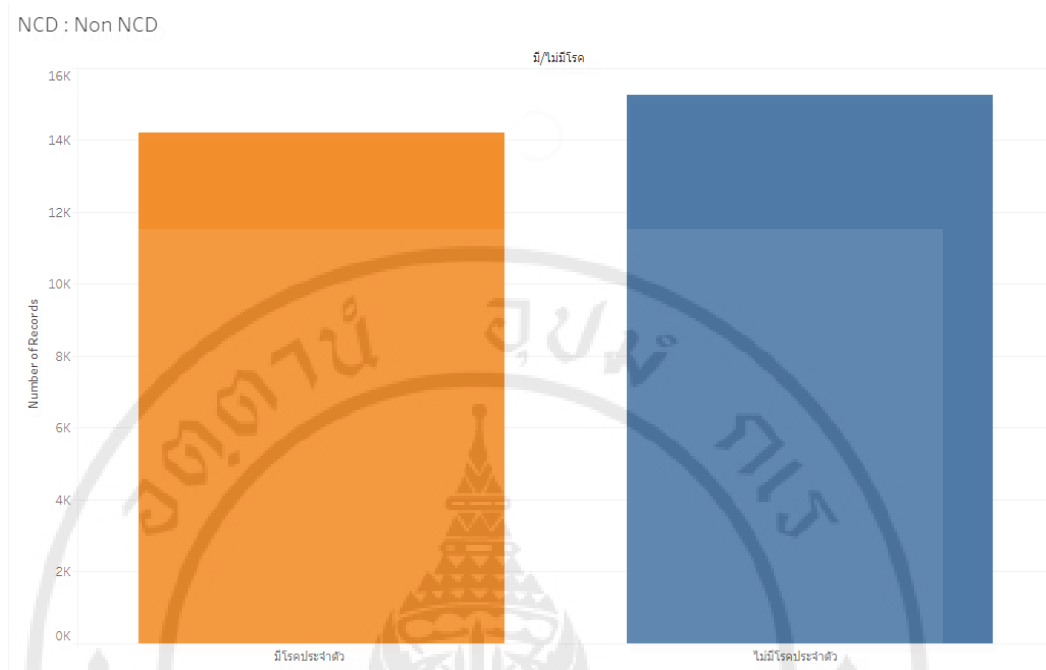
**Figure 4.5** Pie chart display relationship of patient can quit smoking in each region

Pie chart as Figure 4.4 contains 3 pie chart. The first pie chart display patient who can quit smoking spit by region, the information are total of patients and calculate as percentage. Second is pie chart display as patient who can quit smoking of people who can quit smoking in 6 months period. The last one is pie chart display as patient who can quit smoking of people who can quit smoking in 3 months period.



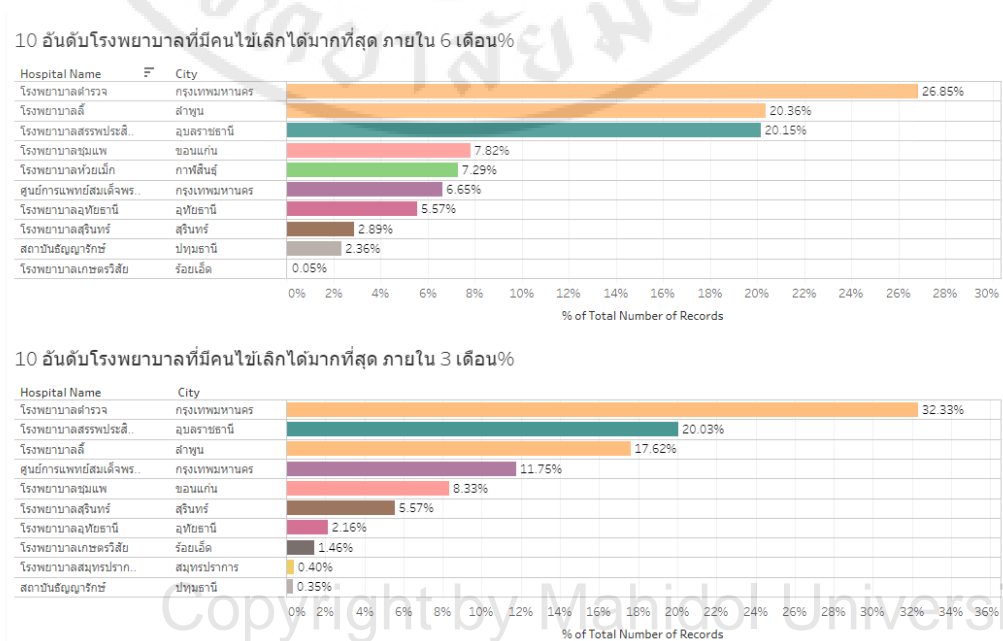
**Figure 4.6** Relationship between pie chart NCDs information and SEX information

As Figure 4.5 display relationship between patients has NCDs and range of ages and sex. Pie chart display all about of NCDs which patient has its. Bar chart display information sex and separate by range of age that range was separate 10 of age. Pie chart and bar chart have relationship between NCDs, sex and age.



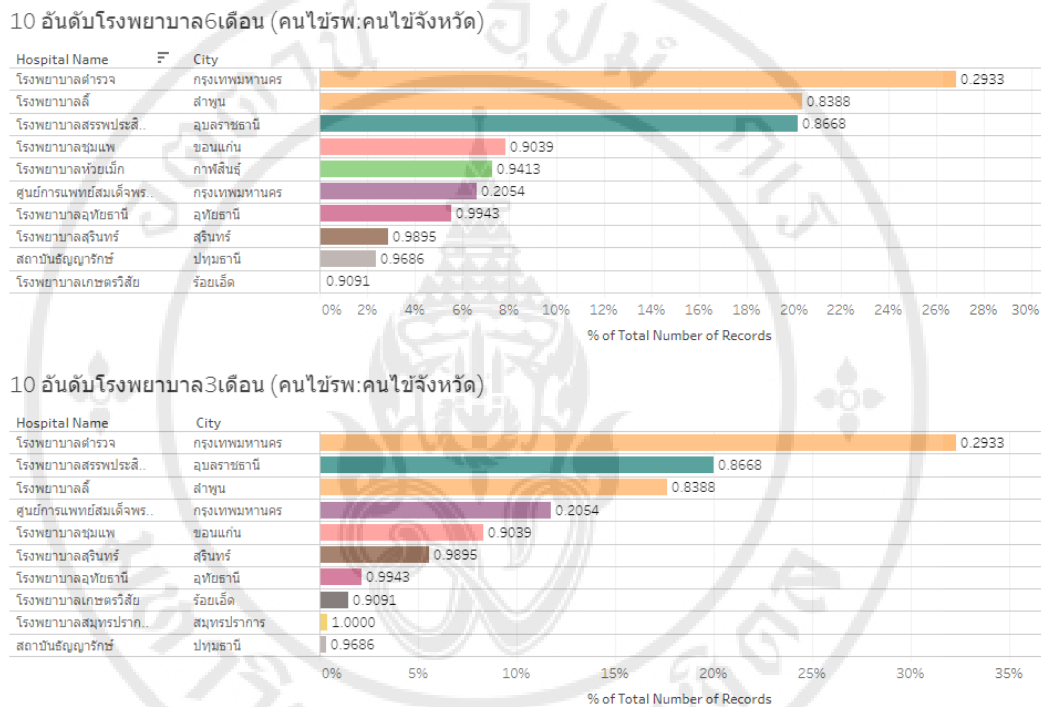
**Figure 4.7** Bar chart compare between patients has NCDs and not have.

That has 2 bar chart is the first one display information of patients has NCDs and the other display information of patients not has NCDs.



**Figure 4.8** Bar graph display top 10 highest hospital that have the most people quit smoking

The information top 10 highest hospital that people have quit smoking show as bar graph. In each bar graph display the people who can quit smoking all of the people and calculate as percentage.



**Figure 4.9** Bar graph display top 10 highest hospital that have the most people quit smoking

The information top 10 highest hospital that people have quit smoking show as bar graph. In each bar graph display the people who can quit smoking all of the people and calculate as percentage.

### 4.3 Evaluation

Evaluated by Project administrator and manager of him. Development of Story board from contain dashboards and sheets used Tableau tool, it was develop of data visualization for decision making support. The information display about

overview of Project KLINIC FHASAI. The purpose of story board with principle on visualization was design by interview and get requirement by project administrator follow as:

- The story board helps to see overview about how many patient quit smoking at least in each 3 months period and 6 months period.
- Information all of patient in CLINIC FHASAI Project who was treatment how many patient has NCDs and what have disease?
- Information about ration between men and women.

**Table 4.1** Evaluation Result

<b>Dimension</b> <b>Storyboard</b>	<b>Management</b>	<b>Visualization</b>	<b>Data connectivity</b>	<b>Decision -making</b>	<b>Average</b>
<b>Storyboard 1</b>	4	3	4	3	<b>3.5</b>
<b>Storyboard 2</b>	4	4	5	4	<b>4.25</b>
<b>Storyboard 3</b>	4	4	5	4	<b>4.25</b>
<b>Storyboard 4</b>	4	4	5	4	<b>4.25</b>
<b>Storyboard 5</b>	4	4	5	4	<b>4.25</b>
<b>Storyboard 6</b>	4	3	5	3	<b>3.75</b>
<b>Storyboard 7</b>	4	4	5	4	<b>4.25</b>
<b>Storyboard 8</b>	4	4	5	4	<b>4.25</b>

## **CHAPTER V**

### **CONCLUSION**

#### **5.1 Conclusions**

In this thesis, the data from Fai Sai Clinic project, a program established under the supervisory authority of the Ministry of Health for smokers having the right therapy. The hospitals and patients in Thailand are participating 552 hospitals and 34851 people. Data are analyzed by using Business Intelligence tools for illustrating and presenting in Data Visualization. The easily overview will be presented for people who do not have the analysis knowledge. The data, from 2017 to 2018, will be regularly organized and managed to the next project. The data consist of hospitals, patients, treatment Information and more information that researching for hospitals to be able to use the Business Intelligence Tool effectively as possible.

#### **5.2 Limitation and future works**

In this research, we have used Business Intelligence Tool to carry out the process in the form of Data Visualizations. The final results will be displayed on the Story Board adopted with the CRIPS-DM theory and has been designed according to each dimension of the project leader. The data were accumulated by interviewing the project controller. Project controllers need all the patients admitted to the project within the project follow-up treatment that how many smokers have quit smoking? They want to know the Fai Sai clinical effectiveness that can effectively help treat people addicted to smoking or not. And also want to know the information as an overview that hospitalized patients with chronic disease before having the treatment. We also have to add another dimension to the story board more attractive and effective management overview in much diversity.

We have taken the completely story board to present project administrator. The comment is satisfied and we have prepared evaluation by including separate 5-

axis there are concluded Management, Visualization, Data connectivity and Data connectivity and have brought to calculate average in each of the criteria, which is satisfied.

In this work, the researchers are just the people who come for treatment only. The data is information that is not yet complete. Cause could be the system to collect data from hospital may not good enough so the initial data for this research is important to analyze because output has inefficient. All the information that researchers are interested in treatment, such as medication, thus making the research analysis only limited information. If we have information that drug treatment or more information, it will be able to work out a variety of research and much more interesting.

In the future, this research can contribute and useful to admin project or users who interested about this project. If there are detailed of the prices of medicines used to treat it can be prepared to demand of drugs and can also contribute to the budget of the project the story board has empowerment of presenting visual and developed storage system is also an important factor. Can make a story board with a lot of quality up and make it work, it is very useful to users.

## REFERENCES

- Dashboardinsight. (2018). From website: <http://www.dashboardinsight.com/articles/digital-dashboards/fundamentals/dashboard-evaluation.aspx>
- DataViz, Inc.. (2018). Data visualization, from website: <https://datavizcatalogue.com/>
- Eberechukwu, Onukwugha, Catherine Plaisant, Ben Shneiderman. (2016). Data Visualization Tools for Investigating Health Services Utilization Among Cancer Patients. From website: <https://www.cs.umd.edu>
- Kelvin KF Tsoi, Felix CH Chan, Hoyee W Hirai, Gary KS Leung, Yong-Hong Kuo Samson, TAI and Helen ML Meng. (2017). Data Visualization Tools for Investigating Health Services Utilization Among Cancer Patients. Proceedings of the 50th Hawaii International Conference on System Sciences. From website: [ieeexplore.ieee.org](http://ieeexplore.ieee.org)
- MAPRIMNAM. (2018). From website: <http://map.rimnam.com/>
- Oluwakemi Ola and Kamran Sedig. (2016). Beyond simple charts: Design of visualizations for big health data. From website: <https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1143&context=fimspub>
- Sharan. (2016, November). Business Intelligence Architecture. From website: <https://analyticstraining.com/business-intelligence-architecture/>
- THAIPUBLICA. (2018). From website: <https://thaipublica.org/2016/02/thaihealth-8-2-2557/>
- Yang, K. and Niu, X. (2016, September). Research on the Spare Parts Inventory. From website: [ieeexplore.ieee.org](http://ieeexplore.ieee.org)

**BIOGRAPHY**

<b>NAME</b>	Thanachaoat Techawattanasirikul
<b>DATE OF BIRTH</b>	16 June 1992
<b>PLACE OF BIRTH</b>	Bangkok, Thailand
<b>INSTITUTIONS ATTENDED</b>	Srinakharinwirot University, 2011-2014; Bachelor of Science ( Computer Science) Mahidol University, 2016-2018; Master of Sciences (Information Technology Management)
<b>HOME ADDRESS</b>	33/ 11, Suksan 6 Village, Kanchanapisek Road, Bangkae, Bangkok 10160 Tel. 095-512-9966 E-mail: ton.thanachaoat@gmail.com
<b>EMPLOYMENT ADDRESS</b>	Advanced Info Service Public Company Limited.1291/1 AIS Tower 2, Phaholyothin Road, Samsen nai, Phayathai, Bangkok 10400 Thailand Tel. 02-029-8196 E-mail: thanacht@ais.co.th