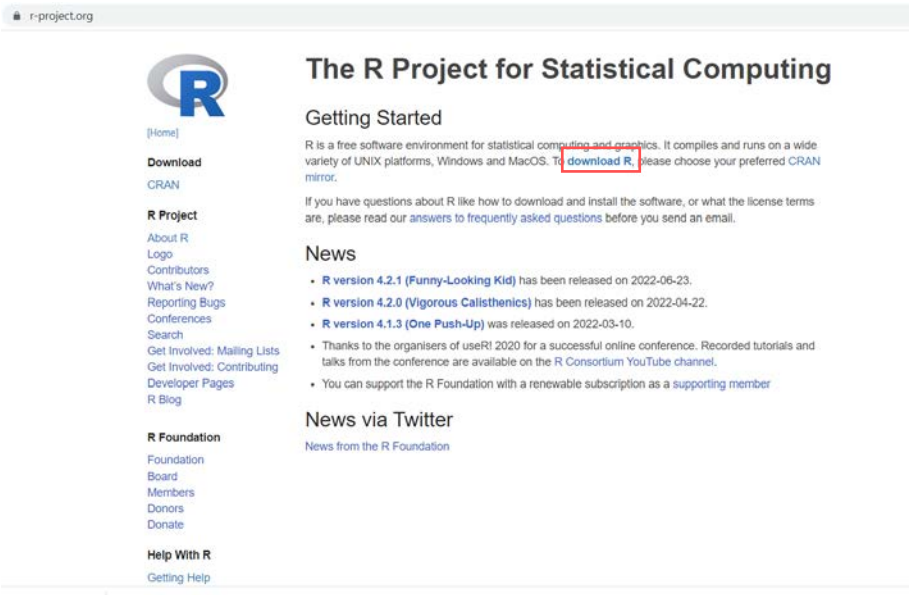
ภาคผนวกที่ 1

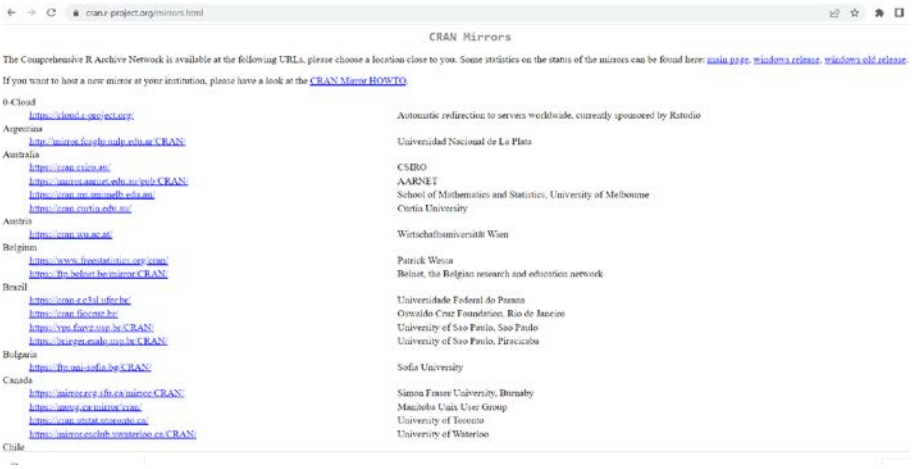
การใช้งานโปรแกรม R และ R – Studio

ขั้นตอนการติดตั้งโปรแกรม R

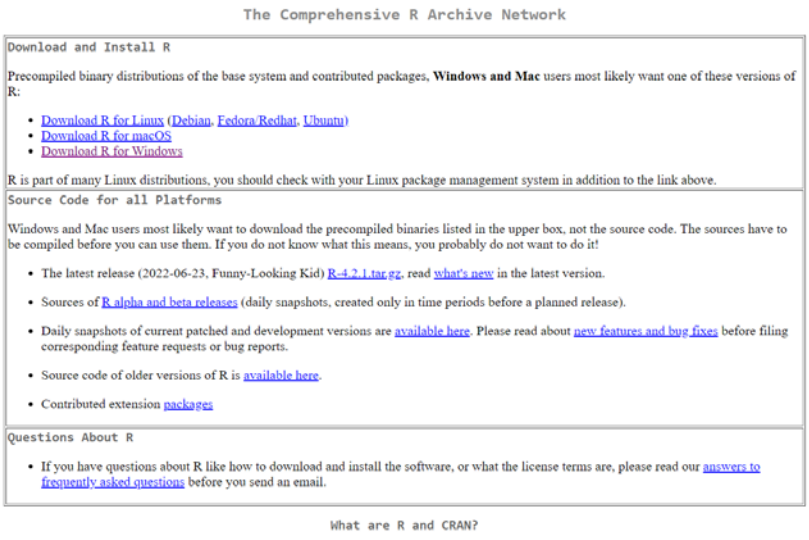
1. เข้า website https://www.r-project.org/ แล้วคลิ๊ก download R



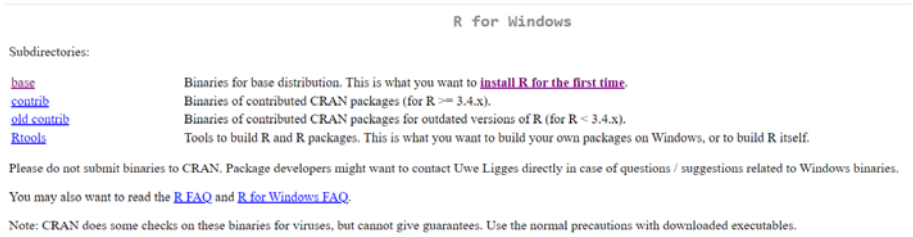
2. เลือก mirror ของ Thailand http://mirrors.psu.ac.th/pub/cran/



3. เลือก Download R for Windows (สำหรับผู้ใช้ Windows) หรือเลือก Download R for macOS (สำหรับผู้ใช้ macbook)



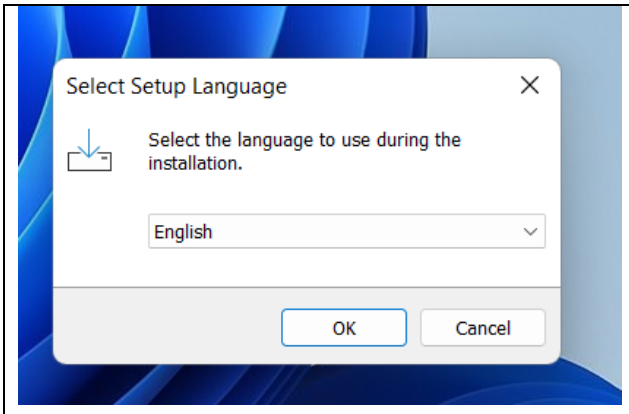
4. เลือก base หรือ install R for the first time.



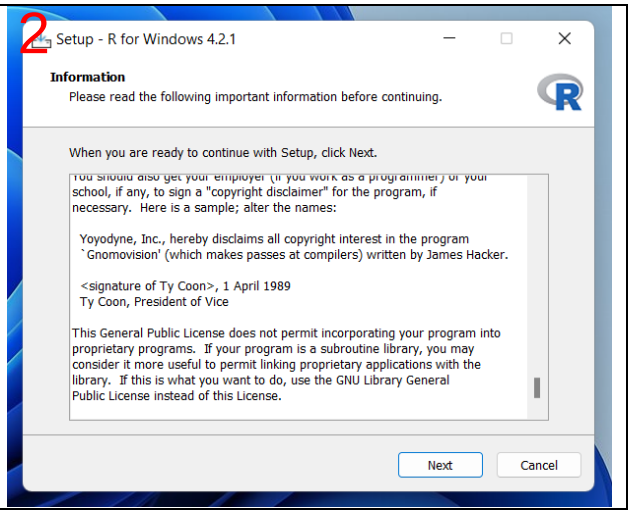
5. เลือก Download R-4.4.1 for Windows รอจนดาวน์โหลดเสร็จสิ้น (R-4.4.1 คือ version ล่าสุด หากใครเคยติดตั้ง R มาก่อนขอเป็น version 4.0 เป็นต้นมา)

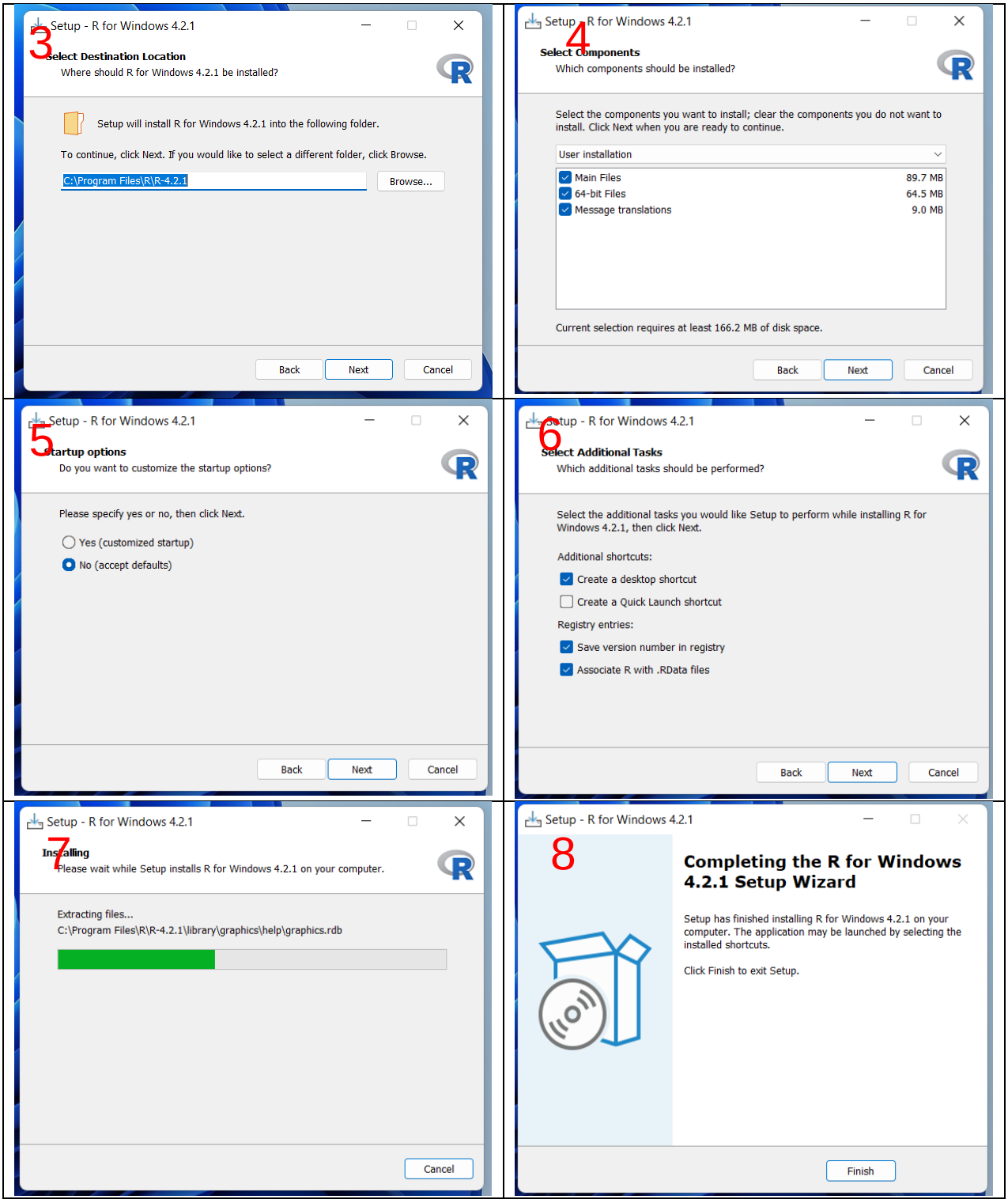
6. เมื่อดาวน์โหลดเสร็จละ ทำการติดตั้งโปรแกรม R คลิ๊ก Next ไปเรื่อยๆ จนเสร็จสิ้น (หมายเหตุ: จากรูปเป็นตัวอย่างการติดตั้งเป็น R version 4.2.1)

1

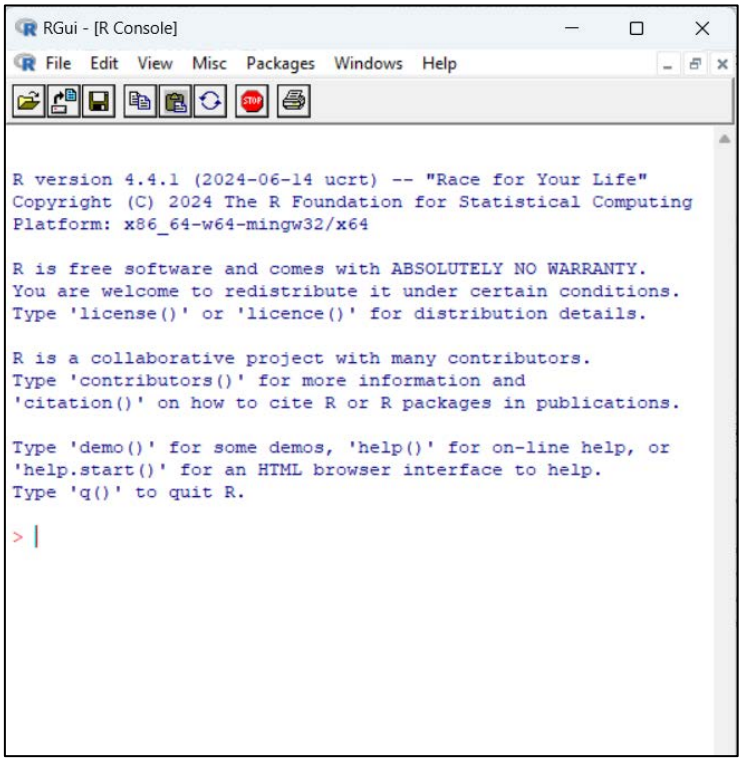


2



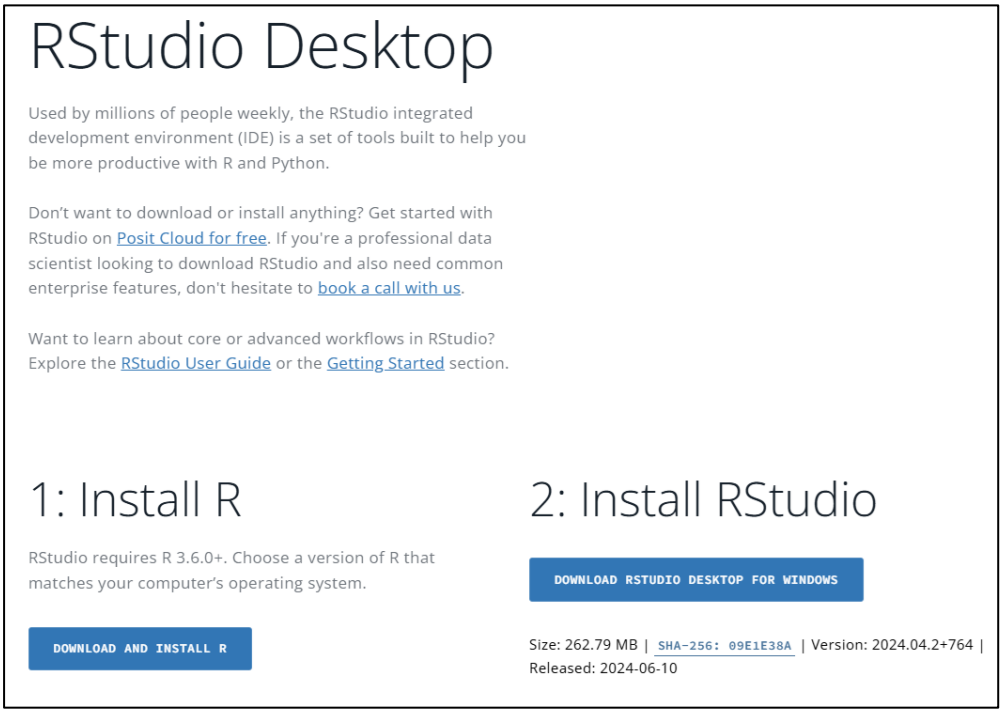


7. เมื่อติดตั้งเสร็จเรียบร้อย แล้วเปิดโปรแกรม R จะได้หน้าตาเป็นตามรูปข้างล่าง บรรทัดแรกจะบอก version ที่ใช้งาน

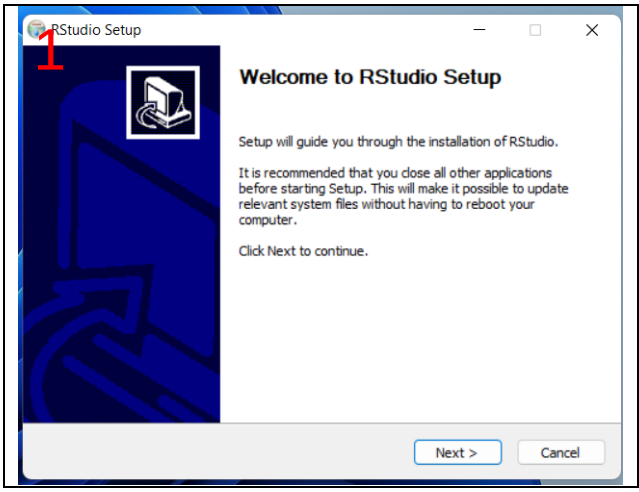

RGui - [R Console]
File Edit View Misc Packages Windows Help
R version 4.4.1 (2024-06-14 ucrt) -- "Race for Your Life"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>

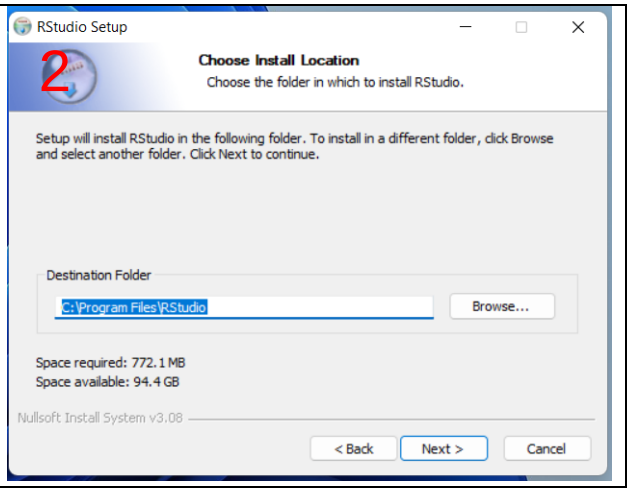
ขั้นตอนการติดตั้งโปรแกรม Rstudio

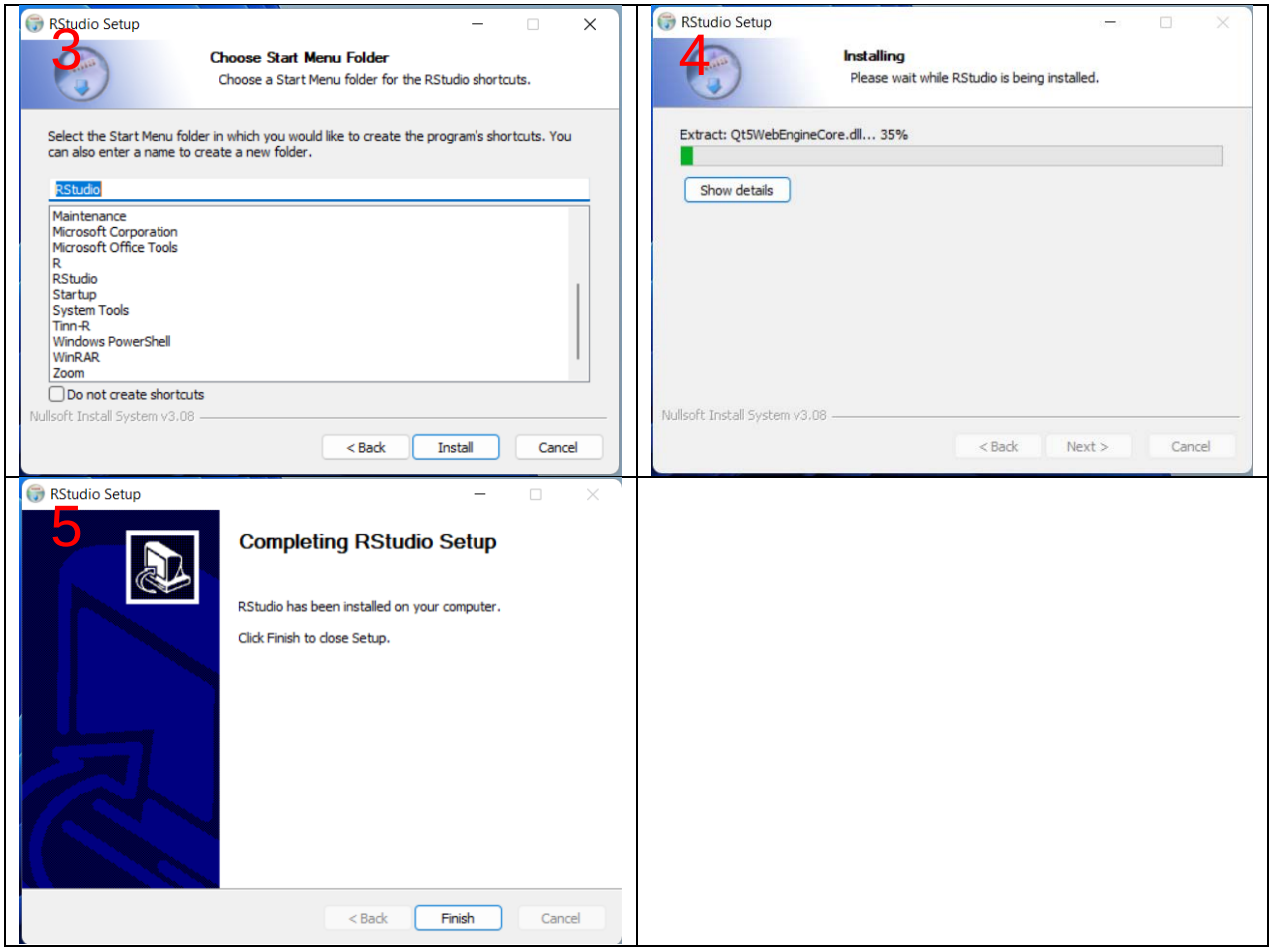
1. เข้า website https://posit.co/download/rstudio-desktop/ แล้วเลือก install RStudio



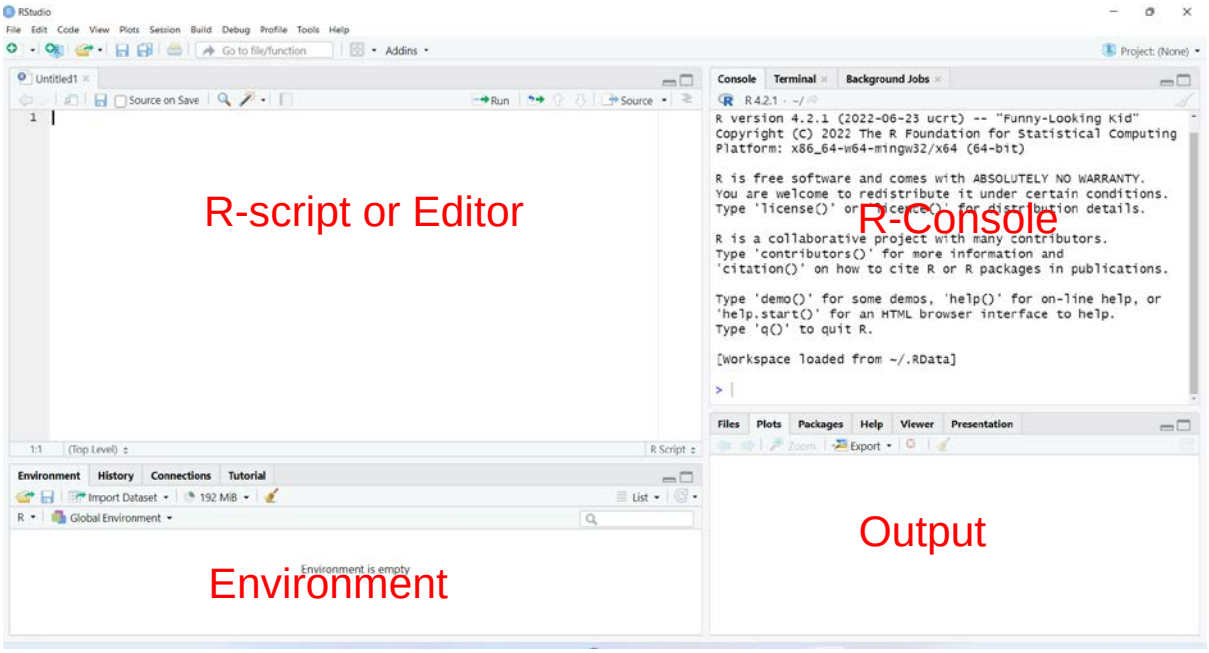
2. เมื่อดาวน์โหลดเสร็จ ทำการติดตั้งโปรแกรม Rstudio คลิ๊ก Next ไปเรื่อยๆ จนเสร็จสิ้น







3. เมื่อติดตั้งเสร็จเรียบร้อย แล้วเปิดโปรแกรม Rstudio จะได้หน้าตาเป็นตามรูปข้างล่าง



หลักการวิเคราะห์ข้อมูลและการใช้งานโปรแกรม R เบื้องต้น

วันจันทร์ที่ 1 กรกฎาคม พ.ศ. 2567

Outline

1. หลักการวิเคราะห์ข้อมูลทางสุขภาพ
2. รู้จักโปรแกรม R และ RStudio
3. การใช้งานโปรแกรม R เบื้องต้น

หลักการวิเคราะห์ข้อมูลทางสุขภาพ

หลักการวิเคราะห์ข้อมูลทางสุขภาพ

1. **การเก็บรวบรวมข้อมูล (Data Collection)**
   - การเลือกแหล่งข้อมูล: ใช้ข้อมูลจากแหล่งที่เชื่อถือได้ เช่น ข้อมูลโรงพยาบาล ข้อมูลจากการสำรวจสุขภาพ เป็นต้น
   - การเก็บข้อมูล: ใช้วิธีการที่เหมาะสมในการเก็บข้อมูล เช่น การสัมภาษณ์ แบบสอบถาม การบันทึกข้อมูลทางการแพทย์ เป็นต้น
2. **การจัดการข้อมูล (Data Management)**
   - การทำความสะอาดข้อมูล (Data Cleaning): กำจัดหรือแก้ไขข้อมูลที่ผิดพลาดหรือไม่สมบูรณ์ เช่น ข้อมูลที่ซ้ำซ้อน ข้อมูลที่หายไป
3. **การวิเคราะห์ข้อมูล (Data Analysis)**
   - การวิเคราะห์เชิงพรรณนา (Descriptive Analysis): ใช้ในการอธิบายลักษณะทั่วไปของข้อมูล เช่น ค่าเฉลี่ย ค่ามัธยฐาน การแจกแจงความถี่ เป็นต้น
   - การวิเคราะห์เชิงอนุมาน (Inferential Analysis): ใช้ในการทดสอบสมมติฐาน เช่น การทดสอบ t-test, ANOVA, การวิเคราะห์การถดถอย เป็นต้น
   - การวิเคราะห์เชิงพยากรณ์ (Predictive Analysis): ใช้ในการทำนายแนวโน้มในอนาคต เช่น การใช้โมเดลการถดถอยเชิงเส้น การใช้การเรียนรู้ของเครื่อง (Machine Learning) เป็นต้น

*ที่มา: ChatGPT

หลักการวิเคราะห์ข้อมูลทางสุขภาพ (ต่อ)

4. การแปลผลและการตีความข้อมูล (Interpretation and Presentation)

- การแปลผลข้อมูล: แปลผลการวิเคราะห์และสรุปข้อมูลที่ได้ เพื่อให้สามารถนำไปใช้ในการตัดสินใจได้
- การนำเสนอผล: ใช้เครื่องมือในการนำเสนอข้อมูล เช่น กราฟ, ตาราง, แผนภูมิ เพื่อให้ข้อมูลที่ซับซ้อนดูเข้าใจง่ายและชัดเจน

5. การประเมินผลและการใช้ข้อมูล (Evaluation and Utilization)

- การประเมินคุณภาพข้อมูลและการวิเคราะห์: ตรวจสอบความถูกต้องและความน่าเชื่อถือของผลการวิเคราะห์
- การใช้ข้อมูลเพื่อการตัดสินใจ: ใช้ข้อมูลเพื่อพัฒนานโยบายสุขภาพ การปรับปรุงบริการสุขภาพ

6. การรักษาความปลอดภัยของข้อมูล

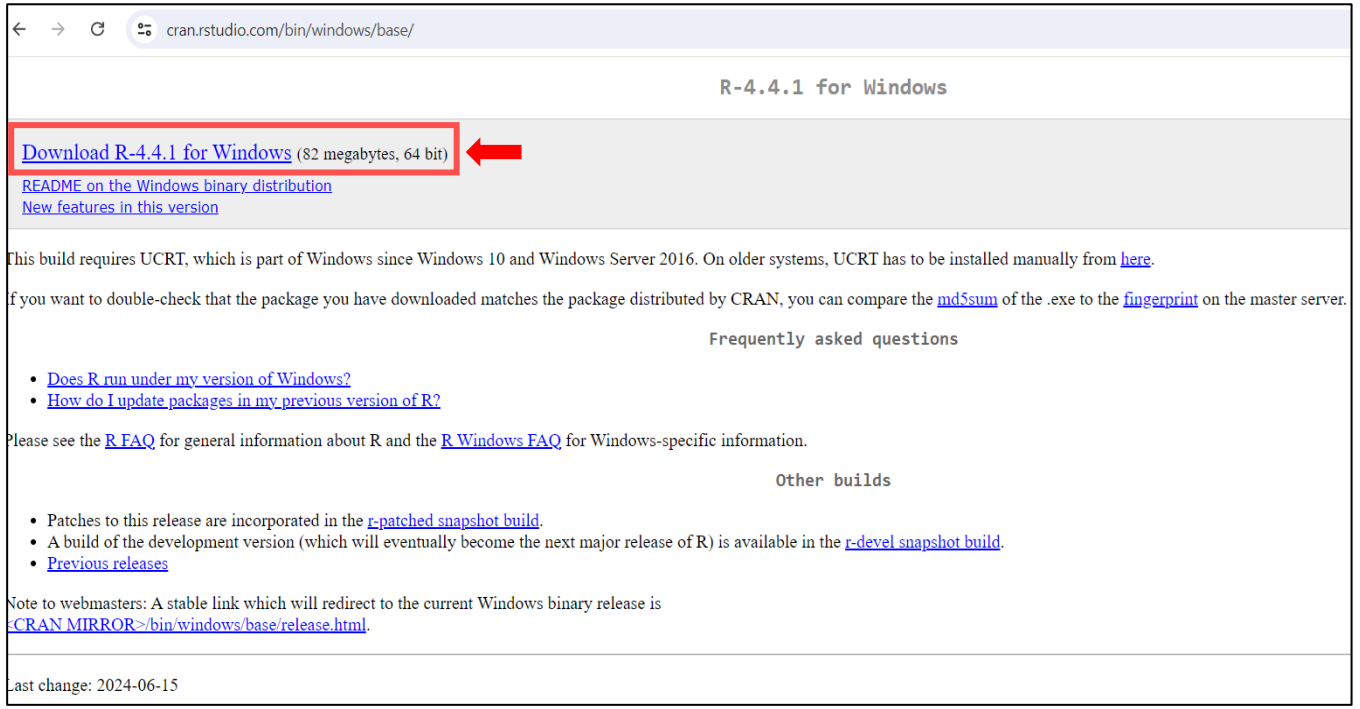
*ที่มา: ChatGPT

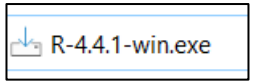
รู้จักโปรแกรม R และ RStudio

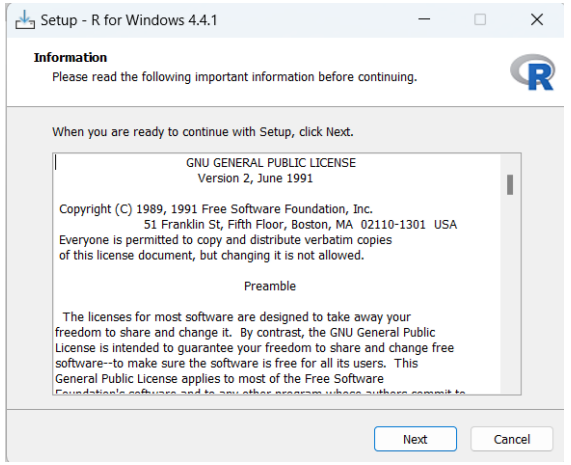
การติดตั้งโปรแกรม R

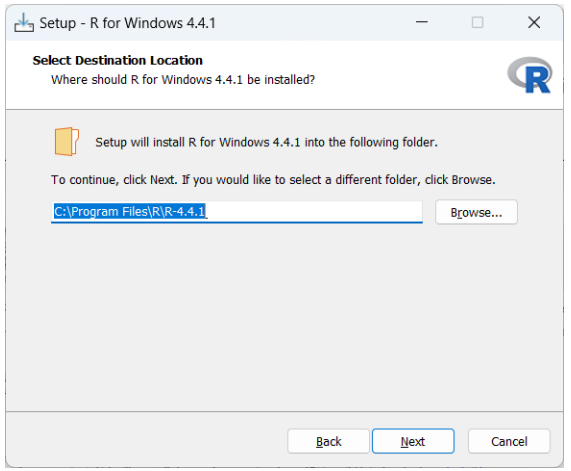
Download R: https://cran.rstudio.com/bin/windows/base/

or Download R: https://www.r-project.org/









*หมายเหตุ: โปรแกรม R มีการอัพเดทเวอร์ชั่นใหม่ค่อนข้างบ่อย แต่เราไม่จำเป็นต้องอัพเดทใหม่ตลอด ยกเว้นกรณีที่บาง package ต้องใช้กับเวอร์ชั่นใหม่ๆ

การติดตั้งโปรแกรม RStudio

Download RStudio: https://posit.co/download/rstudio-desktop/

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

Don't want to download or install anything? Get started with RStudio on Posit Cloud for free. If you're a professional data scientist looking to download RStudio and also need common enterprise features, don't hesitate to book a call with us.

Want to learn about core or advanced workflows in RStudio? Explore the RStudio User Guide or the Getting Started section.

1: Install R

RStudio requires R 3.6.0+. Choose a version of R that matches your computer's operating system.

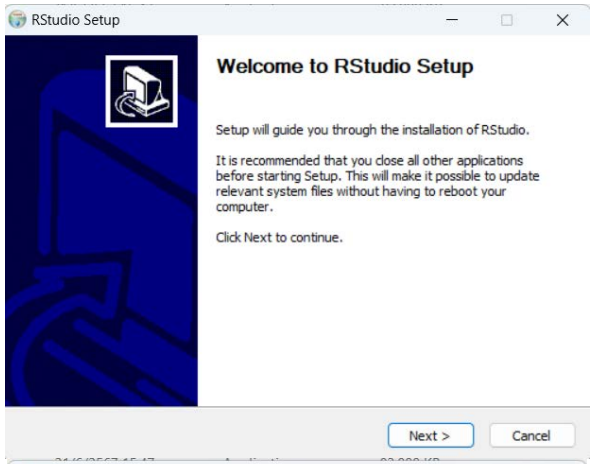
DOWNLOAD AND INSTALL R

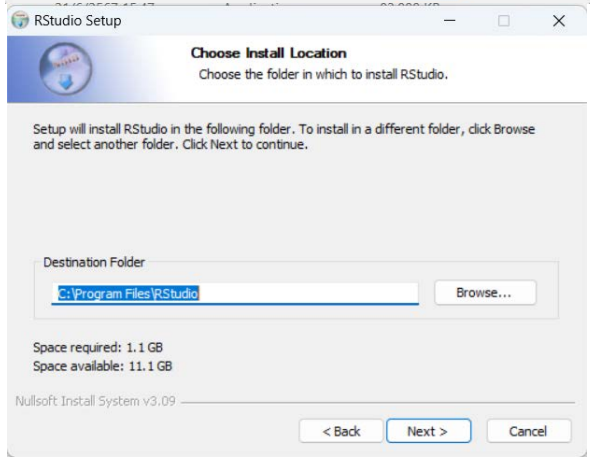
2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 262.79 MB | SHA-256: 09E1E38A | Version: 2024.04.2+764 | Released: 2024-06-10

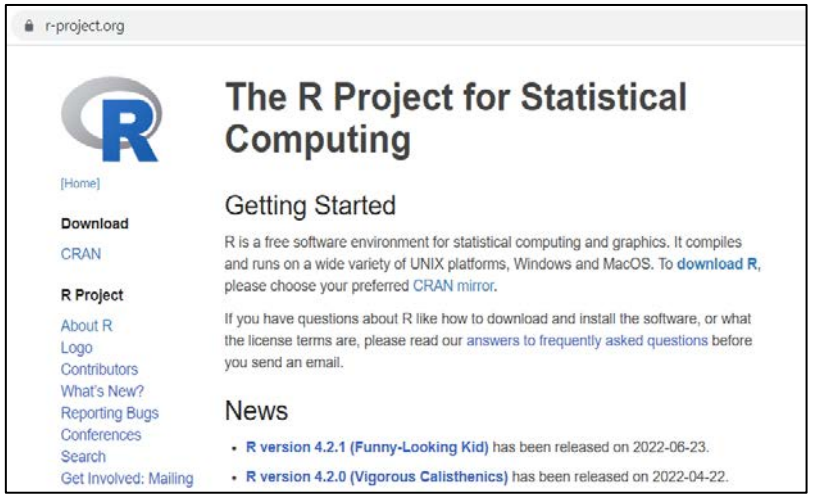
RStudio-2024.04.2-764.exe





*หมายเหตุ: การใช้งาน R-studio ต้องทำการติดตั้ง R ก่อน

โปรแกรม R คืออะไร



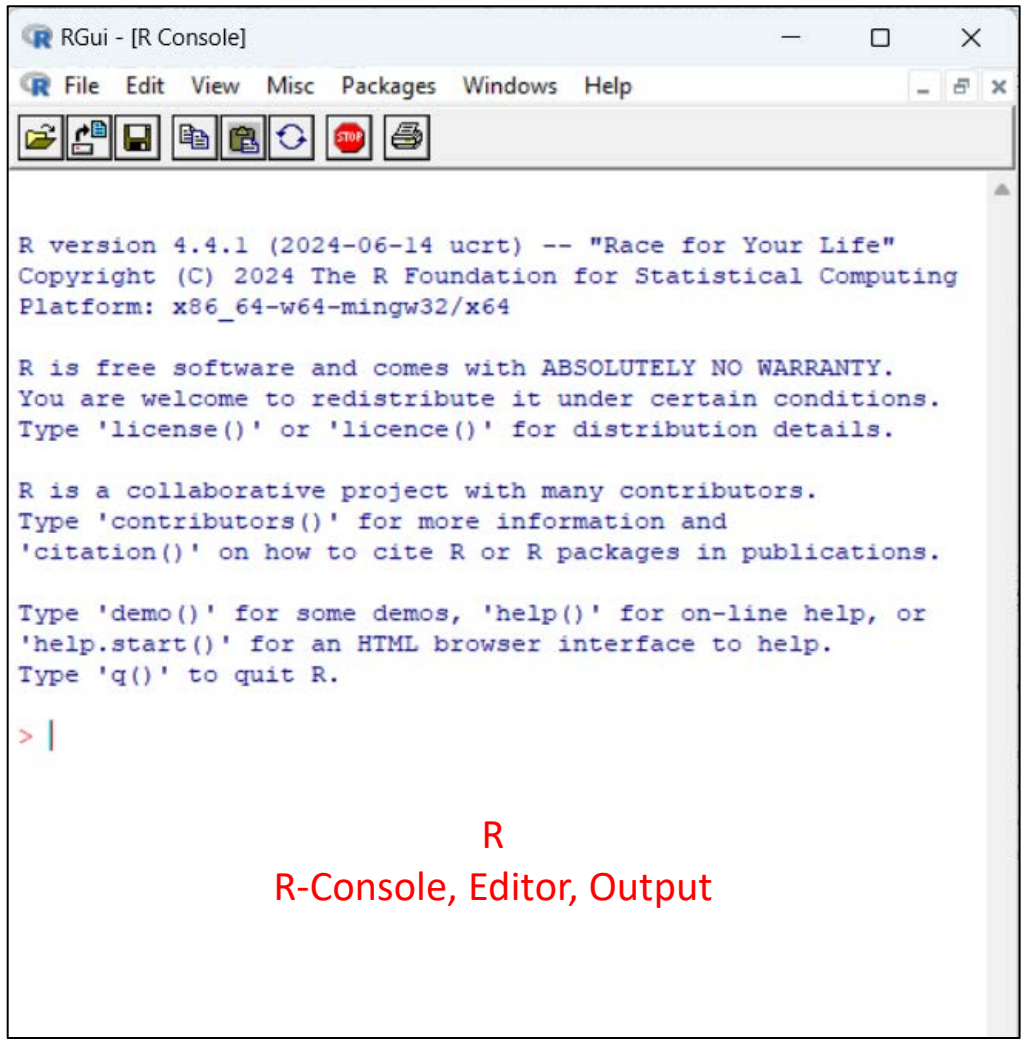
https://www.r-project.org/

https://data-flair.training/blogs/why-learn-r/

- โปรแกรม R เป็นเครื่องมือที่ใช้สำหรับการคำนวณทางสถิติและนำเสนอข้อมูลเป็นกราฟ
- วิเคราะห์ข้อมูลขนาดใหญ่ (Big data) หรืองานที่ซับซ้อนได้
- R เป็นโปรแกรมฟรี ไม่มีค่าใช้จ่าย จึงเป็นที่นิยมในวงวิชาการมากขึ้นเรื่อยๆ
- R เป็นโปรแกรมประเภท open source ทำให้มีการร่วมพัฒนาต่อยอดให้มีความสามารถที่หลากหลายต่อเนื่อง
- ปัจจุบัน R มี package กว่า 10,000 package

ที่มา: การวิเคราะห์ข้อมูลสุขภาพในระบบสาธารณสุขไทยโดยใช้โปรแกรม R

R



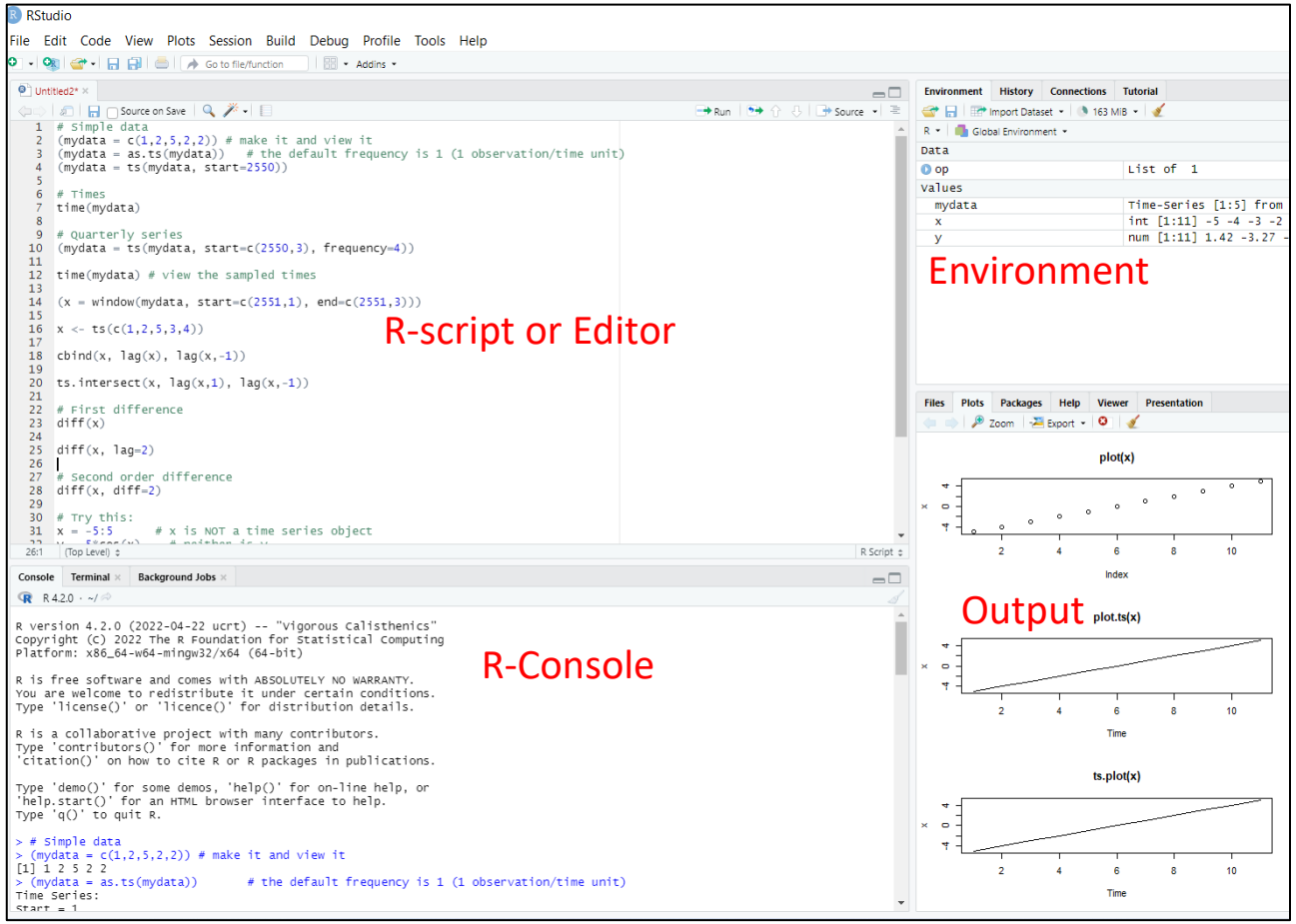
R

R-Console, Editor, Output

Rstudio

Rstudio เป็นโปรแกรมหนึ่งที่ช่วยให้ผู้ใช้งานมีความเข้าใจการทำงานของโปรแกรม R ได้ดีขึ้น และทำให้การทำงานด้วยโปรแกรม R สะดวกขึ้น

*การติดตั้งโปรแกรม Rstudio จะต้องติดตั้งโปรแกรม R ก่อนเสมอ



การใช้งานโปรแกรม R เบื้องต้น

Data Types in R

Numeric	Contains decimal as well as whole numbers
Integer	Contains only whole numbers
Character	Holds character strings
Factor	A vector that can contain only predefined values, and is used to store categorical data.
Logical	Can only take on two values, TRUE or FALSE.

Vector	1 dimension	Sequence of data elements of the same basic type.
Matrix	2 dimensions	Like a Vector but additionally contains the dimension attribute.
Array	2 or more dimensions	Hold multidimensional data. Matrices are a special case of two-dimensional arrays.
Data frame	2 dimensions	Table-like data object allowing different data types for different columns.
List	Collection of data objects, each element of a list is a data object.	

INTEGER:	If the data consists only of whole numbers, integers too may take negative or positive values.
CHARACTER:	If data consists of strings, i.e., words or sentences.
FACTOR:	A vector used to store categorical data which contain only predefined values. They can store both strings and integers.
LOGICAL:	The type of data which can only assume two values, namely, true and false

Source: https://digitaschools.com/data-types-in-r-tutorial/

Vector, Matrix, Data frame, Array และ lists

```
#Vector
age <- c(10,20,30,40,50)
age
   [1] 10 20 30 40 50
sex <- c("male", "male", "female", "female", "male")
sex
   [1] "male"   "male"   "female" "female" "male"
#Matrix
person <- cbind(age,sex)
person
      age  sex
   [1,] "10" "male"
   [2,] "20" "male"
   [3,] "30" "female"
   [4,] "40" "female"
   [5,] "50" "male"
#Data frame
dat <- data.frame(person)
dat
   age    sex
1  10   male
2  20   male
3  30 female
4  40 female
5  50   male
```

Arrays

In an Array, each row is of the same length and each column is also of the same length. So, we say it holds multidimensional rectangular data.

```
a<-array(1:24,dim=c(3,4,2))
a
```

```
, , 1

     [,1] [,2] [,3] [,4]
[1,]    1    4    7   10
[2,]    2    5    8   11
[3,]    3    6    9   12

, , 2

     [,1] [,2] [,3] [,4]
[1,]   13   16   19   22
[2,]   14   17   20   23
[3,]   15   18   21   24
```

Lists

A data structure containing mixed data types is called a List. We can convert any object into a list by using the function **as.list()**. We can also create lists using the function **list()**.

```
n=c(2, 3, 5)
s=c("aa", "bb", "cc", "dd", "ee")
x=list(n, s, 3)
x
```

```
[[1]]
[1] 2 3 5

[[2]]
[1] "aa" "bb" "cc" "dd" "ee"

[[3]]
[1] 3
```

การคำนวณใน R: สามารถใช้คำนวณทางคณิตศาสตร์ได้

```
> 5+3
[1] 8
>
> 5*3
[1] 15
>
> 5/3
[1] 1.666667
>
> 5^3
[1] 125
>
> exp(5)
[1] 148.4132
>
> log(5)
[1] 1.609438
>
> sin(5)
[1] -0.9589243
>
> cos(5)
[1] 0.2836622
>
```

สัญลักษณ์	ความหมาย
==	การเท่ากัน
!=	การไม่เท่ากัน
<	น้อยกว่า
>	มากกว่า
<=	น้อยกว่าหรือเท่ากับ
>=	มากกว่าหรือเท่ากับ
&	และ
|	หรือ

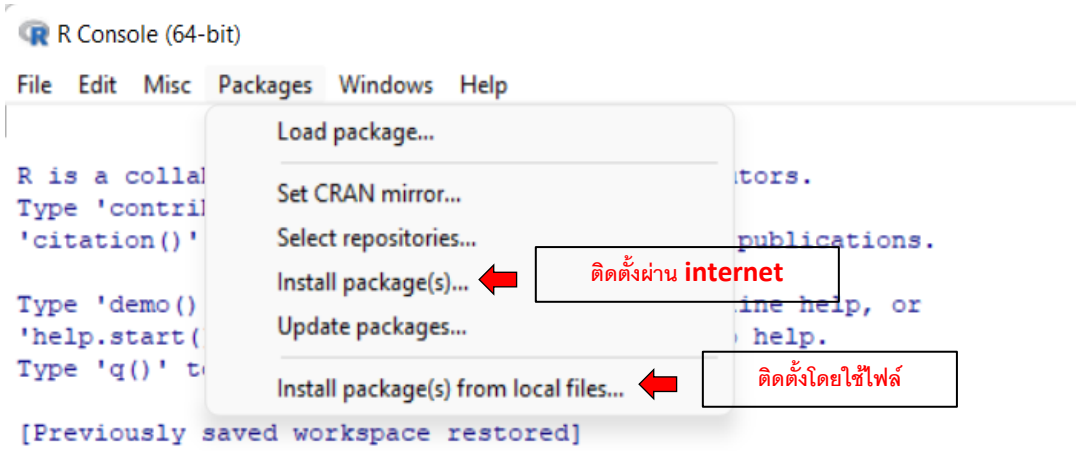
<- หรือ -> หรือ = คือการตั้งค่าให้ตัวแปร

ที่มา: การวิเคราะห์ข้อมูลสุขภาพในระบบสาธารณสุขไทยโดยใช้โปรแกรม R

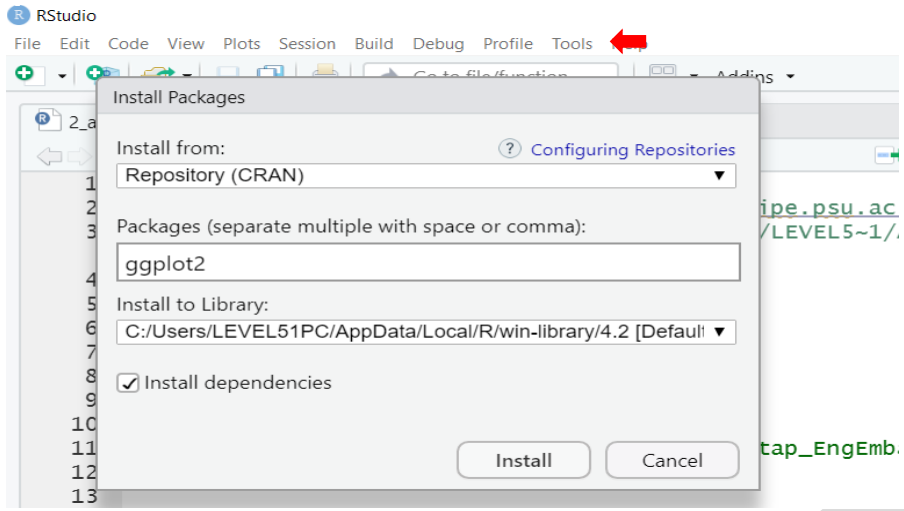
การใช้ package ใน R และ RStudio

package เป็นชุดคำสั่งที่ช่วยในการจัดการข้อมูล วิเคราะห์ข้อมูล สามารถติดตั้งและทำการโหลด package และเรียกใช้ library() โดยติดตั้งเพียงครั้งเดียว หลังจากนั้นใช้ได้ตลอด การติดตั้งด้วยโปรแกรม R ทำได้ 2 วิธี คือผ่าน internet และติดตั้งโดยใช้ไฟล์

การติดตั้ง Package ด้วย R: Packages เลือก install packages และเลือก package ที่เราต้องการติดตั้ง



การติดตั้ง Package ด้วย Rstudio: Tools เลือก install packages เลือก Reposity (CRAN) และพิมพ์ชื่อ package ตัวอย่างนี้เราติดตั้ง ggplot2



การเรียกดูรายละเอียด package

ใช้คำสั่ง help(package="……") เช่น help(package="epiDisplay")

Documentation for package 'epiDisplay' version 3.5.0.2

- DESCRIPTION file.
- Code demos. Use demo() to run them.

Help Pages

A B C D E F G H I K L M N O P R S T V X

-- A --

aggregate.numeric	Summary statistics of a numeric variable by group
aggregate.plot	Plot summary statistics of a numeric variable by group
alpha	Cronbach's alpha
alphaBest	Cronbach's alpha
ANCdata	Dataset on effect of new antenatal care method on mortality
ANCtable	Dataset on effect of new ANC method on mortality (as a table)
Attitudes	Dataset from an attitude survey among hospital staff

-- B --

Bang	Dataset from 1988 Bangladesh Fertility Survey
BP	Dataset on blood pressure and determinants

-- C --

cc	Odds ratio calculation and graphing
cci	Odds ratio calculation and graphing
ci	Confidence interval of probabilty, mean and incidence
ci.binomial	Confidence interval of probabilty, mean and incidence
ci.default	Confidence interval of probabilty, mean and incidence
ci.numeric	Confidence interval of probabilty, mean and incidence
ci.poisson	Confidence interval of probabilty, mean and incidence
clogistic.display	Tables for multivariate odds ratio, incidence density etc
codebook	Codebook of a data frame

aggregate plot {epiDisplay} R Documentation

Plot summary statistics of a numeric variable by group

Description

Split a numeric variable into subsets, plot summary statistics for each

Usage

```
## S3 method for class 'plot'
aggregate(x, by, grouping = NULL, FUN = c("mean", "median"),
    error = c("se", "ci", "sd", "none"), alpha = 0.05, lwd = 1,
    lty = "auto", line.col = "auto", bin.time = 4, bin.method = c("fixed",
        "quantile"), legend = "auto", legend.site = "topright",
    legend.bg = "white", xlim = "auto", ylim = "auto", bar.col = "auto",
    cap.size = 0.02, lagging = 0.007, main = "auto", return.output = FALSE, ...)
```

Details

This function plots aggregated values of 'x' by a factor (barplot) or a continuous variable (time line graph).

When 'by' is of class 'factor', a bar plot with error bars is displayed.

When 'by' is a continuous variable (typically implying time), a time line graph is displayed.

Both types of plots have error arguments. Choices are 'se' and 'sd' for the bar plot and 'ci' and IQR for both bar plot and time line graph. All these can be suppressed by specifying 'error'="none".

'bin.time' and 'bin.method' are exclusively used when 'by' is a continuous variable and does not have regular values (minimum frequency of 'by' <3). This condition is automatically and silently detected by 'aggregate.plot' before 'bin.method' chooses the method for aggregation and bin.time determines the number of bins.

If 'legend = TRUE" (by default), a legend box is automatically drawn on the "topright" corner of the graph. This character string can be changed to others such as, "topleft", "center", etc (see examples).

'cap.size' can be assigned to zero to remove the error bar cap.

Author(s)

Virasakdi Chongsuvivatwong cvirasak@gmail.com

See Also

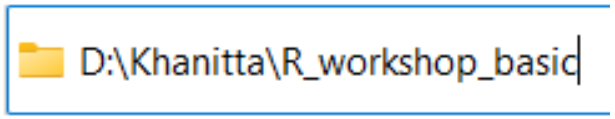
'aggregate.data.frame', 'aggregate.numeric', 'tapply'

การนำเข้าข้อมูล

1. **คำสั่งพื้นฐานก่อนนำเข้าข้อมูลสู่ R**

   คำสั่ง setwd ใช้สำหรับระบุพื้นที่ทำงาน Set Working Directory

   คำสั่ง getwd ใช้สำหรับตรวจสอบพื้นที่ทำงาน Get Working Directory

   ตัวอย่างเช่น 


```
> setwd("D:/Khanitta/R_workshop_basic")
> getwd()
[1] "D:/Khanitta/R_workshop_basic"
```

```
> setwd("D:\\Khanitta\\R_workshop_basic")
> getwd()
[1] "D:/Khanitta/R_workshop_basic"
>
```

2. **นำเข้าข้อมูล** ต้องพิจารณานามสกุล **.xxx** ว่าเป็นไฟล์ลักษณะใด แล้วเลือกนำเข้าให้เหมาะในแต่ละประเภท เช่น .txt หรือ .csv หรือ .dbf หรือ .dta

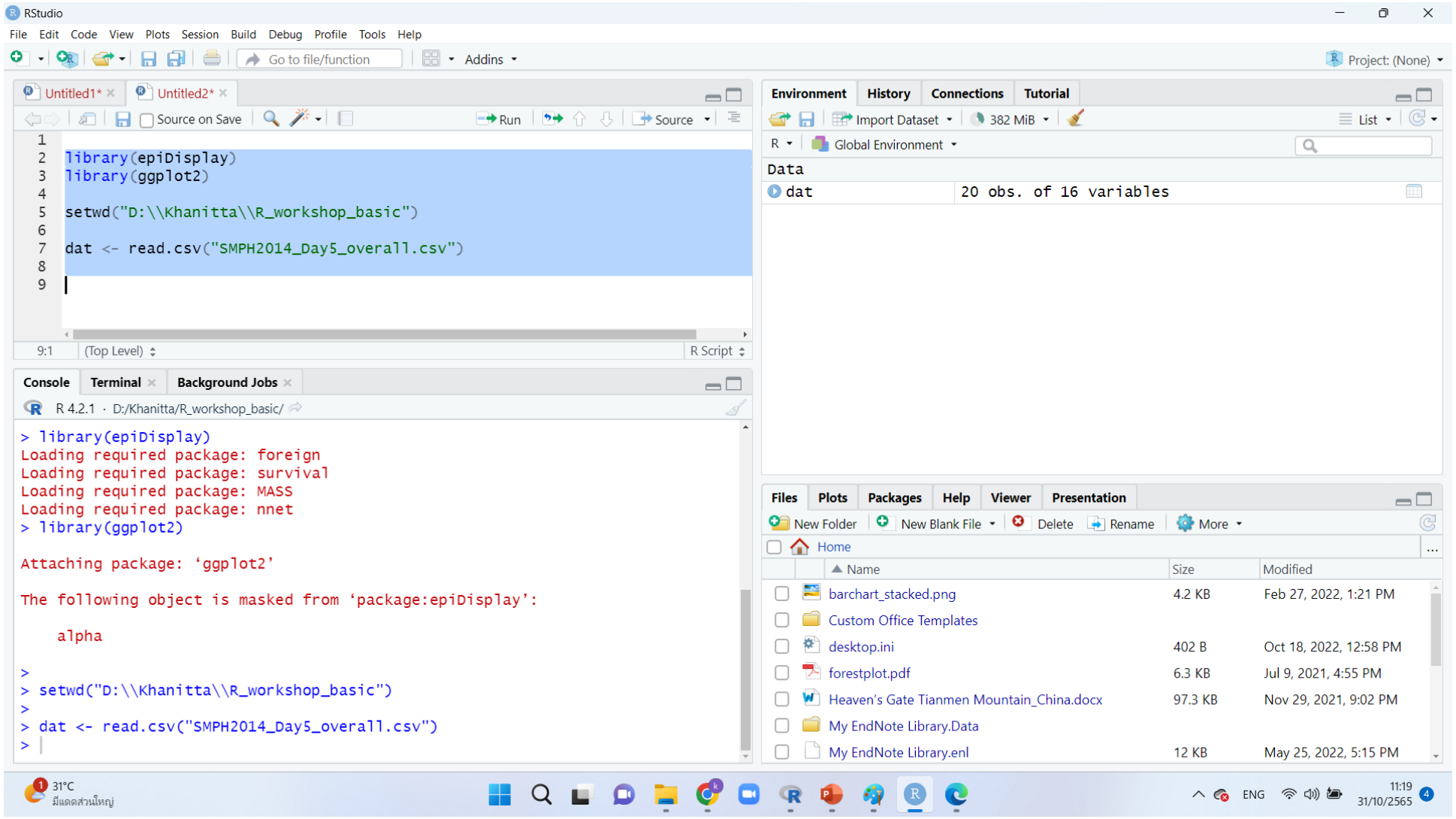
   dat <- read.dta("personnlab.dta")

   dat <- read.csv("personnlab.csv")

   dat <- read.table ("personnlab.txt" , sep=",")

```
library(openxlsx)
dat_prev_f <- read.xlsx("D:\\Khanitta\
\Risk_factor62\\smoking\\data.xlsx",sheet = 1)
```

ตัวอย่างการเริ่มวิเคราะห์ Rstudio



เปิด Rstudio
แล้วมาเริ่มวิเคราะห์ข้อมูลกัน!!!

ชุดข้อมูลที่ทำการวิเคราะห์ประกอบด้วย

1. **ตัวอย่างข้อมูลการตาย**
2. **ตัวอย่างข้อมูลผู้ป่วยใน**

BOD
Burden of Disease Thailand
www.bodthai.net

iHPP
Thailand

FOLLOW US

www.ihppthaigov.net

ihpp.thailand

ihpp_thailand@ihpp.thaigov.net

IhppThailand

ihppchannel

ภาคผนวกที่ 2

แนวคิดพื้นฐานด้านการพยากรณ์โรค

แนวคิดพื้นฐานด้านการพยากรณ์โรค (BASIC CONCEPTS IN DISEASE FORECASTING)

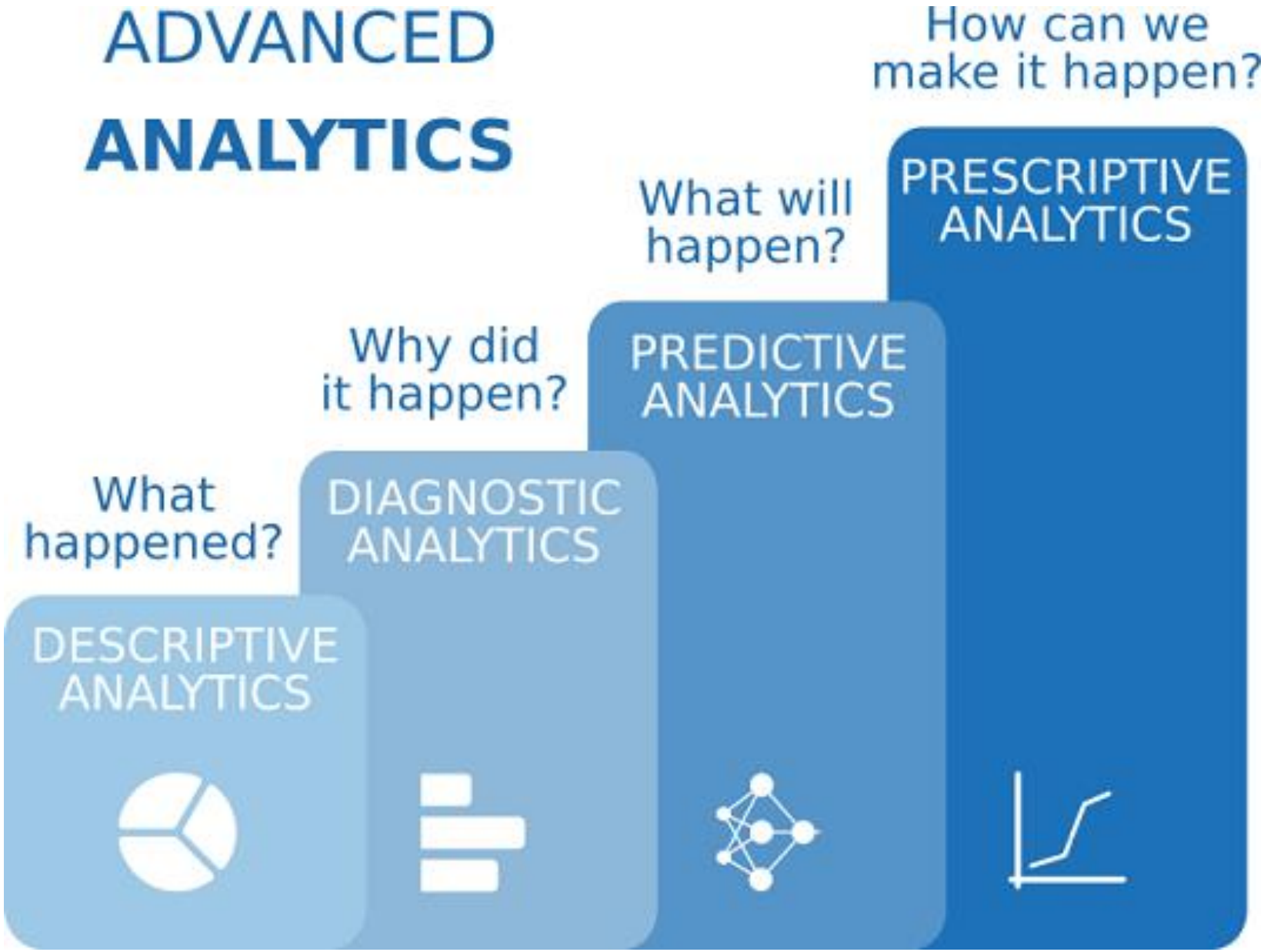
กองระบาดวิทยา กรมควบคุมโรค
มูลนิธิเพื่อการพัฒนานโยบายสุขภาพ ระหว่างประเทศ

2567

Aims of Public Health Research

- DESCRIBE
  - มีผู้ป่วยด้วยโรคปอดอุดกันเรื้อรังมากน้อยเพียงใดในจังหวัดแห่งหนึ่ง
  - มีการสูบบุหรี่เป็นสัดส่วนเท่าไรในประชากรผู้หญิงและผู้ชาย
- EXPLAIN
  - ทำไมผู้ชายจึงสูบบุหรี่มากกว่าผู้หญิง
  - การสูบบุหรี่มือสองเพิ่มความเสี่ยงในการเป็นโรคมะเร็งกระเพาะอาหารหรือไม่
- PREDICT
  - อัตราการเกิดขึ้นของนักสูบบุหรี่หน้าใหม่ในเด็กมัธยมปลายปีหน้าจะเป็นเท่าไร
  - ถ้าสามารถรณรงค์ให้คนในชุมชนเลิกสูบบุหรี่ได้เป็นผลสำเร็จ จำนวนผู้ป่วยโรคหัวใจหลอดเลือดรายใหม่ในปีหน้าจะลดลงเป็นจำนวนเท่าไร
- CONTROL
  - มาตรการที่เหมาะสมสำหรับชุมชน (ภายใต้ข้อจำกัดต่างๆ) คืออะไร

Source: Morgenstern, 2001 (modified)

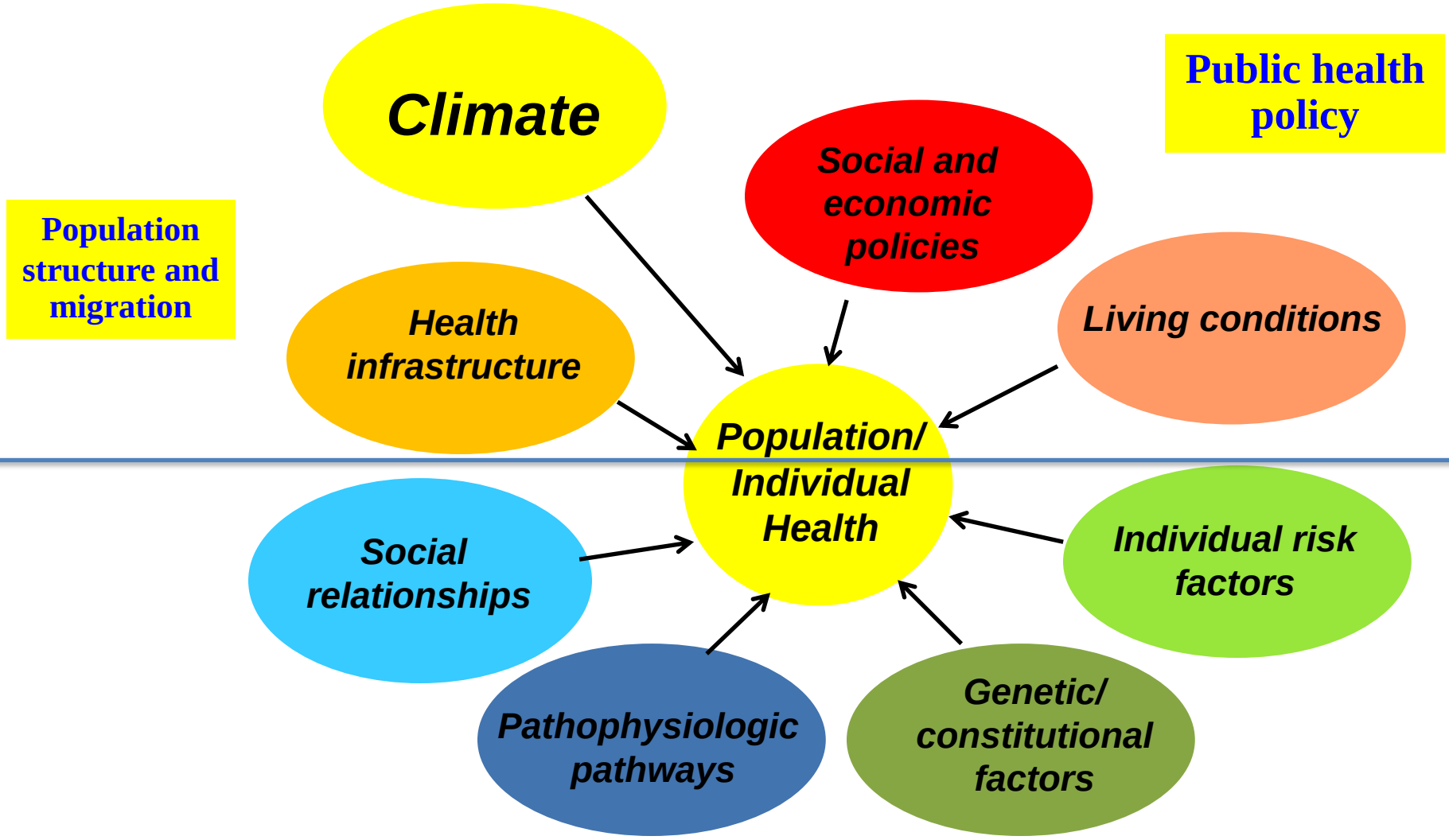


Source: https://www.neuraldesigner.com/blog/what_is_advanced_analytics

Disease has a pattern

- การเกิดโรคมีรูปแบบ (Pattern) ที่อธิบายได้ ตามปัจจัยตัวกำหนด (Determinants)
- การเกิดโรคในชุมชนไม่ใช่กระบวนการสุ่มตัวอย่าง (Random process) แต่จะเกิดมากหรือน้อย หรือไม่เกิดขึ้นเลยในคนบางกลุ่ม บางพื้นที่ บางเวลา
- ต้องใช้ความรู้ทางระบาดวิทยา เพื่ออธิบายปัจจัยกำหนดต่างๆ ที่ทำให้บางคนหรือบางกลุ่มมีโอกาสป่วยมากกว่าปกติ

Determinants of Human Disease



พยากรณ์ หมายถึงอะไร

- **ค่าพยากรณ์ (Forecast)** = a statement of what is judged likely to happen in the future, especially in connection with a particular situation, or the expected weather condition*
- **การพยากรณ์ (Forecasting)** = Process for getting the forecast
- **นักพยากรณ์ (Forecaster)** = One who do forecasting

*Source: Cambridge Advanced Learner's Dictionary

แนวคิดที่เกี่ยวข้องกับการพยากรณ์

- การวางแผนต่างๆล้วนอาศัยการพยากรณ์
- โดยปกติแล้วค่าที่ได้จากการพยากรณ์มักจะมีความคลาดเคลื่อนเกิดขึ้นเสมอ ไม่ได้ถูกต้องแม่นยำอย่างสมบูรณ์แบบ
- การพยากรณ์ที่ดีต้องอาศัย
  - ความรู้ทางวิชาการ (ทั้งในเชิงเนื้อหาและเชิงเทคนิคการพยากรณ์)
  - ประสบการณ์
  - วิจารณญาณ

Source: Ozcan, 2009

Forecasting Assumptions (1)

- The underlying assumption is that past events will continue
- Errors will occur because of the presence of randomness and that actual results are more than likely to be different from those predicted.
- Forecasts of a group of items (aggregate forecasts) tend to be more accurate than those for individual items.
  - For example, forecasts made for a whole province would tend to be more accurate than a district or sub-district forecast because forecasting errors among a group tend to cancel each other

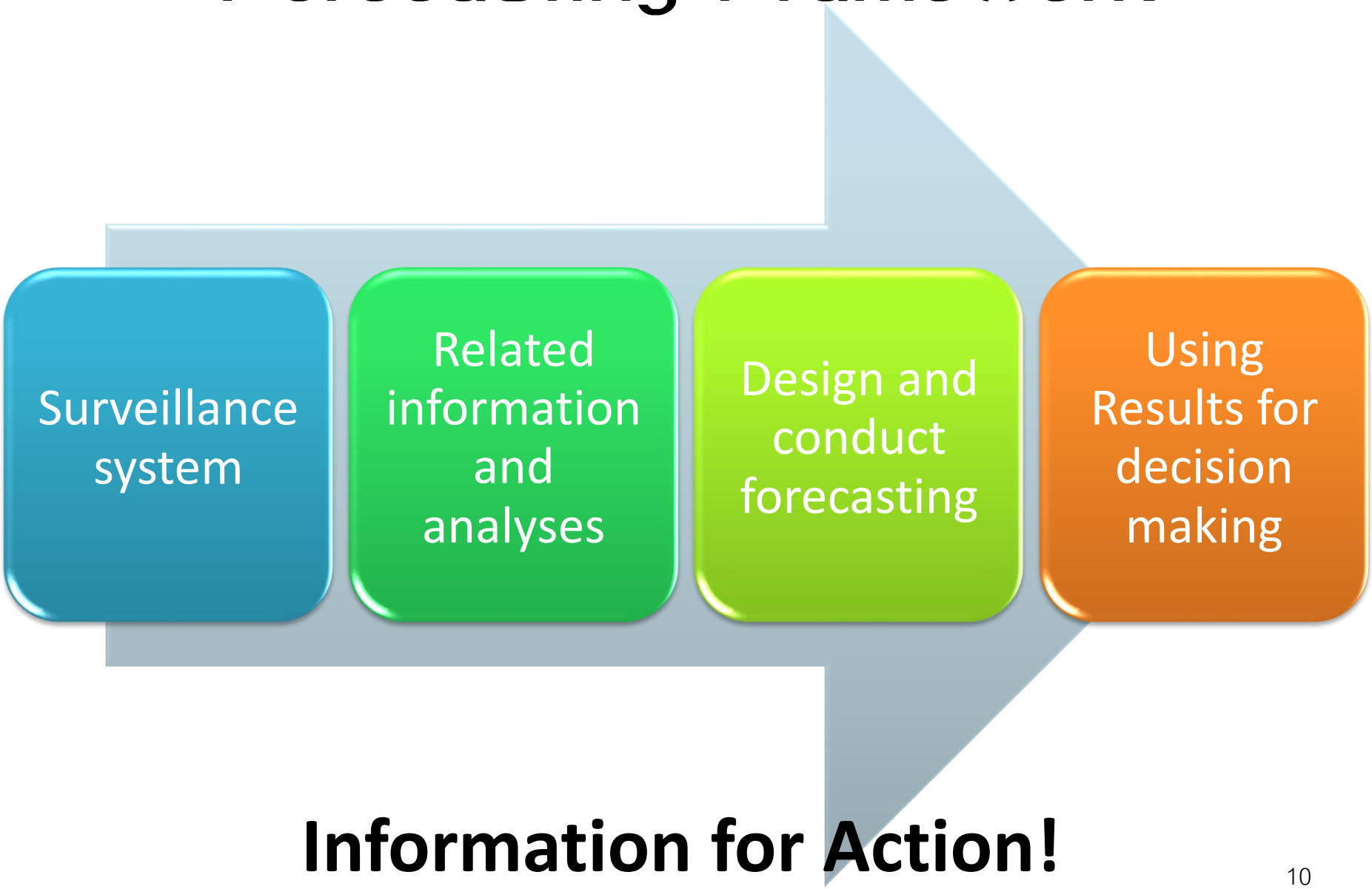
Source: Ozcan, 2009

Forecasting Assumptions (2)

- It is generally accepted that forecast accuracy decreases as the time horizon (the period covered) increases.
  - Short - range forecasts face fewer uncertainties than longer – range forecasts do, so they tend to be more accurate.
- Choosing forecasting horizon depends on goals of organization
  - A flexible health care organization, which responds quickly to changes in demand, makes use of a shorter, more accurate forecasting horizon than do less flexible competitors, who must use longer forecast horizons.

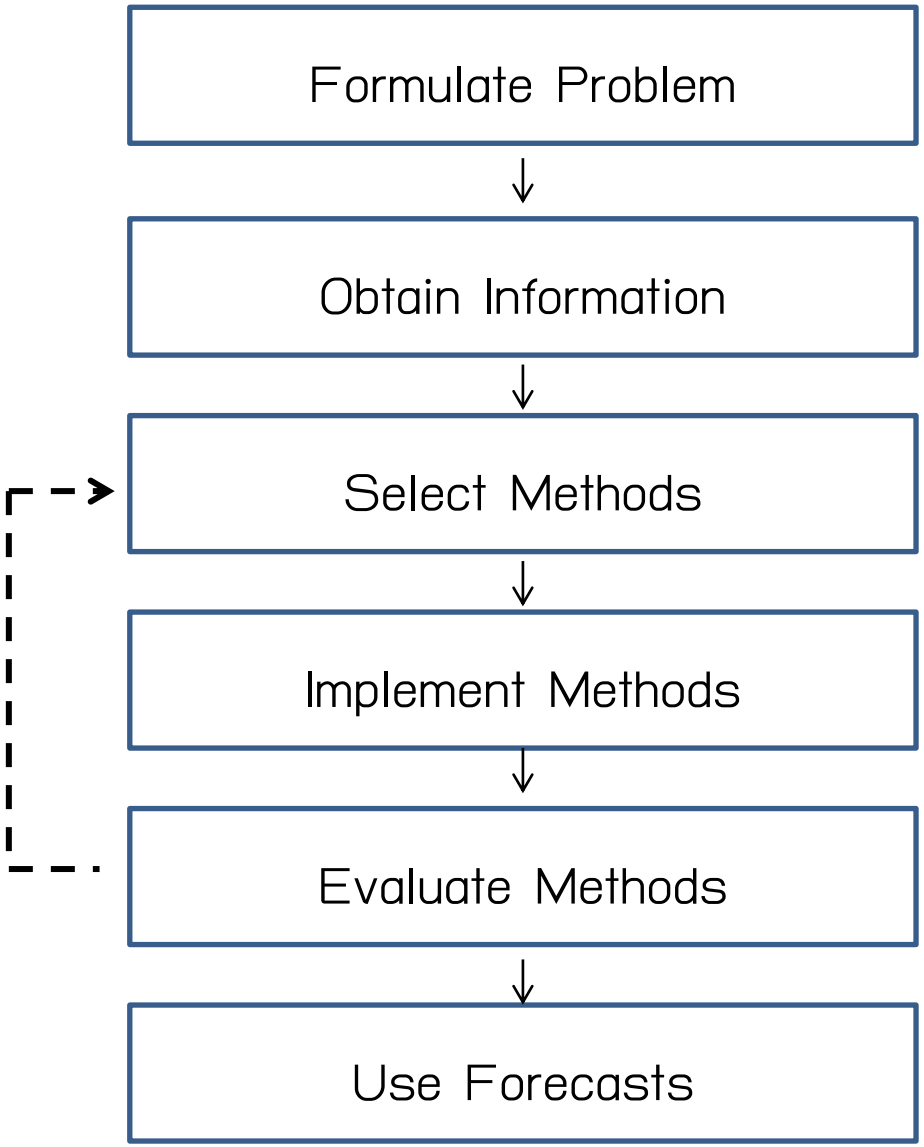
Source: Ozcan, 2009

Forecasting Framework



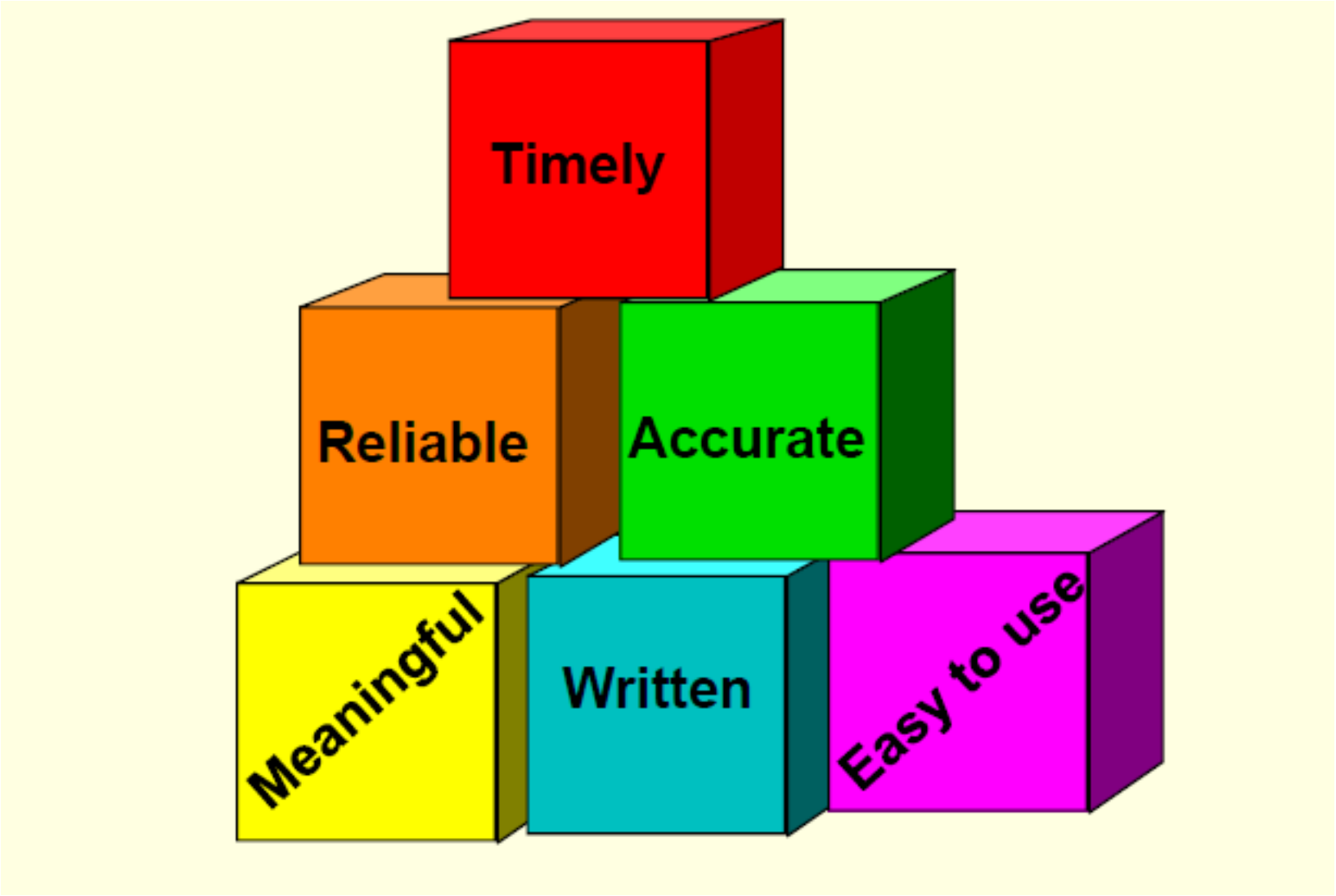
Information for Action!

ขั้นตอนการพยากรณ์



"Forecasting Modelling"

ลักษณะของการพยากรณ์ที่ดี



ปัจจัยที่ต้องนำมาพิจารณา ในการเลือกโรคหรือภัยสุขภาพที่จะนำมาพยากรณ์

Disease Importance

- Frequency
- Severity
- Impact (QOL, Social, economic)
- Public/political concern

Modifiability

- Preventable disease
- Available effective intervention
- Acceptable cost

Predictability

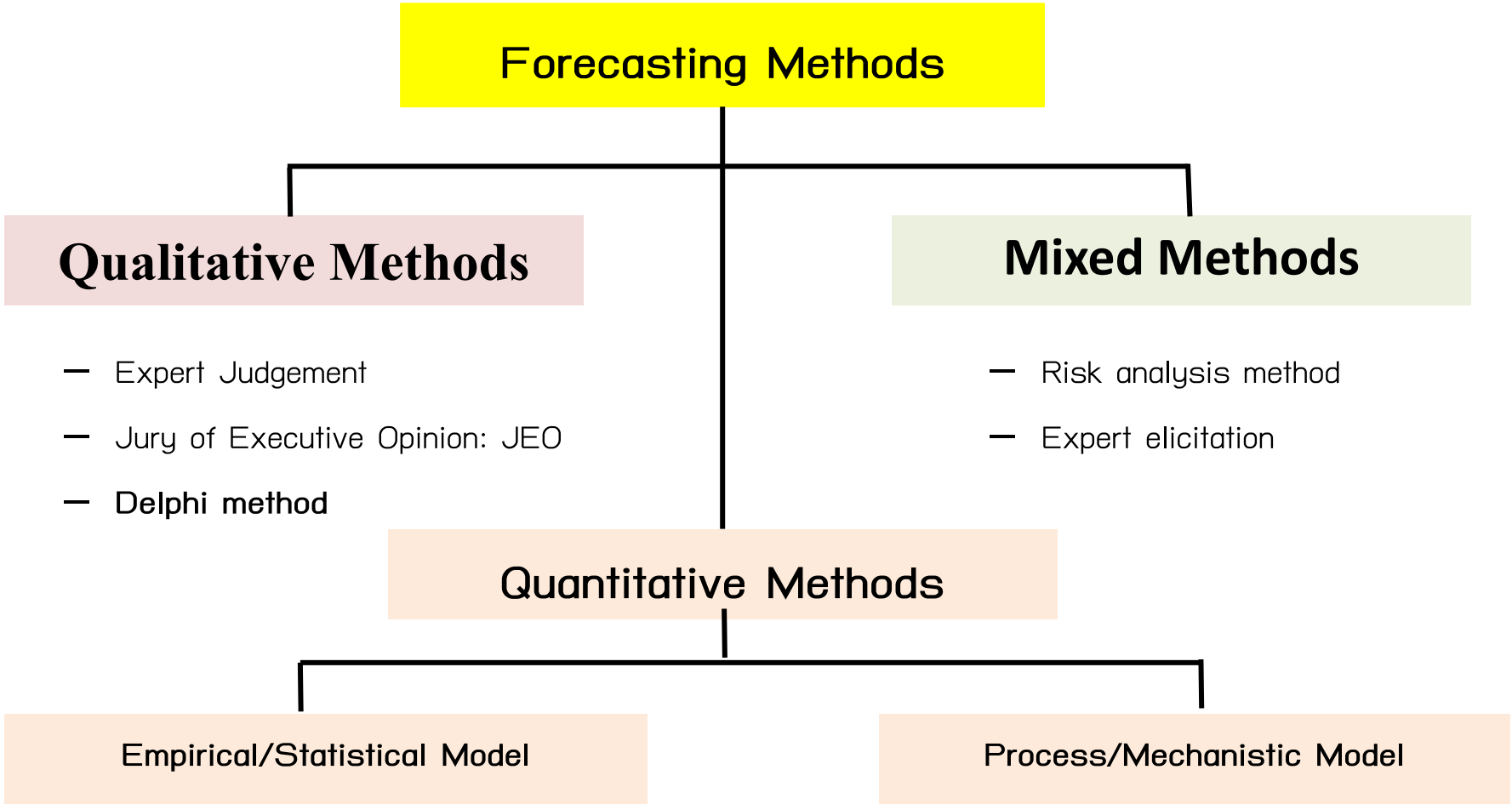
- Available data
- Available technique
- Available technology
- Available experts and experiences (both content and technique)

ปัจจัยที่ต้องนำมาพิจารณา ในการเลือกเทคนิควิธีการพยากรณ์ (1)

- **Forecast form required:** ต้องการค่าพยากรณ์ออกมาเป็นตัวเลขผู้ป่วยหรือเป็นโอกาสที่จะเกิด?
- **Horizon time frame:** พยากรณ์ไปข้างหน้าระยะสั้นหรือยาว?
- **Pattern of data:** ข้อมูลในอดีตมีรูปแบบที่ชัดเจนหรือไม่?
- **Cost ant time available for conducting forecasting:** ค่าใช้จ่ายและเวลาที่ให้สำหรับการดำเนินการพยากรณ์มีมากน้อยเพียงใด?

ปัจจัยที่ต้องนำมาพิจารณา ในการเลือกเทคนิควิธีการพยากรณ์ (2)

- **Accuracy desired:** ต้องการความแม่นยำมากน้อยเพียงใด?
- **Availability of data:** มีข้อมูลอะไรบ้างที่จะนำมาใช้ได้?
- **Ease of operation and understanding:** ความง่ายในการดำเนินการและการอธิบายผลให้คนอื่นเข้าใจ


Forecasting Methods
Qualitative Methods
Expert Judgement
Jury of Executive Opinion: JEO
Delphi method
Mixed Methods
Risk analysis method
Expert elicitation
Quantitative Methods
Empirical/Statistical Model
Simple projection/multiplication
Regression-based model: Linear, Poisson, Age-Period-Cohort Model, etc.
Time series analysis: Decomposition, Smoothing, ARIMA
Machine Learning, Deep Learning
Process/Mechanistic Model
Compartmental (multi-stage) model e.g. SIR model, System dynamics model, Markov model, etc.
Social Network Model

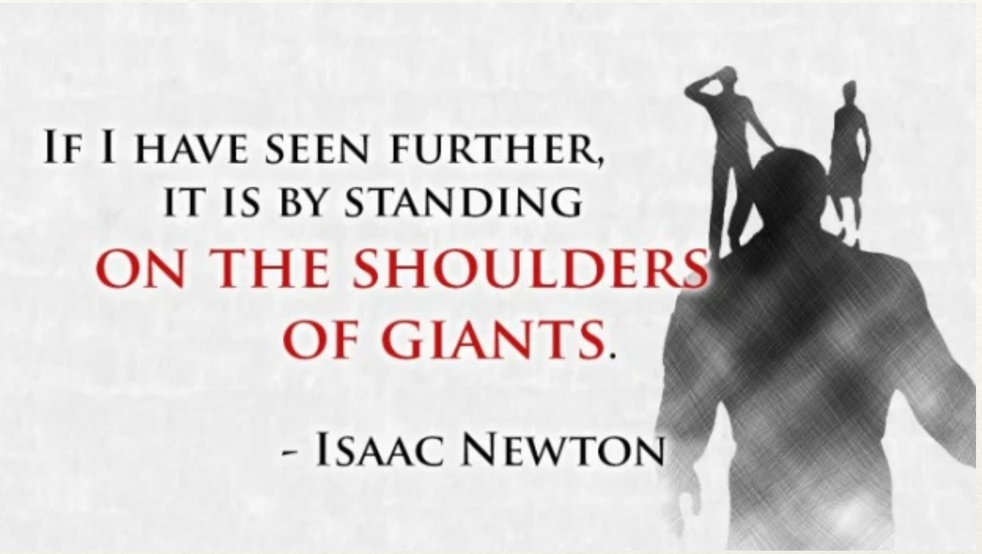
ข้อควรระวังในการพยากรณ์โรค

- การพยากรณ์ด้วยการวิเคราะห์ข้อมูลอนุกรมเวลา อาศัยตัวข้อมูลในอดีตมาคาดการณ์สิ่งที่จะเกิดขึ้นในอนาคต โดยไม่ได้อธิบายว่ามีอะไรเป็นสาเหตุ แต่ตั้งสมมติฐานว่า
  - ข้อมูลในอดีตได้ดูดซับสาเหตุเหล่านั้นแล้ว
  - ในอนาคต เหตุปัจจัยต่างๆจะยังคงเกิดขึ้นและดำเนินไปเหมือนที่ผ่านมา
- ดังนั้น หากมีการเปลี่ยนแปลงของเหตุปัจจัยต่างๆ ก็จะทำให้ผลการพยากรณ์คลาดเคลื่อนได้มาก
- ในการรายงานผลค่าพยากรณ์ ควรทำควบคู่ไปกับการระบุค่าความคลาดเคลื่อน (ที่ได้จากการทำ Backcasting) พร้อมกับความเสี่ยงที่อาจจะมีความคลาดเคลื่อนสูงขึ้นไปอีกหากปัจจัยต่างๆเปลี่ยนแปลง

ภาคผนวกที่ 3

SIR Compartmental Model

และการใช้งานโปรแกรม Vensim



SIR Compartment Model

Phattarin Phongwutthipong, MD
inspired by loads of experts
INCD, DOE

July 2,2024

Outline

- Role of modeling in epidemics
- Model construction
- Epidemic modeling
  - SIR
  - Other variants
- Additional issues
  - System dynamics
    - Causal loop diagram
    - Stock and flow diagram

Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

Role of modeling in Epidemics (1)

- โรคระบาด คือ การระบาดของโรคที่เกิดขึ้นอย่างรวดเร็วและมีจำนวนผู้ป่วยเพิ่มขึ้นอย่างมาก ในกลุ่มประชากรหรือพื้นที่ภาย ในระยะเวลาอันสั้น
  (ตามคำนิยามของ CDC)

- การระบาดของโรคจะเกิดขึ้นเมื่อมีเชื้อโรคและผู้ที่มีความเสี่ยงต่อการติดเชื้ออยู่ในจำนวนที่เพียงพอ และเชื้อโรคสามารถแพร่กระจายจากแหล่งที่มาไปยังผู้ที่มีความเสี่ยงได้อย่างมีประสิทธิภาพ

- การระบาดของโรคอาจเกิดจากปัจจัยต่างๆ เช่น
  - การเพิ่มขึ้นของปริมาณหรือความรุนแรงของเชื้อโรค
  - การนำเชื้อโรค ใหม่เข้ามา ในพื้นที่ที่ไม่เคยมีมาก่อน
  - การเปลี่ยนแปลงของวิธีการแพร่กระจายเชื้อทำ ให้มีผู้ที่มีความเสี่ยงมากขึ้น
  - การเปลี่ยนแปลงของภูมิคุ้มกันของผู้ติดเชื้อ
  - ปัจจัยที่เพิ่มการสัมผัสหรือการแพร่เข้าสู่ร่างกายผ่านช่องทาง ใหม่

Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

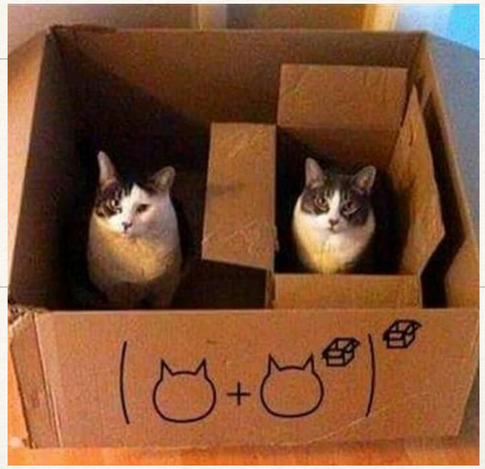
Role of modeling in Epidemics (2)

การจำลองสถานการณ์การระบาดของโรคอาจช่วยตอบคำถามเหล่านี้ได้:

- อะไรคือความเสี่ยงที่จะเกิดการระบาดของโรค?
- ความรุนแรงของการระบาดเป็นอย่างไร?
- จำนวน (และสัดส่วน) ของผู้ติดเชื้อจะมีมากน้อยแค่ไหน?
- จำนวน (และสัดส่วน) ของผู้เสียชีวิตจะมีมากน้อยแค่ไหน?
- การระบาดจะยาวนานแค่ไหน?
- ทุกคนมีความเสี่ยงที่จะติดเชื้อหรือไม่?
- การระบาดจะแพร่กระจายไปไกลแค่ไหน?
- มาตรการแทรกแซงแบบใดบ้างที่จะส่งผลต่อความเสี่ยง ความรุนแรง และระยะเวลาของการระบาด
- การจำลองสถานการณ์โดยใช้โมเดลทางคณิตศาสตร์ เช่น SIR, SEIR และ SEIRS จะช่วยให้สามารถประเมินผลกระทบของปัจจัยต่างๆ ต่อพลวัตของการระบาดได้ ซึ่งจะเป็นประโยชน์ในการวางแผนมาตรการควบคุมโรคและลดผลกระทบที่อาจเกิดขึ้น

Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

Compartment Model



- compartment model หมายถึงโมเดลทางคณิตศาสตร์ที่ใช้ในการศึกษาการแพร่ระบาดของโรคติดเชื้อ โดยแบ่งประชากรออกเป็นหลายกลุ่มหรือหรือ “ช่อง" (compartment” ที่แตกต่างกัน เช่น กลุ่มที่มีความเสี่ยง (Susceptible), กลุ่มที่ติดเชื้อ (Infected), กลุ่มที่หายป่วย (Recovered) เป็นต้น

- โมเดลประเภทนี้ ประชากรจะสามารถเปลี่ยนสถานะไปมาระหว่างกลุ่มต่างๆ ตามอัตราการเปลี่ยนแปลงที่กำหนดไว้ในโมเดล ซึ่งจะถูกอธิบายด้วยสมการเชิงอนุพันธ์ธรรมดา

- ตัวอย่างของ compartment model ที่นิยมใช้ ได้แก่ SIR, SEIR และ SEIS เป็นต้น

- โมเดลเหล่านี้ช่วยให้เข้าใจพลวัตของการแพร่ระบาดของโรคติดเชื้อ และสามารถประเมินผลกระทบของมาตรการควบคุมโรคต่างๆ ได้ จึงถูกนำมาใช้อย่างแพร่หลายในการศึกษาระบาดวิทยาและการตัดสินใจด้านสาธารณสุข

แนวทางการสร้างแบบจำลอง (Model Construction)

1. ค้นหาข้อมูลเกี่ยวกับโรคที่สนใจ (Find out about the disease of interest)
2. กำหนดวัตถุประสงค์ของโมเดล (Determine the objectives of your model)
3. ระบุข้อสมมติในการสร้างโมเดล (State the assumptions –simplify your model)
4. สร้างแผนภาพโมเดล(Sketch a model)
5. เขียนสมการเชิงอนุพันธ์ที่อธิบายพลวัตของโมเดล (Describe a model with set of ordinary differential equations : ODEs)
6. แก้สมการเชิงอนุพันธ์เพื่อหาผลลัพธ์ (Solve equations)
7. แปลผลและวิเคราะห์ผลลัพธ์จากโมเดล (Interpret the model outputs)

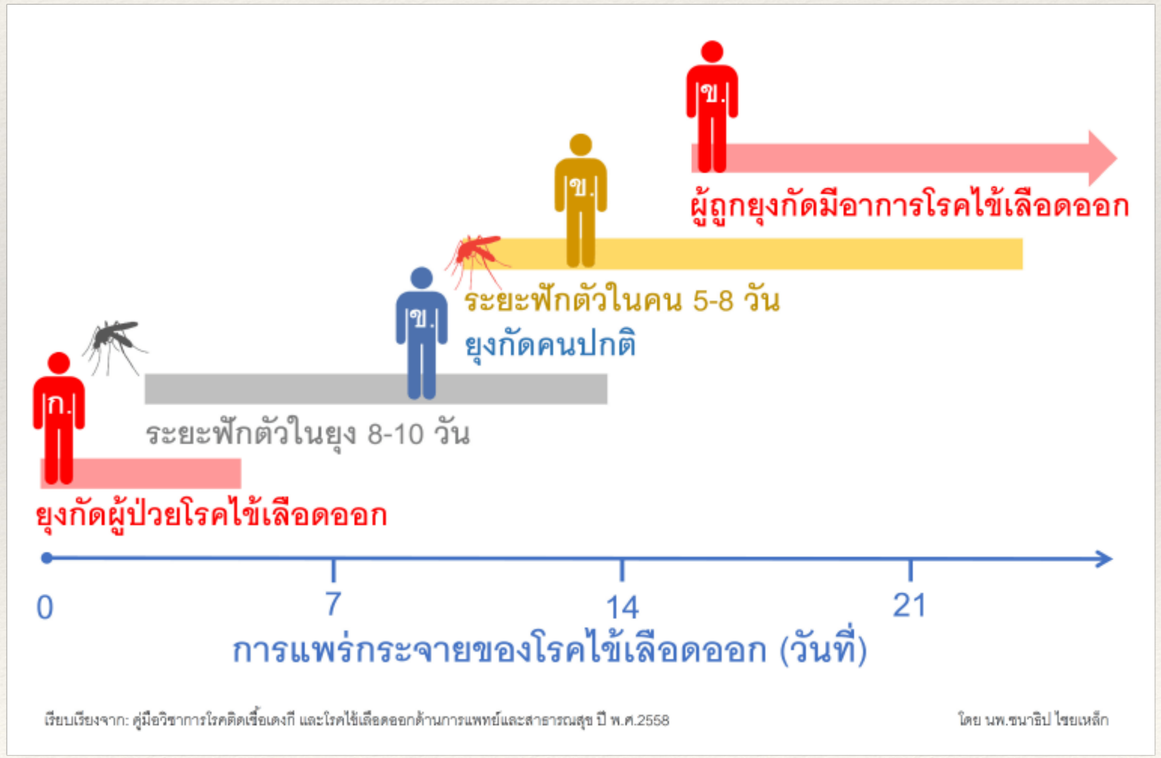
1. ค้นหาข้อมูลเกี่ยวกับโรคที่สนใจ (Find out about the disease of interest)

- ศึกษาลักษณะการติดเชื้อ ระยะฟักตัว ระยะการแพร่เชื้อ และการหายป่วย

ภาพประกอบจาก
- WHO-TDR-Stammers-01
- University of Florida
- Peters, W. and H.M. Gilles. 1997. Tropical Medicine & Parasitology. 4th edition. Mosby-Wolfe, Spain, 248 pp.

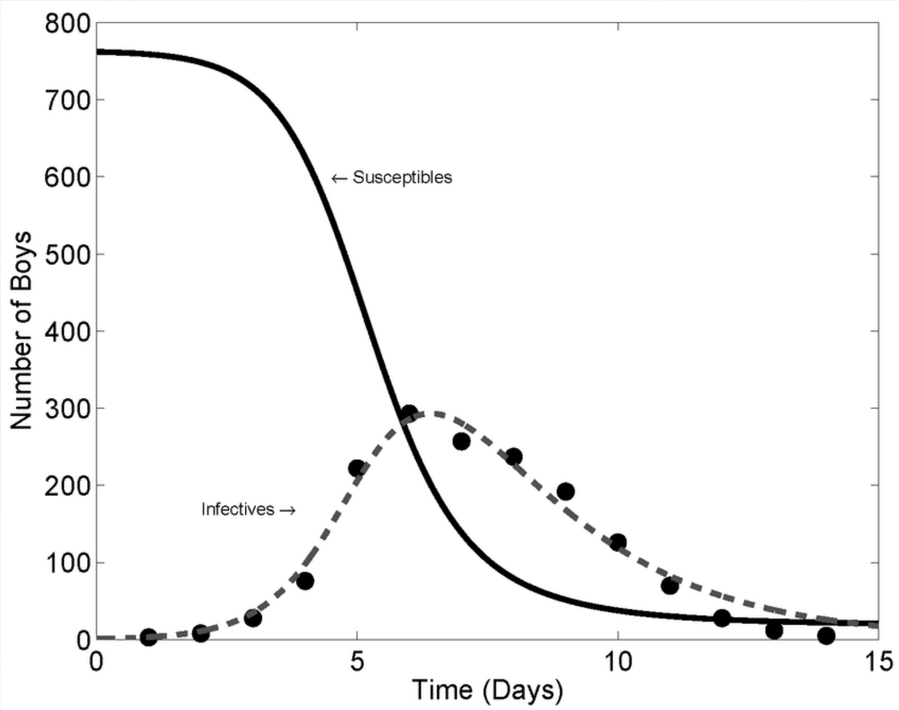
- มีอาการไข้สูง ติดต่อกัน 2-7 วัน
- หน้าแดง ปวดศีรษะ ปวดกล้ามเนื้อ ปวดกระดูก ปวดเบ้าตา บางราย มีปวดท้อง อาเจียน เบื่ออาหาร
- มีจุดแดงเล็กๆตามแขน ขาลำตัวรักแร้อาจมีเลือดกำเดาไหลและเลือด ออกตามไรฟัน

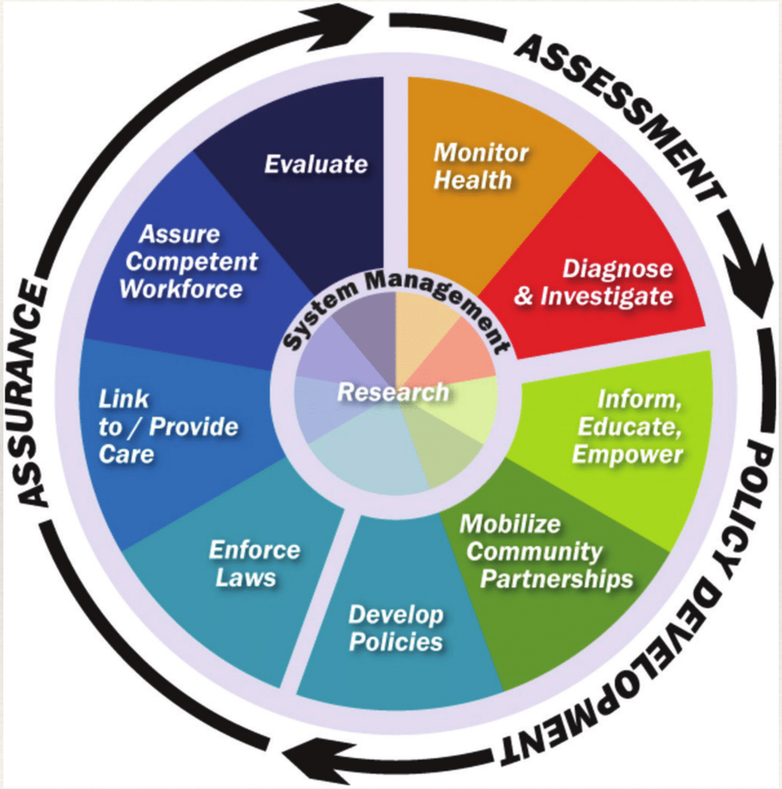
- โรคไข้เลือดออก เกิดจากการติเชื้อไวรัสเดงกี่ (dengue virus) โดยมียุงลายเป็นพาหะนำโรคมาสู่คน มักพบในประเทศเขตร้อนและระบาดในช่วงฤดูฝนของทุกปี อาการมีตั้งแต่ไม่รุนแรงไปจนถึงเสียชีวิตได้



2. กำหนดวัตถุประสงค์ของโมเดล (Determine the objectives of your model)

เช่น ต้องการศึกษาพลวัตของการแพร่ระบาด หรือประเมินผลของมาตรการควบคุมโรค

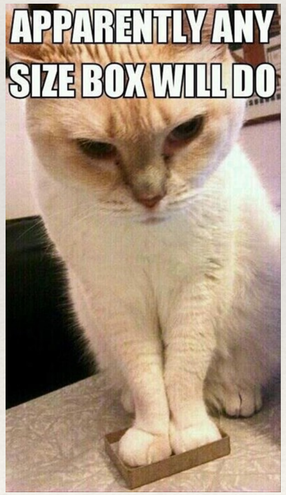




3. ระบุสมมติฐานในการสร้างโมเดล (State the assumptions –simplify your model)

- เพื่อให้โมเดลมีความเรียบง่ายและสามารถวิเคราะห์ได้ เช่น
  - การเปลี่ยนสถานะของประชากรจากกลุ่มหนึ่งไปยังอีกกลุ่มหนึ่งเป็นไปตามอัตราการเปลี่ยนแปลงที่กำหนดไว้ในโมเดล
  - อัตราการติดเชื้อและการแพร่กระจายเชื้อระหว่างกลุ่มประชากรเป็นไปตามกฎของการแพร่กระจายเชื้อ
  - ประชากรมีการผสมกันอย่างสุ่มและสม่ำเสมอ (well-mixed population)
  - ไม่มีการเกิดขึ้นหรือการเสียชีวิตของประชากรในระหว่างการศึกษา (closed population)
  - ไม่มีการเปลี่ยนแปลงของปัจจัยภายนอกที่มีผลต่อการแพร่ระบาด เช่น มาตรการควบคุมโรค
  - ระยะเวลาการฟักตัวของโรคและระยะเวลาการมีอาการป่วยเป็นไปตามการกระจายแบบปกติ
  - ไม่มีการแพร่กระจายเชื้อจากสัตว์สู่คน

- ข้อสมมติเหล่านี้ช่วยให้การสร้างโมเดลมีความเรียบง่ายและสามารถวิเคราะห์ได้ แต่อาจต้องมีการปรับปรุงให้สอดคล้องกับสถานการณ์จริงมากขึ้น



4. สร้างแผนภาพโมเดล (Sketch a model)

- แสดงการเปลี่ยนแปลงระหว่างสถานะต่างๆ ของประชากร เช่น Susceptible, Exposed, Infectious, Recovered
- ตัวอย่างของโมเดลระบาดวิทยาที่ใช้ในการศึกษาการแพร่ระบาดของโรคติดเชื้อ เช่น

1. โมเดล SIR (Susceptible-Infected-Recovered)
   - แบ่งประชากรออกเป็น 3 กลุ่ม: กลุ่มที่มีความเสี่ยง (Susceptible, S), กลุ่มที่ติดเชื้อ (Infected, I) และกลุ่มที่หายป่วย (Recovered, R)
   - บุคคลสามารถเปลี่ยนสถานะระหว่างกลุ่มต่างๆ ตามพลวัตของโรค
   - อธิบายด้วยสมการเชิงอนุพันธ์ธรรมดาที่ควบคุมอัตราการเปลี่ยนแปลงในแต่ละกลุ่ม
2. โมเดล SEIR (Susceptible-Exposed-Infected-Recovered)
   - คล้ายกับ SIR แต่มีกลุ่ม "Exposed (E)" เพิ่มเติม สำหรับบุคคลที่ติดเชื้อแต่ยังไม่มีอาการ
   - นำเข้าช่วงแฝงตัวก่อนที่บุคคลที่ติดเชื้อจะกลายเป็นผู้ติดเชื้อ
   - ใช้สำหรับโรคที่มีระยะฟักตัวที่ยาวนาน เช่น COVID-19
3. โมเดล SEIS (Susceptible-Exposed-Infected-Susceptible)
   - คล้ายกับ SEIR แต่บุคคลที่หายป่วยจะกลับมาอยู่ในกลุ่มเสี่ยงอีกครั้ง โดยไม่มีภูมิคุ้มกัน
   - บุคคลสามารถติดเชื้อได้หลายครั้ง
   - ใช้สำหรับโรคที่การติดเชื้อไม่ก่อให้เกิดภูมิคุ้มกันระยะยาว

Compartment Model

4. โมเดล SEIR ที่มีการแบ่งตามอายุ

- ขยายโมเดล SEIR มาตรฐานโดยแบ่งประชากรออกเป็นหลายกลุ่มอายุ เช่น ความรุนแรงของโรคและกลยุทธ์การให้วัคซีนที่แตกต่างกันตามอายุ

5. โมเดล SEIR ที่ขยายสำหรับ COVID-19

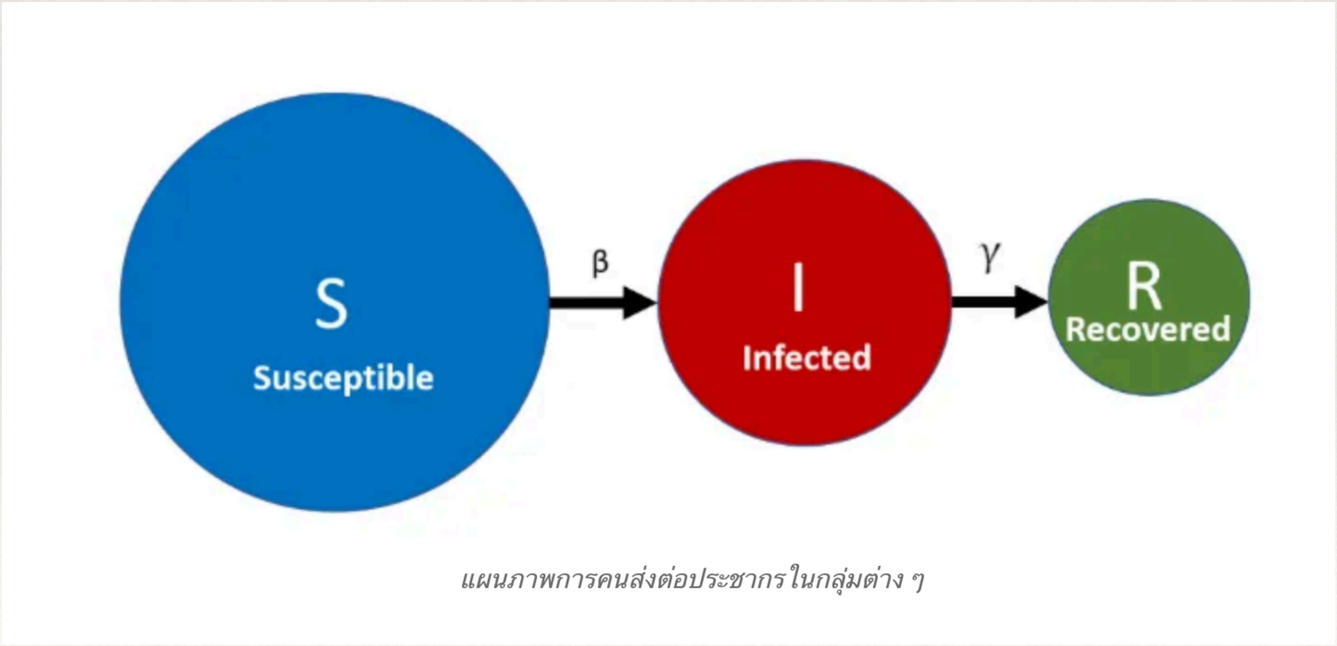
- ขยายโมเดล SEIR เพิ่มเติมโดยรวมกลุ่มอาการแสดงและไม่แสดงอาการ
- จำลองการดำเนินโรคแยกกันระหว่างผู้ที่ได้รับและไม่ได้รับวัคซีน

6. Other compartmental models:

- SIRD model (Susceptible-Infectious-Recovered-Deceased)
- SIRV model (Susceptible-Infectious-Recovered-Vaccinated)
- MSIR model (M class for maternally derived immunity)
- Variable contact rates
- Age structured model
- Stochastic model
- and etc.

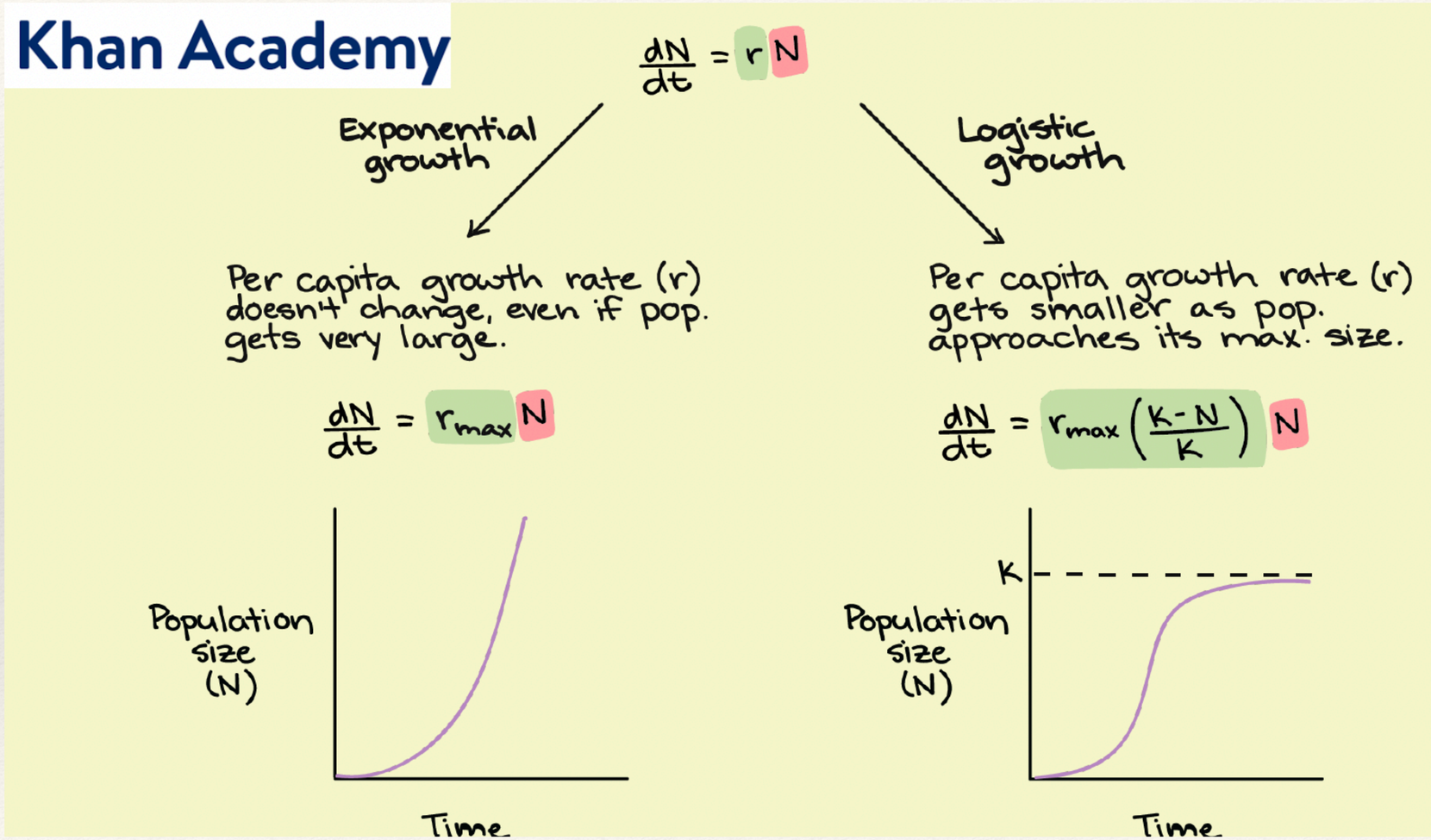
5. เขียนสมการเชิงอนุพันธ์ที่อธิบายพลวัตของโมเดล
(Describe a model with set of ordinary differential equations : ODEs)

- สมการจะแสดงอัตราการเปลี่ยนแปลงของประชากรในแต่ละสถานะ
- ในที่นี้ขอนำเสนอ SIR Model (Susceptible - Infected - Recovered)
  - SIR เป็นแบบจำลองทางคณิตศาสตร์ที่มีความซับซ้อนน้อย สามารถใช้งานกับการทำนายการแพร่ระบาดของโรคต่าง ๆ ได้ ซึ่งตัวแปรและรายละเอียดเกี่ยวกับโมเดล เปลี่ยนแปลงไปตามชนิดของโรคระบาดนั้น ๆ



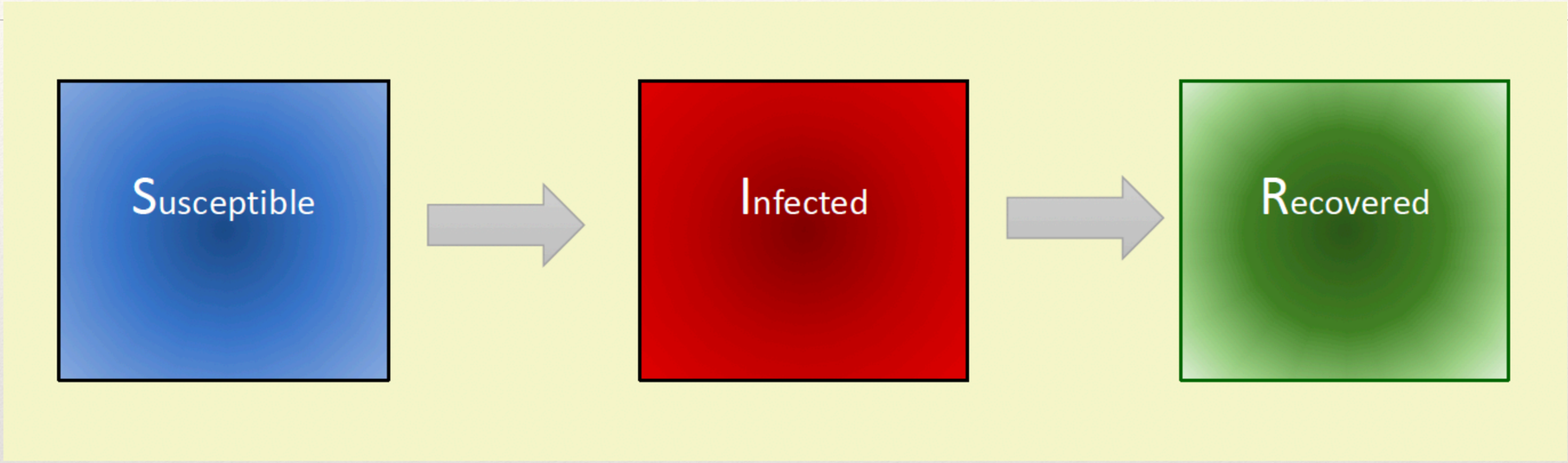
แผนภาพการคนส่งต่อประชากรในกลุ่มต่าง ๆ

Exponential vs Logistic Growth Curve



https://www.khanacademy.org/science/ap-biology/ecology-ap/population-ecology-ap/a/exponential-logistic-growth

SIR Model



- โมเดลจึงแบ่งคนออกเป็นทั้งหมด 3 กลุ่มคือ Susceptible (กลุ่มคนที่มีโอกาสติดเชื้อ), Infectious (กลุ่มคนที่ติดเชื้อและแพร่เชื้อได้) และ Recovery (กลุ่มคนที่รักษาหาย)

- Recovered จะหมายถึงผู้ที่ ติดเชื้อหรือแพร่เชื้อไม่ได้อีกต่อไป จึงรวมไปถึงผู้ที่เสียชีวิต และผู้ติดเชื้อที่โดน Isolate (อย่างแท้จริง) ด้วย

SIR Model

สมการพื้นฐานของแบบจำลอง SIR (Kermack-McKendrick Equations)

DEPARTMENT OF SURGERY
FACULTY OF MEDICINE RAMATHIBODI HOSPITAL

(1) $$\frac{dS(t)}{dt} = -\frac{\beta I(t)S(t)}{N}$$

(2) $$\frac{dI(t)}{dt} = \frac{\beta I(t)S(t)}{N} - \gamma I(t)$$

(3) $$\frac{dR(t)}{dt} = \gamma I(t)$$

- สมการ (๑) บรรยายถึงการเปลี่ยนแปลงของ S(t) ต่อหน่วยเวลา ในประชากรขนาด Nโดย S(t) คือจำนวน Susceptible ณ เวลา t หลังเชื้อเริ่มระบาด และ I(t) คือจำนวนผู้ติดเชื้อ (Infected) ณ เวลา t
- สมการ (๒) บรรยายถึงอัตราการเปลี่ยนแปลงของจำนวน I(t)
- สมการ (๓) บรรยายถึงการเปลี่ยนแปลงของจำนวนกลุ่ม Recovered, R(t)
- และดังนั้น N = S(t) + I(t) + R(t) = ค่าคงที่

SIR Model

$$\frac{dS}{dt} = -\beta IS$$

$$\frac{dI}{dt} = \beta IS - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

สมการของ SIR model *แบบไม่ได้กำหนดจำนวน* population

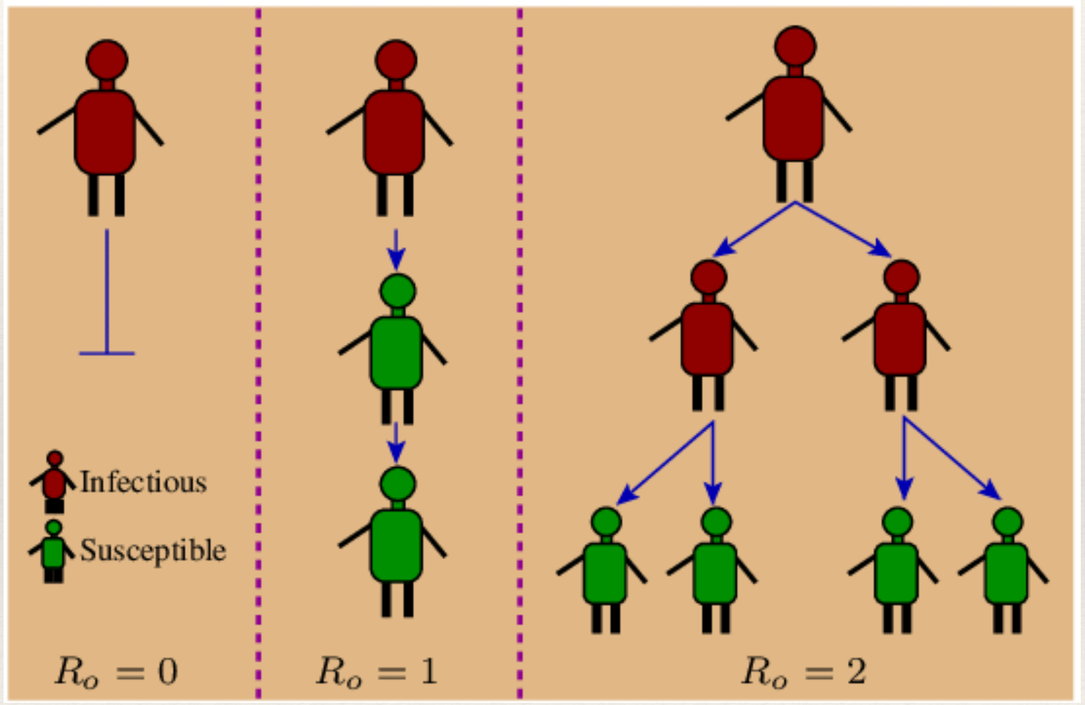
- β (Beta) คือ Transmission rate หรืออัตราการแพร่เชื้อ
สามารถหาได้จากระยะเวลานานเท่าใดที่ทำให้เกิดผู้ป่วยใหม่ขึ้นได้ 1 คน ตัวอย่างเช่น โดยเฉลี่ยแล้วระยะเวลา 7 วันจะมีผู้ติดเชื้อรายใหม่ 1 คน
ดังนั้น β = 1/7 = 0.143 หมายความว่าระยะเวลา 1 วัน มีผู้ป่วยใหม่ 0.143 คน หากพูดเป็นอัตราการแพร่เชื้อ สามารถพูดได้ว่า คนป่วยสามารถแพร่เชื้อให้ 0.143 คนใน 1 วัน เป็นต้น

γ (Gamma) คือ Recovery rate หรืออัตราการหายป่วย สามารถหาได้จากระยะเวลานานเท่าใดที่ทำให้ผู้ป่วย 1 คนหายจากโรค ตัวอย่างเช่น โดยเฉลี่ยแล้วผู้ป่วยใช้เวลา 10 วันในการหายจากโรค ดังนั้น γ = 1/10 = 0.1 หมายความว่าระยะเวลา 1 วัน มีจะผู้หายป่วย 0.1 คน

Basic Reproductive Ratio (R0)

$$\frac{dS(t)}{dt} = -\frac{\beta I(t)S(t)}{N}$$
$$\frac{dI(t)}{dt} = \frac{\beta I(t)S(t)}{N} - \gamma I(t)$$
$$\frac{dR(t)}{dt} = \gamma I(t)$$

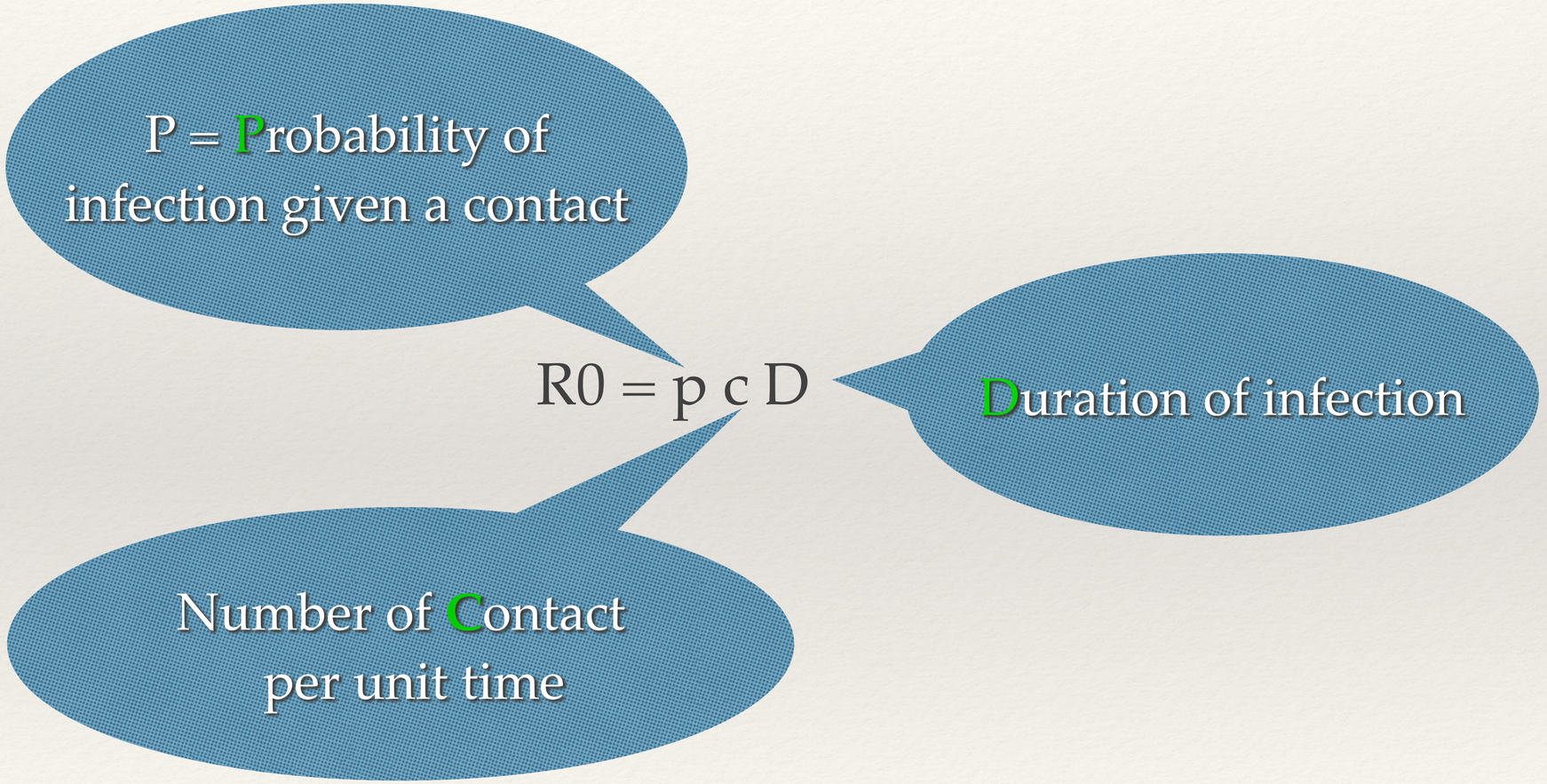
$$\frac{\beta}{\gamma} \equiv r$$



https://www.researchgate.net/figure/An-example-scenario-of-basic-reproduction-number-Ro_fig3_344013981

- Basic reproductive ratio (R0; R naught) คือ ค่าที่แสดงถึงจำนวนเฉลี่ยของผู้ติดเชื้อรายใหม่ที่เกิดจากผู้ติดเชื้อรายเดิมหนึ่งราย ในสภาวะที่ประชากรทั้งหมดยังไม่มีภูมิคุ้มกัน (Susceptible)
- กล่าวอีกนัยหนึ่ง คือ r คือจำนวน Susceptible ที่จะติดเชื้อที่ผู้ติดเชื้อคนหนึ่งๆ จะแพร่ได้ ณ จุดเวลำหนึ่งๆ ต่อจำนวนผู้ติดเชื้อที่หำยจำกโรค ณ จุดเวลาเดียวกัน
- ดังนั้น r = 3 หมายความว่า ณ จุดเวลาหนึ่ง จะมีผู้ติดเชื้อใหม่ 3 ต่อผู้หายจากเชื้อ 1

Basic Reproductive Ratio (R0)



Basic Reproductive Ratio (R0)

ค่า R0 มีความสำคัญในการวิเคราะห์และคาดการณ์การแพร่ระบาดของโรคติดเชื้อ โดยมีความหมายดังนี้:

- ถ้า R0 > 1 แสดงว่าโรคจะแพร่ระบาดและคงอยู่ในประชากร
- ถ้า R0 < 1 แสดงว่าโรคจะค่อยๆ สงบลง และจะไม่แพร่ระบาดต่อไป
- ถ้า R0 = 1 แสดงว่าโรคจะคงอยู่ในระดับคงที่

ค่า R0 ขึ้นอยู่กับปัจจัยต่างๆ เช่น อัตราการติดเชื้อ ระยะเวลาการติดเชื้อ และพฤติกรรมการแพร่กระจายเชื้อ ดังนั้นการลดค่า R0 ให้ต่ำกว่า 1 จึงเป็นเป้าหมายสำคัญในการควบคุมการแพร่ระบาดของโรค

Effective Reproductive Ratio (Re or Rt)

- effective reproductive ratio (Re หรือ Rt) คือ ค่าที่แสดงถึงจำนวนเฉลี่ยของผู้ติดเชื้อรายใหม่ที่เกิดจากผู้ติดเชื้อรายเดิมหนึ่งราย ณ เวลาใดเวลาหนึ่ง โดยคำนึงถึงปัจจัยต่างๆ ที่อาจเปลี่ยนแปลงไป เช่น มาตรการควบคุมโรค ภูมิคุ้มกันของประชากร เป็นต้น

- ค่า Re จะเปลี่ยนแปลงไปตามมาตรการควบคุมโรคและภูมิคุ้มกันของประชากร ดังนั้นการติดตามค่า Re อย่างต่อเนื่องจึงมีความสำคัญในการประเมินประสิทธิผลของมาตรการควบคุมโรคและการวางแผนรับมือกับการระบาด

- ตัวอย่างเช่น ในช่วงการระบาดของ COVID-19 ค่า Re จะลดลงเมื่อมีมาตรการควบคุมที่เข้มงวดมากขึ้น เช่น การเว้นระยะห่าง การสวมหน้ากากอนามัย และการให้วัคซีน ซึ่งจะช่วยลดการแพร่กระจายเชื้อและควบคุมการระบาดได้

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Policy Platform

Assessing the Potential of a Candidate Dengue Vaccine with Mathematical Modeling

WHO-VMI Dengue Vaccine Modeling Group*

- ✓ Allow for immune enhancement (Model a – c)
- ✓ Allow for short-term cross protection (Model c – d)
- ✓ Allow for different infectiousness and transmissibility of secondary infection (Model d and e)

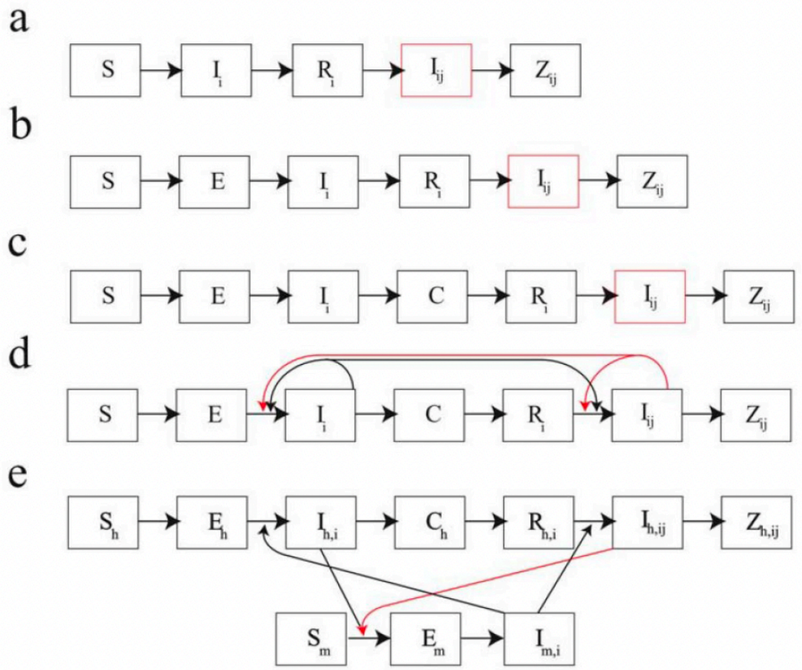
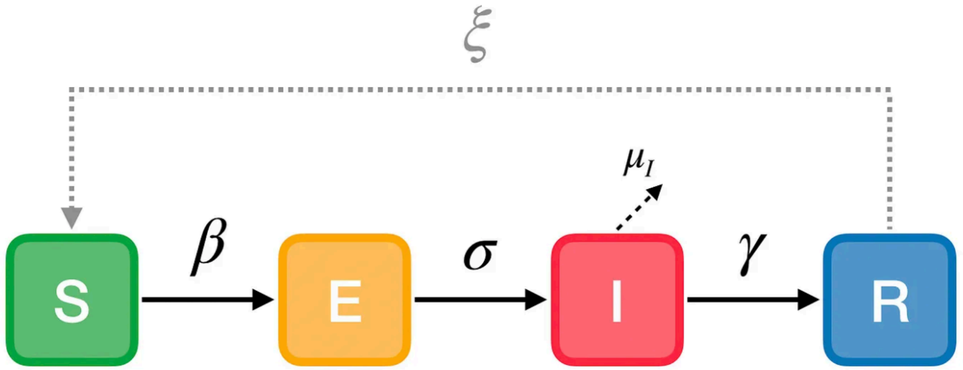


Figure 1. Example structures of dengue models. The disease state space of five alternative

TMHG

SEIR Model



- **σ (Sigma)** คือ Incubation period หรือระยะการฟักตัว ที่ภาครัฐและเอกชน ออกนโยบาย State quarantine 14 วัน ที่มาจากหลักฐานที่ว่าผู้ที่มีความเสี่ยงติดเชื้อ COVID-19 จะแสดงอาการออกมาใน 14 วัน หากไม่มีอาการ สามารถมั่นใจได้ในระดับหนึ่งว่าผู้ที่มีความเสี่ยงนั้น ไม่ได้ติดเชื้อ COVID-19 แต่อาจเห็นได้จากข่าวที่ออกมาช่วงหลังว่า เชื้อ COVID-19 อาจใช้ระยะเวลาฟักตัวนานกว่านั้น

$$\frac{dS}{dt} = \frac{-\beta SI}{N}$$

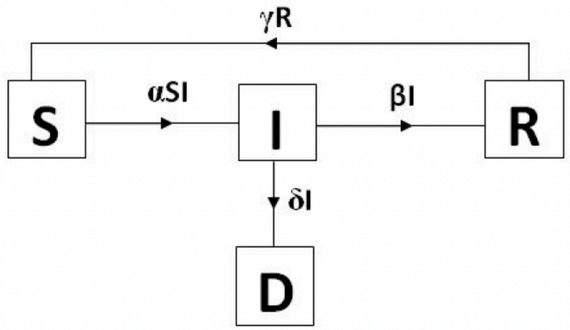
$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

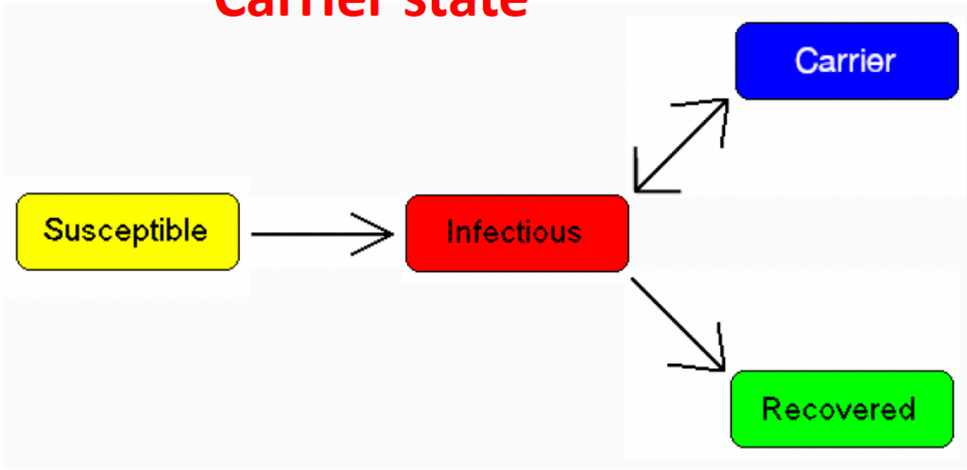
$$\frac{dR}{dt} = \gamma I$$

Example of other model variants

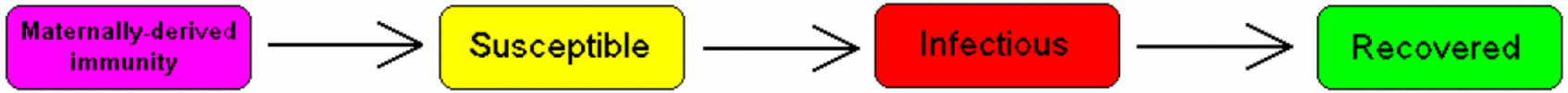
SIRD



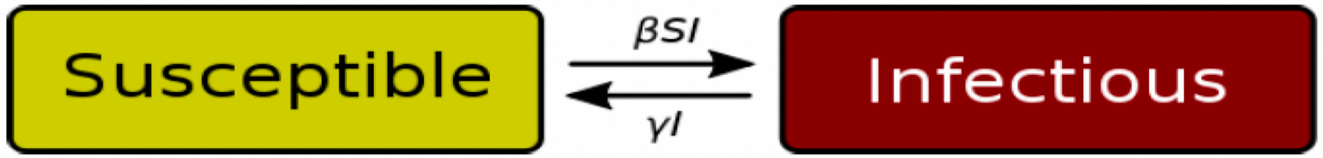
Carrier state



MSIR

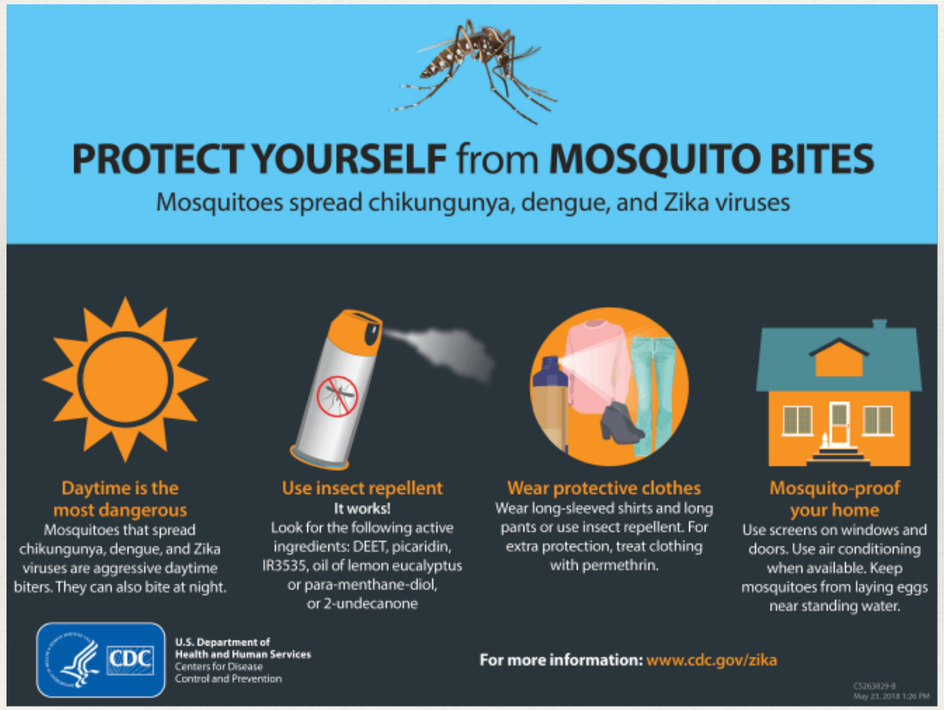


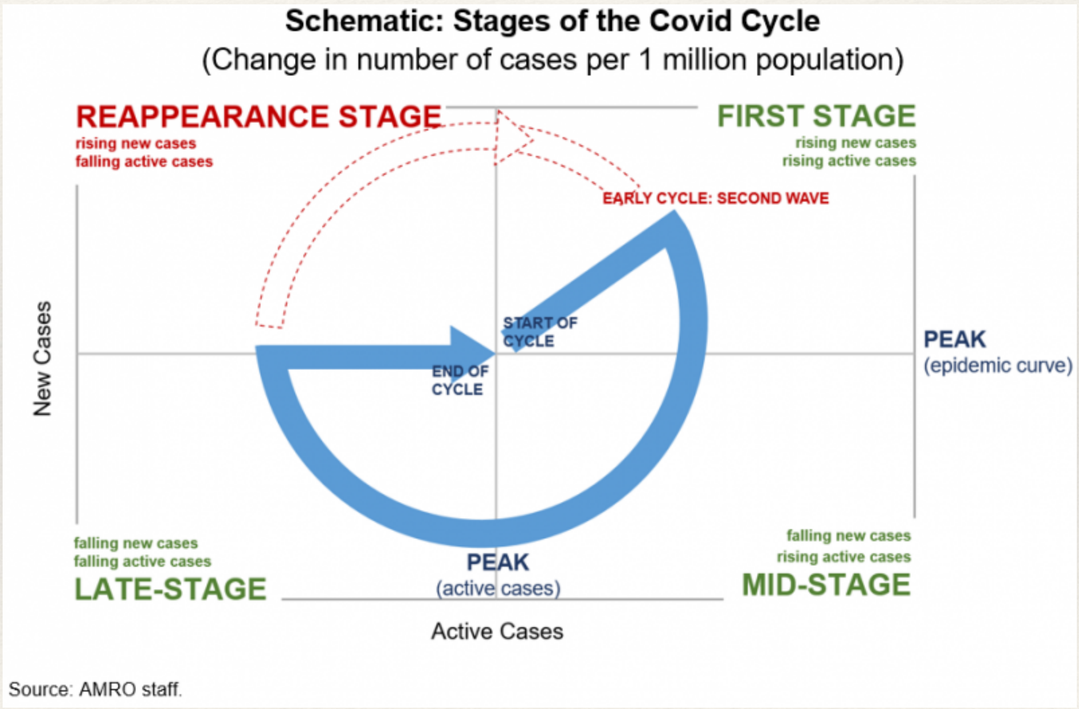
SIS



7. แปลผลและวิเคราะห์ผลลัพธ์จากโมเดล (Interpret the model outputs)

ตีความผลลัพธ์ในเชิงการแพร่ระบาดและประเมินผลของมาตรการควบคุมโรค





Source: AMRO staff.

System Dynamics

- System Dynamics เป็นแนวทางในการทำความเข้าใจพฤติกรรมที่ไม่เป็นเชิงเส้นของระบบที่ซับซ้อน โดยใช้แนวคิดของ "สต็อก" (stock) "กระแส" (flow) และ "ลูปย้อนกลับ" (feedback loop)
- ระบบที่ซับซ้อน (Complex System) หมายถึงระบบที่ประกอบด้วยองค์ประกอบจำนวนมากที่อาจมีปฏิสัมพันธ์กัน เช่น ระบบภูมิอากาศโลก สมองมนุษย์ ระบบไฟฟ้า และระบบขนส่ง
- System Dynamics มีรากฐานมาจากสาขาธุรกิจและวิศวกรรม
- เป็นเครื่องมือที่มีประโยชน์ในการเข้าใจระบบที่ซับซ้อนได้ดีขึ้น และระบุจุดที่สามารถสร้างผลกระทบ (leverage) ต่อระบบ

โดยสรุป System Dynamics เป็นแนวทางในการศึกษาระบบที่ซับซ้อน โดยมุ่งเน้นการทำความเข้าใจพฤติกรรมที่ไม่เป็นเชิงเส้นและการมีปฏิสัมพันธ์ระหว่างองค์ประกอบต่างๆ ในระบบ เพื่อระบุจุดที่สามารถสร้างผลกระทบต่อระบบได้

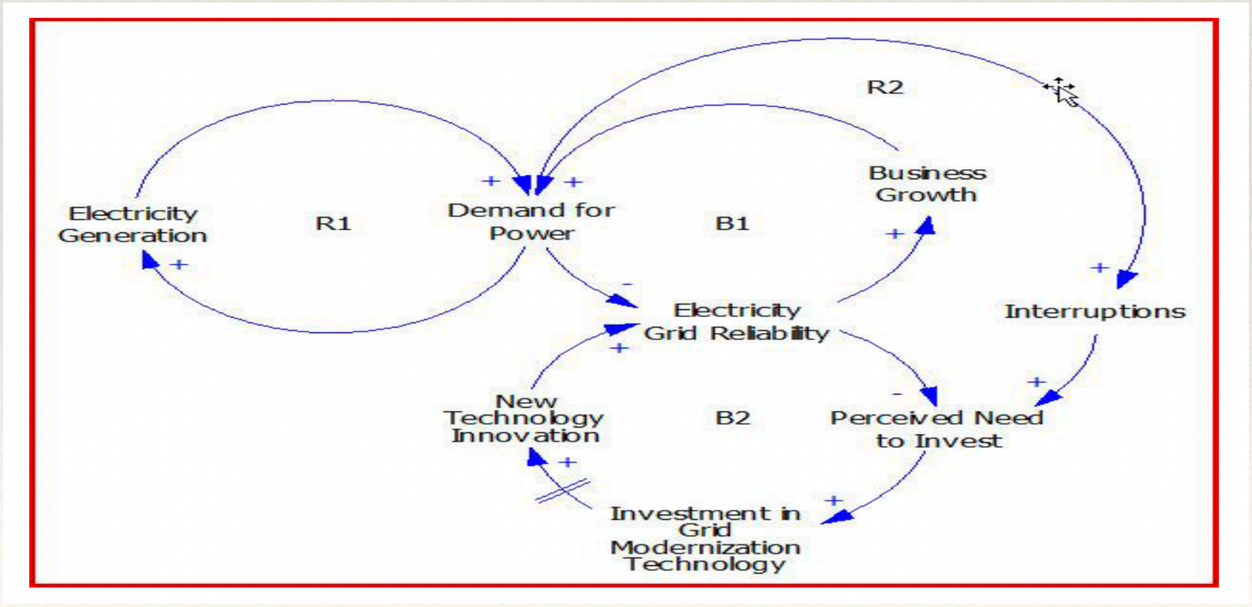
Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

System Dynamics (cont’)

- Causal loop diagram

  - แผนภาพระบบจะแสดงองค์ประกอบหลัก เช่น กลุ่มประชากรที่มีความเสี่ยง, กลุ่มที่ติดเชื้อ, กลุ่มที่หายป่วย และการติดต่อระหว่างกลุ่มเหล่านี้
  - แผนภาพระบบจะแสดงลูปป้อนกลับเชิงบวก (Positive Feedback Loop) ที่ทำให้การแพร่ระบาดขยายตัว และลูปป้อนกลับเชิงลบ (Negative Feedback Loop) ที่ช่วยควบคุมการแพร่ระบาด
  - ลูปป้อนกลับเชิงบวกและเชิงลบอาจเกิดขึ้นพร้อมกันในระบบ และมีอิทธิพลต่อพลวัตของการแพร่ระบาดในทิศทางตรงกันข้าม

โดยสรุป แผนภาพระบบเป็นเครื่องมือสำคัญในการศึกษาและวิเคราะห์การแพร่ระบาดของโรคติดเชื้อ โดยแสดงองค์ประกอบและความสัมพันธ์ในรูปของลูปป้อนกลับเชิงบวกและเชิงลบ ซึ่งจะช่วยให้เข้าใจพฤติกรรมที่ซับซ้อนของระบบได้ดียิ่งขึ้น

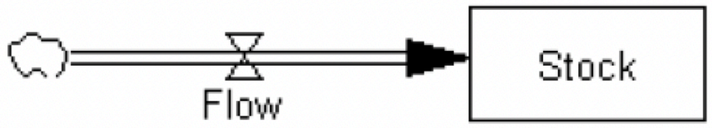


Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

System Dynamics (cont’)

- Stock and flow diagram
  - แผนภาพสต็อกและกระแส (Stock and Flow Diagram) เป็นการแปลงลูปป้อนกลับ (Causal Loop) ให้อยู่ในรูปแบบที่สามารถวิเคราะห์เชิงปริมาณได้อย่างละเอียดยิ่งขึ้น
  - สต็อก (Stock) หมายถึง ตัวแปรที่สะสมหรือลดลงตามเวลา เช่น จำนวนประชากรที่มีความเสี่ยง, จำนวนผู้ติดเชื้อ, จำนวนผู้หายป่วย เป็นต้น
  - กระแส (Flow) หมายถึง อัตราการเปลี่ยนแปลงของสต็อก เช่น อัตราการติดเชื้อ, อัตราการหายป่วย, อัตราการเสียชีวิต เป็นต้น
  - การแปลงลูปป้อนกลับให้อยู่ในรูปแบบของแผนภาพสต็อกและกระแส จะช่วยให้สามารถวิเคราะห์และจำลองพลวัตของการแพร่ระบาดได้อย่างละเอียดและเชิงปริมาณ
  - แผนภาพสต็อกและกระแสจะช่วยให้เข้าใจถึงกลไกการเปลี่ยนแปลงของระบบได้ดีขึ้น และสามารถประเมินผลกระทบของปัจจัยต่างๆ ต่อการแพร่ระบาดได้

โดยสรุป แผนภาพสต็อกและกระแสเป็นเครื่องมือที่ช่วยให้การวิเคราะห์การแพร่ระบาดของโรคติดเชื้อมีความละเอียดและเชิงปริมาณมากขึ้น โดยแสดงถึงการสะสมและการเปลี่ยนแปลงของตัวแปรต่างๆ ในระบบ



– Look very similar to compartmental model!

Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

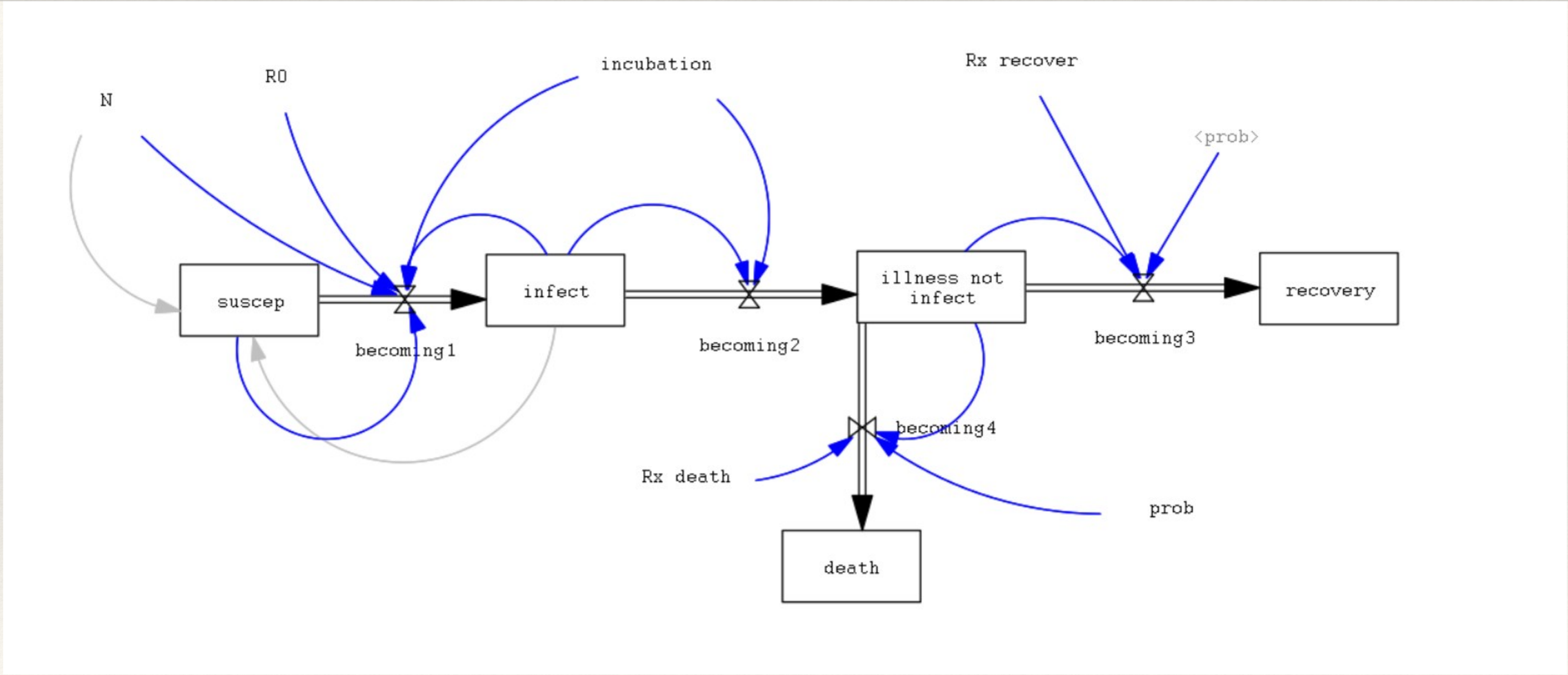
Take Home Messages

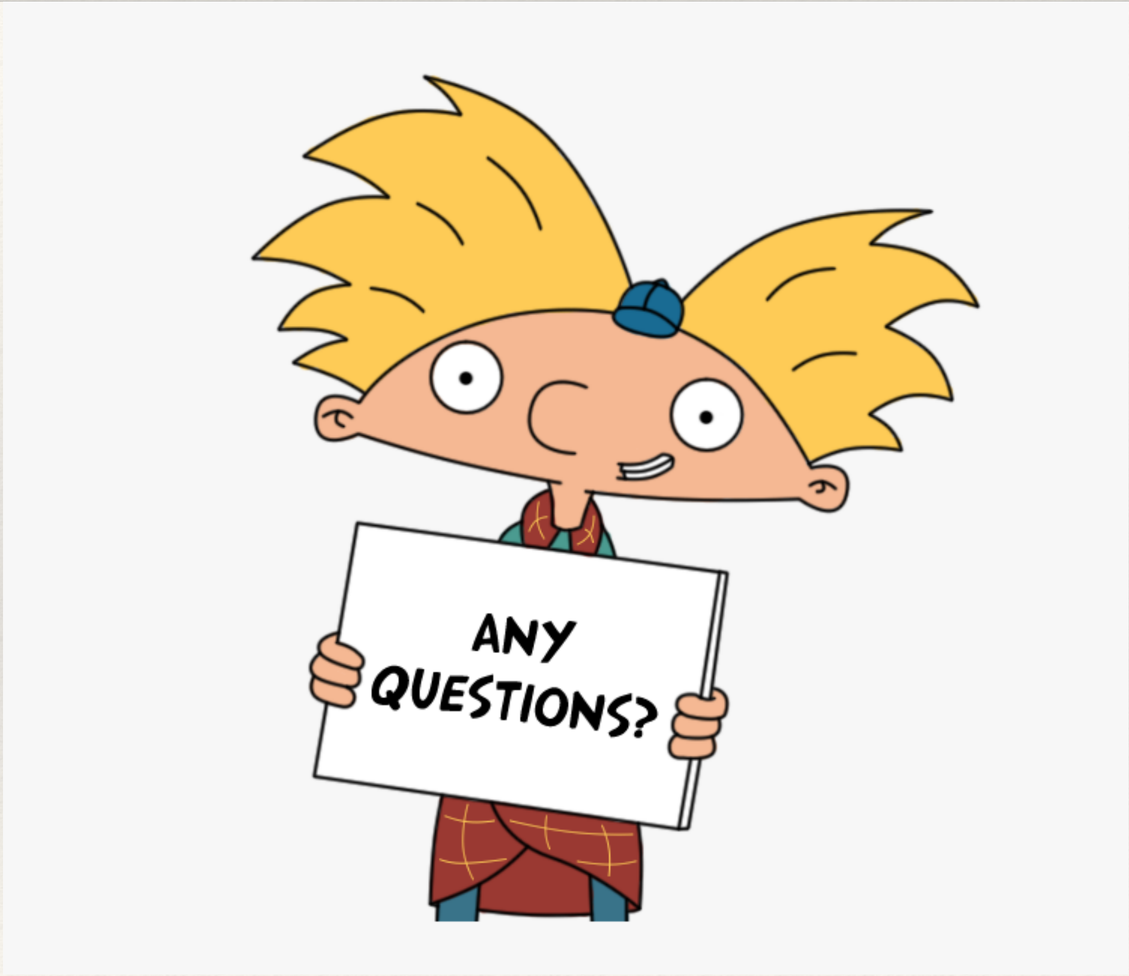
- Compartment model คือการแบ่งประชากรออกเป็นหลายส่วนหรือ "ช่อง" (compartment) ตามพลวัตของการแพร่กระจายเชื้อโรค
- การแบ่งประชากรออกเป็นช่องต่างๆ เช่น กลุ่มที่มีความเสี่ยง (Susceptible), กลุ่มที่ติดเชื้อ (Infected), กลุ่มที่หายป่วย (Recovered) เป็นต้น จะช่วยให้เข้าใจพลวัตของการแพร่กระจายเชื้อได้ดีขึ้น
- ความเข้าใจที่ชัดเจนเกี่ยวกับอัตราการเกิดซ้ำ (reproduction number) และพารามิเตอร์อื่นๆ ที่สำคัญ เช่น ระยะฟักตัว, ระยะหายป่วย, ระยะเวลาติดเชื้อ เป็นสิ่งจำเป็นต่อการวิเคราะห์พลวัตของการแพร่กระจายเชื้อ
- แนวคิดของระบบพลวัต (system dynamics) สามารถนำมาประยุกต์ใช้ในการทำความเข้าใจและวิเคราะห์แบบจำลองแบบแบ่งส่วนทางระบาดวิทยาได้

โดยสรุป แบบจำลองแบบแบ่งส่วนในระบาดวิทยา เป็นแนวคิดที่แบ่งประชากรออกเป็นหลายส่วนตามพลวัตของการแพร่กระจายเชื้อ เพื่อช่วยให้เข้าใจและวิเคราะห์การระบาดของโรคได้ดียิ่งขึ้น

Ref: Rapeepong Suphanchaimat, MD
FETP and IHPP

Example of Dengue infection SIR model

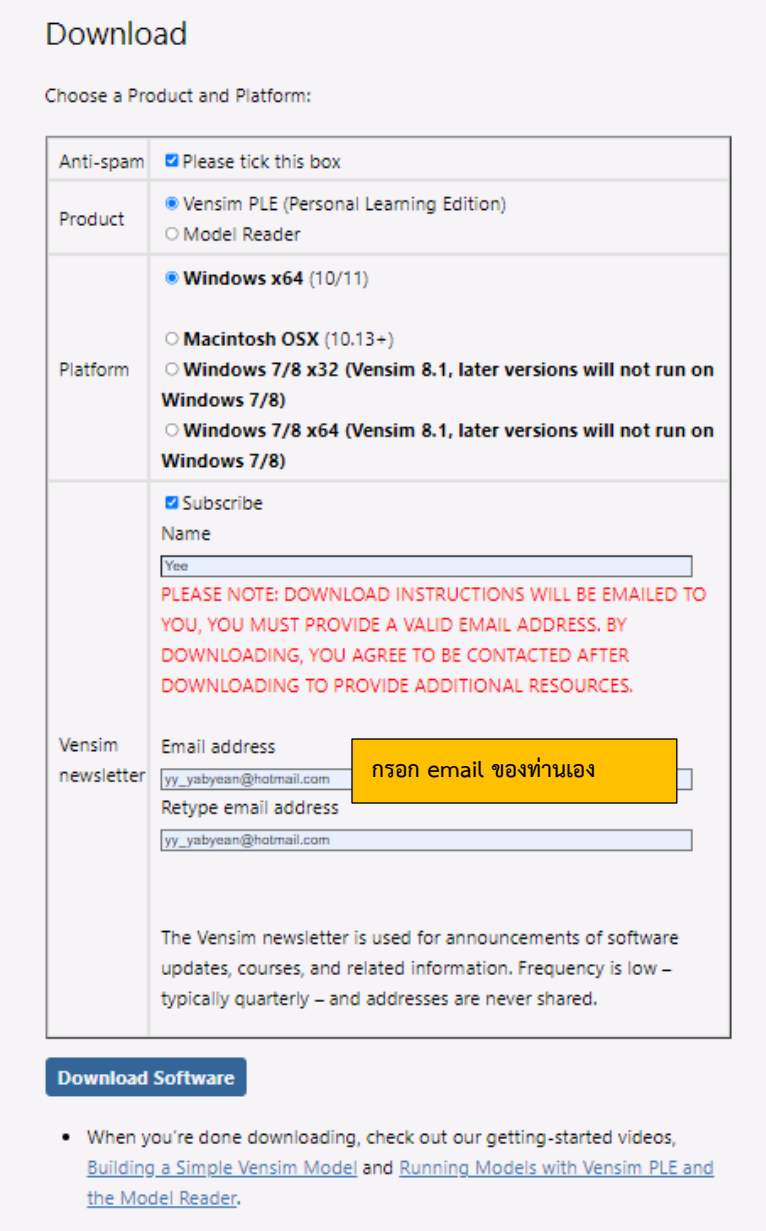



ANY
QUESTIONS?

ขั้นตอนการติดตั้งโปรแกรม vensim

1. **ดาวน์โหลดโปรแกรม vensim**

   https://vensim.com/free-downloads/



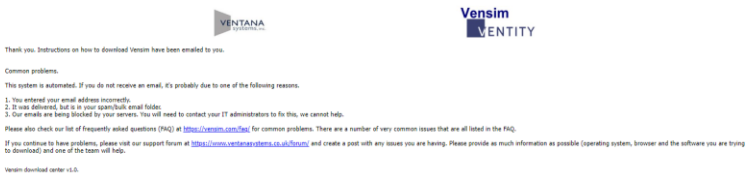
Download

Choose a Product and Platform:

Anti-spam	☑ Please tick this box
Product	◉ Vensim PLE (Personal Learning Edition) ○ Model Reader
Platform	◉ **Windows x64** (10/11) ○ **Macintosh OSX** (10.13+) ○ **Windows 7/8 x32 (Vensim 8.1, later versions will not run on Windows 7/8)** ○ **Windows 7/8 x64 (Vensim 8.1, later versions will not run on Windows 7/8)**
Vensim newsletter	☑ Subscribe Name Yee PLEASE NOTE: DOWNLOAD INSTRUCTIONS WILL BE EMAILED TO YOU, YOU MUST PROVIDE A VALID EMAIL ADDRESS. BY DOWNLOADING, YOU AGREE TO BE CONTACTED AFTER DOWNLOADING TO PROVIDE ADDITIONAL RESOURCES. Email address yy_yabyean@hotmail.com กรอก email ของท่านเอง Retype email address yy_yabyean@hotmail.com The Vensim newsletter is used for announcements of software updates, courses, and related information. Frequency is low – typically quarterly – and addresses are never shared.

Download Software

- When you're done downloading, check out our getting-started videos, Building a Simple Vensim Model and Running Models with Vensim PLE and the Model Reader.



Thank you. Instructions on how to download Vensim have been emailed to you.

Common problems.

This system is automated. If you do not receive an email, it's probably due to one of the following reasons.

1. You entered your email address incorrectly.
2. It was delivered, but is in your spam/bulk email folder.
3. Our emails are being blocked by your servers. You will need to contact your IT administrators to fix this, we cannot help.

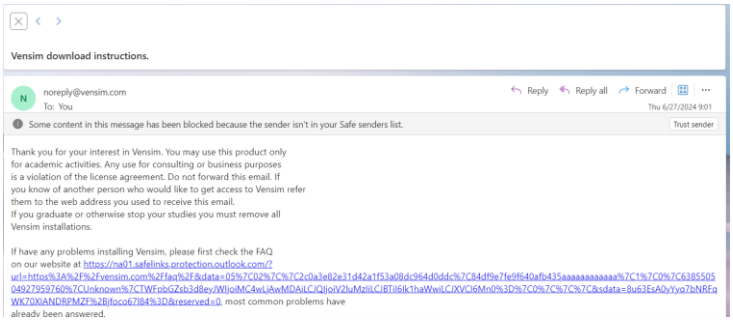
Please also check our list of frequently asked questions (FAQ) at https://vensim.com/faq/ for common problems. There are a number of very common issues that are all listed in the FAQ.

If you continue to have problems, please visit our support forum at https://www.ventanasystems.co.uk/forum/ and create a post with any issues you are having. Please provide as much information as possible (operating system, browser and the software you are trying to download) and one of the team will help.

Vensim download center v1.0.

2. **เมื่อกรอกข้อมูลแล้ว ระบบจะส่งอีเมลไปยังอีเมลที่ท่านกรอกด้านบน**

- โปรด Check อีเมล ของท่าน และกด link เพื่อดาวน์โหลด



Vensim download instructions.

noreply@vensim.com
To: You

Reply Reply all Forward
Thu 6/27/2024 9:01

Some content in this message has been blocked because the sender isn't in your Safe senders list. Trust sender

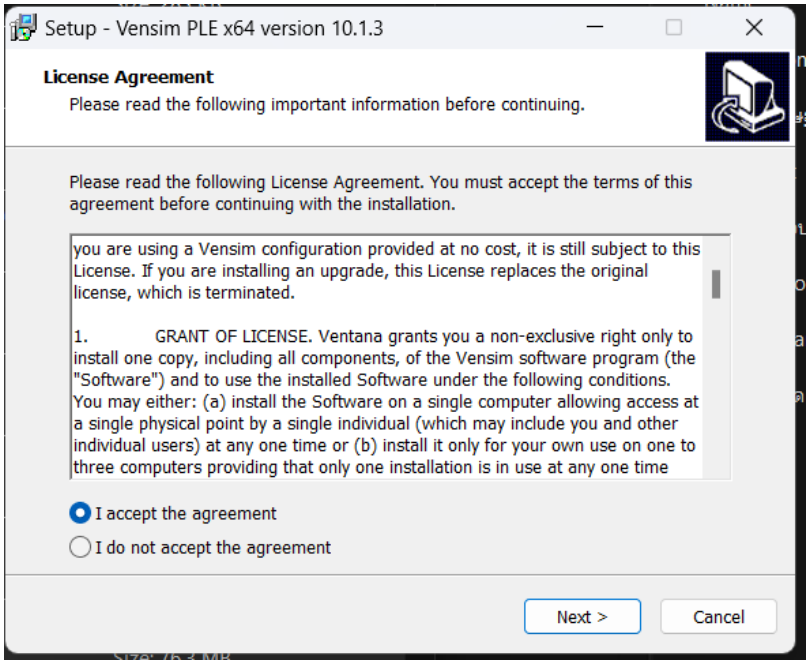
Thank you for your interest in Vensim. You may use this product only for academic activities. Any use for consulting or business purposes is a violation of the license agreement. Do not forward this email. If you know of another person who would like to get access to Vensim refer them to the web address you used to receive this email.
If you graduate or otherwise stop your studies you must remove all Vensim installations.

If have any problems installing Vensim, please first check the FAQ on our website at https://na01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fvensim.com%2Ffaq%2F&data=05%7C02%7C%7C2c0a3e82e31d42a1f53a08dc964d0ddc%7C84df9e7fe9f640afb435aaaaaaaaaaa%7C1%7C0%7C6385505049279597 60%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=8u63EsA0yYyq7bNRFqWK70XIANDRPMZF%2Bjfoco67I84%3D&reserved=0 most common problems have alreadv been answered.

You can download the software by clicking the following link.
NOTE: If you get an error when you click the link, please copy and paste the entire link into a browser.
Some email programs split the link over two lines, if this happens, it will not work.

https://www.vensim.com/php-bin/VSUK/CustomerDownloads_php/FreeDownload.php?x_productid=PLE&x_startdownload=1&x_platform=WIN64&x_name=Yee&x_emailcheck=yy_yabyean@hotmail.com&x_email=yy_yabyean@hotmail.com&x_subscribe=1&x_antispam=1

3. **เลือกดาวโหลดลงเครื่อง**



Setup - Vensim PLE x64 version 10.1.3

License Agreement
Please read the following important information before continuing.

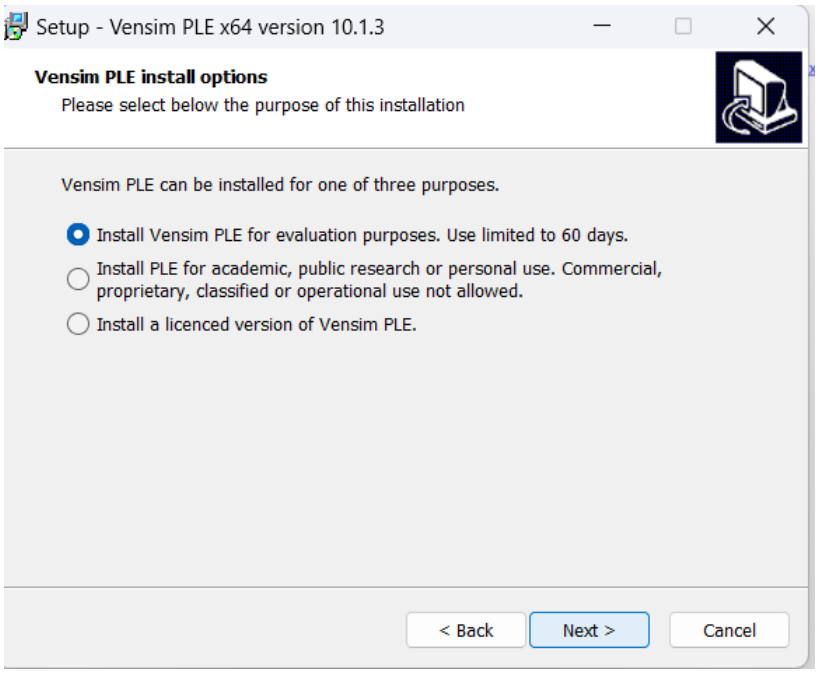
Please read the following License Agreement. You must accept the terms of this agreement before continuing with the installation.

you are using a Vensim configuration provided at no cost, it is still subject to this License. If you are installing an upgrade, this License replaces the original license, which is terminated.

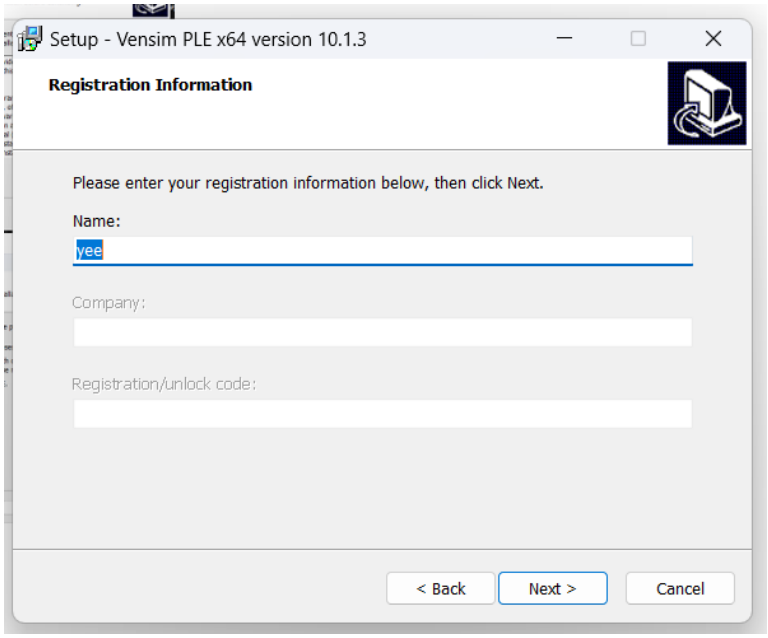
1. GRANT OF LICENSE. Ventana grants you a non-exclusive right only to install one copy, including all components, of the Vensim software program (the "Software") and to use the installed Software under the following conditions. You may either: (a) install the Software on a single computer allowing access at a single physical point by a single individual (which may include you and other individual users) at any one time or (b) install it only for your own use on one to three computers providing that only one installation is in use at any one time

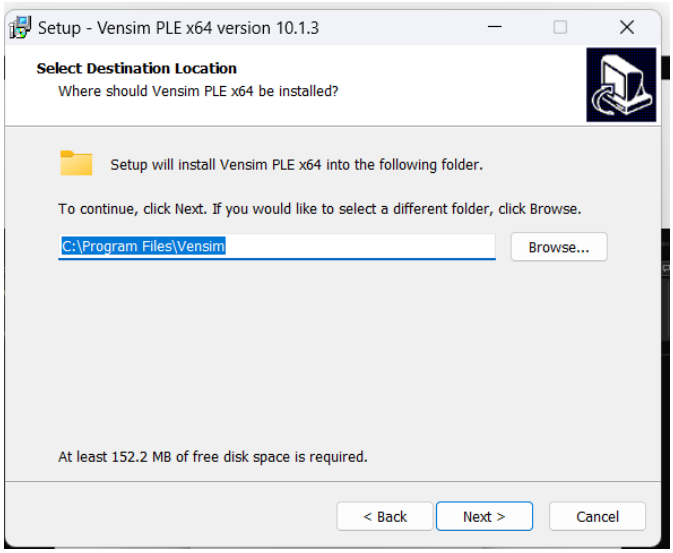
I accept the agreement
I do not accept the agreement

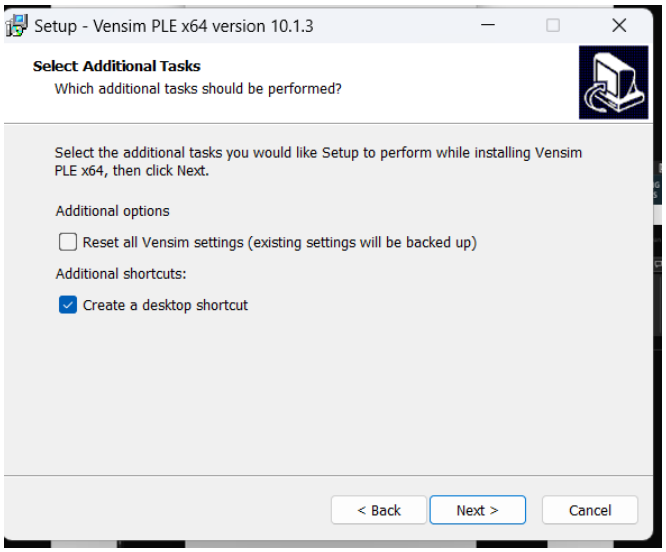
Next > Cancel


Setup - Vensim PLE x64 version 10.1.3
Vensim PLE install options
Please select below the purpose of this installation
Vensim PLE can be installed for one of three purposes.
Install Vensim PLE for evaluation purposes. Use limited to 60 days.
Install PLE for academic, public research or personal use. Commercial, proprietary, classified or operational use not allowed.
Install a licenced version of Vensim PLE.
< Back
Next >
Cancel

4. กรอกชื่อ


Setup - Vensim PLE x64 version 10.1.3
Registration Information
Please enter your registration information below, then click Next.
Name:
yee
Company:
Registration/unlock code:
< Back
Next >
Cancel

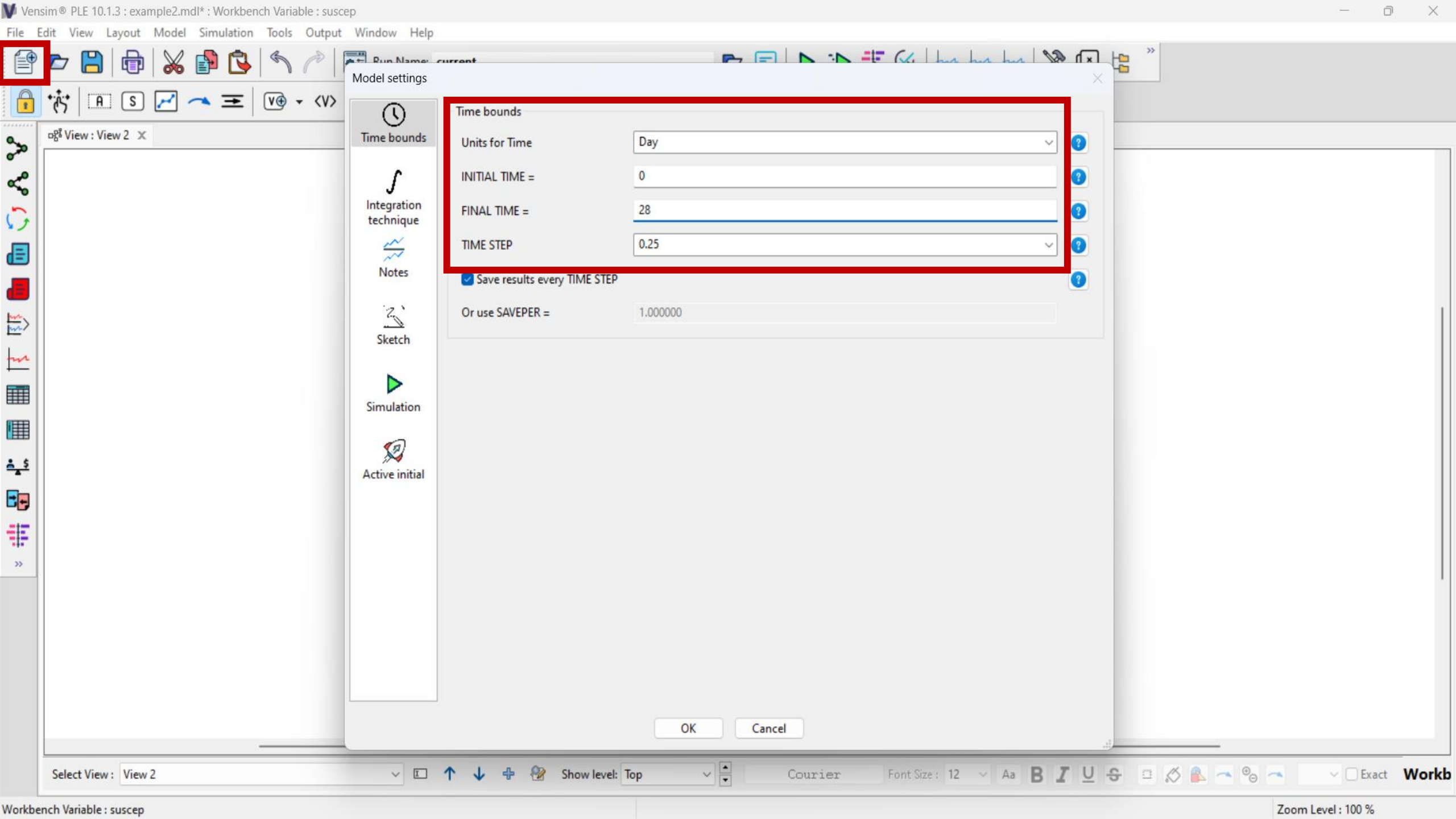

Setup - Vensim PLE x64 version 10.1.3
Select Destination Location
Where should Vensim PLE x64 be installed?
Setup will install Vensim PLE x64 into the following folder.
To continue, click Next. If you would like to select a different folder, click Browse.
C:\Program Files\Vensim
Browse...
At least 152.2 MB of free disk space is required.
< Back
Next >
Cancel


Setup - Vensim PLE x64 version 10.1.3
Select Additional Tasks
Which additional tasks should be performed?
Select the additional tasks you would like Setup to perform while installing Vensim PLE x64, then click Next.
Additional options
Reset all Vensim settings (existing settings will be backed up)
Additional shortcuts:
Create a desktop shortcut
< Back
Next >
Cancel

Vensim PLE
x64

โปรแกรม

vensim


Vensim® PLE 10.1.3 : example2.mdl* : Workbench Variable : suscep
File Edit View Layout Model Simulation Tools Output Window Help
Model settings
Time bounds
Integration technique
Notes
Sketch
Simulation
Active initial
Time bounds
Units for Time
Day
INITIAL TIME =
0
FINAL TIME =
28
TIME STEP
0.25
Save results every TIME STEP
Or use SAVEPER =
1.000000
OK
Cancel
View : View 2
Select View :
View 2
Show level:
Top
Courier
Font Size :
12
Exact
Workb
Workbench Variable : suscep
Zoom Level : 100 %

Vensim® PLE 10.1.3 : New Model : Workbench Variable : Time

File Edit View Layout Model Simulation Tools Output Window Help

Run Name: current

View : View 1

สร้างกล่อง SIR ด้วย ตัว s

จะได้กล่องแบบนี้

Select View : View 1 Show level: Top Courier Font Size : 12 Exact Workb

Workbench Variable : Time

Zoom Level : 100 %

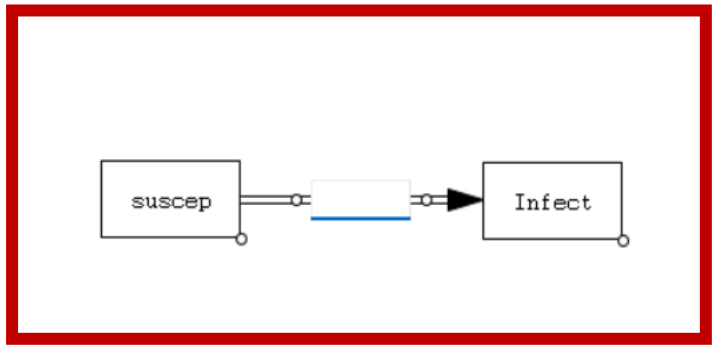
Vensim® PLE 10.1.3 : New Model* : Workbench Variable : Infect

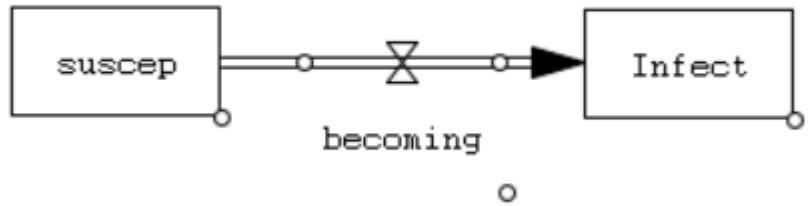
File Edit View Layout Model Simulation Tools Output Window Help

Run Name: current

View : View 1

เชื่อมกล่องด้วยเครื่องหมายนี้





Select View : View 1 Show level: Top Courier Font Size : 12 Top Exact

Workbench Variable : Infect

Zoom Level : 100 %

Vensim® PLE 10.1.3 : New Model* : Workbench Variable : R0

File Edit View Layout Model Simulation Tools Output Window Help

Run Name: current

View : View 1

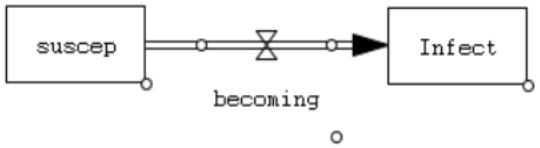
สร้างค่าตัวแปรอื่นๆ เช่น

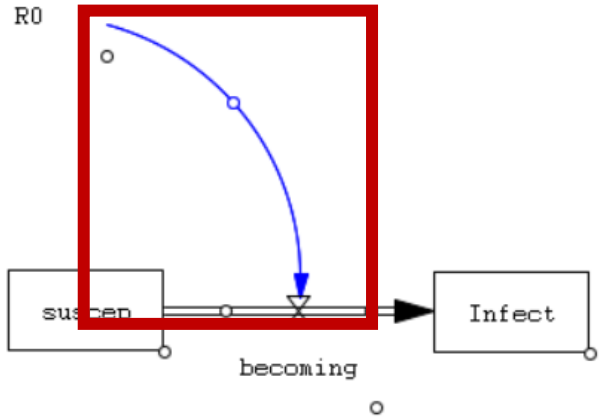
ค่า R0

จำนวนประชากรทั้งหมด

ระยะเวลาฟักตัว

เชื่อมตัวแปรอื่นๆ กับข้อมูลอื่นๆ

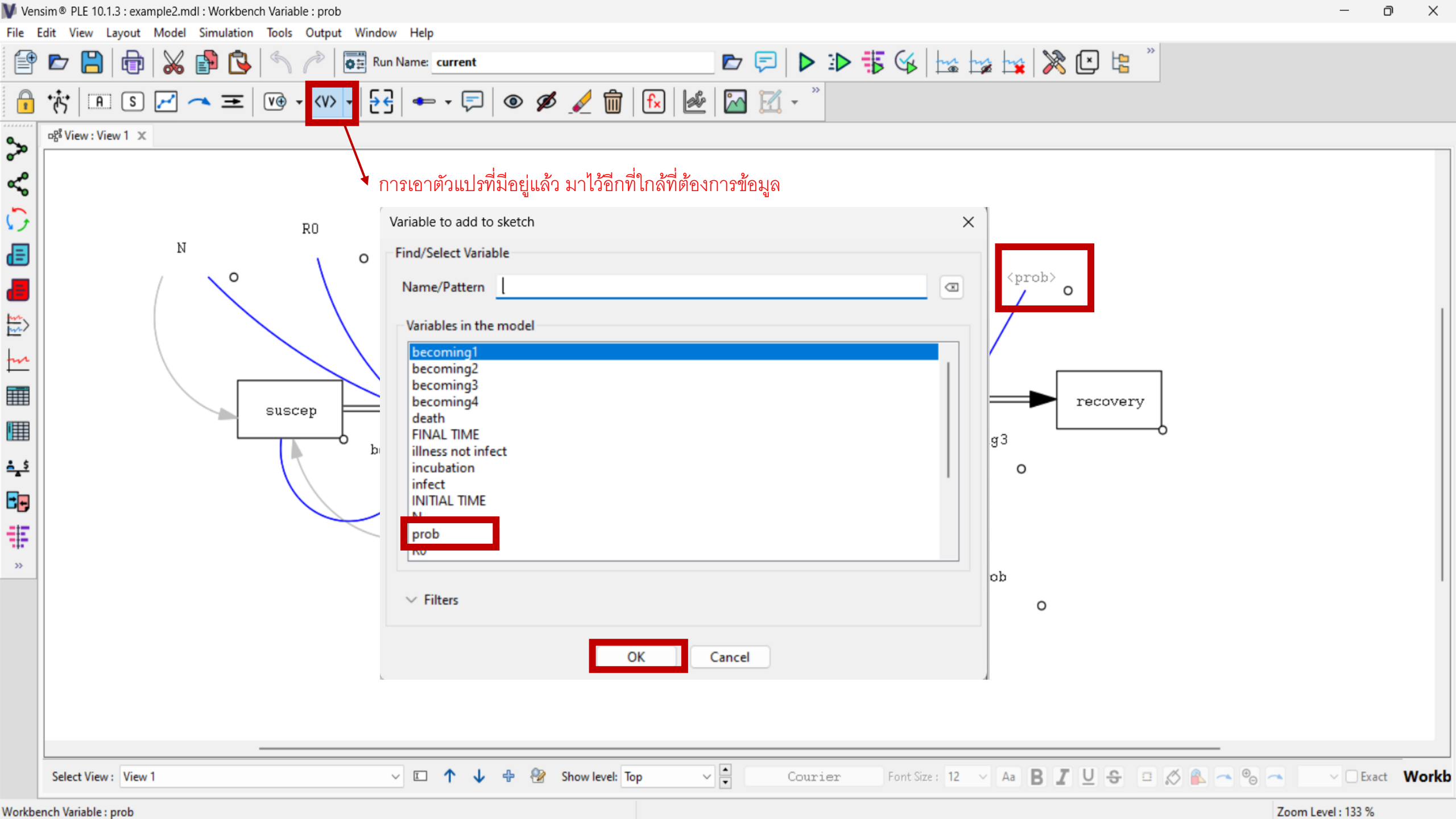


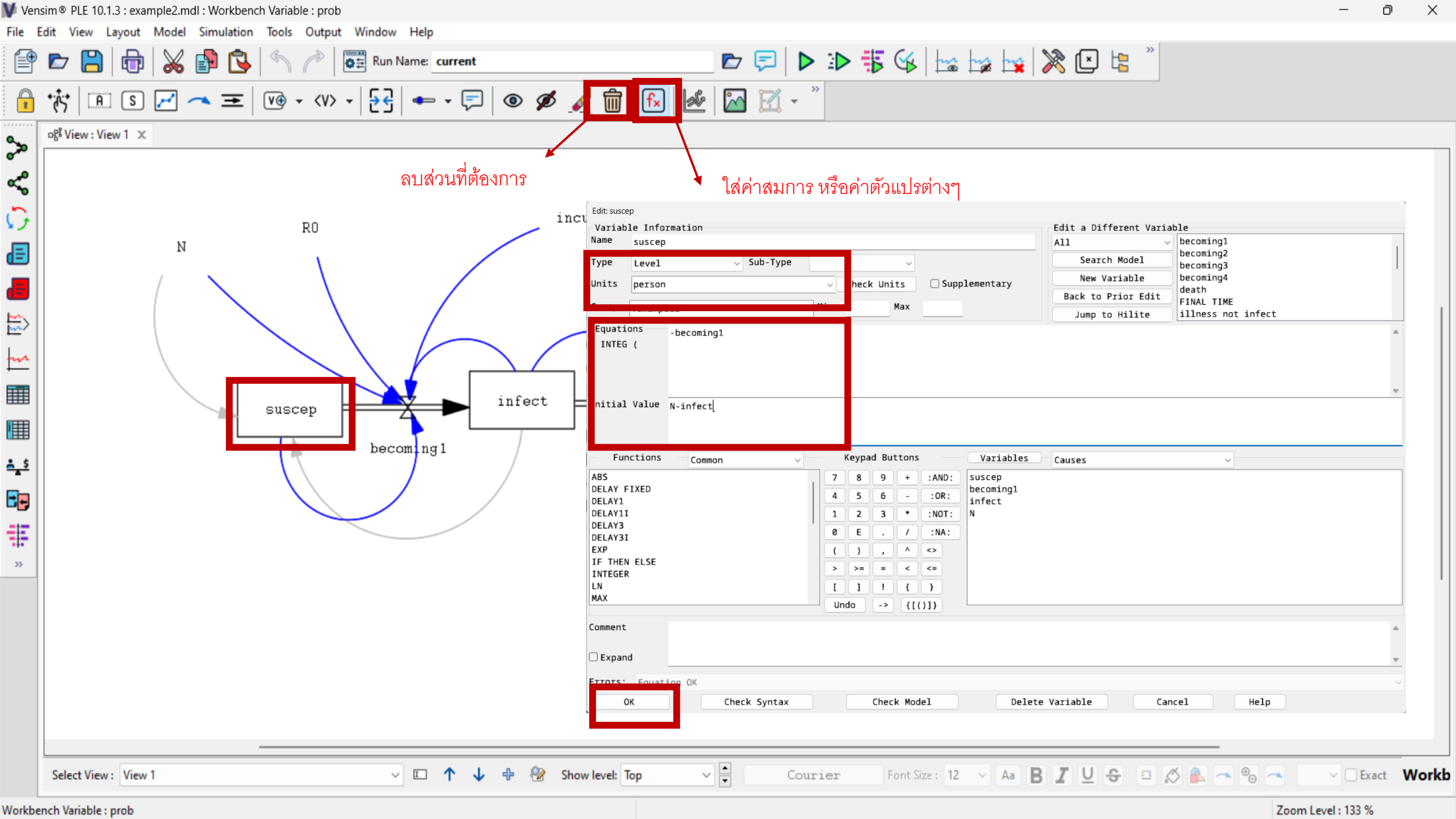


Select View : View 1 Show level: Top Courier Font Size : 12 Top Exact

Workbench Variable : R0

Zoom Level : 100 %


Vensim® PLE 10.1.3 : example2.mdl : Workbench Variable : prob
File Edit View Layout Model Simulation Tools Output Window Help
Run Name: current
การเอาตัวแปรที่มีอยู่แล้ว มาไว้อีกที่ใกล้ที่ต้องการข้อมูล
View : View 1
N
R0
suscep
<prob>
recovery
g3
ob
Variable to add to sketch
Find/Select Variable
Name/Pattern
Variables in the model
becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect
incubation
infect
INITIAL TIME
prob
Filters
OK
Cancel
Select View : View 1
Show level: Top
Courier
Font Size : 12
Exact
Workb
Workbench Variable : prob
Zoom Level : 133 %


Vensim® PLE 10.1.3 : example2.mdl : Workbench Variable : prob
File Edit View Layout Model Simulation Tools Output Window Help
Run Name: current
ลบส่วนที่ต้องการ
ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ
View : View 1
N
R0
suscep
infect
becoming1
Edit: suscep
Variable Information
Name suscep
Type Level
Sub-Type
Units person
Supplementary
Max
Equations
INTEG (
-becoming1
nitial Value
N-infect
Edit a Different Variable
All
Search Model
New Variable
Back to Prior Edit
Jump to Hilite
becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect
Functions
Common
ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX
Keypad Buttons
Variables
Causes
suscep
becoming1
infect
N
Comment
Expand
Errors:
Equation OK
OK
Check Syntax
Check Model
Delete Variable
Cancel
Help
Select View : View 1
Show level: Top
Courier
Font Size : 12
Exact
Workb
Workbench Variable : prob
Zoom Level : 133 %

Edit: N

Variable Information

Name: N

Type: Constant | Sub-Type: Normal

Units: person | Check Units | Supplementary

Group: .example2 | Min | Max | Incr

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

6000

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

Comment

Expand

Errors: Equation OK

OK | Check Syntax | Check Model | Delete Variable | Cancel | Help

Edit: becoming1

Variable Information

Name: becoming1

Type: Auxiliary | Sub-Type: Normal

Units: person/Day | Check Units | Supplementary

Group: .example2 | Min | Max

Edit a Different Variable

All

Search Model | New Variable | Back to Prior Edit | Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

(suscep*infect/N)*R0/incubation

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

incubation
infect
N
R0
suscep

Comment

Expand

Errors: Equation OK

OK | Check Syntax | Check Model | Delete Variable | Cancel | Help

Edit: R0

Variable Information

Name R0

Type Constant Sub-Type Normal

Units Dmnl Check Units Supplementary

Group .example2 Min Max Incr

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

3.44

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables Causes

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: infect

Variable Information

Name: infect

Type: Level Sub-Type:

Units: person/Day Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= INTEG (

becoming1-becoming2

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Initial Value

1000

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}

Undo -> {[()]}

Variables: Causes

infect
becoming1
becoming2

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: incubation

Variable Information

Name: incubation

Type: Constant Sub-Type: Normal

Units: Day Check Units Supplementary

Group: .example2 Min: Max: Incr:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

6

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: illness not infect

Variable Information

Name: illness not infect

Type: Level Sub-Type:

Units: person/Day Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= INTEG (

becoming2-becoming3-becoming4

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Initial Value

0

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

illness not infect
becoming2
becoming3
becoming4

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: becoming2

Variable Information

Name: becoming2

Type: Auxiliary Sub-Type: Normal

Units: person/Day Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

infect/incubation

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

incubation
infect

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: Rx recover

Variable Information

Name: Rx recover

Type: Constant | Sub-Type: Normal

Units: Day | Check Units | Supplementary

Group: .example2 | Min | Max | Incr

Edit a Different Variable

All

Search Model | New Variable | Back to Prior Edit | Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= 3

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

Comment

Expand

Errors: Equation OK

OK | Check Syntax | Check Model | Delete Variable | Cancel | Help

Workbench variable: Rx recover

Edit: becoming3

Variable Information

Name: becoming3

Type: Auxiliary Sub-Type: Normal

Units: person/Day Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

illness not infect*(1-prob)/Rx recover

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

illness not infect
prob
Rx recover

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: becoming4

Variable Information

Name: becoming4

Type: Auxiliary Sub-Type: Normal

Units: person/Day Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

illness not infect*prob/Rx death

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

illness not infect
prob
Rx death

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: Rx death

Variable Information

Name: Rx death

Type: Constant | Sub-Type: Normal

Units: Day | Check Units | Supplementary

Group: .example2 | Min | Max | Incr

Edit a Different Variable

All

Search Model | New Variable | Back to Prior Edit | Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= 7

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

Comment

Expand

Errors: Equation OK

OK | Check Syntax | Check Model | Delete Variable | Cancel | Help

Edit: prob

Variable Information

Name: prob

Type: Constant Sub-Type: Normal

Units: 1 Check Units Supplementary

Group: .example2 Min: Max: Incr:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

=

0.08

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Functions Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables Causes

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Edit: death

Variable Information

Name: death

Type: Level Sub-Type:

Units: person Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= INTEG (

becoming4

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Initial Value

0

Functions: Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables: Causes

death
becoming4

Comment

Expand

Errors: Equation OK

OK Check Syntax Check Model Delete Variable Cancel Help

Workbench variable : death

Edit: recovery

Variable Information

Name: recovery

Type: Level Sub-Type:

Units: person Check Units Supplementary

Group: .example2 Min: Max:

Edit a Different Variable

All

Search Model

New Variable

Back to Prior Edit

Jump to Hilite

becoming1
becoming2
becoming3
becoming4
death
FINAL TIME
illness not infect

Equations

= INTEG (

becoming3

ใส่ค่าสมการ หรือค่าตัวแปรต่างๆ

Initial Value

0

Functions

Common

ABS
DELAY FIXED
DELAY1
DELAY1I
DELAY3
DELAY3I
EXP
IF THEN ELSE
INTEGER
LN
MAX

Keypad Buttons

7	8	9	+	:AND:
4	5	6	-	:OR:
1	2	3	*	:NOT:
0	E	.	/	:NA:
(	)	,	^	<>
>	>=	=	<	<=
[	]	!	{	}
Undo	->	{[()]}		

Variables

Causes

recovery
becoming3

Comment

Expand

Errors: Equation OK

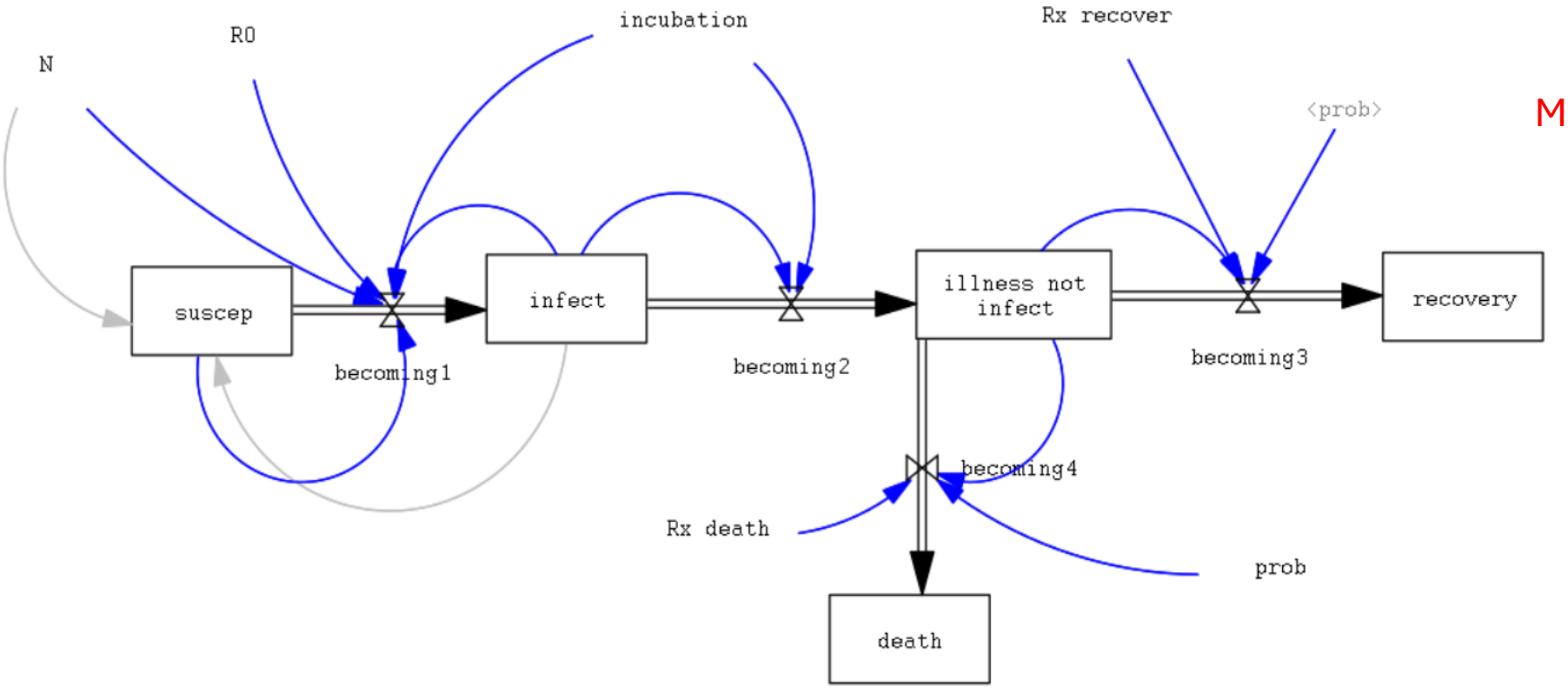
OK Check Syntax Check Model Delete Variable Cancel Help

Vensim® PLE 10.1.3 : example2.mdl : Workbench Variable : recovery

File Edit View Layout Model Simulation Tools Output Window Help

Run Name: current

View : View 1

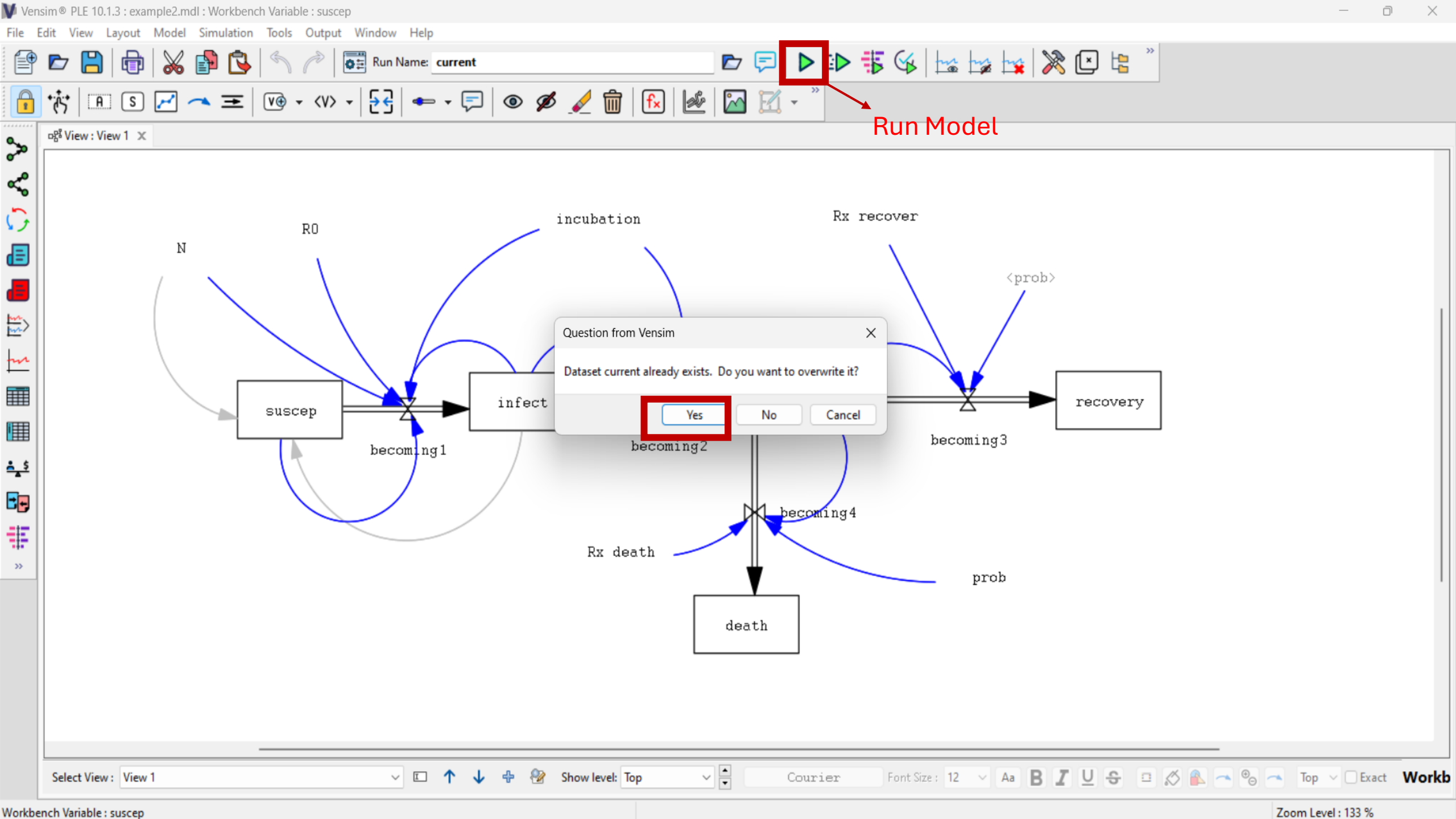


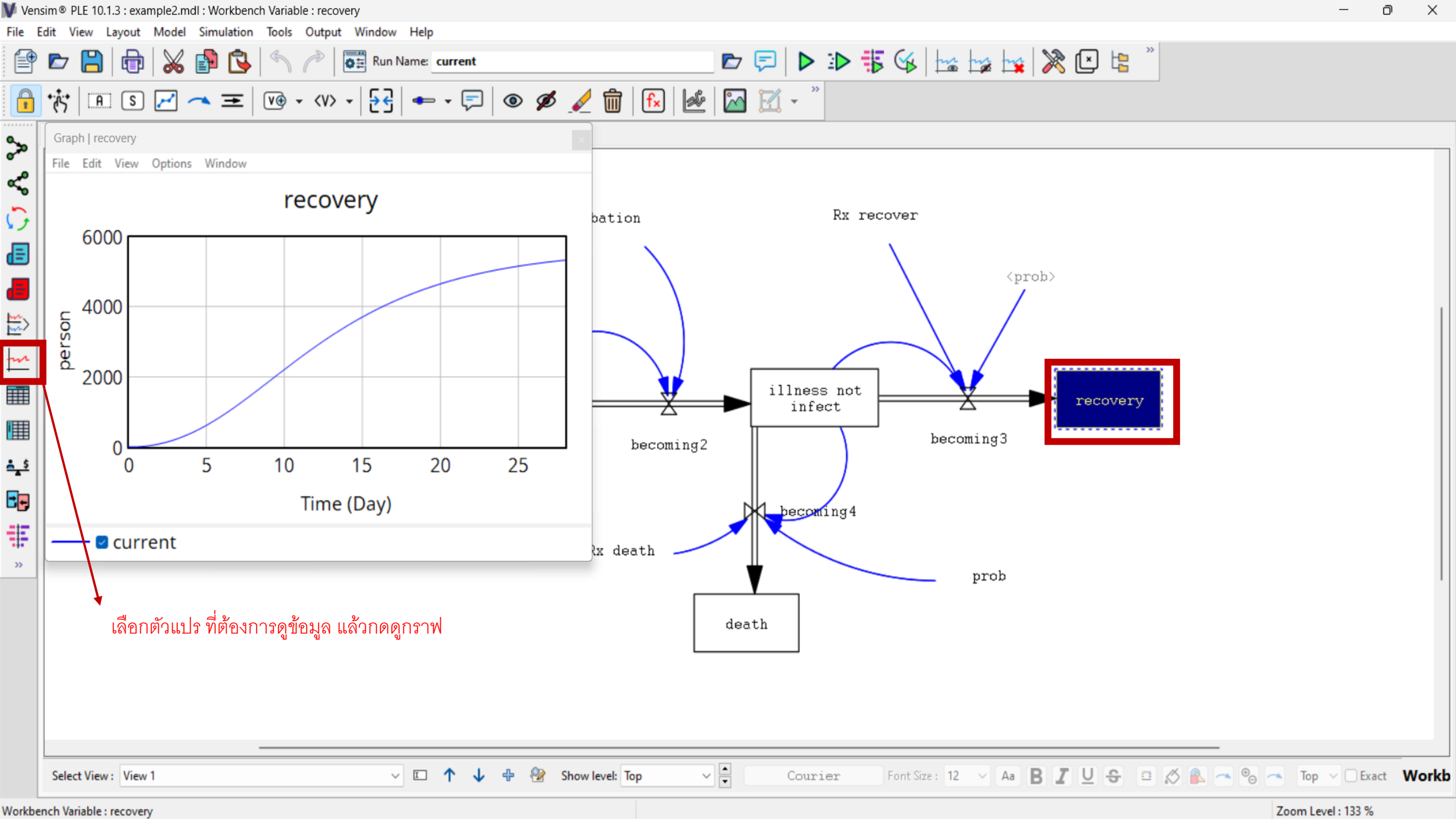
Model ที่เสร็จเรียบร้อย

Select View : View 1 Show level: Top Courier Font Size : 12 Top Exact Workb

Workbench Variable : recovery

Zoom Level : 133 %


Vensim® PLE 10.1.3 : example2.mdl : Workbench Variable : suscep
File Edit View Layout Model Simulation Tools Output Window Help
Run Name: current
Run Model
View : View 1
N
R0
incubation
Rx recover
<prob>
suscep
becoming1
infect
becoming2
becoming3
recovery
becoming4
Rx death
prob
death
Question from Vensim
Dataset current already exists. Do you want to overwrite it?
Yes
No
Cancel
Select View : View 1
Show level: Top
Courier
Font Size : 12
Top
Exact
Workb
Workbench Variable : suscep
Zoom Level : 133 %


Vensim® PLE 10.1.3 : example2.mdl : Workbench Variable : recovery
File Edit View Layout Model Simulation Tools Output Window Help
Run Name: current
Graph | recovery
File Edit View Options Window
recovery
6000
4000
2000
0
person
0
5
10
15
20
25
Time (Day)
current
bation
Rx recover
<prob>
illness not infect
recovery
becoming2
becoming3
becoming4
Rx death
prob
death
เลือกตัวแปร ที่ต้องการดูข้อมูล แล้วกดดูกราฟ
Select View : View 1
Show level: Top
Courier
Font Size : 12
Top
Exact
Workb
Workbench Variable : recovery
Zoom Level : 133 %

ภาคผนวกที่ 4

Time series analysis and forecasting in health science and epidemiology

Time series analysis and forecasting in health science and epidemiology

Rapeepong Suphanchaimat
Division of Epidemiology, Department of Disease Control and
International Health Policy Program,
Ministry of Public Health, Thailand

2024

Presentation outline

- Basic concepts
- Recap on correlation and linear regression
- Autocorrelated disturbances: how to deal with it
- Some useful techniques
  - Smoothing
  - ARIMA
  - etc

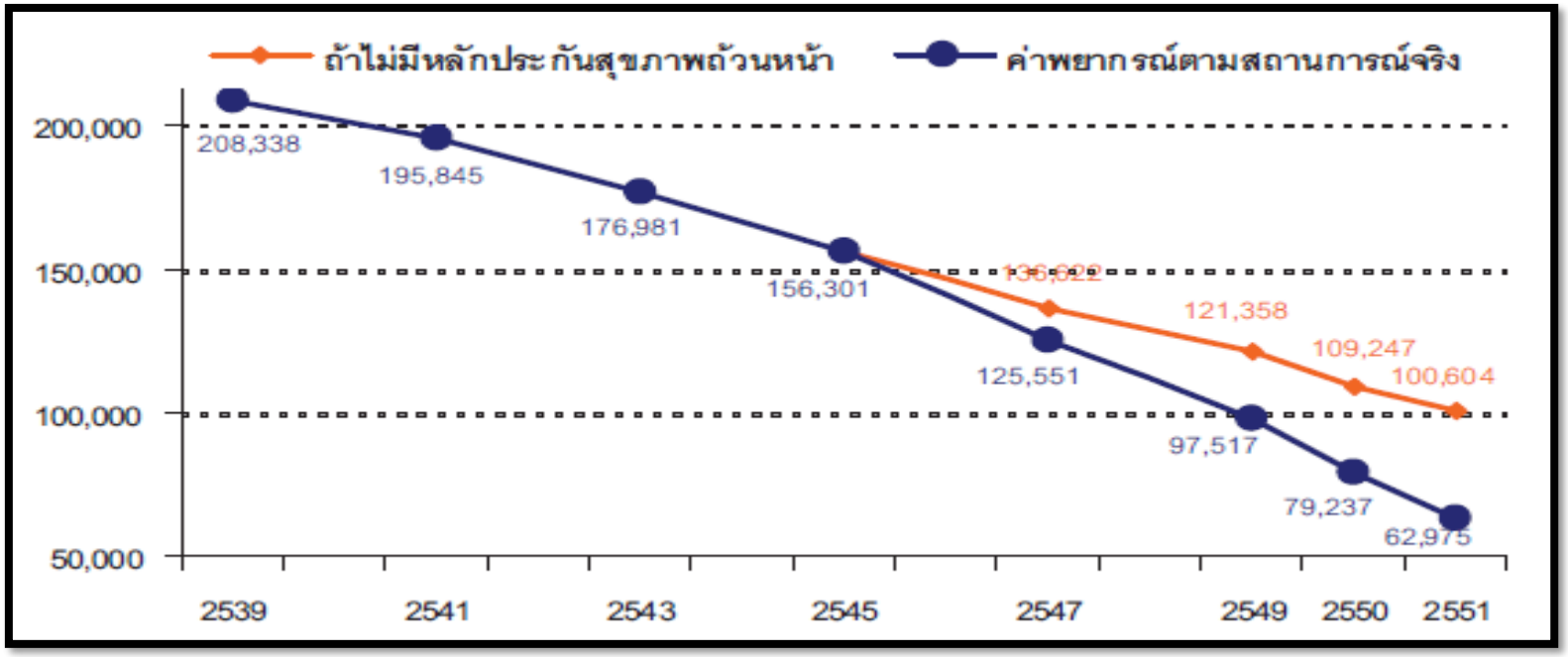
BASIC CONCEPTS

What are time series data?

- Panel data
  - Longitudinal multidimensional data involving measurements over time on the same individuals
- Cross sectional data
  - Unidimensional on time (but possibly covering groups of populations)
- Time series data
  - Unidimensional on panel on a group of individuals (panel members) through multiple time points

Time series analysis' usefulness

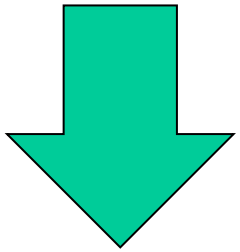
- Better description
- Forecasting
- Explanative
- Measure of (policy) impact



Source: Limwattananon (2011)

Relation to epidemiology

- Objectives/aims of epidemiological study
  - To describe (situation)
  - To explain (risk factors)
  - To predict (any adverse events)



  - To implement control measures

What is forecasting?

- Time series forecasting
  - The use of a model to predict future values based on previously observed values.
- Common assumptions behind forecasting
  - Past events predict the future (and current) events.
  - Series = fit (signal) + residuals
  - Aggregate forecasts tend to be more accurate than those for individual forecasts.
  - Forecast accuracy decreases as the time horizon (the period covered) increases.
  - Choosing forecasting horizon depends on goals of organization

Types of forecasting

- Classification by research methods
  - Qualitative
  - Quantitative
  - Mixed
- Classification by time
  - Multiple steps ahead
  - One step ahead
    - Static
    - Dynamic
- Classification by sets of data involved
  - Univariate
  - Multivariate

Forecasting Methods

Qualitative Methods

- Expert Judgement
- Jury of Executive Opinion: JEO
- Delphi method

Mixed Methods

- Risk analysis method
- Expert elicitation

Quantitative Methods

Empirical/Statistical Model

- Simple projection/multiplication
- Time series analysis: Decomposition, Smoothing, ARIMA
- Regression model: Linear, Poisson, Age-Period-Cohort Model, etc.

Process/Mechanistic Model

- Compartmental (multi-stage) model e.g. SIR model for Infectious diseases, Markov Chain, etc.

Before forecasting, we need to think of…

Disease Importance

- Frequency
- Severity
- Impact (QOL, Social, economic)
- Public/political concern

Modifiability

- Preventable disease
- Available effective intervention
- Acceptable cost

Predictability

- Available data
- Available technique
- Available technology
- Available experts and experiences (both content and technique)

Elements of good forecast

- Compatible
- Meaningful
- Useful time horizon
- Reliable
- Accurate
- Easy to understand and use

Measurements for forecast accuracy (1)

Error (E_t) = Actual value (A_t) – Forecasted value (F_t)

1) Mean Error (ME)

$$= \frac{\sum (E_t)}{n}$$

2) Mean Absolute Error (MAE)

$$= \frac{\sum |E_t|}{n}$$

3) Mean Percent Error (MPE)

$$= \frac{\sum \left(\frac{E_t}{A_t} \right)}{n}$$

4) Mean Absolute Percent Error (MAPE)

$$= \frac{\sum \left| \frac{E_t}{A_t} \right|}{n}$$

Measurements for forecast accuracy (2)

5) Mean Squared Error (MSE)

$$= \frac{\sum (E_t)^2}{n}$$

6) Root-Mean Squared Error (RMSE)

$$= \sqrt{\frac{\sum (E_t)^2}{n}}$$

7) Theil's U

$$= \frac{\sqrt{\sum (A_t - F_t)^2}}{\sqrt{\sum (A_t - A_{t-1})^2}}$$

Sometimes called uncertainty coefficient

Note: Other goodness of fit tests such as AIC and BIC are also useful.

Recap on correlation and regression

Correlation and covariance (1)

- Covariance
  - The covariance between two jointly distributed real-valued random variables X and Y is defined as the expected product of their deviations from their individual expected values.
  - $Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$

- Correlation
  - Like a standardized form of covariance
  - $Corr(X,Y) = E[(X-E[X])(Y-E[Y])]/(\sigma_x\sigma_y)$

Correlation and covariance (2)

BASIS FOR COMPARISON	COVARIANCE	CORRELATION
Meaning	Covariance is a measure indicating the extent to which two random variables change in tandem.	Correlation is a statistical measure that indicates how strongly two variables are related.
What is it?	Measure of correlation	Scaled version of covariance
Values	Lie between $-\infty$ and $+\infty$	Lie between -1 and +1
Change in scale	Affects covariance	Does not affect correlation
Unit free measure	No	Yes

Correlation direction and strength (1)

Perfect positive correlation $r = 1.0$

Strong positive correlation $r = 0.9$

Moderate positive correlation $r = 0.5$

Perfect negative correlation $r = -1.0$

Strong negative correlation $r = -0.9$

Moderate negative correlation $r = -0.5$

Correlation direction and strength (2)

- Denoted as ‘r’
- Direction of association:
  - Positive (up-up, down-down)
  - Negative (up-down, down-up)
  - ~ 0 (independent)
- Strength of association
  - close to 1 or –1 $\Rightarrow$ strong
  - close to 0 $\Rightarrow$ weak
  - guidelines
    - if $|r| \geq .7 \Rightarrow$ say ‘strong’
    - if $|r| \leq .3 \Rightarrow$ say ‘weak’

Linear regression

Regression line equation:

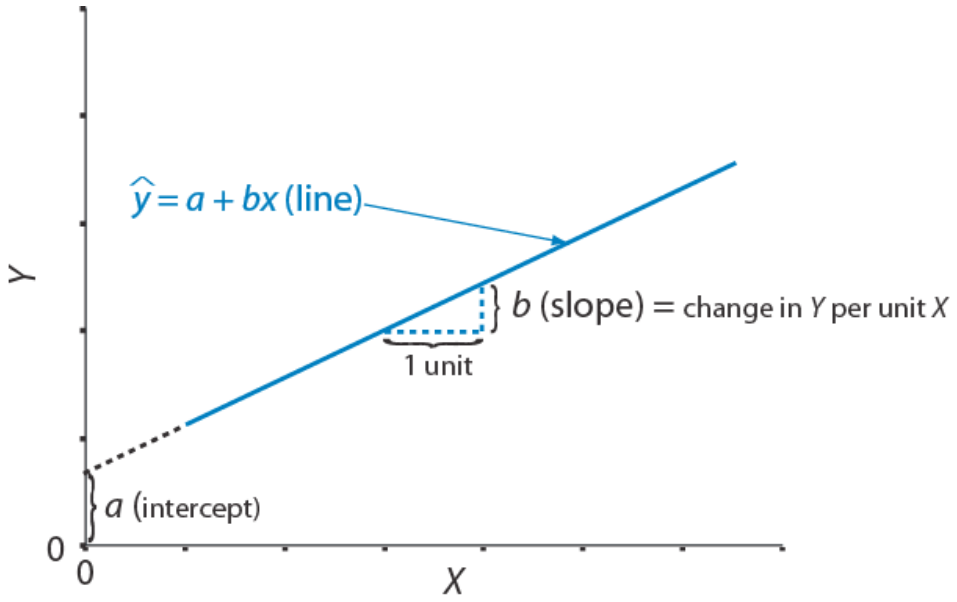
$$\hat{y}_i = \beta_0 + \beta_1 \cdot x_i$$

where

$\hat{y}_i \equiv$ predicted value of Y at x_i

$\beta_0 \equiv$ intercept coefficient
(value of y when x=0)

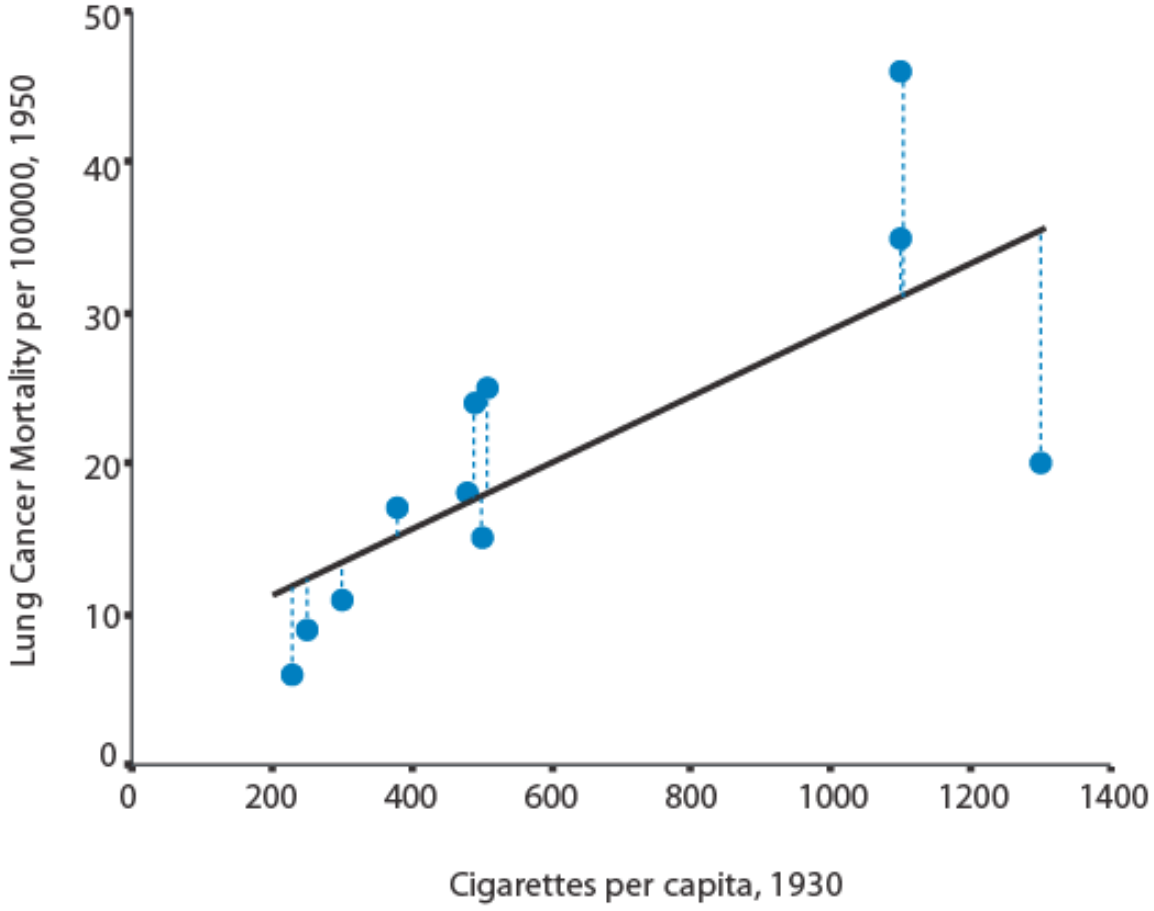
β_1 = slope coefficient
(change in value of y for a one-unit increase of x)



Regression Model: $Y_i = \beta_0 + \beta_1 X_i + e_i$

Least squares line

Residual = distance of data point from regression line (dotted)



The best fitting line is the line that minimizes the residuals (least square approach).

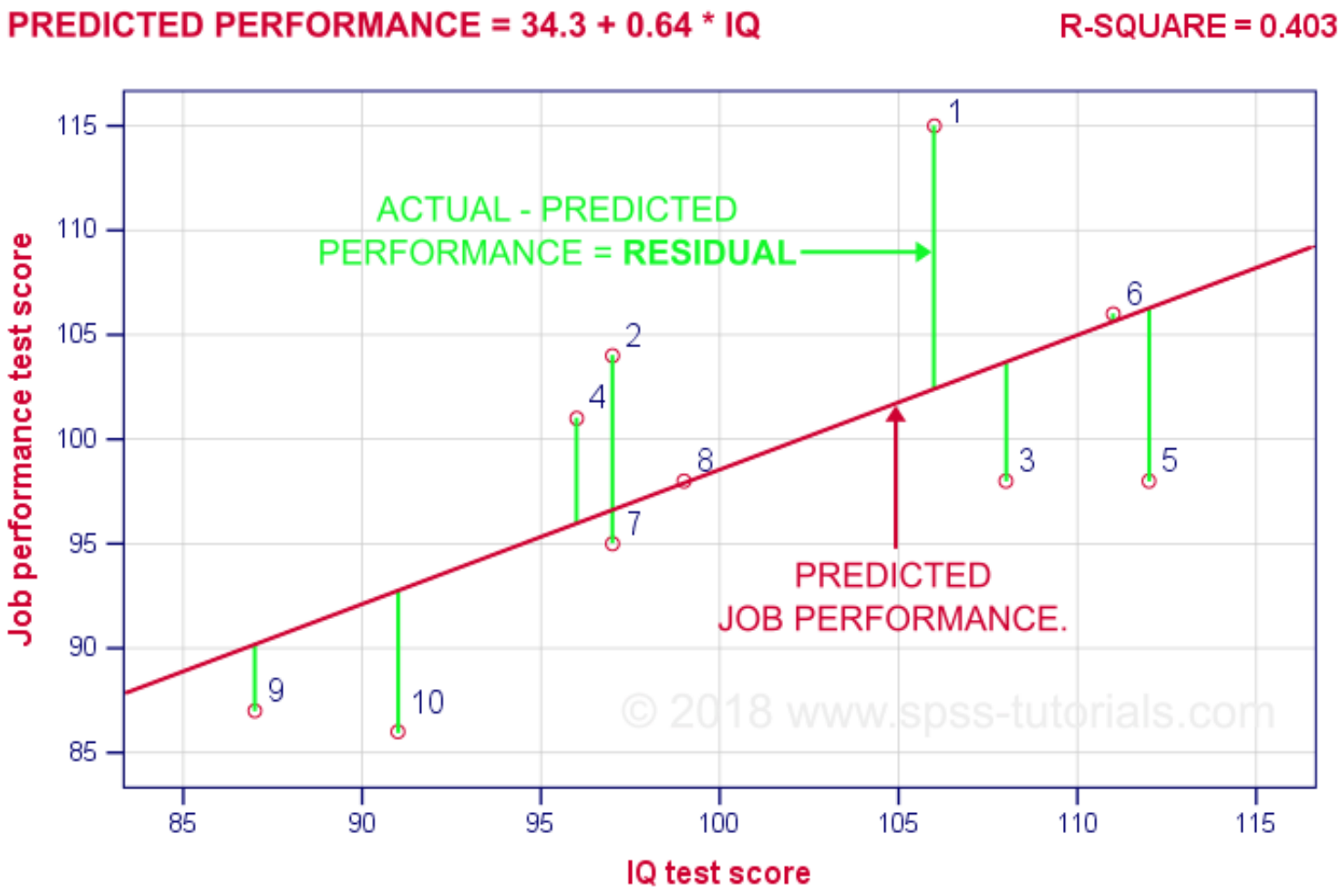
'Ordinary least squares' (OLS)

R^2 = coefficient of determination ➔ To what extent (%) the variance in the response variable is accounted by the model.

Assumption of ordinary least square

1. Exogeneity—Regressor is independent of the errors.
2. Linearity—Mean of the response variable is a linear combination of the parameters (regression coefficients) and the predictor variables.
3. Independence of errors: There is no autocorrelation of error terms.
4. Homoscedasticity: Constant variance of errors is assumed.
5. Normal distribution of errors
6. No perfect multicollinearity

What residuals tell us in regression line

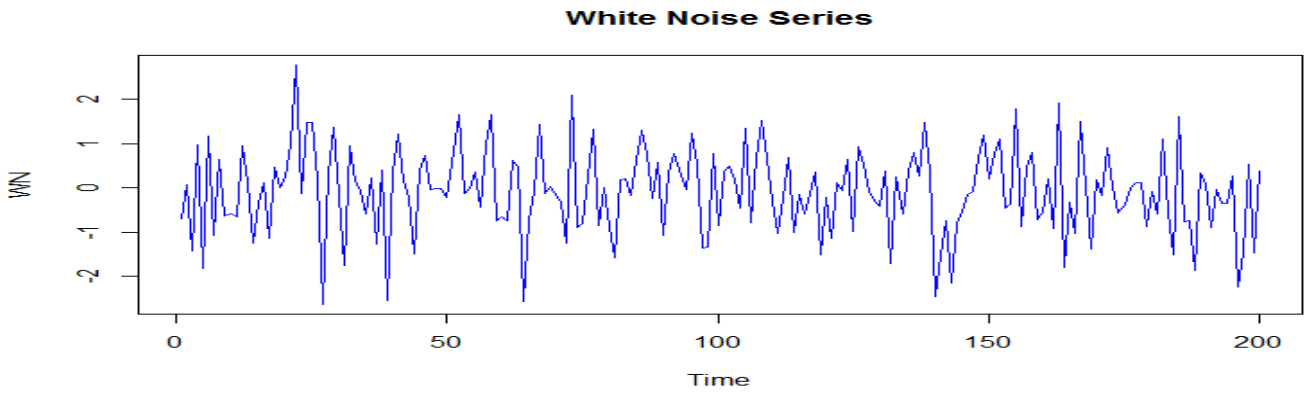


What are residuals?

- Assumed to be independent and identically distributed (i.i.d) if fulfilling all the following criteria
  - $E(e_t) = 0$
  - $E(e_t^2) = \sigma^2$
  - $E(e_t e_h) = 0$ for $\forall t \neq h$
- Other synonyms
  - Errors
  - white noises
  - Innovations
  - Disturbances

White noise

- White noise is an i.i.d. (independent and identically distributed).
- Can be checked by various techniques, eg
  - Autocorrelogram
  - Ljung Box test (portmanteau test)
  - Box–Pierce test



Fundamentals of forecast

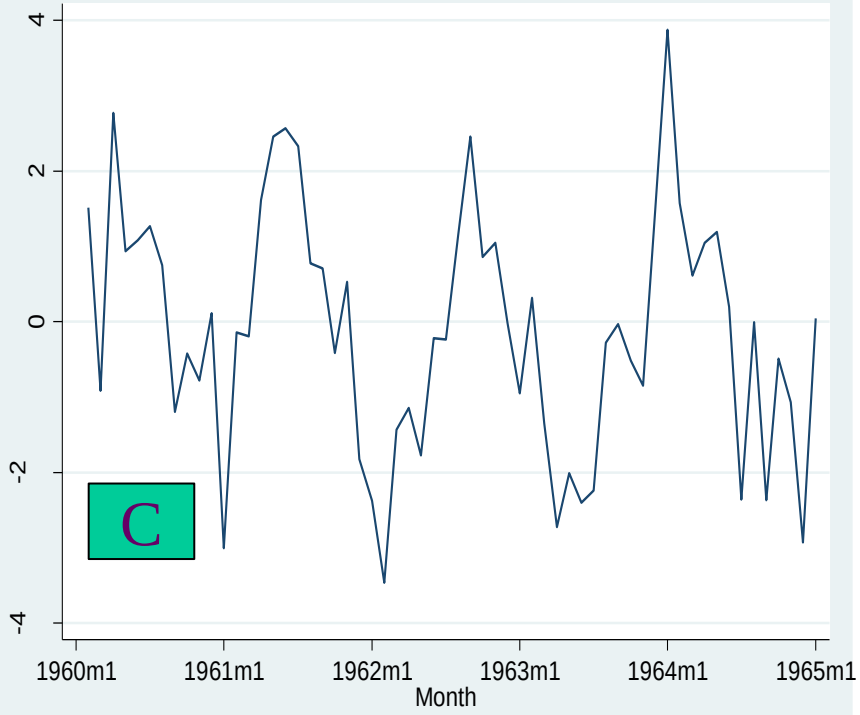
Common components of time series data

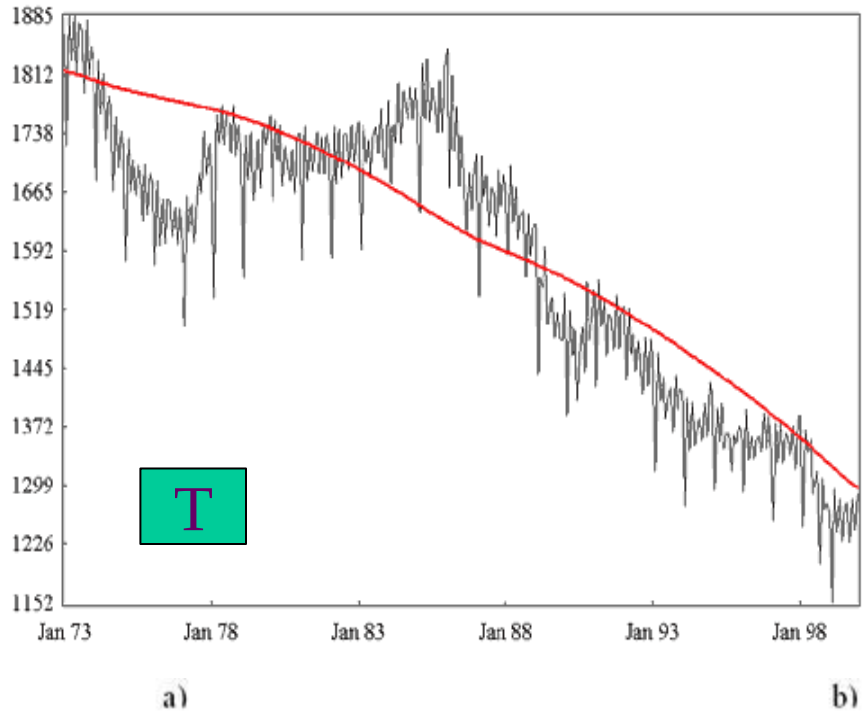
- Trend (T)
  - Persistent systematic tendency for a series to increase or decrease
- Cycle (C)
  - Oscillating around a trend if the series incorporates a trend or around the mean if the series has no trend
  - No need to regular periodic oscillating
- Seasonal (S)
  - Special case of cycle
  - Predictable and periodic
- Random or irregular components (shock) (I/R)

Trend

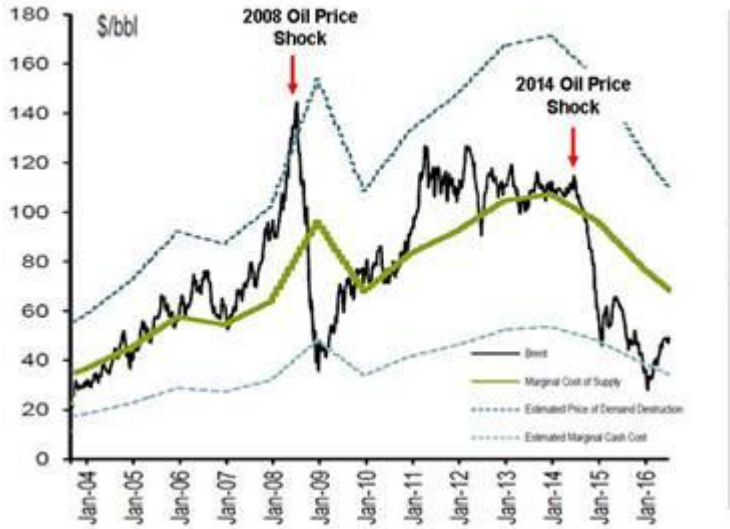
- Deterministic trend
  - $y_t = f(t) + \varepsilon$ ➔ $E(y_t) = \alpha t$; $Var(y_t) = \sigma^2$
  - Smoother is used to estimate f(t).
  - Trend stationary: The mean trend is deterministic. Once the trend is estimated and removed from the data, the residual series is a stationary process.
  - Detrending approach

- Stochastic trend (random walk with drift)
  - $y_t = \alpha + y_{t-1} + \varepsilon_t$ ➔ $E(y_t) = \alpha t$; $Var(y_t) = t\sigma^2$
  - Differencing stationary: Differencing the series *D* times until yielding a stationary process.
  - Differencing approach


C
4
2
0
-2
-4
1960m1
1961m1
1962m1
1963m1
1964m1
1965m1
Month

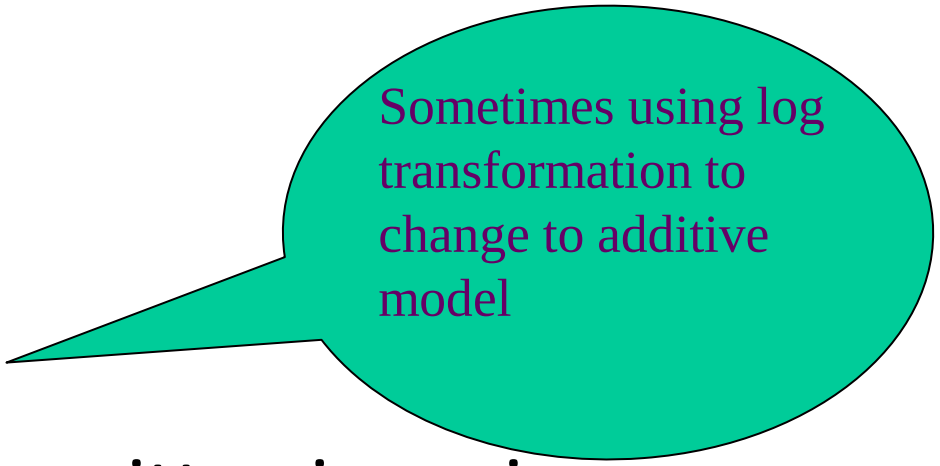

T
1885
1812
1738
1665
1592
1519
1445
1372
1299
1226
1152
Jan 73
Jan 78
Jan 83
Jan 88
Jan 93
Jan 98
a)
b)

Decomposition of airline passengers
Data
Trend component
Seasonal component
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2004
2006
2008
2010
2012
2014
Time
S


$/bbl
2008 Oil Price Shock
2014 Oil Price Shock
180
160
140
120
100
80
60
40
20
0
Jan-04
Jan-05
Jan-06
Jan-07
Jan-08
Jan-09
Jan-10
Jan-11
Jan-12
Jan-13
Jan-14
Jan-15
Jan-16
R

Patterns of time series model

- Additive model
  - Y = T + S + C + R
- Multiplicative model
  - Y = T x S x C x R
- Mixed model (can be written in various forms)
  - Y = T + (S x C x R)
  - Y = (T x C) + (S x R)
  - etc



Smoothing by various smoothers

Smoothing techniques

- Smoothers provide tools for filtering out the noise component in a series, making it easier to isolate the systemic component.

- Most of times, we focus on trend and seasonal effects.

- The smoother does not completely eliminate the residual component from the smoothed series, but it does reduce the variance of the residual component, making the fit more prominent.

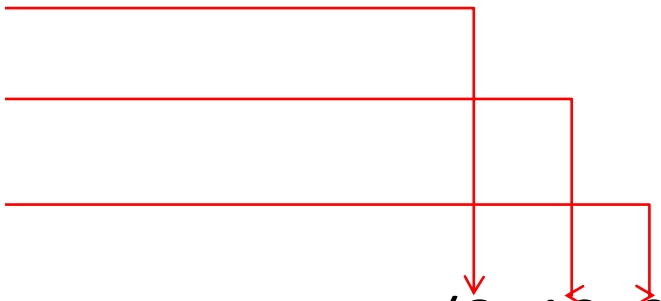
Guide for selecting smoothers

	Season (-)	Season (+)
Trend (-)	• Simple moving average • Weighted moving average (including Hanning) • Non linear filter • Simple exponential	• Holt-Winters without trend
Trend (+)	• Holt-Winters (non-seasonal) • Double exponential	• Holt-Winters with seasonal effect

Example of moving average

$$Moving\ \ Average\ \ (F_t) = \frac{\sum A_t\ \ in\ \ previous\ \ n\ \ periods}{n}$$

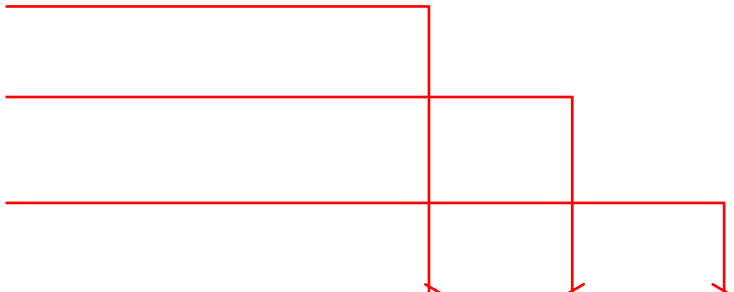
Week	No of cases	Three-week moving average
1	8	
2	10	
3	9	
4	11	(8+10+9)/3 = 9
5	10	(10+9+11)/3 = 10
6	13	(9+11+10)/3 = 10



Example of weighted moving average

$$Weighted\ \ Moving\ \ Average\ (F_t) = \frac{\sum (weight\ \ for\ \ period\ \ n)(Y_t\ \ in\ \ period\ \ n)}{\sum weights}$$

Week	No of cases	Three-week moving average
1	8	
2	10	
3	9	
4	11	(1*8+2*10+3*9)/6 = 9.17
5	10	(1*10+2*9+3*11)/6 = 10.17
6	13	(1*9+2*11+3*10)/6 = 10.17



(double) Exponential (weighted) moving average

- Can be written as
  - $y^*_t = \alpha y_t + (1- \alpha)\, y^*_{t-1}$
- α is called smoothing parameter, ranging from 0 to 1.
- α near 0 → relatively smooth
- α near 1 → relatively volatile
- Can be written in a generalized form as
  - $y^*_t = \alpha y_t + \alpha(1- \alpha)\, y_{t-1} + \alpha\,(1- \alpha)^2 y_{t-2} + \ldots\, \alpha(1- \alpha)^t\, y^*_0$
  - Where y^*_0 = mean of the series at time 0
- This means that for a reasonable long time horizon, the results will be affected by the setting of α than of $y^*_{0.}$
- Double exponential smoothing applies exponential smoothing twice.
  - $Y^{**}_t = \alpha y^*_t + (1- \alpha)\, y^{**}_{t-1}$

Holt-Winters methods (seasonal)

Additive

$$\hat{Y}_{t+n} = a + bn + c_{t+n-s}$$

$$a_t = \alpha(Y_t - c_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

$$c_t = \gamma(Y_t - a_{t+1}) + (1-\gamma)c_{t-s}$$

Estimated mean over t period

Trend component

Seasonal component

Multiplicative

$$\hat{Y}_{t+n} = (a + bn)c_{t+n-s}$$

$$a_t = \alpha\frac{Y_t}{c_{t-s}} + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

$$c_t = \gamma\left(\frac{Y_t}{a_t}\right) + (1-\gamma)c_{t-s}$$

Estimated mean over t period

Trend component

Seasonal component

Value 1 gives more weight to recent data; value close to 0 gives more weight to distant data

Holt-Winters methods (seasonal)

- Alpha captures average value of the dataset. Beta captures trend. Gamma captures season.
- Value 1 gives more weight to recent data; value close to 0 gives more weight to distant data.
- The Additive structure or model is used when the seasonal pattern of data has the same magnitude or is consistent throughout, while the Multiplicative structure or model is used if the magnitude of the seasonal pattern of the data increases over time.

ARIMA model

Autoregressive integrated moving average

- An **autoregressive integrated moving average** (**ARIMA**) model is a generalization of an autoregressive moving average (ARMA) model. To better comprehend the data or to forecast upcoming series points, both of these models are fitted to time series data.

- ARIMA models are applied in some cases where data show evidence of non-stationarity in the sense of mean (but not variance/autocovariance), where an initial differencing step (corresponding to the "integrated" part of the model) can be applied one or more times to eliminate the non-stationarity of the mean function (i.e., the trend).

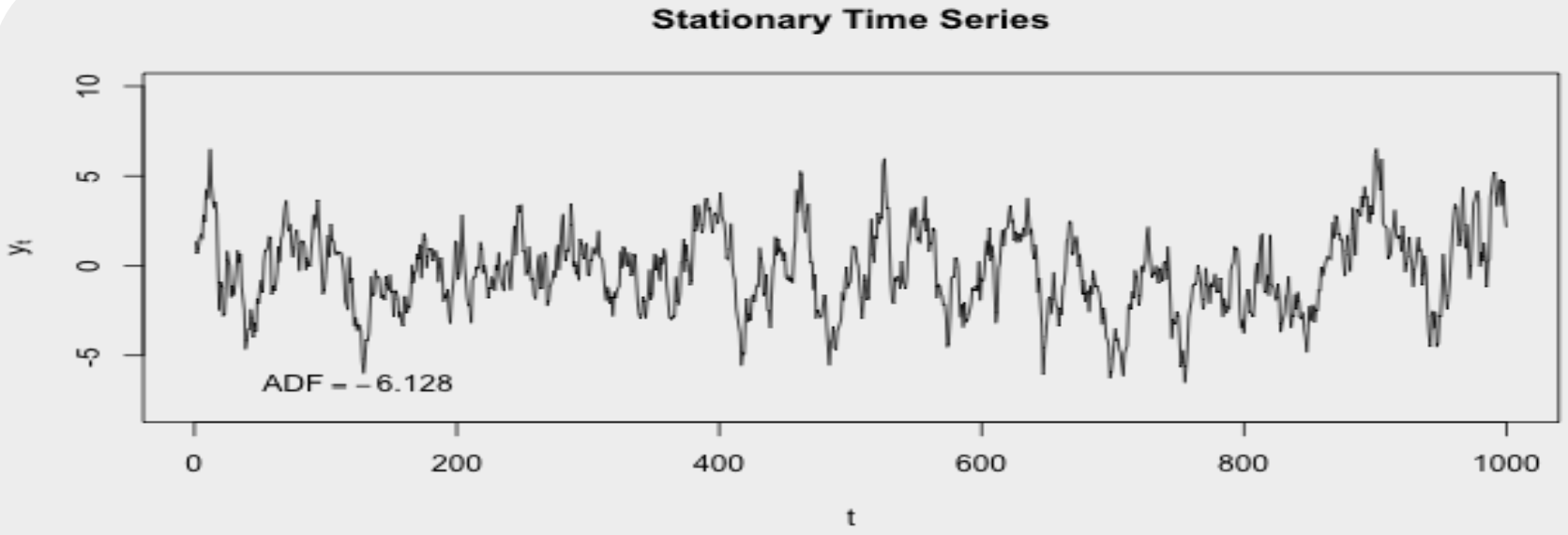
- When the seasonality shows in a time series, the seasonal-differencing could be applied to eliminate the seasonal component.

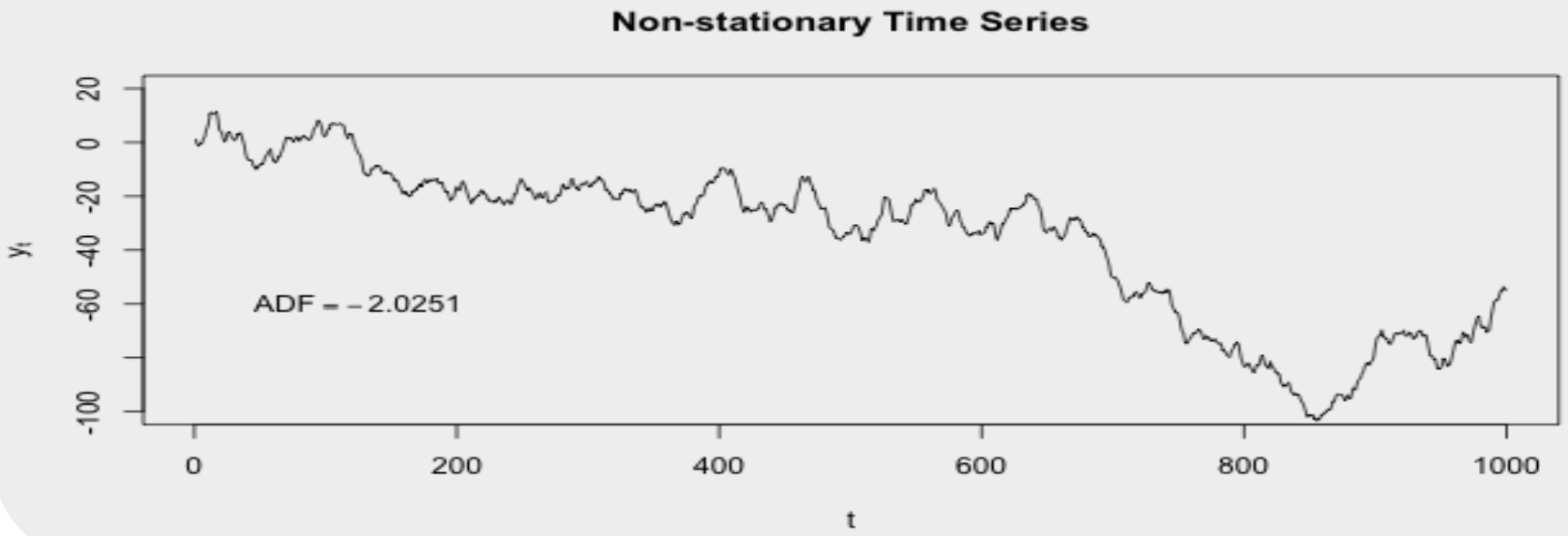
Lag operator

- Consider lag operator
  - $Le_t = e_{t-1}$
  - $L(Le_t) = Le_{t-1} = e_{t-2} = L^2 e_t$
- Lag operator can be viewed as an ordinary algebraic variable, and this allows us to specify the time series model more conveniently.
- Suppose; $y_t = \mu + e_t + \psi_1 Le_t + \psi_2 L^2 e_t + \ldots$ then
  - $y_t = \mu + e_t(1 + \psi_1 L + \psi_2 L^2 + \ldots)$ then
  - $y_t = \mu + \psi(L)e_t$ where $\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \ldots$
- This approach can be applied to other operators.
  - $Fe_t = L^{-1}e_t = e_{t+1}$
  - $De_t = (1-L)e_t = e_t - e_{t-1}$
  - $Se_t = (1-L^s)e_t = e_t - e_{t-s}$

Stationary

- A time series, y_t, is said to be 'stationary' if the following criteria are met:
  - $E(y_t) = \mu$
  - $E(y_t - \mu)^2 = \sigma^2$
  - $E(y_t - \mu)(y_{t-j} - \mu) = f(j) \neq f(t)$.
- A time series, y_t, is said to be 'non-stationary' if at least one of the above criteria is violated.
- Stationary is important because we can apply law of large number, central limit theorem and many other statistic theorems that are applied on independent random variables on stationary random variables.
- Dickey Fuller test or Augmented Dickey Fuller test to test for stationary process (significant p-value means stationary).
- Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests are used for testing a null hypothesis that an observable time series is stationary (non-significant p-value means non-stationary).


Stationary Time Series
y_t
10
5
0
-5
ADF = −6.128
0
200
400
600
800
1000
t


Non-stationary Time Series
y_t
20
0
-20
-40
-60
-100
ADF = −2.0251
0
200
400
600
800
1000
t

ARMA

- ARMA model is a shorthand for autoregressive moving average model, and the model is written as
  - $\phi(L)y_t = \theta(L)e_t$
- For example, an ARMA(2,1) model is written as
  - $(1-\phi_1 L - \phi_2 L^2)y_t = (1-\theta_1 L)e_t$
- For example, an ARMA(2,0) model is written as
  - $(1-\phi_1 L - \phi_2 L^2)y_t = e_t$, aka AR(2)
- For example, an ARMA(0,2) model is written as
  - $(1-\theta_1 L - \theta_2 L^2)e_t = y_t$, aka MA(2)

AR and MA properties

- For instance, let's consider AR(1) process:
  - $y_t = e_t + \phi_1 y_{t-1}$
  - $y_t = e_t + \phi_1 e_{t-1} + \phi_1^2 y_{t-2}$
  - $y_t = e_t + \phi_1 e_{t-1} + \phi_1^2 e_{t-2} + \phi_1^3 e_{t-3} + \ldots$
- Thus we can write AR(q) process as an MA(∞) process and vice versa.
- If $|\phi| < 1$, the AR process is 'stationary' and if $|\theta| < 1$ the MA process is 'invertible'.

What does ϕ tell us in lay term?

- Suppose the model is AR(1), we will find that
  - $Ey_t^2 = \sigma_e^2/(1-\phi_1^2)$ …
- This means that the condition, $|\phi| < 1$, guarantees that the influence of e_{t-j} on y_t diminishes as j increases.
- $|\phi| > 1$ means the impact of disturbances grows without bound (rarely occurs in real world).

Random walk process

- We call the series random walk process if
  - $y_t = y_{t-1} + e_t$ (in other words, $\phi = 1$)
- Let's consider $y_t = y_{t-1} + e_t$, we can prove that..
  - $E(y_t) = y_0$
  - $E(y_t^2) = \infty$ OR $Var(Y) = t\sigma^2$

Violate stationary assumption

- Note that $y_t = y_{t-1} + e_t + \alpha$ ➔ random walk with drift

ARMA and stationary process

- The ARMA models, described above can only be used for stationary time series data.
- However most time series models contain trend and seasonal patterns, which are also non-stationary in nature.
- Thus from application view point, ARMA models are inadequate.
- For this reason the ARIMA model is proposed, which is a generalization of an ARMA model.

ARIMA

- In ARIMA models a non-stationary time series is made stationary by applying finite differencing of the data points.
- ARIMA is a shorthand of autoregressive integrated moving average.
- ARIMA(p,d,q) is given as
  - $\phi(L)(1-L)^d y_t = (L)e_t$
- D is order of differencing until the process is stationary.
- SARIMA is a special case of ARIMA where seasonal component is added (such as $\theta_4 L^4$ for quarterly seasonal effect).➔can be additive and multiplicative.

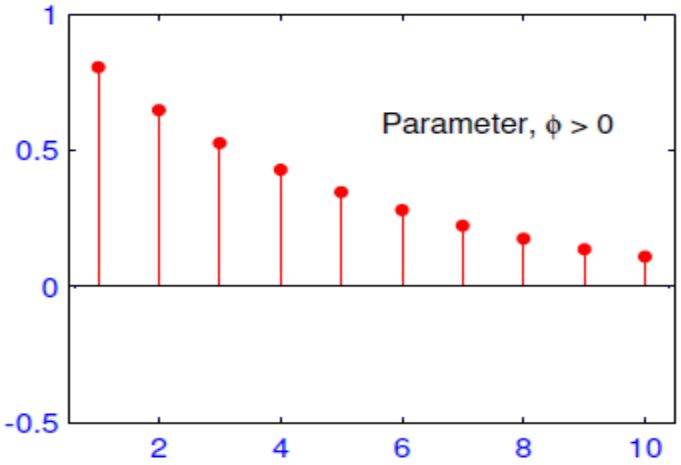
Autocorrelation function (ACF)

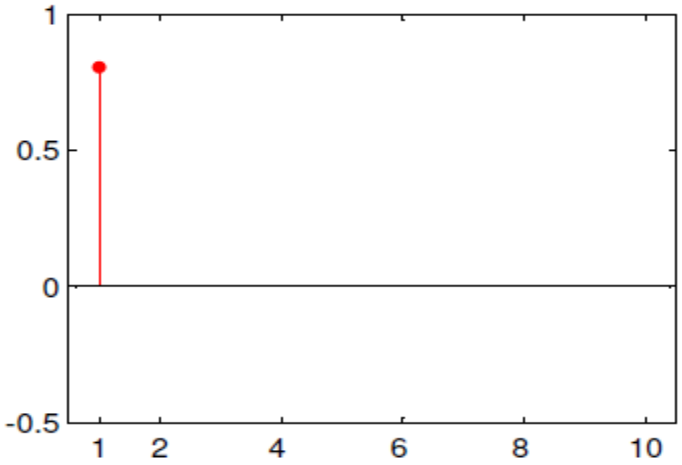
- Assessing autocorrelation with lag time by ‘correlogram’
- A correlogram, also known as an autocorrelation plot, is a plot of the sample autocorrelations r_h versus h (the time lags).
- If the autocorrelations collapse quickly toward 0 (either in the form of exponential decay or damped oscillation), the time series is stationary. We can use 95% CI to determine if the autocorrelations collapse to 0 significantly.
- If the magnitude if autocorrelations declines approximately linearly and does not collapse to 0, at least one more difference is needed.

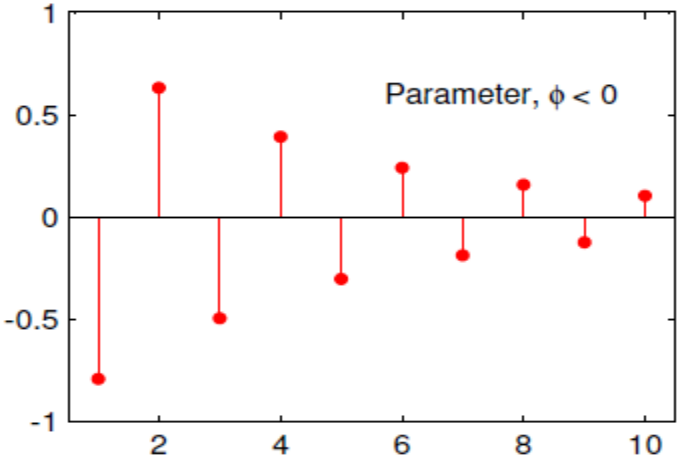
Partial autocorrelation function (PACF)

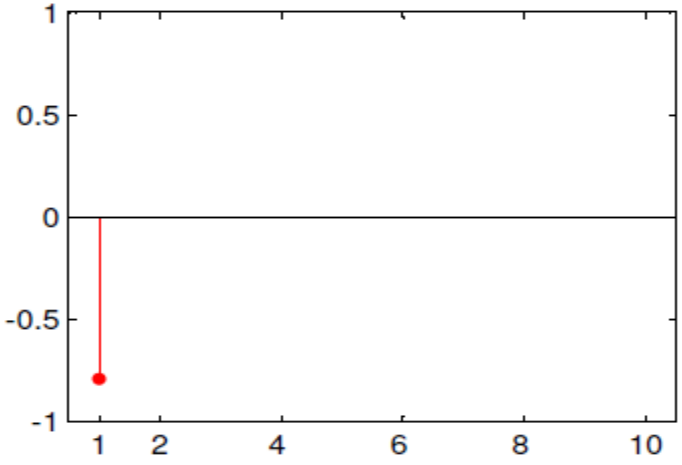
- If we compare an AR(l) with an AR(2), we will find that;
  - In the AR(1) the effect of y_{t-2} on y_t is always through y_{t-1}, and given y_{t-1}, the value of y_{t-2} is irrelevant for predicting y_t;
  - In an AR(2) in addition to the effect of y_{t-2} which is transmitted to y_t through y_{t-1}, there exists a direct effect on y_{t-2} on y_t.
- In general, an AR(p) has direct effects on observations separated by 1, 2, ..., p lags and the direct effects of the observations separated by more than p lags are null.
- Thus the partial autocorrelation coefficient of order k, denoted by ρ^p_k, is defined as the correlation coefficient between observations separated by k periods, when we eliminate the linear dependence due to intermediate values.

ACF and PACF for AR(1)





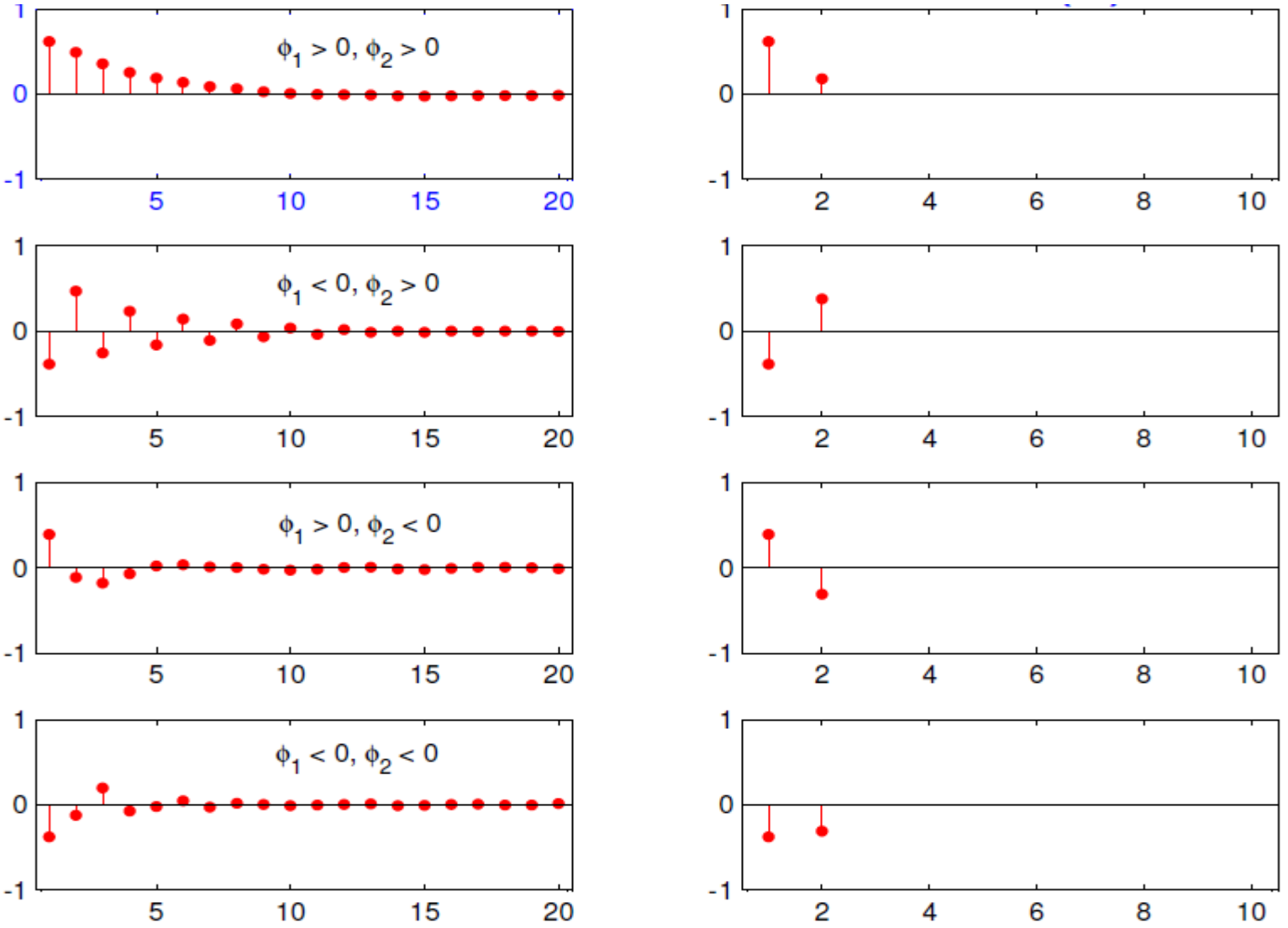




ACF

PACF

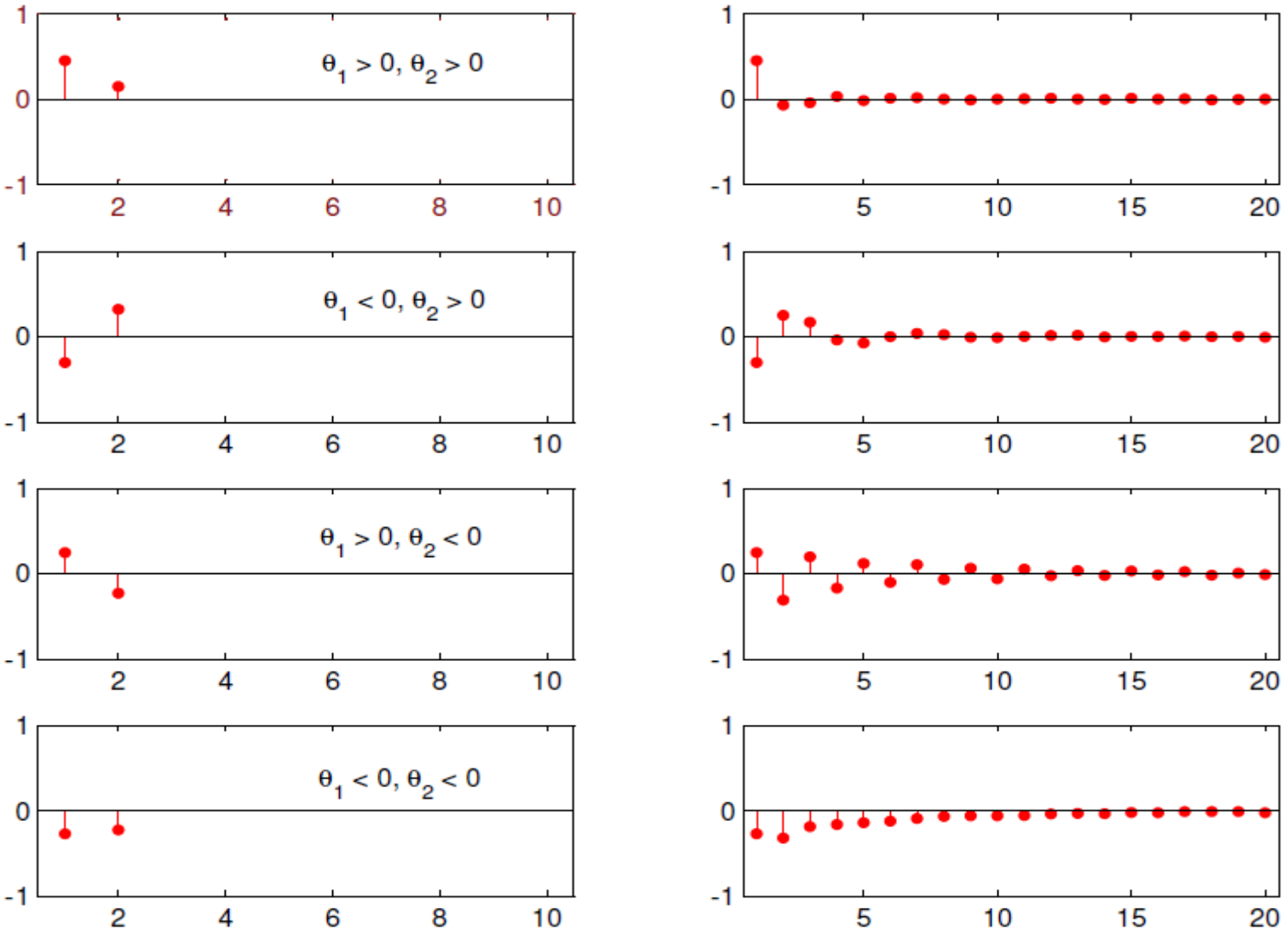
ACF and PACF for AR(2)



ACF

PACF

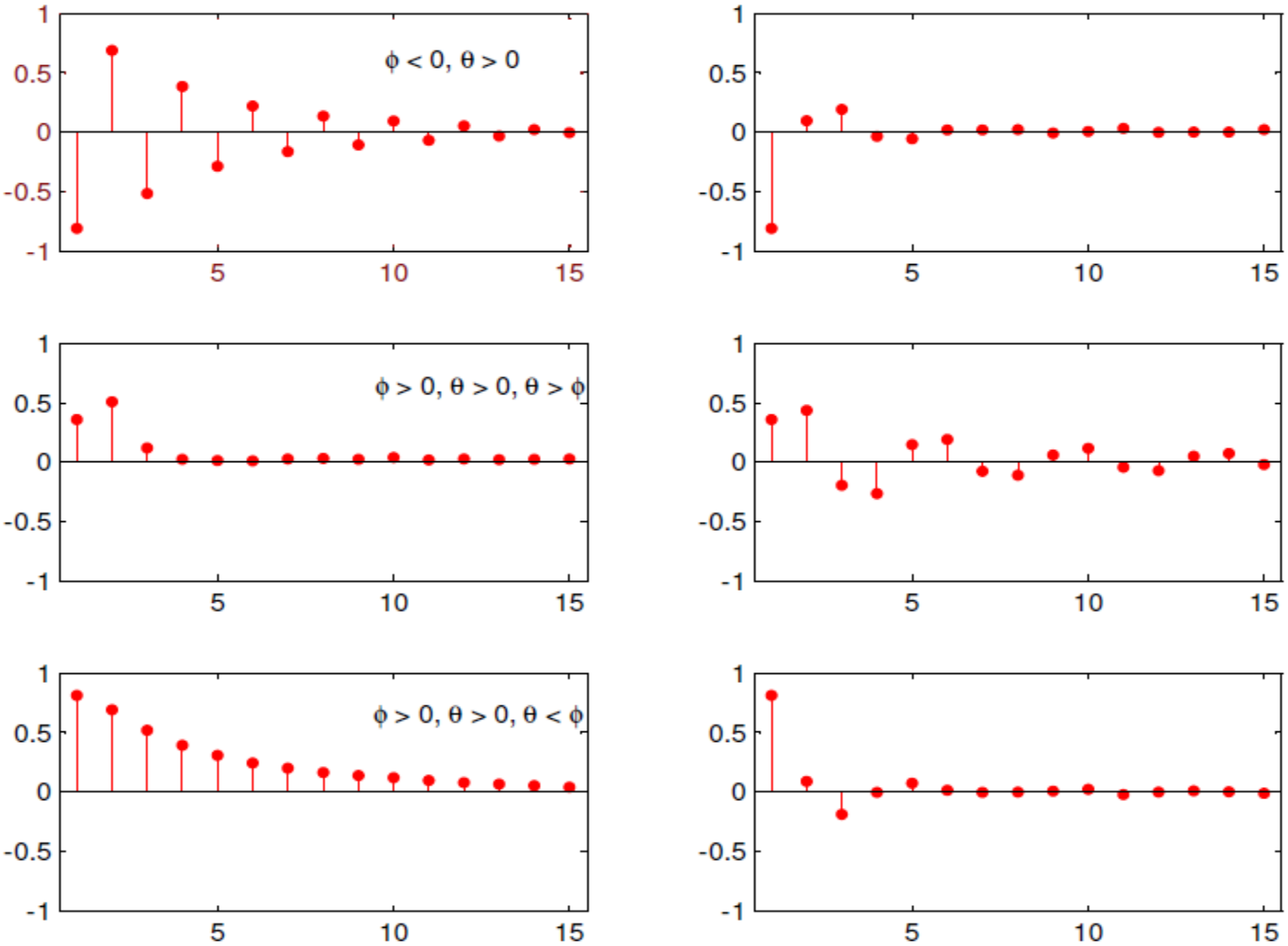
ACF and PACF for MA(2)



ACF

PACF

ACF and PACF for ARMA (1,1)



ACF

PACF

Summary of ACF and PACF

Prozess	ACF ($\rho_X(h)$)	PACF ($\pi_X(h)$)
AR(p)	infinite (dampened exponential or sinusoidal waves	finite $\pi_X(h) = 0$ for $h > p$
MA(q)	finite $\rho_X(h) = 0$ for $h > q$	infinite (dampened exponential or sinusoidal waves
ARMA(p, q)	as AR(p) for $h > q$	as MA(q) for $h > p$

Summary of ARIMA

- ARIMA(0,0,0) = white noise
- ARIMA(0,1,0) = random walk
- ARIMA(1,0,0) = AR(1)
- ARIMA(0,0,1) = MA(1)
- ARIMA(1,0,1) = ARMA(1,1)

More on ARIMA

- ARIMA with seasonal component
  - SARIMA is a designated version of ARIMA model with a seasonal component.
  - SAR(P) process = AR process with P seasonal lag
  - SMA(Q) process = MA process with Q seasonal lag
  - SI(D) process = Seasonal differencing with D seasonal lag
  - We can use the following notation: SARIMA(p,d,q) x $(P,D,Q)_s$
- We can model the residuals with ARIMA

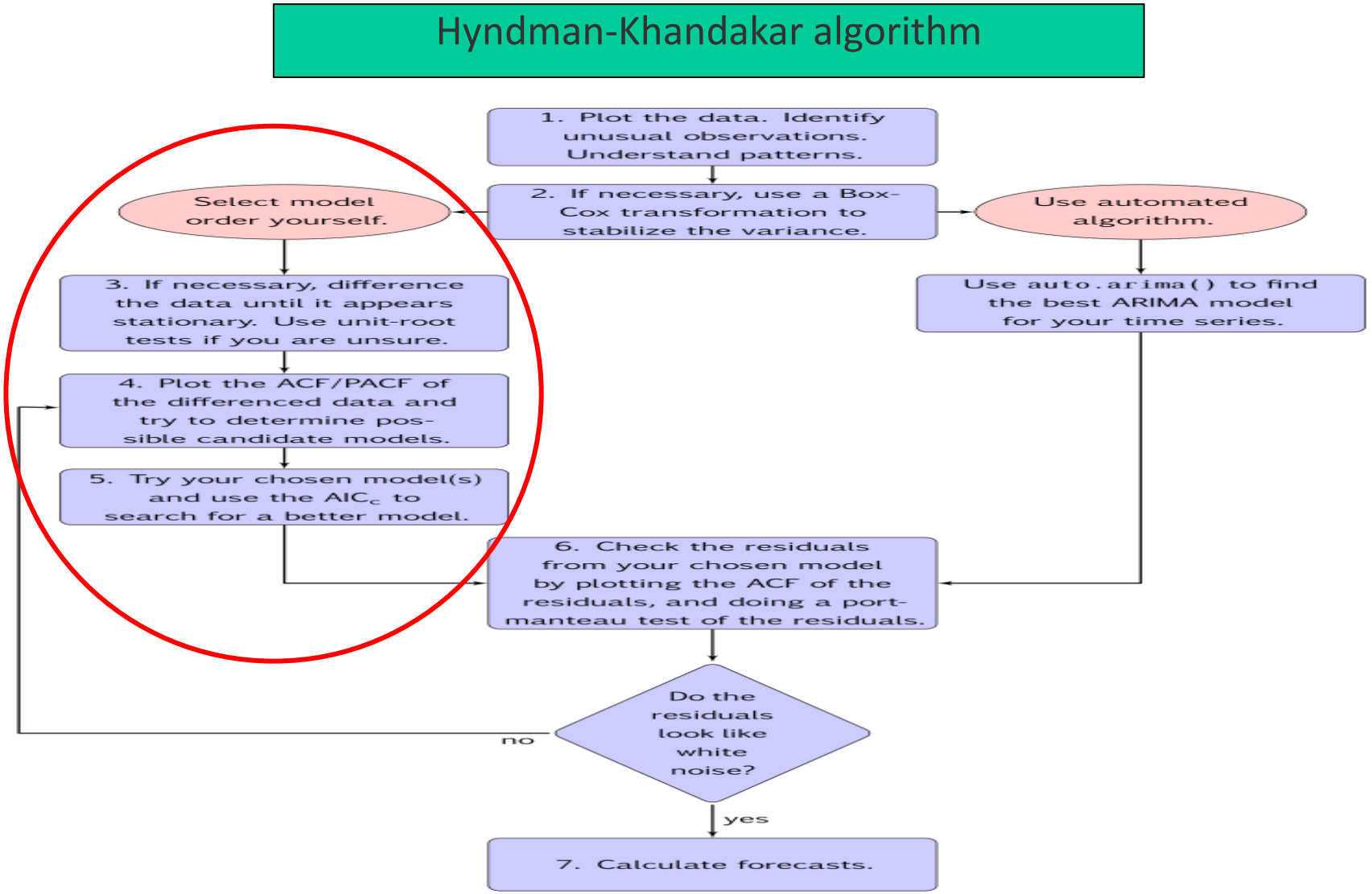
In practice, the steps in executing ARIMA in programming is as follows

When fitting an ARIMA model to a set of (non-seasonal) time series data, the following procedure provides a useful general approach.

1. Plot the data and identify any unusual observations.
2. If necessary, transform the data to stabilize the variance.*
3. If the data are non-stationary, take first differences of the data until the data are stationary.
4. Examine the ACF/PACF and identify p, d, q for ARIMA
5. Try your chosen model(s), and use the AIC to search for a better model.
6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model.
7. Once the residuals look like white noise, calculate forecasts.

*For instance, using a Box-Cox transformation

More on ARIMA (auto.arima in R)



Step by step of forecast

Summary of steps for forecasting

1. Model specification
   - Data visualization
   - Looking for patterns (trend, cycle, seasonal,…)
2. Model fitting and diagnosis
   - Using the whole records
   - Assessing accuracy (MAPE, MAE, etc)
   - Assessing error (autocorrelated disturbances, heteroscedasticity, etc)
3. Model validation
   - Back-testing
   - Assessing accuracy again
4. Model application and forecasting
   - Updating of data
   - Assessing accuracy again (and again)

Box-Jenkins approach

- Box and Jenkins (1970) proposed an iterative approach to time series modeling.
  - Step 1: Identification
    - Is the model ARMA(1,1) or ARMA(2,0)?
    - Assessing ACF and PACF
    - Do we need differencing to make it stationary?
    - Using unit root test ➔ Dickey Fuller test
      - Significant ➔ stationary
      - Non-significant ➔ non-stationary
  - Step 2: Estimation … Let's the software helps us
  - Step 3: Diagnostic checking
    - Overfitting (if now it is AR(2), what about AR(3))? ➔ Any change in goodness of fit
    - Checking for autocorrelation of disturbances ➔Q statistics
  - Step 4: Forecasting (if any)

Points to remember

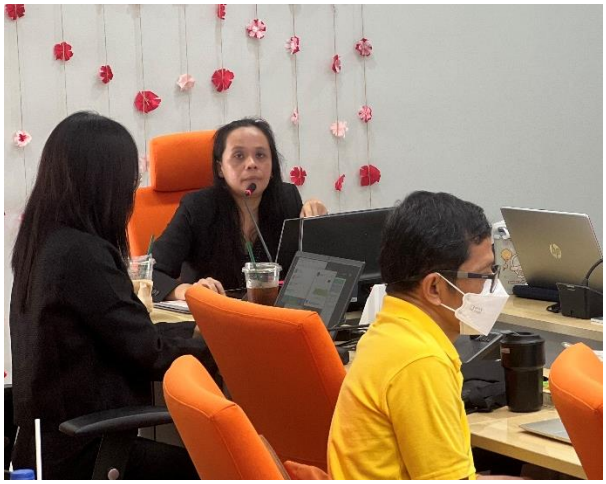
- Box-Jenkins approach is neither foolproof nor definitive. You still have to exercise judgment (and link with theoretical grounds).
- Parsimony *versus* goodness of fit
- Iterative process (remember back-testing!)
- With predict command in Stata or R, forecasting can be performed in many fashions (one step ahead, dynamic, etc)

Thank you
Let's exercise

ภาคผนวกที่ 5

ภาพประกอบการอบรม

กรมอนามัย



กระทรวงสาธารณสุข