



USAID
FROM THE AMERICAN PEOPLE



**Food and Agriculture
Organization of the
United Nations**



The Online R-FETPV 1st Module : Basic Epidemiology and Surveillance Data Analysis

5 April -28 May 2021



Food and Agriculture
Organization of the
United Nations



Basic descriptive and analytic statistics for disease outbreak event

CLIP#3

Sith Premashthira



Food and Agriculture
Organization of the
United Nations



Organizing and Displaying Epidemiologic Data with Tables and Graphs

Can you summarize the age and sex of the case-patients at a glance?

Case No.	Age	Sex
1	9	M
2	39	M
3	29	F
4	10	M
5	55	F
6	11	M
7	9	M
8	7	F
9	17	M
10	10	M

Case No.	Age	Sex
11	10	M
12	6	M
13	9	M
14	40	M
15	40	F
16	10	M
17	11	M
18	43	F
19	71	F
20	9	F

Case No.	Age	Sex
21	38	F
22	34	F
23	9	M
24	10	M
25	6	F
26	11	M
27	9	M
28	41	M
29	6	M
30	11	M

Case No.	Age	Sex
31	10	M
32	31	F
33	8	F
34	9	M
35	10	F
36	11	M
37	38	M
38	11	M
39	7	M
40	16	F

Basic Methods for Organizing and Presenting Data

- Data can be organized through creation of:
 - Tables
 - Graphs
 - Charts

Why organize and present data?

- To summarize when data set has too many records to look at individually
- To become familiar with the data before analysis, and to catch errors
- To look for (and display)
 - Patterns
 - Trends
 - Relationships
 - Exceptions / outliers
- To communicate findings to others

How to organize data

- Identify what data you have
- Use tables and graphs to summarize; catch errors; identify patterns, relationships
- Decide how best to summarize the data to communicate the findings
- Use tables and graphs to communicate the findings effectively

Tables

- Data are arranged in rows and columns
- Quantitative information
- Usually, presents frequency of occurrence of some event or characteristic in different subgroups

Tables

Type of injury by sex, Port-au-Prince field hospital, Haiti,
January 13 – May 28, 2010

	Earthquake -related injury	Other injury	Total
Male	74	259	333
Female	85	151	236
Unknown	3	9	12
Total	162	419	581

Types of Tables

- 1-variable table (frequency distribution)
 - Range of values of a single variable
 - Number of observations with each value
- 2-variable table
 - Counts shown according to 2 variables at once
- 3-variable table
 - Counts shown according to 3 variables at once
- Composite (combination) tables

Example of 1-Variable Table — Tuberculosis Cases by Sex, U.S., 2009

Table 1. Number of Reported Cases of Tuberculosis,
by Sex, United States, 2009

Sex	# Cases
Males	6,990
Females	4,544
Unknown	11
Total	11,545

Example of 1-Variable Table, with Percent Column

Table 2. Number of Reported Cases of Tuberculosis,
by Age, United States, 2009

<u>Age Group (years)</u>	<u># Cases</u>	<u>Percent</u>
≤ 5	401	3.5%
5 – 14	245	2.1%
15 – 24	1,274	11.0%
25 – 44	3,893	33.7%
45 – 64	3,434	29.7%
≥65	2,292	19.9%
Unknown	6	0.1%
Total	11,545	100.0%

Two-Variable Tables

- Shows counts according to two variables simultaneously
- Also called “cross-tab” or contingency tables

Example of Two-variable Table

Table 3. Number of Reported Cases of Tuberculosis, by Age and Sex, United States, 2009

<u>Age Group</u>	<u>Females</u>	<u>Males</u>	<u>Unk</u>	<u>Total</u>
≤ 5	187	214	0	401
5 – 14	119	126	0	245
15 – 24	559	713	2	1,274
25 – 44	1,641	2,247	5	3,893
45 – 64	1,153	2,278	3	3,434
≥ 65	882	1,409	1	2,292
Unknown	3	3	0	6
Total	4,554	6,990	11	11,545

Example of Two-by-Two Table

		Ill	Well		Attack Rate (%)
Drank from stream near Campsite 6?	Yes	18	4	22	81.8%
	No	5	39	44	11.4%
		23	43	66	

Why Tables?

- When too many records, summarize in table (or graph)
- Allow you to identify, explore, understand, and present distributions, trends, relationships, variations, and exceptions in the data
- Tables serve as basis for graphs – always create a table first!

Some Tips for Creating Printed Tables

- Keep it simple
- Should be self-explanatory
- Title (what, where, when) with table number
- Label each row and column clearly and concisely
- Include units of measurement (years, mg/dl, etc.)
- Show totals for rows and columns
- Explain codes, abbreviations, symbols
- Note any exclusions in a footnote
- Note source in a footnote

Graphs

- Display quantitative data using a set of coordinates
- Rectangular graphs (x , y coordinates) most common
- x axis along bottom = method of classification, often time
- y axis along side = frequency, usually number, percent or rate

Graphs: Advantages and Disadvantages

Advantages

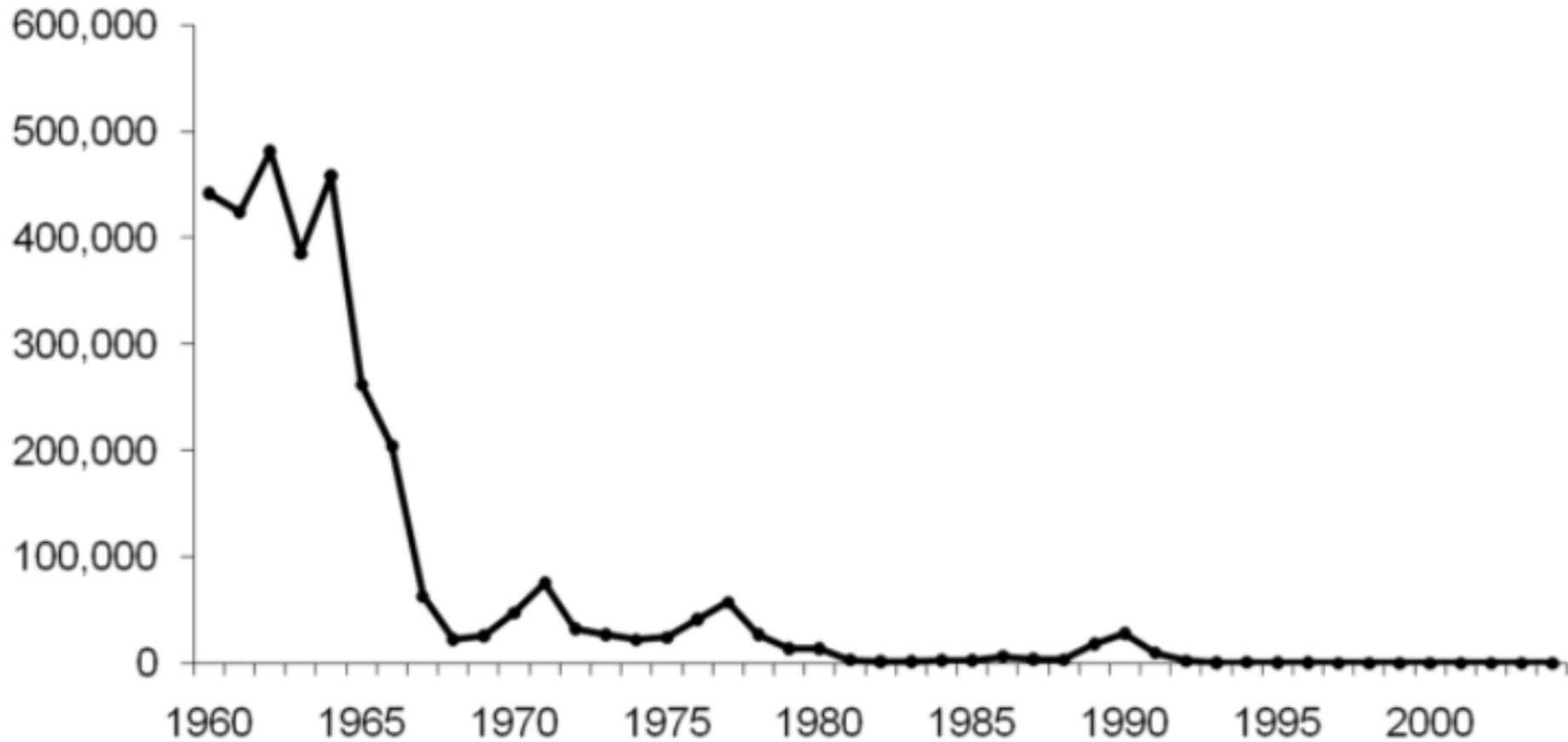
- Easy to understand and interpret
- Reveal patterns in data
 - Useful for generating hypotheses
 - Useful before formal data analysis

Disadvantage

- Loss of detail

Arithmetic Scale Line Graph

Cases

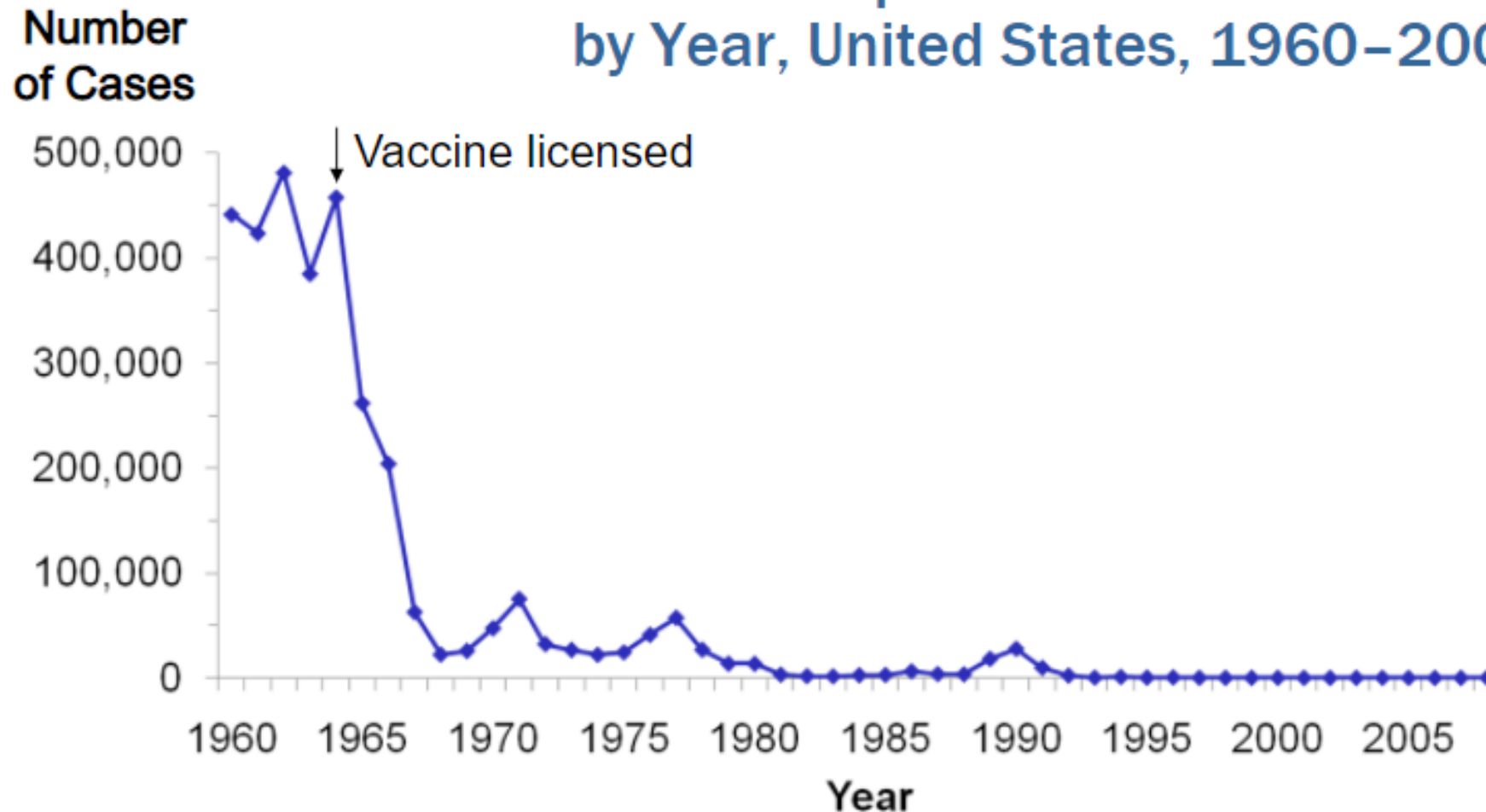


Creating a Line Graph

- Make x-axis longer than y-axis (best ratio 5:3)
- X-axis: Match x-axis scale to intervals used during data collection
- Y-axis:
 - Always start y-axis with 0
 - Identify largest value, round up for maximum Y value
 - Select reasonable intervals for y-axis
- Plot data
- Create title
- Add comments, footnotes

Creating a Line Graph: Add Comments, Footnotes, Source

Number of Reported Cases of Measles by Year, United States, 1960–2008



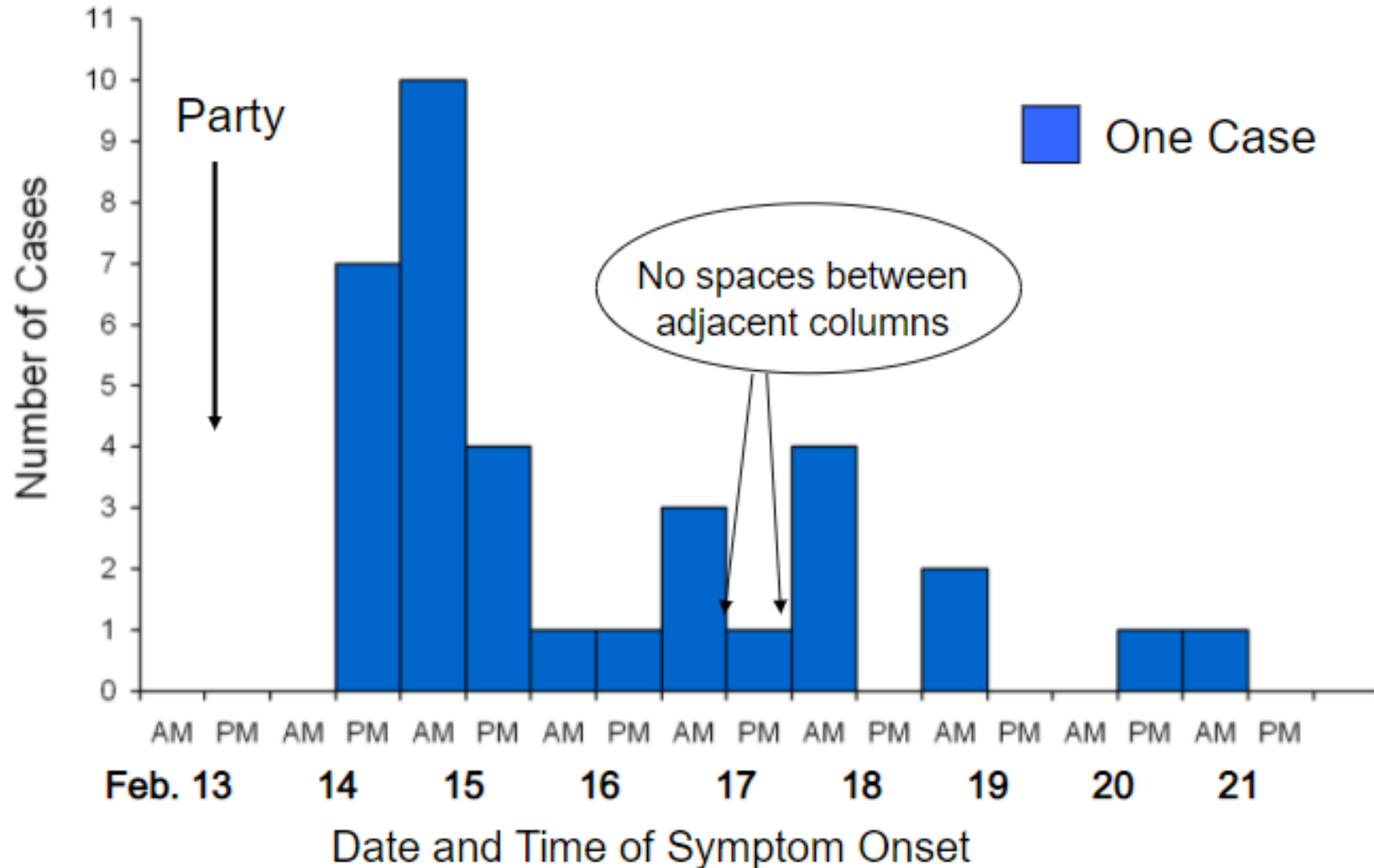
Comments on Arithmetic-Scale Line Graph

- Method of choice for plotting rates over time
- X-axis almost always time (rarely, age)
- Y-axis can be counts, proportions, or rates
 - Y-axis should start with 0
 - Determine largest value of Y needed to plot
 - Round off that number and divide into intervals
- Set distance on either axis represents same quantity anywhere on that axis
- Good for comparing 2 or more sets of data

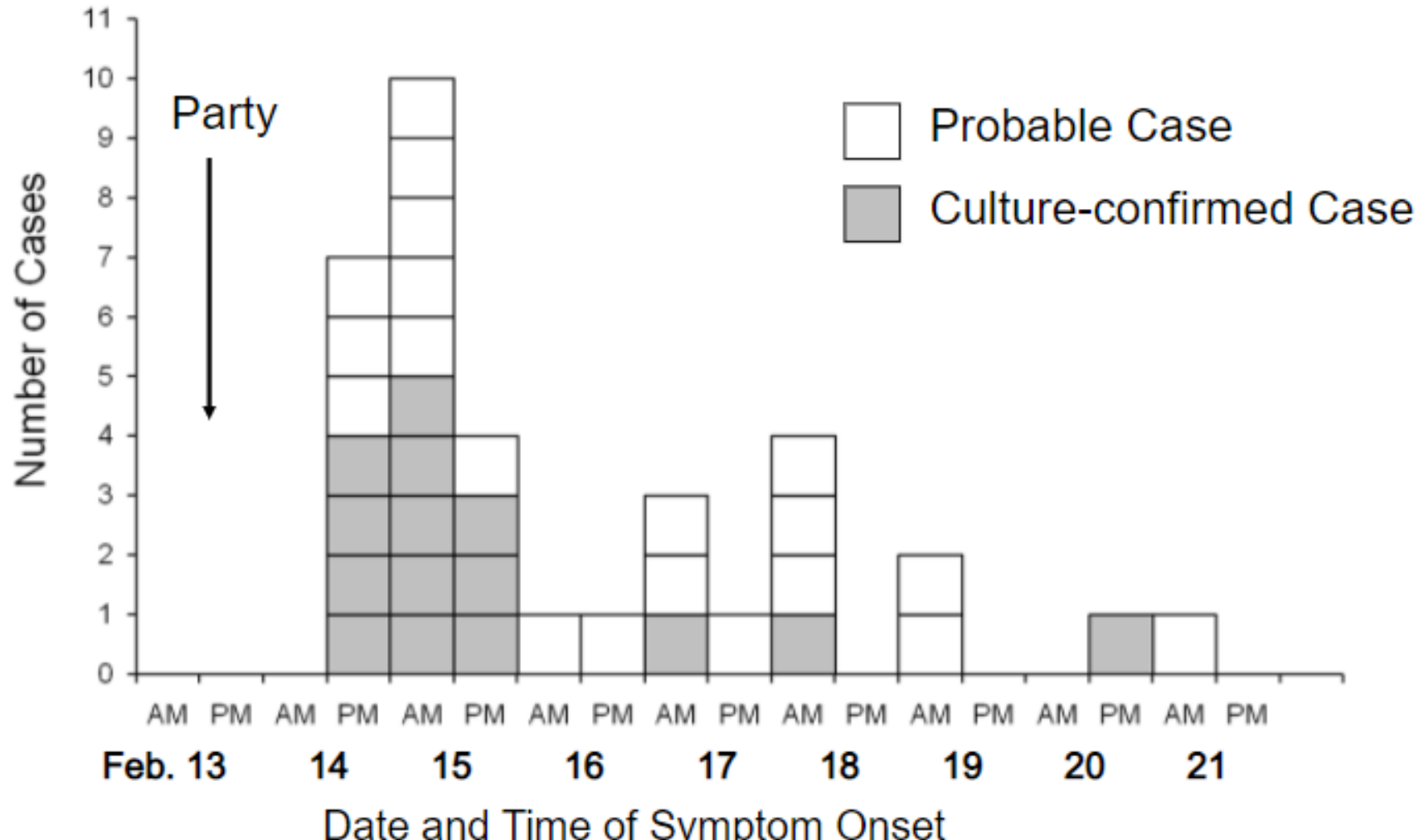
Histogram

- “Epidemic curve” in outbreak investigations
- Frequency distribution of quantitative data
- x axis continuous, usually time (onset or diagnosis date)
- No spaces between adjacent columns, i.e., adjacent columns “touch”
- Easiest to interpret with equal class (x) intervals
- Column height proportional to number of observations in that interval

Number of Cases of *Salmonella* Enteritidis by Date of Onset, Chicago, February 2000



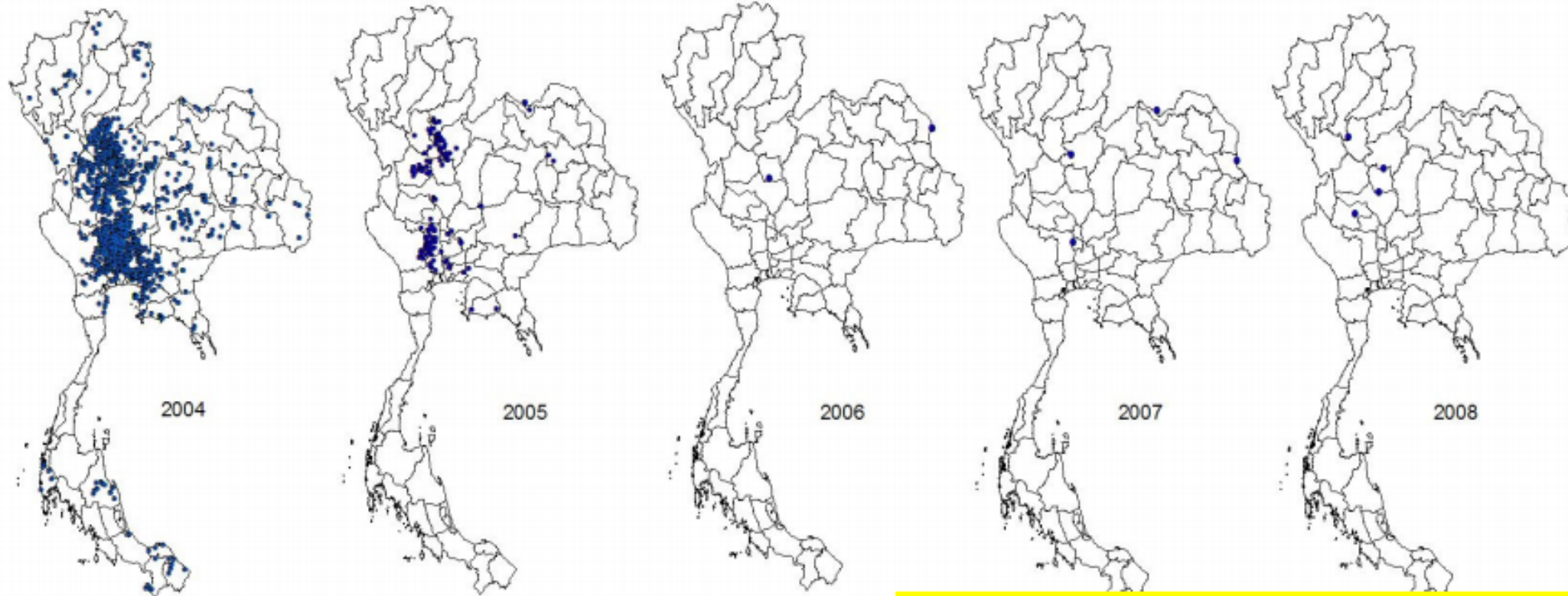
Number of Cases of *Salmonella* Enteritidis by Date of Onset, Chicago, February 2000



Charts

- Display quantitative data using only one coordinate
- Most appropriate for comparing data with discrete categories
- Common types include:
 - Bar charts
 - Pie charts
 - Maps
 - Other

Situation of HPAI in Thailand, 2004-2008

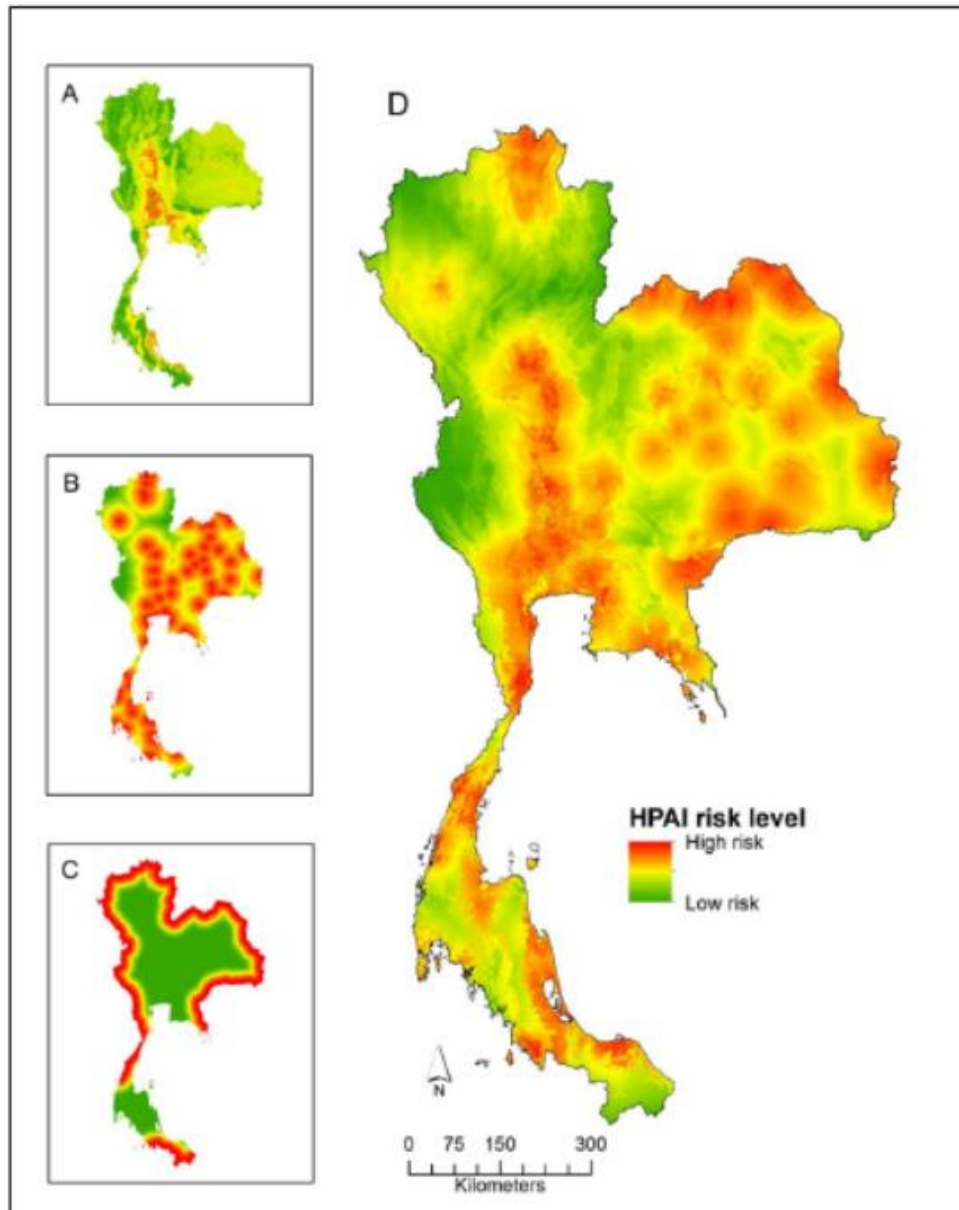


created by AICC,DLD,Thailand

Last HPAI outbreak in November 2008

Figure 5 The risk map of HPAI in Thailand

The risk map of HPAI (D) was analyzed based on Spatial-MCDA model using 3 spatial layers namely; previous HPAI risk map analyzed by Gilbert et al (2008) (A) [3], Distance to migratory bird sanctuary surveyed by DNP (B), and Distance to border (C).



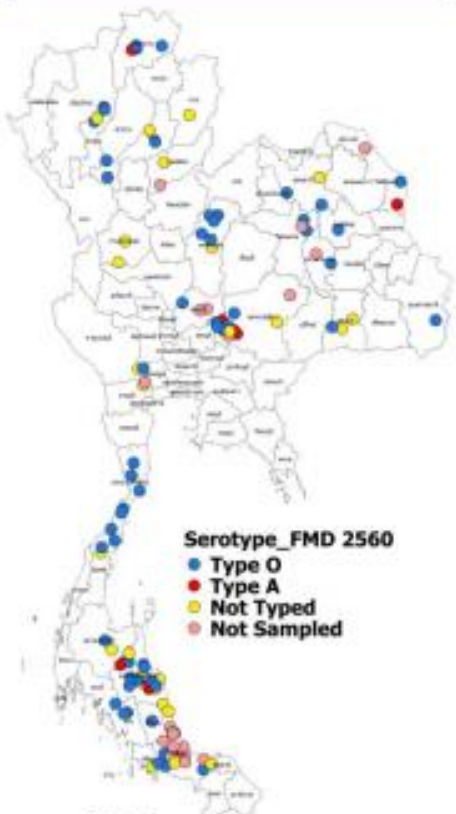
FMD Situation 2017-2020

2017

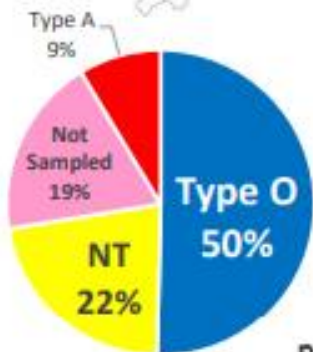
2018

2019

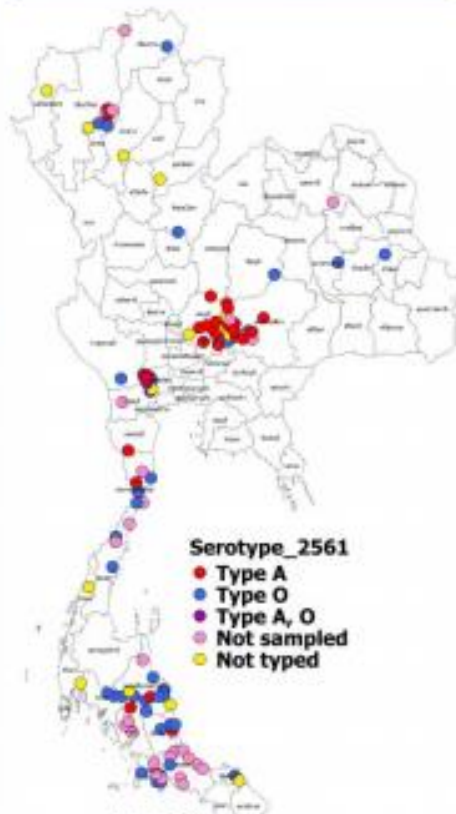
2020



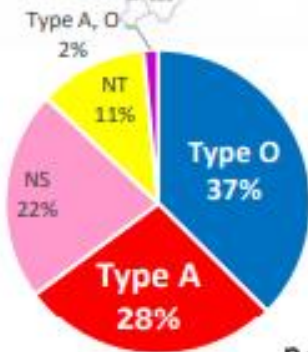
Serotype_FMD 2560
 ● Type O
 ● Type A
 ● Not Typed
 ● Not Sampled



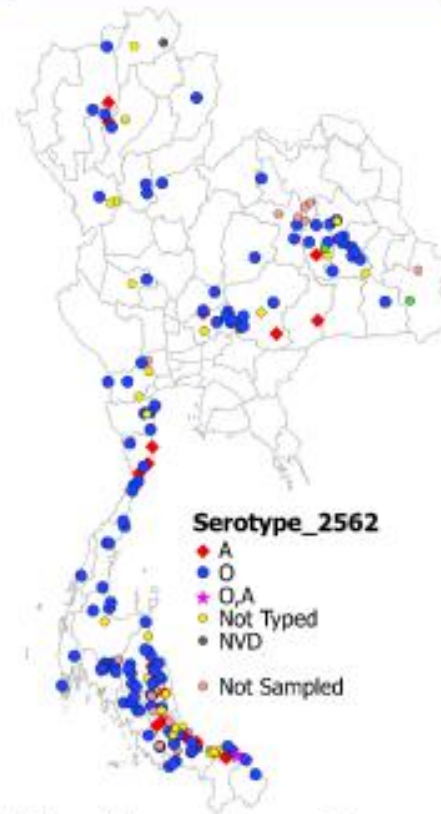
n = 127
32 Provinces



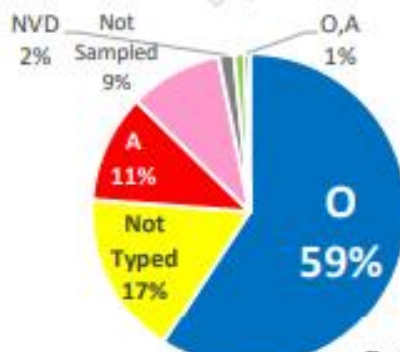
Serotype_2561
 ● Type A
 ● Type O
 ● Type A, O
 ● Not sampled
 ● Not typed



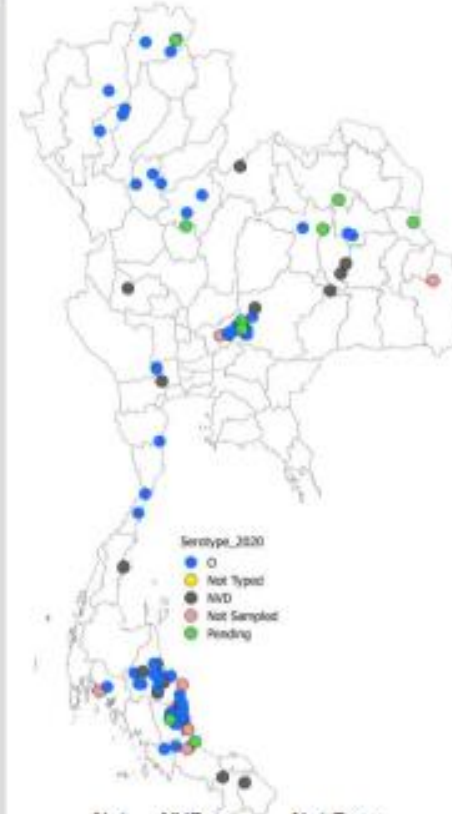
n = 131
30 Provinces



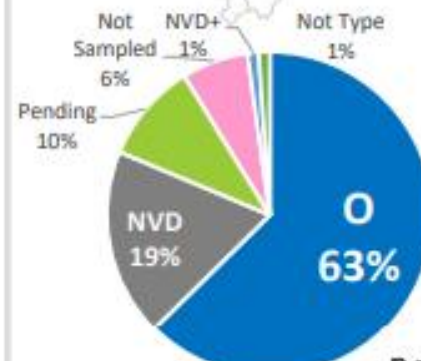
Serotype_2562
 ● A
 ● O
 ● O,A
 ● Not Typed
 ● NVD
 ● Not Sampled



n = 189



Serotype_2020
 ● O
 ● Not Typed
 ● NVD
 ● Not Sampled
 ● Pending



n = 91
Provinces

Some Tips for Creating Printed Graphs

- Should be self-explanatory
- Title (what, where, when) with table number
- Label each axis clearly and concisely
- Include units of measurement (years, mg/dl, etc.)
- In epidemiology, start Y-axis at zero
- Epidemic curve = histogram

Summary

- Data can be organized through the creation of tables, graphs and charts
- The purpose of creating these visual displays
 1. verify and analyze the data
 2. explore patterns and trends
 3. communicate information to others
- An effective figure should be able to be interpreted without any additional information

Summary 2

- Tables can illustrate the number of people with particular characteristics and can provide valuable information about relationships between 2 variables
- Line graphs are useful for showing patterns or trends over some variable, usually time
- Histograms are most commonly used in epidemiology for epidemic curves (cases by time)
- Bar charts provide a visual display of data from a one-variable table, but grouped bar charts can show 2 variables