GETTING TO KNOW DATA (PART 1)

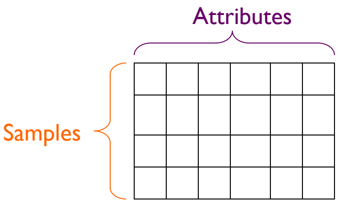
SUPAPORN ERJONGMANEE

DEPARTMENT OF COMPUTER ENGINEERING
FACULTY OF ENGINEERING
KASETSART UNIVERSITY

TYPES OF DATA SETS

- Record
  - Relational records
  - Data matrix, e.g., numerical matrix, crosstabs
  - Document data: text documents: term-frequency vector
  - Transaction data

	team	coach	play	ball	score	game	win	lost	timeout	season
Document 1	3	0	5	0	2	6	0	2	0	2
Document 2	0	7	0	2	1	0	0	3	0	0
Document 3	0	1	0	0	1	2	2	0	3	0

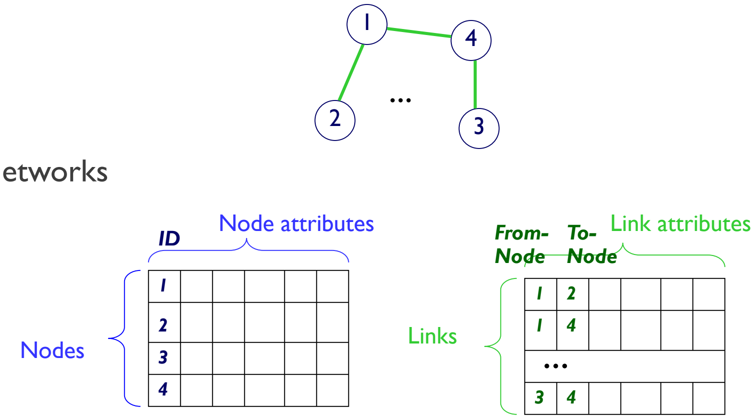


Source:
J. Han, M. Kamber and J. Pei, "Chapter 2 Know Your Data" in Data Mining: Concepts and Techniques, Morgan Kaufmann, July 2011.

3

TYPES OF DATA SETS

- Graph and network
  - World Wide Web
  - Social or information networks
  - Molecular Structures



Node attributes

Nodes

ID					
1					
2					
3					
4					

Link attributes

Links

From-Node	To-Node				
1	2				
1	4				
…					
3	4				

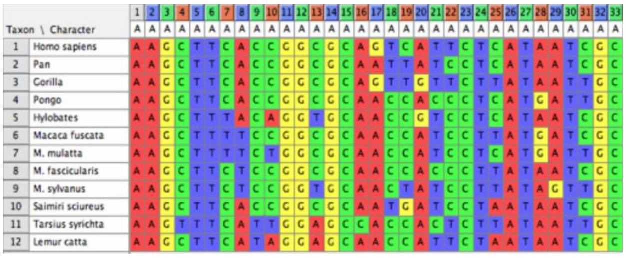
Source:
J. Han, M. Kamber and J. Pei, "Chapter 2 Know Your Data" in Data Mining: Concepts and Techniques, Morgan Kaufmann, July 2011.

3

4

TYPES OF DATA SETS (CONT.)

- Ordered
  - Video data: sequence of images
  - Temporal data: time-series
  - Sequential Data: transaction sequences
  - Genetic sequence data
- Others
  - Spatial data: maps
  - Image data
  - Video data



		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Taxon \ Character		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
1	Homo sapiens	A	A	G	C	T	T	C	A	C	C	G	G	C	G	C	A	G	T	C	A	T	T	C	T	C	A	T	A	A	T	C	G	C
2	Pan	A	A	G	C	T	T	C	A	C	C	G	G	C	G	C	A	A	T	T	A	T	C	C	T	C	A	T	A	A	T	C	G	C
3	Gorilla	A	A	G	C	T	T	C	A	C	C	G	G	C	G	C	A	G	T	T	G	T	T	C	T	T	A	T	A	A	T	T	G	C
4	Pongo	A	A	G	C	T	T	C	A	C	C	G	G	C	G	C	A	A	C	C	A	C	C	C	T	C	A	T	G	A	T	T	G	C
5	Hylobates	A	A	G	C	T	T	T	A	C	A	G	G	T	G	C	A	A	C	C	G	T	C	C	T	C	A	T	A	A	T	C	G	C
6	Macaca fuscata	A	A	G	C	T	T	T	T	C	C	G	G	C	G	C	A	A	C	C	A	T	C	C	T	T	A	T	G	A	T	C	G	C
7	M. mulatta	A	A	G	C	T	T	T	T	C	T	G	G	C	G	C	A	A	C	C	A	T	C	C	T	C	A	T	G	A	T	T	G	C
8	M. fascicularis	A	A	G	C	T	T	C	T	C	C	G	G	C	G	C	A	A	C	C	A	C	C	C	T	T	A	T	A	A	T	C	G	C
9	M. sylvanus	A	A	G	C	T	T	C	T	C	C	G	G	T	G	C	A	A	C	T	A	T	C	C	T	T	A	T	A	G	T	T	G	C
10	Saimiri sciureus	A	A	G	C	T	T	C	A	C	C	G	G	C	G	C	A	A	T	G	A	T	C	C	T	A	A	T	A	A	T	C	G	C
11	Tarsius syrichta	A	A	G	T	T	T	C	A	T	T	G	G	A	G	C	C	A	C	C	A	C	T	C	T	T	A	T	A	A	T	T	G	C
12	Lemur catta	A	A	G	C	T	T	C	A	T	A	G	G	A	G	C	A	A	C	C	A	T	T	C	T	A	A	T	A	A	T	C	G	C

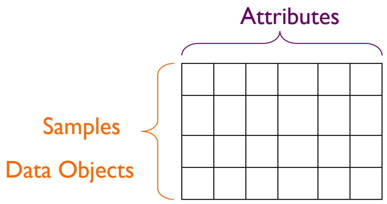
Image source:
https://mesquiteproject.wikispaces.com/file/view/DNAMatrix.gif/518627818/DNAMatrix.gif

Source (edited):
J. Han, M. Kamber and J. Pei, "Chapter 2 Know Your Data" in Data Mining: Concepts and Techniques, Morgan Kaufmann, July 2011.

4

Data Objects

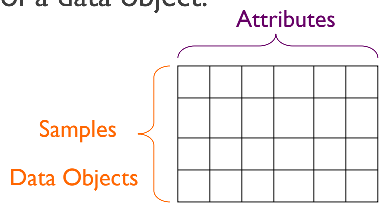
- Database columns -> attributes.
- Database rows -> data objects
  - Data sets are made up of data objects.
  - Also called *samples , examples, instances, data points, objects, tuples*.
  - A **data object** represents an entity.
- Examples:
  - sales database: customers, store items, sales
  - medical database: patients, treatments
  - university database: students, professors, courses



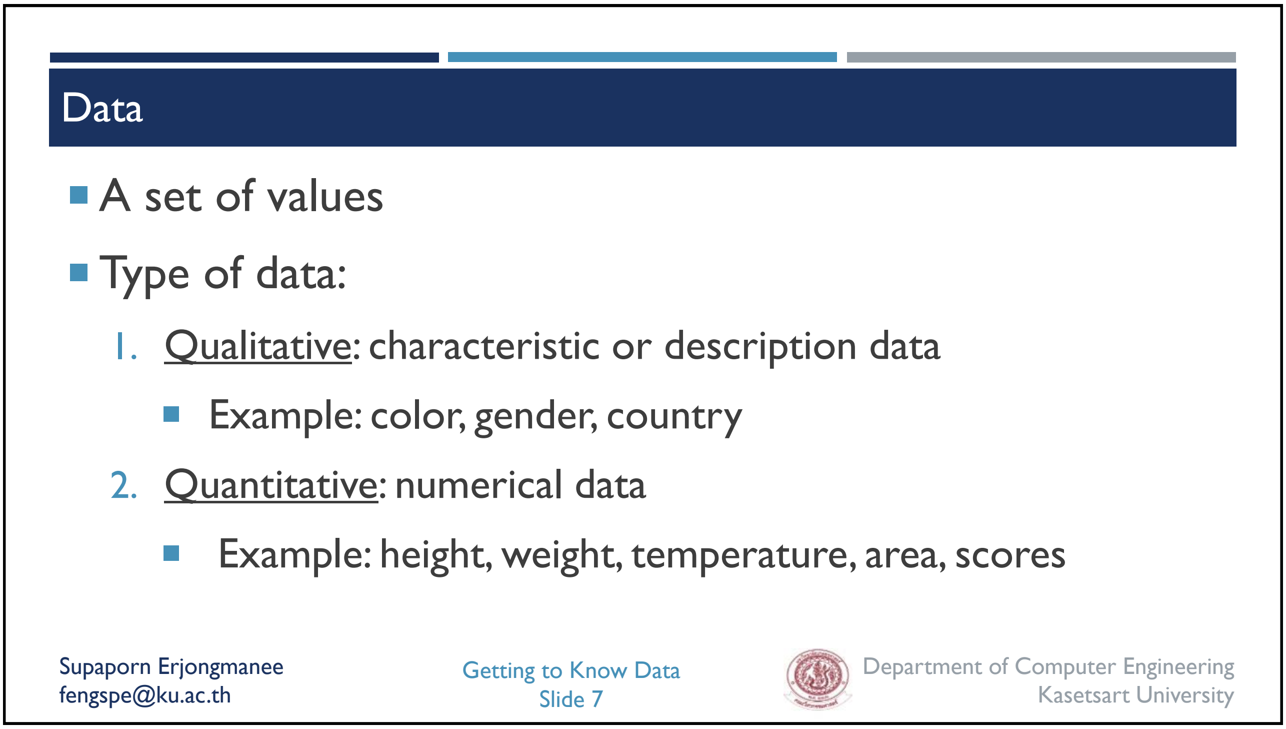
Source (edited):
J. Han, M. Kamber and J. Pei, "Chapter 2 Know Your Data" in Data Mining: Concepts and Techniques, Morgan Kaufmann, July 2011.

Attributes

- **Attribute** (or **dimensions, features, variables**):
  - a data field, representing a characteristic or feature of a data object.
  - *E.g., customer_ID, name, address*
- Types:
  - Qualitative data
  - Quantitative data



Source (edited):
J. Han, M. Kamber and J. Pei, "Chapter 2 Know Your Data" in Data Mining: Concepts and Techniques, Morgan Kaufmann, July 2011.



Data

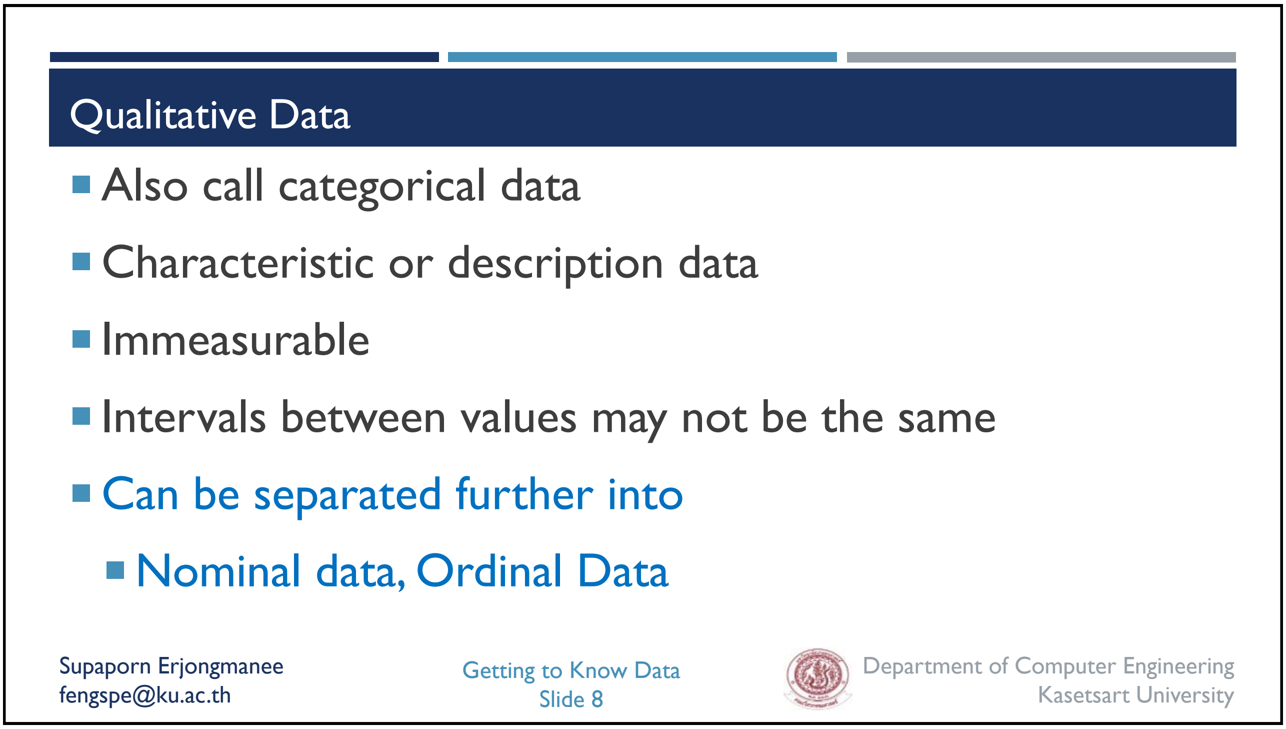
- A set of values
- Type of data:
  1. Qualitative: characteristic or description data
     - Example: color, gender, country
  2. Quantitative: numerical data
     - Example: height, weight, temperature, area, scores

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 7

Department of Computer Engineering
Kasetsart University

7



Qualitative Data

- Also call categorical data
- Characteristic or description data
- Immeasurable
- Intervals between values may not be the same
- Can be separated further into
  - Nominal data, Ordinal Data

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 8

Department of Computer Engineering
Kasetsart University

8

Nominal Data

- Data separated in classes
  - Classes do not always relate to one another
  - Cannot really sort classes (not in order)
- Example:
  - Gender: male, female
  - Regions: America, Asia, Europe
  - Directions: North, East, West, South

Nominal Data Example

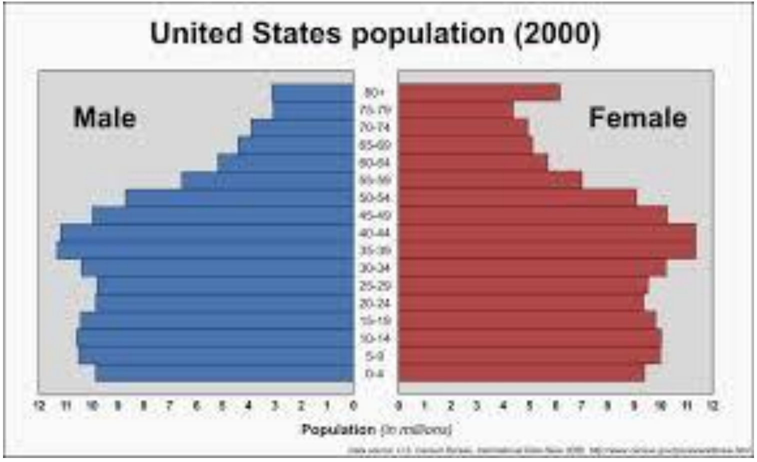


Image source: https://encrypted-tbn1.gstatic.com/images?q=tbn:ANd9GcSOwQW4R12eGZJo71pTF-dqPJb7gVY8fSqMevQFNWw_3izPp_gi

Ordinal Data

- Data with ranks
  - Ranks are not actually numerical values (but some can be converted to numbers)
  - Can be sorted
- Immeasurable
- Intervals between values may not be the same
- Example:
  - Size: small, medium, large
  - Satisfaction degree: best, good, poor, worst

Ordinal Data Example

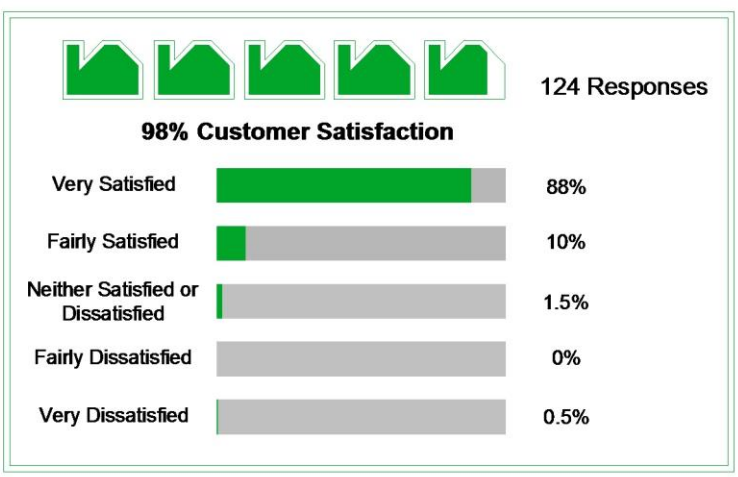
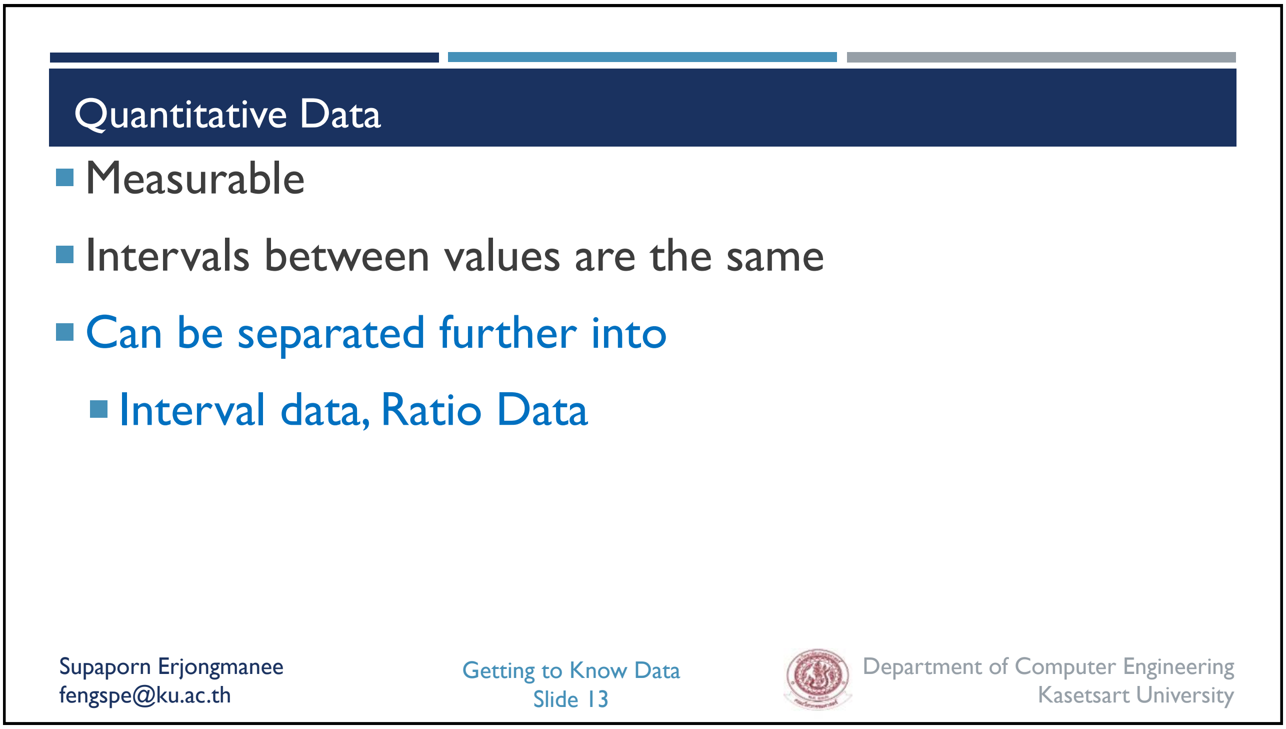


Image source: https://www2.barnsley.gov.uk/media/2624867/Customer%20Satisfaction%20graph,%20page%20content%20for%20detail.jpg



Quantitative Data

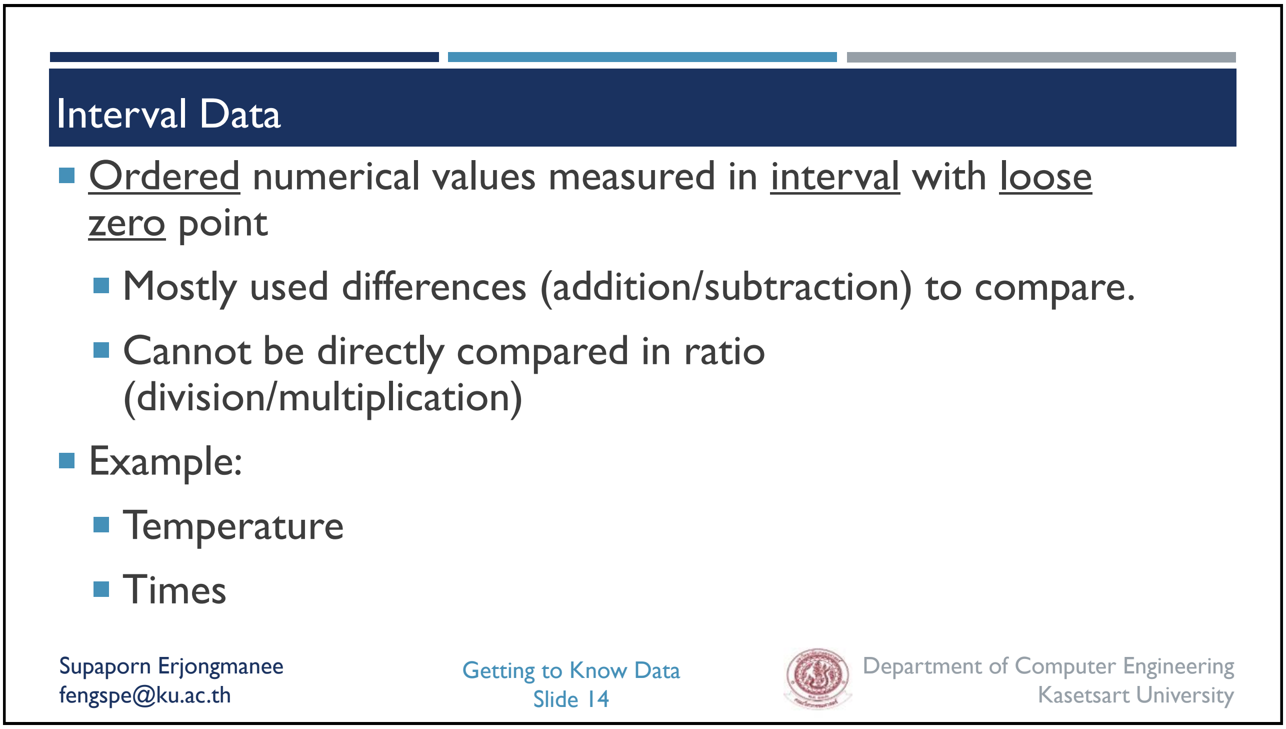
- Measurable
- Intervals between values are the same
- Can be separated further into
  - Interval data, Ratio Data

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 13

Department of Computer Engineering
Kasetsart University

13



Interval Data

- Ordered numerical values measured in interval with loose zero point
  - Mostly used differences (addition/subtraction) to compare.
  - Cannot be directly compared in ratio (division/multiplication)
- Example:
  - Temperature
  - Times

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 14

Department of Computer Engineering
Kasetsart University

14

Interval Data Example

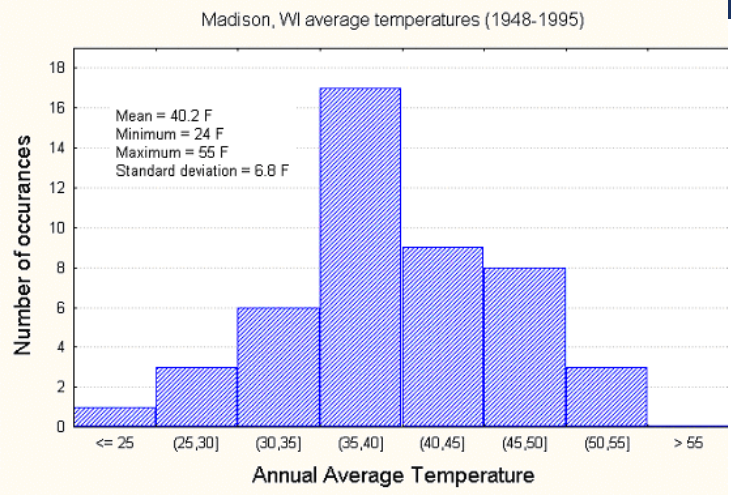
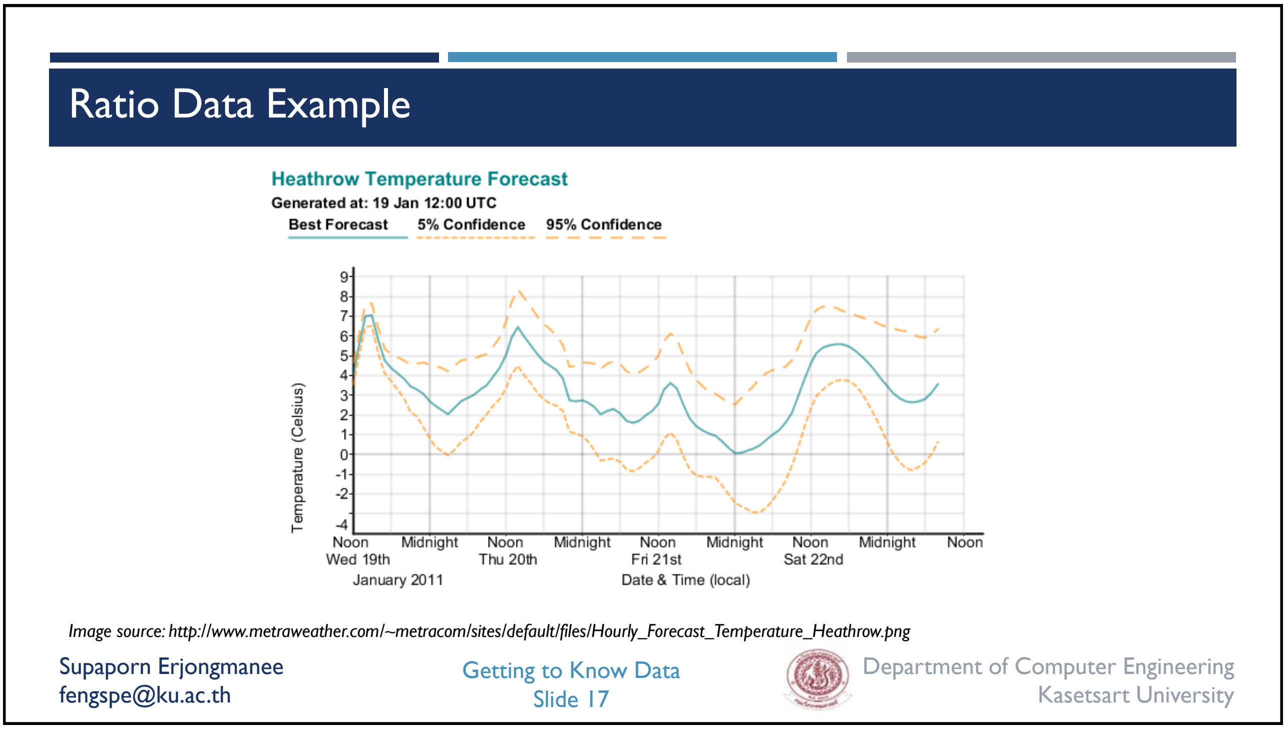


Image source: http://itg1.meteor.wisc.edu/wxwise/AckermanKnox/chap14/madisonhist.gif

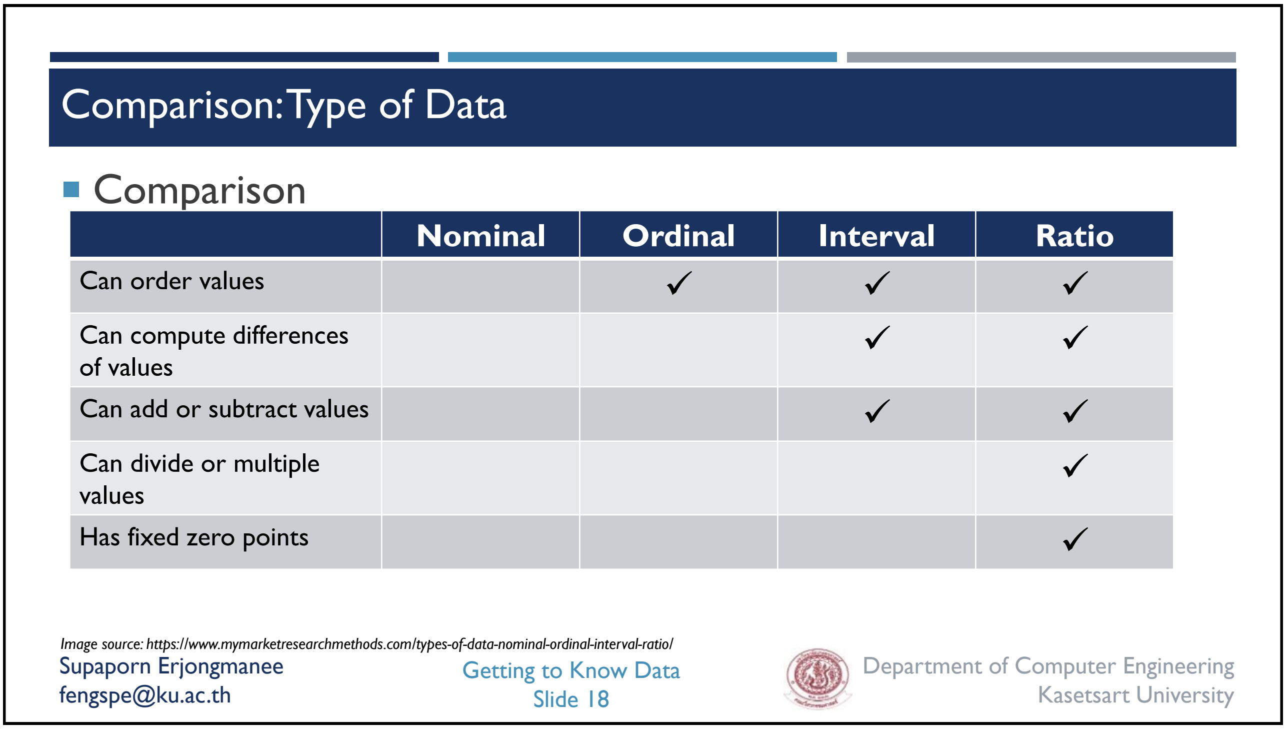
Ratio Data

- Measurable
- Intervals between values are the same.
- Can be computed using
  - Difference (addition/subtraction)
  - Ratio (multiplication/division)
- Example:
  - Weight, Length, Revenue



Ratio Data Example

Image source: http://www.metraweather.com/~metracom/sites/default/files/Hourly_Forecast_Temperature_Heathrow.png

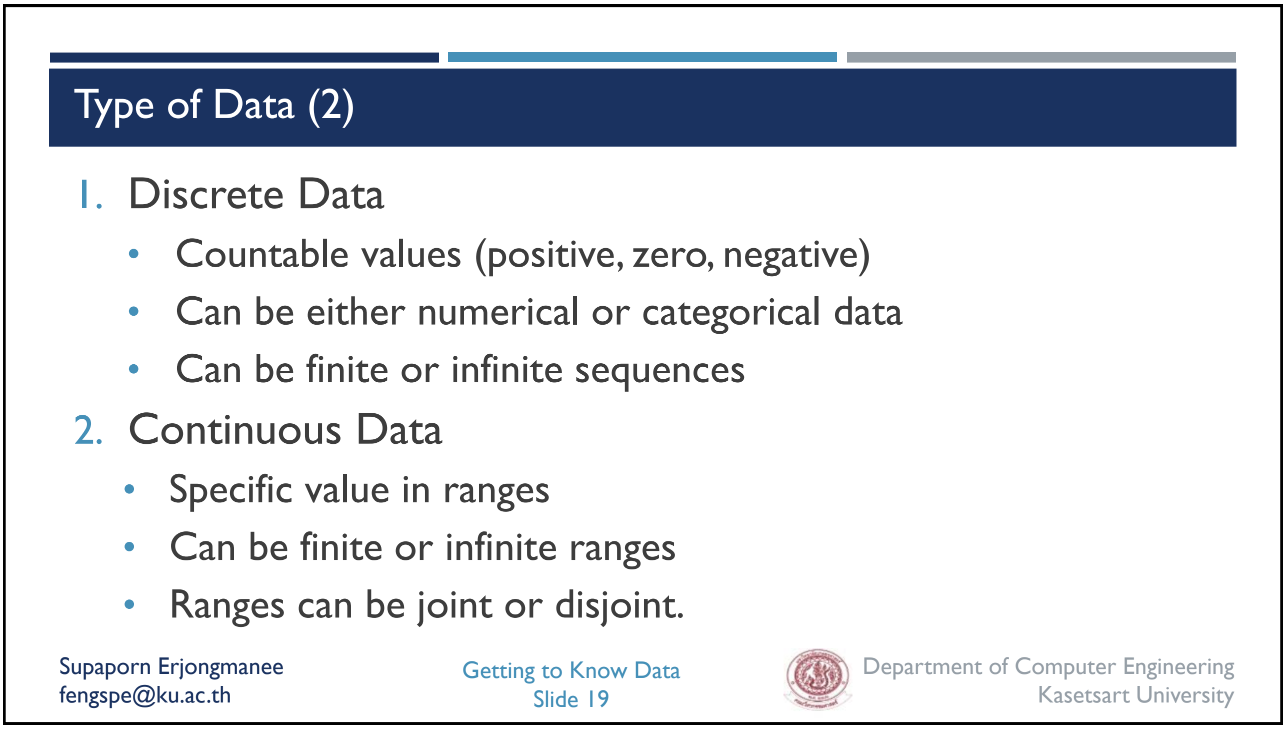


Comparison: Type of Data

- Comparison

	Nominal	Ordinal	Interval	Ratio
Can order values		✓	✓	✓
Can compute differences of values			✓	✓
Can add or subtract values			✓	✓
Can divide or multiple values				✓
Has fixed zero points				✓

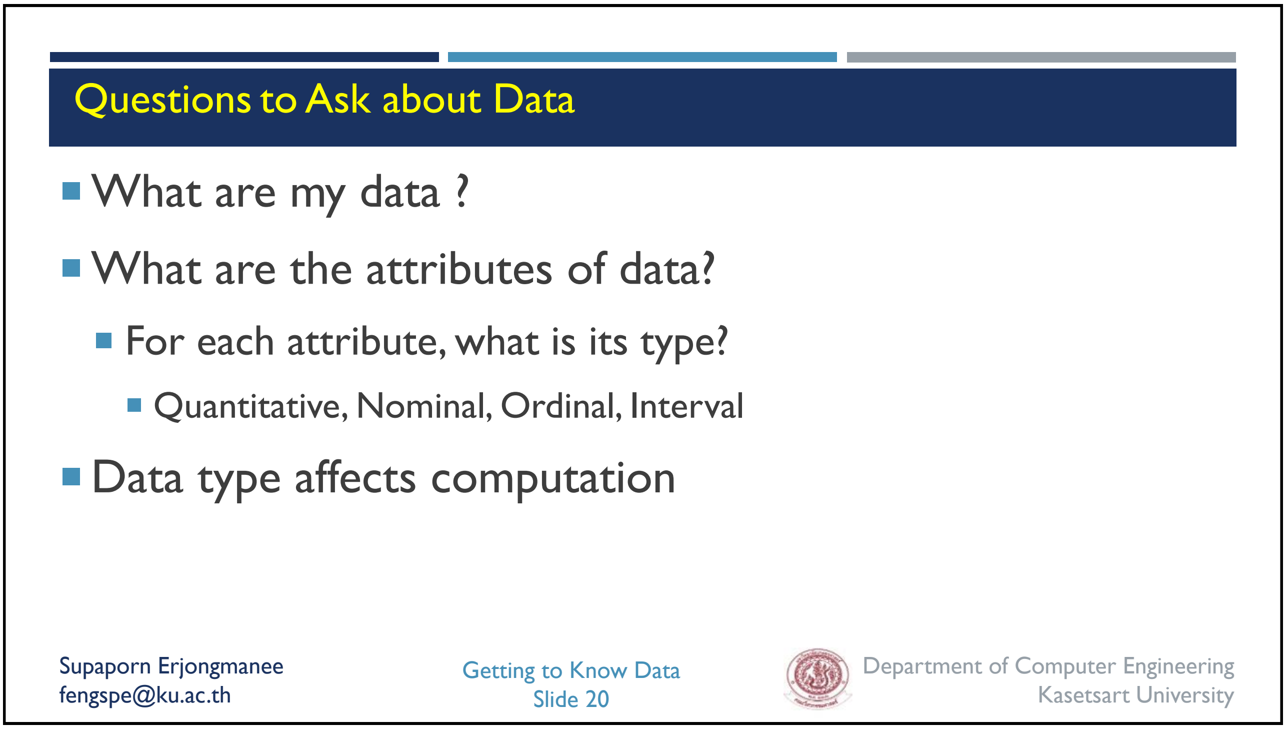
Image source: https://www.mymarketresearchmethods.com/types-of-data-nominal-ordinal-interval-ratio/



Type of Data (2)

1. Discrete Data
   - Countable values (positive, zero, negative)
   - Can be either numerical or categorical data
   - Can be finite or infinite sequences
2. Continuous Data
   - Specific value in ranges
   - Can be finite or infinite ranges
   - Ranges can be joint or disjoint.

19



Questions to Ask about Data

- What are my data ?
- What are the attributes of data?
  - For each attribute, what is its type?
    - Quantitative, Nominal, Ordinal, Interval
- Data type affects computation

20

Measure of Central Tendency

- Mode
- Median
- Mean

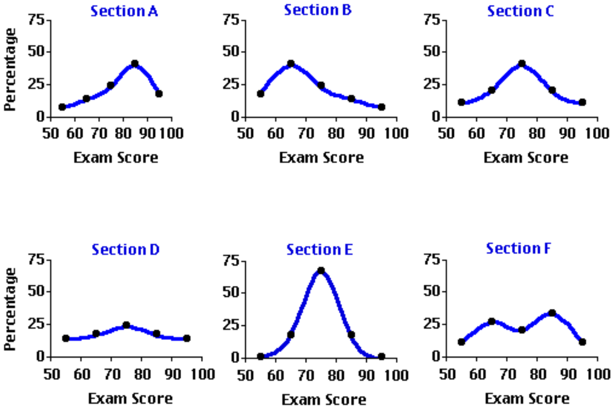
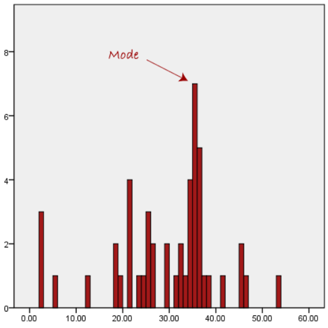
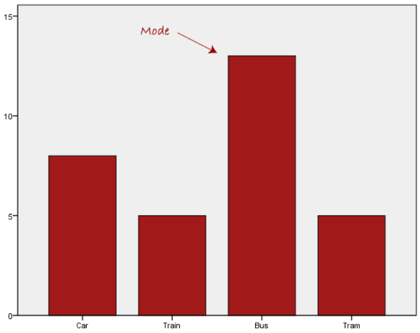


Image source: http://vassarstats.net/textbook/f0203x.gif

Mode

- Most frequent value
- Easily spot as a peak in histogram
- Good for nominal data

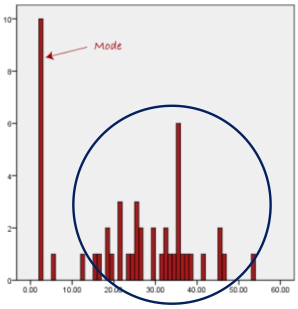


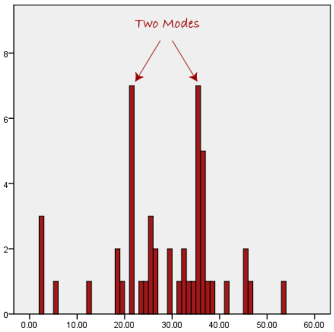


Source: [2]

Mode (cont.)

- Be careful when used with continuous data
  - Difficult to specify detailed value (e.g., 65.3)
- Avoid if mode is not with the majority





Source: [2]

Median

- Middle value in the sorted data
- Can be used with outliers or skewed data

Let n = data size

median = $(\frac{n+1}{2})^{th}$ value

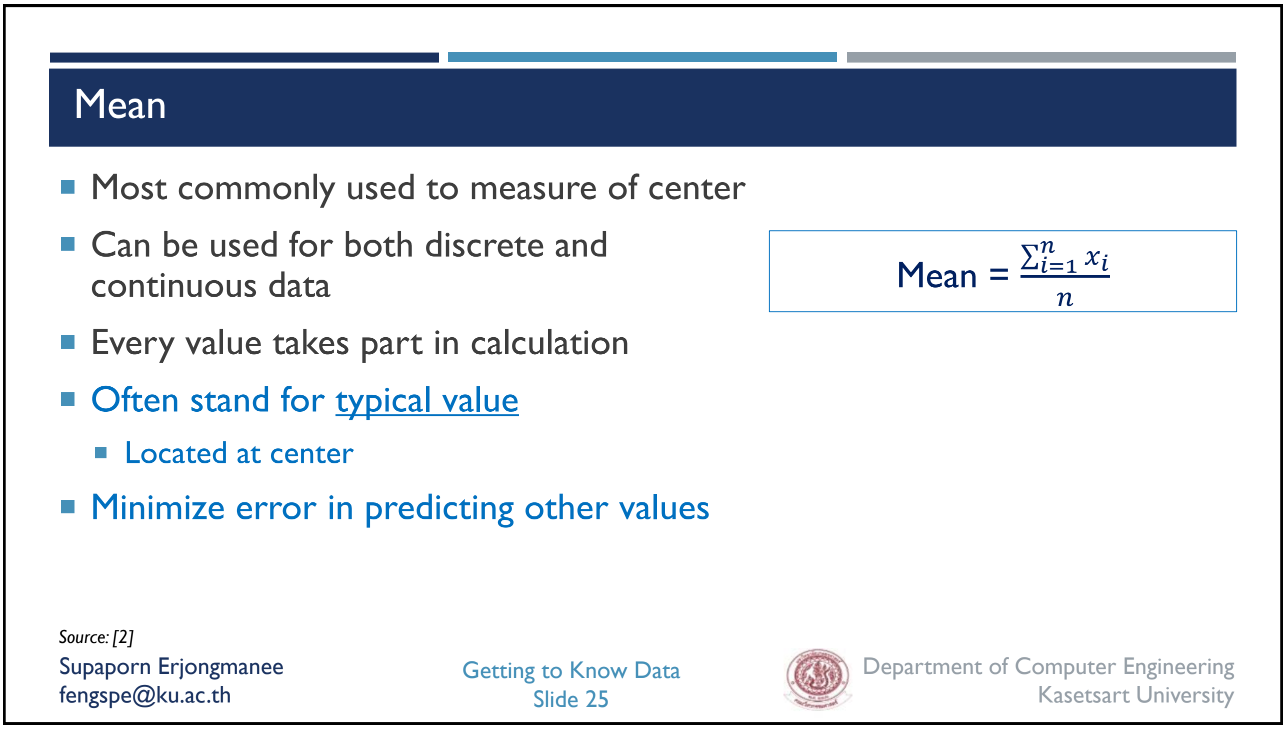
2, 2, 5, 7, 8, 10, 12

Median = 7

0, 2, 5, 6, 7, 8, 8, 8, 9, 10

$$\text{Median} = \frac{7+8}{2} = \frac{15}{2} = 7.5$$

Source: [2]

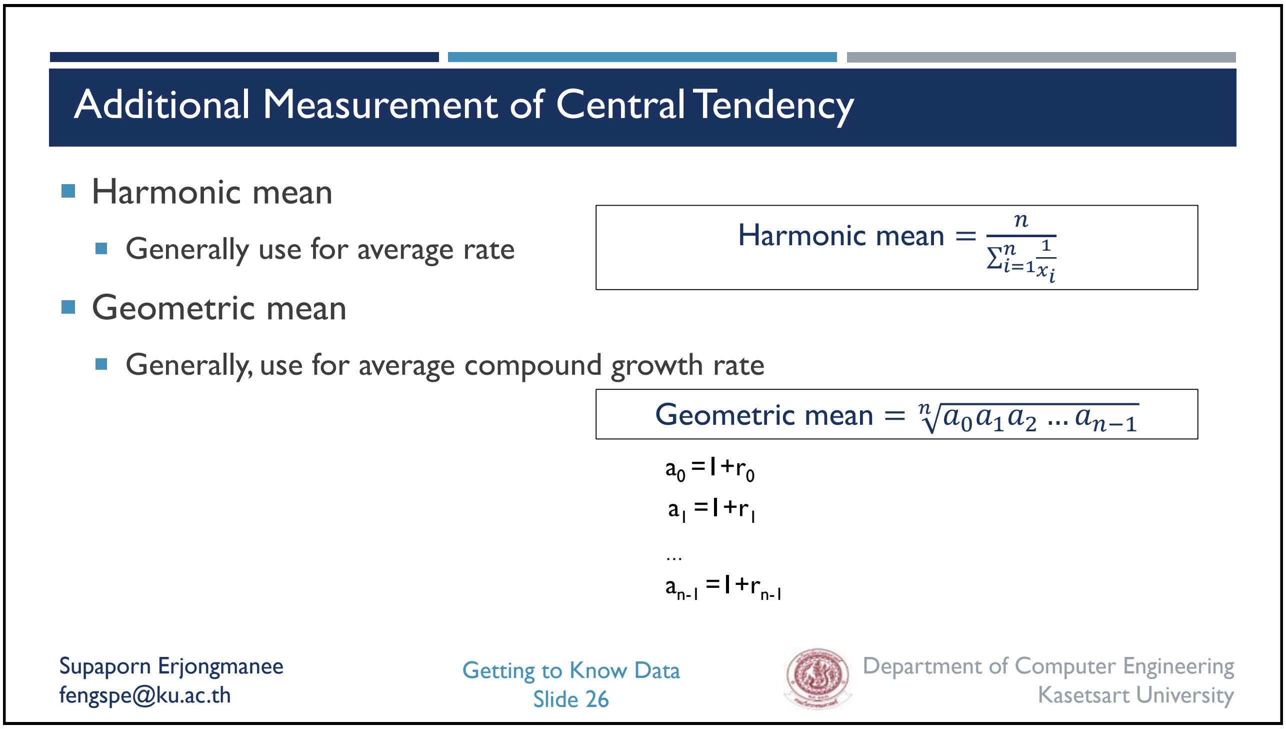


Mean

- Most commonly used to measure of center
- Can be used for both discrete and continuous data
- Every value takes part in calculation
- Often stand for typical value
  - Located at center
- Minimize error in predicting other values

$$\text{Mean} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Source: [2]



Additional Measurement of Central Tendency

- Harmonic mean
  - Generally use for average rate
- Geometric mean
  - Generally, use for average compound growth rate

$$\text{Harmonic mean} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

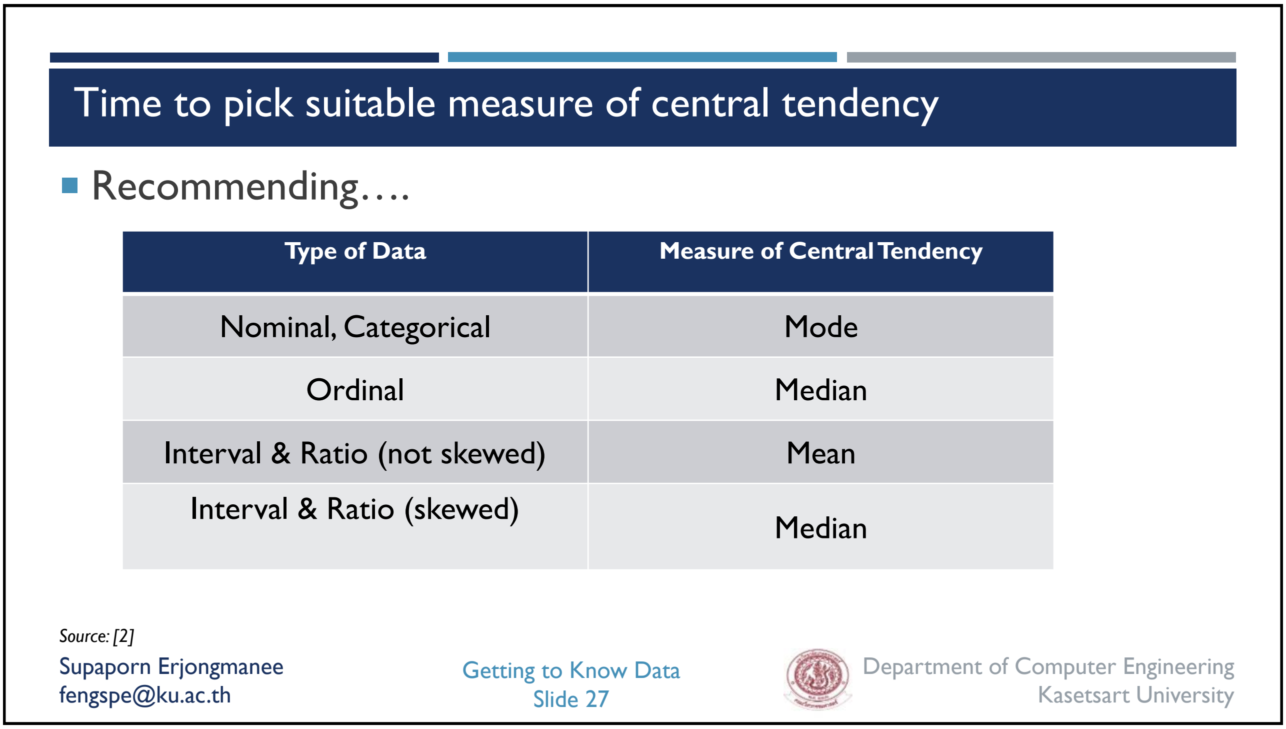
$$\text{Geometric mean} = \sqrt[n]{a_0 a_1 a_2 \ldots a_{n-1}}$$

$a_0 = 1+r_0$

$a_1 = 1+r_1$

…

$a_{n-1} = 1+r_{n-1}$



Time to pick suitable measure of central tendency

- Recommending….

Type of Data	Measure of Central Tendency
Nominal, Categorical	Mode
Ordinal	Median
Interval & Ratio (not skewed)	Mean
Interval & Ratio (skewed)	Median

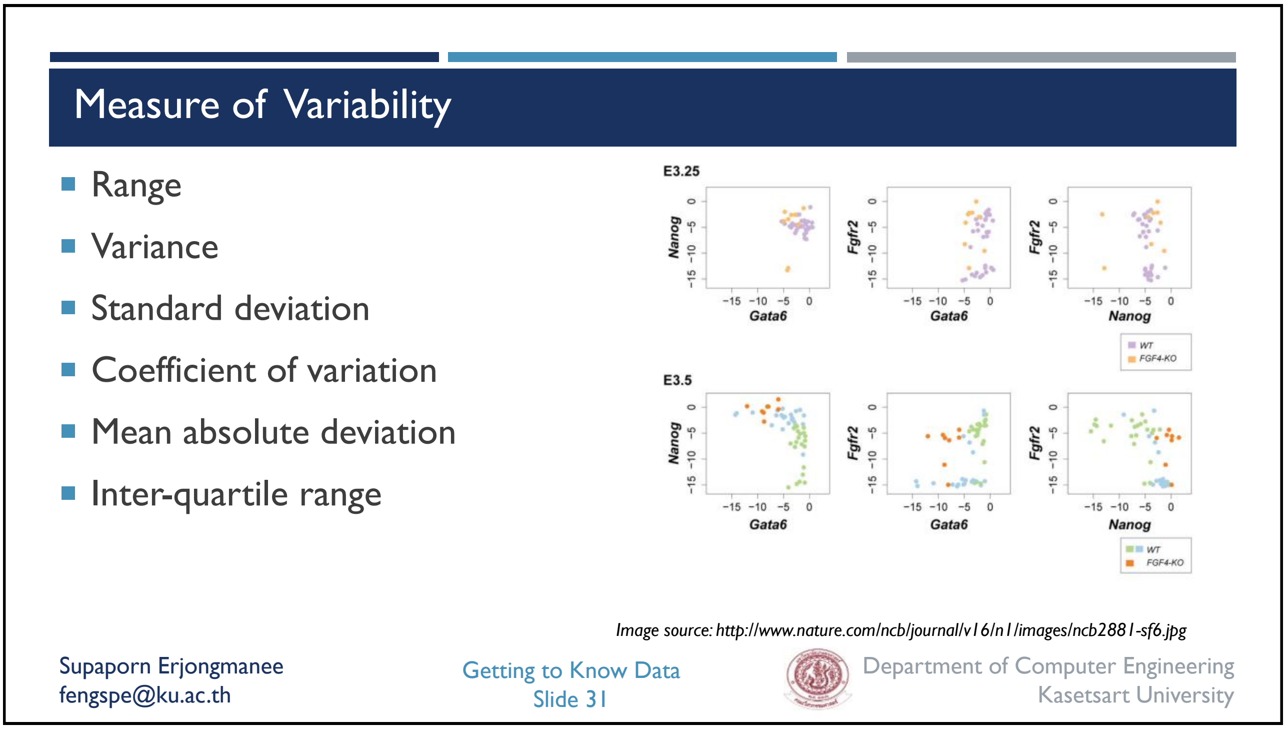
Source: [2]

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 27

Department of Computer Engineering
Kasetsart University

27



Measure of Variability

- Range
- Variance
- Standard deviation
- Coefficient of variation
- Mean absolute deviation
- Inter-quartile range

E3.25
Nanog
Gata6
Fgfr2
Gata6
Fgfr2
Nanog
−15 −10 −5 0
WT
FGF4-KO
E3.5
Nanog
Gata6
Fgfr2
Gata6
Fgfr2
Nanog
−15 −10 −5 0
WT
FGF4-KO

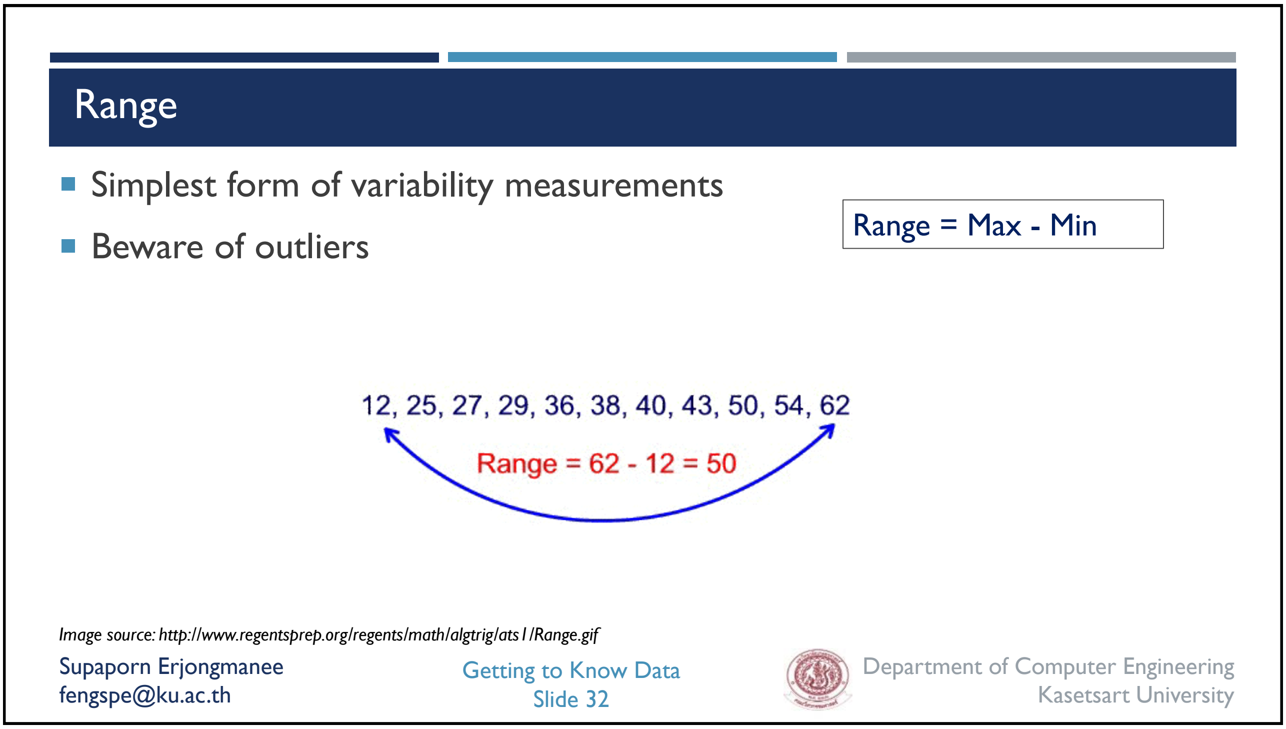
Image source: http://www.nature.com/ncb/journal/v16/n1/images/ncb2881-sf6.jpg

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 31

Department of Computer Engineering
Kasetsart University

31



Range

- Simplest form of variability measurements
- Beware of outliers

Range = Max - Min

12, 25, 27, 29, 36, 38, 40, 43, 50, 54, 62

Range = 62 - 12 = 50

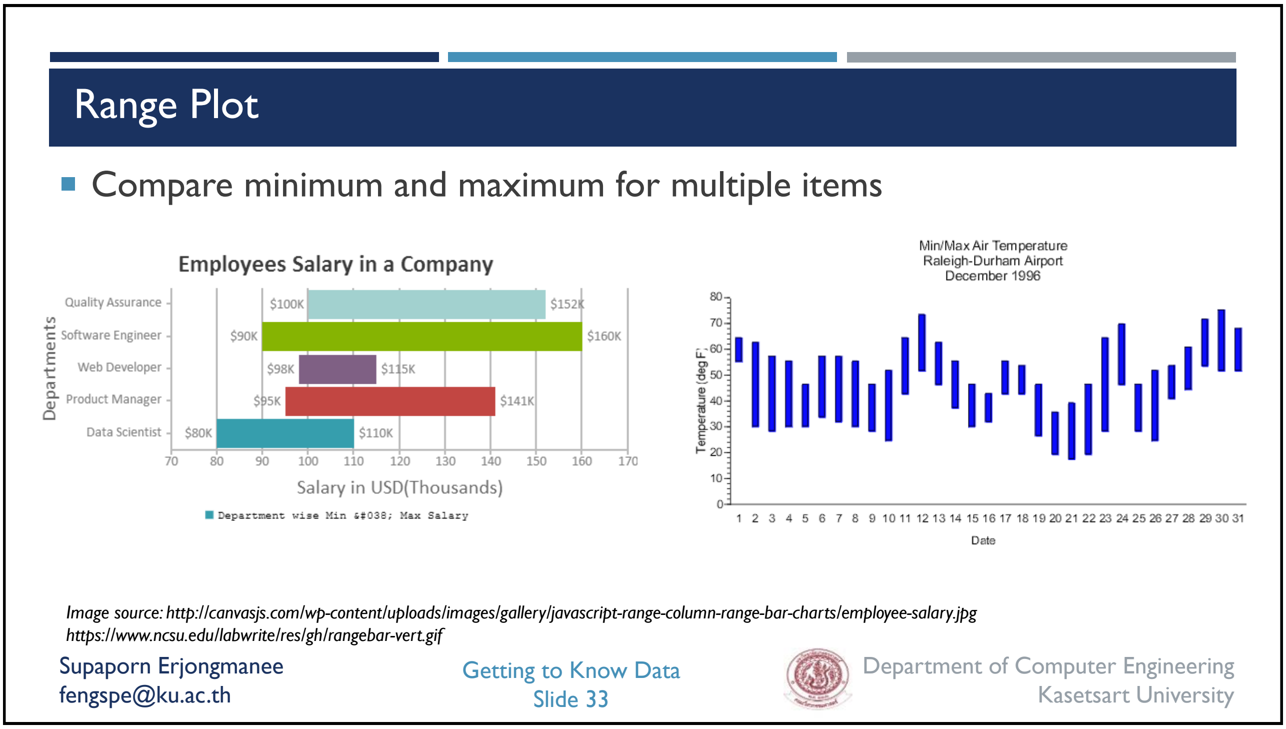
Image source: http://www.regentsprep.org/regents/math/algtrig/ats1/Range.gif

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 32

Department of Computer Engineering
Kasetsart University

32



Range Plot

- Compare minimum and maximum for multiple items

Image source: http://canvasjs.com/wp-content/uploads/images/gallery/javascript-range-column-range-bar-charts/employee-salary.jpg
https://www.ncsu.edu/labwrite/res/gh/rangebar-vert.gif

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 33

Department of Computer Engineering
Kasetsart University

33

Variance & Standard Deviation

- Average *difference (squared distance)* between all data and the mean
- Fit for
  - Continuous data
  - Quantitative data, not categorical data
- Avoid if data are skewed or have outliers
- Unit of variance is squared
- Standard deviation = $\sqrt{variance}$

Population variance

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

Sample variance

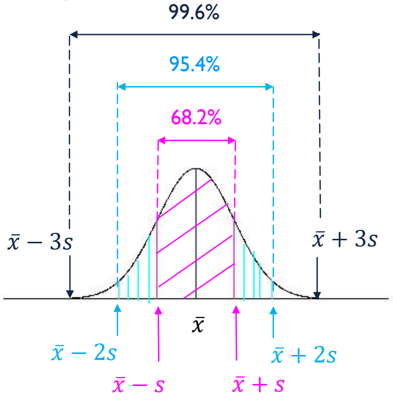
$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(\sum_{i=1}^{n} x_i^2) - n\bar{x}^2}{n-1}$$

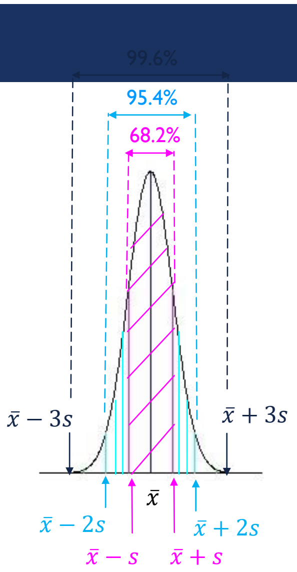
Sample variance (divided by n -1) is unbiased estimate of population variance

Source: [4]

Standard Deviation for Normal Data

- How does standard deviation tell us about spread of normal data?





Coefficient of Variation

- Ratio between standard deviation and mean

$$cv_{student} = \frac{s}{\bar{x}} = \frac{6.22}{174.54} = 0.0356$$

vs.

$$cv_{elephant} = 15.6$$

Elephants have more variability in height than student

Population

$$\text{Coefficient of variation} = \frac{\sigma}{\mu}$$

Sample

$$\text{Coefficient of variation} = \frac{s}{\bar{x}}$$

Mean Absolution Deviation

- Average *absolute distance* between all data and the mean

$$\text{Mean absolution deviation} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

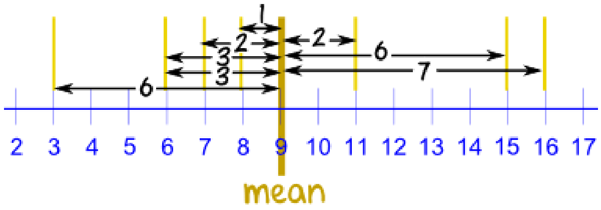
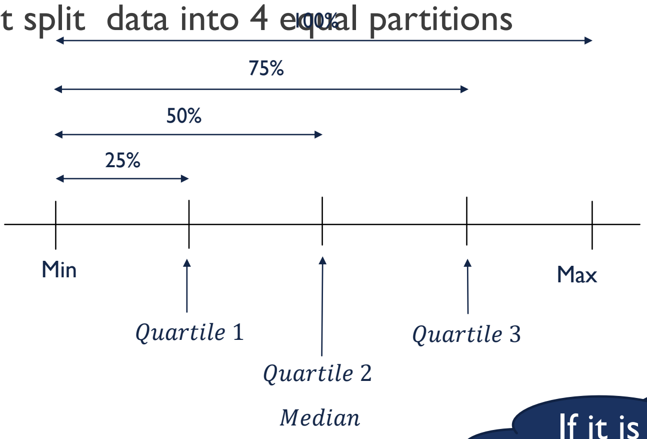


Image source: http://www.mathsisfun.com/data/images/mean-deviation.gif

Quartile

- Values that split data into 4 equal partitions



If it is regular number, call percentile

Interquartile Range (IQR)

- Good to use for data with outlier or skewed
- Not consider all data

Q1 = lower quartile = 25% percentile
Q3 =upper quartile = 75% percentile
IQR = Q3 - Q1

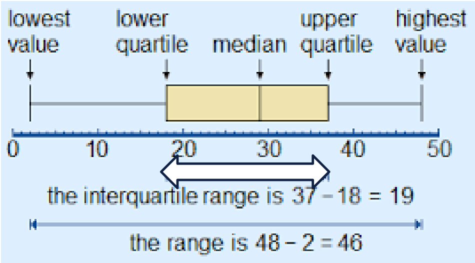


Image source: http://www.dr-aart.nl/Statistiek_bestanden/boxplot1EN.png

SD vs. IQR

- Comparison between normal distribution and box plots

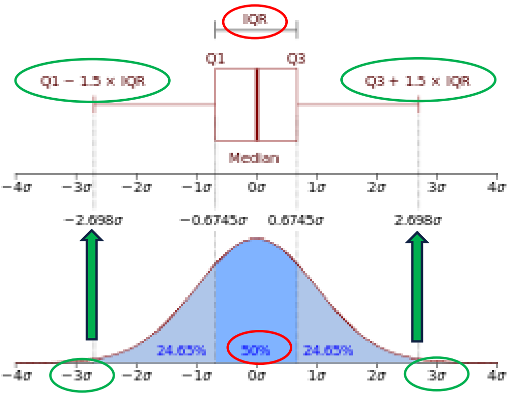
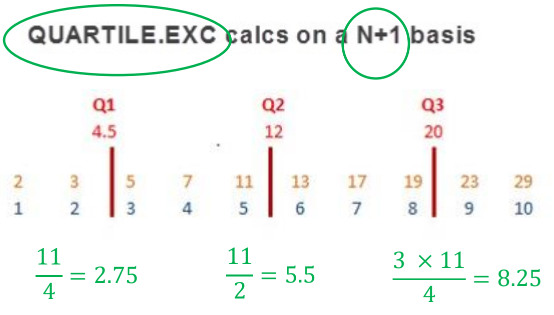


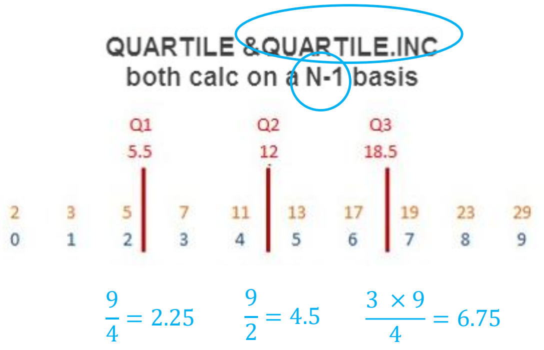
Image source: https://upload.wikimedia.org/wikipedia/commons/thumb/1/1a/Boxplot_vs_PDF.svg/250px-Boxplot_vs_PDF.svg.png

Quartile in Excel

- Quartile.exc vs. Quartile.inc



$\frac{11}{4} = 2.75$ $\frac{11}{2} = 5.5$ $\frac{3 \times 11}{4} = 8.25$



$\frac{9}{4} = 2.25$ $\frac{9}{2} = 4.5$ $\frac{3 \times 9}{4} = 6.75$

Source: http://datapigtechnologies.com/blog/index.php/why-excel-has-multiple-quartile-functions-and-how-to-replicate-the-quartiles-from-r-and-other-statistical-packages/

Which Measurement of Variability to Use?

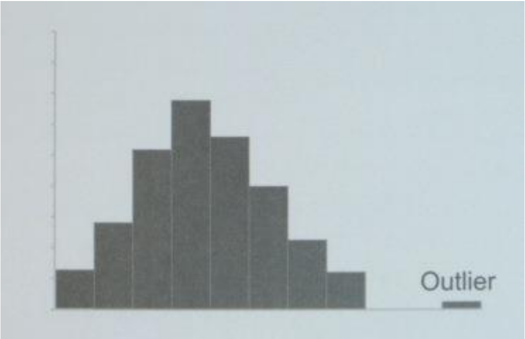
- Range → Easiest to use. Not suitable for data with outliers
- Sample variance
- Sample standard deviation

} Most commonly-used

- Inter quartile range → Good for data with outliers
- Coefficient of variation → Tell more story about the data:
  how std is compared to the mean
  No unit
  Sensitive when mean → 0
  Not suitable for multiple replicates of data

OUTLIERS

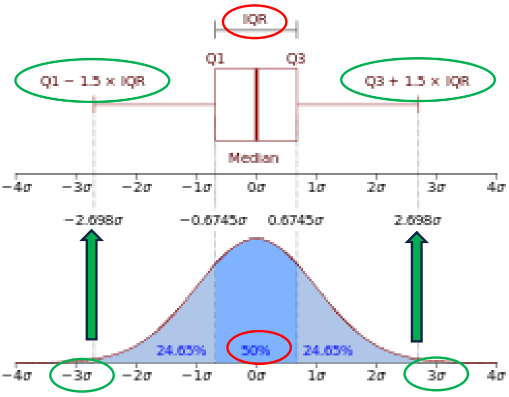
- Out-of-the-norm data
- Threshold is needed to cut outliers



Source: [2]

Outliers

- Outliers can be determined from IQR or SD



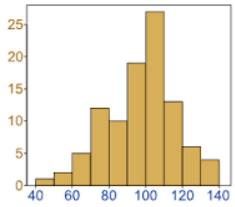
If data value > Q3 + 1.5IQR or
data value < Q1 - 1.5IQR,
we consider such value to be outlier.

If data value > mean + 3SD or
data value < mean -3SD,
we consider such value to be outlier.

Image source: https://upload.wikimedia.org/wikipedia/commons/thumb/1/1a/Boxplot_vs_PDF.svg/250px-Boxplot_vs_PDF.svg.png

Are descriptive statistics enough?

- Descriptive statistics are not answer to everything
- Be careful of outlier and skewed data
- Always GRAPH your data
  - Histogram
  - Boxplot



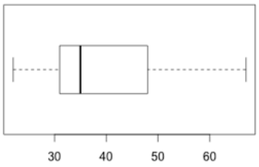


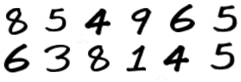
Image sources: https://www.mathsisfun.com/data/images/histogram.gif
http://www.johnquarto.com/wp-content/uploads/2013/09/Boxplot-PartyPeopleAll.png

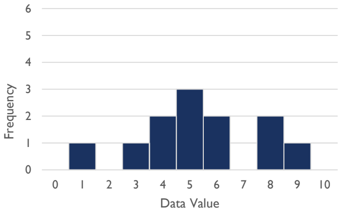
Basic Data Visualization

- Histogram
- Boxplot
- Scatter plot

Histogram

- Specific bar graph representing distribution of data
- x-axis: bins of data values
- y-axis: frequency of data values
- Example:





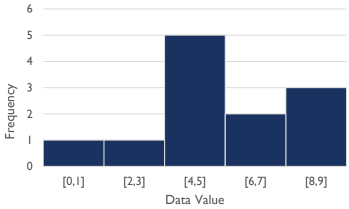


Image source: https://openclipart.org

Box Plots

- Also call box-and-whisker plot
- Use statistical values to plot distribution of data

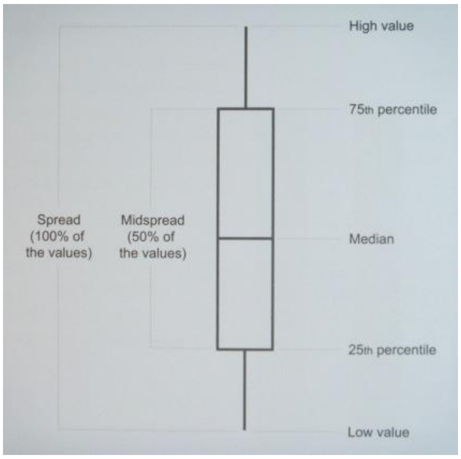
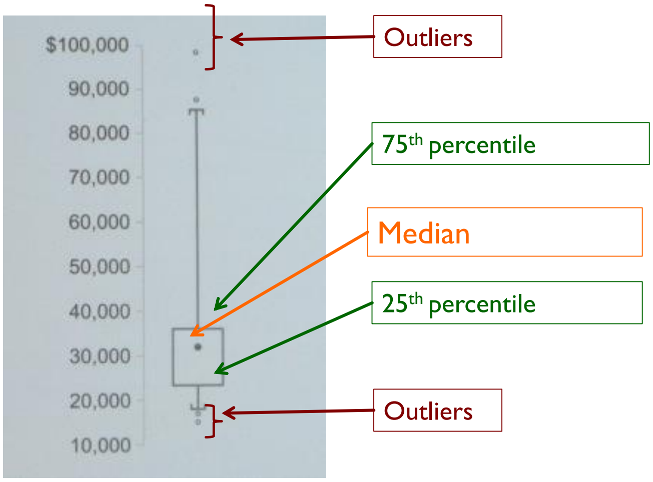


Image source: Figures 10.34 [1]

Box Plots (cont.)



Outliers

75th percentile

Median

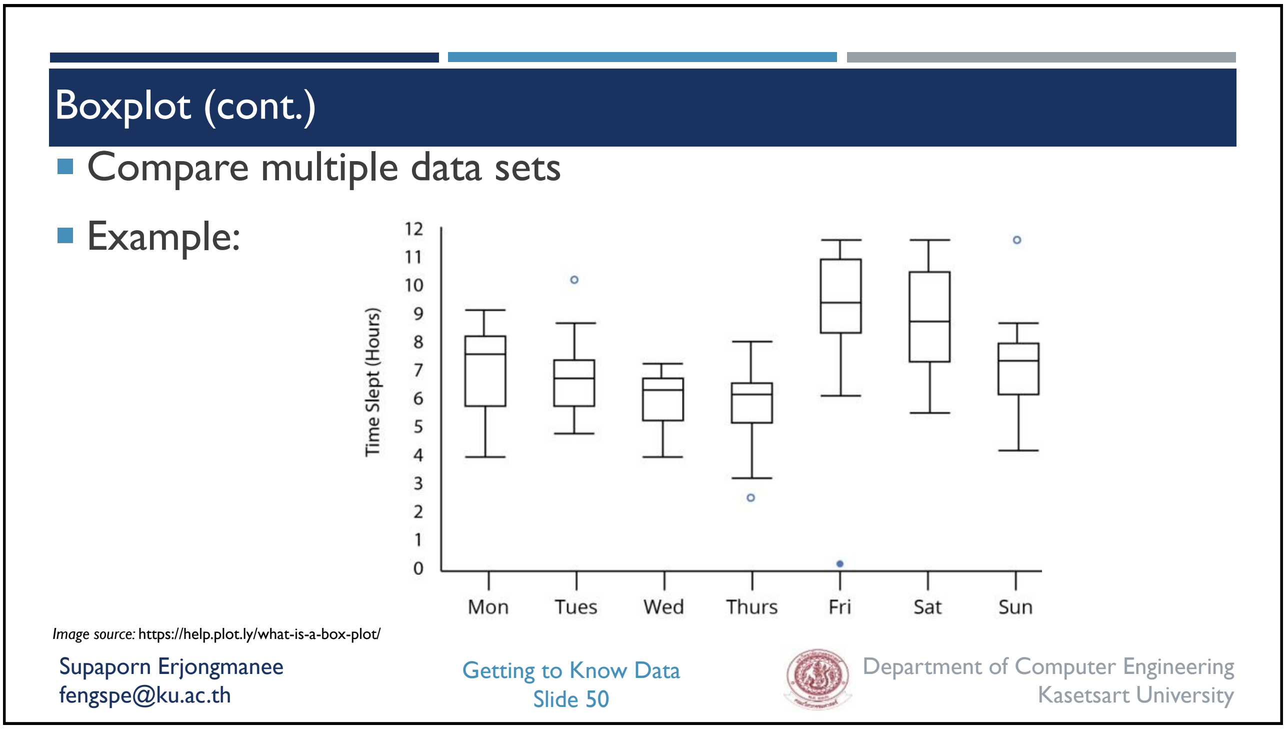
25th percentile

Outliers

- More detailed box plot

Definition of "outliers" must be given.

Image source: Figure 10.38, [1]

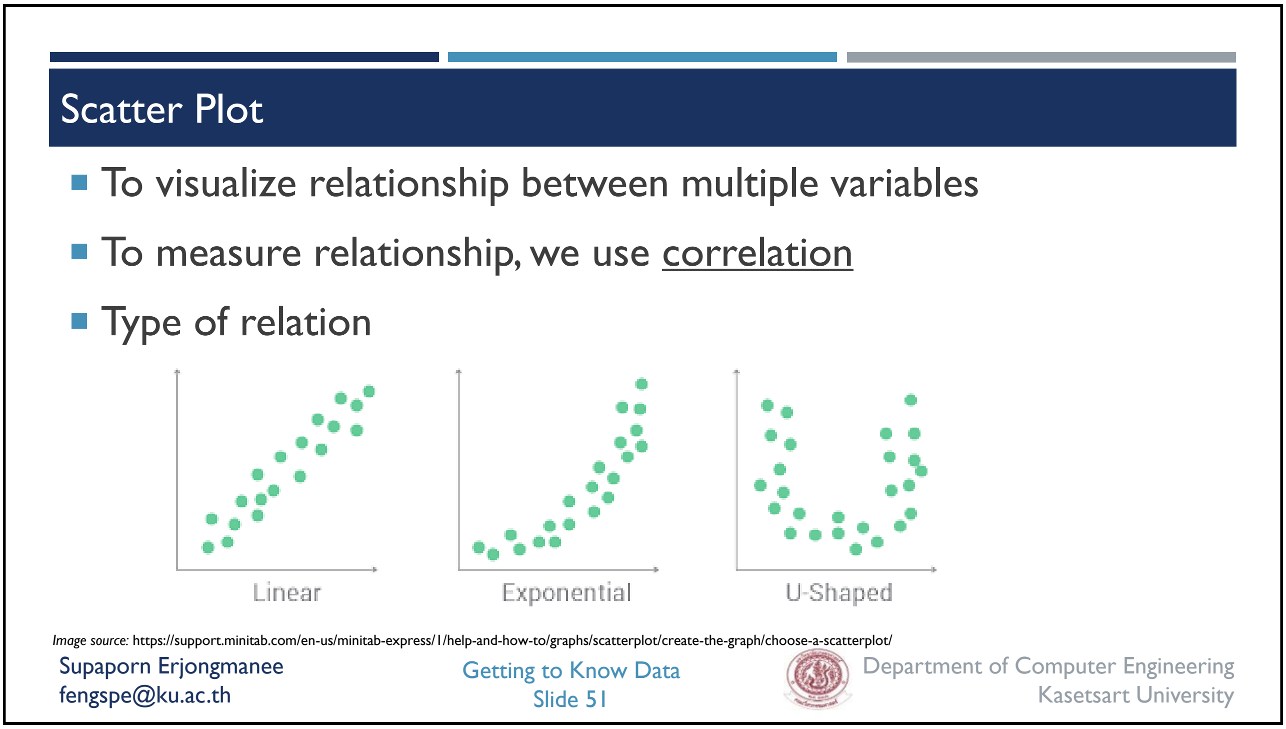


Boxplot (cont.)

- Compare multiple data sets
- Example:

Image source: https://help.plot.ly/what-is-a-box-plot/

50

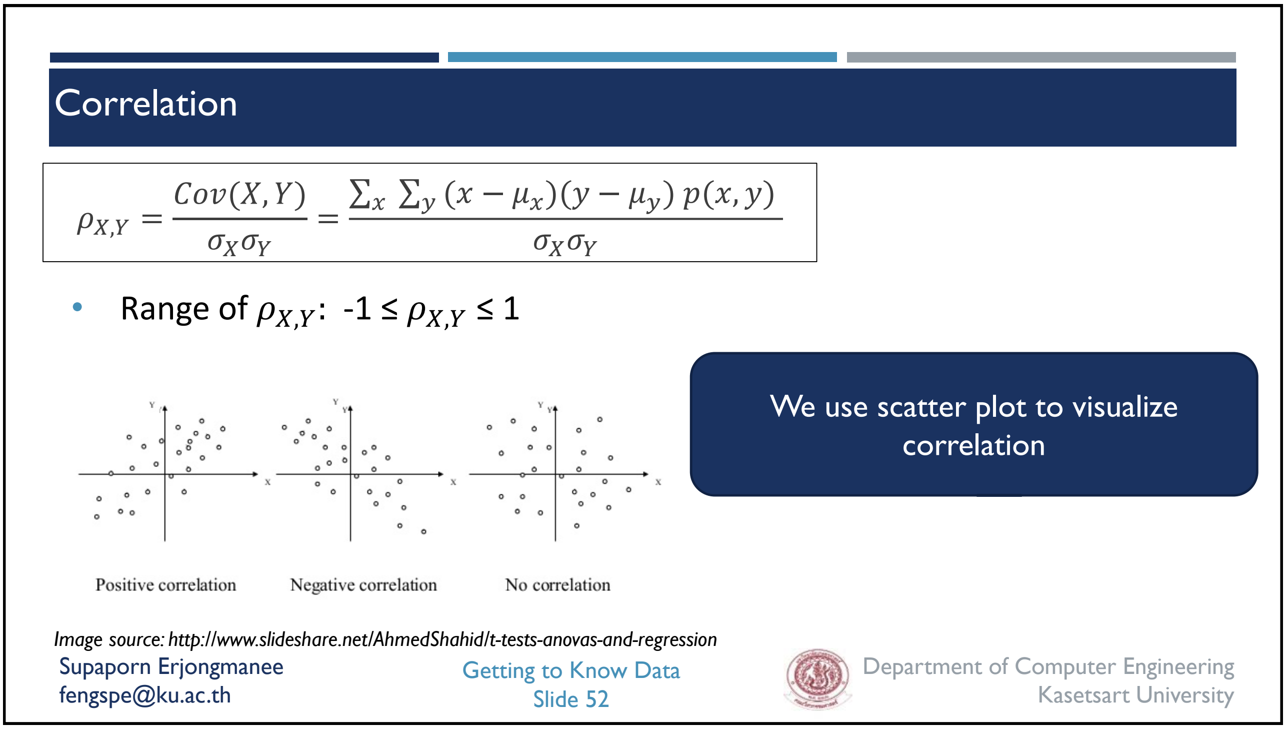


Scatter Plot

- To visualize relationship between multiple variables
- To measure relationship, we use <u>correlation</u>
- Type of relation

Image source: https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/graphs/scatterplot/create-the-graph/choose-a-scatterplot/

51



Correlation

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = \frac{\sum_x \sum_y (x-\mu_x)(y-\mu_y)\,p(x,y)}{\sigma_X\sigma_Y}$$

- Range of $\rho_{X,Y}$: $-1 \leq \rho_{X,Y} \leq 1$

We use scatter plot to visualize correlation

Positive correlation Negative correlation No correlation

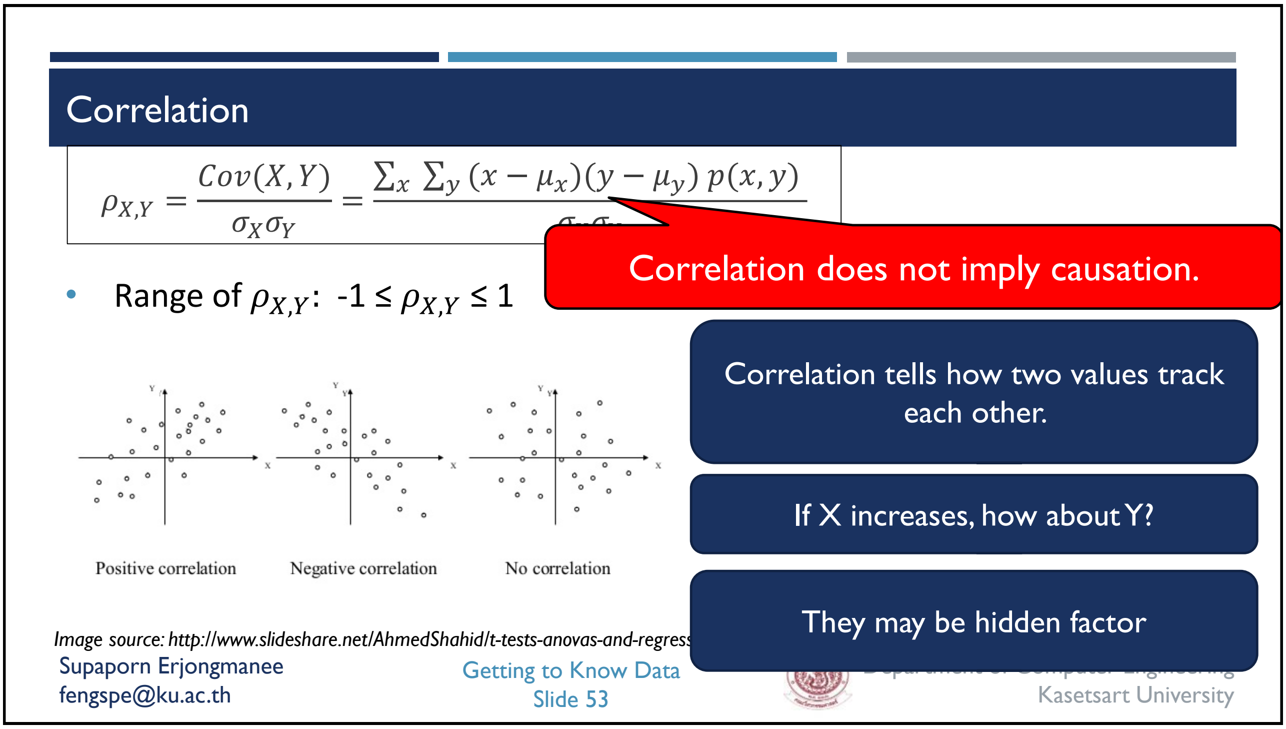
Image source: http://www.slideshare.net/AhmedShahid/t-tests-anovas-and-regression

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 52

Department of Computer Engineering
Kasetsart University

52



Correlation

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = \frac{\sum_x \sum_y (x-\mu_x)(y-\mu_y)\,p(x,y)}{\sigma_X\sigma_Y}$$

- Range of $\rho_{X,Y}$: $-1 \leq \rho_{X,Y} \leq 1$

Correlation does not imply causation.

Correlation tells how two values track each other.

If X increases, how about Y?

They may be hidden factor

Positive correlation Negative correlation No correlation

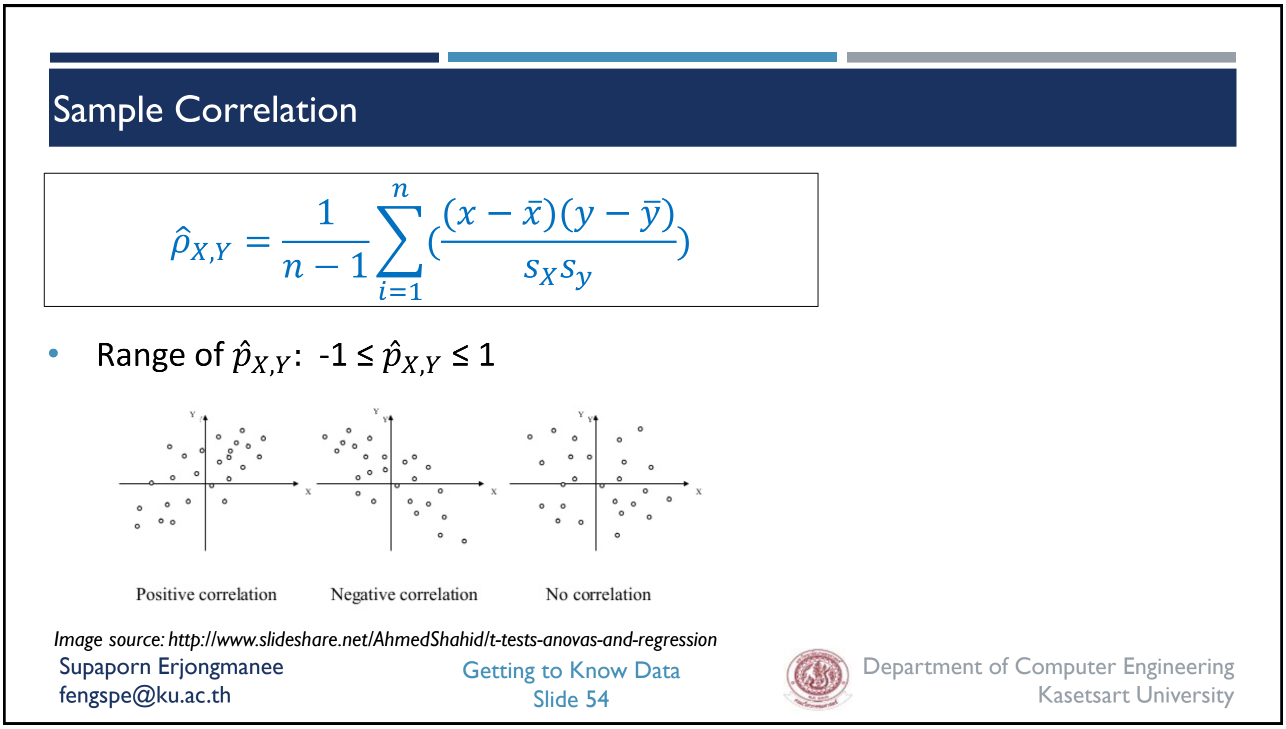
Image source: http://www.slideshare.net/AhmedShahid/t-tests-anovas-and-regress

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 53

Kasetsart University

53



Sample Correlation

$$\hat{\rho}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{(x-\bar{x})(y-\bar{y})}{s_X s_y})$$

- Range of $\hat{p}_{X,Y}$: -1 ≤ $\hat{p}_{X,Y}$ ≤ 1

Positive correlation Negative correlation No correlation

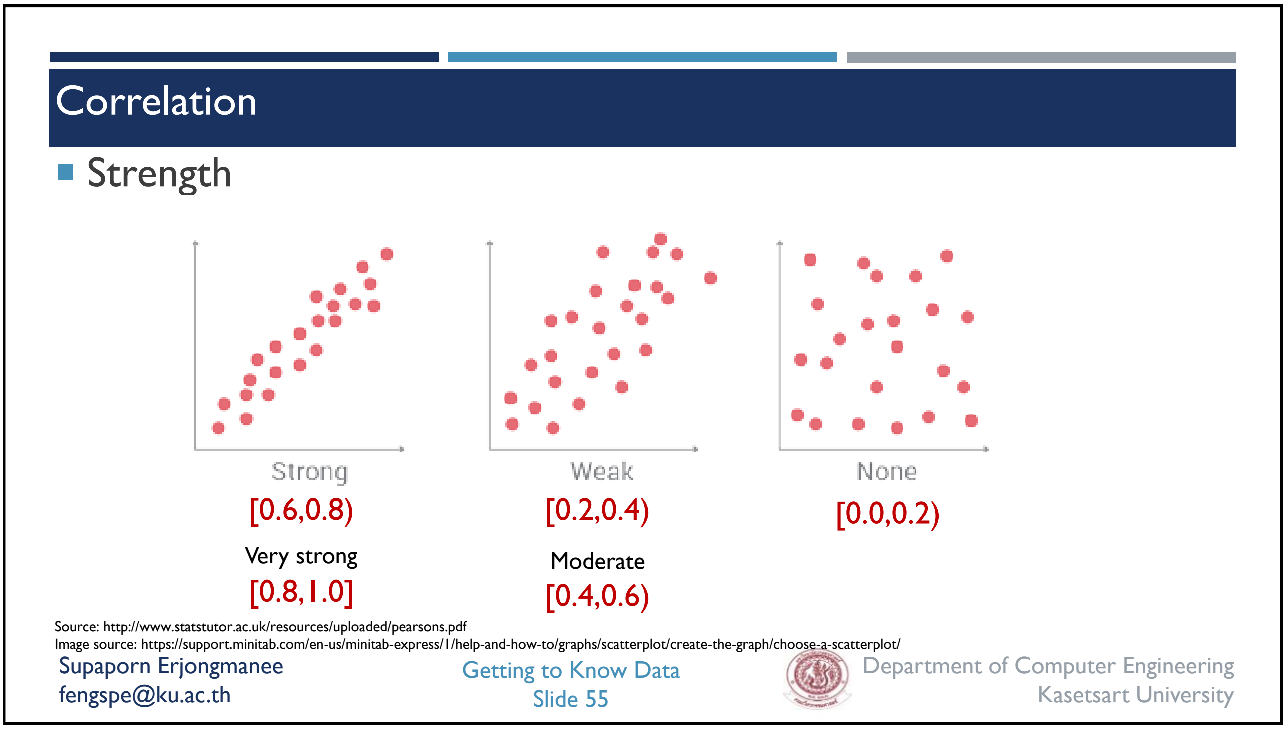
Image source: http://www.slideshare.net/AhmedShahid/t-tests-anovas-and-regression

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 54

Department of Computer Engineering
Kasetsart University

54



Correlation

- Strength

Strong
[0.6,0.8)

Very strong
[0.8,1.0]

Weak
[0.2,0.4)

Moderate
[0.4,0.6)

None
[0.0,0.2)

Source: http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf
Image source: https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/graphs/scatterplot/create-the-graph/choose-a-scatterplot/

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 55

Department of Computer Engineering
Kasetsart University

55

Correlation

- Example: 3 variables (share same X,Y variables)

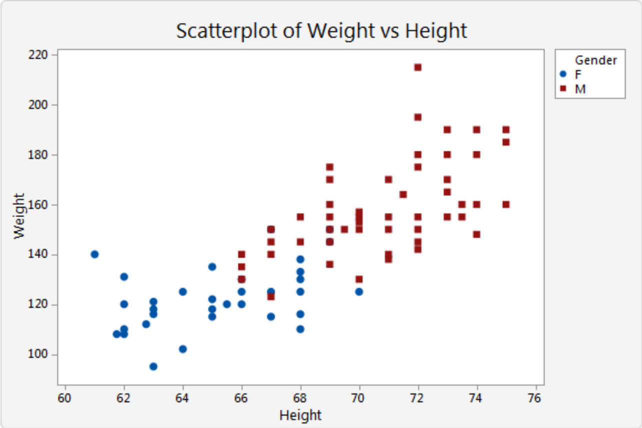
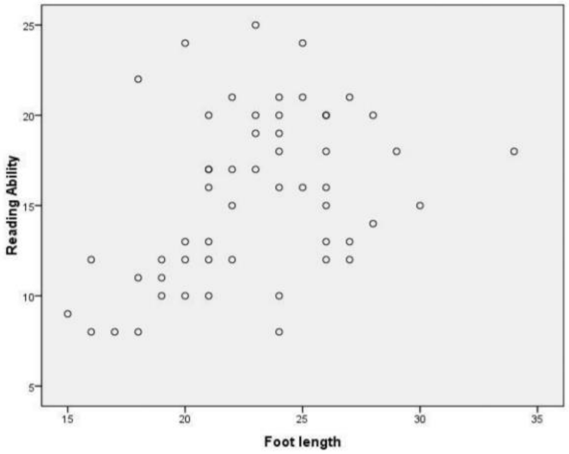


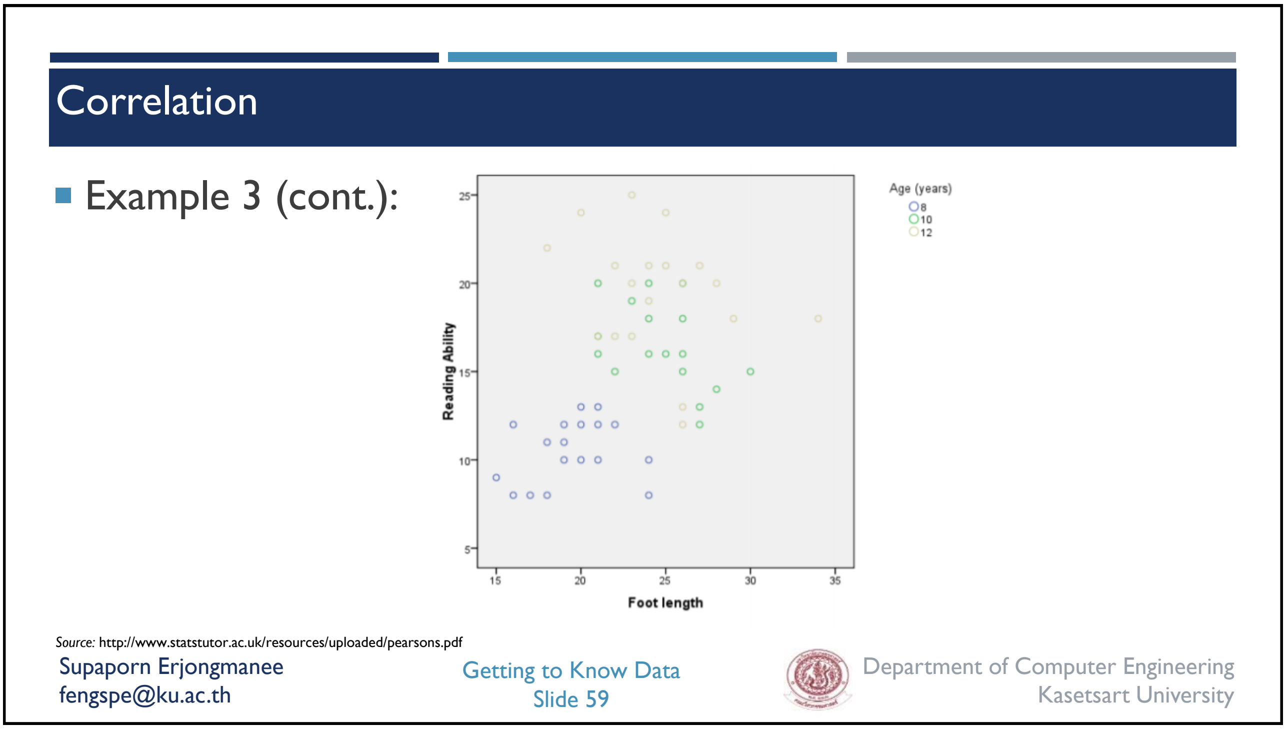
Image source: https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/graphs/scatterplot/create-the-graph/choose-a-scatterplot/

Correlation

- Example 3:



Source: http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf



Correlation

- Example 3 (cont.):

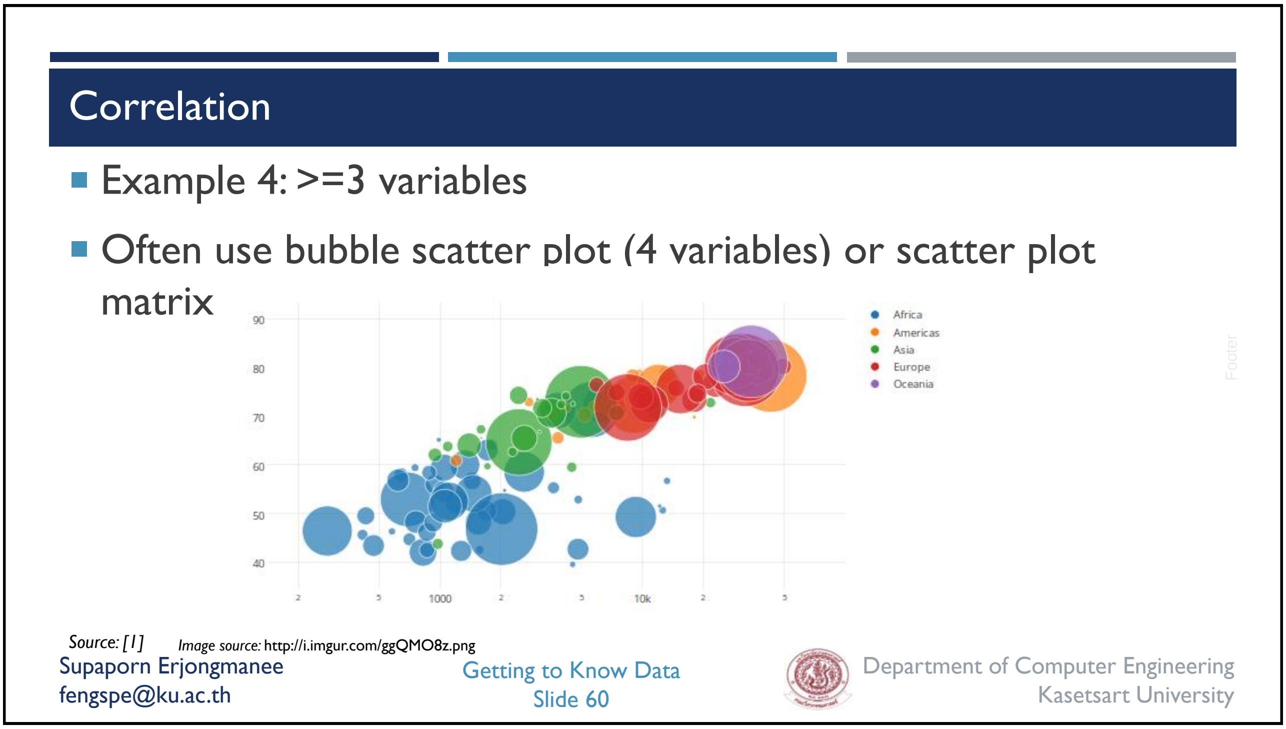
Source: http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 59

Department of Computer Engineering
Kasetsart University

59



Correlation

- Example 4: >=3 variables
- Often use bubble scatter plot (4 variables) or scatter plot matrix

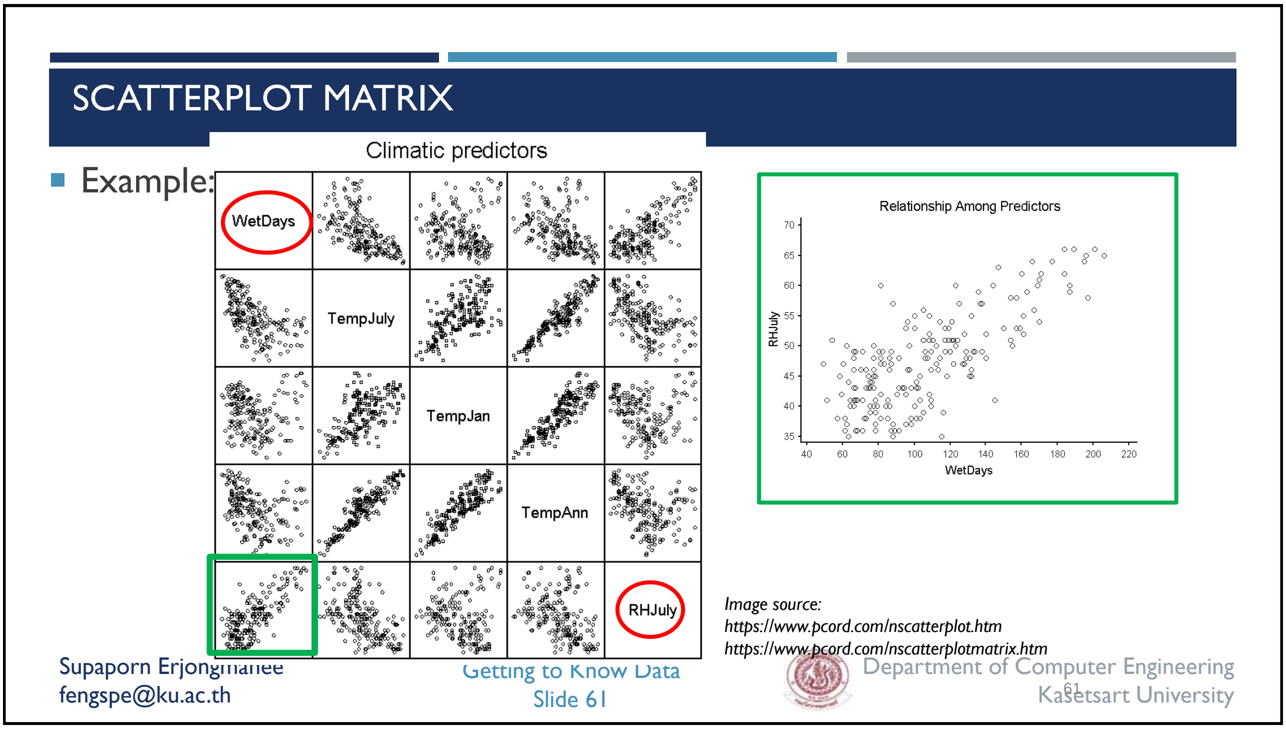
Source: [1] *Image source:* http://i.imgur.com/ggQMO8z.png

Supaporn Erjongmanee
fengspe@ku.ac.th

Getting to Know Data
Slide 60

Department of Computer Engineering
Kasetsart University

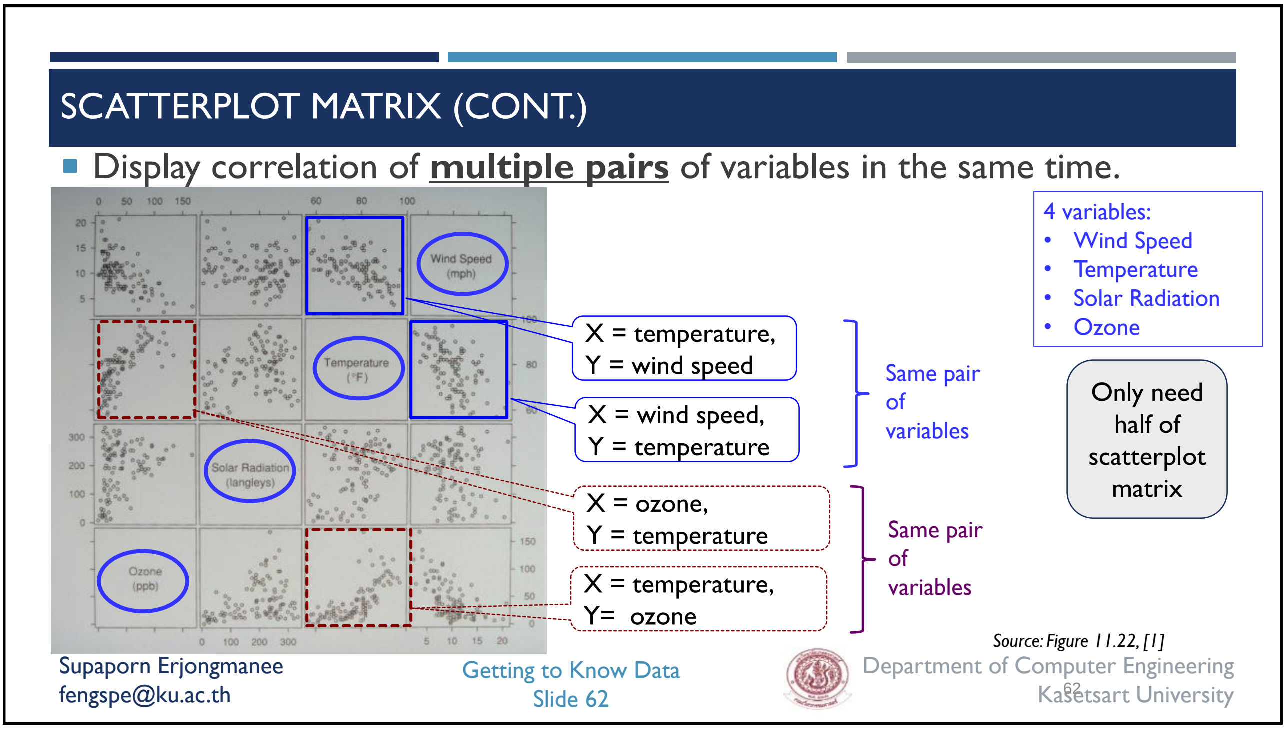
60



SCATTERPLOT MATRIX

- Example:

Image source:
https://www.pcord.com/nscatterplot.htm
https://www.pcord.com/nscatterplotmatrix.htm



SCATTERPLOT MATRIX (CONT.)

- Display correlation of **multiple pairs** of variables in the same time.

X = temperature,
Y = wind speed

X = wind speed,
Y = temperature

Same pair of variables

X = ozone,
Y = temperature

X = temperature,
Y= ozone

Same pair of variables

4 variables:
- Wind Speed
- Temperature
- Solar Radiation
- Ozone

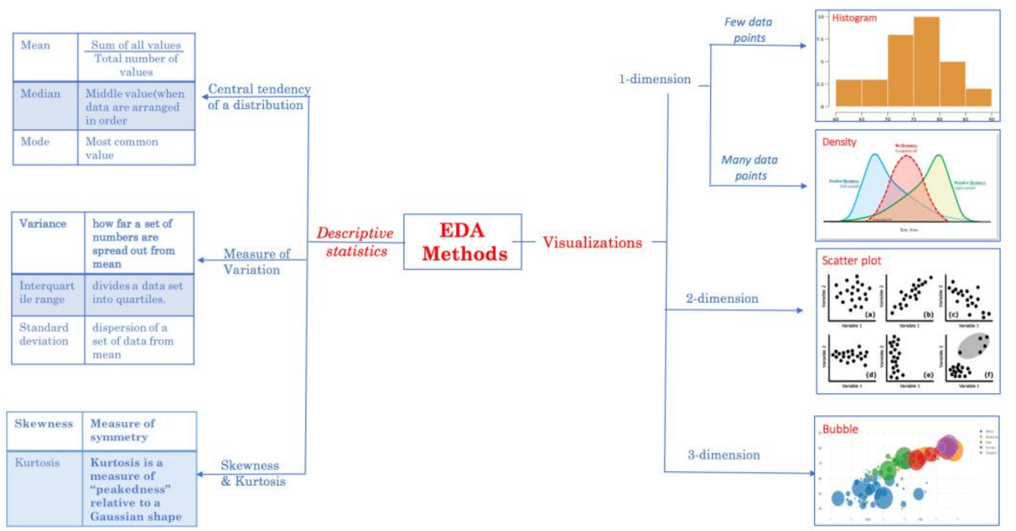
Only need half of scatterplot matrix

Source: Figure 11.22, [1]

Summary

- To first explore data, we can find
  - Outliers
  - Centrality: Mean, Median, Mode
  - Variability: Range, Variance, Standard Deviation, Coefficient of Variation, Mean Absolute Deviation, Interquartile Range
  - Correlation
  - Visualization: Histogram, Boxplot, Scatter Plot

Sum

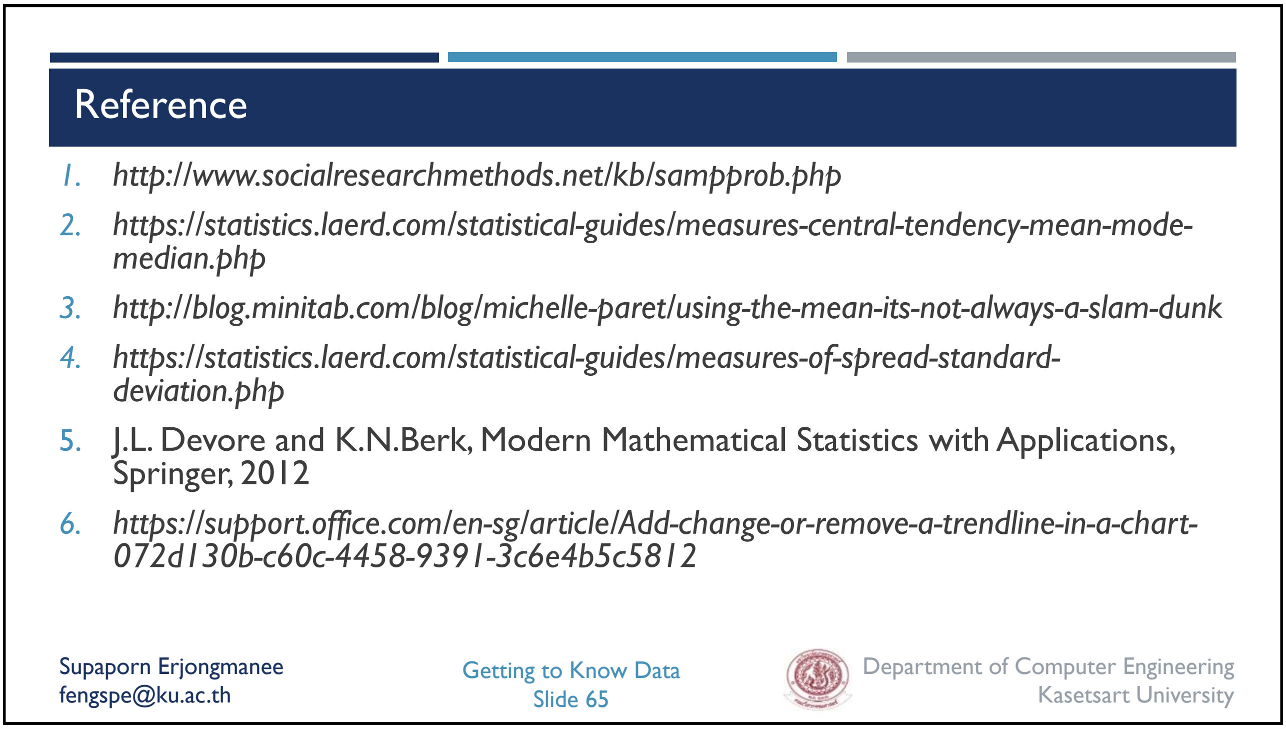


Mean	$\frac{\text{Sum of all values}}{\text{Total number of values}}$
Median	Middle value(when data are arranged in order
Mode	Most common value

Variance	how far a set of numbers are spread out from mean
Interquartile range	divides a data set into quartiles.
Standard deviation	dispersion of a set of data from mean

Skewness	Measure of symmetry
Kurtosis	Kurtosis is a measure of "peakedness" relative to a Gaussian shape

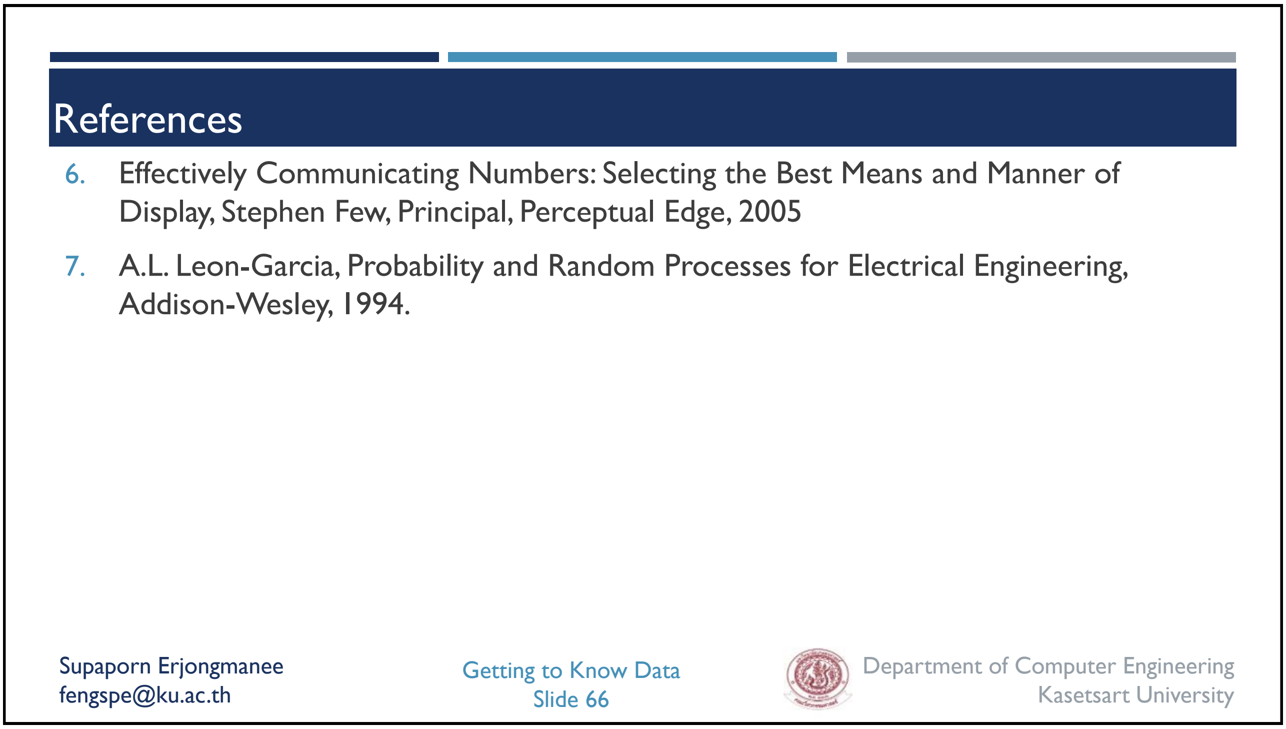
Image Source: https://www.netsolutions.com/insights/data-visualization-a-building-block-of-an-intelligent-enterprise/



Reference

1. *http://www.socialresearchmethods.net/kb/sampprob.php*
2. *https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php*
3. *http://blog.minitab.com/blog/michelle-paret/using-the-mean-its-not-always-a-slam-dunk*
4. *https://statistics.laerd.com/statistical-guides/measures-of-spread-standard-deviation.php*
5. J.L. Devore and K.N.Berk, Modern Mathematical Statistics with Applications, Springer, 2012
6. *https://support.office.com/en-sg/article/Add-change-or-remove-a-trendline-in-a-chart-072d130b-c60c-4458-9391-3c6e4b5c5812*

65



References

6. Effectively Communicating Numbers: Selecting the Best Means and Manner of Display, Stephen Few, Principal, Perceptual Edge, 2005
7. A.L. Leon-Garcia, Probability and Random Processes for Electrical Engineering, Addison-Wesley, 1994.

66