RESEARCH METHODOLOGY
URSP 0010

Results and Data Analysis

Lecturers:

Dr Fatimah Salim (sfatimah.kl@utm.my)

Dr Hazilah Mad Kaidi (hazilah.kl@utm.my)

Dr Fazilah Mohd Hashim(fazilah.kl@utm.my)

8th November 2016

Lecture Room 2, Level 15, Menara Razak

OUTLINE

- Session Objectives
- Type of Research:
- Research Process
- Research Instrument
- Data Gathering
- Validity and Reliability Concept
- Results and analysis-
  - observation and interpretation,
  - error analysis
  - validity and reliability of data
  - result significance of findings,
  - conclusions
  - and recommendation for further work

Session Objectives

To be able to:

- **Present the results** of your collected data.
- Make **good analysis** of the tabulated or graphically presented data.
- Make **effective interpretation** of the data/finding/results.
- **Draw implications or inferences** and generations from the analysis and interpretation of findings.

Type of Research

Application

- Applied Research
- Pure Research

Objectives

- Descriptive Research
- Correlational Research
- Exploratory Research
- Explanatory Research

Enquiry Mode

- Quantitative Research
- Qualitative Research

Type of Research: Application Perspective

- Pure research involves developing and testing theories and hypotheses that are intellectually challenging to the researcher but may or may not have practical application at the present time or in the future.
- Pure research is also concerned with the development, examination, verification and refinement of research methods, procedures, techniques and tools that form the body of research methodology.
- Examples of pure research:
  1. developing a sampling technique that can be applied to a particular situation;
  2. developing a methodology to assess the validity of a procedure;
  3. developing an instrument to measure the stress level in people;
  4. finding the best way of measuring people's attitudes.
- Most of the research in the social sciences is applied.

Type of Research: Objectives Perspective

Type of Research	Main Theme	Example
descriptive research	To describe what is prevalent	• Attitudes of students towards quality of teaching. • Problems faced by new immigrants.
correlational research	To ascertain if there is a relationship	• Relationship between stressful living and the incidence of heart attack. • Effectiveness of an immunisation programme in controlling infectious. • Impact of an advertising campaign on sale of a product.
explanatory research	To explain why the relationship is formed	• Why stressful living results in heart attacks? • How does technology create unemployment / employment?

Type of Research: Objectives Perspective

- The fourth type of research, from the viewpoint of the objectives of a study, is called **exploratory research**.
- This is when a study is undertaken with the objective either to explore an area where little is known or to investigate the possibilities of undertaking a particular research study.
- When a study is carried out to determine its feasibility it is also called a **feasibility study** or a **pilot study**.
- It is usually carried out when a researcher wants to explore areas about which s/he has little or no knowledge.
- A small-scale study is undertaken to decide if it is worth carrying out a detailed investigation. On the basis of the assessment made during the exploratory study, a full study may eventuate.
- Exploratory studies are also conducted to develop, refine and/or test measurement tools and procedures.

Type of Research: Mode of Enquiry Perspective

- Broadly, there are two approaches to enquiry:
    1. the *structured* approach (**quantitative research)**;
    2. the *unstructured* approach (**qualitative research)**.
- In the structured approach:
  - everything that forms the research process – objectives, design, sample, and the questions that you plan to ask of respondents – is predetermined.
  - more appropriate to determine the *extent* of a problem, issue or phenomenon.
  - Is used to find out how many people have a particular perspective, how many people have a particular problem, or how many people hold a particular view.

Type of Research: Mode of Enquiry Perspective

- In the unstructured approach:
  - It allows flexibility in all these aspects of the process.
  - predominantly used to explore its *nature*, in other words, variation /diversity in a phenomenon, issue, problem or attitude towards an issue.
  - For example, if you want to research the different perspectives of an issue, the problems experienced by people living in a community or the different views people hold towards an issue, then these are better explored using unstructured enquiries.

Research Process

Steps in a research project

Project Development	Implementation	Communicating Findings
• Problem definition • Literature review • Objective & hypothesis • Study design	• Sampling • Data collection • Data management • Formal analysis	• Reporting • Publication

Activities

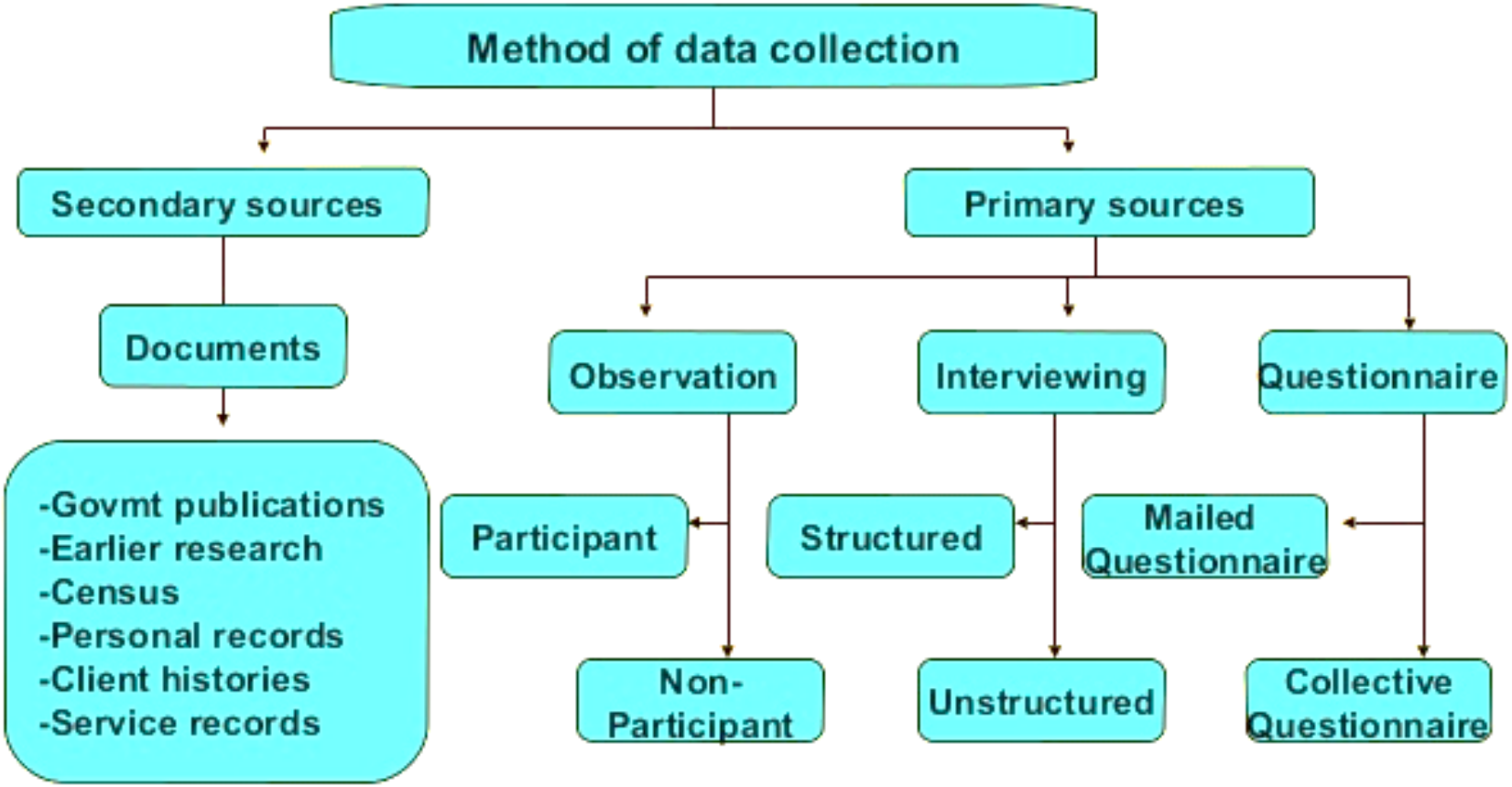
➢ Definition of problem domain & how the specific problem fits in. ➢ Identification of gaps, appropriate methods & theory. ➢ Research will approve or disapprove the hypothesis. ➢ Research strategy to be used, sample size, sampling frame.	➢ Sample selection ➢ Data collection tools ➢ Database development and data cleaning ➢ Exploration, description, modelling & interpretation of statistical outputs	➢ Choice of reporting media & format ➢ Advise on presentation of results ➢ Data sharing media

Research Instruments

- Types of **instruments** used will define the data analysis methods.
- You must begin to **think ahead** about how the information could be organized, analyzed, interpreted and then draw/infer conclusion to be reported to various audiences.

Social Science	Engineering
1. Interview 2. Observation 3. Questionnaire / Survey 4. Document review	1. Experimental 2. Simulation/ Programming/ Algorithm 3. Mathematical modelling

Major Approaches for Information Gathering



Major Approaches for Information Gathering

- None of the methods of data collection provides 100 per cent accurate and reliable information.
- The quality of the data gathered is dependent upon your skill as a researcher (ability to take care of the factors that could affect the quality of your data).
- The choice of a method depends upon the purpose of the study, the resources available and the skills of the researcher.
- There are times when the method most appropriate to achieve the objectives of a study cannot be used because of constraints such as a lack of resources and/or required skills.
- In such situations you should be aware of the problems that these limitations impose on the quality of the data.

Problem using Secondary Sources

- When using data from secondary sources you need to be careful as there may be certain problems with the availability, format and quality of data. The extent of these problems varies from source to source.
- While using such data some issues you should keep in mind are:

Validity and reliability

- The validity of information may vary markedly from source to source. For example, information obtained from a census is likely to be more valid and reliable than that obtained from most personal diaries.

Personal Bias

- The use of information from personal diaries, newspapers and magazines may have the problem of personal bias as these writers are likely to exhibit less rigorousness and objectivity than one would expect in research reports.

Problem using Secondary Sources

Availability of data

- It is common for beginning researchers to assume that the required data will be available, but you cannot and should not make this assumption. Therefore, it is important to make sure that the required data is available before you proceed further with your study.

Format

- Before deciding to use data from secondary sources it is equally important to ascertain that the data is available in the required format. For example, you might need to analyse age in the categories 23–33, 34–48, and so on, but, in your source, age may be categorised as 21–24, 25–29, and so on.

Validity & Reliability Concept

- Validity = *truth* - Does it measure what it intended to measure?
- Reliability = *consistency* in procedures and in reactions of participants
- When reliability and validity are achieved, data are free from systematic errors

Reliability & Validity

Reliability	Validity
• Refer to consistency of measurement / interpretation • Measure what it is supposed to measure. • When measurements yields the same result every time, we say that the measurement is reliable	• Refer to the accuracy of the measurement. • The measurement reflects the real meaning of the concept. • Must consider whether the data collection methods produce information that actually addresses the research question.

The shooting target metaphor

Reliability and validity are closely related to each other!

Data Analysis

- Data analysis has multiple **features** and **approaches** depending on the **type** and **scope** of research.
- It is about manipulating and presenting results based on the data collected.
- Data collection means **gathering information** to address those questions that you have identified.

Data Analysis Guiding Principles

- **Translating** Research Questions / Objectives / Hypotheses into an 'analysis plan'.

- Used our research questions / objectives / hypotheses to **design the study** – experiment or survey, questions to ask and data to collect.

- We use them again to **plan the analysis** – what differences do I need to show, what are my response variables, what types of model may I need to use etc.

- This is often a good time to **draft tables and graphs** which you think will help answer the questions.

Data Analysis Plan

- Data analysis plan is used to:
  - make sure the questions and data collection instrument will **get the information you want**.
  - **align your desired report** with the results of analysis and interpretation.
  - **Improve reliability** (consistent measures over time).

- There are multiple ways to collect information to answer most questions.
- The selection of a method for collecting information must balance several concerns including resources available, credibility, analysis and reporting resources and the skill of the evaluator.

Tips for Organizing Data (1/2)

- Set-up a protocol on **how to received and record the information**.
- **Label all data** immediately as you collect or received it.
- As data are received, check to be sure that the **data is free from error**. You do not want to discover that after all data are collected that there are errors. If data are being transferred in some way, check to be sure that this is done accurately throughout the process.

Tips for Organizing Data (2/2)

- **Back-up** all computer disks containing data.
- Establish a secure place and way to **store all data**. If destroyed or lost, data cannot be replaced. If data are confidential, they should be stored in a locked place, so that only the staff member working with the data has access.
- Set-up a system to **track all data**. This will be your system to check that data are not lost or overlooked as analysis and summarizing are completed.
- Develop a **format for storing and organizing** your data prior to the analysis. For example, you could use spreadsheet program to enter the raw data.

Data Analysis Concept

- In analyzing data, **stress only those important results/unique findings** that give information that could answer your research problem.
- Have to be **consistent and coherent** in your approach as well as **logical** based on certain academic requirements.
- **Sufficient data** should be used **to justify inferences or generalizations**. The implication suggested by the data should be explained and discussed thoroughly.

Evaluation and validation

- Construct validity and reliability
- Statistical analysis
- **Comparative studies** – benchmark, experimental vs. numerical, simulation vs. experimental,
- Comparison with theoretical equation or from previous work

Data Analysis and Interpretation

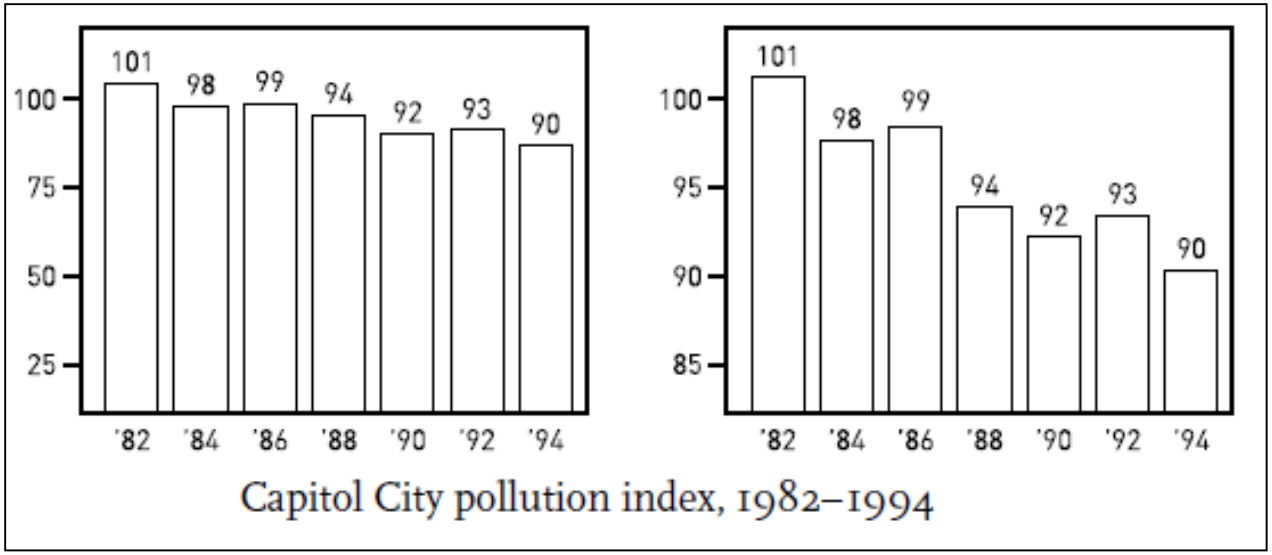
A landscape or the face of a man?

HINT: Nose in the center is also a man sitting and looking at the houses/eyes in the back.

Communicating Data Ethically

- Your graphic must be not only clear and accurate, but **honest**.
- **Do not distort the image** of the data to make your point.

Two line graphs with identical data, yet imply different messages



Capitol City pollution index, 1982–1994

The 0–100 scale in the figure on the left creates a fairly flat slope, which makes the drop in pollution seem small. The vertical scale in the figure on the right, however, begins not at 0 but at 80. When a scale is so truncated, it creates a sharper slope that exaggerates small contrasts.

Guidelines to Avoid Visual Misrepresentation

- **Do not manipulate** a scale to magnify or reduce a contrast.
- Do not use a figure whose image **distorts values**.
- Do not make table or figure **unnecessarily complex** or misleading simple.
- If the table or figure **supports a point, state it**.

Categories of Data Analysis

- ✓ **Narrative (e.g. laws, arts)**
- ✓ **Descriptive (e.g. social sciences)**
- ✓ **Statistical/mathematical (pure/applied sciences/engineering)**
- ✓ **Others**

- Most research analyses, arguably, adopt all the three categories.
- The second and third are, arguably, most popular in pure, applied, and social sciences.

Data Analysis

- **Analyzing data is the process of:**

1. inspecting
2. cleaning
3. transforming
4. modelling
5. coding

With the goal of:

1. highlighting useful information
2. suggesting conclusions
3. support decision-making

Data Presentations Guideline

- Data presentation should be **clear and scholarly done**.
- Analysis of the data refers to **YOUR SKILL** in **describing**, **delineating similarities and differences**, **highlighting the significant findings/data**, and ability to **extract information** or messages out of the presented data.
- Interpretation is the explanation or suggestions inferred from the data, their implications but not conclusions.

Data Presentations

- There are two forms to representing your data/findings in research:
  - **Verbal** : **describe and narrates** to reader what the researcher has done and the results that he/she has obtained.
  - **Symbolic** : **uses graphic representation, tables or statistical values**

Presenting Your Findings Verbally

- When the data are few and simple, readers can grasp them as easily in a sentences as in a table.

In 1996, on average, men earned $32,144 a year, women $23,710, a difference of $8,434.

TABLE 15.1. Male-female salaries ($), 1996

Men	32,144
Women	23,710
Difference	8,434

- If you present more than a few numbers, readers will struggle to keep them straight.

> Between 1970 and 2000, the structure of families changed in two ways. In 1970, 85 percent of families had two parents, but in 1980 that number declined to 77 percent, then to 73 percent in 1990, and to 68 percent in 2000. The number of one-parent families rose, particularly families headed by a mother. In 1970, 11 percent of families were headed by a single mother. In 1980 that number rose to 18 percent, in 1990 to 22 percent, and to 23 percent in 2000. There were some marginal changes among single fathers (headed 1 percent of the families in 1970, 2 percent in 1980, 3 percent in 1990, and 4 percent in 2000). Families headed by no adult remained stable at 3–4 percent.

Presenting Your Findings Verbally

- When the data are few and simple, readers can grasp them as easily in a sentences as in a table.

In 1996, on average, men earned $32,144 a year, women $23,710, a difference of $8,434.

TABLE 15.1. Male-female salaries ($), 1996

Men	32,144
Women	23,710
Difference	8,434

Use of Tables and Graphs

- Tables and graphs are both ways **to organize and arrange data** so that it is more easily understood by the viewer.

- **Tables and graphs are related** in the sense that the information used in tables is frequently also used for basis of graphs.

Presenting Your Findings using Tables

- When designing table, keep the format **clear and simple**.
- **Line up decimal places, note units clearly, use a large enough typeface** and construct a clean orderly arrangement of rows and columns.
- Tables are **numbered consecutively** and table number with caption should be written at the top just right above the table box.
- **Never cut table in two pages**. You may decrease the size of the font to fit the size of page or you may use landscape so that the table can be placed in a page.

Presenting Your Findings using Tables

- **Avoid enlarging the font size** of your table in order to fit in such page.
- **Keep the format clear and simple**. Line up decimal places, note units clearly, use large enough typeface and construct a clean orderly arrangement of rows and columns.

Specific Guidelines for Tables

- Tables with lots of data can seem dense, so organize them to help readers.
- Order the rows and columns by a principle that lets readers quickly find what you want them to see. Do not automatically choose alphabetical order.
- Round numbers to relevant value.
- Sum totals at the bottom of a column or at the end of row, not at the top or left.

Compare Table A and B

Table A

Unemployment in major industrial nations, 1990–2000			
	1990	2001	Change
Australia	6.7	6.5	(.2)
Canada	7.7	5.9	(1.8)
France	9.1	8.8	(.3)
Germany	5.0	8.1	3.1
Italy	7.0	9.9	2.9
Japan	2.1	4.8	2.7
Sweden	1.8	5.1	3.3
UK	6.9	5.1	(1.8)
USA	5.6	4.2	(1.6)

Changes in unemployment rates of industrial nations, 1990–2000

English-speaking vs. non-English-speaking nations

	1990	2001	Change
Canada	7.7	5.9	(1.8)
UK	6.9	5.1	(1.8)
USA	5.6	4.2	(1.6)
Australia	6.7	6.5	(0.2)
France	9.1	8.8	(.3)
Japan	2.1	4.8	2.7
Italy	7.0	9.9	2.9
Germany	5.0	8.1	3.1
Sweden	1.8	5.1	3.3

Frequency Distribution Table

- In the frequency distribution table, the data is first split up into convenient groups (class interval) and the number of items (frequency) which occur in each group is shown in adjacent columns.
- Hence it is a table showing the frequency with which the values are distributed in different groups or classes with some defined characteristics.

Rules for Construction of Frequency Table

- The class interval should not be too large or too small.
- The number of classes to be formed more than 8 and less than 15.
- The class interval should be **equal and uniform** through out the classification.
- After construction of table, **proper and clear heading** should be given to it.
- The base or source of data should be mentioned with the pattern of analysis in the footnote at the end of table.

Frequency Distribution Table

Age distribution of polio patients

Age	Number of patients
0-4	35
5-9	18
10-14	11
15-19	8
20-24	6

Grouped, relative, and cumulative frequency distributions if serum cholesterol levels in 200 men

Interval	Frequency f	relative f	cumulative f
251-260	5	2.5	100.0
241-250	13	6.5	97.5
231-240	19	9.5	91.0
221-230	18	9.0	81.5
211-220	38	19.0	72.5
201-210	72	36.0	53.5
191-200	14	7.0	17.5
181-190	12	6.0	10.5
171-180	5	2.5	4.5
161-170	4	2.0	2.0

Presenting Your Findings using Charts and Diagrams

- Charts and diagrams are useful methods of presenting simple data.
- They have **powerful impact** on imagination of people.
- Gives information at a glance.
- Diagrams are better retained in memory than statistical table.
- However graphs cannot be substituted for statistical table, because the graphs cannot have mathematical treatment where as tables can be treated mathematically.
- Whenever graphs are compared , the difference in the scale should be noted.
- It should be remembered that a lot of details and accuracy of original data is lost in charts and diagrams, and if we want the real study, we have to go back to the original data.

Advantages of Graphs

- **Attractive and effective** presentation of data.
- **Simple and understandable** presentation of data
- Useful in **comparison**
- Useful for **interpretation**
- **Remembrance** for long period
- Helpful in **predictions**
- Universal utility
- Information as well as entertainment
- Helpful in transmission of information
- No need for training

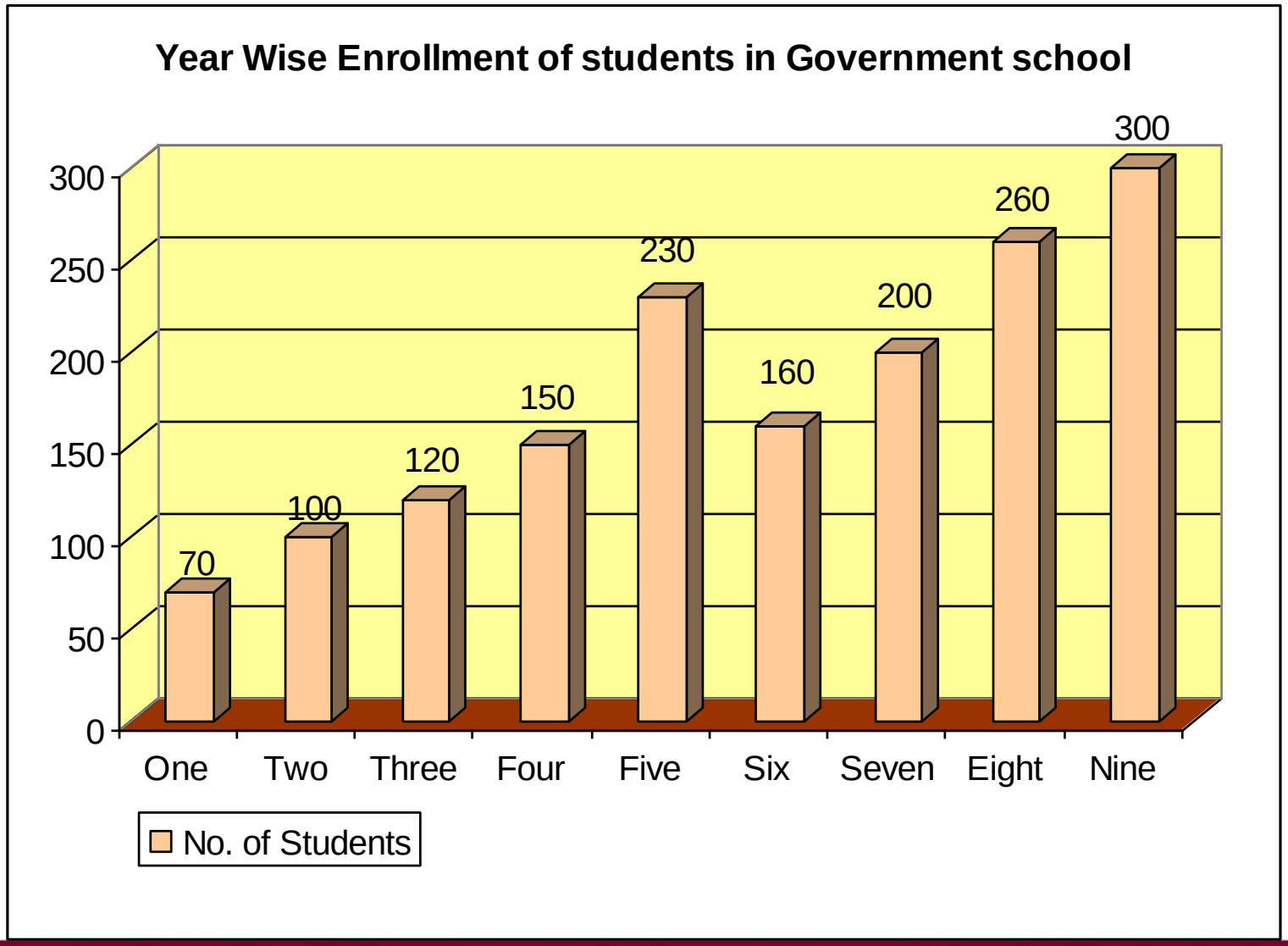
Common Diagrams/Charts

- Pie chart
- Simple bar diagram
- Multiple bar diagram
- Component bar diagram or subdivided bar diagram
- Histogram
- Frequency polygon
- Frequency curve
- O give curve
- Scatter diagram
- Line diagram
- Pictogram
- Statistical maps

Bar Charts

- Bar charts communicate as much by visual impact as by specific numbers.
- Data is presented in the form of rectangular bar of equal breadth.
- The width of the bar and the gaps between the bars should be equal throughout.
- The length of the bar is proportional to the magnitude/ frequency of the variable.
- The bars may be vertical or horizontal.
- Bar chart arranged in no pattern imply no point.
- If possible, group and arranged bars to create an image that matches your message.

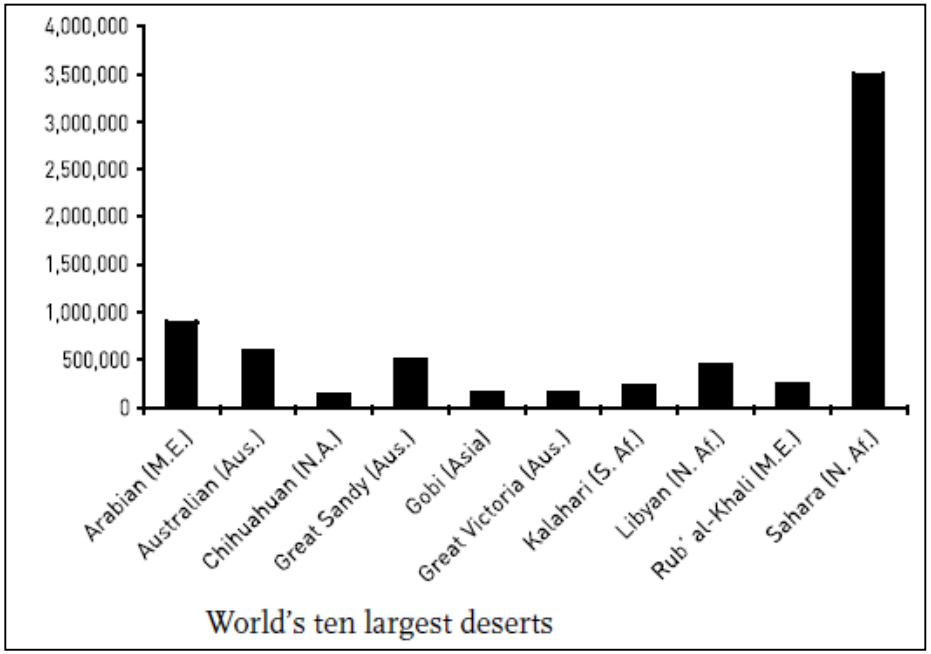
Bar charts


Year Wise Enrollment of students in Government school
300
250
200
150
100
50
0
70
100
120
150
230
160
200
260
300
One
Two
Three
Four
Five
Six
Seven
Eight
Nine
No. of Students

Compare Figure 1 and 2 in the context of the explanatory sentence:

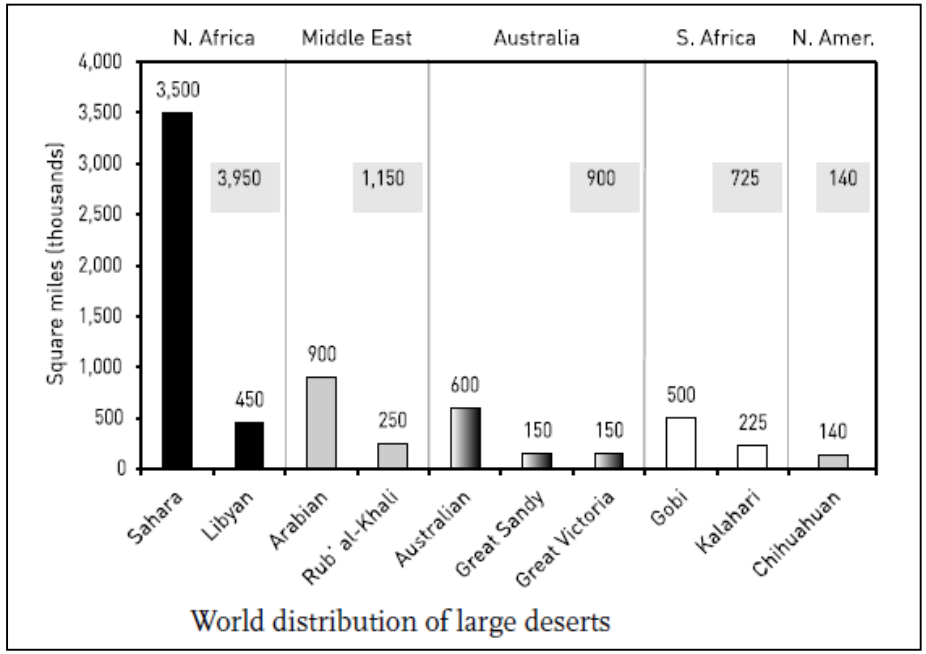
Most of the world's deserts are concentrated in North Africa and the Middle East

Figure 1



World's ten largest deserts

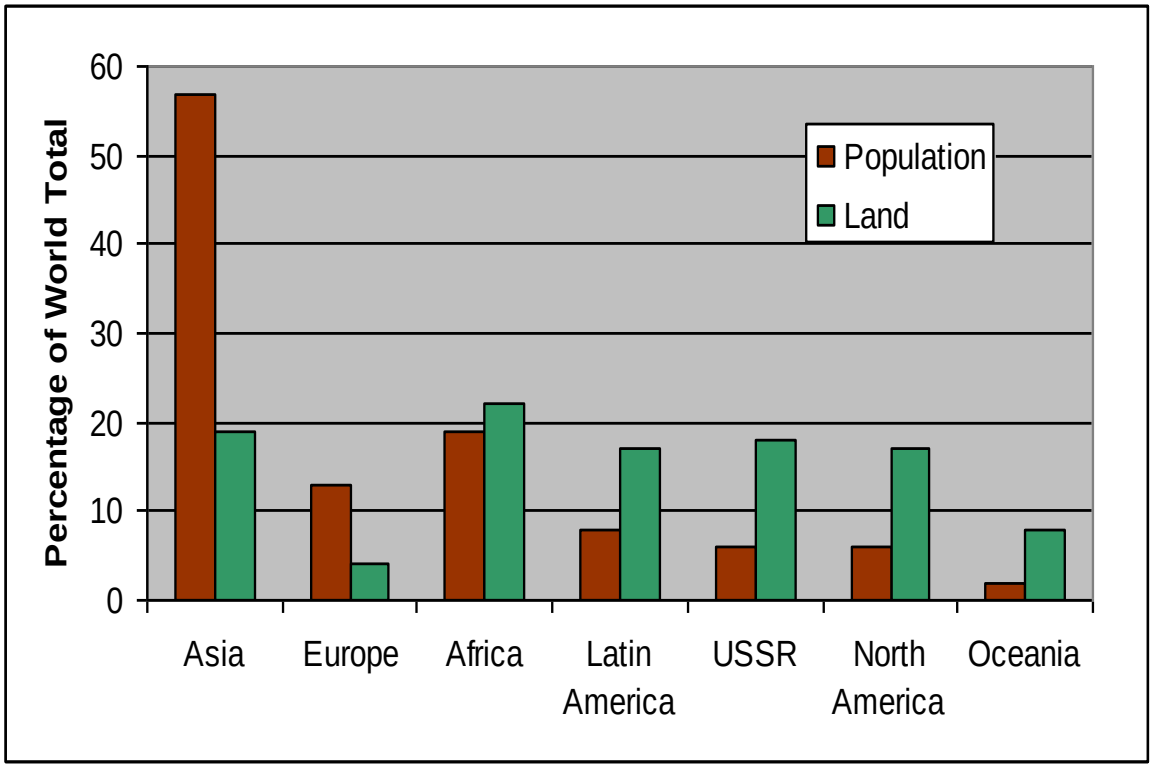
Figure 2



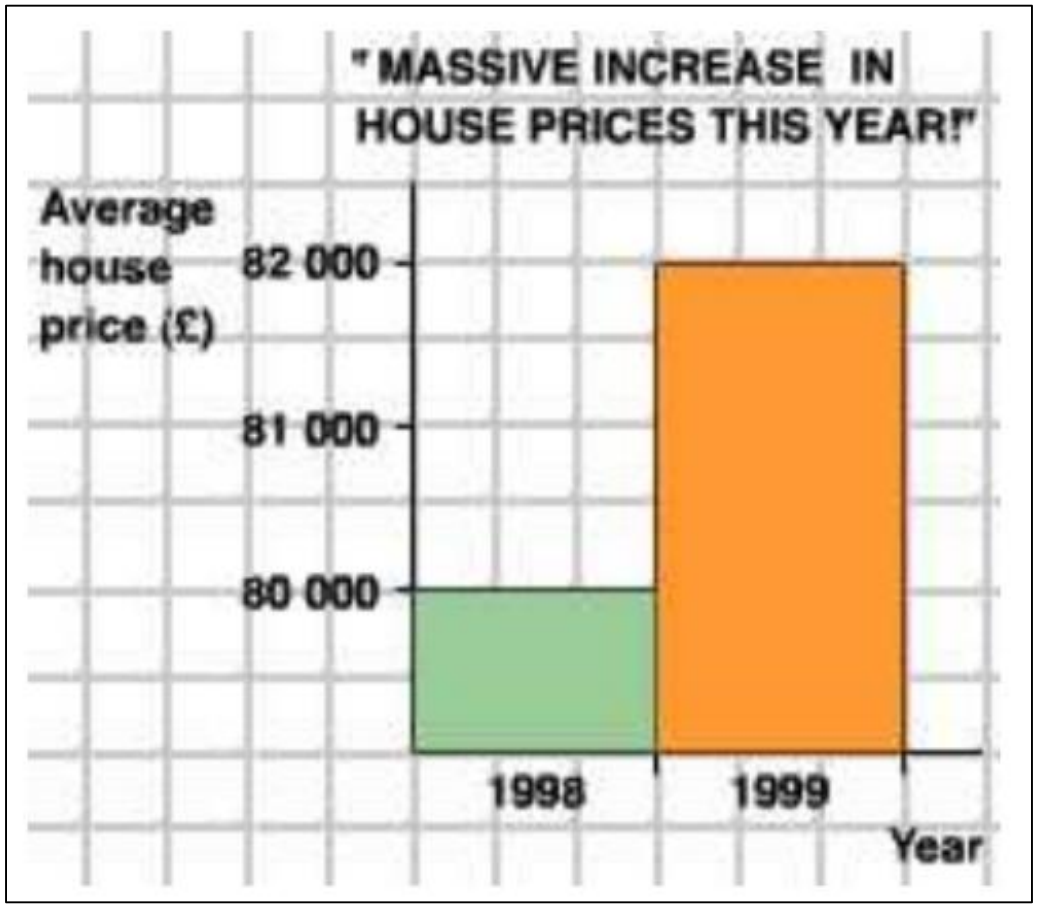
World distribution of large deserts

Multiple Bar Charts

- Also called compound bar charts
- More then one sub-attribute of variable can be expressed

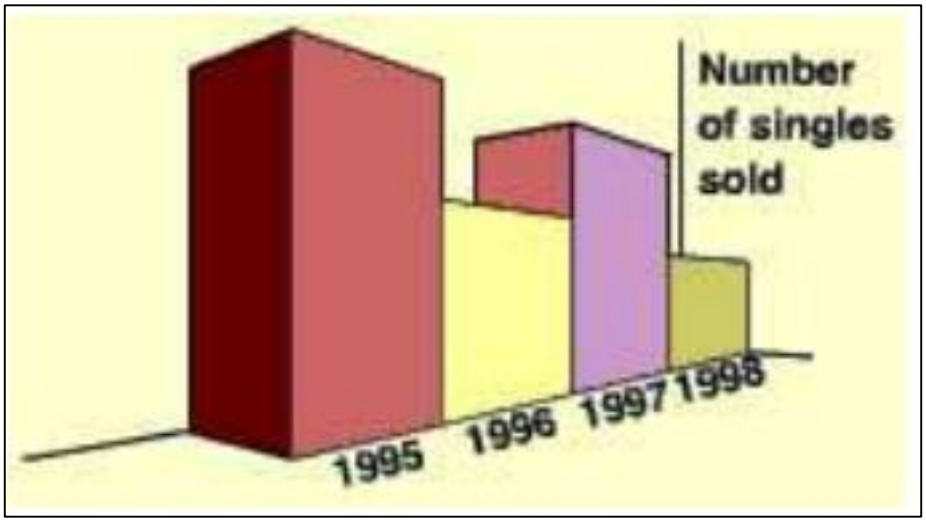


Why does the bar chart below misleading? How should the information be represented?

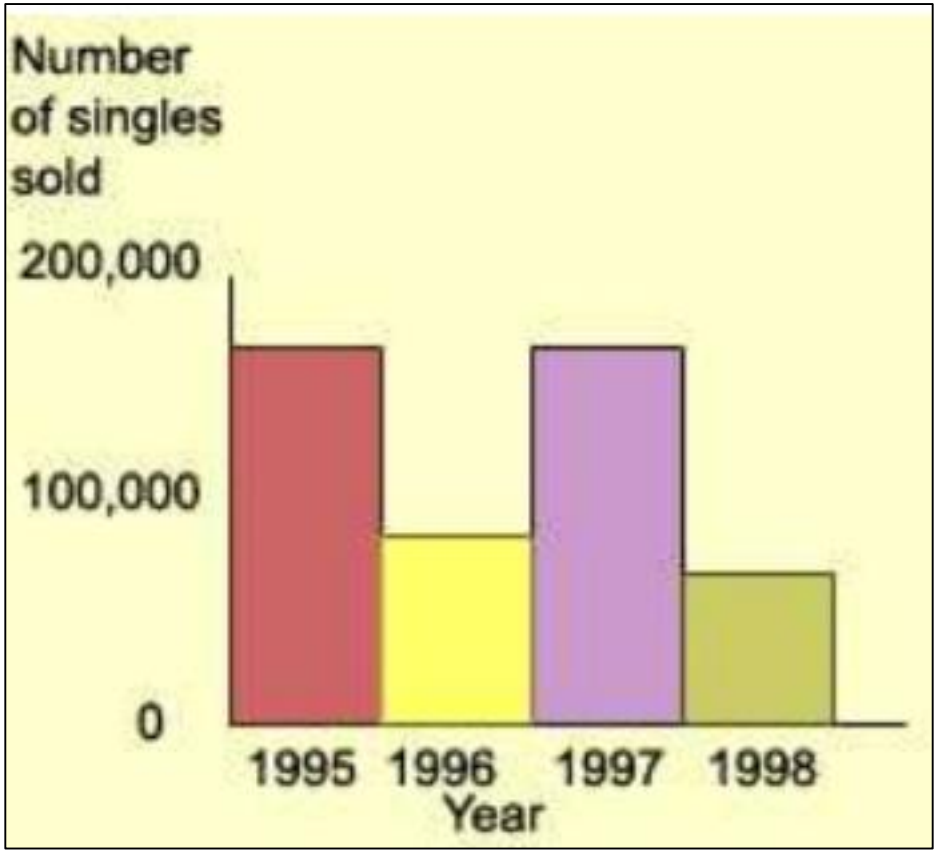


The information can be represented as:

What is wrong with this 3D bar chart?



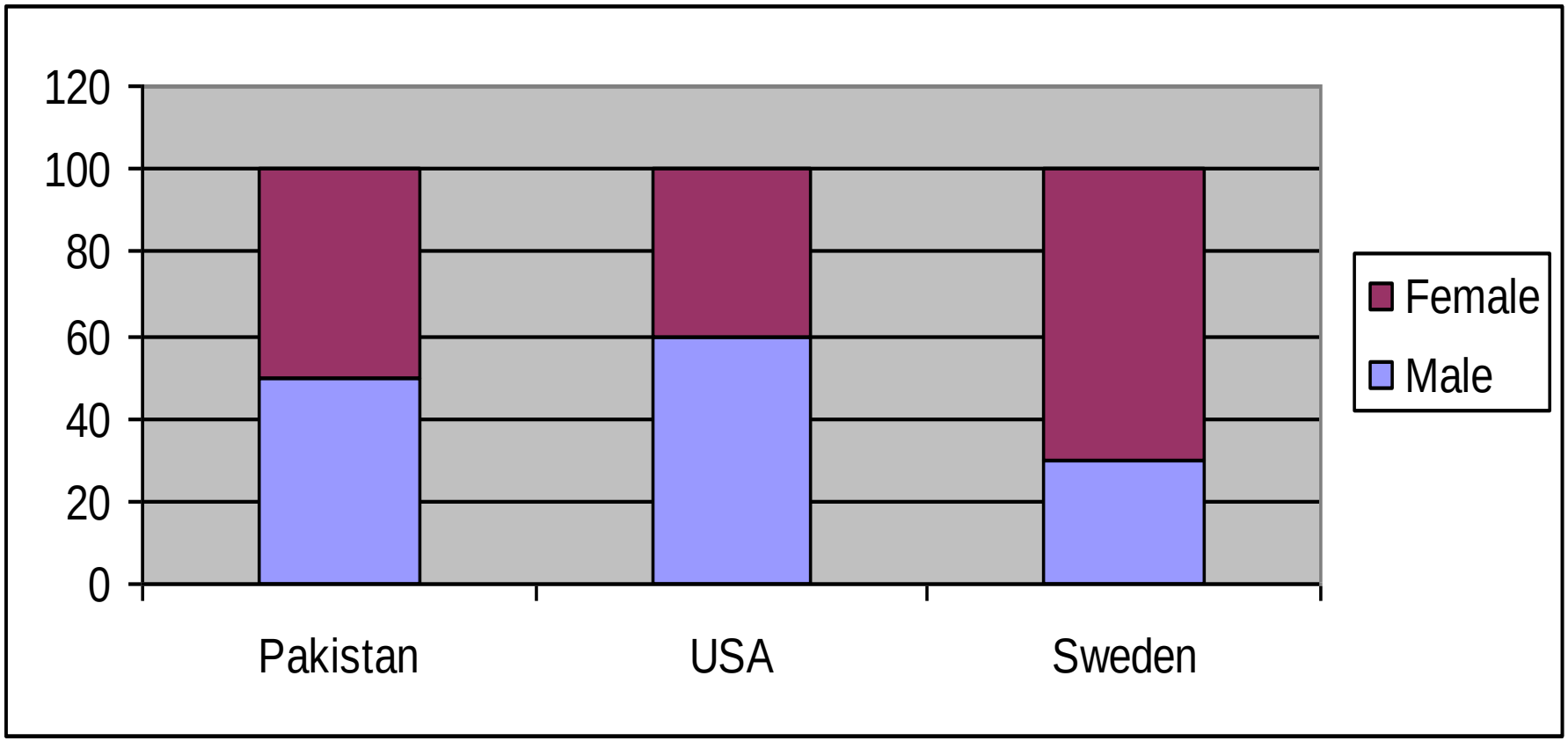
The information can be represented as:



Stacked Bar Charts

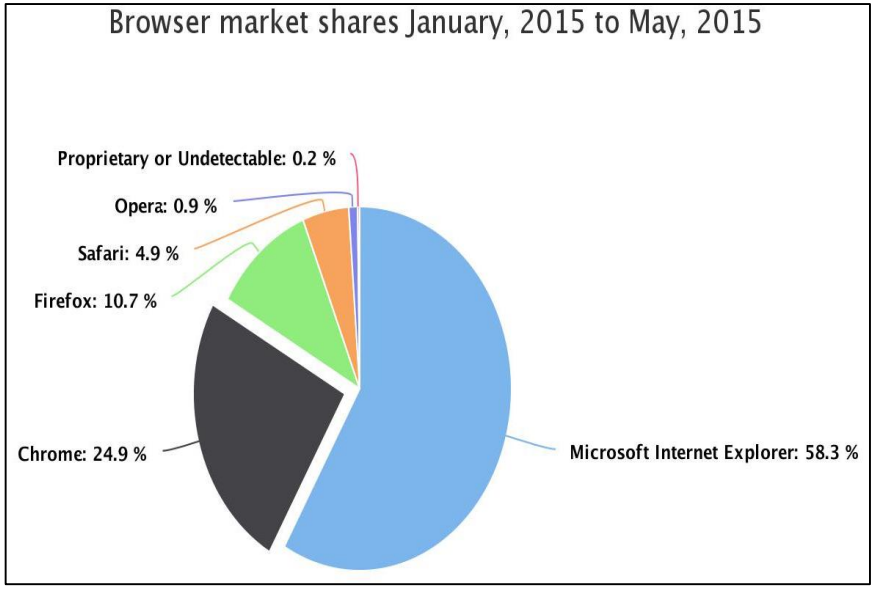
- Use stacked bars only when you want readers to **compare whole values** for different bars rather than their divided segments, because readers cannot easily compare the proportions of segments by eye alone.
- If you do use stacked bars, do this:
  - Arrange segments in a **logical order**. If possible, put the largest segment at the bottom in the darkest shade.
  - Label segments with specific numbers and to assist comparisons, connect corresponding segment with gray lines.

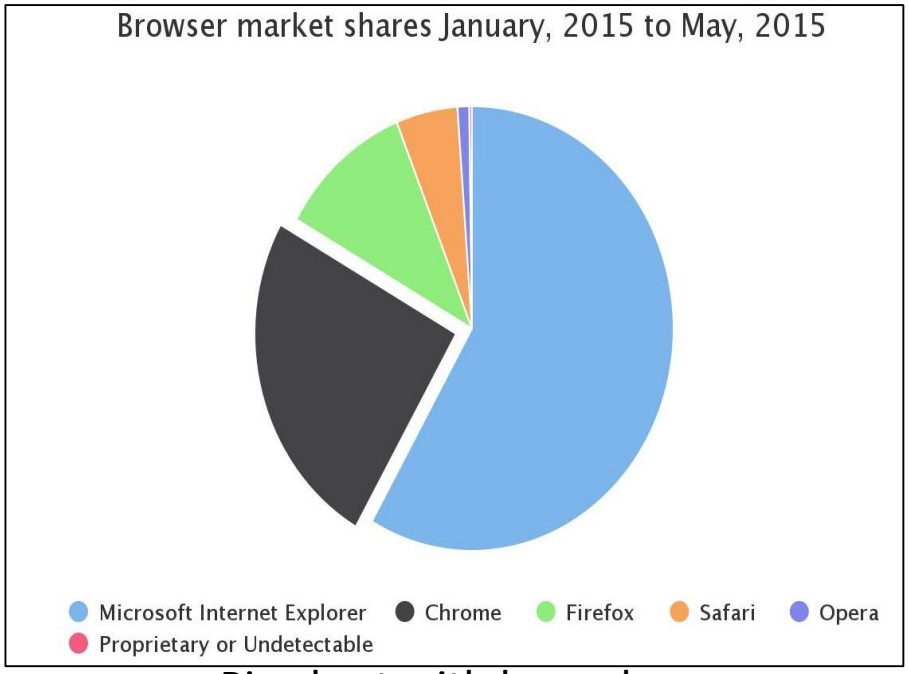
Stacked Bar Chart



Circle Graph / Pie Chart

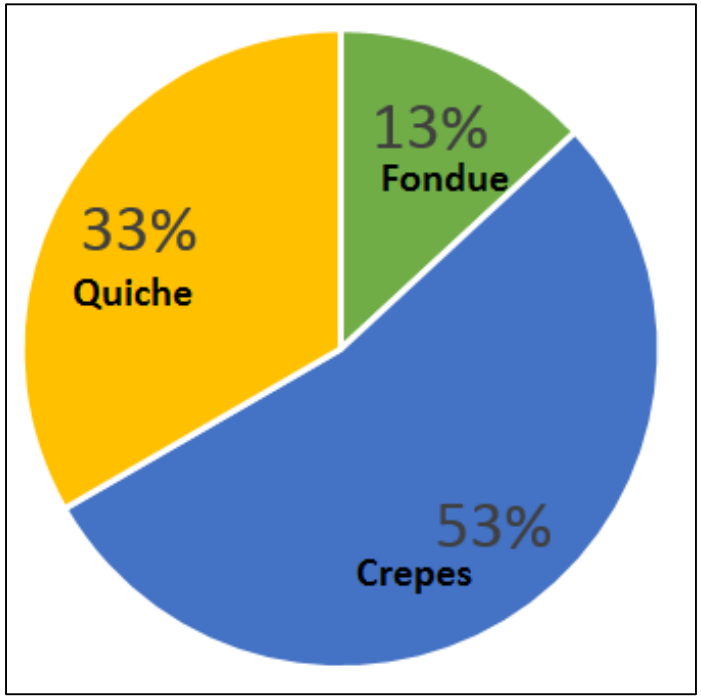
- Definition: A graph that shows data in the form of a circle.
- Advantages:
  - Shows percentages
  - Shows how a total is divided into parts
- Disadvantages:
  - Not best for showing trends





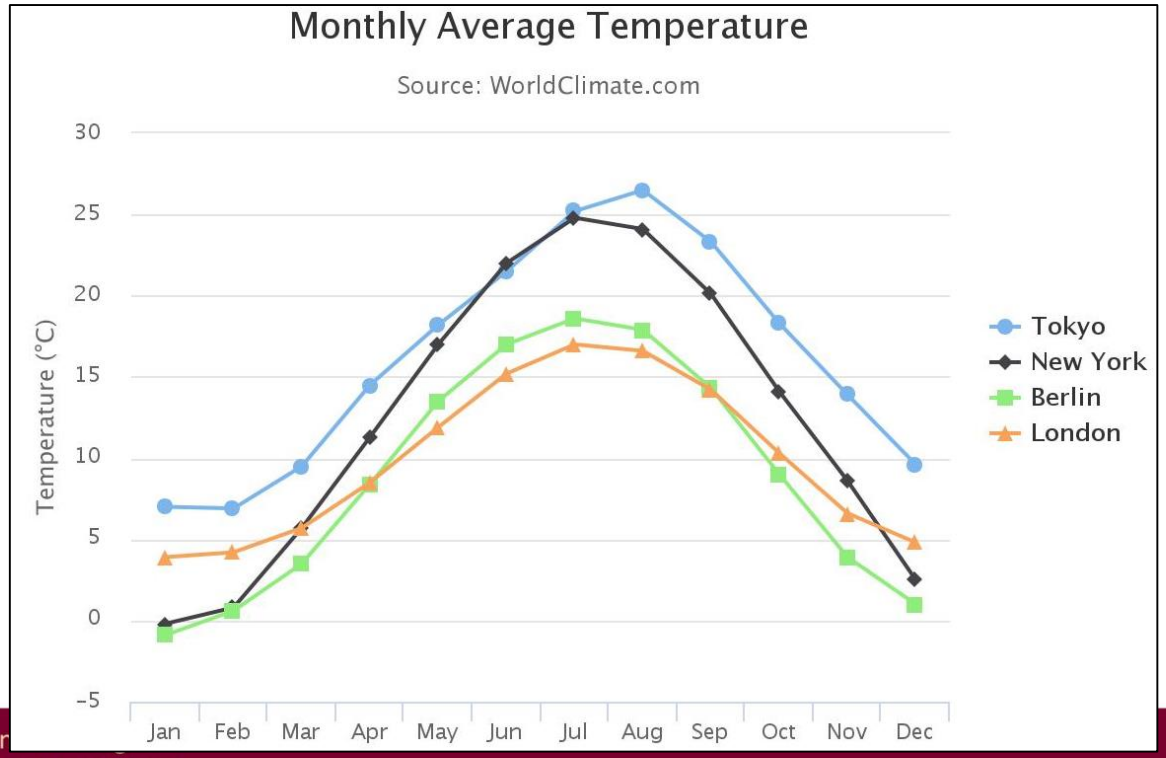
Pie chart with legend

A survey was conducted to determine what food would be served at the French club party.
Explain how the graph misrepresents the data.



Line Graph

- Definition: A graph that shows data in the form of line.
- Advantages:
  - useful in displaying **data** or information that **change continuously over time**. Help you see trends
- Disadvantages:
  - Not easy to compare different categories of data



Example of Line Graph

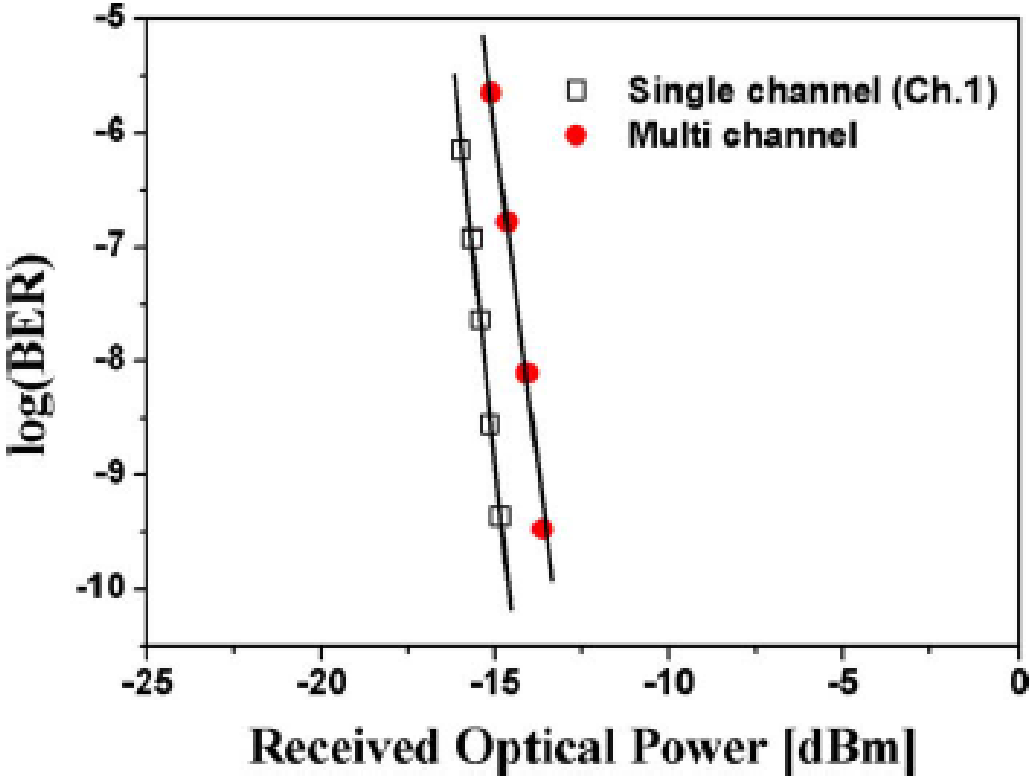


Fig. 5. Measured BER curves for upconversion of single channel and multiple channels.

Example of Line Graph

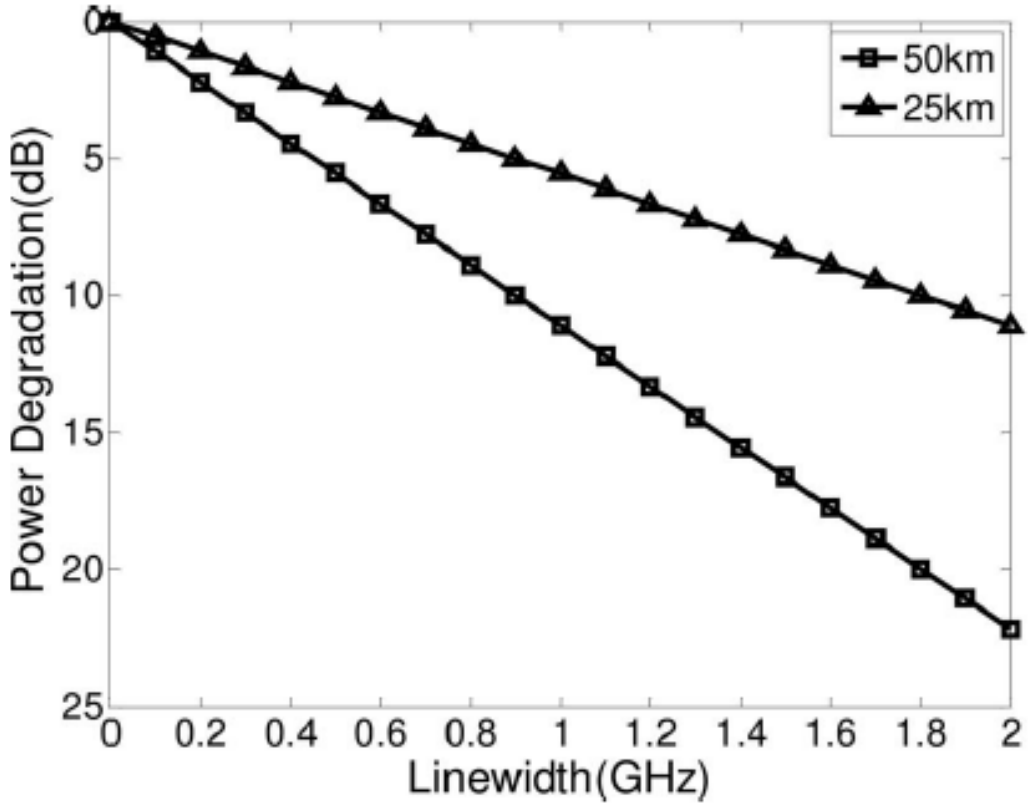
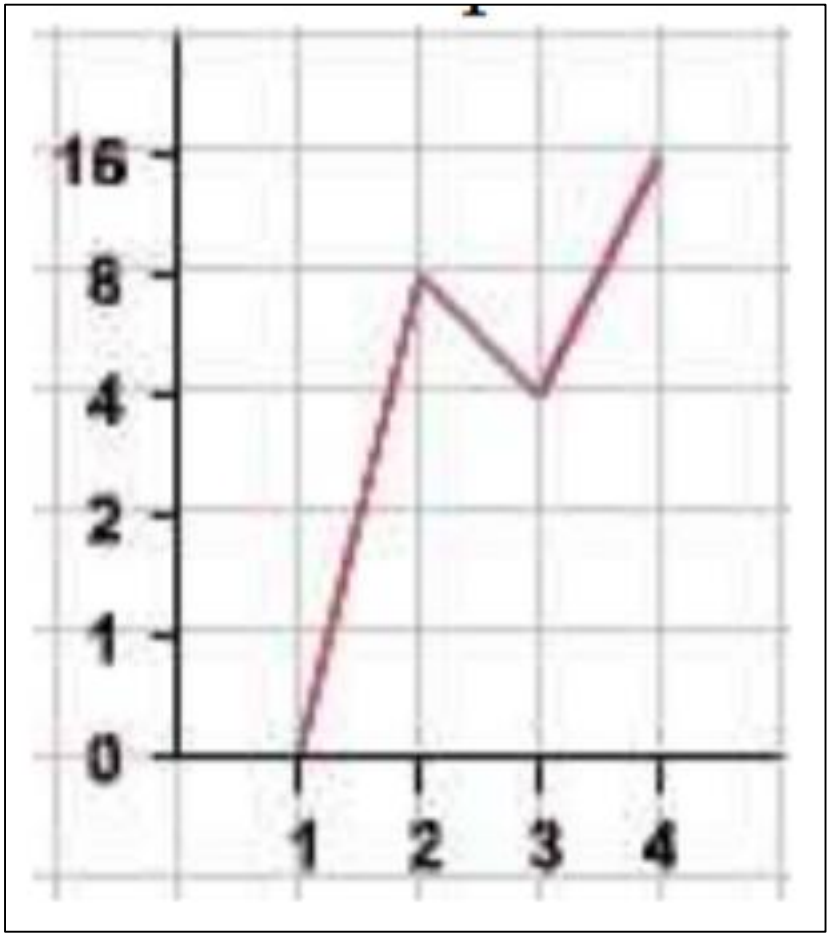


Fig. 3. Millimeter-wave power degradation versus the linewidth of the tunable laser.

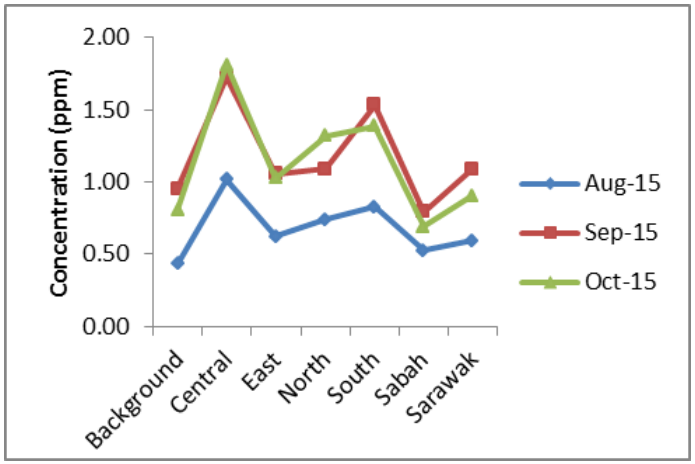
What is wrong with the information represented on this graph?

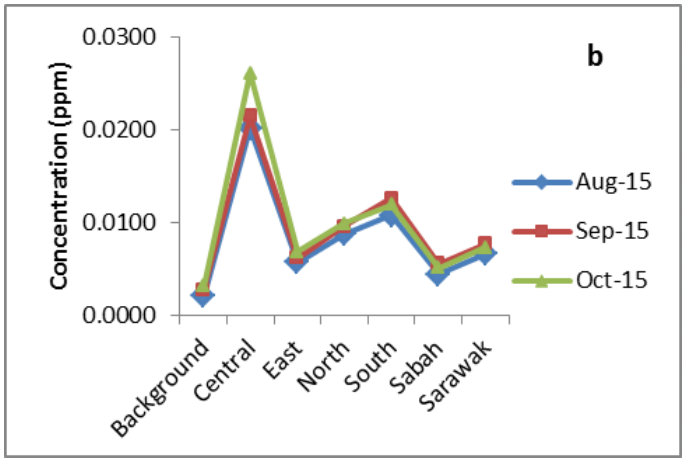


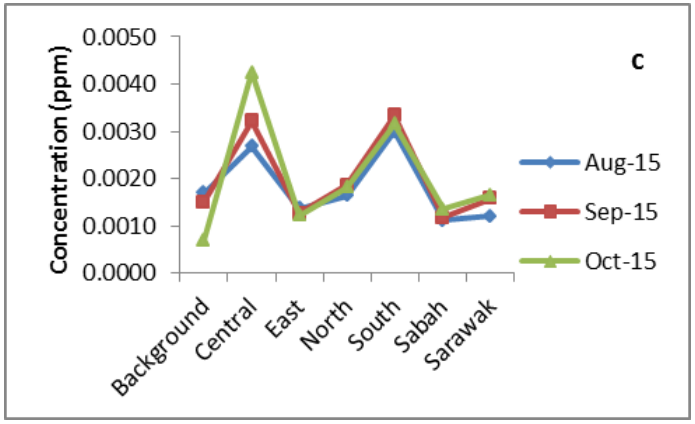
Sample of Line Graph Error

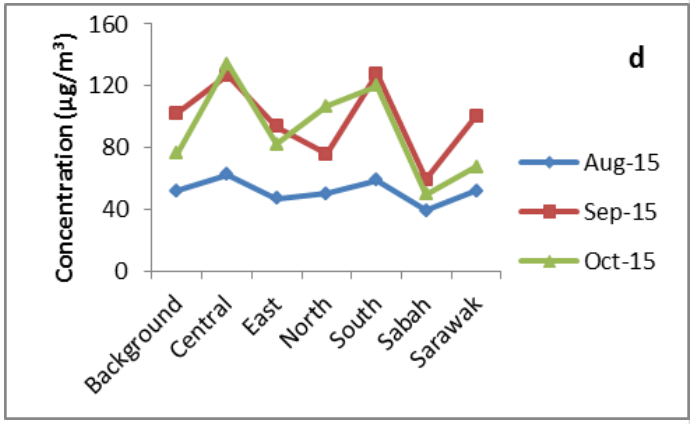
Average concentration of air pollutants during haze episode 2015 for Aug, Sept & Oct 2015 according to various regions

a) carbon monoxide
b) nitrogen dioxide
c) sulfur dioxide
d) PM10









The information can be represented as:

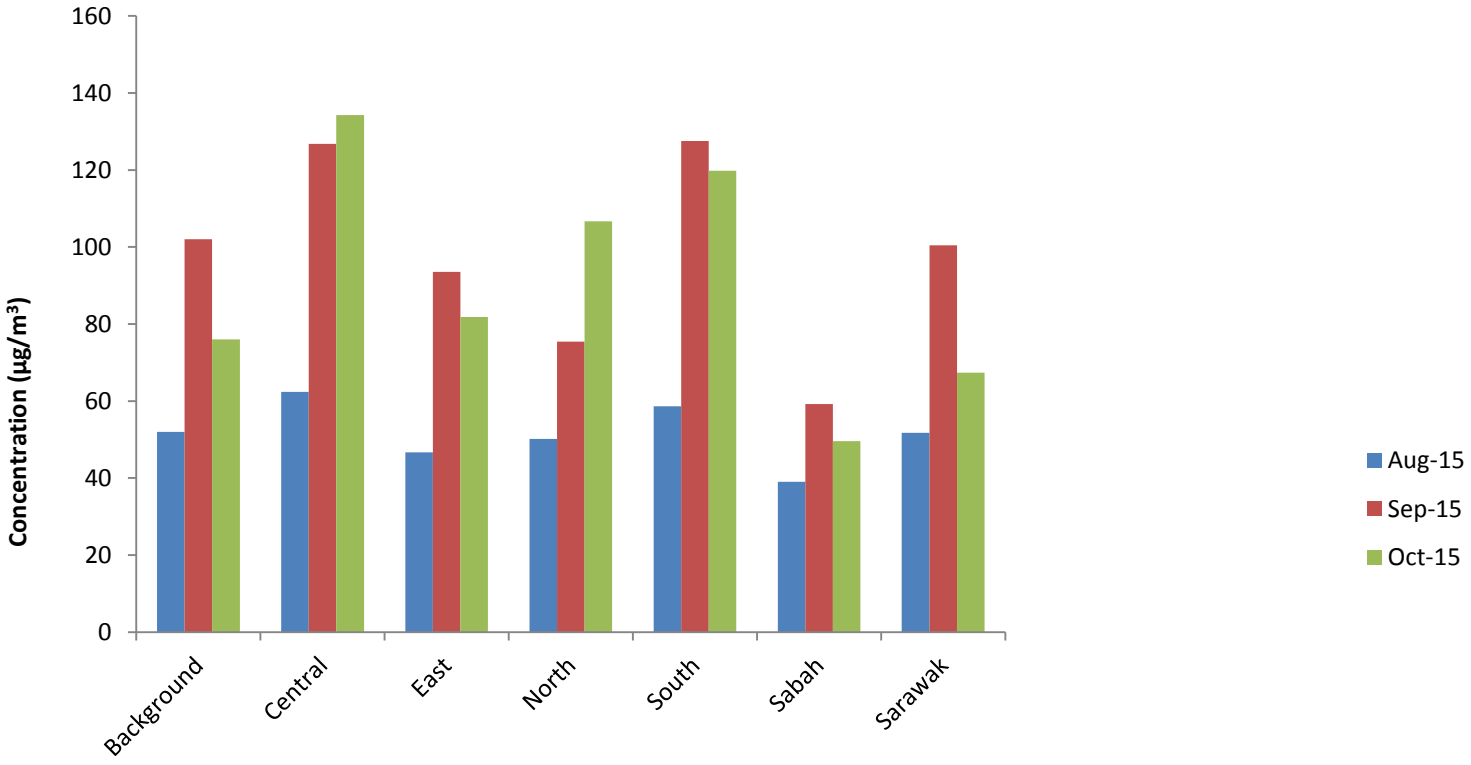
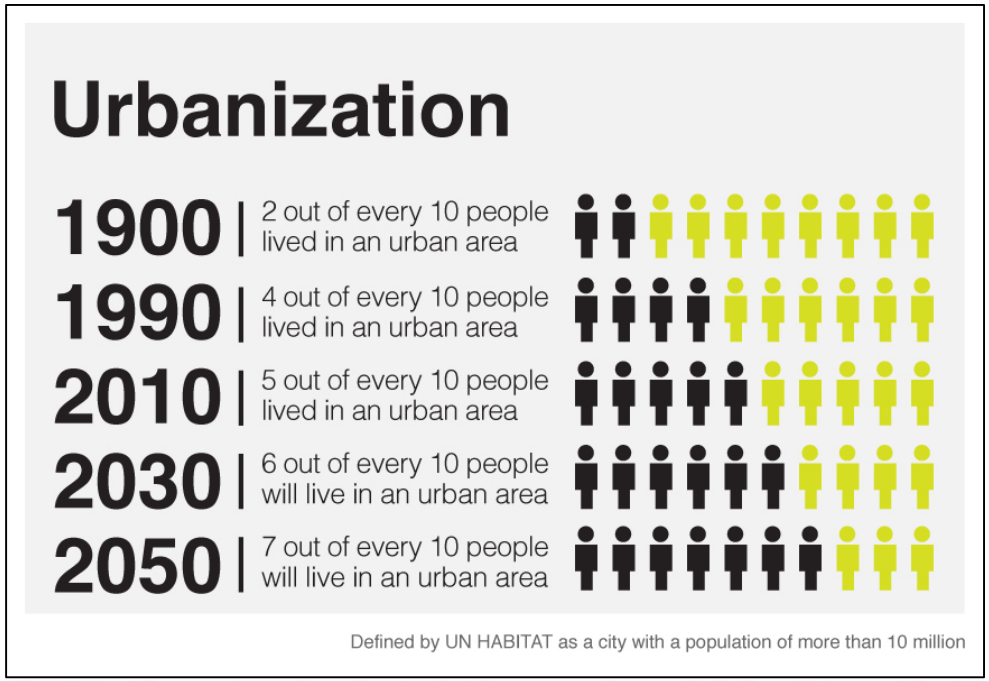


Figure 1. Malaysia: Particulate Matter Concentration by Region or State, August-October 2015

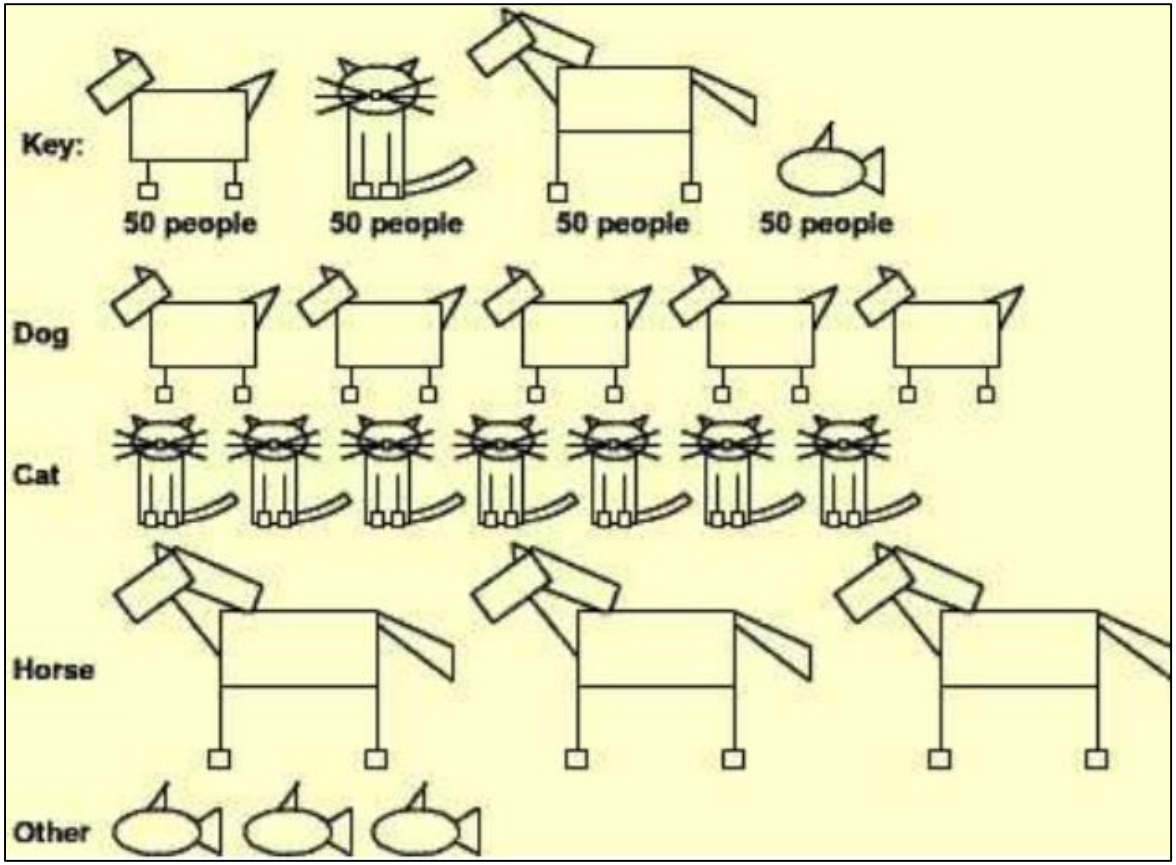
Data Source: DOE; processed by Murnira Othman of UKM

Pictograph

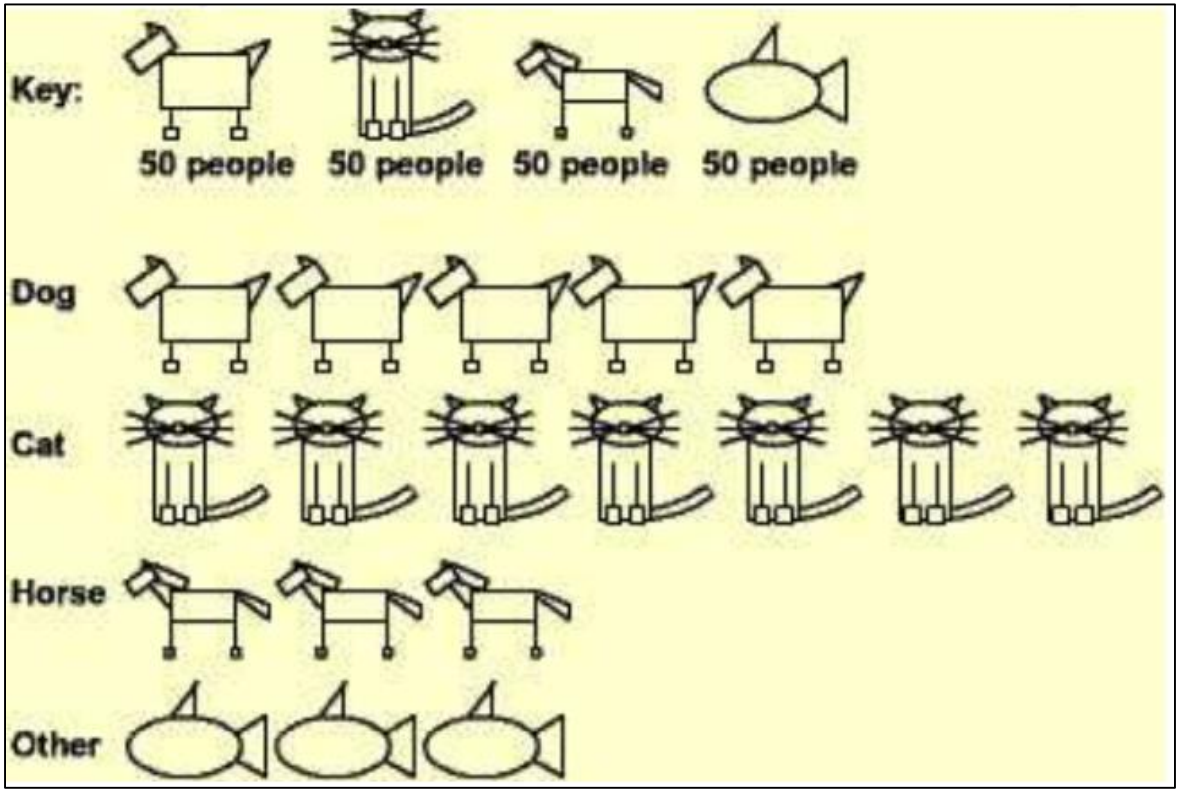
- Definition: A graph that displays data in the form pictures or symbols.
- Advantages:
  - Compare multiple sets of data
  - Visually appealing
- Disadvantages:
  - Hard to read when there are parts of the picture



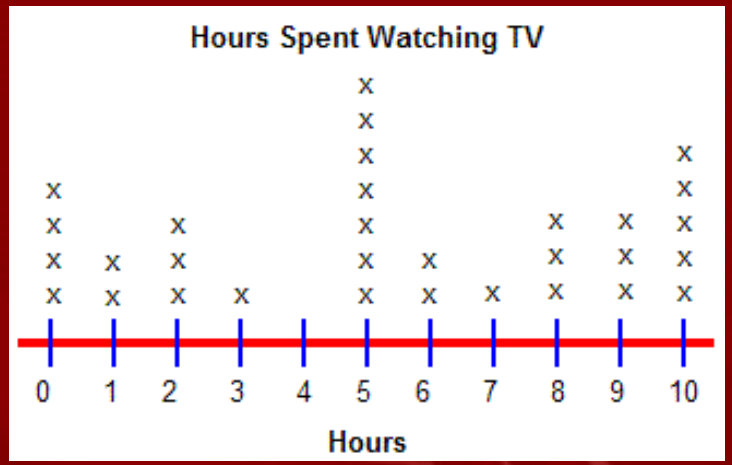
What is wrong with this pictogram showing the number of people who own different types of pets?



The information can be represented as:



Line Plot Graph



Definition: A graph that uses symbols above a number line to display data.

Advantages:

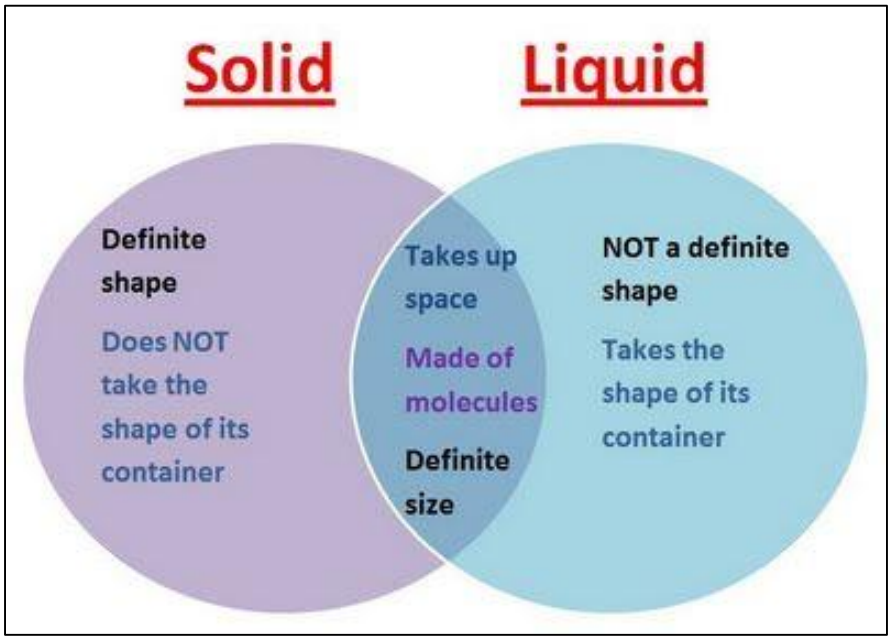
- An easy way to organize data

Disadvantages

- Only good for one group of data and cannot be used to compare different categories of data or show trends.

Venn Diagram

- Definition: Circles that show relationships among sets.
- Advantages:
  - Show comparisons and contrast easily.
- Disadvantages:
  - Does not show trends.



A scatter graph is a type of mathematical diagram using Cartesian coordinates to display values for typically **two variables for a set of data**.

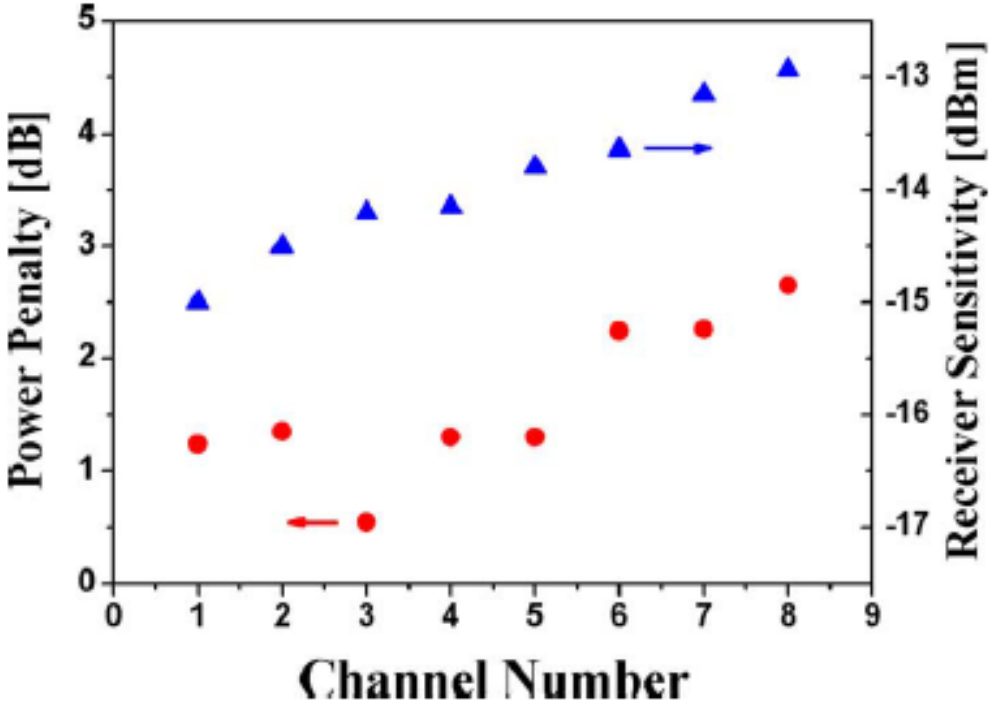
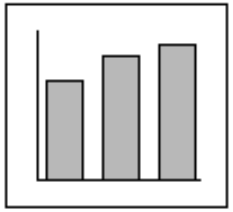
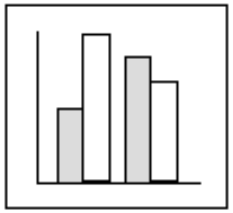
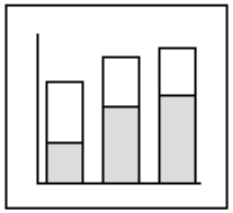
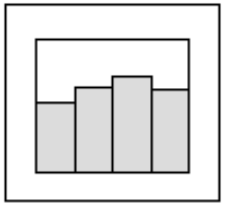


Fig. 6. Receiver sensitivity for single-channel upconverion and power penalty for simultaneous upconversion as a function of the channel number of the optical RF signal.

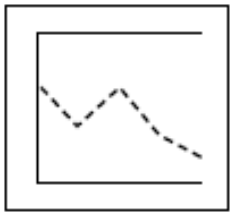
Summary of Visually Data Presentation (1/4)

	Data	Rhetorical Uses
Bar Chart		
	Compares the value of one variable across a series of items called cases (e.g., average salaries for service workers$_{variable}$ in six companies$_{cases}$).	Creates strong visual contrasts among individual cases, emphasizing individual comparisons. For specific values, add numbers to bars. Can show ranks or trends. Vertical bars (called *columns*) are most common, but can be horizontal if cases are numerous or have complex labels.
Bar Chart, Grouped or Split		
	Compares the value of one variable, divided into subsets, across a series of cases (e.g. average salaries$_{variable}$ for men and women service workers$_{subsets}$ in six companies$_{cases}$).	Contrasts subsets within and across individual cases; not useful for comparing total values for cases. For specific values, add numbers to bars. Grouped bars show ranking or trends poorly; useful for time series only if trends are unimportant.
Bar Chart, Stacked		
	Compares the value of one variable, divided into two or more subsets, across a series of cases (e.g. harassment complaints$_{variable}$ segmented by region$_{subsets}$ in six industries$_{cases}$).	Best for comparing totals across cases and subsets *within* cases; difficult to compare subsets across cases (use grouped bars). For specific values, add numbers to bars and segments. Useful for time series. Can show ranks or trends for total values only.

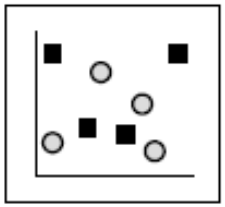
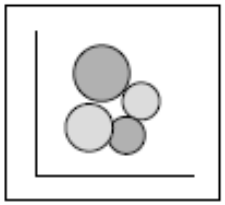
Summary of Visually Data Presentation (2/4)

Histogram		
	Compares two variables, with one segmented into ranges that function like the cases in a bar graph (e.g., service workers$_{\textit{continuous variable}}$ whose salary is \$0–5,000, \$5,000–10,000, \$10,000–15,000, etc.$_{\textit{segmented variable}}$).	Best for comparing segments within continuous data sets. Shows trends, but emphasizes segments (e.g., a sudden spike at \$5,000–10,000 representing part-time workers). For specific values, add numbers to bars.
Image Chart		
	Shows value of one or more variable for cases displayed on a map, diagram, or other image (e.g., states$_{\textit{cases}}$ colored red or blue to show voting patterns$_{\textit{variable}}$).	Shows the distribution of the data in relation to preexisting categories; deemphasizes specific values. Best when the image is familiar, as in a map or diagram of a process.
Pie Chart		
	Shows the proportion of a single variable for a series of cases (e.g., the budget share$_{\textit{variable}}$ of U.S. cabinet departments$_{\textit{cases}}$).	Best for comparing one segment to the whole. Useful only with few segments or segments that are very different in size; otherwise comparisons among segments are difficult. For specific values, add numbers to segments. Common in popular venues, frowned on by professionals.

Summary of Visually Data Presentation (3/4)

	Data	Rhetorical Uses
Line Graph		
	Compares continuous variables for one or more cases (e.g., temperature$_{variable}$ and viscosity$_{variable}$ in two fluids$_{cases}$).	Best for showing trends; deemphasizes specific values. Useful for time series. To show specific values, add numbers to data points. To show the significance of a trend, segment the grid (e.g., below or above average performance).
Area Chart		
	Compares two continuous variables for one or more cases (e.g., reading test scores$_{variable}$ over time$_{variable}$ in a school district$_{case}$).	Shows trends; deemphasizes specific values. Can be used for time series. To show specific values, add numbers to data points. Areas below the lines add no information, but will lead some readers to misjudge values. Confusing with multiple lines/areas.
Area Chart, Stacked		
	Compares two continuous variables for two or more cases (e.g., profit$_{variable}$ over time$_{variable}$ for several products$_{cases}$).	Shows the trend for the total of all cases, plus how much each case contributes to that total. Likely to mislead readers on the value or the trend for any individual case.

Summary of Visually Data Presentation (4/4)

Scatterplot		
	Compares two variables at multiple data points for a single case (e.g., housing sales$_{variable}$ and distance from downtown$_{variable}$ in one city$_{case}$) or at one data point for multiple cases (e.g., brand loyalty$_{variable}$ and repair frequency$_{variable}$ for ten manufacturers$_{cases}$).	Best for showing the distribution of data, especially when there is no clear trend or when the focus is on outlying data points. If only a few data points are plotted, it allows a focus on individual values.
Bubble Chart		
	Compares three variables at multiple data points for a single case (e.g., housing sales,$_{variable}$ distance from downtown,$_{variable}$ and prices$_{variable}$ in one city$_{case}$) or at one data point for multiple cases (e.g. image advertising,$_{variable}$ repair frequency,$_{variable}$ and brand loyalty$_{variable}$ for ten manufacturers$_{cases}$).	Emphasizes the relationship between the third variable (bubbles) and the first two; most useful when the question is whether the third variable is a product of the others. Readers easily misjudge relative values shown by bubbles; adding numbers mitigates that problem.

Statistical Analysis to Analyze Data

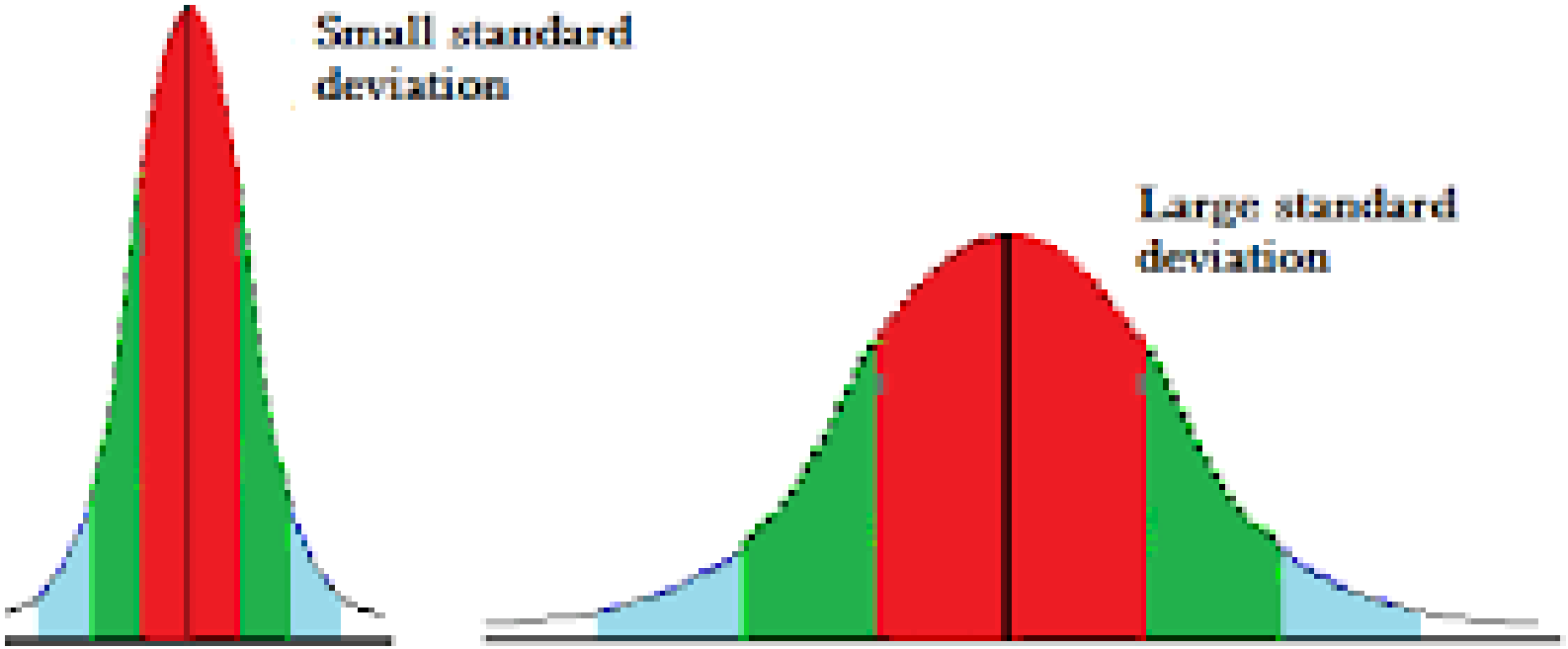
Measure	Central Tendency
Mean	The sum of the numbers in a set of data divided by the number of pieces of data
Mode	The number that occurs most frequently in a set of numbers
Median	The number in the middle of a set of data when the data are arranged in order from least to greatest. When there are 2 middle numbers, the median is the number that is halfway between the two middle numbers

Statistical Analysis to Analyze Data

Measure	Dispersion
Range	The difference between the maximum and minimum value in a numerical data set
Standard Deviation	A measure of dispersion around the mean

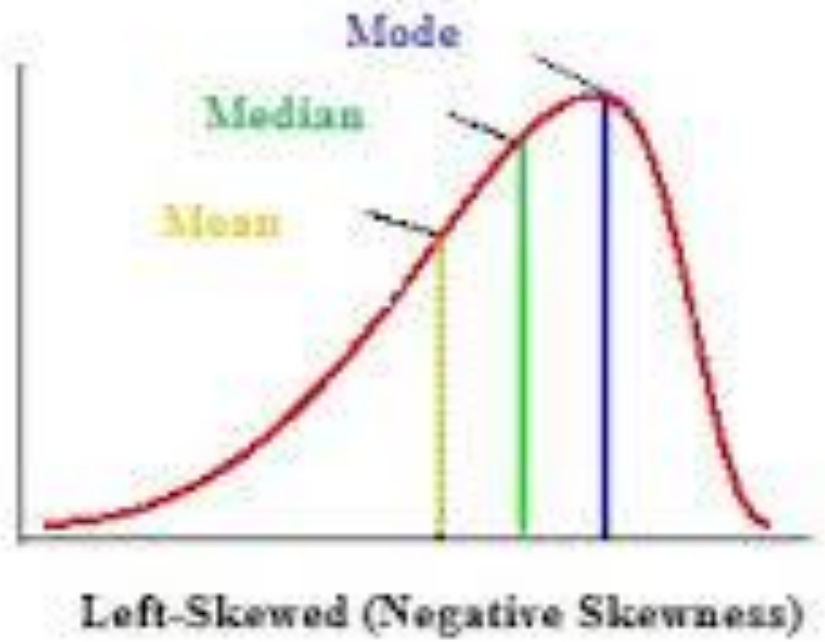
Statistical Analysis to Analyze Data

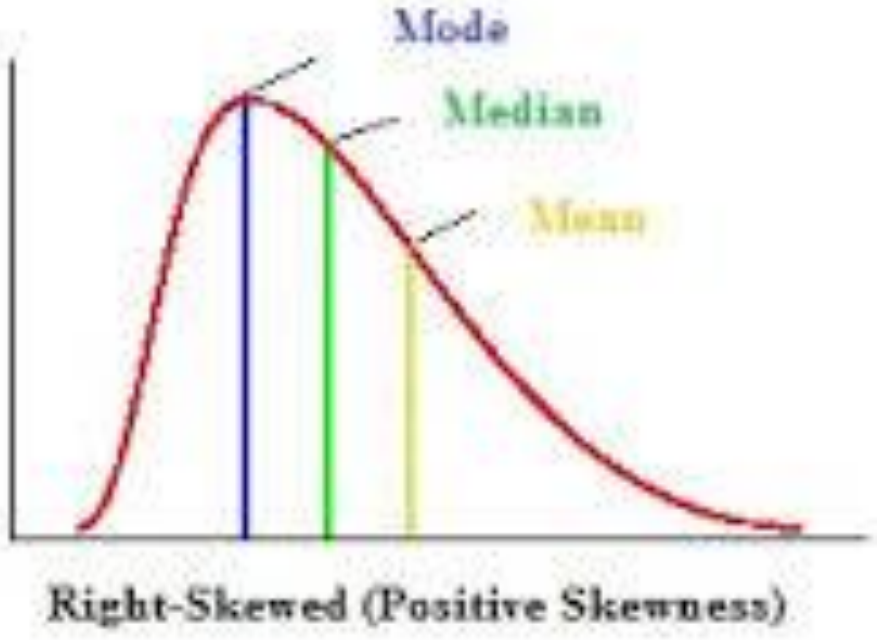
Standard Deviation – is a measure of dispersion around the mean



Statistical Analysis to Analyze Data

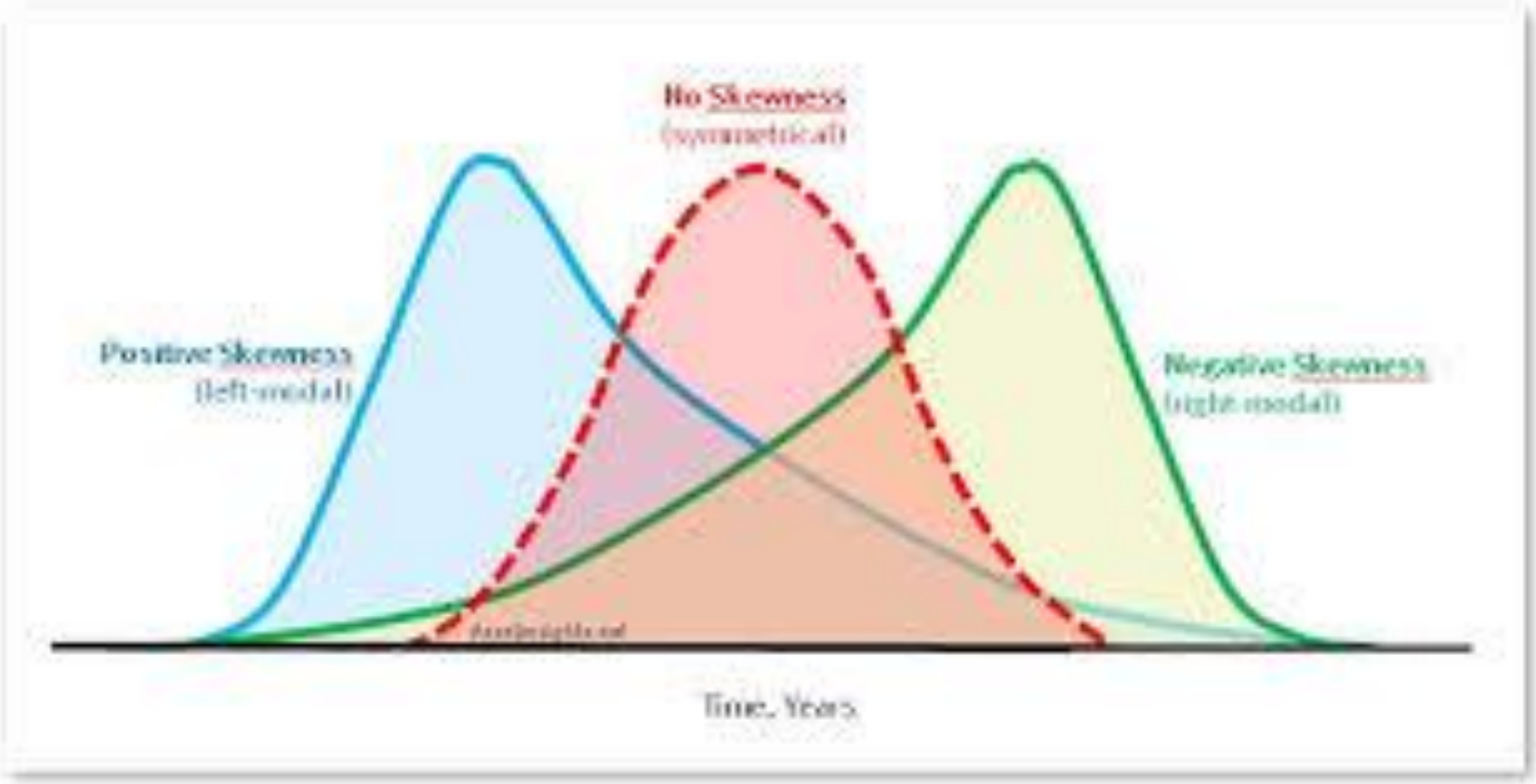
- **Central tendencies and distribution shape**



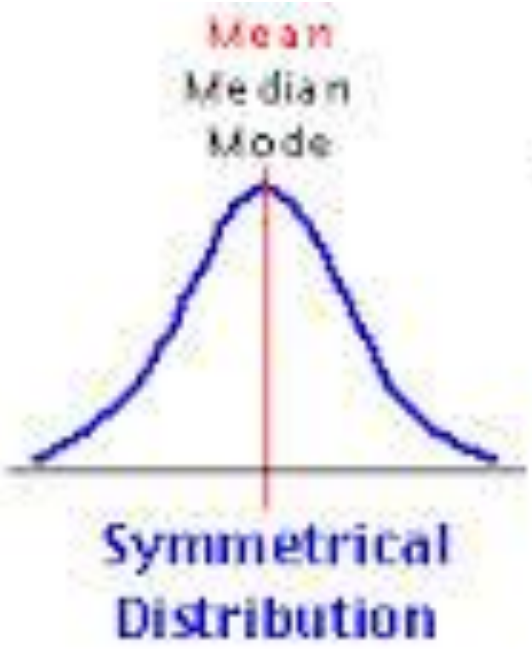


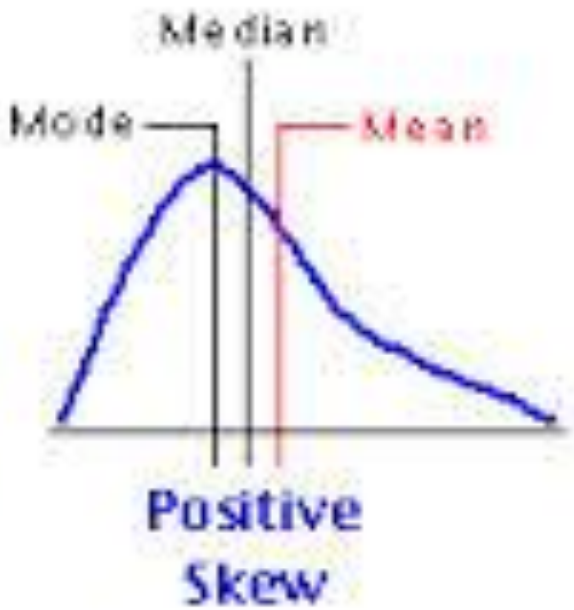
Statistical Analysis to Analyze Data

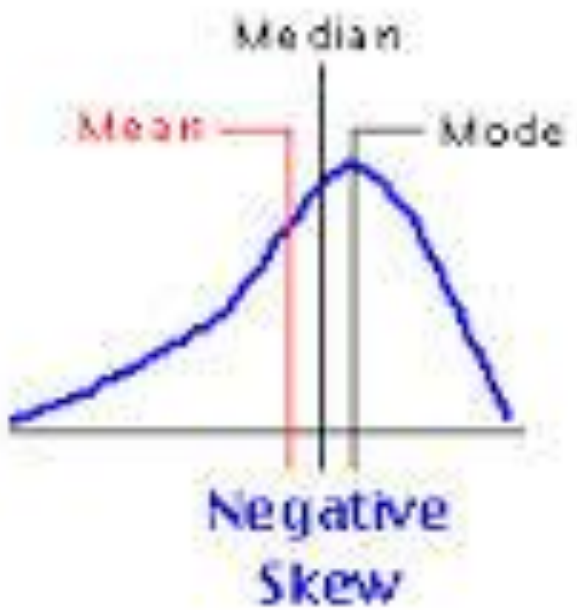
Skewness – is a measure of asymmetric in a distribution.



Statistical Analysis to Analyze Data







Example

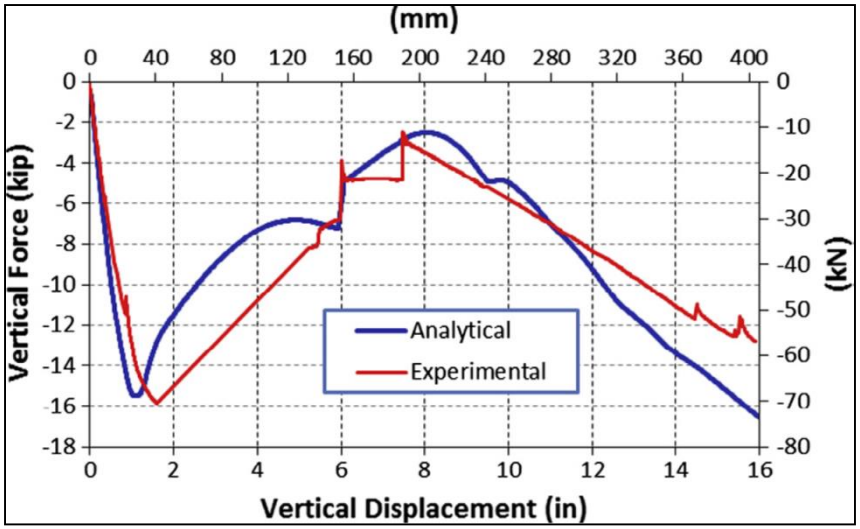
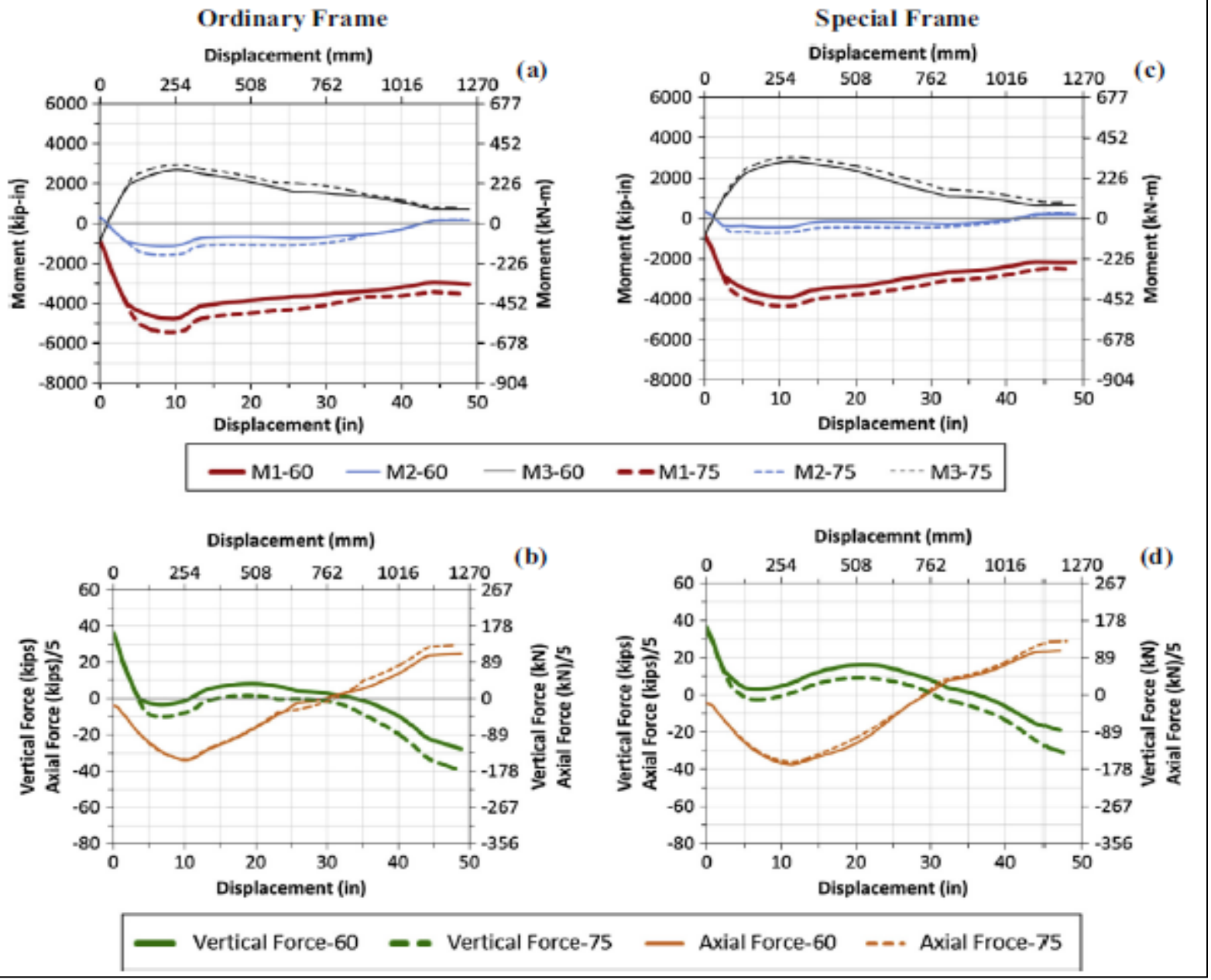


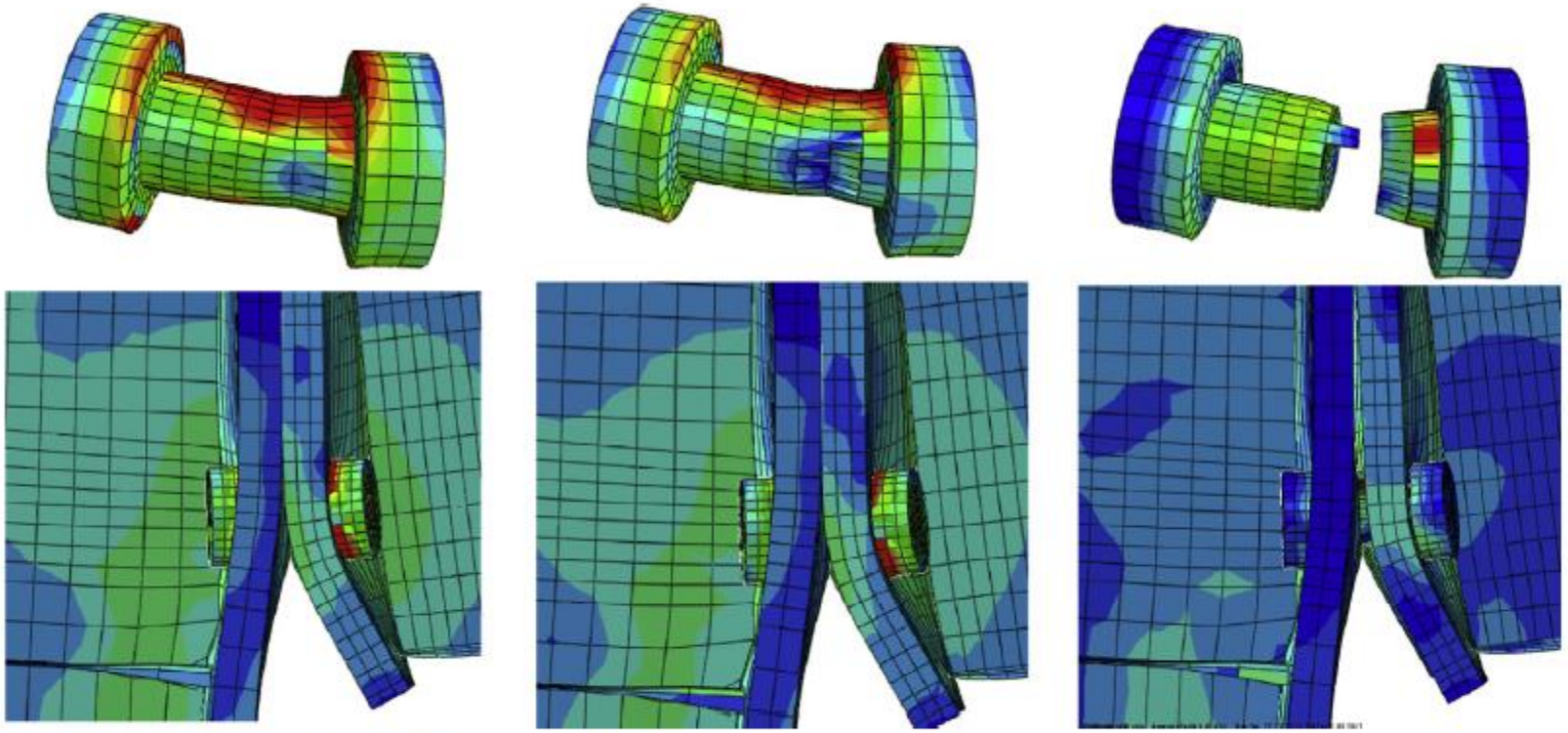
Fig. 2 compares the analytical and experimental force–displacement relationships of the RC beam. A downward force is considered negative. As can be seen, the analytical model captures the three main phases of the force–displacement response observed from experimental results: the first branch up to the peak downward (negative) force is controlled by flexural and compression membrane action, which is followed by a sudden drop in the resisting force due to a combination of the axial compressive force–displacement (P–D) effect and concrete crushing as discussed later in this paper, and finally the third branch in which the vertical load carrying capacity of the beam increases again due to tension membrane (catenary) action.


Ordinary Frame
Special Frame
(a)
(b)
(c)
(d)
Displacement (mm)
Displacement (in)
Moment (kip-in)
Moment (kN-m)
Vertical Force (kips)
Axial Force (kips)/5
Vertical Force (kN)
Axial Force (kN)/5
M1-60
M2-60
M3-60
M1-75
M2-75
M3-75
Vertical Force-60
Vertical Force-75
Axial Force-60
Axial Froce-75

Effects of higher rebar yield strength

Fig. xx compares the internal forces of both special frames and ordinary frames with 60 ksi (414 MPa) and 75 ksi (517 MPa) steel respectively for a 26 ft (7.92 m) beam span. As can be seen by

the moment–displacement curves, **the behavior of the beams is the same up until the point where the bottom rebars at the right end of the beam (section 3 in Fig. 4) yields**. **This happens earlier for the beam with 60 ksi (414 MPa) steel than for the beam with 75 ksi (517 MPa) steel. Following this event, the beam with 75 ksi (517 MPa) steel develops larger moments and reaches equilibrium before the 60 ksi (414 MPa) steel beam. It is noted that the axial force developed in the beam for both 60 ksi (414 MPa) and 75 ksi (517 MPa) is about the same throughout the analysis up until around the time where the steel at the midspan of the beam yields. Overall it can be concluded that both special and ordinary frame beams with 75 ksi (517 MPa) steel behave better in terms of their ability to reach equilibrium than those with 60 ksi(414 MPa) steel beams.**

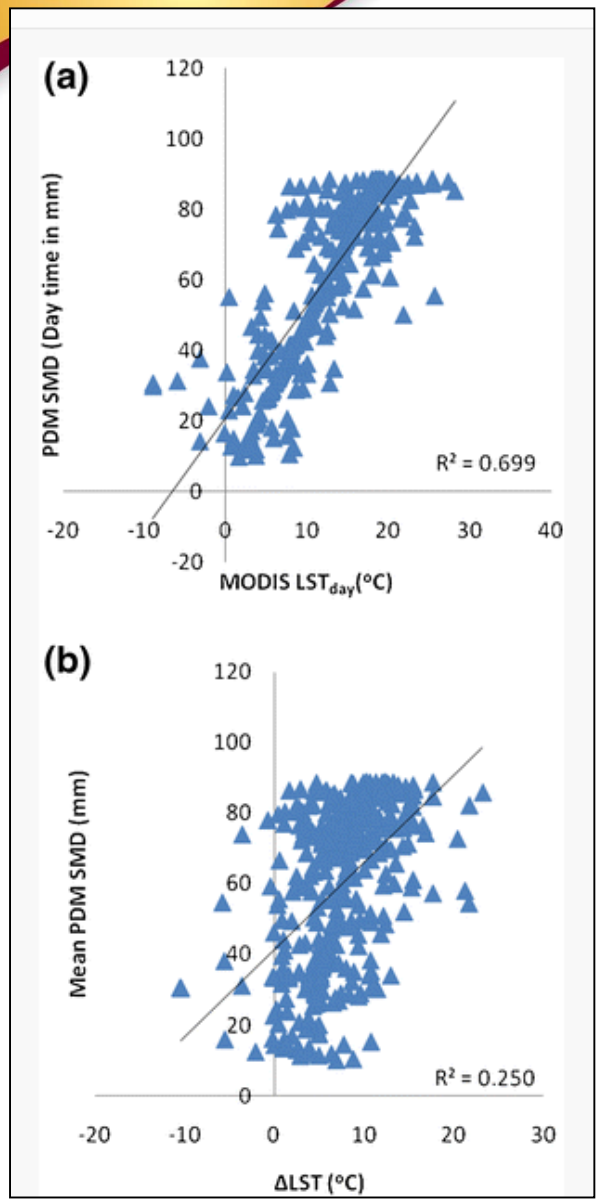


(a) Fracture initiated in bolts (b) Fracture progressed (c) Bolt fractured

Fig. 11. Fracture of bolts at the lower row.

Fig. 11, the damage of the model was also caused by the fracture of the bolts in middle joint. It is found that the quasi-static analysis using explicit dynamic solver could simulate the fracture initiation and evolution until complete failure of the bolts. The fracture on the bolts initiated near the endplate due to the bending deformation of endplate. The failure mode is similar to the experimental results as shown in Fig. 4(f). There is no fracture observed near the bottom flange of the steel beam. This is because the weld was not simulated in the model and the flange was directly connected to the web in the simulation.

Scattergram



As PDM SMD gives an indication of the available storage within the catchment and this is more hydrologically representative for comparison with satellite derived soil moisture products. Hence, we evaluated the MODIS LST products with the PDM retrieved SMD. The two MODIS products that are acquired during the day and night-time are utilized in this research. The R^2has been utilized to examine their performances with the PDM SMD. The analysis of the results reveals that ΔLST does not perform as well as the $LST_{day\text{-}time}$ products as the R^2 statistics observed from the day-time products (0.701) is far better than the statistics derived from the ΔLST products (0.253). Henceforth, only MODIS day-time products are utilized for downscaling the SMOS soil moisture. The representative statistics and related plots are shown in Fig. 4a and b. The main reason behind the poor performance of ΔLST may be attributed to MODIS night-time products, as it suffers more from night dew, relative humidity changes during night or uncertainty in day/night registration of MODIS data (Wan 1999).

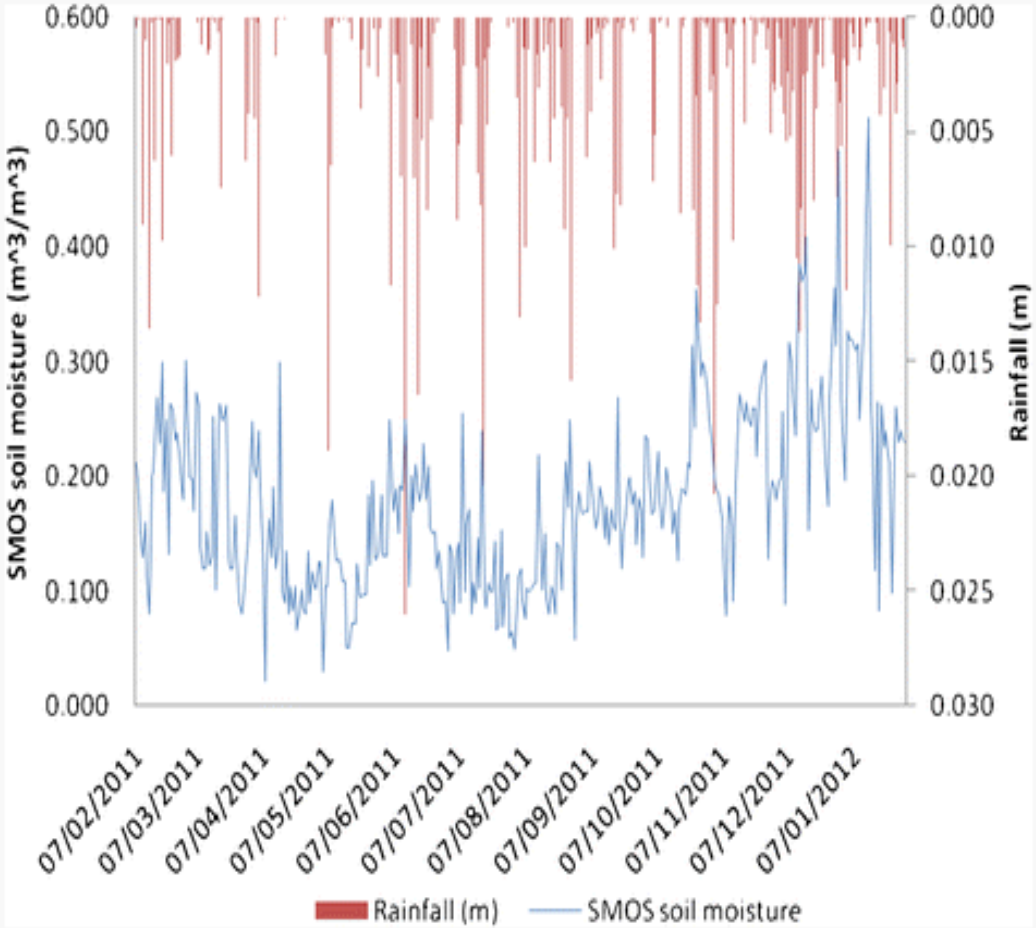
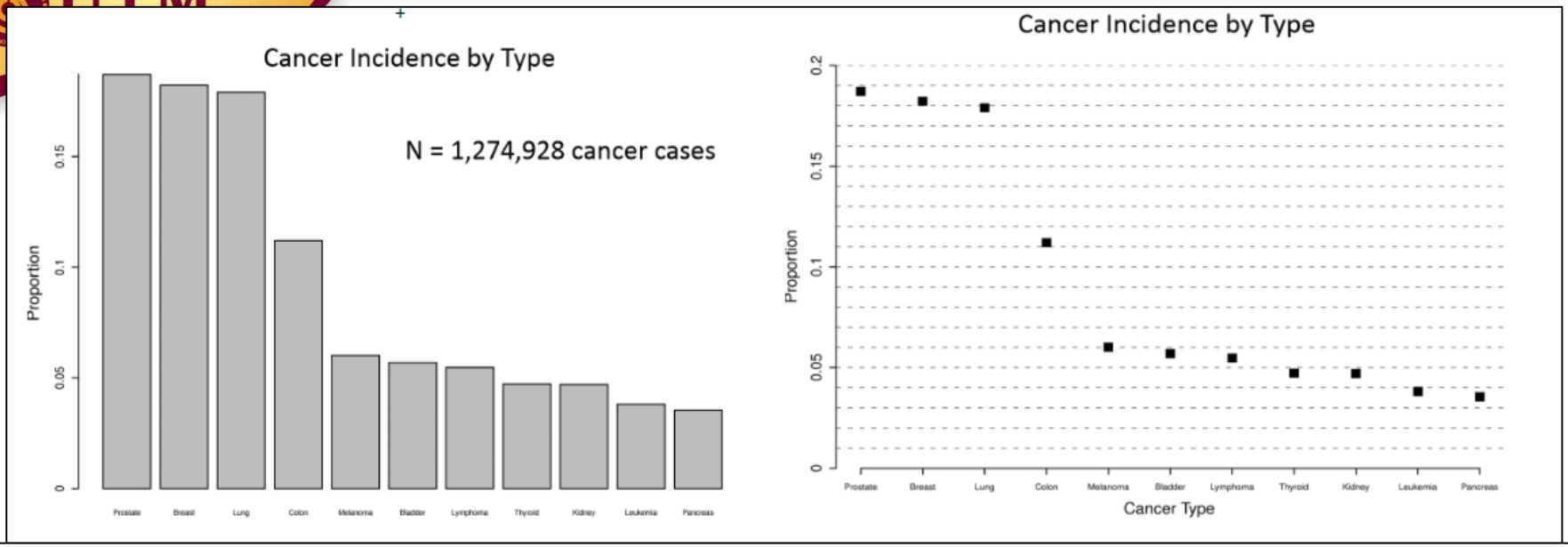
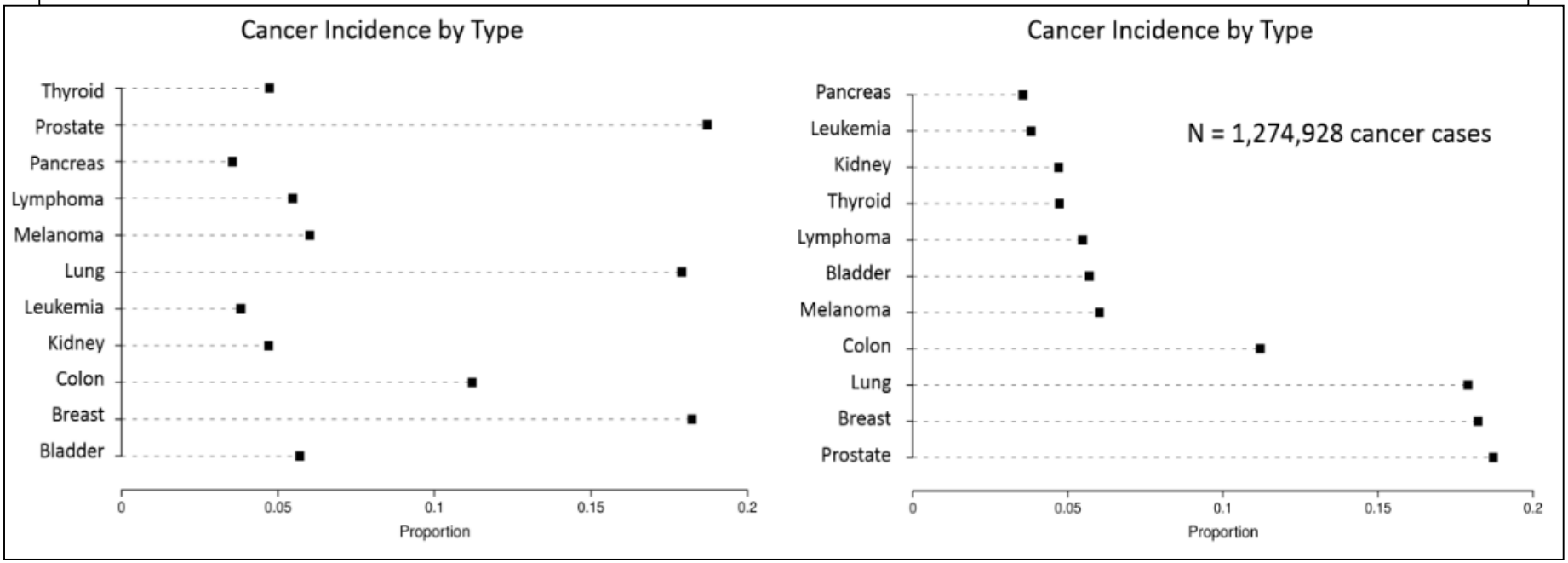


Fig. 6

SMOS soil moisture relation to rainfall

When LST over the terrain decreases in November–December, rainfall wets-up the soil profile– and hence a surging graph can be seen in Fig. 6. The rainfall pattern shows that the period between November–December is the relatively wettest period during the analysis. By contrast, March to May is slightly drier than other months. Generally, maximum rainfall intensities are higher in June than in other periods associated with moderately short storms. LST is low over winter and soils are near to the field capacity until the mid of April in most of the year. Increasing LST after the mid of April or May can lead to a substantial SMD development. Low soil moisture between April to the beginning of August (usually, the driest and warmest period of the year) can be observed in the time series. Interestingly, the LST values during the period April to the mid of August are very high, revealing the strong influence exerted by LST conditions on both surface and subsurface response. In this period the highest land surface temperature is recorded possibly due to low soil moisture record in the April and August period. The higher LST during some periods could be a possible reason that in spite of some high rainfall events, the soil moisture did not fluctuate significantly. It is observed that during the very wet conditions, soil moisture and rainfall on average started to rise at approximately at the same time.





http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/DataPresentation/DataPresentation7.html

At this point you already have your:

1. Collected data
2. Presented your findings/data appropriately

Next is to interpret the data and discuss your results critically

Interpreting Results

- The results are **interpreted** in according to the hypotheses and/or the research problem.

- These are then **discussed critically** in relation to: the existing body of knowledge, agree or disagree with the results of other research studies, and then the conclusions are drawn.

Example of Result Analysis

Result 1

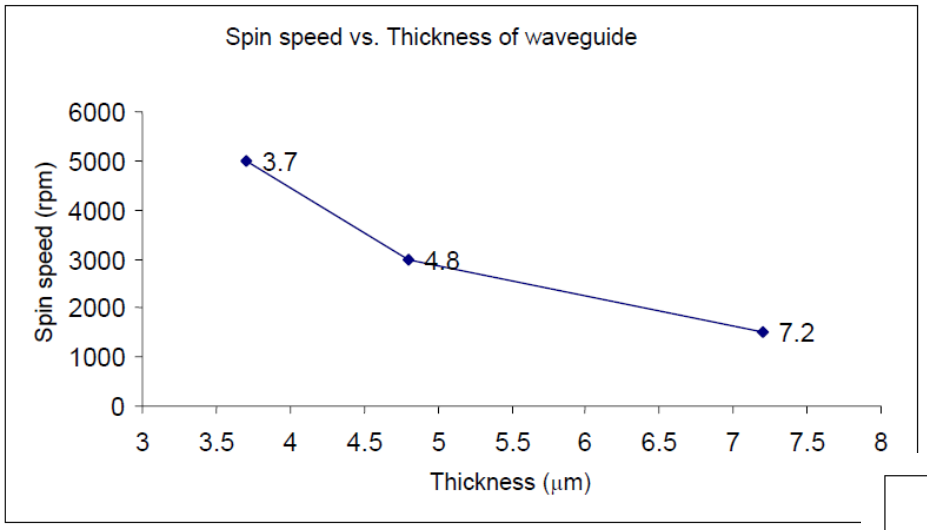


Figure 5.1: Graph of Spin Speed against Thickness

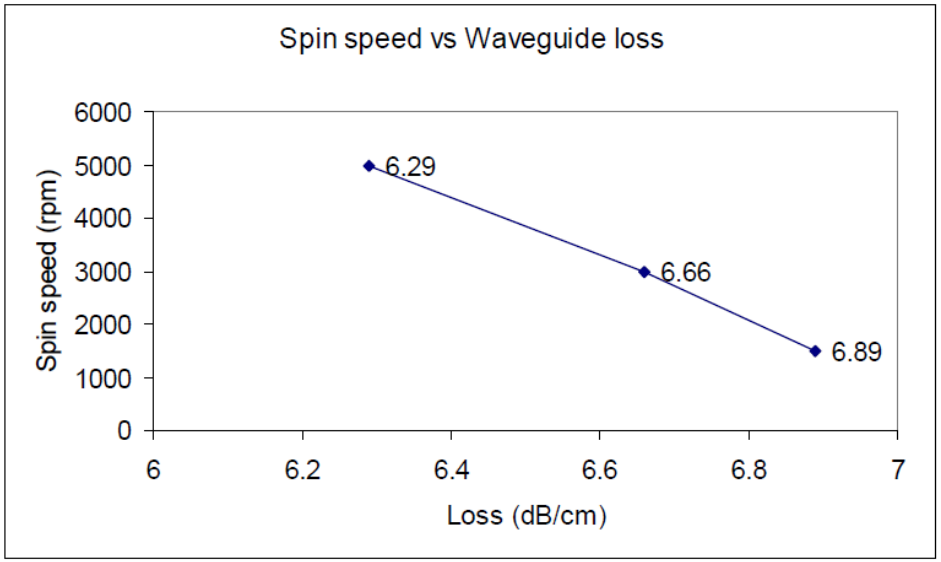


Figure 5.2: Graph of Spin Speed versus loss of waveguide

Discussion 1

Results based on graphs

The spin speed is inversely proportional to the thickness of waveguide and loss. This is due to the mode propagation in the waveguide. The thicker the waveguide, the more modes can go through the waveguide. More modes can cause the light to scatter more and thus cause higher loss. This loss is quite high compared to the measurement that has been done by other researcher. Casey F. Kane et.al [11] reported that measured loss in the waveguide equal to 0.81 dB/cm at 1330nm. These high losses were due to;

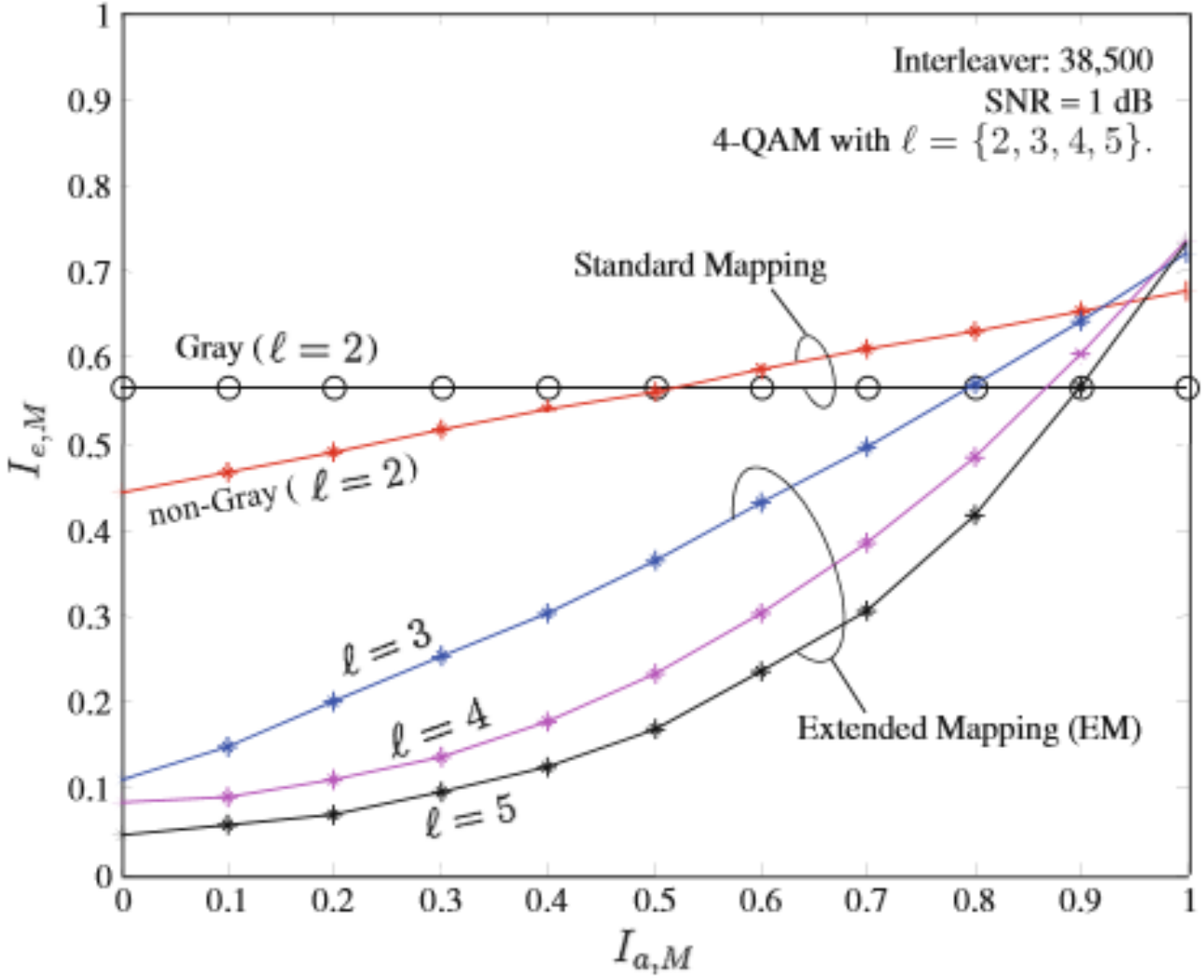
Further explanation (based on theory)

Results comparison from previous work for validation

inferences

i. Inaccuracy of the equipment that has been used,
ii. The light is not fully coupled into the waveguide, because of the small size, it is difficult to make sure that the light is fully guiding into the polymer.
iii. Interference from other light source. Detector should be carried out free from any illumination to ensure that we get the accurate reading.
iv. Difficulty in rotating the rotary stage with minimum scale.

Result 2


Interleaver: 38,500
SNR = 1 dB
4-QAM with $\ell = \{2, 3, 4, 5\}$.
Standard Mapping
Gray ($\ell = 2$)
non-Gray ($\ell = 2$)
$\ell = 3$
$\ell = 4$
$\ell = 5$
Extended Mapping (EM)
$I_{e,M}$
$I_{a,M}$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Discussion 2

Narrating your findings:

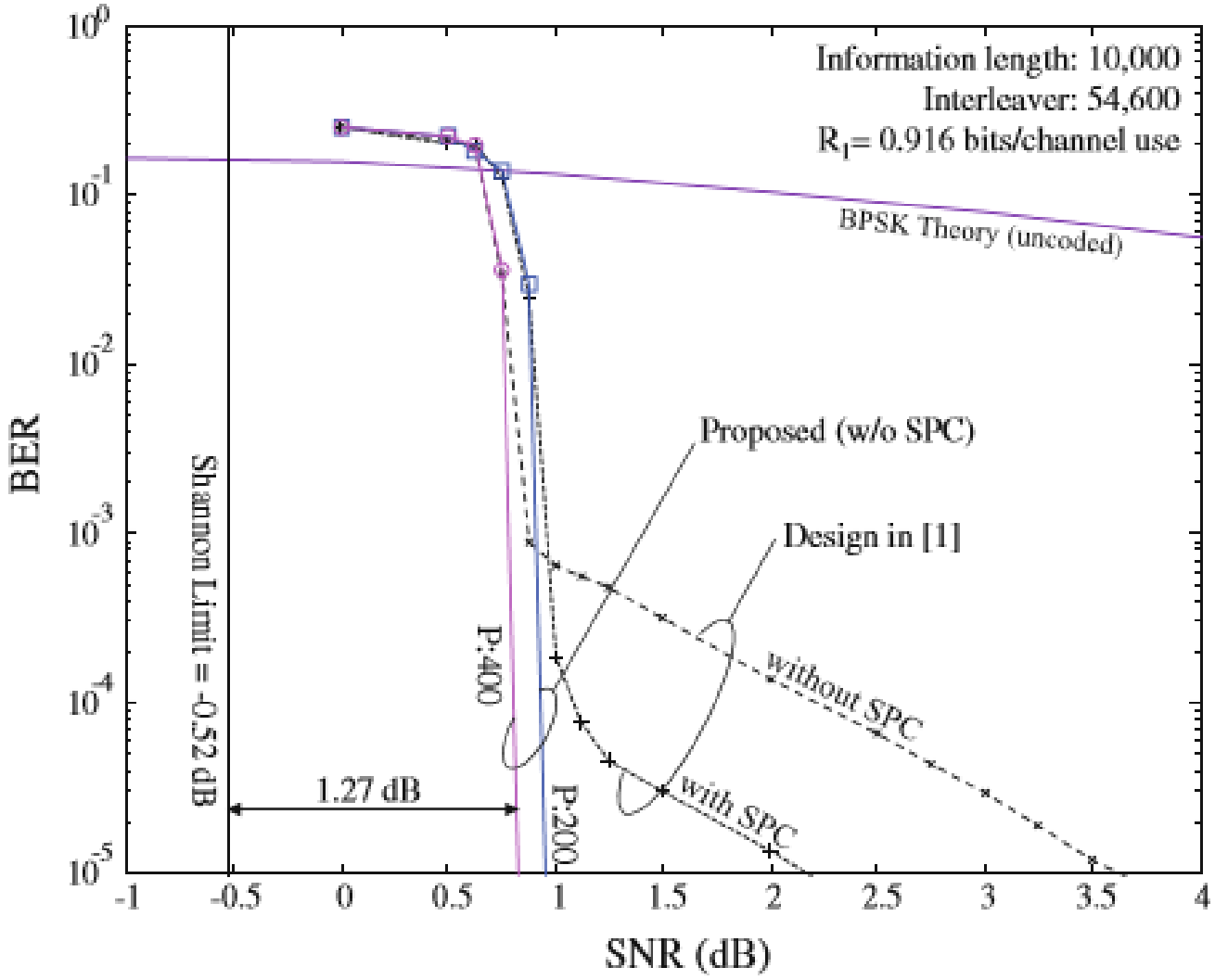
Figure 4 compares for SNR = 1 dB the EXIT functions of 4-QAM demapper for several mapping schemes. The X-axis, $I_{a,M} = I(c; L_{a,M})$, is the *a priori* MI between coded sequence c and *a priori* LLR $L_{a,M}$ provided by the decoder, while Y-axis, $I_{e,M} = I(c; L_{e,M})$, is the MI between c and its corresponding extrinsic LLR $L_{e,M}$ to be forwarded to the decoder.

A mapping with $\ell = 5$ means that each 4-QAM symbol has 5 bits resulting in 8 labeling patterns allocated to each constellation point. **Explanation**

It can be observed from Fig. 4 that the standard Gray mapping is flat, while standard non-Gray mapping (with $\ell = 2$) has *decay*. From the EXIT curves shown in Fig. 4, it is found that standard/non-extended mapping best matches turbo codes [9] and/or convolutional codes with large memory, while the extended mapping ($\ell > 2$) is better matched with the repetition codes such as the code proposed in [1]. The matching analysis between the mapper and encoder is further discussed using EXIT analysis provided by Sect. 3.

Results comparison from previous work for validation

Result 3


Information length: 10,000
Interleaver: 54,600
R_1= 0.916 bits/channel use
BPSK Theory (uncoded)
Proposed (w/o SPC)
Design in [1]
without SPC
with SPC
P:400
P:200
Shannon Limit = -0.52 dB
1.27 dB
BER
SNR (dB)
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4

Discussion 3

Figure 7 shows, for comparison, BER performances with *Case 1* of the proposed and Ref. [1]'s techniques with maximum 50 iterations. The BER curve of EM BICM-ID with SPC exhibits minor improvement, however, the error floor still remains. On the contrary, the proposed technique with ACC and doping ratio of (1:200) or (1:400), can completely eliminate the error floor even without the use of SPC. Assuming the Gaussian codebook, the Shannon limit for $\mathcal{R}_1 = 0.916$ bits/channel use is

$$\begin{aligned} \mathrm{SNR}_{lim}(\mathcal{R}_1) &= 10\log_{10}(2^{\mathcal{R}_1} - 1), \\ &= -0.52 \text{ dB}. \end{aligned} \tag{13}$$

Result 4

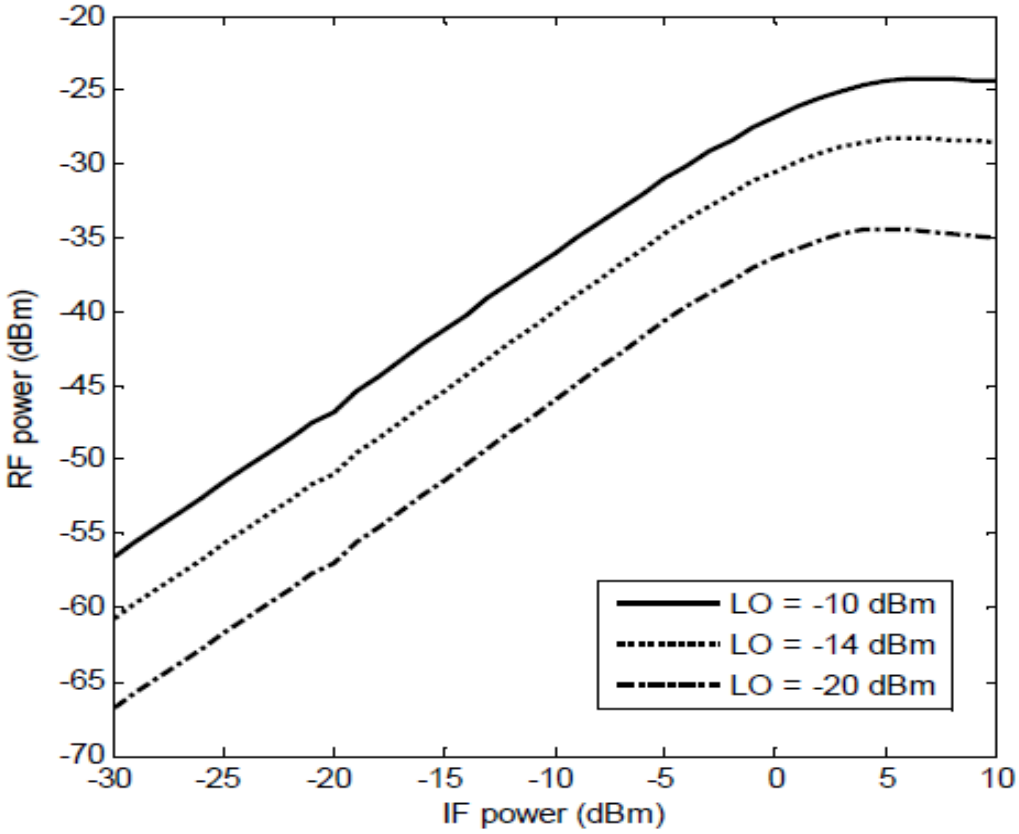
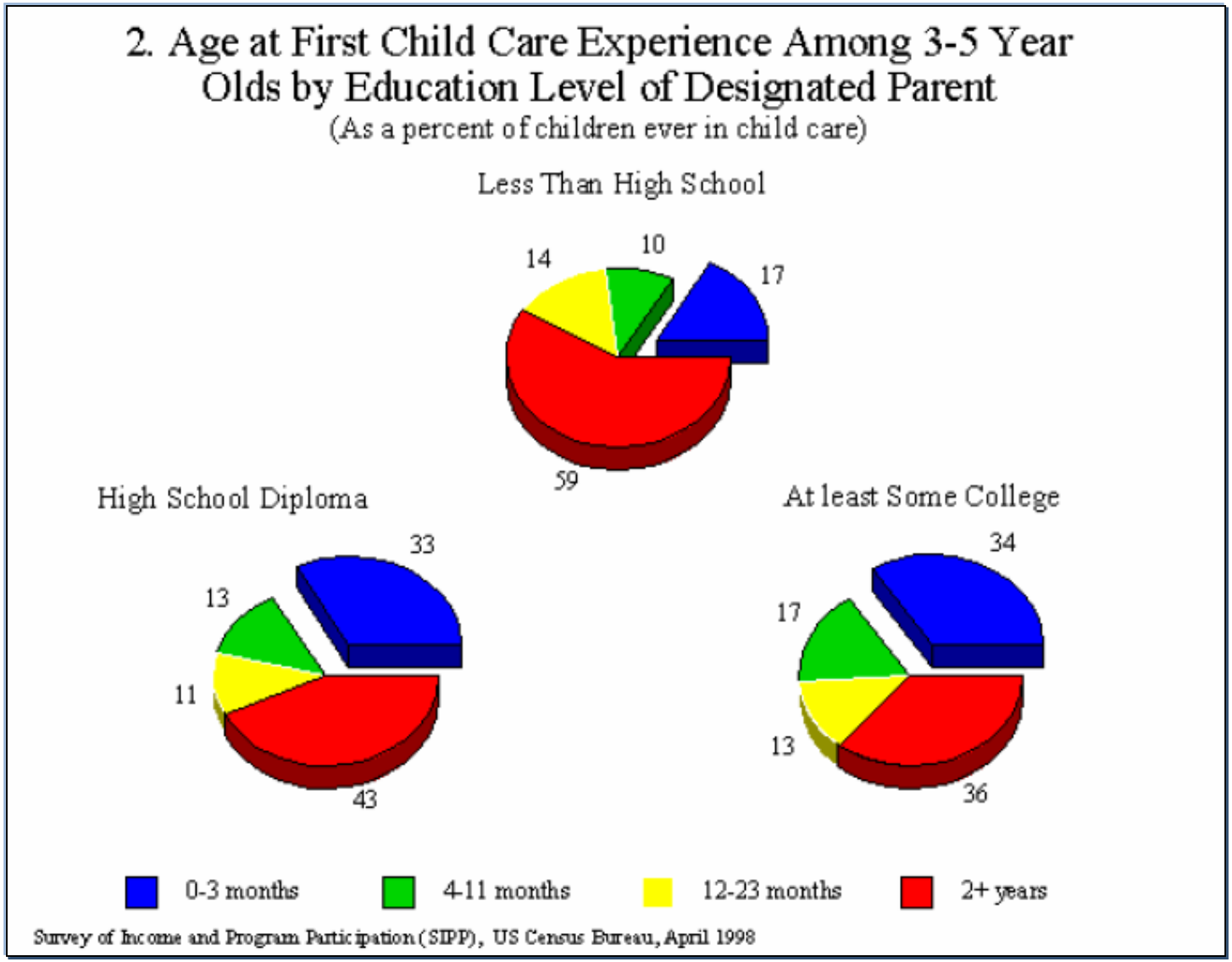


Figure 4.12 : Simulated RF (3.4 GHz) output power as a function of IF optical modulated input power at V_{BE} = 0.74 V and V_{CE} = 15 V

Discussion 4

Figure 4.12 illustrate the simulated upconversion RF output power as a function of the IF optical modulated input power with V_{BE} = 0.74 V and V_{CE} = 15 V for a different LO input power. For the P_{LO} = -10 dBm, the upconversions output signal power were increased from -56 dBm up to -25 dBm at the IF optical modulated input power from -30 dBm until 5 dBm. After the optimum output power was achieved, the RF output power slowly decreased as the IF optically modulated input signal was increased.

Exercise 1: What is the summary?



Exercise 2:
Please create a suitable table for this finding

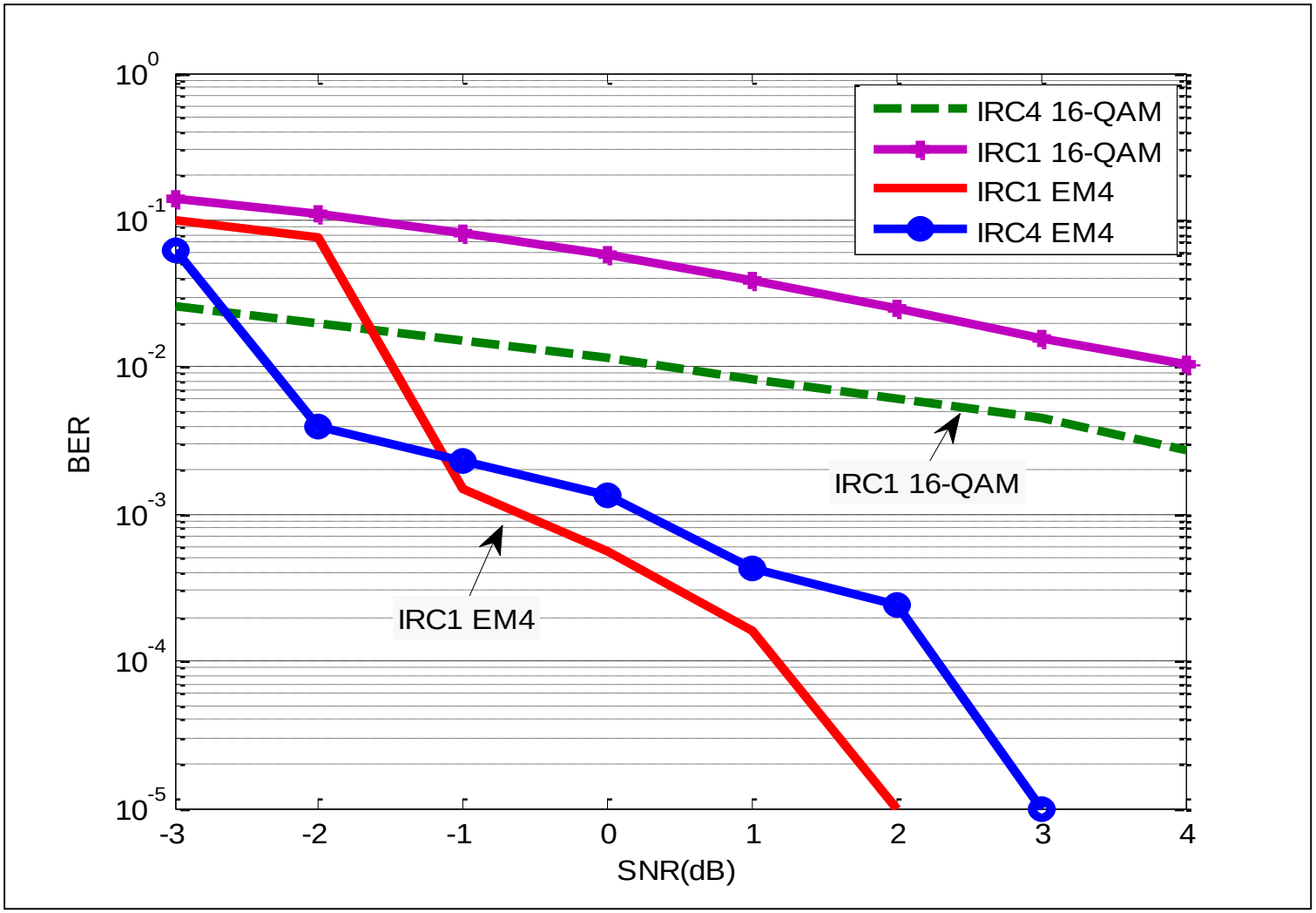


Table for Exercise 2

System	BER at SNR 2dB	Improvement
IRC1 16-QAM	3×10^{-2}	60%
IRC1 EM4	1×10^{-5}	
System	**BER at SNR 3dB**	**Improvement**
IRC4 16-QAM	4.5×10^{-3}	40%
IRC4 EM4	1×10^{-5}	

Evaluate your system performance (depends on field and application)

Signal Processing and Analysis

- Signal to noise ratio (PSNR)
- Mean square error (MSE)

Classification

- Receiver Observer Characteristic Curve (ROC)
- Type I Error, Type II Error
- Sensitivity, Specificity

Comparison

- Mean Absolute Error

Benchmark your Work

- Compare your ideas with well known method/techniques

- Show your ideas improve in certain aspect such as

- ❑ Computational time
- ❑ Reduced noise
- ❑ Fast algorithm
- ❑ Implemented in hardware
- ❑ Overall performance

Discussion

Discussions include:

1. What was learned from your research work.
2. what has been learned from this experiment?
3. From your summarized results, both in written form and visually, using graphs and charts.
4. Do your results support or disprove your hypothesis?
5. If your results do not support your hypothesis, why do you think this is the case?
6. What are the limitations of your study? These limitations are very important to acknowledge.
7. Were there any weaknesses or errors in your study design or data that may have influenced your results?

Results Vs. Discussion

What is the difference between results and discussion?

- **Results** are simply your findings. A results is strictly for narrating your findings, without trying to interpret for evaluate them.
- Done using graphs, figures, and tables.

- **Discussion** means interpreting your results and trying to explain what they mean

Example:

- If you found a some **correlation between two variables** this should be included in your results.
- **Speculating why this correlation exists**, however, belongs in the discussion section.

Significance of Findings

Significance of Findings

- What **contributions** do you think your study make to the **advancement of knowledge** and/or to the solution of some practical or theoretical problems?

- What is the **potential usefulness of the findings** of the study?

Significance of Findings

- How will the findings assist or be of **benefit to the body of knowledge?**

- How would the solution to the problem **influence the theory and practice?**

Significance of Findings

Indicate significance of the study by emphasizing on:

- Theory or **advancing accumulated knowledge**. What implications have these results or findings to existing related theories?
- Current literature
- Model development and implementation
- Given now the findings of the study, **how well** is each of **the significance realized**?


CONCLUSIÓN

Summary & Conclusions

What is **the difference between summary and conclusion**???

The **Summary** provides a **brief recap of the entire study**. Generally, this section summarizes the introduction, problem statement and hypotheses/research questions, literature review, methodology, and findings.

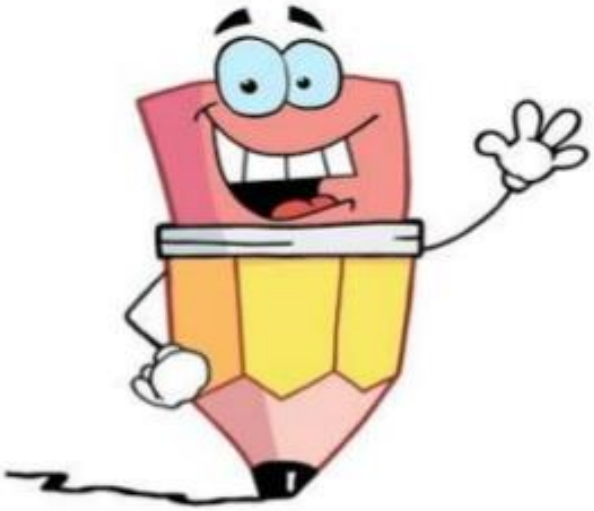
Summary & Conclusions

What is **the difference between summary and conclusion**???

- **Conclusions** are **NOT** just **short descriptions** or a short account **of your findings**.
- Show the **SIGNIFICANCE** of the research for knowledge in the discipline - what is **new and important** about your work. Because you are conveying your findings as current reality hence they are written **basically in the present tense.**
- **Focus on your most important findings**. Use your data and results to justify your conclusions.

Conclusion

Present a strong conclusion



- All your evidence and explanations should build toward a strong ending in which you summarize your view in a clear and memorable way. The conclusion in a persuasive essay might include a call to action.

Conclusion

What is Conclusion?

- Conclusion is summary of your findings
- The purpose of a results section is to ***present and illustrate your findings***.
- Make this section a completely objective report of the results, and use all interpretation for the discussion.

What should Conclusion Contain?

- Summarize your findings in text
  - ✓ Provide a context, such as by describing the question that was addressed by making a particular observation
  - ✓ Describe results of control experiments and include observations that are not presented in a formal figure or table, if appropriate
  - ✓ Analyze your data, then prepare the analyzed (converted) data in the form of a figure(graph), table, or in text form.

CONCLUSIONS

- **Should connect with the research objectives**, its importance and the ways in which it contributes to the knowledge of the discipline.

- In short, construct a much wider and general conclusions of the entire investigation

Exercise

Write a conclusion based on the information given.

OBJECTIVES

- The main objective is to develop error control mechanism for wireless transceiver design system that can achieve low BER at low SNR and closely approached the theoretical limit.

- The specific objectives:

  √ Develop low complexity Turbo FEC

  √ Enhance error detection in the proposed Turbo FEC

  √ Develop Cross-layer ARQ utilizing the enhanced Turbo FEC

CONTRIBUTIONS

1. Development of low complexity of error control using Turbo processing
2. Improvement of error detection using extended mapping
3. Improvement of BER and throughput performance using CLD_ARQ
4. Elimination of error floor in CLD_ARQ using doped-accumulator

Sample of Conclusion

- The integration of IRC as a simple channel coding and Turbo processing in the CLD_ARQ system has significantly reduce the complexity of the Turbo processing design.
- Extended mapping enhances the BER performance in the proposed system.
- CLD_ ARQ system with simple code and simple Turbo processing has been successfully improved the performance of error control mechanism by FEC and feedback transmission. Lower BER at low SNR is significantly achieved.
- Doped-accumulator is able to eliminate the error floor in the CLD_ARQ system.
- **The proposed CLD_ARQ error control is developed to achieve low BER at low SNR and closely to Shannon's limit.**

Limitation and Recommendations

Limitations of the Research

- **Recognising** that your study is still only an in depth study. **Only a very tiny aspect of the field**.
- You may also need **to demonstrate that you appreciate that it is limited**.
- Even within the most comprehensive and large scale study, **there are limitations by virtue of**:
  - The possible scope
  - Methodological restrictions
  - Practical realities
  - All claims and generalisations therefore, have to be tempered by this knowledge, and should be made using cautious language

Recommendation for Future Work

- Final stage/part of your thesis
- Generally accepted as **good practice** to **recommend areas and possibilities for further research and future work** that is indicated by the present project.
- Very few researches give clear-cut results, and most research uncovers more questions than answers.

Recommendation for Future Work

- Should **open up new questions** that can be addressed in the future.
- Recommendations could be **based upon** the:
  - **Results and conclusions**
  - **Researcher's personal opinion**
  - **Integrate previous studies.**


THE REST OF YOUR LIFE...
AVERAGE
MEMORABLE