# Statistics Workshop 1: Data Analysis with SPSS (PART 1)

**Prof. Dr Chua Yan Piaw**

*University of Malaya*

---

## Targets:

1. Concept of Data analysis

2. Prepare Data for Analysis

3. Checking the normality of a data

4. Establishing reliability of a questionnaire

5. Descriptive analysis

6. Inferential analysis

7. Analyse data and report the data

**Research Methods and Statistics Reference Books**

slide with vertical text: RESEARCH STATISTICS / DR CHUA YAN PIAW (UM)

---

**Mastering Research Methods**

Chapter 1: Introduction to Research
Chapter 2: Research Ethics
Chapter 3: Literature Review
Chapter 4: Research Design
Chapter 5: Experimental Study
Chapter 6: Quasi-experimental Study
Chapter 7: Survey Study
Chapter 8: Field Study
Chapter 9: Case Study
Chapter 10: Action Study
Chapter 11: Historical Study
Chapter 12: Probability Sampling Procedures
Chapter 13: Non-probability Sampling Procedures
Chapter 14: Measurement in Research
Chapter 15: Index, Scales and Specific
 Measurement Procedures
Chapter 16: Pilot Study
Chapter 17: Research Instrumentation
Chapter 18: Format of Writing Research Report

CHUA YAN PIAW
**Mastering RESEARCH METHODS**

2011

CHUA YAN PIAW
2nd edition
**MASTERING RESEARCH METHODS**

MASTERING RESEARCH METHODS
2nd Edition
CHUA YAN PIAW

2016

**Mastering Research Statistics**

Chapter 1: Descriptive Statistics
Chapter 2: Inferential Statistics and
         Significance Test
Chapter 3: Qualitative Data Analysis
Chapter 4: Data Preparation for SPSS
         Program
Chapter 5: Reliability of Research
         Instrument
Chapter 6: Chi-Square Tests
Chapter 7: T Tests
Chapter 8: ANOVA Tests
Chapter 9: Correlation Tests
Chapter 10: Multiple Regressions
Chapter 11: Multiple Responses Analysis
Chapter 12: Reporting the Results of Data
         Analysis Based on the
         APA Format

Chua Yan Piaw

**Mastering**
Research
Statistics

Mc Graw Hill Education

**Amazon.com or MPHonline**

# Concept of Data analysis

# Research design and statistics

*Planning a design without the knowledge of statistics, the researcher will find difficult to analyse the data after collecting it.*

# Research Statistics

*Determined by several factors:*

*Factor 1. Type of study*

*Factor 2. Type of measurement scales in an instrument*

*Factor 3. Type of Statistical tests*

*Factor 4. Sample size*

# *Factor 1*

## 1. Types of study:

    **a.  Descriptive study**

    **b.  Inferential study**

---

## *A descriptive study*

•**No sample is drawn from the population, so the results is not generalised. So no inferential statistical test is needed.**

- *In a descriptive study,*

*No sample is drawn from the population.*

*Respondents are the whole population of the study.*

*The results/ findings are not generalised to any subject outside the population .*

---

*Descriptive statistics describes the characteristics (variables) in the population.*

**Frequency**      **Median**

**Percentage**   **Mode**

**Mean**              **Standard deviation**

**Distribution of Scores**

*The results are nearly 100% correct for that population.*

*For example:*
**Shooting performance *of a group of shooters*.**

*1. Collect data (shooting performance) of shooters.*

*2. Calculate the mean score.*

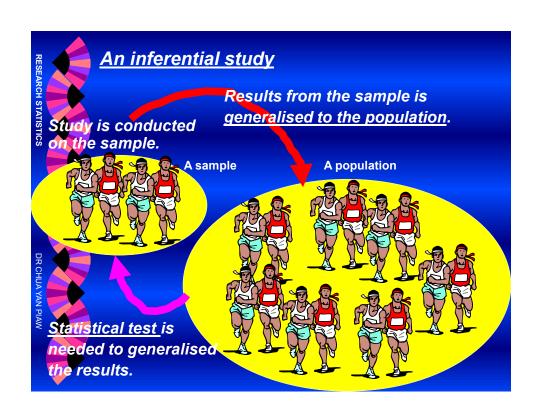*3. It represents the shooting performance of the shooters*
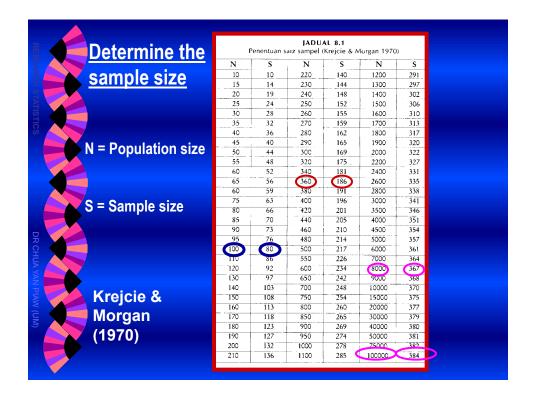
*Did I select a sample?*

*Can I generalize the result to other group? Shooters from other location/country?*

*The result is nearly 100% correct (if the measurement is reliable and is correctly done).*

---

# *B. For an inferential study:*

- *Respondents are subjects selected randomly from a population, to form a sample.*

- *Statistics tests is used to analyse data collected from the sample.*

- *The results / findings are generalised back to the population from where the sample was selected.*

## *An inferential study*

*Results from the sample is generalised to the population.*

*Study is conducted on the sample.*

**A sample**          **A population**

*Statistical test is needed to generalised the results.*

---

## **Determine the sample size**

**N = Population size**

**S = Sample size**

**Krejcie & Morgan (1970)**

**JADUAL 8.1**
Penentuan saiz sampel (Krejcie & Morgan 1970)

| N | S | N | S | N | S |
|---|---|---|---|---|---|
| 10 | 10 | 220 | 140 | 1200 | 291 |
| 15 | 14 | 230 | 144 | 1300 | 297 |
| 20 | 19 | 240 | 148 | 1400 | 302 |
| 25 | 24 | 250 | 152 | 1500 | 306 |
| 30 | 28 | 260 | 155 | 1600 | 310 |
| 35 | 32 | 270 | 159 | 1700 | 313 |
| 40 | 36 | 280 | 162 | 1800 | 317 |
| 45 | 40 | 290 | 165 | 1900 | 320 |
| 50 | 44 | 300 | 169 | 2000 | 322 |
| 55 | 48 | 320 | 175 | 2200 | 327 |
| 60 | 52 | 340 | 181 | 2400 | 331 |
| 65 | 56 | 360 | 186 | 2600 | 335 |
| 60 | 59 | 380 | 191 | 2800 | 338 |
| 75 | 63 | 400 | 196 | 3000 | 341 |
| 80 | 66 | 420 | 201 | 3500 | 346 |
| 85 | 70 | 440 | 205 | 4000 | 351 |
| 90 | 73 | 460 | 210 | 4500 | 354 |
| 95 | 76 | 480 | 214 | 5000 | 357 |
| 100 | 80 | 500 | 217 | 6000 | 361 |
| 110 | 86 | 550 | 226 | 7000 | 364 |
| 120 | 92 | 600 | 234 | 8000 | 367 |
| 130 | 97 | 650 | 242 | 9000 | 368 |
| 140 | 103 | 700 | 248 | 10000 | 370 |
| 150 | 108 | 750 | 254 | 15000 | 375 |
| 160 | 113 | 800 | 260 | 20000 | 377 |
| 170 | 118 | 850 | 265 | 30000 | 379 |
| 180 | 123 | 900 | 269 | 40000 | 380 |
| 190 | 127 | 950 | 274 | 50000 | 381 |
| 200 | 132 | 1000 | 278 | 75000 | 382 |
| 210 | 136 | 1100 | 285 | 100000 | 384 |

## Inferential statistics
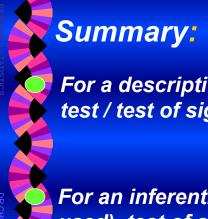
● **Tests of differences**

**T tests**

*ANOVA tests*

*Chi-square tests*

*Mann-Whitney U*

*Kruskal-Wallis test*
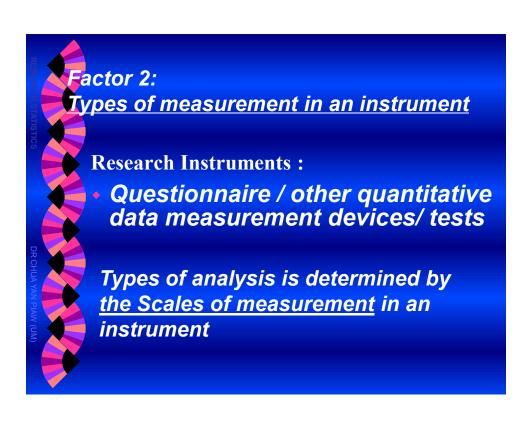
● **Tests of relationship**
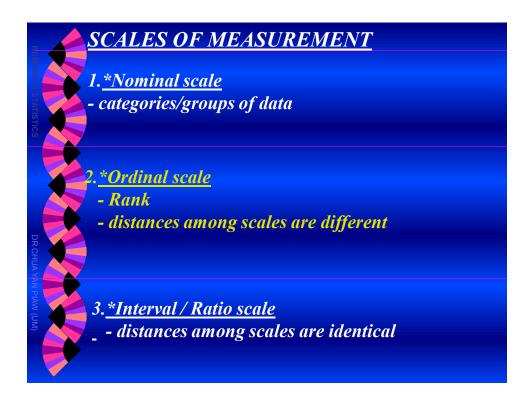
*Pearson r*

*Spearman rho*

*Cramer V*

---

## *Summary:*
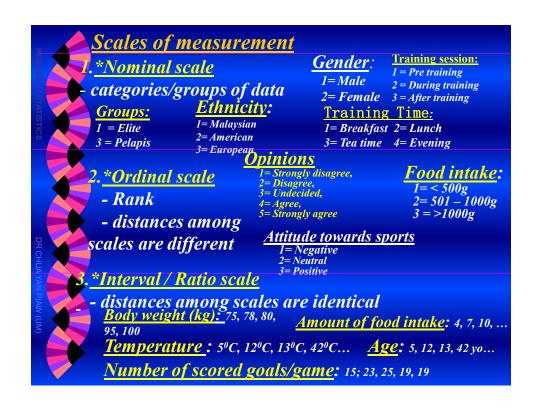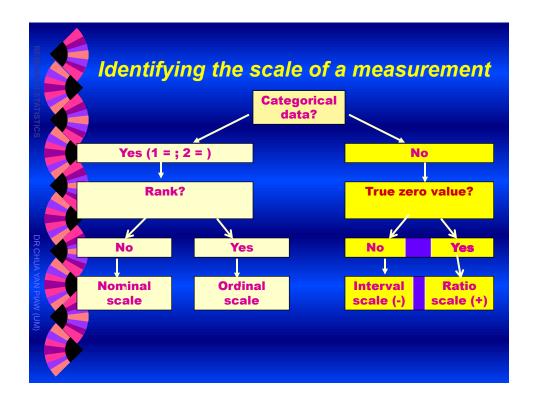
● *For a descriptive study, no inferential test / test of significance is needed.*

● *For an inferential study (a sample is used), test of significance is needed to generalise the result to the population.*

**Factor 2:**
**Types of measurement in an instrument**

**Research Instruments :**

- **Questionnaire / other quantitative data measurement devices/ tests**

**Types of analysis is determined by the Scales of measurement in an instrument**

---

**SCALES OF MEASUREMENT**

**1.*Nominal scale**
**- categories/groups of data**

**2.*Ordinal scale**
  **- Rank**
  **- distances among scales are different**

**3.*Interval / Ratio scale**
  **- distances among scales are identical**

## Scales of measurement

### 1. *Nominal scale

- categories/groups of data

**Groups:**
1 = Elite
3 = Pelapis

**Ethnicity:**
1= Malaysian
2= American
3= European

**Gender:**
1= Male
2= Female

**Training session:**
1 = Pre training
2 = During training
3 = After training

**Training Time:**
1= Breakfast  2= Lunch
3= Tea time   4= Evening

### 2. *Ordinal scale

- Rank
- distances among scales are different

**Opinions**
1= Strongly disagree,
2= Disagree,
3= Undecided,
4= Agree,
5= Strongly agree

**Food intake:**
1= < 500g
2= 501 – 1000g
3 = >1000g

**Attitude towards sports**
1= Negative
2= Neutral
3= Positive

### 3. *Interval / Ratio scale

- distances among scales are identical

**Body weight (kg):** 75, 78, 80, 95, 100

**Amount of food intake:** 4, 7, 10, …

**Temperature :** $5^0C, 12^0C, 13^0C, 42^0C$…

**Age:** 5, 12, 13, 42 yo…

**Number of scored goals/game:** 15; 23, 25, 19, 19

---

## Identifying the scale of a measurement



Categorical data?
- Yes (1 = ; 2 = )
  - Rank?
    - No → Nominal scale
    - Yes → Ordinal scale
- No
  - True zero value?
    - No → Interval scale (-)
    - Yes → Ratio scale (+)

*Identify the scale of measurement.*

1. **The main factors of your performance:**
   A. Food intake factors
   B. Sleep deprivation factors
   C. Psychological factors
   D. Injury factors

2. *Your attitude towards the performance of our national football players is positive.*

   Strongly agree    1   2   3   4   5    Strongly disagree

3. *Do you agree that the main cause of obesity is too much intake of oily foods.*
   1. Yes
   2. No

---

*Please indicate your Body Mass Index with the below scale:*

Less than 10 kg.m²-1 ………………..…1
10 to 20 kg.m²-1 ……….…………………2
21 to 30 kg.m²-1 …………….….………..3
More than 30 kg.m²-1 ……………..……4

**Results**
Table 1. Descriptive Statistics of the Variables

| Variables | Mean | Std. Deviation | N |
|---|---|---|---|
| Performance | 19.60 | 8.88 | 20 |
| Motivation | 4.48 | 0.02 | 20 |
| Confidence | 3.30 | 0.40 | 20 |
| Anxiety control | 3.12 | 0.09 | 20 |
| Mental preparation | 4.00 | 0.05 | 20 |
| Concentration | 3.18 | 0.07 | 20 |
| Cognition | 3.30 | 0.08 | 20 |

Table 1 indicates the descriptive statistics of the variables. The Mean, Standard deviation, as well as the number of the participants, is shown. Performance as the dependent variable has the highest number of mean and standard deviation while the mean variances of the psychological factors are closely associated with one another which indicated that they contribute to the prediction of performance in a relatively similar pattern.

-------------------------------------------------------------------------------------------------------- 171

*JPES* ®     www.efsupit.ro

### Sport team

1. **Selangor**
2. **Perak**
3. **Kedah**

---

*Body Fat Percentage*
**1. Low**
**2. High**

### Physical Sport:

*1.Board games*
*2.Cycling*
*3.Gymnastics*
*4.Shooting sports*

---

| | Calm | 40 | 35 | 30 | 25 | 20 | 15 | 10 | 5 | 0 | -5 | -10 | -15 | -20 | -25 | -30 | -35 | -40 | -45 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 36 | 31 | 25 | 19 | 13 | 7 | 1 | -5 | -11 | -16 | -22 | -28 | -34 | -40 | -46 | -52 | -57 | -63 |
| | 10 | 34 | 27 | 21 | 15 | 9 | 3 | -4 | -10 | -16 | -22 | -28 | -35 | -41 | -47 | -53 | -59 | -66 | -72 |
| | 15 | 32 | 25 | 19 | 13 | 6 | 0 | -7 | -13 | -19 | -26 | -32 | -39 | -45 | -51 | -58 | -64 | -71 | -77 |
| | 20 | 30 | 24 | 17 | 11 | 4 | -2 | -9 | -15 | -22 | -29 | -35 | -42 | -48 | -55 | -61 | -68 | -74 | -81 |
| | 25 | 29 | 23 | 16 | 9 | 3 | -4 | -11 | -17 | -24 | -31 | -37 | -44 | -51 | -58 | -64 | -71 | -78 | -84 |
| Wind (mph) | 30 | 28 | 22 | 15 | 8 | 1 | -5 | -12 | -19 | -26 | -33 | -39 | -46 | -53 | -60 | -67 | -73 | -80 | -87 |
| | 35 | 28 | 21 | 14 | 7 | 0 | -7 | -14 | -21 | -27 | -34 | -41 | -48 | -55 | -62 | -69 | -76 | -82 | -89 |
| | 40 | 27 | 20 | 13 | 6 | -1 | -8 | -15 | -22 | -29 | -36 | -43 | -50 | -57 | -64 | -71 | -78 | -84 | -91 |
| | 45 | 26 | 19 | 12 | 5 | -2 | -9 | -16 | -23 | -30 | -37 | -44 | -51 | -58 | -65 | -72 | -79 | -86 | -93 |
| | 50 | 26 | 19 | 12 | 4 | -3 | -10 | -17 | -24 | -31 | -38 | -45 | -52 | -60 | -67 | -74 | -81 | -88 | -95 |
| | 55 | 25 | 18 | 11 | 4 | -3 | -11 | -18 | -25 | -32 | -39 | -46 | -54 | -61 | -68 | -75 | -82 | -89 | -97 |
| | 60 | 25 | 17 | 10 | 3 | -4 | -11 | -19 | -26 | -33 | -40 | -48 | -55 | -62 | -69 | -76 | -84 | -91 | -98 |

## Scales of measurement

**Interval and Ratio** data are **parametric data** (the data are assumed normal distributed)

*Example:*
*Scores of a group*

*Parametric tests are used:*
*T tests,*
*ANOVA tests*
*Pearson Correlation test,*
*etc*

---

## Scales of measurement

**Nominal** and **Ordinal** data are **non-parametric data** (the data are assumed not normal distributed)

*Level of agreement*
*Strongly disagree - disagree -------- undecided -------- agree - strongly agree*
*1          2                    3                  4          5*

### Non-parametric tests are used:
**Chi-square tests (Nominal/ordinal data)**
**Mann-Whitney test (Ordinal data)**
**Kruskal-Wallis test (Ordinal data)**
**Spearman rho Correlation test (Ordinal data)**
**Cramer V correlation test (Nominal data)**
**etc**

## Example:
### Stress level before the game

**data: Ratio scale** (stress score, exp: 34, 45, 8, 33, 56, 78, 34, 23, 67, 55)

**The Difference od stress between male and female players: use the parametric test - T test**

**data: Ordinal scale**
(1=Very low, 2=Low, 3=Average, 4=High, 5=Very high)

**The Difference of stress between male and female players: use the non-parametric test - Mann-Whitney test**

---

## Factor 3: Sample size

*In inferential statistics, Two assumption are:*

● *1. The data for continouos data (interval / ratio) is normal distributed when n => 30.*
*Each sub-sample: n =>15*

● *2. Data for categorical data are not-normally distribued.*
*Each sub-sample => 5*

# *Question?*

# *Take 5?*