

Conducting Descriptive Statistics

Dr. K. A. Korb
University of Jos

Outline

- Frequency
- Measures of Central Tendency
 - Mode
 - Mean
 - Median
- Measures of Variability
 - Range
 - Variance
 - Standard Deviation

Dr. K. A. Korb
University of Jos

Frequency

- Frequency: Number of times a score occurs.
- Frequencies can either be reported by a table or by a chart.
- Reporting frequency is typically only informative for nominal (categorical) data

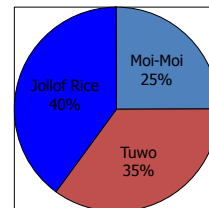
<i>Food is Ready Menu</i>					
	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	Moi-Moi	Moi-Moi	Jollof Rice	Jollof Rice	Tuwo
Week 2	Tuwo	Jollof Rice	Tuwo	Moi-Moi	Jollof Rice
Week 3	Moi-Moi	Tuwo	Jollof Rice	Jollof Rice	Tuwo
Week 4	Jollof Rice	Moi-Moi	Tuwo	Tuwo	Jollof Rice

Menu Availability Frequency Table	
Food	Frequency
Tuwo	7
Moi Moi	5
Jollof Rice	8

Dr. K. A. Korb
University of Jos

Frequency Pie Chart

Percentage of Availability of Menu Options at Food is Ready



Dr. K. A. Korb
University of Jos

Where did UniJos PhD students earn their Masters Degree?

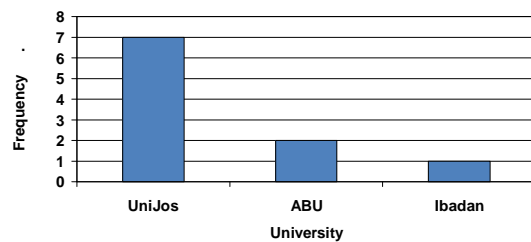
S/N	Uni
1	UniJos
2	ABU
3	UniJos
4	UniJos
5	UniJos
6	UniJos
7	UniJos
8	ABU
9	UniJos
10	Ibadan

Masters Degree Frequency Table	
University	Frequency
UniJos	7
ABU	2
Ibadan	1

Dr. K. A. Korb
University of Jos

Frequency Bar Chart

University where PhD students read their Masters Degree



Dr. K. A. Korb
University of Jos

Types of Statistics

- Three fundamental types of statistics
 1. **Descriptive:** Explains trends in your sample
 - Central Tendency
 - Mode
 - Mean
 - Median
 - Variability
 - Range
 - Standard Deviation
 - Frequencies
 2. **Significance of Means:** Determines whether differences between groups of individuals are large enough to be meaningful
 - t-tests, ANOVA, ANCOVA
 3. **Relationship between Variables:** Compares the relationship between multiple variables within the same sample of individuals
 - Correlation
 - Regression

Dr. K. A. Korb
University of Jos

Descriptive Statistics

- June received a score of 20 on the Extraversion Personality Questionnaire.
 - With just this information, we cannot interpret her score.
 - What does the average person score on the Questionnaire?
 - What is the range of typical scores on the Questionnaire?
- What two things do you need to know to interpret her score?
 - Average: Typical score
 - Average: 30
 - Range of scores
 - Standard Deviation: 10
- Now we can say two things:
 - June has less Extraversion than the typical person because her score of 20 was less than the average score of 30.
 - June is one standard deviation below the mean ($30 - 20 = 10$, the standard deviation), so she has considerably less extraversion than most people.

Dr. K. A. Korb
University of Jos

Central Tendency

- Average: Typical performance
 - Mode: Most frequent score
 - Mean: Sum of scores divided by number of scores
 - Median: Middle score in the distribution
- The next slides give an example of the **Mode**.

Dr. K. A. Korb
University of Jos

What is the typical meal served at the *Food is Ready*?

<i>Food is Ready</i> Menu					
	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	Moi-Moi	Moi-Moi	Jollof Rice	Jollof Rice	Tuwo
Week 2	Tuwo	Jollof Rice	Tuwo	Moi-Moi	Jollof Rice
Week 3	Moi-Moi	Tuwo	Jollof Rice	Jollof Rice	Tuwo
Week 4	Jollof Rice	Moi-Moi	Tuwo	Tuwo	Jollof Rice

- Frequency
 - Tuwo: 7
 - Moi-Moi: 5
 - Jollof Rice: 8
- Jollof Rice is on the menu the most frequently, so the **Mode is Jollof Rice**.

Dr. K. A. Korb
University of Jos

Central Tendency

- Average: Typical performance
 - Mode: Most frequent score
 - Best for nominal (categorical) data
 - Represent by a pie graph
 - Mean: Sum of scores divided by number of scores
 - Median: Middle score in the distribution
- The next slides give explain the **Mean**.

Dr. K. A. Korb
University of Jos

What is the typical price of oranges at the market?

Cost
50
50
55
60
60
65
70
70
70
70

Mean

$$\frac{50+50+55+60+60+65+70+70+70+70}{10} = \frac{620}{10} = 62$$

Mean = 62

We can also find the **Mode**, the most frequent score.

Mode = 70

Dr. K. A. Korb
University of Jos

What is the typical price of oranges at the market?

Cost
50
50
55
60
60
65
70
70
70
70
150

The next person who goes to the market is a bature and they get charged N150 for the oranges. Let's recalculate.

Mean

$$\frac{50+50+55+60+60+65+70+70+70+70+150}{11} = \frac{770}{11} = 70$$

Mean = 70

The mean has jumped from 62 to 70 with just one additional data point.

Dr. K. A. Korb
University of Jos

Price of Oranges

This frequency chart clearly shows that the N150 purchase is an outlier – a data point that is far from the other data points.

Dr. K. A. Korb
University of Jos

Price of Oranges

Cost
50
50
55
60
60
65
70
70
70
70
70

The **Median** is the middle score. First calculate the median for the data without the bature purchase.

Median

↓

50 50 55 60 65 70 70 70 70

Median = 62.5

When arranged from smallest to largest, there are 5 data points to the left and 5 data points to the right. To find the median with an even set of data points, calculate the mean of the two middle numbers – 60 and 65.

Dr. K. A. Korb
University of Jos

Price of Oranges

Cost
50
50
55
60
60
65
70
70
70
70
70
150

Now let's calculate the median with the bature purchase.

Median

↓

50 50 55 60 65 70 70 70 70 150

Median = 65

When arranged from smallest to largest, there are 5 data points to the left and 5 data points to the right of the number 65. The median with an odd set of data points is simply the middle number.

Dr. K. A. Korb
University of Jos

Central Tendency

	Without Bature Purchase	With Bature Purchase
Mean	62	70
Median	62.5	65

- The Mean changed substantially with the outlier bature purchase.
- The median was not strongly influenced by the outlying score.

Dr. K. A. Korb
University of Jos

Central Tendency

- Average: Typical performance
 - Mode: Most frequent score
 - Best for categorical data
 - Represent by a pie graph
 - Mean: Sum of scores divided by number of scores
 - Most mathematically defensible
 - Calculate and report for virtually all numerical data
 - Affected by skew
 - Median: Middle score in the distribution
 - Best for skewed data

Dr. K. A. Korb
University of Jos

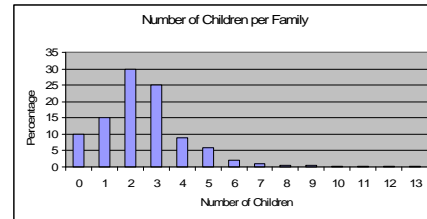
Variability

- **Variability:** The spread of scores
 - Range: Highest and lowest scores in the distribution
 - Variance: Mathematical degree of spread
 - Standard Deviation: Mathematical index of spread in original measurement units
- The next slides give an example of the **Range**.

Dr. K. A. Korb
University of Jos

Variability

- **Range:** Report highest and lowest values
 - *The number of children ranged from 0 to 13.*



Dr. K. A. Korb
University of Jos

Variability

- **Variability:** The spread of scores
 - Range: Highest and lowest scores in the distribution
 - Simply gives readers an idea of the spread of scores.
 - Not mathematically useful for calculating statistics.
 - Variance: Mathematical degree of spread
 - Standard Deviation: Mathematical index of spread in original measurement units
- The next slides give an example of the **Variance**.

Dr. K. A. Korb
University of Jos

Variability

Scores on a 15 point Continuous Assessment		
Score	Deviation	Deviation ²
11	1	1
8	-2	4
7	-3	9
13	3	9
12	2	4
9	-1	1
8	-2	4
12	2	4
Sum	0	36
Divide by 8		4.5

- To Calculate the Variability**
1. Subtract all scores from the mean (deviation)
 - If we summed up the deviation scores, they will *always* add up to 0 because of the mathematical properties of the mean.
 - To solve this problem, we square the deviation.
 2. Square the deviation
 3. Sum the deviation²
 4. Divide by total number of scores
 - 4.5
- Since we are summing up the deviation and dividing by the total number of scores, the variance is effectively the average squared deviation of each score from the mean.

Dr. K. A. Korb
University of Jos
Mean = 10

Variability

- **Variability:** The spread of scores
 - Range: Highest and lowest scores in the distribution
 - Simply gives readers an idea of the spread of scores.
 - Not mathematically useful for calculating statistics.
 - Variance: Mathematical degree of spread
 - The variance is difficult to interpret because it is in squared units of the original data points.
 - Standard Deviation: Mathematical index of spread in original measurement units
- The next slides give an example of the **Standard Deviation**.

Dr. K. A. Korb
University of Jos

Variability

Scores on a 15 point Continuous Assessment		
Score	Deviation	Deviation ²
11	1	1
8	-2	4
7	-3	9
13	3	9
12	2	4
9	-1	1
8	-2	4
12	2	4
Sum	0	36
Divide by 8		4.5
Square Root		2.12

- To Calculate the Standard Deviation**
- Calculate the variance
 - Take the square root of the variance
 - By taking the square root of the variance, the Standard Deviation (SD) is back in the original units.
 - A SD of 2.12 means that the typical person deviates from the mean score of 10 by about 2.12 points.

Dr. K. A. Korb
University of Jos

Variability

- **Variability: The spread of scores**
 - **Range: Highest and lowest scores in the distribution**
 - Simply gives readers an idea of the spread of scores.
 - Not mathematically useful for calculating statistics.
 - **Variance: Mathematical degree of spread**
 - The variance is difficult to interpret because it is in squared units of the original data points.
 - **Standard Deviation: Mathematical index of variability in original measurement units**
 - The Standard Deviation can be interpreted in the original scale.
 - The Standard Deviation is used in many statistical procedures.