## GEOGRAPHY NOTES FORM 2

## STATISTICAL METHDIDS II



## Specific Objectives

By the end of the topic, the learner should be able to:
a. present statistical data using appropriate methods
b. analyse and interpret statistical data
c. explain the advantages and disadvantages of each method of data presentation.

## STATISTICAL METHODS II

## Statistical Presentation, Analysis and Interpretation

Three types of graphs for statistical presentation are considered. They are;

1. Comparative line graph.
2. Comparative bar graph.
3. Divided rectangles.

## Graphs

Two dimensional drawings which show relationships between 2 types of data representing two items also called variables. These are dependent variable which is affected by the other e.g. temperature (on y axis) and independent variable whose change is not affected by the other e.g. altitude (on x axis).

## Steps

1. Draw x and y axis.
2. Choose suitable scale to accommodate the highest and lowest value.
3. Plot the values accurately using faint dots.
4. Join the dots using curved line. If it's a bar graph the dots should be at the middle of the top line. Years should also be at the middle. You should have also decided on the width of the bars.
5. In data without continuity e.g. crop production there should be gaps between bars and for one with continuity e.g. rainfall bars should not have gaps.
6. Draw vertical lines on either side of the dot then draw horizontal line to join them with the dot.
7. Shade uniformly if they are representing only one type of data and differently if representing one type of data.
8. In combined line and bar graph temperature figures are plotted on the right hand side of $y$-axis while rainfall on the left
9. Don't start exactly at zero.
10. Include temperature and rainfall scales.
11. Start where the longest bar ends.

## Learner's Short Notes

## What a Well Drawn Graph Should Have

- Title
- Scale/scales
- Labelled and marked x and y axis starting at zero.
- Key if required e.g. in comparative bar graph.
- Accurately plotted and lines, curves or bars properly drawn.


## 1. Comparative/Group/ Multiple Line Graph

Its consists of a series of lines which are drawn with the use of the horizontal and vertical axes.
Its used where a comparison of two or more variables is necessary. Different lines are usedfor the different variables. The line should be clearly and easily distinguishable from each other. Different colours may be used or different methods of drawing lines may be employed.

The maximum number of lines that can be drawn on one graph is five.


## Learner's Short Notes

## Analysis and interpretation

Once graph has been drawn, the important information it portrays could be extracted and studied for analysis. Some of the information include;
a. The trend- this gives attention to the rise and fall of the value of the variables throughout the period. The trend can be more easily derived from the line graph than from bars or rectangles.
b. A comparison of the variables can be done by looking at the values and determining the greater and the lesser values.
c. The average can be obtained by dividing the total value by the number of years whose data has been provided.
d. The percentage increase can be calculated to show the degree of increase between any two periods.

## Advantages

i) Simple to construct
ii) Suitable when comparing trends or movements
iii) Comparison of items is easy because the graphs are drawn using common axis
iv) It's easy to read exact values from each graph

## Disadvantages

i) Number of items which can be represented are limited
ii) Crossing of lines may make interpretation and comparison difficult and confusing.
iii) Total amount of variable can't be established at a glance.

## Learner's Short Notes

## 2. Comparative Bar Graph

- This method, bars are grouped together for the purpose of comparison. Bars belonging to the same group are drawn side by side, touching each other. They give the impression of totality.
- Space is left between groups of bars. In the first group, bars are arranged in order of magnitude. The bar representing the "others" is drawn last, on the extreme right in the group. The arrangement of bars in the first group determines the position of the bars in all the others groups. But for ease of reading and interpretation, a maximum of five bars is preferred.
- There is no fixed limit to the maximum number of bars that can be grouped together.
- When the bars represent different commodities, they may be shaded or coloured differently


## Advantages

a. Easy to construct
b. Easy to read and interpret
c. Easy to compare similar components within different bars.
d. Gives a good impression of totality.
e. Individual contribution made by each component is clearly seen.
f. Differences in quantity of components are clearly seen.

## Disadvantages

a. Doesn't show trend of components over time.
b. Not easy to compare components where bars are many
c. Not suitable for many components.

## Learner's Short Notes



## 3. Divided Bars or Rectangles

This is one of the most useful methods of statistical presentation but which is employed less frequently than all others. The total quantity is presented by a rectangle. The rectangle is then subdivided to show the consistent parts. Its lack of popular appeal may be attributed to the fact that a circle is more appealing to the eye than a rectangular figure.

There are two types of divided rectangles;

- Simple divided rectangle.
- Compound divided rectangle.


## Learner's Short Notes

## 1. Simple divided rectangle.

The following steps are taken when drawing a simple divided rectangle:
a) Draw a rectangle whose length is proportional to the total value of the data.
b) Choose a suitable scale to represent the dependent variables and place it along the length of the rectangle.
c) Divide the rectangle into consistent parts by comparing the value of each component to the total value of the data and multiply by the length of the rectangle.
d) Label the components parts accordingly or use a key to identify the parts.

## 2. Compound divided rectangle.

- When the data in the question has greater details the divisions of the rectangles may themselves be subdivided. The horizontal axis which is the length of the rectangle is used to show the total quantities while the vertical axis is used to show the subdivisions. These can be scaled in percentage.
- The upper end of the rectangle can be used to label the divisions while the lower end is given a suitable scale whose maximum value is proportional to the total value of the data.


## Production of Sugarcane in 1000 tonnes of 5 major factories in Kenya

| Factory | Production (ooo <br> tonnes) | Length in cm |
| :--- | :--- | :--- |
| Sony | 50 | 0.5 |
| Nzoia | 100 | 1 |
| Chemilil | 200 | 2 |
| Muhoroni | 250 | 2.5 |
| Mumias | 400 | 4 |
| Total | 1000 | 10 |

## Learner's Short Notes

## Reported Visitor Arrivals by Continent for the Year 2000

| Continent | No. of visitors | Length of |
| :--- | :--- | :--- |
| Africa | 153904 | 1.5 |
| America | 77271 | 0.8 |
| Asia | 58784 | 0.6 |
| Europe | 663906 | 6.6 |
| Other | 82672 | 0.8 |
| Total |  | 10.3 |

## Look for a convenient scale say $\mathbf{1 c m}$ rep 100000 visitors

$\Rightarrow \quad$ Draw a divided rectangle 10 cm long to represent the data.
$\Rightarrow \quad$ Show your calculations.
It should have the following:

- Title
- Different shades
- Key
- Width of 2 cm


## Analysis and Interpretation

To get the meaning of

- Factory leading in sugar production is Mumias.
- The $2^{\text {nd }}$ leading is Muhoroni.
- Factory with the lowest production of sugar is Sony.
- Calculation of \%s.


## Learner's Short Notes

## Advantages

a) Easy to construct
b) Easy to compare components because they are arranged in ascending or descending order.
c) Takes less space than when the data is presented using graphs.
d) Each component proportion to the total can easily be seen at a glance.

## Disadvantages

a) Can't be used for a large data.
b) Only one unit of measurement can be used.
c) Difficult to assess values of individual components.
d) The visual impression isn't as good as pie charts.

## Learner's Short Notes

## STATISTICAL METHODS PAST KCSE QUESTIONS.

1. The table below shows petroleum production in thousand barrels per day for countries in the Middle East in April 2006. Use it to answer question (a)

| Country | Production in <br> ‘000" <br> barrels |
| :--- | :--- |
| Iran | 3800 |
| Kuwait | 2550 |
| Qatar | 800 |
| Saudi Arabia | 9600 |
| United Arab | 2500 |
| Emirates | 1900 |
| Iraq |  |

a)
i) What is the difference in production between the highest and the lowest producer ( 1 mk )
ii) What is the total amount of petroleum produced in April 2006 in the region? (1mk)
b) State three conditions that are necessary for the formation of petroleum (3mks)
2. The graph below shows percentage value of some export commodities from Kenya between 1999 and 2003.

Use it to answer questions (a) and (b)

(a)
i) What was the percentage value of the tea exported in the year 2000 ? ( 2 mks )
ii) What was the difference in the percentage values of the horticultural products and coffee exports in 1999 ? (2mks)
iii) Describe the trend of the value of coffee exports from 1999 to 2003 (3mks)
iv) Explain three factors which may have led to the increased export earnings from horticultural produce in Kenya between years 1999 and 2003 ( 6 mks )
v) Give three advantages of using simple line graphs to represent data. (3mks)
(b) State four reasons why Kenya's agricultural export earnings are generally low ( 4 mks )
(c) State five reasons why the common market for Eastern and southern Africa ( 5 mks )
3. (a) Define the following terms
i) Statistics
ii) Statistical data
iii) Statistical methods ( 6 mks )
(b) State two types of statistical data. (2mks)
(c) Write down two types of questionnaires. ( 2 mks )
4. (a) What factors must be considered in selecting methods of data collection. (3mks)
(b) Differentiate between discrete data and continuous data giving relevant examples. (4mks)
5. (a) What is sampling ( 1 mk )
(b) State 3 types of sampling. (3mks)
6. (a) Name two main methods used in analyzing statistical data. (2mks)
(b) What is the significance of statistics in geography? ( 5 mks )
7. (i) Name two types of graphs that you have learnt about. (2mks)
(ii) What are the advantages of using graphs named above in representing statistical data? Give advantages. (4mks)
8. (i) What is a questionnaire?
(ii) State four advantages of using questionnaires in collection of statistical data. ( 4 mks )
(iii) Explain oral interview method. (2mks)
9. Explain the following methods of data recording.

- Tabulation
- Photographing
- Tape recording
- Tallying

10. What is data? (2mks)
11. Marks $72,60,65,70,65,80,65,70,80,84,63,75,63,71,74$

Use the data above to find out mean and mode. ( 4 mks )
12. With the help of data above explain how median is obtained. (3mks)

## MARKING SCHEME

## STATISTICAL METHODS

1. 

(a)
i) $9600-800=8,800,000$ Barrels
ii) 21,150,000 Barrels.
iii) $21,150,000 \backslash 30=705,000$ Barrels
2. Graph
i) $29.3 \%(29-29.5 \%) 29 \%$
ii) $4.75 \% / 4.8 \% / 4.9 \%$
(4.75\%-4.9\%)
(iii) Describe the trend of the value of coffee exports from years 1999 to 2003.
a. The value was generally declining over the five year period.
b. The value was highest in 1999.
c. The decline between 1999 and 2000 was minimal/gradual.
d. The highest drop was between 2000 and 2001
e. There was a minimal drop between 2002 and 2003.
f. The decline between 2001 and 2002 was minimal/gradual.

The value was lowest in 2003.
(iv) Explain three factors which may have led to the increased export earnings from horticultural produce in Kenya between years 1999 and 2003.
a. Improved technology which leads to advanced crop husbandry/increase the volume of fresh horticultural products.
$\beta$. Aggressive promotion of trade abroad leading to a wide/ready market in foreign countries.
c. Improved ways of packaging have made the produce more competitive/ attractive,
d. Improved infrastructure/air/road transport have helped in the quick means of transportation of fresh produce to the market.
e. The declining benefits from traditional agricultural exports leading to the expansion of the areas under horticultural crops.
f. The government has encouraged the formation of organizations that are assisting horticultural farmers.
v) Give three advantages of using simple line graphs to represent data.
a. Give clear visual impression.
b. Easy to construct.
c. Easy to interpret.
d. Can be used to represent a wide variety of variables.
e. Appropriate for comparison
(b) Reasons why Kenya's agricultural export earning generally are low
i) Kenya sells most of her agricultural products in their raw form and they are priced lowly.
ii) International prices keep fluctuating from year to year.
iii) Prices of some commodities are externally determined.
iv) There is competition from other producing countries/from other similar products.
v) Some products are inferior in quality.
vi) There are fixed quarters for some agricultural products.
vii) Decline in quantities of some agricultural exports.
3. (a)
i) Statistics It refers to the art or science that is concerned with the interpretation of numeric information.
ii) Statistical data Refers to the information collected and arranged in a systematic manner.
iii) Statistical methods Refers to the techniques used in collecting, recording, analyzing and presenting data.
(b) Primary data and Secondary data
(c) Closed-ended (rigid) - Open-ended
4. (a)
i) The method should be inexpensive.
ii) Should be time saving
iii) Should give accurate data
iv) Most applicable method
(b) Discreet data refers to the non-continuous data over time given in whole numbers only e.g.
i) Total population in a nation.
ii) Monthly rainfall totals.
iii) No. Of livestock per district
iv) Continuous data can be given in any value including decimals e.g. 1.8 km .
5. Sampling refers to the process by which a representative portion of the whole phenomena under study is analyzed and generalized/ generalization is made.

## Types of sampling

i) Systematic sampling
ii) Stratified sampling
iii) Random sampling
6. (a)
i) Calculation of percentages
ii) Measuring of Central tendency (mean, median and mode)
iii) Frequency distribution
(b)
i. Predicting for future trends.
ii. Showing changes through time
iii. Establishing Geographical relationships
iv. For economic planning
v. For explaining geographical phenomena.
vi. Useful for making comparisons.
7. (i)
a. Simple line graph
b. A combined line and bar graph
c. Simple bar graph
(ii)
a) The simple bar graph
b) Prominent values stick out well
c) Bars are appealing to the eye
d) Easy to draw, read and interpret the data represented
8.
(i) A set of pre-questions which are related to the topic of study.
(ii)
a. Its a source of first hand information
b. The researcher can ask for clarification from the respondent.
c. Similar questions are used for all respondents and comparison can easily be made.
d. When posted, rigid questionnaires reduce fieldwork expenses.
(iii) Interview involves collection of information by asking questions directly and recording the answers given. In this method the researcher established contact with the respondent and agree on time for face to face interview with the respondent. Interviews can also be carried out on telephone with the interviewee.
9. Methods of data recording
i) Tabulation This is recording of data by arranging facts of figures in form of table or list.
ii) Photographing This is done by use of a camera to record geographical information.
iii) Tape recording This can be done when one is collecting data through an oral interview where one uses tape recording device to record conversation.
iv) Tallying Used when the data is collected through counting. One counts and puts a vertical strike for every item counted, on the fifth count one puts a diagonal crossing the four strokes.
10. Data refers to facts and figures collected from the field.
11. The mean is
$72+60+65+70+65+80+65+70+80+84+$ $63+75+63+71+74$
$=1057 / 12=70.47$
12. This is got by arranging the data in an ascending order as follows: $60,63,63,65,65,65,70$, $70,71,72,74,75,80,80,8.4$. The middle number is the median: 70

