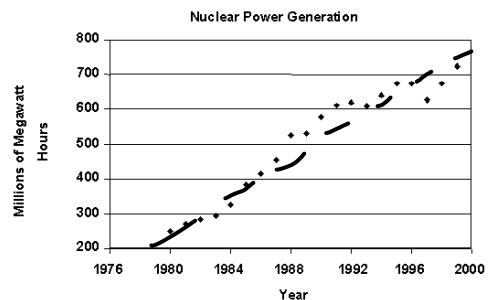


**Unit #4 Test – Critical Analysis of Data**K/U: \_\_\_\_\_  
13Application: \_\_\_\_\_  
12Communication: \_\_\_\_\_  
7TOTAL: \_\_\_\_\_  
32**Knowledge/Understanding: 13 marks****Multiple Choice:** Circle the appropriate answer

1. **Correlation can be described as**
  - a) The number of data points in a survey or study
  - b) The relationship or visible trend within a set of data**
  - c) The sum of the residuals
  - d) None of the above
  
2. **If two variables have no correlation:**
  - a) the points on a scatter plot appear to be random**
  - ~~b)~~ a median-median line must be used to predict trends
  - ~~c)~~ the line of best fit has a negative slope
  - d) the line of best fit has a slope of zero
  
3. **Why does a strong linear correlation not always represent a cause-and-effect relationship?**
  - a) There could be many extraneous variables that haven't been considered
  - b) There could be a poor sample size or sample method chosen
  - c) The relationship could be accidental
  - d) All of the above**
  
4. **To judge the conclusions of a study properly, you need to consider:**
  - a) Information about the analysis of the data
  - b) Any extraneous variables
  - c) **Information about sampling methods**
  - d) All of the above**
  
5. **The horizontal axis of a scatter plot usually represents the:**
  - a) the extraneous variable
  - b) frequency
  - c) the dependent variable
  - d) the independent variable**
  
6. **A control group in an experiment or study is**
  - a) The group where the independent variable is changed
  - b) The same as the experimental group, only smaller
  - c) The group that is in control of the study (including researchers etc.)
  - d) The group where the independent variable is constant compared to the experimental group**

7. What type of correlation exists in the following scatter plot?

- a) weak negative
- b) strong positive**
- c) weak positive
- d) strong negative



8. Two variables have a correlation coefficient of  $r = -0.4$ . This indicates

- a) a moderate negative correlation**
- b) a strong negative correlation
- c) a moderate positive correlation
- d) a strong positive correlation

9. The median-median line:

- a) passes through the median x-value and median y-value of all the data
- b) is created using key points based on medians**
- c) is used when there is no correlation between the variables
- d) is used when you only have 3 data points

10. Given that  $r = 0.73$ , what percent of the variation in  $y$  is due to the variation in  $x$ ?

- a) 0.5329**
- b) 53.29
- c) 0.73
- d) 73

$$r^2 = 0.5329$$

11. A relationship in which the independent variable and the dependent variable are reversed in the process of establishing causality is called:

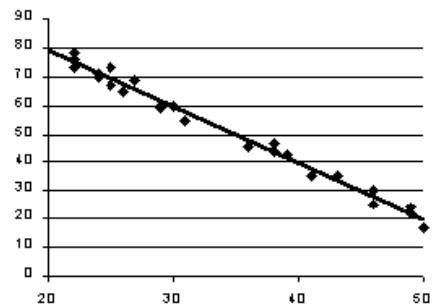
- a) An accidental relationship
- b) A reverse cause-and-effect relationship**
- c) A presumed relationship
- d) A cause-and-effect relationship

12. A residual value is

- a) The same as the coefficient of determination
- b) Always positive
- c) Always negative
- d) The vertical distance between a data point and the line of best fit**

13. The regression line shown would have a correlation coefficient closest to

- a) -0.95**
- b) 0.5
- c) +0.95
- d) 0

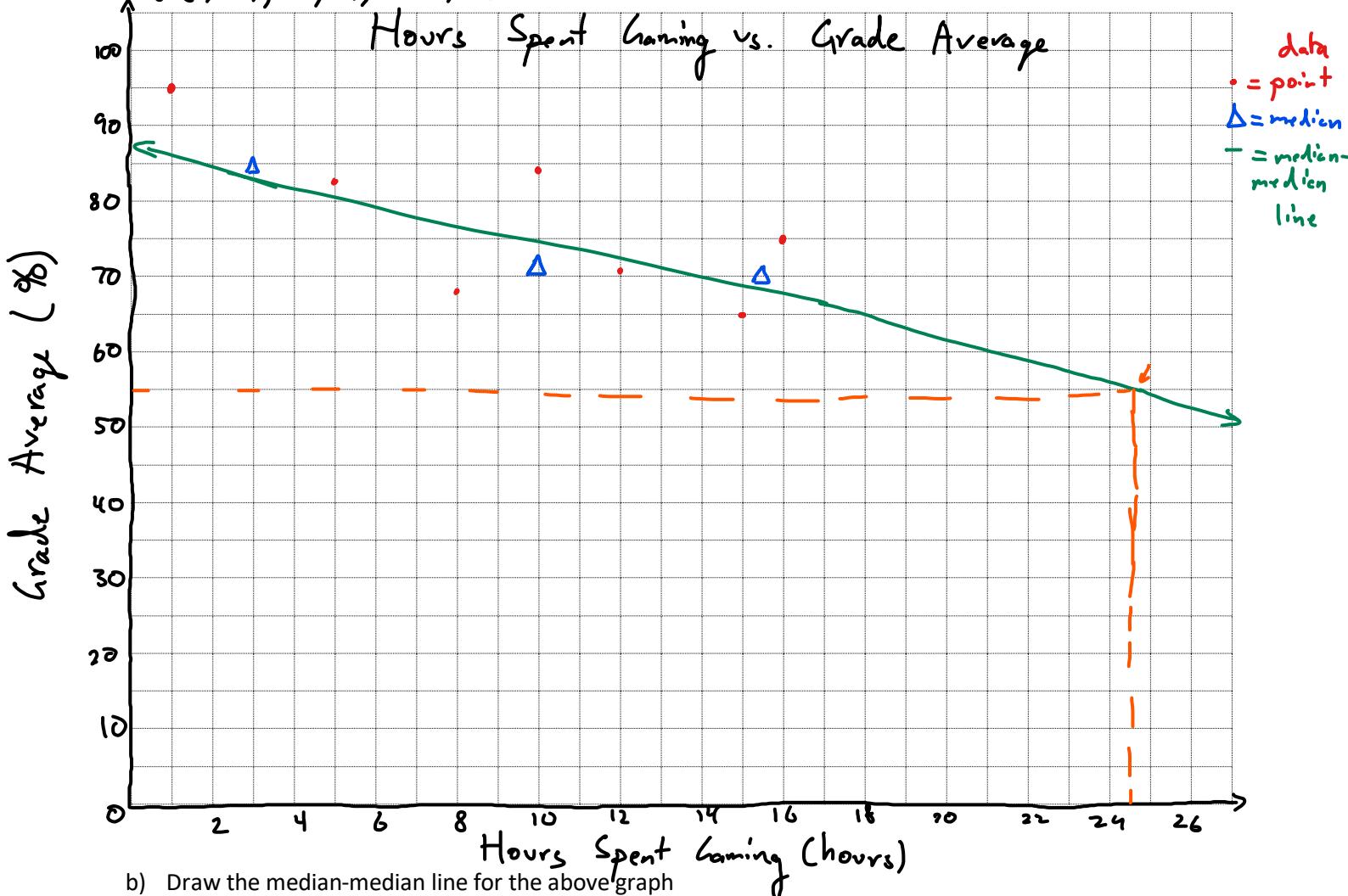


## Application: 12 marks

1. Students in a high school geography class recorded the number of hours spent playing video games and their average in the table below: <8 marks>

| Hours Spent Gaming               | Grade Average (%)              |
|----------------------------------|--------------------------------|
| $med_1 = \frac{1+5}{2} = 3$      | 95                             |
| 5                                | 83                             |
| 8                                | 68                             |
| 10                               | 84                             |
| 12                               | 71                             |
| $med_2 = \frac{15+16}{2} = 15.5$ | 71                             |
| 15                               | 65                             |
| 16                               | 75                             |
| $med_3 = (3, 15.5, 16)$          | $med_3 = \frac{95+83}{2} = 89$ |
| $med_3 = (3, 15.5, 16)$          | $med_3 = \frac{71+75}{2} = 73$ |
| $med_3 = (3, 15.5, 16)$          | $med_3 = (3, 15.5, 16)$        |

- a) Identify the independent variable: Hours Spent Gaming
- b) Identify the dependent variable: Grade Average
- c) Create a scatter plot for this data. Using an app for the scatter plot and median-median line will result in a mark of zero for this question. This work must be done by hand (even if it's messy)



- b) Draw the median-median line for the above graph  
*Line drawn in green*
- c) How many hours would you expect to spend gaming for a grade average of 55%? \_\_\_\_\_ 24.5 h  
*{use the graph to answer this question – there is no need to develop the equation of the line}*

From the graph, a student with a grade average of 55% is expected to spend about 24.5 hours gaming.

2. A high school is testing a new method for teaching math using two classes of the same grade. The period 4 class is taught with an established method and the period 1 class is taught with a new method. When both classes take the same test, the class during period 1 scores somewhat higher.

&lt;2 marks&gt;

Identify:

a) the experimental group \_\_\_\_\_ Period 1 classb) the control group \_\_\_\_\_ Period 4 classc) 2 extraneous variables Time of day when taking the test (Period 1 is in the groggy morning)  
Students' interest, ability and will to learn math

3. Classify the most likely type of causal relationship in the following situations using the following causal relationships (use each type only once):

&lt;2 marks&gt;

*Accidental Relationship  
Reverse Cause and Effect  
Cause and Effect  
Common-cause Factor  
Presumed Relationship*

- a) The population of rabbits increases with the consumer price index.

Accidental relationship

- b) There is a strong positive correlation between a student's score on a physics exam, and their score on a calculus exam.

Common-cause factor

- c) There is a strong correlation found between the number of scholarships received and the number of job offers upon graduation.

Presumed relationship

- d) An increase in coffee consumption results in an increase in insomnia.

Cause and Effect

## Communication: 7 marks

1. What is the purpose of using a control group and an experimental group in a research study? Give an example of a situation where they would be used. *<3 marks>*

**Suppose:** An experiment to test the efficacy of smoking medication A.

**Experimental group:** A random sample of smokers who are given medication A and its treatment plan.

With just an experimental group, it is impossible to determine the efficacy of just the medication A. There may have been external variables that affect each participant's smoking habits.

**Control:** A random sample of smokers who are NOT given medication A.

Studying the difference in the final proportions of both groups helps you isolate the effects of just the medication A. E.g. if only 10 members from both groups get better, ✓ 2. An educational researcher discovers that levels of mathematics anxiety are negatively correlated with attendance in mathematics class. The researcher theorizes that poor attendance causes mathematics anxiety. Is that a positive or negative correlation? Suggest an alternate interpretation of the evidence. *A had no effect.*

*Also reduces the effects of extraneous variables.*

The researcher's observations of levels of math anxiety and attendance imply a negative correlation between the two variables - i.e. high math anxiety is correlated with low attendance while low math anxiety is correlated with high attendance.

There may, however, be a reverse cause and effect relationship between levels of math anxiety and attendance. Students who are very anxious about their math abilities may not feel comfortable or even welcome in the class and would then stop going to math class because of that anxiety. Here, the independent variable and dependent variables have switched from the researcher's study. The independent variable is level of math anxiety and the dependent variable is attendance.