Name:_____Statistics Summer Packet 2018-2019

GNBVT
12th Grade
Summer Statistics Packet

For Extra Help, please visit:
www.khanacademy.com
or
www.stattrek.com

In order to receive CREDIT article & summaries must be attached.

Part 1: Introduction to Statistics

Sta·tis·tics

Etymology: German *Statistik*: study of political facts and figures, from New Latin *statisticus*: of politics, from Latin *status*: state. Date: 1770

1 : a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data *[note: this is for Statistics with a uppercaseS]*

2 : a collection of quantitative data *[note: this is for statistics with a lowercases]*

Source: http://www.merriam-webster.com/dictionary/statistics

Answer the following in complete, well written sentences, to the best of your ability:

1) Before you saw this definition, how would you have defined Statistics? Has your definition changed after reading this?

_____
_____
_____
_____
_____
_____

2) How one collects the data is extremely important. Explain how you would conduct a survey to determine the percentage of GNB Voc-Tech students who are satisfied with the quality of education that they are receiving. Due to resource constraints, however, you will only be able to ask 100 students.

_____
_____
_____
_____
_____
_____

3) You have worked with data before in your science classes, if nowhere else. Provide one example from your life in which you have worked with data. How did you collect it? How did you analyze it? How did you present your findings? What conclusions did you come to?

_____
_____
_____
_____
_____
_____

4) Tell me what you have heard from other people about this class.

_____
_____
_____
_____
_____
_____

5) This class is an elective. So, why did you sign up for it?

1) Find a newspaper or magazine article involving statistics. Attach your article to the back of this packet. Answer the following about your article:

a) *Who* is being studied?

b) *What* about those individuals is being recorded / analyzed (i.e. what are the variables?)? Do you think the variables are categorical or quantitative in nature?

c) *When* was the data collected?

d) *Where* was the data collected (more accurately: what geographical area is associated with the data)?

e) *Why* do you think this data was collected and analyzed?

f) *How* was the data collected and analyzed? In other words, what methods were used?

1.) Write a one-paragraph summary of your article.

Read the following to learn more about early uses of statistics…

Censuses throughout history

Early population counts generally were not concerned with determining the total size of the population or including detailed information about people. Their main goal was to discover who was available for military duty and who held taxable property. These counts usually did not give an accurate number or picture of the population. They often left out large segments of society, such as women and children, men attempting to avoid military service or taxation, and native inhabitants of an area.

The earliest known population counts were made thousands of years ago by the ancient Babylonians, Chinese, and Egyptians. Around 2500 B.C., the Babylonians recorded on clay tablets information about the taxpaying part of the population. These tablets included such data as the number of farm animals, farm products, and households for districts within the kingdom. Tax returns from around 2300 B.C. for parts of ancient China indicate some kind of population count. About 1300 B.C., Egypt was divided into administrative districts. The government registered and counted heads of households and members of the households within these districts.

The fourth book of Bible, the Book of Numbers, describes the census, or numbering, of the tribes in ancient Israel to determine the number of men of fighting age (Numbers 1: 1-46; Numbers 26: 1-51). In 594 B.C., the Greek lawmaker Solon introduced a form of enumeration and registration to reform tax laws in Greece.

The Romans employed census takers known as censors to determine the number of people who were eligible for taxation and military duty. The Roman censor was responsible for officially registering all citizens in a particular area, evaluating their property, collecting revenue, and guarding public morals. Perhaps the best- known Roman census is described in the New Testament story of the birth of Jesus Christ (Luke 2:1-7). This census took place about 5 B.C., when Joseph and Mary traveled to Bethlehem to record their names in a census ordered by the Roman emperor Augustus.

The practice of taking censuses declined in Europe after the fall of the West Roman Empire in A.D. 476. One of the few attempts to count people during the Middle Ages occurred in England in 1086. That year, commissioners sent by William the Conqueror traveled the kingdom and recorded, for tax purposes, the names of all English landowners and the value of their lands and houses, tenants, and servants. The resulting document, known as the Domesday Book, provides historians with a censuslike description of England at that time.

Through the years, with the rise in trade, the growth of towns, and the development of nations, rulers and government officials increasingly recognized the importance of counting people and goods. In 1665, King Louis XIV of France ordered a census in New France, in what is now Quebec, Canada. This census recorded the name of each person, along with such information as age, marital status, occupation, and relationship to the head of the household. The main purpose of this census was to collect information about the colony's progress, rather than to assess how much military service or tax revenue the colonists might provide. Because of this purpose, census historians generally consider the New France enumeration to be the model for modern censuses.
Likewise, in 1703, there was a house-to-house census in Iceland for reasons other than taxation and military service. This census inquired into the effects of economic conditions and natural disasters. The government then used the information to develop programs for economic and social improvement.

A number of European countries undertook censuses of individual cities and provinces in the early 1700's. However, none of these enumerations counted the total population of a nation until 1749. That year, the Swedish government conducted the first national census.

The first modern census —one that was complete, direct, and scheduled to be repeated at regular intervals— was the United States census of 1790. In the 1800's, a number of other countries began taking regular censuses. In 1853, an International Statistical Congress was held in Brussels, Belgium. This conference represented the first attempt to adopt international recommendations and requirements to help in comparing population census data among various countries.

After World War II ended in 1945, censuses became especially important as an aid in planning for the economic reconstruction of countries that had been heavily damaged in the war. In 1946, the United Nations established a separate Population and Statistical Commission, which recognized the need for census statistics. Since then, the United Nations has published a number of principles and recommendations for population and housing censuses to assist countries in the planning of censuses. Following these recommended standards allows for international comparison of collected data. In addition, the United Nations Fund for Population Activities provides many countries with financial and expert assistance for the planning of censuses.

Today, most censuses are proclaimed by a government decree or law and planned and executed by a statistical agency, a permanent or semipermanent census bureau, or both. These census acts or laws require every person to answer the questions to the best of his or her knowledge. Refusal to cooperate can result in a fine or even imprisonment.

Draaijer, Gera. "Census." *World Book Advanced*. World Book, 2011. Web. 5 June 2011.

6) After reading this, provide a conjecture for why the word "Statistics" is rooted in the Latin for "state".

What is Statistics?

by Jordan Neus (from http://www.fiu.edu/~neusj/whatisstatistics.html)

Statistics is becoming increasingly more important in modern society with passing time. We are constantly being bombarded with charts, graphs, and statistics of various types in an attempt to provide us with succinct information to make decisions. Sometimes this information is presented in a manner so as to sway us toward a particular view. As consumers and decision makers we must be aware of this. Which drug should we take? Which car should we buy? Where will the economy go? Who is infected with a particular deadly disease? These are all examples of questions which are usually relegated to the statistician for analysis and dissemination. This lecture will attempt to introduce the beginning to student some of the reasoning behind the necessity of statistical inference.

In order to realistically understand the subject of Statistics it is important to appreciate the rationale behind why and how Statistics is used by the world, at large. That is, why do we need Statistics anyway? This, perhaps, is a bit philosophical, yet I can not over emphasize the need for thinking along these lines. Without proper perspective, Statistics becomes a mere mathematical exercise, diverging from the true nature of the subject.

In order to begin our analysis as to why Statistics is a necessary type of reasoning we must begin by addressing the nature of science and experimentation. A characteristic method used by scientists is to study a relatively small collection of objects, say 2500 people, and a characteristic, say longevity, and through experimentation or observation, draw a conclusion appropriate for the entire class of objects (i.e. people, in general). For example, suppose a study published results suggesting *people who own pets live longer*. Would this mean that all people who own pets are likely to live long lives? Does owning a pet *cause* longevity? Suppose the people in the study, by chance, were on the whole, very healthy people, and therefore lived long lives: Would this invalidate the researcher's assertion that people who own pets live longer? The obvious problem with this type of reasoning is that these issues can never be proved absolutely. This type of scientific reasoning is called **inductive reasoning** and is inherently flawed. One can never study a sample and expect conclusions to hold true for the entire population with absolute certainty. This is exactly why Statistics is needed.

In contrast to the lack of certainty associated with inductive reasoning, the type of logic used in Mathematics is absolutely certain. The mathematician begins with general principles and logically concludes more specific relationships. This type of reasoning from the general to the particular is called **deductive reasoning**. A rather simplistic (but nevertheless correct) example is based on the principle that two numbers can be added in any order, thereby giving the same sum. This is called the axiom of commutativity. An example of deductive reasoning would be to assert that since this holds for any two numbers, surely this must hold for the numbers two and three, in particular. We are, therefore, absolutely certain that $2 + 3 = 3 + 2$, given the axiom of commutativity.

In its applied form, Statistics then becomes a bridge between the inductive uncertainty of science and the deductive certainty of Mathematics. In his classic book, *The Design of Experiments*, Sir Ronald A. Fisher expresses this idea beautifully:

We may at once admit that any inference from the particular to the general must be attended with some degree of uncertainty, but this is not the same as to admit that such inference cannot be absolutely rigorous, for the nature and degree of the uncertainty may itself be capable of rigorous expression.

Statistics, therefore, is the mathematical method by which the uncertainty inherent in the scientific method is rigorously quantified.

7) React to the above pieces in at least three paragraphs (attach to back of packet) :

Part 2: Data and Its Context + Reading Comprehension Involving Statistics

Read the following...

"Teen Automobile Crash Rates Are Higher When School Starts Earlier"

ScienceDaily (June 10, 2010) — Earlier school start times are associated with increased teenage car crash rates, according to a research abstract presented June 9, 2010, in San Antonio, Texas, at SLEEP 2010, the 24th annual meeting of the Associated Professional Sleep Societies LLC.

Results indicate that in 2008 the teen crash rate was about 41 percent higher in Virginia Beach, Va., where high school classes began at 7:20 a.m., than in adjacent Chesapeake, Va., where classes started more than an hour later at 8:40 a.m. There were 65.4 automobile crashes for every 1,000 teen drivers in Virginia Beach, and 46.2 crashes for every 1,000 teen drivers in Chesapeake.

"We were concerned that Virginia Beach teens might be sleep restricted due to their early rise times and that this could eventuate in an increased crash rate," said lead author Robert Vorona, MD, associate professor of internal medicine at Eastern Virginia Medical School in Norfolk, Va. "The study supported our hypothesis, but it is important to note that this is an association study and does not prove cause and effect."

The study involved data provided by the Virginia Department of Motor Vehicles. In Virginia Beach there were 12,916 drivers between 16 and 18 years of age in 2008, and these teen drivers were involved in 850 crashes. In Chesapeake there were 8,459 teen drivers and 394 automobile accidents. The researchers report that the two adjoining cities have similar demographics, including racial composition and per-capita income.

2) Answer the following questions regarding the above excerpt:

a) *Who* is being studied?

_____
_____
_____
_____
_____
_____

b) *What* about those individuals is being recorded / analyzed (i.e. what are the variables?)? Do you think the variables are categorical or quantitative in nature?

_____
_____
_____
_____
_____
_____

c) *When* was the data collected?

_____
_____
_____
_____
_____
_____

d) *Where* was the data collected (more accurately: what geographical area is associated with the data)?

e) *Why* do you think this data was collected and analyzed?

f) *How* was the data collected and analyzed? In other words, what methods were used?

g) Why do you think the authors of the study mentioned that "it is important to note that this is an association study and does not prove cause and effect?"

Part 3: Displaying and Describing Categorical Data

Complete the following survey. For each question, decide whether the question will collect categorical or quantitative data.

	Response	Quantitative or Categorical?
1) What math course did you take last year?	_____	_____
2) What grade did you earn in that course?	_____	_____
3) Choose a random integer from 1 to 25	_____	_____
4) How many siblings do you have?	_____	_____
5) How many cousins do you have?	_____	_____
6) Where did you eat your last meal? (1 = Home, 2 = Restaurant, 3 = Other)	_____	_____
7) How many miles from school do you live?	_____	_____
8) How long have you lived in your home?	_____	_____
9) What sports/activities are you involved in?	_____	_____
10) Do you have a job? (1 = Yes, 2 = No)	_____	_____
11) What is your favorite candy?	_____	_____
12) When is your birthday?	_____	_____

Write three questions which have answers that are categorical in nature and respond to each one.

_____
_____
_____
_____
_____
_____

Write three questions which have answers that are quantitative in nature and respond to each one.

_____
_____
_____
_____
_____
_____

Pick a simple question with simple responses that you would like to ask (e.g. Do you prefer iPhone or Android?; Which is better: Instagram or Snapchat?; Pancakes or Waffles?, etc.)

Ask 30 random people the question, and record their response as well as their gender (try to get a roughly equivalent number of boys and girls):

#	Response to Question	Gender
1		M F
2		M F
3		M F
4		M F
5		M F
6		M F
7		M F
8		M F
9		M F
10		M F
11		M F
12		M F
13		M F
14		M F
15		M F
16		M F
17		M F
18		M F
19		M F
20		M F
21		M F
22		M F
23		M F
24		M F
25		M F
26		M F
27		M F
28		M F
29		M F
30		M F

Summarize your results in a table:

Summarize your findings in one or more graphs:

Does one's gender appear to be independent of how one responds to this question? Explain, and use your results to support your argument.