

Why is this important to statisticians? Using current computing technologies, data sets that are feasible to process simple statistical models in a reasonable amount of time are on the order of terabytes.

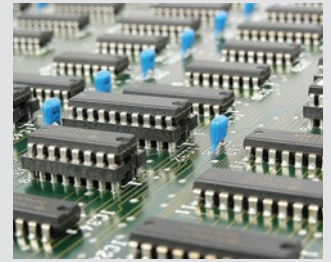
2.1 Exercises

Basic Concepts

1. What is an exabyte?
2. What is 1000 terabytes called?
3. What are the two ways to store data?
4. What is a sensor?
5. List 3 examples of remote sensing technology.
6. Name two things that you use every day that contain sensors.
7. What is Big Data?
8. What are the four attributes of Big Data?
9. List two sources of Big Data in science.
10. List two sources of Big Data in business and industry.
11. On the Apollo mission, what device worked to ensure the safe and precise landing of the lunar module on the moon's surface? Generally speaking, what data sources was this device connected to?

Exercises

12. Compare the size of a zettabyte to a terabyte. Which unit is larger? How much larger?
13. Compare the size of an exabyte to a yottabyte. Which unit is larger? How much larger?
14. Suppose a company collects 50 petabytes of data each year from its customers. If the total available storage capacity of the company is currently two exabytes, how many years of customer data can it store before running out of space?
15. Use Table 2.1.2 and Figure 2.1.1 to determine the year when the world's data storage capacity grew most rapidly between 2010 and 2020.



“The Most Important Master’s Thesis of the 20th Century”

In 1948, Claude Shannon wrote a paper entitled “A Mathematical Theory of Communication” which was the foundational work for a field now called information theory. The paper introduced the term “bit” and demonstrated that a series of bits—1s and 0s, of which eight make a byte—could be used to represent all information. The bit/byte would become the standard unit for data storage and network communication of the future. Shannon’s foundational work in information theory was not his only contribution. His master’s thesis has been called the most important master’s thesis of the 20th century. It showed that electrical switches could be configured to perform Boolean logic functions (i.e., digital logic). Shannon’s work became the foundation of digital circuit design. Digital circuits are the fundamental component of all digital computers and without them we would not have modern computers, nor modern statistics.