Chapter 1

Data Everywhere

Data Everywhere

All slides in this presentations are based on the book Functions, Data and Models, S.P. Gordon and F. S Gordon
ISBN 978-0-88385-767-0

- We are surrounded in everyday life by different data. Here are some important questions regarding data.
  1. How to make decisions based on data?
  2. How to interpret data?
  3. How to use data to answer various questions?

- One variable data; two-variable data; data on more than 2 variables

Two-Variable Data

Independent/Dependent Variables Scale

Two Variable Examples

In many situations we have data that involve 2 different quantities/variables

a) Education level and salary of employees

b) Height and weight of people

c) Size of a land/habitat and number of distinct species who live there

Example of Two Variable Data

Ex.

Number of reptiles and amphibians living on different islands in the Caribbean Sea and the area of these islands in sq. miles

- Two variables – lets give them names

  N - number of species

  A - area

- One of them depends on the other… What do you suggest?

Island	Area (sq. miles)	Number of species
Redonda	1	3
Saba	4	5
Montserrat	40	9
Puerto Rico	3459	40
Jamaica	4411	39
Hispaniola	29,418	84
Cuba	44,218	76

Independent/Dependent Variables

- We have two quantities/*variables;* one depends on the other
- One is the *independent variable*, the other is the *dependent variable*
- *Which is the independent variable? How about the dependent variable? Why?*

Island	Area (sq. miles)	Number of species
Redonda	1	3
Saba	4	5
Montserrat	40	9
Puerto Rico	3459	40
Jamaica	4411	39
Hispaniola	29,418	84
Cuba	44,218	76

Independent/Dependent Variables

- *N* (number of species) is the *dependent variable* and *A* (area of the island) is the *independent variable*
- The independent variable is graphed *horizontally* and the dependent variable - *vertically*

Island	Area (sq. miles)	Number of species
Redonda	1	3
Saba	4	5
Montserrat	40	9
Puerto Rico	3459	40
Jamaica	4411	39
Hispaniola	29,418	84
Cuba	44,218	76

Graphing the data - Scale

The area A extends ***horizontally*** from 1 to 44,218 sq. miles

1. From 0 to 45,000 sq. miles with tick marks every 15,000 sq. miles
2. From 0 to 45,000 sq. miles with tick marks every 9,000 sq. miles
3. From 0 to 50,000 sq. miles with tick marks every every 5,000 sq. miles

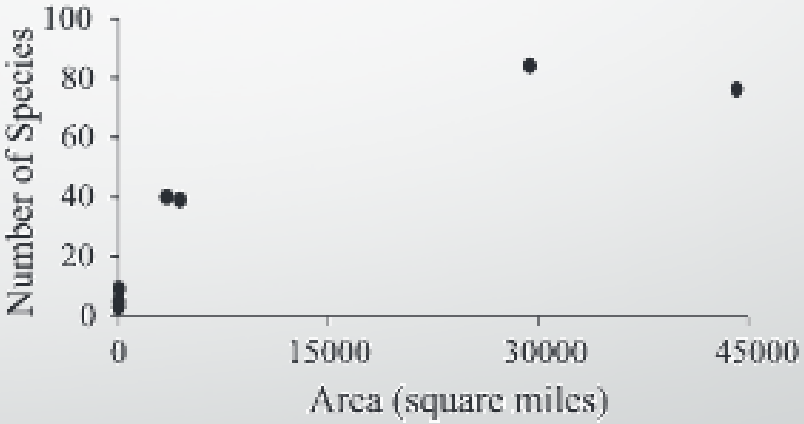
The number of species N extends ***vertically*** from 3 to 84

1. From 0 to 90 with tick marks every 10 species
2. From 0 to 100 with tick marks every 20 species
3. From 0 to 100 with tick marks every 10 species

Scatter Plots

Def - a graph in which the values of two variables are plotted along two axes, the pattern of the resulting points revealing any correlation present.

- Island “Saba” is represented by the point (A, N) = (4, 5)
- Jamaica – by (A, N) = (4 411, 39)
- Cuba – by (A, N) = (44 218, 76)

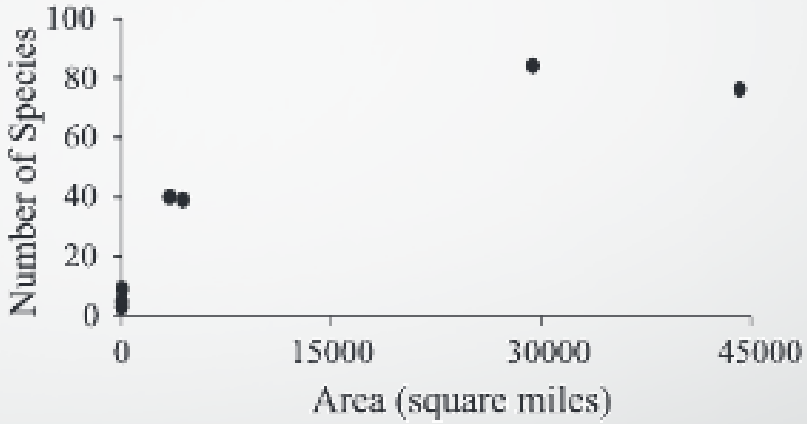


Scatter Plots - Example Continued

- As the Area of the island increases, the Number of species who live on the island also increases
- N grows more and more slowly as A increases
- Want to answer predictive questions, e.g.

  Area of Barbados is 166 sq. miles; estimate the number of species that live there

  Estimate the size of a habitat that will support 25 species of reptiles and amphibians



US population between 1780 and 1900

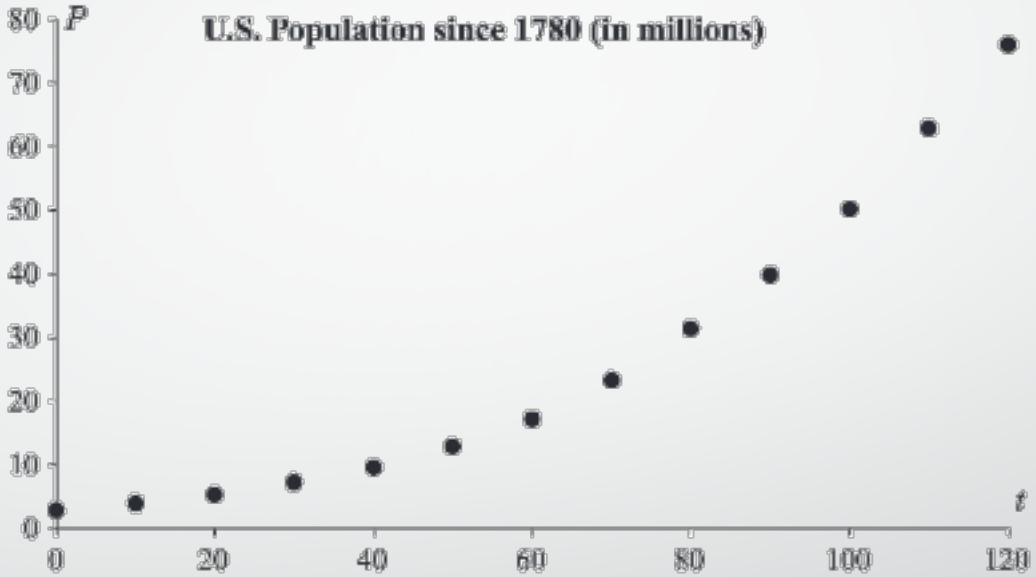
- Which is the independent/dependent variable?
- What scale of the scatter is good to be used?
- State what each variable you use stands for in reality
- Draw the scatterplot
- What predictive questions can be answered later once we get a formula?

Year	Population
1780	2,800,000
1790	3,900,000
1800	5,300,000
1810	7,200,000
1820	9,600,000
1830	12,900,000
1840	17,100,000
1850	23,200,000
1860	31,400,000
1870	39,800,000
1880	50,200,000
1890	62,900,000
1900	76,000,000

Years, Since 1780	Population in millions
0	2.8
10	3.9
20	5.3
30	7.2
40	9.6
50	12.9
60	17.1
70	23.2
80	31.4
90	39.8
100	50.2
110	62.9
120	76.0

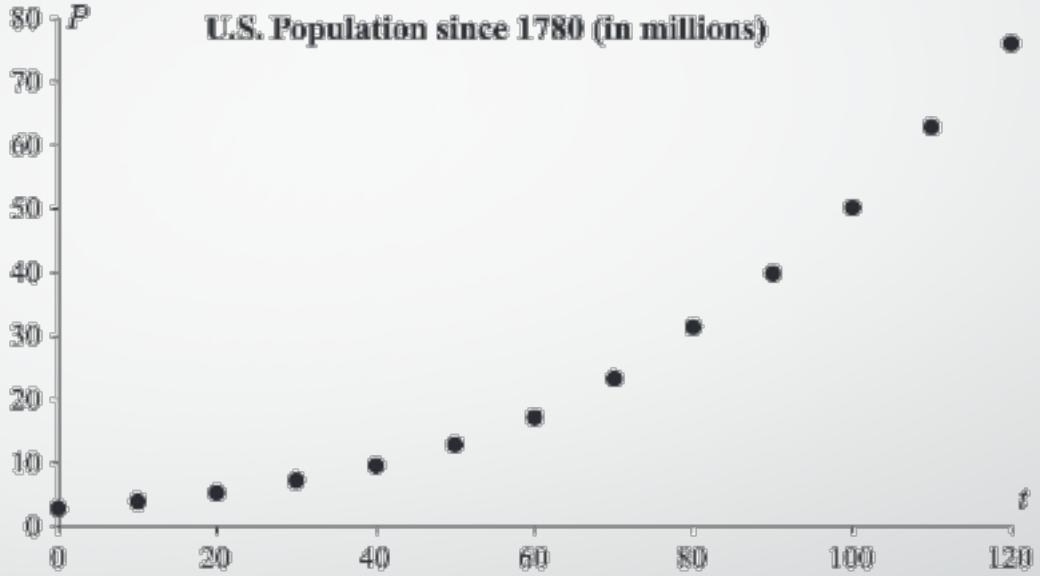
Scatterplot of US population since 1780

Years, Since 1780	Population in millions
0	2.8
10	3.9
20	5.3
30	7.2
40	9.6
50	12.9
60	17.1
70	23.2
80	31.4
90	39.8
100	50.2
110	62.9
120	76.0



Predictive Questions

- What would the population of the U.S. have been in 1920?
- What was the population in 1875?
- When would the U.S. population have reached 100 million?



Independent/Dependent Variable Examples

- In a study to determine whether how long a student sleeps affects test scores; what is the independent variable and what is the dependent variable?

- The independent variable is the length of time spent sleeping while the dependent variable is the test score.

Independent/Dependent Variable Examples

- If you want to know whether caffeine affects your appetite, what is the independent variable and what is the dependent variable?

- the presence/absence of the amount of caffeine would be the independent variable. How hungry you are would be the dependent variable.

Independent/Dependent Variable Examples

- You want to compare different brands of paper towels, to see which holds the most liquid. What is the independent variable and what is the dependent variable?

- The independent variable in your experiment would be the brand of paper towel. The dependent variable would be the amount of liquid absorbed by a paper towel.

Example

Temperature readings were recorded every two hours from midnight to 2:00 P.M. in Dallas on June 5, 2017. The time was measured in hours from midnight.

time	0	2	4	6	8	10	12	14
Temperature	73	73	70	69	72	81	88	91

a. Decide which is the independent variable and which is the dependent variable.

b. Decide on the appropriate scales for the two variables for a scatterplot.

c. What letter will you use for each variable?

d. Draw the scatterplot.

Solution

Temperature readings were recorded every two hours from midnight to 2:00 P.M. in Dallas on June 5, 2017. The time was measured in hours from midnight.

time	0	2	4	6	8	10	12	14
Temperature	73	73	70	69	72	81	88	91

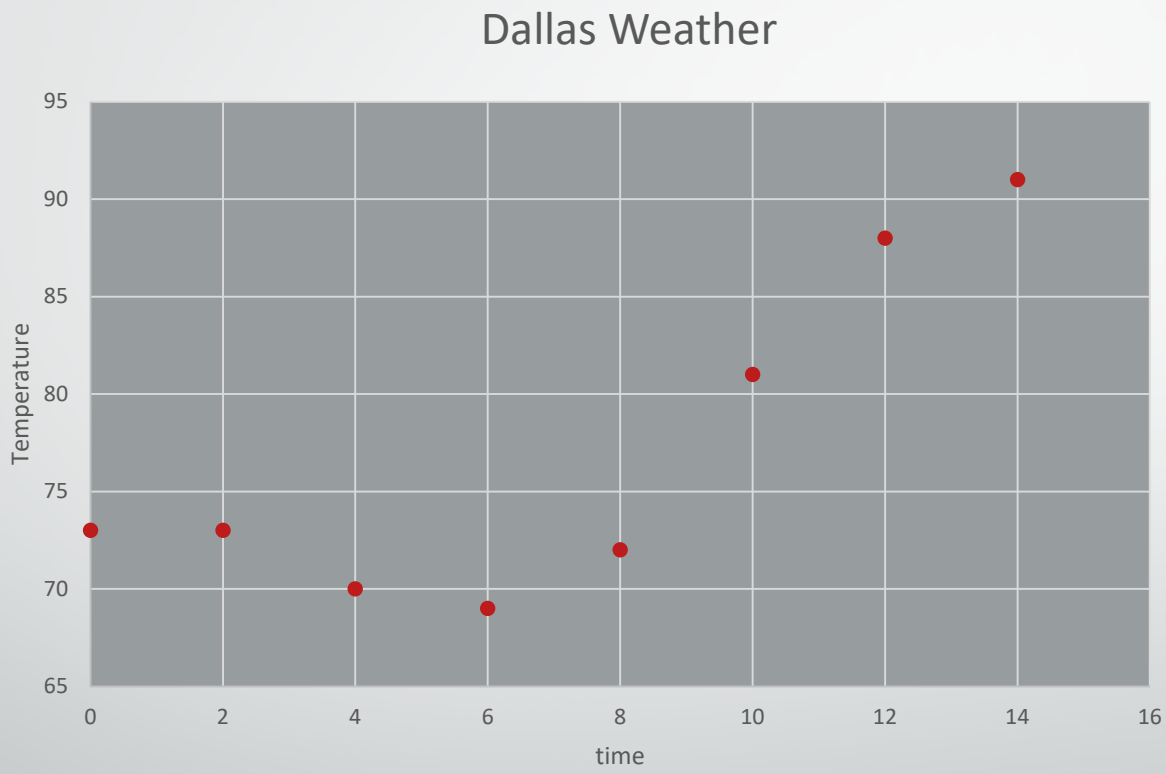
a. Decide which is the independent variable and which is the dependent variable. The Independent variable is time and the dependent variable is Temperature

b. Decide on the appropriate scales for the two variables for a scatterplot. The time scale: 0 – 16 increments of 2 and the Temperature scale: 65 – 95 increments of 5

c. What letter will you use for each variable? We might use t for time and T for Temperature.

d. Draw the scatterplot. (next page)

Solution


Dallas Weather
Temperature
95
90
85
80
75
70
65
0
2
4
6
8
10
12
14
16
time