

Extension

Correlation and Causation

GOAL Understand the difference between causation and correlation.

Key Vocabulary

- correlation



CC.9-12.S.ID.9 Distinguish between correlation and causation.*

You have seen that paired data have a strong positive correlation if the correlation coefficient r is close to 1 and a strong negative correlation if r is close to -1 . But a strong correlation does not necessarily imply cause and effect, or *causation*, between the paired variables.

EXAMPLE 1 Analyze a set of data

COMPUTERS The table shows the number (in millions) of music album downloads and the number (in millions) of individual federal income tax returns filed electronically each year from 2004 to 2008. Analyze the data in terms of correlation and causation.

Year	2004	2005	2006	2007	2008
Album downloads (millions), x	4.6	13.6	27.6	42.5	56.9
Electronic tax returns (millions), y	61.5	68.5	72.8	78.7	89.5

Solution

First, find the correlation coefficient. Because r is close to 1, there is a strong positive correlation. However, an increase in album downloads does not cause an increase in electronic tax returns. These increases are both a result of other factors, such as advances in technology and increased computer usage.

```
LinReg
y=ax+b
a=.4942477768
b=59.84704456
r=.974307673
r=.9870702472
```

PRACTICE

EXAMPLE 1

for Exs. 1–2

In Exercises 1 and 2, analyze the data in terms of correlation and causation.

- BASKETBALL** The table shows the number of minutes played and the number of points scored by 6 college basketball players.

Minutes, x	30	31	33	30	25	18
Points, y	14	13	13	11	7	5

- SALES** The table shows the numbers of cold drinks and hot drinks sold at an outdoor concession stand from June through November.

Hot drinks, x	100	150	200	230	250	275
Cold drinks, y	300	210	175	165	140	125

- REASONING** You want to analyze annual data for music downloads and CD sales for the period 2000–2010 in terms of correlation and causation. What would you expect to find? *Explain.*