

TRANSFORMACE A VYTVÁŘENÍ UMĚLÝCH PROMĚNNÝCH A MULTIPLE RESPONSE SETŮ.

ZUR357 Statistická analýza dat --

19. října 2017

DESKRIPTIVNÍ STATISTIKA

vypovídá o rozložení (distribuci) hodnot v rámci analyzovaného datového souboru

Rozložení dat – absolutní, relativní a kumulativní četnosti

míry centrální tendence:

- Modus (Mo)
 - Nejčetnější hodnota
- Medián (Mdn)
 - Střední hodnota
- Aritmetický průměr (M; X)

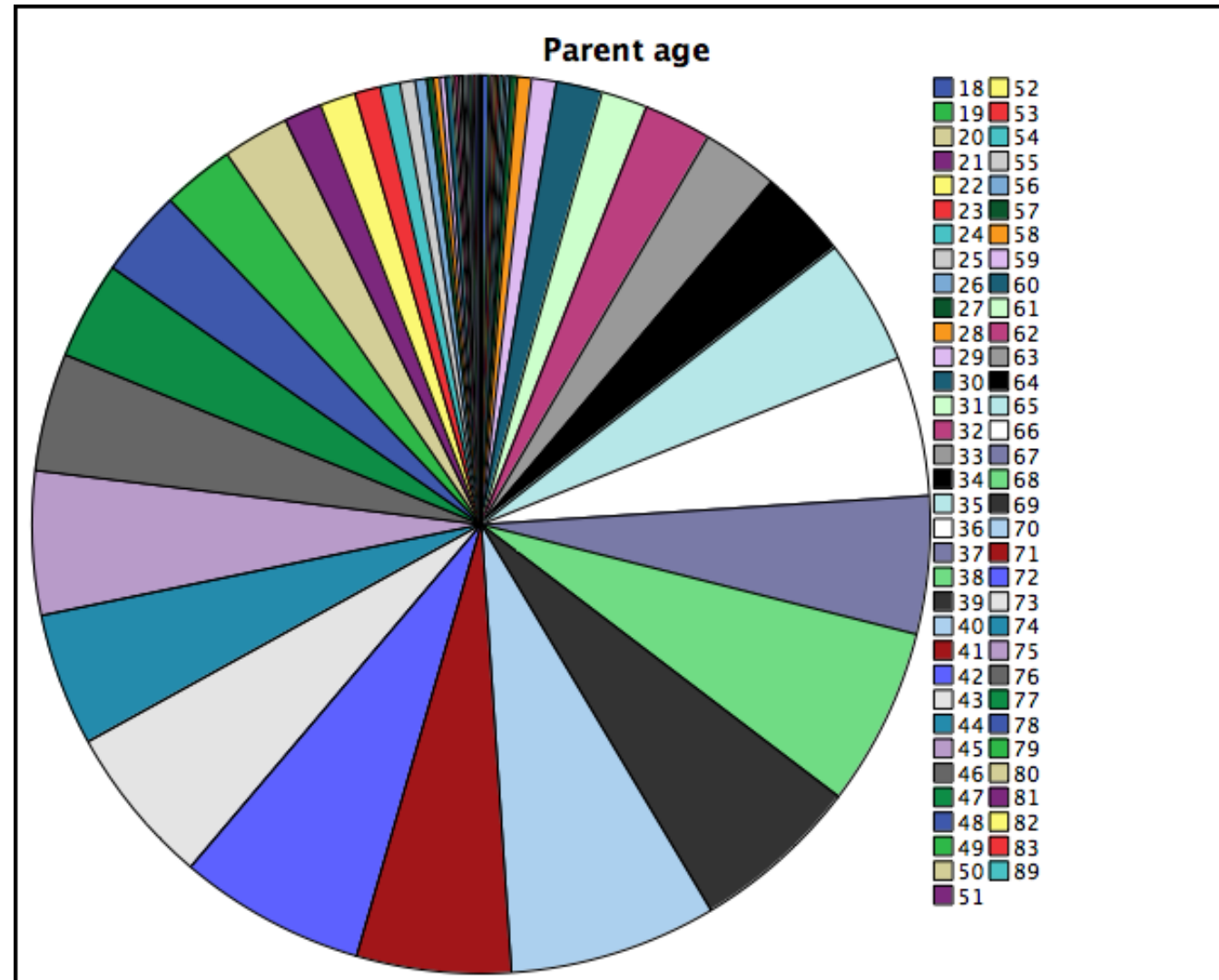
míry disperze:

- Rozpětí (R)
- Rozptyl (variance; S^2)
- Standardní (směrodatná) odchylka (S)

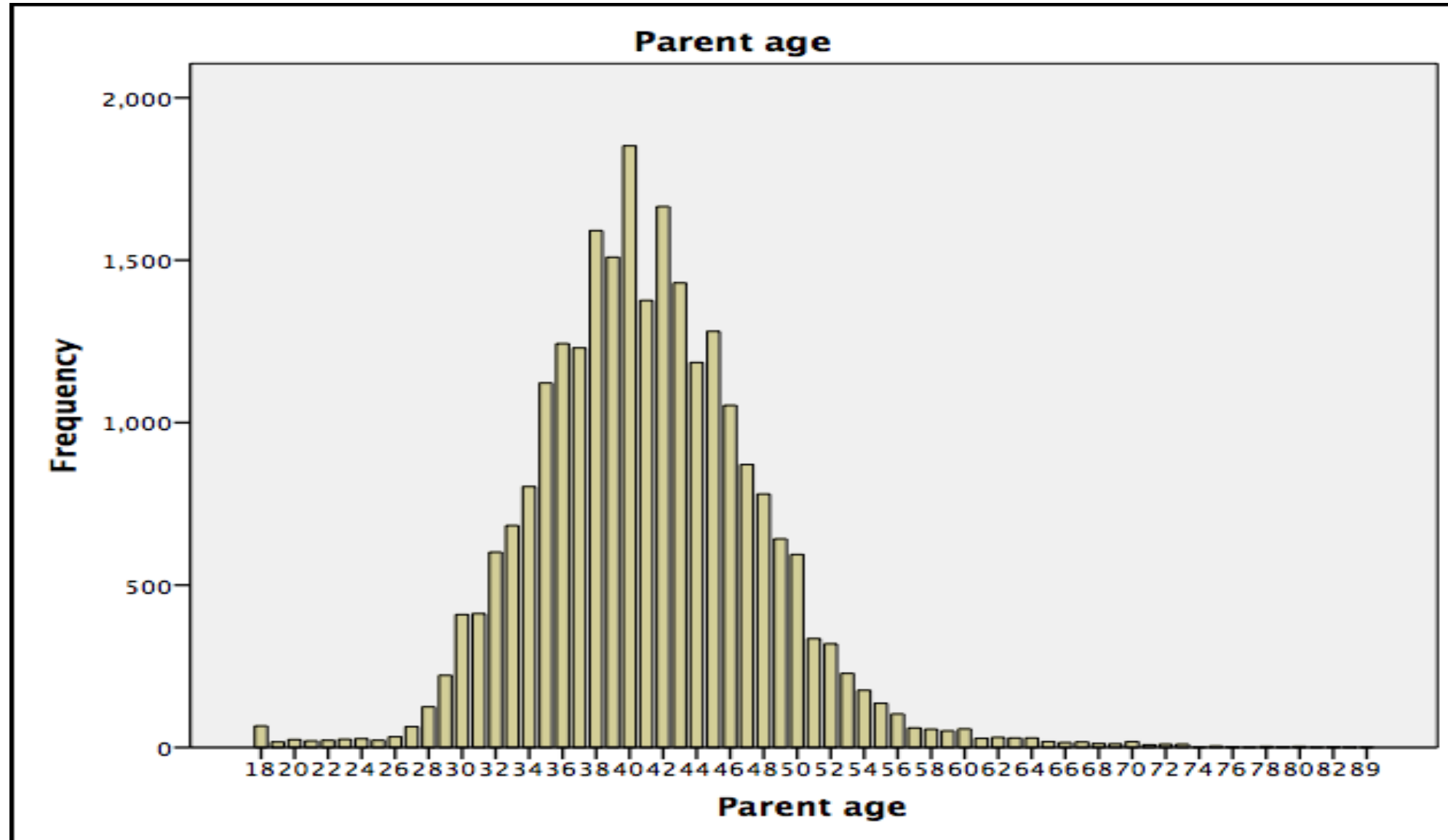
Šikmost a špičatost

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

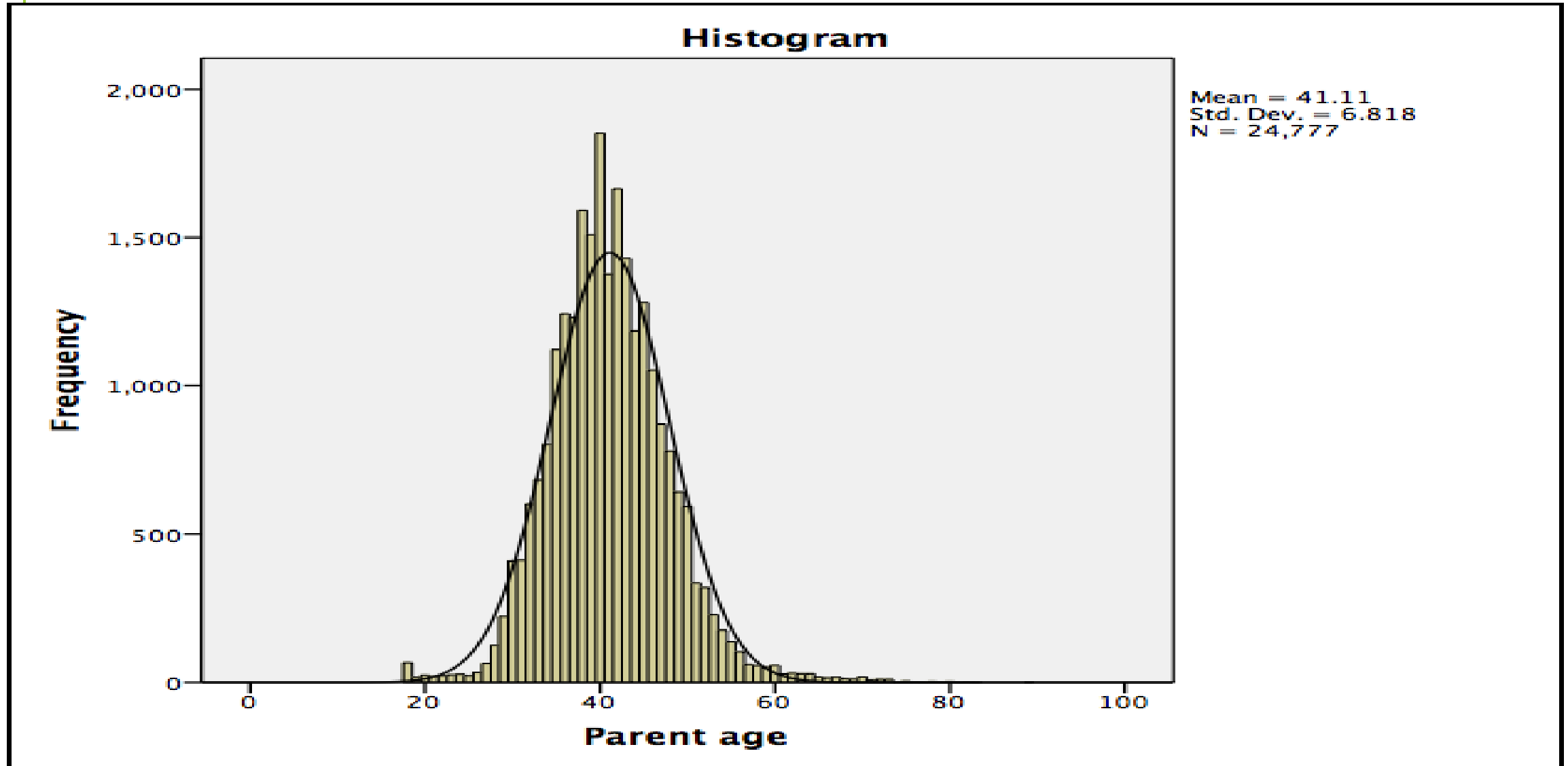
ZOBRAZENÍ DAT – ČIŠTĚNÍ DAT GRAFY



ZOBRAZENÍ DAT – ČIŠTĚNÍ DAT GRAFY



ZOBRAZENÍ DAT – ČIŠTĚNÍ DAT GRAFY



...M BERS

...on for this sad state of affairs is simple: few of us have ever been
...d to design tables and graphs effectively. Some of us have struggled to do
this work well but haven't found useful resources to assist us. Others of us
haven't struggled at all, simply because we haven't seen enough examples of
good design to recognize the inadequacy of our own efforts.

Before saying anything more, let me illustrate what I mean. Take a moment to
examine the graph below.

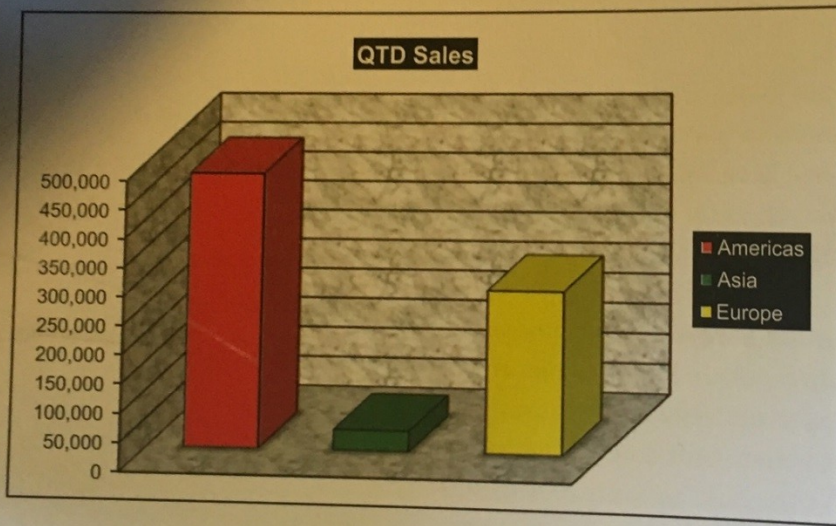


FIGURE 1.1 This is a typical example of a poorly designed business graph. Notice the attempt at artistic flair in the use of color, 3D, and shading of the vertical bars.

This graph has visual impact—it's dramatic, it's colorful, it jumps off the page and demands attention—but to what end? What is the message?

Let's evaluate its effectiveness in light of the objective. Assume that its purpose is to inform a corporation's executives every Monday morning about the current state of quarterly sales, split into three geographical regions: Americas, Europe, and Asia. Given this intention, look at it again. What message could an executive get from this graph? Put yourself in the executive's position. Take a minute or two to examine the graph and interpret its message in this light.

SlicersDicers Sales Compared to Other Product Sales

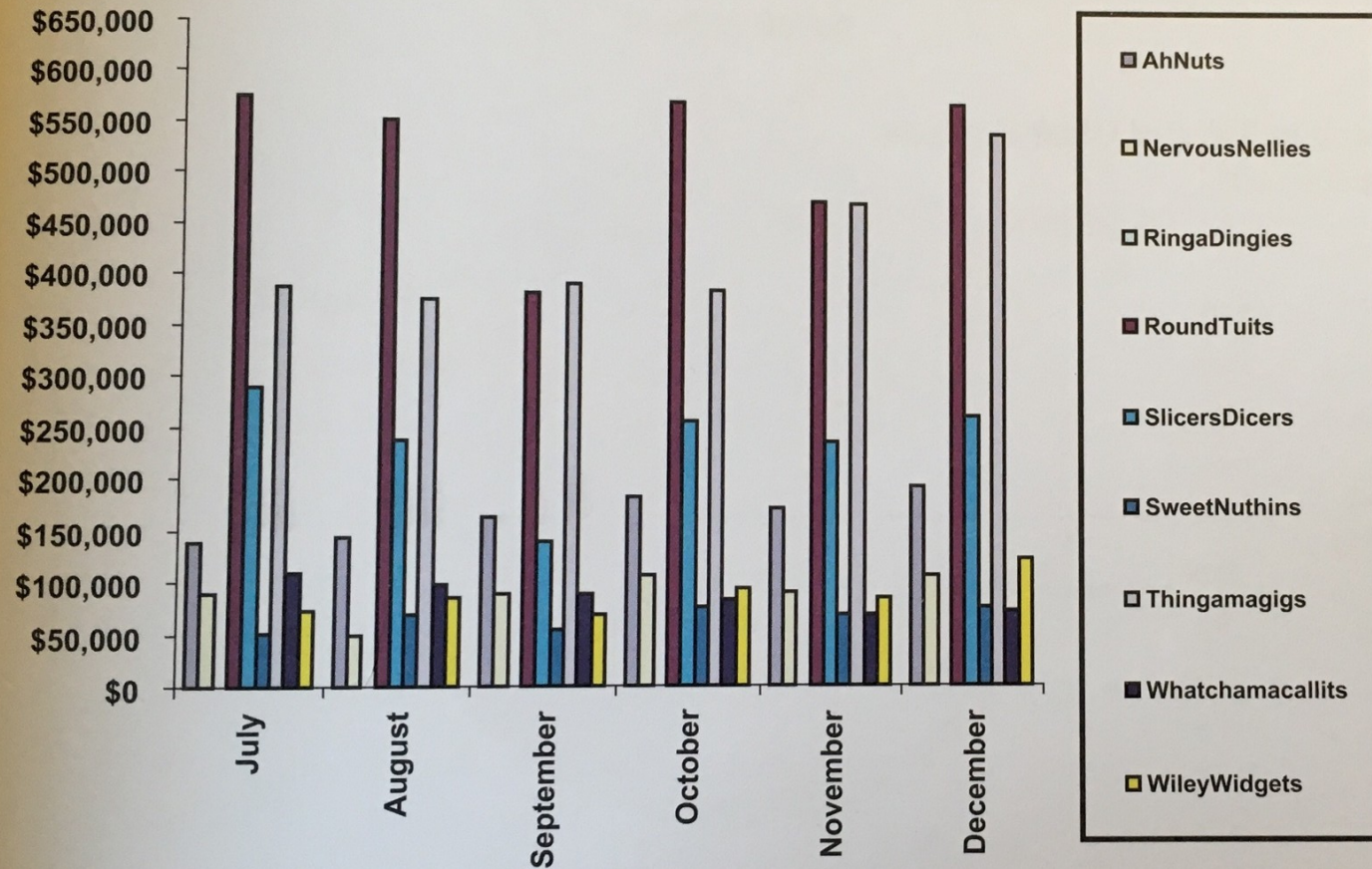


FIGURE 1.5 This is a typical example of a poorly designed vertical bar graph.

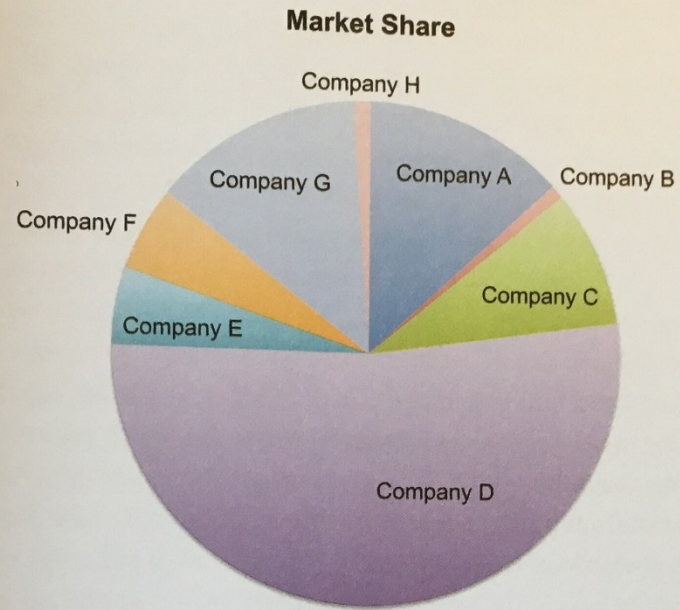


FIGURE 1.7 This is an example of an ineffective style of graph: a pie chart.

Here's the same graph below, but this time it's dressed up with the addition of a 3-D effect:

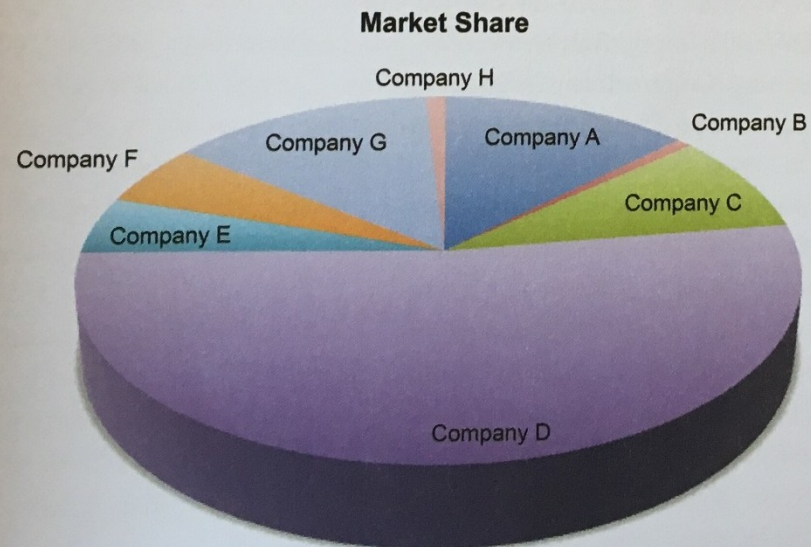
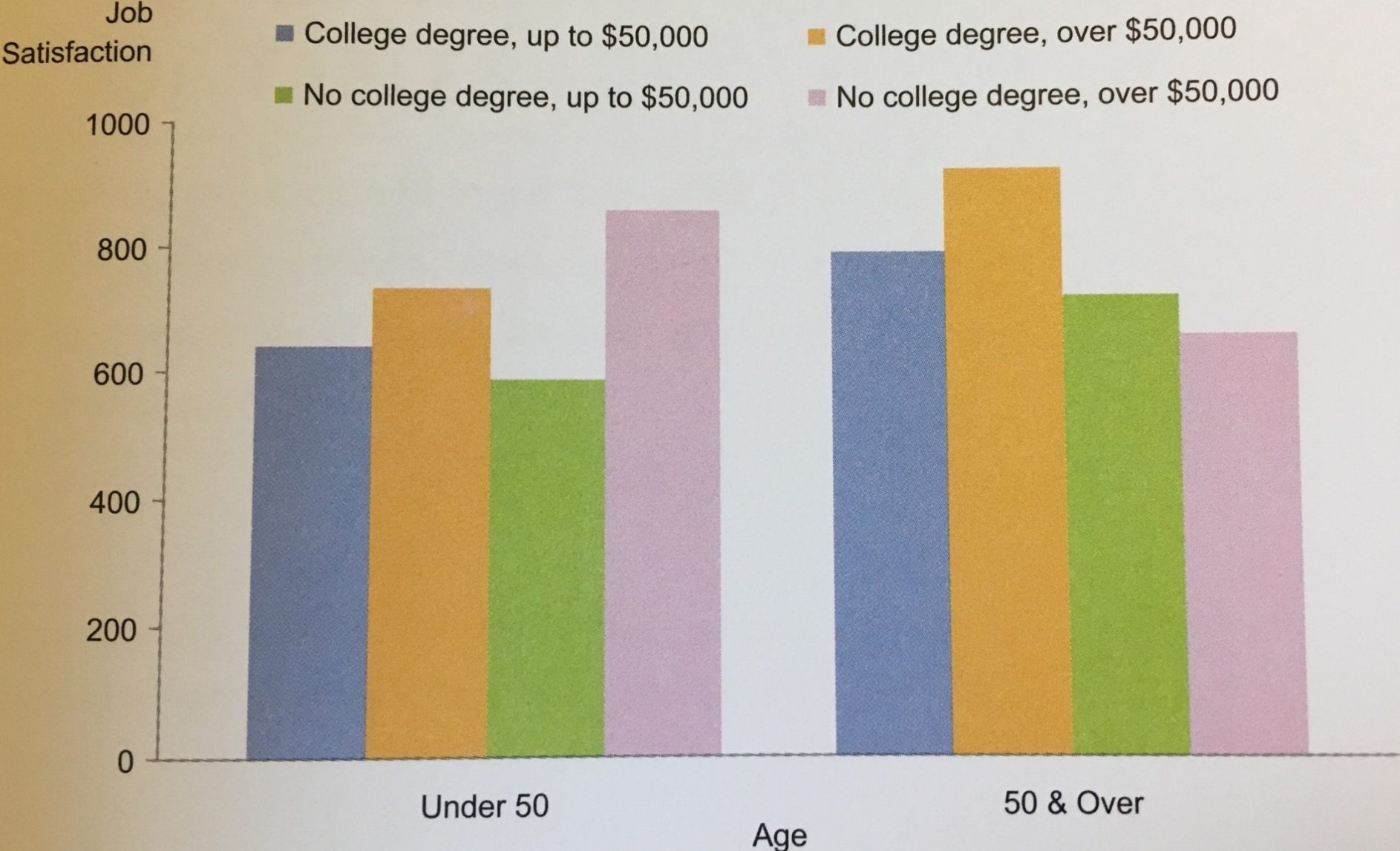
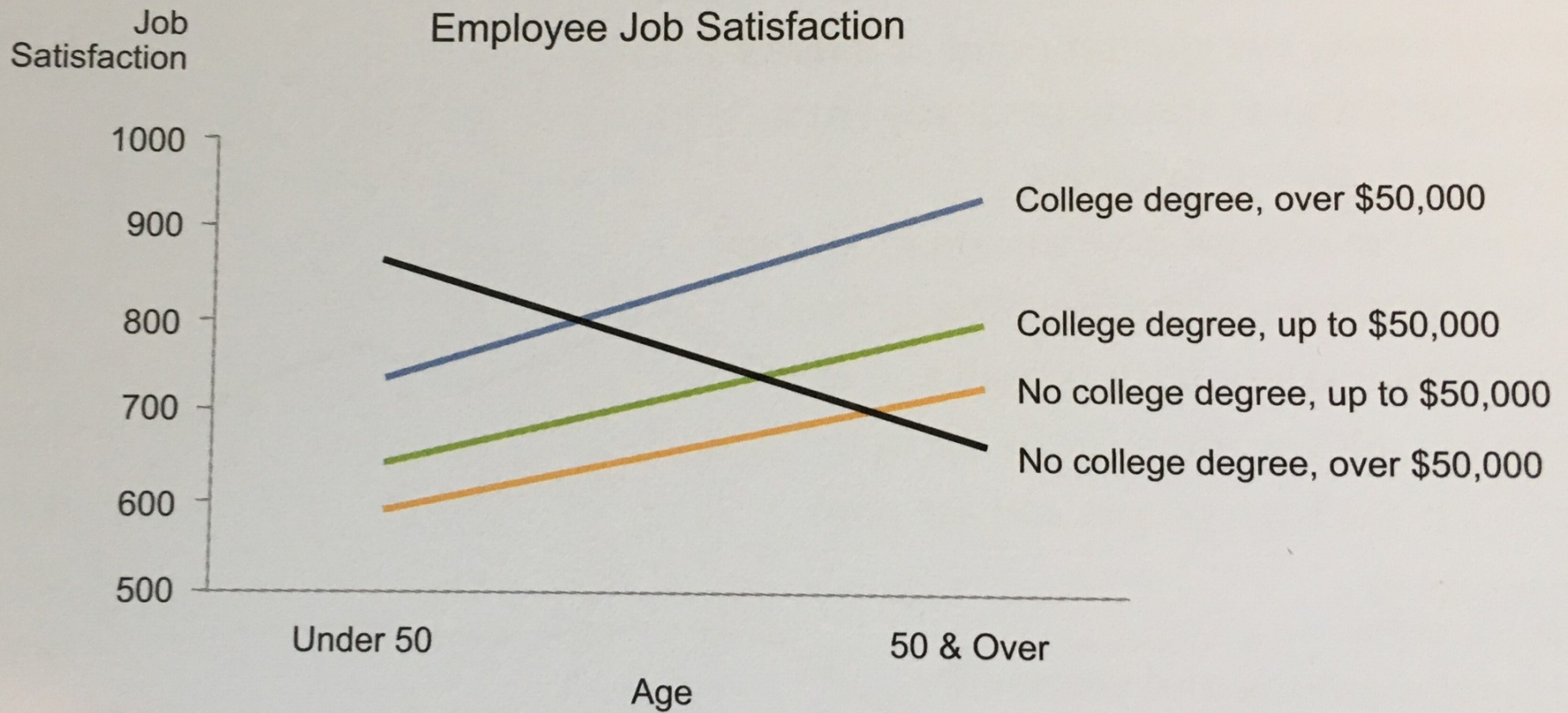


FIGURE 1.8 This is a slight variation of *Figure 1.7*, which adds 3-D perspective. This design is even less effective.

Employee Job Satisfaction





TRANSFORMACE PROMĚNNÝCH

Práce s daty před analýzou

Zjednodušení, redukce informace

Převracení škál

Změna rozložení

Atp.

RECODE

Slučuje hodnoty nominálního znaku do obecnějších kategorií (město → kraj, titul → žánr)

Kategorizuje spojité proměnné (věk → věkové kategorie)

Odstraňuje irelevantní kategorie (nevím, odložená hodnota)

Mění pořadí hodnot:

- I am relieved after I finish a game, because I do not have to play for a wh

144. In the PAST 12 MONTHS, how often, have these things happened to you?

PLEASE TICK ONE BOX ON EVERY LINE

		Very often	Fairly often	Not very often	Never/ almost never
A	I have gone without eating or sleeping because of the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B	I have felt bothered when I cannot be on the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C	I have caught myself surfing when I'm not really interested.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D	I have spent less time than I should with either family, friends or doing schoolwork because of the time I spent on the internet.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E	I have tried unsuccessfully to spend less time on the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

RECODE

Vytvořte kategorie z proměnné věk rodiče

- Do 20 let
- 21 – 30 let
- 31 – 40 let
- 41 – 50 let
- 51 – 60 let
- 61+

Otočte škálu proměnných EIU scale (vyzkoušíme SYNTAX)

Rozdělte EIU skór na tři kategorie (no riks, moderate risk, high risk)

- Substantivní přístup a teoretický přístup
- Distribuční přístup

AUTOMATIC RECODE

Transformace string values do numerických hodnot

CATEGORIZE VARIABLES

Kategorizuje kardinální znaky podle percentilů tak, aby každá kategorie obsahovala přibližně stejný počet případů

COUNT VALUES

Sčítání výskytu určité hodnoty

Např. Které tituly pravidelně čtete (0=nečte, 1 čte)

- Respekt, Reflex, Euro, Týden.....

Kolik strategií používají děti, aby se vypořádali s obtěžc

140. Again, thinking about this time, did you do any of these things?

PLEASE TICK AS MANY BOXES AS NEEDED

- A I stopped using the internet for a while
- B I deleted any messages from the person who sent it to me
- C I changed my filter/contact settings
- D I blocked the person who had sent it to me
- E I reported the problem (eg clicked on a 'report abuse' button, contact an internet advisor or "Internet service provider (ISP)")
- F None of these things
- G Don't know



RANK CASES

Vytvoří novou proměnnou, která určuje pořadí případu podle původní proměnné

COMPUTE

Libovolné operace zadané strukturovaným příkazem

Podmínka IF...

Změna rozložení

Centrování

Sumační indexy

Komplexní operace

144. In the PAST 12 MONTHS, how often, have these things happened to you?

PLEASE TICK ONE BOX ON EVERY LINE

		Very often	Fairly often	Not very often	Never/ almost never
A	I have gone without eating or sleeping because of the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B	I have felt bothered when I cannot be on the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C	I have caught myself surfing when I'm not really interested.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D	I have spent less time than I should with either family, friends or doing schoolwork because of the time I spent on the internet.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E	I have tried unsuccessfully to spend less time on the internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A JDEME DOMŮ..! 😊