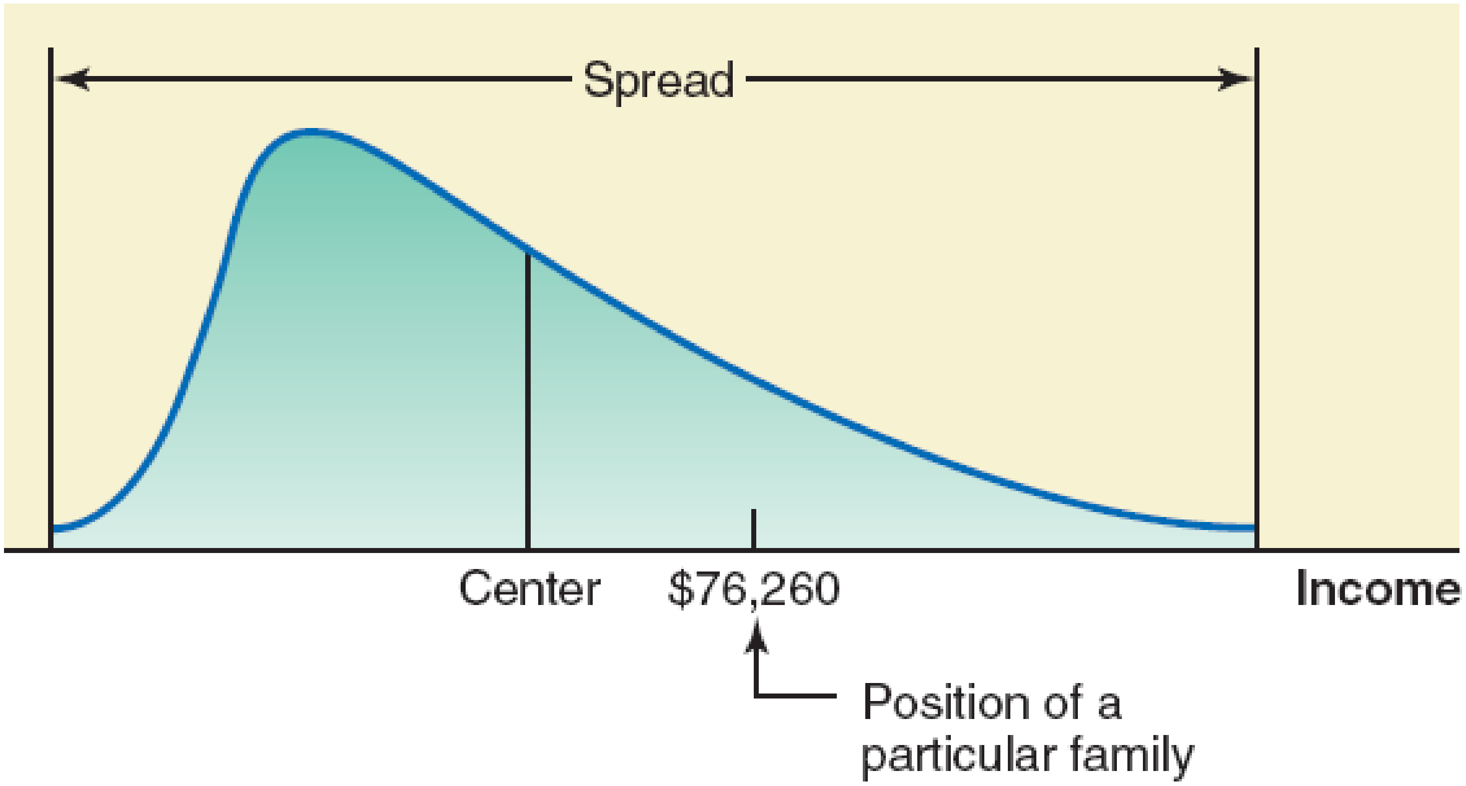
POGLAVLJE 3

NUMERIČKE DESKRIPTIVNE MJERE

3.1 MJERE CENTRALNE TENDENCIJE NEGRUPISANIH PODATAKA

- Aritmetička sredina
- Medijana
- Modus
- Odnosi između aritmetičke sredine, medijane i modusa

Slika 3.1


Spread
Center
$76,260
Income
Position of a
particular family

Aritmetička sredina

Aritmetička sredina negrupisanih podataka se dobija dijeljenjem zbira svih vrijednosti sa brojem tih vrijednosti u seriji podataka. Tako je,

Aritmetička sredina skupa: $$\mu = \frac{\sum x}{N}$$

Aritmetička sredina uzorka: $$\bar{x} = \frac{\sum x}{n}$$

Gdje $\sum x$ je suma svih vrijednosti; *N* je veličina skupa; *n* je veličina uzorka; μ je aritmetička sredina skupa; i $\bar{x}$ je aritmetička sredina uzorka.

Primjer 3-1

Tabela 3.1 pokazuje ukupne novčane donacije (zaokružene na milione dolara) koje je tokom 2010. godine dalo osam američkih kompanija

(*Izvor*: Na osnovu podataka američke Uprave prihoda analizirale *The Chronicle of Philanthropy* (Hronika filantropije) i *USA TODAY*).

Tabela 3.1 Novčane donacije u 2010 od osam američkih kompanija

Kompanija	Novčane donacije (u milionima dolara)
Wal-Mart	319
Exxon Mobil	199
Citigroup	110
Home Depot	63
Best Buy	21
Goldman Sachs	315
American Express	26
Nike	63

Odrediti aritmetičku sredinu novčanih donacija ovih osam kompanija.

Primjer 3-1: Rješenje

$$\sum x = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8$$

$$= 319 + 199 + 110 + 63 + 21 + 315 + 26 + 63 = 1116$$

$$\bar{x} = \frac{\sum x}{n} = \frac{1116}{8} = 139.5 = \$139.5\ miliona$$

Dakle, ovih osam kompanija je u prosjeku doniralo **$139.5 miliona** u 2010. godini u dobrotvorne svrhe.

Primjer 3-2

Dati su podaci o starosti (u godinama) za svih osam zaposlenih jedne male firme:

53 32 61 27 39 44 49 57

Naći prosječnu starost ovih zaposlenih.

Primjer 3-2: Rješenje

Aritmetička sredina skupa je

$$\mu = \frac{\sum x}{N} = \frac{362}{8} = 45.25\ godina$$

Prema tome, prosječna starost svih osam zaposlenih ove firme je **45.25 godina**, ili 45 godina i 3 mjeseca.

Primjer 3-3

Tabela 3.2 sadrži ukupan broj zaplijenjenih domova po službenoj dužnosti u sedam država za vrijeme 2010. godine.

Tabela 3.2 Broj domova izgubljenih po službenoj dužnosti 2010

Država	Broj domova izgubljenih po službenoj dužnosti
Kalifornija	173,175
Ilinois	49,723
Minesota	20,352
Nju Džersi	10,824
Ohajo	40,911
Pensilvanija	18,038
Teksas	61,848

Primjer 3-3

Primijetite da je broj zaplijenjenih domova u Kaliforniji dosta veliki u odnosu na preostalih šest država. Dakle, to je ekstremna vrijednost. Pokazati kako uključivanje ove vrijednosti utiče na vrijednost aritmetičke sredine.

Primjer 3-3: Rješenje

Ako ne uključimo broj zaplijenjenih domova u Kaliforniji (ekstremna vrijednost), prosječan broj zaplijenjenih domova u šest država je

Aritmetička sredina bez ekstremne vrijednosti =

$$\frac{49{,}723+20{,}352+10{,}824+40{,}911+18{,}038+61{,}848}{6} = \frac{201{,}696}{6} = 33{,}616$$

Primjer 3-3: Rješenje

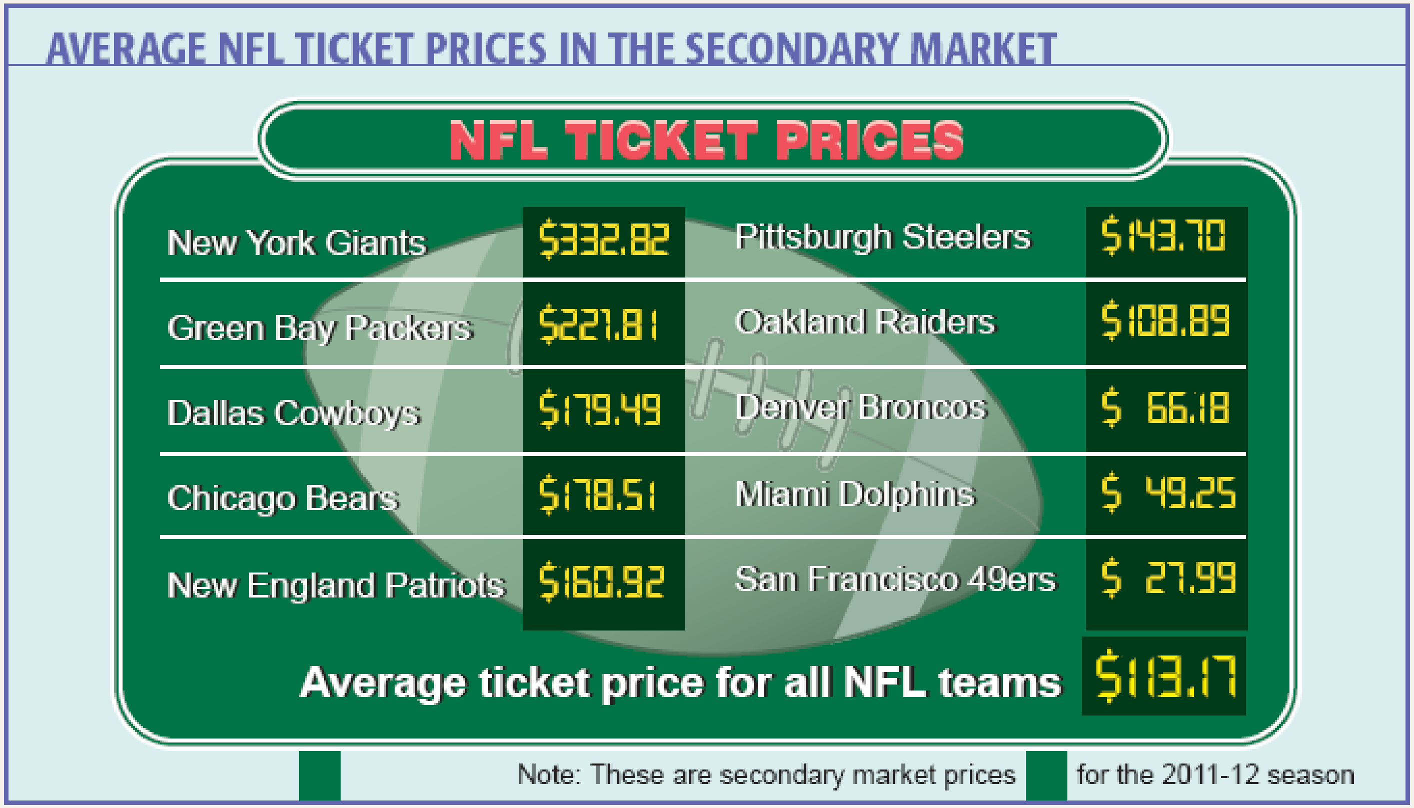
Sada, da bismo vidjeli uticaj ekstremne vrijednosti na aritmetičku sredinu, uključujemo broj zaplijenjenih domova u Kaliforniji i nalazimo prosječan broj zaplijenjenih domova u sedam država. Ovaj prosjek je

Aritmetička sredina sa ekstremnom vrijednošću =

$$\frac{173{,}175+49{,}723+20{,}352+10{,}824+40{,}911+18{,}038+61{,}848}{7} =$$

$$\frac{374{,}871}{7} = 53{,}553$$

Studija slučaja 3-1 Prosječne cijene karata za NFL na sekundarnom tržištu



Data source: seatgeek.com.

Medijana

Definicija

Medijana je jednaka vrijednosti središnjeg člana serije podataka koji su rangirani u rastućem poretku.

Izračunavanje medijane podrazumijeva sledeća dva koraka:

1. Rangiranje podataka od najnižeg ka najvišem.
2. Pronalaženje središnjeg člana. Vrijednost ovog člana jednaka je medijani.

Primjer 3-4

Vratimo se na podatke zaplijenjenih domova u u sedam država koji su dati u Tabeli 3.2 primjera 3.3. Ove vrijednosti su prikazane u donjem redu.

173,175 49,723 20,352 10,824 40,911 18,038 61,848

Odrediti medijanu za ove podatke.

Primjer 3-4: Rješenje

Najprije, rangiramo podatke od najniže do naviše vrijendosti kao što slijedi:

10,824 18,038 20,352 40,911 49,723 61,848 173,175

Pošto imamo sedam podataka u ovoj seriji, središnji član je četvrti član,

10,824	18,038	20,352	40,911	49,723	61,848	173,175
			↑			
			Median			

Dakle, medijana broja zaplijenjeih domova u ovih sedam država je iznosila **40,911** u 2010. godini.

Primjer 3-5

Tabela 3.3 pokazuje ukupne naknade (u milionima dolara) za 2010. godinu, za 12 najplaćenijih izvršnih direktora američkih kompanija.

Tabela 3.3 Ukupne naknade 12 najplaćenijih izvršnih direktora za 2010. godinu

Naći medijanu za ove podatke.

CEO and Company	2010 Total Compensation (millions of dollars)
Michael D. White (DirecTV)	32.9
David N. Farr (Emerson Electric)	22.9
Brian L. Roberts (Comcast)	28.2
Philippe P. Dauman (Viacom)	84.5
William C. Weldon (Johnson & Johnson)	21.6
Robert A. Iger (Walt Disney)	28.0
Ray R. Iran (Occidental Petroleum)	76.1
Samuel J. Palmisano (IBM)	25.2
John F. Lundgren (Stanley Black & Decker)	32.6
Lawrence J. Ellison (Oracle)	70.1
Alan Mulally (Ford Motor)	26.5
Howard Schultz (Starbucks)	21.7

Primjer 3-5: Rješenje

Najprije rangiramo date date ukupne naknade 12 izvršnih direktora na sledeći način:

21.6 21.7 22.9 25.2 26.5 28.0 28.2 32.6 32.9 70.1 76.1 84.5

Ova serija sadrži 12 podataka. Pošto je to paran broj podataka, medijana se računa kao prosjek dvije središnje vrijednosti.

21.6 21.7 22.9 25.2 26.5 $\boxed{28.0 \quad 28.2}$ 32.6 32.9 70.1 76.1 84.5

↑

Median = 28.1

Primjer 3-5: Rješenje

Dvije središnje vrijednosti su šesta i sedma vrijednost u uređenom nizu podataka, i ove dvije vrijednosti iznose 28.0 i 28.2.

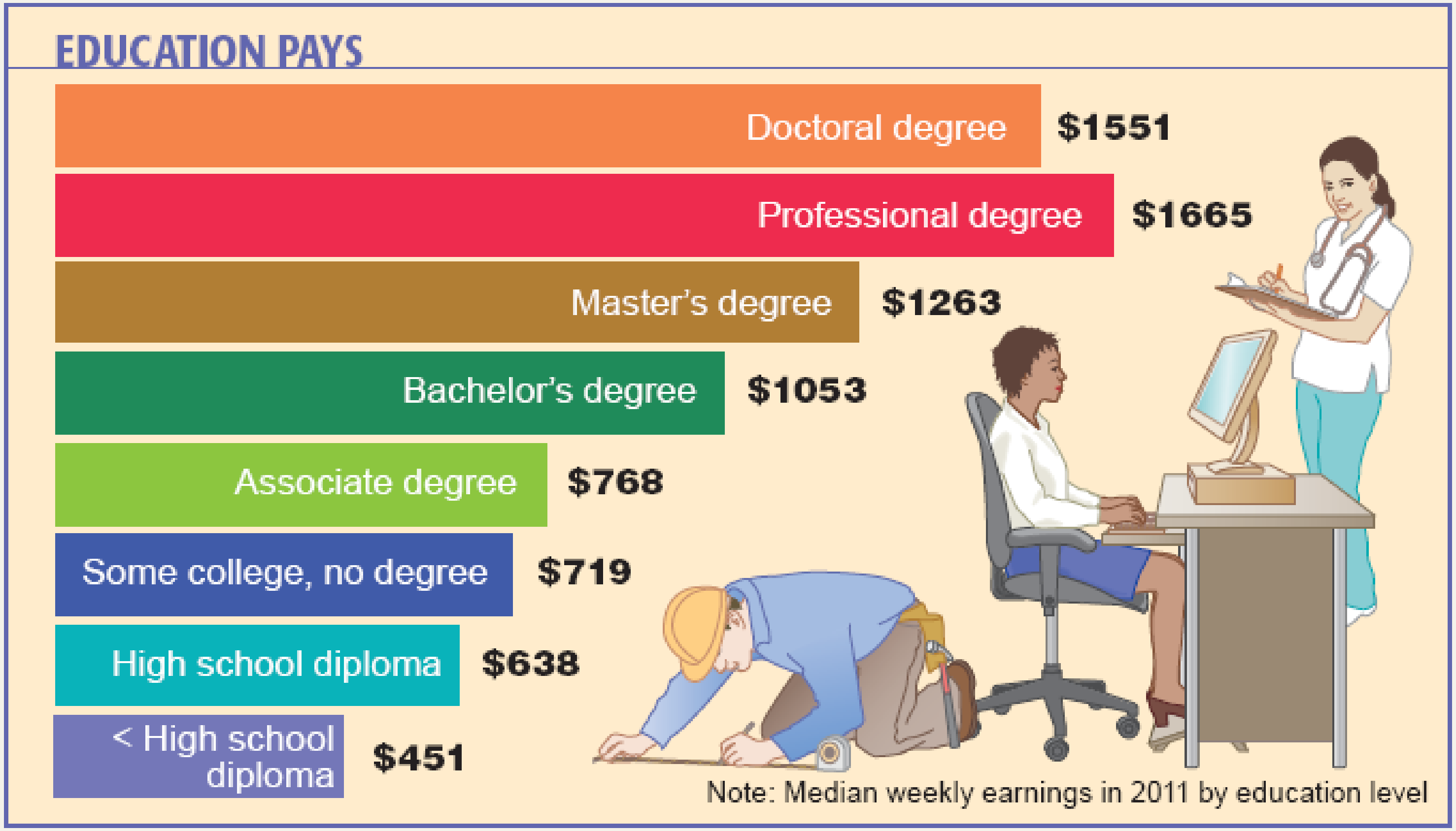
$$\text{Median} = \frac{28.0 + 28.2}{2} = \frac{56.2}{2} = 28.1 = \$28.1 million$$

Dakle, medijana za naknade ovih 12 izvršnih direktora u 2010. godini je **$28.1 milion**.

Medijana

Medijana predstavlja centar (središte) histograma, sa polovinom podataka lijevo od medijane i polovinom podataka desno od medijane. Prednost upotrebe medijane kao mjere centralne tendencije je ta što na nju ne utiču ekstremne vrijednosti. Shodno tome, medijani se daje prednost u odnosu na aritmetičku sredinu kao mjeri centralne tendencije za skupove podataka koji sadrže ekstremne vrijednosti.

Studija slučaja 3-2 Obrazovanje se isplati



Data source: U.S. Bureau of Labor Statistics.

Modus

Definicija

Modus je vrijednost koja se javlja sa najvećom frekvencijom u seriji podataka.

Primjer 3-6

Sledeći podaci predstavljaju brzine (u miljama na sat) osam auotmobila, koji su bili zaustavljeni na autoputu I-95 zbog prekoračenja dozvoljene brzine.

77	82	74	81	79	84	74	78

Odrediti modus.

Primjer 3-6: Rješenje

U ovoj seriji podataka, 74 se pojavljuje dvaput, dok se preostale vrijednosti javljaju samo jednom. Pošto vrijednost 74 ima najveću frekvenciju, ona je modus. Prema tome,

Modus = **74 milja na sat**

Modus

- Najveći nedostatak modusa je u tome što je moguće da jedna serija podataka ili nema ili ima više od jednog modusa, budući da će ta serija imati samo jednu aritmetičku sredinu i samo jednu medijanu.
  - Unimodalna: Serija podataka sa samo jednim modusom.
  - Bimodalna: Serija sa dva modusa.
  - Multimodalna: Serija sa više od dva modusa.

Primjer 3-7 (Serija koja nema modus)

Prošlogodišnja primanja pet slučajno odabranih porodica su bila \$76,150, \$95,750, \$124,985, \$87,490, i \$53,740.

Odrediti modus.

Primjer 3-7: Rješenje

Pošto se svaka vrijednost u ovoj seriji javlja samo jedanput, ova serija **nema modus**.

Primjer 3-8 (Serija sa dva modusa)

Malo preduzeće ima 12 zaposlenih. Njihovo vrijeme dolaska na posao od kuće (zaokruženo na najbliže minute) iznosi 23, 36, 12, 23, 47, 32, 8, 12, 26, 31, 18, i 28, redom.

Odrediti modus za ove podatke.

Primjer 3-8: Rješenje

U datim podacima putovanja na posao 12 zaposlenih, vrijednosti 12 i 23 se javljaju dvaput, dok se sve preostale vrijednosti javljaju samo jednom. Dakle, serija podataka ima dva modusa: **12** i **23 minuta**.

Primjer 3-9 (Serija sa tri modusa)

Godine starosti 10 slučajno izabranih studenata iz grupe su 21, 19, 27, 22, 29, 19, 25, 21, 22 i 30.

Odrediti modus.

Primjer 3-9: Rješenje

Ova serija podataka ima tri modusa: **19**, **21** i **22**. Svaka od ove tri vrijednosti se javlja sa (najvišom) frekvencijom koja iznosi 2.

Modus

Jedna od prednosti modusa je u tome što se može izračunavati i za kvantitativne i za kvalitativne podatke, dok se aritmetička sredina i medijana mogu izračunavati samo za kvantitativne podatke.

Primjer 3-10

Status pet studenata koji su članovi studentskog senata na koledžu su student četvrte godine, student druge godine, student četvrte godine, student treće godine, i student četvrte godine. Odrediti modus.

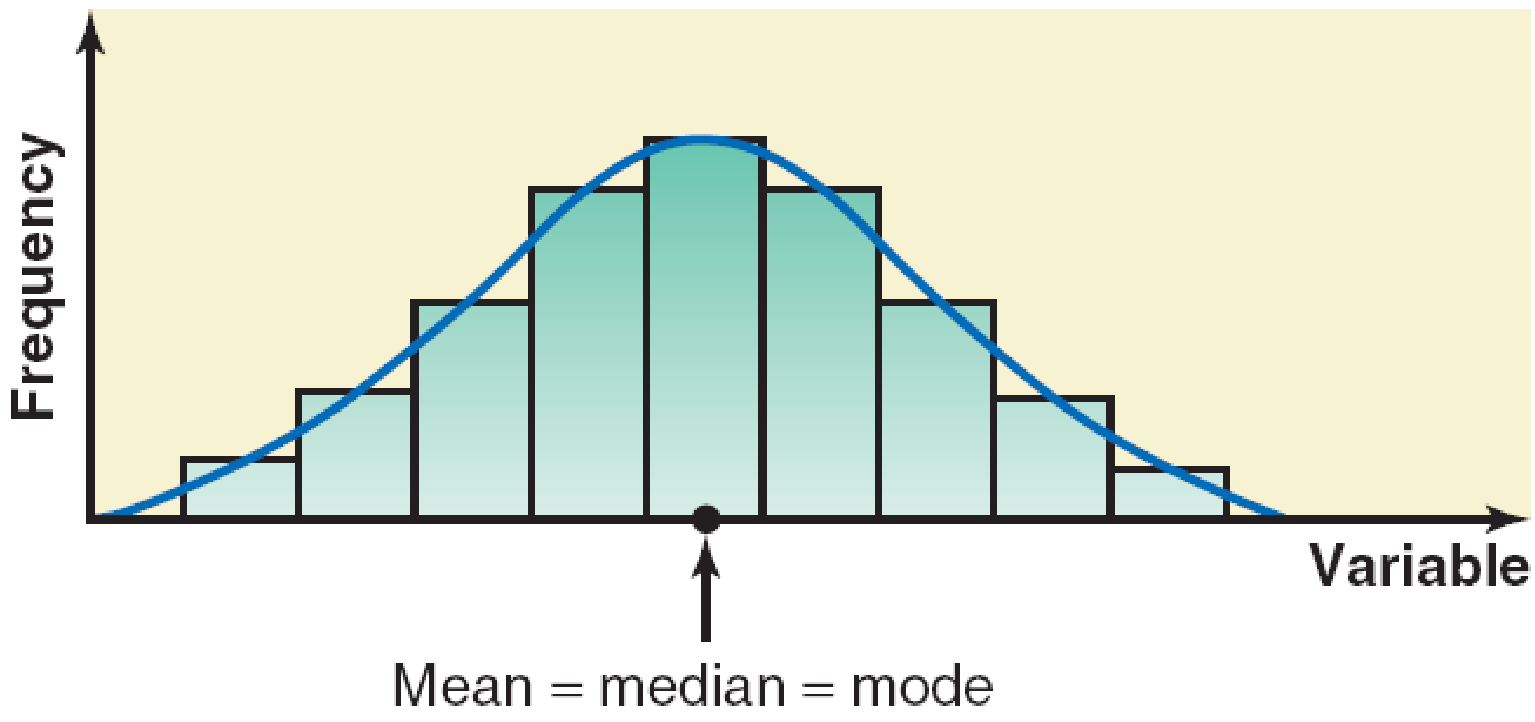
Primjer 3-10: Rješenje

Pošto se kategorija **student četvrte godine** pojavljuje češće od ostalih kategorija, ona je modus ove serije podataka. Za ovu seriju podataka nije moguće izračunati aritmetičku sredinu i medijanu.

Odnosi između aritmetičke sredine, medijane i modusa

1. Za simetrični histogram i simetričnu krivu raspodjele (pogledati sliku 3.2), vrijednosti aritmetičke sredine, medijane i modusa su identične i nalaze se u centru raspodjele.

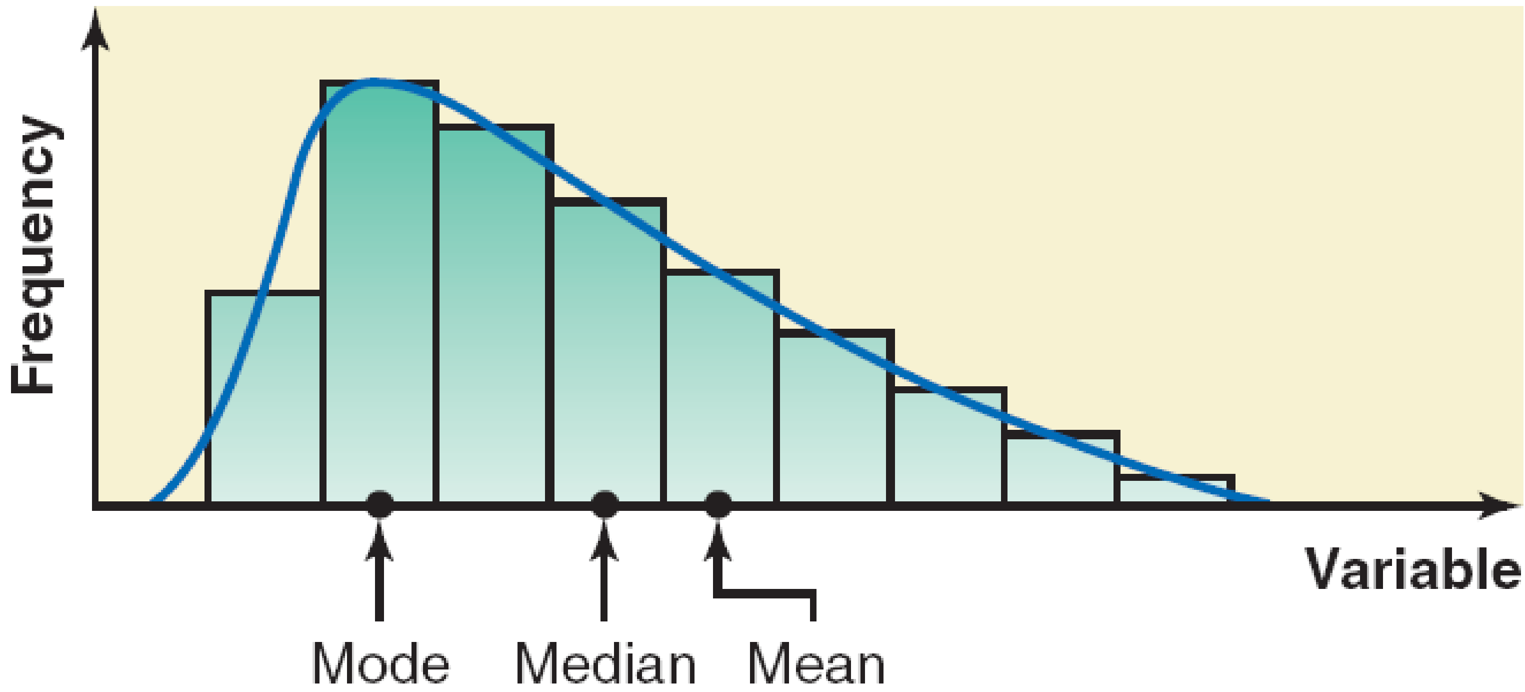
Slika 3.2 Aritmetička sredina, medijana i modus za simetričan histogram i simetričnu krivu raspodjele frekvencija.



Odnosi između aritmetičke sredine, medijane i modusa

2. Kod histograma i krive raspodjele frekvencija koji su asimetrični udesno (pogledati sliku 3.3), aritmetička sredina ima najveću, modus najmanju vrijednost, dok je vrijednost medijane između ove dvije. (Primijetimo da je modus apscisa najviše tačke krive.) Vrijednost aritmetičke sredine je u ovom slučaju najveća jer je osjetljiva na ekstremne vrijednosti koje se javljaju na desnom kraju krive. Ove ekstremne vrijednosti vuku aritmetičku sredinu udesno.

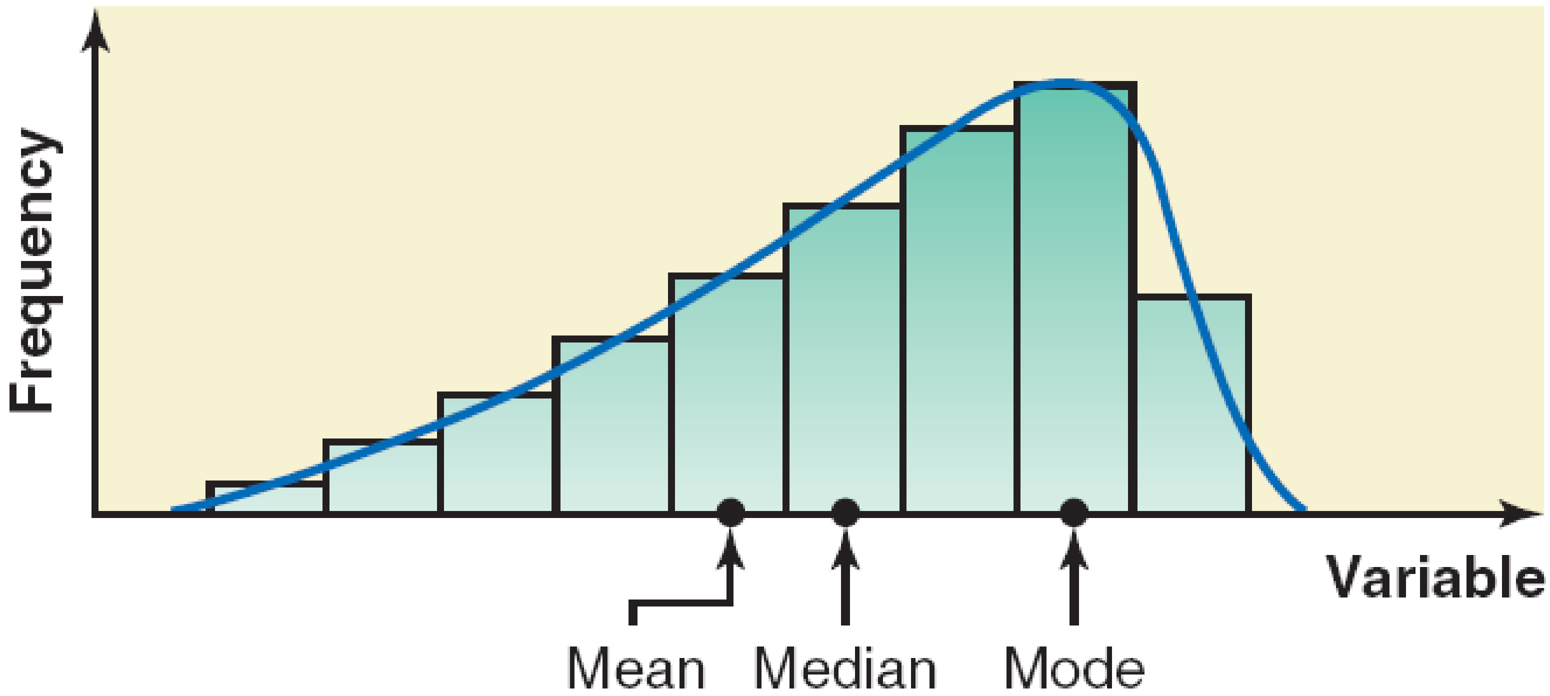
Slika 3.3 Aritmetička sredina, medijana i modus za histogram i krivu raspodjele frekvencija koji su asimetrični udesno.



Odnosi između aritmetičke sredine, medijane i modusa

3. Ako su histogram i kriva raspodjele frekvencija asimetrični ulijevo (pogledati sliku 3.4), aritmetička sredina ima najmanju, a modus najveću vrijednost, dok je vrijednost medijane između ove dvije. U ovom slučaju, ekstremne vrijednosti na lijevom kraju raspodjele vuku aritmetičku sredinu ulijevo.

Slika 3.4 Aritmetička sredina, medijana i modus za histogram i krivu raspodjele frekvencija koji su asimetrični ulijevo.



3.2 MJERE DISPERZIJE NEGRUPISANIH PODATAKA

- Interval varijacije (razmak varijacije)
- Varijansa i standardna devijacija
- Parametri osnovnog skupa i statistike uzorka

Interval varijacije

Određivanje intervala varijacije za negrupisane podatke

Interval varijacije = Najveća vrijednost – Najmanja vrijednost

Primjer 3-11

U tabeli 3.4 su date ukupne površine u kvadratnim miljama četiri države centralnog juga u SAD.

Odrediti interval varijacije ove serije podataka.

Tabela 3.4

Država	Ukupna površina (u kvadratnim miljama)
Arkanzas	53,182
Lujzijana	49,651
Oklahoma	69,903
Teksas	267,277

Primjer 3-11: Rješenje

Interval varijacije = Najveća vrijednost – Najmanja vrijednost

= 267,277 – 49,651

= **217,626 kvadratnih milja**

Dakle, ukupne površine ove četiri države su raspršene u razmaku od 217,626 kvadratnih milja.

Interval varijacije

Nedostaci

- Interval varijacije, kao i aritmetička sredina, ima nedostatak što na njega utiču ekstremne vrijednosti. Prema tome, interval varijacije nije dobra mjera disperzije za seriju podataka koja sadrži ekstremne vrijednosti.
- Za njegovo izračunavanje koriste se samo dvije vrijednosti: najveća i najmanja. Sve ostale vrijednosti serije podataka se zanemaruju. Stoga, interval varijacija nije zadovoljavajuća mjera disperzije.

Varijansa i standardna devijacija

- Standardna devijacija je najčešće korišćena mjera disperzije.

- Vrijednost standardne devijacije pokazuje koliko blizu su vrijednosti serije podataka grupisane oko aritmetičke sredine.

- Uopšteno, manja vrijednost standardne devijacije serije podataka ukazuje da su vrijednosti te serije raspršene veoma malo oko aritmetičke sredine.

Varijansa i standardna devijacija

- Nasuprot tome, veća vrijednost standardne devijacije serije podataka ukazuje da su vrijednosti te serije raspršene u relativno velikom razmaku oko aritmetičke sredine.

- Standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena varijanse.

Varijansa i standardna devijacija

- Varijansa osnovnog skupa se označava sa σ^2 (čitamo kao sigma na kvadrat), a varijansa uzorka se označava sa s^2.

- U skladu sa tim, standardna devijacija osnovnog skupa se označava sa σ, a standardna devijacija uzorka se označava sa s.

Varijansa i standardna devijacija

Osnovne formule za varijansu i standardnu devijaciju negrupisanih podataka

$$\sigma^2 = \frac{\sum (x-\mu)^2}{N} \quad \text{i} \quad s^2 = \frac{\sum \left(x-\overline{x}\right)^2}{n-1}$$

$$\sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}} \quad \text{i} \quad s = \sqrt{\frac{\sum \left(x-\overline{x}\right)^2}{n-1}}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, σ je standardna devijacija osnovnog skupa, i s je standardna devijacija uzorka.

Tabela 3.5

x	$x - \bar{x}$
82	$82 - 84 = -2$
95	$95 - 84 = +11$
67	$67 - 84 = -17$
92	$92 - 84 = +8$
	$\Sigma(x - \bar{x}) = 0$

Varijansa i standardna devijacija

Radne formule za varijansu i standardnu devijaciju negrupisanih podataka

$$\sigma^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N} \quad \text{i} \quad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}$$

$$\sigma = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N}} \quad \text{i} \quad s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, σ je standardna devijacija osnovnog skupa, i s je standardna devijacija uzorka.

Primjer 3-12

Do 2009. godine, putnicima aviokompanija se nije naplaćivao prijavljeni prtljag. Međutim, negdje oko 2009. godine, mnoge američke aviokompanije počele su da naplaćuju naknadu za torbe. Prema Zavodu za statistiku transporta, američke aviokompanije su u 2010. godini prikupile više od 3 milijarde dolara prihoda od naknada za prtljag. Sledeća tabela navodi prihode od ovih naknada za šest američkih aviokompanija za 2010. godinu.

Odrediti varijansu i standardnu devijaciju za ove podatke.

Primjer 3-12

Aviokompanija	Prihodi od naknada za prtljag (u milionima dolara)
United	313
Continental	342
American	581
Delta	952
US Airways	514
Air Tran	152

Primjer 3-12: Rješenje

Označimo sa x prihod od naknade prtljaga (u milionima dolara) jedne aviokompanije. Vrijednosti Σx i Σx^2 su izračunate u Tabeli 3.6.

x	x^2
313	97,969
342	116,964
581	337,561
952	906,304
514	264,196
152	23,104
$\Sigma x = 2854$	$\Sigma x^2 = 1{,}746{,}098$

Primjer 3-12: Rješenje

Korak 1. Izračunati Σx

Zbir vrijednosti prve kolone u Tabeli 3.6 daje 2,854.

Korak 2. Naći Σx^2

Rezultati ovog koraka su prikazani u drugoj koloni Tabele 3.6, što daje sumu 1,746,098.

Primjer 3-12: Rješenje

Korak 3. Odrediti varijansu

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{1,746,098 - \frac{(2,854)^2}{6}}{6-1}$$

$$= \frac{1,746,098 - 1,357,552.667}{5}$$

$$= 77,709.06666$$

Primjer 3-12: Rješenje

Korak 4. Izračunati standardnu devijaciju

Standardna devijacija se računa uzimanjem (pozitivnog) kvadratnog korijena varijanse:

$$s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{77{,}709.06666}$$

$$= 278.7634601 = \$278.76 million$$

Dakle, standardna devijacija prihoda od naknada za prtljag ovih šest aviokompanija u 2010. godini je **$278.76 miliona**.

Dvije napomene

1. Vrijednosti varijanse i standardne devijacije nisu nikada negativne.

2. Jedinice mjere varijanse uvijek predstavljaju kvadrirane jedinice mjere originalnih podataka.

Primjer 3-13

Date su bruto zarade (u hiljadama dolara) za svih šest zaposlenih u jednoj manjoj firmi u 2011. godini.

88.50 108.40 65.50 52.50 79.80 54.60

Izračunati varijansu i standardnu devijaciju za ove podatke.

Primjer 3-13: Rješenje

Neka x predstavlja bruto zarade jednog zaposlenog ove firme u 2011. godini. Vrijednosti $\sum x$ i $\sum x^2$ su izračunate u Tabeli 3.7.

x	x^2
88.50	7832.25
108.40	11,750.56
65.50	4290.25
52.50	2756.25
79.80	6368.04
54.60	2981.16
$\Sigma x = 449.30$	$\Sigma x^2 = 35{,}978.51$

Primjer 3-13: Rješenje

$$\sigma^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N} = \frac{35{,}978.51 - \frac{(449.30)^2}{6}}{6} = 388.90$$

$$\sigma = \sqrt{388.90} = \$19.721 \text{ hiljada} = \$19{,}721$$

Dakle, standardna devijacija bruto zarada svih šest zaposlenih ove firme u 2011. godini je **$19,721**.

Upozorenje

Primijetimo da $\sum x^2$ nije isto što i $(\sum x)^2$. Vrijednost $\sum x^2$ se dobija sabiranjem kvadriranih vrijednosti x. Vrijednost $(\sum x)^2$ se dobija tako što se kvadrira vrijednost $\sum x$.

Parametri osnovnog skupa i statistike uzorka

- Numerička mjera kao što je aritmetička sredina, medijana, modus, interval varijacije, varijansa, ili standardna devijacija izračunata za podatke osnovnog skupa naziva se parametar skupa, ili jednostavno **parametar**.

- Deskriptivna mjera koja se računa za podatke uzorka naziva se statistika uzorka, ili jednostavno **statistika**.

3.3 ARITMETIČKA SREDINA, VARIJANSA I STANDARDNA DEVIJACIJA ZA GRUPISANE PODATKE

- Aritmetička sredina grupisanih podataka
- Varijansa i standardna devijacija grupisanih podataka

Aritmetička sredina grupisanih podataka

Izračunavanje aritmetičke sredine grupisanih podataka

Aritmetička sredina skupa: $\mu = \dfrac{\sum mf}{N}$

Aritm. sredina uzorka: $\bar{x} = \dfrac{\sum mf}{n}$

gdje je m sredina a f frekvencija intervala.

Primjer 3-14

U tabeli 3.8 prikazana je raspodjela frekvencija vremena svakodnevno provedenog u putovanju od kuće do radnog mjesta *svih* 25 zaposlenih u jednom preduzeću (u minutima) .

Izračunati aritmetičku sredinu vremena provedenog u putovanju od kuće do radnog mjesta.

Primjer 3-14

Dnevno vrijeme putovanja na posao (u minutima)	Broj zaposlenih
0 do manje od 10	4
10 do manje od 20	9
20 do manje od 30	6
30 do manje od 40	4
40 do manje od 50	2

Primjer 3-14: Rješenje

Dnevno vrijeme putovanja na posao (u minutima)	f	m	mf
0 do manje od 10	4	5	20
10 do manje od 20	9	15	135
20 do manje od 30	6	25	150
30 do manje od 40	4	35	140
40 do manje od 50	2	45	90
	$N=25$		$\sum mf=535$

Primjer 3-14: Rješenje

$$\mu = \frac{\sum mf}{N} = \frac{535}{25} = \mathbf{21.40\ minutes}$$

Dakle, zaposleni ovog preduzeća u prosjeku provedu **21.40 minuta** dnevno putujući od kuće do radnog mjesta.

Primjer 3-15

Tabela 3.9 sadrži raspodjelu frekvencija broja porudžbina koje su pristizale svakog dana, tokom proteklih 50 dana, u kancelariju jednog preduzeća koje se bavi porudžbinama poštom.

Izračunati prosjek.

Primjer 3-15

Broj porudžbina	Broj dana
10-12	4
13-15	12
16-18	20
19-21	4

Primjer 3-15: Rješenje

Broj porudžbina	f	m	mf
10-12	4	11	44
13-15	12	14	168
16-18	20	17	340
19-21	4	20	280
	n=50		$\sum mf$=832

Primjer 3-15: Rješenje

$$\bar{x} = \frac{\sum mf}{n} = \frac{832}{50} = \mathbf{16.64\ orders}$$

Znači, ovo preduzeće je u periodu od 50 dana u prosjeku primalo **16.64 porudžbina** dnevno.

Varijansa i standardna devijacija grupisanih podataka

Osnovne formule za varijansu i standardnu devijaciju grupisanih podataka

$$\sigma^2 = \frac{\sum f(m-\mu)^2}{N} \quad \text{i} \quad s^2 = \frac{\sum f(m-\bar{x})^2}{n-1}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, a m sredina intervala. U oba slučaja, standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena iz varijanse.

Varijansa i standardna devijacija grupisanih podataka

Radne formule za varijansu i standardnu devijaciju grupisanih podataka

$$\sigma^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{N}}{N} \quad i \quad s^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, a m sredina intervala.

Varijansa i standardna devijacija grupisanih podataka

Radne formule za varijansu i standardnu devijaciju grupisanih podataka

Standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena varijanse.

Standardna devijacija skupa: $\sigma = \sqrt{\sigma^2}$

Standardna devijacija uzorka: $s = \sqrt{s^2}$

Primjer 3-16

Sledeći podaci, koji su preuzeti iz Tabele 3.8 primjera 3-14, predstavljaju raspodjelu svakodnevnog vremena provedenog u putovanju od kuće do posla (u minutima) za svih 25 zaposlenih jednog preduzeća.

Izračunati varijansu i standardnu devijaciju.

Primjer 3-16

Dnevno vrijeme putovanja na posao (u minutima)	Broj zaposlenih
0 do manje od 10	4
10 do manje od 20	9
20 do manje od 30	6
30 do manje od 40	4
40 do manje od 50	2

Primjer 3-16: Rješenje

Dnevno vrijeme putovanja na posao (u minutima)	f	m	mf	m^2f
0 do manje od 10	4	5	20	100
10 do manje od 20	9	15	135	2025
20 do manje od 30	6	25	150	3750
30 do manje od 40	4	35	140	4900
40 do manje od 50	2	45	90	4050
	N=25		$\sum mf$=535	$\sum m^2f$=14,825

Primjer 3-16: Rješenje

$$\sigma^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{N}}{N} = \frac{14{,}825 - \frac{(535)^2}{25}}{25} = \frac{3376}{25} = 135.04$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{135.04} = 11.62 \text{ minutes}$$

Prema tome, standardna devijacija vremena provedenog putujući na posao ovih zaposlenih je **11.62 minuta**.

Primjer 3-17

Sledeći podaci, preuzeti iz Tabele 3.9 primjera 3-15, predstavljaju raspodjelu frekvencija broja porudžbina koje su pristizale svakoga dana, tokom proteklih 50 dana u kancelariju jednog preduzeća koje se bavi porudžbinama poštom.

Izračunati varijansu i standardnu devijaciju.

Primjer 3-17

Broj porudžbina	*f*
10-12	4
13-15	12
16-18	20
19-21	4

Primjer 3-17: Rješenje

Broj porudžbina	f	m	mf	m^2f
10-12	4	11	44	484
13-15	12	14	168	2352
16-18	20	17	340	5780
19-21	4	20	280	5600
	n=50		$\sum mf$=832	$\sum m^2f$=14,216

Primjer 3-17: Rješenje

$$s^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1} = \frac{14{,}216 - \frac{(832)^2}{50}}{50-1} = 7.5820$$

$$s = \sqrt{s^2} = \sqrt{7.5820} = 2.75 \text{ orders}$$

Dakle, standardna devijacija broja porudžbina koje su pristizale u kancelariju ovog preduzeća tokom poslednjih 50 dana je **2.75**.

Empirijsko pravilo

Za raspodjelu u obliku zvona, približno

1. 68% vrijednosti se nalazi u opsegu jedne standardne devijacije od aritmetičke sredine
2. 95% vrijednosti se nalazi u opsegu dvije standardne devijacije od aritmetičke sredine
3. 99.7% vrijednosti se nalazi u opsegu tri standardne devijacije od aritmetičke sredine

Slika 3.9 Ilustracija empirijskog pravila.