

POGLAVLJE 3

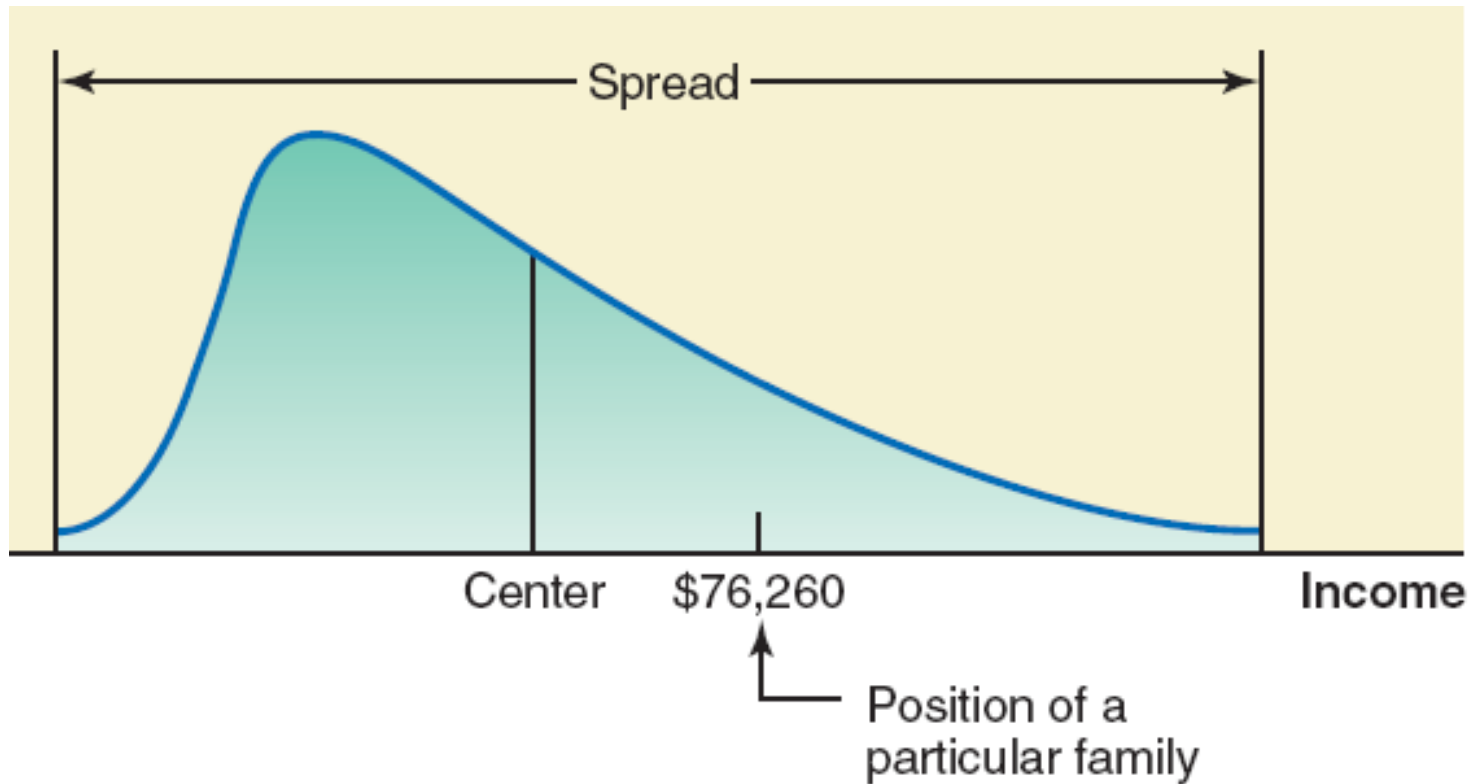


NUMERIČKE DESKRIPTIVNE MJERE

3.1 MJERE CENTRALNE TENDENCIJE NEGRUPISANIH PODATAKA

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

Slika 3.1



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

$$\begin{aligned}\sum x &= x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 \\ &= 319 + 199 + 110 + 63 + 21 + 315 + 26 + 63 = 1116\end{aligned}$$

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

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

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

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

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

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



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 najviše vrijedosti 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 zaplijenjenih 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 28.0 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.

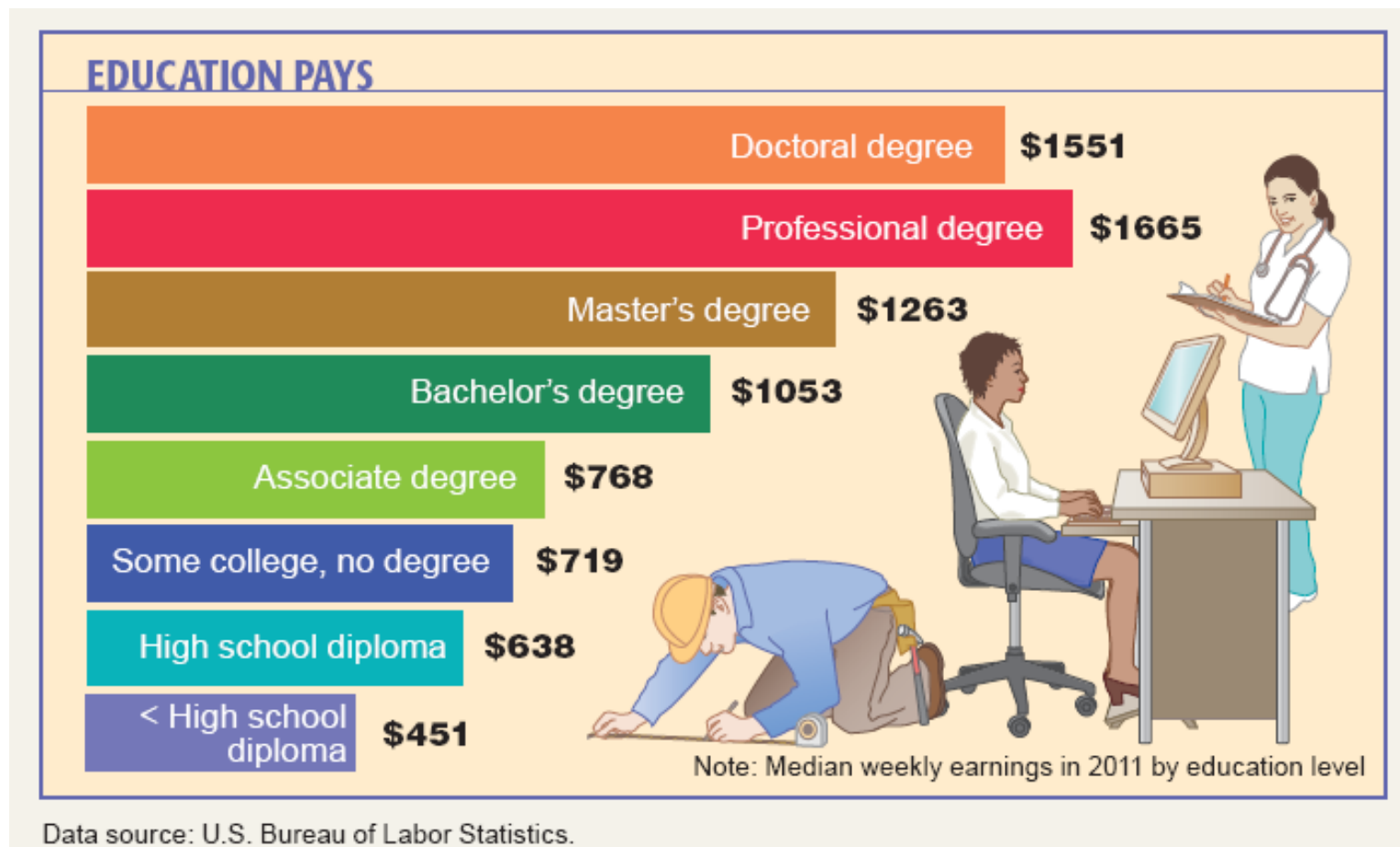
$$\text{Median} = \frac{28.0 + 28.2}{2} = \frac{56.2}{2} = 28.1 = \$28.1 \text{million}$$

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



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.

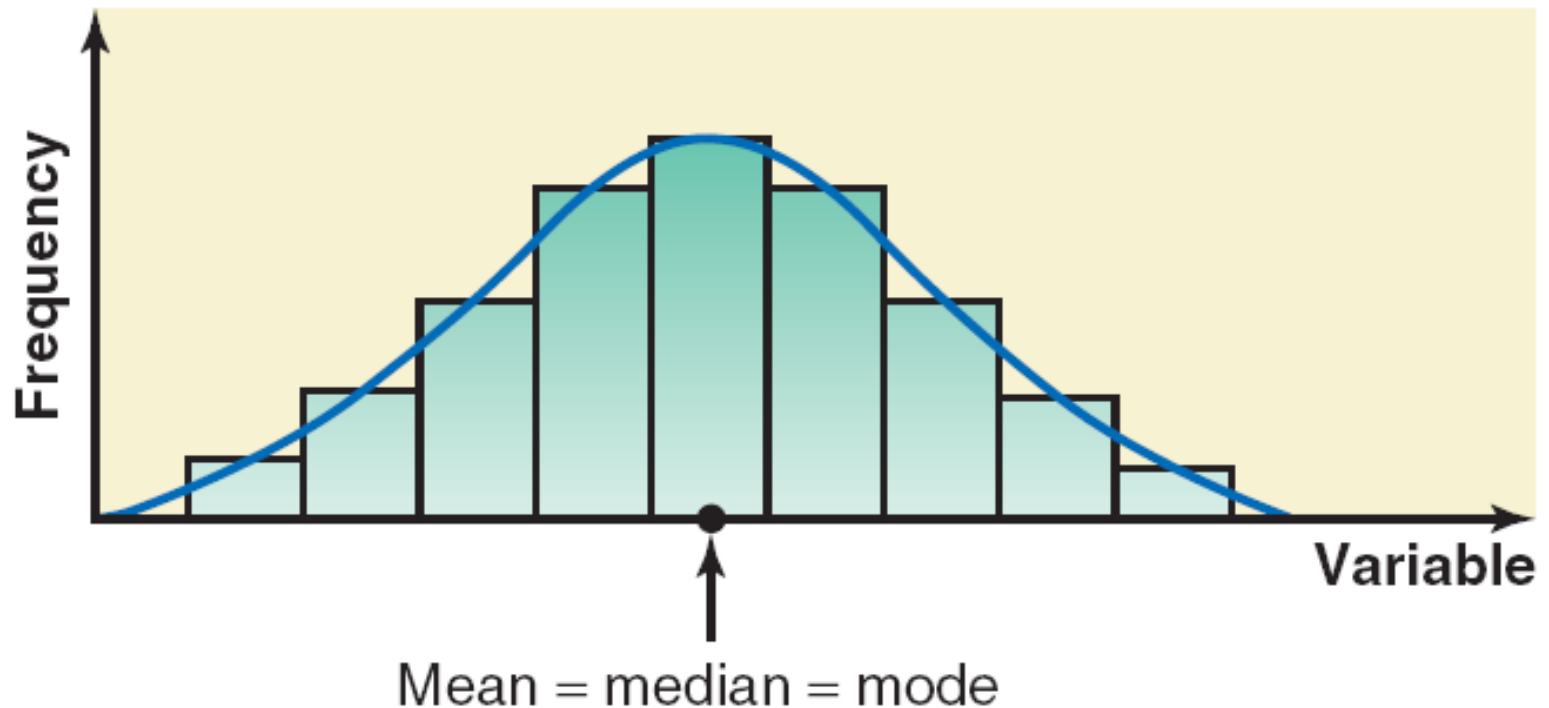
Primjer 3-10: Rješenje

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.

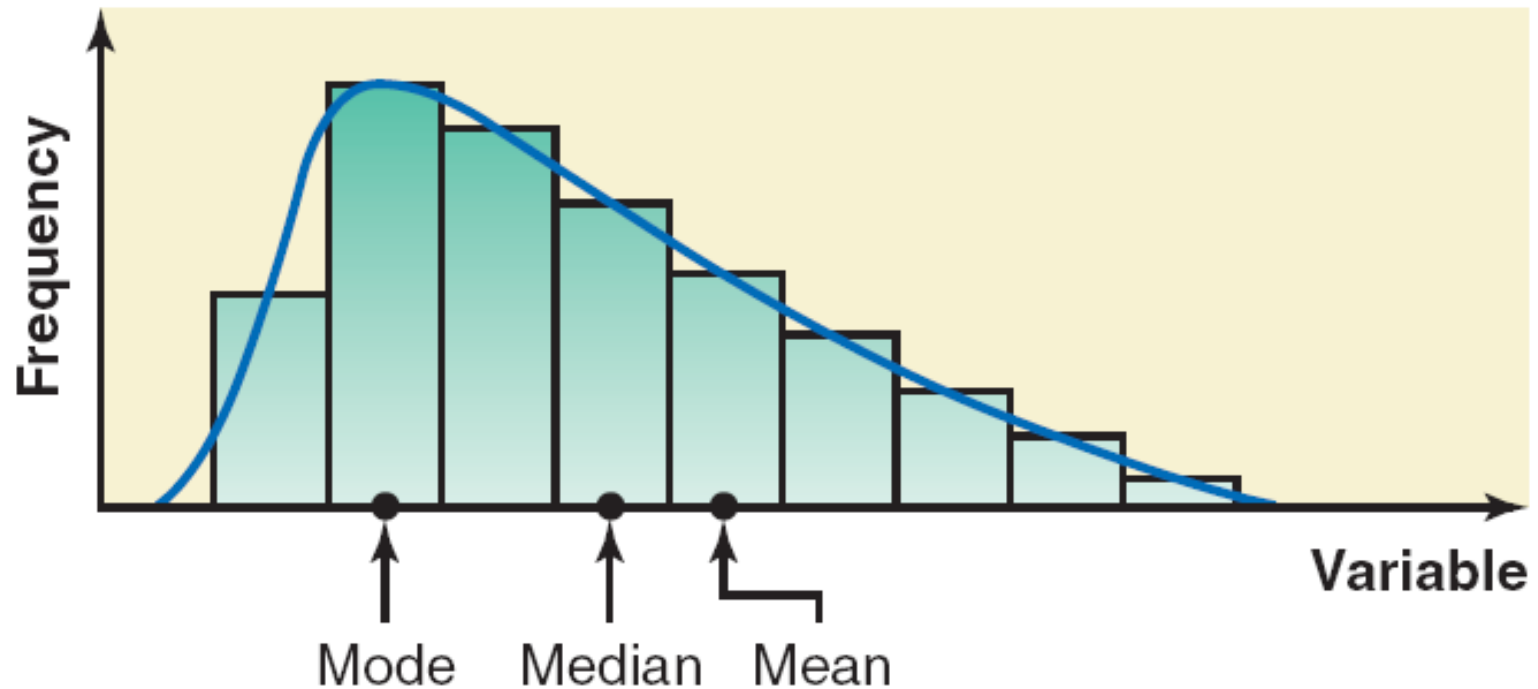
Slika 3.2 Aritmetička sredina, medijana i modus za simetričan histogram i simetričnu krivu raspodjele frekvencija.



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.

