REPUBLIKA SLOVENIJA
MINISTRSTVO ZA ŠOLSTVO IN ŠPORT

Naložba v vašo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad

POSLOVNA INFORMATIKA S STATISTIKO

FRANČIŠKA GREGORC

Višješolski strokovni program: Gostinstvo in turizem
Učbenik: Poslovna informatika s statistiko
Gradivo za 1. letnik

Avtorica:
Frančiška Gregorc, univ. dipl. ekon.
Višja strokovna šola za gostinstvo in turizem Bled

Strokovna recenzentka:
Alenka Jensterle, univ. dipl. ekon.

Lektorica:
Katarina Perger, univ. dipl. prof. slov. in bibl.

```
CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

659.23:004(075.8)(0.034.2)

GREGORC, Frančiška
   Poslovna informatika s statistiko [Elektronski vir] : gradivo za
1. letnik / Frančiška Gregorc. - El. knjiga. - Ljubljana : Zavod
IRC, 2008. - (Višješolski strokovni program Gostinstvo in turizem
/ Zavod IRC)

Način dostopa (URL): http://www.zavod-irc.si/docs/Skriti_dokumenti/
Poslovna_informatika_s_statistiko-Gregorc_2.pdf. - Projekt Impletum

ISBN 978-961-6820-95-0
249661952
```

Izdajatelj: Konzorcij višjih strokovnih šol za izvedbo projekta IMPLETUM
Založnik: Zavod IRC, Ljubljana.
Ljubljana, 2008

Strokovni svet RS za poklicno in strokovno izobraževanje je na svoji 120. seji dne 10. 12. 2009 na podlagi 26. člena Zakona o organizaciji in financiranju vzgoje in izobraževanja (Ur. l. RS, št. 16/07-ZOFVI-UPB5, 36/08 in 58/09) sprejel sklep št. 01301-6/2009 / 11-3 o potrditvi tega učbenika za uporabo v višješolskem izobraževanju.

Gradivo je sofinancirano iz sredstev projekta Impletum 'Uvajanje novih izobraževalnih programov na področju višjega strokovnega izobraževanja v obdobju 2008–11'.
Projekt oz. operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo RS za šolstvo in šport. Operacija se izvaja v okviru Operativnega programa razvoja človeških virov za obdobje 2007–2013, razvojne prioritete 'Razvoj človeških virov in vseživljenjskega učenja' in prednostne usmeritve 'Izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja'.

KAZALO VSEBINE

KAZALO SLIK

PREDGOVOR

Učbenik je namenjen študentom višje strokovne šole za gostinstvo in turizem. Nekateri se s predmetom srečujejo prvič, nekateri so osnove spoznali že v srednji šoli.

Namen predmeta je spoznati osnove statistike, zbiranja, obdelave in prikazovanja statističnih podatkov in pridobivanje informacij z uporabo različnih statističnih metod. Znanje statistične analize podatkov, prikazovanja informacij v tabelah in grafih študentje potrebujejo za uspešen študij, pripravo seminarskih in diplomske naloge, predvsem pa za pripravo poročil in analiz, ki jih bodo sestavljali v praksi. Potrebno je za dobre poslovne odločitve, za pripravo poslovnih načrtov, saj na osnovi analize podatkov in informacij iz preteklosti lahko bolje sklepamo za prihodnost. Znanje statistike študentje uporabljajo in ga dopolnjujejo tudi pri drugih predmetih.

Znanje nam omogoča tudi lažje razumevanje strokovne literature z različnih področij ter spremljanje aktualnih novic s področja gostinstva, turizma in gospodarstva. Tudi v vsakdanjem življenje je potrebno sestavljati vedno več statističnih poročil in izpolnjevati vedno več statističnih obrazcev.

Pri učenju študentje uporabljajo sodobno tehnologijo. Prednosti sodobnih medijev so velike, saj omogočajo hitro pridobivanje podatkov tudi preko interneta, obdelovanje z uporabo sodobnih programov, prikazovanje v tabelah in grafih. Sodobna informacijska tehnologija omogoča večjo hitrost in natančnost pri pridobivanju, obdelovanju in prikazovanju podatkov. Prednost je tudi v hranjenju in spreminjanju podatkov, kar olajša delo tudi pri sestavljanju poročil. Zavedati se je potrebno tudi nevarnosti, saj se podatki lahko izgubijo, poročila popravljajo, novejši programi ne znajo prebirati starejših itd.

Povečanje statičnega znanja zmanjša možnosti zavajanja in zlorabe statističnih podatkov. Statistično neizobraženi imajo mnogokrat negativen odnos do statistike, saj menijo, da podatki ne prikazujejo dejanskega stanja. Bistvo statistike niso podatki v obliki številk, ampak iskanje pomena teh številk.

Snov, ki je teoretično razložena v učbeniku, je dopolnjena na vajah s praktičnimi primeri, ki se rešujejo z računalnikov v programu Excel.

1 UVOD

Beseda STATISTIKA je vsem zelo dobro poznana, saj se uporablja v vsakdanjem življenju. Ima lahko več pomenov:

- sistematično zbrane in prikazane podatke,
- dejavnost, ki se ukvarja z opazovanjem in zbiranjem podatkov,
- veja znanosti, ki razvija teoretične metode (Knežević, 2004).

Področje statičnega zbiranja in prikazovanja podatkov ureja tudi država s številnimi predpisi. Zakon o državni statistiki določa temeljna načela in organizacijo državne statistike, program raziskovanj, metodološke osnove zbiranja, obdelovanja in shranjevanja podatkov. Določa tudi registre, zaščito in posredovanje pridobljenih podatkov.

Dejavnost državne statistike v praksi izvaja **Statistični urad Republike Slovenije**. Deluje po načelu nevtralnosti, objektivnosti, strokovne neodvisnosti, zaupnosti in preglednosti. Pri delu sodeluje s številnimi pooblaščenimi poročevalskimi enotami in izvajalci, ki zanj opravljajo raziskave.

Osnovne naloge Statističnega urada so zbiranje, obdelovanje, shranjevanje in posredovanje podatkov ter pravilno tolmačenje le-teh. Preprečevati mora nepravilno tolmačenje ter se odzvati v primeru napačne uporabe podatkov (Statistični urad RS, 2008).

Program statističnih raziskav izvaja Statistični urad ob sodelovanju s pooblaščenimi izvajalci. Določen je obseg zbiranja, obdelovanje, shranjevanje, analiziranje in izkazovanja podatkov. V programu so opredeljeni pomembni množični pojavi z ekonomskega, demografskega in socialnega področja, problematike varstva okolja itd., roki izvajanja in namen uporabe podatkov (npr. popis prebivalcev vsakih 10 let …).

Statistični urad strokovno sodeluje z mednarodnimi strokovnimi institucijami (Statistična komisija združenih narodov v okviru CANSTAT-a in OECD). Sodeluje tudi v Evropskem statističnem sistemu **EUROSTAT** in z EVROPSKO CENTRALNO BANKO. Upošteva zakonodajo, ki jo predpisuje EU. Ogled na spletni strani: http://epp.eurostat.ec.europa.eu

Statistični urad izdaja številne statistične publikacije. Nekatere se nanašajo na območje Republike Slovenije, nekatere na statistične regije ali občine. Nekatere publikacije vključujejo tudi mednarodne statistične podatke.

Najobsežnejša statistična publikacija, ki se izda vsako leto, je **Statistični letopis Slovenije**. Vsebuje statistične podatke za vsa pomembnejša področja družbenoekonomskega dogajanja (tudi dejavnost gostinstva in turizma). Prvi del sestavljajo podatki po področjih (prebivalstvo, izobraževanje, zaposlenost, življenjska raven, gostinstvo …) za celotno državo, v nadaljevanju so podatki urejeni po statističnih regijah, občinah in v zadnjem delu so podatki po državah.

Publikacija se izdaja v knjižni in v elektronski obliki. Zaradi vedno večje dostopnosti interneta se povečuje uporaba statističnega letopisa v elektronski obliki. Ogled na spletni strani:
http://www.stat.si/
http://www.stat.si/publikacije/pub.asp
http://www.stat.si/publikacije/pub_letopis_prva.asp
http://www.stat.si/letopis/index_vsebina.asp?poglavje=25&leto=2007&jezik=si

Naslovi nekaterih tabel s spletne stani iz poglavja 25 - Turizem:

25.1 Sobe in ležišča po vrstah nastanitvenih objektov in po vrstah krajev, 31. 8.
25.2 Prihodi in prenočitve turistov
25.3 Povprečna doba bivanja po vrstah krajev
25.4 Prihodi turistov po državni pripadnosti
25.5 Prenočitve turistov po državni pripadnosti
25.6 Prihodi in prenočitve turistov po vrstah krajev

Mesečno se izdajajo Mesečni statistični pregledi, ki vključujejo mesečne in četrtletne podatke raziskav.

Vsi bistveni podatke o državi Sloveniji se najdejo v publikaciji Slovenija v številkah, ki je tudi dostopna na spletnem naslovu: http://www.stat.si/publikacije/pub_slovenija.asp…

Izdajajo se tudi rezultati posameznih raziskav za različna področja, npr. letni pregled gibanja cen, statistične informacije in statistični podatki po občinah in regijah itd., ki so dostopni na naslovu: http://www.stat.si/publikacije/pub_regije.asp

Dobro poznavanje ustanov in publikacij, ki jih izdajajo, je koristno, saj nam skrajša čas pridobivanja podatkov in s tem zniža stroške. Zaradi načina njihovega pridobivanja in strokovnosti pri obdelavi je verjetnost, da so podatki dobri, zelo velika.

Naloge:

1. *Oglejte si različne publikacije v knjižni obliki.*
2. *Poiščite spletno stran Statističnega urada RS in si oglejte vsebine.*
3. *Poiščite na spletu publikacije, ki jih izdaja statistični urad, ugotovite razlike!*
4. *V knjižnici si oglejte Statistični letopis.*
5. *Poiščite publikacije, ki jih izdajajo drugi: občina, lokalna turistična organizacija, podjetja….*

2 TEMELJNI POJMI

Za razumevanje in pravilno uporabo statističnih podatkov in sestavo informacij je potrebno poznavanje osnovnih pojmov in statistično terminologijo. Če želimo analizirati uspešnost gostinskega obrata, moramo zbrati ustrezne podatke. Za statistično proučevanje se odločimo z namenom, da bomo spoznali značilnosti pojava, ki nas zanima. V turizmu in gostinstvu nas zanimajo naši gostje. Želimo ugotoviti, od kod prihajajo, kdaj, zakaj, koliko časa ostajajo itd. Vse te informacije potrebujemo za naše pravilne poslovne odločitve o vrsti ponujenih storitev, o višini cen, o odpiralnem času itd.

Zato, da bi odkrili značilnosti pojava, ki nas zanima, moramo najprej opredeliti, kaj je predmet našega proučevanja. Skupino enot, ki jo opredelimo krajevno, časovno in vsebinsko in bo predmet našega preučevanja, imenujemo **populacija**.

Tako bi bila lahko populacija, ki jo proučujemo:
- v hotelu: nočitve v določenem mesecu ali letu, ali gostje v določenem obdobju, ali nemški gostje na določen dan …
- v turistični agenciji: turisti na določenem potovanju, prodane karte določenega izleta, organizirana potovanja v nekem obdobju …
- v restavraciji: prodani meniji na dan, prihodki od prodaje v obdobju, naročene jedi v določenem obdobju …

Natančna opredelitev populacije je zelo pomembna in je odvisna od namena proučevanja.

Najlažje je populacijo opredeliti **časovno**, saj določimo čas opazovanja z datumom, uro ali obdobjem mesec, leto. V hotelu bi opazovali število gostov po dneh, mesecih, letih.

Krajevno opredelimo populacijo z določitvijo geografskih območij ali državnih meja. Pri tako opredeljeni populaciji lahko prihaja do dvoumnosti, saj včasih meje niso jasno določene. Zato je večkrat potrebna dodatna opisna opredelitev, da ne prihaja do dvoumnosti.
Primer: Turist prihaja iz Nemčije (tam živi), rojen je bil v Avstriji in ima Slovensko državljanstvo.

Vsebinska opredelitev je odvisna od namena proučevanja populacije. Določitev je najtežja, saj je potrebno dobro poznavanje pojava in pojmov. Npr. če proučujemo turiste, je potrebno poznati pojem turisti.

Razmislite, kako bi populacijo opredelili v gostinstvu in turizmu!

Populacijo sestavljajo **enote.** Vsaka enota ima značilnosti, ki jih proučujemo. Katere značilnosti bomo proučevali, določimo z opredelitvijo namena raziskave. Enote, ki bi jih proučevali v hotelu, so npr.: gostje na določen dan, ali nemški gostje v mesecu, ali gost, ki prenočujejo 1 noč, stalni gost, gost, ki prihaja z namenom obiska golf turnirja itd.

Enote, ki sestavljajo populacijo, se med seboj razlikujejo po mnogo lastnostih. Gostje v hotelu se razlikujejo po spolu, starosti, narodnosti, namenu, času in trajanju obiska itd. Zato je nujna natančna opredelitev lastnosti enot, ki bodo predmet našega proučevanje. Te lastnosti, ki jih bomo opazovali, imenujemo **spremenljivke**. Določimo jih glede na cilj in namen raziskovanja. Lastnosti, ki nas zanimajo, so npr.: spol, starost, narodnost, namen obiska, čas bivanja, pogostost obiska, itd. Npr. v hotelu bi želeli izbrati pravo glasbo za goste, zato nas zanima starost, narodnost, namen obiska gostov … Glede na te lastnosti bi sprejeli boljšo odločitev o zvrsti glasbe, ki bi jo ponudili gostom (narodno-zabavna, koncertni večer ali disko …).

Spremenljivka ima pri vsaki enoti določeno **vrednost**. Ta je lahko izražena s **številko** ali z **opisom.** Npr. starost gostov v hotelu - vrednost spremenljivke je določena s številko, z letnico rojstva ali leti starosti in ima lahko veliko vrednosti.

Številske vrednosti so lahko **zvezne**, če ima vrednost lahko poljubno vrednost, ali **diskretne,** kjer je vrednost izražena le v celi vrednosti. Npr. račun, ki ga plača gost za prenočitev (zvezna številska spremenljivka, saj lahko zavzema tudi manjšo vrednost od 1 EUR). Gostje v hotelu lahko prenočujejo 1 noč ali 2 noči ali več. Ena oseba je en gost ne glede na težo, starost (diskretne številske vrednosti, saj lahko zavzemajo le cele vrednosti).

Vrednost spremenljivke je lahko določena opisno in ima lahko le 2 vrednosti, npr. gostje v hotelu po spolu (M/Ž), ali pa veliko vrednosti, npr. narodnost, državljanstvo, izobrazba …

Značilnosti opisnih spremenljivk imajo tudi spremenljivke, ki so izražene s številkami, npr. telefonske številke, davčna številka, št. bančne ali kreditne kartice. Te vrednosti so izražene s številkami in ne s števili, ki bi jih statistično obdelovali. Iz njih ne moremo izračunati parametrov.

Primerjajte in ugotovite razlike med temeljnimi pojmi, ki ste jih spoznali pri statistiki in pri ekonomskih predmeti!

Parametri so lastnosti populacije, ki jih s proučevanjem enot želimo ugotoviti. Osnovni namen in cilj statističnega proučevanja je ugotavljanje parametrov pojava. Enote populacije proučujemo zato, da bi na osnovi njihovih lastnosti odkrili značilnosti celotne populacije, kar je pomembno za naše poslovne odločitve. Mere, s katerimi izražamo in merimo te lastnosti, imenujemo statistični parametri. Zelo poznani parametri so: minimum, maksimum, povprečje, število enot itd.

Npr. v hotelu želijo ugotoviti, kako dolgo se zadržujejo gostje, zato na osnovi podatkov, pridobljenih pri vsaki enoti populacije (gostu), izračunajo povprečno dobo bivanja. Ugotavljamo tudi zasedenost, število gostov, število nočitev, narodnost gostov, namen obiska itd.

Brez poznavanja osnovnih pojmov, kot so populacija, enota, spremenljivka, parametri, ne moremo pravilno zbrati in obdelati statističnih podatkov. Ker so kakovostni podatki bistveni za kakovost končnega rezultata statistične analize, je dobro poznavanje pojmov izrednega pomena.

*Naloge***:**

1. Na spletu poiščite podatke, določite populacijo, enote, parametre, spremenljivke.
2. V statističnem letopisu – poglavje turizem – izpišite tabele in določite: spremenljivke, rednosti (številske, opisne).
3. Izberite primere, kjer so številske vrednosti določene zvezno ali diskretno.
4. Za izbrani primer ugotovite, kateri parametri bi se lahko izračunali in razložite zakaj.

3 ZBIRANJE, OBDELAVA IN PRIKAZOVANJE PODATKOV

Za vsak pojav, ki je predmet proučevanja, je potrebno pridobiti podatke. Ker je zbiranje podatkov zamudno, strokovno zahtevno in povezano s stroški, naj vsakdo, ki podatke za statistično obdelavo potrebuje, pred pričetkom dela razmisli, ali je to delo sploh potrebno. V poglavju 1 so omenjene institucije, ki podatke zbirajo in objavljajo.

Za zbiranje podatkov se odločimo šele, ko je natančno določen namen preučevanja in nedvoumno krajevno, časovno in vsebinsko opredeljena populacija in enote. Zato je poznavanje osnovnih pojmov, razloženih v poglavju 2, nujno.

Namen obdelave je preglednost zbranih podatkov, ki bi omogočili hitro ugotovitev lastnosti preučevane populacije. V fazi zbiranja je bilo pridobljenih veliko podatkov, saj je bila opazovana populacija, ki jo sestavlja veliko enot, ki so bile opazovane po več spremenljivkah, ki imajo lahko veliko vrednosti. Npr. v hotelu je na dan 200 gostov, ki prihajajo iz različnih držav, ostajajo različno število dni, namen obiska je različen.

Zaradi preglednosti je izbira oblike prikazovanja podatkov zelo pomembna. Običajno se oblike dopolnjujejo, ker ima vsaka svoje prednosti in pomanjkljivosti. V analizah in poročilih zaradi večje nazornosti in hitre berljivosti besedilu dodajamo tabele in grafe.

3.1 ZBIRANJE PODATKOV

Statistični raziskovalec mora pred pričetkov pridobivanja podatkov preveriti, ali niso podatki že zbrani in bi jih lahko iz objavljenih statističnih virov povzel. V gostinstvu in turizmu so podatki zbrani v različnih statističnih in drugih publikacijah. Zbirajo in objavljajo jih tudi podjetja v internih in drugih glasih in tudi v številnih evidencah, ki jih vodijo v podjetju (v recepciji, računovodstvu ...). Zato je nujno pred pričetkom zbiranja podatkov poznati in preveriti **sekundarne vire.** Šele, če teh virov ni na razpolago ali niso zadostni, se odločimo, da podatke zbiramo sami. Tako zbiranje podatkov imenujemo **primarni vir.**

Če je zbiranje podatkov nujno, se je potrebno odločiti o organizacijsko-tehnični izvedbi pridobivanja, obdelave in prikazovanja le-teh. Nujno je predvideti tudi stroške raziskovanja in razpoložljiva finančna sredstva.

Raziskava, pri kateri uporabljamo statistične metode, ima več faz dela: opredelitev raziskovalnega problema, načrtovanje raziskave, uresničevanje načrta in pisanje poročila o raziskavi. Npr. hotel v recepciji zbira podatke o številu gostov in nočitev za vsak dan. Na osnovi zbranih podatkov lahko dobimo odgovore na številna vprašanja, kot so: kakšna je splošna značilnost teh podatkov, kakšni so v povprečju, kako so raztreseni okoli povprečja, kakšno mesto ima vsak podatek posebej in med drugimi (je večji, manjši, koliko), kakšen je odnos navedenih rezultatov do drugih rezultatov (istovrstnih ali drugih)? Za pridobitev teh in podobnih odgovorov je potrebno zbrane podatke urediti, obdelati in prikazati.

Kot v gostinstvu ni dobre jedi brez kakovostnih surovin, tudi v statistiki ni dobrih informacij brez kakovostnih podatkov. Zato je faza zbiranja podatkov zelo pomembna. Zbrati je potrebno **popolne** in **pravilne** podatke za vse opazovane enote populacije. Zbiramo lahko podatke pri vseh enota populacije **(popolno)** ali le za del enot **(delno).**

Ker je v podjetju običajno število enot, ki sestavljajo populacijo manjše se velikokrat odločajo za popolno opazovanje, ki ga izvedejo s **popisom** ali sprotnim spremljanjem dogodkov ali **registracijo** (npr. prijava gosta v recepciji).

Če je število enot, ki sestavljajo populacijo, veliko ali gre za dolgo časovno obdobje, se odločamo tudi za delno opazovanje, ki ga imenujemo **vzorčenje** (anketiramo goste hotela v določenem obdobju). Parametre le ocenimo, saj podatki niso pridobljeni pri vseh enotah. Vzorci so lahko **slučajni,** če so izbrani po načelu enake verjetnosti za izbiro (vsak deseti obiskovalec), ali **neslučajni**, če jih izbiramo po subjektivni presoji, kot je določitev tipičnih predstavnikov (predstavniki gostov po starosti, enako število žensk in moških).

Razmislite, kdaj bi se odločili za posamezno obliko in zakaj? Poiščite odgovore pri predmetu Poslovno sporazumevanje in vodenje!

Podatke lahko opazuje in pridobiva **neposredno** oseba, ki organizira zbiranje podatkov, ali **posredno**, ko to delo organiziramo preko drugih (Košmelj, 1993).

Razmislite, kakšne so prednosti in slabosti posamezne oblike opazovanja? Poiščite odgovore pri predmetu Poslovno sporazumevanje in vodenje!

Pred pričetkom opazovanja je potrebno pripraviti **vprašalnike**. Ti so odvisni od namena raziskav in načina obdelave. Upoštevati moramo sposobnosti in pripravljenosti tistih, ki jih izpolnjujejo. Vprašanja naj bodo kratka in razumljiva, ne prezahtevna in ne preveč osebna. Vprašalnik ne sme biti preveč dolg. Pomembna je izvedba v primernem **času**. Vprašalnik mora vsebovati navodila za izpolnjevanje in pojasnjen cilj in namen raziskave, s katero mora biti anketiranec seznanjen.

Razmislite, kako bi neprimeren čas opazovanja lahko povzročil napake!
Poiščite več informacij o oblikovanju vprašalnikov pri predmetu Poslovno sporazumevanje in vodenje!

Za statistično obdelavo so najprimernejši vprašalniki v obliki statističnih **obrazcev**. Obrazci se izpolnjujejo za vsako enoto. Na zastavljena vprašanja se odgovarja z obkroževanjem, vpisovanjem v za to vnaprej pripravljena polja, izbiranjem med več možnimi odgovori itd. Možno jih je hitreje izpolnjevati, obdelovati in kontrolirati napake. Odgovori so oblikovani na različne načine. Možni so odgovori da in ne, obkroževanje številk pred vprašanji, opisni odgovori itd. Prednosti imajo odgovori, ki dajejo številske podatke, ki jih je lažje obdelovati in iz njih izračunati parametre.

Razmislite, kakšne so prednosti in slabosti obrazcev! Kdaj bi se odločili za tako obliko in zakaj?

Zaradi pomanjkljivih navodil, nerazumljive ali dvoumne opredelitve populacije, nejasnih vprašanj, površnosti pri pripravi in izpolnjevanju vprašalnika, nepravilnem času opazovanja itd. lahko prihaja do številnih **napak**. **Slučajne** napake niso problematične, saj se učinek pri velikem številu enot izravna. Nepravilne rezultate dobimo, če nastanejo **sistematične** napake, ki so posledica napačno oblikovanih vprašanj, ki vodijo do napake, ki ima pri vseh enotah enak učinek.

Da bi bilo statistično opazovanje uspešno, je potrebna skrbna priprava. Zbiranja podatkov je temelj za kakovostne statistične informacije zato je nujna dobra organizacija dela, iskanje podatkov iz sekundarnih virov, ki jih je potrebno poznati in pridobiti, zbrati iz primarnih virov z opazovanjem celotne populacije ali vzorčenjem, sestaviti ustrezne vprašalnike, izvesti zbiranje podatkov s sposobnimi popisovalci, v ustreznem času in kraju in pravočasno odkrivati in preprečevati napake.

Naloge:

1. Navedite primer napake zaradi nejasno opredeljene starosti anketiranca.
2. Navedite primere dvoumno opredeljene populacije (gostje, turisti, narodnost, državljanstvo, starost ...).
3. Oblikujte primer vprašalnika.
4. Analizirajte primer vprašalnika! (Ugotovite dobro opredeljena in slaba vprašanja.)
5. Na spletu najdite in poglejte oblike statističnih obrazcev.
6. Poiščite primere statističnih obrazcev na praksi.

3.2 OBDELAVA PODATKOV

Zbrane **podatke** moramo **urediti** in **prikazati** v pregledni obliki. Urejamo jih lahko na različne načine. Lahko jih razporejamo v **ranžirne vrste** ali v **skupine** po vrednosti določene spremenljivke. Če ima spremenljivka manjše število vrednosti, lahko za vsako vrednost opredelimo svojo skupino (gostje po spolu: M, Ž - dve skupini). Če ima spremenljivka preveč možnih vrednosti in bi bilo število skupin preveč veliko, kar bi zmanjšalo preglednost, takrat skupine združujemo po sorodnosti. Npr. za starost gostov ne določamo skupine za vsako leto starosti, ampak za daljša obdobja (do 18. let, od 19 do 29 ali daljše obdobje). Paziti je potrebno na določitev primernega števila skupin, ki še omogoča preglednost in še pridejo do izraza značilnosti proučevanega pojava.

Pri opredeljevanju skupin morata biti izpolnjena dva pogoja:
- **enoličnost** skupine, ki je dosežena tako, da so meje skupine določene tako jasno, da ni nikoli dvoma, kam spada posamezna enota,
- **enovitost (homogenost)** skupine, ki je odvisna od cilja, saj naj skupina vključuje sorodne enote.

Razmislite, ali je opredelitev skupine »mladi« dobra in zakaj?

Pri **opisnih spremenljivkah** opredelitev skupin včasih ni enostavna. Če nas zanima spol hotelskih gostov, je opredelitev skupin lahka in zagotovitev enoličnosti in enovitosti enostavna, saj sta možnosti le dve in so meje jasno določene.

Če nas zanima, od kod prihajajo hotelski gostje, skupine lahko določamo po narodnosti, državljanstvu, državi rojstva, državi stalnega bivališča itd. Ti podatki se pri posamezni enoti lahko med seboj razlikujejo in ob nenatančni opredelitvi skupin lahko prihaja do napak, ki prikažejo nepravo sliko značilnosti proučevanega pojava. Veliko težav je pri določanju skupin, če želimo pridobiti podatke o gostih glede izobrazbe, zaposlitve, poklica itd.

Če so vrednosti spremenljivke povezane s časom, npr. čas bivanja, leto rojstva, določamo **skupine po časovnih enotah,** npr. datum, število dni, mesec, leto ali daljše obdobje.

Pri **številskih spremenljivkah skupine imenujemo razredi**. Zagotovitev enoličnosti je lažja, saj meje postavimo s številkami. Upoštevati moramo posebnosti, ki so odvisne od vrste številske spremenljivke. Poznavanje temeljnih pojmov o zveznih in diskretnih vrstah številskih spremenljivk je nujno, da pri določanju mej ne pride do napak, ki bi povzročile neizpolnitev pogoja enoličnosti skupine. Določitev meja je odvisna tudi od načina zaokroževanja številskih vrednosti. **Zaokrožujemo** jih lahko na najbližjo celo vrednost ali na najvišjo celo vrednost. Pri cenah in vrednostih, izraženih v denarnih enotah, običajno uporabljamo zaokroževanje na najbližjo celo vrednost. Številke do 5 zaokrožimo navzdol in nad 5 navzgor. Ta način zaokroževanja se običajno uporablja v matematiki. Pri starosti gostov uporabljamo zaokroževanje na najvišjo celo vrednost, to je vedno navzdol z dopolnjeno starostjo.

Preverite načine zaokroževanje števil pri predmetu matematika!

Pri določitvi števila razredov izhajamo iz enovitosti razredov. Število je odvisno od cilja raziskave. V razred so vključene vse enote, ki izpolnjujejo pogoje, določene z **zgornjo** in **spodnjo mejo** razreda. Razliko med mejami imenujemo **širina razreda**.

Pravilno oblikovanje razredov je pogoj za frekvenčno porazdelitev, ki je obravnavana v poglavju 6.

Vse enote, ki sestavljajo celotno populacijo, razporedimo po kriterijih, ki smo jih določili v skupine ali razrede. Tako razvrstitev skupin imenujemo **statistična vrsta**. Če so podatki razvrščeni po krajevni, časovni ali stvarni opredelitvi opazovane populacije, dobimo krajevne, časovne ali stvarne vrste. Pri časovnih vrstah se podatki lahko nanašajo na posamezni trenutek (**momentne** vrste) ali časovno obdobje (**intervalne** vrste).

Npr. če bi hotelske goste razporedili po državah od koder prihajajo, dobimo krajevno vrsto, če jih razporedimo glede na razlog prihoda ali izobrazbo, oblikujemo stvarno vrsto, če se podatki o številu gostov nanašajo na datum, so momentne časovne vrste, če spremljamo nočitve po mesecih, oblikujemo časovne intervalne vrste.

Razmislite, kako bi na primeru hotelskih gostov opredelili skupine, razrede! Kako bi zagotovili enoličnost in enovitost? Koliko skupin bi predlagali za posamezne primere?

Podatke celotne populacije obdelujemo individualno za vsako enoto ali po skupinah. **Ročno obdelavo** podatkov je danes nadomestila uporaba sodobne informacijske tehnologije. Včasih so podatke ročno prepisovali in s črtkanjem vnašali v vnaprej pripravljeno obdelovalno tabelo. Tabelo so sestavljale vse spremenljivke in njihove vrednosti. Z vnosom črtic ali pik v ustrezno polje so po končanem vnosu ugotavljali predvsem število enot.

Za obdelavo danes uporabljamo **sodobno informacijsko tehnologijo**, ki omogoča hitro in natančno vnašanje podatkov, zbiranje, razporejanje teh v skupine in izračunavanje različnih parametrov. Prednosti so v hitrosti vnosa, saj lahko uporabljamo tudi optično čitanje podatkov, kar omogočajo tudi tako pripravljeni statistični vprašalniki. Poveča se tudi natančnost vnosa, razporejanja v skupine in obdelava, zato tudi zaradi sprotne kontrole nastaja manj napak. Velike prednosti obdelave so tudi zaradi možnosti kopiranja in shranjevanja. Že oblikovane podatke lahko preoblikujemo in uporabimo v novih izračunih.

Za obdelavo lahko uporabljamo **različne obstoječe računalniške programe**. Uporaba je odvisna od obsega statistične obdelave. Kjer so potrebne obsežnejše statistične obdelave, uporabljamo specializirane statistične programe. Za obdelavo in prikazovanje podatkov za potrebe manjših podjetij zadošča uporaba programa Excel in podobnih programov preglednic. Ti programi so cenovno dostopni uporabnikom in je pridobitev znanja uporabe enostavno.

Za uporabnika informacij je zelo pomembno, da so podatki, ki so zbrani, tako obdelani, da ohranijo pravilnost in pridobijo preglednost, ki omogoča hitro analizo za dobre poslovne odločitve. Zato je potrebno znanje in uporaba sodobne informacijske tehnologije, poznavanje osnovnih statističnih pojmov, načinov razporejanja enot v skupine ali razrede ter postopkov oblikovanja pravilnih statističnih vrst.

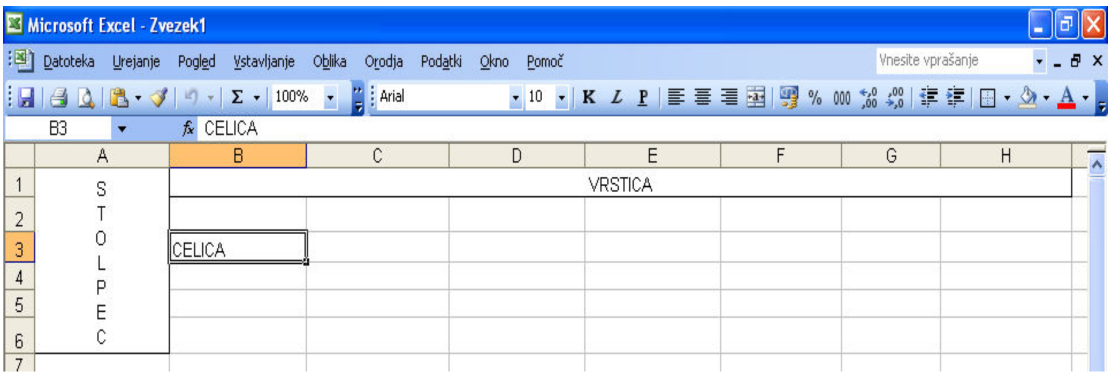
Naloge:

1. Za primer izbranih podatkov iz statističnega letopisa analizirajte statistično vrsto in ugotovite, kako je oblikovana, kako so postavljene meje! Kakšne vrste bi še lahko oblikovali iz teh podatkov?
2. Kakšne parametre lahko ugotovimo iz primera statistične vrste?
3. Na spletu pridobljene ali na praksi zbrane podatke razporedite v skupine, razrede in oblikujte statistično vrsto (krajevno, stvarno, časovno).
4. Ugotovite, ali so skupine enolične in enovite, obrazložite zakaj.
5. Razložite kako so določene meje razredov?
6. Analizirajte kakšni so podatki v časovni vrsti (momentni, intervalni)?

3.3 PRIKAZOVANJE PODATKOV

3.3.1 Tabele

Največkrat podatke prikazujemo v tabelah, ki jih ne sestavljamo ročno, ampak za oblikovanje uporabljamo sodobno **informacijsko tehnologijo** in programe, kot so **Excel** in drugi, ki so namenjeni prikazovanju podatkov v tabelah in računanju z njimi. Program omogoča enostavno vnašanje podatkov v polja, ki so označena s stolpci in vrsticami.



Slika 1: Excel
Vir: Microsoft Office Excel 2003

Glede na obseg podatkov in cilj raziskave so tabele različno velike. Imajo lahko veliko stolpcev in vrstic. Pri ročno oblikovanih tabelah najprej tabelo oblikujemo in nato vanjo vnašamo želene podatke. Vrsti red dela pri računalniško pripravljenih tabelah je obraten. Podatke najprej vnesemo, nato jih oblikujemo, določamo velikost in postavitev tabele v poročilo.

Ponovite pravila oblikovanja tabel v programu Excel!

Prednost računalniško pripravljenih tabel je tudi možnost kopiranja dela ali celotne tabele, enostavno popravljanje, manjšanje ali povečevanje obstoječe tabele, hranjenje in enostavno grafično prikazovanje podatkov. Pri vnašanju podatkov, oblikovanju, kopiranju in shranjevanju moramo poznati in upoštevati pravila, ki jih določa program, s katerim delamo. Zato mora vsak uporabnik dobro poznati izbrani računalniški program.

Tudi računalniško pripravljena tabela mora imeti **bistvene sestavine**:
Naslov, iz katerega je razvidno, kaj v tabeli prikazujemo, ki naj bi bil jasen in kratek.
Vir, ki pojasni, od kod smo pridobili podatke. Vir je lahko dodan naslovu ali ga dodamo pod tabelo.
Glavo tabele, ki jo predstavlja prva vrstica in pove vsebino podatkov v stolpcih.
Čelo tabele, ki je prva kolona in daje informacije o vsebini podatkov v vrsticah.
Zbirni stolpec in **zbirno vrstico** ki sta v statističnih tabelah prva vrstica in prvi stolpec.
Merske enote, denarne enote, zaokrožene številke (npr. v tisoč) in podatki, ki se ponavljajo, se navedejo nad tabelo.
Pod tabelo so lahko dodani znaki, ki pojasnjujejo posamezne podatke v tabeli (npr. podatek je ocenjen, ni podatka …).

Tabela ali preglednica, kot pove že beseda, mora bralcu prikazati podatke tako **pregledno**, da lahko hitro pride do informacije o značilnosti proučevanega pojava. Zato tabela **ne** sme biti **prevelika** (število stolpcev in vrstic največ v velikosti vidnega polja ekrana). Vpisane **številke** naj ne bodo prevelike (odvisno od potrebe po natančnosti podatkov), zato jih večkrat **zaokrožujemo**. V polja **ne** vpisujemo podatkov, ki se **ponavljajo**, kot so merske enote, denarne enote, ki jih vpisujemo nad ali v glavo tabele. Tabeli preglednost povečuje tudi oblikovanje, zato je pomembno, kako **oblikujemo** številke (enako število decimalnih mest, podpisovanje), **glavo in čelo**. Širino kolon in vrstic prilagajamo vnesenim podatkom in velikosti tabele. Pomembnejše podatke lahko poudarimo ali obarvamo. Tabeli povečamo preglednost in zmanjšamo verjetnost napake pri prebiranju tudi z **risanjem črt**, ki ločujejo kolone in vrstice.

	GLAVA TABELE			
ZBIRNA VRSTICA →	ZBIRNI STOLPEC↓			
		↓		
VRSTICA →		S		
		T		
Č		O		
E		L		
L		P		***polje***
O		E		
		C		

Slika 2: Tabela
Vir: Lastni

Glede na število statističnih vrst in število spremenljivk, ki jih prikazujemo v eni tabeli, so tabela lahko:

-**enorazsežne ali enostavne**, če prikazujejo le eno statistično vrsto in so podatki prikazani za eno spremenljivko, npr. gostje hotela po državah,

-**sestavljene**, ki prikazujejo več statističnih vrst hkrati in so vse oblikovane po vrednosti iste spremenljivke, npr. gostje hotela in nočitev po državah od koder prihajajo – ista krajevna vrsta, kjer sta dve spremenljivki - nočitve in število gostov (lahko bi tabelo razstavili na dve ločeni enostavni tabeli),

-**večrazsežna ali kombinacijska** tabela, kjer statistično vrsto prikazujemo po več spremenljivkah hkrati (Košmelj, 1993). Npr. podatke o številu gostov prikazujemo hkrati po dveh spremenljivkah, kot sta domači in tuji. Goste v hotelu prikazujemo lahko po več spremenljivkah, npr. določimo več starostnih skupin. Značilnost teh tabel je možnost seštevanja podatkov, npr. gostje po spolu (moški, ženske, skupaj).

Ker so tabele najbolj pogosto uporabljena oblika prikazovanja statističnih podatkov, je poznavanje pravil sestavljanja zelo pomembno. Pravilno sestavljena in dobro oblikovana tabela omogoča hitro pridobitev dobrih in natančnih podatkov in informacij. Vsako poročilo ali analiza vsebuje tudi tabelarični prikaz, saj tabela omogoča prikazovanje podatkov poljubno točno in celovito, odvisno od potrebe in namena poročila. Zato je dobro poznavanje vsebine nujno.

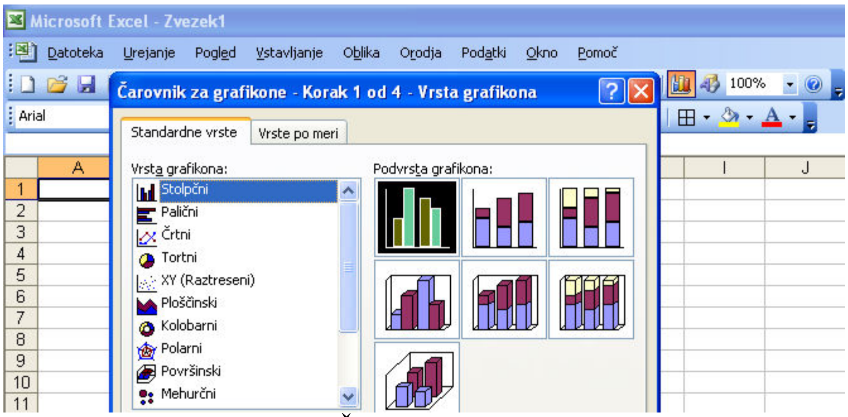
Naloge:

1. Ponovite pravila vnašanja podatkov, oblikovanja tabel, popravljanja, kopiranja in shranjevanja tabel v programu Excel.
2. Oglejte si tabele v statističnem letopisu in ugotovite vrsto.
3. Poiščite primer tabele in analizirajte, ali ima vse potrebne sestavine (glava, čelo ...).
4. Sestavite in oblikujte enostavno, sestavljeno in kombinirano tabelo.

3.3.2 Grafikoni

Grafikoni so dopolnitev prikazovanja podatkov s tabelami. Z njimi lahko lastnosti opazovanih pojavov poudarimo. So manj točni kot prikazi v tabelah, vendar zaradi nazornosti dobra osnova za hitro analizo. Bralec poročila jih hitro opazi, ker so enostavni in na pregleden način posredujejo osnovne informacije. Zato so nujni sestavni del vsakega poročila in poslovnega načrta. Da bi dosegli ta cilj, moramo pridobiti znanje za pravilno oblikovanje grafikonov.

Grafikonov ne rišemo več ročno, ampak za oblikovanje uporabljamo sodobno računalniško tehnologijo in različne programe. Poznati je potrebno pravila za risanje in oblikovanje grafikonov za izbrani program. V programu Excel, v katerem izdelamo tabelo, je risanje grafikonov enostavno.



Slika 3: Čarovnik za grafikone
Vir: Microsoft Office Excel 2003

Za pravilno izbiro vrste grafikona je potrebno teoretično znanje statistike, saj vsak grafikon ni primeren za prikazovanje vseh podatkov. Na izbiro ustrezne vrste grafa vpliva cilj raziskave.

Pred risanjem grafikona moramo pravilno **pripraviti podatke** v tabeli, da si delo olajšamo. V grafikonu lahko prikažemo vse podatke iz tabele ali le del, ki jih želimo analizirati. Zaradi nazornosti je potrebno paziti, da prikazanih **podatkov** v enem grafikonu **ni preveč**. Zato se večkrat odločamo, da vseh podatkov ne prikažemo v enem, ampak jih prikazujemo v več grafikonih. Pri risanju grafa je potrebno paziti na **oblikovanje**, ki prikaže dejansko značilnost proučevanega pojava. Zaradi nepoznavanja statistične teorije in pravil risanja grafov ali namerno lahko prikažemo nerealno značilnost pojava.

Vsak grafikon mora imeti:
- **naslov,** iz katerega je razvidna vsebina prikaza,
- **legendo,** če je v tabeli prikazanih več podatkov,
- izbrana mora biti **prava vrsta grafa** in
- prikazani **pravi podatki**.

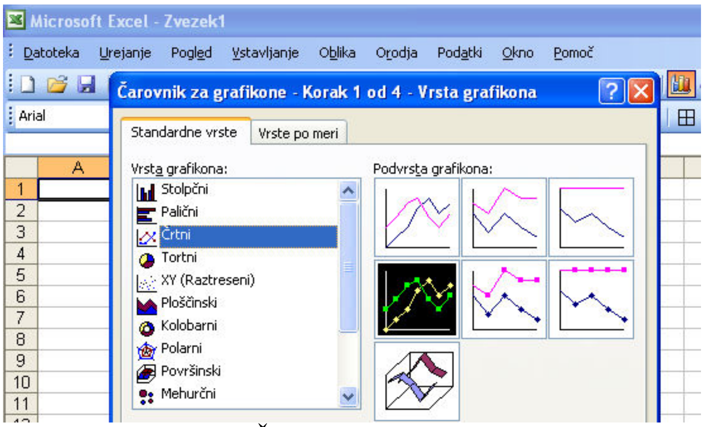
Obnovite pravila risanja grafov v programu Excel in statistično teorijo poglavja 2 in 3.

Poznamo več **vrst grafikonov**. V statističnih obdelavah najpogosteje uporabljamo linijske grafikone, prikaze podatkov s stolpci in krogi. V gostinstvu in turizmu se pogosto uporabljajo tudi polarni grafikoni. Državna statistika pogosto prikazuje podatke tudi s kartogrami in prikazi s figurami.

Oglejte si grafikone v Statističnem letopisu!

Linijski grafikon je prikaz v pravokotnem koordinatnem sistemu s točkami, ki so povezane v daljice. Obe osi, abscisna in ordinatna, morata vsebovati oznake podatkov, ki se prikazujejo, in merske enote. Primerni so za prikazovanje številskih in časovnih statističnih vrst. Prednost grafa je v nazornosti prikaza nihanja ali gibanja pojava. Če je v grafu prikazanih več statističnih vrst, omogočajo tudi primerjavo. Primerni so tudi za risanje trendne črte, ki prikazuje smer gibanja.

Paziti je potrebno na izbiro pravilnih intervalov na obeh oseh, da ne pride do napačnega prikaza, saj različno merilo lahko vpliva na prikazano velikost rasti, npr. izbrani ožji intervali na abscisni osi prikazujejo manjšo rast istega pojava, kot če izberemo večje. Pri trenutnih (momentnih) vrstah podatke vnašamo nad točke, na katere se podatki nanašajo, na razmičnih (intervalnih) vrstah so podatki vpisani na sredini razmika.

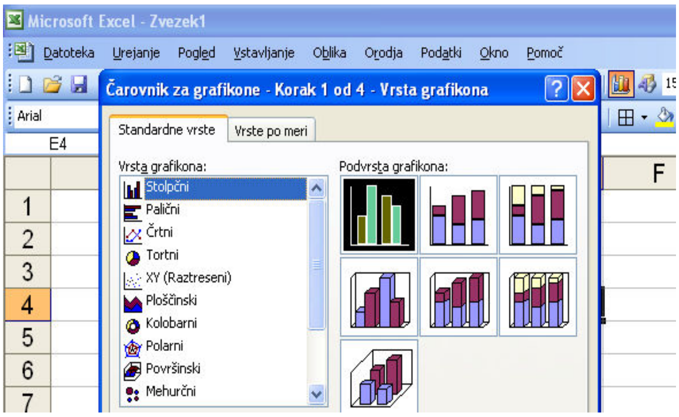


Slika 4: Črtni (linijski) grafikoni
Vir: Microsoft Office Excel 2003

Na vajah so z računalnikov v programu Excel prikazani primeri risanja pravilnih linijskih grafov. Pred vključitvijo narisanega grafa v poslovno poročilo mora biti ta tudi ustrezno oblikovan. Če je v grafu prikazanih več linij, je legenda nujna. Zaradi preglednosti večkrat vrišemo tudi mrežne črte, poudarimo posamezne podatke, dodamo podatke iz tabele itd.

V različnih statističnih publikacijah si oglejte grafične prikaze podatkov! Primerjajte jih in ugotavljajte razlike!

Prikazi s stolpci se uporabljajo za prikazovanje opisnih, predvsem krajevnih, statističnih vrst. Včasih jih uporabljamo tudi za časovne vrste, predvsem, če želimo poudariti primerjavo in manj gibanje.



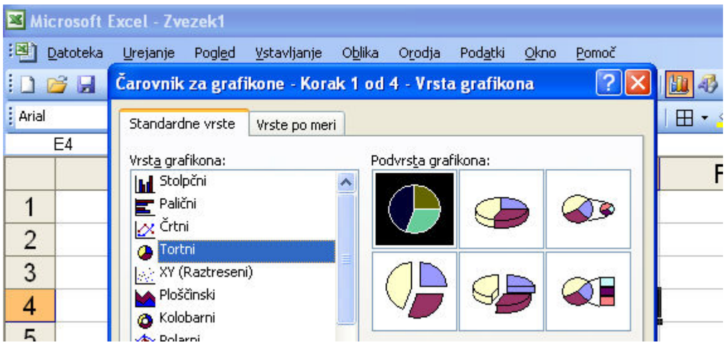
Slika 5: Stolpčni grafikoni
Vir: Microsoft Office Excel 2003

V programu Excel so prikazane različne oblike stolpcev. Izbira je odvisna od namena in cilja statistične analize. Stolpci so lahko enako ali različno široki, postavljeni navpično ali vodoravno. Včasih je smiselno stolpce razporediti po velikosti. Če v istem grafu prikazujemo dve ali več statističnih vrst, ki se vsebinsko povezujejo, so lahko stolpci postavljeni drug ob drugem, lahko se tudi delno prekrivajo ali je med njimi razmik. Izberemo obliko, ki nazorneje prikaže značilnost pojava, ki je namen raziskave. Npr. če bi želeli prikazati **primerjavo** domačih in tujih gostov v hotelu po mesecih, bi izbrali stolpčni graf, kjer bi se stolpca za domače in tuje goste dotikala, da bi bila primerjava velikosti lažja. Stolpčni grafi so primarni tudi za prikazovanje sestave ali **strukture** pojava in primerjave struktur. Odvisno od cilja raziskave lahko izbiramo enako visoke ali različno visoke strukturne stolpce. Če želimo primerjati samo strukturo in zanemarimo velikost pojava, izberemo strukturne stolpce, ki so enako visoki, če želimo poudariti oboje, velikost in sestavo pojava, izberemo različno velike strukturne stolpce. Npr. če nas zanima le delež domačih in tujih gostov, ne zanima nas, koliko jih je skupaj, izberemo stolpce, ki so enako visoki.

Ob nepoznavanju statistične teorije in izbiri nepravilne vrste grafa lahko pride do napačne analize. Npr. hotel ima danes 50 gostov od tega 50 % enodnevnih in 50 % gostov, ki ostanejo več dni, naslednji dan ima 100 gostov in od tega zopet 50 % gostov ostaja 1 noč in 50 % več noči. Pri izbiri grafa, ki prikazuje samo strukturo, ne bo razvidna velikost pojava in bosta stolpca enako visoka. Če želimo poudariti velikost in strukturo, pa izberemo stolpčni graf, kjer bosta stolpca različno visoka.

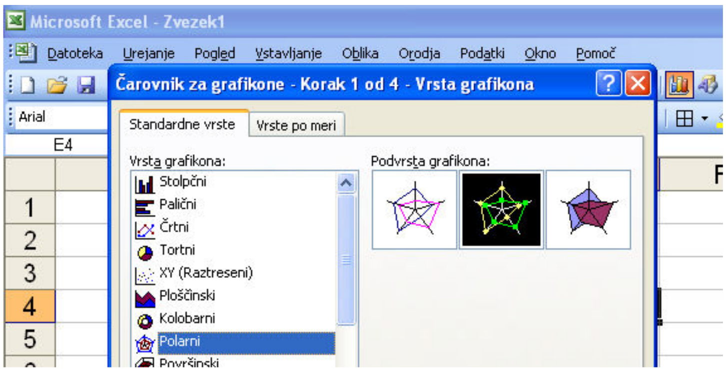
Vsak, tudi stolpčni graf, mora imeti naslov, če je potrebno, legendo in oznake podatkov.

Krožni grafi se največkrat uporabljajo za prikazovanje sestave ali strukture pojava. Krog prikazuje celoto, ki je razdeljena na izseke, ki predstavljajo posamezne dele. Večja ploščina izseka nazorno prikazuje del, ki ga sestavlja večje število enot. Manj primerni so za primerjave. Primerjave so mogoče, saj velikost kroga ponazarja velikost pojava. Ob primerjavi statističnih vrst, ponazorjenih s krogi, lahko različno veliki krogi, postavljeni drug ob drugem, ponazarjajo tudi velikost. Za primerjavo velikosti in strukture je možna izbira tudi **kolobarjev**, kjer so krogi postavljeni drug v drugem.



Slika 6: Krožni (tortni) grafikoni
Vir: Microsoft Office Excel 2003

V turizmu se pogosto uporabljajo tudi **polarni grafi**, saj so primerni za prikazovanje sezonskih nihanj. V krogu je prikazano nihanje pojava z oddaljenostjo od središča. Primerni so tudi za prikazovanje podatkov iz anket, kjer so odgovori razdeljeni v a, b, c. Polarni grafikoni imajo različno število osi glede na število prikazanih spremenljivk (Sagadin, 2003).



Slika 7: Polarni grafikoni
Vir: Microsoft Office Excel 2003

Geografska razširjenost pojava se pogosto ponazarja s **kartogrami**. Kartograme imenujemo geografske karte, v katere so vrisane krajevne statistične vrste. Če označimo število enot, ki so prikazane v obliki stolpcev, krogov ali drugih znakov, se imenujejo **diagramske** karte. Npr. zemljevid Slovenije z vrisanimi podatki o količini padavin (stolpčni prikaz), onesnaženost zraka, … Intenzivnost in gostoto pojava lahko prikazujemo na geografski karti tudi z barvami ali posebnimi znaki. Takrat govorimo o **pravem kartogramu**, npr. zemljevid Slovenije z vrisanimi občinami in z barvami označene občine. Legenda prikazuje, kaj posamezna barva predstavlja, npr. občine Slovenije - obmorski kraji bi bili označeni z isto barvo, gorski kraji z drugo barvo itd.

Možni so tudi prikazi s figurami, ki se imenujejo **piktogrami**. Obravnavane pojave predstavljamo s predmeti ali slikami. Uporabljeni znaki morajo biti vsebinsko povezani s pojavom (slike avtomobilov, oseb, telefonov …). Velikost znaka lahko ponazarja tudi velikost pojava, npr. v geografskih kartah na ta način označujejo število prebivalcev itd.

Dobro poročilo, analiza in poslovni načrt mora vsebovati ne le tekstovni del, ampak tudi pravilne in ustrezno oblikovane tabele ter prikaze značilnosti pojava z ustrezno izbranimi grafikoni, ki morajo biti pravilno oblikovani. Zato je znanje prikazovanja podatkov s tabelami in grafikoni nujna za vse, ki poslovna poročila sestavljajo ali jih uporabljajo.
Paziti je potrebno na izbiro prave vrste tabele (enostavno, sestavljeno, kombinacijsko), ki ima vse sestavne dele (naslov, glavo, čelo, zbirno vrstico, opombe) in je tako oblikovana, da je pregledna ne le tistemu, ki jo je sestavil, ampak predvsem tistim, katerim je namenjena. Natančnost prikazanih podatkov v tabelah dopolnimo s preglednejšimi prikazi v grafikonih. Paziti moramo na izbiro prave vrste grafikona (linijski, stolpčni, krožni, polarni ...), sestavne dele, pravilno prikazane podatke ter na ustrezno oblikovanje.

Naloge:

1. Iz podatkov sestavite in oblikujte v programu Excel enostavno tabelo in narišite grafikon.
2. Sestavljeno tabelo popravite, kopirajte in shranite in narišite grafikon.
3. Poiščite tabelo s podatki in analizirajte, kateri grafikoni bi bili primerni za prikazovanje in zakaj?
4. Iz podatkov tabele narišite ustrezne vrste grafikonov (v programu Excel):
- linijski grafikon, stolpčni (primerjalni in strukturni), vodoravni stolpci, polarni graf, krožni graf, analizirajte razlike.

4 RELATIVNA ŠTEVILA

Za razumevanje in ugotavljanje značilnosti pojava običajno ne zadošča le nazoren prikaz zbranih in urejenih podatkov v tabelah in grafikonih. Zbrani podatki so le osnova za statistične analize, s katerimi želimo ugotoviti lastnosti pojava.

Če v hotelu želimo informacije o zasedenosti hotela, ne zadošča, da zberemo in v tabeli in grafikonu prikažemo podatke o številu gostov v hotelu po posameznih dneh. Do informacij o zasedenosti, sezonskih nihanjih itd. pridemo, če podatke med seboj primerjamo. Boljšo sliko o naši uspešnosti dobimo, če se primerjamo tudi z drugimi. Tako kot športnik ne more pridobiti informacije o tem, kako je dober, dokler se na tekmovanju ne primerja z ostalimi, tako tudi podjetje potrebuje primerjave. Pomen posameznega podatka se za analizo pojava poveča, če ga primerjamo z drugimi podatki.

Možne se različne vrste primerjav in nekatere so prikazane v poglavju relativnih števil. Pri izbiri relativnega števila izhajamo iz smiselnosti izračunavanja. Zaradi velikih možnosti uporabe, enostavnega izračunavanja in enostavne razlage se relativna števila v statističnih analizah zelo pogosto uporabljajo. Z njimi lahko analiziramo notranjo sestavo pojava, intenzivnost pojavljanja, spremembe v času in razširjenost v prostoru. S tem razložimo pomembne lastnosti opazovanega pojava. Omogočajo tudi primerjave med istovrstnimi in raznovrstnimi pojavi.

Za sestavljanje osnovnih poslovnih poročil je uporaba znanja relativnih števil nujna. Kljub enostavnosti izračunavanja lahko prihaja do napak pri izbiri vrste relativnega števila in izračunavanju, še pogosteje pa do napačne razlage dobljenih rezultatov. Le poznavanje statističnih metod omogoča pravilno izbiro statističnega kazalca. Prednosti posameznega relativnega števila omogoča doseči cilj statistične analize.

4.1 VRSTE

Najpogosteje ugotavljamo **razliko** med pojavoma (matematično odštevamo) ali izračunavamo **razmerja** (matematično delimo).

Primerjamo lahko **istovrstne** podatke (nočitve v hotelu po mesecih) ali **raznovrstne** (število nočitev in goste, da dobimo povprečno dobo bivanja).

Spremembe pojava lahko izrazimo s **številom** ali **z %**, npr. danes imamo v hotelu 20 gostov več kot včeraj (izračunali smo razliko) ali danes imamo v hotelu za 10 % več gostov.

Razliko lahko izračunavamo le med istovrstnimi podatki. Pridobljeni rezultat je lahko pozitiven, negativen ali 0 in je izražen v isti merski enoti kot podatki proučevanega pojava. Če predznaka pri rezultatu ne upoštevamo, dobimo **absolutno** razliko.

Razmerja lahko izračunavamo med istovrstnimi in raznovrstnimi podatki. Ko primerjamo istovrstne podatke, se merske enote okrajšajo in dobimo **neimenovano** število. Pri računanju razmerij med raznovrstnimi podatki merske enote ostanejo in dobimo **imenovana** števila.

Ponovite pravila računanja, ki ste jih spoznali pri matematiki!

Ker lahko primerjamo različne podatke med seboj, lahko izračunamo različne vrste relativnih števil. Poznamo:

- **Strukture**, ki so razmerja med istovrstnimi podatki, kjer del primerjamo s celoto.
- **Indekse**, ki so razmerja med istovrstnimi podatki, pomnožena s 100.
- **Stopnje**, ki so primerjave istovrstnih podatkov, s katerimi izražamo gibanje pojava.
- **Statistične koeficiente,** ki so razmerja med raznovrstnimi podatki, ki so v vsebinski zvezi (Šadl, 2001).

4.2 STRUKTURE

V gostinstvu in turizmu nas zanimajo mnoge značilnosti opazovanih pojavov, ki se lahko izrazijo s strukturami, zato se to relativno število zelo pogosto uporablja. Zanima nas struktura gostov v hotelu po spolu, starosti, državah, namenu obiska, trajanju bivanja, stalnosti itd. Za ugotavljanje teh značilnosti ni zadosti poznati le podatke, ki jih pridobimo v recepciji, npr. število gostov, ki pri nas ostajajo več dni, ampak je zaradi primerljivosti bolje, če izračunamo strukturne deleže. Npr. več nam pove podatek, da je bilo v preteklem mesecu 50 % gostov, ki so le prenočevali, kot če bi poznali število teh gostov.

Ponovite matematična pravila računanja strukture!

Struktura ali sestava je primerjava podatka za del s podatkom za celoto. Je razmerje med istovrstnimi podatki in nam pove, kakšna je notranja zgradba ali sestava proučevanega pojava. Lahko je izražena v deležu, odstotku ali odtisočku.

$$\textbf{Strukturni delež} = \frac{\textbf{del}}{\textbf{celota}}$$

$$P_j = \frac{Y_j}{Y}$$

P_j - strukturni delež
Y_j - del
Y - celota

$$\textbf{Strukturni odstotek} = \frac{\textbf{del}}{\textbf{celota}} \times 100$$

$$P_j\% = \frac{Y_j}{Y} \times 100$$

$$\textbf{Strukturni odtisoček (promila)} = \frac{\textbf{del}}{\textbf{celota}} \times 1000$$

$$P_j‰ = \frac{Y_j}{Y} \times 1.000$$

Če vse strukturne deleže seštejemo, dobimo celoto. Vsota strukturnih deležev je 1, vsota strukturnih % je 100 % in vsota strukturnih odtisočkov je 1000 ‰ .

Najpogosteje se strukture izražajo s strukturnimi odstotki. Če so deleži zelo majhni pogosto izberemo strukturne odtisočke, npr. v odtisočkih (promilih) izražamo delež alkohola v krvi.

Na vajah se na praktičnih primerih izračunavajo strukture iz podatkov različnih statističnih vrst v programu Excel. Za izračunavanje je potrebno pravilno oblikovati formule, da se te lahko kopirajo za celotno statistično vrsto, za katero se značilnost pojava izrazi s strukturnimi deleži, % ali ‰. Ko izračunamo vsoto struktur, dobimo celoto pojava.

Če je opazovani pojav razčlenjen le po vrednosti ene spremenljivke, izračunamo **enorazsežne** strukture. Npr. v hotelu bi za goste določenega dne ugotovili države, iz katerih prihajajo in izračunali strukturo. Skupno število gostov bi predstavljalo celoto, število gostov iz posamezne države pa del. Delež bi bil največji tam, kjer bi bilo število enot največje. Tako izračunani strukturni deleži so dobra informacija o zgradbi pojava, saj hitro lahko analiziramo kateri del je največji, najmanjši in kako so med seboj različni. Če strukturne deleže prikažemo v tabeli ali grafu, je še nazorneje. Za prikazovanje bi uporabili enostavno tabelo. Podatki so primerni tudi za primerjavo nihanja pojava, npr. spremljamo, kako se giblje delež gostov iz posamezne države v nekem obdobju.

Če opazovani pojav razčlenjujemo hkrati po vrednostih več spremenljivk, izračunavamo **večrazsežne** strukture. Zaradi preglednosti najpogosteje izračunavamo strukture hkrati le za dve spremenljivki in sestavimo **dvorazsežno** strukturo. Npr. za goste hotela po državah bi hkrati ugotavljali tudi dobo bivanja in izračunavali delež gostov posamezne države, ki bivajo eno noč ali več noči. Za prikazovanje uporabljamo kombinacijske tabele.

Za grafično prikazovanje struktur najpogosteje uporabljamo **krožni grafikon** in prikazovanje s strukturnimi stolpci. V programu Excel na osnovi podatkov, prikazanih v tabeli, enostavno narišemo grafikon, ki mora vsebovati vse sestavine pravilno oblikovanega grafikona. Prednost krožnega prikazovanja je v nazornosti prikazovanja deležev, saj je jasno in enostavno prikazana celota s krogom in deleži s ploščino izseka. Večji izsek predstavlja večji delež. Sodobna tehnologija omogoča tudi grafično oblikovanje, ki še poudari značilnost, ki jo želimo predstaviti, npr. izseke poudarimo, pobarvamo, premikamo.

Prikazovanje s **strukturnimi stolpci** je zelo primerno za primerjavo struktur predvsem pri časovnih vrstah. Enorazsežno strukturo prikažemo s strukturnim stolpcem, tako da višina stolpca prikazuje celoto, ki je razdeljen na dele glede na deleže. Če želimo pojave primerjati med seboj le po strukturi, zanemarimo pa velikost pojava, izberemo prikaz s stolpci, ki so **enako visoki**, saj so vsote strukturnih deležev povsod enake. V kolikor želimo primerjati velikost in strukturo pojava hkrati, izberemo strukturne stolpce, ki so **različno visoki**, saj višina stolpca ponazarja velikost pojava, deleži pa so razvidni iz delov, na katere je razdeljen posamezni stolpec. Izbira ustreznega grafičnega prikaza je odvisna od cilja analize. Na ustrezno izbiro vpliva tudi poznavanje statistične teorije.

Pomen izračunavanja strukture pri statističnih analizah je zelo velik, saj struktura omogoča dober vpogled v notranjo zgradbo pojava. Obseg pridobljenih informacij se še poveča, če izračunamo strukture z upoštevanjem dveh spremenljivk. Izbrati moramo populacijo, ki jo proučujemo in določiti, kaj predstavlja celoto in kaj so deli. Odločiti se moramo za izbiro izračunavanja strukturnih deležev, odstotkov ali odtisočkov. Lahko izračunamo enorazsežne, dvorazsežne ali večrazsežne strukture. Možne so tudi primerjave danega pojava za različna območja in v gostinstvu in turizmu primerjave struktur danega pojava v različnih časovnih obdobjih. Tako se ugotavljajo spremembe v času in odkrivajo morebitne razvojne tendence opazovanega pojava. Za ustrezno oblikovanje in prikazovanje struktur v tabelah in grafikonih je nujno poznavanje programa Excel ali ustreznih drugih programov. Izbiramo med krožnimi in stolpčnimi grafikoni.

*Naloge:*

1. Poiščite tabelo s podatki, iz katere bi lahko izračunali strukturo.
2. Na primeru izračunajte enorazsežno in večrazsežno strukturo.
3. Izračunane strukturne deleže preračunajte v strukturne odstotke in odtisočke.
4. Strukture prikažite v krožnem grafu.
5. Strukture prikažite v stolpčnem grafu – strukturnem, kjer so stolpci enako visoki in strukturnem, kjer so stolpci različno visoki, primerjajte in ugotovite razlike.

4.3 INDEKSI

Indeksi nam dobro prikazujejo relativne spremembe pojava. Prednosti indeksov so tudi v tem, da so **neimenovana števila**, ker nam omogočajo primerjavo med raznovrstnimi podatki. Če so podatki izraženi v različnih merskih enotah, neposredna primerjava ni mogoča, ampak je mogoča primerjava le na podlagi relativnih števil, še najpogosteje z indeksi.

Indeksi so relativna števila, ki so razmerja med dvema istovrstnima podatkoma, ki sta pomnožena s 100. Indeks je neimenovano število in je **vedno pozitivno**. V števcu je primerjalni podatek, v imenovalcu pa osnova ali baza.

$$\textbf{Indeks} = \frac{\textbf{primerjalni podatek}}{\textbf{osnova ali baza}} \times \mathbf{100}$$

Splošni obrazec: $I_{j/0} = \dfrac{Y_j}{Y_0} \times 100$

Indeks ima vrednost večjo od 100, če je primerjalni podatek večji od osnove.
Indeks doseže vrednost enako 100, če sta podatka enaka.
Indeks ima vrednost manjšo od 100, če je primerjalni podatek manjši od osnove.

I > 100	porast pojava
I < 100	upad pojava
I = 100	pojav je nespremenjen

Indeksi se razložijo tudi tako, da od indeksa odštejemo 100 in dobimo relativno razliko $Dj/_0$, ki jo izrazimo v odstotkih in pomeni:

- če je indeks večji od 100, je razlika pozitivna in pomeni porast pojava,
- če je indeks manjši od 100, je razlika negativna in pomeni upad pojava,
- če je indeks 100, ni razlike in pomeni, da se pojav ni spremenil.

Osnova indeksa je lahko fiksna ali spremenljiva. Indekse s stalno ali fiksno osnovo imenujemo tudi **bazni** indeksi. Najpogostejši indeksi s **spremenljivo** osnovo so verižni indeksi.

Indekse, ki so izračunani za en sam pojav, imenujemo **enostavni** indeksi, če jih izračunavamo za več pojavov, izračunamo **skupinske** indekse.

Če se indeksi izračunavajo iz dveh istovrstnih podatkov, ki se nanašajo na dve geografski območji, izračunamo **krajevne indekse**. Za primerjavo vzamemo za osnovo kraj, s katerim se želimo primerjati, ali podatke za geografsko območje, ali skupino krajev, ali celotno državo. Npr. podatke za podjetje bi primerjali z istovrstnimi podatki za Bled, za Gorenjsko, za gorske kraje, za Slovenijo itd. Tako bi lahko primerjali povprečno dobo bivanja, zasedenost hotela itd.

Indekse običajno izračunamo iz opazovanih podatkov, lahko pa tudi iz relativnih števil.

Ponovite matematična pravila računanja indeksov.

Za ocenjevanje značilnosti pojava ne zadošča izračunavanje le enega indeksa, ampak indekse za celotno časovno vrsto. Iz tako izračunane **indeksne vrste** lahko z analizo ugotavljamo razlike med območji pri krajevnih indeksih in razlike med časovnimi obdobji pri časovnih indeksnih vrstah. Npr. izračunamo indeks, kjer primerjamo povprečno dobo bivanja turistov našega podjetja z več kraji (gorskimi, obmorskimi, Bledom, Bohinjem, Kranjsko goro …) ali indeksi za več let.

Iz vsebinskega vidika je zelo pomembna pravilna izbira **osnove za primerjavo**.

Pri **krajevnih indeksih** izberemo kot osnovo za primerjavo običajno kraj, v katerem je naša lokacija, npr. če je hotel na Bledu, se primerja s podatki Bleda. Lahko je osnova tudi občina, regija, država ali drugačna geografska opredelitev, ki je smiselna, npr. če smo na Bledu, se primerjamo s podatki občine Bled, s podatki za Gorenjsko, podatki za več občin, s katerimi imamo skupne značilnosti, npr. Bled, Bohinj, Kranjska gora, ali s podatki, kot so npr. gorski kraji.

Pri **časovnih indeksih** za osnovo izberemo tisto časovno obdobje ali trenutek, ko ni bilo kakih posebnih dogodkov, ki bi prikazovali nerealne vrednosti, npr. leto vojne in osamosvajanja Slovenije ni dobro izbrati za osnovo za izračunavanje indeksov, saj izračunani indeksi ne bodo prikazali pravilne značilnosti pojava. Za osnovo lahko vzamemo tudi podatke povprečja za daljše časovno obdobje. Časovne indekse izračunavamo tudi za analizo dinamike pojava. To izražamo z izračunavanjem verižnih indeksov, koeficientov dinamike in stopnje rasti.

Največkrat za **osnovo** izberemo povprečno vrednost, lahko pa tudi **mediano**, **najmanjšo** vrednost, **največjo** vrednost ali kak **drug podatek**, ki se nam zdi primeren za primerjavo glede na cilj statistične analize. Npr. če bi v hotelu število gostov na določen dan primerjali z različnimi osnovami, bi izračunali različne indekse in vsak izračun bi nam dal drugačne informacije. Če bi se primerjali z minimumom, bi izračunali vse indekse, ki bi dosegli vrednost nad 100 in bi prikazovali porast pojava, če bi si izbrali maksimum za osnovo, bi izračunali indekse, ki bi bili vsi manjši od 100 in bi prikazovali upad pojava, če bi izbrali osnovo povprečje, bi izračunali nekaj indeksov z vrednostjo nad 100, ki bi izražali vrednosti večje od povprečja in nekaj pod 100, ki bi prikazovali podpovprečno velikost pojava. Napačna izbira osnove, lahko vodi v zavajanje ali v nepravilne analize in slabe poslovne odločitev. Če bi želeli prikazati, kako dobri in uspešni smo, se bomo primerjali z najslabšim, če želimo poudariti slabosti se primerjamo z najboljšim. Lahko za osnovo izbiramo tudi konkurenčna podjetja, npr. na Bledu hotel Toplice primerja podatke s hotelom Golf.

Poiščite in razložite izračunane indekse pri ostalih ekonomskih predmetih!

4.3.1 Časovni indeksi

Značilnost časovnih indeksov je primerjava dveh istovrstnih podatkov, ki se nanašata na **dva različna časovna trenutka ali intervala** (obdobja). Največkrat se uporabljajo za prikazovanje dinamike oziroma gibanje pojava v daljšem časovnem obdobju. Prikazana časovna vrsta že sama prikazuje dinamiko, z izračunom indeksom pa jo še poudarimo. Iz časovnih vrst lahko izračunamo indekse s **stalno osnovo** in **verižne indekse**.

Pri indeksih s **stalno osnovo** vsak člen časovne vrste primerjamo s členom, ki smo ga določili za osnovo.

$$I_{j/0} = \frac{Y_j}{Y_0} \times 100$$

Yj - vrednost členov v časovni vrsti za obdobje ali trenutek
Y_0 - vrednost člena, določenega za osnovo

Preračunavanje indeksov s stalno osnovo **na novo osnovo** je potrebno, kadar primerjamo dinamiko več pojavov, za katere poznamo podatke indeksnih vrst z različnimi osnovami, ne poznamo pa osnovnih podatkov za izračun želenega indeksa na izbrano osnovo. Preračunamo jih tako, da vsak indeks indeksne vrste delimo z indeksom tistega obdobja (meseca, leta), ki smo ga določili za novo osnovo.

Če je dinamika pojava prikazana z indeksno vrsto, ne pa z osnovno časovno vrsto, lahko iz **indeksov izračunamo podatke osnovne vrste**, če poznamo vrednost vsaj enega člena. Če indeks s stalno osnovo za posamezno obdobje pomnožimo z znanim podatkom časovne vrste, izračunamo vrednosti časovne vrste za ostala obdobja.

$$I_{j/0} = \frac{Y_j}{Y_0} \times 100 \quad \textit{in iz tega}: \; Y_j = \frac{Y_0 \times I_{j/0}}{100}$$

Y_j - podatki za ostala obdobja časovne vrste
Y_0 - znani podatek časovne vrste
$I_{j/0}$ - indeks s stalno osnovo

Pri računanju s programom Excel moramo paziti, kako oblikujemo formulo, da to lahko kopiramo za celotno časovno vrsto. Zato je nujno poznavanje pravil oblikovanja in kopiranja formul v programu Excel in statistične teorije za izbiro prave osnove pri indeksih in prave vrste relativnih števil. Z analizo dobljenih rezultatov ugotavljamo značilnosti pojava. Indeksi nad 100 kažejo na porast pojava, pod 100 upad, če se pojav v primerjavi z osnovo ni spremenil, je vrednost indeksa 100.

Ponovite matematična pravila preračunavanja indeksov!

Za ugotavljanje relativnih sprememb med zaporednimi časovnimi trenutki ali obdobji izračunavamo **verižne indekse**. Ker pri izračunavanju ne upoštevamo stalne osnove, ampak se ta spreminja, jih imenujemo indeksi s premično osnovo. Izračunamo jih tako, da podatek za pojav v posamičnem trenutku ali obdobju primerjamo s podatkom za isti pojav v predhodnem časovnem obdobju ali intervalu.

$$V_j = \frac{Y_j}{Y_{j-1}} \times 100$$

Verižni indeks izračunamo tako, da podatek za časovni trenutek ali interval danega obdobja delimo z istovrstnim podatkom za predhodno obdobje in kvocient pomnožimo s 100.

Za prvo obdobje, ki je prikazano v časovni vrsti, verižnega indeksa ni mogoče izračunati, ker nimamo predhodnega podatka, s katerim bi se lahko primerjali. Običajno se računajo na eno decimalno mesto.

Verižne indekse lahko preračunamo v indekse s stalno osnovo. Formula za izračun se razlikuje glede na obdobje izračunavanja.

Za obdobje **pred** izhodiščnim obdobjem jih preračunamo tako, da indeks s stalno osnovo izhodiščnega obdobja delimo z verižnim indeksom tega obdobja in izračunani količnik pomnožimo s 100.

Indeks s stalno osnovo **za** izhodiščno obdobje in vsa naslednja obdobja izračunamo tako, da indeks s stalno osnovo za izhodiščno obdobje pomnožimo z verižnim indeksom naslednjega obdobja in dobljeni rezultat delimo s 100. Pri preračunavanju z računalnikom moramo paziti na pravilno oblikovanje formule zaradi kopiranja na celotno časovno vrsto.

Indeksi **pred** izhodiščnim obdobjem	$\frac{I_{08/08}}{V_{08}} \times 100$ in $\frac{I_{07/08}}{V_{07}} \times 100$
Indeksi **za** izhodiščnim obdobjem	$\frac{I_{08/08} \times V_{09}}{100}$ in $\frac{I_{09/08} \times V_{10}}{100}$

Verižne indekse preračunamo v vrednosti, če poznamo vrednost izhodiščnega člena časovne vrste tako, da ga pomnožimo z verižnim indeksom in delimo s 100. Če ne poznamo vrednosti izhodiščnega člena, moramo najprej verižne indekse preračunati na indekse s stalno osnovo na obdobje znanega podatka. Tako pridobljena vrednost omogoča izračun vseh vrednosti časovne vrste.

$$V_j = \frac{Y_j}{Y_{j-1}} \times 100 \ \ in\ iz\ tega : Y_j = \frac{Y_{j-1} \times V_j}{100}$$

Y_j - vrednost novega člena
V_j - verižni indeks novega člena
Y_{j-1} - vrednost izhodiščnega člena

Pri računanju verižnih indeksov in preračunavanju s programom Excel je potrebno paziti na pravilno oblikovanje formul, kar omogoča kopiranje formule za celotno časovno vrsto.

Določiti je potrebno ustrezno število decimalnih mest. Običajno se indeksi izračunavajo na eno decimalno mesto.

Analiza izračunanih verižnih indeksov za statistično vrsto nam pove gibanje in spremembe v opazovanem obdobju. Če je vrednost verižnega indeksa 100, se opazovani pojav v tekočem obdobju v primerjavi s predhodnim ni spremenil, če je indeks nad 100, se je pojav glede na predhodno časovno obdobje povečal, če je indeks pod 100, se je pojav zmanjšal.

Indekse v poslovnih poročilih in statističnih analizah zelo pogosto uporabljamo, predvsem kadar želimo ugotoviti dinamiko pojava. Ker nas pri poslovnih odločitvah zanima gibanje pojava v preteklosti zaradi ocenjevanja pričakovanj v prihodnosti, je znanje indeksov nujno.

Izračunavanje **verižnih indeksov** je smiselno za pojave, za katere je značilno periodično ali sezonsko nihanje. To je nihanje pojava na krajša časovna obdobja, npr. dan, teden, mesec, leto. Med periodičnimi nihanji so še posebej pomembna sezonska nihanja, ki se izračunavajo za obdobje enega leta. Prav ta nihanja so zelo značilna za gostinstvo in turizem. Zato se verižni indeksi zelo pogosto izračunavajo. Glede na obseg prikazanih podatkov v statistični vrsti se lahko z njimi izražajo nihanja v dnevu (npr. če prikazujemo podatke o prodaji v restavraciji po urah), v tednu (npr. če spremljamo število gostov v hotelu po dneh), v mesecu (npr. če spremljamo prodajo v restavracijo po tednih), v letu (npr. če spremljamo število nočitev po mesecih) ali nihanja v daljšem obdobju, če se podatki nanašajo na več let.

Glede na namen statistične obdelave izberemo ustrezne podatke, ki jih bomo analizirali. Za poslovne odločitve nas zanimajo različne informacije zato izračunavamo tudi indekse z različnimi **stalnimi osnovami**.

Naloge:

1. *Poiščite krajevne in ostale statistične vrste, iz katerih lahko izračunavamo indekse.*
2. *Poiščite podatke časovne vrste na spletu.*
3. *Iz pridobljenih podatkov izračunajte verižne indekse in analizirajte dobljene rezultate.*
4. *Ugotovite možne osnove za izračun indeksov s stalno osnovo in primerjajte razlike.*
5. *Izračunajte indekse s stalno osnovo, minimum, maksimum, povprečje in določen izbrani podatek in analizirajte dobljene rezultate.*
6. *Na primeru preračunajte indekse s stalno osnovo na novo izbrano stalno osnovo.*
7. *Na primeru iz indeksov s stalno osnovo izračunajte vrednosti za celotno časovno vrsto.*
8. *Verižne indekse preračunajte v indekse s stalno osnovo.*
9. *Iz verižnih indeksov izračunajte vrednosti za celotno časovno vrsto.*
10. *Grafično prikažite indekse s stalno osnovo in verižne indekse! Izberite ustrezno vrsto grafa in ga oblikujte.*

4.3.2 Kazalci rasti

Kazalci rasti so relativna števila, ki, podobno kot verižni indeksi, prikazujejo relativne spremembe pojava v času. Uporabljajo se predvsem pri proučevanju sprememb med dvema zaporednima časovnima obdobjema. Kazalci rasti so:

- razlika, $\mathbf{D_j}$,
- koeficient rasti (dinamike), $\mathbf{K_j}$,

- verižni indeks, **V_j,**
- stopnja rasti (relativna razlika), **S_j**.

Ponovite matematična pravila računanja kazalcev dinamike!

Razlika, D_j, je izračunana iz dveh istovrstnih podatkov, ki se nanašata na dva zaporedna časovna trenutka ali obdobja.

$D_j = Y_j - Y_{j-1}$

Od podatka za dano obdobje odštejemo istovrstni podatek predhodnega obdobja, npr. od števila gostov danes odštejemo število gostov včeraj ali od števila nočitev tekočega meseca odštejemo število nočitev preteklega meseca. Izračunana razlika je **pozitivna**, če pojav raste, **negativna**, če pojav upada, ali je enaka **0,** če se pojav ni spremenil. Razlika je imenovano število.

Pri računanju razlike v programu Excel moramo paziti na pravilno oblikovanje formule, da lahko to kopiramo in izračunamo razliko za celotno časovno vrsto.

Razlike nikoli ne moremo izračunati za prvo opazovano obdobje, ker podatka pred prvim obdobjem nimamo.

Koeficient rasti ali koeficient dinamike, K_j, je relativno število, ki je razmerje med dvema istovrstnima podatkoma, ki se nanašata na dva zaporedna časovna trenutka ali obdobja.

$$K_j = \frac{Y_j}{Y_{j-1}}$$

Podatek za dano časovno obdobje ali trenutek primerjamo z istovrstnim podatkom predhodnega obdobja ali trenutka. Je zelo podoben verižnemu indeksu, ki se izračuna tako, da isto razmerje še pomnožimo s 100. Koeficient rasti ima lahko vrednost večjo od 1, če pojav raste, manjšo od 1, če pojav upada, če pojav ostaja nespremenjen, je vrednost koeficienta dinamike 1. Enako kot verižni indeks tudi koeficient prikazuje relativne spremembe od člena do člena v časovni vrsti. Običajno ga izračunamo na tri decimalna mesta.

Pri računanju koeficientov z računalnikom morate paziti na pravilno oblikovanje formul, ki omogoča kopiranje teh na celotno časovno vrsto, in na določitev ustreznega števila decimalnih mest.

Stopnja rasti, S_j, je razmerje med razliko in stanjem v predhodnem časovnem obdobju ali trenutku. Izražena je v %. Izraža relativno razliko ali odstotno spremembo dveh zaporednih členov časovne vrste. Največkrat jo uporabljamo za razlago dinamike proučevanega pojava. Tudi izračunane verižne indekse ali koeficiente dinamike običajno razlagamo s stopnjo rasti.

$$S_j = \frac{Y_j - Y_{j-1}}{Y_{j-1}} \times 100$$

Stopnja rasti je **pozitivna**, če se pojav poveča, **negativna,** če se pojav zmanjša, ali je enaka **0**, če se pojav ni spremenil.

Stopnja rasti se nanaša na relativno spremembo enega pojava in je izračunana iz dveh istovrstnih podatkov, ki se lahko nanašata na dva zaporedna časovna trenutka ali na dva zaporedna časovna intervala. Za izračunane stopnje v demografski statistiki je značilno, da se lahko izračunajo tudi iz istovrstnih podatkov, ki se nanašajo na isti časovni interval.

Pri računanju stopnje rasti z računalnikom moramo paziti na pravilno oblikovanje formule, ki omogoča kopiranje za celotno časovno vrsto. Zaradi primerljivosti z verižnim indeksom jo običajno izračunamo na eno decimalno mesto.

Spremembe v času oz. dinamiko pojava lahko torej proučujemo in izražamo z različnimi kazalci. Vsak izraža spremembe na svoj način. Lahko izračunavamo razliko ali razmerja, ki izražajo relativno spremembo pojava. Če poznamo vrednost za enega izmed njih, lahko iz zvez med njimi ugotovimo tudi vrednost za ostale. Iz obrazcev za izračunavanje lahko razberemo sledeče povezave med kazalci dinamike.

$$\mathbf{V_j - 100 = S_j} \qquad \mathbf{S_j + 100 = V_j}$$

$$\mathbf{K_j = \frac{V_j}{100}} \qquad \mathbf{V_j = K_j \times 100}$$

To povezavo dokazujejo tudi vrednosti posameznih kazalcev dinamike, ki prikazujejo naraščanje, upadanje ali stagniranje proučevanega pojava (Blejec, 1973).

Kazalec dinamike		**Pojav narašča**	**Pojav upada**	**Pojav stagnira**
Verižni indeks	$\mathbf{V_j}$	$V_j > 100$	$V_j < 100$	$V_j = 100$
Koeficient dinamike	$\mathbf{K_j}$	$K_j > 1$	$K_j < 1$	$K_j = 1$
Stopnja rasti	$\mathbf{S_j}$	$S_j > 0$	$S_j < 0$	$S_j = 0$

Za **grafično prikazovanje** indeksov najpogosteje uporabljamo linijski grafikon in stolpce. Pazimo pri označevanju x osi, saj indekse, ki se nanašajo na intervalno časovno vrsto, označimo na sredini intervala, točko za indekse, računane za trenutno časovno vrsto, pa vrišemo nad opazovani časovni trenutek. Pri stolpčnih grafih se nazornost prikaza poveča, če stolpce, ki prikazujejo indekse nad 100, obrnemo navzgor, stolpce, ki prikazujejo indekse pod 100, pa navzdol. Lahko izberemo prikaz z vodoravno ali navpično postavljenimi stolpci. Ti so lahko razporejeni po časovni vrsti ali pa jih razporedimo po vrednosti indeksa padajoče ali naraščajoče.

V programu Excel izberemo pravo vrsto grafa, prave podatke in pazimo na oblikovanje in vse sestavine, ki jih mora imeti vsak graf.

Znanje izračunavanja indeksov in grafično prikazovanje teh je zelo pomembno, saj ni statistične analize brez primerjave podatkov z izbrano osnovo ali primerjave gibanja pojava. Zaradi lastnosti neimenovanih števil, ki omogočajo tudi primerjavo raznovrstnih podatkov, se v praksi stalno uporabljajo. Pravilno je potrebno izbrati osnovo, se odločiti za izračun indeksov s stalno osnovo (baznih) ali verižnih indeksov (spremenljiva osnova). Za prikazovanje dinamike lahko izberemo tudi druga relativna števila: koeficiente, stopnje ali razliko. Paziti je potrebno na izbiro ustreznega relativnega števila in za pravilno razlago dobljenih rezultatov. Poznati je potrebno tudi medsebojne povezave različnih relativnih števil, ki omogočajo tudi preračunavanje in izražanje značilnosti pojava na različne načine. Da znamo razbrati in pravilno razložiti značilnosti naraščanja, upadanja ali stagnacije opazovanega pojav, je nujno znanje statistične teorije. Za pravilno grafično prikazovanje podatkov je potrebno dobro znanje ustreznega računalniškega programa in statistične teorije.

*Naloge:*

1. Iz podatkov izbrane časovne vrste izračunajte vse kazalce dinamike, koeficiente, verižne indekse, stopnje, razliko in primerjajte dobljene rezultate.
2. Preračunajte na primeru verižne indekse v koeficiente dinamike in stopnje! Primerjajte rezultate.
3. Preračunajte na izbranem primeru koeficiente dinamike v verižne indekse in stopnje ter primerjajte rezultate.
4. Preračunajte na izbranem primeru stopnje v verižne indekse in koeficiente dinamike ter primerjajte rezultate.

4.4 STATISTIČNI KOEFICIENTI

Pri analizi opazovanega pojava je zelo pomemben parameter, s katerim ugotavljamo značilnosti pojava, tudi statistični koeficient. Za sprejemanje dobrih poslovnih odločitev se v gostinstvu in turizmu zelo pogosto izračunavajo različni statistični koeficienti. Za celovito analizo pojava ne zadoščajo le primerjave istovrstnih podatkov, kar omogočajo ostala relativna števila, ampak tudi primerjave raznovrstnih podatkov. Analitična vrednost pridobljenih rezultatov z izračunavanjem statističnih koeficientov je zelo velika, saj omogoča primerjavo različnih podatkov, ki pa morajo biti v medsebojni zvezi. Izražajo intenzivnost dogajanja in omogočajo primerjave intenzivnosti v času in prostoru (npr. zasedenost hotela, povprečna doba bivanja gostov). Pomembni so za izražanje uspešnosti podjetja in primerjavo s konkurenco, geografsko ali časovno (npr. ekonomičnost, rentabilnost, donosnost, produktivnost ...). Izračunavajo se pogosto v demografski statistiki, npr. rodnost, umrljivost.

S statističnimi koeficienti se srečujete pri vseh ekonomskih in strokovnih predmetih, saj so odlični pri ugotavljanju in analizi uspešnosti podjetja in posameznih poslovnih odločitvah. Statistični koeficienti, ki se najpogosteje izračunavajo v gostinstvu in turizmu, so, poleg že omenjenih, povprečna doba bivanja gostov, zasedenost hotela, prihodek na 1 gosta ali na 1 nočitev.

Izračunamo jih kot razmerje med dvema **raznovrstnima podatkoma**, ki morata biti v medsebojni zvezi. Primerjava mora biti vsebinsko **smiselna.** Oba primerjana podatka se morata nanašati **na isti** trenutek ali isto **časovno** obdobje in sta **enako krajevno opredeljena**.

$$\boldsymbol{K = \frac{Y}{X}} \quad \textit{oziroma} \quad \boldsymbol{K = \frac{Y}{X} \times E}$$

K - statistični koeficient
Y - podatek, ki ga primerjamo
X - podatek, ki smo ga izbrali za osnovo primerjave

E - 1, 10, 100, 1.000 ali 100.000, odvisno od tega na koliko enot računamo koeficient

Koeficienti so navadno imenovana števila. Takrat navajamo tudi mersko enoto, v kateri je merjen pojav, iz katerih je izračunan koeficient. Npr. povprečna doba bivanja je izražena v številu dni. Včasih je tudi neimenovano število, če sta podatka, ki jih primerjamo, izražena v istih merskih enotah, lahko je izražen tudi v %.

Ponovite koeficiente, ki ste jih spoznali pri ekonomskih predmetih (ekonomičnost, produktivnost, rentabilnost...).

Izračunani koeficient je lahko izražen na enoto podatka v imenovalcu ali na 10, 100, 1000 ali več enot. Npr. število gostov na enega natakarja ali število upravnih delavce na zmogljivost hotela 100 postelj.

Če je smiselno, lahko v koeficientu tudi zamenjamo podatek iz števca in imenovalca. Tako izračunamo obratni ali **recipročni** koeficient glede na prvega. Npr. koeficient obračanja zalog, ki nam pove, kolikokrat se povprečne zaloge surovin ali blaga obrnejo v enoti časa, lahko izrazimo z recipročnim koeficientom, ki kaže dolžino trajanja enega obrata v dnevih ali povprečni čas skladiščenja surovin. Ali produktivnost, ki jo izrazimo s številom izdelkov ali storitev na delavca, lahko izrazimo z recipročnim koeficientom, ki izraža čas delavca za izdelavo izdelka ali opravljanje storitve.

Podatke lahko neposredno primerjamo le, če se oba nanašata na isti časovni trenutek ali interval. Če ta pogoj ni izpolnjen, moramo podatke najprej prilagoditi, tako da **trenutne podatke spremenimo v intervalne**. Poznati moramo vsaj nekaj trenutnih podatkov, ki se nanašajo na obravnavani časovni interval. Iz njih izračunamo povprečje. Tako je koeficient razmerje, ki je izračunano iz intervalnega podatka in povprečja trenutnih podatkov opazovanega intervala. Npr. pri izračunavanju produktivnosti upoštevamo povprečno število zaposlenih ali povprečno število opravljenih ur v nekem obdobju, saj se število zaposlenih in število opravljenih ur lahko zelo razlikuje.

Zaradi primerjav izračunanih statističnih koeficientov z drugimi se ti običajno izračunavajo za **časovno obdobje enega leta**. Npr. koeficient obračanja zalog, povprečna doba bivanja, rentabilnost, profitna stopnja itd. Izbira dolžine obdobja je odvisna od cilja raziskave. Pojav lahko opazujemo in zbiramo intervalne podatke za obdobje ure, dneva, tedna, meseca, leta ali več let. Dolžina časovnega intervala vpliva na izračunano vrednost koeficienta. Zato mora biti navedeno obdobje, na katerega se nanaša izračunani koeficient. Npr. urna produktivnost, mesečna ekonomičnost, letna rentabilnost. V ekonomskih analizah se pogosto izhaja iz koeficientov, izračunanih za krajša časovna obdobja, npr. število prodanih kav na uro, število obiskovalcev na dan, povprečna prodaja na teden, povprečna mesečna vrednost prodaje na natakarja itd.

Na praksi ugotovite, katere koeficiente bi bilo smiselno izračunati!

Koeficiente **grafično prikazujemo** s stolpčnimi in linijskimi grafikoni. Pri stolpčnih grafikonih je višina stolpca sorazmerna z velikostjo koeficienta. Lahko izberemo prikaz z navpičnimi ali vodoravnimi stolpci, ki jih zaradi nazornosti lahko razporedimo tudi po velikosti.

Zaradi pogoste uporabe statističnih koeficientov v ekonomskih analizah je potrebno dobro poznati statistično teorijo, način izračunavanja in ekonomske pojme. Pomembna je izbira opazovanega obdobja, trenutne podatke večkrat preračunati na intervalne in izbrati podatke, ki jih je smiselno primerjali. Lahko izračunamo statistične koeficiente ali recipročne koeficiente. V analizi jih je potrebno pravilno razložiti in grafično prikazati.

Naloge:

1. Poiščite statistične koeficiente, ki ste jih računali pri ekonomskih predmetih.
2. Poiščite tabelo s podatki in ugotovite, kateri statistični koeficienti bi bili smiselni.
3. V izbrani tabeli ugotovite, ali so izpolnjeni pogoji o enaki krajevni opredelitvi podatkov, ki jih primerjate.
4. V izbrani tabeli ugotovite podatke, ki se nanašajo na trenutek in podatke, ki se nanašajo na obdobje.
5. Iz izbranih podatkov izračunajte statistične koeficiente in recipročne koeficiente, če so ti smiselni!.

4.5 RAZLIKA IN RELATIVNA RAZLIKA ZA RELATIVNA ŠTEVILA

Tudi za relativna števila pogosto izračunavamo razliko in relativno razliko. Paziti je potrebno predvsem na razlago dobljenih rezultatov.

Ker je večina relativnih števil **neimenovanih, se razlika med njimi izrazi v točkah**. Izračunamo jo tako, da od relativnega števila odštejemo istovrstno relativno število. Razlika strukturnih odstotkov je izražena v odstotnih točkah, pri indeksih v indeksnih točkah, pri stopnjah v odstotnih točkah. S primerjavo koeficientov pa razliko izrazimo v isti merski enoti kot primerjani relativni števili.

Razmislite, katere razlike bi bilo smiselno izračunati!

Relativna razlika je pri vseh relativnih številih, imenovanih in neimenovanih, izražena v **odstotkih**.

Čeprav je matematično enostavno izračunati razliko in relativno razliko, je pomembna predvsem pravilna razlaga dobljenih rezultatov, zato je potrebno dobro znanje poglavja relativnih števil. Glede na cilj statistične obdelave se odločamo za izračunavanje ustrezne vrste razlike.

Naloge:

1. Iz izbranih podatkov izračunajte razliko ter relativno razliko, razložite in primerjajte dobljene rezultate.
2. Iz relativnih števil izračunajte razliko ter relativno razliko, razložite in primerjajte dobljene rezultate.
3. Iz izbranih podatkov izračunajte koeficiente, razliko ter relativno razliko koeficientov, razložite in primerjajte dobljene rezultate.

5 RANGI IN KVANTILI

Za pravilno ugotavljanje lastnosti preučevane populacije in hitro izračunavanje parametrov, je zelo pomemben pravilen in pregleden prikaz zbranih podatkov. Raziskovalec mora zato izbrati ustrezno obliko prikaza. Pri tem so mu v veliko pomoč sodobni mediji Npr. v hotelu nas zanima, kdaj je bilo največ gostov, kdaj najmanj, kaj pomeni podatek o današnjem številu gostov glede na ostale dni itd. Če se želi prikazati položaj posamezne enote med ostalimi enotami populacije, se lahko oblikujejo ranžirne vrste.

5.1 RANŽIRNA VRSTA

Odgovore na ta in podobna vprašanja lahko pridobimo tudi iz prikaza podatkov, ki so razvrščeni v ranžirno vrsto. **Ranžirno vrsto** podatkov dobimo, če podatke razvrstimo po velikosti od najmanjšega do največjega ali obratno. V ranžirni vrsti ima vsak podatek neko zaporedno mesto, ki je določeno glede na vrednost. Mesto, ki ga zavzema, imenujemo **rang**. Rang ima vse vrednosti od 1 do N (števila enot proučevane populacije). Pri naraščajoči vrsti je prvi podatek, ki ima rang 1, najslabši, sledijo boljši podatki. Čim večji rang ima podatek, boljši je rezultat. Obratno je pri padajočih ranžirnih vrstah. To je zelo enostavna in hitra oblika ureditve rezultatov. Ročno urejanje je zamudno, zato se danes uporabljajo računalniški programi, ki omogočajo enostavno, hitro in natančno oblikovanje ranžirnih vrst. Zato je potrebno poznati v programu predviden ukaz, ki to omogoča.

Poiščite različne ranžirne vrste (športni rezultati, telefonski imeniki, slovarji ...)!

Iz ranžirne vrste hitro ugotovimo najmanjši in največji rezultat. Rang posameznega podatka nam da dodatno informacijo o vsaki enoti. Opazimo tudi podatek, ki se pojavlja najpogosteje, in hitro najdemo mesto opazovanega podatka, ki nas zanima.

5.2 RANGI

Pri navajanju **absolutnega ranga**, to je zaporednega mesta posameznega podatka oz. enote v ranžirni vrsti, je potrebno vedno navesti tudi število enot populacije, da ne prihaja do zavajanja in napačne razlage pojava. Zato je potrebno poznati **slabost absolutnega ranga**. Npr. športnik je dosegel 3. mesto na tekmovanju (absolutni rang 3). Če ne navedemo števila tekmovalcev, sklepamo, da je zelo uspešen. Šele, ko ob podatku o doseženem 3. mestu navedemo število tekmovalcev, je možna realna ocena uspešnosti. Če so tekmovali trije, je doseženo tretje mesto slabo, saj je bil zadnji, če jih je tekmovalo 100 ali več je zelo dobro.

Analizirajte ranžirne vrste npr. športnih rezultatov!

Pri zveznih numeričnih podatkih določimo absolutni ranga 1, ki vključuje razmik vrednosti od 0,5 do 1,5. To lastnost je potrebno upoštevati pri računanju kvantilnega ranga.

Kvantilni rang določa položaj posamezne enote relativno, to je v odnosu do skupnega števila enot. V praksi se zelo pogosto uporablja, ker odpravlja slabost absolutnega ranga. Izračunamo ga kot razmerje med absolutnim rangom in številom enot populacije. Ker v razmerju upoštevamo spodnjo mejo absolutnega ranga, ki je zvezni numerični podatek, upoštevamo pri izračunu popravek 0,5 (Sagadin, 2003).

Relativni rang P imenujemo kvantilni rang (kvantum - količina).

$$P = \frac{R - 0,5}{N}$$

Dobljeni rezultat nam pove, koliko odstotkov enot ima manjše in koliko večje vrednosti od izbrane enote.

Če poznamo kvantilni rang, lahko izračunamo absolutnega.

$$R = N \times P + 0,5$$

Nekateri kvantili imajo posebna imena in se pogosteje uporabljajo. To so:

1. **Kvartili** (Q), ki populacijo razdelijo na 4 po velikosti (po obsegu ali številu enot) enake dele. Populacija je razdeljena na četrtine (N/4).
    a. Prvi kvartil - Q1 - določa kvantilni rang P = 0,25 kar pomeni, da je 25 % enot manjših ali enakih tej vrednosti, 75 % enot pa je večje vrednosti.
    b. Drugi kvartil - Q2 - s kvantilnim rangom P = 0,50 je vrednost, od katere ima 50 % enot manjše ali enake vrednosti, 50 % enot pa večje vrednosti. Sem spada tudi mediana (razložena v poglavju srednjih vrednosti).
    c. Tretji kvartil - Q3 - s kvantilnim rangom P = 0,75 je vrednost, od katere je 75 % enot manjših ali enakih, 25 % pa večjih.

2. **Decili** (D), ki populacijo razdelijo na 10 po velikosti enakih delov - na desetine (N/10).
    a. Prvi decil - D1 - je kvantil s kvantilnim rangom 0,10, kar pomeni, da je 10 % enot manjših ali enakih tej vrednosti, 90 % pa večjih.
    b. Drugi decil - D2 - določa rang P = 0,20, kar pomeni, da je 20 % enot manjših in 80 % večjih.
    c. Peti decil se ujema z drugim kvartilom oziroma z mediano.
    d. Deseti decil zajema zadnje enote.

3. **Centili** (C), ki populacijo razdelijo na 100 po velikosti enakih delov – na stotine (N/100).
   Deseti centil se ujema s prvim decilom, dvajseti z drugim itd (Šadl, 2001).

Vrednost posamezne spremenljivke lahko izrazimo z različnimi kvantili. V poglavju srednjih vrednosti bo omenjena MEDIANA, ki jo lahko izrazimo s Q2 = D5 = C50 in se najpogosteje izračunava. Predstavlja vrednost, od katere ima polovica enot manjšo ali enako vrednost, druga polovica enot pa je večjih.

Izračunavanje kvantilov je zelo **uporabno**, saj nam omogoča različne izračune. Za znano vrednost posamezne enote se z izračunom kvantilnega ranga izrazi položaj te enote relativno glede na celotno populacijo, kar pove delež enot, ki so manjše oz večje.

Sestavite različne ranžirne vrste. Analizirajte rezultate!

Npr. v hotelu je na določen dan prenočilo 70 gostov. Če želimo ugotoviti, kaj to pomeni v tem mesecu, izračunamo kvantilni rang, ki pove koliko odstotkov enot (v našem primeru število dni) v tem mesecu je manjših od podatka in koliko odstotkov enot večjih. Za izračun je potrebno oblikovati ranžirno vrsto, določiti absolutni rang in izračunati kvantilni rang.

Iz znanega kvantilnega ranga lahko poiščemo ustrezno vrednost enote v opazovani populaciji. Npr. za izbrani mesec nas v hotelu zanima številu gostov, ko je bilo 50 % dni več gostov in 50 % dni manj gostov v hotelu. Izračunati želimo mediano ali Q2 ali D5 ali C50. Na osnovi izračunanega ranga v ranžirni vrsti razberemo želeni rezultat.

Ranžirna vrsta je primerna za prikazovanje majhnega števila rezultatov, do približno 20. Če je podatkov preveč, je prikaz preveč podroben, preobsežen in si težko ustvarimo sliko o splošni značilnosti proučevane populacije. Enostavno in pregledno je tudi grafično prikazovanje. Najpogosteje za prikaz izberemo stolpce ali linijski grafikon.

Ker se v delih podjetja (restavraciji, kuhinji, recepciji) pogosto obdeluje manjše število podatkov zaradi enostavnosti obdelave in prikazovanja z uporabo računalniških programov, se ta oblika statistične obdelave v praksi zelo pogosto uporablja. Za pravilno uporabo je potrebno poznati osnovne značilnosti oblikovanja ranžirne vrste, določanja absolutnega in kvantilnega ranga. Glede na cilj analize se odločimo za izračunavanje mediane, kvartilov, decilov ali centilov, ki jih je potrebno pravilno razložiti.

Naloge:

1. *Izbrane podatke v tabeli razporedite v ranžirno vrsto od največje do najmanjše vrednosti in obratno! Razložite prikazane podatke.*
2. *Ranžirno vrsto prikažite v stolpčnem (vodoravni in navpični stolpci) in linijskem grafikonu!*
3. *Določite absolutni rang.*
4. *Za podatke izračunajte relativni rang P ali kvantilni rang.*
5. *Za izbrane podatke izračunate kvartilni, decilni in centilni rang.*
6. *Primerjajte izračune in analizirajte razlike med rangi.*

6 FREKVENČNA PORAZDELITEV

Vsako podjetje, ki se odloča za statistično obdelavo, v fazi zbiranja pridobi veliko podatkov, saj vsako opazovano populacijo sestavlja veliko enot, ki so opazovane po več spremenljivkah, te pa imajo lahko veliko vrednosti. Npr. restavracijo na dan obišče 200 gostov, ki naročajo različne vrste jedi in pijač, plačujejo na različne načine itd. Zato moramo množico zbranih podatkov urediti in prikazati v pregledni obliki. Urejamo jih lahko na različne načine, če je število enot manjše, lahko z ranžirnimi vrstami, če je večje, se pogosteje poslužujemo razporejanja v frekvenčne porazdelitve.

Frekvenčna porazdelitev je razporeditev enot v skupine po vrednosti neke spremenljivke. Glede na vrste spremenljivk so frekvenčne porazdelitve za opisne in za številske spremenljivke. Namen sestavljanja je izboljšanje preglednosti, hitra analiza in lažje ugotavljanje značilnosti proučevanega pojava.

Ker so v statističnih poročilih praviloma prikazani podatki v tabelah v obliki frekvenčnih porazdelitev, je poznavanje problematike razvrščanja enot v frekvenčno porazdelitev nujno. Zato je potrebno dobro poznavanje osnovnih pojmov, ki so razloženi v uvodnih poglavjih. Na oblikovanje frekvenčne porazdelitve vpliva opazovani pojav in lastnosti, ki jih želimo ugotoviti.

Glede na vrsto spremenljivk poznamo frekvenčne porazdelitve za opisne in za številske spremenljivke.

6.1 FREKVENČNE PORAZDELITVE ZA OPISNE SPREMENLJIVKE

Najenostavneje je določati skupine, če ima spremenljivka manjše število vrednosti, saj lahko **za vsako vrednost določimo svojo skupino**. V gostinstvu in turizmu so primeri, ko lahko skupine določamo enostavno po vrednostih. Npr. razporejanje enot v skupine po spolu, saj lahko enostavno določimo 2 skupini (moški, ženske), ali državi, od koder prihajajo gostje itd.

Razmislite, kako vi oblikujete frekvenčne porazdelitve za opisne spremenljivke!

Kadar spremenljivke lahko zavzemajo veliko število vrednosti, v skupino **združujemo sorodne vrednosti**, da tako zmanjšamo število skupin, kar povečuje preglednost. Npr. spremljanje prihoda gostov v Slovenijo po vrstah turističnega kraja. Po sorodnosti so v posamezno skupino vključeni različni kraji. Skupine so razporejene: Ljubljana, zdraviliški kraji, obmorski kraji, gorski kraji, drugi turistični kraji, drugi kraji.

Problematika razvrščanja enot v skupine je odvisna od vrste spremenljivke. Pri opisnih spremenljivkah je nujno izpolniti pogoja enoličnosti in enovitosti za pravilno oblikovanje skupin. Ker pri opisnih spremenljivka opredelitev skupin včasih ni enostavna, je nujno dobro poznavanje poglavja 3.2. Neustrezno oblikovanje skupin vodi do napak pri razporejanju enot, neustrezne frekvenčne porazdelitve in nepravilne analize značilnosti pojava.

Naloge:

1. Glede na pridobljeno znanje v poglavju 3.2 razmislite, kako bi za primere iz gostinstva in turizma pravilno oblikovali skupine in sestavili frekvenčne porazdelitve.
2. Kako bi opredelili skupine, če bi želeli pridobiti podatke o narodnosti gostov, ki prenočujejo v hotelu? Ali bo to državljanstvo, ki ga izkazujejo z dokumentom? Ali bo to država rojstva? Ali bo to država stalnega bivališča? Ali bo to država začasnega bivališča?
3. Kako bi opredelili skupine, če bi želeli pridobiti podatke o izobrazbi? Ali bo to stopnja, ki pa je lahko v različnih državah različno določena? Ali bo to dokončana šola? Poklic, ki ga opravlja?
4. Če bi želeli ugotovit starostno strukturo gostov v hotelu. Ali bi bila določitev dveh skupin stari, mladi primerno? Zakaj? Kakšne skupine predlagate? Kako bi zagotovili enoličnost in enovitost?
5. Kakšne skupine predlagate, če bi želeli v frekvenčni porazdelitvi prikazati zadovoljstvo gostov s kakovostjo storitev? Analizirajte in razložite trditev?

6.2 FREKVENČNE PORAZDELITVE ZA ŠTEVILSKE SPREMENLJIVKE

To so frekvenčne porazdelitve, v katere so razporejene enote v skupine po vrednostih številskih spremenljivk. Te skupine imenujemo **razrede**. Tudi razredi morajo izpolnjevati temeljne pogoje enoličnosti in enovitosti. Frekvenčna razporeditev tudi pri številskih spremenljivkah mora zagotoviti preglednost prikazanih podatkov. Poznati in upoštevati moramo posebnosti in značilnosti številskih spremenljivk.

6.2.1 Meje razredov

Za številske spremenljivke z manjšim številom vrednosti opredelimo **skupine za posamezne vrednosti**. Npr. za ocene od 1 do 5 bi opredelili pet skupin. Kadar imajo številske spremenljivke veliko število vrednosti, oblikujemo razrede. Razrede med seboj razmejimo z določitvijo **spodnje in zgornje meje**. Te morajo biti določene tako nedvoumno, da je izpolnjen pogoj **enoličnosti**, kar pomeni, da morajo biti tako jasno določene, da vsako enoto lahko razporedimo le v en razred. Pri določanju mej se mora upoštevati značilnost številskih spremenljivk. Te so lahko zvezne ali diskretne in so lahko zaokrožene ali pa so cele vrednosti. Zato je potrebno poznati značilnosti številskih spremenljiv, ki so razložene v poglavju 3.

Za **zvezne spremenljivke** meje razredov lahko določamo na dva načina. K številčni oznaki meje dodamo besedo **pod** ali **nad.** Razlika med opredeljenima načinoma označevanja je mejna vrednost, ki je lahko vključena v spodnji ali zgornji razred (Šadl, 2001).

Primer: Razredi so lahko določeni:

I. možnost	**II.** možnost	**III.** možnost
Znesek v EUR	Znesek v EUR	Znesek v EUR
pod 20 20 do pod 40	do 20 nad 20 do 40	do 20 20 do 40

Skušajte določiti za vse možne načine določanja mej, prikazanih v zgornji tabeli, kam bi vključili prodajo v vrednosti 20 EUR.

I. primer: prodaja v vrednosti 20 EUR bi bila vključena v 2. razred,
II. primer: prodaja v vrednosti 20 EUR bi bila vključena v 1. razred,
III. primer: meje razredov niso jasno določene, zato bi bila vrednost 20 lahko vključena v oba razreda.

Za **diskretne** spremenljivke, ki lahko zavzemajo le cele vrednosti, lahko meje določamo s sredino prvega in sredino zadnjega razmika. Ker so vrednosti lahko le cela števila, jih ne označujemo s sredino, ampak navajamo le cele vrednosti.

Npr. za število gostov ali za število nočitev določamo meje razredov:

do 100
101 do 200
201 do 300

Zaradi večje nazornosti pri prikazovanju včasih razredom meja ne določimo. Takrat govorimo o **odprtih** razredih. Običajno se določajo za zelo razpršene majhne ali največje vrednosti opazovane populacije. Zaradi odprtih razredov ne moremo izračunati vseh parametrov za opazovani pojav, ker ne moremo izračunati sredine razreda. Širino razreda večkrat na osnovi predvidevanj o meji odprtega razreda ocenimo. Npr. gostje ostajajo v hotelu 1 noč, 2 noči, 3 noči, 1 teden, več. Zadnji razred je odprt navzgor.

6.2.2 Število in širina razredov

Da bi ugotovili pravo značilnost opazovanega pojava je potrebna odločitev opazovalca o **številu** in **širini** razredov. Znanje statistične teorije in cilja raziskave je temelj za pravilno odločitev. Izbira se med večjim številom razredov, kar omogoča prikazovanje natančnejših podatkov, ali manjšim številom razredov, kar omogoča večjo preglednost. Preozki razredi in preveliko število zaradi prevelikih podrobnosti moti preglednost. Če so preširoki in jih je premalo, se zaradi pregrobega grupiranja preveč podrobnosti izgubi. Običajne širine so 2, 3, 5, 10. Optimalno število razredov je okoli 10. Ko določimo širino razreda, lahko izračunamo število razredov tako, da delimo razliko med največjo in najmanjšo vrednostjo, ki se pojavlja v populaciji, z izbrano širino razreda. Dobljeni količnik zaokrožimo na naslednje višje celo število.

Razmislite, kako bi določili pravo število razredov v različnih primerih!

Izbira **enake širine** razredov je boljša zaradi možne neposredne primerjave podatkov in enostavnejšega izračunavanja statističnih parametrov. Izberemo jih lahko, kadar ni velikih razlik med vrednostmi opazovane spremenljivke. Če so vrednosti zelo različne, predvsem če se vrednosti manjšega števila enot zelo razlikujejo od vrednosti večine enot v populaciji, moramo oblikovati razrede z **neenako** širino. Včasih pri oblikovanju razredov z neenakimi širinami izhajamo iz stalnega razmerja med spodnjo in zgornjo mejo razreda. Pri neenako širokih razredih izračunamo **gostoto frekvence, g_j,** tako, da frekvenco danega razreda delimo s širino razreda (Šadl, 2001).

$$g_j = \frac{f_j}{d_j}$$

Razmislite, kdaj je smiselno oblikovati enake ali različne širine razredov!

Širino razreda izračunamo kot razliko med spodnjo in zgornjo mejo razreda.

$$\mathbf{dj = y_{j,max} - y_{j,min}}$$

Pri **diskretnih** spremenljivkah in pri številskih spremenljivkah, ki se **zaokrožujejo na najbližjo** vrednost, pri izračunavanju širine upoštevamo **popravek** za zveznost, ki ga od spodnje meje oštevamo in k zgornji meji prištevamo.

Sredina razreda je predstavnik vseh vrednosti v razredu, saj v frekvenčni porazdelitvi z razporeditvijo enot v razrede njihovih posamičnih vrednosti ne poznamo. Izračunamo jo kot povprečje med zgornjo in spodnjo mejo razreda.

$$\mathbf{y_j = \frac{(y_{j,max} + y_{j,min})}{2}}$$

Ponovite matematična pravila računanja in zaokroževanja števil!

Pri računanju je potrebno paziti na vrsto spremenljivke in način zaokroževanja (Košmelj, 1999).

	Zvezne	***Diskretne***	***Zaokrožena na najbližjo vrednost***	*Zaokrožene na največjo celo vrednost*
Meje	250 do pod 255	**250 do 254**	**250 do 254**	250 do 254
dj – širina (min-max)	5	**5**	**5**	5
Yj - sredina	(250+255)/2= 252,5	**252**	**252**	252,5

6.2.3 Sestavljanje frekvenčnih porazdelitev

Frekvenčno porazdelitev individualnih vrednosti oblikujemo tako, da po vrsti, vodoravno ali navpično, napišemo posamezne vrednosti, poleg vsake pa njeno frekvenco. Frekvence povedo število enot populacije, ki so dosegle določen rezultat. Pri tako oblikovani frekvenčni porazdelitvi so frekvence nizke, lahko pa se nekatere vrednosti v opazovani populaciji sploh ne pojavijo in frekvence ni. Vrsta je dolga in nepregledna. Zato se običajno oblikujejo **frekvenčne porazdelitve grupiranih vrednosti** spremenljivk.

Razmislite, kako oblikujete frekvenčne porazdelitve ročno. Ponazorite način oblikovanja!

Ročno razporejanje enot, ki je zamudno in je verjetnost napak velika, danes nadomeščajo programi, ki omogočajo hitro in natančno razporejanje velikega števila enot v razrede. Namesto risanja tabel, v katere so se ročno s črtkanjem vnašali podatki za vsako enote, to delo opravi računalnik. Zato je potrebno poznati ustrezne **računalniške programe** in statistično teorijo. Za celotno populacijo, ki je sestavljena iz velikega števila enot, bomo računalniško hitro in pravilno sestavili frekvenčno porazdelitev le, če pravilno pripravimo podatke. Podatki morajo biti pravilno vneseni in prikazani v navpični statistični vrsti. Glede na najmanjšo in največjo vrednost določimo število razredov, širino razredov in postavimo nedvoumne meje. Z izbiro ustrezne funkcije (FREQUENCY) računalnik razvrsti enote v razrede, jih prešteje in izpiše frekvenco, to je število enot v razredu. Če seštejemo vse frekvence v vseh razredih, dobimo število enot celotne populacije. Velika grška črka sigma (Σ) je simbol za vsoto (Kragelj, 1999).

6.2.4 Analiza frekvenčne porazdelitve

Za pravilno analizo podatkov, prikazanih v frekvenčni porazdelitvi, je potrebno poznati pojme.

Frekvenca je število enot v posameznem razredu. Označimo jo s **fj**. Vsota frekvenc je enaka številu enot N v opazovani populaciji.

$$\Sigma f_j = N$$

Relativna frekvenca izraža delež enot v posameznem razredu. Izračunamo jo kot razmerje med frekvenco in številom enot celotne populacije. Izračuna se enako kot strukturni delež. Vsota relativnih frekvenc je celota in je izražena z 1.

$$f_j^\circ = \frac{f_j}{N}$$

Ponovite matematična pravila izračunavanja in oblikovanja formul za izračunavanje z računalnikom!

Če frekvence postopno seštevamo, dobimo **kumulativo frekvenc F_j.**

$$F_j = F_{j-1} + f_j$$

Z uporabo računalnika je izračunavanje kumulative enostavno in hitro, če pravilno oblikujemo formulo in jo kopiramo za celotno tabelo. Kumulativa frekvenc nam pove število enot, ki imajo manjšo vrednost od zgornje meje razreda.

Če postopno seštevamo relativne frekvence, dobimo **kumulativo relativnih frekvenc,** ki pove delež enot, ki imajo manjšo vrednost od zgornje meje razreda, na katerega se kumulativa nanaša.

$$F_j^\circ = F_{j-1}^\circ + f_j^\circ$$

6.2.5 Grafično prikazovanje frekvenčne porazdelitve

Grafično prikazujemo frekvenčne porazdelitve z linijskimi ali stolpčnimi grafikoni. Prikaze s stolpci imenujemo **histogram**, prikaze z linijskim grafom pa **poligon**.

Histogram je sestavljen iz stolpcev (**pravokotnikov**), katerih širina ustreza širini razredov, višina pa frekvenci posameznega razreda. Prikazan je v prvem desnem zgornjem kvadrantu koordinatnega sistema. Na abscisni (vodoravni) osi x so prikazane meje razredov za prikazane vrednosti spremenljivke, na ordinatni (navpični) y osi pa lestvica frekvenc razredov. Če se vrednosti na abscisni osi ne pričnejo z 0, to označimo z prekinjeno abscisno osjo. Včasih na desni strani označimo tudi skalo za relativne frekvence. Pri risanju grafikona je pomembno razmerje med širino in višino. Pri nepravilnem razmerju nam grafikon ne prikaže realnega naraščanja pojava. Višina stolpca pri največji frekvenci naj bo od 60 % do 80 % širine slike. Zaradi nazornosti prikaza se odločamo med risanjem vodoravnih in navpičnih mrežnih črt, risanjem stolpcev ali le zunanjega obrisa pravokotnika. Če želimo poudariti značilnost pojava, je v vsaki statistični analizi potrebno ne le izbrati pravi grafikon, ampak ga tudi ustrezno oblikovati.

V statističnih publikacijah si oglejte grafično prikazovanje frekvenčnih porazdelitev!

Tudi pri prikazovanju frekvenčnih porazdelitev z **neenakimi širinami razredov** mora biti **ploščina stolpca** sorazmerna s številom enot, ki so vključene v posamezni razred kot pri histogramih z enako širokimi razredi. Da to dosežemo, na ordinatni osi upoštevamo gostoto frekvenc in ne število enot. Poligon prikazuje stolpce, ki niso enako široki, njihova višina je enaka gostoti frekvence, ploščina pa frekvenci.

Linijski graf, s katerim prikazujemo frekvenčno porazdelitev, se imenuje **poligon**. Je mnogokotnik, ki je narisan tako, da so nanesene točke na sredini razmika posameznega razreda, katerega ordinata (oddaljenost od abscisne osi) je sorazmerna frekvenci razreda. Poligon mora biti označen s sklenjeno črto, zato označimo razred, ki je pred najnižjim in za najvišjim razredom. Ker sta frekvenci praznih razredov 0, sta ustrezni točki na abscisni osi na sredini praznih razredov. Običajno pomožne črte niso narisane. Poligon je boljši prikaz od histograma, ker se bolj približa stvarni razmestitvi rezultatov. Histogram predpostavlja, da so rezultati enakomerno razmeščeni v posameznem razredu. Poligon pa prikazuje večjo gostoto rezultatov na tisti strani razrednega razmika, ki je bližji mestom z največjo gostoto v celotni frekvenčni porazdelitvi. Primernejši so tudi za primerjavo dveh ali več frekvenčnih porazdelitev, ki so prikazane v istem grafikonu.

Grafično prikazovanje nazorneje prikazuje gostitev pojava v posameznih razredih kot prikaz podatkov v tabeli. Pri risanju grafikonov z računalniškimi programi moramo poznati pravila označevanja podatkov, izbire prave vrste grafikona in pravilnega oblikovanja.

Kumulative frekvenc in **relativnih frekvenc prikazujemo v linijskem grafikonu**. Na abscisni osi so prikazane vrednosti spremenljivke in so razvidne širine razredov, na ordinatni osi pa kumulative frekvenc ali kumulative relativnih frekvenc. Označene točke kumulativ frekvenc so povezane v daljico, ki nam kaže enote, ki imajo vrednost do zgornje meje razredov. Če v istem grafikonu prikazujemo kumulativo frekvenc in relativnih frekvenc hkrati, imamo dve skali, na levi strani skalo za kumulativo frekvenc, na desni strani pa skalo kumulative relativnih frekvenc. Grafični prikaz kumulative frekvenc imenujemo tudi **ogiva**, ki ima pogosto obliko črke S. Čeprav lahko iz grafikona razberemo le približno vrednost, je njegov analitični pomen velik, saj hitro in jasno prikaže značilnosti opazovanega pojava in je odlično dopolnilo tabelaričnemu prikazu (Šadl, 2001).

6.2.6 Oblike frekvenčnih porazdelitev

V analizah nas predvsem zanima, kakšna je razporeditev opazovanih enot in kolikšna je gostitev rezultatov. Te značilnosti nazorneje prikazujejo grafični prikazi frekvenčne porazdelitve kot tabelarični.

Opazovane enote so razporejene **simetrično**, če je ena gostitev na sredini in frekvence enakomerno padajo od središča gostitve na levo in desno. Simetrični razporeditvi, ki je zvonaste oblike, pravimo tudi normalna porazdelitev. Več o njej bo razloženo v poglavju 7 Variabilnost.

Razporeditvi z enim središčem gostitve rezultatov pravimo **enovrhna** ali **unimodalna**. Lahko sta dve središči gostitve (**bimodalna**) in tudi več gostitev (**polimodalna**) frekvenčna porazdelitev. Porazdelitve imajo lahko tudi obliko črke J ali U (Trstenjak, 2001).

V poslovnih poročilih večinoma prikazujemo zbrane podatke s frekvenčnimi porazdelitvami. Glede na cilj analize in vrsto spremenljivk (opisne ali numerične) oblikujemo skupine ali razrede. Zagotoviti moramo enovitost in enoličnost skupin in razredov s pravilno določenimi mejami in ustrezno širino razredov (enako ali različno). Za določanje parametrov izračunavamo sredino in gostoto razredov. Za nazornost je pomembna izbira ustreznega števila skupin ali razredov ter pravilni grafični prikaz s poligonom ali histogramom. Za razlago je potrebno poznavanje pojmov frekvenca, relativna frekvenca, kumulativa frekvenc in kumulativa relativnih frekvenc.

Naloge:

1. Oblikujte skupine za primere gostov ali nočitev v hotelu.
2. Oblikujte razrede za primere gostov ali nočitev v hotelu! Določite meje za zvezne in nezvezne numerične spremenljivke! Izračunajte širino in sredino razredov.
3. Oblikujte meje, izračunajte sredino in širino razredov za primer številskih spremenljivk, ki se zaokrožujejo na največjo in na približno vrednost.
4. Na primeru izračunajte gostoto frekvence pri neenako širokih razredih in razložite razliko med frekvenco in gostoto.
5. Na primeru določite frekvence, relativne frekvence in analizirajte dobljene rezultate.
6. Na primeru izračunajte kumulative frekvenc in relativnih frekvenc in analizirajte razlike in dobljene rezultate.
7. Grafično prikažite podatke v histogramu in poligonu! Primerjajte grafična prikaza.
8. Iz primera izračunajte kumulative frekvenc in jih grafično prikažite! Ogiva.

7 SREDNJE VREDNOSTI

Srednje vrednosti se v statističnih analizah zelo pogosto uporabljajo. Čeprav so vrednosti spremenljivk pri večini opazovanih enotah različne, za opisovanje značilnosti proučevane populacije praviloma ne navajamo vseh enot in vseh vrednosti, ampak izmed njih izberemo neko vrednost, ki bo predstavnik vseh opazovanih enot. So enostavne za izračunavanje in lahko dobro predstavljajo vse enote opazovane populacije. Srednja vrednost naj bi bila takšna, da se večina opazovanih enot od nje ne bi veliko razlikovala. Izbiramo med več vrstami srednjih vrednosti glede na cilj in namen analize. Za vsako analizo presodimo, katera srednja vrednost je najprimernejša za prikaz parametrov. Obstaja nevarnost, da se ne prikažejo prave značilnosti populacije, saj se s prikazom vrednosti le ene enote, ki predstavlja vse enote, ne prikazujejo posebnosti posameznih enot in njihove različnosti. Zato je dobro, da v analizi izračunane srednje vrednosti dopolnimo z drugimi izračuni, kot so relativna števila in variabilnost. Vsem zelo dobro poznana in v vsakdanjem življenju ne le v statističnih analizah uporabljena srednja vrednost je aritmetična sredina, ki jo bolj poznamo kot povprečje. Povprečje je zelo pogosto uporabljena osnova za primerjav,. npr. povprečna zasedenost v hotelu, povprečna doba bivanja, povprečno število gostov, povprečna plača, povprečne cene, povprečna količina padavin, povprečna onesnaženost zraka itd.

V gostinstvu in turizmu večina analiz in primerjav temelji na izračunavanju srednjih vrednosti, zato je poznavanje načinov izračunavanja, uporabnosti, lastnosti in slabosti izračunanih rezultatov nujno za pridobivanje ustreznih informacij za dobre poslovne odločitve. Izračunavanje je z uporabo računalniških programov olajšano, saj se izognemo težjim matematičnim operacijam. Zato pa je potrebno znanje ustreznih programov, ki omogočajo izračun in grafično prikazovanje.

Poznamo več vrst srednjih vrednosti, ki imajo različne lastnosti in vsaka od njih na svoj način odkriva tisto, kar je tipično za opazovane enote. Srednje vrednosti imenujemo tudi tipične predstavnike, ki so dobri le, če se posamične enote ne razlikujejo veliko od njene osrednje vrednosti. Izračunavamo jih lahko le za številske spremenljivke. Največkrat se uporabljajo: aritmetična sredina, mediana, modus, harmonična sredina in geometrijska sredina.

7.1 ARITMETIČNA SREDINA - POVPREČJE

Najbolj znana in največkrat uporabljena srednja vrednost je **aritmetična sredina ali povprečje**. Izračunamo jo iz vseh opazovanih enot tako, da vsoto posameznih vrednosti številskih spremenljivk delimo s številom enot opazovane populacije.

$$\textbf{Aritmetična sredina} = \frac{\textbf{vsota posamičnih vrednosti}}{\textbf{število enot opazovane populacije}}$$

$$\bar{y} = M = \frac{1}{N}\sum_{i=1}^{N} y_i$$

M - aritmetična sredina
y_i - vrednost ene enote
N - število enot opazovane populacije

Aritmetična sredina kot predstavnik enot bi bila enaka vrednosti opazovanih enot ob predpostavki, da bi vse enote opazovane populacije dosegle enako vrednost.

Y = M x N

Npr. v hotelu bi v enem tednu imeli vsak dan enako število gostov, to je 70. Aritmetična sredina bi bila tudi 70 in bi bila enaka vrednosti posamične enote. Običajno pa populacijo sestavljajo enote, ki dosegajo različno vrednost, zato se aritmetična sredina od njih bolj ali manj razlikuje. Npr. v hotelu imamo v treh dne število gostov 10, 15, 20, če te številke enakomerno porazdelimo, dobimo število 15 za vsak dan. Aritmetična sredina je 15.

Zato lahko sklepamo, da če bi bile vrednosti spremenljivk pri vseh enotah opazovane populacije enake, bi bila enaka tudi aritmetična sredina, sicer je aritmetična sredina različna.

Ponovite pravila računanja srednjih vrednosti!

Iz posamičnih vrednosti je izračun aritmetične sredine enostaven. Pri večjem številu opazovanih enot je ročno izračunavanje zamudno, zato je prednost uporaba računalniških programov, ki omogočajo hitro izračunavanje z uporabo **formule ali funkcije.** V programu Excel je to funkcija **AVERAGE**.

V gostinstvu in turizmu se aritmetična sredina zelo pogosto izračunava in izraža povprečje za različne parametre in različna obdobja. Izračunavamo povprečno število gostov na dan, nočitev v določenem obdobju, zasedenost hotela, povprečno dobo bivanja, povprečne plače zaposlenih, povprečno porabo na gosta ali nočitev itd.

Če so vrednosti številskih spremenljivk razvrščene v **frekvenčno porazdelitev**, so individualne vrednosti nepoznane, zato aritmetično sredino **ocenimo**. Za posamezni razred poznamo frekvenco, to je število enot, ki imajo vrednosti med zgornjo in spodnjo mejo razreda, ne poznamo pa posamičnih vrednosti. Zato lahko pri izračunu upoštevamo le sredino razreda kot predstavnika vrednosti vseh enot, ki so vključene v posamezni razred.

$$\mathbf{M^* = M = \frac{1}{N}\sum_{j=1}^{k} y_j f_j}$$

M* - ocenjena aritmetična sredina
N - število enot
y_j - sredina razreda
f_j - frekvenca razreda

$$\text{Ocenjena aritmetična sredina} = \frac{\text{vsota ocenjenih vrednosti}}{\text{število opazovanih enot}}$$

Ocenjena vrednost razreda = sredina razreda x frekvenca tega razreda (ponder)

Ponovite uporabo funkcij v programu Excel!

Aritmetična sredina iz frekvenčne porazdelitve, ki se imenuje tehtana (ponderirana) aritmetična sredina, se izračuna tako, da:

- izračunamo sredino vsakega razreda,
- sredine razredov pomnožimo s frekvencami v vsakem razredu,

- produkte seštejemo in delimo s številom enot populacije.

Tako izračunana aritmetična sredina se nekoliko razlikuje od aritmetične sredine, izračunane iz posamičnih vrednosti.

Naloge:

1. Izračunajte iz primera aritmetično sredino iz individualnih vrednosti z uporabo formule in funkcije AVERAGE! Primerjajte dobljena rezultata.
2. Iz primera frekvenčne porazdelitve izračunajte ocenjeno aritmetično sredino.
3. Posamične podatke razporedite v frekvenčno porazdelitev in izračunajte aritmetično sredino iz individualnih podatkov in iz frekvenčne porazdelitve.
4. Primerjajte dobljena rezultata aritmetične sredine in ocenjene aritmetične sredine.

7.2 MEDIANA

Mediana ali središčna vrednost je tista srednja vrednost, od katere je polovica enot (50 %) opazovane populacije večjih in polovica enot manjših. Mediana leži točno na sredini ranžirne vrste, ki prikazuje po velikosti razporejene podatke. Rang nam pove mesto mediane v ranžirni vrsti. Ustreza kvantilnemu rangu P = 0,50 in je enaka drugemu kvartilu. Njena prednost je v možnosti določanja tudi pri frekvenčni razporeditve z odprtimi razredi. Slabost je neupoštevanje vrednosti majhnih in večjih enot opazovane populacije.

Računamo jo lahko za posamezne vrednosti in za frekvenčne porazdelitve.

7.2.1 Mediana iz posameznih vrednosti

Za določitev mediane ročno je potrebno vrednosti številskih spremenljivk razporediti v ranžirno vrsto, to je po velikosti od najmanjše do največje ali obratno. S tem določimo vsaki enoti zaporedno mesto. To mesto imenujemo rang, ki je označen z R. Rang ima v ranžirni vrsti vrednosti od 1 do N (števila enot opazovane populacije). Oblikovanje ranžirne vrste z uporabo računalniških programov je hitro.

Rang mediane lahko izračunamo:

$$\mathbf{R = \frac{(N+1)}{2}}$$

Ker je mediana tisto število, ki stoji točno na sredini vseh vrednosti, smo izračunali sredino tako, da smo številu enot prišteli 1 in delili z 2. Rang lahko izračunamo tudi iz znanega kvantilnega ranga za mediano P = 0,50 po obrazcu:

$$\boldsymbol{R = N \times P + 0{,}5}$$

$$\boldsymbol{P = \frac{R - 0{,}5}{N}}$$

P = kvantilni rang

Če je število opazovanih enot populacije **liho,** je mediana enaka številu, ki stoji točno na sredini ranžirne vrste. Če je število opazovanih enot **sodo,** mediana ni celo število, ampak povprečje, izračunano iz vrednosti tistih dveh števil, ki sta najbližja izračunani sredini v ranžirni **vrsti.**

Razmislite, kje ste določili mediano in kako!

Hitro in enostavno je določanje mediane z uporabo ustreznih računalniških programov. V programu Excel je to **funkcija MEDIAN.** Za pravilno določitev je potrebno poznati pravila uporabe te funkcije. Podatkov ni potrebno oblikovati v ranžirne vrste, niti računati ranga. Potrebno pa je paziti na pravilno označitev podatkov, iz katerih želimo, da računalnik izbere mediano. Ker v podjetju praviloma obdelujejo manjše število podatkov, je določanje mediane na ta način zelo pogosto.

V grafikonu ranžirne vrste lahko mediano določimo tudi grafično.

7.2.2 Mediana v frekvenčni porazdelitvi

Pri populacijah z velikim številom enot je kljub možnosti razporejanja z računalnikom sestavljanje ranžirne vrste zamudno, predvsem pa nepregledno. Ker nimamo posamičnih podatkov, prikazanih v ranžirni vrsti, ampak so ti razporejeni v frekvenčni razporeditvi, mediane ne moremo določiti točno, ampak lahko izračunamo **oceno** za mediano. Pri izračunavanju ne izhajamo iz individualnih vrednosti, ampak iz kumulative frekvenc. Rang, kjer so enote razporejeno v frekvenčno razporeditev, izračunamo enako kot za posamezne vrednosti. Podatek za rang nam nadomesti kumulativa, ki nam pove število enot, ki imajo manjšo vrednost od zgornje meje razreda. Po kumulativi frekvenc ugotovimo, v katerem razredu je enota, ki ustreza izračunanemu rangu. To je razred, kjer je kumulativa frekvenc večja od ranga. Imenuje se **medialni razred**. Mediana je ocenjena vrednost med zgornjo in spodnjo mejo medialnega razreda in jo izračunamo po obrazcu:

$$\boldsymbol{Me = y_{j,min} + d_j \times \frac{R - F_{j-1}}{f_j}}$$

Mediana (ocenjena) = spodnja meja medialnega razreda + širina razreda x rang – kumulativa frekvenc pred medialnim razredom / frekvenca medialnega razreda

Oceno za mediano lahko določimo tudi iz grafičnega prikaza kumulative frekvenc, kjer mediana ustreza kumulativi relativnih frekvenc 0,5.

Naloge:

1. Iz posamičnih podatkov oblikujte ranžirno vrsto, izračunajte rang in določite mediano za primer sodega in lihega števila enot. Ugotovite razliko.
2. Iz posamičnih podatkov prve tabele izračunajte mediano z uporabo funkcije MEDIAN in primerjajte dobljene rezultate.
3. Podatke prikažite v linijskem grafikonu in določite mediano grafično.

4. Posamične podatke (iz prve tabele) razporedite v frekvenčno porazdelitev, izračunajte rang, določite medialni razred in izračunajte ocenjeno vrednost za mediano. Primerjajte dobljene rezultate.
5. Podatke frekvenčne razporeditve prikažite v grafikonu in ocenite mediano.

7.3 MODUS

Modus je srednja vrednost, ki kaže gostitev pojava tako, da opazujemo katera je tista vrednost, ki se v opazovani populaciji največkrat pojavi. Npr. če opazujemo podatke o številu gostov v hotelu po dnevih, je modus številka, ki se je največkrat ponovila.

Iz posameznih vrednosti ročno določamo modus tako, da pregledamo vse podatke in preštejemo, katera vrednost se največkrat ponovi. Pri velikem številu opazovanih enot in pri majhnih razlikah v vrednostih posameznih enot je to delo zamudno in možnost napak velika. Zato danes uporabljamo računalnike, ki z različnimi programi omogočajo hitro in natančno določanje modusa. V programu Excel uporabljamo **funkcijo MODE**. Za pravilno določanje potrebujemo znanje programa in statistične teorije. Ker za pripravo analiz in poslovnih načrtov v podjetjih najpogosteje poznamo posamične vrednosti proučevanega pojava, je določanje modusa z uporabo funkcije MODE zelo uporabno.

V statistični vrsti se modus lahko določi le, če se iste vrednosti podatkov večkrat pojavijo. Če so posamične vrednosti enot različne, **modusa** v opazovani populaciji **ni**. Npr. če opazujemo število gostov v hotelu v enem tednu ali v enem mesecu, je verjetnost, da imamo vsak dan različno število gostov. Takrat modusa ni mogoče določiti, ker se opazovane številske vrednosti ne ponovijo.

Razmislite, kako bi določili modus iz podatkov, razporejenih v različnih seznamih, npr. športni rezultati!

Če ne razpolagamo z individualnimi podatki, ampak s frekvenčnimi porazdelitvami, **modus ocenimo**. Določamo ga na osnovi frekvence razredov. Izračunamo ga lahko le iz frekvenčne porazdelitve z enako širokimi razredi. Če širine razredov niso enako široke, po frekvencah ne moremo določati gostitve vrednosti. Razred, v katerem je frekvenca največja, imenujemo **modalni razred.** V modalnem razredu izračunamo ocenjeno vrednost modusa z obrazcem:

$$Mo = y_{j,min} + d_j \times \frac{f_j - f_{j-1}}{2f_j - f_{j-1} - f_{j+1}}$$

Mo - ocenjena vrednost modusa
$y_{o,min}$ - spodnja meja modalnega razreda
d_o - širina razreda
f_o - frekvenca modalnega razreda
f_{-1} - frekvenca pred modalnim razredom
f_{+1} - frekvenca za modalnim razredom

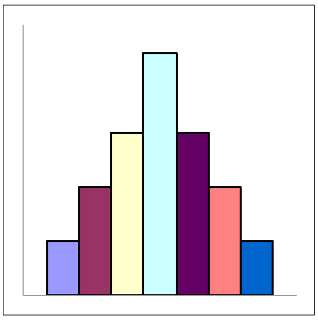
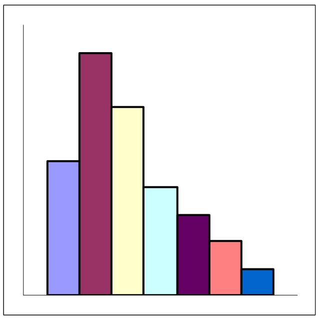
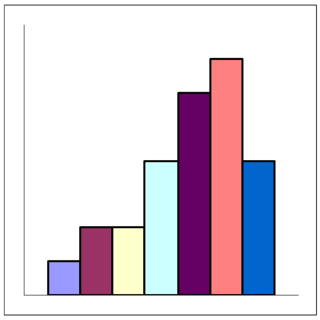
Vrednost za modus lahko ocenimo tudi v histogramu frekvenčne porazdelitve.

Naloge:

1. V statistični vrsti določite modus ročno in z uporabo funkcije MODE.
2. Podatke statistične vrste prikažite v frekvenčni porazdelitvi, določite modalni razred in ocenite modus. Ocenjeno vrednost modusa primerjajte z dejanskim modusom.
3. Izberite statistično vrsto, v kateri modusa ni mogoče določiti.

7.4 ODNOS MED ARITMETIČNO SREDINO, MEDIANO IN MODUSOM

Tipični predstavnik enot opazovane populacije je lahko aritmetična sredina, mediana ali modus. Za te tri najpogosteje izračunane srednje vrednosti je potrebno dobro poznati razlike in njihovo pravilno rabo, saj imajo lahko izračunani rezultati iz istih podatkov enake, podobne ali zelo različne vrednosti. To je odvisno od značilnosti opazovane populacije in oblike frekvenčne porazdelitve. Za porazdelitev enot, ki so **simetrične z eno gostitvijo (unimodalne), so vse tri srednje vrednosti enake**. Pri asimetričnih porazdelitvah se med seboj razlikujejo (Šadl, 2001).

PORAZDELITEV		
SIMETRIČNA Od modusa ko je gostitev enot največja, je upadanje gostitve v obe smeri enakomerno.	**ASIMETRIČNA V DESNO** Upadanje gostitve je od modusa v desno počasnejše.	**ASIMETRIČNA V LEVO** Upadanje gostitve je od modusa v levo počasnejše
Mo = Me = M	Mo <Me <M	Mo > Me >M
		

Slika 8: Primerjava porazdelitev
Vir: Lastni

Pri izbiri srednje vrednosti upoštevamo zahtevo, da mora srednja vrednost izražati tisto, kar je tipično za pojav in da mora biti njena vrednost čim bližja vrednosti večine opazovanih enot. Problema izbire pri simetričnih porazdelitvah ni, saj dosežejo vsi trije predstavniki enako vrednost. Pri **asimetričnih porazdelitvah** se dobljeni rezultati razlikujejo. V takih porazdelitvah sta boljša predstavnika vseh opazovanih enot **mediana in modus**, ker aritmetična sredina vključuje vse vrednosti tudi skrajne, ki lahko vplivajo na izračun prevelike ali premajhne aritmetične sredine.

Z uporabo računalnika in ustreznih programov izračunavanje srednjih vrednosti ni več zamudno, zato se za boljšo analizo opazovanih enot priporoča izračunavanje **vseh treh srednjih vrednosti**. Iz **primerjave** dobljenih rezultatov lahko sklepamo tudi o značilnosti razporeditve enot. Za pravilno izračunavanje je potrebno poznati formule in računalniški program. Za pravilno analizo pridobljenih rezultatov je potrebno poznati statistično teorijo o značilnostih aritmetične sredine, mediane in modusa. Čeprav se najpogosteje uporablja povprečje ali aritmetična sredina, je za pravilne poslovne odločitve pomembno izračunavanje tudi ostalih tipičnih predstavnikov zaradi njihovih prednosti.

Naloge:

1. Iz primera izračunajte in določite M, Me, Mo! Primerjajte dobljene rezultate! Analizirajte razlike in določite vrsto porazdelitev.
2. Iz istega primera z računalnikom in funkcijami izračunajte M, Me, Mo. Primerjajte postopek in dobljene rezultate.

7.5. HARMONIČNA SREDINA – H

Zelo pomembna je odločitev, katera srednja vrednost je pravilni predstavnik vseh vrednosti opazovanih enot. Če ne poznamo statistične teorije, lahko pridemo do napačnih rezultatov in neustreznih poslovnih odločitev, kljub matematično pravilno izračunanemu rezultatu. Čeprav se aritmetična sredina, predvsem zaradi enostavnosti, v statistiki največ uporablja, pa ni vedno primerna za prikazovanje povprečja. Uporabljamo jo za izračunavanje povprečja, če poznamo številske vrednosti opazovanih enot, npr. število gostov, število nočitev, vrednost prodaje, zneske stroškov, cene itd. Za izračunavanje povprečja iz relativnih števil pa ni primerna, saj ne da pravilnih rezultatov. V statistični analizi se uporablja **harmonična sredina predvsem za računanje povprečij iz strukturnih odstotkov** (deležev ali odtisočkov) in **statističnih koeficientov**.

Harmonična sredina je enaka recipročni vrednosti aritmetične sredine in je izračunana iz recipročnih vrednosti spremenljivk.

$$H = \frac{N}{\sum_{i=1}^{N} \frac{1}{y_i}}$$

Ponovite pravila računanja, ki ste jih spoznali pri matematiki!

Zaradi reševanja dvojnih ulomkov je ročno izračunavanje harmonične sredine dokaj zahtevno, zato so jo mnogi nadomeščali z izračunavanjem aritmetične sredine. Dobljeni rezultati niso enaki. Pri računanju povprečja iz istih podatkov po formuli **aritmetične sredine je dobljeni rezultat večje od rezultata harmonične sredine**. Z uporabo računalnikov in ustreznih programov je izračun enostaven, hiter in natančen. V programu Excel se uporablja **funkcija HARMEAN.**

Ob poznavanju statistične teorije ter z znanjem uporabe računalniškega programa lahko hitro in enostavno izračunamo pravilne rezultate, ki so osnova za ustrezne analize in dobre poslovne odločitve.

Naloge:

1. Izberite statistične podatke na spletu.
2. Z računalnikom izračunajte aritmetično in harmonično srednjo vrednost! Primerjajte dobljena rezultata! Razložite kateri je pravilen in zakaj?
3. Iz statističnih koeficientov izračunajte harmonično srednjo vrednost in razložite rezultat.

4. Iz strukturnih odstotkov izračunajte harmonično srednjo vrednost in razložite rezultat.

7.6 GEOMETRIJSKA SREDINA - G

Geometrijska sredina za številske spremenljivke je N-ti koren iz produkta vseh N vrednosti.

$$\overline{V} = \sqrt[N]{V_1 \times V_2 \times V_3 \times ..... \times V_N}$$
$$\overline{K} = \sqrt[N]{K_1 \times K_2 \times K_3 \times ... \times K_N}$$

Pri izračunavanju je potrebno poznati zahtevno matematično operacijo logaritmiranja. Zato se mnogokrat ni uporabljala, ampak so jo nadomestili z izračunavanjem aritmetične sredine. Dobljena rezultata, čeprav izračunana iz istih podatkov, vendar zaradi uporabe drugačnih matematičnih postopkov, se razlikujeta. Aritmetična sredina je večja od geometrijske.

Ponovite pravila računanja, ki ste jih spoznali pri matematiki!

Geometrijska ali geometrična srednja vrednost se uporablja za izračunavanje povprečja iz relativnih števil, kot so **verižni indeksi, koeficienti rasti in povprečna stopnja rasti**. Povprečne stopnje rasti ne moremo izračunati neposredno. Ker so v nekaterih obdobjih lahko enake 0 ali negativne, zato jih moramo najprej pretvoriti v verižne indekse ali koeficiente rasti. Iz teh šele lahko izračunamo povprečje, ki ga pretvorimo v povprečno stopnjo rasti.

Izračunani povprečni koeficienti (verižni indeksi, stopnje rasti) nam služijo za predvidevanja gibanja pojava v prihodnosti ob predpostavki, da se razmere ne bodo spremenile. Npr. če je bila v zadnjih petih letih povprečna rast nočitev 5%, lahko predvidimo enako gibanje tudi za naslednje obdobje.

Z uporabo računalnika in ustreznih računalniških programov je izračunavanje hitro, enostavno in natančno. V programu Excel za izračunavanje uporabljamo **funkcijo GEOMEAN**.

Če iz istih podatkov izračunamo povprečje po treh matematičnih postopkih (aritmetični, geometrijski, harmonični sredini), dobimo različne rezultate. Odnos med povprečji izračunanimi iz istih podatkov, je: **H<G<M**, kar pomeni, da je harmonična sredina najmanjša in aritmetična največja.

Zaradi različnih rezultatov za M, H in G, ki jih izračunamo iz istih podatkov, je pomembno izbrati pravo vrsto povprečja. Na izbiro vpliva vrsta podatkov, iz katerih želimo izračunati povprečje. Kljub enostavnosti izračuna z računalnikom je pomembno dobro znanje statistike, da znamo ne le izbrati pravo vrsto predstavnika, ampak predvsem pravilno razložiti dobljeni rezultat.

Naloge:

1. Iz podatkov (verižnih indeksov, koeficientov dinamike) z računalnikom izračunajte geometrijsko srednjo vrednost.
2. Iz podatkov o številu nočitev izračunajte M, G, H, in primerjajte dobljene rezultate! Kateri je pravilen in zakaj?

3. Iz verižnih indeksov izračunajte M, G, in H in primerjate dobljene rezultate! Kateri je pravilen?

8 VARIABILNOST

Srednje vrednosti so v statistični analizi zelo pomembni parametri, vendar imajo pomanjkljivost, ker osvetljujejo populacijo le z enega vidika. Z njimi odkrivamo tisto, kar je tipično za pojav, zanemarijo pa posebnosti posameznih enot in njihovo različnost. Zato je potrebna dopolnitev prikaza, ki omogoča analizo različnosti enot in sprememb v času.

Značilnost vseh pojavov, ki jih proučujemo s statističnimi metodami je, da so sestavljeni iz enot, ki se po opazovanih lastnostih med seboj razlikujejo v opazovanem časovnem obdobju. Spreminjajo se tudi od obdobja do obdobja. Npr. gosti hotela na določen dan se razlikujejo v mnogih lastnostih, razlikujejo se tudi gosti hotela v lanskem in letošnjem letu itd. Zato je potrebno stalno statistično opazovanje. Glede na cilje določimo termine, ko podatke zberemo in evidentiramo.

Ugotavljanje **razlik** in spremljanje sprememb opazovanih enot omogočajo različne statistične metode. Vsak parameter odkriva te lastnosti na drugačen način. **Relativna števila** odkrivajo s primerjavami relativne razlike med enotami. Npr. strukturni deleži gostov po narodnosti prikazujejo različnost. Verižni indeksi nočitev po mesecih prikazujejo spreminjanje in različnost. Indeksi števila gostov z osnovo minimum, maksimum ali povprečje prikazujejo različnost enot.

Razliko med enotami lahko razberemo tudi iz frekvenčne porazdelitve. **Frekvence** po razredih so lahko zelo različne ali skoraj enake. Opazimo eno ali več gostitev ali enakomerno razporeditev, npr. v mesecu avgustu je število nočitev največje, v novembru najmanjše itd. Iz poligona in histograma nazorno razberemo različnost ali sorodnost opazovanih enot.

Tudi **srednje vrednosti**, ki predstavljajo vrednost okoli katere se gostijo ostale vrednosti, lahko predstavljajo dobro informacijo za odkrivanje **razlik** med enotami. Če poznamo vsaj še en podatek poleg srednje vrednosti, že lahko s primerjavo ugotavljamo razlike in razpone med največjo in najmanjšo vrednostjo enote. Npr. povprečno število gostov v hotelu v tekočem mesecu je 80, če ta podatek primerjamo z današnjim, že lahko sklepamo o različnosti.

To različnost ali variabilnost lahko ugotavljamo s primerjavo le dveh ali več podatkov. Najboljši so rezultati, ki jih izračunamo iz vrednosti vseh spremenljivk.

Kako se vrednosti spremenljivk med seboj razlikujejo, ugotavljamo z **merami variabilnosti**. Mere variabilnosti so parametri, s katerimi lahko analiziramo variabilnost samo pri številskih spremenljivkah. Večja kot je razlika med vrednostmi opazovanih enot, večja je variabilnost. Za ugotavljanje razlik lahko upoštevamo le dve skrajni vrednosti, to je najmanjšo in največjo ali vrednosti vseh opazovanih enot. Mera variabilnosti, izračunana samo iz dveh vrednosti, je variacijski razmik,izračunana iz vseh vrednosti pa varianca in standardni odklon.

8.1 VARIACIJSKI RAZMIK

Najbolj enostavna mera variabilnosti je **variacijski razmik ali razpon**, ki ga lahko izračunamo samo, če poznamo posamične vrednosti opazovanih spremenljivk. V frekvenčnih porazdelitvah, kjer posamezne vrednosti niso razvidne, te mere variabilnosti ni mogoče izračunati.

Variacijski razmik je razlika med največjo in najmanjšo vrednostjo opazovanih enot.

$$VR = y_{max} - y_{min}$$

Razmislite, kdaj izražati lastnosti populacije z razliko!

Večji variacijski razpon prikazuje večjo variabilnost opazovanih enot. Ker pri izračunu upoštevamo le dva podatka in zanemarimo razmestitev enot znotraj, ta parameter lahko ni dobra mera variabilnosti. Slabosti se pokažejo predvsem, kadar so skrajne vrednosti izredno majhne ali izredno velike. Npr. pri ugotavljanju variacijskega razmika o številu nočitev v Slovenije je leto 1991 izjemno, zato bi bil izračunan variacijski razmik, ki naj bi prikazoval variabilnost števila nočitev v Sloveniji v obdobju med letom 2008 in 1991, nerealno velik. To slabost lahko odpravimo tudi tako, da za izračunavanje razmika ne upoštevamo skrajnih vrednosti ali da upoštevamo le del populacije, npr. izključimo 10 % enot ali 10 % najmanjših in 10 % največjih vrednosti (DECILNI RAZMIK).

Ker je izračunavanje enostavno in hitro tudi brez uporabe sodobne računalniške opreme, se zelo pogosto uporablja. Npr. v podjetju izračunamo razliko med najvišjo in najnižjo plačo zaposlenih.

Variacijski razmik iz opazovanih podatkov lahko izračunamo tudi z računalnikom. Najprej s pomočjo **funkcij MAX (največja vrednost) in MIN (najmanjša vrednost)** iz množice podatkov izberemo največjo in najmanjšo vrednost opazovanih spremenljivk, nato oblikujemo **formulo** za izračun razlike.

Variacijski razmik dopolnjuje podatke srednjih vrednosti in drugih tipičnih predstavnikov enot.

8.2 VARIANCA

Varianca je najpomembnejša mera variabilnosti, ki se poleg aritmetične sredine najpogosteje uporablja v statističnih analizah. Prednost variance pred variacijskim razmikom je v tem, da pri izračunavanju upoštevamo vrednosti številskih spremenljivk vseh opazovanih enot. Za vsako enoto izračunamo njeno razliko od povprečja M. Ker imajo nekatere enote večje, druge pa manjše vrednosti od povprečja, bi bil rezultat vsote razlik nič, zato moramo za izračunavanje variance te razlike kvadrirati, da dobimo pozitivne vrednosti, ki jih lahko seštejemo. Dobljeni rezultat delimo s številom opazovanih enot.

Varianca je izračunana kot povprečje kvadratov odklonov vseh številskih spremenljivk od aritmetične sredine (Šadl, 2001).

$$VAR = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - M)^2$$

Računanje variance **ročno** je zelo zamudno in zahtevno, saj je potrebno podatke kvadrirati.

Izračunavanje variance z **računalnikom** je hitro in natančno, saj lahko v ustreznih računalniških programih uporabimo že oblikovane funkcije za izračun. V programu Excel je to funkcija **VARP**, ki izračunava varianco na podlagi podatkov za celotno populacijo. To funkcijo lahko uporabljamo, če poznamo posamične podatke celotne populacije, če izračunavamo varianco na osnovi vzorca, izberemo funkcijo **VAR**.

Slabost variance je, ker je izražena v merski enoti na kvadrat.

Pri izračunavanju **variance iz frekvenčne porazdelitve** ne poznamo posamičnih vrednosti, zato jih nadomestijo sredine razredov. Zato izračunamo razlike med sredinami razredov in aritmetično sredino. Te razlike kvadriramo in pomnožimo s frekvenco posameznega razreda. Produkte seštejemo in delimo s številom enot.

$$VAR = \sigma^2 = \frac{1}{N}\sum_{j=1}^{k} f_j (y_j - M)^2$$

Za varianco, izračunano iz frekvenčne porazdelitve, je potrebno izračunati in upoštevati **popravek** (Sheppardov popravek).

$$\sigma^2_{cor} = \sigma^2 - \frac{d^2}{12}$$

8.3 STANDARDNI ODKLON

Če varianco korenimo, izračunamo **standardni odklon**.

$$\sigma = SD = \sqrt{\sigma^2}$$

Prednost standardnega odklona je v tem, da je izražen v isti merski enoti kot opazovana številska spremenljivka. Zahteven matematični postopek korenjenja, ki je potreben pri ročnem izračunavanju standardnega odklona, nadomestimo z enostavnim in hitrim računanjem z uporabo računalnika. Računalniški programi imajo že pripravljene funkcije za izračunavanje standardnega odklona. V programu Excel sta to funkciji **STDEV** za ocenitev standardnega odklona vzorca in **STDEVP** za izračun standardnega odklona iz podatkov za celotno populacijo.

Standardni odklon nam pove gostitev pojava okrog srednje vrednosti. V **simetrični porazdelitvi** je standardni odklon približno ena šestina variacijskega razmika.

VR = 6 σ

Razmislite, kje se uporablja simetrična razporeditev! Ugotovite, pri katerih predmetih je omenjena.

Simetrična razporeditev, ki jo imenujemo tudi **normalna porazdelitev** enot, je poznana tudi pod imenom **Gaussova krivulja**. Zanjo je značilno, da imajo aritmetična sredina, mediana in modus enako vrednost. Enote se gostijo okoli srednje vrednosti in enakomerno padajo levo do najmanjše in desno do največje vrednosti. Na intervalu (M - 3σ) do (M + 3σ) je zajeta celotna populacija. Zunaj intervala se nahaja zelo malo vrednosti (Šadl, 2001).

8.4 RELATIVNE MERE VARIABILNOSTI

Večkrat želimo izračunano variabilnost posameznega pojava primerjati z variabilnostjo drugih pojavov. Primerjava je mogoča le, če sta variabilnosti obeh opazovanih pojavov izraženi v istih merskih enotah. Primerjava mer variabilnosti različnih opazovanih pojavov vedno ni mogoča in smiselna, kljub temu da so izražene v enaki merski enoti.

Primerjavo variabilnosti različnih pojavov pa nam omogočajo izračunane relativne mere variabilnosti.

Izračunamo jih s primerjavo mer variabilnosti z ustreznimi srednjimi vrednostmi. Običajno se izražajo v odstotkih, kar je prednost pri primerjavi. Največkrat se izračunavata **koeficient variabilnosti in relativni variacijski razmik.**

8.4.1 Koeficient variabilnosti

Koeficient variabilnosti **je razmerje med standardnim odklonom in aritmetično sredino.**

$$KV\% = \frac{\sigma}{M} \times 100$$

Večja variabilnost je izražena z večjim koeficientom.

Ponovite poglavje 3 Relativna števila!

8.4.2 Relativni variacijski razmik

Najbolj enostavna relativna mera variabilnosti se imenuje relativni variacijski razmik, ki je **razmerje med variacijskim razmikom in neko srednjo vrednostjo**, npr. povprečjem ali mediano.

$$\mathbf{R\% = \frac{VR}{M} \times 100}$$

Večja variabilnost je izražena z večjim relativnim variacijskim razmikom. Ker je količnik pomnožen s 100, je izražen v %.

8.5 NORMALNA PORAZDELITEV

Normalna porazdelitev je porazdelitev enot na osnovi teoretičnih predpostavk. Vnaprej določimo nekatere statistične parametre, kot so aritmetična sredina, modus, mediana in standardni odklon. Normalna porazdelitev je znana pod imenom Gaussova porazdelitev ali krivulja.

Značilnosti normalne porazdelitve so:
enake srednje vrednosti:

$$\mathbf{M = Me = Mo}$$

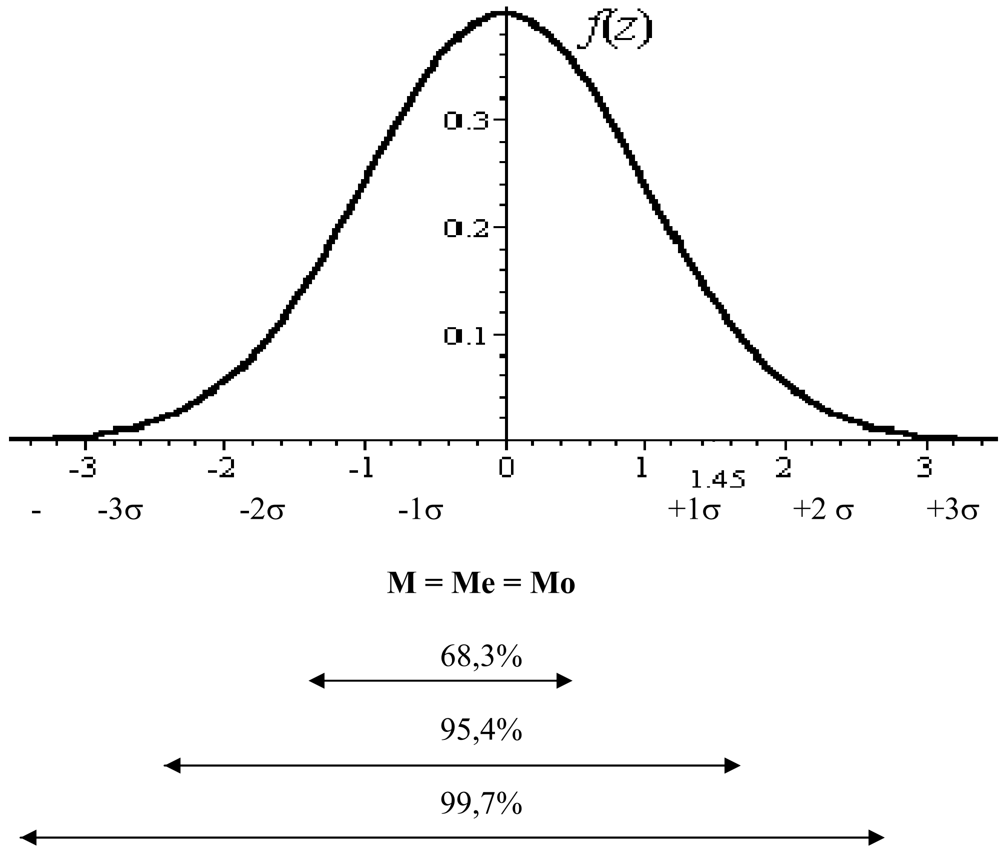
Aritmetična sredina = mediani = modusu

porazdelitev je **simetrična**:

$$\mathbf{VR = y_{max} - y_{min} = 6\ \sigma}$$

Variacijski razmik je 6 kratni standardni odklon

Razmik enot od povprečja (Šadl, 2001):
M - σ do M + σ se nahaja 68,3 % vseh vrednosti spremenljivke (**2/3 vseh enot**)
M - 2σ do M + 2σ se nahaja 95,4 % vseh vrednosti spremenljivke
M - 3σ do M + 3σ se nahaja 99,7 % vseh vrednosti spremenljivke



Slika 9: Normalna porazdelitev, Gaussova krivulja

8.6 PODOBNOSTI STVARNIH PORAZDELITEV Z NORMALNO

Porazdelitve, ki so oblikovane na osnovi empirično pridobljenih podatkov, imenujemo stvarne porazdelitve. Te so lahko zelo podobne normalni ali pa se od nje razlikujejo.

Ko zberemo podatke o številu nočitev ali gostov v hotelu, nas zanima, kako dejanski podatki odstopajo od normalne porazdelitve. Zanima nas, ali je dejanska porazdelitev simetrična, asimetrična v levo ali desno, je unimodalna, bimodalna, kako je sploščena oziroma koničasta.

Podobnost porazdelitve, ki smo jo ugotovili iz dejanskih podatkov, pridobljenih s statistično raziskavo in normalno porazdelitvijo, ugotavljamo na več načinov.

1. Izračunamo srednje vrednosti, iz katerih lahko ocenimo simetričnost.

M = Me = Mo – populacija je **simetrična**
M > Me > Mo – asimetrična v **desno**
M < Me < Mo – asimetrična v **levo**

Najenostavnejši in najhitrejši način določanja asimetričnosti porazdelitve je z grafičnim prikazom opazovanih enot. V poligonu ali histogramu prikazana simetrična porazdelitev je enakomerno padajoča levo in desno od sredine in ima zvonasto obliko. Znana je pod imenom Gaussova krivulja. Asimetričnost opazimo s sliko, kjer je črta razpotegnjena v eni smeri. Glede na smer daljše črte označimo asimetričnost. Za grafično prikazovanje in pravilno analizo je potrebno poznati poglavje 6.4.

Ponovite poglavje 7 Srednje vrednosti!

2. Izračunamo stopnjo asimetrije z izračunom koeficienta asimetrije, ki je izračunan lahko na osnovi modusa ali mediane. Izračunani koeficienti so osnova za sklepanje o smeri in velikosti asimetričnosti.

a. koeficient, izračunan na osnovi **modusa**

$$KA_{Mo} = \frac{M - Mo}{\sigma}$$

b. koeficient, izračunan na osnovi **mediane**

$$KA_{Me} = \frac{3(M - Me)}{\sigma}$$

Če je dobljeni rezultat **negativen,** je porazdelitev asimetrična v **levo,** saj sta modus in mediana večja od aritmetične sredine.

Če je dobljeni rezultat **pozitiven,** je porazdelitev asimetrična v **desno,** saj sta modus in mediana manjša od aritmetične sredine.

Čim večji je izračunani koeficient, večja je jakost asimetrije. Teoretična meja izračunanega koeficienta je +3 in -3 (Blejec, 1979).

3. Izračunamo odstotek enot v razmiku + - standardni odklon.

Za dejanske podatke opazovanih enot izračunamo aritmetično sredino in standardni odklon. Izračunamo vrednost, ki jo dosega enota, ki je od aritmetične sredine oddaljena 1-krat (+, -) standardni odklon. Za te vrednosti izračunamo kvantilni rang. Izračunamo razliko med kvantilnima rangoma in jo primerjamo s celotnim številom opazovanih enot. Pri normalni porazdelitvi je delež vseh enot, ki izpolnjuje ta pogoj, 68,3%. Čim bližje je dobljeni rezultat temu podatku, bolj je opazovana populacija podobna normalni porazdelitvi.

Opazovanje variabilnosti opazovane populacije je v statistični analizi zelo pogosto, zato je potrebno dobro poznati pojme in mere variabilnosti. Izbira mer je odvisna od cilja analize. Z uporabo računalnikov se izračun poenostavi, zato se vedno bolj uporabljajo mere, ki vključujejo podatke vseh enot. Vendar je za pravilno izbiro in analizo potrebno dobro poznavanje statistične teorije in imeti zadostno računalniško znanje. Potrebno je poznati pojme, kot so: variabilnost, variacijski razmik, varianca, standardni odklon, normalna porazdelitev, Gaussova krivulja, asimetrične porazdelitve. Pri izračunavanju mer variabilnosti je potrebno znanje nekaterih formul in funkcij v programu Excel. Največkrat zaradi nazornosti in enostavnosti v analizah variabilnost prikazujemo grafično, zato je potrebno podatke pravilno prikazati v tabelah in izbrati pravo vrsto grafa, ki ga je potrebno oblikovati tako, da je možna pravilna analiza.


Naloge:

1. Iz izbranih podatkov oblikujte graf in analizirajte variabilnost (v programu Excel).
2. Iz podatkov izračunajte tipične predstavnike in ocenite variabilnost. Uporabite formule in funkcije (M, Me, Mo).
3. Iz podatkov izračunajte variacijski razmik in analizirajte variabilnost! Ugotovite prednosti in slabosti tako izračunane mere variabilnosti.
4. Iz podatkov izračunajte varianco in standardni odklon! Pri izračunavanju uporabljajte funkcije v Excelu! Analizirajte dobljene rezultate.
5. Iz podatkov izračunajte koeficient variabilnosti in relativni variacijski razmik! Analizirate dobljene rezultate! Ocenite variabilnost opazovane populacije.
6. Primerjajte dobljene rezultate z normalno porazdelitvijo! Kako odstopajo na osnovi primerjave M, Me, Mo?
7. Za podatke ocenite simetričnost oz. asimetričnost in ugotovite, v katero smer je asimetrična porazdelitev! Kaj bi izračunali? Kakšne rezultate dobimo?

9 ANALIZA ČASOVNIH VRST

Vsak dan se v podjetju sprejemajo številne odločitve. Za večje in pomembnejše poslovne odločitve sestavljamo poslovne načrte. Za dobre poslovne odločitve je potrebno pridobiti ustrezne podatke in jih spremeniti v dobre informacije. Mnogo dobrih podatkov pridobimo z opazovanjem pojavov v preteklosti. Analiza teh nam omogoča napovedovanje razvoja v prihodnosti. Napovedovanje ni povsem zanesljivo, saj so na pretekle rezultate vplivali številni dejavniki, ki pa se v prihodnosti lahko spremenijo. Kljub temu so statistične analize podatkov, ki temeljijo na opazovanju preteklih obdobij, temelj za sprejemanja poslovnih odločitev. Zato je potrebno poznati slabosti in prednosti pri uporabi le-teh.

9.1 ČASOVNE VRSTE

Časovne vrste se oblikujejo s ponavljanjem opazovanja enote ali pojava v enakih časovnih enotah. V podjetju so te časovne enote običajno leto, mesec, teden, dan, lahko pa tudi krajše, kot so delovna izmena, ura. Redko so daljše od enega leta. Časovne vrste prikazujejo spremembe oziroma dinamiko pojava v opazovanem obdobju.

Razmislite, katere podatke bi lahko prikazali v časovni vrsti! Navedite primere iz prakse.

Namen oblikovanja je odkrivanje lastnosti spreminjanja pojava v preteklosti. Analiza nam omogoča predvidevanja v prihodnosti. Napovedi so bolj točne za krajše obdobje, saj se dejavniki, ki vplivajo na opazovani pojav, manj spreminjajo.

Za analizo dinamike opazovanega pojava se uporabljajo številne statistične metode. Enostavno analizo omogočajo relativna števila (poglavje 4) in povprečja (poglavje 6), ki omogočajo ugotavljanje in primerjavo razlik. Hitro, a nekoliko bolj grobo, predstavo o časovni vrsti pridobimo z grafičnim prikazom podatkov. Običajno nam linijski graf nazorno predstavi nihanje, gibanje in smer razvoja. Za enostavne in hitre analize so prav grafični prikazi najprimernejši. Z uporabo računalniških programov je ta metoda zelo enostavna in nazorna.

Ponovite poglavje 4 in 6!

Za natančnejšo analizo časovne vrste je potrebno spoznati **dejavnike**, ki povzročajo kratkoročna in dolgoročna **nihanja**. Ta so zelo velika v gostinstvu in turizmu, zato je znanje ekonomike turizma nujno za dobro analizo podatkov, prikazanih v časovnih vrstah.

Ti dejavniki so lahko **ekonomsk**i (nove investicije v podjetju), **naravni** (vpliv sezone, vremena), **institucionalni** (davčna zakonodaja) in **drug**i (osamosvojitvena vojna).

Nekateri povzročajo **kratkoročna** nihanja (prosti čas konec tedna) drugi **dolgoročna** (letni čas) zato oblikujemo različne časovne vrste za različna časovna obdobja.

Dopolnite znanje o vzrokih dinamike pri ekonomskih predmetih!

Za daljše opazovano časovno obdobje prikazani podatki (vsaj 10 let) omogočajo ocenjevanje **trenda, to je smeri razvoja.** Večini opazovanih pojavov v preteklosti lahko določimo smer razvoja, ki je lahko naraščanje ali upadanje, le redki so pojavi, ki se ne spreminjajo. Npr.

izboljševanje rentabilnosti, zasedenosti ali upadanje obiska nemških gostov, krajšanje povprečne dobe bivanja itd.

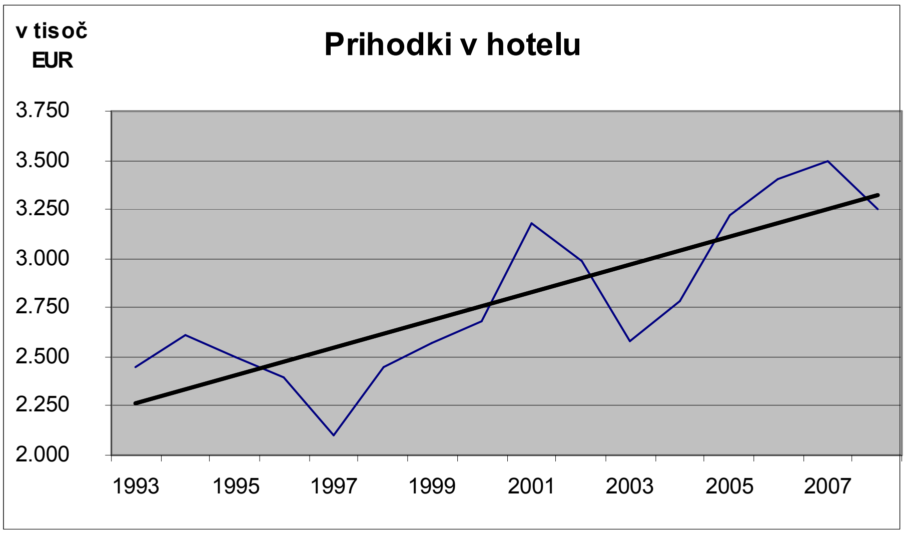
Za krajša enako dolga opazovana obdobja (mesec, dan ...) prikazani podatki omogočajo ocenjevanje **periodičnih nihanj.** V gostinstvu in turizmu so ta nihanja, ki jih imenujemo tudi **sezonska nihanja,** zelo pogosta. Npr. število gostov v poletnih mesecih, prodaja v soboto, obisk gostov v določenih urah ...

V podjetjih, kjer prikazujejo časovne vrste za krajše obdobje, običajno ne analizirajo **cikličnih nihanj**, ki so značilna za ekonomske pojave daljšega obdobja in so razložena pri predmetu Ekonomika, ter **iregularnih** nihanj, ki so slučajna (npr. naravne katastrofe, vojne).

9.2 DOLOČANJE TRENDA

Iz opazovanih in prikazanih podatkov časovne vrste za zelo dolgo obdobje najmanj 10 let je opazen trend ali smer razvoja. Ugotavljamo ga iz podatkov časovne vrste, kjer opazujemo gibanje vrednosti številskih enot v tabeli ali iz grafičnega prikaza.

Najenostavnejše je določanje trenda **z grafičnim** prikazom podatkov časovne vrste v linijskem grafikonu. V programu Excel v že narisanem grafu izberemo možnost dodajanja **trendne črte**. Ta se prikazani časovni vrsti prilega, tako da se podatki osnovne časovne vrste od trenda odklanjajo navzgor in navzdol. Trend je lahko **premica** ali krivulja. Za enostavne analize običajno uporabljamo premico, ki prikazuje **linearni** trend. Gibanje prikazuje enakomerno, čeprav je dejansko spreminjanje običajno neenakomerno, v nekaterih obdobjih celo obratno, kot prikazuje trend. Npr. če spremljamo gibanje prihodkov, je lahko trend naraščanje, kljub temu da je v posameznih letih prišlo do upada.



Slika 10: Vrisana trendna črta
Vir: Lastni

Natančnejše je **izračunavanje** trenda iz statistične vrste. V programu Excel nam to omogoča uporaba **funkcije TREND**. Pridobljeni rezultat s to funkcijo nam da podatek, ki omogoča napovedovanje pričakovanih rezultatov v prihodnosti. Če je rezultat pozitiven, ga prištevamo zadnjemu podatku, če je negativen, ga odštevamo.

9.3 ANALIZA PERIODIČNIH NIHANJ

Za gostinstvo in turizem so periodična nihanja zelo značilna, zato jih pogosto ugotavljamo in analiziramo. Pojavi se lahko ponavljajo po urah, izmenah, dnevih, tednih, mesecih (npr. število gostov, nočitev, prodaja pijače, promet v restavraciji). Najdaljše obdobje za ponavljajoče se periodično nihanje je **eno leto**. Če se nihanja ponavljajo na eno leto, jih imenujemo sezonska nihanja. Časovno obdobje, v katerem se nihanje ponovi, se imenuje **perioda**. Ta je razdeljena na več krajših obdobij, npr. opazovana perioda (leto) je razdeljena na več krajših obdobij (mesecev, trimesečij). Glede na dolžino opazovanega obdobja lahko v periodi določamo različno število obdobij.

Razmislite, kateri so vzroki periodičnih nihanj! Dopolnite znanje pri ekonomskih predmetih.

Najenostavnejši način prikazovanja periodičnega nihanja je z **grafičnim prikazom**. V njem za periodo, ki je razdeljena na ustrezno število krajših enot, prikažemo podatke v linijskem ali stolpčnem grafikonu. Velikost stolpcev prikazuje nihanje pojava.

Še boljše je izračunati indekse z osnovo povprečje in jih prikazati v grafikonu. Program Excel omogoča indekse nad 100, ki prikazujejo nadpovprečne vrednosti, prikazati naraščajoče in stolpce, ki prikazujejo podpovprečne vrednosti, kjer so indeksi manjši od 100, padajoče. V že oblikovanem grafikonu premaknemo x os tako, da so stolpci obrnjeni navzgor in navzdol, kar poveča nazornost prikaza.

Zaradi značilnosti dejavnosti gostinstva in turizma, kjer so periodična nihanja zelo pogosta, je statistična analiza nihanja nujna. Zato je potrebno dobro poznavanje ekonomske teorije za prepoznavanje dejavnikov, ki vplivajo na časovno vrsto, in izbrati pravo časovno obdobje opazovanja in prikazovanja podatkov. Glede na cilj analize izberemo obdobje in statistične metode. Za enostavne analize zadošča izračunavanje s formulami in funkcijami v programu Excel, s katerimi lahko izračunamo periodično nihanje in trend. Pri analizi se najpogosteje poslužujemo grafičnega prikazovanja. Zato je potrebno poznavanje Excela, pravilne izbire vrste grafikona in ustrezno oblikovanje, kot je risanje trendnih črt, obračanje stolpcev.

Naloge:

1. *Izberite podatke časovne vrste za različna obdobja, oblikujte tabele in analizirajte periode.*
2. *Podatke časovne vrste prikažite v linijskem grafikonu in vrišite trendno črto! Analizirajte grafični prikaz.*
3. *Za podatke časovne vrste izračunajte trend s funkcijo TREND! Analizirate dobljeni rezultat.*
4. *Podatke časovne vrste za obdobje enega leta prikažite v stolpčnem grafikonu – primerjalni stolpci! Analizirajte dobljene rezultate.*
5. *Podatke časovne vrste preračunajte v indekse z osnovo povprečje. Prikažite indekse v stolpčnem grafikonu. Stolpce obrnete pri vrednosti 100. Analizirate dobljene rezultate.*

VIRI IN LITERATURA:

Košmelj, B. *Statistika*. Ljubljana: DZS,d.d., l999.
Košmelj, B., in drugi. *Statistični terminološki slovar*. Ljubljana: Statistično društvo Slovenije, 1993.
Košmelj, B. *Vaje iz statistike*. Ljubljana: Ekonomska fakulteta, 1979.
Blejec, M. *Statistične metode za ekonomiste*. Ljubljana: Univerza v Ljubljani, 1973.
Blejec, M. *Uvod v statistiko*. Ljubljana: Ekonomska fakulteta, 1979.
Šadl, M., Terseglav Klemenc, M. *Zbirka vaj iz statistike*. Ljubljana: Statistično društvo Slovenije, 1996.
Šadl, M. *Statistika za komercialiste*. Murska Sobota: Ekonomska šola, Višja strokovna šola, 2001.
Šadl, M. *Statistika-učbenik za 4. letnik srednješolskega programa ekonomski tehnik*. Celovec: Mohorjeva/Hermagoras, 2004.
Šadl, M. *Zbirka vaj iz statistike*. Murska Sobota: Ekonomska šola, Višja strokovna šola, 2002.
Trstenjak, I. *Statistika za srednje ekonomske šole - delovni zvezek z osnovami teorije*. Ljubljana: Modrijan, 2001.
Sagadin, J. *Statistične metode za pedagoge*. Maribor: Obzorja, 2003.
Knežević, M. *Statistika skoraj brez matematike*. Portorož: Univerza na Primorskem, Turistica, Visoka šola za turizem, 2004.
Knežević, M. *Statistika z uporabo računalnika*. Portorož: Univerza na Primorskem, Turistica, Visoka šola za turizem, 2006.
Kragelj, I., Ušaj, T. *Popotovanje z Excelom v poslovni svet-Vaje*. Ljubljana: DZS, 1999.
Spletna stran Statističnega urada RS (online) (povzeto 15. 9. 2008). Dostopno na naslovu: http://www.stat.si
Microsoft Office Excel 2003

Projekt **Impletum**
Uvajanje novih izobraževalnih programov na področju višjega strokovnega izobraževanja v obdobju 2008–11

Konzorcijski partnerji:

Operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo RS za šolstvo in šport. Operacija se izvaja v okviru Operativnega programa razvoja človeških virov za obdobje 2007–2013, razvojne prioritete Razvoj človeških virov in vseživljenjskega učenja in prednostne usmeritve Izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja.