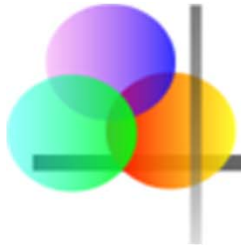




Бизнис статистика

Вовед



Статистика: порано и сега

- **Порано.** Иако статистички податоци биле собирани и пред илјадници години тие обично вклучувале само броење на луѓе или имот за олеснување на оданочувањето. Потоа, статистиката се трансформира од релативно пасивно собирање и опишување на податоци во сè поактивна и корисна деловна алатка, која овозможува да се извлечат заклучоци од собраните информации и да се донесат соодветни одлуки.
- **Денес.** Статистиката и нејзините примени се составен дел од нашиот живот. Во многу различни области, како политика, медицина, образование, бизнис ..., човечките активности се мерени и водени од статистиката.



Примери

- Следењето на пазарот доведува до заклучоци за продажба кои им помагаат на компаниите да одлучат кои производите да бидат задржани, отфрлени или променети.
- Авто осигурителните компании собираат податоци за возраста, типот на возилото и несреќите, и овие статистики ги водат компаниите кон екстремно високи премии за тинејџери кои поседуваат или возат автомобили со висока моќ.
- Според Меѓународната асоцијација за млечни производи, сладоледот и сродните замрзнати десерти се консумираат во повеќе од 90% од домаќинствата во САД. Најпопуларен вкус е ванилата со 30% од продажбата, чоколадниот е со 10% од продажбата.



Основни концепти

Статистиката е наука која се занимава со:

1. прибирање, организација, групирање и анализа на податоци,
2. изведување на заклучоци кои важат за целокупното множество податоци, иако набљудувањата се направени на само дел од податоците.



Што е статистика?

- Математичка гранка за собирање на податоци и трансформирање на броеви во корисни информации за донесување на одлуки.
- Методи за обработка и анализа на броеви.
- Методи за намалување на несигурноста својствена за донесување одлуки.
- Ние ќе ја разгледуваме повеќеслојната улога на статистиката: како дескриптор на информации, алатка за анализа, средство за доаѓање до заклучоци и помош за донесување одлуки.



Зошто да се учи статистика?

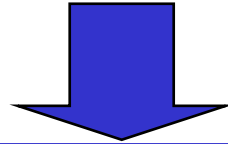
Статистика се користи за:

- Правилно да се презентираат и опишат податоците и информациите,
- Да се подготват заклучоци за големи групи на поединци или предмети, користејќи информации собрани од подгрупи на поединци или предмети,
- Да се направат сигурни предвидувања за понатамошните активности,
- Подобрување на бизнис процесите ...

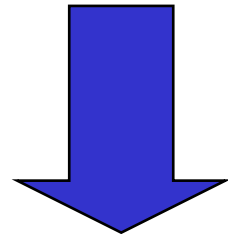


Растот и развојот на статистиката се должат на

**Потребата на владите да собираат
податоци за нивните жители**



Развојот на теоријата на веројатност



Предноста на компјутерите



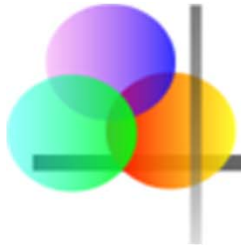
Полиња на примена

<ul style="list-style-type: none">■ Економија<ul style="list-style-type: none">■ Предвидување■ Демографија	<ul style="list-style-type: none">■ Инженерство<ul style="list-style-type: none">■ Производство■ Материјали
<ul style="list-style-type: none">■ Спорт<ul style="list-style-type: none">■ Индивидуални и тимски перформанси	<ul style="list-style-type: none">■ Бизнис<ul style="list-style-type: none">■ Потреби на потрошувачите■ Финансиски трендови



Бизнис статистика

- Статистичките алатки се применуваат во многу области:
бизнис, образование, психологија, земјоделство, економија, ...
- Кога податоците кои се анализираат доаѓаат од **економски науки** и **менаџмент**, се користи терминот **бизнис статистика** за да се разликува оваа примена на статистичките алатки и концепти од примените во другите области.



Податоци

- Основниот материјал во статистиката се податоците.
- Податоците може, но и не мора да се броеви. Броевите се резултат на броење или мерење.

Примери:

- Колку студенти дипломирале минатата година на ФИНКИ (броење).
- Колку е висината на секој од дипломираните студенти (мерење).



Податоци

- Статистички податоци се однородни објекти кои имаат една или повеќе заеднички карактеристики.
- На пример, на множеството од сите деца во Скопје на десет годишна возраст, може да се испитаат следните карактеристики: висина, тежина, успех на училиште, пол, националност, боја на очи и сл.
- На множеството компјутери на ФИНКИ, може да се испитуваат карактеристиките: брзина на процесорот, големина на RAM меморијата, капацитет на дискот и сл.



Популација и обележје

- Во статистиката, множеството еднородни објекти или резултати од некоја операција, кои имаат една или повеќе заеднички карактеристики се нарекува *популација*. Заедничката карактеристика се нарекува *обележје*.
- Всушност, популацијата се состои од сите објекти или поединци кои се од интерес на истражувањето и за кои треба да се донесат одредени заклучоци.
- Секој елемент на популацијата се нарекува *единка* или *индивидуа*. Може да биде објект или резултат на некоја операција која е предмет на истражувањето.



Примери

- Во наведените примери, за популацијата “деца во Скопје”, обележја се: висината, тежината, успехот, полот, националноста, бојата на очите и сл., а за популацијата “компјутери на ФИНКИ”, обележја се: брзина на процесорот, големина на RAM меморијата, капацитет на дискот, итн.
- Може да се воочи дека вредноста што ја добива обележјето е променлива величина и таа не може со сигурност да се предвиди за конкретна единица од популацијата. Затоа, ќе сметаме дека секое обележје е променлива.



Обележје

Примери:

- Плата на вработени во ИТ секторот,
- Број на положени испити по првата година на студии,
- Број на освоени поени на испит,
- Висина на студенти,
- Пол на студенти,
- Месечна каматна стапка,
- Просечна дневна температура,
- Количество на PM10 честички во воздухот,
- Време поминато на интернет во текот на денот,



Зошто се собираат податоци?

- Маркетинг аналитичарот треба да ја процени ефективноста на новата телевизиска реклама.
- Фармацевтскиот производител треба да утврди дали новиот лек е поефективен од оние што се користат во моментот.
- Директорот на производство сака да го следи процесот на производство за да дознае дали квалитетот на производот што се произведува е во согласност со стандардите на компанијата.
- Ревизорот сака да ги ревидира финансиските трансакции на компанијата со цел да утврди дали компанијата е во согласност со општо прифатените сметководствени принципи.



Извори на податоци

Примарни
Собирање на
податоци

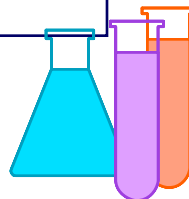
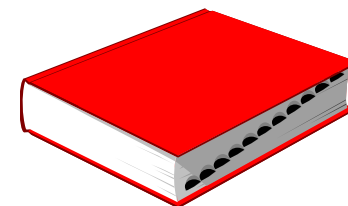
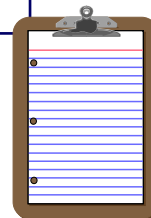
Секундарни
Компилација на
податоци

Набљудување

Анкета

Експерименти

Печатени или
електронски





Извори на податоци

- **Примарни извори:** Податоците ги собира оној што ги користи истите за анализа
 - Податоци од политичко истражување
 - Податоци добиени од експеримент
 - Податоци добиени со набљудување

- **Секундарни извори:** Лицето што врши анализа на податоци не е тој што ги собира податоците
 - Анализа на податоците од пописот
 - Испитување на податоци од печатените списанија или податоци објавени на интернет



Извори на податоци

Бараме соодветни податоци кои ќе послужат како основен материјал за нашето истражување. Податоците може да ги најдеме од повеќе извори:

1. Рутински записи.

Во студентската служба се води евиденција за студентите во поглед на запишани предмети, положени испити, извршени уплати

2. Анкети.

Податоци можат да се добијат и од анкети ако се однесуваат на одговор на одредено прашање.



Извори на податоци

3. Експерименти.

Честопати податоците потребни да се одговори на некое прашање можат да се добијат само со реализација на некој експеримент.

4. Надворешни извори.

Податоците потребни за да се одговори на одредено прашање може да постојат во облик на:

- публикувани податоци,
- комерцијални банки на податоци, или
- во истажувачка литература,

односно некој веќе некаде го поставил истото прашање.



Извори на податоци

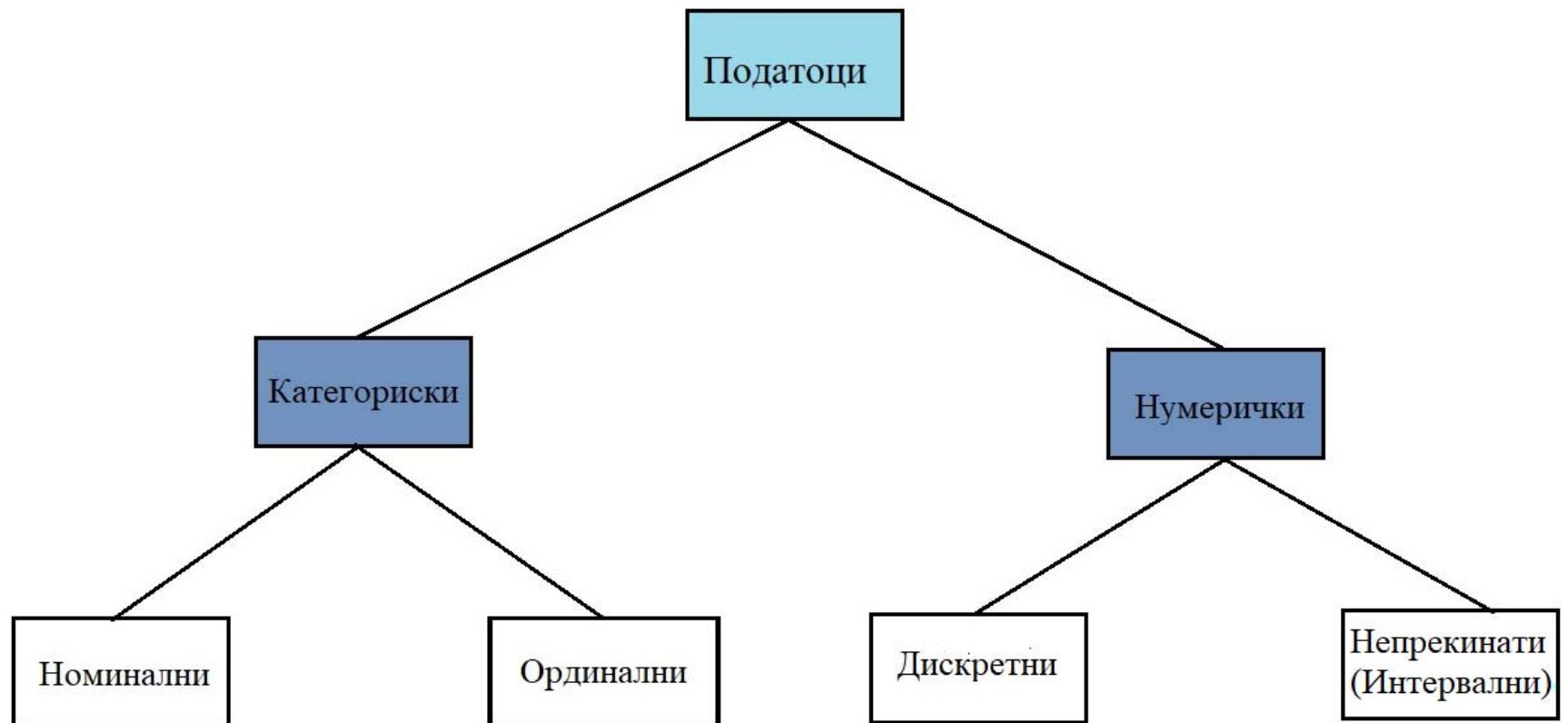
Пример (извор на податоци – експеримент):

За да се процени дали воведувањето на нов метод на учење ги подобрува конгнитивните способности на учениците се прави експеримент на следниот начин:

- Една група ученици учат со стариот метод.
- Друга група ученици учат со новиот метод.
- На крајот се тестираат на исти тестови и се споредуваат резултатите.



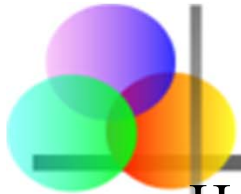
Видови податоци





Видови податоци

- Податоците може да бидат *нумерички* или *категориски*.
 - Нумеричките податоци се податоци кои се изразуваат во броеви. Овие податоци се поврзани со одредени мерења (како на пример, мерење на висината, тежината, коефициентот на интелигенција или крвниот притисок на една личност) или се поврзани со броење (како на пример, бројот на акции што ги поседува едно лице, број на студенти во една група, и сл.) Нумеричките податоци уште се нарекуваат и *квантитативни* податоци.



Нумерички податоци

- Нумеричките податоци можат понатаму да се поделат на два вида: дискретни и непрекинати.
 - Дискретни податоци се добиваат како резултат на броење. Множеството на можни вредности за овие податоци може да биде конечно или преброиво. На пример, бројот на глави во 100 фрлања на монета може да прими вредност од 0 до 100 (конечен случај), но бројот на фрлања додека не се паднат 100 глави може да прими вредност од 100 до бесконечност (преброив случај).
 - Непрекинатите податоци се добиваат со мерења. Нивните можни вредности можат да се опишат само со употреба на интервали на реалната права. На пример, времето на работа на една машина е во интервал $[0, T]$, каде T е максималниот век на работа на машината, а може да биде и бесконечност. Значи, овие податоци се од бесконечно непреброиво множество.



Категориски податоци

- Со категориски податоци се претставуваат карактеристики како што се пол, националност, боја на очи, брачен статус, роден град и сл.
- На категориските податоци може да им се доделат нумерички вредности (на пример, „1“ да означува маж и „2“ да означува жена), но тие броеви немаат математичко значење. Категориските податоци уште се нарекуваат и **квалитативни** податоци.



Категориски податоци

- Категориските податоци може да бидат *номинални* или *ординални*.
 - Номиналните вредности претставуваат дискретни единици и се користат за обележување (лабелирање) на променливите и тие немаат квантитативна вредност.
 - Броевите кои ги претставуваат податоци може да се користат само за класифицирање или категоризирање.
 - На пример, броевите за идентификација на вработените (или број на индекс) се номинални податоци. Некои други видови на обележја со кои често се добиваат податоци на номинално ниво се пол, религија, етничка припадност, географска локација и место на раѓање.
 - Да воочиме дека номиналните податоци немаат редослед. Затоа, ако би го смениле редоследот на неговите вредности, значењето не би се променило.



Категориски податоци

- Ординалните податоци може да се користат за рангирање или подредување.
 - На пример, супервизорот може да оцени тројца вработени со рангирање на нивната продуктивност со броевите од 1 до 3 (најпродуктивен, најмалку продуктивен, и некаде помеѓу). Но, не може да каже дека разликите во продуктивноста меѓу работниците рангирани со 1 и 2; и со 2 и 3 се нужно исти, т.е., растојанијата претставени со последователни броеви не се секогаш еднакви.
 - Ординална скала обично се користи за мерење на ненумерички карактеристики како среќа, задоволство на клиентите и сл.



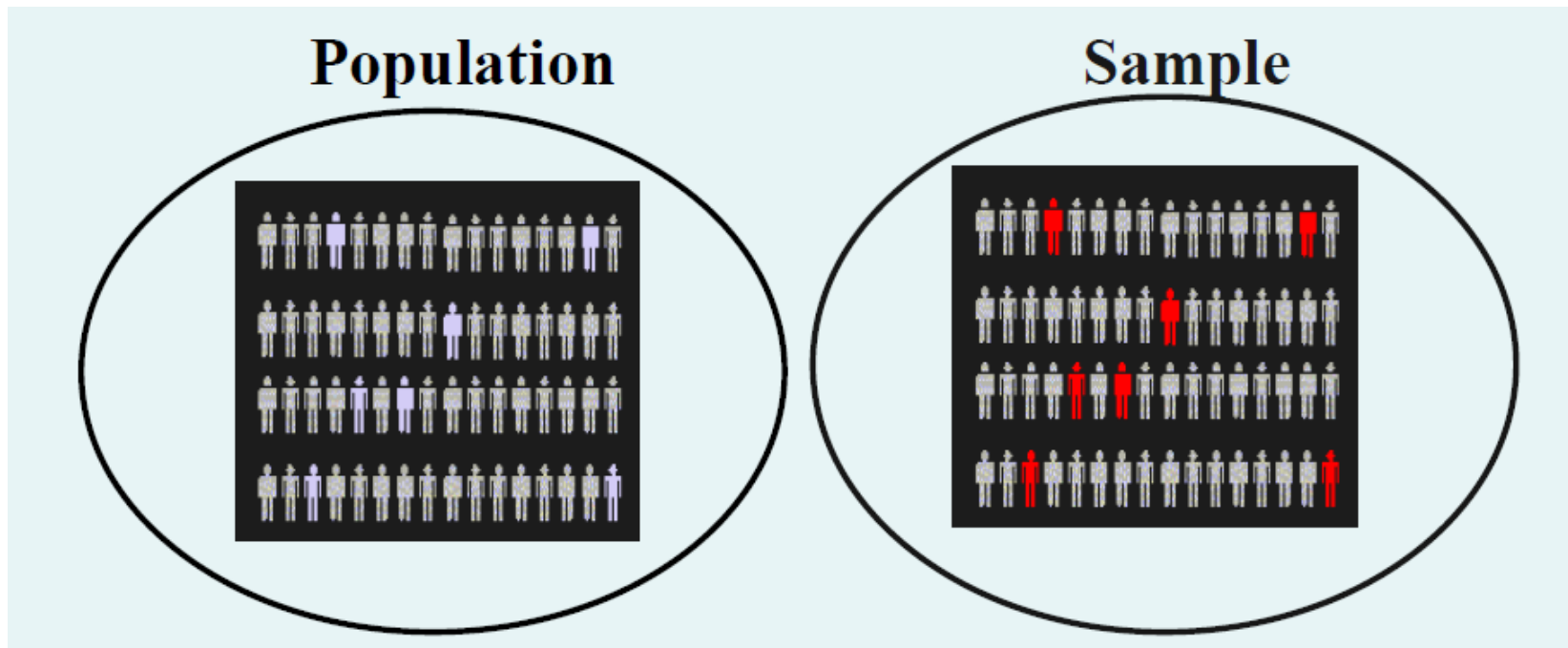
Примерок

- Најчесто, при вршење на статистички испитувања, невозможно е да се испитаат сите елементи од популацијата.
- Се случува испитувањата да траат долго, да бидат поврзани со трошење на големи финансиски средства или уште повеќе, при испитување на одредени објекти да дојде до нивно уништување.
- Затоа, се избира дел од популацијата на кој се вршат сите испитувања и притоа треба да се донесат заклучоци кои ќе важат за целата популација.



Примерок

- Делот од популацијата на кој се вршат потребните испитувања се вика *примерок*.
- Бројот на елементи во примерокот се вика *обем на примерокот* и се означува со n .





Избор на примерокот

- Заклучоците кои ќе се донесат врз основа на примерокот треба да важат за целата популација.
- Затоа, потребно е примерокот да биде репрезентативен, т.е. да претставува мини-модел на популацијата.
- За да се постигне тоа, потребно е секој елемент од популацијата да има еднакви шанси да влезе во примерокот.



Избор на примерок: пример

- Отстапувањето од принципот на случаен избор може да доведе до сериозни грешки. Да го разгледаме следниот реален пример.
- Американското списание “Литературен зборник” спровело анкета за испитување на јавното мислење околу претстојните претседателски избори во САД во 1936. Кандидати биле Ф.Рузвелт и А.Лендон.
- За избор на примерок, списанието употребило телефонски именик, претплатници на часописи, списоци на членови на некои клубови. Биле избрани 10 милиони адреси и на сите им била испратена анкета со прашања околу кандидатите. Списанието потрошило големи финансиски средства за испраќањето на картичките и обработката на податоците. Биле добиени одговори од 2,4 милиони. Било објавено дека за претседател ќе биде избран А.Лендон.
- Резултатот од изборите потврдил дека прогнозата била погрешна.

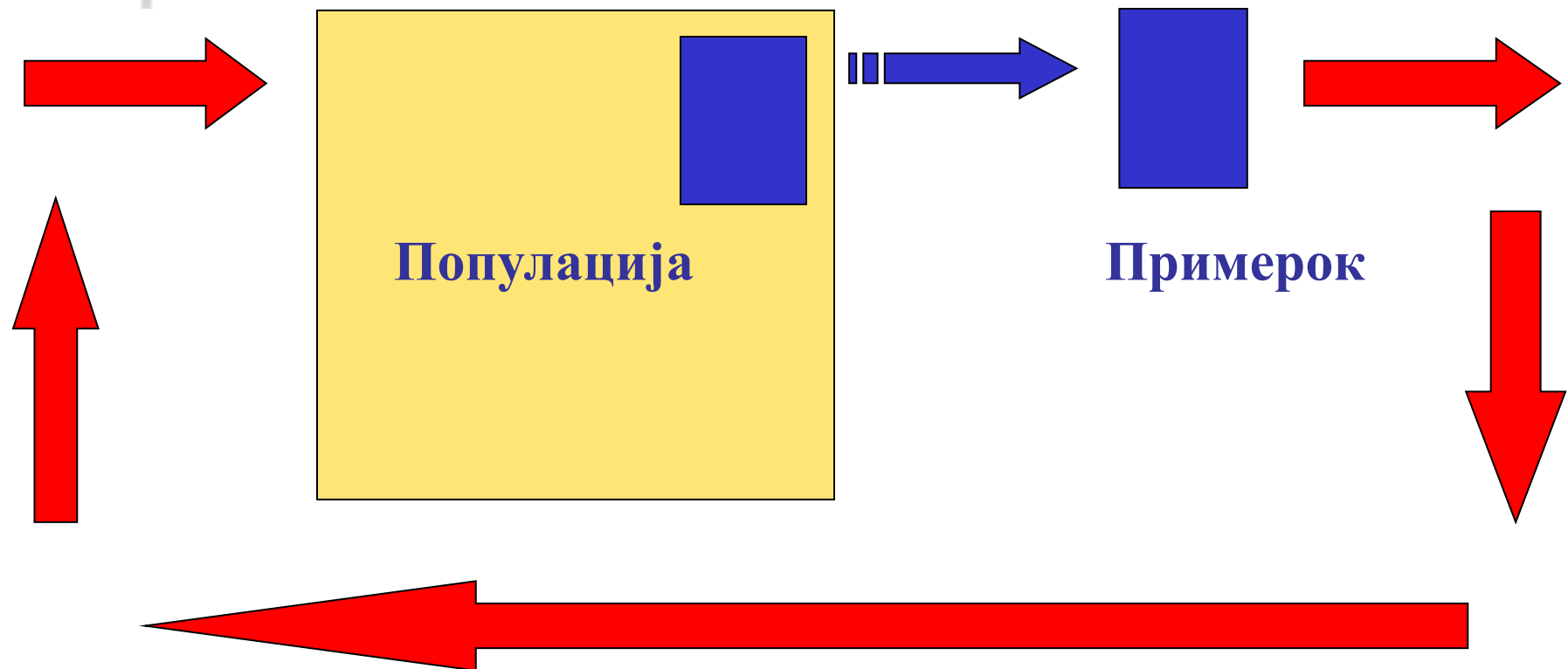


Избор на примерок: пример

- Каде било погрешено?
- Во овој случај, направени се две грешки.
 - Прво, во тој период на владеење на економска криза, телефони имале само најбогатите.
 - Второ, на анкетата одговориле најмногу деловните луѓе, кои, всушност, имале навика да одговараат на писма, а сите тие го подржувале Лендон.
- Од друга страна, двајца социолози, околу истото прашање, направиле анкета со 50 илјади испитаници и дошле до вистинскиот одговор.
 - Причина за тоа е што тие не само што правилно го составиле примерокот, туку тргнале од тоа што општеството е составено од различни социјални слоеви и најголем дел од припадниците на еден слој подржуваат ист кандидат.
 - Така, со помош на проучувањето по социјални слоеви биле донесени заклучоци кои важат за целата држава.



Популација и примерок

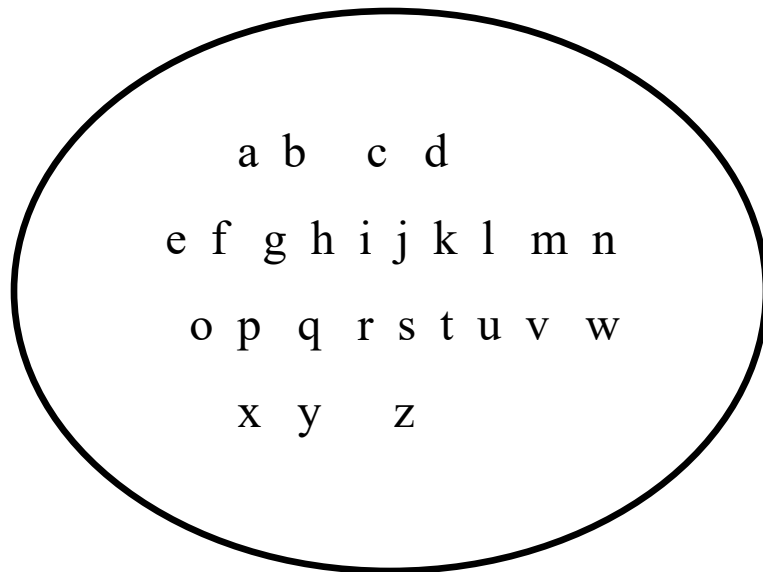


Заклучоците изведени од примерокот треба да важат за целата популација.

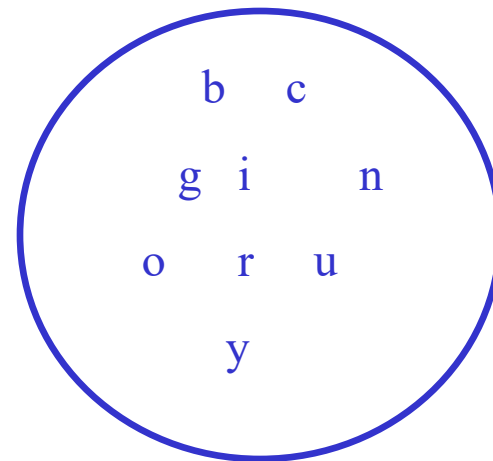


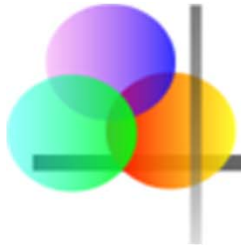
Популација и примерок

Популација



Примерок





Примери на популации

- Сите жители на Македонија (обележје – име)
- Сите жители на Скопје (обележје – висина на приходи)
- Сите домаќинства во државата (обележје - годишна потрошувачка на струја)
- Сите дипломирани инженери во Македонија (обележје – пол)
- Сите матуранти во Македонија (обележје – поени освоени на државната матура)



Пример

- Популација – Студентите на УКИМ.
- Примерок – 100 студенти од прва година.



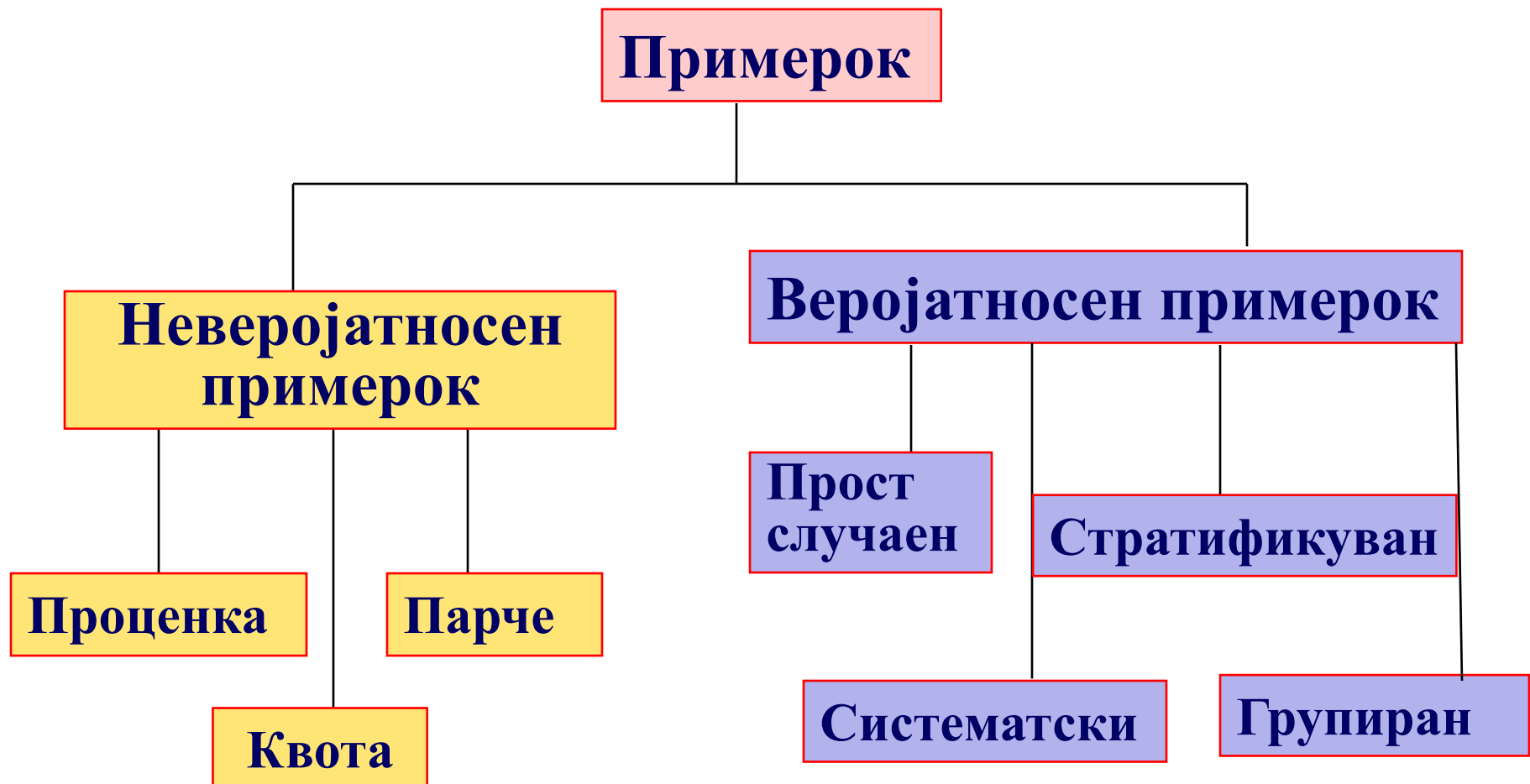
Примерок

Земање на примерок се состои од:

- Собирање на случајни податоци од голема популација, на пример, мерење на висината на случајно избрани возрасни мажи.
- Запишување на резултатите од експерименти, на пример, мерење на брзината на спуштање фајлови од интернет.
- Во општ случај ќе претпоставуваме дека популација е бесконечна (или голема). Исто така, ќе претпоставиме дека набљудувањата се независни. Резултатите од било кој експеримент не влијаат на другите експерименти.



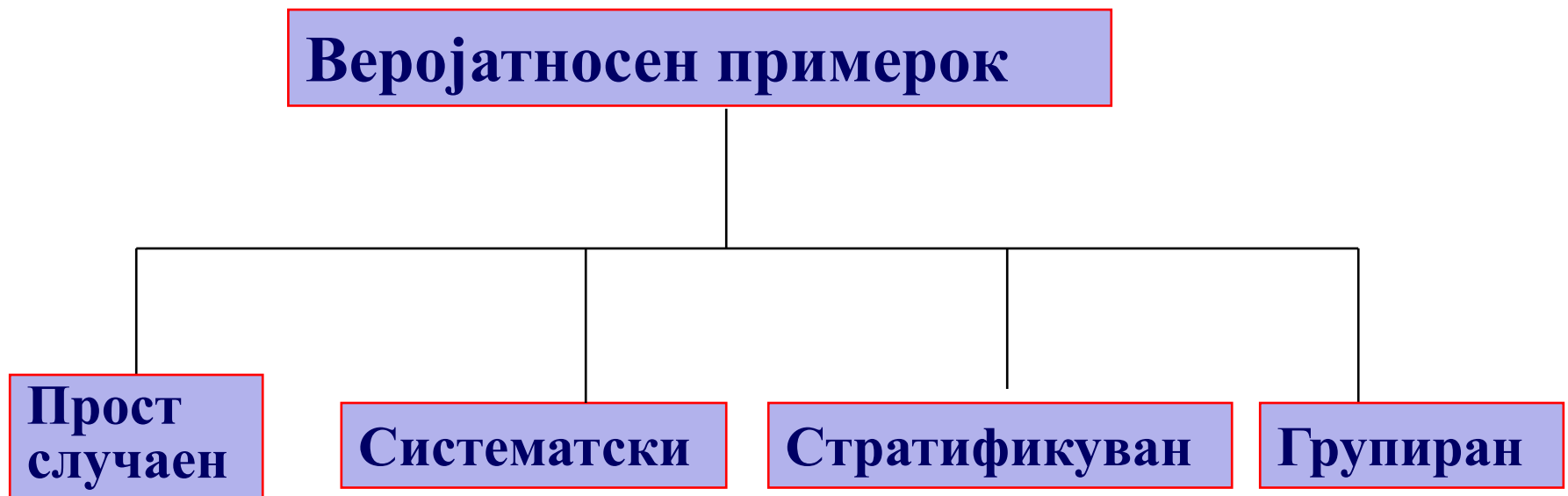
Методи на избор на примерок





Веројатносен примерок

- Единките во примерокот се избрани врз основа на познати веројатности





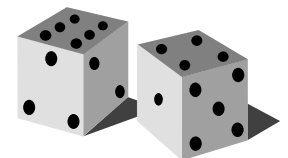
Прост случаен примерок

Прост случаен примерок е примерок во кој:

- Секоја единица од популацијата е случајно избрана,
- Секоја единица од популацијата има еднаква шанса да биде избрана,
- Секое можно множество од n единици има исти шанси да биде избрано.

Изборот може да биде со враќање или без враќање.

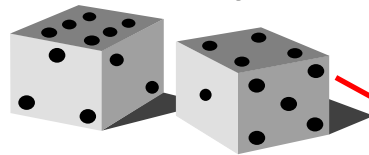
Примероци добиени со користење на табела на случајни броеви или компјутерски програм за генерирање на броеви.





Систематски примерок

- Се определува големината на примерокот – n .
- Се разделува популацијата од N единици во групи од k единици: $k = N/n$.
- На случаен начин се избира една единица од 1^{та} група.
- Понатаму се зема секоја k -та единица.

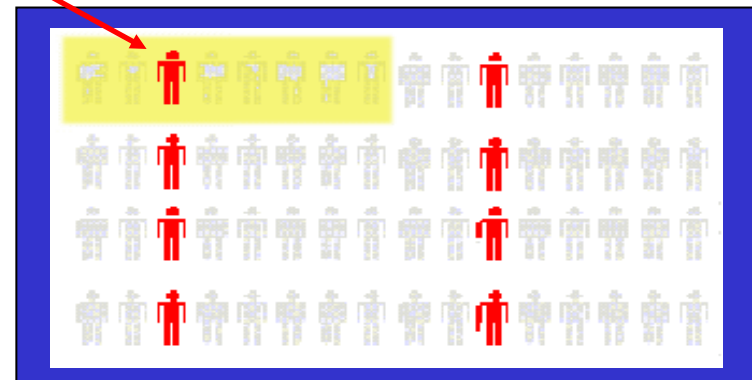


$$N = 64$$

$$n = 8$$

$$k = 8$$

Прва група





Систематски примерок - пример

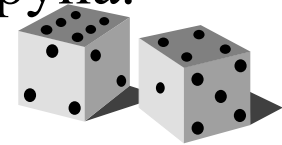
- Истражувач сака да ги тестира производителите во Тексас. Тој има финансиска поддршка за испитувања на примерок од 1.000 компании (n). Во базата на производители во Тексас дадени се вкупно околу 17.000 производители (N) по азбучен ред. Тогаш вредноста на k е 17 ($17.000 / 1.000$) и истражувачот во својот примерок ја зема секоја 17-та компанија од базата.
- Дали истражувачот започнува со првата компанија во базата или со 17-тата или со некоја помеѓу? За избор на појдовната точка може да се користи табела или генератор на случајни броеви за да се избере вредност помеѓу 1 и k како појдовна точка. Потоа се избира секој k -ти.



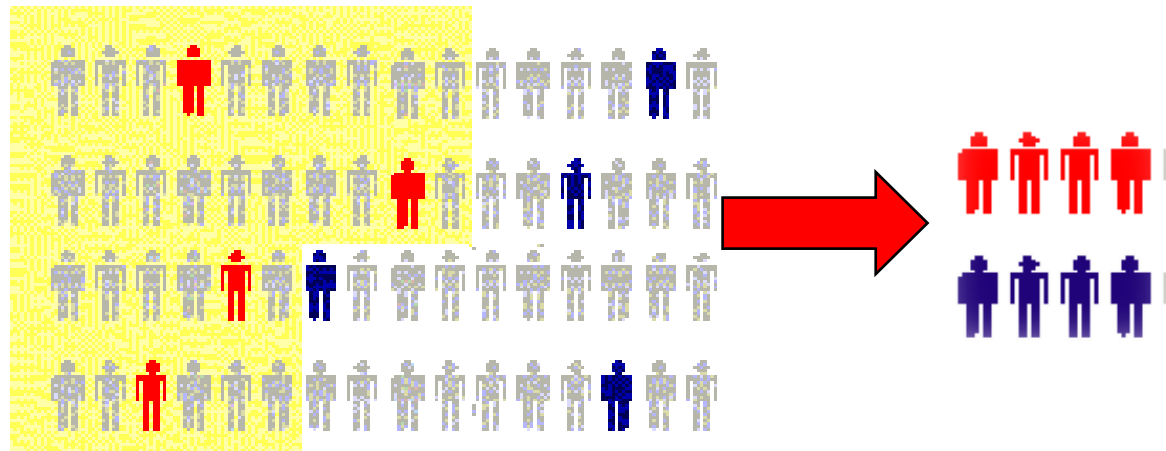
Стратификуван примерок

- Популацијата се разделува во две или повеќе групи според некоја заедничка карактеристика.

- Се бира прост случаен примерок од секоја група.



- Добиените примероците се комбинираат во еден примерок.





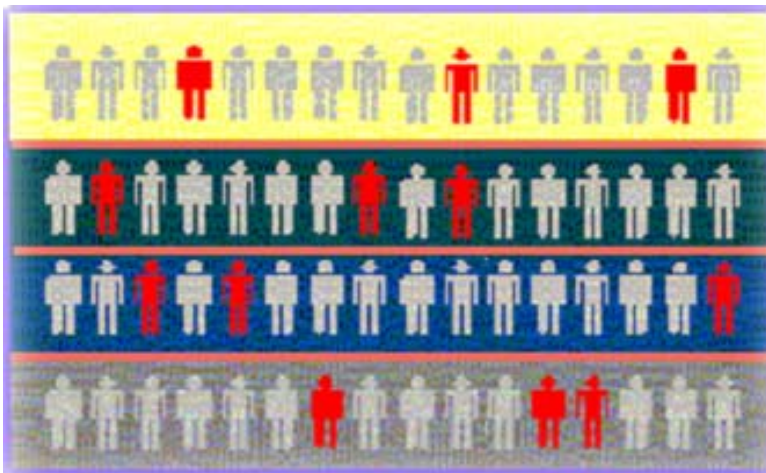
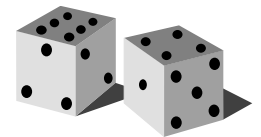
Стратификуван примерок

- Стратификуваниот примерок може да биде пропорционален или непропорционален. Пропорционален стратификуваниот примерок се добива кога процентот на примерокот земен од секоја група е пропорционален на застапеноста на групата во целата популација.
- На пример, да претпоставиме дека се испитуваат гласачите во Бостон и примерокот е стратификуван според религијата: католици, протестанти, евреи и други. Ако популацијата на Бостон е 90% католици и ако се земе примерок од 1.000 гласачи, потребно е во примерокот да бидат земени околу 900 католици за да се постигне пропорционална стратификација. Ако бројот на католици не е приближно 900, тогаш стратификацијата би била непропорционална. Учеството на другите религии во примерокот треба, исто така, да ги следи процентите на населението.



Бирање на примерок со кластерирање

- Популацијата се разделува во неколку “кластери,” секој репрезентативен за популацијата. За разлика од стратификуваниот примерок, каде што групите внатре се хомогени, тука кластерите имаат тенденција да бидат внатрешно хетерогени.
- Во теорија, секој кластер содржи широк спектар на елементи и е мини модел, или микрокосмос, на популацијата.
- Од секој кластер се бира прост случаен примерок.
- Примероците се комбинираат заедно.



**Популација
разделена
во 4
кластери.**



Бирање на примерок со кластерирање - Пример

- Еден пример за деловни истражувања што користат кластерирање е маркетинг тестирање на нови производи. Често при ваквите тестирања, САД се поделени во кластери за тестирање на пазарот.





Предности и недостатоци

- Прост случаен примерок и систематски примерок
 - Едноставни за користење.
 - Може да не е репрезентативен за основните карактеристики на популацијата.
- Стратификуван примерок
 - Порепрезентативен примерокот од простиот случаен примерок, но поскап.
 - Бидејќи делови од целиот примерок се земени од различни подгрупи на популацијата обезбедува претставници на единици од целата популација.
- Примерок со кластерирање
 - Поекономично
 - Помалку ефикасен (потребен е поголем примерок за да се добие истото ниво на прецизност)



Дескриптивна и аналитичка статистика

- Статистика се дели на две главни гранки: дескриптивна статистика и аналитичка статистика.
- **Дескриптивна статистика:**
Множество од методи кои се користат за собирање, систематизирање и организирање на податоците од примерокот.
- **Аналитичка (индуктивна) статистика:**
Множество на методи кои се користат за изведување на заклучоци за популацијата врз база на податоците од примерокот.



Дескриптивна и аналитичка статистика

- Дескриптивна статистика
 - Графички, нумерички и други процедури за собирање, организирање и процесирање на податоците за да се трансформираат во информација.
- Аналитичка (индуктивна) статистика
 - Методи кои даваат основа за предвидувања, прогнозирања и проценувања кои се користат за трансформирање на информацијата во знаење.

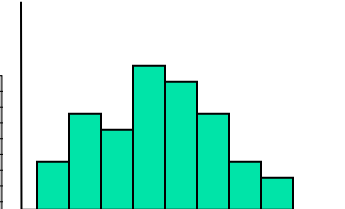
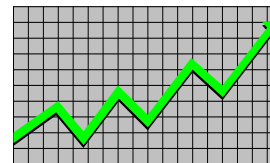


Дескриптивна статистика

- Собирање на податоци
 - На пример: анкета

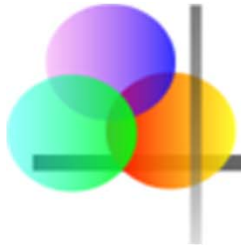


- Прикажување на податоци
 - Табели и графици



- Систематизирање

- На пример: Просек на податоците =
$$\frac{\sum X_i}{n}$$



Аналитичка статистика

- Оценување

- Да се процени просечната тежина на популацијата од просечната тежина на примерокот.



- Тестирање на хипотези

- Да се тестира тврдењето дека просечната тежина е 65 кг.



Процеси на изведување заклучоци за популацијата врз база на резултатите на **примерокот**