

STATISTICKÉ METODY V GEOGRAFII

Kdybych měl poslední den života, chtěl bych ho strávit na přednášce ze statistiky –
- je tak nekonečně dlouhá

Metody studia geografie

Úvod + geografie jako věda
statistické metody v geografii

Podmínky úspěšného zakončení:

Průběžný test - středoškolské znalosti zeměpisu, 20 otázek, min. 12

Získání zápočtu ze cvičení,
následně

Písemná zkouška, teorie, příklady

statistika

– definice

– pojetí

statistika - definice

***Statistika* je vědní obor zabývající se zkoumáním jevů, které mají hromadný charakter.**

***Statistika* je v určitém smyslu jazykem pro shromažďování, zpracování, rozbor, hodnocení a interpretaci hromadných jevů**

Hromadný jev ve statistice

Statistika se zabývá **hromadnými jevy** tj. jevy, které se vyskytují u souboru lidí, věcí, událostí buď v kvantitativní formě nebo i kvalitativní formě převoditelné na číselnou

hromadné jevy – příklady:

u souboru lidí hromadný jev: věk osob, váha, dosažené vzdělání **studenti dopíší další příklady (např. viz. statistické ročenky)....**

Co je typické pro statistiku

- Zkoumá **hromadné** jevy.
- Zabývá se proměnlivými - **variabilními** - vlastnostmi.
- Pracuje s čísly
- Používá výpočetní techniku.

statistika – dva významy - dvě pojetí

– vědní disciplína

- předmět zkoumání : stav a vývoj číselně vyjádřených hromadných jevů

– praktická činnost

- (zaznamenání, třídění, shrnování číselných údajů o skutečnostech, *udělám si statistiku*)

STATISTIKA jako vědní disciplína

- statistika popisná
- matematická statistika

Statistika jako praktická činnost

- **Činnost** Statistická evidence (např. sběr údajů, třídění, sumarizace apod.), ,
- **Instituce**, která tuto evidenci provádí (např. ČSÚ, ministerstva aj.)
- **výsledek** - Souhrn údajů o nějaké skutečnosti (statistika nezaměstnanosti, ročenka meteorologických pozorování atd.)

Základní etapy statistického zpracování dat

Zjišťování – zpracování – analýza - prezentace

- **1. Zjišťování/ Sběr údajů** - shromáždění a zaznamenání údajů, jejich kontrola aj.,
 - periodicita sběru:
- **2. Zpracování** – uspořádání, sumarizace,
- **3. Analýza** - výpočet charakteristik, měření závislostí, srovnávání, měření dynamiky
- **4. Prezentace** výsledků - tabulkové či grafické vyjádření a slovní zhodnocení výsledků předcházejících etap.

Základní dělení statistických údajů

- podle zdroje ODKUD?— **primární a sekundární,**
- podle periodicity zjišťování JAK ČASTO? — **průběžné, periodické a jednorázové,**
- podle časového hlediska KDY? ZA JAK DLOUHO?— **okamžikové a intervalové**
- podle reálnosti situace OPRAVDU?— **skutečné a simulované,**
-

Co statistika „umí“

- Zjišťovat
- Popisovat struktury
- Shrnovat dílčí ukazatele v čase a prostoru
- Srovnávat agregované ukazatele v čase nebo prostoru
- Měřit závislosti

... a co statistika „neumí“:

Statistika selhává, pokud:

- Nemá k dispozici adekvátní číselné údaje
- Nemá-li k dispozici dostatečně rozsáhlý soubor případů
- Není-li v datech přítomna proměnlivost (variabilita).

Statistika a výpočetní technika

- Výpočetní technika zasahuje do všech etap statistického zpracování dat.
- Exploze výpočetní techniky umožňuje provádět výpočty, které byly dříve nerealizovatelné (z důvodů velkého objemu dat, pracnosti, ...).
- Na druhou stranu však roste nebezpečí výběru nesprávného postupu.

Výhody počítačového zpracování I.

- **Přesnost a rychlost**

-

- **Univerzálnost**

-

- **Grafika**

- **Flexibilita**

- **Nové veličiny:** *Snadno lze vytvářet nové veličiny pomocí požadovaných transformací.*

- **Velikost datových souborů:**

- **Snadný přenos dat:**

...ale

Nevýhody počítačového zpracování I.

Chyby v softwaru.

Ne všechny statistické programy jsou spolehlivé.

Univerzálnost.

Může vést k výběru nevhodné metody zpracování.

Je velmi důležité, aby každý, kdo používá statistický software, si byl vědom úrovně svých statistických znalostí a užíval pouze ty metody, kterým rozumí. Pozor na používání neznámých statistických metod.

Černá skříňka.

.

Špatná data plodí špatné závěry.

.

Vymezení základních statistických pojmů

statistika - definice

***Statistika* je vědní obor zabývající se zkoumáním jevů, které mají hromadný charakter.**

Hromadné jevy:

jevy, které

které se vyskytují u souboru lidí, věcí, událostí

A jsou

výsledkem působení velkého množství příčin,

Příklady: kvalita vody – chem. složení, emise, produkce
odpadů,

zaměstnanost,

novorozenecká úmrtnost,

HMOTNOST NOVOROZENÝCH DĚTÍ

zatížení osob hlukem.....

Statistická jednotka: je to určitý jev či prvek, který je předmětem statistického šetření a pro který se zjišťují údaje

Statistická jednotka musí být přesně vymezena na počátku vlastního šetření a to z hlediska **věcného, časového, prostorového. (CO, KDY, KDE)**

Příklady:

stat. jednotka – občan, novorozenec, rodina, dům,, obec, výrobní podnik,

Co:

Kde:

Kdy:

Statistický znak:

je to určitá vlastnost statistické jednotky, kterou se snažíme postihnout.

Tzv. **shodné (společné) znaky** vymezují příslušnost statistické jednotky k určitému statistickému souboru.

Ostatní jsou znaky **proměnlivé (variabilní)**.

Příklady: **shodný znak – novorozenost**

**proměnlivé znaky – váha, délka,
jméno, národnost.....**

stat. jednotka
novorozenec

Statistické znaky lze dělit na znaky

- A) **prostorové**

místo narození: Brno

- B) **časové**

datum: 30.9.2011

- C) **věcné:**

- 1. **kvalitativní:**

pohlaví: muž

- **alternativní**

- **možné**

národnost: česká

- 2. **kvantitativní:**

- **spojité**

délka v cm: 55

- **diskrétní/nespojité**

Doplňte další příklady

Statistické znaky můžeme získat :

- **přímo** – (např. měřením, zvážením) tj.
.....**data**
- **nepřímo** (výpočtem), (znaky odvozené) tj.
.....**data**

Statistický soubor:

skupina statistických jednotek stejného druhu (věcně, prostorově a časově vymezených)

Je to množina všech prvků, které jsou předmětem daného statistického zkoumání.
Každý z prvků je statistickou jednotkou.

.

Prvky tvořící statistický soubor mají:

určité společné vlastnosti - tzv. **shodné - identifikační znaky**

- **sledované znaky** – tyto znaky statisticky šetříme

Příklad:

statistický soubor Novorozenci v ČR

Shodný - identifikační znak: novorozenost

sledovaný znak: váha, živý, pohlaví

Statistický soubor: Občané v produktivním věku

Shodný - identifikační znak:

Sledovaný znak:

Statistický soubor můžeme podle různých hledisek dále dělit:

Statistický soubor

- **jednorozměrný**
- **vícerozměrný**

Příklady

(váha dítěte), 1 –rozm.: 3650, 2100, 1200, 3500, 4100, 2800

dvourozm. (váha; délka), 3650, 55;

! jako dvojice! 2100, 47;

1200, 36,

3500, 50

Statistický soubor **základní a výběrový**

Výběrový soubor

je podmnožinou základního souboru. Je vytvořen ze statistických jednotek, vybraných podle určitého hlediska.

Př. Novorozenci v Jihomoravském kraji

Reprezentativní výběr:

Pokud zkoumaný výběr dobře odráží strukturu celého zkoumaného souboru, nazýváme jej reprezentativním výběrem.

Př. šetření průzkum volebních výsledků, peoplemetry

Rozsah statistického souboru:

počet statistických jednotek v souboru:

N – rozsah základního souboru

n – rozsah výběrového souboru

Grafické znázornění jevů

Grafické znázornění jevů

- **Graf – definice**
- – kresba podle pravidel znázorňující kvalitativní a kvantitativní informace

- **Základní prvky grafického znázornění:**
- 1.Název, příp. podnázev
- 2.vlastní kresba
- 3.stupnice a její popis (rovnoměrná, nerovnoměrná)
- 4.legendy/klíč
- 5.zdroj údajů
- vysvětlivky, poznámky,

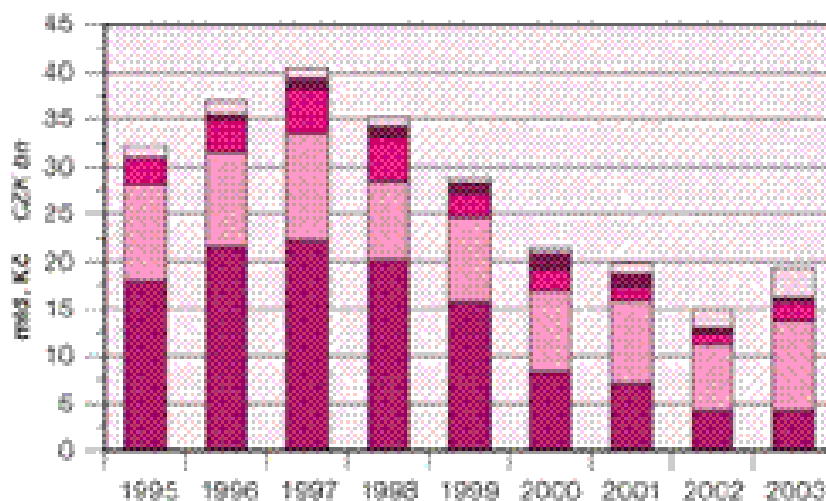
Graf – ukázka

3. ŽIVOTNÍ PROSTŘEDÍ

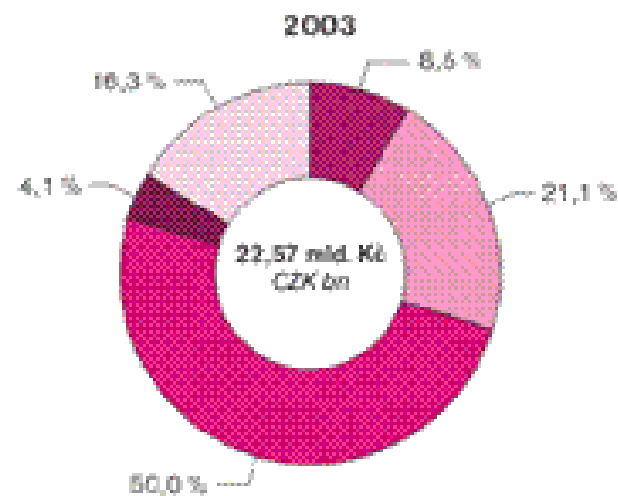
3. ENVIRONMENT

OCHRANA ŽIVOTNÍHO PROSTŘEDÍ ENVIRONMENTAL PROTECTION

POŘÍZENÉ INVESTICE
VALUE OF FIXED ASSETS ACQUIRED



NEINVESTIČNÍ NÁKLADY
NON-INVESTMENT EXPENDITURES



OCHRANA OVZDUŠÍ A KLIMATU
AIR POLLUTION CONTROL AND CLIMATE PROTECTION

OCHRANA KRAJINY A BIODIVERSITY PROTECTION
LANDSCAPE AND BIODIVERSITY PROTECTION

NAKLÁDÁNÍ S ODPADNÍMI VODAMI
WASTEWATER MANAGEMENT

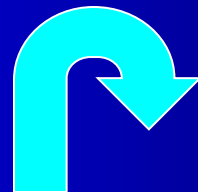
OSTATNÍ
OTHER

NAKLÁDÁNÍ S ODPADY
WASTE MANAGEMENT

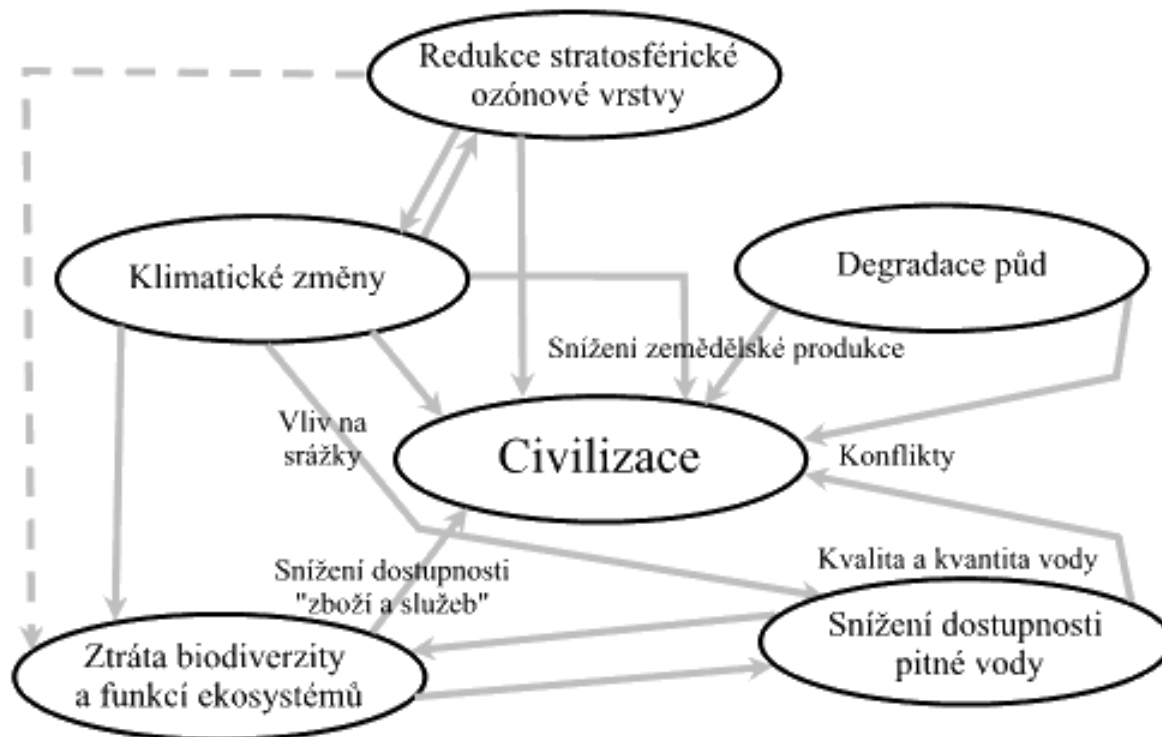
Český statistický úřad, 2003

Typy grafů

- **schéma** – znázorňuje strukturu a vztahy jevu či procesu
- **Příklad**
- **diagram** – znázorňuje kvantitativní údaje o souboru
 - sloupcový, bodový, plošný atd.
- **příklad**
- **statistická mapa** – prostorové rozložení prvku v podkladové mapě



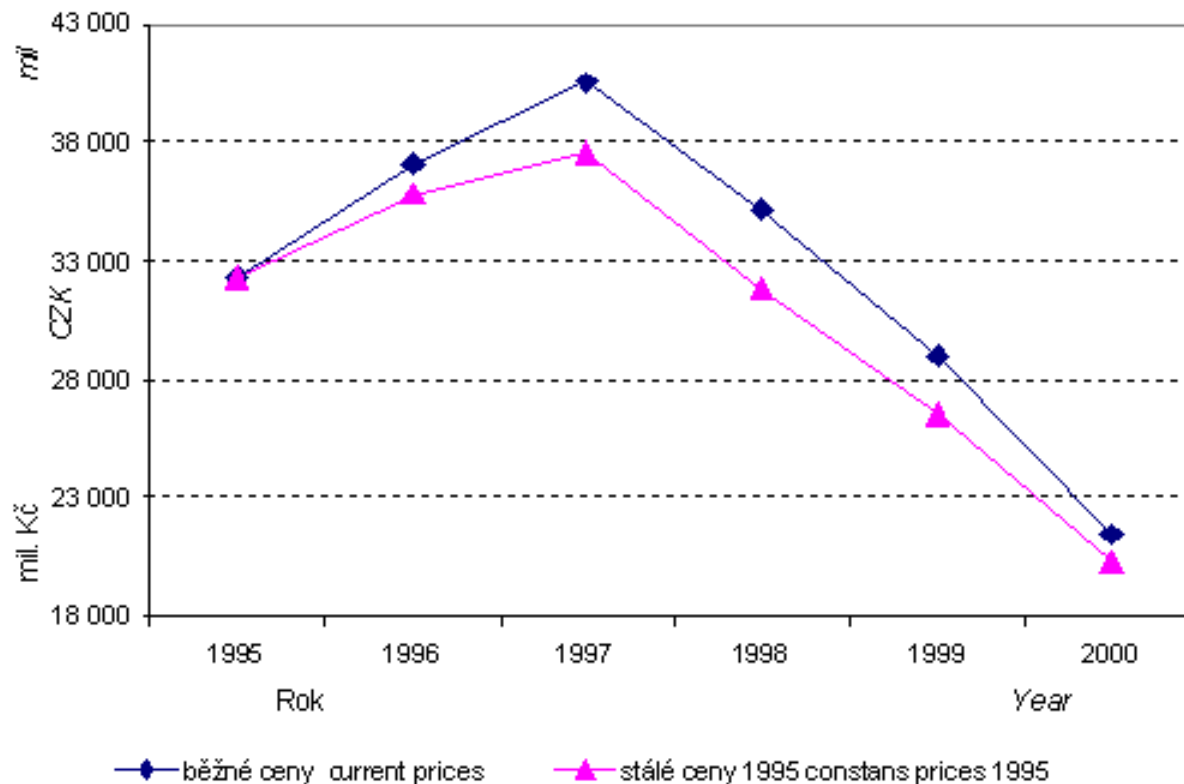
schéma



Diagram

Celkové investice na ochranu životního prostředí v běžných a stálých cenách roku 1995

INVESTMENT FOR ENVIRONMENT POLLUTION CONTROL PROJECTS: CURRENT PRICES AND 1995 CONSTANT PRICES



Diagram_ - věkové složení obyvv., tzv.pyramida života

4. OBYVATELSTVO

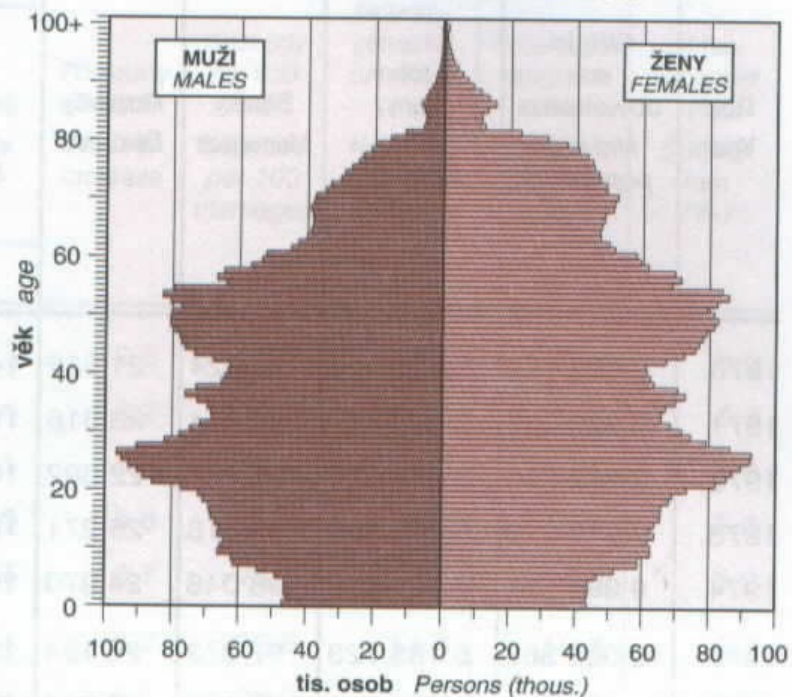
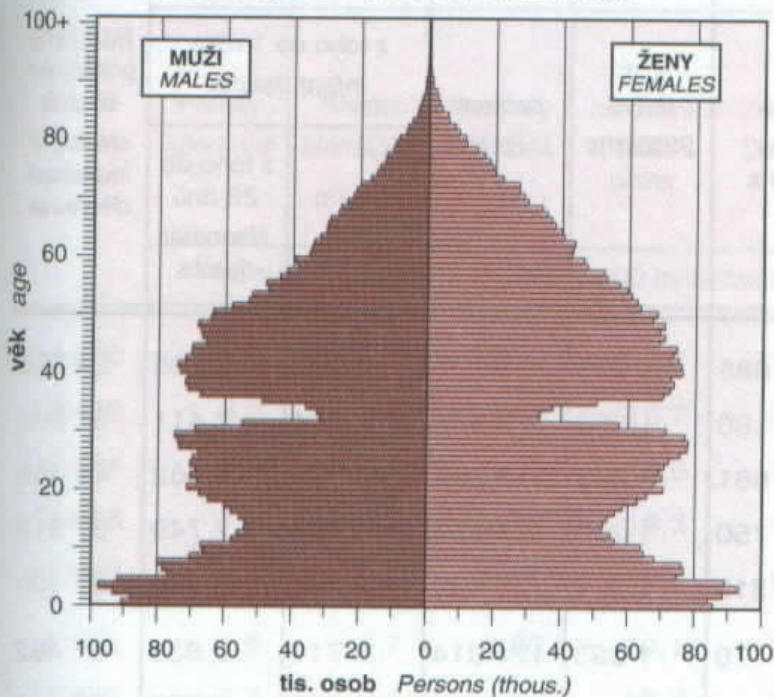
4. POPULATION

VĚKOVÉ SLOŽENÍ OBYVATELSTVA

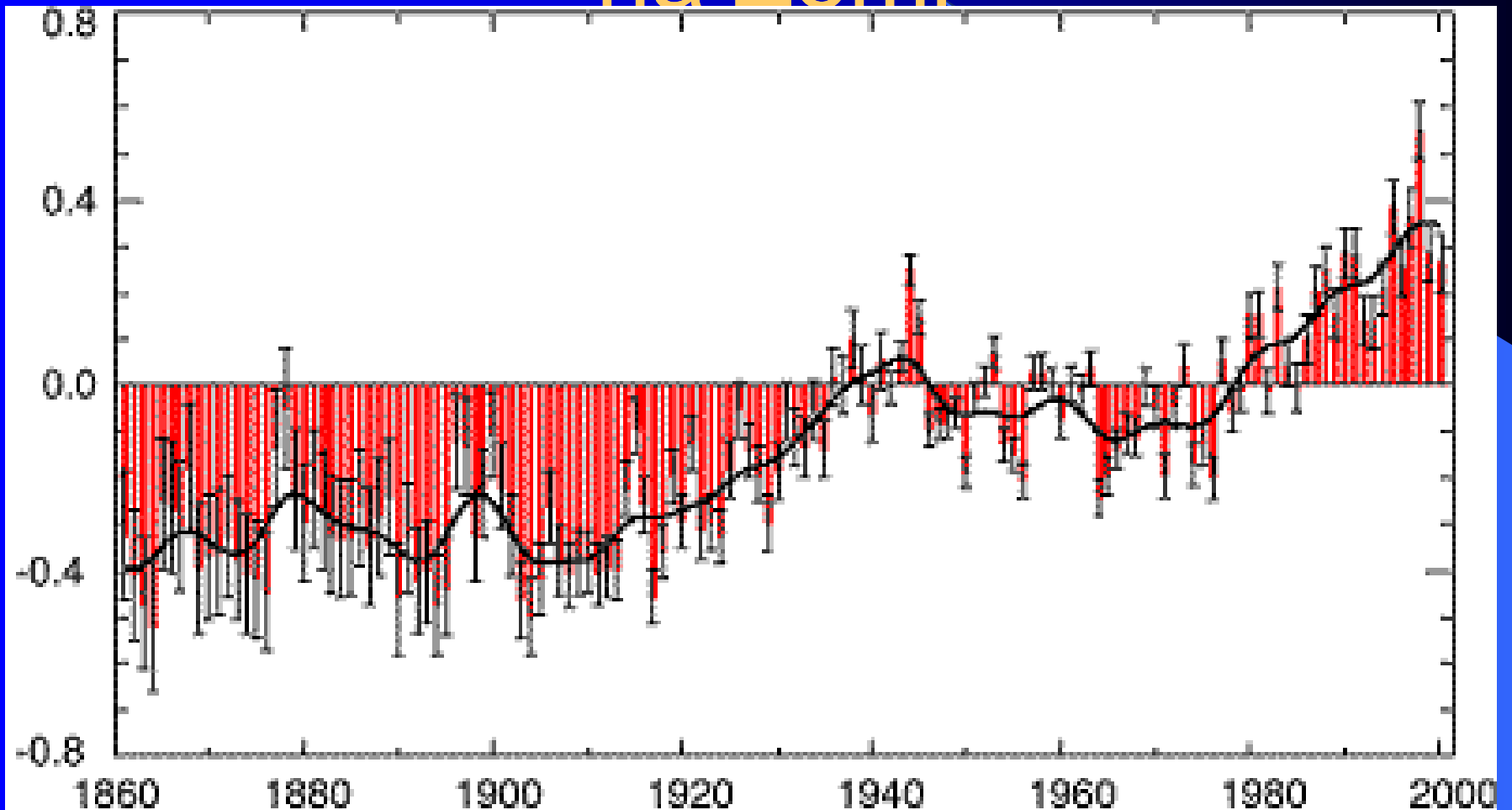
POPULATION BY AGE

k 31. 12. 1950 31 DECEMBER 1950

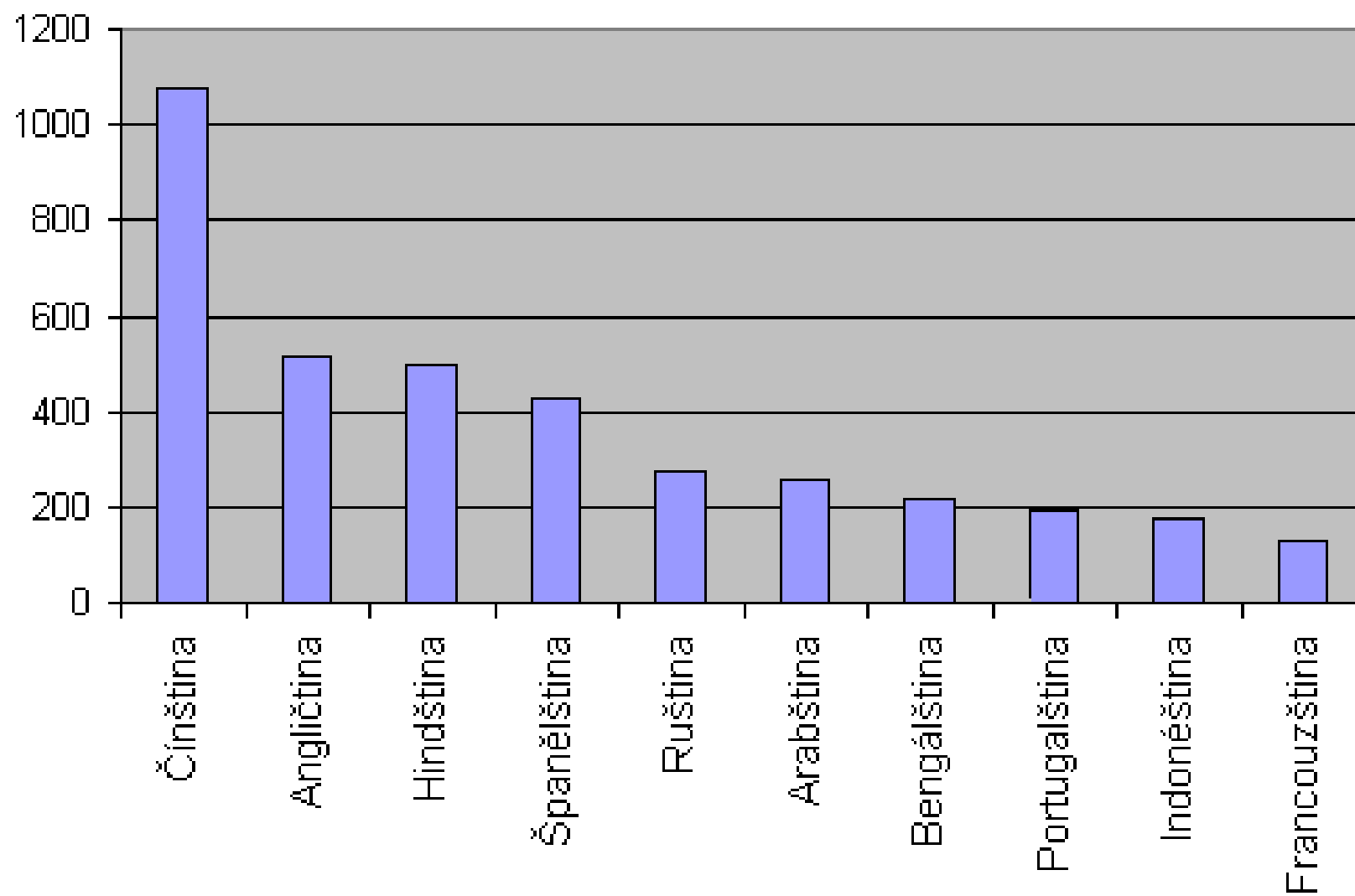
k 31. 12. 2000 31 DECEMBER 2000



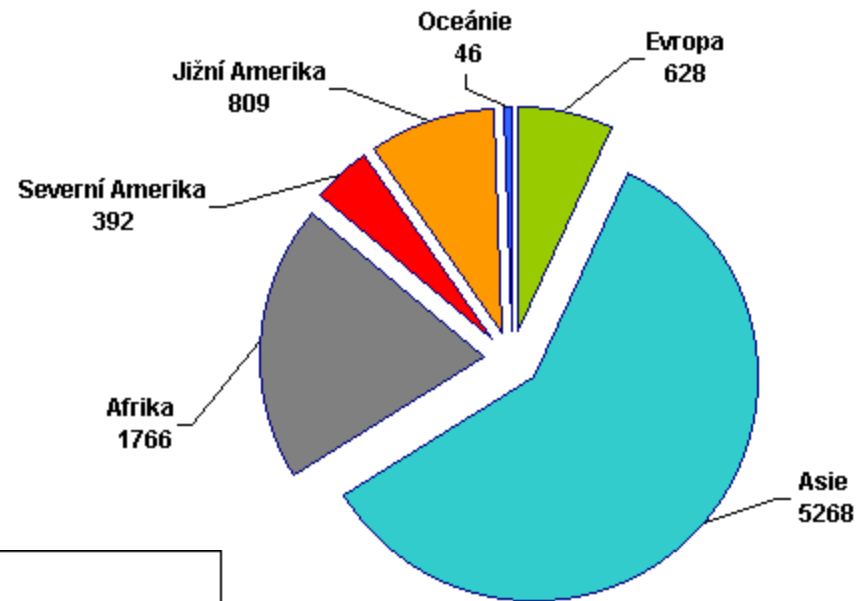
Odchylka od průměrné teploty na Zemi



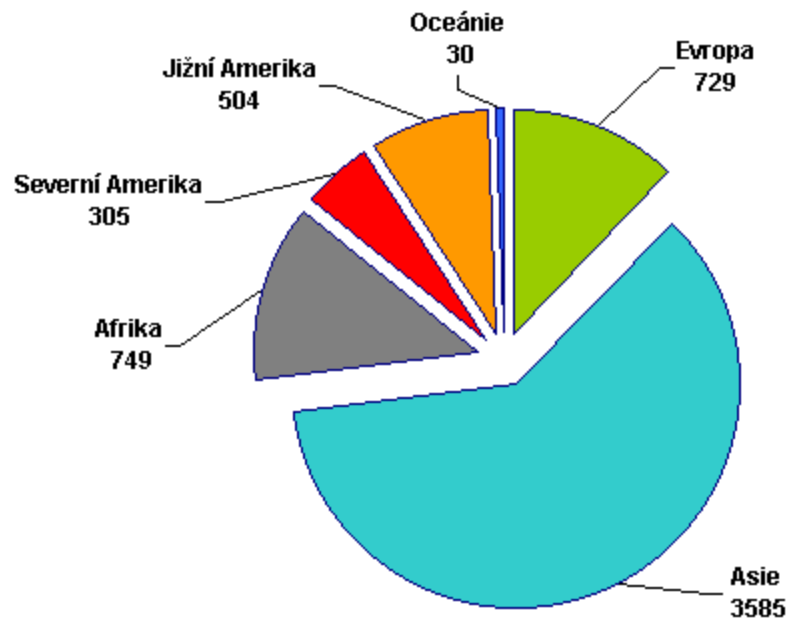
Nejrozšířenější světové jazyky v mil. uživatelů (1998)



Světová populace v roce 2050 (mil.)



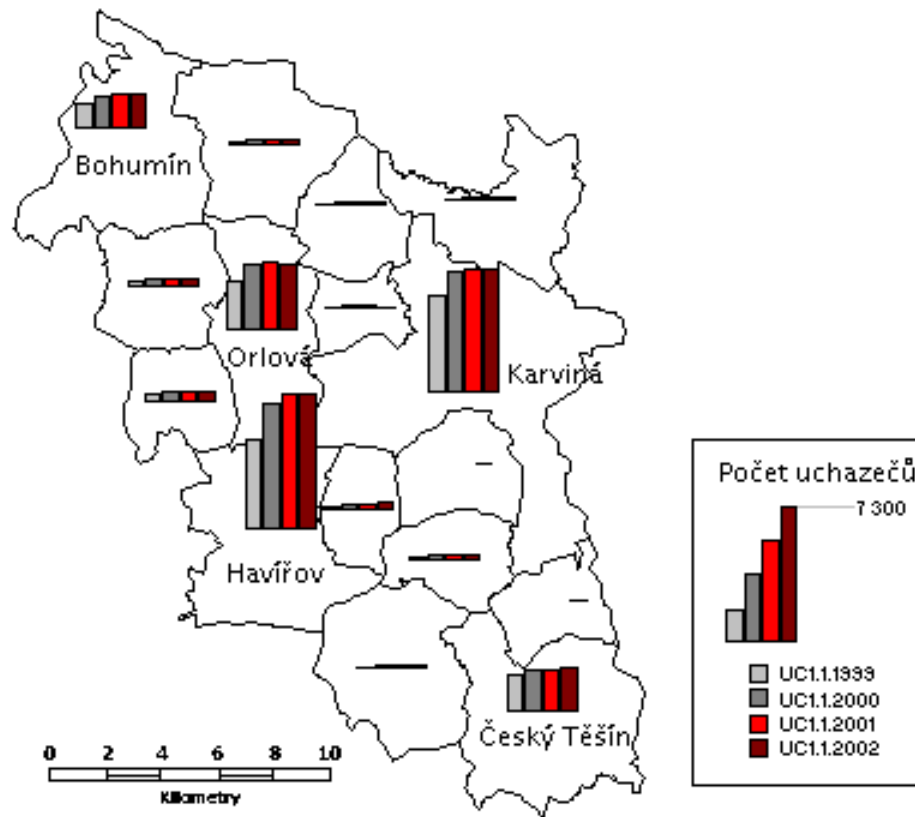
Světová populace v roce 1998 (mil.)



Statistická mapa

Vývoj počtu uchazečů od 1.1.1999 do 1.1.2002

obce okresu Karviná



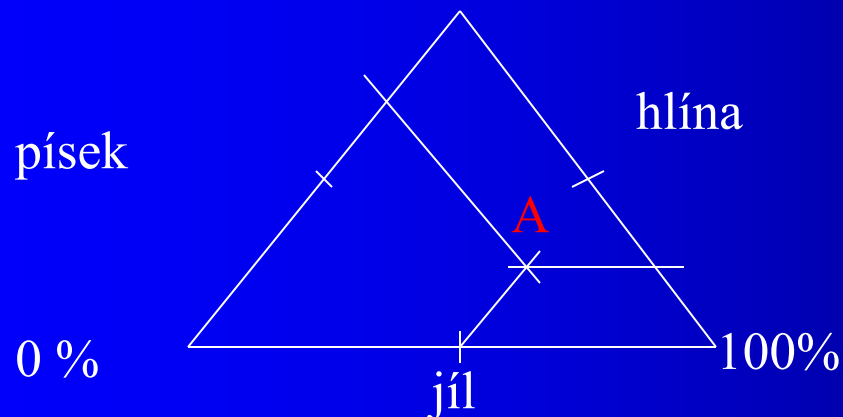
Použití grafických papírů při studiu geografických jevů

Grafický papír usnadňuje vynášení prvků do grafu.

- Milimetrový papír – rovnoměrné stupnice, čáry se jeví v původní, nezkrácené podobě
- Polologaritmický papír – kombinace dvou sítí – rovnoměrné a logaritmické
- Pravděpodobnostní papír – kombinace rovnoměrné a pravděpodobnostní stupnice

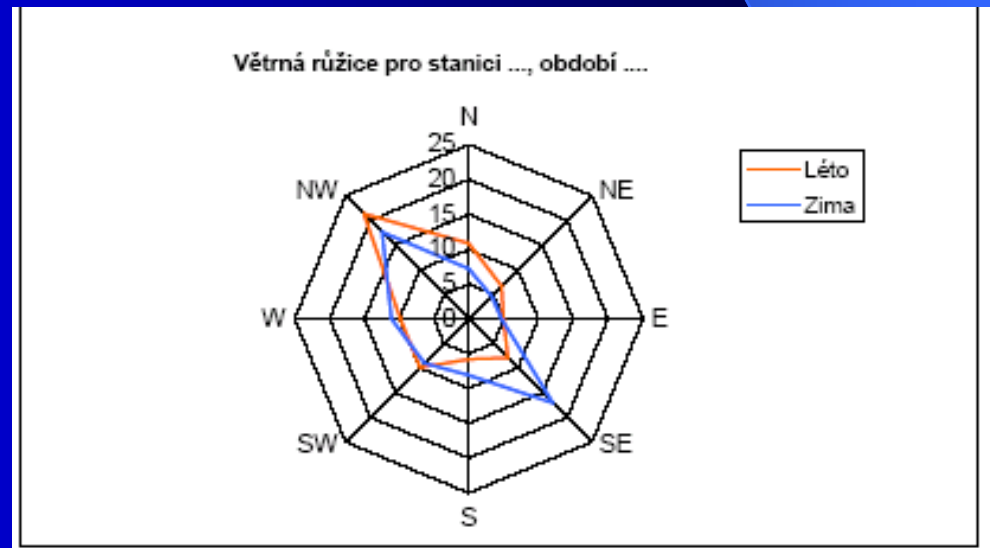
Sítě

- Trojúhelníková síť – znázorňování jevů o třech prvcích, které mají vždy konstantní součet
- např. půdní druhy
- půda A:: 50 % jílu, 25% hlíny, 25%, písku
- př. **Vzdělání**



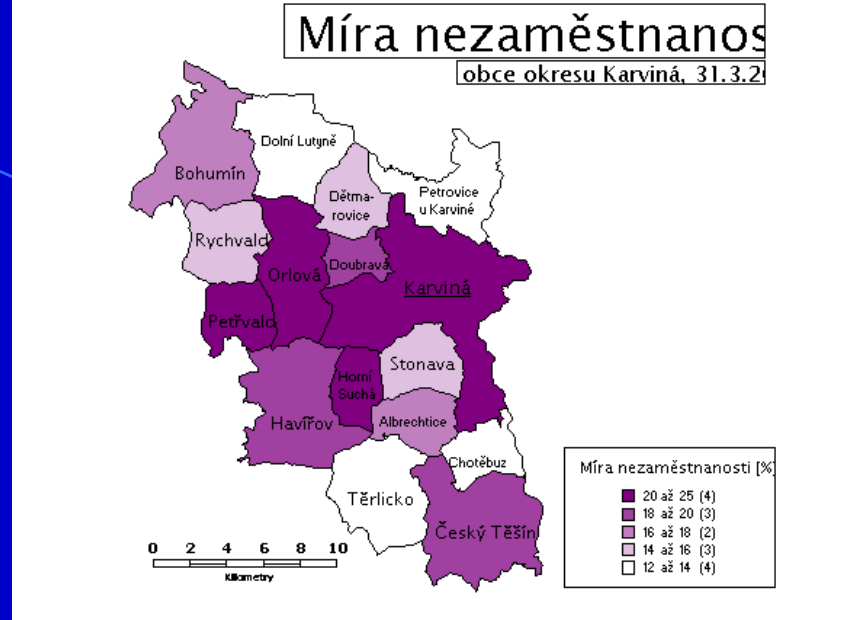
Sítě

- Kruhová (radiální) síť – kombinace soustředných kružnic a přímek procházejících středem kružnice
- pro grafické znázorňování opakujících se jevů, struktury jevů
- Příklad
- roční chod teploty
- směry větru



statistická mapa:
kartogram
kartodiagram

kartogram

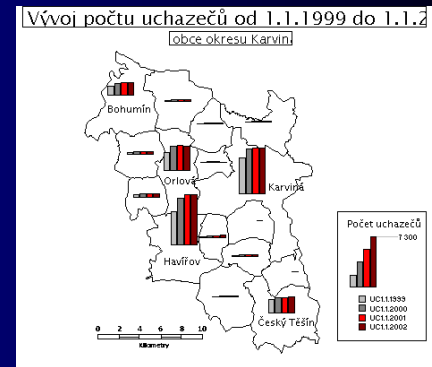


Kartogram je obrysová kartografická kresba územních celků, ve kterých jsou grafickým způsobem (barevný odstín, rast) plošně znázorněna statistická data týkající se různých geografických jevů (lidnatost, využívání ploch apod.)

Kartogramy lze rozdělit podle územního dělení
na:

- kartogramy s geografickými hranicemi
- kartogramy s geometrickými hranicemi

Kartodiagram

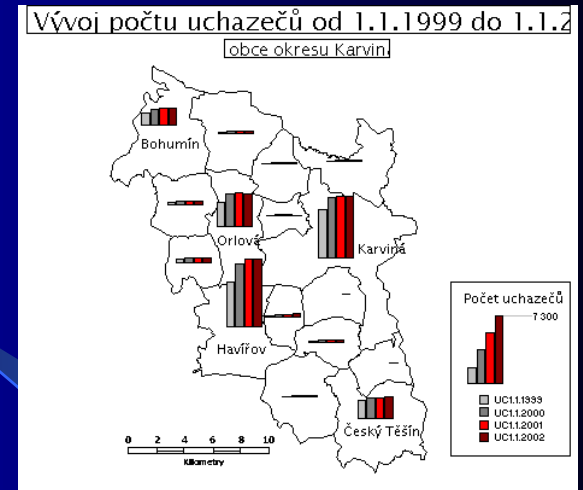


Kartodiagramy jsou diagramy vložené do mapové kostry, kterou tvoří dílčí územní celky.

Jejich údaje se vztahují na celé území jednotky, kde leží

(rozdíl od metody **lokalizovaných diagramu** – údaj vztahující se k urč. bodu – např. chod roční srážek na meteorolog. stanici)

Kartodiagramy



Vkládanými diagramy mohou být:

- Spojnicové diagramy pro vyjadřování časových řad
- sloupcové diagramy (sloupce, věkové pyramidy apod.)
- různě dělené geometrické značky

Grafické metody analýzy geografických jevů

- 1. znázornění intenzity jevu v prostoru
- a) absolutními metodami
- * značková metoda (velikost značky odpovídá velikosti jevu)
- * bodová metoda (počet prvků....velikost jevů)
- b) relativními metodami (např. šrafovaní-hustota obyv.)

př

Grafické metody analýzy geografických jevů

- 2. znázornění **struktury** jevu v prostoru
- využití výsečových grafů
- *pouze strukturu vyjádříme výsečovými grafy se stejným poloměrem
- *strukturu a velikost celku (výsečový graf + velikost poloměru odp. velikosti jevu)

Náležitosti statistické mapy

Obsah mapy tvoří všechny objekty, jevy a jejich vztahy, které jsou v mapě kartograficky znázorněny

Základní údaje tvoří

- Název mapy - stručně a výstižně charakterizuje zobrazené území, druh mapy
lze i název hlavní a vedlejší)
- Mapový rámeček – „vlastní mapa“
- Měřítko v číselné, grafické nebo slovní formě
- Legenda (vysvětlivky) – podávají výklad použitých mapových značek a ostatních kartografických vyjadřovacích prostředků včetně barevných a velikostních stupnic, legenda musí být:
 - Úplná, logicky uspořádaná, přehledná a zapamatovatelná, **POZOR na intervaly, na barevnou škálu**
- Autoři

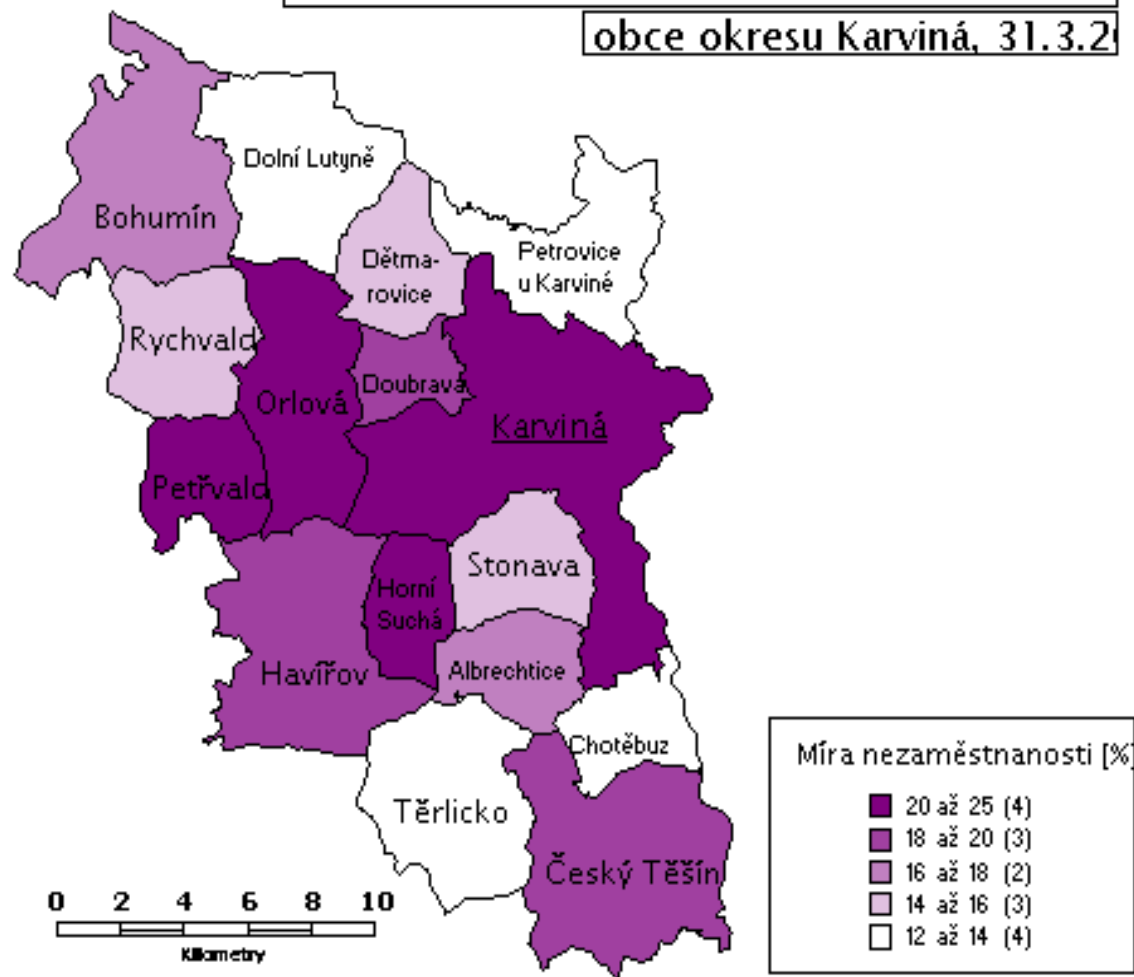
Dalšími údaji mohou být :

- vyznačení severu nebo směrová růžice, souřadnicový systém, přehled použitých mapových podkladů, datum, ke kterému se obsah mapy vztahuje
- obrázky, grafy, tabulky, text

Hledejme chyby

Míra nezaměstnanosti

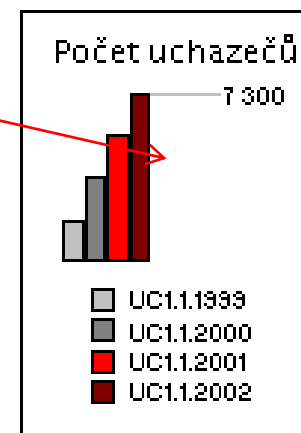
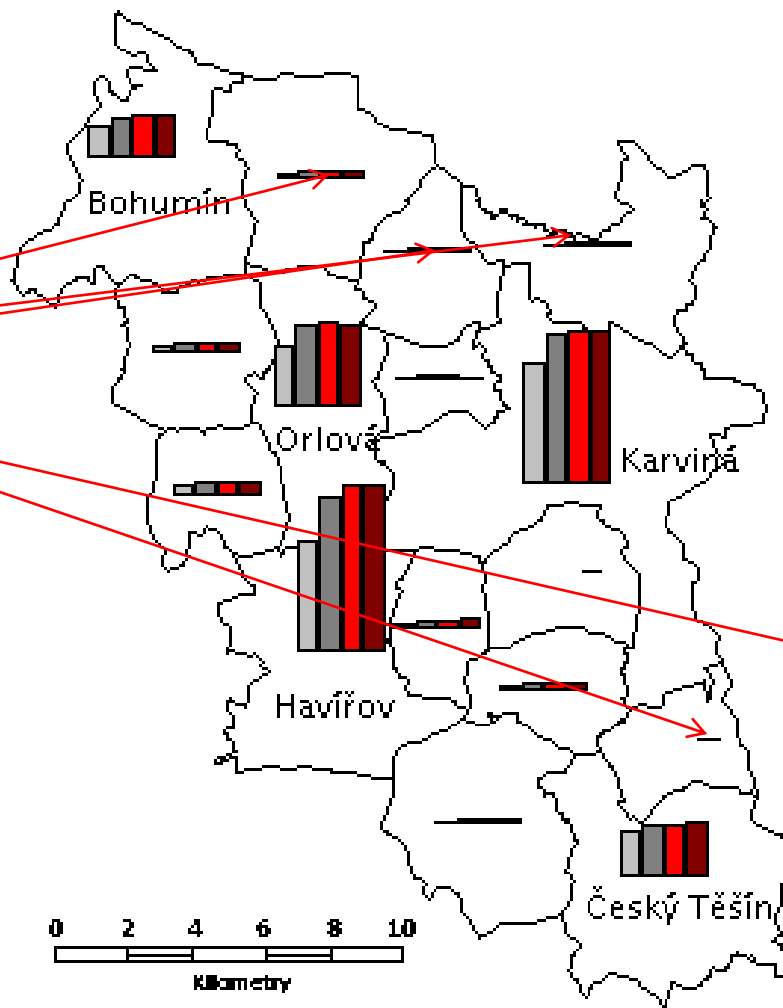
obce okresu Karviná, 31.3.2002



Hledejme chyby

Vývoj počtu uchazečů od 1.1.1999 do 1.1.2002

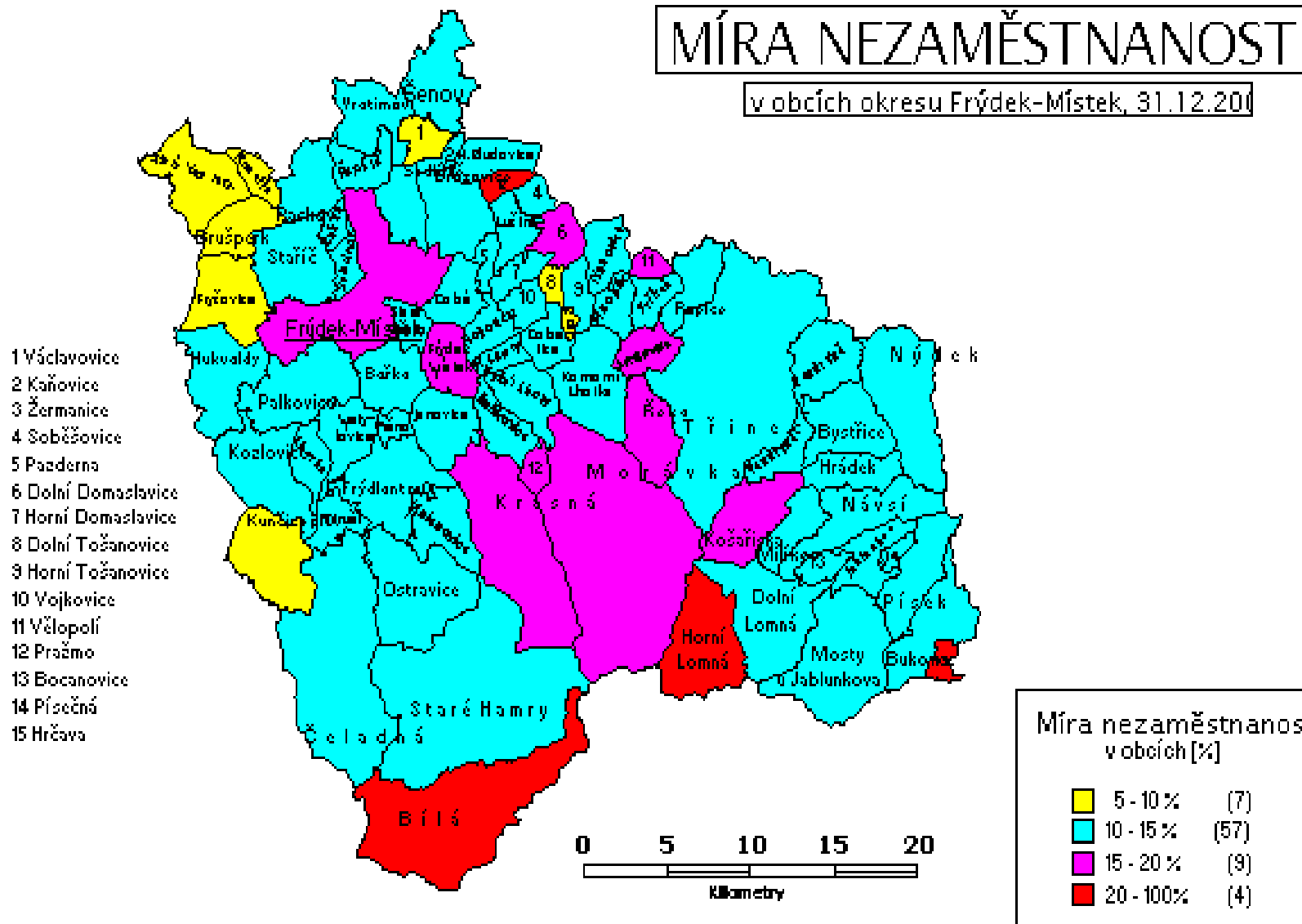
obce okresu Karviná



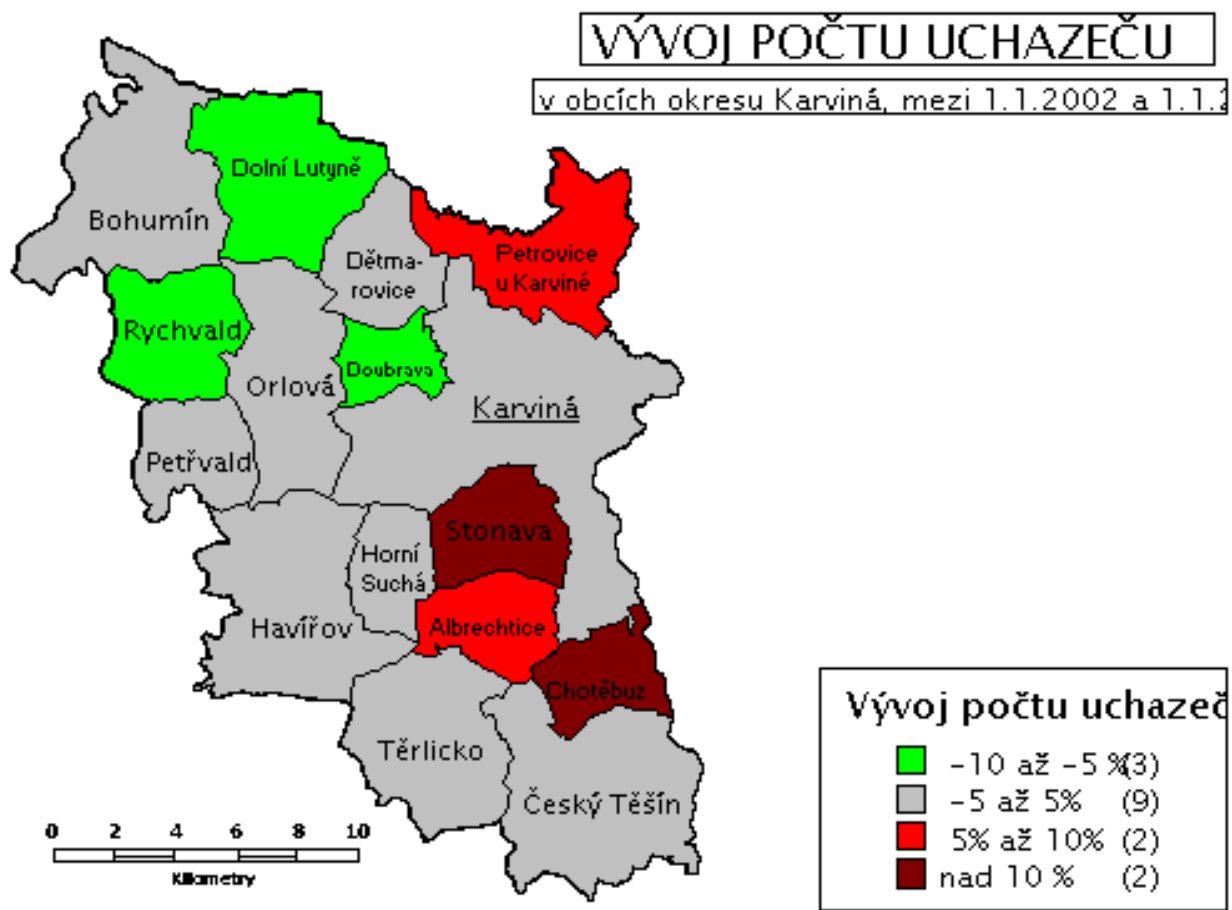
studenti samostatně

MÍRA NEZAMĚŠTNANOSTI

v obcích okresu Frýdek-Místek, 31.12.2011



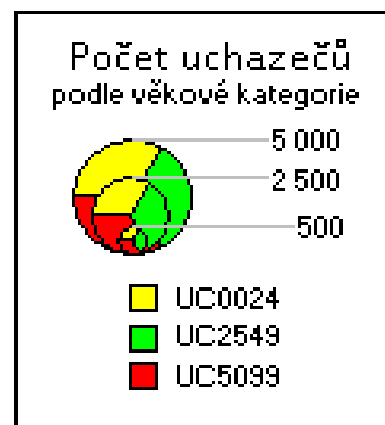
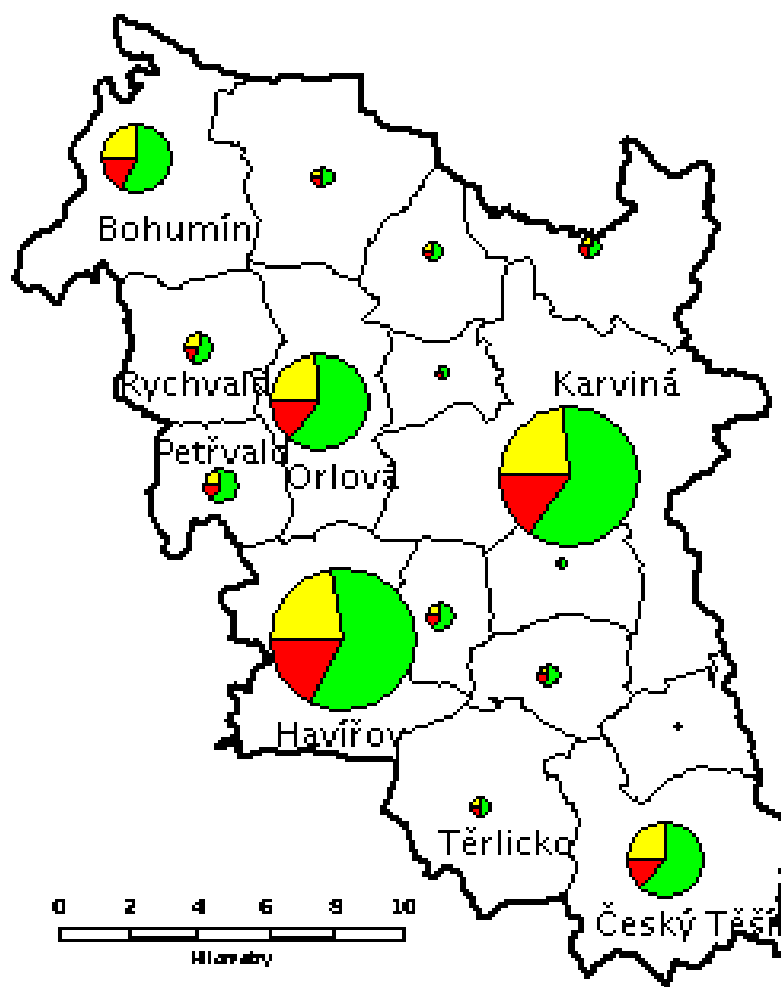
Hledejme chyby



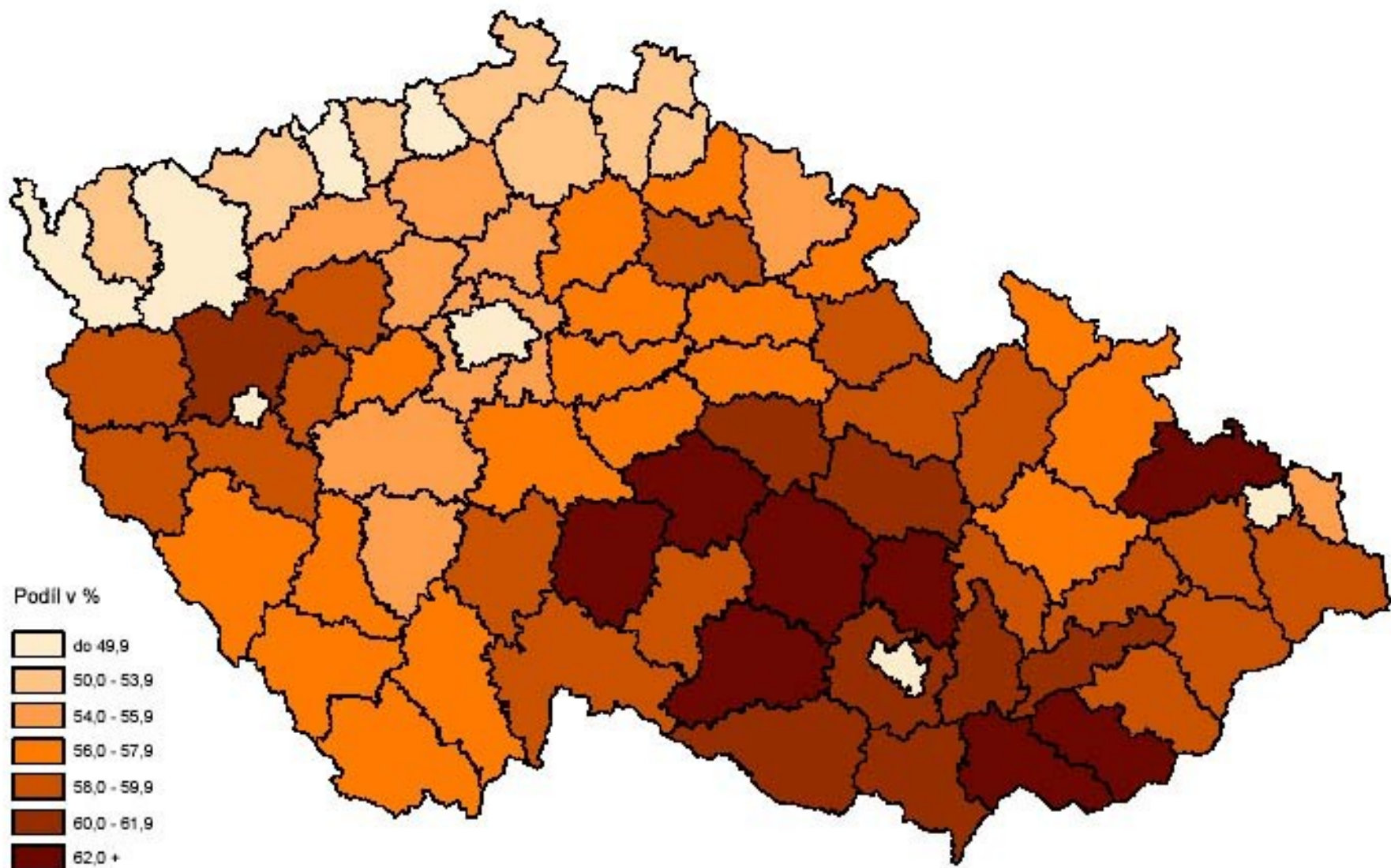
Jak byla vymezena st.
jednotka?

Vel .stupnice?:

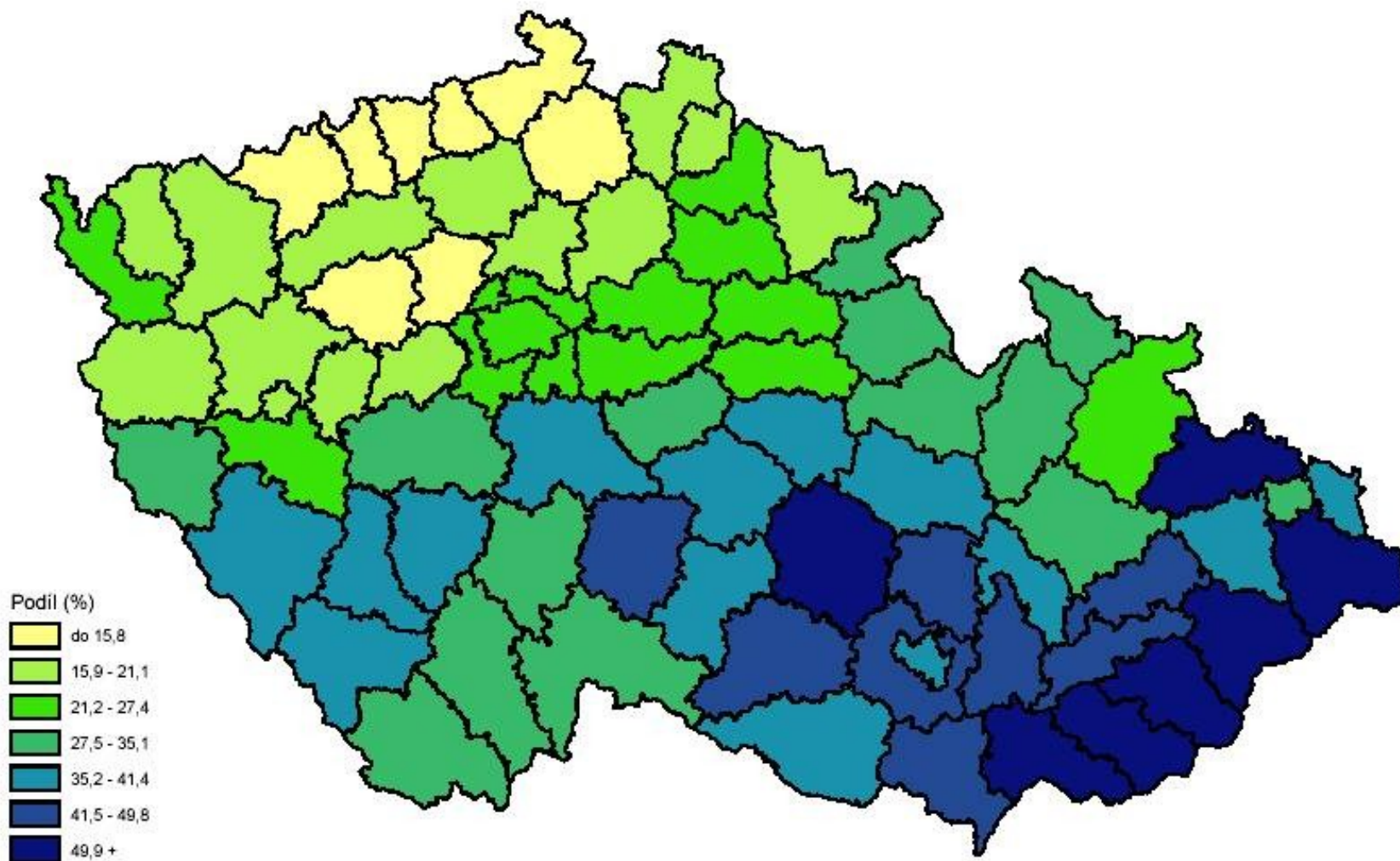
Barevnost?:



Podíl úplných rodin z úhrnu domácností - SLDB 2001



Podíl obyvatel s náboženským vyznáním - SLDB 2001

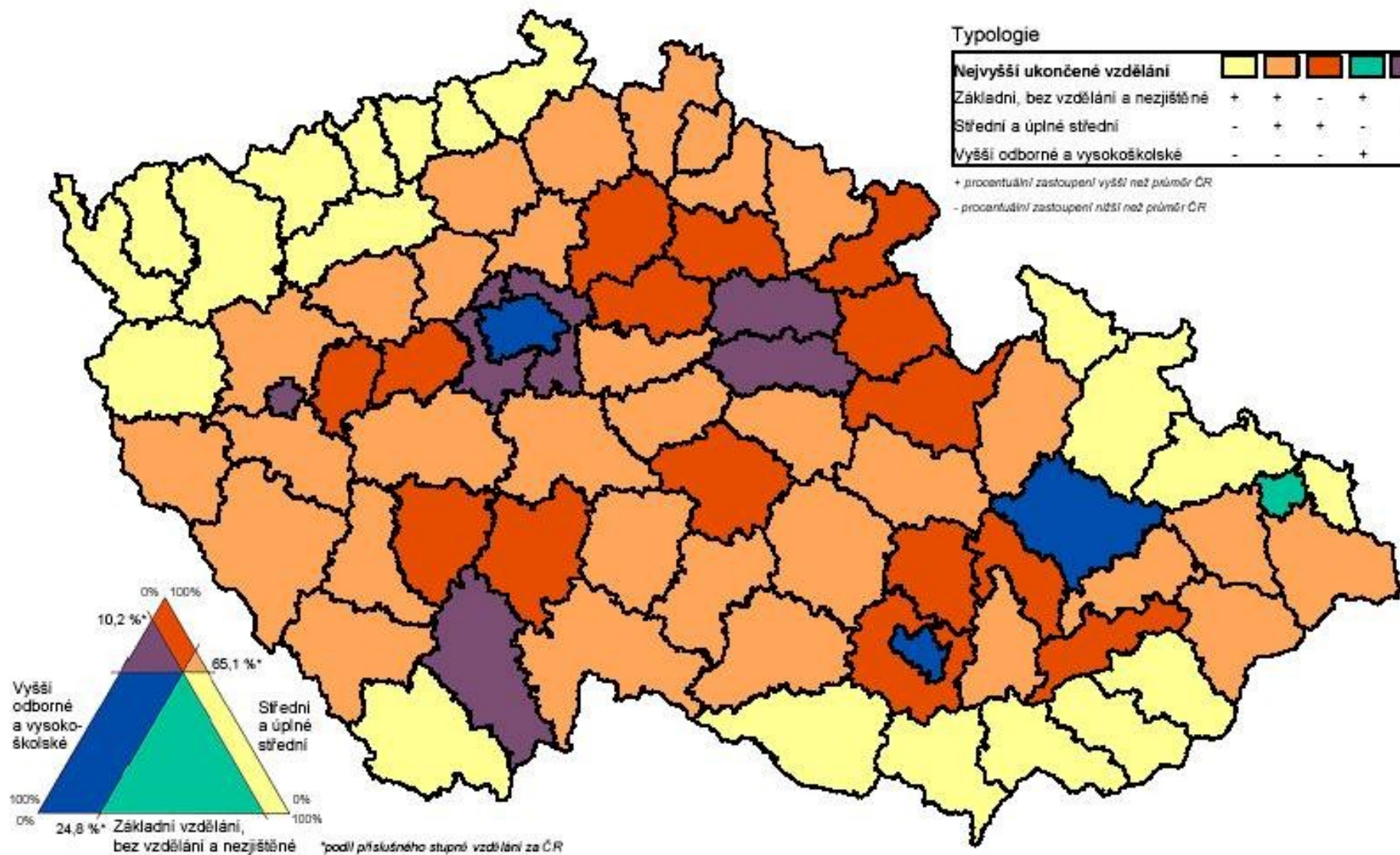


Typologie okresů ČR dle nejvyššího ukončeného vzdělání obyvatel ve věku 15 a více let - SLDB 2001

Typologie

Nejvyšší ukončené vzdělání	Yellow	Light Orange	Dark Orange	Teal	Purple
Základní, bez vzdělání a nezjištěné	+	+	-	+	-
Střední a úplné střední	-	+	+	-	+
Vyšší odborné a vysokoškolské	-	-	-	+	+

+ procentuální zastoupení vyšší než průměr ČR
 - procentuální zastoupení nižší než průměr ČR



Izolinie – konstrukce a vlastnosti

- Izolinie – čáry, které v grafu spojují body se stejnou intenzitou (velikostí, hodnotou) jevu
- získávají se metodou prostorové interpolace hodnot vynesných do grafu
- plynulé čáry
- izobary, izotermy, vrstevnice atd.
- Konstrukce izolinie - příklad

Rozdělení četností

Absolutní, relativní kumulované četnosti

- četnost – počet výskytu určité hodnoty v souboru, frekvence hodnoty
- rozdělení četností – počty prvků s určitými hodnotami statistického znaku, obvykle pro nespojitě hodnoty
- skupinové rozdělení četností - počty prvků s hodnotami statistického znaku, které patří do určitého intervalu, obvykle pro spojité hodnoty

skupinové rozdělení četností

- roztrídíme statistické jednotky podle velikosti jejich statistického znaku do intervalů
- interval – hranice, dolní a horní mez, šířka (délka)

zásady:

- vymezené hranice pro jednoznačné zařazení prvků
- obvykle stejná šířka
- přiměřený počet intervalů

Četnosti

- absolutní četnost – počet jednotek v intervalu
- relativní četnost – podíl četností na rozsahu souboru
- kumulovaná četnost – počet jednotek s hodnotami menšími nebo rovny horní hranici intervalu
- příklad

Tab.S Skupinové rozdělení četností,
ukázka – příklad váha 50 novorozenců v JMK

Interval	střed	abs. č.	relativ. č.	kumul. abs.	kumul. relat.
500 - 1000	750				
1001 - 1500	1250				
1501 - 2000	1750				
atd.					
		50	100%		

Grafické znázornění rozdělení četností

- histogram
- polygon
- čára kumulovaných četností

čára kumulovaných četností – součtová čára,
graf kumulované četnosti, vždy k horní hranici intervalu

Histogram

Histogram – sloupcový diagram,
šířka sloupce – šířka intervalu, výška sloupce - četnost

náčrt

Polygon

Polygon – spojnicový diagram,
hodnoty četnosti se vynáší ke středům intervalu

náčrt

Čára kumulovaných četností

čára kumulovaných četností – součtová čára,
graf kumulované četnosti, vždy k horní hranici
intervalu

náčrt

histogram – věkové složení obyvatelstva, věková struktura, pyramida života

4. OBYVATELSTVO

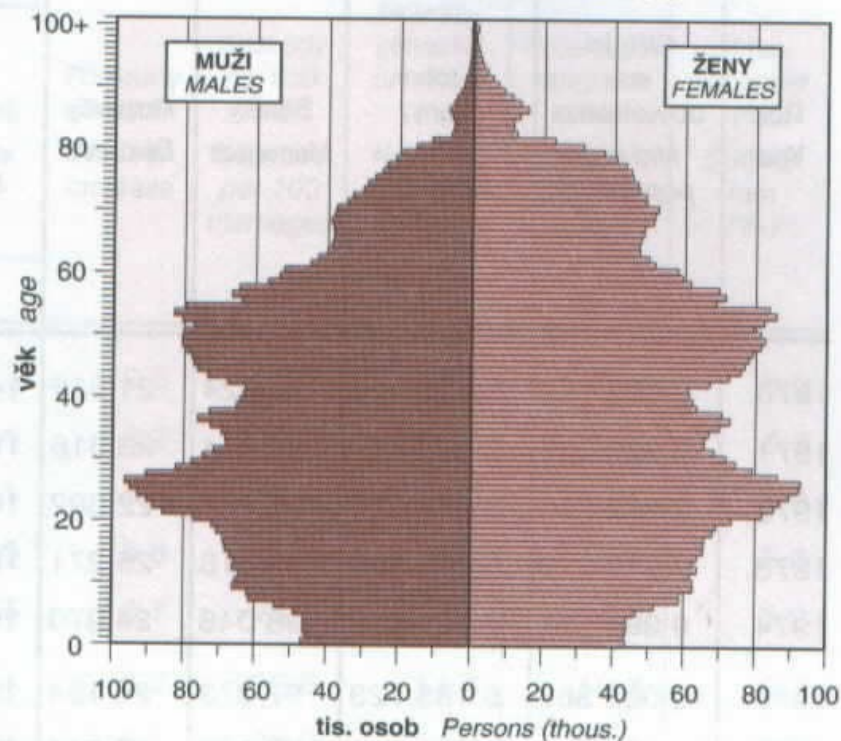
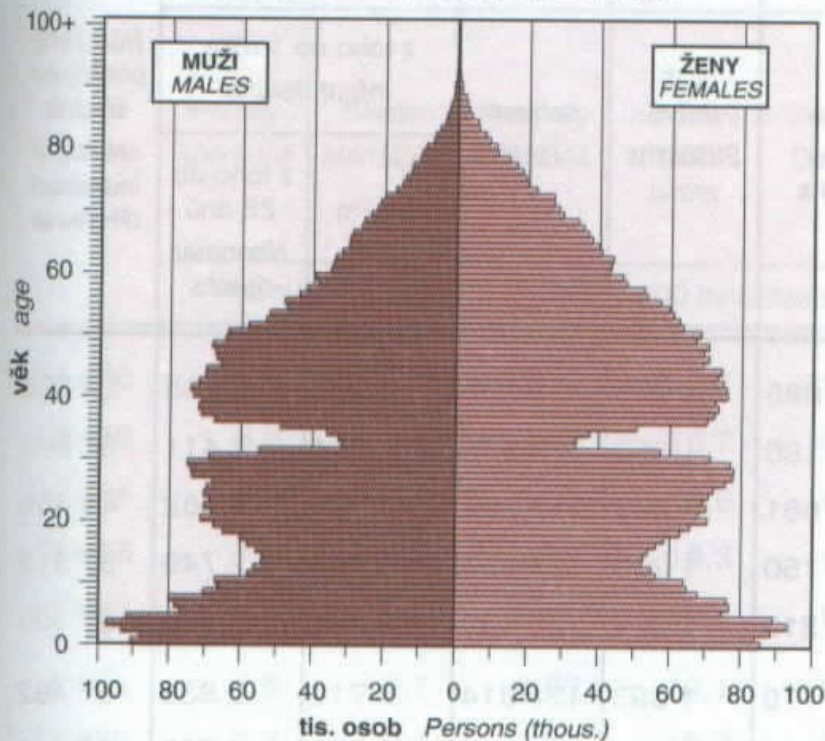
4. POPULATION

VĚKOVÉ SLOŽENÍ OBYVATELSTVA

POPULATION BY AGE

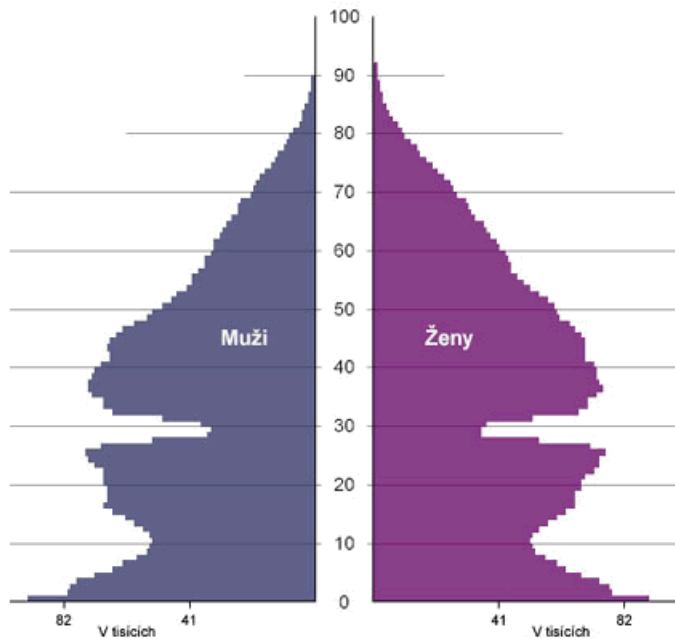
k 31. 12. 1950 31 DECEMBER 1950

k 31. 12. 2000 31 DECEMBER 2000



Věková skladba obyvatelstva: 1946

Česká republika



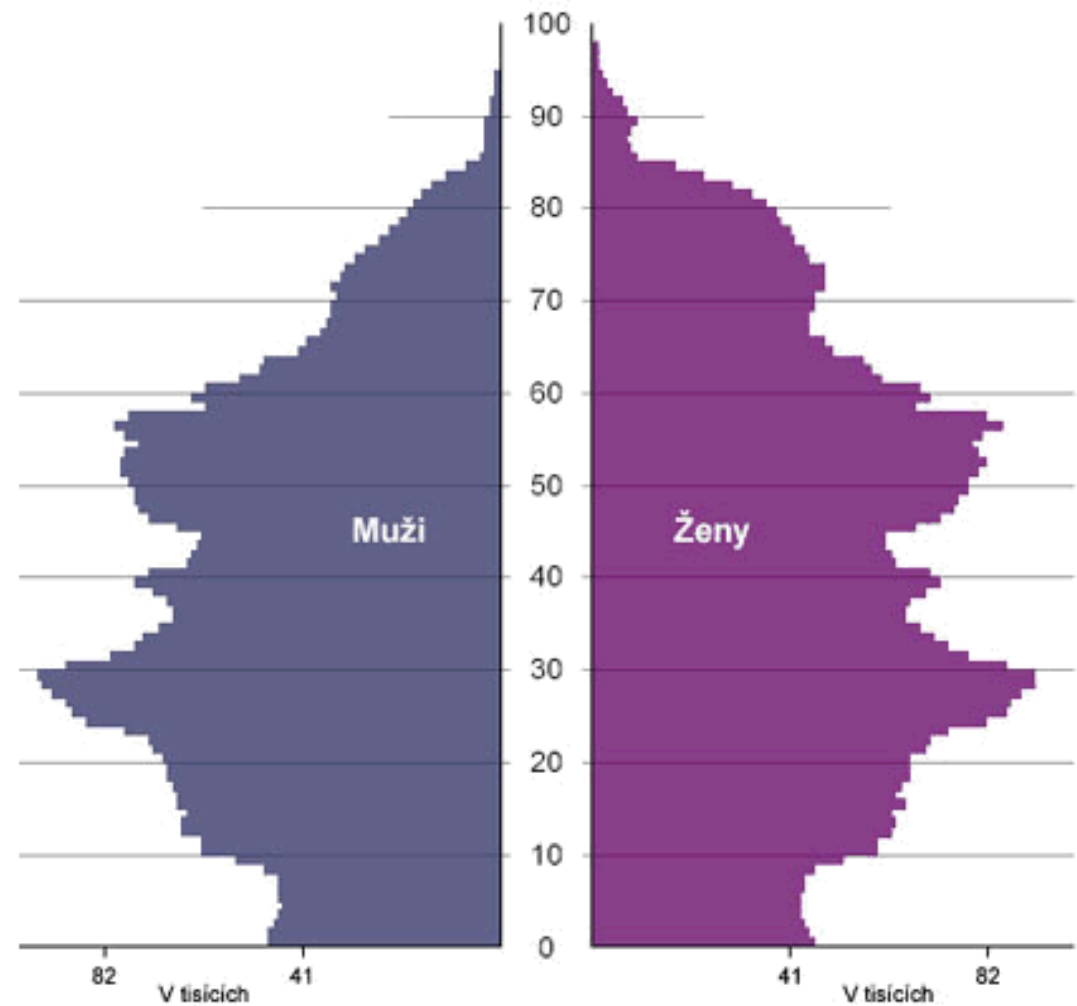
Věková skladba obyvatelstva: 1985

Česká republika

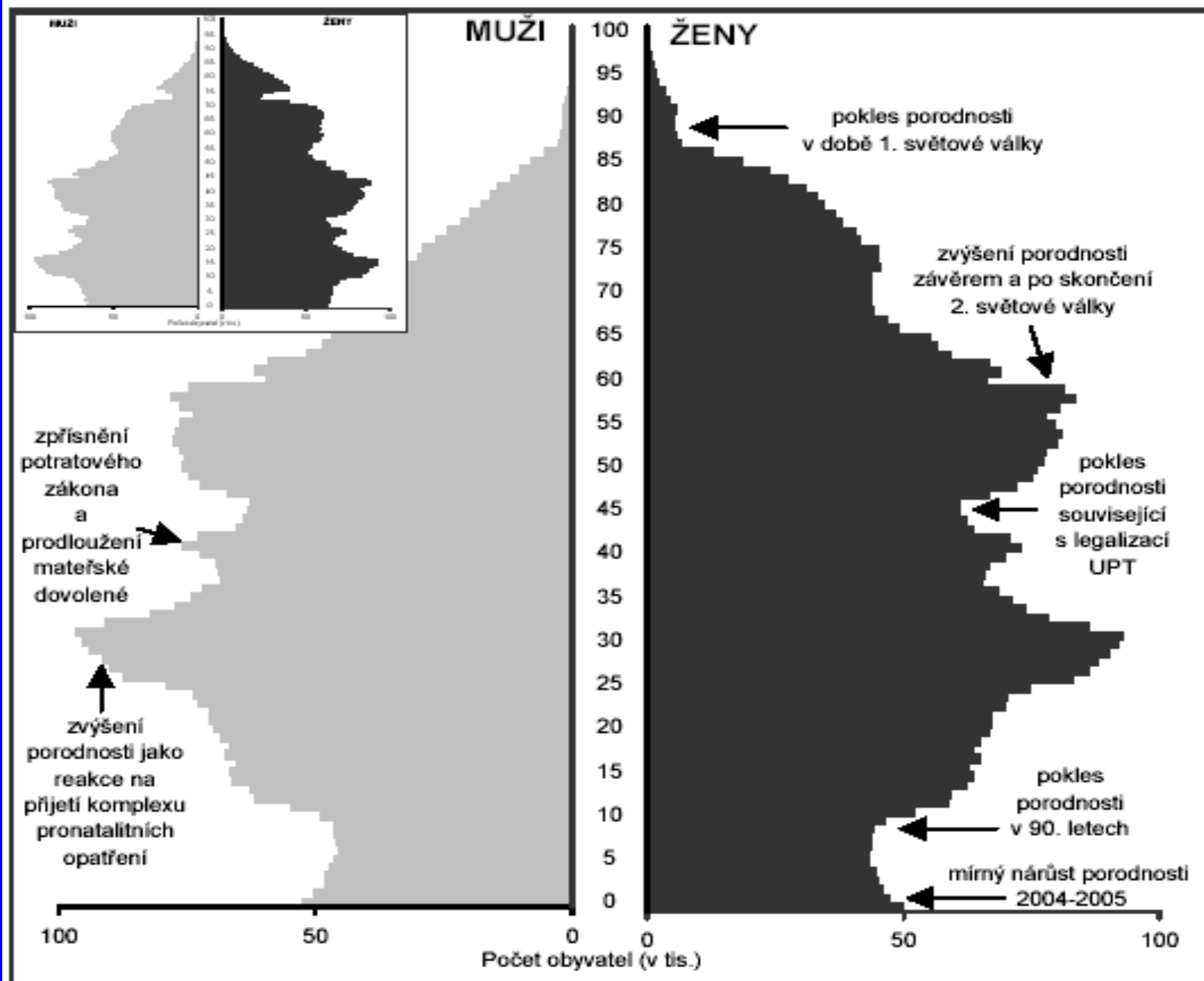


Věková skladba obyvatelstva: 2003

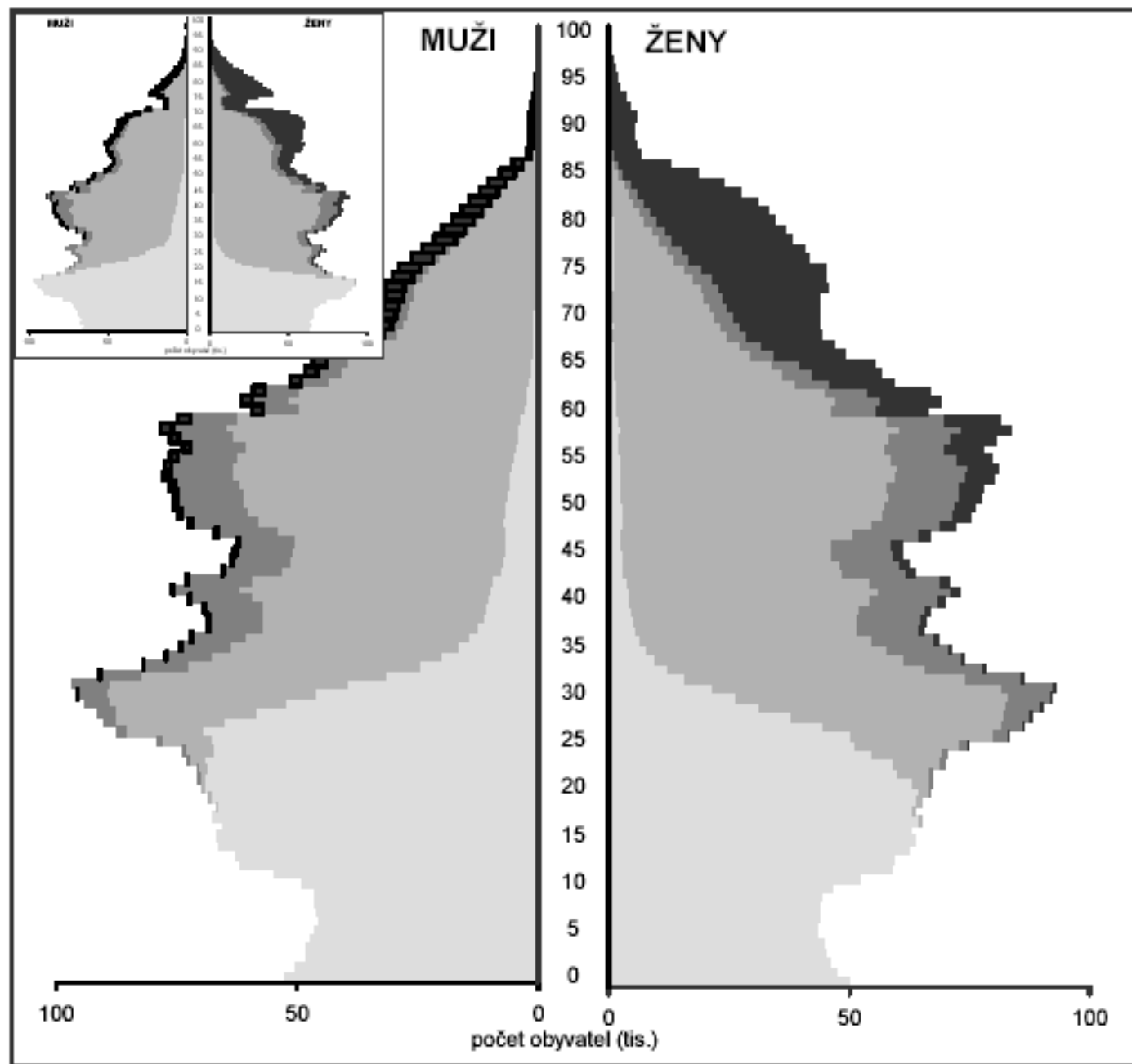
Česká republika



Obr. 1.1 Věkové složení obyvatel, 31.12.2005 a 31.12.1990 (malý obrázek)



Obr. 1.4 Složení obyvatel podle věku a rodinného stavu, 31.12.2005 a 31.12.1990 (malý obrázek)



Základní statistické charakteristiky

Základní statistické charakteristiky

Popisná statistika

- základní statistické charakteristiky „popisují“ statistický soubor
- a) charakteristiky úrovně – tzv. střední hodnoty
- b) charakteristiky variability
- c) charakteristiky asymetrie a špičatosti

Střední hodnoty

- Místo jednotlivých hodnot u jednorozměrného statistického souboru používáme často střední hodnoty
- Střední hodnoty umožňují porovnávání souborů

Střední hodnoty

- aritmetický průměr (+ vážený aritm. průměr, geometrický průměr, harmonický průměr)
- modus
- aritmetický střed
- medián a kvantily
- geografický medián

Aritmetický průměr

- nejčastěji používaná st. charakteristika
- typický a netypický průměr
- (jedno a více vrcholová rozdělení četností)
- typický aritm. průměr – jednovrcholové rozdělení četností + blízký nejčetnější hodnotě

Obr.

Vážený aritmetický průměr

- při výpočtu množství srážek v povodí – váha – plocha území
- v klimatologii – výpočet denního průměru teplot ze tří měření

Př. výpočtu

Modus

- modus - nejčetnější hodnota kvantitativního znaku ve studovaném souboru
- významný především u souboru nespojitých veličin
- modální interval – interval zahrnující největší počet jednotek, závisí však na stanovení hranic intervalů
- rozdělení s více mody – polymodální rozdělení

příklad

Aritmetický střed

- Aritm. střed je polovina součtu min. a max. hodnoty znaku v souboru
- pokud soubor obsahuje extrémní hodnoty, je aritmetický střed značně zkreslující charakteristika

příklad

Medián

- Medián – tzv. prostřední hodnota,
- je to prvek řady uspořádané v neklesajícím pořadí (od nejm. po největší), který ji dělí na dvě poloviny, které mají menší a větší hodnotu znaku
- POZOR: soubor je třeba vždy uspořádat
- pořadí prvku (kolikátý prvek to je, hodnota prvku je medián!) určují vzorce :
- pro řadu s lichým počtem prvků $(n+1)/2$,
- pro řadu o sudém počtu je medián průměr z hodnot mezi prvky na $(n/2)$ a $(n/2+1)$ místě
- **Příklad**

Kvantily

- Medián je kvantil dělící soubor na dvě poloviny dle předch. pravidel

obdobně

- kvartily – na čtvrtiny, x_{25} , x_{50} , x_{75} ,

- decily

- percentily

kvantily obecně široké použití ve statistice a v geografii

příklad

Geografický medián

- Geografický medián je čára dělicí plochu, kde se jev vyskytuje tak, aby hodnota jevu byla v obou plochách stejná

Charakteristiky variability

- variační rozpětí
- kvantilové odchyly
- průměrné odchyly
- rozptyl
- směrodatná odchylka
- variační koeficient

Variační rozpětí

- rozdíl největší a nejmenší hodnoty sledovaného statist. znaku
- $R = X_{\max} - X_{\min}$
- jednoduchá charakteristika
- podléhá extrémním hodnotám, které mohou být i chybami

příklad

Průměrné odchylky

- průměrná odchylka je definována jako aritmetický průměr odchylek jednotlivých hodnot znaku od vybrané střední hodnoty (tj. od aritmetického průměru, mediánu, modu apod.)

Kvantilové odchytky

- Založeny na kladných odchytkách jednotlivých sousedních kvantilů
- např. kvartilová odchytka
- decilová odchytka
- percentilová odchytka

Střední diference

- je def. jako aritmetický průměr absolutních hodnot všech možných rozdílů jednotlivých hodnot sledovaného znaku
- v praxi vhodná pouze pro malé soubory

Příklad

Rozptyl a směrodatná odchylka

- nejdůležitější charakteristiky variability
- Rozptyl s^2 z n hodnot znaku x je průměr druhých mocnin odchylek jednotlivých hodnot znaku od aritmetického průměru
- směrodatná odchylka s je mírou měnlivosti hodnot souboru kolem aritmetického průměru
- je druhou odmocnina rozptylu

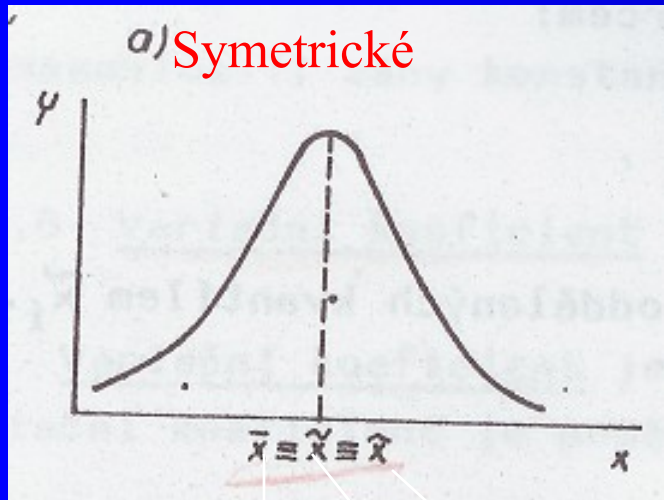
Variační koeficient

- je častou používanou relativní mírou variability
- je definován jako poměr směrodatné odchylky k aritmetickému průměru

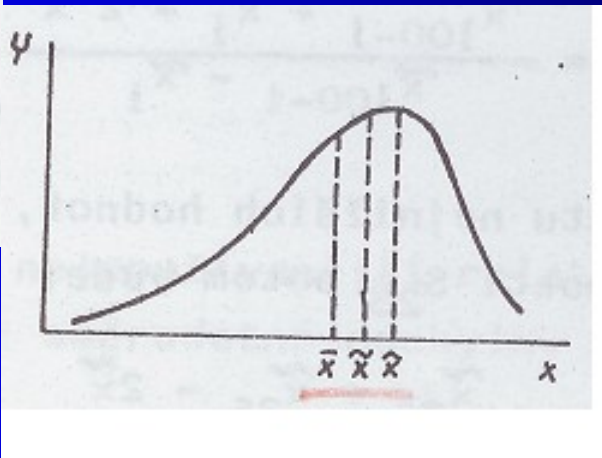
Charakteristiky asymetrie

- Charakteristiky asymetrie (míry šikmosti) jsou čísla dávající představu
 - souměrnosti tvaru rozdělení četností
- míra šikmosti pro souměrné rozdělení je nula
- pro nesouměrné je kladná nebo záporná

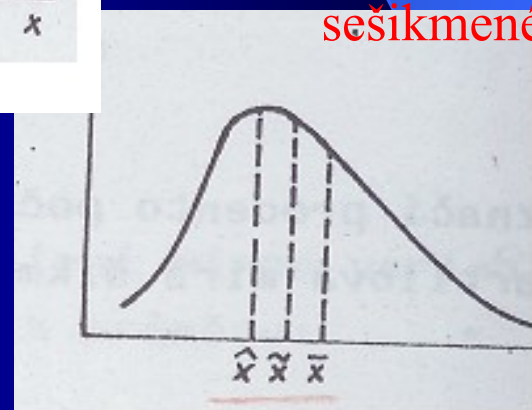
Charakteristiky asymetrie



Záporně sešikmené



Kladně sešikmené



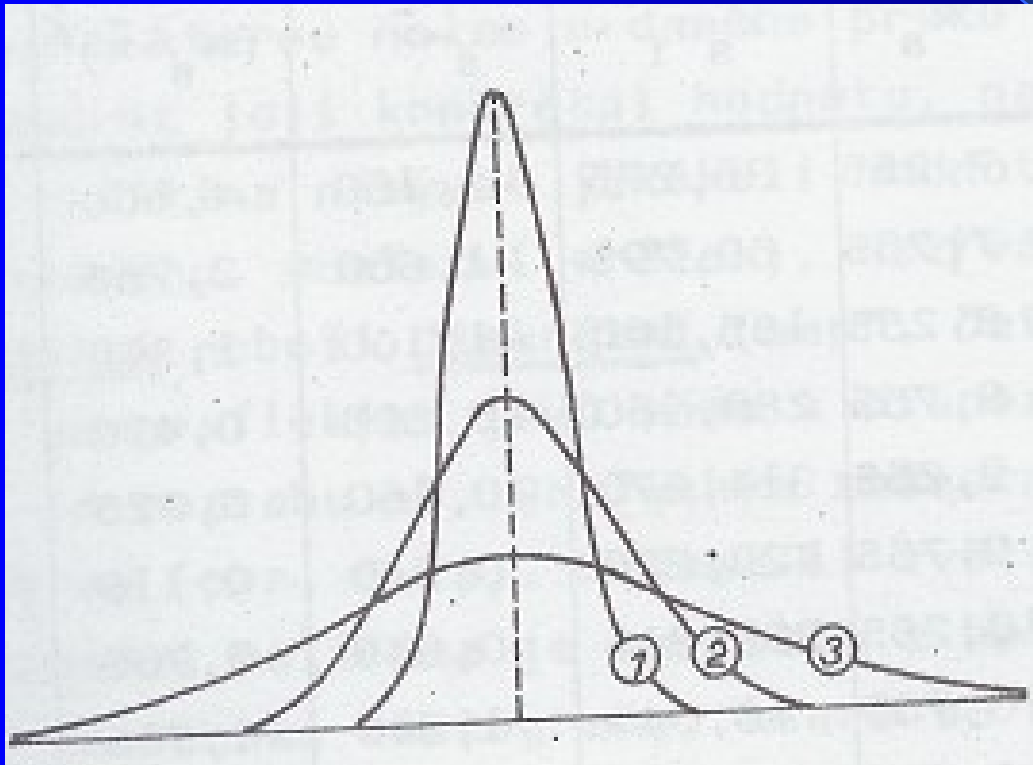
ar. průměr, medián, modus

charakteristiky špičatosti

- Charakteristiky špičatosti (míry špičatosti) jsou čísla charakterizující koncentraci prvků souboru v blízkosti určité hodnoty znaku

Obr. Špičaté, normální a ploché rozdělení

charakteristiky špičatosti



1 – špičaté

2 – normální

3 – ploché

rozdělení

STATISTICKÉ METODY V GEOGRAFII



Karl Friedrich
Gauss
1777-1855

Teoretická rozdělení

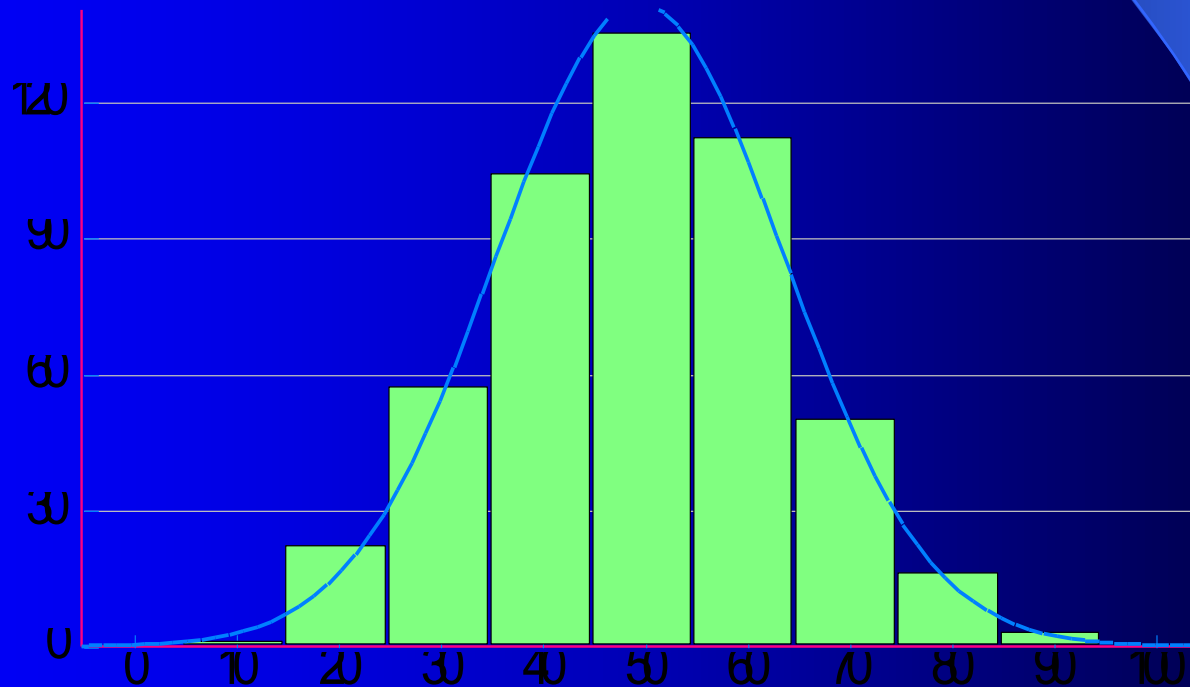
Základní pojmy

- náhodná veličina spojitá
- Může teoreticky nabývat nekonečného množství hodnot z určitého intervalu např. teplota)

- náhodná veličina nespojitá
- Nabývá jen konečného množství hodnot urč. intervalu. Např. počet měsíců s teplotou nad...)
- Každé hodnotě je možno přiřadit pravděpodobnost jejího výskytu, součet všech dílčích pravděpodobností je 1

Teoretická rozdělení

- histogram – grafické znázornění četností
- rozsah souboru se blíží k nekonečnu + náhodná veličina je spojitá
- – frekvenční funkce / hustota pravděpodobnosti

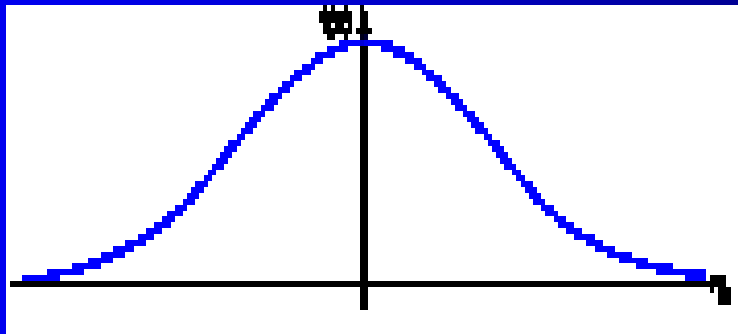


- kumulativní relativní četnost tj. součtová čára -
- distribuční funkce
- obr.

Normální rozdělení / Gaussovo, Laplaceovo- Gaussovo

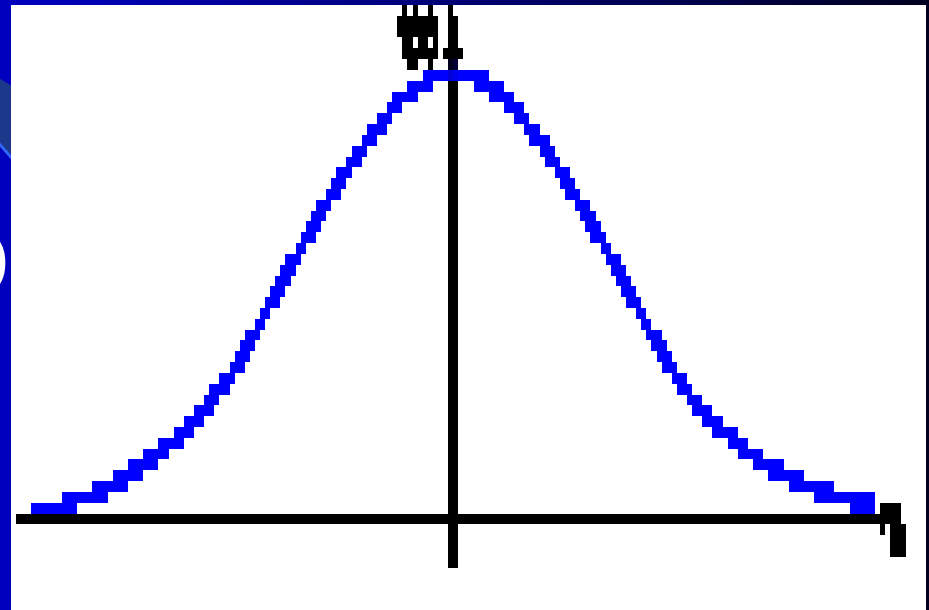
- Normální rozdělení se univerzálně používá k aproximaci (k přibližnému vyjádření) rozdělení pravděpodobnosti velkého množství náhodných veličin (v biologii, technice, ekonomii atd.)

Hustota pravděpodobnosti normálního rozdělení je symetrická zvonovitá **Gaussova křivka**.



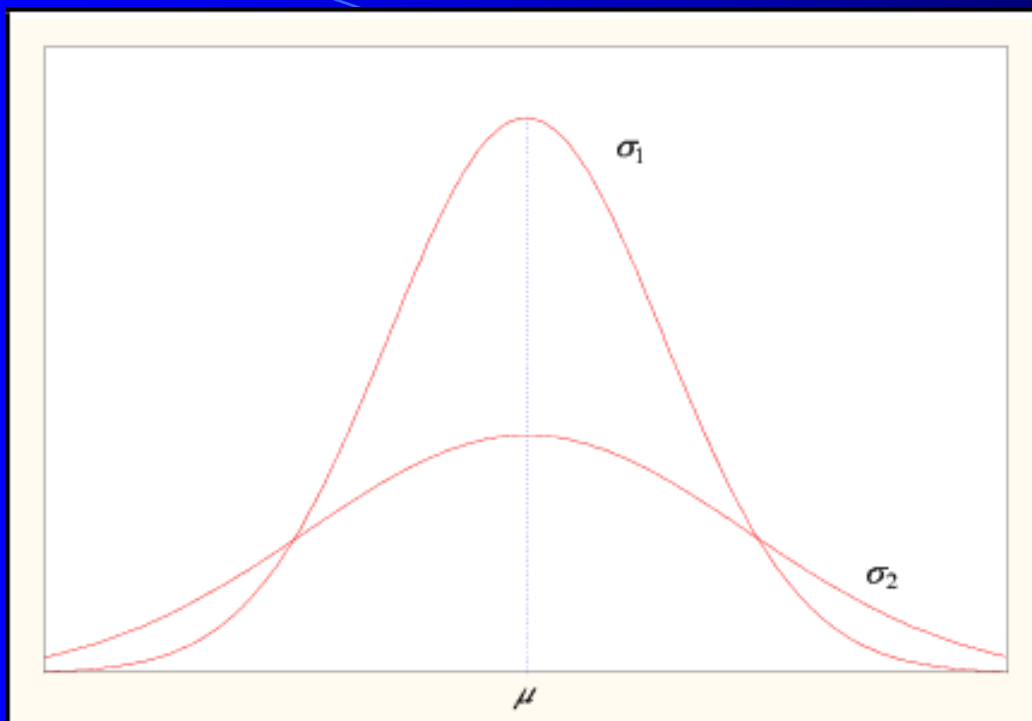
Normální rozdělení

- Zvonovitý tvar
- Souměrný
- Šikmost 0, špičatost 0
- Asymptoticky se blíží 0



- Normální rozdělení s parametry:
- stejný průměr, různé směrodatné odchyly
- čím větší odchylna , tím „plošší tvar rozdělení

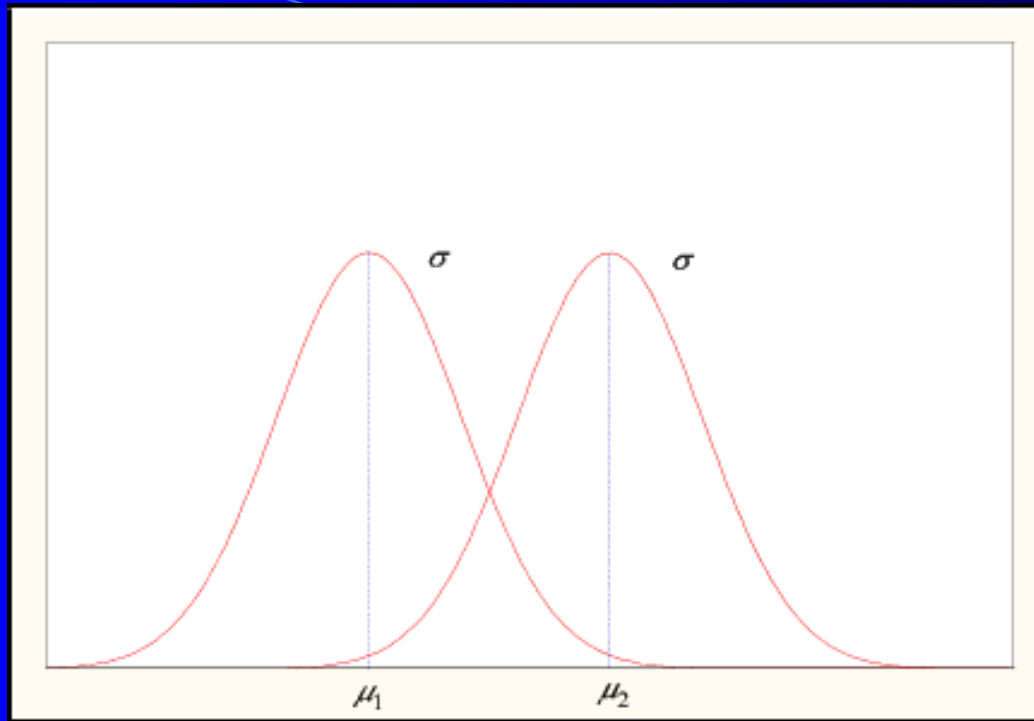
Načrtni obr s oběma křivkami



- Normální rozdělení s parametry:
- stejný průměr, různé směrodatné odchylky
- čím větší odchylka , tím „plošší tvar rozdělení

- Normální rozdělení
- různé průměry, stejná směrodatná odchylka

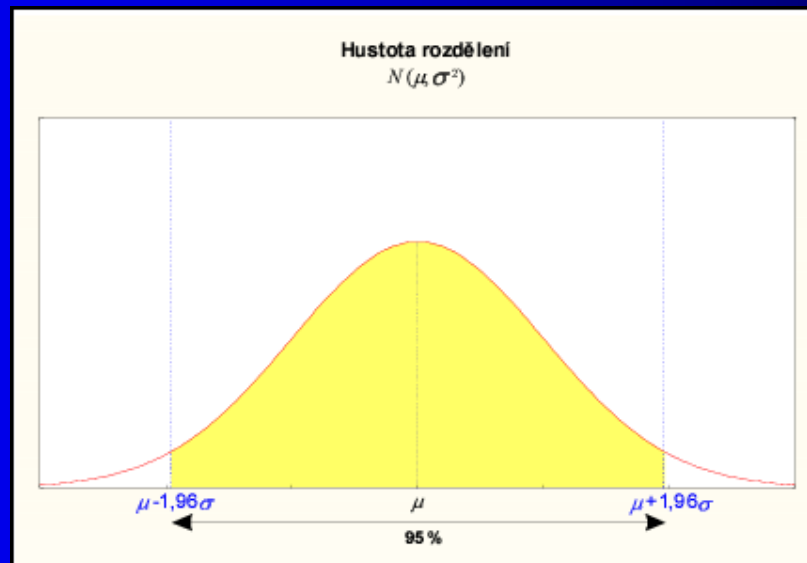
Načrtni obr s oběma křivkami



- Normální rozdělení
- různé průměry, stejná směrodatná odchylka

Normální rozdělení / Gaussovo pokračování

- Normální křivka a osa x vymezují plochu 100%,
- tj. lze stanovit pravděpodobnosti, s nimiž leží hodnoty v určitém intervalu,
- hranice intervalu tvoří průměr a násobky směrodatné odchylky
- obr.



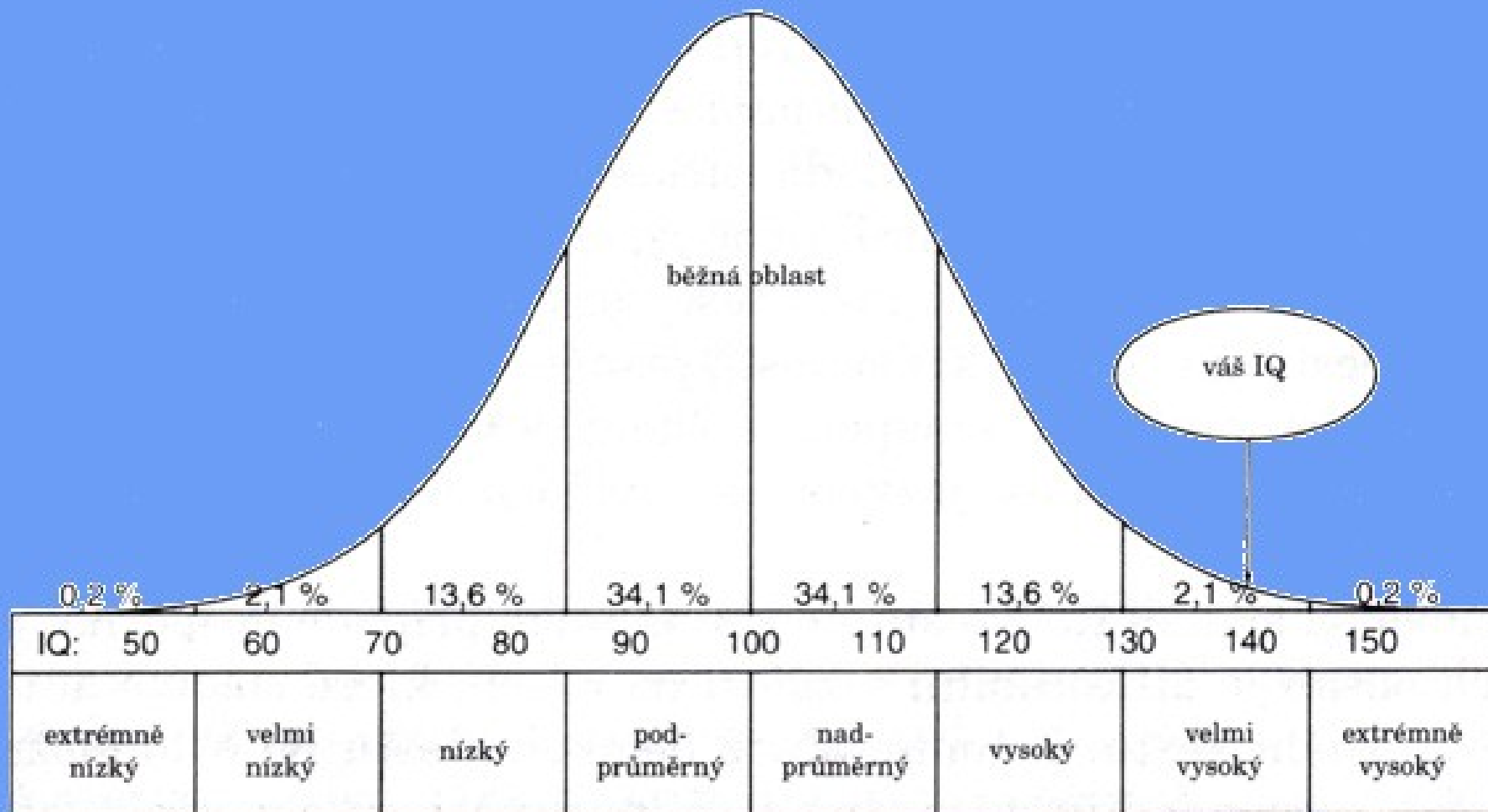
V normálním rozdělení:

- **68, 27% leží v intervalu:**
- **(průměr + - směr. odchylka)**

- **95% leží v intervalu:**
- **(ar. průměr +- 1,96 směr. odchylky)**

- **99% leží v intervalu:**
- **(ar. průměr +- 2,576 směr. odchylky)**

Normální rozdělení pro IQ



imbecilita

debilita

Lehká d.

průměr

vynikající

genialita

idiocie

IQ (v bodech) stupeň inteligence procento zkoumaných
případů (v %)

méně než 20	idiocie	0,1
20 - 49	imbecilita	0,5
50 - 69	debilita	1,9
70 - 79	tzv. lehká debilita	5,0
80 - 89	podprůměrná	14
90 - 109	průměrná	48
110 - 119	nadprůměrná	18
120 - 139	vynikající	11
140 a více	genialita	1,5



Příklady

Př.1

- Populace má v daném testu průměr 100, směrodatnou odchylku 15.
- **Vypočítejte hranice intervalů, v kterém se nachází 68 % populace.**

**68, 27% leží v intervalu:
(průměr + - směr. odchylka)**

Příklad

- Výška v populaci chlapců ve věku 3,5 - 4 roky má normální rozdělení s průměrem 102 cm a směrodatnou odchylkou 4,5 cm.
- Vypočítejte hranice intervalu hodnot výšky, ve kterých se nachází
- A) 68%
- B) 95%
- C) 99%
- příslušné populace

V normálním rozdělení:

- 68, 27% leží v intervalu:
(průměr + - směr. odchylka)
- 95% leží v intervalu:
(ar. průměr +- 1,96 směr. odchylky)
- 99% leží v intervalu:
(ar. průměr +- 2,576 směr. odchylky)

Příklad 3

- zadání:
- Výška v populaci chlapců ve věku 3,5 - 4 roky má normální rozdělení s průměrem 102 cm a směrodatnou odchylkou 4,5 cm.
- Spočtete, jaké procento chlapců v uvedeném věku má výšku menší nebo rovnou 93 cm.

Řešení 3

- Pravděpodobnost, že výška nabude hodnoty menší nebo rovné 93 cm, je vyjádřena hodnotou **distribuční funkce F(93)** pro **parametry normálního rozdělení 102;4,5**

Microsoft Excel

Formula bar: `=NORMDIST(93;102;4,5;pravda)`

NORMDIST

X	93	= 93
Střed_hodn	102	= 102
Sm_odch	4,5	= 4,5
Součet	pravda	= PRAVDA

= 0,022750062

Vrátí hodnotu normálního součtového rozdělení pro zadanou střední hodnotu a směrodatnou odchylku.

Součet je logická hodnota: součtová distribuční funkce = PRAVDA, hromadná pravděpodobnostní funkce = NEPRAVDA.

Výsledek = 0,022750062 = 2,27%

OK Storno

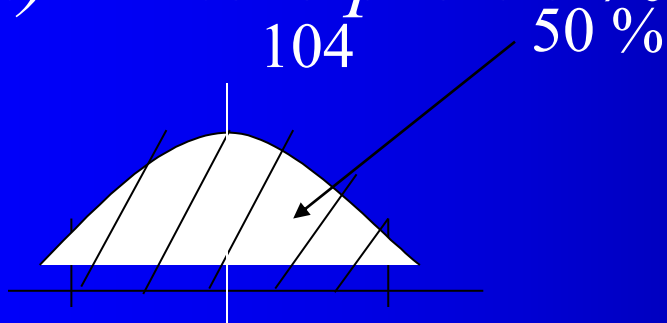
Odpověď: 2,27 % chlapců ve věku 3,5 – 4 roky je menších než 93 cm

Příklad 4

- Psychologickými testy bylo zjištěno, že hodnota IQ populace je náhodnou veličinou s normálním rozdělením, jehož střední hodnota je 104 a směrodatná odchylka 8.
- Určete hodnotu IQ, kterou podle uvedených pravděpodobnostních předpokladů:
 - meze, ve kterých bude 50% populace,
 -

Řešení 4

- a) *meze pro 50 % mužské populace*



Hledáme dolní a horní meze intervalu (hodnot IQ),
ve které se bude nacházet 50% mužské populace, tj 1. a 3. kvartil

Řešení 2a)

Excel, statistická funkce inverzní k e Gauss. - NORMINV

Microsoft Excel

Formule: **=NORMINV(0,25;104;8)**

NORMINV

Prst	0,25	= 0,25
Střední	104	= 104
Sm_odch	8	= 8

= 98,60407707

Vrátí inverzní funkci k distribuční funkci normálního součtového rozdělení pro zadanou střední hodnotu a směrodatnou odchylku.
Prst je pravděpodobnost normálního rozdělení, číslo mezi 0 a 1 včetně.

Výsledek = 98,60407707

OK Storno

Microsoft Excel

Formule: **=NORMINV(0,75;104;8)**

NORMINV

Prst	0,75	= 0,75
Střední	104	= 104
Sm_odch	8	= 8

= 109,3959229

Vrátí inverzní funkci k distribuční funkci normálního součtového rozdělení pro zadanou střední hodnotu a směrodatnou odchylku.
Prst je pravděpodobnost normálního rozdělení, číslo mezi 0 a 1 včetně.

Výsledek = 109,3959229

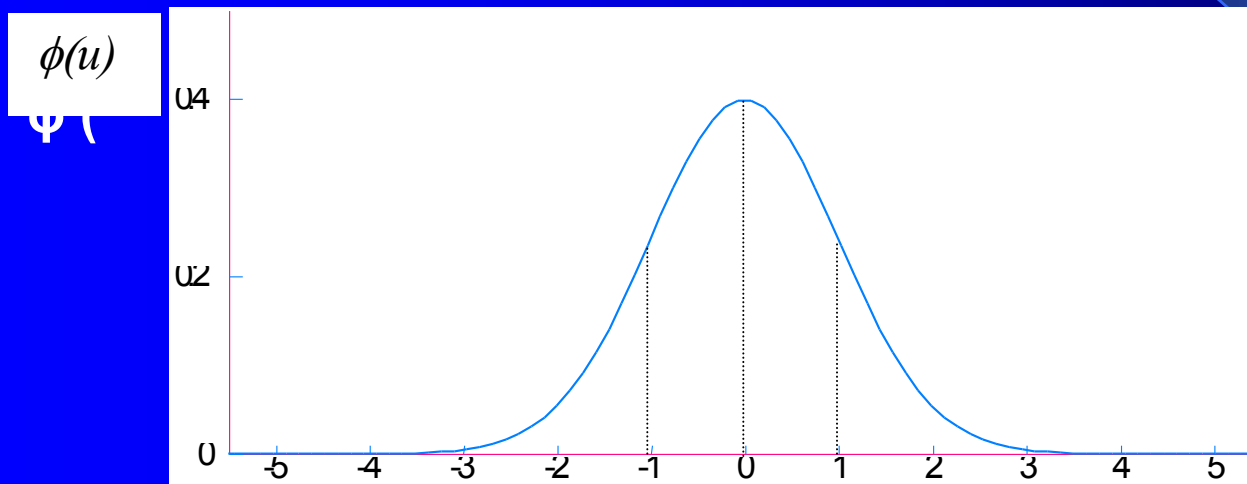
OK Storno

Podle parametrů daného normálního rozdělení 50 populace má IQ v intervalu 98,6 a 109,4.

- Pro normované normální rozdělení zavedeme označení $N(0, 1)$.

Normování hodnoty: od hodnoty se odečte aritmetický průměr, výsledek (tj. odchylka) se dělí směr. odchylkou

Hustota pravděpodobnosti normovaného normálního rozdělení:



Tabulkové vyjádření vybraných hodnot hustoty pravděpodobnosti

u	0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5
$\phi(u)$	0,399	0,352	0,242	0,130	0,054	0,018	0,004	0,001

Tabulkové vyjádření vybraných hodnot distribuční funkce

u	0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5
$\Phi(u)$	0,500	0,691	0,841	0,933	0,977	0,994	0,999	0,999

Domů Vložení Návrh Animace Prezentace Revize Zobrazení Doplnky

Sešit1 - Microsoft Excel

Vložení Rozložení stránky Vzorce Data Revize Zobrazení Doplnky

Písmo Zarovnaní Číslo Podmíněně formátování Formátovat jako tabulku Styly Buňky

=NORMDIST(102;102;0,5;nepravda)

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	průměr	102		0,02275	distribuční pro 93											
	směr odch	4,5		0,5	distribuční pro 102											
				1	pro 120											
				98,9648	první kvartil											
				105,0352	třetí kvartil											
				102	medián											

Argumenty funkce

NORMDIST

X 102 = 102

Střed_hodn 102 = 102

Sm_odch 0,5 = 0,5

Součet nepravda = NEPRAVDA

= 0,797884561

Vrátí hodnotu normálního součtového rozdělení pro zadanou střední hodnotu a směrodatnou odchylku.

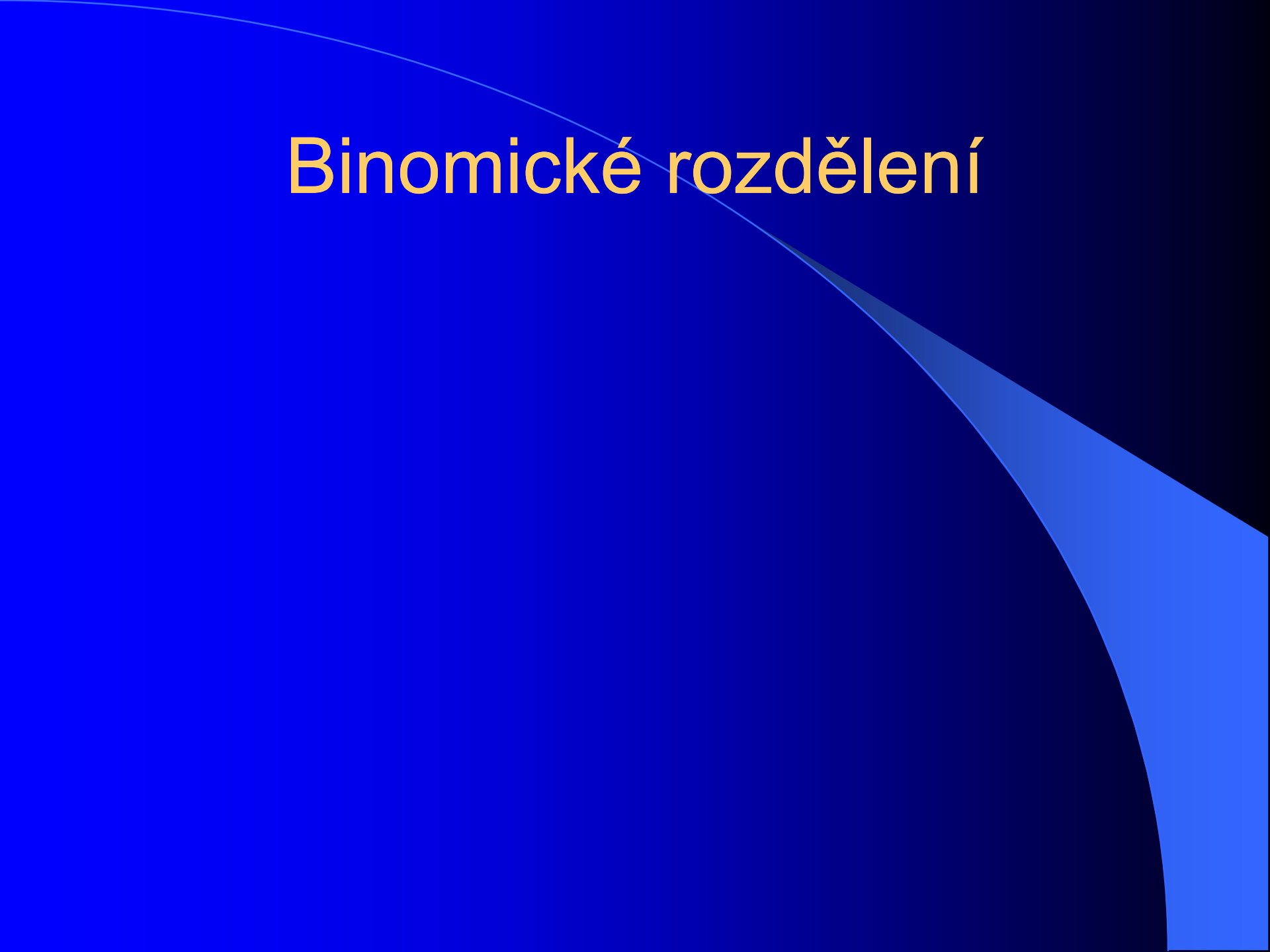
Sm_odch je směrodatná odchylka rozdělení, kladné číslo.

Výsledek = 0,797884561

[Nápověda k této funkci](#)

OK Storno

Binomické rozdělení



Binomické rozdělení

- pro diskrétní náhodné proměnné,
- které mohou nabývat pouze dvou hodnot (např. ano, ne)
- pravděpodobnost, že nastane alternativa ANO označme π
- pravděpodobnost, že nastane NE ... $q = 1 - \pi$), protože
- platí $\pi + q = 1$ (100 %)
- k výpočtu se používá binomický rozvoj

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$P(X = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad \text{pro } k = 0, 1, 2, \dots, n.$$

Příklad 1 – binomické rozdělení

- Předpokládejme, že pravděpodobnost narození dívky je 0,49.
- Jaká je pravděpodobnost toho, že mezi třemi dětmi v rodině je právě jedna dívka?

Řešení 1

Tabulka3: Parametry binomického rozdělení v příkladu

Pokus	Úspěch	Neúspěch	Pravděpodobnost úspěchu	Počet pokusů	Počet úspěchů
				n	k
narození dítěte	dívka	chlapec	0,49	počet dětí	počet dívek

Řešení 1

Jak je vidět z tabulky, počet narozených dívek v rodině je náhodná veličina s binomickým rozdělením. Pravděpodobnost, že mezi třemi dětmi je právě jedna dívka tedy vypočteme jako

$$P(X = 1) = \binom{3}{1} 0,49^1 \cdot 0,51^2 = 3 \cdot 0,127 = 0,38. \diamond$$

$$\binom{3}{1} = \frac{3!}{1!2!} = \frac{3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 1} = 3.$$

Pravděpodobnost, že ze tří dětí bude jedna dívka, je 38%.

Microsoft Excel

Formula bar: `=BINOMDIST(1;3;0,49;NEPRAVDA)`

BINOMDIST

Úspěch	1	= 1
Pokusy	3	= 3
Prst_úspěchu	0,49	= 0,49
Počet	NEPRAVDA	= NEPRAVDA

Result: = 0,382347

Vrátí hodnotu binomického rozdělení pravděpodobnosti jednotlivých veličin.

Počet je logická hodnota: součtová distribuční funkce = PRAVDA, hromadná pravděpodobnostní funkce = NEPRAVDA.

Výsledek = 0,382347

Buttons: OK, Storno

Příklad 2

Jaká je pravděpodobnost, že v rodině s 8 dětmi jsou právě 3 dívky? Pravděpodobnost narození dívky je 0,49.

Řešení

binomický rozvoj:

$$P(k = 3) = \binom{8}{3} 0,49^3 \cdot 0,51^5 = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} \cdot 0,118 \cdot 0,035 = 0,23. \diamond$$

Pravděpodobnost, že v rodině s 8 dětmi jsou tři dívky, je 0,23, tj. 23 %.

Příklad 2, binomické rozdělení

- Vypočítejte pravděpodobnost, se kterou se vyskytne určitý počet měsíců v roce hodnocených jako „suché“.
- Konkretizace:
 - oblast Oxford,
 - období 1851 – 1943, tj. 1116 měsíců
 - **Suchý měsíc** - tj. méně srážek v měsíci než je dlouhodobý průměr tohoto měsíce.
 - 617 měsíců hodnocených jako suché
 - 499 – vlhké měsíce

Řešení 2

„úspěch“	„neúspěch“	Pravděpodobnost suchého měsíce	Pravděpodobnost vlhkého měsíce	Počet měsíců	Počet suchých měsíců
suchý	vlhký	$\pi = 617/1116$ $\pi = 0,553$	$q = 499/1116$ $q = 0,447$ $(q = 1 - \pi)$	$n = 12$	$k = 0$ až 12

Řešení

- Ručně pomocí binomického rozvoje
- s podporou např. Excel

Řešíme dílčí příklady, tj. jaká je pravděpodobnost, že v roce se vyskytne

- žádný suchý měsíc, tj- $k = 0$
- Jeden suchý měsíc, tj. $k = 1$
- Atd.
- všechny měsíce suché, $k = 12$

Řešení 2

k	f(x)
0	0,000
1	0,000945
2	0,006428
3	0,026507
4	0,073785
5	0,146051
6	0,21
7	0,223
8	0,172
9	0,095
10	0,035
11	0,0079
12	0,0008

Microsoft Excel

Binomická funkce: **BINOMDIST**

Úspěch 5 = 5

Pokusy 12 = 12

Prst_úspěchu 0,553 = 0,553

Počet nepravda = NEPRAVDA

Formule: =BINOMDIST(5;12;0,553;nepravda)

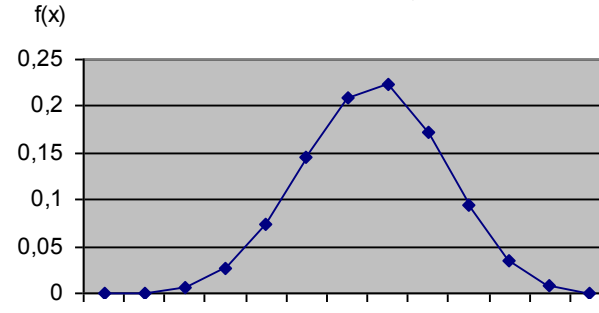
Výsledek: = 0,146050652

Vrátí hodnotu binomického rozdělení pravděpodobnosti jednotlivých veličin.

Úspěch je počet úspěšných pokusů.

OK Storno

Pravděpodobnost počtu suchých měsíců v roce,
Oxford, 1851 - 1943



—●— pravděpodobnost

počet měsíců

Screenshot of a Windows XP desktop environment showing Microsoft Excel and a data selection dialog box.

Excel Worksheet Data:

k	f(x)
0	0,000
1	0,000945
2	0,006428
3	0,026507
4	0,073785
5	0,146051
6	0,21
7	0,223
8	0,172
9	0,095
10	0,035
11	0,0079
12	0,0008

Excel Chart: A line chart titled 'f(x)' showing the probability distribution. The x-axis ranges from 0 to 12, and the y-axis ranges from 0 to 0.25. The chart shows a smooth curve peaking at x=7.

Dialog Box: Vybrat zdroj dat (Select Data Source)

Oblast dat grafu: =List1!\$F\$1:\$G\$14

Popis legendy (řady): f(x)

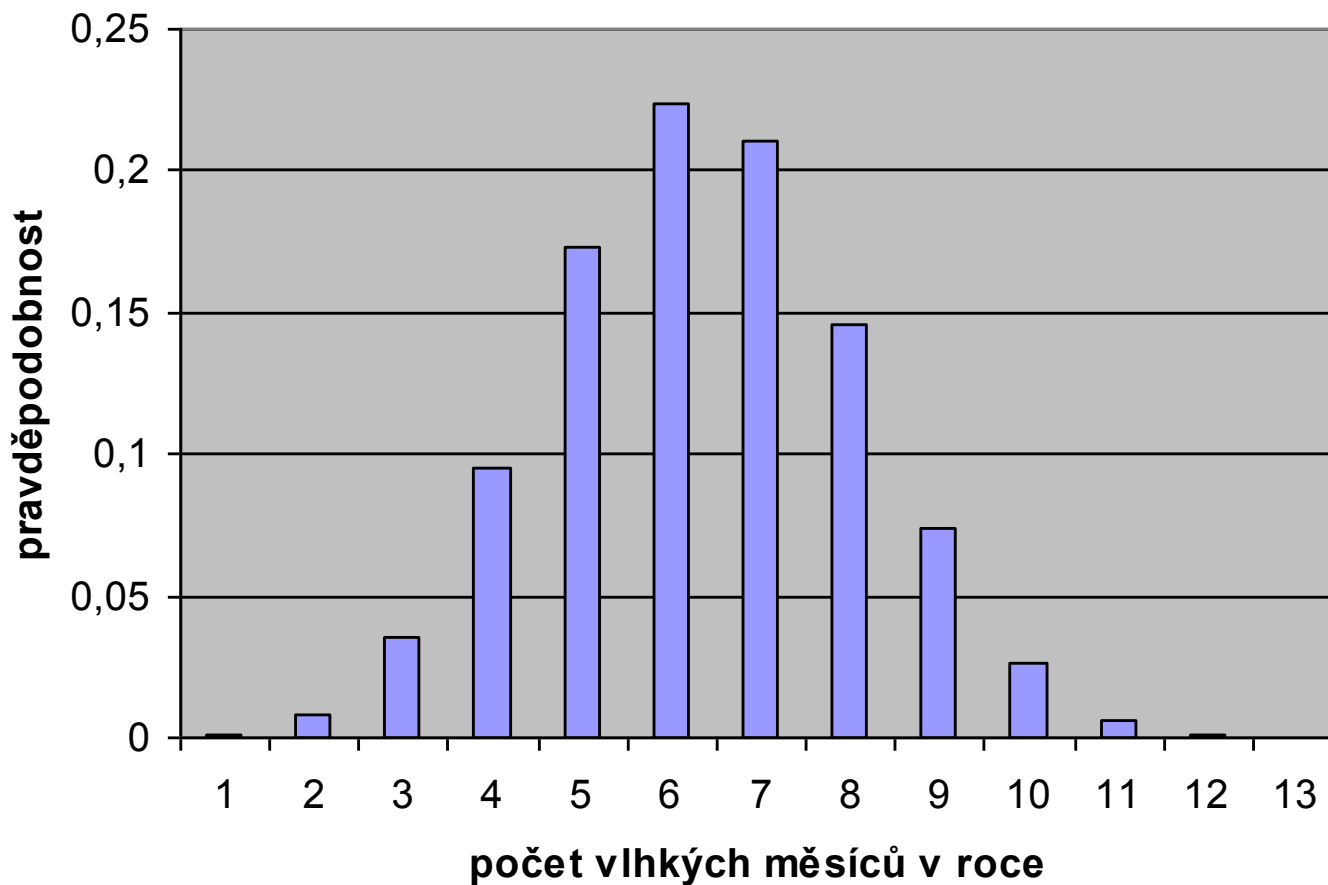
Popisky vodorovné osy (kategorie): 0, 1, 2, 3, 4

Buttons: Přepnout řádek/sloupec, Přidat, Upravit, Odebrat, Upravit, Sčítané a prázdné buňky, OK, Storno

Jak bude vypadat situace pro „vlhke“ měsíce?

Binomické rozdělení

Pravděpodobnost výskytu vlhkého měsíce
v oblasti Oxfordu v letech 1851 - 1943



Poisson - příklad

Poissonovo rozdělení

- – pro rozdělení vzácných případů
- (zimní bouřka, výskyt mutace apod.).

- Je-li pravděpodobnost nějaké výjimečné události (např. určité mutace genu) relativně malá a rozsah výběru poměrně velký, pak **Poissonovo rozdělení v podstatě splývá s binomickým**, ale je mnohem výhodnější pro počítání .

Poisson - příklad

- Předpokládejme, že v určité populaci krys se vyskytuje albín s pravděpodobností
- $p = 0,001$, ostatní krys jsou normálně pigmentované.
- Ve vzorku 100 krys náhodně vybraných z této populace určete pravděpodobnost, že vzorek
 - a) neobsahuje albína,
 - b) obsahuje právě jednoho albína.

Řešení

- určete pravděpodobnost, že vzorek
- neobsahuje albína,

Microsoft Excel

Formula bar: `=BINOMDIST(0;100;0,001;NEPRAVDA)`

BINOMDIST

Úspěch	0	= 0
Pokusy	100	= 100
Prst_úspěchu	0,001	= 0,001
Počet	NEPRAVDA	= NEPRAVDA

= 0,904792147

Vrátí hodnotu binomického rozdělení pravděpodobnosti jednotlivých veličin.

Úspěch je počet úspěšných pokusů.

Výsledek = 0,904792147

OK Storno

Pravděpodobnost, že neobsahuje albína, je 90,47 %

Řešení 3

Microsoft Excel

Binomická funkce: **BINOMDIST**

Úspěch 1 = 1

Pokusy 100 = 100

Prst_úspěchu 0,001 = 0,001

Počet NEPRAVDA = NEPRAVDA

Výsledek = 0,090569784

Vrátí hodnotu binomického rozdělení pravděpodobnosti jednotlivých veličin.

Počet je logická hodnota: součtová distribuční funkce = PRAVDA, hromadná pravděpodobnostní funkce = NEPRAVDA.

OK Storno

Pravděpodobnost, že 100 členná populace krys bude obsahovat albína, je 9 %.

Další rozdělení

Pearsonova křivka III. typu

- Na empirické rozdělení mnoha statistických souborů s nimiž v geografii pracujeme, **nelze aplikovat normální rozdělení.**
- Platí to například v těch případech, kdy studovaná náhodná veličina **nemá teoreticky zdůvodněnou možnost nabývat nekonečných hodnot** nebo je-li omezena konečnými čísly
V takovýchto případech lze aplikovat na studovaný soubor některou ze dvanácti křivek Pearsonova systému.

Pearsonova křivka III. typu

- Pearsonova křivka III. typu
- - obvykle pro veličiny s omezeným množstvím hodnot, které může nabývat
- - z křivky lze např. vyčíst pravděpodobnost se kterou bude hodnota sledovaného statistického znaku dosažena
- v hydrologii se počítá Pearsonova křivka ve variantě součtová čára četností jako
- tzv. čára překročení

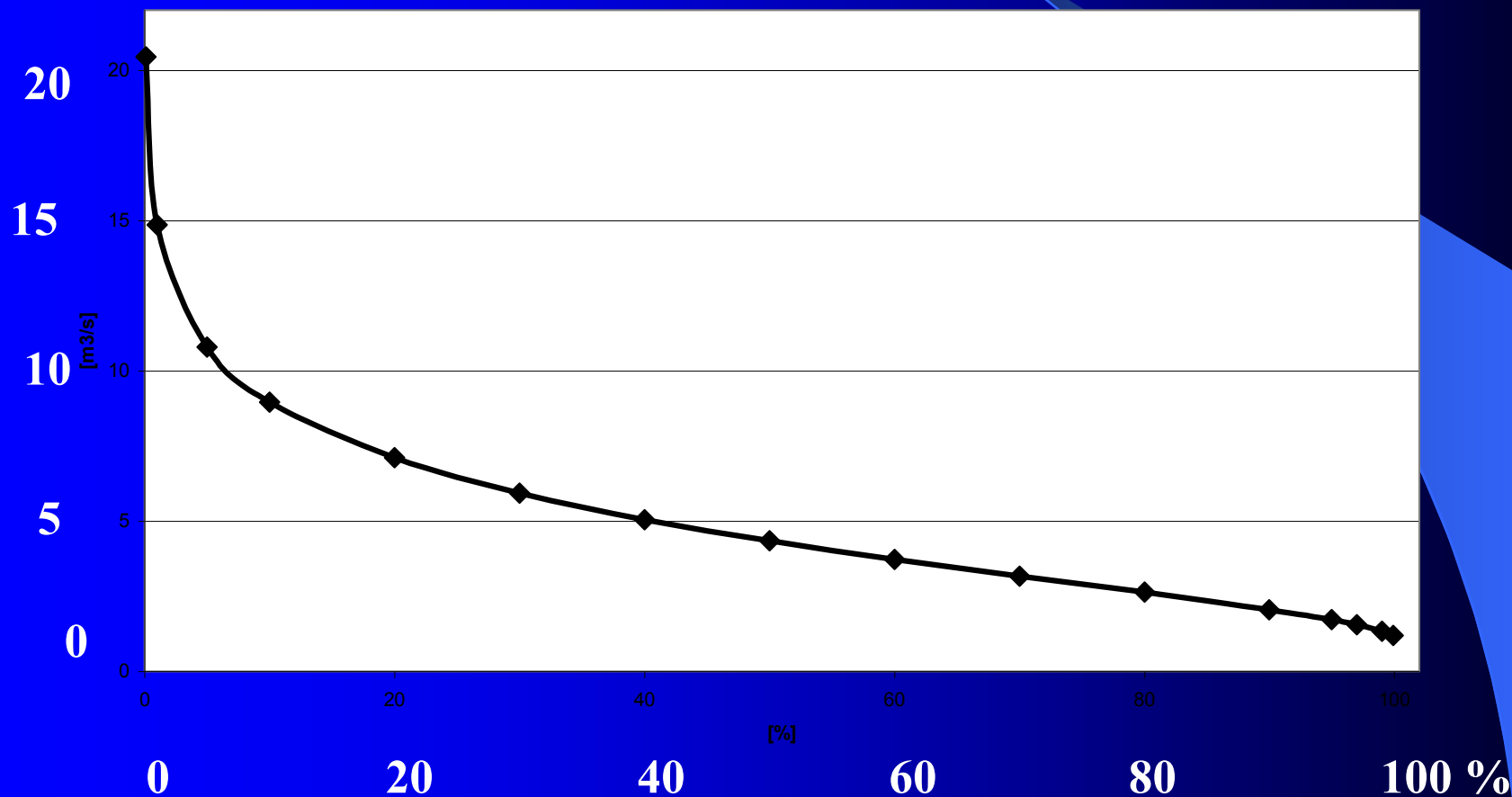
- **příklad**
- Konstrukce čáry překročení z průměrných ročních průtoků vodního toku Lažánka za říjen 2002.

den	průtok Qd (m ³ /s)	den	průtok Qd (m ³ /s)
1	2,99	16	2,98
2	2,84	17	4,64
3	2,75	18	12,2
4	3,22	19	7,73
5	3,55	20	4,38
6	12,2	21	3,41
7	9,12	22	3,85
8	3,82	23	3,47
9	3,55	24	3,36
10	3,23	25	3,51
11	2,89	26	12,2
12	3,25	27	10,3
13	3,79	28	6,2
14	3,05	29	4,15
15	3,05	30	5,75
		31	5,1

Křivka překročení průměrných ročních průtoků , Lažanka, říjen 2002

Křivka překročení průměrných ročních průtoků vodního toku Lažanka za říjen 2002

m³/s



Odhady parametrů intervaly spolehlivosti

Základní pojmy

- základní soubor,
- statistický soubor
- výběrový soubor
- náhodný výběr
- k základnímu jednomu souboru lze získat více výběrových, různé charakteristiky

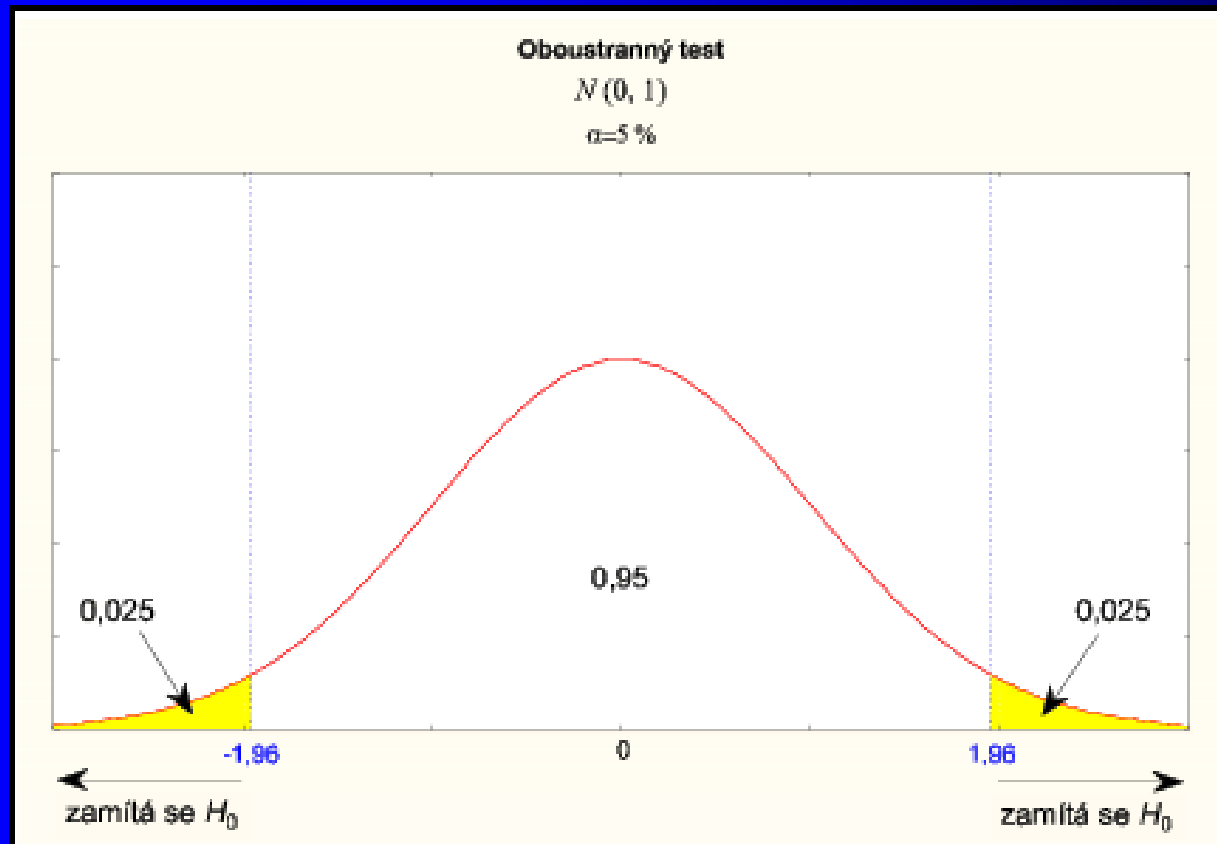
Základní pojmy

- **reprezentativnost výběru** – kvalita výběru
- **prostý náhodný výběr** (s opakováním a bez opakování)
- **oblastní náhodný výběr** (výběr z každé dílčí části)
- **systematický náhodný výběr** (podle pravidla, které nesouvisí se sledovaným znakem, např. sledovaný znak - počet obyvatel obce, seřadit obce podle abecedy a vybrat vždy každou pátou obec)

Intervaly spolehlivosti

- normální rozdělení,
- Statistický soubor s norm rozdělením (X, s)
- Jeho výběrový soubor bude mít norm rozdělení s param $(x, s/\sqrt{n})$,
- Interval spolehlivosti – pro zvolený koeficient spolehlivosti (pravděpodobnost, že tam X padne) (např. 95 %)
- vypočítáme interval, ve kterém s touto pravděpodobností leží X .

- **Obrázek: Oboustranný test H_0**



- provedeme-li výběr o rozsahu n a spočteme \bar{x} , pak průměr X leží s pravděpodobností 0,95 ve vzdálenosti menší než $1,96 s \sqrt{n}$ od \bar{x} ,
- tj. v intervalu s krajními body
- $(\bar{x} - 1,96 s \sqrt{n}, \bar{x} + 1,96 s \sqrt{n})$... *interval spolehlivosti pro průměr.*
- *koeficient spolehlivosti* $P = 0,95$
- (tj. *hladinu významnosti* $\alpha = 0,05$)

- lze použít intervaly spolehlivosti např.
- pro 95 % ($\mu \pm 1,960\sigma$),
- pro 99 % ($\mu \pm 2,576\sigma$), tj. širší! interval

- hodnoty, které leží mimo interval, v tzv. kritickém oboru se považují za nepřijatelné, jejich odchylky od průměru za významné

Závislost náhodných veličin

Závislost náhodných veličin

- Do jaké míry závisí změna prvku jednoho statistického souboru změnu prvku druhého statistického souboru?
- Jak podmiňuje změna prvku x změnu prvku y ?
- Jak těsně na sobě závisí prvky dvourozměrného statistického souboru?
- Např.
 - vztahy teplota a nadm. výška,
 - srážky a odtok v povodí
 - váha a výška člověka,

Vztahy náhodných veličin

- Jednostranné (nezávislá hodnota \underline{x} jednoho stat. souboru podmiňuje hodnotu y druhého stat. souboru)
- Vzájemné (nelze rozlišit závislou a nezávislou proměnou)

Vztahy náhodných veličin

- Podle stupně závislosti
- Funkční (pevnou)
- (určité hodnotě x odpovídá jediná hodnota y , vztah x a y lze tedy vyjádřit mat. funkcí),
- *např.*
- *Konkrétní teplotě odpovídá jedna hodnota stupně nasycení vodní párou*

Vztahy náhodných veličin

- Statistická
- (jediné hodnotě x odpovídá více hodnot y , hodnoty y mají své rozdělení s průměrem, tento průměr hodnot y je i pro různá x shodný)



Vztahy náhodných veličin

- **Korelační**
- Se změnou hodnot x se mění soubory hodnot y , které mají své rozdělení a různých průměrech
- *např. pro určitou těl výšku existuje více hodnot hmotnosti, které budou mít normální rozdělení,*
- *různým výškám odpovídají hmotnosti s normálním rozdělením, ale s různým průměrem*
- Př. Pro 170 cm existuje norm. rozdělení hmotností o průměru 68 kg, pro 180 cm opět normální rozdělení hmotností s průměrem 76 kg

Korelační závislost

- Určení těsnosti korelační závislosti
- (jak těsný je vztah mezi výškou a hmotností člověka)
- **Korelace** je druh závislosti mezi prvky dvou souborů
- **Regresní čára** znázorňuje graficky tuto korelační závislost

Př. lineární regrese

- Vypočítejte koeficient korelace pro vztah délky slunečního svitu a teploty na datech meteorol. stanice Tuřany, 2002

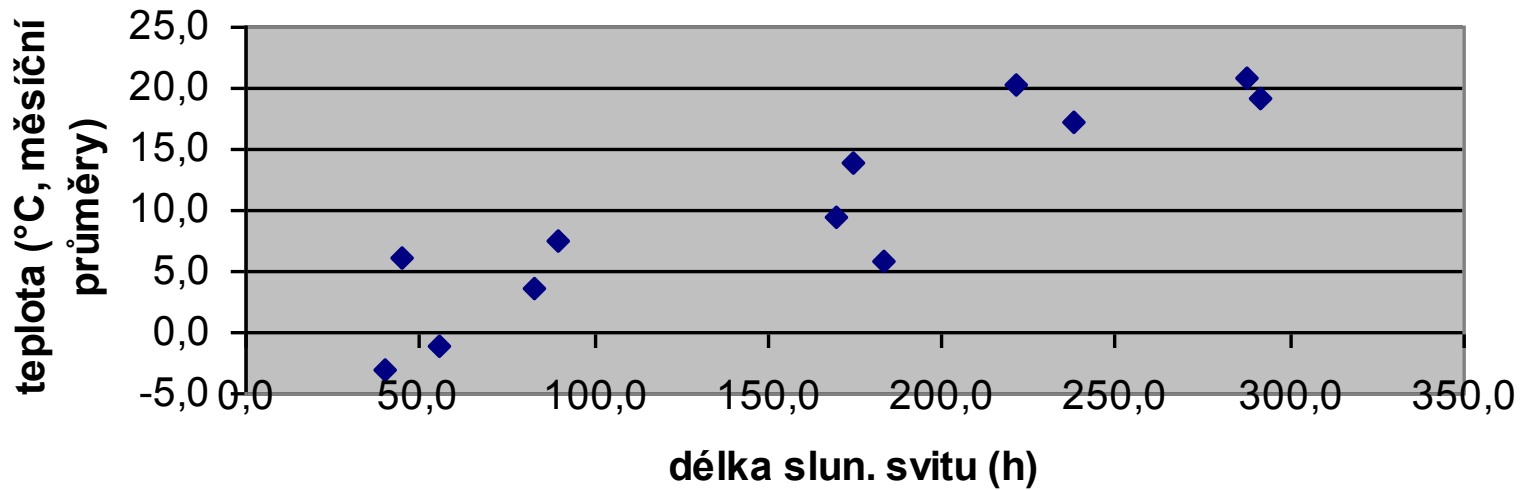
Délka slun. svitu (h)	55,6	82,7	183,4	169,5	238,3	291,4	288,0	221,2	174,5	89,4	44,7	40,3
Teplota (°C)	-1,2	3,6	5,8	9,4	17,1	19,1	20,9	20,4	14,0	7,6	6,0	-3,1

t	-1,2	3,6	5,8	9,4	17,1
h	55,6	82,7	183,4	169,5	238,3

corel h, t 0,920888

corel t, h 0,920888

Závislost teploty na délce slunečního svitu, Brno, 2002



Výpočet koeficientu regrese b :

Excel, funkce CORREL, POLE1 - hodnoty délka slun. Svitů,
Pole2 - hodnoty teploty

The screenshot shows the Microsoft Excel interface. A dialog box for the CORREL function is open, displaying the formula `=CORREL(C17:N17;C18:N18)` and the calculated result `= 0,903991059`. The dialog also shows the data ranges for Pole1 (C17:N17) and Pole2 (C18:N18).

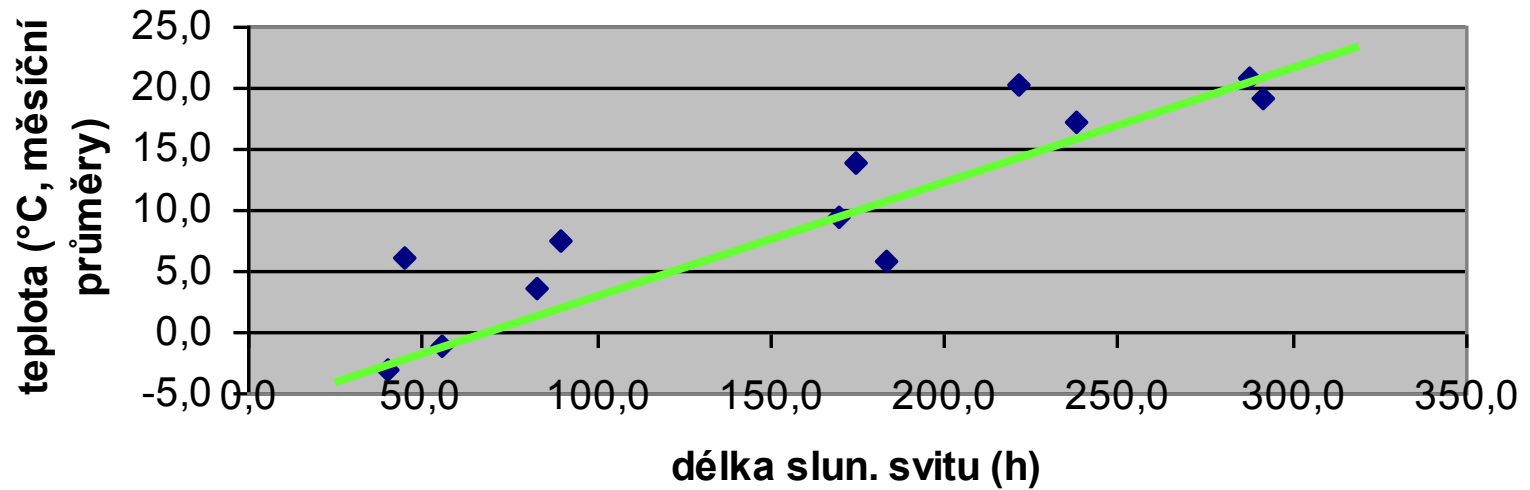
The spreadsheet data is as follows:

		I	J	K	L	M	N	O					
10													
11													
12													
13	teplota	-1,2	3,6	5,8	9,4	17,1	19,1	20,9	20,4	14,0	7,6	6,0	-3,1
14	úhrn sráže	8,1	21,3	21,0	28,6	45,8	81,7	58,0	91,2	39,2	71,9	48,2	46,0
15													
16	délka slun.	55,6	82,7	183,4	169,5	238,3	291,4	288,0	221,2	174,5	89,4	44,7	40,3
17	teplota	-1,2	3,6	5,8	9,4	17,1	19,1	20,9	20,4	14,0	7,6	6,0	-3,1
18													
19	korelace t/s		0,656547										
20													
21	korelace t/d		0,903991										
22													
23	korelace s/d		0,461355										
24													
25													
26	regresní přímka												
27													
28													
29													
30													

The scatter plot, titled "Závislost teploty na délce slunečního svitu, Brno, 2002", shows the relationship between monthly solar radiation (x-axis) and monthly average temperature (y-axis). The data points are as follows:

Month	Solar Radiation (délka slun.)	Temperature (teplota)
I	55,6	-1,2
II	82,7	3,6
III	183,4	5,8
IV	169,5	9,4
V	238,3	17,1
VI	291,4	19,1
VII	288,0	20,9
VIII	221,2	20,4
IX	174,5	14,0
X	89,4	7,6
XI	44,7	6,0
XII	40,3	-3,1

Závislost teploty na délce slunečního svitu, Brno, 2002



lineární regresní čára - Přidat spojnici trendu

Časové řady

Bazické a řetězové

Z - diagram

časová řady – základní pojmy

- **statistická řada**
- posloupnost hodnot znaku uspořádaných podle určitého hlediska
- časová řada
- statistická řada upořádaná podle času
- časová řada=dynamická=chronologická = vývojová

Sestavování časových řad

Cíl – získat porovnatelná čísla

- dodržovat zásady:
 - stejně dlouhá časová období
 - (přepočít na „standardizovaný“ měsíc se 30 dny, přepočít na počet shodný počet pracovních dní v měsíci p
 - stejně velká území, příp. stejná úroveň (shodná rozloha, povodí řádu toku, administrativní jednotka)
 - stejně jednotky

- časová řada OKAMŽIKOVÁ

- sleduje se hodnoty znaku k určitému okamžiku
- např. počet obyvatel ČR k 31.12. 2000, 2001,

- časová řada INTERVALOVÁ

- sleduje se hodnota znaku v intervalu , období
- denní úhrn srážek, průměrná denní teplota, měsíční těžba...
 - pouze k této řadě se vztahuje požadavek stejného intervalu zvláště u sledování ekonomických ukazatelů

Analýza časových řad

- cíle analýzy:
 - zjistit hlavní rysy průběhu časových řad a analyzovat je
- podle průběhu časové řady:
- stacionární nebo s trendem
- s periodickým opakováním výkyvů nebo bez výkyvů
- všechny možné kombinace

Charakteristiky časových řad

přírůstky a indexy

- přírůstky:
- **absolutní přírůstek** – rozdíl hodnot po sobě následujících („druhá“ – „první“)
- $X_i - X_{i-1}$
- **relativní přírůstek**
- podíl $x_i - x_{i-1} / x_{i-1}$

Řetězové a bazické indexy

- **bazický index**
- podíl $x_i / x_z * 100$,
- x_z - první „ základní „ hodnota časové řady
- změny k jedné základní (bazické) hodnotě

- **řetězový index** (koeficient růstu)
- podíl $x_i / x_{i-1} * 100$
- podíl v procentech po sobě následujících hodnot
- (změny např. z měsíce na měsíc“ – řetězení)

Klouzavé úhrny

- zvláštní typ součtové čáry
- vhodné pro porovnávání dvou či více řad hodnot za po sobě následující období
- např. kolísání ročního chodu srážek
- postup viz. např. skripta Brázdil. a kol. str. 147

měsíc	1	2	3	4	5	6	7	8	9	10	11	12
prům úhrn srážek;2002; mm	8,1	21,3	21	29	45,8	81,7	58	91,2	39,2	71,9	48,2	46
prům úhrn srážek;2003, mm	26,6	4,3	4,1	22	92,8	59,8	66,1	37	24,3	58,5	32,4	54, 3

KLOUZAVÝ ÚHRN	482, 6	454, 9	48 6	52 1	58 6	56 5	57 3	51 8	50 4	49 0	474, 3	48 3
--------------------------	-----------	-----------	---------	---------	---------	---------	---------	---------	---------	---------	-----------	---------

Klouzavý úhrn, vždy součet 12 měsíčních hodnot, tj. daný měsíc plus +11 předchozích

LEDNOVÁ HODNOTA – SOUČET „NOVÝ“ LEDEN + 11 předchozích měsíců

ÚNOROVÁ HODNOTA – SOUČET „NOVÝ“ LEDEN + ÚNOR
+STARÉ OSTATNÍ MĚSÍCE

Z - diagramy

- GRAFICKÉ ZNÁZORNĚNÍ
 - řada běžných hodnot,
 - součtová čára,
 - řada klouzavých úhrnů

- společné body Z - diagramu(tj. spol. hodnoty)
 - výchozí bod součtové č. a řady běžných hodnot
 - poslední hodnota součtové čáry a poslední hodnota klouzavého úhrnu

Cvičení 12_07_casove_rady_indexy.doc - Microsoft Word

Microsoft Excel

Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda

D15

z diagram, klouz.uhrn.xls

	A	B	C	D	E
1					
2			měsíc	1	2
3			prům úhrn srážek;2002; mm	8,1	21,3
4			prům úhrn srážek;2003, mm	26,6	4,3
5					
6					
7			průměrných úhrnů srážek Brno, 2003		
8					
9			měsíc	1	2
10			MĚSÍČNÍ PRŮMĚRY	26,6	4,3
11			KUMULOVANÝ SOUČET	26,6	30,9
12			KLOUZAVÝ PRŮMĚR	482,6	454,9
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					

zdrojová data

Oblast dat: Řada

Řady

MĚSÍČNÍ PRŮMĚRY
KUMULOVANÝ SOUČET
Řada3

Název:

Hodnoty:

Přidat Odstranit

Popisky osy X (kategorie):

Storno < Zpět Další > Dokončit

List1 List2 List3

Start Automatiké tvarv JARO_06 Cvičení_10_testo... Cvičení 12_07_ca... Microsoft Excel Microsoft Excel Darina Foltýnová... CS 10:59

Z - diagramy

Z - diagram průměrných úhrnů srážek (mm), Brno, 2003

