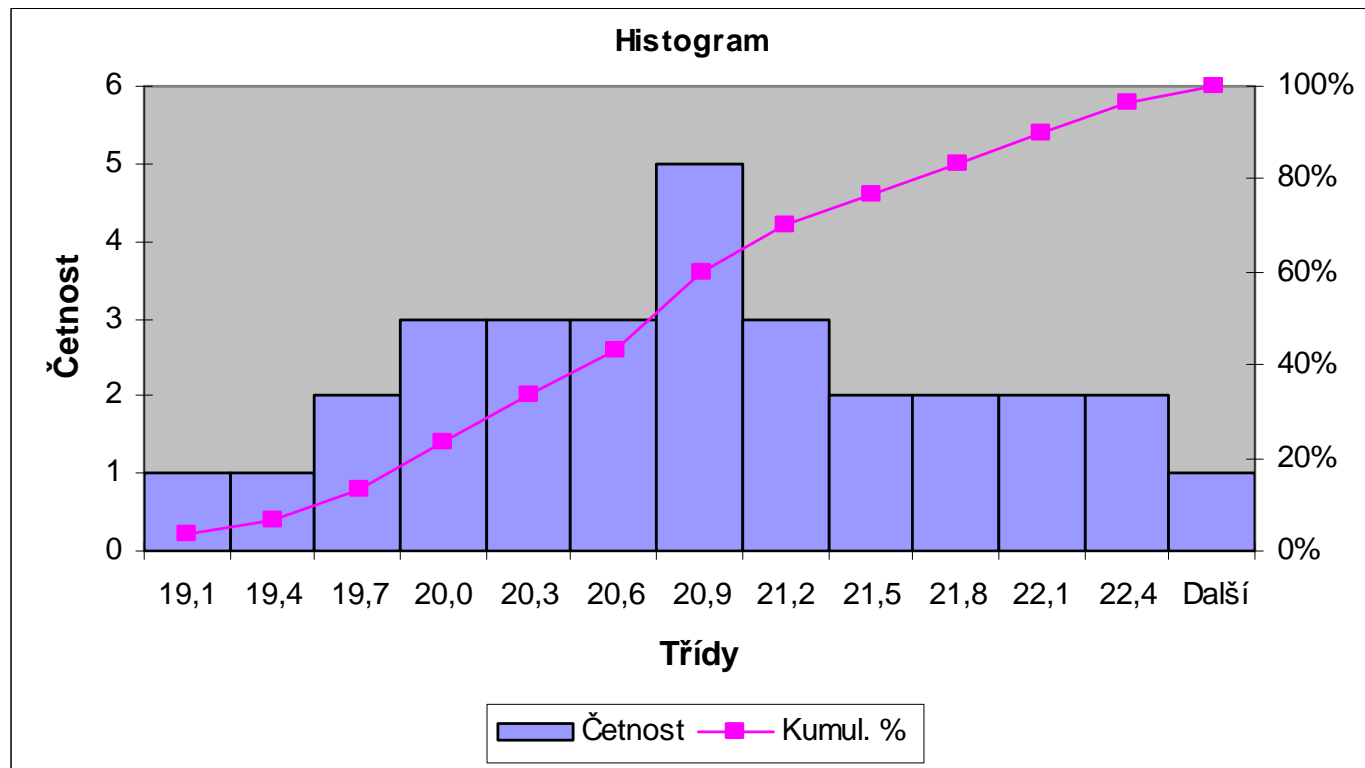
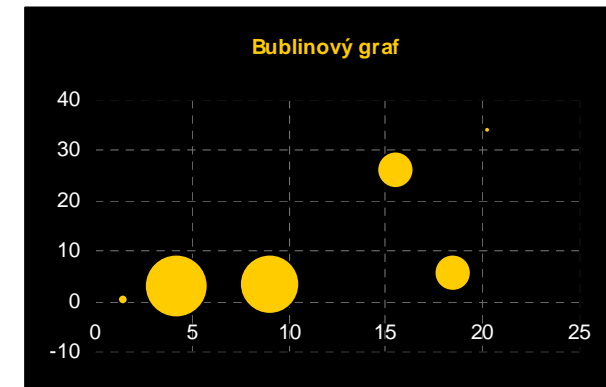


Výrobní produkce divizí Ice Cream



Statistika

Poloha planet



IN2009

Střední hodnoty

Udávají střed celé skupiny údajů, kolem kterého všechny hodnoty kolísají (analogie těžiště).

- **Aritmetický průměr** - vznikne součtem hodnot statistického souboru dělených jejich počtem

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- **V Excelu se užívá pro průměr funkce: PRŮMĚR(ČÍSLO1;ČÍSLO2)**

Střední hodnoty

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	=									

Vložit funkci

Funkce: naposledy použité vše finanční datum a čas matematické statistické vyhledávací databáze text logické informační

Název funkce: HODINA SUMIF **PRŮMĚR** NÁHČÍSLO MIN KDYŽ SMODCH.VÝBĚR ODMOCNINA SVYHLEDAT VVYHLEDAT

PRŮMĚR(číslo1;číslo2;...)
Vrátí průměrnou hodnotu (aritmetický průměr) argumentů. Argumenty mohou být čísla nebo názvy, matice nebo odkazy, které obsahují čísla.

OK Storno

X **✓** **=** =PRŮMĚR(C3:L3)

B	C	D	E	F	G	H	I	J	K	L
n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	=PRŮMĚR(C3:L3)									

Střední hodnoty

- **Medián** – je prostřední hodnota v uspořádaném statistickém souboru a to prostřední v pořadí hodnot uspořádaných podle velikosti
- tzn.: polovina hodnot výběru je menší nebo rovna mediánu, polovina hodnot je větší nebo rovna mediánu
- pokud není prostřední hodnota určena jednoznačně /sudý počet hodnot/, je za medián brán průměr dvou prostředních hodnot

Střední hodnoty

- Příklad 1:

Mějme dána následující čísla:

1, 3, 2, 2, 4, 4, 2, 2, 5, 1, 2, 1

Čísla uspořádáme vzestupně:

1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5

Protože hodnot proměnných je 12, je **medián** roven aritmetickému průměru šesté a sedmé hodnoty:

$x_{50} = 2$ (přesný výpočet uvedeme v kap. **Kvantily**)

- V Excelu se využívá funkce **MEDIAN(ČÍSLO1;ČÍSLO2)**

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	=									

Vložit funkci [?] [X]

Funkce: Název funkce:

naposledy použité	LOGLINREGRESE
vše	LOGLINTREND
finanční	LOGNORMDIST
datum a čas	MAX
matematické	MAXA
statistické	MEDIAN
vyhledávací	MIN
databáze	MINA
text	MODE
logické	NEGBINOMDIST
informační	NORMDIST

MEDIAN(číslo1;číslo2;...)
Vrátí medián, střední hodnotu množiny zadaných čísel.

[?] [OK] [Storno]

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	=MEDIAN(C3:L3)									

Střední hodnoty

- **Modus** – je nejčastější hodnota znaku, nebo-li ta hodnota, která se ve statistickém souboru vyskytuje nejčastěji
- např. z příkladu 1:
1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5 je vidět, že je modus roven 2
- **V Excelu** využíváme funkce:
MODE(ČÍSLO1;ČÍSLO2)

Modus

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	2,05									
modus	=									

Vložit funkci

Funkce: Název funkce:

- naposledy použité
- vše
- finanční
- datum a čas
- matematické
- statistické**
- vyhledávací
- databáze
- text
- logické
- informační

- LOGNORMDIST
- MAX
- MAXA
- MEDIAN
- MIN
- MINA
- MODE**
- NEGBINOMDIST
- NORMDIST
- NORMINV
- NORMSDIST

MODE(číslo1;číslo2;...)
Vrátí hodnotu, která se v matici nebo v oblasti dat vyskytuje nejčastěji.

OK Storno

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	2,05									
modus	=MODE(C3:L3)									

MODE

Číslo1: C3:L3 = {2,11;2,01;2,09;2,02;2,03;2,03;2,1}

Číslo2: = matice

= 2,11

Vrátí hodnotu, která se v matici nebo v oblasti dat vyskytuje nejčastěji.

Číslo1: číslo1;číslo2;... je 1 až 30 čísel nebo názvů, matic nebo odkazů obsahujících čísla, jejichž modus chcete zjistit.

Wynodek = 2,11

OK Storno

Shrnutí středních hodnot

- **Modus** je nejčastěji se vyskytující hodnota v množině hodnot
- **Medián** je prostřední hodnota uspořádaného souboru
- **Průměr** je průměrná hodnota

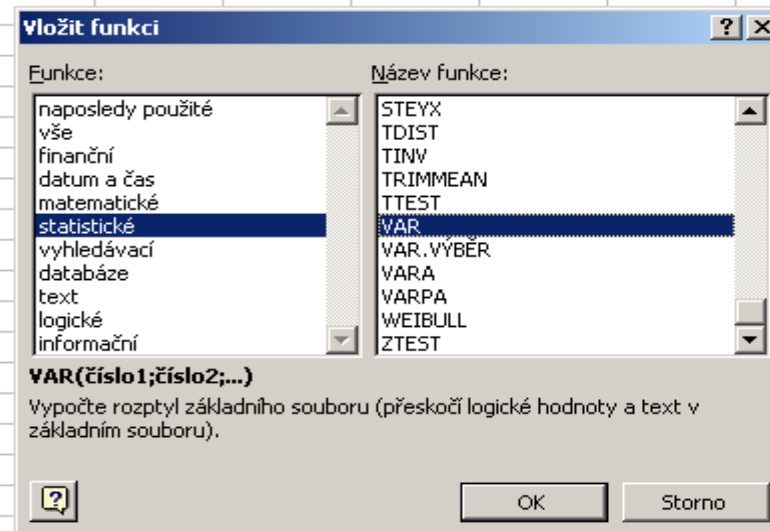
Funkce popisující odchylky jednotlivých hodnot od aritmetického průměru

- **Rozptyl s_x^2** – je definován jako součet kvadratických odchylek od průměru, děleným rozsahem stat. souboru
- **V Excelu se užívá funkce: VAR(číslo1;číslo2)**

$$s_x^2 = \frac{1}{n} \sum_{i=1}^n \left(x_i - \bar{x} \right)^2$$

Rozptyl

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	2,05									
modus	2,11									
rozptyl	=									



VAR													
A	B	C	D	E	F	G	H	I	J	K	L		
	n	1	2	3	4	5	6	7	8	9	10		
	naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11		
	průměrná hodnota	2,06											
	medián	2,05											
	modus	2,11											
	rozptyl	=VAR(C3:L3)											

- **Směrodatná odchylka s_x** – vypočítá se jako druhá odmocnina z rozptylu

$$s_x = \sqrt{s_x^2}$$

- **V Excelu se užívá funkce **SMODCH(ČÍSLO1;ČÍSLO2)****

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	2,05									
modus	2,11									
rozptyl	0,00136									
směrodatná odchylka	=									

Vložit funkci

Funkce: Název funkce:

naposledy použité	QUARTIL
vše	RANK
finanční	RKQ
datum a čas	SKEW
matematické	SLOPE
statistické	SMALL
vyhledávací	SMODCH
databáze	SMODCH.VYBER
text	STANDARDIZE
logické	STDEVA
informační	STDEVPA

SMODCH(číslo1;číslo2;...)

Vypočte směrodatnou odchylku základního souboru, který byl zadán jako argument (přeskočí logické hodnoty a text).

OK Storno

Směrodatná odchylka s

SMODCH =SMODCH(C3:L3)

	A	B	C	D	E	F	G	H	I	J	K	L	M
2		n	1	2	3	4	5	6	7	8	9	10	
3		naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11	
4		průměrná hodnota					2,06						
5		medián					2,05						
6		modus					2,11						
7		rozptyl					0,00136						
8		směrodatná odchylka	=SMODCH(C3:L3)										

SMODCH

Číslo1 C3:L3 = {2,11;2,01;2,09;2,02;2,03;2,03;2,1;2,05;2,05;2,11}

Číslo2 = číslo

= 0,036878178

Vypočte směrodatnou odchylku základního souboru, který byl zadán jako argument (přeskočí logické hodnoty a text).

Číslo1: číslo1;číslo2;... je 1 až 30 čísel nebo odkazů obsahujících čísla, které odpovídají základnímu souboru.

?

Výsledek = 0,036878178

OK Storno

Rozptyl a směrodatná odchylka určují, jak se liší jednotlivé hodnoty od střední hodnoty

Variační koeficient v_x

- Charakterizuje přesnost měření či výsledku

$$v_x = \frac{s_x}{\bar{x}} \cdot 100$$

- Uvádí se v %

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,11	2,01	2,09	2,02	2,03	2,03	2,1	2,05	2,05	2,11
průměrná hodnota	2,06									
medián	2,05									
modus	2,11									
rozptyl	0,001									
směrodatná odchylka	0,037									
variační koeficient	=08/04*100									

Kvantil x_p – (p-procentní kvantil)

- kvantil je hodnota znaku, pro kterou platí, že nejméně p-procent prvků má hodnotu menší nebo rovnu x_p a (100-p) procent prvků je větších nebo rovno x_p (viz medián)
- používají se tyto kvantily:

Medián x_{50} , dolní kvartil x_{25} ,

horní kvartil x_{75} , decily ($x_{10}, x_{20}, \dots, x_{90}$)

a percentily (x_1, x_2, \dots, x_{99})

Kvantil

- Pro nás nejdůležitější:

Medián, dolní a horní kvartil a percentil

- Vrátime se k příkladu 1:

Máme dána čísla: 1, 3, 2, 2, 4, 1, 4, 2, 2, 5, 1, 2

Čísla uspořádáme vzestupně:

1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5

Počet naměřených hodnot je tedy 12.

Výpočet kvantilů bez využití MS Excel

- Poté, co si hodnoty uspořádáme vzestupně dle velikosti, příslušný kvantil získáme jako pořadí **k**-té hodnoty, vypočtené ze vztahu:

$$\mathbf{k = (\text{počet hodnot}) \cdot (\text{úroveň kvantilu „p“}) / 100}$$

kde číslo **k** zaokrouhlujeme nahoru

naše řada: 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5

- **Dolní kvartil x_{25} :**

$(12.25)/100 = 3$, a je tedy roven číslu na

3. místě: 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5

$$x_{25} = 1$$

- **Medián x_{50} :** (máme sudý počet (12) hodnot, je roven aritmetickému průměru 6 a 7 hodnoty)

$$x_{50} = (2+2)/2 = 2$$

- **Horní kvartil x_{75} :**

$(12.75)/100 = 9$ a je tedy $x_{75} = 3$

Kvantily v MS Excel

- dolní kvartil se označuje jako kvartil 1
- medián jako kvartil 2
- horní kvartil jako kvartil 3
- využíváme funkce **QUARTIL**
- opět označíme danou oblast (pole)
- do druhého řádku doplníme číslo kvartilu, který chceme spočítat

Dolní kvartil

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,01	2,02	2,03	2,03	2,05	2,05	2,09	2,1	2,11	2,11
dolní kvartil	=QUARTIL(C14:L14;1)									
medián										
horní kvartil										
percentil										

QUARTIL

Pole C14:L14 = {2,01;2,02;2,03;2,03;2,05;2,05;2,09;2,1;2,11;2,11}

Kvartil 1 = 1

= 2,03

Vrátí hodnotu kvartilu množiny dat.

Kvartil je číslo: minimální hodnota = 0, první kvartil = 1, medián = 2, třetí kvartil = 3, maximální hodnota = 4.

Wysledek = 2,03

OK Storno

Percentil

n	1	2	3	4	5	6	7	8	9	10
naměřeno	2,01	2,02	2,03	2,03	2,05	2,05	2,09	2,1	2,11	2,11
dolní kvartil					2,03					
medián					2,05					
horní kvartil					2,0975					
percentil	=PERCENTIL(C14:L14;0,85)									

využíváme funkce:

PERCENTIL(OBLAST;HODNOTA PERCENTILU)

Mezikvartilová odchylka

- je definována jako polovina z rozdílu horního a dolního kvartilu $\frac{1}{2} \cdot (x_{75} - x_{25})$
- užívá se v kombinaci s ostatními charakteristikami zejména pro rozlišení, jaká je variabilita (správnost) či koncentrace hodnot kolem středu



Získané hodnoty statistických souborů se zpracovávají do grafů – tzv. histogramů

Zkuste z uvedeného grafu

- 1) vytvořit tabulku
- 2) určit medián a modus počtu telat
- 3) vypočítat směrodatnou a mezikvartilovou odchylku
- 4) určit variační koeficient