

Cvičení 9:

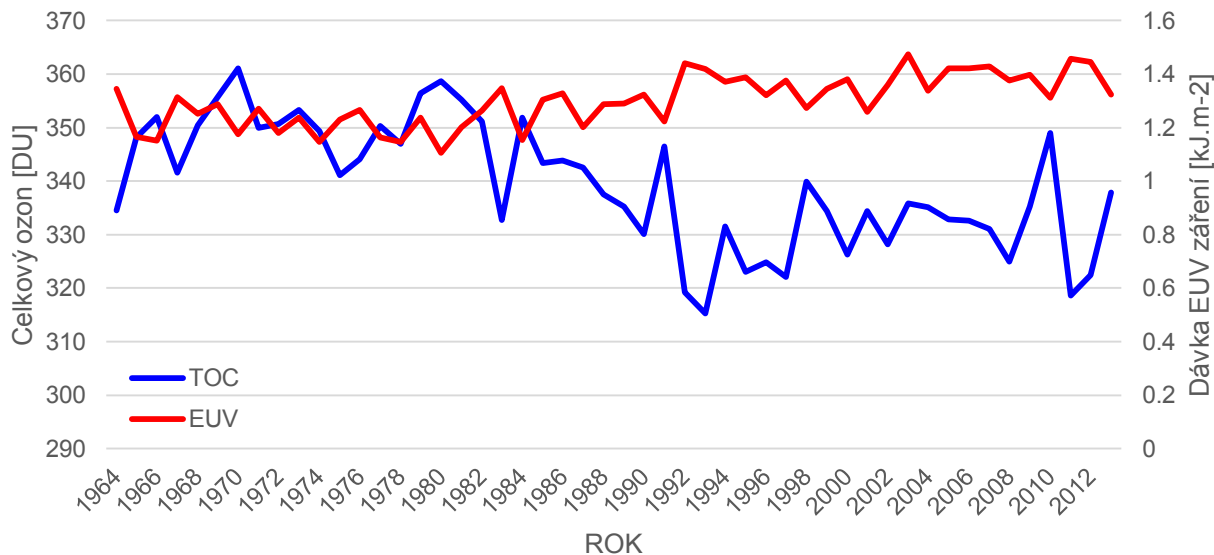
Korelační počet

Řešený příklad

- Máme k dispozici časové řady průměrných měsíčních teplot vzduchu ze stanic Praha-Klementinum a z další stanice (každý jinak, viz cvičení 3) v období 1961–1990
- **Hodnotíme těsnot vztahu časových řad v jednotlivých měsících**

POZOR! Data použitá v této prezentaci nejsou z vašich souborů!
(vaše výsledky se budou lišit)

V této prezentaci jsem použila průměrný měsíční sloupec ozonu (TOC) a měsíční průměr dávek erytmového UV záření (EUV) v Hradci Králové v období 1964–2013



Příprava dat

V Excelu spojíme data z Klementina a data z vaší stanice:

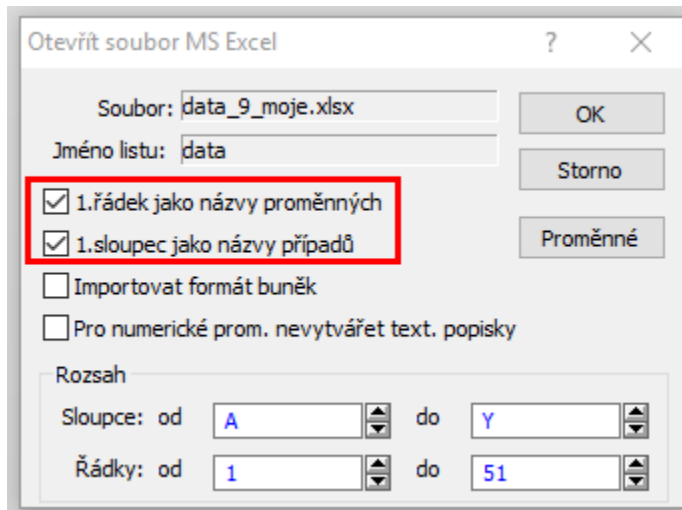
- Nejprve zařadíme průměrné měsíční teploty vzduchu v jednotlivých měsících prosinec z Klementina
- Pak průměrné měsíční teploty vzduchu v jednotlivých měsících z vaší stanice

Pomocí názvů sloupců odlišíme obě stanice (například K_I–K_XII a X_I–X_XII).

| Klementinum | | | | | | | | | | | | | Vaše stanice | | | | | | | | | | | |
|-------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|--------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| ROK | EUV_I | EUV_II | EUV_III | EUV_IV | EUV_V | EUV_VI | EUV_VII | EUV_VIII | EUV_IX | EUV_X | EUV_XI | EUV_XII | TOC_I | TOC_II | TOC_III | TOC_IV | TOC_V | TOC_VI | TOC_VII | TOC_VIII | TOC_IX | TOC_X | TOC_XI | TOC_XII |
| 1964 | 0,179974 | 0,356656 | 0,711408 | 1,511748 | 2,43797 | 3,074028 | 3,271717 | 2,042632 | 1,575531 | 0,619524 | 0,208202 | 0,118004 | 317,7444 | 366,0477 | 375,1449 | 399,733 | 381,7572 | 334,9465 | 330,902 | 320,0077 | 297,0871 | 299,525 | 278,2199 | 313,6707 |
| 1965 | 0,137116 | 0,407793 | 0,761907 | 1,209954 | 1,739884 | 2,575216 | 2,278059 | 2,288295 | 1,398464 | 0,802003 | 0,195963 | 0,133211 | 355,8482 | 392,8182 | 411,7521 | 399,4212 | 394,6654 | 354,3509 | 338,4226 | 323,1715 | 302,2555 | 280,4872 | 310,9139 | 317,8539 |
| 1966 | 0,177767 | 0,304865 | 0,576994 | 1,40001 | 2,063123 | 2,588159 | 2,319597 | 2,00734 | 1,395266 | 0,613312 | 0,186012 | 0,12698 | 368,4351 | 359,6764 | 437,876 | 389,4628 | 381,5495 | 359,5801 | 347,0565 | 323,133 | 298,2721 | 297,613 | 322,5045 | 337,7947 |
| 1967 | 0,187008 | 0,363015 | 0,705281 | 1,440096 | 2,181035 | 3,010552 | 3,067458 | 2,26115 | 1,295015 | 0,765858 | 0,260232 | 0,14669 | 375,3519 | 375,222 | 396,8692 | 390,776 | 354,6419 | 356,4386 | 334,3727 | 323,1058 | 299,9739 | 280,1351 | 285,0695 | 329,242 |
| 1968 | 0,187039 | 0,264098 | 0,834394 | 1,654746 | 2,2573 | 2,88119 | 2,74441 | 2,06195 | 1,266742 | 0,509654 | 0,206115 | 0,115716 | 397,6726 | 388,5881 | 403,8872 | 388,4379 | 364,1008 | 352,4603 | 345,429 | 338,0397 | 304,4317 | 291,1258 | 273,3157 | 358,0552 |
| 1969 | 0,189075 | 0,308297 | 0,798552 | 1,543321 | 2,429858 | 2,536952 | 3,002874 | 1,897278 | 1,619469 | 0,660724 | 0,22237 | 0,181683 | 377,1346 | 425,2395 | 387,4682 | 407,0307 | 382,0771 | 376,7341 | 342,8925 | 342,7658 | 299,6012 | 290,806 | 308,8439 | 333,6124 |
| 1970 | 0,175009 | 0,344565 | 0,734452 | 1,089132 | 1,986624 | 2,677383 | 2,693881 | 2,017053 | 1,472566 | 0,544789 | 0,182574 | 0,118412 | 350,5972 | 439,2741 | 432,9776 | 428,3402 | 401,898 | 377,1708 | 350,2298 | 333,4172 | 316,3271 | 299,1919 | 302,3892 | 308,3112 |
| 1971 | 0,19948 | 0,302383 | 0,760524 | 1,485522 | 2,239979 | 2,359993 | 3,023039 | 2,474801 | 1,276187 | 0,736048 | 0,211476 | 0,102562 | 361,1175 | 395,2399 | 413,9916 | 394,4481 | 376,3285 | 371,0598 | 343,2615 | 320,8688 | 323,7148 | 289,9958 | 304,1017 | 309,0793 |
| 1972 | 0,158178 | 0,324904 | 0,846914 | 1,270992 | 1,827704 | 2,751645 | 2,569716 | 2,184841 | 1,197267 | 0,651593 | 0,224615 | 0,122714 | 362,5859 | 378,5138 | 398,0024 | 384,6324 | 381,2078 | 369,1317 | 345,4323 | 321,1367 | 324,7106 | 311,2608 | 306,1498 | 325,818 |
| 1973 | 0,135191 | 0,265279 | 0,841022 | 1,151628 | 2,225787 | 2,684124 | 2,394877 | 2,529736 | 1,51107 | 0,638891 | 0,228056 | 0,159092 | 346,5172 | 395,5076 | 389,7574 | 441,8863 | 380,3753 | 363,4373 | 357,6481 | 330,0543 | 297,16 | 308,8013 | 311,1532 | 321,1358 |
| 1974 | 0,109201 | 0,301083 | 0,940837 | 1,584764 | 1,830976 | 2,324384 | 2,19069 | 2,262814 | 1,404861 | 0,424036 | 0,212874 | 0,108397 | 356,9497 | 388,6736 | 356,0806 | 397,5356 | 392,9527 | 378,6317 | 345,3622 | 325,2585 | 298,1865 | 333,9936 | 313,7065 | 308,6931 |
| 1975 | 0,150524 | 0,480166 | 0,69051 | 1,453044 | 2,336462 | 2,302567 | 2,716415 | 2,257237 | 1,541888 | 0,475558 | 0,19161 | 0,108795 | 331,6476 | 354,8721 | 389,5059 | 397,8799 | 367,8606 | 363,3538 | 342,2355 | 326,436 | 297,6793 | 302,0874 | 311,1888 | 309,4442 |
| 1976 | 0,175404 | 0,366247 | 0,74363 | 1,582629 | 2,345674 | 3,071949 | 2,653105 | 2,110928 | 1,171979 | 0,606522 | 0,200638 | 0,13956 | 339,543 | 361,8966 | 388,3589 | 389,6413 | 375,2479 | 358,005 | 350,7081 | 337,9542 | 317,7005 | 286,9479 | 293,5088 | 330,6922 |
| 1977 | 0,188713 | 0,271606 | 0,781985 | 1,363001 | 2,253637 | 2,654123 | 2,454106 | 1,816457 | 1,185403 | 0,583758 | 0,23325 | 0,121708 | 371,7511 | 405,9634 | 379,9724 | 405,9881 | 378,7614 | 369,754 | 356,9448 | 336,0832 | 310,4425 | 280,5106 | 304,4414 | 308,161 |
| 1978 | 0,140808 | 0,345199 | 0,800823 | 1,399811 | 1,957199 | 2,703918 | 2,419856 | 1,963066 | 1,053719 | 0,572754 | 0,197572 | 0,139124 | 357,079 | 382,1564 | 375,6613 | 404,5841 | 391,6171 | 375,5233 | 356,6637 | 335,7665 | 307,8985 | 286,2845 | 293,9647 | 299,3793 |
| 1979 | 0,218947 | 0,474283 | 0,662355 | 1,220023 | 2,586885 | 2,876765 | 2,188351 | 2,07424 | 1,436948 | 0,753269 | 0,186686 | 0,112966 | 376,0037 | 388,6261 | 403,0612 | 420,9151 | 384,9258 | 356,7382 | 365,0452 | 333,1463 | 305,9254 | 301,1263 | 299,4961 | 342,9332 |
| 1980 | 0,2238 | 0,358331 | 0,696904 | 1,154496 | 2,274024 | 2,241338 | 1,935255 | 2,128194 | 1,298084 | 0,554907 | 0,222191 | 0,133682 | 365,0534 | 369,619 | 404,0586 | 435,4673 | 401,5971 | 379,6615 | 372,6677 | 333,2912 | 306,7745 | 304,7599 | 307,2586 | 325,1172 |
| 1981 | 0,205336 | 0,353923 | 0,738802 | 1,458267 | 2,204575 | 2,741693 | 2,234932 | 2,194939 | 1,299369 | 0,556073 | 0,208746 | 0,161927 | 367,9843 | 391,4613 | 381,8723 | 396,6101 | 401,7871 | 353,1156 | 351,7566 | 335,0359 | 309,7714 | 313,7574 | 314,6744 | 346,929 |
| 1982 | 0,209317 | 0,342135 | 0,797074 | 1,51294 | 2,256021 | 2,628305 | 2,669816 | 2,157218 | 1,51767 | 0,612247 | 0,263108 | 0,127247 | 363,6782 | 403,9848 | 421,0748 | 400,0812 | 385,0302 | 374,683 | 347,3886 | 337,5854 | 301,3233 | 301,4526 | 282,397 | 298,0455 |
| 1983 | 0,145605 | 0,542652 | 0,744404 | 1,406095 | 2,27256 | 3,007175 | 3,338035 | 2,144722 | 1,421219 | 0,689378 | 0,245039 | 0,167967 | 317,828 | 365,3226 | 362,5812 | 378,7586 | 362,0597 | 354,1018 | 330,9323 | 332,5369 | 302,9067 | 282,8215 | 291,1421 | 315,0891 |
| 1984 | 0,171168 | 0,342469 | 0,80673 | 1,417609 | 1,979892 | 2,444198 | 2,302798 | 2,233483 | 1,125061 | 0,629158 | 0,229048 | 0,123943 | 370,3273 | 378,0132 | 392,2486 | 409,586 | 387,3626 | 376,3639 | 352,0071 | 332,3935 | 317,3145 | 298,032 | 289,0194 | 321,0972 |
| 1985 | 0,215689 | 0,383164 | 0,669503 | 1,491833 | 2,338564 | 2,250378 | 3,163519 | 2,472395 | 1,464995 | 0,721328 | 0,274596 | 0,118935 | 382,8794 | 395,0804 | 381,0657 | 376,6055 | 368,6385 | 366,3144 | 331,998 | 311,3922 | 298,7056 | 282,2347 | 308,2645 | 321,5131 |
| 1986 | 0,163551 | 0,485416 | 0,85634 | 1,588074 | 2,384627 | 2,934426 | 2,753217 | 2,178658 | 1,390641 | 0,769725 | 0,254873 | 0,120603 | 383,6018 | 402,7928 | 366,4311 | 393,0576 | 364,7488 | 353,5982 | 343,4127 | 314,9754 | 296,2819 | 287,3906 | 297,7495 | 326,6789 |

Import dat

1. Připravený soubor v Excelu uložíme a zavřeme
2. Importujeme do programu STATISTICA:
 - Soubor – Otevřít
 - 1. řádek jako názvy proměnných
 - 1. sloupec jako názvy případů



Korelační analýza

- **Korelační počet** se zabývá analýzou síly závislosti statistických znaků
- **Závislost by měla být vždy logicky zdůvodnitelná**
- **Korelační koeficient r_{xy}**
 - Relativní míra závislosti statistických znaků (lze použít k porovnání síly závislosti dvou a více dvojic výběrových souborů), nabývá hodnot od -1 do 1
 - Vypočítá se jako podíl kovariance (absolutní míra závislosti, nelze ji použít k porovnání závislosti více dvojic výběrových souborů) a součinu směrodatných odchylek obou souborů

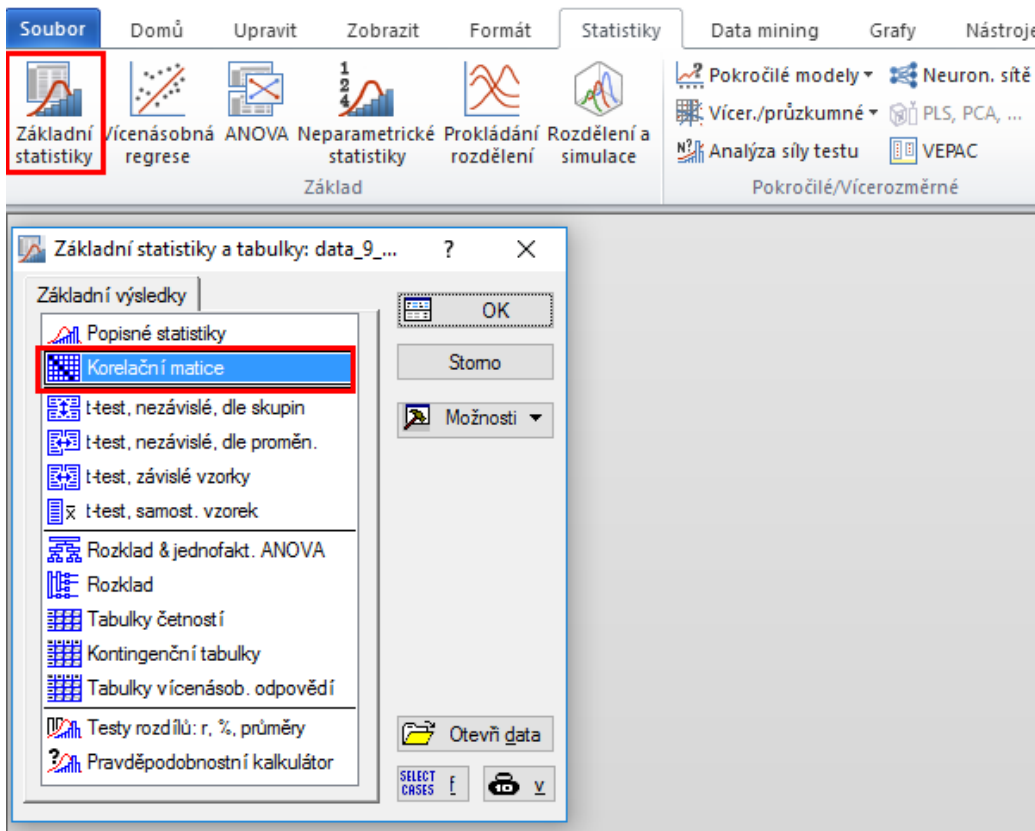
$r_{xy} = 0$ → soubory jsou nekorelované (neexistuje mezi nimi lineární závislost)

$r_{xy} > 0$ → soubory jsou kladně korelované (s růstem hodnot veličiny Y rostou hodnoty veličiny X a naopak)

$r_{xy} < 0$ → soubory jsou záporně korelované (s růstem hodnot veličiny Y klesají hodnoty veličiny X a naopak)
- **Podmínky použitelnosti**
 - Normální rozdělení výběrových souborů (a jeho dvourozměrnost)
 - Lineární vztah
 - **před výpočtem korelačního koeficientu je vhodné ověřit normalitu rozdělení!**

Korelační matice

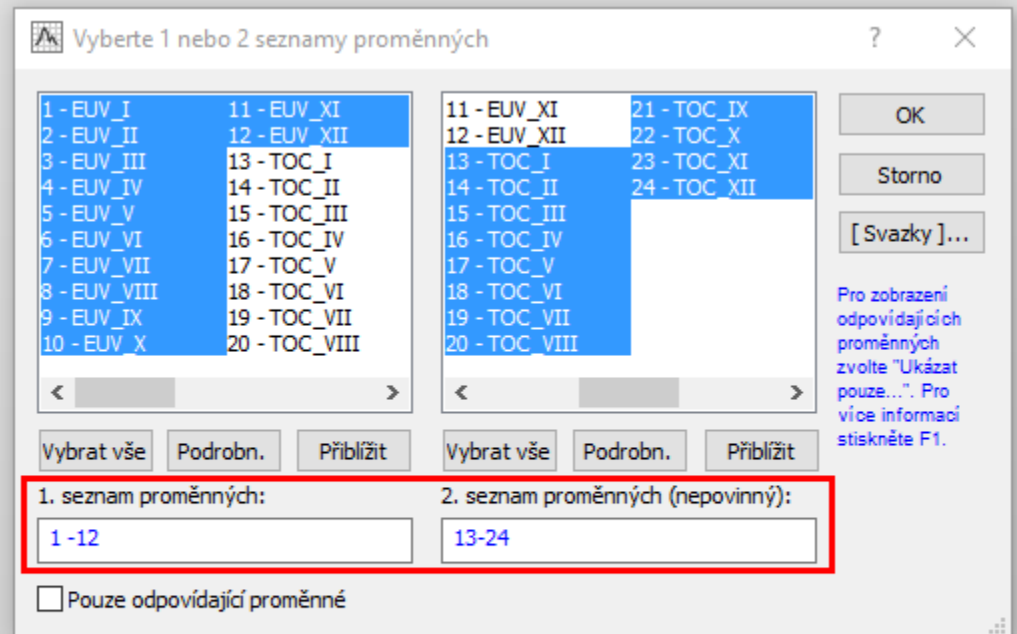
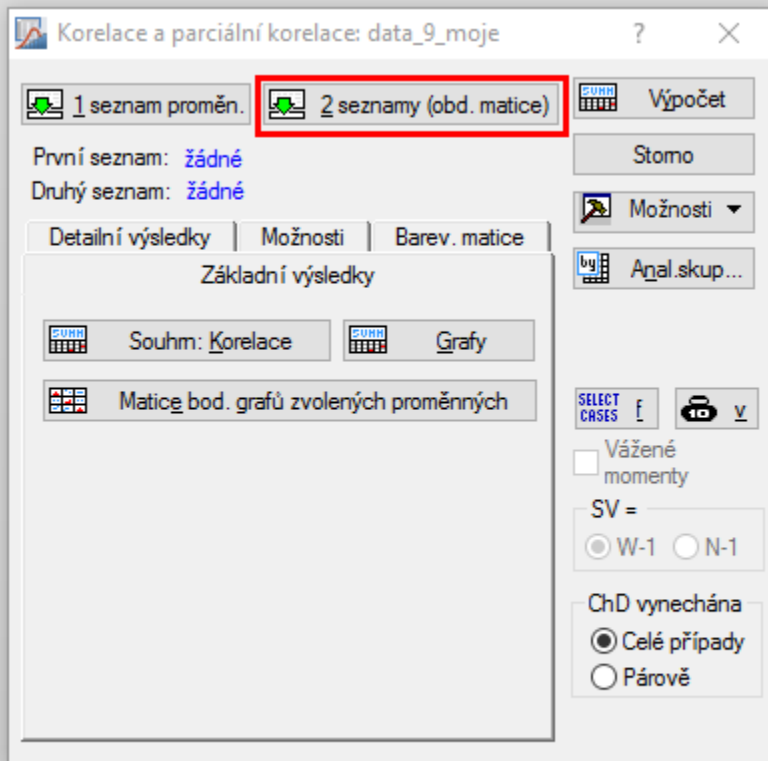
1. Vypočítáme **korelační matici**: Statistika – Základní statistiky – Korelační matice – OK



Korelační matice

2. Vybereme proměnné: 2 seznamy (obd. matice)

- do 1. seznamu zadáme proměnné 1–12 (Klementinum)
- do 2. seznamu zadáme proměnné 13–24 (vaše stanice)



Korelační matice

3. Výpočet a interpretace korelační matice

- Zajímají nás jen korelační koeficienty **na diagonále** (leden s lednem, únor s únorem...)

| Proměnná | Korelace (data_9_moje) | | | | | | | | | | | |
|----------|--|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | Označ. korelace jsou významné na hlad. $p < ,05000$ N=50 (Celé případy vynechány u ChD) | | | | | | | | | | | |
| | TOC_I | TOC_II | TOC_III | TOC_IV | TOC_V | TOC_VI | TOC_VII | TOC_VIII | TOC_IX | TOC_X | TOC_XI | TOC_XII |
| EUV_I | -0,057278 | -0,109529 | 0,107405 | -0,193316 | -0,153347 | -0,194305 | -0,101334 | -0,016109 | -0,159290 | -0,239753 | -0,112295 | 0,026991 |
| EUV_II | -0,282171 | -0,345164 | -0,359157 | -0,271172 | -0,306675 | -0,361116 | -0,319525 | -0,263717 | -0,161553 | -0,237939 | 0,007628 | -0,224960 |
| EUV_III | -0,235188 | -0,224020 | -0,580039 | -0,359678 | -0,380460 | -0,367197 | -0,289825 | -0,215235 | -0,266172 | -0,109910 | -0,055538 | -0,268400 |
| EUV_IV | -0,229296 | -0,307070 | -0,395798 | -0,752794 | -0,625322 | -0,576217 | -0,550763 | -0,555038 | -0,318088 | -0,247027 | -0,349358 | -0,017459 |
| EUV_V | -0,474518 | -0,475756 | -0,510196 | -0,504130 | -0,731448 | -0,675025 | -0,441857 | -0,540524 | -0,334044 | -0,430726 | -0,313835 | -0,085801 |
| EUV_VI | -0,301869 | -0,299735 | -0,422434 | -0,415668 | -0,533383 | -0,722474 | -0,416602 | -0,385045 | -0,227261 | -0,315469 | -0,393182 | -0,071371 |
| EUV_VII | -0,267180 | 0,017975 | -0,169167 | -0,358625 | -0,414375 | -0,368092 | -0,732985 | -0,227495 | -0,192326 | -0,412376 | -0,304370 | -0,262845 |
| EUV_VIII | -0,302973 | -0,400414 | -0,442486 | -0,516640 | -0,599205 | -0,513973 | -0,457580 | -0,809434 | -0,251697 | -0,129951 | -0,195118 | -0,181984 |
| EUV_IX | -0,141946 | 0,144923 | -0,086057 | -0,116210 | -0,082491 | -0,201939 | -0,276543 | -0,111812 | -0,775355 | 0,027495 | -0,086193 | -0,202409 |
| EUV_X | 0,008296 | 0,059603 | 0,018238 | -0,127591 | -0,132775 | -0,125033 | -0,214447 | -0,123733 | -0,131099 | -0,555998 | -0,111546 | -0,060611 |
| EUV_XI | 0,108925 | 0,087306 | -0,152339 | -0,284282 | -0,369055 | -0,152106 | -0,225145 | -0,384695 | -0,353187 | -0,196108 | -0,468401 | -0,087179 |
| EUV_XII | -0,047986 | 0,127584 | -0,182308 | 0,116921 | 0,021858 | 0,018446 | -0,047137 | 0,169706 | -0,002540 | -0,006748 | 0,128396 | -0,067975 |

$r_{xy} = 0$ → soubory jsou nekorelované (neexistuje mezi nimi lineární závislost)

$r_{xy} > 0$ → soubory jsou kladně korelované (s růstem hodnot veličiny Y rostou hodnoty veličiny X a naopak)

$r_{xy} < 0$ → soubory jsou záporně korelované (s růstem hodnot veličiny Y klesají hodnoty veličiny X a naopak)

Korelační matice

3. Výpočet a interpretace korelační matice

| Proměnná | Korelace (data_9_moje) | | | | | | | | | | | |
|----------|--|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | Označ. korelace jsou významné na hlad. p < ,05000 N=50 (Celé případy vynechány u ChD) | | | | | | | | | | | |
| | TOC I | TOC II | TOC III | TOC IV | TOC V | TOC VI | TOC VII | TOC VIII | TOC IX | TOC X | TOC XI | TOC XII |
| EUV_I | -0,057278 | -0,109529 | 0,107405 | -0,193316 | -0,153347 | -0,194305 | -0,101334 | -0,016109 | -0,159290 | -0,239753 | -0,112295 | 0,026991 |
| EUV_II | -0,282171 | -0,345164 | -0,359157 | -0,271172 | -0,306675 | -0,361116 | -0,319525 | -0,263717 | -0,161553 | -0,237939 | 0,007628 | -0,224960 |
| EUV_III | -0,235188 | -0,224020 | -0,580039 | -0,359678 | -0,380460 | -0,367197 | -0,289825 | -0,215235 | -0,266172 | -0,109910 | -0,055538 | -0,268400 |
| EUV_IV | -0,229296 | -0,307070 | -0,395798 | -0,752794 | -0,625322 | -0,576217 | -0,550763 | -0,555038 | -0,318088 | -0,247027 | -0,349358 | -0,017459 |
| EUV_V | -0,474518 | -0,475756 | -0,510196 | -0,504130 | -0,731448 | -0,675025 | -0,441857 | -0,540524 | -0,334044 | -0,430726 | -0,313835 | -0,085801 |
| EUV_VI | -0,301869 | -0,299735 | -0,422434 | -0,415668 | -0,533383 | -0,722474 | -0,416602 | -0,385045 | -0,227261 | -0,315469 | -0,393182 | -0,071371 |
| EUV_VII | -0,267180 | 0,017975 | -0,169167 | -0,358625 | -0,414375 | -0,368092 | -0,732985 | -0,227495 | -0,192326 | -0,412376 | -0,304370 | -0,262845 |
| EUV_VIII | -0,302973 | -0,400414 | -0,442486 | -0,516640 | -0,599205 | -0,513973 | -0,457580 | -0,809434 | -0,251697 | -0,129951 | -0,195118 | -0,181984 |
| EUV_IX | -0,141946 | 0,144923 | -0,086057 | -0,116210 | -0,082491 | -0,201939 | -0,276543 | -0,111812 | -0,775355 | 0,027495 | -0,086193 | -0,202409 |
| EUV_X | 0,008296 | 0,059603 | 0,018238 | -0,127591 | -0,132775 | -0,125033 | -0,214447 | -0,123733 | -0,131099 | -0,555998 | -0,111546 | -0,060611 |
| EUV_XI | 0,108925 | 0,087306 | -0,152339 | -0,284282 | -0,369055 | -0,152106 | -0,225145 | -0,384695 | -0,353187 | -0,196108 | -0,468401 | -0,087179 |
| EUV_XII | -0,047986 | 0,127584 | -0,182308 | 0,116921 | 0,021858 | 0,018446 | -0,047137 | 0,169706 | -0,002540 | -0,006748 | 0,128396 | -0,067975 |

Interpretujeme všechny měsíce (ukázka pro ozon a EUV záření):

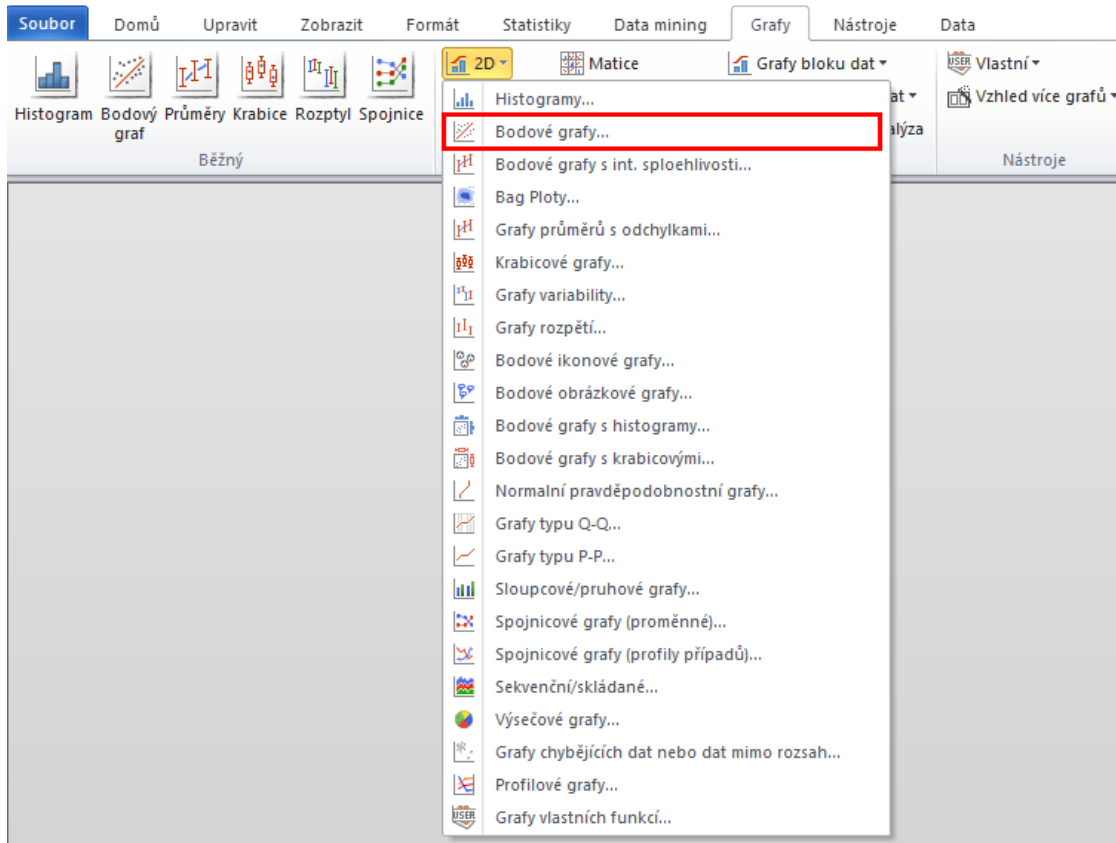
V lednu a v prosinci neexistoval mezi průměrným měsíčním množstvím ozonu a průměrnými dávkami EUV záření statisticky významný vztah.

Ve všech měsících s výjimkou ledna a prosince bylo průměrné měsíční množství ozonu statisticky významně záporně korelováno s průměrnými dávkami EUV záření.

Nejsilnější vztah mezi průměrným měsíčním množstvím ozonu a průměrnými dávkami EUV záření byl zaznamenán v srpnu, kdy korelační koeficient činil -0,81. Nejslabší vztah byl zaznamenán v lednu ($r_{xy} = -0,06$).

Korelační pole

- Pouze pro měsíc s nejvyšší a nejnižší hodnotou korelačního koeficientu
- Korelační pole = bodový graf
 1. Grafy – 2D – Bodové grafy



Korelační pole

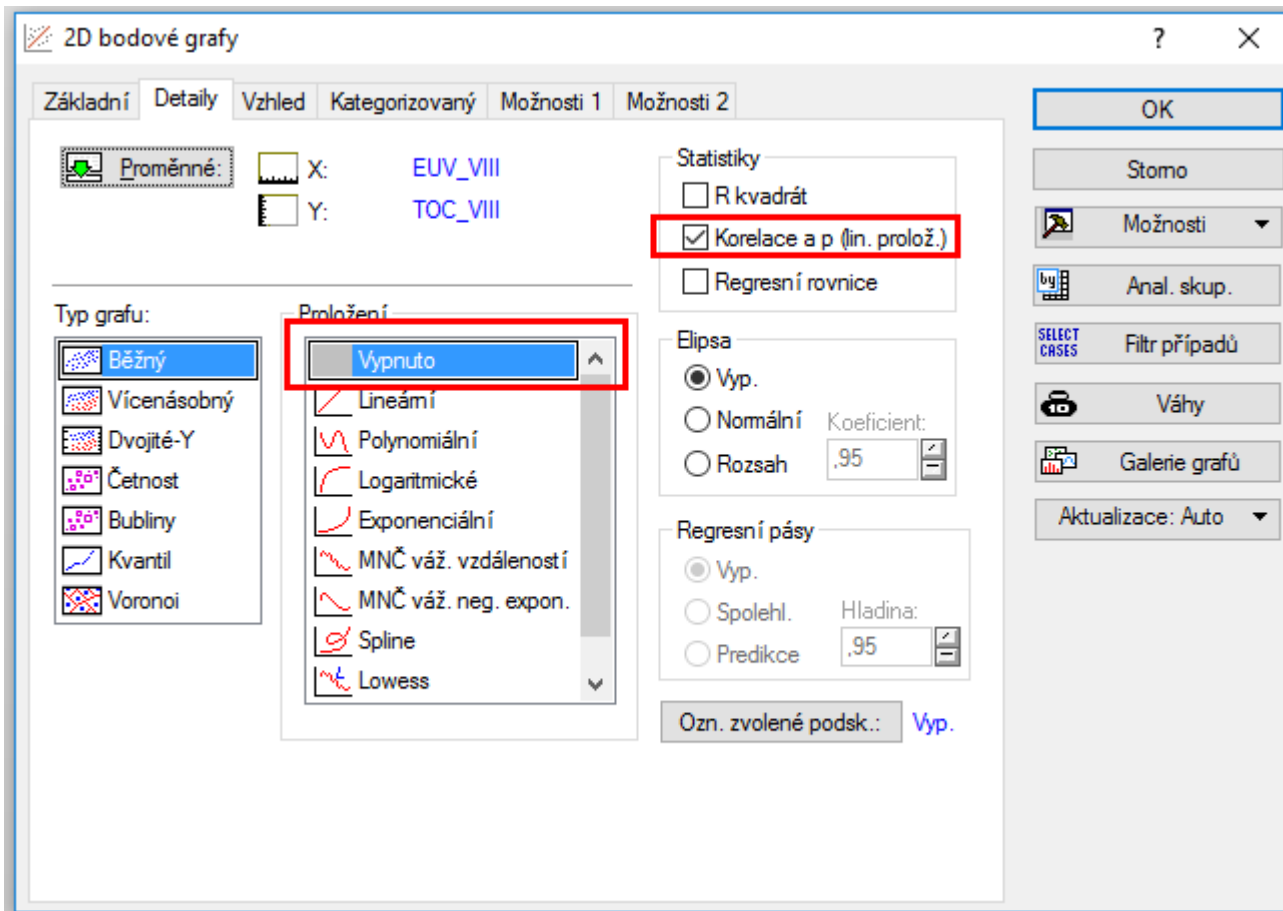
2. Vybereme proměnné

- X: Klementinum (měsíc s nejvyšším korelačním koeficientem)
- Y: vaše stanice (měsíc s nejvyšším korelačním koeficientem)

The image shows two overlapping software windows. The background window is titled "2D bodové grafy" and has tabs for "Základní", "Detaily", "Vzhled", "Kategorizovaný", "Možnosti 1", and "Možnosti 2". The "Základní" tab is active, showing a "Proměnné:" section with a red box around the variable selection icon. Below it, X and Y axes are both set to "žádné". The "Typ grafu:" section has "Běžný" selected. The "Regresní pásy" section has "Vyp." selected and "Hladina:" set to ".95". The "Typ proložení:" section has "Lineární" checked. The foreground window is titled "Vyberte proměnné pro bodový graf" and displays a list of variables in two columns. The first column lists variables from 1 to 10 (EUV_I to EUV_X), and the second column lists variables from 11 to 20 (EUV_XI to TOC_VIII). The variable "8 - EUV_VIII" is highlighted in blue. Below the list, there are input fields for X and Y, with "8" entered for X and "20" entered for Y. The "Vybrat vše" button is visible for both columns.

Korelační pole

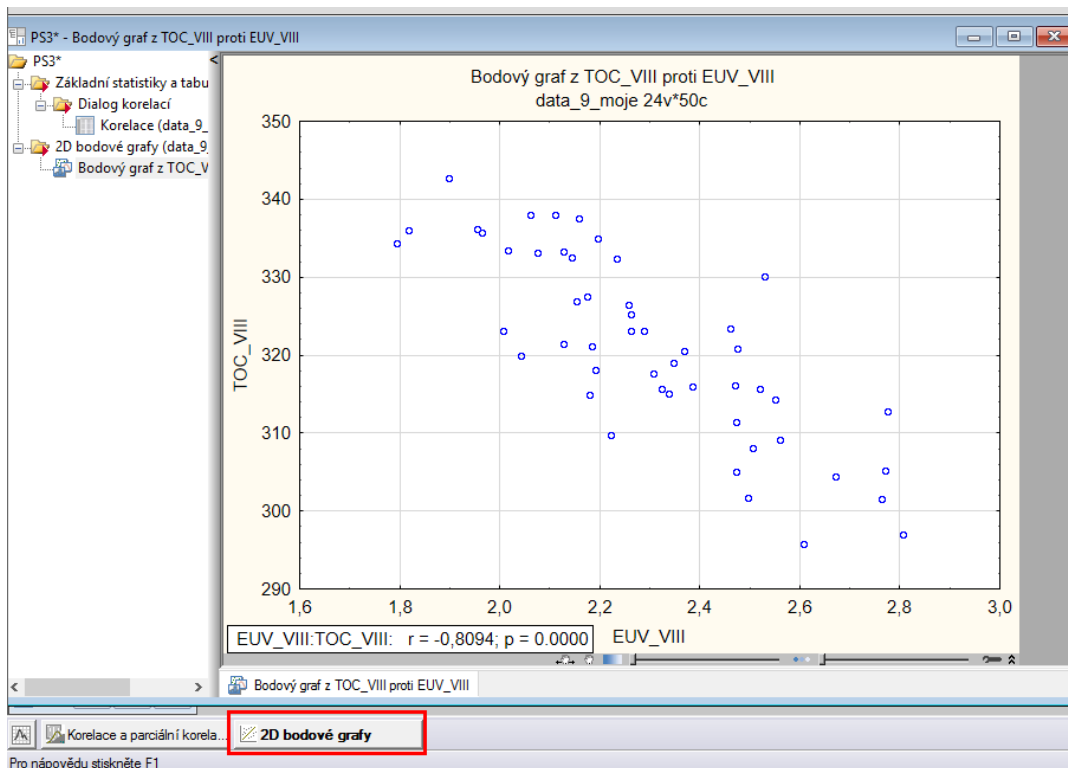
3. Na kartě Details vypneme proložení a zvolíme výpočet korelačního koeficientu a příslušné p-hodnoty



Korelační pole

4. Necháme si vykreslit bodový graf (OK)

5. Postup zopakujeme pro měsíc s nejnižším korelačním koeficientem (stačí se přepnout na okno 2D bodových grafů a změnit proměnné)



2D bodové grafy

Základní | Detaily | Vzhled | Kategorizovaný | Možnosti 1 | Možnosti 2

Proměnné:

X: EUV_VIII
Y: TOC_VIII

Statistiky

R kvadrát
 Korelace a p (lin. prolož.)
 Regresní rovnice

Elipsa

Vyp.
 Normální Koeficient: .95
 Rozsah

Regresní pásy

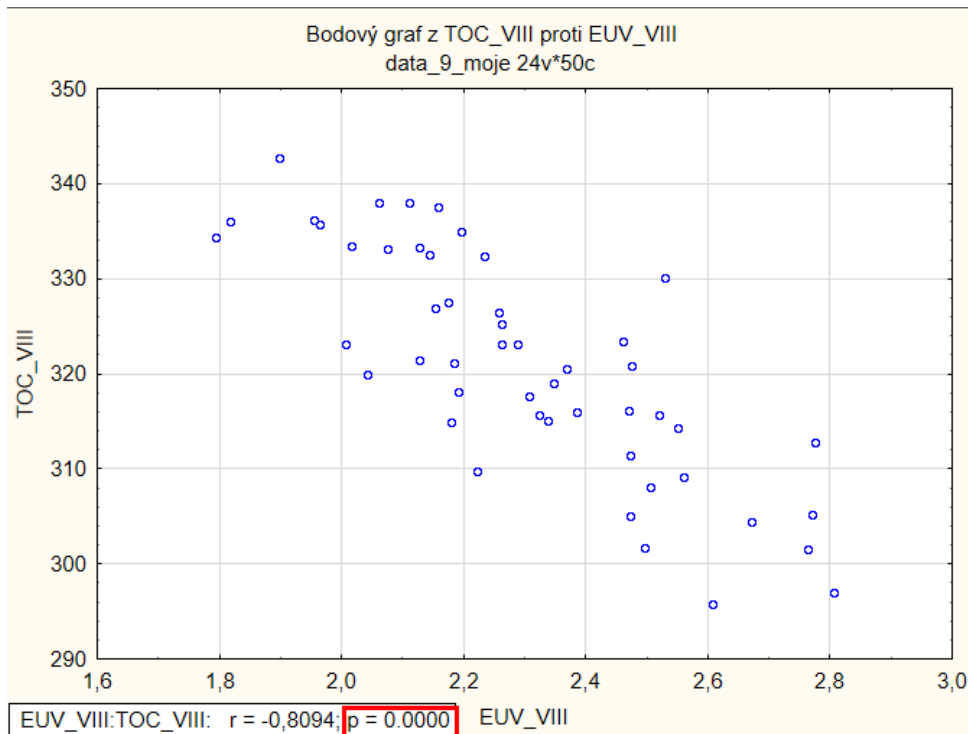
Vyp.
 Spolehl. Hladina: .95
 Predikce

Ozn. zvolené podsk.: Vyp.

OK
Storno
Možnosti
Anal. skup.
Filtr případů
Váhy
Galerie grafů
Aktualizace: Auto

Korelační pole

6. Oba vykreslené grafy popíšeme (kladná, záporná, nebo žádná korelace?) a interpretujeme p-hodnotu



Nulová hypotéza:

Mezi výběrovými soubory neexistuje statisticky významný vztah.

Hladina významnosti = 0,05

Pomocí **p-hodnoty**

(označeno červeně)

zjistíme **platnost nulové hypotézy:**

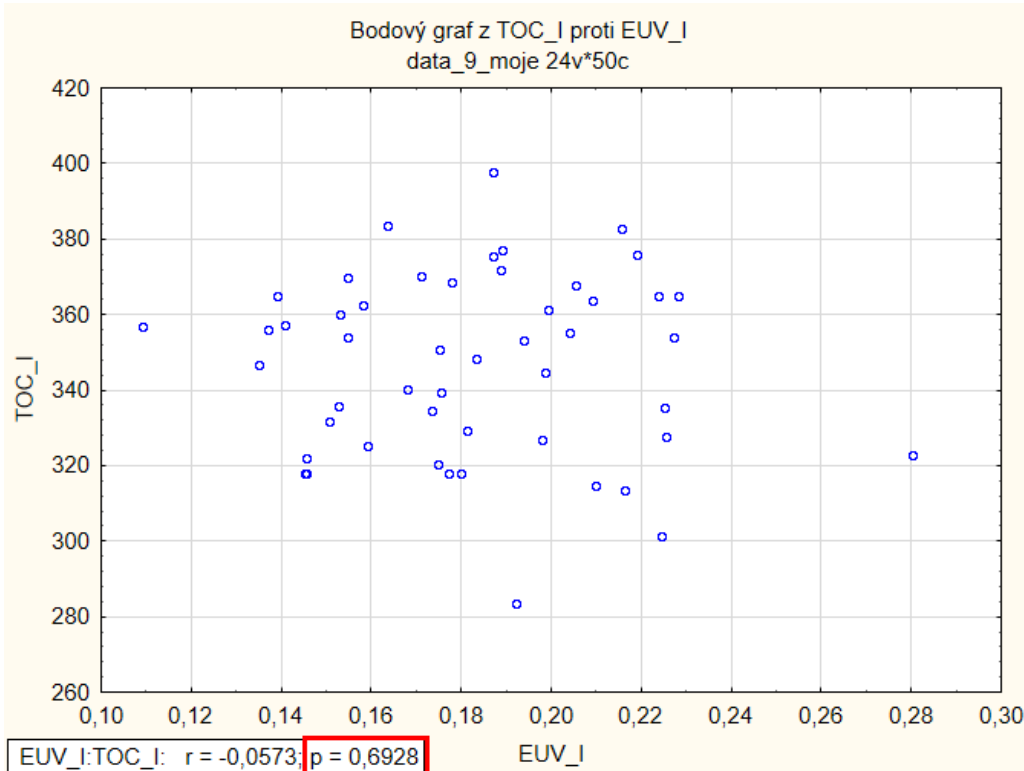
$p > 0,05 \rightarrow H_0$ přijímáme

$p < 0,05 \rightarrow H_0$ zamítáme

Interpretace (ukázka): Mezi průměrným měsíčním sloupcem ozonu a dávkami EUV záření existuje v srpnu velmi silná záporná korelace, která je statisticky významná na hladině významnosti 0,05. Nulovou hypotézu na hladině významnosti 0,05 zamítáme.

Korelační pole

6. Oba vykreslené grafy popíšeme (kladná, záporná, nebo žádná korelace?) a interpretujeme p-hodnotu



Nulová hypotéza:

Mezi výběrovými soubory neexistuje statisticky významný vztah.

Hladina významnosti = 0,05

Pomocí **p-hodnoty**

(označeno červeně)

zjistíme platnost nulové hypotézy:

$p > 0,05 \rightarrow H_0$ přijímáme

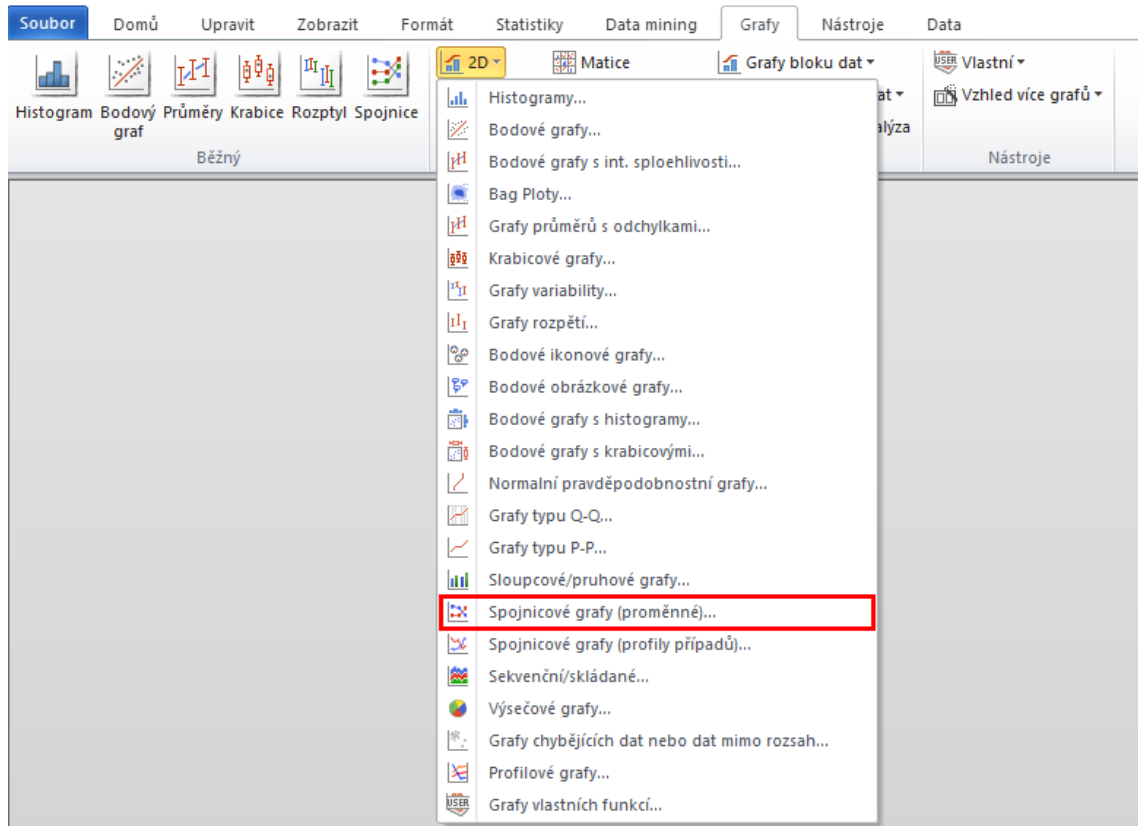
$p < 0,05 \rightarrow H_0$ zamítáme

Interpretace (ukázka): Mezi průměrným měsíčním sloupcem ozonu a dávkami EUV záření existuje v lednu velmi slabá záporná korelace, která není statisticky významná na hladině významnosti 0,05. Nulovou hypotézu na hladině významnosti 0,05 přijímáme.

Graf průběhu teplot

- Časová řada teploty vzduchu pro měsíce s nejvyšším a nejnižším korelačním koeficientem na obou stanicích
 - Vše do jednoho grafu (spojnicový graf)

1. Grafy – 2D – Spojnicové grafy (proměnné)



Graf průběhu teplot

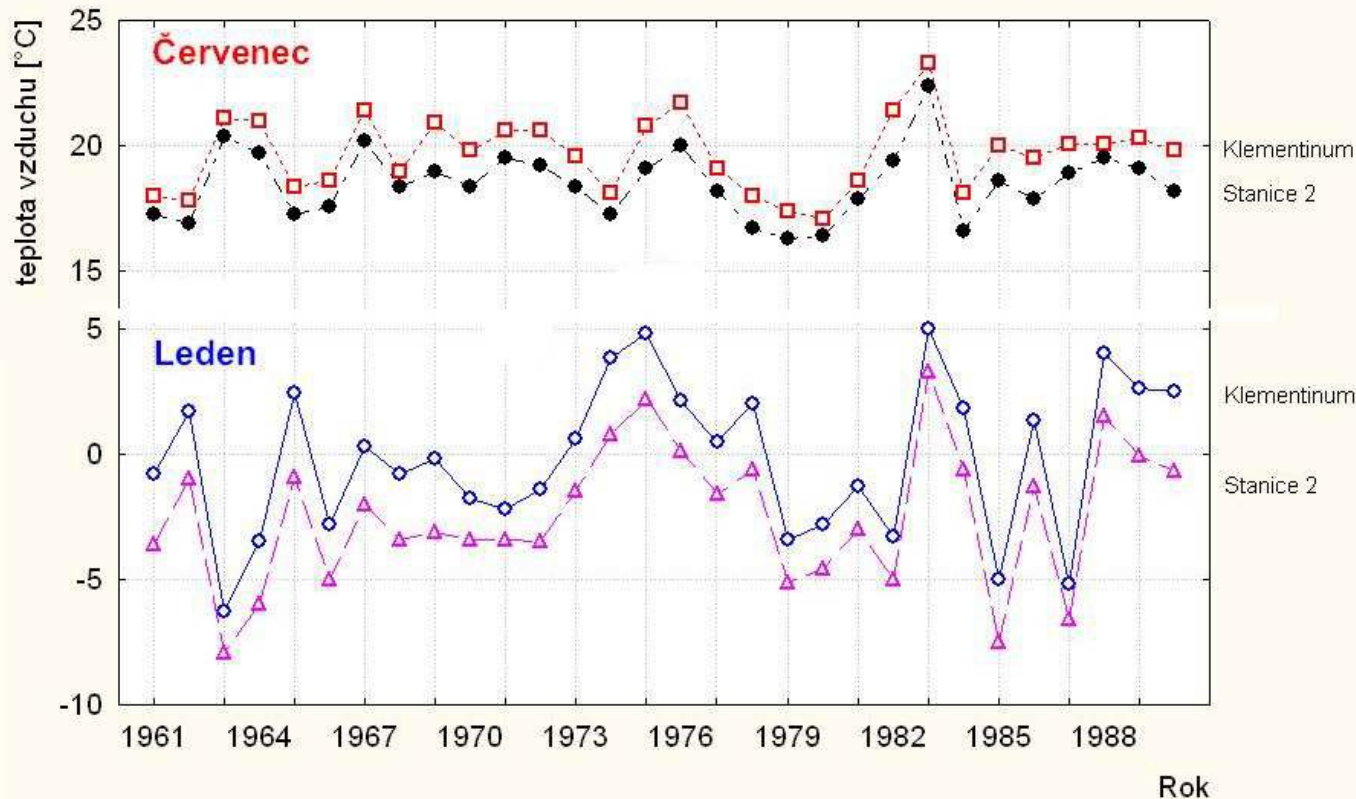
2. **Vybereme proměnné** (teplota vzduchu z Klementina a z vaší stanice pro měsíc s nejvyšším a nejnižším korelačním koeficientem, celkem 4 proměnné)

3. Přepneme typ grafu na **vícenásobný**

The image shows two overlapping dialog boxes from a software application. The left dialog, titled "2D spojnicové grafy", has tabs for "Základní", "Detaily", "Vzhled", "Kategorizovaný", "Možnosti 1", and "Možnosti 2". The "Základní" tab is active, showing a "Proměnné:" field with a dropdown menu set to "žádné". Below it, the "Typ grafu:" section has two radio buttons: "Běžný" and "Vícenásobný", with "Vícenásobný" selected. There are also checkboxes for "Ignorovat ChD" (unchecked) and "Ignorovat body mimo" (checked). The right dialog, titled "Vyberte proměnné pro spojnicový graf", contains a list of 24 variables: 1 - EUV_I, 2 - EUV_II, 3 - EUV_III, 4 - EUV_IV, 5 - EUV_V, 6 - EUV_VI, 7 - EUV_VII, 8 - EUV_VIII, 9 - EUV_IX, 10 - EUV_X, 11 - EUV_XI, 12 - EUV_XII, 13 - TOC_I, 14 - TOC_II, 15 - TOC_III, 16 - TOC_IV, 17 - TOC_V, 18 - TOC_VI, 19 - TOC_VII, 20 - TOC_VIII, 21 - TOC_IX, 22 - TOC_X, 23 - TOC_XI, and 24 - TOC_XII. Variables 1, 8, 13, and 20 are highlighted in blue. At the bottom, the "Proměnné:" field contains the text "1 8 13 20".

Graf průběhu teplot

4. Popíšeme graf (shody, rozdíly v průběhu teplot)



TIP
Po dvojkliknutí na osu můžeme zadat její název

Ukázka je ze zadání cvičení od prof. Dobrovolného

Vyšší teploty vzduchu v Klementinu znamenají i vyšší teploty vzduchu na Stanici 2 (průměrné lednové i červencové teploty vzduchu jsou kladně korelovány). V lednu, kdy je korelace slabší, jsou patrné výraznější rozdíly v průběhu teplot v Klementinu a na Stanici 2.

Do protokolu

- Zadání
- **Korelační matice** (1 tabulka) s vyznačenou diagonálou (nebo tabulka s korelačními koeficienty pro leden až prosinec) + její interpretace
 - Statisticky významné korelační koeficienty vyznačte (podbarvení, jiná barva písma...)
- **Korelační pole** pro měsíc s nejvyšším a nejnižším korelačním koeficientem + popis závislosti a interpretace p-hodnoty (2 grafy)
- **Graf průběhu teplot** + popis (1 graf)
- Závěr

Zdroje

- BUDÍKOVÁ, Marie. Jednoduchá korelační analýza (přednáška). Brno: Masarykova univerzita, 2016.
- DOBROVOLNÝ, Petr. Z1069 Statistické metody a zpracování dat: VI. Korelační a regresní počet (přednáška) Brno: Masarykova univerzita, 2016.
- Veterinární a farmaceutická univerzita Brno: Biostatistika (2016). Dostupný z WWW: <<http://cit.vfu.cz/statpotr/potr/prednasky.htm>>.

Pár slov ke cvičení 7

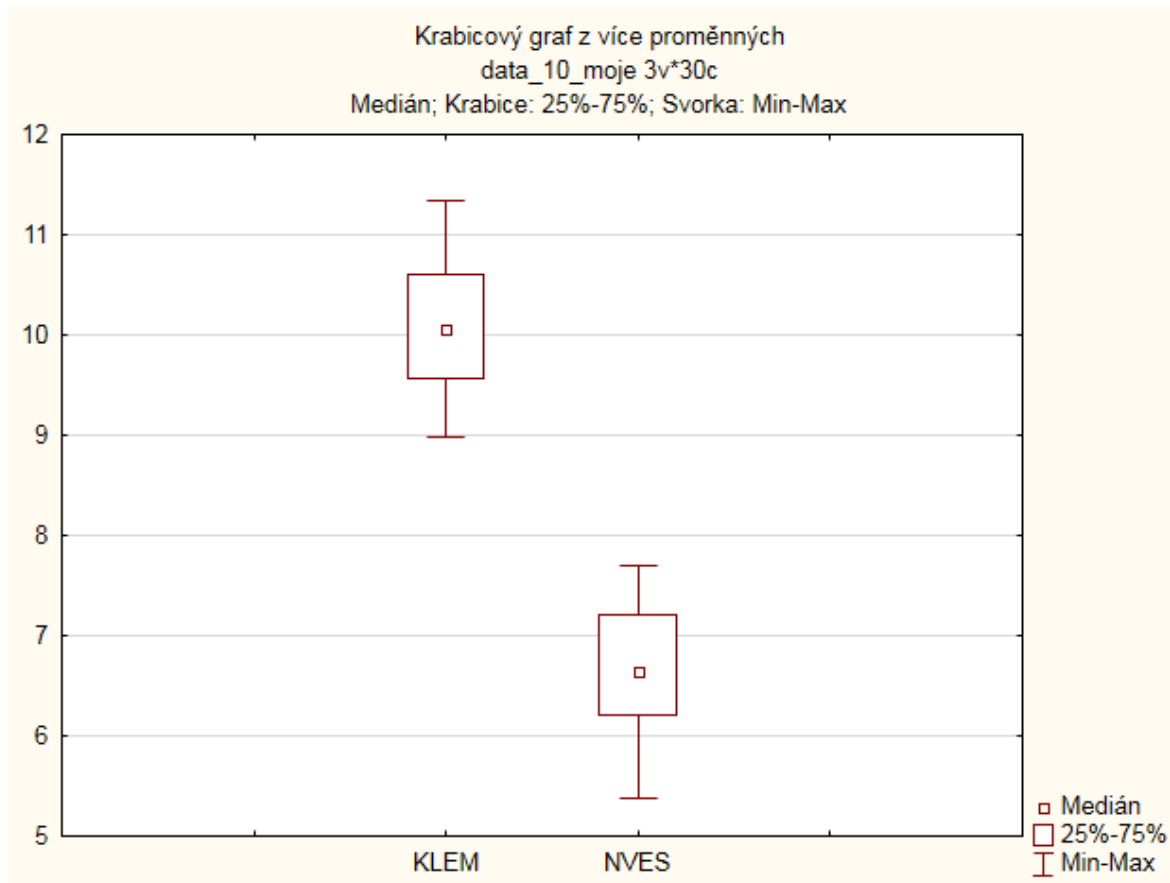
ČASTÉ CHYBY

- Hladina významnosti x hladina spolehlivosti
 - **Hladina významnosti α** : pravděpodobnost chyby I. druhu (zamítnutí H_0 , když ve skutečnosti platí); nejčastější: 0,05; 0,01; 0,10; respektive 5 %, 1 %, 10 %
 - **Hladina spolehlivosti $1 - \alpha$** : nejčastější jsou 0,95; 0,99; 0,90; respektive 95 %, 99 %, 90 %
- Statisticky významně lepší průměr – prosím NE.
- Zdroj: vlastní tvorba
 - Působí to poněkud legračně; běžně se má zato, že kde není výslovně uveden zdroj, je to vlastní myšlenka/nápad/zpracování autora práce
- Chybná data – je důležité vědět, co vlastně testujeme...
 - Vráceno všem, u kterých jsem si toho všimla
 - Prosím buď přepočítat pro hodnoty podle zadání; nebo zjistit, s čím jste pracovali, a následně opravit popis tabulky a grafu a závěr podle toho, co jste doopravdy testovali (například průměrná lednová/červencová teplota vzduchu na stanici X v období Y–Z atd.)

Pár slov ke cvičení 7

NENÍ GRAF JAKO GRAF

- V programu STATISTICA lze udělat více typů krabicových grafů
- Vždy je **nutné zahrnout legendu**



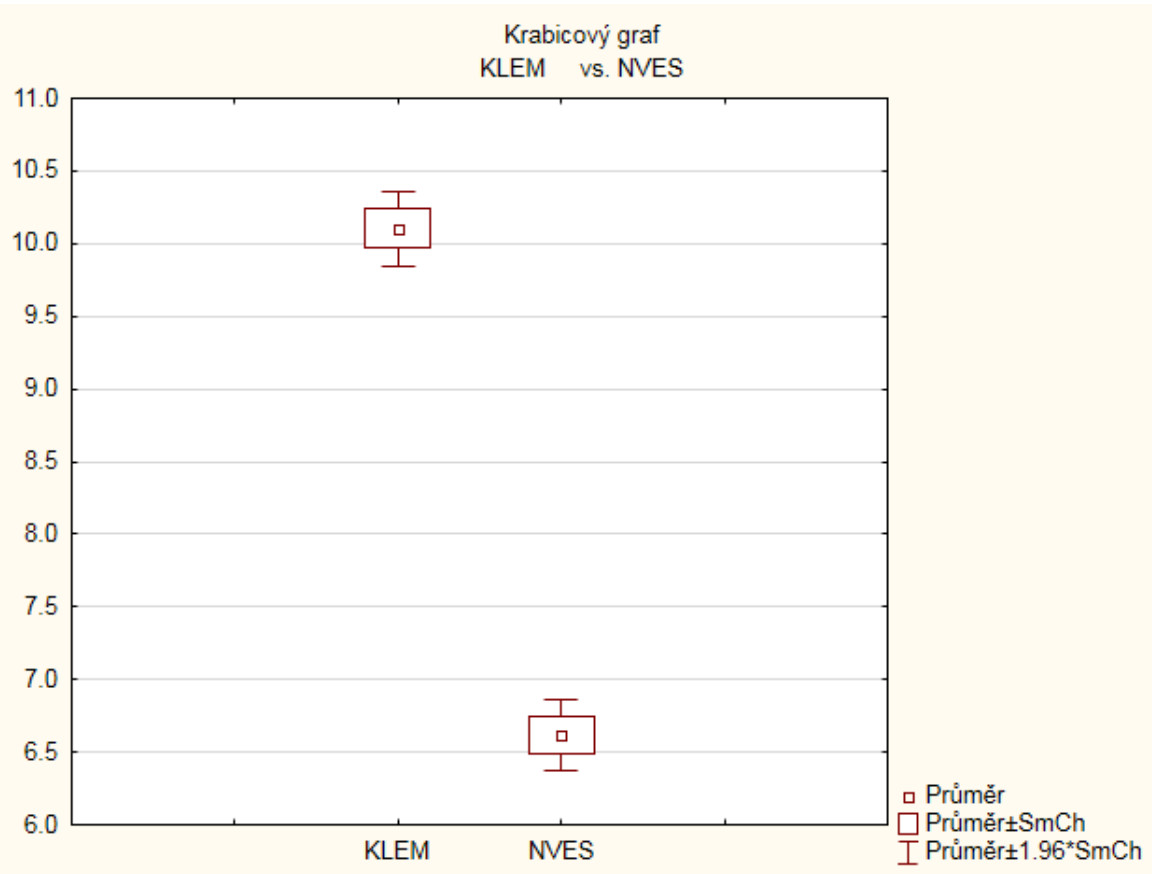
Medián – kvartily – min-max

- Jednoduchá interpretace
- Vhodný pro sledování základních statistických charakteristik souborů (míry úrovně a variability)
- Umožňuje i přidání průměru a posouzení normality rozdělení

Pár slov ke cvičení 7

NENÍ GRAF JAKO GRAF

- V programu STATISTICA lze udělat více typů krabicových grafů
- Vždy je **nutné zahrnout legendu**



Průměr – směrodatná chyba průměru

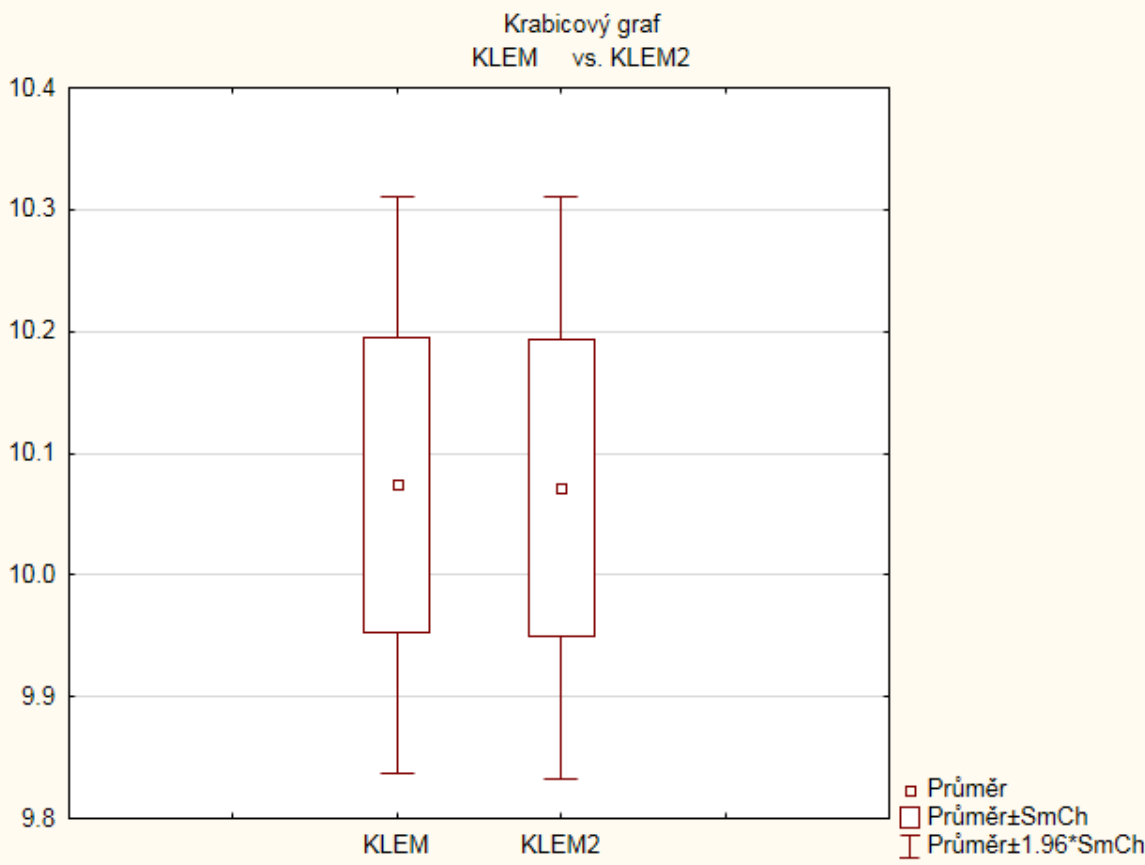
- Nejde o směrodatnou odchylku výběrových souborů
- Směrodatná chyba průměru je **směrodatná odchylka všech možných výběrových průměrů vypočítaných z jednoho základního souboru při velikosti vzorku n**

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Více najdete v kapitole 5 zde: <http://cit.vfu.cz/statpotr/potr/teorie/predn1/variabil.htm>

Pár slov ke cvičení 7

Využití krabicového grafu s průměrem a směrodatnou chybou průměru: hodnocení t-testu pro nezávislé vzorky



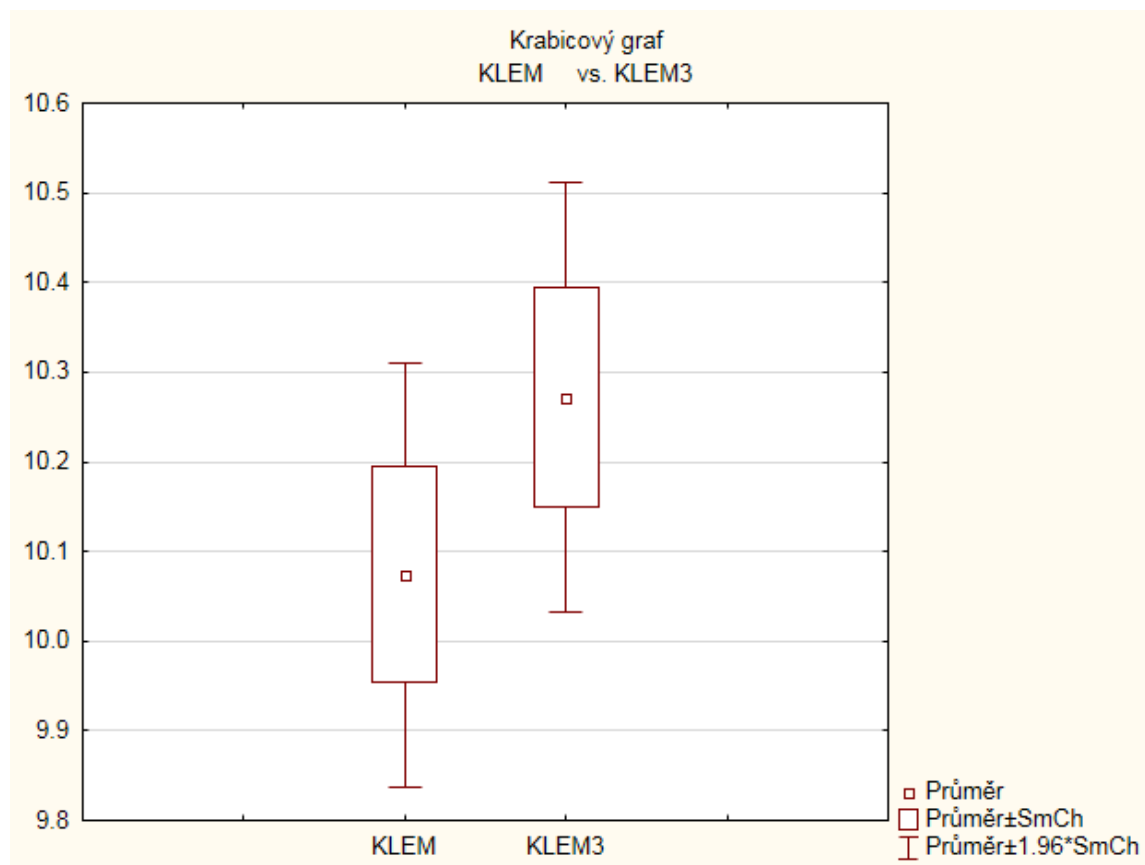
Pokud se svorky nepřekrývají s krabičkami, mezi soubory je statisticky významný rozdíl na hladině významnosti 0,05.

V tomto případě činila p-hodnota testového kritéria párového t-testu **0,93**.

Zkoušela jsem pouze t-test se společným odhadem rozptylů, u samostatného odhadu rozptylů to může fungovat trochu jinak a u párového t-testu taky.

Pár slov ke cvičení 7

Využití krabicového grafu s průměrem a směrodatnou chybou průměru: hodnocení t-testu pro nezávislé vzorky

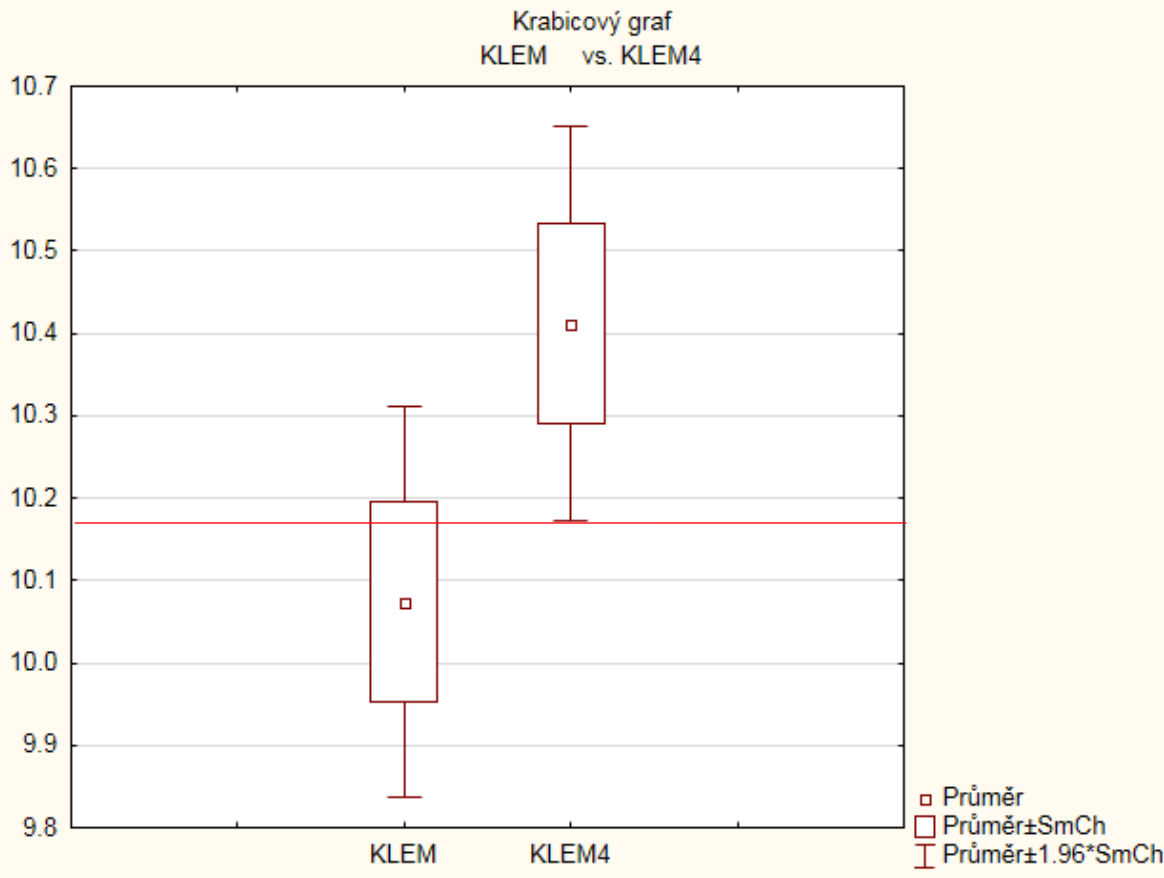


Pokud se svorky nepřekrývají s krabičkami, mezi soubory je statisticky významný rozdíl na hladině významnosti 0,05.

V tomto případě činila p-hodnota testového kritéria t-testu **0,67**.

Pár slov ke cvičení 7

Využití krabicového grafu s průměrem a směrodatnou chybou průměru: hodnocení t-testu pro nezávislé vzorky

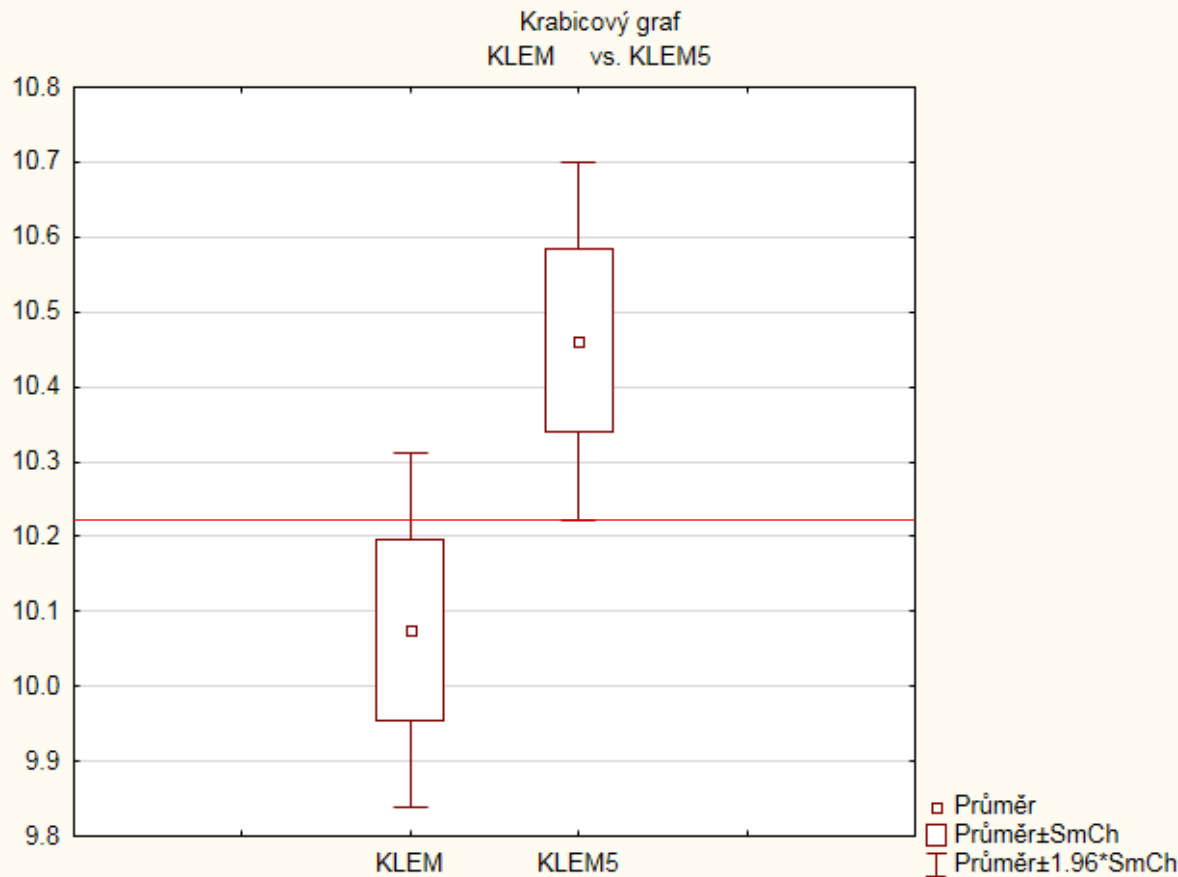


Pokud se svorky nepřekrývají s krabičkami, mezi soubory je statisticky významný rozdíl na hladině významnosti 0,05.

V tomto případě činila p-hodnota testového kritéria t-testu **0,054**.

Pár slov ke cvičení 7

Využití krabicového grafu s průměrem a směrodatnou chybou průměru: hodnocení t-testu pro nezávislé vzorky



Pokud se svorky nepřekrývají s krabičkami, mezi soubory je statisticky významný rozdíl na hladině významnosti 0,05.

V tomto případě činila p-hodnota testového kritéria t-testu **0,028**.