

Grafické zobrazenie údajov

Grafické zobrazenie údajov je prvým a často dôležitým krokom pri analýze údajov. Správne grafické znázornenia môžu upozorniť na javy a vzťahy, ktoré iba v číselnej forme nemusia byť jasné. Následne zvolené a použité metódy môžu byť preto pri absencii základných grafických zobrazení aplikované nesprávne.

Grafy kvalitatívnych premenných

Základom tvorby väčšiny grafov a grafických zobrazení kvalitatívnych premenných je zostavenie jednoduchej frekvenčnej tabuľky pozorovaní. V tabuľke 1 uvádzame frekvenčnú tabuľku počtu laktácií kráv holštajnského plemena podľa poradia laktácie.

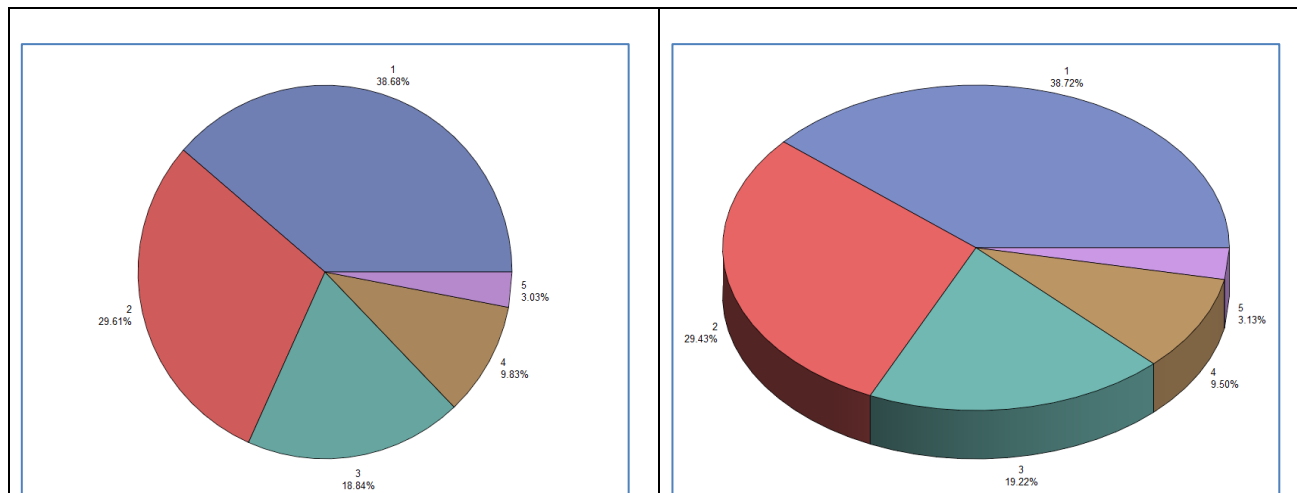
Tab. 1

Poradie laktácie	Frekvencia	%	Kumulatívna frekvencia	Kumulatívne %
1	1314	38.68	1314	38.68
2	1006	29.61	2320	68.30
3	640	18.84	2960	87.14
4	334	9.83	3294	96.97
5	103	3.03	3397	100.00

Koláčové grafy

Koláčové grafy sú najčastejšie využívané na zobrazenie relatívnych frekvencií malého počtu hodnotených kategórií. Neodporúčajú sa ale, pokiaľ máme pomerne veľké množstvo jednotlivých kategórií. Koláčové grafy môžu byť tiež mäťúce, keď sa používajú na porovnanie výsledky dvoch rôznych prieskumov alebo dvoch rôznych experimentov. Ak sú koláčové grafy založené na malom počte pozorovaní, môže byť tiež nesprávne označovať jednotlivé časti grafu percentami (aj malá zmena početnosti v tomto prípade, môže v relatívnom vyjadrení predstavovať veľkú zobrazovanú zmenu).

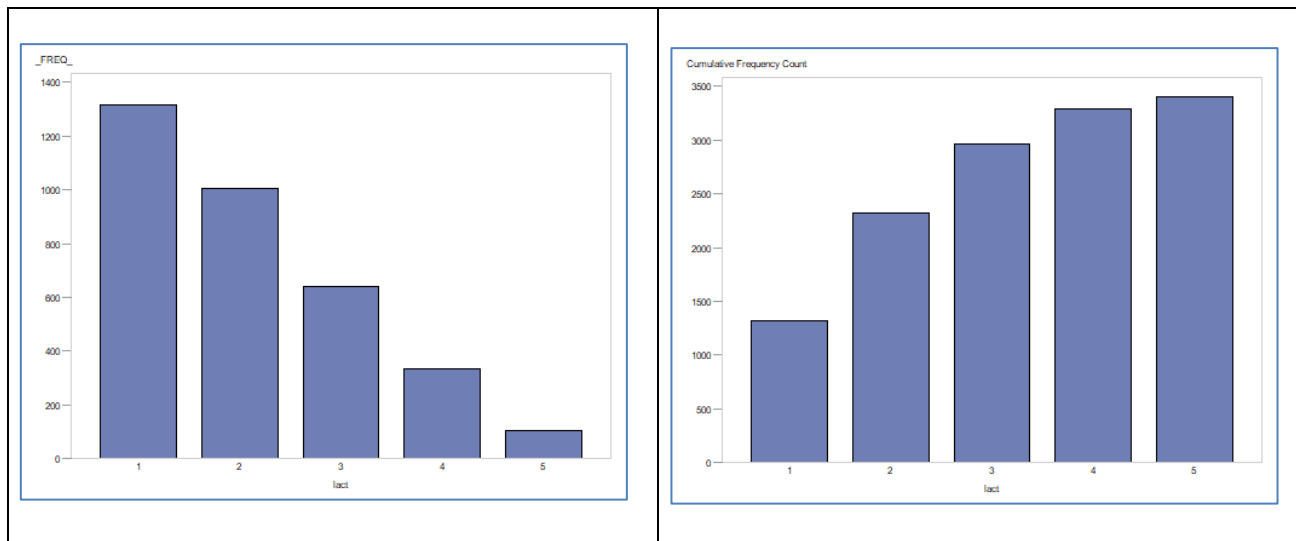
Príklady koláčových grafov:



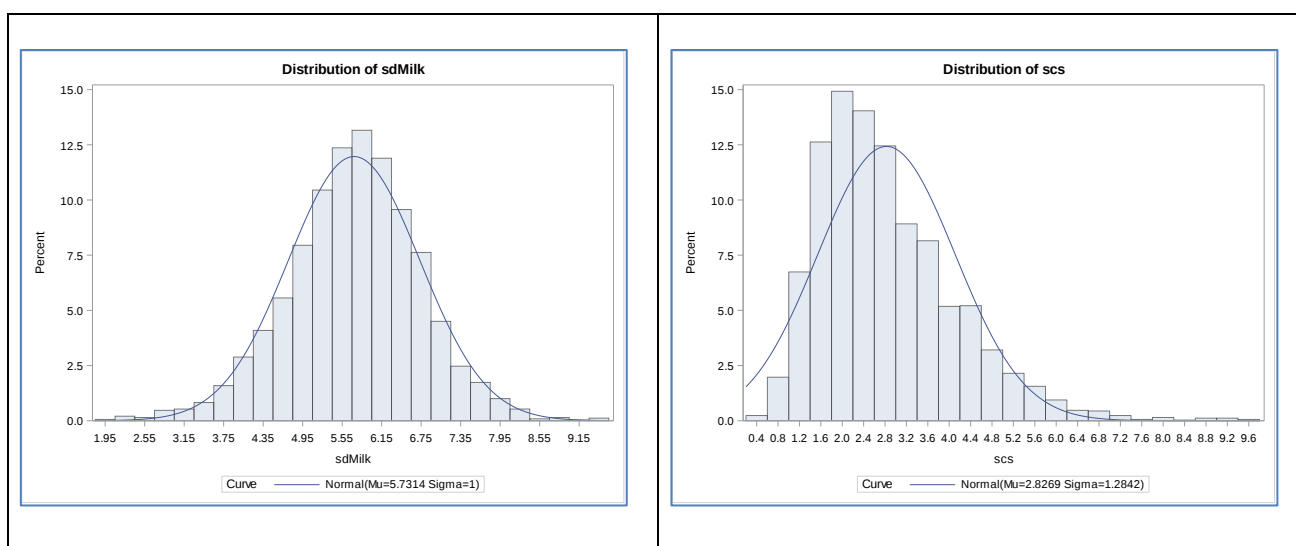
Stĺpcové grafy

Na znázornenie frekvencií rôznych kategórií možno použiť aj stĺpcové grafy. Stĺpcové grafy sú najviac používanými grafmi pri prezentácii výsledkov. Na osi Y sa zvyčajne zobrazuje počet pozorovaní v každej kategórii, nie percento pozorovaní v každej kategórii, ako je typické v koláčových grafoch. Mierne rozdielne stĺpcové grafy sú používané aj pri grafickom znázornení rozdelenia početnosti v rámci distribučnej analýzy vo forme histogramov. Stĺpcové grafy sú často vynikajúce na ilustráciu rozdielov medzi dvoma distribúciami. Histogramy sú spravidla zobrazované kontinuálne bez medzier medzi jednotlivými kategóriami.

Príklady stĺpcových grafov (frekvencie a kumulatívne frekvencie):



Príklady stĺpcových grafov využívaných pri distribučnej analýze:



Koláčové a stĺpcové grafy môžu byť účinnými metódami kvalitatívneho zobrazenia údajov. Stĺpcové grafy sú lepšie, ak existuje viac ako len niekoľko kategórií, alebo sú názornejšie pre porovnanie dvoch alebo viacerých distribúcií (rozdelení početností).

Grafy kvantitatívnych premenných

Kvantitatívne premenné sú premenné merané na číselnej škále. Výška, hmotnosť, doba odozvy nejakého javu, časová premenná, teplota, mlieková úžitkovosť, priemerný denný prírastok, prepočítané normálne skóre ukazovateľa sú príklady kvantitatívnych premenných. Kvantitatívne premenné sa odlišujú od kategorických (niekedy nazývaných

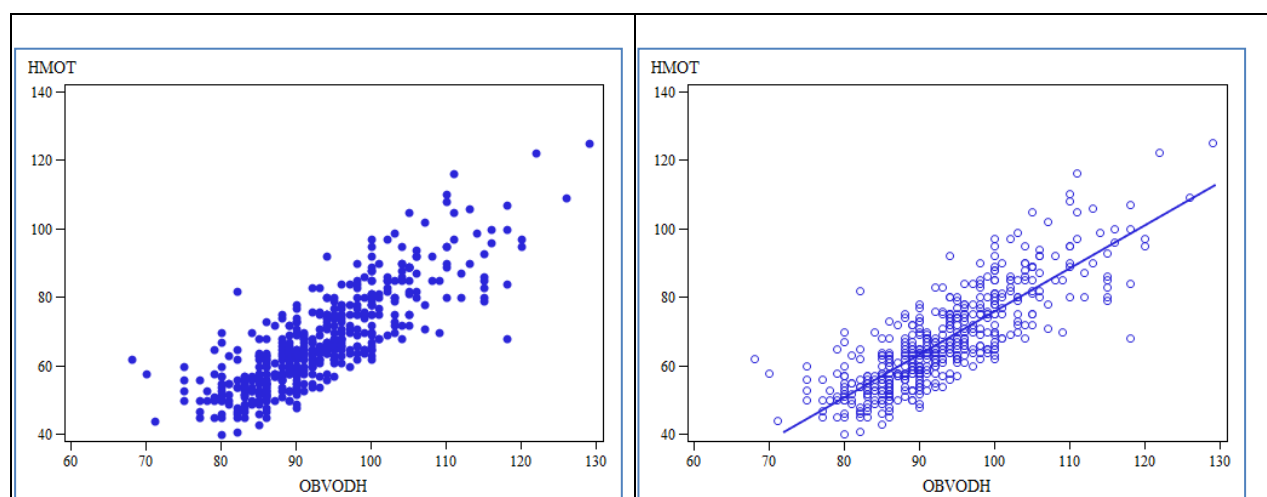
kvalitatívnych premenných) premenných. Existuje mnoho typov grafov, ktoré možno použiť na zobrazenie kvantitatívnych premenných.

Medzi najčastejšie používané grafy pri analýzach kvantitatívnych premenných patria: bodové grafy, čiarové grafy, stĺpcové grafy, krabicové grafy, grafy vo forme histogramov, grafy vo forme zložitejších polynómov.

Niektoré typy grafov sú najvhodnejšie pre malé až stredné množstvá údajov, zatiaľ čo iné, ako napríklad histogramy, sú najvhodnejšie pre veľké množstvo dát. Typy grafov, ako sú krabicové grafy, sú dobré pre zobrazovanie rozdielov medzi distribúciami v rámci rozdelenia početnosti znakov. Umožňujú názorne zobrazovať samotný rozptyl resp. variabilitu hodnotených znakov. Na znázornenie vzťahu medzi dvoma premennými sa používajú bodové a čiarové grafy (resp. krivkové grafy pri analýze nelineárnych vzťahov).

Bodové grafy

Bodové grafy zohrávajú mimoriadny význam pri posudzovaní a analýze vzťahov medzi jednotlivými dvojicami kvantitatívnych premenných. Môžu obsahovať prakticky všetky hodnotené pozorovania danej vlastnosti, čím umožňujú odhaliť chybové merania, alebo vybočujúce merania, ktoré môžu výrazne ovplyvniť samotnú analýzu. V prípade veľkých dátových súborov s veľkým počtom pozorovaní sa môžu ale stať neprehľadnými.



Často krát sa bodový graf dopĺňa priamkou, resp. krivkou, ktorá lepšie vyjadruje samotný analyzovaný vzťah. Z obľuby sa používajú tzv. trendové priamky, alebo krivky.

Bodové grafy môžu byť ľahšie pochopiteľné ako konvenčné stĺpcové grafy alebo histogramy.

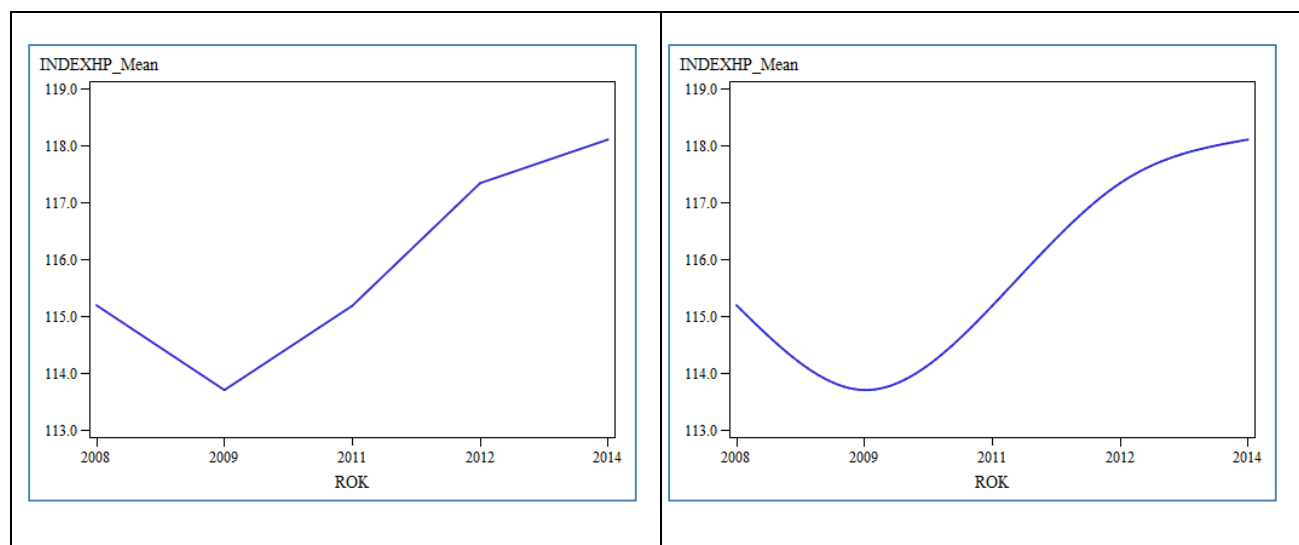
Bodový graf umožňuje presnejšie sledovať vzťah medzi dvoma bodmi ako stĺpcové grafy a je mimoriadne flexibilný.

Nevhodne zvolené mierky jednotlivých osí v rámci bodových grafov môžu neobjektívne zvýrazniť, alebo naopak potlačiť vyjadrenie a popis vzťahu medzi premennými.

Bodové grafy spolu so zobrazenými priamkami (krivkami) nachádzajú hlavné uplatnenie pri korelačných a regresných analýzach kvantitatívnych údajov a vlastností.

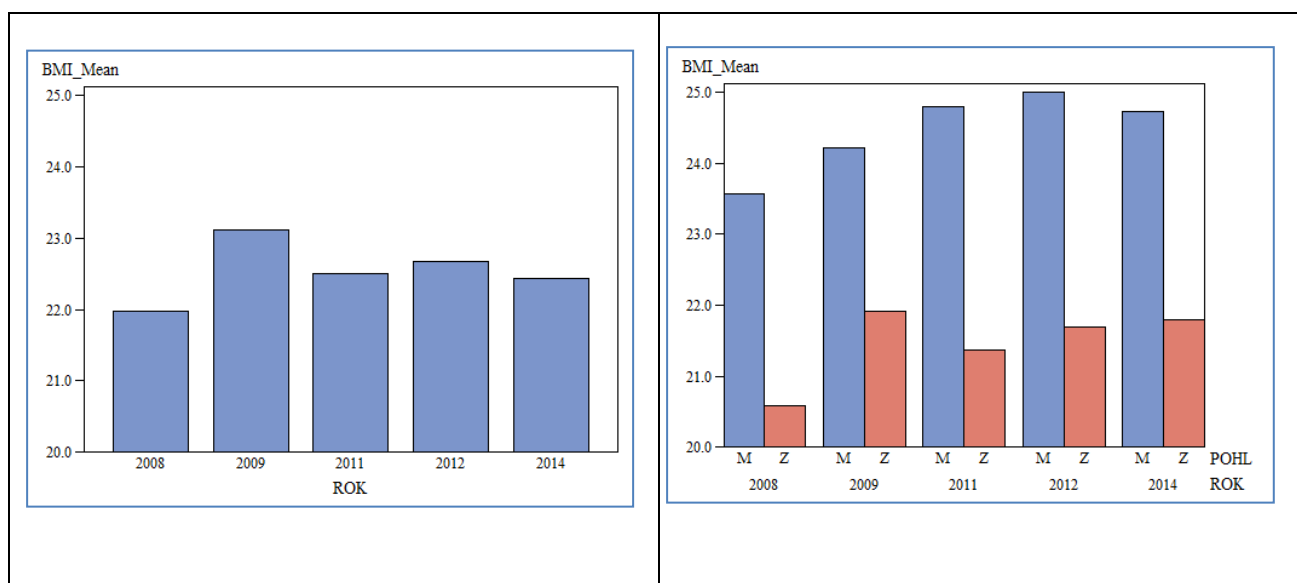
Čiarové grafy

Význam zostavovania čiarových grafov je hlavne pri popise a znázornení hodnôt kvantitatívnych premenných v závislosti na prebiehajúcej časovej premennej (rok, mesiac, deň, minúta, sekunda). Čiarový graf slúži k zobrazeniu trendov v priebehu času alebo medzi kategóriami. Napríklad, ak chceme zobraziť, ako sa priemerná hodnota indexu obvod hrudníka / obvod pásu menila za niekoľko rokov, môžeme použiť čiarový graf. Čiarové grafy sú ideálne pre základné posudzovanie a hľadanie lineárnych resp. nelineárnych závislostí medzi kvantitatívnymi premennými.



Stĺpcové grafy

Využitie stĺpcových grafov pri hodnotení kvantitatívnych premenných je rozdielne ako ich použitie pri znakoch kvalitatívnych, kde nás zaujímajú zvyčajne frekvencie jednotlivých kategórií. Stĺpcové grafy v tomto prípade najčastejšie znázorňujú priemerné hodnoty analyzovaných kvantitatívnych znakov v početne vhodne zvolených triedeniach. Stĺpcové grafy nachádzajú uplatnenie aj v rámci skupinových porovnaní. Ako príklad uvádzame stĺpcový graf priemerných hodnôt BMI (Body Mass Index) v triedení podľa pohlavia študentov.

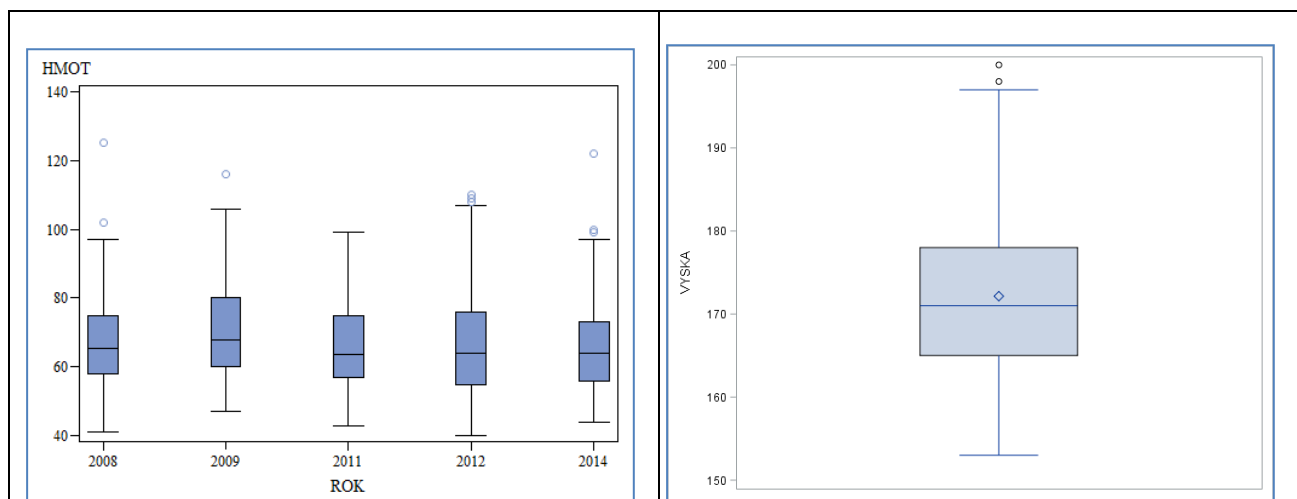


Krabicové grafy

Krabicové grafy sa často využívajú v popisnej štatistike na znázornenie viacerých popisných charakteristík súčasne (charakteristiky polohy aj charakteristiky variability). Krabicové grafy sú užitočné na identifikáciu odľahlých hodnôt a pre porovnanie jednotlivých rozdelení početností. Zobrazenie krabicových grafov môže byť vo forme paralelného zobrazenia (ak súčasne porovnávame viaceré kategórie, alebo triedenia), alebo vo forme individuálneho zobrazenia, ak hodnotíme jednu vlastnosť za celý výberový, alebo základný súbor údajov.

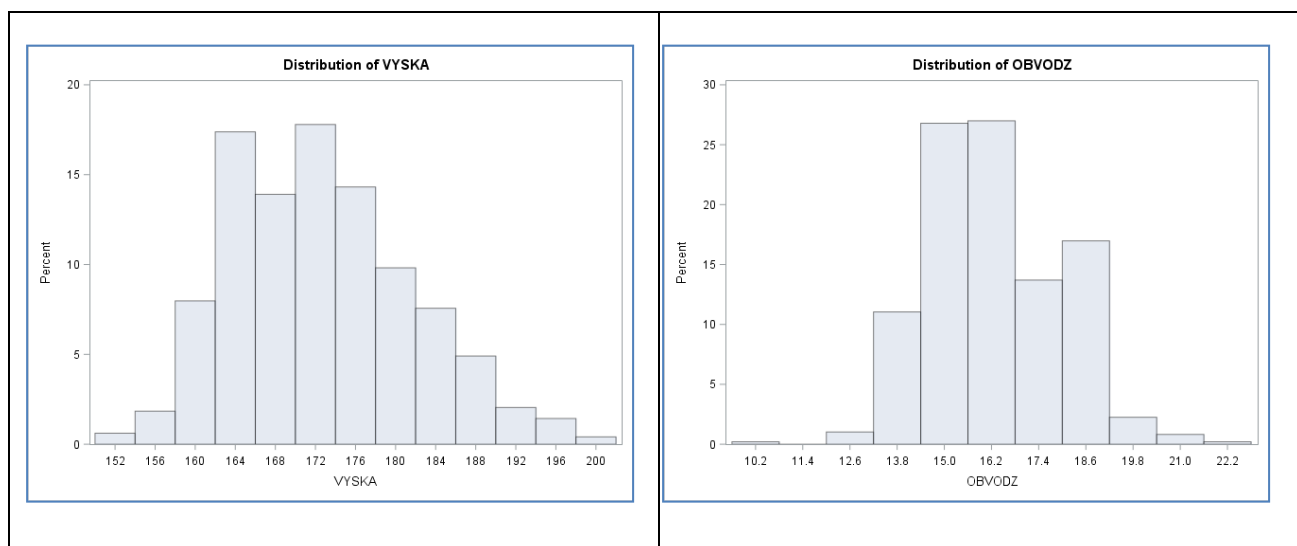
V zobrazeniach krabicových grafov je možné nastaviť: $\pm$ násobky medzikvartilového rozsahu čo vykreslí zvislé čiary od rámčeka po najextrémnejší bod v rámci stanovených medzikvartilových rozsahov,

extrémy najvyšších a najnižších hodnôt, alebo jednotlivé percentily (1. a 99. percentil, 5. a 95. percentil, 10. a 90. percentil, 25. a 75. percentil).



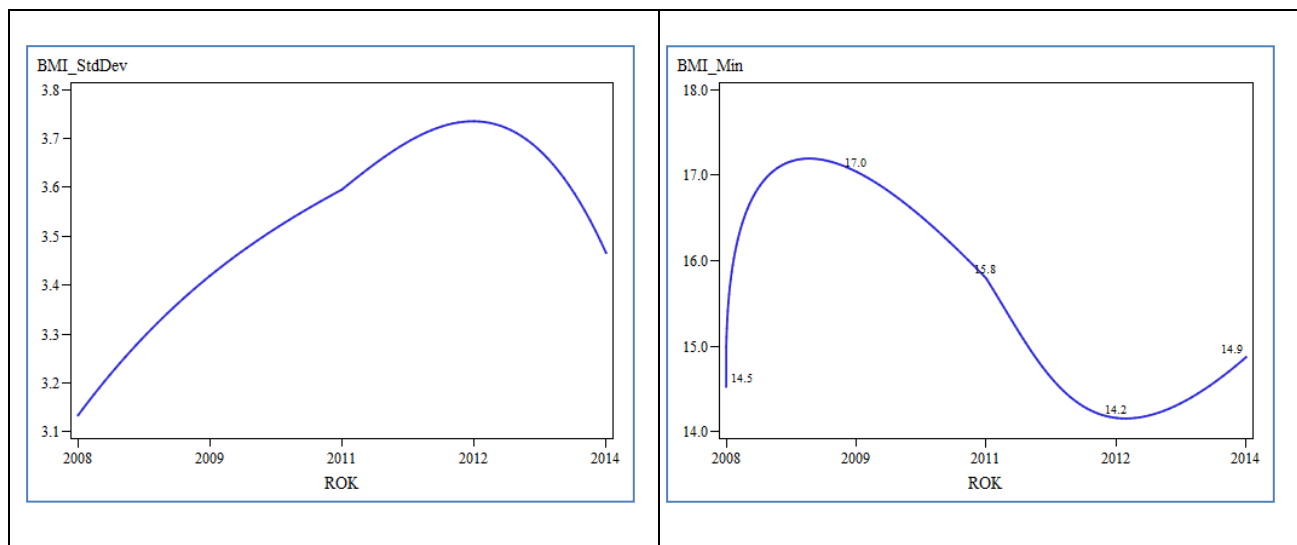
Histogramy

Histogramy sú grafickým vyjadrením tvaru rozloženia početnosti kvantitatívneho ukazovateľa. Majú význam hlavne pri veľkom počte pozorovaní. Ich použitie v rámci početne malých údajových súborov ale môže byť menej presné. V biologických experimentoch sú histogramy najčastejšie zostavované z dôvodu skúmania normality údajov (potvrdenie alebo zamietnutie normálneho rozdelenia početnosti).



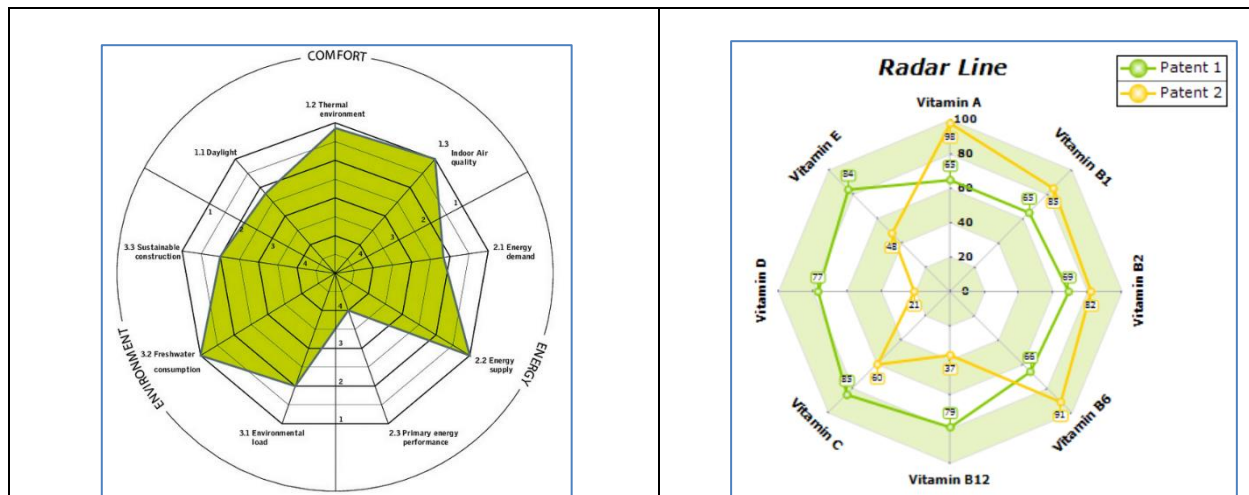
Grafy vo forme polynómov

Grafy vo forme polynómov sú v skutočnosti zložitejšie čiarové grafy využívané hlavne pri nelineárnych závislostiach. V nasledujúcich obrázkoch uvádzame **Legendreove polynómy**, ktoré vyhladzujú priebeh skúmaného javu, alebo vlastnosti na základe zložitejších nelineárnych vzťahov. Legendreove polynómy sú často využívané pri modelovaní laktačných kriviek hovädzieho dobytku v rámci genetických hodnotení mliekovej úžitkovosti. Uvádzame príklady ich použitia pri modelovaní vývoja priemerných a minimálnych hodnôt indexu BMI v ľudskej populácii.



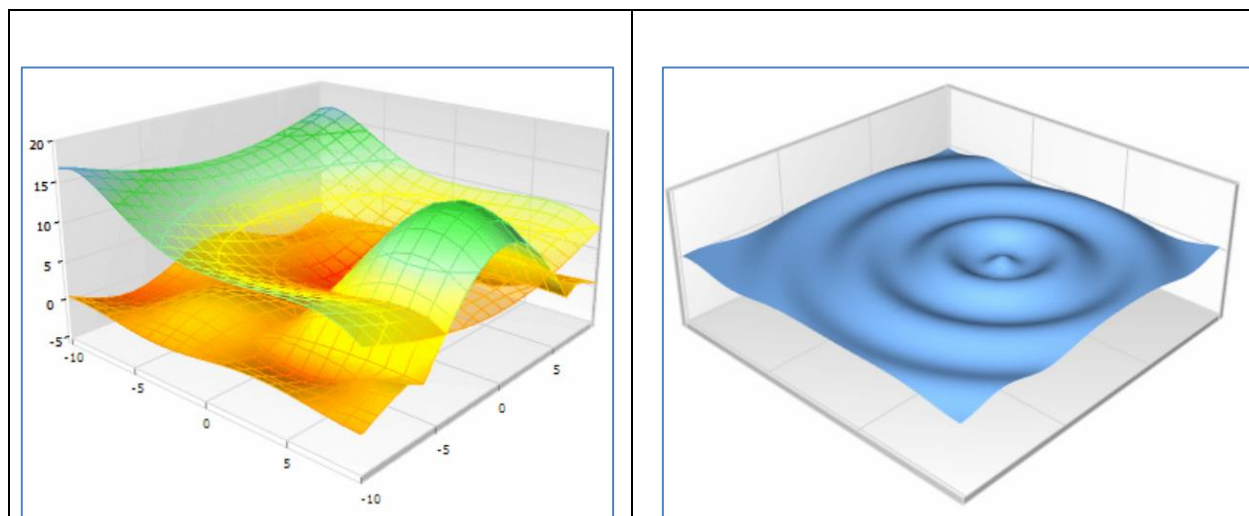
Radarové grafy

Radarové grafy vyjadrujú relatívnu polohu údajových bodov voči stredu. Počet údajových bodov, alebo skupina údajových bodov určuje počet osí hodnôt a hodnota bodu potom určuje jeho vzdialenosť od stredu. Každá skupina má vlastnú os hodnôt, ktorá vychádza zo stredu. Poradie jednotlivých osí je v smere otáčania hodinových ručičiek. Všetky hodnoty v rovnakom údajovom rade sú spojené čiarami. Pri záporných hodnotách je najnižšia hodnota v strede osí. Radarové grafy sú používané pre znázornenie symetrie a pravidelnosti údajov jedného alebo niekoľkých údajových radov. Majú široké využitie hlavne v technických disciplínach, ale nachádzame ich aj v medicínskom a biologickom výskume.



Povrchové grafy

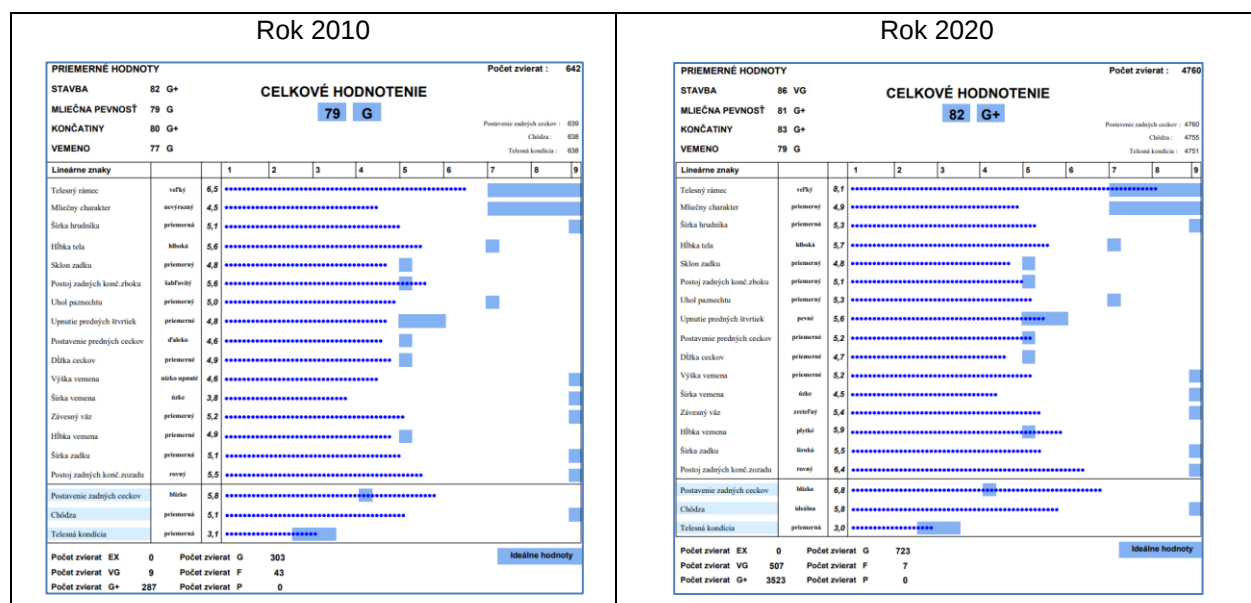
Povrchové grafy vhodné pre nájdenie optimálnych kombinácií dvoch údajových radov. Rovnaké farby a vzorky v grafe predstavujú rovnaké rozsahy hodnôt. Zdrojová údajová tabuľka musí obsahovať aspoň dva rady. Jednotlivé hodnoty sú zobrazené vo formách vrstiev. Existujú, ale aj špeciálne povrchové jednovrstvové grafy.



Špeciálne grafy (grafy exteriérových profilov)

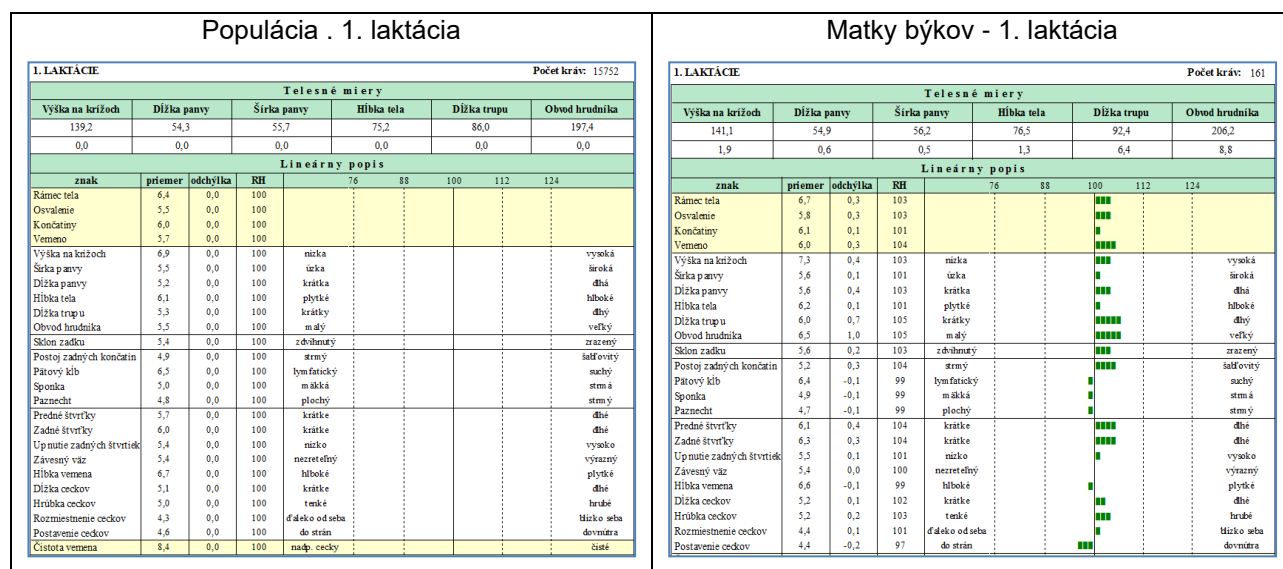
Pri hodnotení exteriéru zvierat sa z obľubou zostavujú grafy tzv. exteriérových profilov, ktoré vyjadrujú vývoj exteriéru (zovňajšku) konkrétneho jedinca, alebo skupiny jedincov oproti nastaveným priemerným resp. ideálnym hodnotám. V nasledujúcom obrázku 10 uvádzame grafické vyjadrenie zmeny exteriérových profilov lineárneho

hodnotenia exteriéru kráv holštajnského plemena v Slovenskej republike za posledných 10 rokov (reálny stav v hodnotenej populácii kráv oproti ideálnym hodnotám).



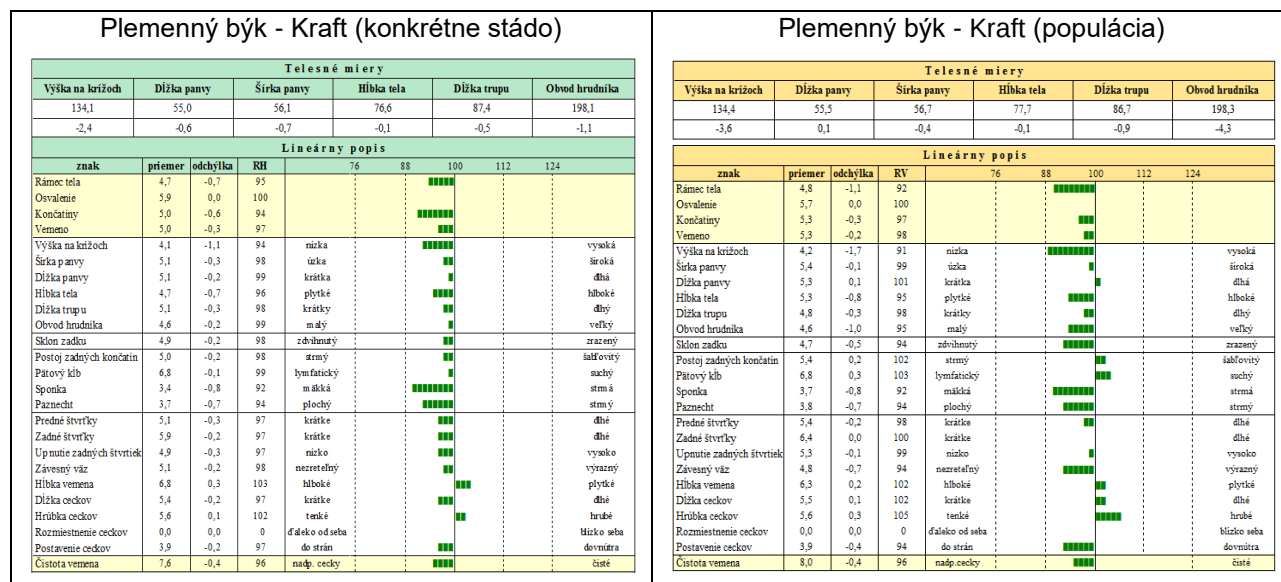
Obr. 10 Zmena exteriérových profilov (zdroj: databáza projektu KEGA)

Iným typom grafov exteriérových profilov sú exteriérové profily v populácii kráv slovenského strakatého plemena pri ktorých je znázornená odchýlka hodnotených znakov oproti aktuálnemu priemeru týchto znakov vo vybranej časti populácie (obrázok 11).



Obr. 11 Exteriérový profil matiek býkov (zdroj: databáza projektu KEGA)

V obrázku 12 uvádzame grafické porovnanie exteriérového profilu plemenného býka na úrovni konkrétneho stáda a na celopopulačnej úrovni.



Obr. 12 Exteriérový profil býka Kraft (zdroj: databáza projektu KEGA)

Grafické vyjadrenie exteriérových profilov znázorňuje drobné odchýlky porovnania vývoja exteriéru hodnotených dcér plemenníka Kraft po druhej laktácii na úrovni konkrétneho stáda v porovnaní s jeho výsledkami na úrovni celej populácie.

<https://www.asb.sk/architektura/rodinne-domy-architektura/pasivne-domy/prvy-certifikovany-aktivny-dom-na-svete/attachment/9485-5b36831eb54e4>

<https://www.nevron.com/products-reporting-services-chart-gallery-charting-types-radar-chart.aspx>