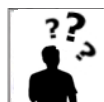


# Soutěž: Nahrad'te nuly

Newsletter Statistica ACADEMY



Téma: Možnosti softwaru  
Typ článku: Soutěž

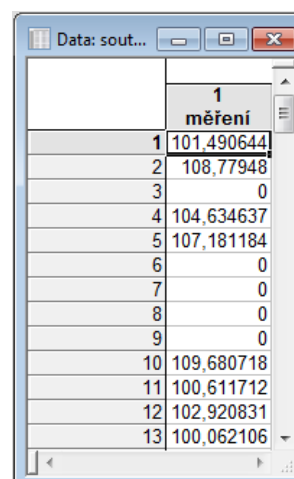
*Jistě máte rádi soutěže, při kterých se můžete něco naučit nebo třeba objevit nečekané funkcionality softwaru. Tahle soutěž patří do této kategorie.*

Máme k dispozici následující **soubor s jednou proměnnou**. V tomto souboru hodnoty 0 znamenají, že jsme data nenaměřili. Jediné, co tedy chceme s tímto souborem udělat, je nahradit hodnoty 0 chybějícími hodnotami.

Vaším úkolem je najít a zaslat nám dva postupy, které dokáží tento úkol provést. Je tedy potřeba udělat daný úkol dvakrát pokaždé využitím trochu jiných funkcionalit, přičemž jednoduché ruční vymazání hodnot nepočítáme.

Věřte, že možných řešení je - jak už jste si u Statisticy mohli zvyknout - více než jen dvě.

Vaši odpověď nám zašlete na adresu [soutez@statistica.cz](mailto:soutez@statistica.cz) nejpozději do **6.12.2015**.



|    | 1 měření   |
|----|------------|
| 1  | 101,490644 |
| 2  | 108,77948  |
| 3  | 0          |
| 4  | 104,634637 |
| 5  | 107,181184 |
| 6  | 0          |
| 7  | 0          |
| 8  | 0          |
| 9  | 0          |
| 10 | 109,680718 |
| 11 | 100,611712 |
| 12 | 102,920831 |
| 13 | 100,062106 |

## Podmínky soutěže

Výherce získává voucher na kurz dle vlastního výběru. Voucher je platný jeden rok od vystavení a je vázán pouze na jméno vylosovaného výherce. Voucher slouží jako volný vstup pro výherce na jeden z nabízených kurzů na stránkách <http://www.statsoft.cz/sluzby/1-kurzy-skoleni/>. Výherce je povinen přihlásit se na jím vybraný kurz nejméně 14 dní před termínem konání vybraného kurzu. Podmínkou účasti je, že vybraný kurz bude otevřen a nebude plně obsazen. Voucher se nevztahuje na výcvik Six Sigma Green Belt.

## Minulá soutěžní otázka – správná odpověď

Správná odpověď [na předchozí otázku](#) je b). Test se souvisí se jmény Aspinová – Welch.

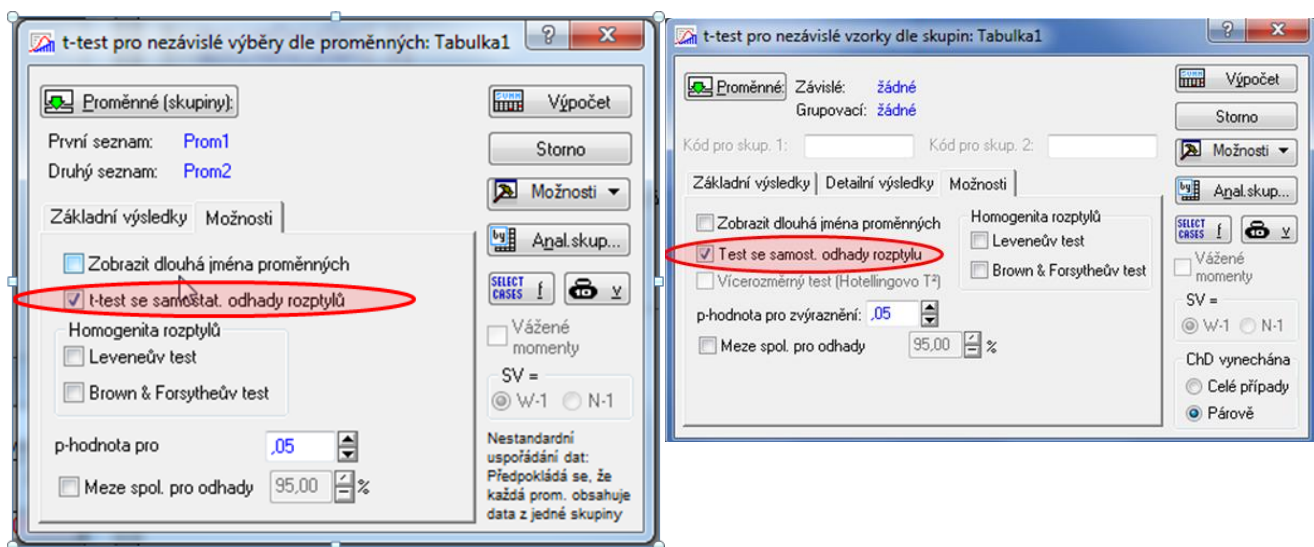
Trochu informací o tomto testu:

**Aspinové-Welschův test** (v nápovědě Statistica zmiňován jako Welschův t-test)

- *Předpoklady:* 2 výběry z normálního rozdělení  $N(\mu, \sigma_i^2)$ , různé rozsahy, nezávislé.
- *Popis:* Alternativa dvouvýběrového t testu pro data s nesterjnými rozptyly, počítá odhady rozptylu pro výběry zvlášť, pro stupně volnosti t rozdělení se využívá speciální vzoreček.
- *Vzorec:* testová statistika:  $T^* = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S}$ ,  $S = \sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}$ ,  $\bar{X}$  a  $\bar{Y}$  jsou výběrové průměry,  $S_X^2$  a  $S_Y^2$  jsou výběrové rozptyly. Úprava pro nesterjné rozptyly pak spočívá a tom, vůči čemu se testuje: Test zamítá, pokud  $|T^*| \geq t_f(\alpha)$ , kde  $t_f(\alpha)$  je kritická hodnota t rozdělení s  $f$  stupni volnosti  $f = \frac{s^4}{\frac{S_X^4}{(m-1)m^2} + \frac{S_Y^4}{(n-1)n^2}}$ .

Více informací viz: Jiří Anděl – Statistické metody, str. 98.

- *Příkaz v R:* t.test(..., var.equal = FALSE)
- Kde jej najdeme v softwaru:



## Minulá soutěžní otázka – výherce

Správně odpovědělo mnoho z Vás, vylosovaným výhercem je pan **Petr Snopek**. Gratulujeme výherci volného vstupu na jeden z našich kurzů.