

Analýza a vizualizácia dát

6. mája 2024

VBA

- ▶ Visual basic
- ▶ programovanie v E
- ▶ automatizácia, zautomatizovanie často opakovaných činností
- ▶ vytváranie funkcií
- ▶ jedná sa o objektovo orientované programovanie

zapnutie

- ▶ *File* → *Options* → *Customize Ribbon* → *Developer*
- ▶ *Developer* záložka
- ▶ *Record Macro*

cvičenie

- ▶ vytvorte (nahrajte) pomocou *Record Macro* jednoduché makro, ktoré napíše reťazec do bunky A1 a zmení font písma, farbu pozadia a veľkosť písma
- ▶ sledujte, čo spraví vytvorené makro na novom hárku
- ▶ upravte makro, aby v ňom nezostali nepotrebné príkazy

cvičenie

- ▶ vytvorte makro, ktoré vypíše číslo e do vyskakovacieho okna pomocou *MsgBox*
- ▶ definovanie premennej: *Dim*
- ▶ typy premenných:
 - *Integer* - celé číslo, krátky rozsah
 - *Long* - celé číslo, dlhý rozsah
 - *Single* - desatinné číslo, krátky rozsah
 - *Double* - desatinné číslo, dlhý rozsah
 - *String* - reťazec
 - *Variant* - hocičo, zaberá veľa miesta v pamäti
 - *Range* - bunka, riadok, stĺpec, súbor buniek (aj 3D)

procedúry

- ▶ *Option Explicit* - na začiatku kódu pred procedúrami, núti deklarováť premenné
- ▶ premenné sa dajú deklarováť vrámci procedúry alebo globálne
- ▶ procedúry sa dajú písať v každom hárku lokálne, v zošite globálne, do modulov

prístup k bunkám

- ▶ *Range("A1"), Range("A1:A10")*
- ▶ *ActiveCell.Value* - hodnota v aktívnej bunke
- ▶ *ActiveCell(1,2)* - bunka v pozícii (1,2) vzhľadom na aktívnu bunku (aktívna bunka má pozíciu (0,0))
- ▶ *.Offset(rowOffset:=3, columnOffset:=3).Activate* - aktivuje bunku v 3. stĺpci napravo a v 3. riadku dolu od momentálne aktívnej bunky
- ▶ *Cells(1,1).Select* - selektuje bunku "A1"
- ▶ *Columns("A:F"), Rows("1:10")*
- ▶ *.FormulaR1C1* používa *R[-1]C[2]* vo formulách (bunka o 1 hore a 2 napravo)

cvičenie

- ▶ vytvorte makro, ktoré zmení font hlavičky na *Times New Roman* a zmení hlavičku, aby bola boldom
- ▶ pokiaľ ste makro v predchádzajúcom kroku nahrali, vymažte z kódu všetky zbytočné príkazy
- ▶ vytvorte makro, ktoré automaticky nastaví optimálnu šírku všetkých stĺpcov a riadkov tabuľky
- ▶ vytvorte makro, ktoré automaticky nastaví optimálnu šírku všetkých stĺpcov a riadkov celého hárku

pozor

- ▶ čo sa zmení cez VBA, to sa nedá vrátiť, oplatí sa mať zálohované pôvodné dáta, resp. kopírovať hárky
- ▶ pri zložitejších makrách pomáha pred nahrávaním zapísať si najskôr, čo chceme nahráť

- ▶ *Selection.End(xlToRight), xlDown, xlUp, xlToLeft*
- ▶ *Selection.End(xlToRight)*

button

- ▶ *Developer* → *Insert* → *Button*
- ▶ *Form Controls* sú jednoduchšie, pri *ActiveX Controls* sú vznešenejšie možnosti napr. po vizuálnej stránke tlačidla (a trochu iné možnosti)
- ▶ môžeme mu priradiť makro, ktoré sa spustí po stlačení

samostatné cvičenie

- ▶ vytvorte makro, ktoré doplní do dát ďalší riadok (vypýta si dáta), pričom počet riadkov dát je neznámy/variabilný

samostatné cvičenie

- ▶ vytvorte makro *Hlavicka*, ktoré vytvorí v aktívnom hárku hlavičku obsahujúcu demografické údaje
 - ID
 - Meno
 - Priezvisko
 - Pohlavie
 - Veknastaví farbu pozadia ako bledomodrú (*Range("A1").Interior.ColorIndex*) a následne nastaví aktívnu bunku ako prvú bunku ďalšieho riadka
- ▶ vytvorte makro *ZistiUdaje*, ktoré pomocou príkazu *InputBox* zistí tieto údaje a vypíše ich do aktívneho riadku a následne nastaví aktívnu bunku ako prvú bunku ďalšieho riadka

samostatné cvičenie

- ▶ vytvorte makro, ktoré doplní do dát ďalší riadok, ktorý bude zhodný s posledným riadkom dát (pričom počet riadkov dát je neznámy/variabilný)

decisions

- ▶ *if...then...end if*
- ▶ *if...then...else...end if*
- ▶ *if...then...elseif...then...elseif...then...end if*

loops

- ▶ *for...next*
- ▶ *for each...next*
- ▶ *while...wend*
- ▶ *do while...loop*
- ▶ *do until...loop*

cvičenie

- ▶ prekódovať *muž/žena* v dátach na *0/1*
- ▶ zmeniť pozadie všetkých buniek s *NA* v dátach na červené
- ▶ vytvorte nový hárok, do ktorého sa skopírujú upravené dáta tak, aby:
 - *kod* bol celé číslo
 - *pohlavie* bolo kódované *0/1*
 - *vyska, vzdelanie, rok narodenia, vek* zostali nezmenené
 - bol pridaný nový stĺpec *aktualny.vek*, ktorý by dopočítal aktuálny vek respondenta pričom predpokladáme, že jeho dátum narodenia je 1.1. daného roka

samostatné cvičenie

- ▶ vytvorte nový hárok, kde sa skopírujú upravené dáta tak, aby:
 - sa prekopírovali iba respondenti narodení po roku 1970 - *kod* bol *Pozorovanie* + číslo respondenta
 - *pohlavie* bolo farebne odlišené

Objekty

- ▶ Application, Workbook, Worksheet, Range Objects
- ▶ vzťahujú sa na ne metódy, ktoré vracajú objekty, napr.
 - *Workbooks.Open Filename:="NovyWB.xlsx"*
 - *Worksheets("Sheet1").Select*
 - *Worksheets("Sheet1").Range("A2:A3").Value=0*

cvičenie

- ▶ vytvorte makro, ktoré vytvorí nový workbook s názvom *NovyWB.xlsx*
- ▶ vytvorte makro, ktoré tento nový pracovný zošit otvorí
- ▶ vytvorte makro, ktoré vytvorí nový hárok s názvom *NovyHarok*

Funkcie

- ▶ vo VBA nie sú implementované všetky funkcie, ktoré poznáme z E
- ▶ vieme ich zavolať
- ▶ *WorksheetFunction.Max("A1:A10")*

cvičenie

- ▶ vytvorte makro, ktoré identifikuje maximum pre každý stĺpec a zmení štýl formátovania buniek s touto hodnotou zo štýlu *Normal* na štýl *Good*

cvičenie

- ▶ vytvorte makro, ktoré
 - vytvorí nový hárok s názvom *NajdiNA* a zaradí ho ako prvý v poradí
 - následne prehľadá všetky hárky okrem prvého a vyhladá v nich *NA* hodnoty
 - zaznamená do hárku *NajdiNA* čísla pacientov s chýbajúcou hodnotou do stĺpca, kde tá hodnota chýba

array

- ▶ *Dim arr()* - nešpecifikovanej veľkosti
- ▶ *Dim arr(6)* - veľkosti 6, ale index začína od 0, teda sa doňho zmestí 7 hodnôt
- ▶ hocijaký typ
- ▶ *Dim arr(3,4)* - viacrozmerný, 3 riadky, 4 stĺpce
- ▶ fcie vyťahujúce sa na array napr.: *LBound, UBound, Split, Join, IsArray, Erase,...*

štandardizovanie dát

- ▶ dáta sa nachádzajú približne v okolí ich priemeru

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

- ▶ ich "rozkákanosť" okolo neho je popísaná (výberovým) rozptylom

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

- ▶ veľa štatistických metód využíva tzv. štandardizované dáta, kedy je priemer 0 a výberový rozptyl 1, čo docielime transformáciou

$$\frac{x_i - \bar{x}}{s}$$

cvičenie

- ▶ vytvorte makro, ktoré si ako vstup vypýta cez *InputBox* dáta, ktoré následne štandardizuje a vypíše do stĺpca, ktorého pozíciu si vypýta

procedúry spustené akciou

- ▶ *Sub Workbook_Open*
- ▶ *Workbook* - rozhranie
- ▶ *Open* - udalosť
- ▶ iné udalosti: *NewSheet (ByVal Sh As Object), BeforeSave (ByVal SaveAsUI As Boolean, Cancel As Boolean),...*

cvičenie

- ▶ vytvorte procedúru, ktorá vypíše "Vitajte!" pri spustení

funkcie

- ▶ *Function ... End Function*
- ▶ dajú sa volať aj v iných workbookoch
- ▶ uložiť ako *Excel Add-in*
- ▶ *Developer* → *Excel Add-ins* → *Browse*

cvičenie

- ▶ vytvorte funkciu, ktorá bude počítať BMI

samostatné cvičenie

- ▶ vytvorte funkciu, ktorá prevedie centimetre na stopy a otestujte ju na dátach
- ▶ vytvorte funkciu, ktorá prevedie kilogramy na libry a otestujte ju na dátach
- ▶ upravte procedúru *Standardizuj*, ktorá štandardizuje dáta, aby si nepýtala číslo stĺpca, ale aby vložila nový stĺpec vedľa štandardizovaného a tam vypísala dáta (ostatné stĺpce sa posunú doprava)