

ANALYZING BIG TABLES AT SPREADSHEET EXCEL

Agneša GAŠPERANOVÁ

Abstract: The electronic spreadsheet is basic tool for business modeling, simulation and decision making. Important part of spreadsheet using is processing large tables, especially computing subtotals. This paper presents the course of creating subtotals, witch author used for teach-in Excel at academic year 2009/2010 at The Faculty of National Economy, University of Economics in Bratislava.

Key words: electronic spreadsheet, subtotals, pivot tables, database functions

ANALÝZA ROZSIAHLÝCH TABULIEK V EXCELI

Resumé: Tabuľkový kalkulátor je základný pracovný nástroj dátovej analýzy, ktorý sa využíva pre ekonomické modelovanie, simuláciu, podporu rozhodovania. Významnou oblasťou využívania tabuľkového kalkulátora je spracovanie veľkých tabuliek, hlavne výpočet medzisúčtov. Článok opisuje kurz tvorby medzisúčtov, ktorý autorka využívala na cvičeniach Excel v akademickom roku 2009/2010 na Národohospodárskej fakulte Ekonomickej univerzity v Bratislave.

Kľúčové slová: tabuľkový kalkulátor, medzisúčty, kontingenčná tabuľka, databázové funkcie

1 Úvod

Tabuľkový kalkulátor Excel je jedným z najdôležitejších aplikácií na riešenie úloh v oblasti administratívy a manažmentu. Študenti Národohospodárskej fakulty Ekonomickej univerzity v Bratislave absolvujú v druhom semestri štúdiu cvičenie, v ktorom by mali získať zručnosti pokročilého alebo aspoň mierne pokročilého používateľa tabuľkového kalkulátora Excel. Nestačí však, aby študenti zvládli použitie konkrétnych nástrojov, musia pochopiť ich význam a možnosti využitia. Skúsenosti pri testovaní ukázali, že úspešnosť študentov závisí nielen od ich vedomostí, ale ja od toho, ako je úloha sformulovaná. Výsledky sú lepšie, ak zadanie obsahuje určitý návod na riešenie. Študenti, ktorí ovládajú konkrétny pracovný nástroj, vyriešia zadanie bez väčších problémov. Dokazuje to vyhodnotenie dvoch testov na spracovanie rozsiahlych tabuliek, ktoré študenti absolvovali v letnom semestri akademického roku 2008/2009. Prvý test na zistenie medzisúčtov obsahuje v zadani používaný prostriedok (kontingenčná tabuľka s opisom štruktúry tabuľky), v druhom teste návod chýba (analyzujte údaje, zistite súčet), úlohu môže študent riešiť ľubovoľným nástrojom, ale musí ho nájsť sám. Podmienkou úspešnosti bolo získanie viac ako 50 % bodov. Testy riešilo 117 študentov, prvý test úspešne vyriešilo 78 %, druhý iba 48 % študentov.

Výučba pokročilých nástrojov tabuľkového kalkulátora nastoľuje niekoľko problémov. Prvý problém je hľadanie vhodných príkladov na demonštráciu, druhý problém je nájdenie vhodného nástroja na riešenie problému, prípadne porovnanie nástrojov, ak sa úloha dá riešiť rôznymi prostriedkami, čo je typické hlavne pri spracovaní rozsiahlych tabuliek.

Klasický postup pri výučbe je vytvorenie vhodného príkladu pre každý nástroj osobitne, čo je jednoduchšie, lebo tabuľka musí splniť menej podmienok. V akademickom roku 2009/2010 sme zmenili prístup pri vyučovaní nástrojov na analýzu a spracovanie údajov, kládli sme väčší dôraz na význam použitých prostriedkov, snažili sme vysvetliť princíp, nielen postup riešenia. Poukázali sme aj na možnosť získania výsledkov rôznymi postupmi, ak to bolo v úlohe možné, analyzovali sme alternatívne riešenia, porovnávali jednotlivé prostriedky.

Vytvorenie demonštračných tabuliek bolo náročnejšie, pretože sme ich spracovali rôznymi nástrojmi. Museli sme vytvoriť veľké tabuľky s obsahom, ktorý je vhodný na demonštráciu rôznych nástrojov dátovej analýzy. Záverečné testovanie ukázalo, že výsledky študentov sa zlepšili, úlohu úspešne vyriešilo 60 % študentov, aj keď si študenti museli zvoliť postup sami, v zadani chýbal návod na riešenie.

V článku uvádzame príklady, ktoré tvoria štvrtinu rozsahu semestrálnej látky cvičenia.

2 Kumulatívne výpočty- minikurz Excel

Tabuľkový kalkulátor poskytuje účinné nástroje na spracovanie rozsiahlych tabuliek. Analýza údajov je dôležitou stránkou využívania informácií, uložených v tabuľkovej forme. Funkcie a databázové nástroje umožnia získať prehľadné informácie jednoducho a rýchlo. V článku uvedieme príklady na vyhľadávanie informácií a efektívne vytváranie tzv. medzisúčtov v tabuľkách. Medzisúčty systémov na spracovanie hromadných údajov označujú prostriedky na získanie kumulatívnych údajov, okrem súčtov napríklad aj maximum, minimum, početnosť a priemer hodnôt v skupine údajov tabuľky, pričom skupinu tvoria riadky so zhodnou hodnotou kľúčovej položky, stĺpca.

Využívané prostriedky

Prostriedky a postupy, používané pri analýze údajov, uložených v tabuľke, závisia od množiny skúmaných informácií a od požiadaviek, ktoré na výsledky kladieme. Ponuka prostriedkov v tabuľkovom procesore je bohatá, tie isté výsledky môžeme niekedy získať rôznymi postupmi. V článku použijeme na riešenie úloh databázové nástroje (filter, medzisúčty, kontingenčné tabuľky), ale aj databázové funkcie a vybrané štatistické a matematické funkcie. Pokúsime sa ukázať rôzne postupy riešenia tej istej úlohy, porovnať ich náročnosť, výhody a nevýhody pre konkrétny príklad.

Úlohy riešime na demonštračnom príklade, využívame databázovú tabuľku, ktorá obsahuje prehľad výkonov servisnej organizácie, ktorá zabezpečuje služby pre zákazníkov v desiatich krajinách Európskej únie. Faktúry vybraných služieb sledujeme v priebehu 12 mesiacov, od 1.5.2009 do 30.4.2010, tabuľka obsahuje 696 riadkov. Našou úlohou bude spočítať výšku faktúr podľa rôznych hľadísk (krajina, predajca, dátum). Prvé riadky tabuľky sú na obrázku 1.

krajina	meno	dátum	kód služby	faktúra
DE	Decker	3.5.2009	BC 369521	485,00 €
AT	Kanev	4.5.2009	HZ 258369	98,00 €
HU	Karacs	5.5.2009	GY 147963	686,00 €

Obr. 1: Štruktúra tabuľky

3 Možnosti sčítovania údajov v tabuľke

V texte uvedieme niekoľko príkladov sčítovania údajov. Počet, priemer, maximum, minimum získame podobne, väčšinou stačí zvoliť pri definovaní zadania vhodnú funkciu, respektíve upraviť argumenty funkcie. Položka,

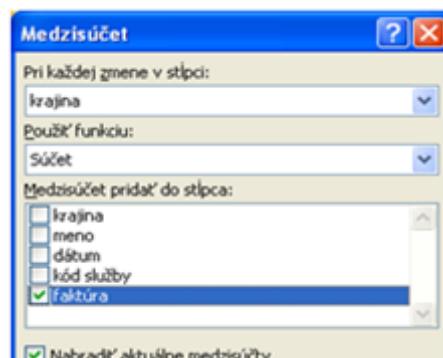
ktorá určí skupinu, je kľúč výberu riadkov na tvorbu medzisúčtu. Výsledkom bude súčet hodnôt vybraného stĺpca tabuľky z tých riadkov, ktoré majú rovnakú kľúčovú položku. Príklady riešime v Excel 2007.

Označenie rozsiahlej tabuľky je niekedy problematické. Keď tabuľku označíme pred začatím operácie, pomôžeme si klávesovými skratkami, presunieme kurzor do tabuľky a stlačíme kombináciu klávesov Ctrl+A, alebo Ctrl+*. Databázové nástroje nevyžadujú označenie tabuľky, stačí pred výberom príkazu presunúť kurzor do tabuľky, Excel po zadaní príkazu rozšíri označenie na celú tabuľku. Ťažkosti môžu nastať, ak databáza je argument funkcie, vtedy tieto pomôcky nefungujú. Riešenie je použitie názvov, ktoré spravujeme skupinou nástrojov *definované názvy* v príkaze vzorce. Po označení tabuľky príkaz *definovať názov* pomenuje celú tabuľku, príkaz *vytvoriť názov z výberu* pomenuje stĺpce a/alebo riadky tabuľky. V našom príklade *databáza faktúr* je názov tabuľky, menovky stĺpcov označujú stĺpce tabuľky (hodnoty stĺpca bez menovky). Výhodou názvov je fakt, že sú to absolútne odkazy, sú jednoznačným označením oblasti, platné v hárku alebo v celom zošite, čo je náš prípad.

Súčet údajov podľa jednej kľúčovej položky získame rôznymi spôsobmi, uvedieme najčastejšie používané.

Medzisúčty

Súčet faktúr podľa jednotlivých krajín ľahko získame databázovým nástrojom *medzisúčty* príkazu *údaje*, postup má ale jedno úskalie, výsledky sú správne iba vtedy, ak je tabuľka usporiadaná podľa kľúčovej položky, v našom prípade to je krajina.



Obr. 2: dialógové okno medzisúčty

Skupiny tvorí položka krajina, pri každej zmene hodnoty v stĺpci krajina sa do tabuľky doplní medzisúčet. Zvolili sme funkciu súčet, medzisúčet pridáme do stĺpca faktúra (pre každú

krajinu sa spočítajú hodnoty faktúr), označenie medzisúčtu sa vloží do stĺpca krajina, súčet faktúr do stĺpca faktúra (pozri obrázok 3). Výpočet realizuje Excel použitím matematickej funkcie *Subtotal(function_num,ref)*. Prvý argument je číslo funkcie (priemer je 1, počet 2, súčet 9), druhý argument je odkaz na súvislú oblasť riadkov s rovnakou kľúčovou hodnotou. Funkciu môžeme vložiť do tabuľky aj priamo, ale nástroj medzisúčty je efektívnejší. V prípade potreby môžeme doplniť počet faktúr, keď opakujeme postup tvorby medzisúčtov, použijeme funkciu Počet a v dialógovom okne zrušíme označenie nahradit' medzisúčty.

1	2	3	A	B	C	D	E
	1	krajina	meno	dátum	kód služby	faktúra	
+	74	AT Celkom				51 840,00 €	
+	111	CZ Celkom				20 553,00 €	
+	196	DE Celkom				58 329,00 €	
+	233	FR Celkom				26 083,00 €	
+	320	GB Celkom				64 649,00 €	
+	477	HU Celkom				110 365,00 €	
+	514	IT Celkom				23 696,00 €	
+	611	SK Celkom				65 217,00 €	
+	696	SR Celkom				59 304,00 €	
+	707	SW Celkom				7 848,00 €	
-	708	Celkový súčet				487 884,00 €	

Obr. 3 Súčet faktúr jednotlivých krajín

Funkcia Sumif

Funkcia *Sumif* sčítuje údaje definovaného stĺpca riadku, v ktorom hodnota kľúčovej položky vyhovuje podmienke. Výhodou funkcie je, že tabuľka nemusí byť usporiadaná, nevýhodou naopak je fakt, že pre každú hodnotu kľúčovej položky musíme zostrojiť funkciu. Vhodným umiestnením vzorcov a úpravou odkazov však môžeme vytvoriť kopírovateľné vzorce a tým zjednodušiť postup. V našom príklade sme ako argumenty funkcie použili názvy stĺpcov, $=Sumif(krajina;K2;faktúra)$ je vzorec v bunke K3, kde *krajina* je názov prvého stĺpca tabuľky, *faktúra* je názov stĺpca s hodnotou faktúr. Názov je absolútny odkaz, vzorec môžeme skopírovať do ostatných buniek riadku.

	J	K	L	M	N	O	P	Q	R	S	T
1											
2	krajina	HU	SR	CZ	AT	DE	FR	GB	IT	SK	SW
3	faktúry	110 365	59 304	20 553	51 840	58 329	26 083	64 649	23 696	65 217	7 848

Obr. 4: Súčet faktúr, vypočítaný funkciou Sumif

Funkcia $=Countif(krajina;K2)$ vypočíta počet, $=Averageif(krajina;K2;faktúra)$ priemernú hodnotu faktúr krajiny. Spomenuté funkcie nenahradia medzisúčty, nakoľko neexistujú podmienené funkcie všetkých medzisúčtov, riešením môžu byť databázové funkcie.

Funkcia Dsum

Vybrané matematické a štatistické funkcie, majú svoju obdobu databázových funkcií. Výhodou týchto funkcií je ich jednoduchosť, každá databázová funkcia má rovnakú štruktúru: *Dfunkcia(database,field,criteria)*. Databázová funkcia spracuje celú databázu (prvý argument je odkaz na oblasť, alebo názov tabuľky), hľadá riadky tabuľky, ktoré spĺňajú kritériá (tretí argument funkcie) a spracuje hodnoty vo zvolenom stĺpci (druhý argument), v príklade je to číslo stĺpca, ale môžeme použiť názov v tvare reťazca („faktúra“) alebo adresu bunky, v ktorej je nadpis stĺpca. V príklade hodnotu faktúr spočíta funkcia *Dsum*.

Nevýhodou databázových funkcií môže byť, že pre každú hodnotu kľúčového stĺpca musíme vytvoriť vlastnú oblasť kritérií. V príklade sú jednoduché kritériá, výber určuje iba jedna hodnota, tak sme výberové podmienky vytvorili skopírovaním do súvislej oblasti nad oblasťou, kam sme vkladali funkciu. V prvom argumente je názov tabuľky, do bunky K12 sme uložili vzorec $=DSUM(databáza_faktúry;5;K8:K9)$, potom sme ho skopírovali do ostatných buniek.

	J	K	L	M	N	O	P	Q	R	S	T
7											
8	kritériá	krajina	krajina	krajina	krajina	krajina	krajina	krajina	krajina	krajina	krajina
9		HU	SR	CZ	AT	DE	FR	GB	IT	SK	SW
10											
11	krajina	HU	SR	CZ	AT	DE	FR	GB	IT	SK	SW
12	faktúry	110 365	59 304	20 553	51 840	58 329	26 083	64 649	23 696	65 217	7 848

Obr. 5: Súčet faktúr, vypočítaný funkciou Dsum

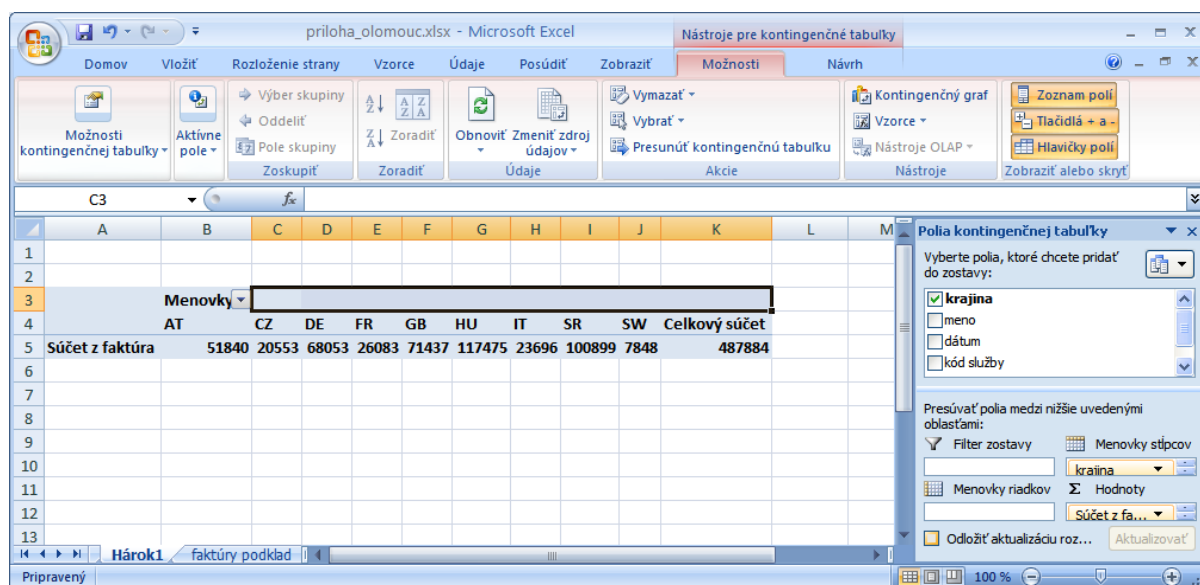
Kritérium môže byť ľubovoľne zložené, čo na jednej strane rozširuje množinu riešiteľných príkladov, na druhej strane zvyšuje prácnosť a tým časovú náročnosť vytvárania vzorcov.

Kontingenčná tabuľka

Efektívnym nástrojom získania medzisúčtov je kontingenčná tabuľka. Vytvorenie a úprava kontingenčnej tabuľky je otázkou niekoľkých minút. Kurzor presunieme do databázovej tabuľky, potom v príkaze *Vložiť kontingenčnú tabuľku* definujeme umiestnenie tabuľky, čím otvoríme okno *Polia kontingenčnej tabuľky*. Kritickým bodom návrhu je umiestnenie položiek databázy do vhodnej oblasti polí. Kľúčové položky budú polia riadkov alebo stĺpcov (v našom príklade je jediná stĺpcová položka, *krajina*), údajové položky tvoria spracované hodnoty (v príklade pole *faktúra*). Implicitné nastavenie funkcie spracovania pre číselné hodnoty je súčet, pre ostatné hodnoty počet. Keď kurzor je v oblasti údajov kontingenčnej tabuľky, môžeme upraviť vlastnosti spracovaných údajov

prípadne zmeniť funkciu spracovania. Príkaz *nastavenie poľa* je v kontextovej ponuke ale aj v zozname príkazov názvu položky v časti *definovania oblastí* okna *polia kontingenčnej tabuľky*. Množinu zobrazených údajov môžeme upraviť, ak do oblasti *filter zostavy* presunieme kľúčovú položku, podľa možnosti položku, ktorá nemá príliš veľa hodnôt a určíme filter zostavy, čím najčastejšie upravujeme podmienky na vytvorenie zostavy pre tlač. Podobne môžeme upraviť podmienky výberu pre riadkové a stĺpcové položky. Výhodou kontingenčnej tabuľky je jednoduchá modifikácia tabuľky,

premiestnenie alebo rozšírenie spracovaných polí dáva úplne iný pohľad na pôvodnú tabuľku. Tabuľka na obrázku 6 je podobná riešeniu pomocou funkcií. Výsledok sa bude podobat' medzisúčtom, ak pole *krajina* presunieme do oblasti riadkov. Počet faktúr v medzisúčtoch zobrazíme pridaním ďalšej úrovne medzisúčtov, v kontingenčnej tabuľke jednoducho vložíme do oblasti hodnôt ľubovoľné ďalšie pole a v prípade potreby zmeníme funkciu na počet. V našom príklade je najvhodnejšie pole *krajina*, vtedy systém zvolí funkciu počet.



Obr. 6: Súčet faktúr, vypočítaný kontingenčnou tabuľkou

Pridaním ďalšieho poľa do riadkových položiek vytvoríme dvojrozmernú tabuľku, ktorú funkcie ani medzisúčty neumožnia vytvoriť.

Tabuľka v našom prípade obsahuje dátum vystavenia faktúr. Tento údaj v prehľadných výstupoch je príliš podrobný, potrebujeme väčšinou mesačné, prípadne štvrťročné a ročné výsledky. Medzisúčty a Sumif neumožnia vytvoriť takýto prehľad bez pomocných výpočtov v tabuľke. Pomocou databázových funkcií by sme mohli úlohu vyriešiť tak, že definujeme výberové kritériá pre každý štvrťrok, potom vytvoríme vzorce (funkcia Dsum), ročné výsledky získame sčítaním štvrťročných výsledkov alebo priamo, ak ako výberové kritérium použijeme rok.

Rozčlenenie časových medzisúčtov podľa krajín je ešte zložitejšie. Pridanie ďalšej úrovne medzisúčtov databázy v upravenej tabuľke je pomerne jednoduché, ak je tabuľka vhodne usporiadaná, ale výsledok je hierarchicky členený, čo niekedy nestačí.

Kontingenčná tabuľka umožní zoskupenie údajov. Do riadkových položiek tabuľky vložíme dátum, kurzor presunieme na ľubovoľný dátum v kontingenčnej tabuľke, v paneli nástrojov *možností kontingenčnej tabuľky* alebo príkazom *Zoskupiť* v kontextovej ponuke otvoríme okno na tvorbu skupiny (štvrtrok, rok). Výsledok bude dvojrozmerná tabuľka, v ktorej sa zobrazia medzisúčty v jednotlivých rokoch a štvrtrokoch (pozri obrázok 7).

Súčet z faktúra	Me											
Menovky riadkov	AT	CZ	DE	FR	GB	HU	IT	SK	SR	SW	Celkový súčet	
2009	9372	5216	16320	7240	14298	26339	9672	20836	11402	2021	122716	
štvt2	98	3312	4671	3144	5582	13402	7573	9925	5386	231	53324	
štvt3	8359	1739	5676	2700	4208	8572	1374	6189	5657	1034	45508	
štvt4	915	165	5973	1396	4508	4365	725	4722	359	756	23884	
2010	42468	15337	42009	18843	50351	84026	14024	44381	47902	5827	365168	
štvt1	25872	11152	23950	12652	30827	55319	10537	26031	32458	3188	231986	
štvt2	16596	4185	18059	6191	19524	28707	3487	18350	15444	2639	133182	
Celkový súčet	51840	20553	58329	26083	64649	110365	23696	65217	59304	7848	487884	

Obr. 7: Súčet faktúr krajín, zoskupené podľa rokov a štvrtí rokov.

Pomocou databázovej funkcie síce vytvoríme tabuľku, ale počet kritérií rastie, bude to počet štvrtí rokov krát počet krajín, a funkcie pravdepodobne nebudú kopírovateľné. Túto istú úlohu vyriešime v kontingenčnej tabuľke za minútu. Stačí vložiť položku krajina do oblasti menovky stĺpcov, dvojrozmerná kontingenčná tabuľka obsahuje súčet faktúr, zoskupené podľa rokov a v tom podľa štvrtí rokov, rozčlenené podľa krajín. Medzisúčty rokov a celkové súčty stĺpcov a riadkov môžeme podľa potreby schovať alebo zobrazovať.

Kontingenčné tabuľky sú ako z príkladov vidieť efektívne a ľahko použiteľné, ale majú svoje nedostatky. Jednoduché úlohy je často lepšie riešiť pomocou funkcií, hlavne vtedy, ak výsledky chceme použiť ako operandy alebo argumenty ďalších výpočtov. Nevýhodou kontingenčnej tabuľky je aj to, že pri vytvorení kontingenčnej tabuľky systém pripraví podklady pre tvorbu tabuľky, tieto pracovné výpočty sú náročné na pamäť. Veľkosť pracovného zošita sa zväčší približne o tretinu, aj keď výhodou je, že kontingenčné tabuľky, vytvorené z tej istej tabuľky majú spoločný pracovný priestor. Ďalšie kontingenčné tabuľky sa vytvárajú na základe existujúceho pracovného priestoru, čo šetrí priestor a čas, tento fakt je ale aj nevýhodou, pretože zmeny v tabuľke sa v kontingenčnej tabuľke neprejavujú automaticky, dokonca ani v tej, ktorú zostrojíme po úprave tabuľky, kým funkcie a medzisúčty na úpravu tabuľky reagujú okamžite. Kontingenčné tabuľky musíme aktualizovať príkazom *obnoviť* v kontextovej ponuke kontingenčnej tabuľky, pričom ale stačí aktualizovať jednu kontingenčnú tabuľku, bez ohľadu na tom, koľko sme ich vytvorili.

4 Záver

Spracovanie a analýza rozsiahlych tabuliek je náročnou úlohou. Excel poskytuje rôzne

prostriedky na ich spracovanie, efektívnosť spracovania a výsledok závisí od toho, či použijeme vhodný nástroj. Riešenie konkrétnej úlohy je často možné rôznymi prostriedkami, čo je na jednej strane užitočné, používateľ si môže vybrať postup, ktorý mu vyhovuje, na druhej strane bohatá ponuka prostriedkov vyžaduje väčšie úsilie a viac času na ich zvládnutie.

V článku sme uviedli typické príklady na využívanie tabuľkového procesora na spracovania informácií a získanie podkladov na vyhodnotenie údajov a ekonomické modelovanie. Potrebné informácie môžeme získať často rôznymi prostriedkami, záleží na konkrétnej situácii, ktorú alternatívu si zvolíme. Prostriedky tabuľkového procesora Excel sú bohaté a každou verziou sa rozširujú, je skoro nemožné všetky zvládnuť, to ani nebolo účelom tohto článku. Chceli sme sa iba podeliť s čitateľmi o skúsenosti s využívaním menej známych ale o to užitočnejších prostriedkov Excelu a ukázať príklady, ktoré by im mohli byť užitočné pri riešení vlastných problémov.

5 Literatúra:

- [1] BROŽ, M.: Microsoft Excel pro manažery a ekonomy. 5. Vyd. Praha : Computer Press, 2006. ISBN 8025113078
- [2] CSAPO, B.: Formális és nem-formális tanulás során szerzett tudás integrálása (Integrácia vedomostí, získaných formálnym a neformálnym štúdiom). In: Iskolakultúra, roč. 2006, 2006, č. 2. ISSN 1215-5233
- [3] DOUCEK, P. - NOVOTNÝ, O. - PECÁKOVÁ, I. - VOŘÍŠEK, J.: Lidské zdroje v ICT – Analýza nabídky a poptávky po IT odbornících v ČR. 1. Vyd. Praha : Professional publishing, 2007. ISBN 978-80-86946-51-1.
- [4] HUBA, M. – PIŠŮTOVA, GERBER, K.: Základy e-vzdělávání. Bratislava : STU, 2007. ISBN 978-80-89316-00-7.

- [5] HUNTER, M.: Účinné vyučování v kostce. Praha : Portál, 1999.
- [6] KOMPOLTOVÁ, S. Pedagogika. Bratislava : Vydavateľstvo Ekonóm, 2009. ISBN 978-80-225-2805-4.
- [7] RALSTON, A. - REILLY jr., E.: Encyclopedia of Computer Science and Engineering. New York : Van Nostrand Reinhold Company Inc., 1983.
- [8] GAŠPERANOVÁ, A., LEDVÉNYI, S., ŠČESNÝ, R., ŠÓŠ, I.: *Informatika - praktikum*. Bratislava : Vydavateľstvo Ekonóm 2006. ISBN 80-225-2236-8.

RNDr. Agneša Gašperanová
**Katedra aplikovanej informatiky a výpočtovej
techniky**
Národohospodárska fakulta
Ekonomická univerzita v Bratislave
Dolnozemska 1
85235 Bratislava
Slovensko
tel +421-02/6729 1274
E-mail: agnes.gasperanova@euba.sk
www pracoviska: www.euba.sk