

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Luka Lampret

**Skladišče EMG podatkov s sprotnim
analitičnim prikazom v okolju
Excel 2010**

DIPLOMSKO DELO
VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

Ljubljana, 2012

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Luka Lampret

**Skladišče EMG podatkov s sprotnim
analitičnim prikazom v okolju
Excel 2010**

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: prof. dr. Blaž Zupan
SOMENTOR: doc. dr. Gaj Vidmar

Ljubljana, 2012



Št. naloge: 00154/2011

Datum: 01.09.2011

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **LUKA LAMPRET**

Naslov: **SKLADIŠČE EMG PODATKOV S SPROTNIM ANALITIČNIM PRIKAZOM
V OKOLJU EXCEL**

EMG DATA WAREHOUSE WITH EXCEL-BASED VISUAL ANALYTICS

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija prve stopnje

Tematika naloge:

Delo se navezuje na dolgoletne slovenske raziskave električne aktivnosti gladkega mišičja maternice med nosečnostjo, kjer je nastala obsežna in kompleksna podatkovna zbirka meritev na živalskem modelu (pri brejih ovcah). Diplomsko delo naj bi pokazalo, da je moč učinkovito rešitev za upravljanje in analitično grafično prikazovanje tovrstnega podatkovja izdelati v okolju Microsoft Excel 2010. Vključevalo naj bi uporabo celotne palete zmogljivosti Excela - od neprogramskih razvojnih elementov, kot so dinamične formule in kontrolniki, preko programske kode v okolju VBA do novosti v različici 2010, zlasti iskrnih črt (tj. grafikonov sparkline) in dodatka PowerPivot. V diplomskem delu naj bi bile predstavljene tudi možnosti spletne objave in drugih oblik omrežne interaktivne uporabe izdelanega podatkovno-analitičnega središča na podlagi Microsoftovih tehnologij.

Mentor:

prof. dr. Blaž Zupan

Dekan:

prof. dr. Nikolaj Zimic

Somentor:

doc. dr. Gaj Vidmar



Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .

IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Luka Lampret, z vpisno številko **63020094**, sem avtor diplomskega dela z naslovom:

Skladišče EMG podatkov s sprotnim analitičnim prikazom v okolju Excel 2010

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom prof. dr. Blaža Zupana in somentorstvom doc. dr. Gaja Vidmarja
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 15. marec 2012

Podpis avtorja:

Zahvaljujem se mentorju prof. dr. Blažu Zupanu in somentorju doc. dr. Gaju Vidmarju za mentorstvo in nasvete pri izdelavi diplomskega dela. Somentorju se posebej zahvaljujem za vso pomoč. Hvala tudi družini in prijateljem, ki so mi stali ob strani ter pretrpeli mojo odsotnost.

Kazalo

Povzetek

Abstract

1	Uvod	1
2	Uporabljena orodja in tehnologije	3
2.1	Orodja	3
2.1.1	Microsoft Excel 2010	3
2.1.2	MySQL	4
2.1.3	MySQL Workbench	5
2.1.4	PowerPivot	6
2.2	Tehnologije	7
2.2.1	VBA	7
2.2.2	SQL	8
3	Razvoj okolja za interaktivno podatkovno analitiko v programu Excel	9
3.1	Uporaba MSSQL/MySQL ali Excel za vir podatkov	9
3.2	Izdelava prikaza uspešnosti EMG meritev	12
3.3	Risanje iskrnih črt	15
3.3.1	Nastavitvena forma	15
3.3.2	Izrisani graf	22

KAZALO

4	Uporaba dodatka PowerPivot	27
4.1	Uvoz podatkov	27
4.2	Praktična uporaba	35
5	Sklepne ugotovitve	37

Seznam uporabljenih kratic in simbolov

ANSI (angl. *American National Standards Institute*)

DAX (angl. *Data Analysis Expressions*)

EMG (angl. *electromyography*); elektromiografija

IDE (angl. *Integrated Development Environment*); integrirano razvojno okolje

MF (angl. *Median Frequency*); medianska frekvenca

ODBC (angl. *Open Database Connectivity*)

OLAP (angl. *Online Analytical Processing*); sprotni analitični prikaz

RDBMS (angl. *Relational Database Management Systems*)

RMS (angl. *Root Mean Square*); efektivna vrednost

SQL (angl. *Structure Query Language*); strukturiran poizvedovalni jezik

VBA (angl. *Visual Basic for Applications*)

VBE (angl. *Visual Basic Editor*)

Povzetek

Namen diplomskega dela je pokazati, da se da kakovostno in učinkovito izdelati rešitev za upravljanje in raziskovalno-analitično grafično prikazovanje podatkov v okolju Excel 2010. Uporabili smo obsežno in kompleksno podatkovno zbirko meritev električne aktivnosti gladkega mišičja maternice med nosečnostjo na živalskem modelu (pri brejih ovcah). Delo sestavljata dva sklopa. Prvi sklop vključuje celotno paleto zmogljivosti zadnje verzije Excela – od zahtevnejših formul preko kontrolnikov in programske kode v okolju VBA do iskernih črt (grafikonov sparkline) – za izrisovanje meritev po uporabnikovi izbiri. V drugem sklopu smo za odpravljanje omejitev in dosego interaktivnosti v Excelovih grafikonih uporabili dodatek PowerPivot, kar nam je omogočilo združitve vseh podatkov v eno datoteko in enostavno manipulacijo s temi podatki. Tako smo dobili sprotni analitični prikaz (OLAP), ki upošteva vse odvisnosti med podatki in na učinkovit način omogoča, da jih kakovostno prikažemo z uporabo Excelovih grafikonov.

Ključne besede:

elektromiografija, prikaz podatkov, interaktivna vizualizacija, elektronske preglednice, VBA, Excel, PowerPivot, OLAP

Abstract

The aim of the thesis was to demonstrate that a state-of-the-art solution for management and analytical visualization of data can be implemented based on Excel 2010. We used a large and complex database of measurements of electrical activity in uterine smooth muscles during pregnancy in an animal model (pregnant sheep). The thesis comprises two parts. The first part fully exploits the functionalities of the latest version of Excel – ranging from advanced formulas through controls and VBA code to sparklines – for graphing measurements according to the user’s demand. In the second part, we used the PowerPivot add-in to overcome the limitations and implement interactivity in Excel’s charts, which enabled us to combine all the data in one file and manipulate them effortlessly. In this way, we produced an OLAP application that takes into account all the relations within the data and efficiently implements high-quality visual analytics using Excel’s charts.

Key words:

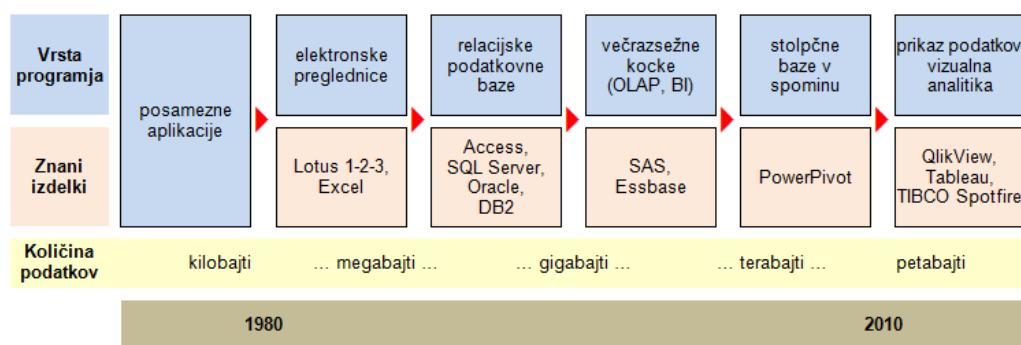
electromyography, data display, interactive visualization, spreadsheet, VBA, Excel, PowerPivot, OLAP

Poglavje 1

Uvod

Analiza podatkov je praviloma iterativen postopek, v katerem podamo poizvedbo in dobimo odgovor, oblikujemo naslednjo poizvedbo, ki temelji na odgovoru, in tako nadaljujemo. Obseg in vsebino poizvedbe tako praviloma ožamo oziroma izboljšujemo na podlagi predhodnega znanja in povratnih informacij [1].

Novejši razvoj računalniških orodij za sprotno analizo podatkov oziroma analitične poizvedbe poenostavljeno povzema slika 1.1. Hkrati s tem lahko sledimo razvoju prikaza podatkov (informacijske vizualizacije) oziroma vizualne podatkovne analitike kot samostojne discipline [2, 3], ki so ji temelje postavili raziskovalci s področja statistične grafike [4, 5]. V skladu z iterativno naravo poizvedb po podatkovnih zbirkah je vodilno načelo vizualne podatkovne analitike “pregled (angl. *overview*), izbira in presejanje (angl. *zoom & filter*), podrobnosti na zahtevo (angl. *details-on-demand*)” [6].



Slika 1.1: Razvoj računalniških orodij za sprotno analizo podatkov (prirejeno po <http://apandre.wordpress.com>)

V diplomskem delu smo si zadali nalogo pokazati, da se da kakovostno in učinkovito izdelati rešitev za upravljanje in raziskovalno-analitično grafično prikazovanje podatkov v okolju Excel 2010. Uporabili smo obsežno in kompleksno podatkovno zbirko meritev električne aktivnosti gladkega mišičja maternice med nosečnostjo na živalskem modelu (pri brejih ovcah) [7, 8]. Najprej smo razvili interaktivno okolje v Excelu, ki omogoča prikaz uspešnosti meritev in izris iskrne črte (angl. sparkline) [9] za vsako posamezno meritev na posamezni ovci. To nam omogoča podrobno pregledovanje in analizo podatkov, kjer lahko ugotovimo morebitne zakonitosti in trende, ki se pojavijo med različnimi ovcami.

V drugem delu smo s pomočjo Excelovega dodatka PowerPivot združili vse podatke v OLAP (angl. *Online Analytical Processing*) kocko, ki vsebuje v naprej pripravljene agregate podatkov. Podatkovno kocko sestavlja več povezanih tabel, kjer je v osrednjem delu tabela podatkov (ang. *Fact Table*), okoli nje pa so povezane dimenzijske tabele. Vsebinsko teh tabel potem s pomočjo vrtilnih tabel (angl. *Pivot Table*) in grafov spremenimo v učinkovito in enostavno orodje za iskanje ključnih podatkov. Z uporabo vgrajenih Excelovih razdelilnikov in grafikonov, vezanih na OLAP kocko, smo lahko na ta način implementirali interaktivne grafične prikaze.

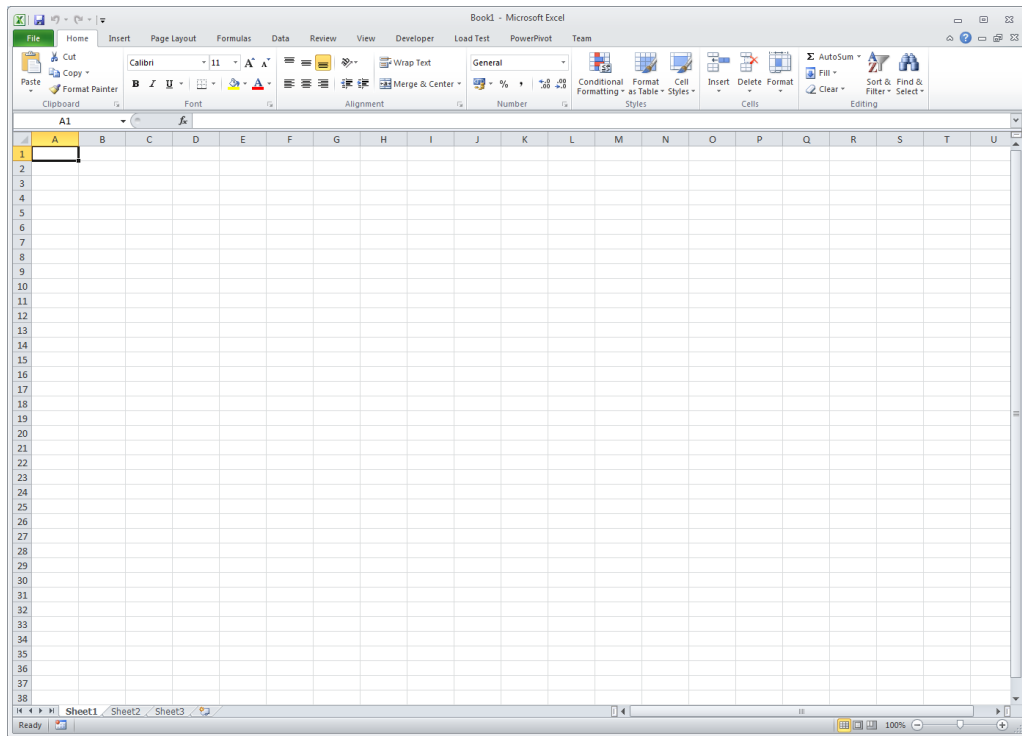
Poglavje 2

Uporabljena orodja in tehnologije

2.1 Orodja

2.1.1 Microsoft Excel 2010

Excel (slika 2.1) je računalniški program, ki je del vodilnega pisarniškega programskega paketa Microsoft Office. Je daleč najbolj razširjena elektronska preglednica – po celem svetu jo uporablja več kot pol milijarde ljudi, med katerimi je vsaj desetina zahtevnejših uporabnikov. Njegove glavne prednosti so intuitiven uporabniški vmesnik, zmogljiva orodja za računanje in izdelavo grafov ter praktično neomejena razširljivost z makri v programskem jeziku VBA. Največkrat se uporablja za urejanje preglednic, izdelavo grafikonov in obdelave podatkov z vrtilnimi tabelami. Vgrajenih ima na stotine funkcij in množico podatkovnih orodij. Zaradi vsega navedenega je odlično in univerzalno orodje za izobraževanje, finance, inženiring in mnoga druga področja [10, 11].



Slika 2.1: Microsoft Excel 2010

2.1.2 MySQL

MySQL je večuporabniška, formalno definirana in centralno nadzorovana zbirka podatkov. Je klasična (relacijska) zbirka podatkov, ki lahko deluje kot samostojen podatkovni strežnik ali pa v kombinaciji z drugimi tehnologijami. Poizvedbe v podatkovni bazi MySQL se izvajajo z jezikom SQL (angl. *Structured Query Language* – strukturiran poizvedovalni jezik).

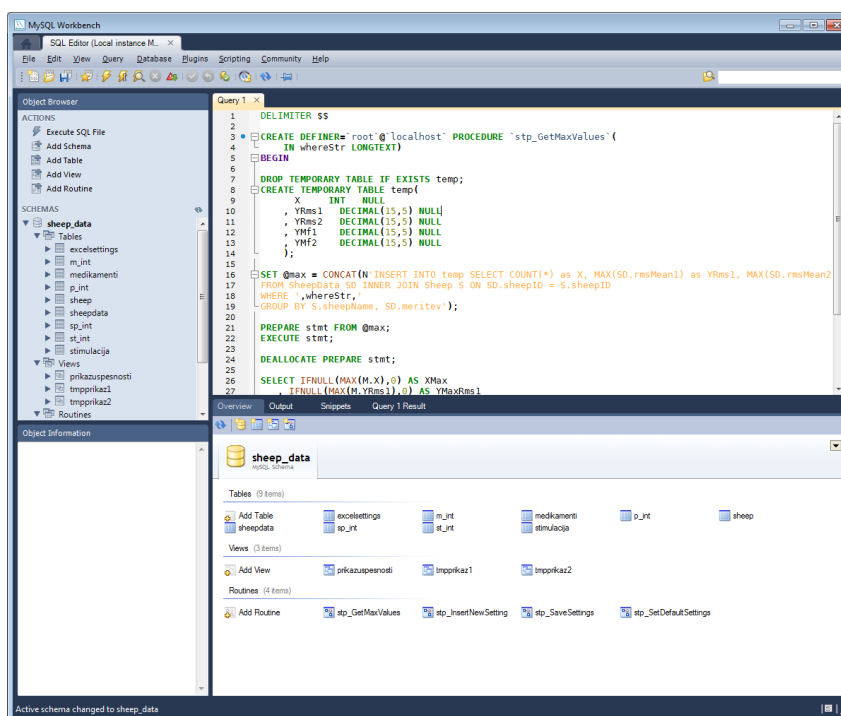
Pred približno desetimi leti je MySQL v jezikih C in C++ razvilo švedsko podjetje MySQL AB in ga ponudilo trgu pod pogoji uporabe javne licence GNU (<http://www.gnu.org>). Postal je najbolj priljubljena baza odprtega tipa in njegova uporaba po zaslugi visoke zanesljivosti, hitrega delovanja, prilagojenosti in nenehnega razvoja še vedno raste. Trenutno je nameščen na več kot 8 milijonih sistemov in deluje na več kot 20 operacijskih sistemih (Linux, Windows, OS/X, HP-UX, AIX, Netware, idr.).

Strežnik MySQL lahko upravljamo na dva načina: z grafičnim uporabniškim vmesnikom (MySQL Workbench, MyAdmin, MySQL Front idr.) ali z ukazno vrstico [12].

2.1.3 MySQL Workbench

MySQL Workbench (slika 2.2) je orodje za vizualno oblikovanje, načrtovanje, modeliranje, ustvarjanje in upravljanje podatkovnih baz. V glavnem ga uporabljajo razvijalci, arhitekti in bazni administratorji.

MySQL Workbench vključuje SQL Editor (urejevalnik SQL), ki ga uporabljamo za ustvarjanje, izvajanje in optimizacijo poizvedb SQL. Vsebuje tudi Object Browser (objektni brskalnik), ki omogoča takojšen dostop do objektov in podatkovnih baz. Poleg tega nam z vizualnimi orodji omogoča urejanje uporabniških pravic na strežniku, kot tudi upravljanje z nastavitvami strežnika [13].



Slika 2.2: Običajni pogled na MySQL Workbench

2.1.4 PowerPivot

PowerPivot za Excel 2010 (slika 2.3) je orodje za analizo podatkov, ki nam ponuja veliko računalniško moč neposredno znotraj programa Excel. Sestavljen je iz dodatka za Excel 2010 in pogona VertiPaq, ki v pomnilniku stisne in upravlja na milijone vrstic podatkov, kar prinaša zelo hitro delovanje. S tem dodatkom lahko v Excelu velike količine podatkov z veliko hitrostjo pretvorimo v smiselne informacije in tako pridobimo želene odgovore. Podatke lahko uvozimo iz različnih virov, kot so besedilna datoteka, Excelova datoteka, accessova baza, SQL strežnik itn.

S funkcijami v Excelu 2010, kot so vrtilne tabele, razčlenjevalniki in druge znane funkcije za analizo, lahko interaktivno poizvedujemo in analiziramo podatke brez posebnega predhodnega usposabljanja. Uporabimo lahko tudi izraze za analizo podatkov (DAX – Data Analysis Expressions) in izvedemo zmogljive prilagoditve podatkov. Sledimo relacijam med tabelami kot v zbirki podatkov, določimo izračunane stolpce in merske enote ter združimo na milijone vrstic [14].

PowerPivot za Excel 2010 je brezplačen in ga lahko prenesemo s spletnega mesta www.powerpivot.com.

SheepID	Meritew	Datum	Izbajanje meritew	Minuta meritew	S.	Sp. minut	M.	M. minut	S.	S. minut	B.	P. minut
14	104	11.3.1999	699	61	0	1	0					
14	104	11.3.1999	699	62	0	2	0					
14	104	11.3.1999	699	63	0	3	0					
14	104	11.3.1999	699	64	0	4	0					
14	104	11.3.1999	699	65	0	5	0					
14	104	11.3.1999	699	66	0	6	0					
14	104	11.3.1999	699	67	0	7	0					
14	104	11.3.1999	699	68	0	8	0					
14	104	11.3.1999	699	69	0	9	0					
14	104	11.3.1999	699	70	0	10	0					
14	104	11.3.1999	699	71	0	11	0					
14	104	11.3.1999	699	72	0	12	0					
14	104	11.3.1999	699	73	0	13	0					
14	104	11.3.1999	699	74	0	14	0					
14	104	11.3.1999	699	75	0	15	0					
14	104	11.3.1999	699	76	0	16	0					
14	104	11.3.1999	699	77	0	17	0					
14	104	11.3.1999	699	78	0	18	0					
14	104	11.3.1999	699	79	0	19	0					
14	104	11.3.1999	699	80	0	20	0					
14	104	11.3.1999	699	81	0	21	0					
14	104	11.3.1999	699	82	0	22	0					
14	104	11.3.1999	699	83	0	23	0					
14	104	11.3.1999	699	84	0	24	0					
14	104	11.3.1999	699	85	0	25	0					
14	104	11.3.1999	699	86	0	26	0					
14	104	11.3.1999	699	87	0	27	0					
14	104	11.3.1999	699	88	0	28	0					

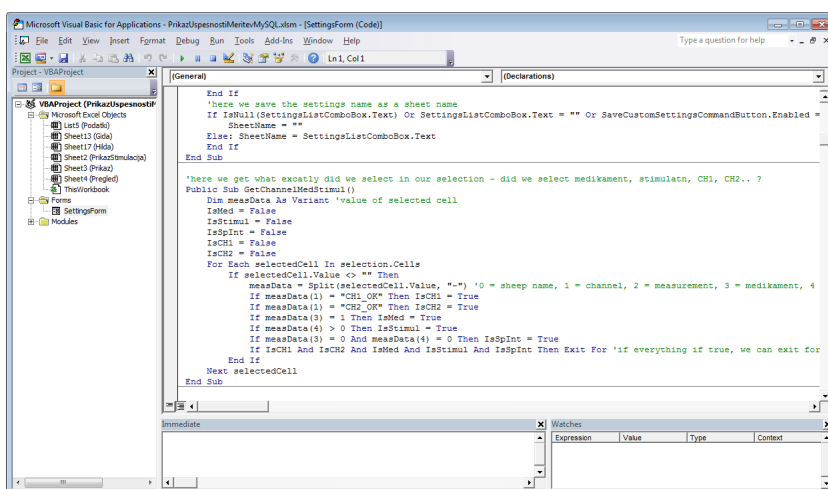
Slika 2.3: Dodatek PowerPivot

2.2 Tehnologije

2.2.1 VBA

Visual Basic for Applications (VBA) (slika 2.4) je izvedba Microsoftovega programskega jezika Visual Basic. Gre za razširljiv programski jezik, ki je sestavljen iz nabora ukazov in lahko neposredno dela z objekti v delovni datoteki. To pomeni, da VBA pozna stvari, kot so delovni zvezki, delovni listi, celice, grafikoni itn. Temelji na objektih (angl. *object-based*), vendar ne gre za pravi objektno usmerjen jezik, in na dogodkih, se pravi, da se nobena vrstica kode ne izvede, dokler se ne zgodi nek dogodek.

Povezan je z integriranim razvojnim okoljem (IDE – angl. *Integrated Development Environment*) oz. VBE (angl. *Visual Basic Editor*), vgrajenim v večino programov v zbirki Microsoft Office (v našem primeru Microsoft Excel). To je za uporabnika zelo prijazno okolje, kjer razvijamo, testiramo in spreminjamo makroje, ki jih uporabljamo v Excelu. Ko uporabimo te makroje, razvite v VBE, postanejo del delovnega zvezka. Taki delovni zvezki so shranjeni v datoteki s končnico *xlsm* [15].



Slika 2.4: VBE v Excelu, kjer pišemo makroje v jeziku VBA

2.2.2 SQL

SQL ali strukturirani povpraševalni jezik (angl. *Structured Query Language*) se uporablja za komunikacijo z bazo podatkov. V skladu z ANSI (American National Standards Institute), je standardni jezik v relacijskih sistemih za upravljanje z bazami podatkov (RDBMS – Relational Database Management Systems). Izjave SQL se uporabljajo za opravljanje nalog, kot so posodobitev podatkov v bazi, ali pa pridobivanje podatkov iz baze. Nekateri pogosti relacijski sistemi za upravljanje baz podatkov, ki uporabljajo SQL, so: MySQL, Oracle, Sybase, Microsoft SQL Server, Access in Ingres. Čeprav večina teh sistemov uporablja ukaze SQL, ima večina od njih tudi svoje dodatne razširitve, ki se običajno uporabljajo le v njihovih sistemih. Vendar pa se lahko standardni ukazi SQL (slika 2.5), kot so “Select”, “Insert”, “Update”, “Delete”, “Create” in “Drop”, uporabijo za izpolnitev skoraj vseh zahtev, ki jih moramo izpolniti z bazo podatkov [16].

```
SELECT name,
        title,
        price,
        price * 0.06 AS sales_tax
FROM    book
WHERE   price > 100.00
ORDER  BY title;

INSERT INTO my_table
        (field1,
         field2,
         field3)
VALUES  ('test',
        'N',
        NULL);

UPDATE my_table
SET    field1 = 'updated value'
WHERE  field2 = 'N';

DELETE FROM my_table
WHERE  field2 = 'N';

CREATE TABLE my_table
(
    my_field1 INT,
    my_field2 VARCHAR(50),
    my_field3 DATE NOT NULL,
    PRIMARY KEY (my_field1, my_field2)
);

ALTER TABLE my_table ADD my_field4 NUMBER(3) NOT NULL;

TRUNCATE TABLE my_table;

DROP TABLE my_table;
```

Slika 2.5: Primeri osnovnih SQL poizvedb

Poglavje 3

Razvoj okolja za interaktivno podatkovno analitiko v programu Excel

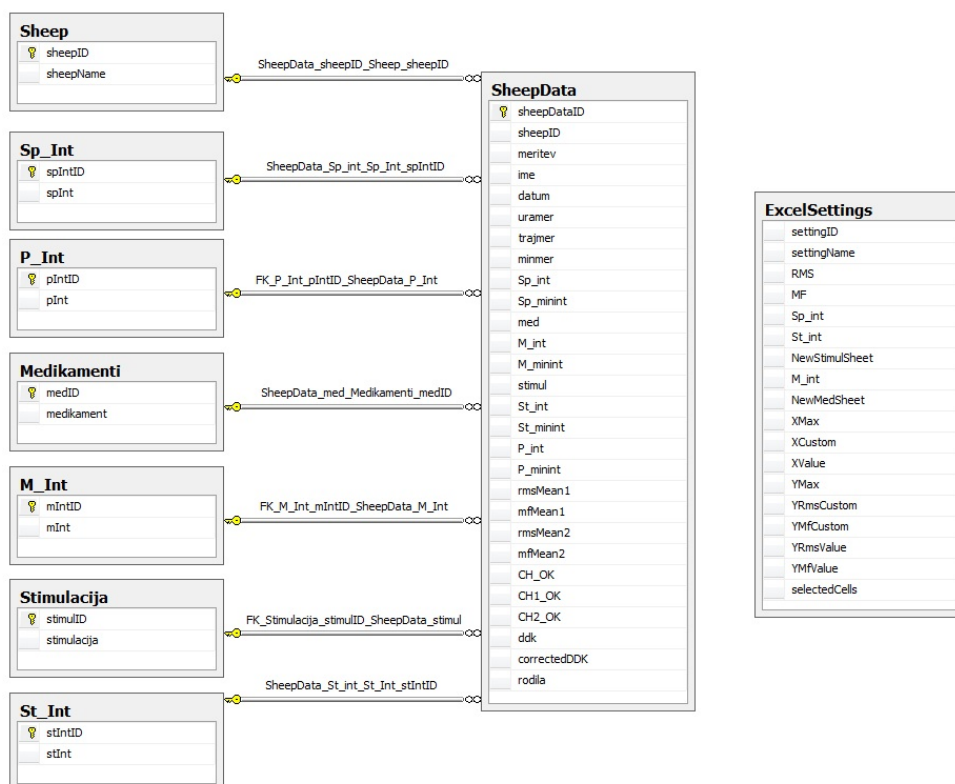
Pri razvoju okolja za interaktivno podatkovno analitiko smo se oprli na predhodne raziskave in obstoječo rešitev [17]. Ta prikaz smo izboljšali in pohitrili ter dodali nove funkcionalnosti, kjer lahko za posamezno meritev izrišemo graf v obliki iskrnih črt.

3.1 Uporaba MSSQL/MySQL ali Excel za vir podatkov

Naš prvotni vir podatkov je bilo 21 Excelovih datotek po ena za vsako ovco, ki smo jih morali najprej združiti v eno datoteko. Pri tem združevanju podatkov je prišlo tudi do nepričakovanih zapletov, saj nekateri podatki niso bili pravilni oziroma popolni. Ponekod je bilo treba popraviti zamik podatkov, drugod datum in uro itn. Ko smo podatke združili, jih je bilo potrebno kopirati v bazo.

Za postavitev in izgradnjo baze smo se odločili uporabiti MSSQL (Microsoft SQL Server). Baza sama je zgrajena iz devetih tabel, relacije med

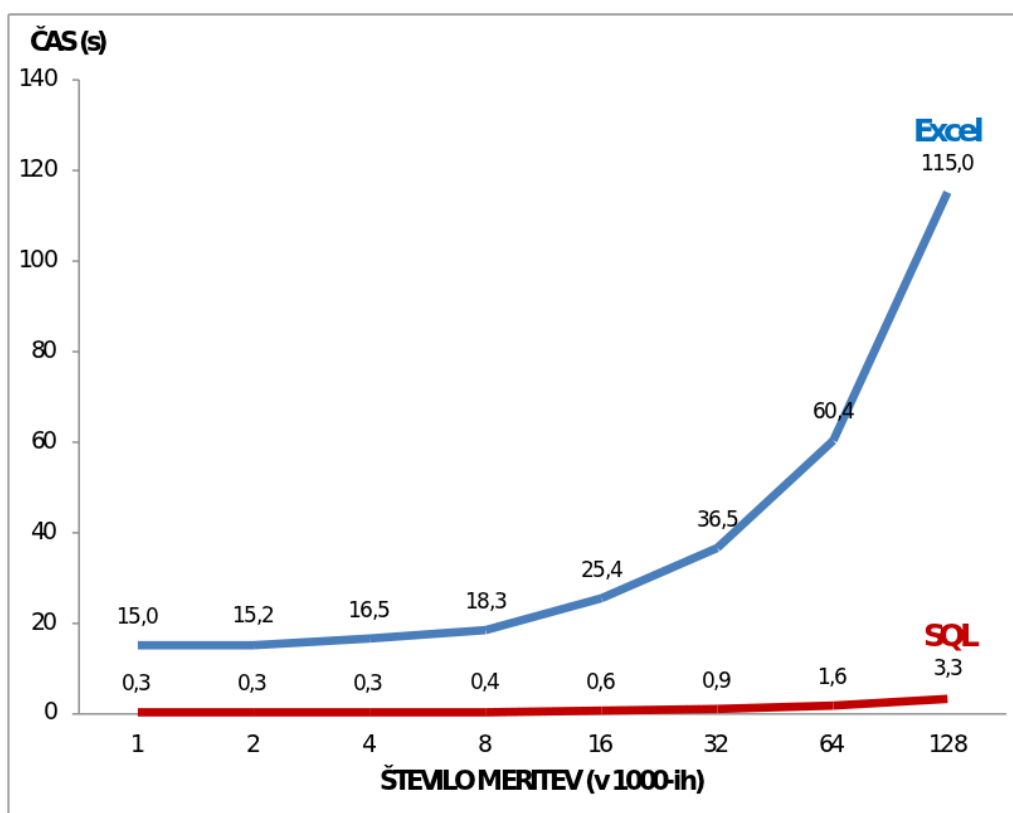
njimi so vidne na sliki 3.1. Poleg teh tabel so v bazi prisotne še štiri shranjene procedure (angl. *Stored Procedure*) in en pogled (angl. *View*).



Slika 3.1: Diagram relacij podatkovne baze

Procedure *stp_InsertNewSetting*, *stp_SetDefaultSettings* in *stp_SaveSettings* se uporabljajo pri shranjevanju nastavitev, *stp_GetMaxValues* pa pri računanju maksimalnih vrednosti meritev. Pogled *PrikazUspesnosti* nam pomaga pri pridobivanju ustreznih podatkov za izdelavo prikaza uspešnosti EMG meritev.

Ker smo sprva želeli, da bi se interaktivno okolje uporabljalo samo s pomočjo Excela, smo razvili tudi to možnost. Za povezavo med Excelovima datotekama smo uporabili ponudnika (angl. *Provider*) Microsoft.ACE.OLE-DB.12.0, ki pa se je izkazal za prepočasnega, saj se je pri večjem številu poizvedb čas, potreben za izris, povečal čez mejo, ki bi bila še sprejemljiva za interaktivnost (slika 3.2).



Slika 3.2: Primerjava hitrosti pridobivanja podatkov med Excelom in SQL strežnikom

Na koncu smo se odločili, da bomo namesto MSSQL uporabili odprtokodno rešitev MySQL, kjer je glavna prednost predvsem v tem, da ne predstavlja dodatnih stroškov. Seveda prehod ni potekal brez težav, saj se sintaksa ponekod razlikuje in MySQL ne omogoča gnezdenih poizvedb, kot jih omogoča MSSQL. Zato so bile potrebne manjše spremembe pri pogledu *PrikazUspešnosti*.

Za delovanje interaktivnega okolja je potrebno poleg strežnika MySQL, ki ga lahko brezplačno prenesemo z naslova <http://www.mysql.com/downloads/mysql/>, namestiti še MySQL Workbench (<http://www.mysql.com/downloads/workbench/>) in dva ponudnika (Microsoft.ACE.OLEDB.12.0 ter MySQL ODBC 5.1 Driver). Vse potrebne datoteke so priložene na zgoščenki.

Po namestitvi vseh naštetih programov je potrebno ustvariti bazo oziroma

t.i. shemo. To naredimo z uporabo orodja MySQL Workbench, kamor skopiramo vsebino iz priložene datoteke “BazaMySQL.sql” in skripto izvršimo. S tem se nam ustvari potrebno ogrodje, kamor bomo skopirali naše podatke iz Excelove datoteke. Preden se lotimo kopiranja, pa je potrebno v bazo vnesti še nekatere obvezne podatke, ki so potrebni za delovanje interaktivnega okolja. Te najdemo v datoteki “Obvezni podatki-MySQL.sql” v obliki stavkov “Insert”, ki jih – enako kot prej – prekopiramo v orodje Workbench in izvršimo.

Ko imamo pripravljeno podatkovno bazo, lahko prekopiramo podatke iz naše Excelove datoteke “VseOvceMySQL.xlsm”. Ko odpremo to datoteko, moramo najprej nastaviti pravilno ime strežnika MySQL in podatkovne baze skupaj z uporabniškim imenom in geslom, potrebnima za dostop do maloprej ustvarjene sheme, ter polno potjo do te datoteke. Če podatki niso pravilni, nas na to opozori opozorilno okno. Po pravilno vpisanih nastavitvah kliknemo na gumb “Kopiraj podatke v bazo”, ki najprej pobriše stare podatke iz tabele in nato kopira vse podatke v to prazno tabelo. Da se vsi podatki skopirajo, je potrebno počakati vsaj tri minute, saj je podatkov za več kot 300.000 vrstic.

3.2 Izdelava prikaza uspešnosti EMG meritev

Za začetek smo avtomatizirali pridobivanje podatkov, ki so potrebni za prikaz meritev. To smo storili z že omenjenim pogledom (angl. *View*) v bazi, ki se imenuje *PrikazUspešnosti*. Ta pogled nam združi podatke o posameznih ovcah in njihovih meritvah. Tako nam za vsako meritev, ki je bila opravljena na določeni ovci, vrne ime ovce, meritev, datum opravljanja meritve, število dni do kotitve, uspešnost meritve na rogu in vratu ovce, datum kotitve ter ali so bili ob tem merjenju dodani medikamenti in ali je bila ovca dodatno stimulirana. Da te podatke pridobimo, moramo najprej pravilno izpolniti podatke o strežniku in podatkovni bazi, kot smo to storili pri kopiranju podatkov v bazo, in nato s klikom na gumb “Osveži podatke” osvežimo stare

podatke. Na tem zavihku se nahaja tudi legenda, ki pove, kaj pomenijo barve na izdelanem prikazu uspešnosti (slika 3.3).

Ovca	Meritev	Datum	dkl	Rog OK	Vrat OK	rodila	Medikamenti	Stimulacija
2 B	1	30.8.1995	-90	2	2	28.9.1995	0	0
3 B	2	3.7.1995	-87	2	2	28.9.1995	0	0
4 B	3	4.7.1995	-86	2	2	28.9.1995	0	0
5 B	4	5.7.1995	-85	2	2	28.9.1995	0	0
6 B	5	7.7.1995	-83	0	0	28.9.1995	0	0
7 B	6	7.7.1995	-83	0	0	28.9.1995	0	0
8 B	7	10.7.1995	-80	2	2	28.9.1995	0	0
9 B	8	11.7.1995	-79	2	2	28.9.1995	0	0
10 B	9	12.7.1995	-78	2	2	28.9.1995	0	0
11 B	10	14.7.1995	-76	2	2	28.9.1995	0	0
12 B	11	19.7.1995	-71	2	2	28.9.1995	0	0
13 B	12	1.8.1995	-58	2	2	28.9.1995	0	0
14 B	13	9.8.1995	-58	0	2	28.9.1995	0	0
15 B	14	11.8.1995	-48	2	2	28.9.1995	0	0
16 B	15	21.8.1995	-38	2	2	28.9.1995	0	0
17 B	16	25.8.1995	-34	2	2	28.9.1995	0	0
18 B	17	28.8.1995	-31	2	2	28.9.1995	0	0
19 B	18	1.9.1995	-27	1	1	28.9.1995	0	0
20 B	19	5.9.1995	-23	2	2	28.9.1995	0	0
21 B	20	12.9.1995	-16	0	2	28.9.1995	0	0
22 B	21	15.9.1995	-13	0	2	28.9.1995	0	0
23 B	22	19.9.1995	-9	0	0	28.9.1995	0	0
24 B	23	22.9.1995	-6	0	0	28.9.1995	0	0
25 B	24	26.9.1995	-2	0	2	28.9.1995	0	0
26 B	25	29.9.1995	0	2	2	28.9.1995	0	0
27 B	26	29.9.1995	1	2	2	28.9.1995	0	0
28 B	27	12.10.1995	14	2	0	28.9.1995	0	0
29 D	1	8.8.1996	-56	0	0	3.10.1996	0	0
30 D	2	8.8.1996	-56	0	0	3.10.1996	0	0
31 D	3	9.8.1996	-55	0	0	3.10.1996	0	0
32 D	4	9.8.1996	-55	0	0	3.10.1996	0	0
33 D	5	10.8.1996	-54	0	0	3.10.1996	0	0
34 D	6	10.8.1996	-54	0	0	3.10.1996	0	0
35 D	7	11.8.1996	-53	0	0	3.10.1996	0	0
36 D	8	11.8.1996	-53	0	0	3.10.1996	0	0
37 D	9	12.8.1996	-52	0	0	3.10.1996	0	0
38 D	10	12.8.1996	-52	0	0	3.10.1996	0	0
39 D	11	12.8.1996	-52	0	0	3.10.1996	0	0
40 D	12	12.8.1996	-52	2	2	3.10.1996	0	0
41 D	13	12.8.1996	-52	2	2	3.10.1996	0	0
42 D	14	13.8.1996	-51	2	2	3.10.1996	0	0
43 D	15	13.8.1996	-51	2	2	3.10.1996	0	0
44 D	16	14.8.1996	-50	0	0	3.10.1996	0	0
45 D	17	14.8.1996	-50	2	2	3.10.1996	0	0
46 D	18	15.8.1996	-49	2	2	3.10.1996	0	0

Slika 3.3: Legenda s podatki o uspešnosti meritev

V naslednjem koraku razvoja rešitve smo nekoliko spremenili, dopolnili in pohitrili risanje uspešnosti meritev. Risanje izvedeno s klikom na gumb “Nariši uspešnost meritev”. Na sliki 3.4 vidimo grafični prikaz uspešnosti meritev, ki so jih izvedli pri ovcah.

Prikaz je sestavljen iz več delov. V zgornji vrstici vidimo vsa imena ovc, pod njimi datume prve meritve, ki so jo izvedli pri tej ovci, pod datumi pa se nahajata kanala, na katerih so bile opravljene meritve (rog = kanal 1 in vrat = kanal 2). Na levem robu vidimo številke, ki predstavljajo posamezne dneve. Ob kliku na dan se nam označi celotna vrstica, da lahko hitreje vidimo, katere meritve spadajo pod tisti dan. Ovca ima lahko na dan eno ali več meritev, če pa tisti dan ni bilo meritve, je puščen prazen prostor. Vsaka meritev je posebej obrobjena, tako da če imamo na določen dan npr. štiri meritve, imamo za tisti dan štiri pravokotnike na kanalu 1 in štiri na kanalu 2.

The screenshot displays an Excel spreadsheet titled 'PrikazUspešnostiMeritevMySQLJfsm - Microsoft Excel'. The spreadsheet contains a table with 19 columns representing different individuals and multiple rows of data. The columns are labeled: B, D, Gida, Hilda, I, Jolanda, Kaja, Liza, Liza2, Mimi, Mimi2, Nika, Pika, Quin, Rita, Sasa, Wilma, Xantia, Yaris, Zala, and Ajda. Each column contains a date in the first row. The data cells are color-coded: white (good), yellow (partial), and grey (poor). Some cells are highlighted in light green or light blue. The Excel interface includes the ribbon (File, Home, Insert, Page Layout, Formulas, Data, Review, View, Developer, PowerPivot) and the status bar at the bottom.

Slika 3.4: Pregled uspešnosti EMG meritev

Meritve so pobarvane glede na uspešnost: če je celica bela, pomeni, da je bila celotna meritev v redu; če je rumena, je bila meritev le delno v redu; pri sivih celicah pa je celotna meritev slaba (zunanji šum, okvare elektrod ipd.). Poleg te osnovne razdelitve, ki je bila prisotna v originalnem prikazu, smo dodali še dve, ki nam prikazeta, pri katerih meritvah so uporabili medikamente in pri katerih so ovce električno stimulirali. Medikamente smo označili s svetlo zeleno, stimulacijo pa z modro barvo. Dan, na katerega je ovca rodila, smo označili z debelo rdečo črto, da je hitro razvidna iz celotnega pregleda meritev.

3.3 Risanje iskrnih črt

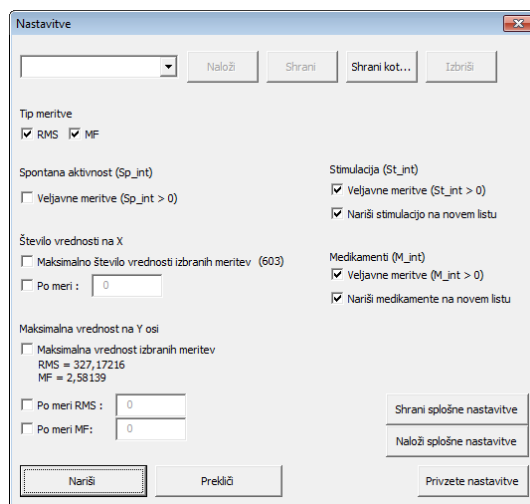
Ena izmed bistvenih novih funkcionalnosti za implementacijo je bilo risanje iskrnih črt za vsako posamezno meritev. Za vsako izbrano celico je bilo torej potrebno narisati (mikro)graf, ki prikazuje vrednosti te meritve.

Postopek za risanje iskrnih črt je sledeč:

- najprej označimo zelene celice oz. meritve;
- kliknemo na gumb “Nariši” v zgornjem levem robu trenutnega delovnega lista;
- prikaže se nastavitvena forma, kjer si lahko izberemo zelene nastavitve;
- narišemo s klikom na gumb “Nariši”.

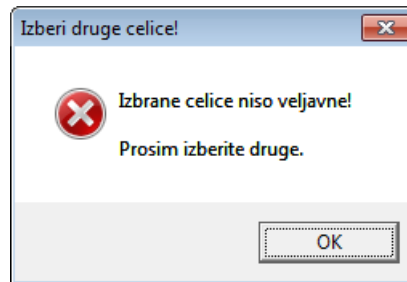
3.3.1 Nastavitvena forma

Nastavitvena forma (slika 3.5) se vedno prikaže, preden se narišejo iskrne črte za izbrane meritve. Omogoča veliko funkcionalnosti – od izbire tipa meritve in veljavnosti meritev do shranjevanja vseh nastavitvev, ki smo si jih izbrali.



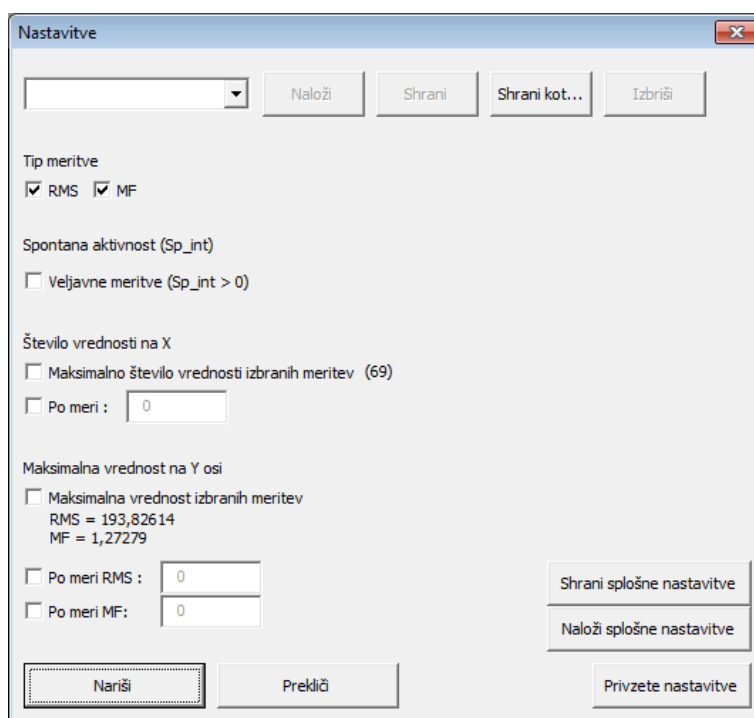
Slika 3.5: Nastavitvena forma

Preden se nastavitve sploh odprejo, se že preveri, ali smo izbrali pravilne celice oz. če je bila naša izbira celic v veljavnem območju (tj. kjer se sploh nahajajo meritve). Ob nepravilni izbiri celic se nam pokaže opozorilno okno (slika 3.6).



Slika 3.6: Opozorilno okno pri izbiri napačnih celic za izris

Če so bile izbrane celice v dovoljenem območju, se sprehodimo čez vse vrednosti izbranih meritev in si zapomnimo, za kakšne meritve gre (sponatana aktivnost, medikamenti, stimulacija) in kateri kanal (rog, vrat ali oba) smo izbrali. Glede na to ali smo izbrali meritve, kjer so bili dodani medikamenti ali je bila takrat stimulacija, se dodajo oz. odstranijo nastavitve za ta primera (slika 3.7).



Slika 3.7: Nastavitvena forma brez nastavitve za stimulacijo in medikamente

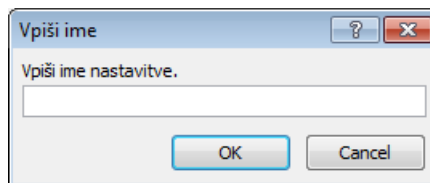
Poleg tega se izračuna maksimalno število vrednosti izbranih meritev, tj. katera izmed izbranih meritev ima največ vrednosti, in maksimalna vrednost pri teh izbranih meritvah. Tukaj se izračunajo maksimalne vrednosti tako za RMS (efektivna vrednost, angl. *Root Mean Square*) kot za MF (median-ska frekvenca, angl. *Median Frequency*), ki se glede na izbiro tipa meritve ustrezno skrivata, če ne želimo prikazati teh meritev.

Forma omogoča dva načina shranjevanja nastavitve skupaj z označenimi celicami. Tako se ob prikazu forme naložijo splošne nastavitve. Te nastavitve se bodo vedno naložile, ko se nam odpre forma. Lahko jih shranimo s klikom na gumb “Shrani splošne nastavitve” in ponovno naložimo s klikom na “Naloži splošne nastavitve”. Če pa želimo ponastaviti splošne nastavitve, to storimo s klikom na “Privzete nastavitve”, ki so fiksno zapisane v podatkovni bazi in se jih preko te forme sicer ne da spreminjati.

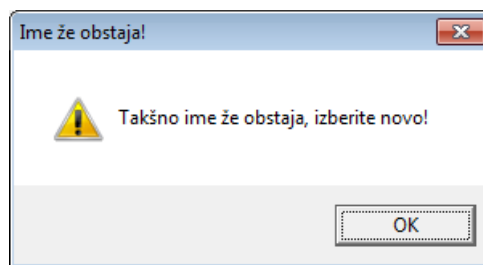
Drugo shranjevanje je shranjevanje nastavitve in celic, ki smo jih izbrali,

pod lastnim imenom, ki se ne sme ponavljati. Postopek takšnega shranjevanja je sledeč:

- Kliknemo na gumb “Shrani kot”. Prikaže se nam okno za vpisovanje imena nastavitve (slika 3.8). Ime mora biti enolično, drugače nas aplikacija na to opozori, kot to vidimo na sliki 3.9.

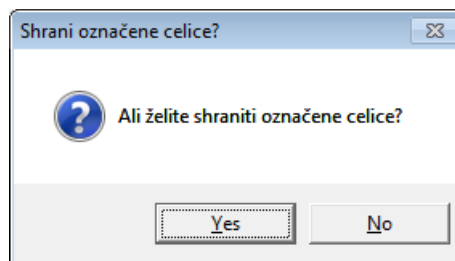


Slika 3.8: Okno za vpisovanje imena nastavitve, ki jo želimo shraniti



Slika 3.9: Opozorilno okno za že obstoječe ime nastavitve

- Ko vpišemo enolično ime, nas aplikacija vpraša, če želimo shraniti označene celice, ki jih lahko shranimo ali pa ne (slika 3.10).



Slika 3.10: Okno, ki nas vpraša, če želimo shraniti označene celice

Ko nastavitve uspešno shranimo, se nam v kombinirni izvlečni listi (angl. *combobox*) izbere pravkar shranjena nastavitve. Ko se iz seznama izbere katerakoli nastavitve, se omogočijo gumbi “Naloži”, “Shrani” in “Izbriši”. Pri vse treh nas aplikacija vpraša, če res želimo izvesti to akcijo nad izbrano nastavitvijo, pri nalaganju in shranjevanju pa nas vpraša še po celicah, če jih želimo naložiti oz. shraniti.

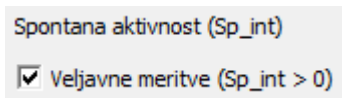
Za izbiro tipa meritve imamo na voljo že prej omenjena parametra RMS in MF. Kot vidimo na sliki 3.11, mora vedno biti označen vsaj en tip, torej imamo na voljo tri kombinacije. Ob spreminjanju tipa se nam ustrezno omogoči/onemogoči in skrije/prikaže nastavitve glede maksimalni vrednosti na osi y.

The image shows three panels illustrating different combinations of RMS and MF checkboxes and their corresponding maximum values on the Y-axis.

- Panel 1 (Left):** Both RMS and MF are selected. The maximum value on the Y-axis is 308,16676 for RMS and 2,02475 for MF. Input fields for 'Po meri RMS' and 'Po meri MF' are both set to 0.
- Panel 2 (Middle):** Only RMS is selected. The maximum value on the Y-axis is 308,16676 for RMS. Input fields for 'Po meri RMS' and 'Po meri MF' are both set to 0.
- Panel 3 (Right):** Only MF is selected. The maximum value on the Y-axis is 2,02475 for MF. Input fields for 'Po meri RMS' and 'Po meri MF' are both set to 0.

Slika 3.11: Kombinacije med tipoma meritev

Spontana aktivnost (slika 3.12) je določena vsaki meritvi. Lahko ima vrednost 0, kar pomeni, da ta meritev ni veljavna, vrednost 1 pomeni jutranjo meritev, 2 večerno meritev, 3 pa predstavlja meritev 120 do 60 minut pred porodom. Pri nastavitvah imamo možnost izbrati vse meritve ali pa samo veljavne, torej tiste ki imajo vrednost večjo od nič. V vsakem primeru se pri spremembi te nastavitve v ozadju ponovno izračuna maksimalno število izbranih meritev (os x) in maksimalna vrednost izbranih meritev (os y), saj se je s tem spremenila izbira meritev, ki ustrezajo nastavitvam.

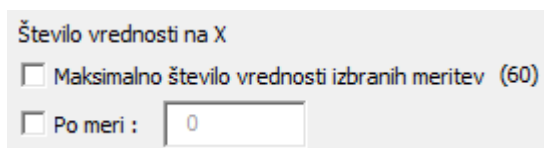


Slika 3.12: Nastavitev "Spontana aktivnost"

V naši rešitvi večinoma rišemo po več grafov naenkrat in jih med seboj primerjamo. Da pa je primerjava med različnimi iskrnimi črtami smiselna, je potrebno imeti razmerje na osi x in y na vseh grafih enako.

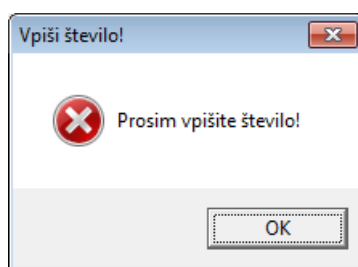
Za vrednosti na osi x imamo tri različne možnosti (slika 3.13):

- če ne izberemo maksimalnega ali po meri, ima vsaka meritev svoje število vrednosti na osi x;
- če izberemo "Maksimalno število vrednosti izbranih meritev", je število vrednosti na osi x enako številu v oklepajih, ki je bilo pred tem že izračunano;
- če izberemo "Po meri", lahko vnesemo poljubno pozitivno celo število.



Slika 3.13: Nastavitev "Število vrednosti na X"

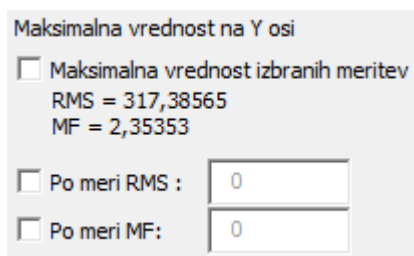
Izbiri "Maksimalno" in "Po meri" ne moreta biti izbrani hkrati. Pri izbiri "Po meri" moramo biti pazljivi, da res vnesemo število in ne črke; če se zmotimo, nas aplikacija na to opozori (slika 3.14). Podobno sporočilo dobimo, če vpišemo negativno število. Pri vpisu decimalnega števila pa se to samodejno zaokroži na najbližje celo število.



Slika 3.14: Okno, ki nas opozori na napačen vnos števila

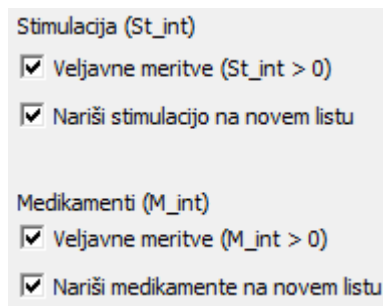
Pri maksimalni vrednosti na osi y imamo podobno izbiro kot pri osi x (slika 3.15).

- Če ne izberemo ničesar, ima vsak graf svojo maksimalno vrednost, kar ni najboljša izbira, saj tako ne moremo učinkovito primerjati grafov.
- Pri izbiri “Maksimalne vrednosti izbranih meritev” se maksimalna vrednost na grafu nastavi na že prej izračunane vrednosti in glede na to kateri tip meritve imamo označen, so tiste vrednosti ustrezno prikazane. Poleg tega se nam onemogoči možnost “Po meri”.
- Pri izbiri “Po meri” je zelo podobno, kot pri osi x, kjer se preverja pravilnost vnosa, namesto zaokroževanja decimalnih števil pa se decimalna pika zamenja z decimalno vejico, ker je ta uporabljena za nastavitvev pri grafih. Ob izbiri katerekoli vrednosti “Po meri” se onemogoči izbira maksimalne vrednosti.



Slika 3.15: Nastavitev “Maksimalna vrednost na Y osi”

Pri izbiri meritev lahko izberemo tudi tiste, kjer so bili dodani medikamenti ali pa je bila izvajana stimulacija (slika 3.16). Ker je teh meritev razmeroma malo v primerjavi z običajnimi, smo se odločili, da teh nastavitev ne prikazujemo, če te celice niso označene. V obeh primerih lahko označimo samo veljavne meritve, kjer se ob kliku, kot pri spontani aktivnosti, ponovno izračunajo maksimalne vrednosti za os x in y. Poleg te izbire imamo tu na voljo še možnost risanja teh grafov na ločenem delovnem listu, se pravi, da so na svojem listu samo tiste meritve, kjer so bili dodani medikamenti, in na svojem tiste, ko so bile ovce stimulirane.



Slika 3.16: Nastavitve za stimulacijo in medikamente

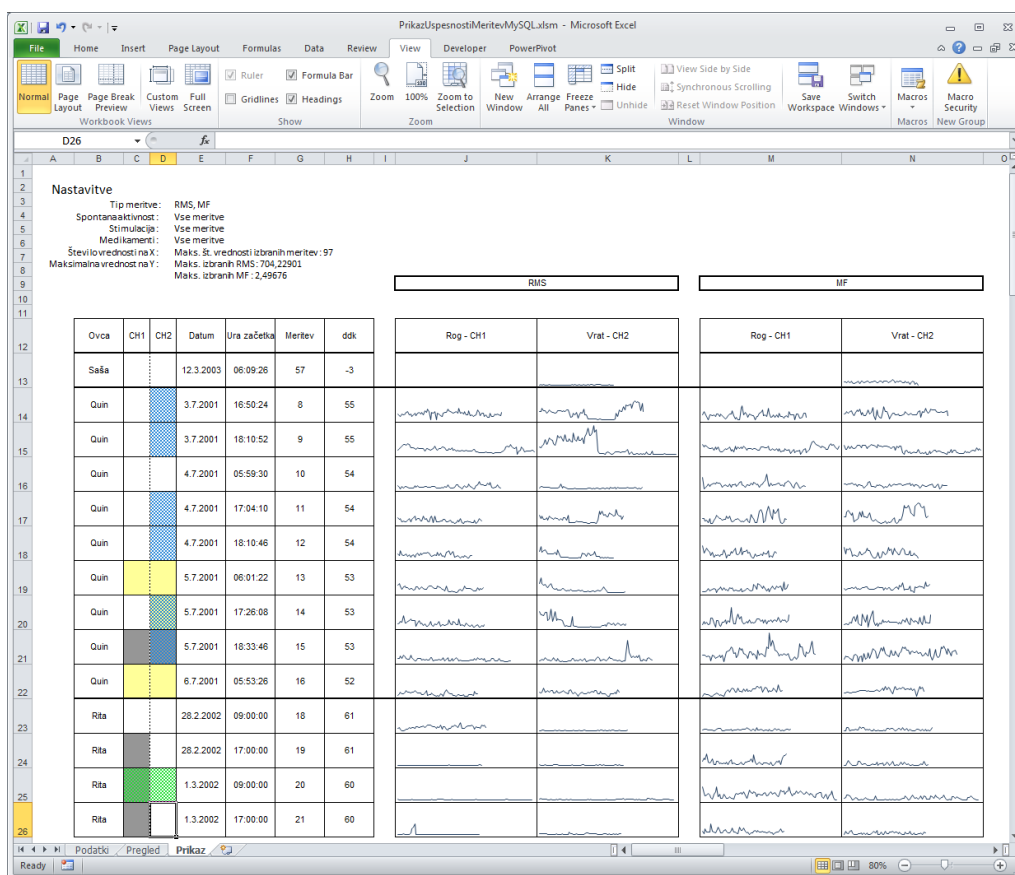
Na formi sta še dva gumba, in sicer “Prekliči”, ki nastaviteno formo zapre in ne narišemo ničesar, in “Nariši”, ki nariše naše izbrane meritve glede na nastavitve.

3.3.2 Izrisani graf

Primer izrisanega grafa vidimo na sliki 3.17. Najprej lahko opazimo, da imamo v zgornjem levem kotu delovnega lista našete vse nastavitve, ki smo jih izbrali za trenutni izris. Ker smo izbrali tako RMS in MF meritve kot oba kanala, lahko vidimo dva stolpca, RMS in MF, ki imata oba po dva kanala.

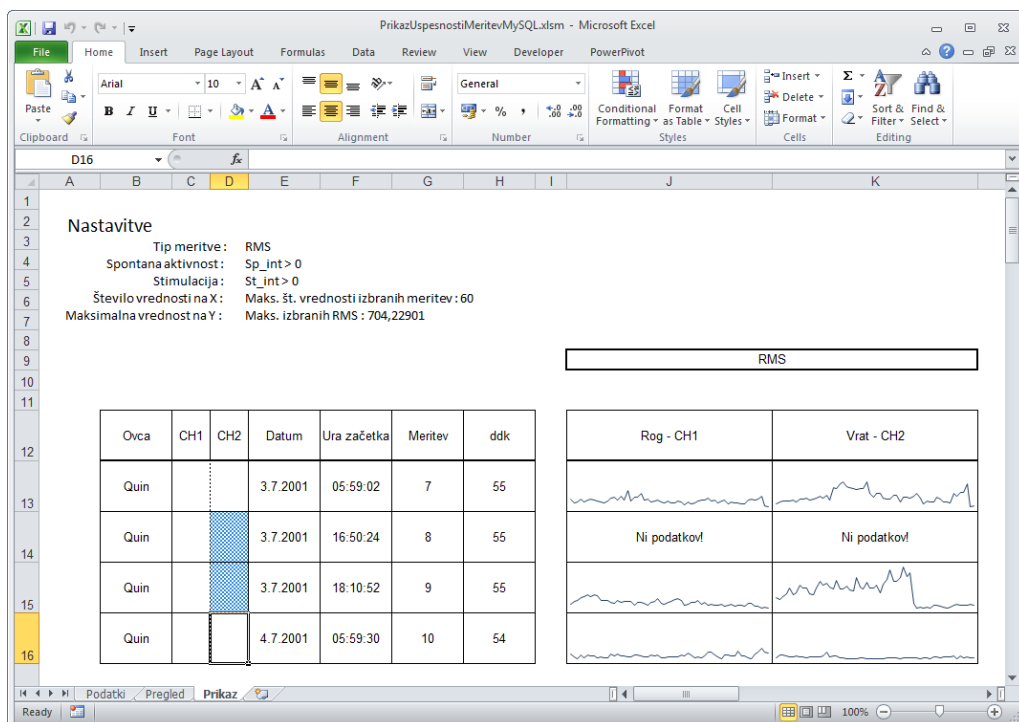
Za vsako meritev, tj. vsako vrstico, imamo na voljo:

- ime ovce, ki ji meritev pripada,
- barvo kanala, torej uspešnost meritve,
- datum meritve,
- uro začetka meritve,
- številko meritve,
- koliko dni je še do kotitve (negativne vrednosti pomenijo število dni po kotitvi).



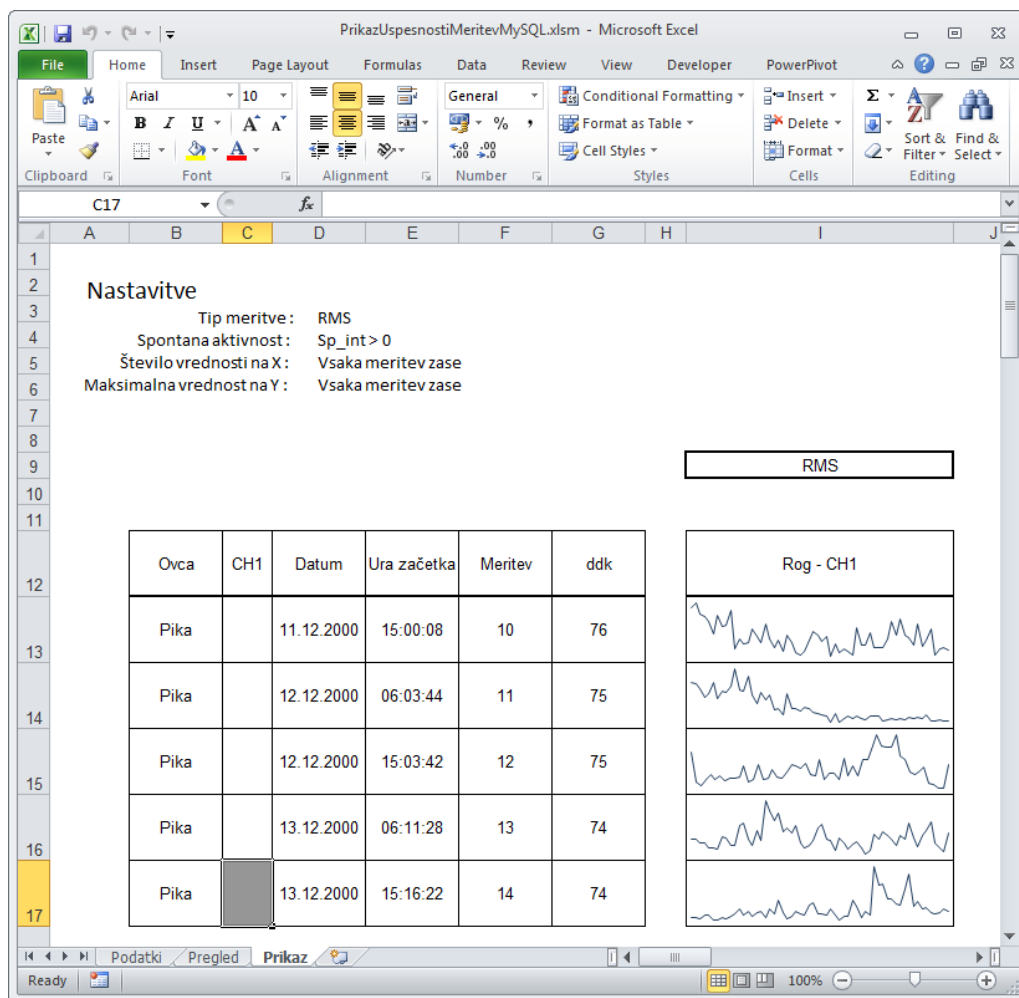
Slika 3.17: Prikaz izrisanih grafov iskrnih črt

Če smo pri izbiri meritev izbrali več ovc, potem so te ločene z odebeljeno črto. Pri izbiri nastavitve lahko naletimo na kombinacijo, ki za izbrano meritev ne bo vrnila ustreznih podatkov. V tem primeri se namesto iskrne črte izpiše “Ni podatkov!”. To lahko vidimo na sliki 3.18.



Slika 3.18: Primer izrisa iskrnih črt, kjer ni podatkov

Celotno risanje teh grafov poteka dinamično. Kot vidimo na sliki 3.19, imamo lahko narisane samo meritve RMS na kanalu 1 brez meritev MF in kanala 2.



Slika 3.19: Primer izrisa samo meritev RMS na kanalu 1

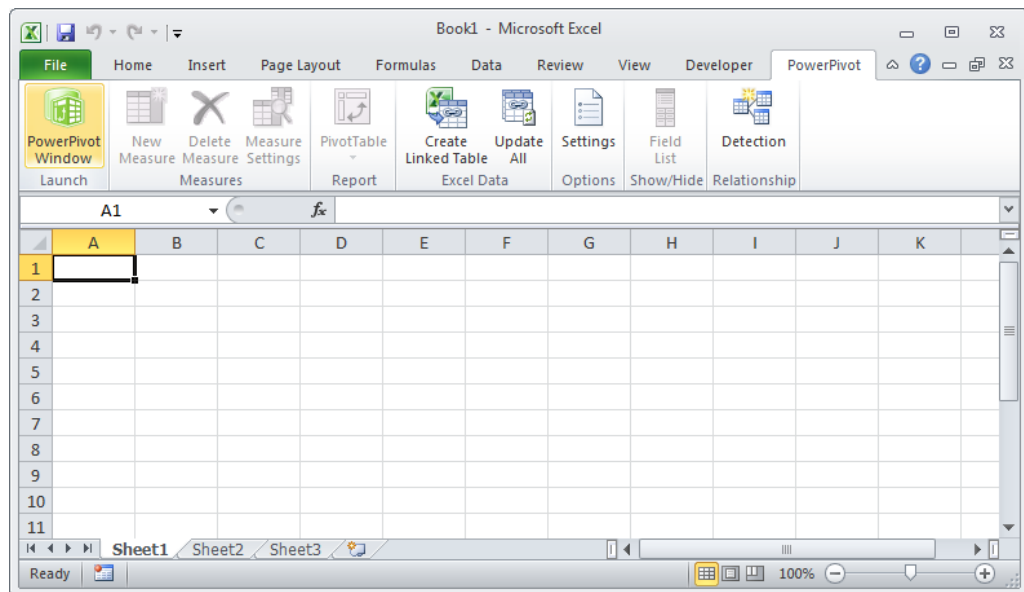
Poglavje 4

Uporaba dodatka PowerPivot

Opisano interaktivno okolje omogoča vpogled v vrednosti vsake posamezne meritve brez odvisnosti od drugih podatkov. Da bi prikazali čim več odvisnosti med podatki, smo se odločili uporabiti dodatek PowerPivot, ki enostavno zgradi OLAP kocko. Ta kocka vsebuje vse medsebojno odvisne podatke, ki jih potem vrača glede na izbrane filtre. Ko imamo želene podatke “na kupu”, jih z Excelom zlahka spremenimo v grafikone.

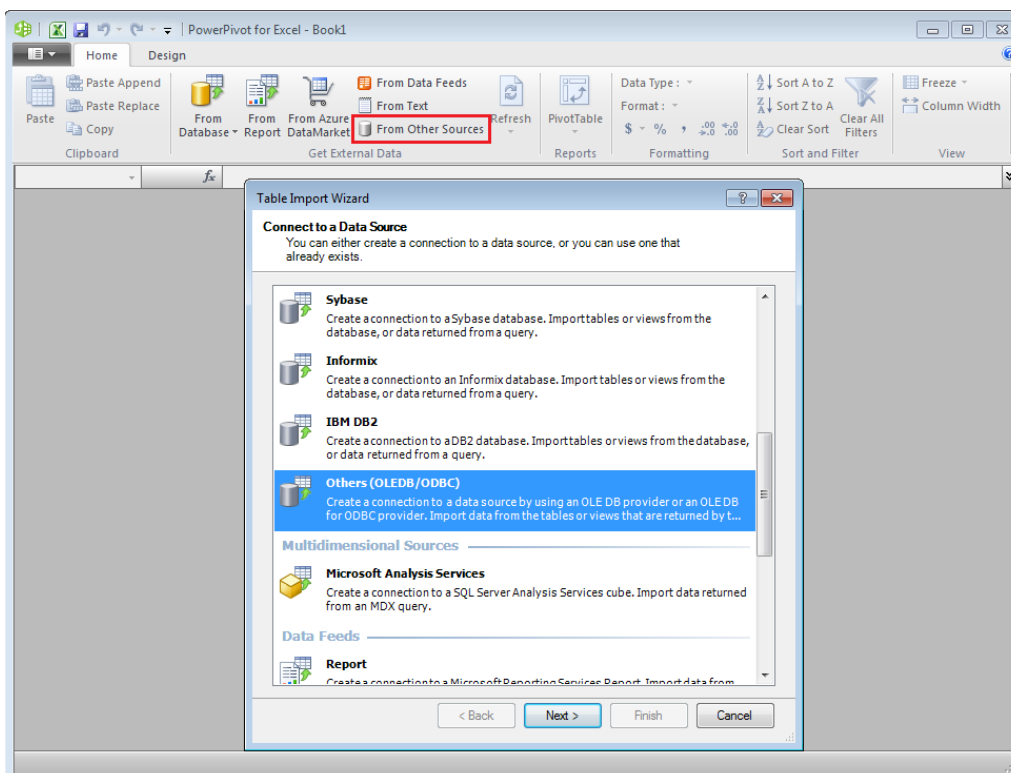
4.1 Uvoz podatkov

Za uvoz podatkov iz podatkovne baze MySQL je potrebno v operacijskem sistemu Windows najprej nastaviti vir podatkov ODBC (Open Database Connectivity). To storimo tako, da v nadzorni plošči (angl. *Control Panel*) izberemo skrbniška orodja (angl. *Administrative Tools*), kjer odpremo “Viri podatkov (ODBC)”. Tukaj dodamo nov zapis, kjer izberemo “MySQL ODBC 5.1 Driver”, izpolnimo potrebna polja za podatkovni strežnik in potrdimo. Ko nastavimo ta vir podatkov, lahko namestimo dodatek PowerPivot. Po namestitvi dodatka PowerPivot je bilo najprej potrebno uvoziti podatke iz podatkovne baze. To smo storili tako, da smo v Excelu kliknili na zavihek “PowerPivot” in nato na “PowerPivot Window”, s katerim odpremo okno za urejanje podatkov (slika 4.1).



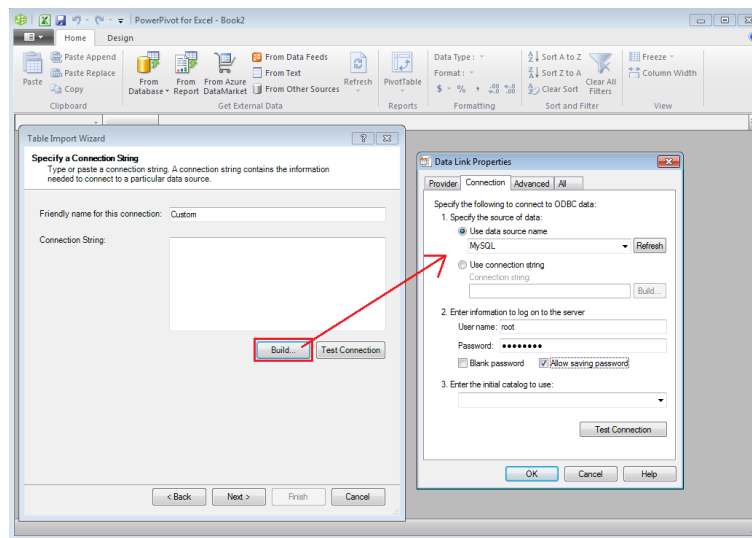
Slika 4.1: Dodatek PowerPivot v Excelu 2010

Nato smo izbrali možnost "From Other Sources", kar je odprlo čarovnika za uvoz podatkov, in za izvor podatkov izbrali "Others (OLEDB/ODBC)", kot je razvidno iz slike 4.2.



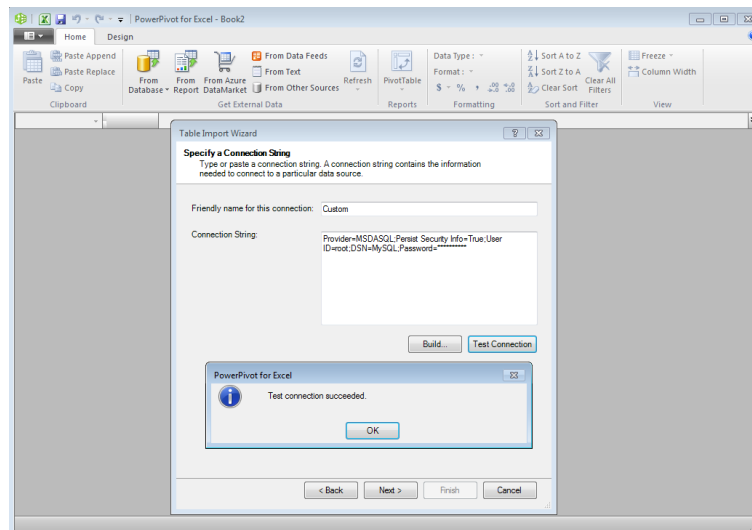
Slika 4.2: Izbira tipa podatkovnega vira

Po kliku na “Next” se nam je odprlo okno za izdelavo nastavitve o povezavi do podatkovne baze. Kot vidimo na sliki 4.3, se ob kliku na “Build” odpre okno z nastavitvami, ki jih je potrebno ustrezno izpolniti. Iz prve kombinirne izvlečne liste izberemo povezavo do podatkovne baze, ki smo jo določili pri namestitvi, v našem primeru “MySQL”. Izpolnimo še uporabniško ime in geslo ter obkljukamo shranjevanje gesla “Allow saving password”. Možnosti “Initial catalog” ne izberemo.



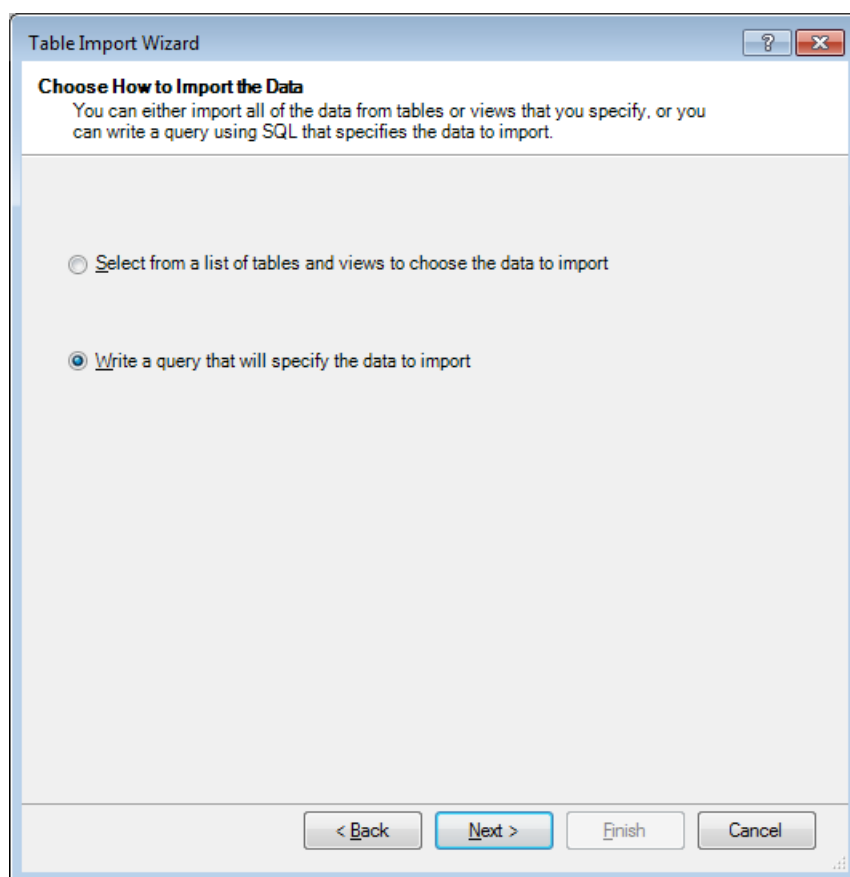
Slika 4.3: Nastavitve za dostop do MySQL

Po potrditvi nastavitve lahko povezavo še testiramo s klikom na “Test connection” in (kot vidimo na sliki 4.4) dobimo obvestilo o uspešni povezavi do podatkovne baze.



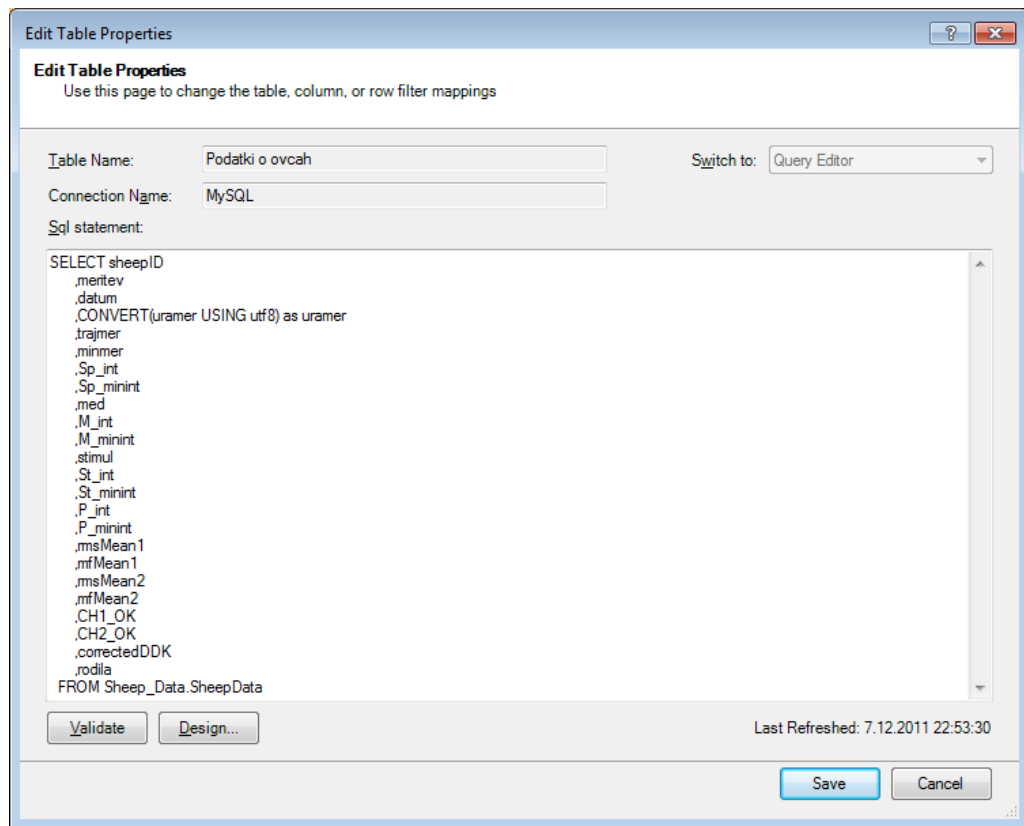
Slika 4.4: Testiranje povezave do strežnika

Kliknemo “Naprej” in pridemo do dveh možnosti (slika 4.5), kjer bomo izbrali tabele, ki jih bomo uvozili. Prva možnost, da bi sami “naklikali” tabele, žal trenutno ne deluje, ker ODBC gonilnik ne vrne prave sintakse in posledično pride do napake. Tako moramo za vse tabele, ki jih želimo uvoziti, ročno napisati poizvedbene stavke (angl. *query*).



Slika 4.5: Izbira med samodejnim in ročnim vnosom tabel

Dobro je, da so poizvedbe napisane v celoti – z imeni vseh stolpcev, ki jih potrebujemo. Ne smemo pa pozabiti pred imenom tabele napisati ime sheme/baze, kjer se ta nahaja (slika 4.6).



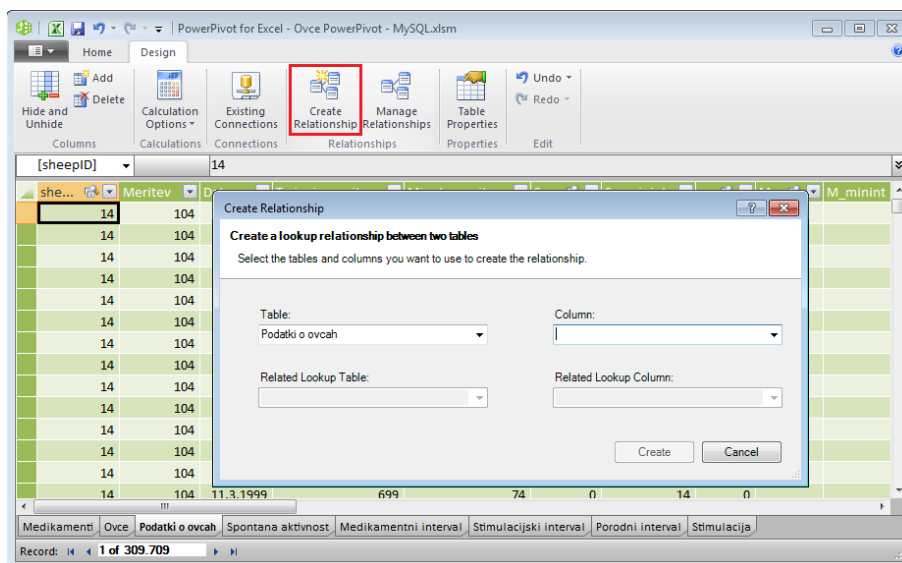
Slika 4.6: Primer napisane poizvedbe

Če pride do sprememb podatkov v podatkovni bazi, lahko podatke za PowerPivot enostavno osvežimo s klikom na “Refresh” brez ponovnega nastavljanja (slika 4.7).

she...	Meritve	Datum	Trajanje meritve	Minuta meritve	S...	Sp_minint	m...	M...
14	104	11.3.1999	699	61	0	1	0	
14	104	11.3.1999	699	62	0	2	0	
14	104	11.3.1999	699	63	0	3	0	
14	104	11.3.1999	699	64	0	4	0	
14	104	11.3.1999	699	65	0	5	0	
14	104	11.3.1999	699	66	0	6	0	
14	104	11.3.1999	699	67	0	7	0	
14	104	11.3.1999	699	68	0	8	0	
14	104	11.3.1999	699	69	0	9	0	
14	104	11.3.1999	699	70	0	10	0	
14	104	11.3.1999	699	71	0	11	0	
14	104	11.3.1999	699	72	0	12	0	
14	104	11.3.1999	699	73	0	13	0	
14	104	11.3.1999	699	74	0	14	0	
14	104	11.3.1999	699	75	0	15	0	
14	104	11.3.1999	699	76	0	16	0	
14	104	11.3.1999	699	77	0	17	0	
14	104	11.3.1999	699	78	0	18	0	
14	104	11.3.1999	699	79	0	19	0	
14	104	11.3.1999	699	80	0	20	0	
14	104	11.3.1999	699	81	0	21	0	
14	104	11.3.1999	699	82	0	22	0	
14	104	11.3.1999	699	83	0	23	0	
14	104	11.3.1999	699	84	0	24	0	
14	104	11.3.1999	699	85	0	25	0	
14	104	11.3.1999	699	86	0	26	0	
14	104	11.3.1999	699	87	0	27	0	
14	104	11.3.1999	699	88	0	28	0	
14	104	11.3.1999	699	89	0	29	0	
14	104	11.3.1999	699	90	0	30	0	

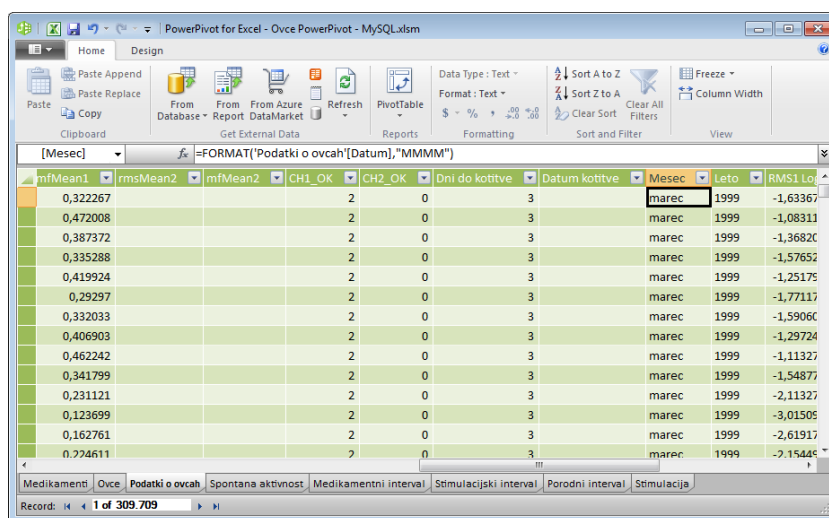
Slika 4.7: Uvoženi podatki z enostavnim osveževanjem

Po končanem uvozu podatkov oz. tabel v PowerPivot je potrebno nastaviti odvisnosti med tabelami. To storimo tako, da na zavihku “Design” kliknemo na gumb “Create Relationship”, kot lahko vidimo na sliki 4.8. Nato si izberemo tabelo in stolpec, ki ga povežemo s stolpcem iz druge tabele, ki vsebuje odvisne podatke glede na prvo izbiro.



Slika 4.8: Nastavljanje relacij med tabelami

Po nastavljenih relacijah smo – s pomočjo izrazov DAX – ustvarili dodatne stolpce, ki predstavljajo dodatne filtre pri prikazovanju podatkov na grafikonih. Primer vidimo na sliki 4.9, kjer smo naredili stolpec “Mesec”, ki vsebuje ime meseca iz datuma opravljene meritve.

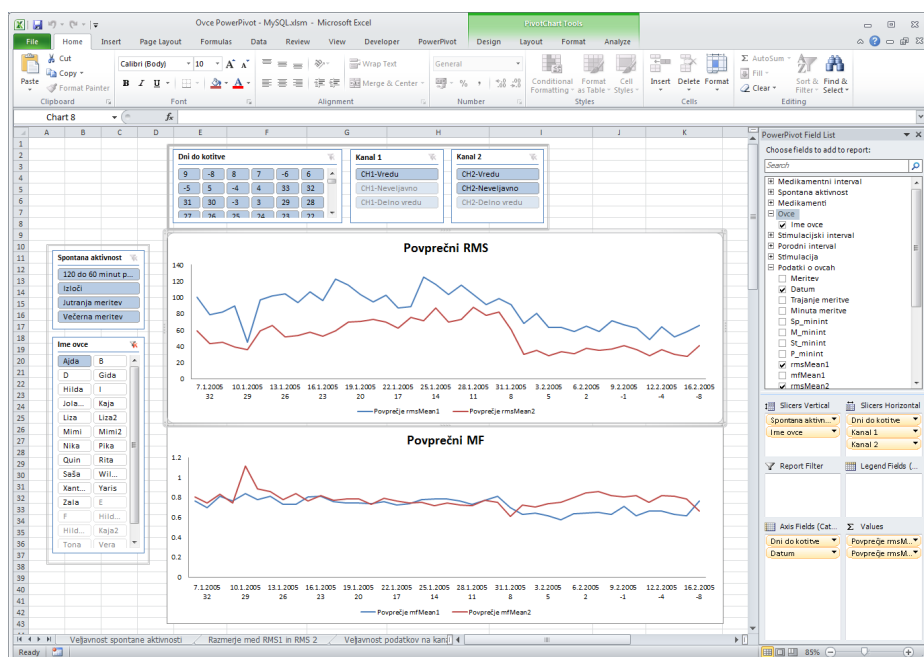


Slika 4.9: Uporaba DAX izrazov za ustvarjanje novim stolpcev

4.2 Praktična uporaba

Ko smo nastavili vse potrebno in imamo vse podatke v dodatku PowerPivot, pride na vrsto uporaba podatkov. Uporabili smo PowerPivotove vrtilne tabele, ki omogočajo enostavno izbiro stolpcev s seznama in njihovo razporejanje v zelena območja, npr. na os x, kot prikazane vrednosti ali med filtre.

Na sliki 4.10 vidimo primer, kjer smo za graf “Povprečni RMS” za os x izbrali vrednosti “Dni do kotitve” in “Datum”. Vrednosti predstavljajo povprečni RMS na kanalu 1 in kanalu 2. Imamo pa tudi razčlenjevalnike (angl. *slicers*), ki predstavljajo filtre. Na navpičnem imamo “Spontano aktivnost” in “Ime ovce”, na vodoravnem pa “Dni do kotitve”, “Kanal 1” ter “Kanal 2”. Ti razčlenjevalniki so še posebej priročni, če imamo več grafov, ki so odvisni od istih filtrov, saj jih lahko povežemo z njimi, kot smo v našem primeru storili z grafom “Povprečni MF”.



Slika 4.10: Grafa povprečnega RMS in MF s pomočjo vrtilnih tabel in razčlenjevalnikov

Poglavje 5

Sklepne ugotovitve

Pri izdelavi diplomske naloge smo z uporabo podatkovne baze MySQL, Microsoftove elektronske preglednice Excel in njegovega dodatka PowerPivot ustvarili dve interaktivni okolji za vizualno podatkovno analitiko, ki sta enostavni za uporabo in hkrati zmogljivi.

V prvem delu naloge smo iz množice podatkov enaindvajsetih ovac ustvarili pregled vseh uspešnih in neuspešnih meritev, ki vključuje tudi meritve z dodanimi medikamenti ali stimulacijo. Preko preglednice smo omogočili izbiro vsake posamezne meritve in nato njen izris s pomočjo iskrnih črt, pri čemer ima uporabnik na voljo številne nastavitve. V drugem delu naloge pa smo s pomočjo dodatka PowerPivot vse podatke združili v OLAP kocko, ki nam s pomočjo vrtilnih tabel in grafikonov omogoča enostaven in hiter prikaz najrazličnejših odnosov med podatki.

Ena izmed pomankljivosti naše rešitve je vsekakor odvisnost od oddaljene (strežniške) podatkovne baze. Dobro bi bilo, če bi lahko uporabili preprostejšo namizno rešitev, kot je npr. Accessova baza, a v tem primeru bi se zopet soočili s problemi glede hitrosti in nabora funkcionalnosti. Podatkovna baza MySQL namreč omogoča uporabo shranjenih funkcij in procedur.

Kot manjšo izboljšavo bi v prihodnosti lahko dodali možnost vpisovanja posamezne vrstice v podatkovno bazo, saj tako po spremembi podatkov ne bi bilo potrebno dolgo čakati na osveževanje. Več prostora za izboljšave ob-

stoječe rešitve je v grafičnih prikazih v okviru drugega dela (tj. povezanih z OLAP kocko), ki jih bo potrebno pred mednarodno javno objavo, ki jo načrtujemo, še skrbneje izbrati in dodelati, da bodo dosledno ustrezali načelom kakovostnega prikaza podatkov [18, 19].

Možnosti nadaljnega razvoja vidimo najprej v premiku izdelanih interaktivnih okolij na internet. To bi se dalo doseči s pomočjo Microsoftovega programa SharePoint, a zanj potrebujemo MSSQL strežnik. Kot alternativa bi lahko služilo gostovanje PowerPivot okolij (<http://hostedpowerpivot.com/>).

Nadaljnji korak bi lahko bila nadgradnja oziroma integracija interaktivne vizualizacije z zahtevnimi analitičnimi metodami [20, 21]. Lahko bi uporabili združevanje v skupine, analizo časovnih vrst in druge metode, ki so dostopne kot Microsoftove tehnologije Analysis Services v okviru MSSQL, lahko pa bi se naslonili na integracijo s statističnimi paketi, kot sta odprtokodna R (<http://www.r-project.org/>) in gretl (<http://gretl.sourceforge.net/>).

Literatura

- [1] J.M.Hellerstein in sod. (1999) Interactive data analysis: the control project. Computer 32 (8), 51-59.
- [2] R.Mazza (2009) Introduction to information visualization. London: Springer.
- [3] S.Few (2010) Now you see it: simple visualization techniques for quantitative analysis. Oakland, CA: Analytics Press.
- [4] W.S.Cleveland (1993) Visualizing data. Summit: Hobart Press.
- [5] E.R.Tufte (1998) The visual display of quantitative information, 16th printing. Cheshire: Graphics Press.
- [6] B.Shneiderman (1996) The eyes have it: a task by data type taxonomy for information visualizations. Proceedings of the IEEE Symposium on Visual Languages, str. 336-343, Washington: IEEE Computer Society Press.
- [7] D.Rudel (1996) Elektromiografska aktivnost materničnega vratu kot možen pokazatelj njegove zrelosti (doktorska disertacija). Ljubljana: Univerza v Ljubljani, Medicinska fakulteta.
- [8] B.Leskošek (2002) Elektromiografija maternice med gravidnostjo pri človeku in ovci (doktorska naloga). Ljubljana: Univerza v Ljubljani, Medicinska fakulteta.
- [9] E.R.Tufte (2006) Beautiful evidence. Cheshire: Graphics Press.

-
- [10] Microsoft Corporation (2012) Uvod v Excel 2010. Dostopno na:
<http://office.microsoft.com/sl-si/excel-help/uvod-v-excel-2010-HA010370218.aspx>
- [11] Wikipedia, The Free Encyclopedia (2011). Microsoft Excel. Dostopno na:
http://en.wikipedia.org/wiki/Microsoft_Excel
- [12] (2011) Kaj je MySQL? Dostopno na:
<http://www.e-uspeh.com/pomoc/kaj-je-mysql.htm>
- [13] Wikipedia, The Free Encyclopedia (2012) MySQL Workbench. Dostopno na:
http://en.wikipedia.org/wiki/MySQL_Workbench
- [14] D.M.Oppenheimer (2009) Introducing PowerPivot. Dostopno na:
<http://blogs.office.com/b/microsoft-excel/archive/2009/10/22/introducing-powerpivot.aspx>
- [15] Wikipedia, The Free Encyclopedia (2012) Visual Basic for Applications. Dostopno na:
http://en.wikipedia.org/wiki/Visual_Basic_for_Applications
- [16] Wikipedia, The Free Encyclopedia (2012) SQL. Dostopno na:
<http://en.wikipedia.org/wiki/SQL>
- [17] G.Vidmar, B.Leskošek, D.Rudel (2007) Visualizing time-course and efficacy of in-vivo measurements of uterine EMG signals in sheep. Lect Notes Comput Sci 4370, 183-188.
- [18] N.B.Robbins (2005). Creating more effective graphs. New York: Wiley-Interscience.
- [19] D.M.Wong (2010) The Wall Street Journal guide to information graphics: the dos and don'ts of presenting data, facts, and figures. New York: W. W. Norton.

-
- [20] J.Zhao, F.Chevalier, E.Pietriga, R.Balakrishnan (2011) Exploratory analysis of time-series with ChronoLenses. *IEEE Trans Vis Comp Graph* 17(12), 2422-2431.
- [21] W.Aigner, S.Miksch, H.Schumann, C.Tominski (2011) *Visualization of time-oriented data*. London: Springer.