Postup přípravy dat

Stáhnout okrsková data:

http://volby.cz/opendata/ps2010/PS2010_data_dbf.zip http://volby.cz/opendata/ps2010/PS2010ciselniky2010.zip http://volby.cz/opendata/ps2013/ps2013_dbf_20131026.zip http://volby.cz/opendata/ps2013/PS2013ciselniky20131021.zip http://volby.cz/opendata/kz2008/kz2008_data_dbf.zip http://volby.cz/opendata/kz2012/kz2012_data_dbf.zip

Pro každé volby je postup stejný, liší se jen pojmenování souborů. Proto bude v následujícím popisu vždy uváděn nejprve název v datech o krajských volbách a následně v závorce název pro parlamentní volby.

Otevřete soubor KZT6 (PST4) v excelu. Klikněte pravým na název souboru a zvolte "otevřít v programu".

Smažte zbytečné sloupce: TYP_FORM, OPRAVA, CHYBA, KC_1, KC_2, ZAKRSTRANA

Přejmenujte ostatní sloupce – do názvu přidejte rok (např. VOLICI_2008 namísto VOL_CELKEM)

Podle okresů odfiltrujte kraje:

Středočeský: 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112 Jihočeský: 3101, 3102, 3103, 3104, 3105, 3106, 3107 Plzeňský: 3201, 3202, 3203, 3204, 3205, 3206, 3207 Karlovarský: 4101, 4102, 4103 Ústecký: 4201, 4202, 4203, 4204, 4205, 4206, 4207 Liberecký: 5101, 5102, 5103, 5104 Královéhradecký: 5201, 5202, 5203, 5204, 5205 Pardubický: 5301, 5302, 5303, 5304 Vysočina: 6101, 6102, 6103, 6104, 6105 Jihomoravský: 6201, 6202, 6203, 6204, 6205, 6206, 6207 Olomoucký: 7101, 7102, 7103, 7104, 7105 Zlínský: 7201, 7202, 7203, 7204 Moravskoslezský: 8101, 8102, 8103, 8104, 8105, 8106

	- 7	• (° -	Ŧ						vył	perte zálo	žku data a	a klepně	te na fi	tr
Soul	bor	Domů	Vlože	ní I	Rozlože	ní strár	nky V	zon e Da	ta Revize	Zobrazer	IÍ (BBYY P	DF Transfor	mer+	
apli Acce	A kace ess	Z webu te	Z Z extu zd	jiných Irojů ▼	Existuj připoje	jící Al ení	ktualizov vše *	Deripoje	ení Azosti t odkazy	Z A Seřadit	Filtr	mazat užít znovu řesnit	Text do C sloupců d	debrat uplicity
		Načíst e	xterní d	ata		6 0	VDEC	Připojení		Seřa	dit a filtrovat			Da
		A1	•	<u> </u>	J	Ger O	KRES							
1		B	C	D	E WD		G	H		J	K	L	M	
$\frac{1}{2}$	2101	513/82	1	150	66	66	66	_CELK						
2	2101	529303	1	966	222	222	332							
4	2101	529303	2	834	337	337	332							
5	2101	529303	3	633	253	252	250							
6	2101	529303	4	719	292	292	291							
Z ap Ac	A olikace	e Z webu	Z textu	Z jinj zdroj	ých E jů * pi	xistujío řipojer	í Aktu ní v	alizovat še •	Vlastnosti Upravit odka	Z↓ Z Z↓ Se	řadit Filtr	To Pou	žít znovu esnit	Text do C sloupců c
		Naci	st exteri	ni data		-		Pripo	ojeni		Seradit a	filtrovat		
		E2		eledu tedu	o nak	≴ڑ ریالہ	66							
	1	A		Dait		JUK	- <u> </u>	E	F	G	Н	1		J
1	OK	RE	BEC	• OI	KRSE 1	- VO	L_S 💌	VYD_C -	ODEV2 -	PL_HL_	ELK			
Åz↓	Seřa	a <u>d</u> it od ne	jmenšíh	io k ne	jvětším	u		66	66	66				
Z A↓	S <u>e</u> řa	adit od ne	jvětšího	k nejr	nenším	u		333	333	332				
	Seřa	adi <u>t</u> podle	barvy				•	337	337	332				
1	Vur	nazat filtr :						253	252	250				
	<u>v</u> yn	aust so all		_				292	292	291				
	Filtr	ovat podi	e barvy					262	262	261				
	<u>F</u> iltr	y čísel					•	282	282	279				
	Hle	edání					Q	308	308	307				
		4207	~	. v	ybert	te ok	(resy)	patřící	225	225				
	1	5101		d	Io vai	šeho	kraje	, zde	273	273				
		✓ 5102		v	ybrái	n Pa	rdubio	ký kraj,	260	254				
		5104		a	odk	likně	te Ok	(57	56				
		5201			-			223	223	222				
		5202						290	290	289				
		5203			- ()		-	316	316	315				
					V			372	372	371				
				OK		Sto	orno	318	317	314				
				2				252	252	251				
20		2101	52930	03	1	8	863	357	357	356				
-		A 4 6 4			-	~	0.40	5.05	505	500				

Výběr označte (CTRL+A), a nakopírujte do nového okna (CTRL+N, CTRL+V)

Nové okno uložte jako xls nebo xlsx dokument např. pod názvem UCAST_ROK (*pojmenujte si to, jak chcete, tento název ale bude používán dále v pokynech*)

Otevřete soubor KZT6p (PST4p) v excelu. Klikněte pravým na název souboru a zvolte "otevřít v programu".

Smažte zbytečné sloupce:TYP_FORM, OPRAVA, CHYBA, KC_1, KC_2, HLASY_01, ..., HLASY_70, KC_3, KC_4, KC_5, POSL_KAND, KC_SUM

Použijte filtr na výběr okresů v kraji.

Použijte filtr na výběr strany.

(Čísla stran naleznete v souboru KZRKL - v datech za poslanecké volby je v číselnících)

Vždy vyberte jen jednu stranu, výběr nakopírujte do nového okna a uložte pod názvem STRANA_ROK.

Otevřete soubory v SPSS

ta Untitle	ed1 (Data	Set0] -	IBM SPSS Sta	itistics Data	a Editor	-	
<u>File</u> d	it <u>V</u> iew	<u>D</u> ata	Transform	Analyze	Direct Marketing	<u>G</u> raphs	
New				*		2 H	
				ĸ	🔁 Data		
:e Oper	n Data <u>b</u> as	se		•	Syntax		
Read	l Text Da	ta				var	
🔄 Open Data		-	-	012			×
Look in:	z2008 data	a dbf			a fa 11: 1=		
Kscm08 x		_		_			
kscm 08.	xlsx						
ucast_08.	xlsx		-	3			
🖉 zaklad08.)	dsx			-			
				vyberte,	že chcete otevřít	data v	
				excelu, z	volte soubor k ot jej odkliknutím "	evření a 'open" –	
						open	
File name:							 <u>O</u> pen
Files of type:	Excel (*.x	ls, *.xlsx	*.xlsm)	7			Paste
Encoding:	Excel (*.x	ls, *.xlsx	*.xlsm)	V			Cancel
	Lotus (*.w	/^) ~)					Help
	dBase (*.)	dbf)					

Po otevření tabulek začněte se spojováním.

Nejprve je nutné seřadit data v každé tabulce podle sloupců OKRSEK a OBEC

	🕼 *Untitled	12 [DataSet1] - IBM SPS	SS Statistics Data Edito	or	_				-		
l	<u>F</u> ile <u>E</u> dit	<u>View</u> <u>Data</u> <u>Transfo</u>	orm <u>A</u> nalyze Direc	t <u>M</u> arke	eting <u>G</u> raphs	Utilities	Add- <u>o</u> ns	Window	<u>H</u> elp		
l	🔁 H	🖨 🛄 🖛	A K		H	*5		st 🛄		Ø	
l	1 : OKRSE	К 1					_				
L		OKRES	OBEC				EZNIAM	VYD_0	OBALKY	1	ODEV
	1	5101	513890		Cut		klikněte	pravým	na	6	
	2	5101	514161		Copy •		název sl	oupce O	KRSEK	!3	
ł	3	5101	514195		Paste		a zvolte	"Sort		36	
	4	5101	514276		Clear		Ascendi	ng". Toté	éž pak)7	
L	5	5101	530387		🔛 Insert Var	iable	provedt	e ve slou	ipci	4	
	6	5101	544337		Sort Asce	nding	obec			8ز	
l	7	5101	544337		Cart	anding	47			41	
L	8	5101	546232		Son Dest	ending	07	7		49	
l	9	5101	546259		Descriptiv	es Statist	ics 54	ŧ		62	
l	10	5101	546267		Spelling		21			73	
l	11	5101	546275		1		81			47	
	Data View	Variable View									

Připojování začněte od tabulky UCAST_ROK (tabulky STRANA_ROK totiž obvykle neobsahují všechny obce v kraji, ale jen ty kde strana získala alespoň jeden hlas)

- Onutied		isetti - Ibivi SPSS Statistics Data Euritor						
<u>F</u> ile <u>E</u> dit	View	Data Iransioni Analysic Direct Marketing	<u>G</u> raphs	<u>U</u> tilities	Add- <u>o</u> ns	Window	<u>H</u> elp	
		Define Variable Properties	A A	*	4	S 🗰		
1: OBEC		24 Set Measurement Level for Unknown				<u> </u>		
		🔚 Copy Data Properties	EK	VOL S	SEZNAM	VYD	OBALKY	ODE
1		New Custom Attribute	okně s	otevřen	ým soub	orem UC	AST_ROK	
2		🗟 Define Dates V	nabídc	e "Data'	' zvolte p	oložku '	'Merge Fil	es"
3		Hefine Multiple Response Sets	poté "/	Add Vari	iables"			
4		Validation ►			200	,	51	
5		🔡 Identify Duplicate Cases	1		107	7	44	
6		🔝 Identify Unusual Cases	1		117	/	42	
7		📆 Compare Datasets	1		309)	139	
8		Sort Cases	1	[7	597	7	250	
9		Sort Variables	1	V	334	l	138	
10		Transpose	1	V	129	,	58	
11		Merge Files		Cases	147	1	41	
12		Restructure		Variables	180	r)	100	
14		Aggregate	Add 1	vanables.	1597	7	721	

Add Variables to Untitled2[DataSet1]			X
Select a dataset from the list of open datasets o	from a file to merge with th	ne active dataset	
Untitled3[DataSet2]			
◎ <u>An</u> external SPSS Statistics data file			
			Browse
Non-SPSS Statistics data files must be opened	in SPSS Statistics before th	ney can be used as par	t of a merge.
Continue	Cancel Help		
E402 E44477	4 4507	704	

Add Variables from DataSet2	alle e 2 a 🖬 -	
Excluded Variables:		New Active Dataset:
OBEC(+) OKRES(+) OKRSEK(+)	Zaškrněte políčka "Match Cases key Variables" a "Cases are sort " OBEC a OKRSEK označte a vložte do okna Key Variables Renzie	on ed s jej CKRES(*) VOL_SEZNAM(*) VYD_OBALKY(*) ODEVZ_OBAL(*) PL_HL_CELK(*) KSTRANA(+) POC_HLASU(+) OBEC(*) OKRSEK(*)
Capes are sorted in orde	r of key variables in both datasets	Key <u>V</u> ariables:
 Both files provide cases Non-active dataset is key Active dataset is keyed taget 	ved table able	
Add Variables from DataSet2		×
Excluded Variables:		New Active Dataset:

OKRES(+)	¥	OKRES(*) VOL_SEZNAM(*) VYD_OBALKY(*) ODEVZ_OBAL(*) PL_HL_CELK(*) KSTRANA(+) POC_HLASU(+)
Rename Match cases on key variables Cases are sorted in order of key variables in both datasets Both files provide cases Non-active dataset is keyed table Active dataset is keyed table	*	Key Variables: OBEC OKRSEK
Indicate case source as variable: source01 (*)=Active dataset		
(+)=DataSet2	Help	

Opakujte pro všechny další tabulky

Pozor – v Moravskoslezském kraji v roce 2011 vznikla obec Libhošť, ve zlínském pak v roce 2013 obce Krhová a Poličná

Spojování tedy začněte se soubory za poslanecké volby 2013

Spojený soubor uložte jako excel.

V excelu je nutné provést souhrn – seřaďte data dle obce.

Soi	ubor	Doi	mů Vl	ožení	Rozložení	stránky	Vzorce	Data	Revize	Zobraz	ení	ABBYY	PDF Trans	former+				
ap Ac	A likace cess	Z webu Na	Z textu číst exter	Z jiných zdrojů v ní data	Existující připojeni	Aktua vš	izovat e ▼ Připoje	řipojení lastnosti Ipravit odk	AZ AZY	Seřadit Se	Filtr řadit a	▼ 🐼 V ∏o Pi IV U a filtrovat	ymazat oužít znov I přesnit	u Text c sloup	do Odebra ců duplicit	at Ověře y dat atové n	ení Slouč ástroje	it Citliv
		B2		- (e)	f _x	5134	82											
		Δ	R	(-		E	E	G					_	V			
L	OKR	ES	OBEC	OKR	SEL Seřad	it				1214							2	X
2		2101	5134	32														
2		2101	52930	13	₽. 2	<u>P</u> řidat ú	roveň 🗙	<u>O</u> dstranit	úroveň	🕒 <u>К</u> орі́г	rovat ú	iroveň		Možn	osti	🗸 Data	obsa <u>h</u> uji	záhlavi
1		2101	52020	12	Slow	noc			Ď	azoní				Pořad	· · · · ·			
-		2101	52930	22	Sořa	dit nodle	CORE C		A.							1	*****	
2		2101	52930	J3	Jeru	uit poule	OBEC			loanoty			·		ejmensino	k nejvet	simu	
þ		2101	52930	03														
7		2101	52930	03														
3		2101	52930	03														
)		2101	52930	03														
0		2101	52930	03														
1		2101	52930	03														
2		2101	52030	13	1													
2		2101	52930	22														
3		2101	52930	J3											0	К	Sto	rno
4		2101	52930	03	1				_								_	-
S	oubor	Domů	Vložení	Rozložení	stránky Vzo	rce Dat	a evize	Zobrazení	ABBYY PDF	Transformer+					-	~		
3		*				Dřipoje 🗋	ní 👌 🔀		隊 Vyma	zat	-+		jì 👬) *i	1	📰 🐴 Z	obrazit pod	robnosti
Za	plikace	z	Z Z jiný	ch Existujíc	í Aktualizovat	Vlastno	sti 🏹 Se	řadit Filtr	la Použí	t zno v záložce	e "Dat	a" klikn	ěte na "So	ouhrn"	o delete So	uhrn	krýt podrob	nosti
A	ccess	webu	textu zdroji	ů ∗ připojen	ií vše *	Se Upravit	odkazy	Cožodit o f	¥ Upřes	snit sloupe	ů duplicit	ty dat • Datavá nástr	analýz	a * *		200000		-
				£	OKRES	npojeni	Souhrny	Seraultai	IIII Ovat	? X		Jacove nasu	oje			JSHOVA		101
			B	C	D	F				-				~		0	p	0
1	OKR	ES OE	BEC OK	RSEK VO	L_SEZNA VYE	OBALK		ve slouper.			D8 71	poiozce olte OBI	u kazde i FC	zmeny ve	sioupci	510 V	V10	TOP10
2	8	3104 5	00259	1	1589	620	Použít funkci:			<u> </u>	19,					17	193	48
3	8	3106 5	00291	1	1001	434	Součet				70	použiite	funkci sou	ičet ¹¹³	70	17	70	14
4	8	3106 5	00291	2	1129	624	Při <u>d</u> at souhrn d	o sloupce:			69			.3	69	17	69	11
5		7102 5	00496	1	701	258	OBEC				35	35	35	61	35	17	35	
0	-	7102 5	00496	2	592	238	OKRSEK	AM			35	35 ⊿7	35	58 79	35	17	35 47	
8	-	7102 5	00496	4	919	397	YD_OBALK	Ŷ	-	=		označte v	všechny sl	oupce kt	eré chcete	7	-47	1:
9		7102 5	00496	5	672	292	PL_HL_CEL	K			289 s	sečíst (tj.	všechny)	kromě ol	bce a okrs	ku 7	289	56
10) 7	7102 5	00496	6	641	236	Nahradit ak	tuální souhrny			123					7	122	29
11	1 7	7102 5	00496	7	999	477	Konec strán	iky mezi skupin	ami		58 0	odklikně	te ok	±	55	_7	56	17
12	2	7102 5	00496	8	899	405	Celkový <u>s</u> ou	ihrn pod dat,			63	63	62	117	63	17	62	11
13		102 5	00496	9	/59	354	Odebrat vše	ОК		Storno	118	114	111	287	118	17	111	28
14		7102 5	00496	10	621	361		J. J.	1		30	30	30	121	30	17	30	10
16	5	7102 5	00496	12	999	451	451	448	676	316	66	66	65	119	66	17	65	11

Po několika hodinách bude souhrn dokončen.

Označte všechna data (CTRL + A) a zkopírujte je do nového listu jako čísla.

Soubor	Do	mů ∨l	ožení	Rozložení s	tránky V	zorce Da	ata Re	vize	Zobraze	ení	АВВҮҮ Р	DF Transforn
	*				ß	Dipoje	ení osti	₹↓	A Z A	Y	家 Vy 長 Po	mazat užít znovu
Z aplikace Access	Z webu	Z J textu	Z jiných zdrojů ∗	Existující připojení	Aktualizova vše *	ent 🕬 Upravi	it odkazy	Z A↓	Seradit	Filtr	ゾ Up	přesnit
	NL C-UI	číct outor	of data	× 100	==	Připojení			Seř	adit a f	iltrovat	
		ori - 11	• A	А 📆 🗸	% 000 ≞≞≝							
	в	I≣	🗞 - A	- <u> </u>	° 400 ≪ 400 ≪	E	F		G		Н	I
1	v)			kliknutí	pravým t	tlačítke	em n	nyši do			
2		vyjmo <u>u</u> t			první bu	iňky			-			
3		Kop <u>i</u> rova	at		klepnut	í na složk	a 123 🕯	v po	ložce			
4		Možnos	ti vložení:		možnos	ti vložení	í	-				
5		12	$3 f_x$	* • %	63							
6		Vložit	/	1								
7		Ho	dnoty (H)									
8		<u>V</u> ložit ko	pirované l	bunky								
9		Od <u>s</u> trani	it									
10		Vy <u>m</u> azat	obsah									

Posledním krokem je nutnost zbavit se dat, ze kterých je souhrn vypočítán

Sou	ibor Dom	nů Vložení	Rozl	ožení	stránky	Vzorce	Data	Revize	Zobrazení	ABBY
ж					2	P 👰	Připojení /lastnosti	≵↓	A Z A	T to
Z ap Acc	likace Z cess webu	Z Z Jinj textu zdro	ých Ex iů ≠ při	istujic innien	i Aktualizo ní vše v)vat . ∞ (Jpravit od	lkazy X↓	Seřadit F	iltr 📝
	Nač	íst externí data	a '	Calib	ori - 11	т А́ А	x 🖫 -	% 000 🖷	Seřad	it a filtrov
	C1	• (*		B	I 🗏 🖄	- A -	· 🖂 - 🐧	,00, 00, 00, 00,		
	А	в 🕻	С		D	L	ik prav	- vím na n	ázev slou	
1	OKRES	OBEC	OKRSE	ŵ	Vyjmo <u>u</u> t		ik piav	slounce		pce
2	8104	500259			Kop <u>í</u> rovat	۹۳ -V	ložit bi	iňky - tí	n se nřid	á
3		500259 Ce	lkem	<u></u>	Možnosti	vložer	rázdný -	sloupec	in se prio	u
4	8106	500291			123	fx	<u>*</u> 8 %	SIGUPCC	427	9
5	8106	500291			Vlažit iinal			23	622	8
6		500291 Ce	lkem		VIOZIL JINAK				1049	18
7	7102	500496		•	Vložit <u>b</u> uňl	ky		58	254	6
8	7102	500496			Od <u>s</u> tranit			37	237	7
9	7102	500496			Vy <u>m</u> azat o	bsah		32	231	6
10	7102	500496			Formát bu	něk		94	394	5
11	7102	500496			čomat Du	IIEK		92	290	6
10	7100	E00406			Sirk <u>a</u> sloup	oce		26	225	C

Soι	ubor Dor	nů Vložení Rozložer	stránky Vzorce Data Revize Zobrazení ABBYY PDF Transformer+	
	A		Připojení ŽI ŽI Vymazat Vlastnosti Vlastnosti Valence Valence	
Zap	likace Z	Z Z jiných Existují	Aktualizovat Sveravit odkazy A Seřadit Filtr Text do Ogebrat Ověření Sloučit Citlivos	stní Si
ACI	Na	číst externí data	Připojení Seřadit a filtrovat Datové nástroje	-
	B1	1. vyber	sloupec obec klikněte na "Text so	
			sloupců", v nabídce vyberte	
1		BC	pevná šířka a klikněte na	C 5 5 1
2	8104	500259	Průvodce převodem textu do sloupců (1/3) "další" a po té na "dokončit"	
3	010	500259 Celkem		
4	8106	500291	Průvodce prevodem textu zjistil, ze data jsou oddelena.	
5	8106	500291	Zvolte datový typ, který datům odpovídá nejlépe, a potom kliknéte na tlačitko Další.	
6		500291 Celkem	Zdrojový datový typ	
7	7102	500496	Vyberte typ souboru, ktery datům nejlepe odpovida:	
8	7102	500496		
9	7102	500496		
10	7102	500496		
11	7102	500496		
12	7102	500496		
13	7102	500496	Nanied vybranych dat:	
14	7102	500496		
15	7102	500496	3500259 Celkem	
16	7102	500496	4 500291	
1/	/102	500496	6500291 Celkem	
18	7102	500496		
19	7102	500496	Storno Zočt Dakovčit	
20	7102	500496		
21	7102	500490	16 070 440 440 447 062 257 15 15	

Nakonec seřaďte data podle nového sloupce, v tomto případě sloupec C a smažte řádky, které slovo celkem neobsahují

Spočítejte relativní zisky stran (STRANA_ROK/PL_HLASY_ROK*100)

K výslednému souboru můžete v spss připojit data ze Sčítání lidu, případně další data.