



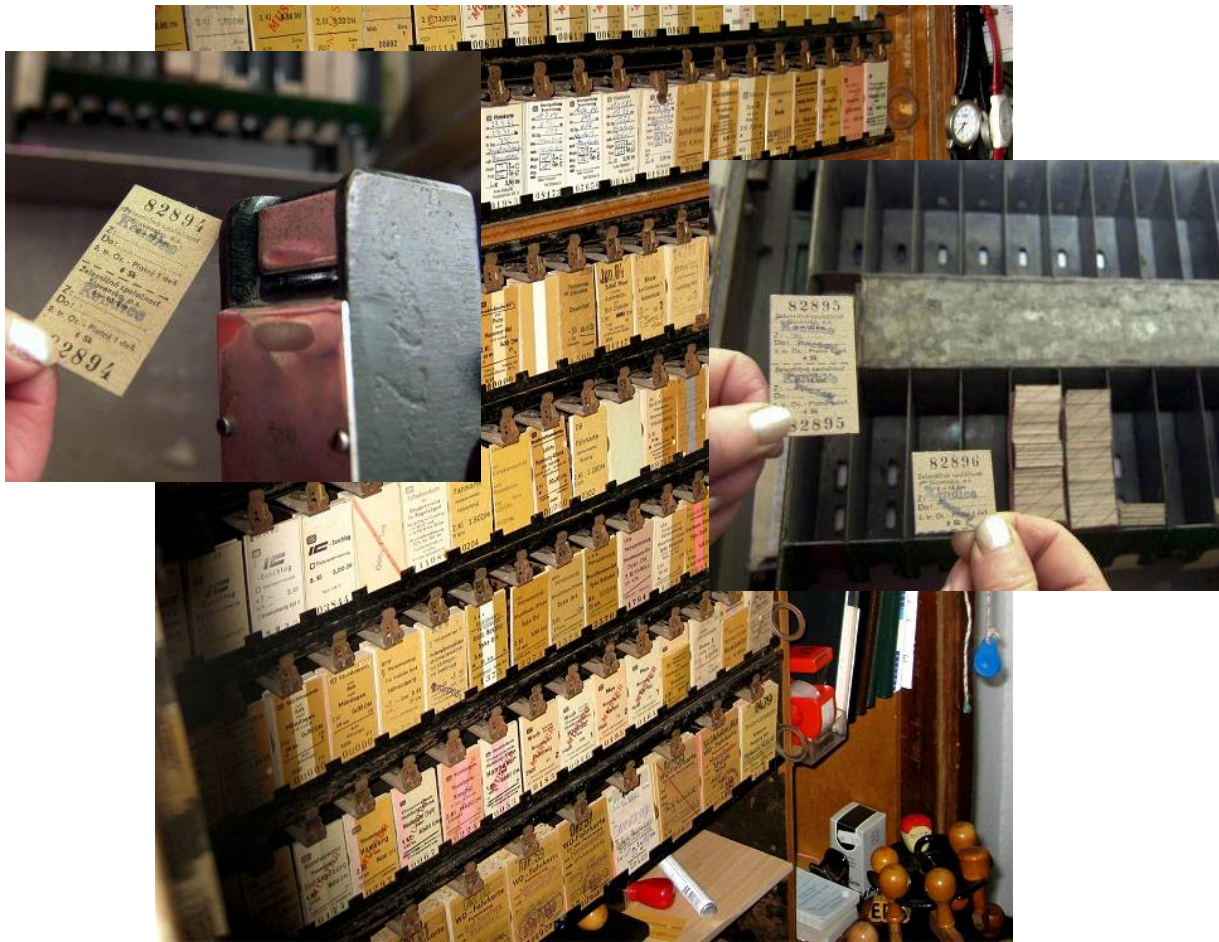
**Hewlett Packard  
Enterprise**



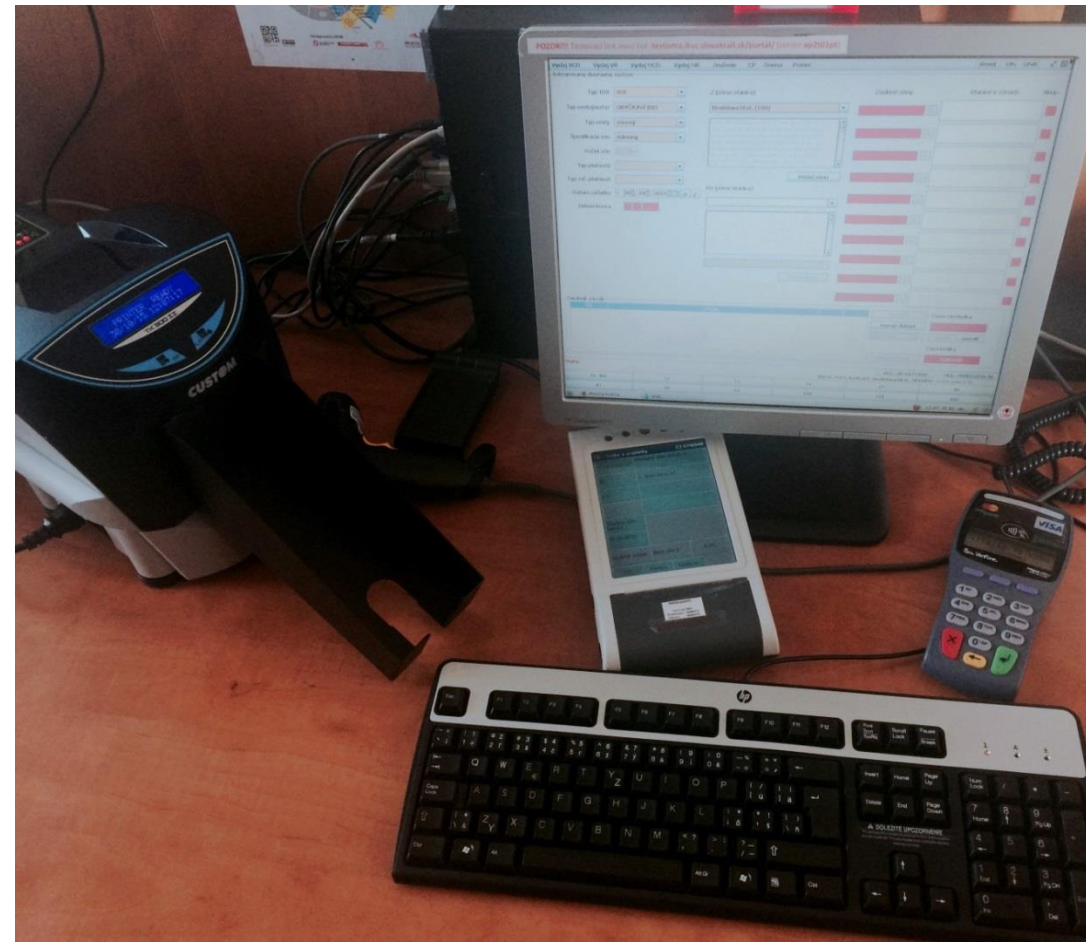
# **Big Data rýchlo a za prijateľné investície**

Július Slíž, Riaditeľ odboru stratégie ZSSK

# História



80-te roky 20. storočia



Od 2007

### Predaj cestovných dokladov v predajnom mieste (miesto+relácia+druh)

Od: 01.07.2006  
Do: 27.07.2006  
Relácia: Košice - Sabinov

Strana: 1

Druh cestovného	Počet dokladov	Počet osôb	Výkon	Tržba	Tržba
		(v osobách)	(osobokm)	vrátane DPH (Sk)	bez DPH (Sk)
1.Obyčajný	12	24	648	888	746
2.Polovičný	0	0	0	0	0
3.Študent	5	10	270	185	155
4.Žiacky týž.lístok	0	0	0	0	0
5.Žiacky mes.lístok	0	0	0	0	0
6.Žiacky týž.lístok -15	0	0	0	0	0
7.Žiacky mes.lístok -15	0	0	0	0	0
8.Štatistický	4	8	216	40	34
9.ĽP	1	2	54	37	31
10.ĽP-S	1	3	54	37	31
11.Junior 40	1	2	54	45	38
12.Klasic 25	1	2	54	56	47
13.Senior 40	2	4	108	90	76
14.Trať týž.lístok	0	0	0	0	0
15.Trať mes.lístok	0	0	0	0	0
16.Trať štvrť.lístok	0	0	0	0	0
17.Skupina do 15 rokov	0	0	0	0	0
18.Skupina nad 15 rokov	0	0	0	0	0
19.TEŽ	0	0	0	0	0
20.Sieťové TEŽ (SČL TEŽ + Podiel TEŽ zo	0	0	0	0	0
21.OŽ	0	0	0	0	0
22.Sieťové OŽ (SČL OŽ + Podiel OŽ zo SČL	0	0	0	0	0
23.TREŽ	0	0	0	0	0

Od: 15.07.2006  
Do: 26.07.2006  
Stanice: Siet' ZSSK  
Cez: Všetky PPS

## PPS CADCA(GR)

Druh cestovného	Počet predan. dokladov	Počet prepr. cest. v osobách	Podiel z celk. pr. cest.	Prepr. výkony v osobokm	Podiel dosiah. výkon. v %	Priem. prepr. vzdial. v km	Tržba za cestovné Inkaso v pokladnici v Sk	Tržba za cestovné podiel ZSSK v Sk	Tržba za cestovné provízia pre ZSSK v Sk
Spolu malý pohr.styk	0	0	0,00	0,00	0	0,00	0,00	0,00	0,00
Spolu sused.preprava s kartou "Z"	0	0	0,00	0,00	0	0,00	0,00	0,00	0,00
skupina +15	22	22	7,8	12 676	9,8	576,18	20 091	235,00	28,00
skupina -15	8	16	5,67	5 702	4,41	356,38	6 391	62,00	10,00
Spolu susedská preprava bez karty "Z"	30	38	13,48	18 378	14,21	483,63	26 482	296,00	38,00
City Star dospelý	8	16	5,67	5 702	4,41	356,38	6 391	62,00	10,00
City Star dieťa	8	16	5,67	5 702	4,41	356,38	6 391	62,00	10,00
City Star spolu	8	16	5,67	5 702	4,41	356,38	6 391	62,00	10,00
obyčajný	22	22	7,8	12 676	9,8	576,18	20 091	235,00	28,00
dieťa	2	2	0,71	642,00	0,5	321	554,00	9,00	1,00
CIV-TCV spolu	24	24	8,51	13 318	10,3	554,92	20 645	243,00	29,00
Suma za PPS	78	110	39,01	48 802	37,74	443,65	66 300	724,00	98,00

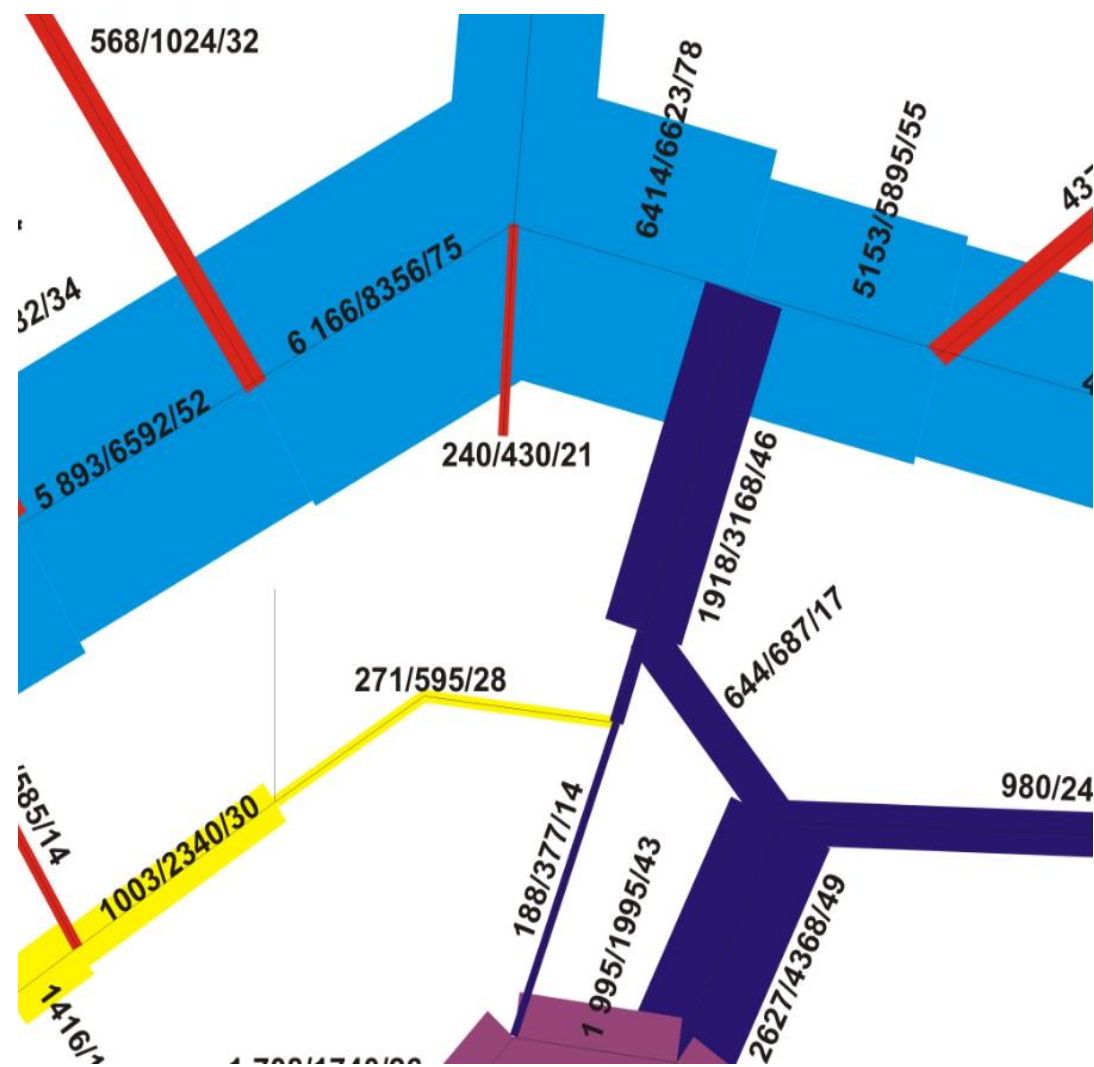
## PPS HORNI LIDEC(GR)

Druh cestovného	Počet predan. dokladov	Počet prepr. cest. v osobách	Podiel z celk. pr. cest.	Prepr. výkony v osobokm	Podiel dosiah. výkon. v %	Priem. prepr. vzdial. v km	Tržba za cestovné Inkaso v pokladnici v Sk	Tržba za cestovné podiel ZSSK v Sk	Tržba za cestovné provízia pre ZSSK v Sk
1. Malý pohr.styk									
obyčajný	1	1	2,56	51	0,25	51	78	1	0
Spolu malý pohr.styk	1	1	2,56	51	0,25	51	78	1	0
2. Sused.preprava s kartou "Z"									
Spolu sused.preprava s kartou "Z"	0	0	0	0	0	0	0	0	0
3. Sused.preprava bez karty "Z"									

## Vytaženost' traťových segmentov

**Predajné miesto:** Sieť ZSSK  
**Od - do:** 01.06.2006 - 30.06.2006  
**Dopravca:** Všetci dopravcovia  
**Trať:** Všetky trate  
**Druh cestovného:** Všetky druhy cestovného

Číslo trate	Traťový segment	Počet vypravených cestujúcich	Počet km	Prepravený výkon v osobokm	Tržba cestovné s DPH	Tržba cestovné bez DPH	Tržba príplatok s DPH	Tržba príplatok bez DPH
1101	Bratislava hlavná st.-Kúty Gr.	4 115	71	80 953	158 729	133 386	12 795	10 752
1141	Kúty-Skalica na Slovensku Gr.	780	28	13 992	15 539	13 058	95	80
1201	Bratislava hlavná st.-Žilina	12 811	207	593 577	579 975	487 374	41 631	34 984
1231	Trenčianska Teplá-Nemšová Gr.	20	12	104	109	92	6	5
1271	Žilina-Čadca Gr.	146	37	1 019	2 641	2 219	108	90
1301	Bratislava hlavná st.-Szob Gr.	16 965	156	500 809	510 772	429 220	25 113	21 104
1302	Bratislava hlavná st.-Šurany	16 803	83	455 511	464 980	390 740	23 033	19 356
1321	Bratislava hlavná st.-Rajka Gr.	847	32	4 807	5 335	4 483	212	178
1322	Bratislava hlavná st.-Kittsee Gr.	847	19	4 807	5 335	4 483	212	178
1331	Leopoldov-Galanta	1 497	30	19 328	21 871	18 379	589	495
1351	Nové Zámky-Komárom Gr.	152	33	3 528	3 538	2 973	120	100
1401	Nové Zámky-Prievidza	5 478	140	109 737	109 964	92 407	5 218	4 385
1411	Leopoldov-Lužianky	2 780	28	38 927	37 653	31 641	2 754	2 314
1501	Nové Zámky-Zvolen osobná st.	5 655	130	191 409	181 470	152 496	14 160	11 899
1511	Nové Zámky-Zlaté Moravce	3 103	51	25 718	25 639	21 546	1 177	989
1521	Štúrovo-Levice	102	52	3 707	4 090	3 437	113	95
1601	Zvolen osobná st.-Košice	955	268	79 461	70 346	59 114	3 532	2 968
1621	Lučenec-Utekáč	97	41	1 352	1 569	1 319	10	8
1601	Košice Čapá Gr.	204	26	707	1 237	1 035	0	0



# Idea

## Nové otázky na ktoré hľadáme odpoveď

Koľko cestujúcich cestuje zo stanice A do stanice B?

Aké tržby stratím na celej sieti ak zruším vedľajšiu trať?

O koľko mi narástli tržby ak vlak zastavím v stanici X?

V ktoré dni mám vyšší dopyt?

Ako sa zmenili prúdy cestujúcich,  
keď som zaviedol priame spojenie?

Kedy ponúknem lepšiu cenu,  
aby som prilákal cestujúcich?

S ktorým dopravcom cestoval nakoniec cestujúci?

## Výzvy a zmeny na trhu

Bezplatná preprava

Nástup ďalších dopravcov

Jeden lístok na celú sieť

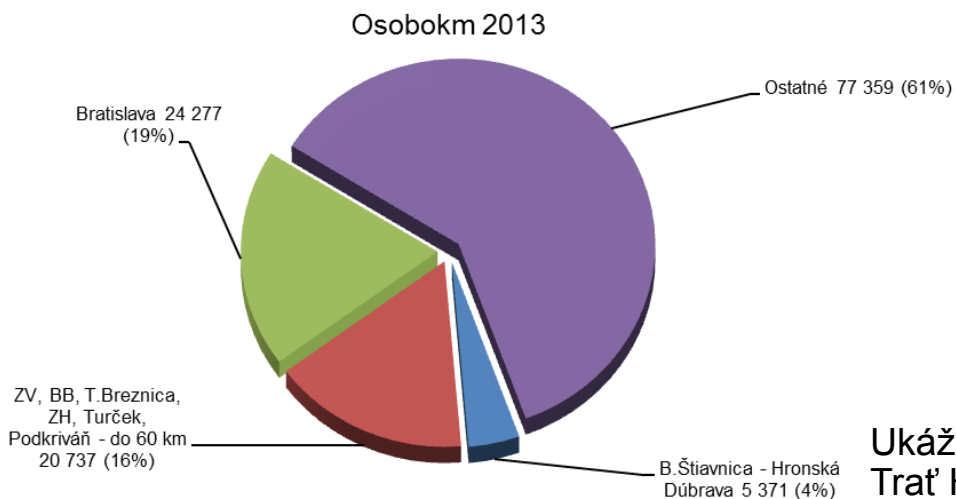
Integrované systémy

Taktový grafikon – vlaky chodia aj v čase  
nižšieho dopytu

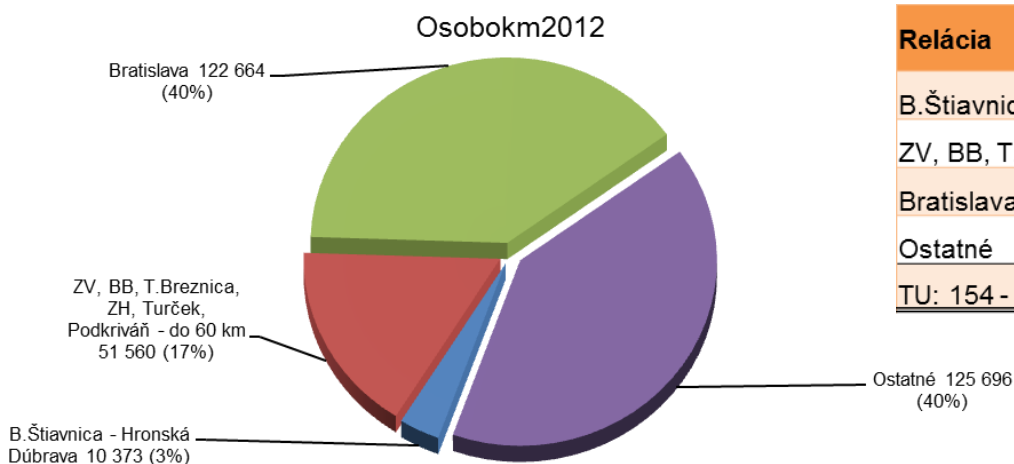
# Riešenie

## Analýza zdrojových dát

- Export dát zo systému KVC bez úprav vo formáte CSV
- Spracovanie veľkého objemu dát v programe FileMaker
- FileMaker použitý na výber dát
- Výstup spracovaný v programe Excel pomocou kontingenčných tabuliek
- Náročnosť spracovania cca 5 hodín



Ukážka analýzy dát za porovnateľné krátke časové obdobie roka 2013 a 2012  
Trať Hronská Dúbrava – Banská Štiavnica



Relácia	Osoby 13	Osoby 12	Tržba 13	Tržba 12	Tržba rozdiel
B.Štiavnica - Hronská Dúbrava	369	695	248,58 €	448,39 €	-45%
ZV, BB, T.Breznica, ZH, Turček, Podkriváň - do 60 km	660	1 469	782,43 €	1 960,67 €	-60%
Bratislava	235	815	806,53 €	3 975,33 €	-80%
Ostatné	339	613	1 867,70 €	3 674,67 €	-49%
TU: 154 - celá sieť	1 603	3 592	3 705,24 €	10 059,06 €	-63%

---

# Dáta

**205.000.000** záznamov  
od roku 2010

**300.000.000**  
zaznamenaných ciest

**100.000** unikátnych  
relácií Z-DO

**Až 139** polí v  
každom zázname

# Preukázanie schopností

## Možnosť vyskúšať riešenie

- Poskytnutie HW formou virtual machine
- Poskytnutie licencie Vertica na obdobie testu
- Trial licencie Tableau pre spracovanie výstupov
- Zostavenie štruktúry dát a ich import z produkčného predajného servera v priebehu týždňa

## Výstupy

- Dáta pre analyzovanie vstupu konkurencie na trať Praha – Košice pripravené v priebehu dopoludnia
- Rýchla analýza vývoja cestujúcich po zavedení bezplatnej prepravy – riadenie kapacít na základe dopytu
- Analýza vplyvu bezplatnej prepravy na obsadenie IC vlakov

## Záver:

- niekoľkonásobné zvýšenie produktivity analytického tímu,
- na základe detailnejšieho rozboru dát exaktný popis príčin zmeny počtu cestujúcich,
- nové druhy analýz pre riadenie kontrolných zamestnancov,
- analýzy pre plánovanie predajného personálu,
- skrátenie doby na prípravu analýz a návrh opatrení z týždňov na dni až hodiny



---

# Zdroje dát

Predajný systém KVC

Prenosné osobné pokladne

Odpočty zahraničných železníc

Prevádzkový systém EVOD

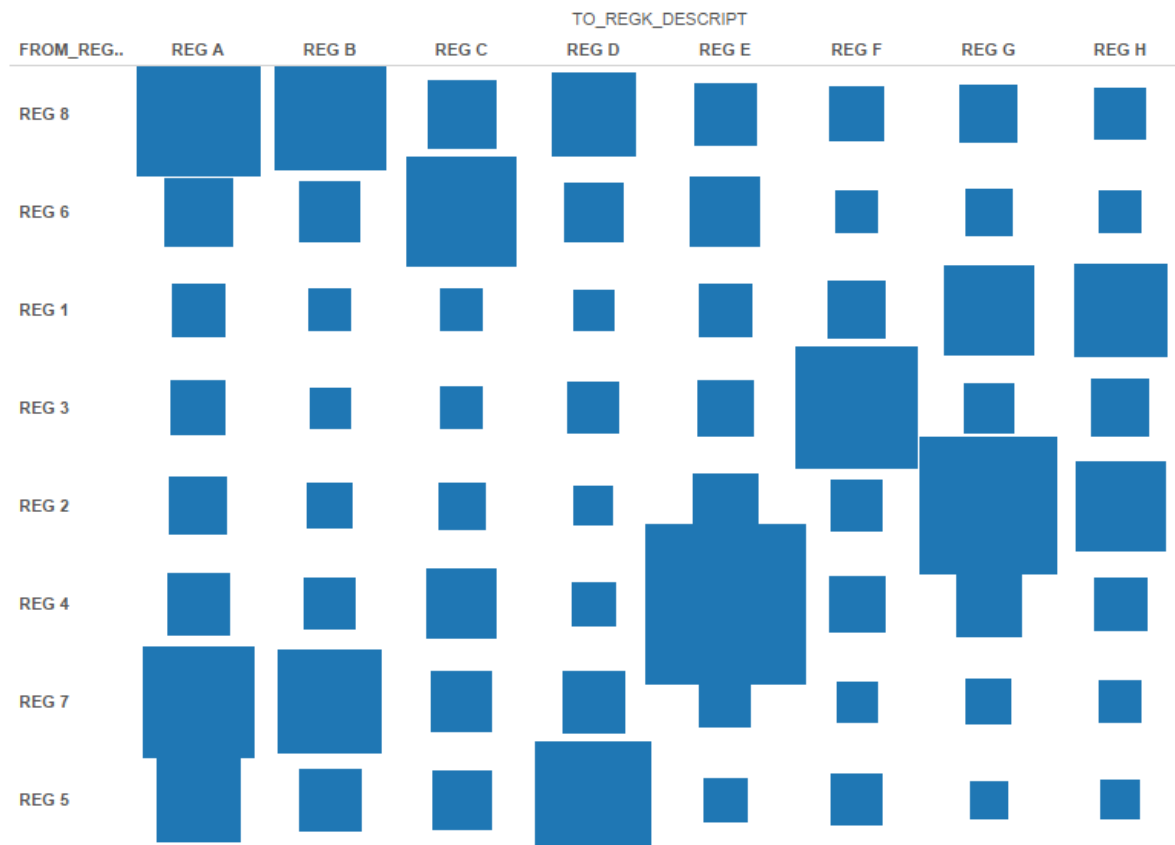
Prevádzkový informačný systém ŽSR

Cestovný poriadok

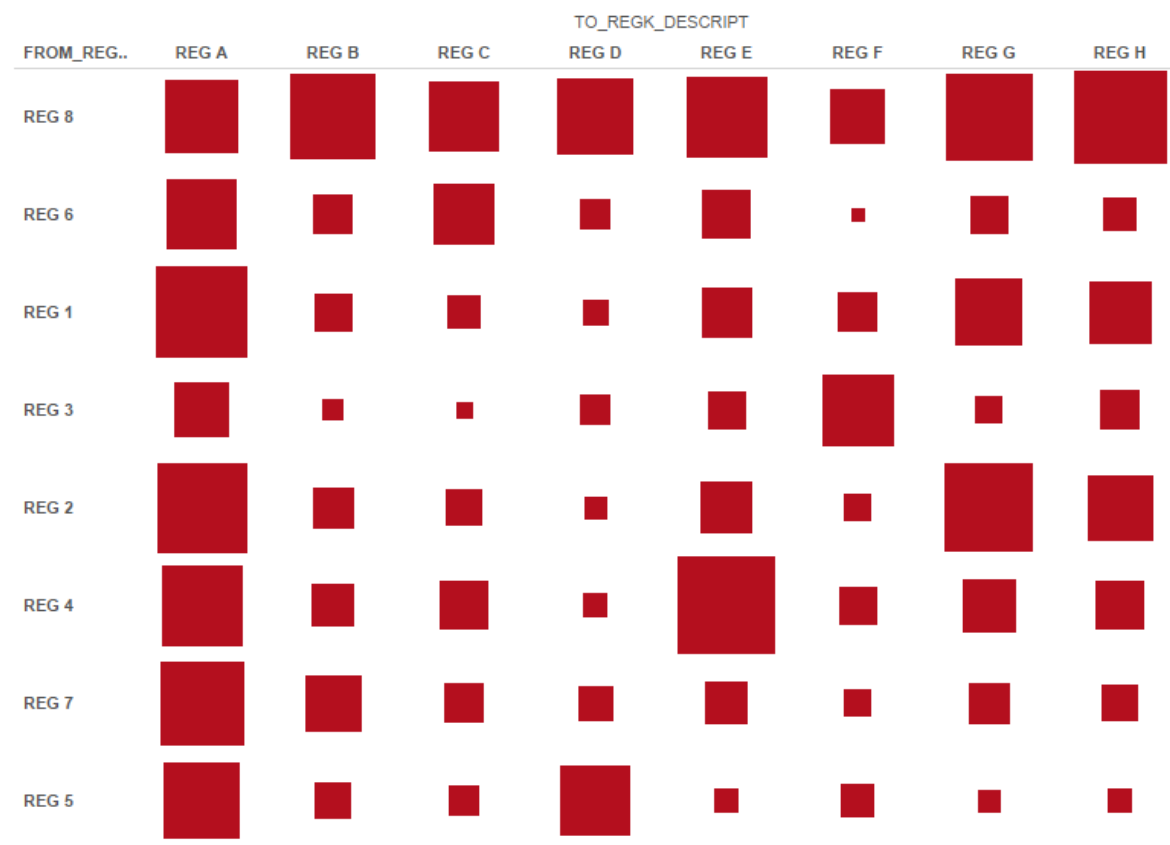
---

# Výsledok - samoobslužnosť a kontrola

# Ukážky výstupov



cestujúci



tržby

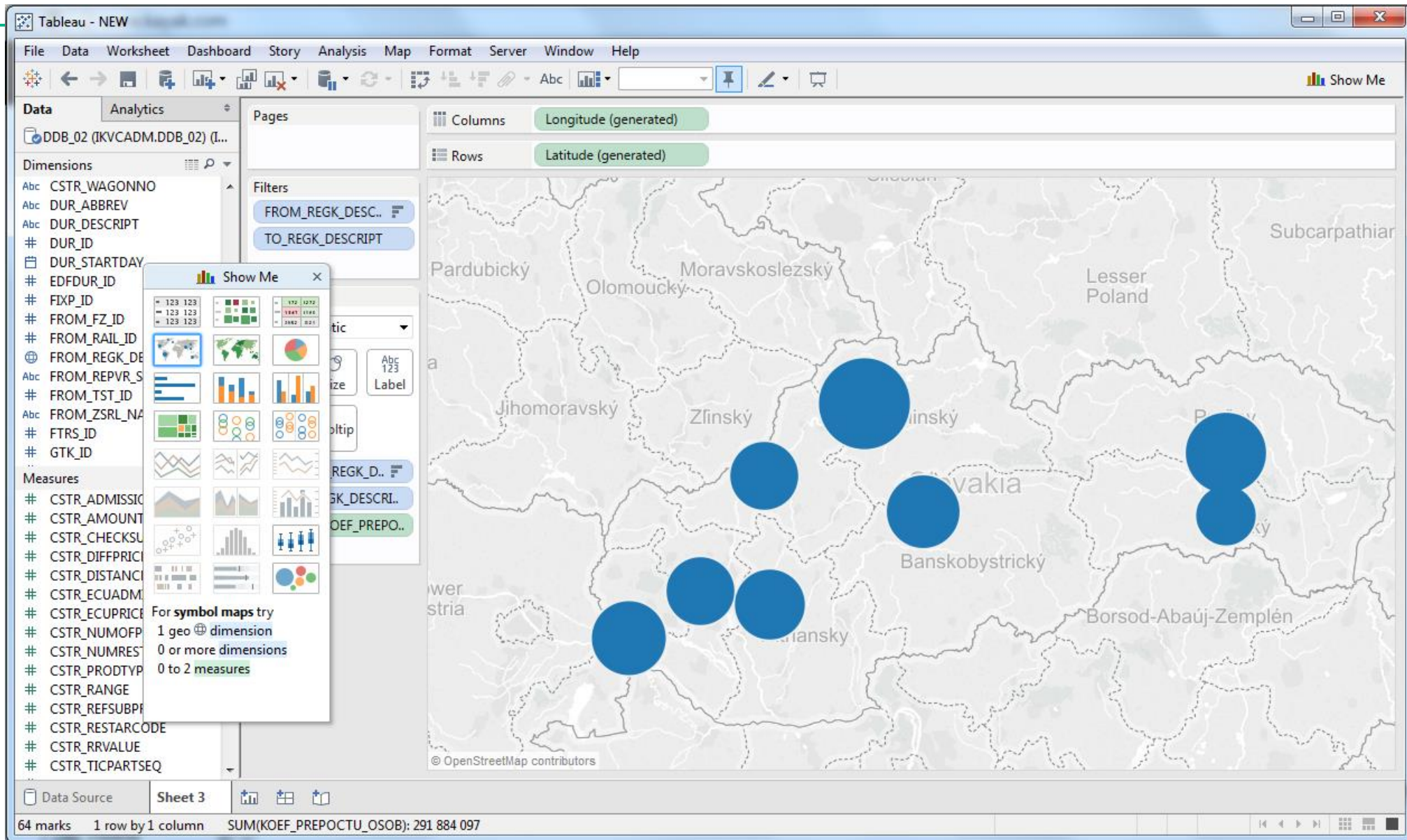


Tableau - NEW

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Analytics

Data Source: DDB\_02 (IKVCADM.DDB\_02) (I...)

Dimensions:

- CSTR\_REIMBYESNO
- CSTR\_RESNO
- CSTR\_RESSMOKEYESNO
- CSTR\_RESSTATUS
- CSTR\_SELLDATETIM
- CSTR\_STARTDATE
- CSTR\_STRITEMNO
- CSTR\_TIMESTAMP
- CSTR\_TRAINNO
- CSTR\_VARCHAR
- CSTR\_WAGONNO
- DUR\_ABBREV
- DUR\_DESCRIPT
- DUR\_ID
- DUR\_STARTDAY

Measures:

- CSTR\_ADMISSIONAMOUNT
- CSTR\_AMOUNT
- CSTR\_CHECKSUM
- CSTR\_DIFFPRICE
- CSTR\_DISTANCE
- CSTR\_ECUADMISIONAM...
- CSTR\_ECUPRICE
- CSTR\_NUMOFPERSON
- CSTR\_NUMREST
- CSTR\_PRODTYPEID
- CSTR\_RANGE
- CSTR\_REFSUBPRODSEQNO
- CSTR\_RESTARCODE
- CSTR\_RRVALUE
- CSTR\_TICPARTSEQ

Columns: TO\_REGK\_DESCRIPT

Rows: FROM\_REGK\_DESCRIPT..

Filters: CSTR\_STARTDATE

Marks: Square

Color, Size, Label (Abc 123)

Detail, Tooltip

SUM(KOEF\_PREPO...)

SUM(KOEF\_PREPO...)

SUM(KOEF\_PREPOCTU\_O...)

1 671 14 955 583

FROM_REG..	REG A	REG B	REG C	REG D	REG E	REG F	REG G	REG H	REG I
REG 8	35 125	2 906 043	2 366 057	347 785	249 998	22 528	43 178	43 728	266 079
REG 7	16 138	2 389 214	9 602 566	187 203	694 487	23 885	79 414	79 248	421 076
REG 6	19 504	350 175	180 145	6 968 534	289 876	180 672	31 677	27 434	268 495
REG 1	28 920	253 955	405 383	267 433	567 118	1 795 189	786 782	5 061 833	7 039 683
REG 5	4 938	247 601	696 626	277 929	14 955 583	41 099	1 009 147	182 482	581 261
REG 3	81 144	42 472	82 791	29 964	1 030 158	462 167	5 405 633	515 915	796 714
REG 2	2 230	40 389	77 666	25 324	183 631	608 714	520 369	4 497 163	5 142 679
REG 9	1 967 525	37 550	15 297	19 821	3 462	5 707	99 832	2 794	27 583
REG 4	1 671						838 575		1 828 590

Filter [CSTR\_STARTDATE]

Relative dates, Range of dates, Starting date, Ending date, Special

Range of dates

31. 10. 2010 1. 12. 2013

2. 10. 2006 28. 11. 2015

Show: Only relevant values Include Null Values

Reset OK Cancel Apply

81 marks 9 rows by 9 columns SUM(KOEF\_PREPOCTU\_OSOB): 93 130 262

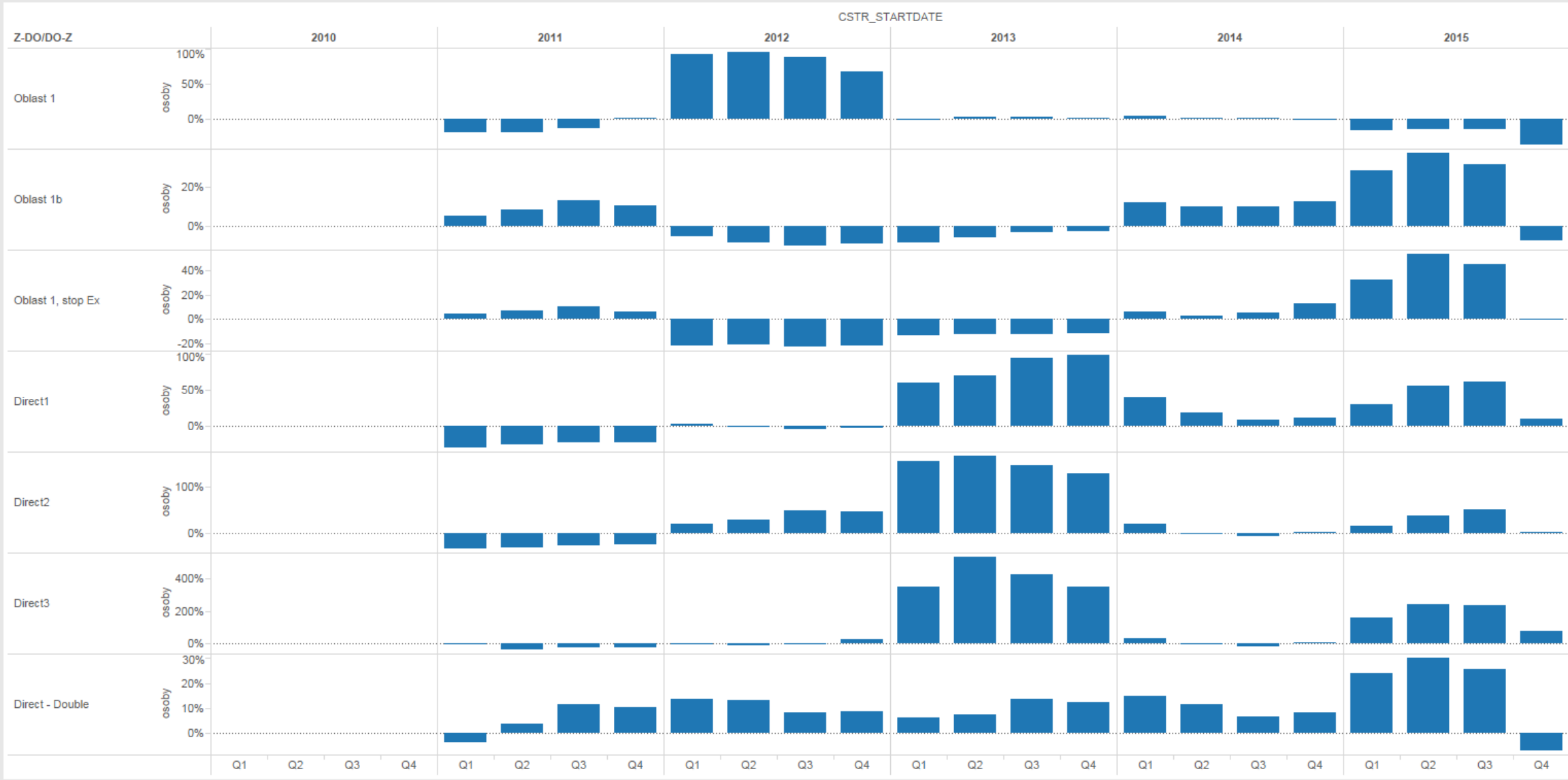
Show Me

For symbol maps try  
1 geo dimension  
0 or more dimensions  
0 to 2 measures

Filters  
Z-DO/DO-Z

Marks  
Bar  
Color Size Label  
Detail Tooltip

Show Me  
filled maps try  
geo dimension  
or more dimensions  
or 1 measure



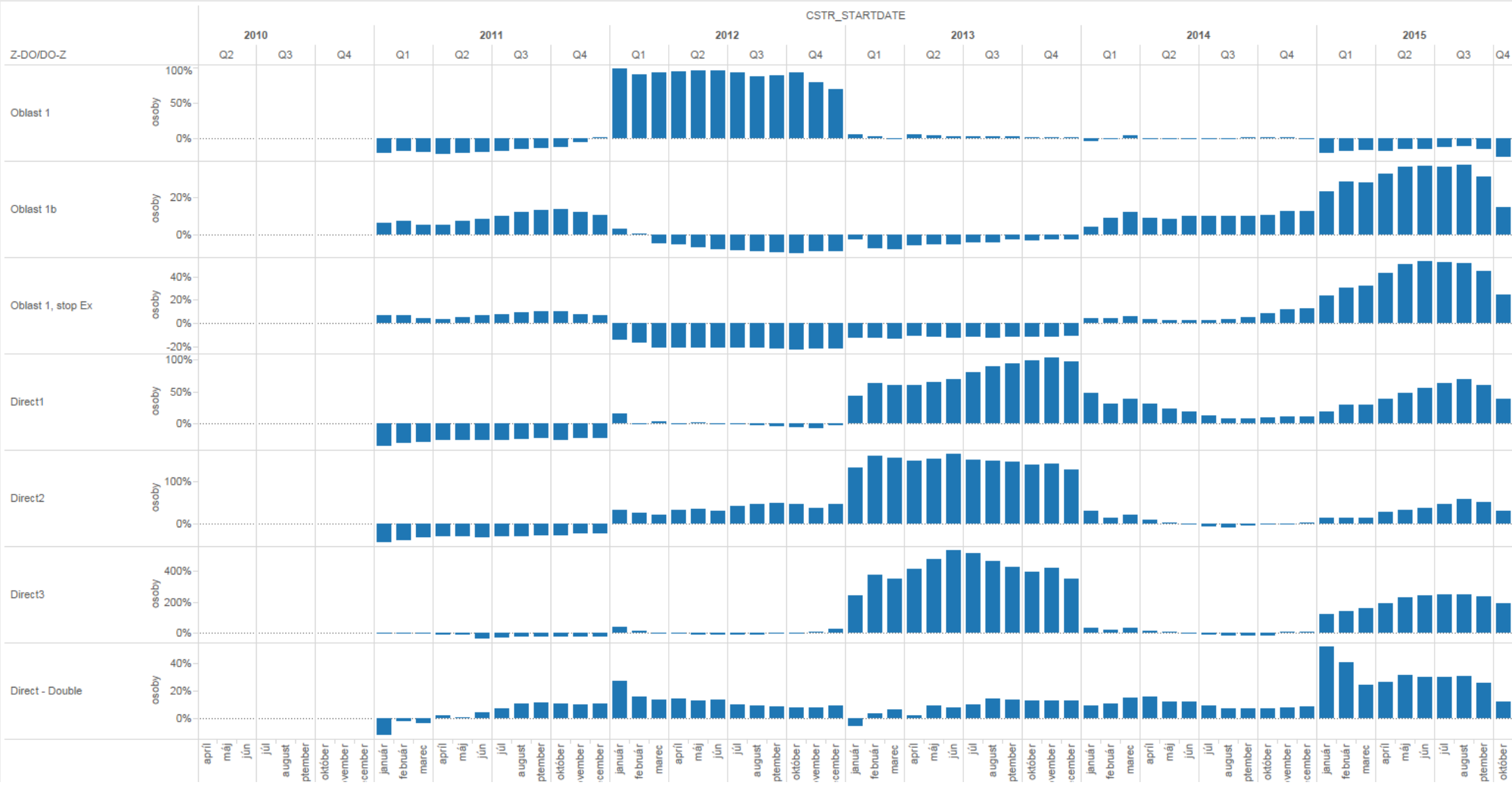
**Filters**  
Z-DO/DO-Z

**Marks**  
Bar

Color Size Label  
Detail Tooltip

**Show Me**

lines (discrete) try  
date  
or more dimensions  
or more measures





**Hewlett Packard  
Enterprise**



**Ďakujem  
za pozornosť'**