

A BDF website elemzése SPSS CLEMENTINE WEB MINING segítségével

Zsiros Péter

Kezdőlap - Windows Internet Explorer

http://www.bdf.hu/Default.aspx

Fájl Szerkesztés Nézet Kedvencek Eszközök Súgó

Kezdőlap

BERZSENYI DÁNIEL FŐISKOLA

Berzsenyi Dániel Főiskola,
Szombathely Károlyi Gáspár tér 4.
www.bdf.hu

Kezdőlap Bemutakozunk Elérhetőségek Nemzetközi kapcsolatok Intranet

Kezdőlap

Hírek, sajtószoza	<h3>Kiemelten fontos híreink</h3> <h4>Vezetői és oktatói pályázatok</h4> <p>Beadási határidő: 2007. június 15. 12.00 óra</p> <h4>Európa-tanulmányok szakirányú képzési program</h4> <p>A Berzsenyi Dániel Főiskola Európa-tanulmányok Tanszéke <i>új formában, megújult tartalommal</i> indítja 2007 szeptemberében Európa-tanulmányok szakirányú továbbképzési szakját diplomával rendelkezők részére, igény esetén angol nyelven is. A képzés 3 féléves, melyben neves külföldi és hazai előadók, szakértők is részt vesznek A tandíj 170000 Ft/félév. Jelentkezési határidő 2007. augusztus 31. További információ: 94/504-581 www.ises.hu</p> <h4>Szakvizsgás képzések a BDF-en!</h4> <p>Bővebb információk: http://sts.bdf.hu/pszk/build/szakvizsgák.aspx</p> <h4>Oktatási és Kulturális Minisztérium - Hírek</h4> <ul style="list-style-type: none">(2007. 06. 13.) Vita a Botticelli-freskóról - Hiller: biztosítani kell a restaurálás feltételeit(2007. 06. 13.) Hiller István beiktatta az operaház új vezetőit(2007. 06. 13.) Magyar-olasz filmkoprodukciónak megállapodás(2007. 06. 13.) Rangos nemzetközi elismerést kapott a Sulinet Digitális Tudásbázis(2007. 06. 13.) Az OKM közleménye a tankönyvekről(2007. 06. 13.) Az OKM közleménye a tarnabodi pedagógus-elbocsátások miatt(2007. 06. 13.) Négy festőnő - kiállítás a Római Magyar Akadémián(2007. 06. 13.) Bogyay Katalin szerint a kulturális párbeszéd a leghatékonyabb eszköz a konfliktusok kezelésére
Rendezvénynaptár	
Szervezeti felépítés	
Főiskolai telefonkönyv	
Pályázatok	
Képzések	
NEPTUN	
Hallgatóinknak	
Felvételizőknek	
Szolgáltatások	

Partnerünk



Berzsenyi Rádió FM 98.8 MHz
A fülek szállodája



Szláv Történelmi és
Filológiai Társaság

Megbízható helyek 100%

Start BDF_site zsp Stream1 - Cle... 2 Internet ... BDF website ... Microsoft Exc... Microsoft Po... HU 10:18

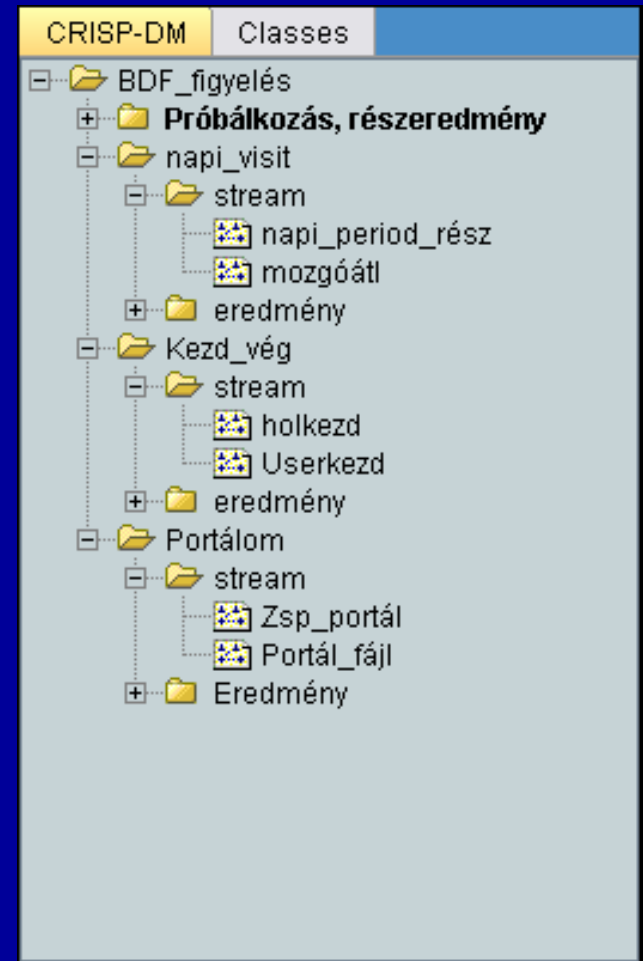
Az elemzés kiindulópontja, célok



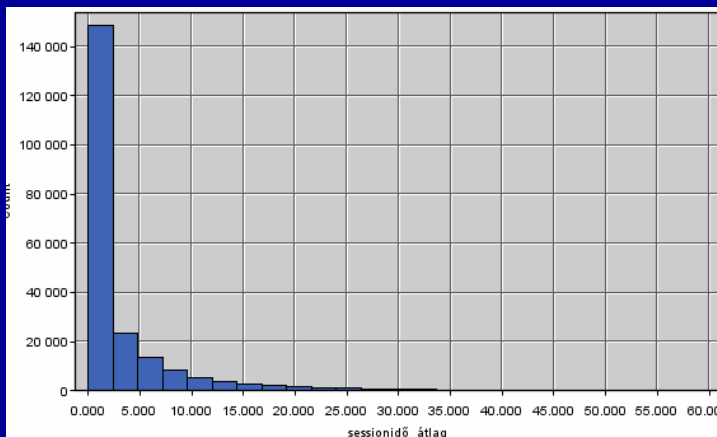
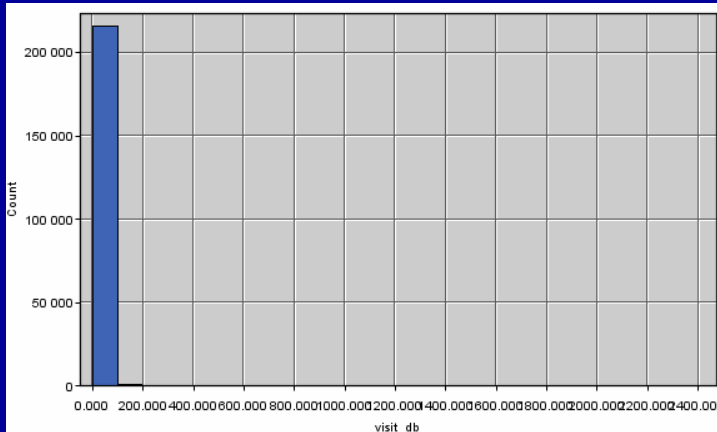
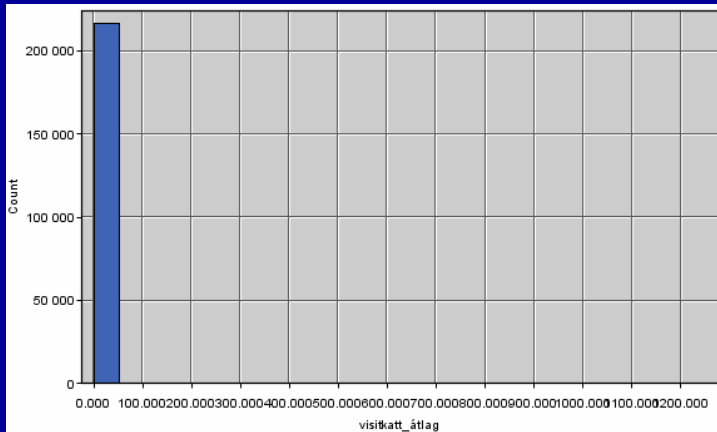
- Google analízis: heti hullámváz (Grujber Zoltán) ↗
- Log fájlok vizsgálata: külső és belső IP címek – web mining után ezek jelölése egy külső mezőben
- Regisztrált felhasználók nevei – ezek jelölése
- Eseménydefinícióként minden főlapon szereplő hely, részletesebben saját intézetem eseményei+saját honlap
- Információgyűjtés saját honlapom látogatóiról
- *Igények eltérnek a kereskedelmi célú demostreamektől -> streamépítés*
- *Clementine hatékonyságának bemutatása*

Bemenő adatok, projekt

- 2006.08.28 és 2007.05.21. közötti logfájlok
- Az eseménydefinícióknak megfelelő rekordok száma: 2,9 millió
- Visitek száma: 712 ezer
- Userek száma: 217 ezer
- Téli szünetben belső leállítás van, de kívülről elérhető a website
- *BDF_figyel projekt, a lényege 6 db stream*



Általános uservizsgálat



Összkatt	
Statistics	
Mean	13.733
Min	1
Max	31898
Range	31897
Standard Deviation	135.496

visitkatt_átlag	
Statistics	
Mean	4.327
Min	1.000
Max	2812.000
Range	2811.000
Standard Deviation	9.937

Összidő	
Statistics	
Mean	18.064
Min	0.000
Max	88962.133
Range	88962.133
Standard Deviation	339.128

sessionid_átlag	
Statistics	
Mean	3.589
Min	0.000
Max	713.517
Range	713.517
Standard Deviation	10.938

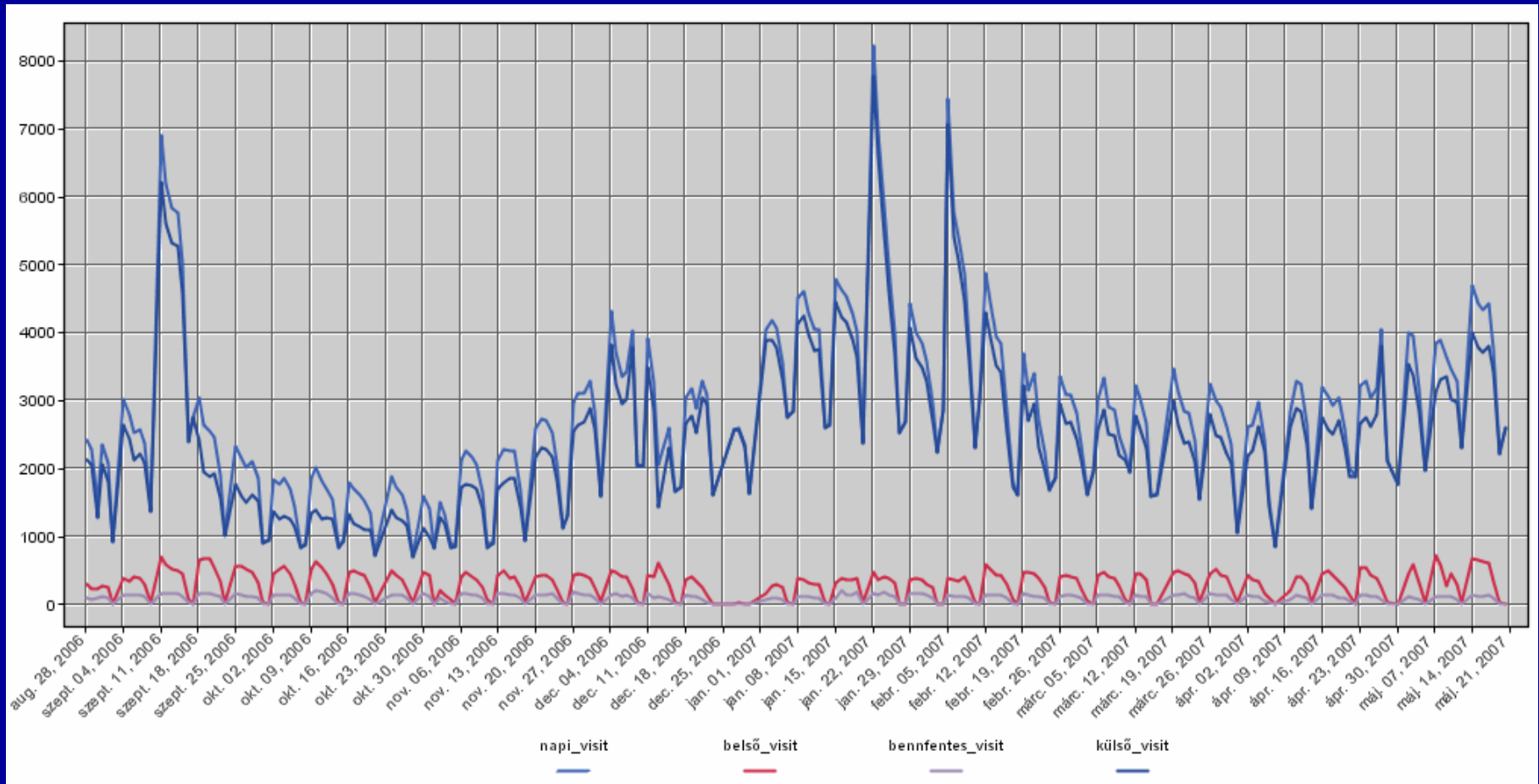
visit_db	
Statistics	
Mean	3.273
Min	1
Max	2422
Range	2421
Standard Deviation	18.556

• **Az idők percben értendők!**

• **Itt csak azokra a userekre, akiknek egy óránál rövidebb az átlagos session-idejük.**

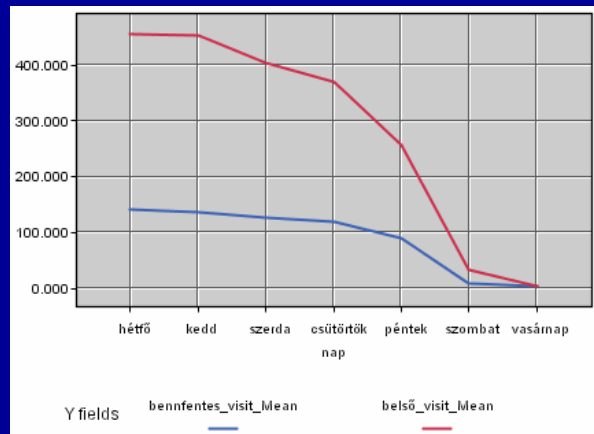
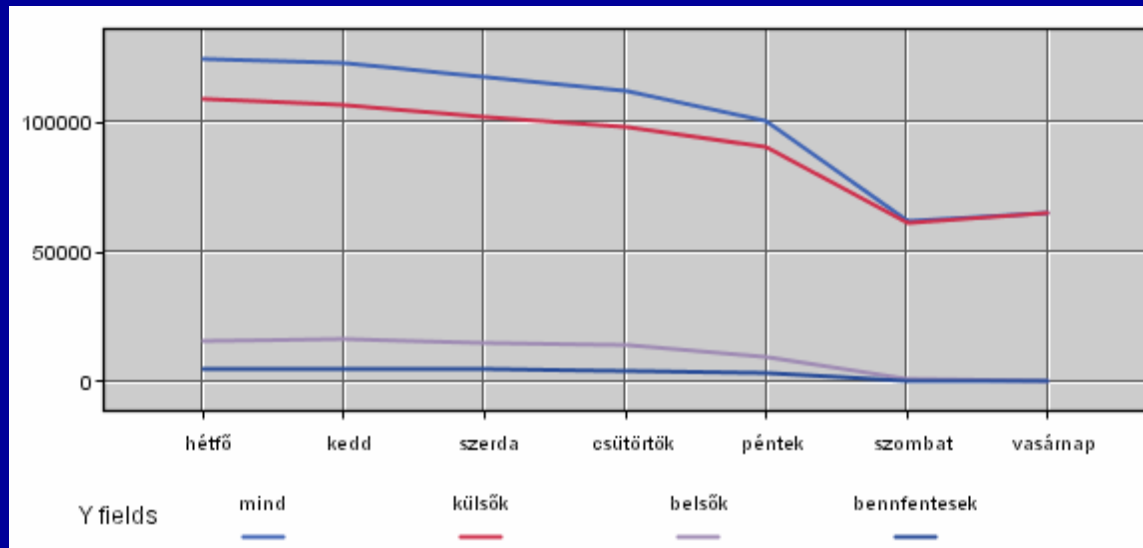
1. Heti periodicitással kapcsolatos vizsgálatok

1.1 A visitek heti periódusú változásában mely típusú látogatók milyen mértékben vesznek részt?



1. Heti periodicitással kapcsolatos vizsgálatok

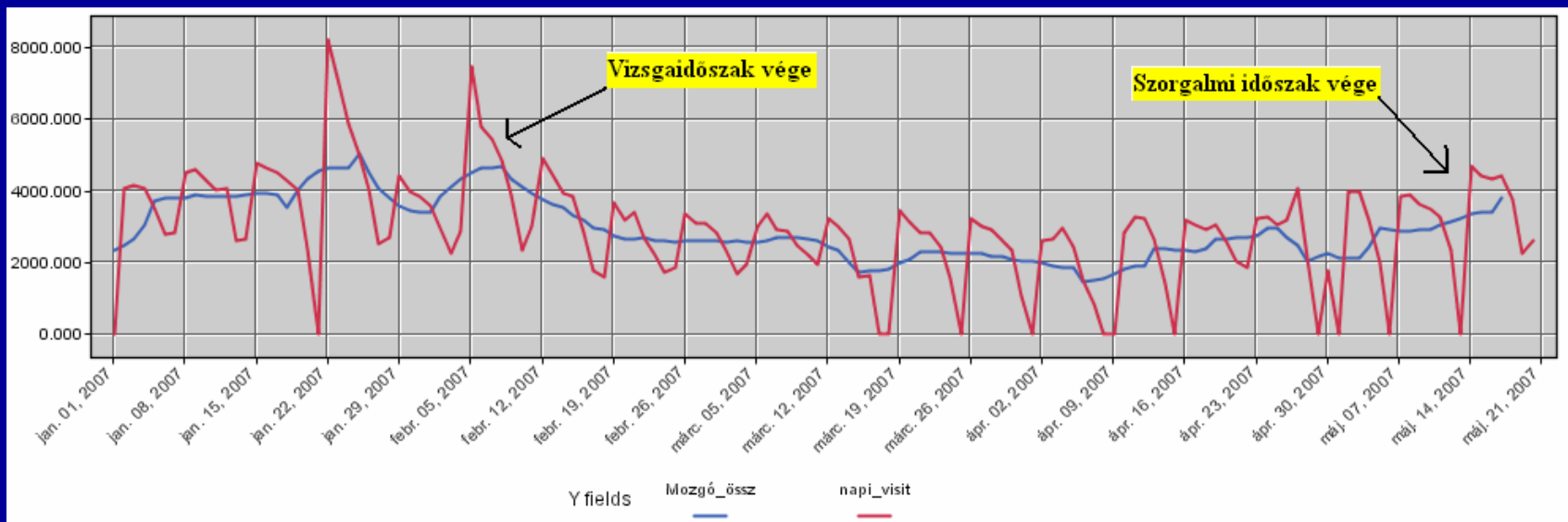
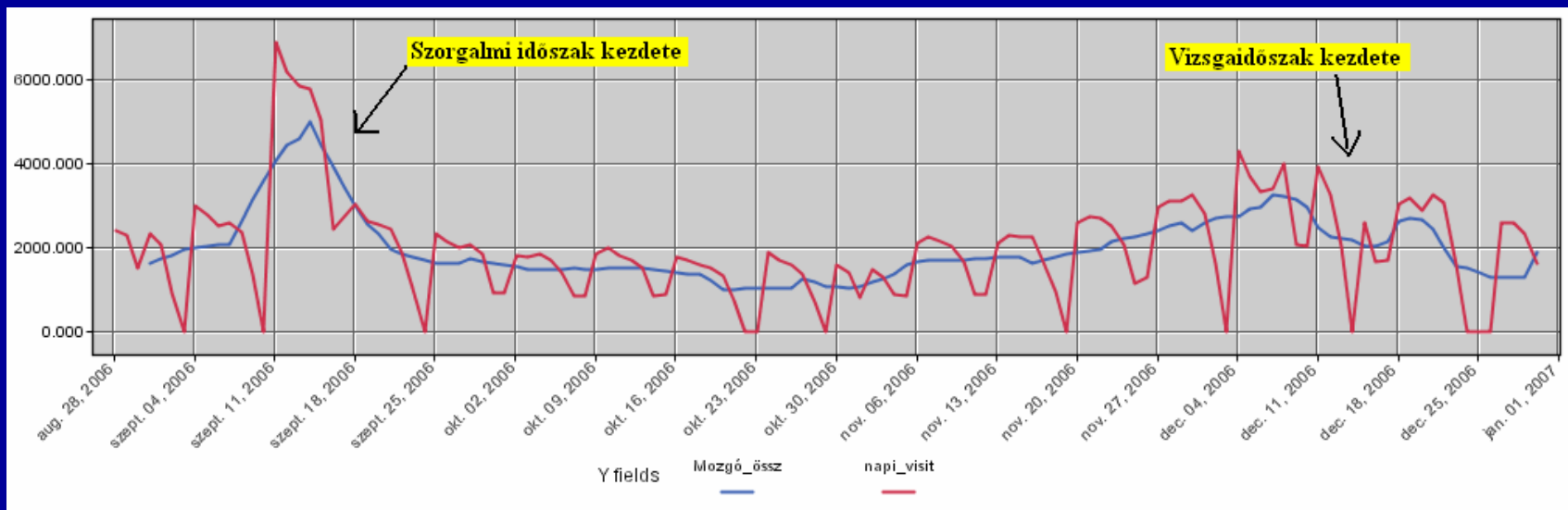
1.2 Szezonális adatok meghatározása



nap	mind	külsők	belsek	bennfentes...
hétfő	124819	109289	15533	4816
kedd	122701	106384	16323	4873
szerda	117512	102207	15307	4809
csütörtök	112465	98022	14445	4468
péntek	100396	90612	9784	3370
szombat	62212	60934	1279	315
vasárnap	65365	65297	70	132

1. Heti periodicitással kapcsolatos vizsgálatok

1.3 Simítás mozgóátlagolással, trendfigyelés



2. Visitek kezdeteinek, végeinek vizsgálata

- 2.1 Mi a visitek legelső eseménye?

	Event Name	visittkezdb	ittkezdbelső	ittkezdkülső	ittkezdbennf	ittkezdoösszhossz	ittkezdatlaghossz
1	HomePage	388880	57022	331858	17557	2237000.517	5.752
2	Neptun	112334	1802	110532	41	339697.967	3.024
3	HOK	38810	1722	37088	311	222027.033	5.721
4	Felveteli	21867	221	21646	47	78743.183	3.601
5	Informatika	17506	5709	11797	353	151344.167	8.645
6	EgySzervezetiEgyseg	16106	281	15825	64	51002.033	3.167
7	PSZK	9258	74	9184	202	57119.967	6.170
8	Hirek	8488	934	7554	243	38766.433	4.567
9	Hallgatoknak	7573	1251	6322	80	44754.300	5.910
10	TTK	6987	148	6839	56	55121.217	7.889
11	MSZI	6796	136	6660	59	44103.633	6.490
12	Események	6685	256	6429	71	32124.417	4.805
13	FKK	5184	47	5137	170	18967.483	3.659
14	Karrier	4676	136	4540	67	36931.033	7.898
15	HallgatoknakPalyazatok	3825	59	3766	29	26560.300	6.944
16	SzKI	3776	128	3648	74	18406.983	4.875
17	Intranet	3114	2848	266	3107	17276.500	5.548
18	EN	2998	21	2977	7	24466.933	8.161
19	IariTemekEsFormatervezo	2881	15	2866	9	24764.400	8.596
20	TMK	2838	25	2813	14	23442.817	8.260
21	BTK	2394	164	2230	57	25952.333	10.841
22	SzervezetiEgysegLista	2360	576	1784	79	14378.100	6.092
23	MuszakiMenedzser	1789	22	1767	7	18004.150	10.064
24	Rekreacio	1180	1	1179	0	7952.400	6.739
25	Erasmus	1153	40	1113	17	13431.167	11.649
26	Andragogia	1123	3	1120	0	18928.783	16.856
27	HallgatoknakDokumentumok	1096	63	1033	54	11043.750	10.076
28	Tanito	1074	1	1073	0	9366.083	8.721
29	Elerhetosegek	1049	49	1000	2	5328.983	5.080
30	KepzesLista	1020	126	894	25	6493.317	6.366
31	TermektervezoMuszakiMenedz...	1001	5	996	4	9562.783	9.553
32	Romanisztika	985	2	983	0	7562.650	7.678
33	MuszakiInformatikaiMemokAs...	946	2	944	0	6578.200	6.954
34	Kornyezettan	944	1	943	0	9221.133	9.768
35	Foldrajz	939	2	937	0	8213.600	8.747
36	InformatikusKonytaros	930	3	927	1	9709.667	10.441
37	GazdalkodasiMenedzsmnt	930	5	925	1	11884.283	12.779
38	IntezmenyiKommunikator	917	0	917	0	12759.617	13.915
39	Matematika	914	0	914	0	7005.833	7.665
40	IdegenforgalmiSzakmenedzser	914	18	896	9	8261.700	9.039
41	GazdalkodasiMenedzserAszis	903	5	898	1	8827.883	9.776

2. Visitek kezdeteinek, végeinek vizsgálata

- 2.2 Van-e releváns különbség a külső, a belső és a bennfentes visitek kezdetei között?

	Event Name	lttkezd_belső	lttkezd_külső	lttkezd_bennf
1	HomePage	57022	331858	17557
2	Neptun	1802	110532	41
3	HOK	1722	37088	311
4	Felveteli	221	21646	47
5	EgySzervezetiEgyseg	281	15825	64
6	Informatika	5709	11797	353
7	PSZK	74	9184	202
8	Hirek	934	7554	243
9	TTK	148	6839	56
10	MSZI	136	6660	59
11	Esemenyek	256	6429	71
12	Hallgatoknak	1251	6322	80
13	FKK	47	5137	170
14	Karrier	136	4540	67
15	HallgatoknakPalyaz...	59	3766	29
16	Szki	128	3648	74
17	EN	21	2977	7
18	IariTemekEsFormat...	15	2866	9
19	TMK	25	2813	14
20	BTK	164	2230	57
21	SzervezetiEgysegLi...	576	1784	79
22	MuszakiMenedzser	22	1767	7
23	Rekreacio	1	1179	0
24	Andragogia	3	1120	0
25	Erasmus	40	1113	17
26	Tanito	1	1073	0
27	HallgatoknakDoku...	63	1033	54
28	Elerhetosegek	49	1000	2
29	TermektervezoMusz...	5	996	4
30	Romanisztika	2	983	0
31	MuszakiInformatikai...	2	944	0
32	Kornyezettan	1	943	0
33	Foldrajz	2	937	0
34	InformatikusKonyta...	3	927	1
35	GazdalkodasiMene...	5	925	1
36	IntezmenyiKommun...	0	917	0
37	Matematika	0	914	0
38	GazdalkodasiMene...	5	898	1
39	IdegenforalmiSzak	18	896	9

	Event Name	lttkezd_belső	lttkezd_külső	lttkezd_bennf
	HomePage	57022	331858	17557
	Neptun	1802	110532	41
	HOK	1722	37088	311
	Felveteli	221	21646	47
	EgySzervezetiEgyseg	281	15825	64
	Informatika	5709	11797	353
	PSZK	74	9184	202
	Hirek	934	7554	243
	TTK	148	6839	56
	MSZI	136	6660	59
	Esemenyek	256	6429	71
	Hallgatoknak	1251	6322	80
	FKK	47	5137	170
	Karrier	136	4540	67
	HallgatoknakPalyaz...	59	3766	29
	Szki	128	3648	74
	EN	21	2977	7
	IariTemekEsFormat...	15	2866	9
	TMK	25	2813	14
	BTK	164	2230	57
	SzervezetiEgysegLi...	576	1784	79
	MuszakiMenedzser	22	1767	7
	Rekreacio	1	1179	0
	Andragogia	3	1120	0
	Erasmus	40	1113	17
	Tanito	1	1073	0
	HallgatoknakDoku...	63	1033	54
	Elerhetosegek	49	1000	2
	TermektervezoMusz...	5	996	4
	Romanisztika	2	983	0
	MuszakiInformatikai...	2	944	0
	Kornyezettan	1	943	0
	Foldrajz	2	937	0
	InformatikusKonyta...	3	927	1
	GazdalkodasiMene...	5	925	1
	IntezmenyiKommun...	0	917	0
	Matematika	0	914	0
	GazdalkodasiMene...	5	898	1
	IdegenforalmiSzak	18	896	9

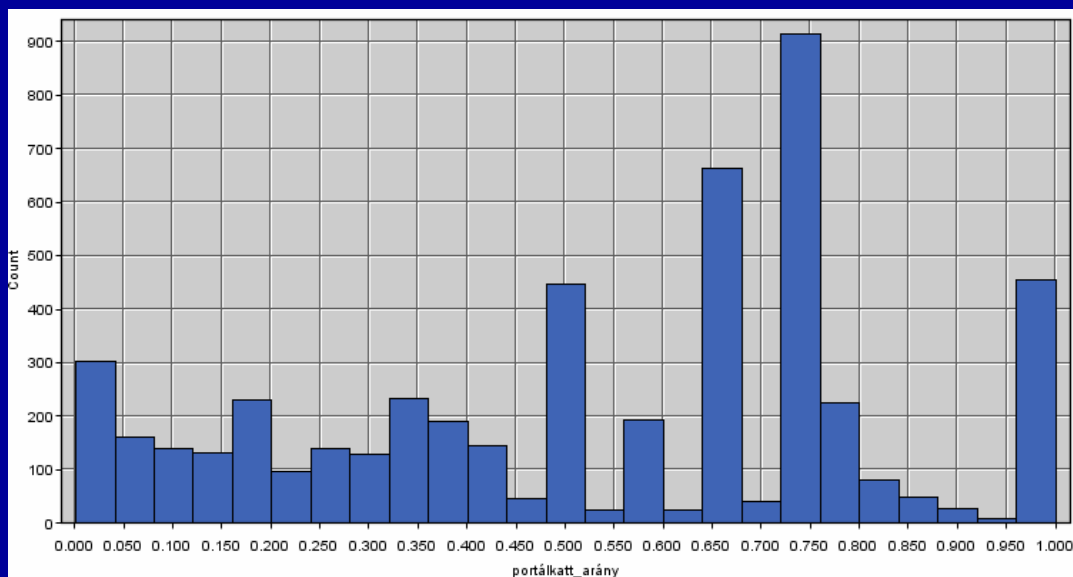
	Event Name	lttkezd_belső	lttkezd_külső	lttkezd_bennf
	HomePage	57022	331858	17557
	Intranet	2848	266	3107
	GT	246	162	379
	Informatika	5709	11797	353
	HOK	1722	37088	311
	Hirek	934	7554	243
	PSZK	74	9184	202
	FKK	47	5137	170
	SSS	165	562	82
	Hallgatoknak	1251	6322	80
	SzervezetiE...	576	1784	79
	Szki	128	3648	74
	Esemenyek	256	6429	71
	Karrier	136	4540	67
	EgySzerveze...	281	15825	64
	MSZI	136	6660	59
	BTK	164	2230	57
	TTK	148	6839	56
	Hallgatkna...	63	1033	54
	Felveteli	221	21646	47
	Neptun	1802	110532	41
	Zsp	80	448	37
	Hallgatkna...	59	3766	29
	KepzesLista	126	894	25
	Erasmus	40	1113	17
	TMK	25	2813	14
	IariTemekE...	15	2866	9
	Idegenforga...	18	896	9
	MuszakiMen...	22	1767	7
	EN	21	2977	7
	Magunkrol	29	802	4
	Termekterve...	5	996	4
	Haloterv	12	660	3
	Elerhetoseg...	49	1000	2
	Szlavisztika	3	888	2

3. Saját portálom vizsgálata

- 3.1 Azokat a visiteket figyeltem, melyek tartalmaztak a portálon való kattintást. Az ilyen visitekben kiszámítottam a portálon való kattintások számát és az összes kattintások számát, továbbá ezek arányát, valamint e visitek sessionhosszát.

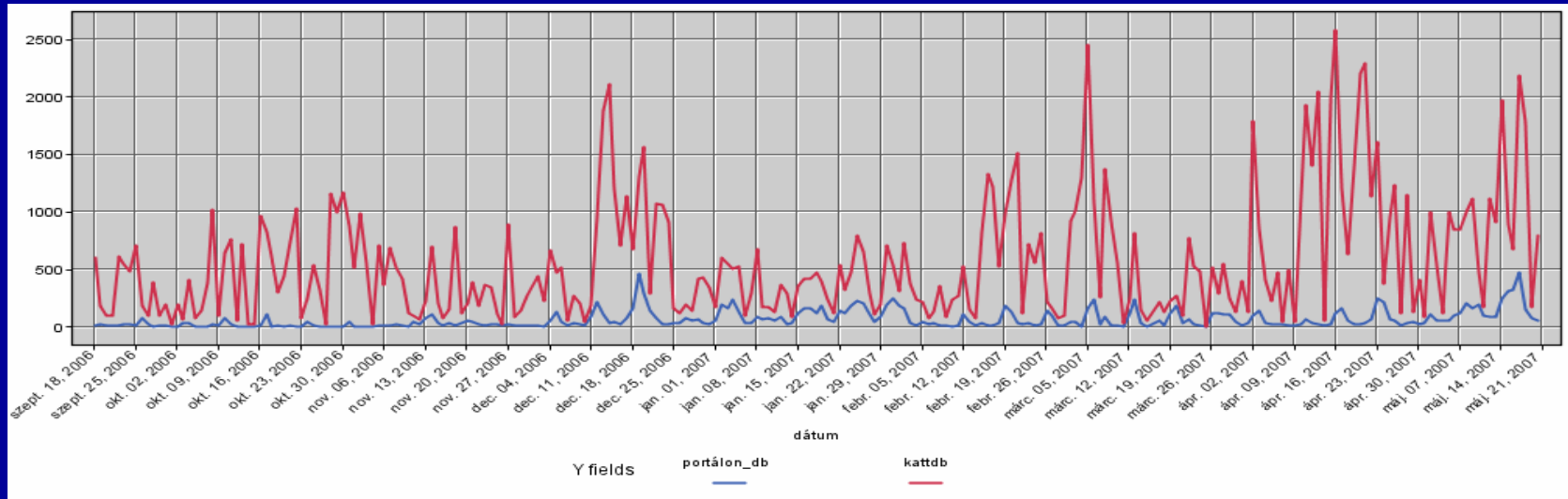
Katt%	
Statistics	
Mean	0.534
Min	0.001
Max	1.000
Variance	0.083
Standard Deviation	0.288
Pearson Correlations	
portálon_katt	0.086 Strong
portálon_katt	
Statistics	
Mean	3.192
Min	1.000
Max	124.000
Variance	12.624
Standard Deviation	3.553
Pearson Correlations	
összkatt_db	0.159 Strong
összkatt_db	
Statistics	
Mean	27.478
Min	1
Max	1741
Variance	10739.535
Standard Deviation	103.632
Pearson Correlations	
portálon_katt	0.159 Strong
session_hossza	
Statistics	
Mean	39.781
Min	0.000
Max	719.983
Variance	19584.903
Standard Deviation	139.946
Pearson Correlations	
portálon_katt	0.143 Strong

Statisztikai számításokat végeztem a Clementine-nal, hogy mennyi e statisztikai változók átlaga, szórása, minimuma és maximuma, illetve korrelációt számítottam e változók és a visitek portálon való kattintásainak száma között.



3. Saját portáлом vizsgálata

3.2 A portálon való kattintást tartalmazó visitek esetében ábrázoltam a portálon való napi kattintások és az összes napi kattintások számát a dátum függvényében.

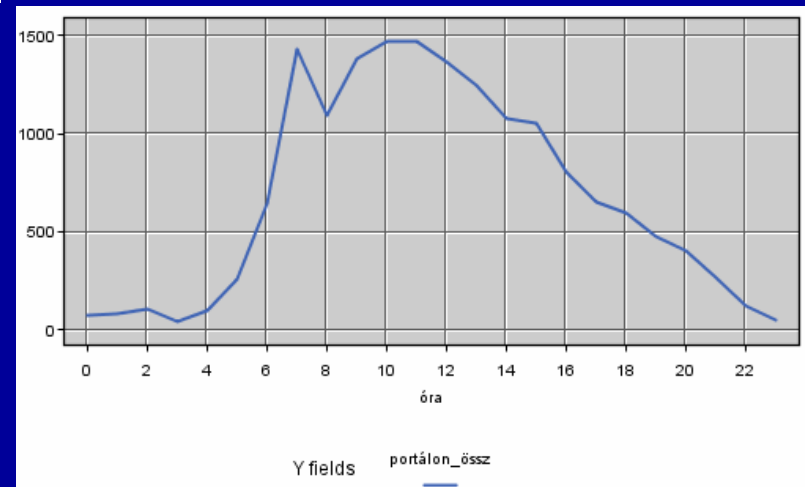
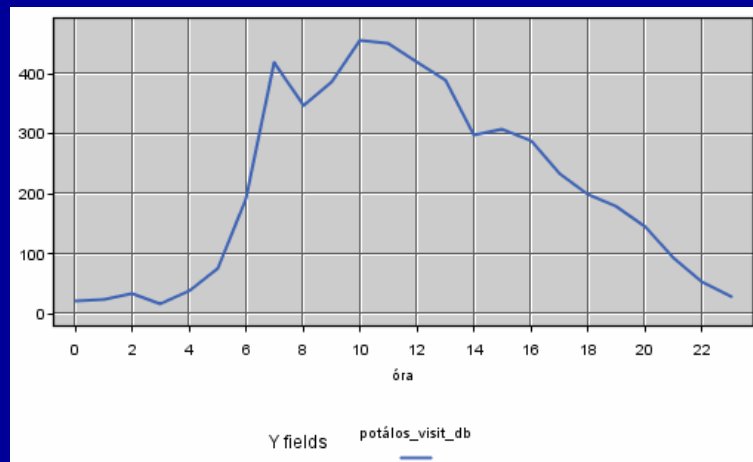
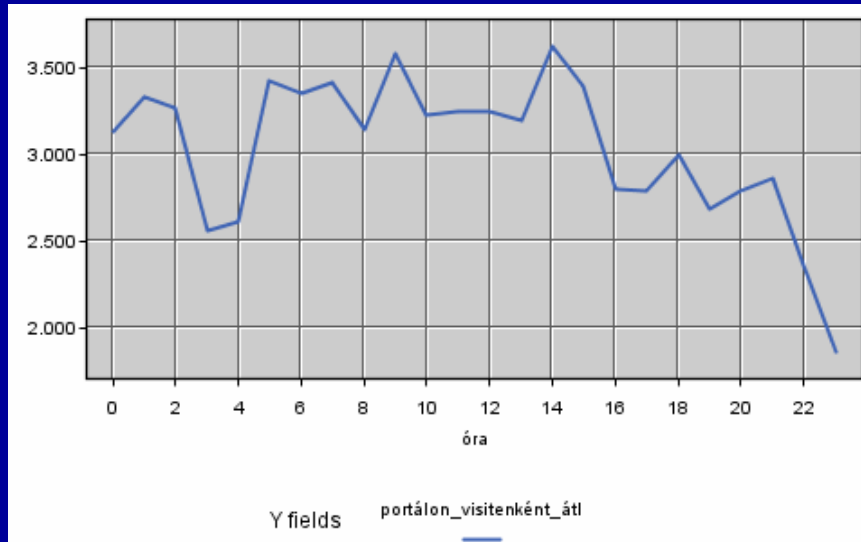


visitbeli portálon-kattintások száma a visitkezdés időpontja függvényében



3. Saját portáлом vizsgálata

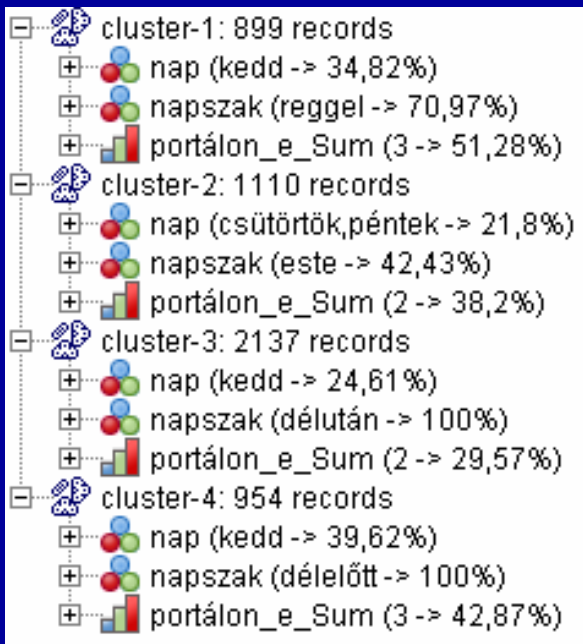
3.3 A portálon való kattintást tartalmazó visitek esetében vizsgáltam a portálon való kattintások számát, arányát - aszerint, hogy milyen napszakban történt.



3. Saját portálom vizsgálata

3.4a) Visitek Twostep klaszterezése időpont és kattintási mennyiség szerint

3.4b) Kétféle C5.0 döntési fa, cél, hogy egy visitben mennyi kattintás a portálon. A „mennyi kattintás” fogalmát is Twostep modellel alakítottam ki.



```

Event Name = Szki [Mode: nemkatt] => nemkatt
Event Name = SzabadBolcseszlet [Mode: nemkatt] => nemkatt
Event Name = Szenatus [Mode: nemkatt] => nemkatt
Event Name = SzervezetiEgységLista [Mode: nemkatt] => nemkatt
Event Name = Szlavisztika [Mode: nemkatt] => nemkatt
Event Name = TMK [Mode: nemkatt] => nemkatt
Event Name = TTK [Mode: nemkatt] => nemkatt
Event Name = Tanito [Mode: nemkatt] => nemkatt
Event Name = Targy [Mode: nemkatt] => nemkatt
Event Name = TermektervezoMuszakaiMenedzserAszisztens [Mode: nemkatt] => nemkatt
Event Name = TestNeveloEdzo [Mode: nemkatt] => nemkatt
Event Name = Tortenelem [Mode: nemkatt] => nemkatt
Event Name = Zsp [Mode: alig_katt]
  bennfentes = T [Mode: sok_katt] => sok_katt
  bennfentes = F [Mode: alig_katt]
    belső = T [Mode: alig_katt] => alig_katt
    belső = F [Mode: kevés_katt] => kevés_katt
    
```

```

visit_katt_db > 488 [Mode: sok_katt]
  session_hossz <= 719,967 [Mode: sok_katt] => sok_katt
  session_hossz > 719,967 [Mode: kevés_katt] => kevés_katt
Event Name = TermektervezoMuszakaiMenedzserAszisztens [Mode: nemkatt] => nemkatt
Event Name = TestNeveloEdzo [Mode: nemkatt] => nemkatt
Event Name = Tortenelem [Mode: nemkatt] => nemkatt
Event Name = Zsp [Mode: alig_katt]
  visit_katt_db <= 3 [Mode: kevés_katt]
    visit_katt_db <= 2 [Mode: kevés_katt] => kevés_katt
    visit_katt_db > 2 [Mode: kevés_katt]
      session_hossz <= 0,833 [Mode: kevés_katt] => kevés_katt
      session_hossz > 0,833 [Mode: kevés_katt]
        session_hossz <= 13,267 [Mode: kevés_katt] => kevés_katt
        session_hossz > 13,267 [Mode: alig_katt] => alig_katt
  visit_katt_db > 3 [Mode: alig_katt]
    visit_katt_db <= 5 [Mode: alig_katt] => alig_katt
    visit_katt_db > 5 [Mode: sok_katt]
      bennfentes = T [Mode: sok_katt] => sok_katt
      bennfentes = F [Mode: sok_katt]
        visit_katt_db <= 8 [Mode: alig_katt]
          session_hossz <= 0,667 [Mode: alig_katt]
            session_hossz <= 0,367 [Mode: kevés_katt] => kevés_katt
            session_hossz > 0,367 [Mode: alig_katt] => alig_katt
          session_hossz > 0,667 [Mode: alig_katt]
            nap = csütörtök [Mode: alig_katt]
              belső = T [Mode: sok_katt] => sok_katt
              belső = F [Mode: alig_katt] => alig_katt
            nap = hétfő [Mode: alig_katt] => alig_katt
            nap = kedd [Mode: sok_katt] => sok_katt
            nap = péntek [Mode: alig_katt] => alig_katt
            nap = szerda [Mode: alig_katt] => alig_katt
    
```

Results for output field portál_típ

Comparing \$C-portál_típ with portál_típ

Correct	350 907	99,48%
Wrong	1 828	0,52%
Total	352 735	

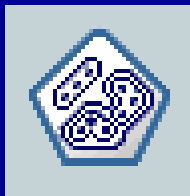
Results for output field portál_típ

Comparing \$C-portál_típ with portál_típ

Correct	351 271	99,58%
Wrong	1 464	0,42%
Total	352 735	

4. Összefoglalás

- Fő célom az volt, hogy az SPSS Clementine Web Mining adta lehetőségek sokféleségét egy konkrét példán bemutassam.
- Tapasztalatom és végső következtetésem, hogy a legbonyolultabb website-okkal kapcsolatos kérdéshez is lehet arra választ adó streamet írni. Ez arra mutat, hogy a Web Mininggal való kutatásokban a kérdések feltevésén lesz a fő hangsúly.
- Mivel az oktatásban való felhasználási lehetőségeket kutatom, megpróbáltam bemutatni, hogy érdemes elszakadni az elsősorban kereskedelmi célú beépített demostreamektől.
- A rendszergazdák felé hasznos jelzést adhatok, a napszakokban és hétvégeken való használat, a leglátogatottabb portálok, a visitkezdetek megoszlása a szakember kezében minőségbiztosítási szemmel nézve is hasznos információ.
- Saját portálom vizsgálata mutatja, milyen felhasználói szokásaik vannak hallgatóimnak. Ez olyan tanulási szokásokra (pl. ZH előtti kampánytanulás, oktatófájlok hirtelen tömeges letöltése) utal, amin talán változtathatnak a szokásaikkal való szembesítés révén.
- A kutatást a jövőben újabb adatokkal és újabb kérdésekkel is érdekes lesz folytatni.



Köszönöm a figyelmet!

