

PxEdit: rakenteiset tiedostot

Yleistä	2
Perusmuoto	3
Perusmuoto koodeilla	7
Taulukkomuoto	8
Avainsanalohko	9
Kielitunniste	10
Monikielinen taulukko	11
Hierarkkinen muuttuja	12
Esimerkki	13
Peräkkäistiedostoista	16
Lisäohjeita	17

12.11.2019

Yleistä

PxEditioniin voidaan siirtää taulukoita teksti- tai Excel-tiedostoista, suoraan Excelistä tai leikepöydän avulla (esimerkiksi LibreOffice-taulukoista).

Taulukoilla tulee olla px-muunnokseen soveltuva **rakenne**.

Rakenteisen taulukon tulee sisältää välttämättömät tiedot:

- taulukon otsikko,
- muuttujien nimet,
- muuttujien arvotekstit (arvot, nimikkeet) ja
- numeerinen dataosa.

Kaikkia taulukon data-arvoja ei välttämättä tarvita syötetiedoissa.

Puuttuvat data-arvot korvataan erikseen määritellyllä **täytekoodilla**.

Taulukon muuttujien tai muuttujien arvotekstien ei tarvitse olla missään määrättyssä järjestyksessä ja jopa toisteiset datasoluvitteet sallitaan (tällöin viimeistä löytyvää viitettä käytetään kyseiselle datasolulle).

PxEdit varoittaa käyttäjää toisteisista datasoluvitteistä.

Huomaa lisäksi, että:

- tekstitiedostoissa solut erotetaan joko sarkainmerkeillä tai puolipisteillä
- pilkkujen tai välilyöntien käyttöä ei suositella
- datasyötteessä hyväksytään desimaalierottimena joko pilkku tai piste
- tuhaterottimia **ei** tule käyttää
- datasoluissa voi olla myös PxWebissä käytettäviä pistekoodoja tai niille mahdollisesti määriteltyjä korvaavia esitysmuotoja
- ellei yksittäisen datasolun arvoa pystytä muuntamaan luvuksi tai pistekoodiksi, se korvataan täytekoodilla.

PxEdit ei varoita, jos yksittäisen luvun muuntaminen ei onnistu.

Seuraavissa esimerkeissä käytetään pääasiassa Excel-tilaukkoesimerkkejä, joista tilaukko rakenne käy selkeästi ilmi. Samat periaatteet pätevät myös tekstitiedostojen käsittelyssä.

12.11.2019

Perusmuoto

Yksinkertaisimmillaan siirtotaulukossa on **ensimmäisellä rivillä** (kulmasolussa) taulukon otsikko. Seuraavalla rivillä ovat muuttujien nimet. Muuttujien **alla** ovat riveittäin jokaisen muuttujan arvot ja lopuksi niitä vastaava numeerinen data-arvo.

Muuttujien nimien alisen osan voi mieltää koostuvan sarakkeista, joilla ovat ensin eri muuttujien arvotekstit ja viimeisellä sarakkeella aina kunkin rivin muuttuja-arvo-yhdistelmää vastaava data-arvo.

Tämän tyyppinen tiedosto lienee luontevin relaatiokannoista tuotettavien taulukoiden siirrossa.

Oheisessa taulukossa on neljä muuttujaa; jokaisen muuttujan arvotekstit on annettu kullakin rivillä erikseen:

	A	B	C	D	E	F	G
1	Yöpymiset ja saapuneet vieraat asuinmaittain						
2	Vuosi	Maakunta	Maa	Majoitusliikkeen tyyppi	Tiedot		
3	2018	Ahvenanmaa	Suomi	Kaikki majoitusliikkeet	Saapuneet, lkm	93944	
4	2018	Ahvenanmaa	Suomi	Kaikki majoitusliikkeet	Yöpymiset, lkm	167165	
5	2018	Ahvenanmaa	Suomi	Hotellit	Saapuneet, lkm	65540	
6	2018	Ahvenanmaa	Suomi	Hotellit	Yöpymiset, lkm	108485	
7	2018	Ahvenanmaa	Eurooppa	Kaikki majoitusliikkeet	Saapuneet, lkm	107457	
8	2018	Ahvenanmaa	Eurooppa	Kaikki majoitusliikkeet	Yöpymiset, lkm	221656	
9	2018	Ahvenanmaa	Eurooppa	Hotellit	Saapuneet, lkm	62454	
10	2018	Ahvenanmaa	Eurooppa	Hotellit	Yöpymiset, lkm	105396	
11	2018	Ahvenanmaa	Aasia	Kaikki majoitusliikkeet	Saapuneet, lkm	276	
12	2018	Ahvenanmaa	Aasia	Kaikki majoitusliikkeet	Yöpymiset, lkm	528	
13	2018	Ahvenanmaa	Aasia	Hotellit	Saapuneet, lkm	143	
14	2018	Ahvenanmaa	Aasia	Hotellit	Yöpymiset, lkm	258	
15	2018	Ahvenanmaa	Oseania	Kaikki majoitusliikkeet	Saapuneet, lkm	133	
16	2018	Ahvenanmaa	Oseania	Kaikki majoitusliikkeet	Yöpymiset, lkm	257	
17	2018	Ahvenanmaa	Oseania	Hotellit	Saapuneet, lkm	.	
18	2018	Ahvenanmaa	Oseania	Hotellit	Yöpymiset, lkm	.	
19	2018	Ahvenanmaa	Afrikka	Kaikki majoitusliikkeet	Saapuneet, lkm	..	
20	2018	Ahvenanmaa	Afrikka	Kaikki majoitusliikkeet	Yöpymiset, lkm	..	
21	2018	Ahvenanmaa	Afrikka	Hotellit	Saapuneet, lkm	..	
22	2018	Ahvenanmaa	Afrikka	Hotellit	Yöpymiset, lkm	..	
23	2018	Ahvenanmaa	Amerikka	Kaikki majoitusliikkeet	Saapuneet, lkm	787	
24	2018	Ahvenanmaa	Amerikka	Kaikki majoitusliikkeet	Yöpymiset, lkm	1547	

12.11.2019

Muuttujien arvot voidaan antaa myös hierarkkisesti.

Tyhjän kentän sisältönä käytetään viimeksi annettua arvotekstiä.

Jokaisella muuttujalla tulee aina olla **ensimmäinen** arvoteksti annettuna.

	A	B	C	D	E	F	G
1	Yöpymiset ja saapuneet vieraat asuinmaittain						
2	Vuosi	Maakunta	Maa	Majoitusliikkeen tyyppi	Tiedot		
3	2018	Ahvenanmaa	Suomi	Kaikki majoitusliikkeet	Saapuneet, lkm	93944	
4					Yöpymiset, lkm	167165	
5				Hotellit	Saapuneet, lkm	65540	
6					Yöpymiset, lkm	108485	
7			Eurooppa	Kaikki majoitusliikkeet	Saapuneet, lkm	107457	
8					Yöpymiset, lkm	221656	
9				Hotellit	Saapuneet, lkm	62454	
10					Yöpymiset, lkm	105396	
11			Aasia	Kaikki majoitusliikkeet	Saapuneet, lkm	276	
12					Yöpymiset, lkm	528	
13				Hotellit	Saapuneet, lkm	143	
14					Yöpymiset, lkm	258	
15			Oseania	Kaikki majoitusliikkeet	Saapuneet, lkm	133	
16					Yöpymiset, lkm	257	
17				Hotellit	Saapuneet, lkm	.	
18					Yöpymiset, lkm	.	
19			Afrikka	Kaikki majoitusliikkeet	Saapuneet, lkm	..	
20					Yöpymiset, lkm	..	
21				Hotellit	Saapuneet, lkm	..	
22					Yöpymiset, lkm	..	
23			Amerikka	Kaikki majoitusliikkeet	Saapuneet, lkm	787	
24					Yöpymiset, lkm	1547	
25				Hotellit	Saapuneet, lkm	437	

Seuraavassa esimerkissä on myös perusmuotoinen taulukko, mutta nyt kaikkia data-arvoja ei ole mukana:

	A	B	C	D	E	F
1	Yöpymiset ja saapuneet vieraat asuinmaittain					
2	Vuosi	Maakunta	Maa	Majoitusliikkeen tyyppi	Tiedot	
3	2018	Ahvenanmaa	Suomi	Kaikki majoitusliikkeet	Saapuneet, lkm	93944
4					Yöpymiset, lkm	167165
5				Hotellit	Saapuneet, lkm	65540
6					Yöpymiset, lkm	108485
7			Eurooppa	Kaikki majoitusliikkeet	Saapuneet, lkm	107457
8					Yöpymiset, lkm	221656
9				Hotellit	Saapuneet, lkm	62454
10					Yöpymiset, lkm	105396
11			Aasia	Kaikki majoitusliikkeet	Saapuneet, lkm	276
12					Yöpymiset, lkm	528
13				Hotellit	Saapuneet, lkm	143
14					Yöpymiset, lkm	258
15			Oseania	Kaikki majoitusliikkeet	Saapuneet, lkm	133
16			Amerikka	Kaikki majoitusliikkeet	Saapuneet, lkm	787
17					Yöpymiset, lkm	1547
18				Hotellit	Saapuneet, lkm	437
19					Yöpymiset, lkm	849
20		Etelä-Karjala	Suomi	Kaikki majoitusliikkeet	Saapuneet, lkm	263543
21					Yöpymiset, lkm	476844
22				Hotellit	Saapuneet, lkm	232938
23					Yöpymiset, lkm	387283
24			Eurooppa	Kaikki majoitusliikkeet	Saapuneet, lkm	104988
25					Yöpymiset, lkm	202720

Taulukko näyttää PxEditiin siirrettynä seuraavan kaltaiselta, otsikointi on oletusarvoisesti hierarkkinen ja täytekoodina on kaksi pistettä:

[1] Yöpymiset ja saapuneet vieraat asuinmaittain muuttujina Vuosi, Maakunta, Ma...

TiedostoMuokkaaIkkunaKieli

2018	Ahvenanmaa	Suomi	Kaikki majoitu...	Saapuneet, lkm	93944
				Yöpymiset, lkm	167165
			Hotellit	Saapuneet, lkm	65540
				Yöpymiset, lkm	108485
		Eurooppa	Kaikki majoitu...	Saapuneet, lkm	107457
				Yöpymiset, lkm	221656
			Hotellit	Saapuneet, lkm	62454
				Yöpymiset, lkm	105396
		Aasia	Kaikki majoitu...	Saapuneet, lkm	276
				Yöpymiset, lkm	528
			Hotellit	Saapuneet, lkm	143
				Yöpymiset, lkm	258
		Oseania	Kaikki majoitu...	Saapuneet, lkm	133
				Yöpymiset, lkm	..
			Hotellit	Saapuneet, lkm	..
				Yöpymiset, lkm	..
		Amerikka	Kaikki majoitu...	Saapuneet, lkm	787
				Yöpymiset, lkm	1547
			Hotellit	Saapuneet, lkm	437
				Yöpymiset, lkm	849
		Afrikka	Kaikki majoitu...	Saapuneet, lkm	..
				Yöpymiset, lkm	..
			Hotellit	Saapuneet, lkm	..
				Yöpymiset, lkm	..
Etelä-Karjala	Suomi	Kaikki majoitu...	Saapuneet, lkm	263543	
			Yöpymiset, lkm	476844	
	Hotellit	Saapuneet, lkm	232938		
		Yöpymiset, lkm	387283		
Eurooppa	Kaikki majoitu...	Saapuneet, lkm	104988		
		Yöpymiset, lkm	..		

<456 riviä x 1 saraketta = 456 lukua>

Veli-Matti Jantunen

12.11.2019

Perusmuoto koodeilla

Siirrettävän taulukon yhteydessä voidaan antaa muuttujien arvotekstien lisäksi näitä vastaavat **luokituskoodit**. Koodit ja tekstit annetaan **samalla** muuttujanimellä, jolloin ensimmäisellä (vasemmanpuoleisella) sarakkeella annetaan koodit ja seuraavalla tekstit.

	A	B	C	D	E	F	G	H	I
1	Yöpymiset ja saapuneet vieraat asuinmaittain								
2	Vuosi	Maakunta	Maakunta	Maa	Maa	Majoitusliikkeen tyyppi	Majoitusliikkeen tyyppi	Tiedot	
3	2018	MK21	Ahvenanmaa	A002	Suomi	01	Kaikki majoitusliikkeet	Saapuneet, lkm	93944
4	2018	MK21	Ahvenanmaa	A002	Suomi	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	167165
5	2018	MK21	Ahvenanmaa	A002	Suomi	02	Hotellit	Saapuneet, lkm	65540
6	2018	MK21	Ahvenanmaa	A002	Suomi	02	Hotellit	Yöpymiset, lkm	108485
7	2018	MK21	Ahvenanmaa	B001	Eurooppa	01	Kaikki majoitusliikkeet	Saapuneet, lkm	107457
8	2018	MK21	Ahvenanmaa	B001	Eurooppa	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	221656
9	2018	MK21	Ahvenanmaa	B001	Eurooppa	02	Hotellit	Saapuneet, lkm	62454
10	2018	MK21	Ahvenanmaa	B001	Eurooppa	02	Hotellit	Yöpymiset, lkm	105396
11	2018	MK21	Ahvenanmaa	B004	Aasia	01	Kaikki majoitusliikkeet	Saapuneet, lkm	276
12	2018	MK21	Ahvenanmaa	B004	Aasia	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	528
13	2018	MK21	Ahvenanmaa	B004	Aasia	02	Hotellit	Saapuneet, lkm	143
14	2018	MK21	Ahvenanmaa	B004	Aasia	02	Hotellit	Yöpymiset, lkm	258
15	2018	MK21	Ahvenanmaa	B005	Oseania	01	Kaikki majoitusliikkeet	Saapuneet, lkm	133
16	2018	MK21	Ahvenanmaa	B005	Oseania	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	257
17	2018	MK21	Ahvenanmaa	B005	Oseania	02	Hotellit	Saapuneet, lkm	.
18	2018	MK21	Ahvenanmaa	B005	Oseania	02	Hotellit	Yöpymiset, lkm	.
19	2018	MK21	Ahvenanmaa	B006	Afrikka	01	Kaikki majoitusliikkeet	Saapuneet, lkm	..
20	2018	MK21	Ahvenanmaa	B006	Afrikka	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	..
21	2018	MK21	Ahvenanmaa	B006	Afrikka	02	Hotellit	Saapuneet, lkm	..
22	2018	MK21	Ahvenanmaa	B006	Afrikka	02	Hotellit	Yöpymiset, lkm	..
23	2018	MK21	Ahvenanmaa	B007	Amerikka	01	Kaikki majoitusliikkeet	Saapuneet, lkm	787
24	2018	MK21	Ahvenanmaa	B007	Amerikka	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	1547
25	2018	MK21	Ahvenanmaa	B007	Amerikka	02	Hotellit	Saapuneet, lkm	437
26	2018	MK21	Ahvenanmaa	B007	Amerikka	02	Hotellit	Yöpymiset, lkm	849
27	2018	MK09	Etelä-Karjala	A002	Suomi	01	Kaikki majoitusliikkeet	Saapuneet, lkm	263543
28	2018	MK09	Etelä-Karjala	A002	Suomi	01	Kaikki majoitusliikkeet	Yöpymiset, lkm	476844
29	2018	MK09	Etelä-Karjala	A002	Suomi	02	Hotellit	Saapuneet, lkm	232938

Koodeja ei tarvitse antaa kaikille muuttujille.

PxEditin sisäänlukurutiini käsittelee jokaista mahdollista arvo-koodi-paria erillisenä, jolloin samoja arvotekstejä (vaikkapa *Vyörytyserät* tai *Yhteensä*) voi olla samassa muuttujassa useita, mikäli näillä on **eri** koodit.

Px-standardin mukaan arvotekstien tulee **aina** olla uniikkeja.

PxEdit sallii toisteisten arvotekstien käsittelyn jättäen käyttäjän vastuulle näiden muuntamisen uniikeiksi esimerkiksi lisäämällä luokituskoodin arvotekstien eteen.

Käyttäjää varoitetaan aina toisteisista koodeista tai arvoista.

Veli-Matti Jantunen

12.11.2019

Taulukkomuoto

Taulukkomuotoisessa esityksessä käytetään sekä rivi- että sarakemuuttujia. Dataosa on tällöin monisarakkeisena taulukkona. Perusmuotoon verrattuna taulukossa tulee olla myös sarakemuuttujien kuvaus, jossa muuttujien nimet annetaan otsikkorivin jälkeen **riveittäin** ensimmäisellä sarakkeella ja muuttujien arvotekstit ovat omilla datasarakkeillaan.

	A	B	C	D	E	F	G	H	I
1	Yöpymiset ja saapuneet vieraat asuinmaittain								
2	Majoitusliikkeen tyyppi					01		02	
3	Majoitusliikkeen tyyppi					Kaikki majoitusliikkeet		Hotellit	
4	Tiedot					Saapuneet, lkm	Yöpymiset, lkm	Saapuneet, lkm	Yöpymiset, lkm
5	Vuosi	Maakunta	Maakunta	Maa	Maa				
6	2018	MK21	Ahvenanmaa	A002	Suomi	93944	167165	65540	108485
7				B001	Eurooppa	107457	221656	62454	105396
8				B004	Aasia	276	528	143	258
9				B005	Oseania	133	257	.	.
10				B006	Afrikka
11				B007	Amerikka	787	1547	437	849
12		MK09	Etelä-Karjala	A002	Suomi	263543	476844	232938	387283
13				B001	Eurooppa	104988	203739	87810	156255
14				B004	Aasia	8608	17919	8294	10060
15				B005	Oseania	81	183	54	128
16				B006	Afrikka	31	52	.	.
17				B007	Amerikka	767	1870	689	1564
18		MK14	Etelä-Pohjanmaa	A002	Suomi	400576	676854	296880	489928
19				B001	Eurooppa	9278	23344	7807	18865
20				B004	Aasia	1105	5867	852	2897
21				B005	Oseania	57	267	51	261
22				B006	Afrikka
23				B007	Amerikka	351	882	312	766
24		MK10	Etelä-Savo	A002	Suomi	297541	522275	225006	346213
25				B001	Eurooppa	47462	142661	22006	51886
26				B004	Aasia	3143	6679	2233	3634
27				B005	Oseania	285	629	128	265
28				B006	Afrikka	55	158	49	133
29				B007	Amerikka	929	1969	596	1084

Minkä tahansa mielekkään taulukkoesityksen voi muuntaa helpohkosti PxEditin hyväksymään rakenteiseen taulukkomuotoon. Esityksen muunnos moniulotteiseksi px-tilaukiksi perustuu **tyhjien** solujen olemassa oloon.

Taulukon otsikko-osan muotoilussa tulee olla tarkkana, koska väärin tulkittu taulukon muotoilu saattaa tuottaa tarpeettoman suuren tulos-tilaukun.

Avainsanalohko

Taulukon yhteydessä voidaan antaa myös taulukkoon liittyviä px-avainsanoja, paitsi DATA, KEYS, LANGUAGES, PARTITIONED, VALUES ja kaikki hierarkia-avainsanat. Taulukon kieliasetus (LANGUAGE) kannattaa antaa mieluiten otsikon viereisessä solussa.

Avainsanalohko erotetaan varsinaisesta taulukosta kahdella tyhjällä rivillä.

Taulukkokohtaiset avainsanat annetaan riveittäin, avainsana ensin ja sisältö viereisessä solussa.

119	2018	MK02	Varsinais-Suomi	B007	Amerikka	5087
120						
121						
122	COPYRIGHT	YES				
123	UNITS	lkm				
124	NOTE	taulukon alaviiteteksti				
125						

Muuttujakohtaiset avainsanat tarvitsevat vastaavan muuttujannimen omalle rivilleen; tämä osoitetaan ohjainkoodilla `variablename`.

125				
126	NOTE		maakunta-muuttujan alaviite	
127	TIMEVAL			A
128	variablename		Maakunta	Vuosi
129				

Huomaa, että TIMEVAL-asetukseen riittää pelkkä jaksotuskoodi, PxEdit asettaa muuttujalistaa ja jaksotusta vastaavan aikasarjausekkeen, jos mahdollista.

Arvokohtaiset avainsanat tarvitsevat vielä vastaavan arvotekstin tai -koodin, joka osoitetaan ohjainkoodilla `valuetext`.

129					
130	VALUENOTE		Uudenmaan alaviite	Varsinais-Suomen alaviite	
131	variablename		Maakunta	Maakunta	
132	valuetext		Uusimaa	MK02	
133					

Solukohtaisten avainsanojen määrittäminen on sitten jo hiukan monimutkaisempaa: ohjainkoodeina käytetään nyt taulukon muuttujien nimiä.

134	CELLNOTE					Hotelleihin saapuneiden suomalaisten alaviite
135	Vuosi					2018
136	Maakunta				*	
137	Maa				Suomi	
138	Majoitusliikkeen tyyppi				Hotellit	
139	Tiedot				Saapuneet, lkm	
140						

Usean eri tasoisen avainsanan antaminen samalla on mahdollista.

Avainsanojen sisällöt tarkistetaan kevyemmin kuin taulukkoa avattaessa, joten avainsanalohkoa laadittaessa on syytä olla huolellinen.

Muuttujien ja arvotekstien käsittelyssä kirjainkokoja ei oteta huomioon.

Avainsanalohko voidaan antaa myös käännettynä, jolloin avainsanat ja ohjainkoodit ovat omalla rivillään ja niitä vastaavat tiedot ovat omilla sarakkeillaan.

Kielitunniste

Taulukon otsikon viereisessä solussa voi antaa taulukon kielikoodin. Tämä välitetään taulukon LANGUAGE-avainsanaksi.

	A	B	C	D	E	F
1	Yöpymiset ja saapuneet vieraat asuinmaittain	fi				
2	Majoitusliikkeen tyyppi			Kaikki majoitusliikkeet		Hotellit
3	Tiedot			Saapuneet, lkm	Yöpymiset, lkm	Saapuneet, lkm
4	Vuosi	Maakunta	Maa			
5	2018	Ahvenanmaa	Suomi	93944	167165	655

Veli-Matti Jantunen

12.11.2019

Monikielinen taulukko

Taulukon otsikon viereisessä solussa annetaan taulukon kielikoodit pilkuilla erotettuna listana. Eri kieliset tekstit annetaan samalla **pääkielisellä** muuttujanimellä pääkielisen sarakkeen tai rivin jälkeen kielikoodien ilmaisemassa järjestyksessä.

1	Yöpymiset ja saapuneet vieraat asuinmaittain	fi,en					
2	Vuosi	Maakunta	Maakunta	Maa	Maa	Majoitusliikkeen tyyppi	Majoitusliikkeen tyyppi
3	2018	Ahvenanmaa	Åland	Suomi	Finland	Kaikki majoitusliikkeet	All accommodation establis
4	2018	Ahvenanmaa	Åland	Suomi	Finland	Kaikki majoitusliikkeet	All accommodation establis
5	2018	Ahvenanmaa	Åland	Suomi	Finland	Hotellit	Hotels
6	2018	Ahvenanmaa	Åland	Suomi	Finland	Hotellit	Hotels
7	2018	Ahvenanmaa	Åland	Eurooppa	Europe	Kaikki majoitusliikkeet	All accommodation establis
8	2018	Ahvenanmaa	Åland	Eurooppa	Europe	Kaikki majoitusliikkeet	All accommodation establis
9	2018	Ahvenanmaa	Åland	Eurooppa	Europe	Hotellit	Hotels
10	2018	Ahvenanmaa	Åland	Eurooppa	Europe	Hotellit	Hotels
11	2018	Ahvenanmaa	Åland	Aasia	Asia	Kaikki majoitusliikkeet	All accommodation establis

Muuttujien nimiä ei voi suoraan antaa kielittään itse taulukossa, vaan siihen tulee käyttää avainsanalohkoa sekä VARIABLES- ja languagecode-tunnisteita.

458	2018	Varsinais-Suom	Southwest	Amerikka	America	Hotellit	Hc
459							
460							
461	languagecode	en					
462	CONTENTS	Nights spent and arrivals by country of residence					
463	VARIABLES	Year,Region,Country,Type of accommodation establishment,Information					
464							

Kaikille muuttujille ei tarvitse antaa erillistä kielitietoa silloin, kun arvotekstit ovat samat (esimerkiksi *Vuosi*-muuttujalle). Kaikki eri kieliset arvotekstit tarvitaan kuitenkin silloin, jos halutaan välittää muuttujalle arvokoodeja.

Veli-Matti Jantunen

12.11.2019

Hierarkkinen muuttuja

Rakenteisessa taulukossa voidaan antaa hierarkkisen muuttujan **hierarkiatasot** (päätaso = 0) omalla sarakkeellaan (tai rivillään) kooditietoa ennen.

Hierarkiasarakkeessa voidaan antaa myös vastaava hierarkiakoodisto, jolloin PxEdit yrittää päätellä käytettävän hierarkiarakenteen tämän perusteella.

1	Tuottajahintaindeksi tuotteittain				
2	Vuosi			2010	2011
3	Toimiala	Toimiala	Toimiala		
4	0	SSS	Yhteensä	96,7	102,
5	1	A	A MAA-, METSÄ- JA KALATALOUSTUOTTEET
6	2	01	01 Maatalous- ja riistataloustuotteet sekä niihin liittyvät palvelut
7	3	01.1	01.1 Yksivuotiset viljelykasvit
8	4	01.11	01.11 Viljakasvit (pl. riisi), palkokasvit ja öljysiemenet
9	4	01.13	01.13 Vihannekset ja melonit, juurekset ja mukulat
10	3	01.2	01.2 Monivuotiset viljelykasvit
11	4	01.23	01.23 Sitruhedelmät
12	4	02.20	02.20 Raakapuu
13	2	03.0	03.0 Kala ja muut kalastustuotteet; vesiviljelytuotteet; kalastusta avustavat palvelut
14	3	03.00	03.00 Kala ja muut kalastustuotteet; vesiviljelytuotteet; kalastusta avustavat palvelut
15	1	B	B KAIVOSTOIMINTA JA LOUHINTA	93,2	100,
16	2	07	07 Metallimalmit	118,3	127,
17	3	07.2	07.2 Värimetallimalmit	118,3	127,
18	4	07.20	07.20 Muut värimetallimalmit ja pikasteet	118,3	127,

Veli-Matti Jantunen

12.11.2019

Esimerkki

Muokataan oheinen taulukko siirtoa varten:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Kotieläimet ja kalastus															
2																
3		hevostia					muuleja					aaseja				
4		1 000					1 000					1 000				
5		2000	2001	2002	2003	2004	2000	2001	2002	2003	2004	2000	2001	2002	2003	2004
6	Afganistan	104	104	104			30	30	30			920	920	920		
7	Alankomaat	118	120	121	126	120										
8	Albania	63	67	65	63	65	21	22	20	19	22	120	105	96	88	105
9	Algeria	44	43	44	44	44	43	43	43	43	43	178	170	170	170	
10	Andorra															
11	Angola	1	1	1	1	1						5	5	5	5	5
12	Antigua ja Barbuda	1	1	1	1	1						1	1	2	2	2
13	Arabiemiirikunnat	0	0	0	0	0										
14	Argentiina	3 600	3 600	3 650	3 655	3 655	180	180	180	185	185	95	95	95	98	98
15	Armenia	12	11	12	12	13	0	0	0	0	0	2	3	3	3	3
16	Australia	220	220	220	220	220						2	2	2	2	2
17	Azerbaidžan	61	64	66	68	68	0	0	0	0	0	36	38	39	38	38
18	Bahama															
19	Bahrain															
20	Bangladesh															
21	Barbados	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
22	Belgia	31	31	31	33	33										
23	Belize	5	5	5	5	5	4	4	5	5	5					
24	Benin	1	1	1	1	1						1	1	1	1	1
25	Bhutan	28	28	29	28	28	10	10	10	10	10	18	18	18	18	18
26	Bolivia	322	322	323	323	323	81	81	82	82	82	631	631	632	635	635
27	Bosnia ja Hertsegovina	18	18	18	18	18										
28	Botswana	33	33	33	33	33	3	3	3	3	3	330	330	330	330	330
29	Brasilia	5 832	5 801	5 900	5 901	5 901	1 348	1 346	1 350	1 351	1 350	1 242	1 239	1 250	1 250	1 260

Otsikkorivi on jo kunnossa. Poistetaan otsikon alta turha tyhjä rivi ja lisätään rivi rivimuuttujia varten:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Kotieläimet ja kalastus															
2		hevostia					muuleja					aaseja				
3		1 000					1 000					1 000				
4		2000	2001	2002	2003	2004	2000	2001	2002	2003	2004	2000	2001	2002	2003	2004
5																
6	Afganistan	104	104	104			30	30	30			920	920	920		
7	Alankomaat	118	120	121	126	120										
8	Albania	63	67	65	63	65	21	22	20	19	22	120	105	96	88	105
9	Algeria	44	43	44	44	44	43	43	43	43	43	178	170	170	170	
10	Andorra															
11	Angola	1	1	1	1	1						5	5	5	5	5
12	Antigua ja Barbuda	1	1	1	1	1						1	1	2	2	2

Seuraavaksi kirjoitetaan sarakemuuttujien nimet ensimmäiseen sarakkeeseen ja rivimuuttujien nimet omalle rivilleen vierekkäin. Yksikkötiedot lisätään jokaisen arvotekstin perään sulkeisiin, poistetaan yksikkörivi ja lisätään kielitieto. Nyt taulukon otsikkorakenne on kunnossa:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Kotieläimet ja kalastus	fi														
2	Tieto	hevosia (1000)					muuleja (1000)					aaseja (1000)				n
3	Vuosi	2000	2001	2002	2003	2004	2000	2001	2002	2003	2004	2000	2001	2002	2003	2004
4	Valtio															
5	Afganistan	104	104	104			30	30	30			920	920	920		
6	Alankomaat	118	120	121	126	120										
7	Albania	62	67	65	62	65	21	22	20	19	22	120	105	96	88	105

Excel-taulukoissa kannattaa aina tarkistaa taulukkoalueen oikea reuna ja alapuoli mahdollisten selitetekstien, alaviitteiden yms. varalta. Excel-taulukot saattavat joissain tapauksissa olla hankalia käsiteltäviä, esimerkiksi sisäinen *kuukausi/vuosi*-esitysmuoto ei muunnu oikein, sarakkeiden tai rivien poisto ei onnistu pivottoinnin tahi taulukon muotoilujen takia, samaan koneeseen asennetut eri versiot saattavat estää taulukon siirron tms.

Seuraavaksi tarkistetaan taulukon alaosa.

[illegible]

Lisätään yksi rivi taulukon alle erottamaan taulukko-osa alahuomautuksista. Luodaan avainsanalohko muuntamalla alaviitteet vastaaviksi px-avainsanoiksi. Päiväys tulee muotoilla px-standardin mukaiseksi.

[illegible]

Taulukon siirrossa kopioidaan vain taulukon dataosa, joten Excel-muotoilut (mm. kirjasinlajit ja -koot, väritykset sekä prosentti- ja valuuttakentät) eivät juuri vaikuta siirtoon. Tulostaulukon desimaaliasetukset on hyvä tarkistaa siirron jälkeen.

Esimerkkitaulukko siirrettynä PxEditioniin:

[1] Kotieläimet ja kalastus muuttujina Valtio, Tieto ja Vuosi									
Tiedosto Muokkaa Ikkuna Kieli									
	hevosia (1000)					muuleja (1000)			
	2000	2001	2002	2003	2004	2000	2001	2002	2003
Afganistan	104.0	104.0	104.0	30.0	30.0	30.0	
Alankomaat	118.0	120.0	121.0	126.0	120.0	
Albania	63.0	67.0	65.0	63.0	65.0	21.0	22.0	20.0	19
Algeria	44.0	43.0	44.0	44.0	44.0	43.0	43.0	43.0	43
Andorra	
Angola	1.0	1.0	1.0	1.0	1.0	
Antigua ja Barbuda	0.5	0.5	0.5	0.5	0.5	
Arabiemiirikunnat	0.4	0.4	0.4	0.4	0.4	
Argentiina	3600.0	3600.0	3650.0	3655.0	3655.0	180.0	180.0	180.0	185
Armenia	12.0	11.0	12.0	12.0	13.0	0.1	0.2	0.2	0
Australia	220.0	220.0	220.0	220.0	220.0	
Azerbaidžan	61.0	64.0	66.0	68.0	68.0	0.1	0.1	0.1	0

193 riviä x 45 saraketta = 8 685 lukua

fi

12.11.2019

Peräkkäistiedostoista

Joistain järjestelmistä, erityisesti vanhoista keskuskonesovelluksista, ulos saatava datamassa koostuu peräkkäistiedoston tietueista. Tällaisen peräkkäistiedoston tietuekuvauksen purkaminen muuttujiksi voidaan tehdä **tietuepiteen** avulla.

Otsikkoriviä seuraavalle riville kirjoitetaan sarakepeite, jossa jokainen tietueen kenttä peitetään merkillä, jota toistetaan tarvittava määrä. Jos tietueessa on kenttiä, välimerkkejä tms. joita ei haluta ottaa mukaan, nämä peitetään välilyönneillä. Datakenttä on aina viimeinen, ja sen merkitsemiseen riittää dataosan ensimmäisen merkin sijainnin osoittaminen (vaihtuvamittaisten dataosien pituutta ei voi tiedoston alkuosan perusteella päätellä). Peräkkäistiedostoja luettaessa siirretään tietueista löytyneet muuttujien arvotekstit sekä muuttujien arvoiksi että koodeiksi, koska yleensä peräkkäistiedostoissa käsitellään koodeja.

Seuraavassa tiedostossa on kuusi muuttujaa, puolipistettä ei oteta mukaan:

Tiedosto	Muokkaa	Muotoile	Näytä	Ohje
Valmistuneet asunnot 3/2003				
AAAAABCCCDEF G				
19941004401;0				
19941004402;0				
19941005101;5				
19941005102;570				
19941005161;4				
19941005162;454				
19941005171;1				
19941005172;116				
19941006101;2				
19941006102;254				
19941006161;2				
19941006162;254				
19941009101;2				
19941009102;215				
19941009151;1				
19941009152;95				
19941009161;1				
19941009162;120				
19941010101;3				
19941010102;426				

PxEditioniin tulevat nyt ”muuttujanimiksi” AAAA, B, CCC, D, E ja F; dataosan alku on merkitty G:llä. Nämä tulee tietenkin korvata oikeilla nimillä (AAAA näyttäisi olevan vuosi, B neljännes ja CCC kuntakoodi).

Peräkkäistiedostojen yhteydessä ei voi käyttää avainsanalohkoja.

12.11.2019

Lisäohjeita

Taulukoista luettavista muuttujien nimistä poistetaan tarvittaessa väli-lyönnit alusta ja lopusta, muuttujien arvoista (kuten luokitusnimikkeistä) ja koodista vain lopusta; näin siksi, että joissain järjestelmissä käytetään alkuvälejä osoittamaan muuttujien sisäistä hierarkiaa.

PxEdition luettavan taulukon kokoa rajoittaa vain käytössä oleva tallennustila.

Taulukkomuodon, erityisesti perusmuodon hionnassa seuraavat lisätiedot voivat olla avuksi:

- rivien *lopussa* olevia tyhjiä kenttiä ei tarvitse ilmaista oikealla määrällä peräkkäisiä kenttäerottimia (esimerkiksi otsikkorivillä)
- muuttujien arvojärjestys määritellään pelkästään arvotekstien löytymäjärjestyksen mukaan; järjestyksen voi aina korjata PxEditionin lajittelutoiminnalla, mutta haluttaessa voi muuttujan arvot antaa oikeassa järjestyksessä heti tulostaulukon *alussa* ilman datasoluja
- jos taulukon yhteydessä on avainsanalohko taikka alaviitteitä tai muita siirtotaulukkoon kuulumattomia tietoja, ne tulee erottaa vähintään kahdella tyhjällä rivillä; samoin mahdollinen avainsanalohko tulee erottaa mahdollisista muista tiedoista vähintään kahdella tyhjällä rivillä
- Excel-tilukossa voi olla dataosan (EI otsikko- tai avainsanalohkon) vieressä kommentti- tai muita soluja, jotka ohitetaan taulukkoa luettaessa
- Excel-tilukon solukommentit siirretään vastaaviksi px-alaviitetiedoiksi
- tekstitiedostoon voidaan upottaa **solukohtaisia** alaviitetietoja (tästä on lisätietoja PxEditionin ohjetiedostossa).