

Skripta ārējās tirdzniecības datu izguvei- “Intrastat_datu_izguve” apraksts

| | |
|--------------------------|--|
| Datums | 2018-02-26 |
| Apraksts | Veic ārējās tirdzniecības datu izguvi no CSP datubāzes 3, 4, 5, 6, 7 un 8 zīmēs par norādītām valstīm, precēm, ārējās tirdzniecības plūsmām un periodiem. Dati tiek saglabāti XLSX datu formātā. Skriptā tiek izmantot PX datņu funkcionalitāti https://www.scb.se/Grupp/OmSCB/API/API-description.pdf |
| Nepieciešamās R pakotnes | "data.table" "openxlsx" "gdata" "pxR" "RCurl" |
| Licence | EUPL |
| Autors | Juris Breidaks |
| Uzturētājs | juris.breidaks@csb.gov.lv |

R funkcijas izsaukšana R programmā

Intrastat_datu_izguve (filepath = ceļš, px_downloaded = FALSE, first_year = NULL, last_year = NULL, codes_digits = 8)

R funkcijas argumenti

filepath – ceļš, kurā norāda, kur atrodas konfigurācijas datne.

Ceļu norāda formātā ceļš/datnes nosaukums.txt

px_downloaded – ja arguments ir FALSE, tad tiek lejuplādēti uzdoto periodu 4 vai 8 zīmju ārējās tirdzniecības dati no CSP datubāzes, ievietojot tos konfigurācijas datnes direktorijas apakšdirektorijā “4zim” vai “8zim”. Ja arguments ir TRUE, tad lejupielāde nenotiek.

first_year – norāda no kura gada tiks lejuplādēti ārējās tirdzniecības dati, ja datu filtrēšana norādīts periods “ALL”.

last_year – norāda līdz kuram gadam tiks lejuplādēti ārējās tirdzniecības dati, ja datu filtrēšana norādīts periods “ALL”.

codes_digits – norāda cik zīmēs tiks lejuplādēti ārējās tirdzniecības dati. Iespējamie varianti - 3, 4, 5, 6, 7 un 8 zīmes.

Apraksts

Skripts veic ārējās tirdzniecības datu izguvi no CSP datubāzes 3, 4, 5, 6, 7 un 8 zīmēs par valstīm, precēm, ārējās tirdzniecības plūsmām un periodiem. Dati tiek saglabāti XLSX datu formātā.

Lai sekmīgi izgūtu ārējās tirdzniecības datus no CSP datubāzēm jābūt:

- 1) instalētai R programma;
- 2) instalētai R pakotnēm:

```
install.packages(c("data.table", "openxlsx", "gdata", "pxR", "RCurl"))
```

- 3) Izveidotai *izguves konfigurācijas datnei*, kurā norāda visus atlasē un grupēšanas parametrus vienai vai vairākām tabulām.
- 4) Izpildīta komanda, kas izpilda funkciju `Intrastat_datu_izguve`

```
source("http://www.csb.gov.lv/sites/default/files/dazadi/intrastat_datu_izguve.R")
```

Izguves konfigurācijas datnes saturs

Izguves konfigurācijas datne ir teksta fails ar paplašinājums `.txt`. Tajā definē izvades datu detalizāciju, datu atlasē- prece, valsts un periods, kā arī definē XLSX tabulas skatu. Vienā pieprasījumā var izveidot vairākas tabulas.

Izvades sagataves datnei ir jāievēro noteikta struktūra.

Atslēgvārdi

Datu detalizācija un XLSX tabulas skats `<breakdown> ... </breakdown >`
Preču kodu atlase `<codes> ... </codes>`

Valstu atlase `<countries> ... </countries>`

Periodu atlase `<periods> ... </periods>`

Ja preču, valstu vai periodu atlase nav nepieciešama, starp atslēgvārdiem ievieto "all".

Piemērs:

```
< codes >  
all  
</ codes >
```

Datu detalizācija un XLSX tabulas skats `<breakdown> ... </breakdown >`

Datu grupēšanā var norādīt vairākas pazīmes, piemēram valsts, gads, ārējās tirdzniecības plūsmas u.c., skatīt tabulu.

Datu grupēšanā pieļaujamās detalizācijas:

| | |
|--------------|--|
| Mainīgais | Apraksts |
| year | Gads |
| month | Mēnesis |
| quarter | Ceturksnis |
| country | Valsts nosaukums |
| country_code | Valsts kods |
| ALL | Visas valstis (AD, AE, AF, AG, AI, AL, AM, AN, AO, AR, AT, AU, AZ, BA, BB, BD, BE, BF, BG, BH, BI, BJ, BM, BO, BR, BS, BU, BV, BW, BY, BZ, CA, CD, CF, CG, CH, CI, CK, CL, CM, CN, CO, CR, CS, CU, CV, CY, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ER, ES, ET, FI, FJ, FO, FR, GB, GD, GE, GH, GI, GL, GM, GN, GQ, GR, GT, GW, HK, HN, HR, HT, HU, ID, IE, IL, IN, IO, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KM, KN, KP, KR, KW, KY, KZ, LA, LB, LC, LI, LK, LR, LT, LU, LY, MA, MC, MD, MG, MH, MK, ML, MM, MN, MO, MP, MR, MT, MU, MX, MY, MZ, NA, NC, NE, NG, NI, NL, NO, NP, NR, NZ, OM, PA, PE, PF, PH, PK, PL, PR, PT, PY, QA, QQ, QR, QU, QV, RO, RU, RW, SA, SC, SD, SE, SG, SI, SK, SL, SN, SR, ST, SV, SY, TG, TH, TJ, TM, TN, TO, TR, TT, TW, TZ, UA, UG, UM, US, UY, UZ, VC, VE, VG, VI, VN, VU, YE, YU, ZA, ZM, ZR, ZW) |
| EUROPE | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK, MK, TR, IS, NO, CH, LI, BY, MD, RU, UA, AX, FO, GG, IM, JE, SJ, AL, AD, BA, GI, VA, ME, SM, RS, MC |
| EU28 | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK |
| EU27 | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK |
| EU15 | BE, CZ, DK, DE, EE, IE, EL, ES, FR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, SI, SK, FI, SE, UK |
| EU12 | BE, DK, DE, IE, EL, ES, FR, IT, LU, NL, PT, UK |
| EU12 | BE, DK, DE, IE, EL, ES, FR, IT, LU, NL, PT, UK |
| EU10 | "BE, DK, DE, IE, EL, FR, IT, LU, NL, UK" |
| EAA_NON | IS, LI, NO |
| EFTA | CH, IS, LI, NO |
| CIS | AM, AZ, BY, GE, KZ, KG, MD, RU, TJ, TM, UA, UZ |
| POL | AS, CK, NU, PF, PN, TK, TO, TV, WF, WS" |
| OCE | AU, NZ, NF, FJ, NC, PG, SB, VU, GU, KI, MH, FM, NR, MP, PW, AS, CK, PF, NU, PN, WS, TK, TO, TV, WF |
| ASI_W | AE, AM, AZ, BH, GE, IL, IQ, JO, KW, LB, OM, PS, QA, SA, SY, YE |
| ASI_S_E | BN, KH, ID, LA, MY, MM, PH, SG, TH, TL, VN |
| ASI_S | AF, BD, BT, IN, IR, MV, NP, PK, LK |
| ASI_E | CN, HK, MO, KP, JP, MN, KR |
| ASI_C | KZ, KG, TJ, TM, UZ |
| AME_S | AR, BO, BR, CL, CO, EC, FK, GF, GY, PY, PE, SR, UY, VE |
| AME_C | BZ, CR, SV, GT, HN, MX, NI, PA |
| AME_N | BM, CA, GL, PM, US |
| CRB | AI, AG, AW, BS, BB, VG, KY, CU, DM, DO, GD, GP, HT, JM, MQ, MS, AN, PR, BL, KN, LC, MF, VC, TT, TC, VI |

| | |
|-------------|--|
| AME | AI, AG, AW, BS, BB, VG, KY, CU, DM, DO, GD, GP, HT, JM, MQ, MS, AN, PR, BL, KN, LC, MF, VC, TT, TC, VI, BZ, CR, SV, GT, HN, MX, NI, PA, AR, BO, BR, CL, CO, EC, FK, GF, GY, PY, PE, SR, UY, VE, BM, CA, GL, PM, US |
| AFR_W | BJ, BF, CV, CI, GM, GH, GN, GW, LR, ML, MR, NE, NG, SH, SN, SL, TG |
| AFR_W | BJ, BF, CV, CI, GM, GH, GN, GW, LR, ML, MR, NE, NG, SH, SN, SL, TG |
| AFR_S | BW, LS, NA, ZA, SZ |
| AFR_N | DZ, EG, LY, MA, SD, TN, EH |
| AFR_E | BI, KM, DJ, ER, ET, KE, MG, MW, MU, YT, MZ, RE, RW, SC, SO, UG, TZ, ZM, ZW |
| AFR_C | BI, CF, CG, CM, GA, GQ, RW, ST, TD |
| AFR | BI, KM, DJ, ER, ET, KE, MG, MW, MU, YT, MZ, RE, RW, SC, SO, UG, TZ, ZM, ZW, AO, CM, CF, TD, CG, CD, GQ, GA, ST, DZ, EG, LY, MA, SD, TN, EH, BW, LS, NA, ZA, SZ, BJ, BF, CV, CI, GM, GH, GN, GW, LR, ML, MR, NE, NG |
| Goods_code | Preču kods |
| Goods_name | Preču nosaukums |
| Mervienības | Preču mērvienības |
| Virziens | Plūsma |

Piemērs:

<breakdown>

```
year + month + quarter + country + country_code + EUROPE + EU28 + EU27 +
EU15 + EU12 + EAA_NON + EFTA + CIS + goods_code + goods_name ~ mervienības
+ virziens
```

</breakdown>

Preču kodu atlase <codes> ... </codes>

Preču kodu filtrēšanā izmanto ārējās tirdzniecības preču kombinēto nomenklatūru, kuru var atrast <http://www.csb.gov.lv/dokumenti/klasifikācijas-30827.html>.

Ja nepieciešams filtrēt pēc vairākiem preču kodiem, tad visus kodus sarindo vienu zem otra.

Piemērs:

```
<codes>
12101000
12102010
</codes>
```

Vienā pieprasījumā var norādīt tikai 3, 4, 5, 6, 7 un 8 zīmju preču kodus.

Valstu atlase <countries> ... </countries>

Valstu filtrēšanā izmanto valstu un teritoriju nomenklatūru Kopienas ārējās tirdzniecības statistikas un dalībvalstu savstarpējās tirdzniecības statistikas vajadzībām, kuru var atrast

<http://www.csb.gov.lv/dokumenti/klasifikācijas-30827.html>, un papildus šādus saīsinājumus, kas dod iespēju iegūt ārējās tirdzniecības datus valstu grupās:

| Saīsinājums | Paskaidrojums | Valstu kodi |
|-------------|---|--|
| ALL | Visas valstis | AD, AE, AF, AG, AI, AL, AM, AN, AO, AR, AT, AU, AZ, BA, BB, BD, BE, BF, BG, BH, BI, BJ, BM, BO, BR, BS, BU, BV, BW, BY, BZ, CA, CD, CF, CG, CH, CI, CK, CL, CM, CN, CO, CR, CS, CU, CV, CY, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ER, ES, ET, FI, FJ, FO, FR, GB, GD, GE, GH, GI, GL, GM, GN, GQ, GR, GT, GW, HK, HN, HR, HT, HU, ID, IE, IL, IN, IO, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KM, KN, KP, KR, KW, KY, KZ, LA, LB, LC, LI, LK, LR, LT, LU, LY, MA, MC, MD, MG, MH, MK, ML, MM, MN, MO, MP, MR, MT, MU, MX, MY, MZ, NA, NC, NE, NG, NI, NL, NO, NP, NR, NZ, OM, PA, PE, PF, PH, PK, PL, PR, PT, PY, QA, QQ, QR, QU, QV, RO, RU, RW, SA, SC, SD, SE, SG, SI, SK, SL, SN, SR, ST, SV, SY, TG, TH, TJ, TM, TN, TO, TR, TT, TW, TZ, UA, UG, UM, US, UY, UZ, VC, VE, VG, VI, VN, VU, YE, YU, ZA, ZM, ZR, ZW |
| EUROPA | Eiropas valstis | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK, MK, TR, IS, NO, CH, LI, BY, MD, RU, UA, AX, FO, GG, IM, JE, SJ, AL, AD, BA, GI, VA, ME, SM, RS, MC |
| EU28 | Eiropas Savienības 28 valstis | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, HR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK |
| EU27 | Eiropas Savienības 27 valstis | BE, BG, CZ, DK, DE, EE, IE, EL, ES, FR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK |
| EU25 | Eiropas Savienības 25 valstis | BE, CZ, DK, DE, EE, IE, EL, ES, FR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, SI, SK, FI, SE, UK |
| EU15 | Eiropas Savienības 15 valstis | BE, DK, DE, IE, EL, ES, FR, IT, LU, NL, AT, PT, FI, SE, UK |
| EU12 | Eiropas Savienības 12 valstis | BE, DK, DE, IE, EL, ES, FR, IT, LU, NL, PT, UK |
| EU10 | Eiropas Savienības 10 valstis | BE, DK, DE, IE, EL, FR, IT, LU, NL, UK |
| EEA_NEU | EEA-valstis izņemot ES valstis | IS, LI, NO |
| EFTA | Eiropas Brīvās tirdzniecības asociācija | CH, IS, LI, NO |
| CIS | Neatkarīgo valstu savienība (bijušā Padomju Savienība, izņemot BAL) | AM, AZ, BY, GE, KZ, KG, MD, RU, TJ, TM, UA, UZ |

Ja ir nepieciešams filtrēt pēc vairākiem valstu kodiem vai valstu grupām, tad visus kodus sarindo vienu zem otra.

Piemērs:

```
<countries>  
all  
</countries>
```

Periodu filtrēšana <periods> ... </periods>

Periodu filtrēšanā datuma formāts ir yyyy/mm.

Ja nepieciešams izvadīt visu periodu datus lieto "ALL".

Piemērs:

```
<periods>  
all  
</periods>
```

Ja nepieciešams datus izvadīt no-līdz lieto "yyyy/mm-yyyy/mm"

Piemērs:

```
<periods>  
2006/01-2016/12  
</periods>
```

Funkcijas izmantošanas piemēri

Izveidot XLSX tabulu, kurā ir informācija par ārējās tirdzniecības preces - **Kafija, grauzdēta vai negrauzdēta, ar kofeīnu vai bez kofeīna; kafijas pupiņu čaumalas un apvalki; kafijas aizstājēji, kas satur kafiju (0901)** importu un eksportu par visām valstīm periodā no 2006. gada janvāra līdz 2016. gada decembrim.

Izvades sagataves datnes sagatavošana - kafija4.txt

```
<breakdown>  
year + country ~ mervienibas + virziens  
</breakdown>
```

```
<codes>  
0901  
</codes>
```

```
<countries>  
ALL  
</countries>
```

```
<periods>  
2006/01-2016/12  
</periods>
```

Darbības R programmā

```
install.packages(c("data.table", "openxlsx", "gdata", "pxR", "RCurl"))  
  
rm(list = ls())
```

```
gc()

source("http://www.csb.gov.lv/sites/default/files/dazadi/intrastat_datu_izguve.R")

#Lai izgūtu informāciju par kafiju 4 zīmēs no Centrālās statistikas pārvaldes mājas lapas, ir nepieciešams sagatavot txt izguves datni par kafiju un pēc tam izpildīt sekojošu komandu

Intrastat_datu_izguve(filepath = "C:/kafija4.txt",
                      px_downloaded = FALSE)
```

Izveidot XLSX tabulu, kurā ir informācija par ārējās tirdzniecības preces - **apiņu rogas, kas nav smalcinātas, ne arī pulverī vai granulās (12101000) un apiņu rogas, maltas, pulverveidā vai granulās, ar augstāku lupulīna saturu; lupulīns (12102010)** importu un eksportu par visām valstīm periodā no 2000. gada janvāra līdz 2017. gada decembrim.

Izvades sagataves datnes sagatavošana - Apini8.txt

```
<breakdown>
year + month + quarter + country + country_name + goods_code + goods_name ~
mervienibas + virziens
year + month + quarter + country + goods_code + goods_name ~ mervienibas +
virziens
year + month + country + country_name ~ mervienibas + virziens
year + goods_code + goods_name ~ mervienibas + virziens
</breakdown>

<codes>
12101000
12102010
</codes>

<countries>
ALL
</countries>

<periods>
ALL
</periods>
```

Darbības R programmā

```
install.packages(c("data.table", "openxlsx", "gdata", "pxR", "RCurl"))

rm(list = ls())

gc()

source("http://www.csb.gov.lv/sites/default/files/dazadi/intrastat_datu_izguve.R")

Intrastat_datu_izguve(filepath = "C:/Apini8.txt",
                      px_downloaded = FALSE,
                      first_year = 2000,
                      last_year = 2017)
```