

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
MATEMATINĖS STATISTIKOS KATEDRA

Žana Grigianec

**KLaidų Ieškojimo Įmonių Pramoninių Gaminių Gamybės
Ataskaitose Automatizavimas**

Magistro baigiamasis darbas

Vilnius, 2009

Darbo vadovas:

Docentas, matematikos m. dr. Vytautas Kazakevičius

Recenzentas:

Docentas, matematikos m. dr. Rimantas Eidukevičius

Registracijos Nr.:

Darbo gynimo data:

TURINYS

ĮVADAS.....	4
1. ĮMONĖS PRAMONINIŲ GAMINIŲ GAMYBOS TYRIMO (P-12) APRAŠYMAS ...	7
1.1 Tyrimo raida.....	7
1.2 Tyrimo tikslai	7
1.3 Tyrimo atlikimo metodai.....	8
1.4 Statistinių duomenų surinkimas	9
1.5 Statistinių duomenų pirminis redagavimas.....	13
1.6 Statistinės informacijos pateikimas vartotojams	16
2. STATISTINIŲ DUOMENŲ KOKYBĖS KONTROLĖ	18
2.1 Imties formavimas	18
2.2 Paklaidos	19
2.3 Duomenų apdorojimas	21
2.4 Duomenų aprašymas	23
2.5 Statistinių duomenų antrinis redagavimas (siūlomas modelis).....	26
3. LOGISTINĖ REGRESIJA	33
IŠVADOS	36
SANTRAUKA	37
REZIUMĖ	38
LITERATŪROS SĄRAŠAS	39
PRIEDAI.....	40
1 Priedas	40
2 Priedas	42
3 Priedas	46
4 Priedas	48

IVADAS

Lietuvos Respublikos statistikos įstatymas, nustatantis oficialiosios statistikos¹ organizavimo bendruosius principus bei teises ir pareigas kaip teikiantiems, taip tvarkantiems ir naudojantiems statistinius duomenis², pirmo skyriaus 5 straipsnyje numato vieną iš pagrindinių oficialią statistiką tvarkančių institucijų ir įstaigų uždavinių - apdoroti, apibendrinti ir analizuoti statistinius duomenis, rengti statistinę informaciją³.

Statistikos departamentas prie Lietuvos Respublikos Vyriausybės (toliau – Statistikos departamentas) – institucija, koordinuojanti valstybės oficialiąją statistiką bei įgyvendinanti bendrą valstybės politiką statistikos metodologijos ir organizavimo srityse. Tai institucija, kuri vadovaudamasi tarptautinių organizacijų metodologiniais principais, renka, apdoroja, analizuoja ir skelbia statistikos duomenis apie šalies ekonominius, socialinius, demografinius ir aplinkos pokyčius. Siekdamas užtikrinti statistinės informacijos ir savo veiklos kokybę, Statistikos departamentas vadovaujasi Europos statistikos sistemos (ESS) kokybės deklaracijos ir Europos statistikos praktikos kodekso nuostatomis. 2007 m. Statistikos departamente įdiegta kokybės vadybos sistema, atitinkanti standarto ISO 9001:2000 reikalavimus (žiūr.: <http://www.stat.gov.lt/lt/pages/view/?id=1954&PHPSESSID=61f1824ff9d2af5c5cbda0a05d86c83a>).

Taigi, šio darbo tikslas – gerinti Pramonės statistikos skyriaus duomenų kokybę, t.y. automatizuoti įmonių pramoninių gaminių gamybos ataskaitose loginių klaidų paiešką, siekiant mažesnėmis sąnaudomis gauti tikslesnius rezultatus. Darbe panaudoti naujausi šių metų Statistikos departamento Pramonės statistikos skyriaus duomenys, gauti iš įmonės pramoninių gaminių gamybos mėnesinės ataskaitos P-12.

Statistikos departamento Pramonės skyriaus atliekamas tyrimas yra mėnesinio periodiškumo. Iš atsiskaitančių įmonių gaunamos informacijos kiekis yra didelis (virš 10 000 eilučių kas mėnesį). Surinktų duomenų tikrinimui ir jų paskelbimui skiriamos vos kelios dienos. Nuodugnus

Pagrindinės sąvokos:

¹ Oficialioji statistika - valstybės ir vietos savivaldos institucijų ir įstaigų valstybės reikmėms skirtų statistinių duomenų apie ekonominius, demografinius procesus, socialinius veiksmus ir visuomeninius bei aplinkos pokyčius rinkimo, tvarkymo ir statistinės informacijos skelbimo pagal Oficialiosios statistikos darbų programą sistema. [LR Statistikos įstatymo pakeitimo įstatymas]

² Statistiniai duomenys - respondentų pateikti bei iš tvarkomųjų, organizacinių dokumentų, registrų ir informacinių sistemų gauti duomenys statistinėms suvestinėms sudaryti. [LR Statistikos įstatymo pakeitimo įstatymas]

³ Statistinė informacija - informacija, gauta apdorojus statistinius duomenis. [LR Statistikos įstatymo pakeitimo įstatymas]

⁴ Statistinių duomenų apdorojimas - statistinių duomenų tikrinimas, sumavimas, grupavimas, klasifikavimas, statistinių lentelių ir grafikų rengimas. [LR Statistikos įstatymo pakeitimo įstatymas]

⁵ Statistiniai tyrimai - statistinių duomenų rinkimas, kaupimas, apdorojimas, analizė, statistinės informacijos rengimas ir teikimas. [LR Statistikos įstatymo pakeitimo įstatymas]

⁶ Statistinių duomenų tvarkymas - statistinio tyrimo ataskaitų bei anketų, taisyklių joms pildyti rengimas, jų teikimas respondentams, rinkimas ir rengimas apdorojimui. [LR Statistikos įstatymo pakeitimo įstatymas]

kiekvienos ataskaitos patikrinimas ir klaidų taisymas dėl laiko stokos yra neįmanomas.

Skirtinguose duomenų perdavimo etapuose yra įdiegti įvairių klaidų rūšių nustatymo kontroliniai mechanizmai. Todėl siekiant gerinti duomenų kokybę ir užtikrinti maksimalų klaidų nustatymo procentą, kontrolių skaičius turi būti nuolat didinamas. Gaminių gamybos tyrimo tikslas - atlikti statistinį įmonių pramoninių gaminių gamybos tyrimą ir turėti išsamia, naujausią ir patikimą informaciją apie pramonės produkciją, nustatyti pagrindinių gaminių gamybos, pardavimo apimtį ir jų pokyčius per tiriamąjį laikotarpį natūrine ir vertine išraiška.

Visi darbe pateikti skaičiavimai buvo atlikti statistinio paketo SAS, Oracle ir MS Excel programų pagalba. Pagrindinėje darbo dalyje pateikiamos pagrindinės logines klaidas „išgaudančios“ programos ir jų rezultatai. O prieduose sukeltos papildomos kelių mėnesių kainas ar gamybos apimtis lyginančios programos.

Darbo tikslas - gerinant Statistikos departamento Pramonės skyriaus duomenų kokybę, automatizuoti klaidų paiešką pramoninėse Gaminių gamybos ataskaitose, siekiant sumažinti su statistinės informacijos tikrinimu susijusias sąnaudas bei padidinti teikiamos informacijos patikimumą.

Uždaviniai:

- Išanalizuoti ir aprašyti P-12 tyrimą, jo tikslus, gaunamus rezultatus ir duomenis;
- Išanalizuoti ir aprašyti statistinių duomenų galimas paklaidų priežastis;
- Išanalizuoti Gaminių gamybos tyrime pasitaikančias klaidas;
- Pritaikyti logistinės regresijos metodą klaidų atsiradimo įtakojantiems faktoriams nustatyti;
- Pasiūlyti paprastą ir greitą Gaminių gamybos tyrime loginių klaidų ieškojimo būdą.

1. ĮMONĖS PRAMONINIŲ GAMINIŲ GAMYBOS TYRIMO (P-12) APRAŠYMAS

1.1 Tyrimo raida

Gamybos statistika vaidina svarbų vaidmenį rengiant ir teikiant visuomenei aktualią, kokybišką informaciją apie gaminių gamybą.

Duomenys reikalingi ne tik Europos komisijai, bet ir kiekvienai Europos Ekonominės Bendrijos (EEB) sąjungos šaliai, tos šalies vyriausybei pramonės ir prekybos politikos įvertinimui. Bendri EEB rodikliai leidžia nustatyti kiekvienos šalies pramonės veiklos rūšių efektyvumą. Ši informacija domina taip pat ir pačias įmones, kurios vis dažniau kreipiasi informacijos į Statistikos departamentą (SD) norėdamos žinoti savo vietą pramonėje bei įvertinti prekių rinkas. Šių tikslų vedamas, EUROSTAT'as (angl. *The Statistical Office of the European Communities*), glaudžiai bendradarbiaudamas su Europos Sąjungos šalių statistikomis, sukūrė gamybos statistinę sistemą - PRODCOM (iš pranc. k. – *Bendrijos produkcija*).

Lietuvoje prekių gamybos apimtys pradėtos rinkti pagal Prodcom sąrašą nuo 2002 metų. Nuo 1996m. iki 2002 m. gaminių gamybos statistikoje buvo naudojama Kombinuotoji muitų tarifų nomenklatūra (KN), kuri yra labiau pažįstama ir prieinama įmonėms, kadangi ji naudojama prekyboje su užsieniu. Siekiant patenkinti Lietuvos informacijos vartotojus, buvo sukurta nacionalinė Prodcom versija PGPK (Produktų, gaminių ir paslaugų klasifikatorius). Nacionalinė versija buvo praplėsta naujais tik Lietuvai būdingais ir svarbiais gaminių kodais, be to buvo įvestas papildomas rodiklis - pagamintos produkcijos kiekis natūriniais vienetais. Šias prekių grupes sudarė tik Lietuvoje gaminami gaminiai (pvz.: obuolių sūriai, midaus trauktinės ir pan.), kai kurios PGPK dalos buvo susmulkintos į kelis kodus (pvz.: bendras pavadinimas „dešros“ išskirtas į „rūkytos dešros“ ir „virtos dešros“), išskirti svarbūs Lietuvai gaminiai (pvz.: ruginiai miltai, ruginė duona). O pats gaminių statistikos tyrimas tiria ne tik pagamintos produkcijos struktūrą ir apimtį, bet taip pat ir parduotos produkcijos vidinius pardavimus bei eksportą. Kai tuo tarpu Prodcom sistema remiasi tik parduotos produkcijos kiekiais natūrine ir vertine išraiška (Įmonės pramoninių gaminių gamybos tyrimo metodika).

1.2 Tyrimo tikslai

Tyrimo tikslas – nustatyti pagrindinių gaminių gamybos ir pardavimo apimtį ir jų pokyčius per ataskaitinį laikotarpį (natūrine ir vertine išraiška).

Gaminių statistika turi atitikti EB ir Lietuvos nacionalinius interesus. Tai yra trumpojo laikotarpio statistika, kuri tenkina užsakovų norus turėti operatyvią bei išsamią gaminių informaciją.

Kaip jau buvo paminėta Gaminių statistika renkama pagal Eurostat'o sukurtą vieningą gamybos statistikos sistemą - Prodcom sąrašą. O galutiniai tyrimo rezultatai kasmet išsiunčiami į Eurostat'ą, kur po kruopštaus patikrinimo duomenys kartu su kitomis Europos šalimis talpinami nemokomoje bazėje.

1.3 Tyrimo atlikimo metodai

Pagal Statistikos departamento Pramonės statistikos skyriaus aprašymą pateikiamą metodinėje medžiagoje: "P-12 tyrimo populiacija (visuma) – visų rūšių ir nuosavybės formų įmonės, vykdančios pramoninę veiklą. Tyrimo vienetas – įmonė, klasifikuojama pagal EVRK 2 red.(2 priedas). Vykdomo tyrimo sritis apima Europos Bendrijos (EB) ekonominės veiklos rūšių klasifikatoriaus B (Kasyba ir karjerų eksploatavimas) ir C (Apdirbamoji gamyba) sekcijose išvardytas veiklos rūšis".

Imtis kasmet sudaroma naujai iš EVRK 2 red. klasės lygmens imant daugiausiai pardavimų vertinę išraišką turinčias įmones, kurių pardavimų suma sudarytų ne mažiau nei 90% visų šios klasės metinės pardavimų apimties. Siekiant įvertinti 100% gamybą, įtraukiamos visos veikiančios pramonės įmonės iš statistinio ūkio subjektų registro sąrašo ir įvertinamos jų metinės pajamos. Paskaičiuojamos 90% pajamų nuo kiekvienos 4 ženklų EVRK 2 red. veiklos rūšies. Jei metinės įmonių pajamos didesnės nei 250000 Lt., tai beveik visos veiklos rūšys tenkina reikalavimus. Jei pagal veiklos rūšį surenkama mažiau nei 90%, įtraukiamos įmonės, kurių metinės pajamos mažesnės nei 250000 Lt. (Įmonės pramoninių gaminių gamybos tyrimo metodika).

Gaminių gamybos tyrimui sudaroma imtis (ėmimas be grąžinimo su tikimybėmis, proporcingomis įmonės gaunamoms pajamoms), į kurią įtraukiamos visų rūšių įmonės, vykdančios pramoninę veiklą, kuriose metinės pajamos iš pagamintos produkcijos didesnės nei 0,25 mln. Lt. Taip pat patenka nepramoninės įmonės, vykdančios pramoninę veiklą, kurių metinės pajamos iš pramoninės veiklos didesnės kaip 0,5 mln. Lt. Jei įmonė netenkina šių sąlygų, tačiau gamina svarbią Lietuvos ūkiui produkciją, ši įmonė įtraukiama į tyrimą. Tik Lietuvai būdingus ar svarbius gaminius galima atpažinti pagal gaminio kodo (koduojamo pagal Nace 2 red. nacionalinę versiją - PGPK klasifikatorių) devinto ženklo nenulinę reikšmę. Tyrimas apima tik Lietuvos teritorijoje gaminamą produkciją.

Įmonių atrinkimo schema:

O_1	O_2	...	O_N
p_1	p_2	...	p_N
q_1	q_2	...	q_N

O_i – i-toji įmonė;

p_i – i-tosios įmonės patekimo į imtį tikimybė; q_i – i-tosios įmonės nepatekimo į imtį tikimybė;

$p_1 \neq p_2 \neq \dots \neq p_N$; $q_1 \neq q_2 \neq \dots \neq q_N$.

Statistinių duomenų apdorojimui naudojamos programos Oracle aplinkoje. Užklausos, suvestinės, kontrolės ir palyginimai formuojami SQL ir SAS pagalba, taip pat naudojamas MS Excel (Pivot Table) ir MS Access.

1.4 Statistinių duomenų surinkimas

Lietuvoje Gaminių statistiką reprezentuoja mėnesinė P-12 ataskaita, kurios pagalba renkami šie rodikliai:

- Pagamintos produkcijos kiekis natūriniais vienetais (kg, vnt., m ir kt.);
- Parduotos produkcijos kiekis natūriniais vienetais iš viso (kg, vnt., m ir kt.);
- Parduotos produkcijos kiekis natūriniais vienetais vidaus rinkoje (kg, vnt., m ir kt.);
- Parduotos produkcijos vertė iš viso (tūkst. lt);
- Parduotos produkcijos vertė vidaus rinkoje (tūkst. lt).


Gaminiams klasifikuoti naudojama PRODCOM gaminių sąrašo nacionalinė versija – Produktų, gaminių ir paslaugų klasifikatorius (toliau PGPK 2008) (Įmonės pramoninių gaminių gamybos tyrimo metodika) (2 priedas).

Statistinio ūkio subjektų registro duomenimis 2008 metų pabaigoje Lietuvoje buvo užregistruota 8064 vykdančios pramoninę veiklą įmonės, iš kurių tyrimui buvo atrinktos 3618 įmonių. Be minėtų, į tyrimą buvo įtrauktos 165 nepramoninės įmonės, bet didelę pajamų dalį gaunančios iš pramoninės veiklos, bei 155 įmonių (mėsos, mėsos produkcijos bei pieno, pieno produkcijos gamintojų) duomenys imami iš administracinių šaltinių, t.y. iš Žemės ūkio ir maisto produktų rinkos informacinės sistemos (ŽUMPRIS).

Klausimynas sudarytas popierinėje laikmenoje. Nuo 2006 m. sausio mėnesio parengta Įmonės pramoninių gaminių gamybos klausimyno (P-12) elektroninė forma (ABBY eFormFiller programos pagalba). Taip respondentams sudarytos sąlygos duomenis teikti jiems priimtinesniu ir patogesniu būdu. Mažinant respondentams našta, t.y. norėdami palengvinti ataskaitos pildymą ir sumažinti pildymui sugaištamą laiką, buvo sugalvota paruošti įmonėms

personifikuotas ataskaitas su išpausdintais ne tik pagrindiniais įmonės duomenimis (įmonės JAR kodas, pavadinimas bei pagrindinis veiklos kodas su pavadinimu), bet ir gaminių kodais su pavadinimais ir matavimo vienetais. Taigi, įmonėms, kurios dalyvauja tyrimuose ne pirmus metus, atsižvelgiant į jų ankstesnių laikotarpių ataskaitose nurodytus gaminamų gaminių kodus, siunčiamuose popieriniuose klausimynuose yra išpausdinami gaminių kodai pagal ataskaitiniais metais aktualiąją PGPK versiją. Elektroninę ataskaitos formą pildantiems respondentams taip pat sudarytos palankios sąlygos, nes yra sukurta statistinių ataskaitų elektroninio surinkimo sistema (žiūr.: <https://e-formos.stat.gov.lt/alfa/AW/>). Tai saugesnis, greitesnis ir patogesnis duomenų teikimo būdas, nes registruojantis į sistemą vietoj įmonės kodo naudojami Vartotojo ID ir slaptažodis. Tai leidžia vienai įmonei turėti kelis už statistinę atskaitomybę atsakingus darbuotojus. Sistemoje galima rasti visas elektroninių statistinių ataskaitų teikimo ir pildymo naujienas: konkrečios įmonės (įmonei užsiregistravus ir prisijungus prie sistemos) statistinės ataskaitos, kurias ūkio subjektas privalo atsiskaityti SD per einamuosius metus, teikimo kalendorių, elektroninių ataskaitų formų šablonus. Taip pat respondentas turi galimybę peržiūrėti ankstesnes ataskaitas ir perduoti naujas.

Įmonės pramoninių gaminių gamybos mėnesinės ataskaitos elektroninė forma



Forma patvirtinta Statistikos departamento prie Lietuvos Respublikos Vyriausybės generalinio direktoriaus 2007 m. lapkričio 14 d. įsakymu Nr. DĮ-240

(Įmonės, įstaigos pavadinimas)

Statistikos departamento prie LRV
teritorinei statistikos įstaigai

ĮMONĖS PRAMONINIŲ GAMINIŲ GAMYBOS M. MĖN. ATASKAITA
P-12 MĖNESINĖ

0000 - 00 - 000 Nr.
(Užpildymo data)

Įmonės kodas registre

Adresas, veikla nurodomi tik tuo atveju, jei jie pasikeitė

Adresas

Pagrindinės veiklos kodas (EVRK)

Pagrindinės veiklos pavadinimas (EVRK)


Pateikiama: 7 kalendorinę dieną ataskaitiniam mėnesiui pasibaigus

Pateikia: visų rūšių įmonės, vykdančios pramoninę veiklą

Prieš pildydami ataskaitą perskaitykite paaiškinimus.

Garantuojamas gautų duomenų konfidencialumas

Kviečiame pildyti elektroninę ataskaitos formą.
Pildymo tvarką rasite interneto svetainėje
<http://www.stat.gov.lt>
Statistinių ataskaitų formos



1 974 153 673 123

2


Įmonės kodas registre

P-12 mėnesinė
 mėn.

I. GAMINIŲ GAMYBA IŠ NUOSAVŲ MEDŽIAGŲ IR ŽALIAVŲ, PUSFABRIKAČIŲ

Statistiniai duomenys pateikiami pagal Produktų, gaminių ir paslaugų klasifikatorių (PGPK II dalis) (žr. www.stat.gov.lt → Klasifikatoriai, PRODCOM 2008 m. nacionalinis)

A. Gaminio kodas pagal PGPK II dalį		Parduota natūriniais vienetais		Parduota to meto kainomis be PVM ir akcizų, tūkst. Lt	
1 iš viso	2 iš viso	3 iš to skaičiaus Lietuvos rinkoje	4 iš viso	5 iš to skaičiaus Lietuvos rinkoje	
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>
A <input type="text"/>	1 <input type="text"/>	2 <input type="text"/>	3 <input type="text"/>	4 <input type="text"/>	5 <input type="text"/>



5 974 185 455 1785

5

Įmonės kodas registre

P-12 mėnesinė
 mėn.

II. GAMINIŲ GAMYBA IŠ UŽSAKOVŲ MEDŽIAGŲ BEI ŽALIAVŲ, PUSFABRIKAČIŲ

Statistiniai duomenys pateikiami pagal Produktų, gaminių ir paslaugų klasifikatorių (PGPK II dalis) (žr. www.stat.gov.lt → Klasifikatoriai, PRODCOM 2008 m. nacionalinis)

A Gaminių kodas pagal PGPK II dalį	Atliktų pramoninių darbų vertė, tūkst. Lt	
1 Pagaminia natūriniais vienetais ataskaitinį mėnesį	2 iš viso	3 iš to skaičiaus Lietuvos rinkai
A <input type="text"/>	<input type="text"/>	<input type="text"/>
A <input type="text"/>	<input type="text"/>	<input type="text"/>
A <input type="text"/>	<input type="text"/>	<input type="text"/>
A <input type="text"/>	<input type="text"/>	<input type="text"/>
A <input type="text"/>	<input type="text"/>	<input type="text"/>
A <input type="text"/>	<input type="text"/>	<input type="text"/>

Prasome nurodyti, kiek laiko sugaišote statistiniams duomenims rengti ir ataskaitai pildyti: val. min. Pildoma ataskaitoje už balandžio mėn.


Įmonės vadovas arba jo įgaliotas asmuo

(Parašas) (Vardas ir pavardė)

(Ataskaitą užpildžiusio asmens vardas ir pavardė) (Telefonas) (Faksas)

(El. paštas)

Pasiteirauti:
 Statistikos departamente: tel.: (8 5) 236 4700, 236 4709, 236 4709, 236 4740, 236 4797
 Vilniaus TSV (Vilniaus apskrities respondentams): tel. (8 5) 272 4235; faks. (8 5) 273 0264; el. p. VilniausTSV@stat.gov.lt
 Kauno TSV (Kauno, Marijampolės, Alytaus apskričių respondentams): tel. (8 37) 31 2522; faks. (8 37) 31 2593; el. p. KaunoTSV@stat.gov.lt
 Klaipėdos TSV (Klaipėdos, Tauragės apskričių respondentams): tel. (8 46) 34 6270; faks.: (8 46) 34 0812, 34 0850; el. p. KlaipėdosTSV@stat.gov.lt
 Panevėžio TSV (Panevėžio, Utenos apskričių respondentams): tel. (8 45) 52 3441; faks. (8 45) 46 0248; el. p. PanevezioTSV@stat.gov.lt
 Šiaulių TSV (Šiaulių, Telsių apskričių respondentams): tel. (8 41) 52 5038; faks. (8 41) 52 5010; el. p. SiauliuTSV@stat.gov.lt


5 974 185 55225 6

I atranką patekusioms įmonėms klausimynų formos siunčiamos metų pradžioje. Septintą kalendorinę dieną, ataskaitiniam laikotarpiui pasibaigus, respondentai užpildytus mėnesinio periodiškumo statistinius klausimynus (P-12) pateikia faksu / paštu teritorinėms statistikos valdyboms (TSV) arba elektroniniu būdu.

Įmonės pramoninių gaminių gamybos klausimynas (P-12) turi dvi dalis.

I dalis: Gaminių gamyba iš nuosavų medžiagų ir žaliavų, pusfabrikačių. Šioje dalyje duomenis pateikia įmonės, gaminančios produkciją iš nuosavų medžiagų ir žaliavų, pusfabrikačių – visas medžiagas įmonė nusipirko pati iš savo apyvartinių lėšų; įmonė, turinti medžiagas ir žaliavas, pusfabrikačius, tačiau pati negaminanti produkcijos, o turimas medžiagas ir žaliavas, pusfabrikačius perduodanti produkcijos gamybai sutartiniu pagrindu Lietuvos teritorijoje esančioms įmonėms. Tokio pobūdžio įmonės ataskaitų I dalyje nurodo tik parduotos produkcijos apimtį natūriniais vienetais ir vertę.

II dalis: Gaminių gamyba iš užsakovų medžiagų bei žaliavų, pusfabrikačių. Įmonė, gaminanti produkciją tik iš užsakovų medžiagų bei žaliavų, pusfabrikačių, pildo tik II dalį, nurodydama visą pagamintos produkcijos apimtį natūrine išraiška ir / ar atliktų pramoninių darbų vertę.

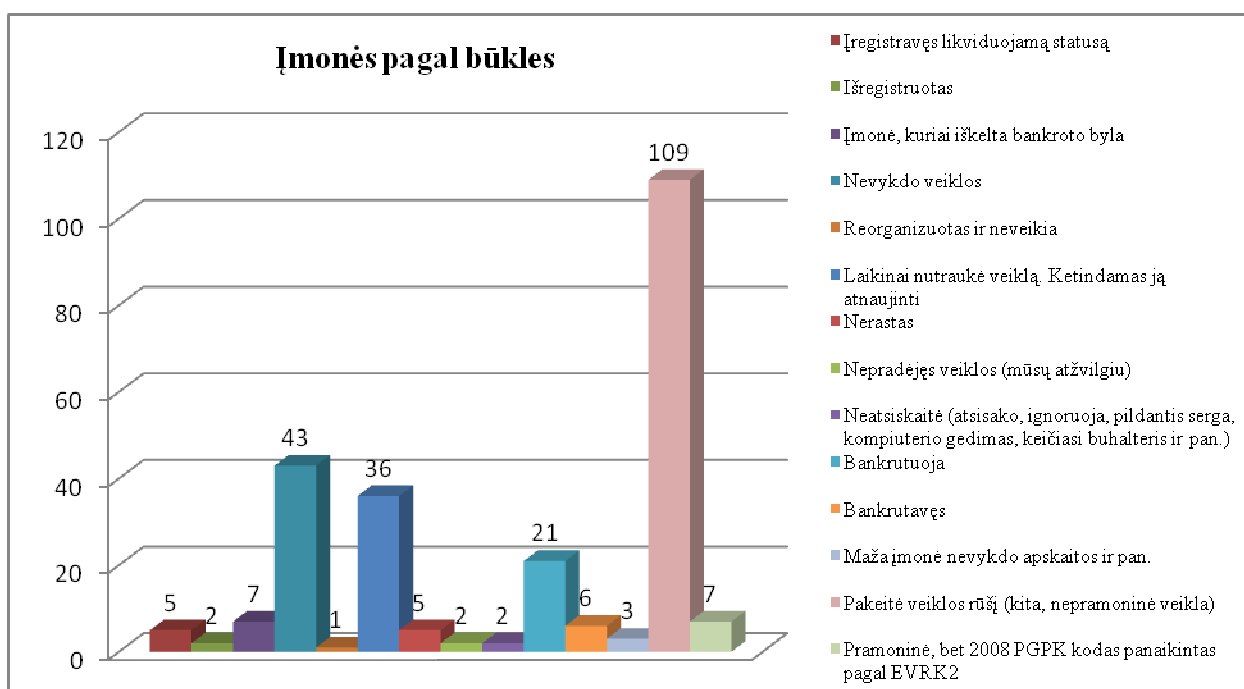
Įmonės, gaminančios produkciją tiek iš savo turimų medžiagų ir žaliavų, pusfabrikačių, tiek iš užsakovo medžiagų ir žaliavų, pusfabrikačių (dalį medžiagų įmonė nusipirko pati, o už

užsakovo medžiagas ir žaliavas, pusfabrikačius įmonė nemoka) I dalyje nurodo tik iš savo turimų medžiagų, žaliavų, pusfabrikačių pagamintos produkcijos apimtį ir paduotos produkcijos apimtį bei vertę tik tos produkcijos dalies, kuri pagaminta iš įmonės nuosavų žaliavų. II dalyje šios įmonės nurodo tik iš užsakovo žaliavų pagamintos produkcijos apimtį natūriniais vienetais ir atliktų darbų vertę (Įmonės pramoninių gaminių gamybos tyrimo metodika).

1.5 Statistinių duomenų pirminis redagavimas

Ataskaitos surenkamos, tikrinamos ir pirminiai statistiniai duomenys į duomenų bazę yra įvedami TSV darbuotojų, kurie ir atlieka tą pirminę vizualinę ir loginę kontrolę. Statistinių duomenų įvedimo programoje (Oracl'o aplinkoje) jau yra įdiegtos loginių klaidų kontrolės, tikrinamos ikiataskaitinio ir ataskaitinio laikotarpio kainos pagal gaminius. Programoje numatytas identifikacinis langelis, kuriame užrašomas ataskaitą įvedusio asmens prisijungimo vardas, bei pastaboms skirtas langelis. TSV darbuotojai taip pat tikrina savo apskrities atsiskaitymo pilnumą, lygindami su atrinktų įmonių sąrašū nustato ir užfiksuoja įmonės būklę. Pastarosios fiksuojamos vidiniame P-12 registre ne tik dėl vidinės tvarkos ir apskaitos. Toks būklių fiksavimas padeda atnaujinti ir patikslinti Statistinio ūkių subjektų registro duomenis. Tokią funkciją nuolat atlieka Pramonės statistikos skyriaus darbuotojai.

Jau per du mėnesius nuo metų pradžios užfiksuota 249 įmonių būklių pasikeitimų.



1 pav. Įmonių pasikeitusių būklių suvestinė

Statistinių duomenų įvedimo programoje tikrinamos tokios klaidos:

- Jeigu gaminio kodas B skiltyje suvestas neteisingai, tai suvedamo gaminio pavadinimas ir matavimo vienetai (A ir C skiltyse atitinkamai) neatsiras, t.y. pavadinimas „iššoka“ automatiškai, jei suvedamas teisingas 10 skaitmenų gaminio kodas;
- Pagrindinė agregavimo lygio taisyklė – A lygio kodas turi būti lygus B lygio kodų sumai. Tai tikrinama visoms penkioms ataskaitos skiltims (gamybai, pardavimams, pardavimams vidaus rinkoje, pinigams ir pinigams vidaus rinkoje);
- Jeigu pardavimai viso lygūs pardavimams vidaus rinkoje, tai ir pinigai viso turi būti lygūs pinigams vidaus rinkoje, t.y. jeigu skiltys $2=3 \rightarrow 4=5$;
- Jeigu pardavimai yra, tai turi būti užpildytos ir skiltys, reikalaujančios vertinės pardavimų išraiškos (tūkstančiais litų), t.y. jei skiltys $2 \neq 0$ ir / arba $3 \neq 0 \rightarrow 4 \neq 0$ ir / arba $5 \neq 0$;
- Analogiškai, jeigu pinigai yra, tai turi būti užpildytos su pardavimais susijusios skiltys, t.y. jei $4 \neq 0$ ir / arba $5 \neq 0 \rightarrow 2 \neq 0$ ir / arba $3 \neq 0$;
- Jeigu pardavimai iš viso = 0, tai ir pardavimai vidaus rinkoje turi būti = 0;
- Analogiškai su pinigais, jeigu pinigai iš viso = 0, tai ir pinigai vidaus rinkoje = 0;
- Jeigu gaminio kodo matavimo vienetai yra „-“, tai pardavimai nurodomi tik vertine išraiška, t.y. skiltys $1=0$, $2=0$ ir $3=0$.

Gaminio:
A – pavadinimas;
B – kodas;
C – matavimo vienetai.

Ikiataskaitinio mėnesio gamyba

„A“ lygio kodas

„B“ lygio kodas

1 Gaminų gamyba iš nuosavų medžiagų ir žaliavų, pusfabrikacijų										
A	B	C	1	2	3	4	5	Vidutinės kainos VISO		
								atask.	ikiatask.	
Audinių iš sintetinių plūšų arba iš si	1330195000	-				225	213	0.00	0.00	
Antklodės ir kelioniniai priedai iš sintefini	1392115000	vnt.	281	763	293	293	6.	10.	27.30	28.76
Lovos skalbiniai iš medvilnės (išskyrus	1392125300	kg	38	279	129	128	3.	3.	23.26	27.94
Lovos skalbiniai iš austinį tekstilės me	1392125900	kg	224	200	224		6.		26.79	31.00
Patalpų dekoravimui skirti dirbiniai, įskai	1392166000	-					6.		0.00	0.00
Patalmės renkamos su kitais užpildais	3103129000	vnt.	14595	21954	16444	16444	163.	163.	9.91	11.77
Antklodės, kitokios nei su pūky ar plauk	1392249910	vnt.	265	1252	548	548	29	29	52.92	72.07
Pagalvės, kitokios nei su pūky ar plauk	1392249920	vnt.	14330	20702	15896	15896	134.	134.	18.43	9.20
Ciuziniai (išskyrus spyruoklinius, ar sklyto	3103129000	vnt.	53	501	147	147	37.	37.	251.70	48.60
								VIDAUS atask. ikiatask.		
								251.70	48.60	
								EKSPORTO atask. ikiatask.		
								0.00	0.00	

Klaida

Klaida

Klaida

Gaminio 3103129000 ataskaitinio ir ikiataskaitinio mėnesių kainų lyginimas (vidaus rinkoje) - klaida

Gaminio ataskaitinio ir ikiataskaitinio mėnesių vidutinių kainų lyginimas

1 – gamyba;
2 – pardavimai viso;
3 – pardavimai vid. rinkoje;
4 – pinigai viso;
5 – pinigai vidaus rinkoje

2 pav. Įmonės pramoninių gaminių gamybos mėnesinės ataskaitos statistinių duomenų įvedimo programos langas.

Duomenys, užpildyti elektroniniu būdu (elektroninė ataskaita pateikiama 1.4 skyrelyje), tiesiogiai patenka į tyrimo duomenų bazę. Iš pirmo žvilgsnio atrodytų, kad elektroniniu būdu siunčiamos ataskaitos išvengia pirminės kontrolės, bet tai klaidinga – elektroninė ataskaitos forma sudaryta tokiu pat principu kaip ir statistinių duomenų įvedimo programa naudojama SD ir TSV darbuotojų, t.y. joje tikrinamas loginis ryšys tarp skilčių bei gaminio kodo tikslumo (kad gaminys būtų užkoduotas pagal tuo metu galiojantį PGPK klasifikatorių). Skirtumas tas, kad popierines ataskaitas tikrina Statistikos darbuotojai, o elektronines - patys respondentai suaktyvindami tikrinimo mygtuką. Jeigu formoje yra klaidų, pastaboms skirtoje skiltyje šalia geltonų vėliavėlių, nurodančių įvesto gaminio tikslų pavadinimą su matavimo vienetais, išmetamos raudonos vėliavėlės, nurodančios tiksliai klaidų vietas (skiltis). Ataskaita automatiškai išsaugojama duomenų bazėje tik tokiu atveju, jei joje nėra klaidų. Tokio statistinių duomenų perdavimo būdo trūkumai – sistemos trukdžiai: saugomos ataskaitos su klaidom ar nepriimamos jokios ataskaitos.

Taigi, TSV darbuotojai suvedinėdami duomenis neturėtų pasitikėti vien tik įvedimo programoje įdiegtomis kontrolėmis, nes šalia dažniausiai daromų agregavimo lygio klaidų („A“

lygio duomenys yra agreguoti, bet nėra lygūs „B“ lygio duomenų sumai; arba „B“ lygio kodo(-ų) išvis nėra) galima išskirti tokius atvejus, kurių įvedimo programa netikrina. Pavyzdžiui, pardavimai (pinigai) „iš viso“ neturi viršyti pardavimų (pinigų) vidaus rinkoje. Dėl pastebėtų klaidų (manomų klaidų) ar trūkstamų reikšmių bendraujama tiesiogiai su respondentais. Kiekvienas toks atvejis yra išsiaiškinamas. Tokiais atvejais dažniausiai skambinama telefonu, nors kartais naudojamas ir trūkstamos informacijos ekspertinis įvertinimas. Pastarasis atliekamas turint kitų administracinių šaltinių, tokiu kaip energetikos skyrius, užsienio prekyba, mėnesinis pramonės įmonės veiklos tyrimas (P-11) ir mokesčių inspekcija, informacija apie įmonę ir ankstesnių laikotarpių duomenis. Tada, atsižvelgiant į ataskaitinio mėnesio darbo dienas, trūkstami duomenys įvertinami pagal vidutines tos įmonės kainas ir gaminamų gaminių struktūrą. Duomenų tikrinimo metu dar lyginami ataskaitinio ir ikiataskaitinio laikotarpio gaminių gamybos apimtys natūrine išraiška, duomenų kodavimo pagal PGPK tikslumas. Kodavimas tai teisingas gaminio kodo parinkimas. Gaminius įmonės koduoja pagal nurodytą PGPK klasifikatorių. Jei įmonės negali parinkti gaminio kodo, tai atlikti padeda TSV ir SD pramonės skyriaus darbuotojai. Darbą palengvina eksportuojamų gaminių Kombinuotos nomenklatūros kodai, kurių pagalba gaminiai perkoduojami į Prodcom kodus.

Tolesnę kontrolę atlieka Statistikos departamento pramonės statistikos skyriaus darbuotojai: duomenų tikrinimas su skyriuje atliekamų kitų tyrimų duomenimis, su kitų skyrių informacija, duomenų įvertinimas veikusiomis ir palyginamosiomis kainomis Respublikos lygyje.

1.6 Statistinės informacijos pateikimas vartotojams

Visus statistinės informacijos vartotojus galima būtų suskirstyti į dvi pagrindines grupes – tai vidaus ir išorės vartotojai. Vidaus vartotojai – tai Statistikos departamento skyriai tokie kaip nacionalinių sąskaitų, energetikos, kainų, statybos ir inovacijų, suvestinės, įmonių bei žemės ūkio skyriai. Pagrindiniais išorės vartotojais yra ministerijos, bankai, ambasados, gamybinės ir prekybinės asociacijos, konsultacinės firmos, įmonės, studentai, verslininkai bei tarptautinės organizacijos.

Gaminių statistika skaičiuojama pagal įmones ir gaminius, pagal pramonės veiklos rūšis, pagal periodus, pagal teritorijas (rajonus, apskritis) ir viešai skelbiama:

- Metinis leidinys „Gaminių gamyba“,
- Metinis leidinys „Lietuvos statistikos metraštis“,
- Leidinys „Lietuvos ekonominė ir socialinė raida“,
- Leidinys „Lietuvos apskritys“,
- Leidinys „Pramonės darbo rezultatai“,

- Duomenų paruošimas JT „Industrial Commodity Statistics Yearbook“ (Jungtinių tautų metinis leidinys),
- Duomenų perdavimo formato (GESMES) paruošimas ir jų perdavimas,
- Pranešimai spaudai,
- Statistinė informacija SD skyriams ir išorės vartotojams pagal užklausas,
- Statistinė informacija duomenų bazei, t.y. statistikos departamento tinklapyje www.stat.gov.lt.

2. STATISTINIŲ DUOMENŲ KOKYBĖS KONTROLĖ

2.1 Imties formavimas

Panagrinėkime iš N elementų sudarytą baigtinę populiaciją¹ $U = [u_1, \dots, u_N]$ arba trumpiau $U = \{1, 2, \dots, N\}$. Imtį³ iš baigtinės populiacijos žymėsime $i = \{u_{i1}, u_{i2}, \dots, u_{in}\}$ arba trumpiau $\mathbf{i} = \{i_1, i_2, \dots, i_n\}$, čia i_1, i_2, \dots, i_n – imčiai priklausančių populiacijos U elementų eilės numeriai. Imties elementų skaičius n vadinamas imties dydžiu (Krapavickaitė, Plikusas, 2005).

Išskirtinis atsitiktinės (tikimybinės) imties bruožas – kiekvienos visumos (populiacijos), iš kurios imami objektai tyrimui, vienetas turi apibrėžtą galimybę (nebūtinai lygią atskiriems visumos vienetams) patekti į imtį (Martišius, Kėdaitis, 2004).

Tikimybinis ėmimas – tai imties išrinkimo būdas, kuris, tiesiogiai renkant elementus iš populiacijos, tenkina tokias sąlygas:

- Nusakomos visos galimos skirtingos imtys $\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_V \subset U$, kurias iš turimos populiacijos galima išrinkti taikant nagrinėjamą imčių išrinkimo procedūrą. Tai reiškia, kad galima tiksliai nusakyti, kurie populiacijos elementai priklauso imčiai \mathbf{i}_1 , kurie priklauso imčiai \mathbf{i}_2 ir t.t.

- Kiekvienai galimai imčiai \mathbf{i}_v priskiriama žinoma imties išrinkimo tikimybė $P(\mathbf{i}_v) = p(\mathbf{i}_v) > 0$, $v=1, 2, \dots, V$, taip, kad $\sum_{v=1}^V p(\mathbf{i}_v) = 1$.

- Kiekvienas populiacijos elementas priklauso bent vienai galimai imčiai, t.y. yra teigiama tikimybė, kad jis bus išrinktas.

- Imtis išrenkama taikant tokią atsitiktinę procedūrą, kai kiekviena iš galimų imčių \mathbf{i}_v išrenkama su nurodyta tikimybe $p(\mathbf{i}_v)$, $v=1, \dots, V$.

Imtis, kuri gaunama, taikant imties išrinkimo procedūrą, tenkinančią visus keturis minėtus reikalavimus, vadinama tikimybine imtimi. Tik praktikoje ji beveik niekada nerenkama išrašant visas galimas imtis \mathbf{i} ir priskiriant joms išrinkimo tikimybes $p(\mathbf{i})$. Tokių imčių būna paprastai labai daug ir tai ne visada įmanoma. Todėl dažniausiai imtys renkamos nuosekliai, vienas po kito išrenkant populiacijos elementus, kol gaunama reikiamo dydžio imtis (Krapavickaitė D., Plikusas A., 2005).

Taigi, šiame darbe nagrinėjama tikimybinė imtis, o tiksliau, ėmimas su nelygiomis tikimybėmis, kurios yra tiesiogiai proporcingos įmonės gaunamoms pajamoms. Baigtinė populiacija $U=8064$ (statistinio ūkio subjektų registro duomenys), o išrinktos imties dydis $n=3938$.

Pagrindinės sąvokos:

¹Baigtinė populiacija arba populiacija – baigtinė individų ar objektų aibė.

²Populiacijos elementai – populiaciją sudarantys objektai.

³Imtis – bet koks populiacijos elementų rinkinys.

Norėdami apskaičiuoti k-tojo populiacijos elemento išrinkimo tikimybę $p_k = \frac{z_k}{\sum_{i=1}^N z_i}$,

kai $k=1,2, \dots, N$ ir $\sum_{k=1}^N p_k = 1$, pasinaudojame papildomu kintamuoju su teigiamomis reikšmėmis.

Mūsų atveju z_k yra k-tosios įmonės pajamos, o $\sum_{i=1}^N z_i$ - visų įmonių pajamos.

Iš imties duomenų gauto parametro įverčio reikšmė visada daugiau ar mažiau skiriasi nuo tikrosios parametro reikšmės. Jeigu pažymėtume ε paklaidos dydį, α parametro reikšmė ir a iš imties duomenų apskaičiuoto įverčio reikšmė, tai skirtumas formaliai gali būti užrašytas taip: $\alpha = a + \varepsilon$, kur $\varepsilon = \varepsilon_S + \varepsilon_A$, čia ε_S - sisteminė klaida, o ε_A - atsitiktinė klaida (S.A.Martišius, 2004).

2.2 Paklaidos

Statistiniai įverčiai gali būti netikslūs dėl daugybės priežasčių, kurios ne visada būna tiesiogiai susijusios su pasirinktu ėmimo planu ar taikomo įvertinio kokybe. Kaip aiškina imčių metodų specialistai, dažnai paklaidų šaltiniais tampa ėmimo sąrašo klaidos, neatsakymai į apklausas, matavimo klaidos bei technines, arba duomenų apdorojimo, klaidos.

Imčiai išrinkti dažnai naudojami ėmimo sąrašai – tiriamos populiacijos elementų sąrašai. Ėmimo sąrašas gali būti netikslus: jis gali būti nepilnas arba (ir) perteklinis, t.y. jame gali trukti kai kurių tiriamos populiacijos elementų arba gali būti tokių elementų, kurie tiriamai populiacijai nepriklauso.

Neatsakymų priežastys būna įvairios. Į imtį išrinktas populiacijos elementas gali būti nepasiekiamas dėl ėmimo sąrašo klaidų arba jis paprasčiausiai gali nenorėti (arba negalėti) dalyvauti apklausoje. Dar visai tikėtina, kad užpildytas klausimynas gali pasimesti (ar pamesti klausėjas) arba tiesiog jis gali būti netinkamas tolesnei analizei dėl jame esančių klaidų.

Klaidos, kurias padaro klausėjas ar atsakytojas, kurios atsiranda dėl klausimyno trūkumų ar apklausos vykdymo būdo (apklausa telefonu, paštu, internetu ar pasitelkiant klausėjus), vadinamos matavimo klaidomis.

Technines, arba duomenų apdorojimo, klaidos – tai duomenų įvedimo, kodavimo, redagavimo klaidos.

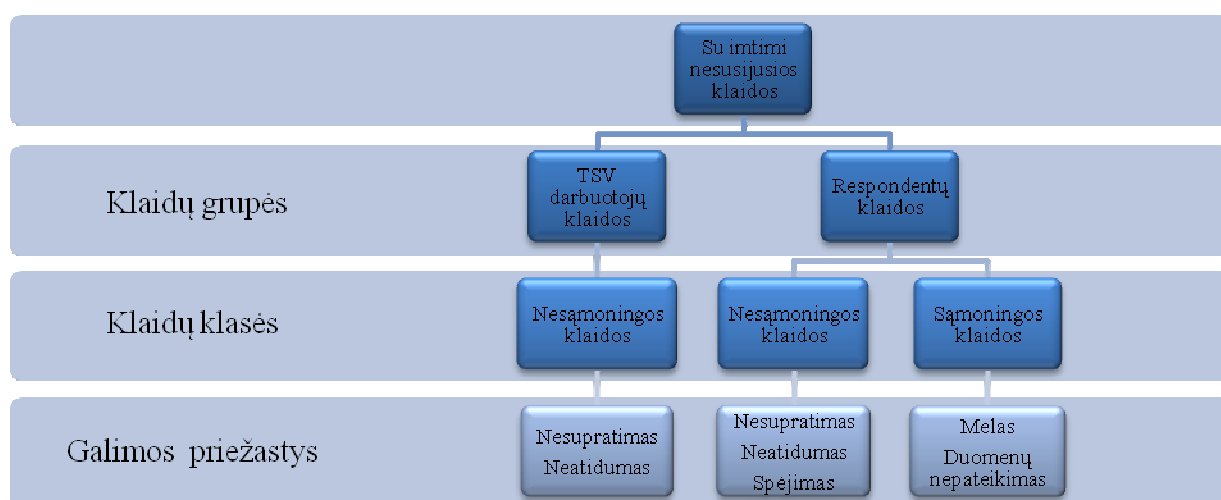
Visos tyrimo metu atsirandančios paklaidos nulemia skelbiamų duomenų kokybę – turi įtaką įverčio tikslumui. Todėl visų čia minimų paklaidų stengiamasi išvengti arba bent jau sumažinti jų skaičių atidžiau planuojant ir atliekant tyrimą.

Paklaidas pagal jų kilmės šaltinius galima suskirstyti į tris dideles grupes:

- a) *Ėmimo paklaidos*. Jų atsiradimo priežastis pats atsitiktinio ėmimo metodas, nes, jį taikant, stebimos tik imties elementų kintamojo reikšmės, o išvados daromos apie visą populiaciją.
- b) *Su ėmimu nesusijusios*, kitų tyrimo sudėtinių dalių nulemtos paklaidos. Jų šaltiniai: ėmimo sąrašas, klausimynas, klausėjas, klausiamasis, apklausos būdas, neatsakymai į apklausas.
- c) *Duomenų apdorojimo*, arba techninės klaidos. Tai duomenų įvedimo, kodavimo, redagavimo klaidos. Jos būdingos bet kokiam statistinio duomenų apdorojimo procesui (Krapavickaitė, Plikusas, 2005).

Dauguma su imtimi nesusijusių klaidų atsiranda duomenų rinkimo etape. Ir jeigu V.Kėdaitis (2003) savo metodinėse priemonėse išskyrė dvi pagrindines ne su imtimi susijusias klaidų grupes tokias kaip *apklausėjų* ir *respondentų* klaidos, tai mūsų atveju šios dvi grupės yra: respondentų padarytos klaidos ir TSV (teritorinė statistikos valdyba) darbuotojų duomenų įvedimo metu padarytos klaidos. Skirtumas tas, kad mūsų antrai klaidų grupei (TSV darbuotojų padarytų klaidų grupė) nebūdingos tyčia padarytos klaidos, sudarančios sąmoningų klaidų klasę, kurią kartu su nesąmoningai, ne savo noru, padarytomis klaidomis, V. Kėdaitis išskiria kiekvienam klaidų tipui.

Visos su imties formavimu nesusijusios klaidos gali būti skirstomos į sekančias grupes:



3 pav. Klaidų tipai

Prie sąmoningai darančių klaidas respondentų priskiriami tie, kurie dėl tam tikrų priežasčių pateikia neteisingą informaciją arba išvis jos nepateikia, t.y. ignoruoja bet kokius priminimus ir net perspėjimus dėl duomenų nepateikimo atveju grėšiančios administracinės nuobaudos. Tokio atsainaus elgesio priežastis gali būti noras apginti savo privatumą, kvalifikuotų darbuotojų trukumas, tokios informacijos reikalingumo nesuvokimas arba tiesiog požiūris į pačią duomenų rinkimo įstaigą - Statistikos departamentą. (pvz. atsisako teikti informaciją nurodytais matavimo vienetais, nes patys apskaitą veda visiškai kitais ir perskaičiuoti atsisako)

Nesąmoningos respondentų klaidos atsiranda tada kai atsakymas nėra teisingas, bet jis tiki, kad „sako“ tiesą. Respondentas pildo ataskaitą nesuprasdamas ir/arba nesilaikydamas instrukcijų – pateiktų aprašų, arba kai jis nėra tikras dėl savo atsakymo teisingumo ir bando spėti (pats populiariausias nesąmoningos klaidos pavyzdys yra, kai buhalteris neteisingai sukoduoja savo įmonės produkciją). Ir žinoma tokios kaip gaminio kodo neįrašymo klaidos padaromos per neatidumą ir skubėjimą.

Kadangi visiškai išvengti klaidų yra neįmanoma, todėl įvedamų kontrolių tikslas yra bent jau sumažinti jų kiekį ir išvados daromą įtaką. Pavyzdžiui, nesąmoningų, su tyrimu susijusių asmenų padaromų klaidų kiekį sumažinti galima būtų atidžiau renkantis darbuotojus ir, žinoma, organizuojant apmokymus, susitikimus, kurių metu būtų pristatytas pats tyrimas, jo tikslas, paaiškintos ataskaitos pildymo taisyklės, užakcentuotos sudėtingesnės (problemiškiausios) vietos. Tokių atsargumo priemonių imtasi ir Statistikos departamente. Kiekvienais metais sausio mėnesį organizuojami TSV ir pramonės statistikos skyriaus darbuotojų susitikimai, kurių metu aptariami praeitų metų įvykiai, primenamos ataskaitos pildymo, duomenų suvedimo, tikrinimo ir dokumentavimo taisyklės, o taip pat informuojama apie galimus ataskaitos arba klasifikatoriaus ir su tuo susijusių pasikeitimų bei apmokama naudotis atnaujinta ir patobulinta duomenų įvedimo programine įranga.

2.3 Duomenų apdorojimas

Duomenų apdorojimas apima nuo duomenų rinkimo iki galutinio rezultatų pateikimo etapus: informacijos rinkimas; duomenų kodavimas; duomenų įvedimas į kompiuterį; kontrolinis duomenų tikrinimas įvedimo metu; mikroredagavimas; pakartotinė atsakytojų apklausa, radus duomenyse įtartinų vietų; praleistų reikšmių įrašymas; makroredagavimas; galutinių rezultatų pateikimas.

Duomenų apdorojimo klaidos – tai duomenų kodavimo bei įvertinimo klaidos, kurios atsiranda ir dėl nekokybiško klausimyno, kai neteisingai suprantamas klausimas, ir dėl to, kad klausėjas galbūt turėjo įtakos tiriamojo atsakymui arba to atsakymo nesuprato. Ir kaip taisyklė, tokių klaidų visada būna.

Atsiradusios klaidos duomenyse taisomos visuose duomenų apdorojimo etapuose tarp jų įvedimo ir galutinių rezultatų pateikimo. O procedūrų, padedančių nustatyti ir ištaisyti atsiradusias klaidas, visuma vadinama duomenų redagavimu.

Mikroredagavimas - tai klaidų paieška, neatsižvelgiant į tai, kokią įtaką jos turi rezultatams. Jos metu nustatomas atskirų reikšmių suderinamumas ir ryšių tarp atskirų parametrų logiškumas. Tikrinama, ar kiekvienas duomenų įrašas tenkina redagavimo taisyklių sąlygas, kurios

sudaromos remiantis klausimynu, duomenų analize bei ekspertų nuomone. Redagavimo taisyklėmis nustatomos galimos kintamųjų reikšmių ribos ir tikrinama, ar turimos reikšmės neišeina už jų (Krapavickaitė, Plikusas, 2005).

Redagavimo procesas turi būti paprastas, pigus ir greitas, o tai reiškia, kad jis turėtų būti kiek galima daugiau automatizuotas. Redaguojant duomenis negalima persistengti, nes turi būti išsaugota kaip galima daugiau nepakeistų reikšmių. Redagavimą galima būtų suskirstyti į būtiną ir pasirinktinį. Būtiną redagavimo metu nustatomos neabejotinos klaidos, negalimos, praleistos ar nesuderinamos reikšmės. Pasirinkto redagavimo metu nustatomos galimos klaidos – reikšmės, kurios netenkina tam tikrų reikalavimų: už nustatytų ribų, reikšmingai didesnės ar mažesnės, nei kitos to rodiklio reikšmės. Neabejotinos klaidos privalo būti taisomos. Galimas klaidas sunkiau ištaisyti, jų įtaka rezultatams mažesnę nei neabejotinų ir jų ištaisymas brangesnis (Recommended Practices for Editing and Imputation in Cross-Sectional Business Surveys, 2007).

Redagavimo metu nustatomos sisteminės, reikšmingos, atsitiktinės klaidos, išsiskiriančios bei praleistos reikšmės.

Praleistas reikšmes galima būtų suskirstyti į praleistas atskiro įrašo reikšmes ir viso klausimyno nepateikimas. Pastarosios Gaminių gamybos tyrime išgaudamos lyginant su pradiniu gamybos tyrimo atrankos sąrašu bei su TSV darbuotojų pagalba. Tuo tarpu praleistų atskirų reikšmių identifikavimas įeina į loginių klaidų tikrinimo programas.

Sisteminės klaidos yra pačios pavojingiausios, nes turi reikšmingą poveikį galutiniams rezultatams, bet ir sunkiausiai nustatomos. Geriausias tokios klaidos pavyzdys yra neteisingas gaminių kodavimas. Jeigu neteisingai sukoduoto gaminio neužkliūva nei gamybos apimtys, nei vidutinė vieneto kaina, tai tokia klaida atrandama tik atsitiktinumo dėka.

Klaidos, kai vienas ar keli klaidingi stebėjimai gali nulemti bendrą įvertį ir jo tikslumą, vadinamos reikšmingomis (esminėmis) klaidomis. Norint tokias klaidas nustatyti reikia žinoti galimų reikšmių intervalus. Pramonės statistikoje tokios klaidos pavyzdžiu galėtų būti didelės įmonės, savo keturženklėje veikloje sudarančioje vos ne 50% visos gamybos, įvelta klaida. Kai tarkim vietoj reikalaujamų kg, ji gamybą pateikia vienetais.

Išskirtinės reikšmės – tai reikšmės, kurios netenkina kintamajam taikomo modelio. Jos kaip ir reikšmingos klaidos turi būti nustatytos, įvertintos ir koreguotos norint išvengti įverčių poslinkio ar didelės variacijos.

Atsitiktinės klaidos atsiranda dėl nenumatytų priežasčių. O jų vidurkis artimas vidutinei kintamojo reikšmei (Recommended Practices for Editing and Imputation in Cross-Sectional Business Surveys, 2007).

2.4 Duomenų aprašymas

Kaip jau buvo anksčiau minėta, duomenys (du duomenų rinkiniai už 2009 metų vasario mėnesį) gauti iš Statistikos departamento Pramonės statistikos skyriuje atliekamo „Įmonės pramoninių gaminių gamybos tyrimo P-12“. Paimti vieno mėnesio duomenys: „pradinis“ duomenų rinkinys – statistiniai duomenys iki antrinio duomenų redagavimo atliekamo jau Statistikos departamento darbuotojų ir „galutinis“ duomenų rinkinys – duomenys po redagavimo. Kadangi ataskaitų „gavimo - įvedimo procesas“ nėra griežtai terminuotas, todėl vis naujai gaunamos ataskaitos suvedinėjamos (arba automatiškai įrašomos, jei tai yra elektroninės ataskaitos) į pagrindinę duomenų bazę tol kol tokios yra. Teoriškai ataskaitų siuntimo ir suvedimo procesas turėtų užtrukti ne ilgiau kaip nuo 7 iki 18 (blogiausiu atveju iki 20) mėnesio dienos. Ir apytiksliai jau 18 mėnesio dieną Pramonės statistikos skyriaus darbuotojai pilnu pajėgumu turėtų tikrinti gautus duomenis prieš juos paskelbiant „Lietuvos ekonominėje ir socialinėje raidoje“ 23-25 dienomis. Tačiau realybė yra truputi kitokia: įmonės vėluoja persiųsti ataskaitą, TSV darbuotojai nespėja laiku visko suvesti, o SD darbuotojams lieka vos kelios dienos duomenų antriniam redagavimui. Būtent dėl šios priežasties ir turi būti sukurtos modernios, daug laiko neužimančios ir nereikalaujančios „rankinio“ darbo klaidų tikrinimo programos.

Taigi, mūsų gautų dviejų duomenų masyvu dydžiai yra skirtingi, kovo 17 dienos – 8606 eilutės, o kovo 30 dienos – 10808 įrašai. Matome, kad skirtumas sudaro apie 20%, net 2202 įrašai. Jeigu 2008 metais vidutinis vienos ataskaitos pildomų gaminių skaičius buvo 4 gaminiai, tai gauname, kad laiku nepateikė ir nebuvo suvesta vidutiniškai 550 ataskaitų.

IM_KOD	IM_PAV	APSKR	SAV	KAS_IVEDE	DALIS	PROD_KOD	PROD_MTV	PROD_PAV	PROD_LYGIS	PAG_NATURA	PARD_VISO	PARD_VID	PINIGAI_VISO	PINIGAI_VID
1000001	-	3	55	El.forma	1	1610103200	m ²	Spygliuočių mediena, kurios storis didesnis kaip 6 mm, išilgai perpjauta arba perskelta, nudrožta arba be žievės, sujungta galais arba šlifauta		13	13	13	7,7	7,7
1000002	-	6	29	El.forma	2	2229916000	-	Plastinės dalys ir reikmenys visoms sausumos kelių transporto priemonėms (išskyrus lokomotyvus arba riedmenis)					52,5	52,5
1000003	-	10	79	TSV_VLN8	1	2511236000	kg	Kitos statybinės konstrukcijos iš geležies, ketaus arba plieno		700	700	700	4,2	4,2
1000004	-	3	21	El.forma	1	2511236000	kg	Kitos statybinės konstrukcijos iš geležies, ketaus arba plieno		1	1	1	16,4	16,4
1000005	-	8	78	TSV_SL7	1	3101120000	vnt.	Mediniai biuro baldai	A	35	35	35	15,3	15,3
1000005	-	8	78	TSV_SL7	1	3101120050	vnt.	Kiti mediniai įstaigos baldai, spintos ir lentynos, kurių aukštis didesnis kaip 80 cm	B	35	35	35	15,3	15,3
1000006	-	5	73	El.forma	1	1041260000	kg	Neapdorotas rapsų, rapsukų ar garstyčių aliejus ir jo frakcijos (išskyrus chemiškai modifikuota)		366000	2484450	2445690	913,4	891,4

4 pav. „Pradinis“ duomenų rinkinys MS Excel formatu

Čia IM_KOD - septinių ženklų įmonės kodas, IM_PAV - įmonės pavadinimas, APSKR – apskritis (nuo 1 iki 10), SAV – savivaldybė (nuo 11 iki 94), KAS_IVEDE – įvedusio darbuotojo vartotojo vardas, DALIS – ataskaitos dalis (1 arba 2), PROD_KOD – pildomo gaminio 10

ženklų kodas pagal tuo metu galiojančią PGPK, PROD_MTV – gaminio matavimo vienetai, PROD_PAV – gaminio tikslus pavadinimas, PROD_LYGIS – gaminio agregavimo lygis (A arba B), PAG_NATURA – ataskaitinio mėnesio pagamintos produkcijos apimtys natūriniais vienetais, PARD_VISO – ataskaitinio mėnesio parduotos iš viso produkcijos apimtys natūriniais vienetais, PARD_VID – ataskaitinio mėnesio parduotos vidaus rinkoje produkcijos apimtys natūriniais vienetais, PINIGAI_VISO – ataskaitinio mėnesio parduotos iš viso produkcijos apimtys tūkstančiais litų, PINIGAI_VID – ataskaitinio mėnesio parduotos vidaus rinkoje produkcijos apimtys tūkstančiais litų.

Klausimynus užpildo patys respondentai, kurie turi galimybę pasirinkti jiems patogiausią statistinių duomenų pateikimo būdą: užpildytus klausimynus siųsti paštu, faksu, elektroniniu paštu ar internetu. Šis duomenų surinkimo būdas dažniausiai naudojamas vykdant verslo statistinius tyrimus, kada respondentas yra įmonė.

IM_KO	IM_PAV	APSK	SAV	KAS_IVED	DALI	PROD_KO	PROD_MT	PROD_PAV	PROD_LYGI	PAG_NATUR	PARD_VISO	PARD_VI	PINIGAI_VISO	PINIGAI_VI
1000000	-	2	19	El.forma	2	3312199000	-	Kitų bendros paskirties mašinų, nenurodytų kitoje vietoje, remontas ir techninė priežiūra					13,4	13,4
1000001	-	3	55	El.forma	1	1610103200	m²	Spygliuočių mediena, kurios storis didesnis kaip 6 mm, išilgai perpjauta arba perskelta, sudrožta arba be žievės, sujungta galais arba šlifauta		13	13	13	7,7	7,7
1000002	-	6	29	El.forma	2	2229916000	-	Plastikinės dalys ir reikmenys visoms sausumos kelių transporto priemonėms (išskyrus lokomotyvus arba riedmenis)					52,5	52,5
1000004	-	3	21	El.forma	1	2511236000	kg	Kitos statybinės konstrukcijos iš geležies, kietaus arba plieno		1000	1000	1000	16,4	16,4
1000008	-	10	13	TSV_VLNS	1	3101120000	vnt.	Mediniai biuro baldai	A	12	12	12	6,7	6,7
1000008	-	10	13	TSV_VLNS	1	3101120010	vnt.	Mediniai įstaugos baldai, kiti stalai, kurių aukštis ne didesnis kaip 80 cm	B	9	9	9	,4	,4
1000008	-	10	13	TSV_VLNS	1	3101120040	vnt.	Kiti mediniai įstaugos baldai, spintos ir lentynos, kurių aukštis ne didesnis kaip 80 cm	B	2	2	2	2,4	2,4
1000008	-	10	13	TSV_VLNS	1	3101120050	vnt.	Kiti mediniai įstaugos baldai, spintos ir lentynos, kurių aukštis didesnis kaip 80 cm	B	1	1	1	3,9	3,9

5 pav. „Galutinis“ duomenų rinkinys MS Excel formatu

Pateiktuose duomenų fragmentuose yra išskirtos eilutės – tai klaidos pavyzdys. Įmonės siunčiama elektroninė ataskaita iškart pateko į duomenų bazę. Klaida (neteisingai suvestos gamybos ir pardavimų apimtys) pastebėta antrinio redagavimo metu lyginant gaminių kainas. Ši klaida taip pat būtų pastebėta lyginant 3 paskutinių mėnesių gamybą pagal įmonės ir gaminių kodus.

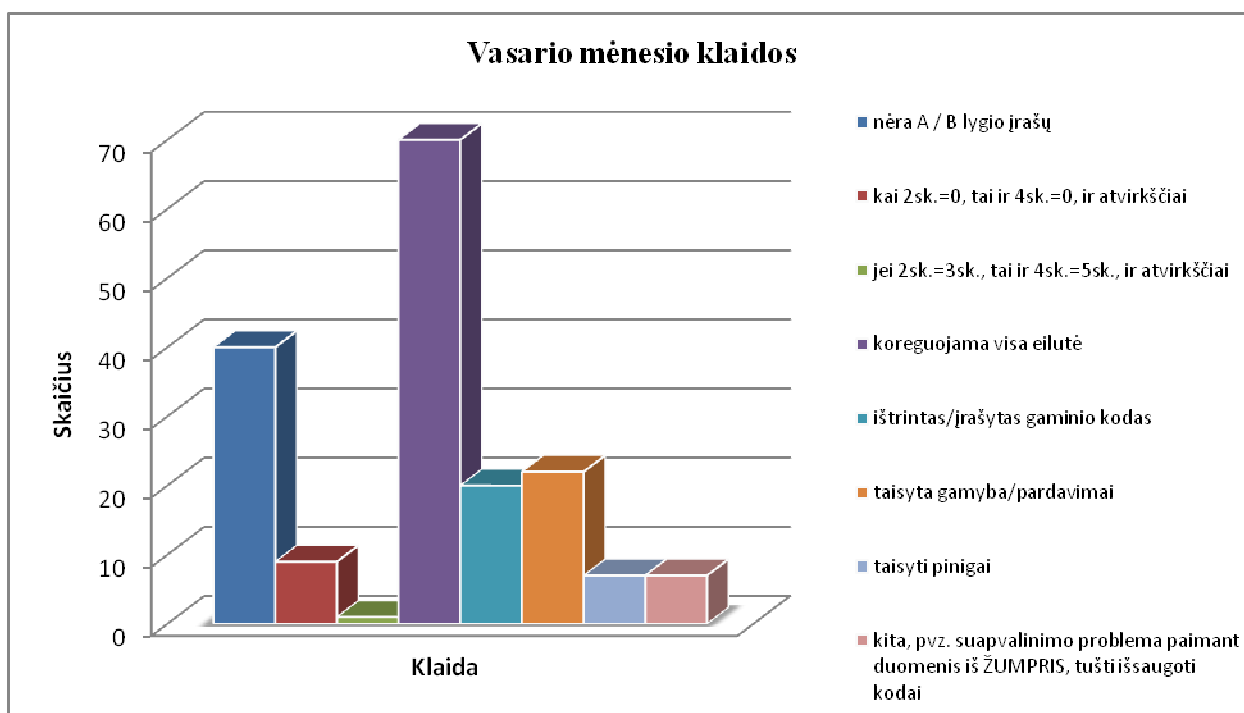
R.	IM_KOD	PAY	IM_VEIKLA	PROD_KOD	PROD_PAV	MTV	DALIS	PROD_LYGIS	METAI	MENUSO	PAG_NATURA	PAG_SAY_REIK	PARD_VISO	PARD_VID	PARD_EXP	TLT_VISO	TLT_VID	TLT_EXP	VNT_K	VNT_K_VID
1	1000004	UAB "	301100	2511236000	Kitos stabybinės konstrukcijos	kg	1	2009	1		1029 0	1029	1029	0	33,7	33,7	0	32,75	32,75	
2	1000004	UAB "	301100	2511236000	Kitos stabybinės konstrukcijos	kg	1	2009	2		1 0	1	1	0	16,4	16,4	0	16400	16400	
3	1000004	UAB "	301100	2511236000	Kitos stabybinės konstrukcijos	kg	1	2009	3		4765 0	4765	4765	0	23,7	23,7	0	4,97	4,97	
4	1000004	UAB "	301100	2599293100	Kopėdės ir laiptėliai iš geležies vnt.		1	2009	1		3 0	3	3	0	5,9	5,9	0	1966,67	1966,67	

6 pav. Įmonės (su fiktyviu įmonės kodu 1000004) trijų mėnesių duomenų palyginimas

Dažniausiai pasitaikančios klaidos yra:

- Pinigai suvesti litais, o ne tūkstančiais litų, kaip reikalaujama ataskaitoje;
- Sunkiau randamos klaidos – tai blogo kodavimo klaidos, kai neteisingai sukoduojami gaminiai;
- Žymus šuolis gamyboje (analogiškai ir su pardavimais), t.y. lyginant gamybą ar pardavimus su praeitais mėnesiais pastebimas vienas nereikalingas įsivėlęs skaitmuo;
- Prirašyti nereikalingi nuliai;
- Supainiotos dalys (gamyba iš nuosavų ar užsakovų žaliavų);
- Supainioti kodai, ypač B lygio kodai, kai jų yra keli. Kartais juos paprasčiausiai sukeičia vietomis.

Kadangi statistinių duomenų įvedimo programoje numatyta tik klaidų skaičiaus fiksavimo (nepriklausomai nuo klaidos tipo) sistema, todėl vasario mėnesio duomenis man teko stebėti papildomai (visi kiti mėnesiai analogiškai). Man pavyko išskirti tokias pagrindines klaidų, tiksliau koregavimų, rūšis.

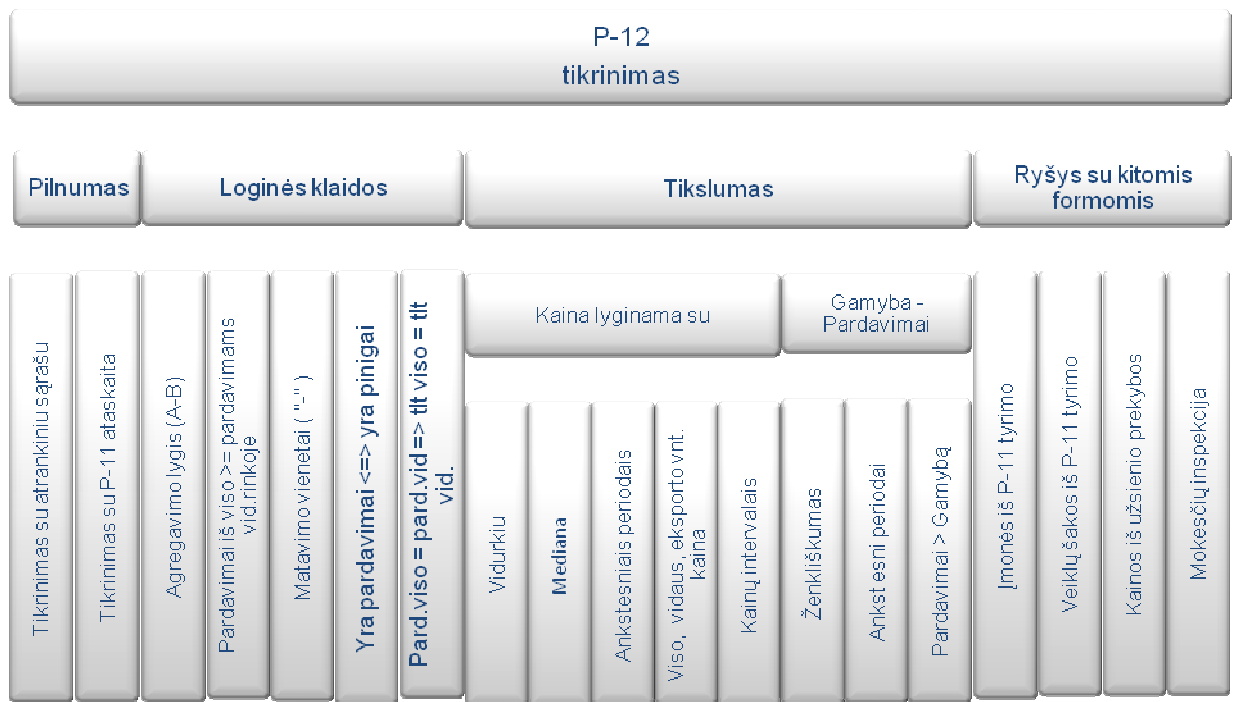


7 pav. Vasario mėnesį atlikti koregavimai

Iš šių 176 atrastų klaidų, 140 jų buvo padarytos respondentų, 29 klaidos buvo jų pačių patikslinami duomenys, o 13 klaidų įvėlė operatorės. Beveik 88% klaidų buvo rasta elektroninėse ataskaitose. O juk elektroniniu būdu dabar atsiskaito apie 33% ir tikėtina, kad tas skaičius vis augs. Todėl tai dar kartą patvirtina loginių klaidų paieškos automatizavimo poreikį.

2.5 Statistinių duomenų antrinis redagavimas (siūlomas modelis)

Siekiant sumažinti klaidų skaičių užpildytuose ataskaitose, pramonės statistikos skyriaus specialistai atlieka duomenų mikroredagavimą pagal parodytą schemą:



8 pav. P-12 formos duomenų tikrinimo schema

Remiantis šita schema buvo sudarytos tokios tikrinimo programėlės. Kiekvienos dalies rezultatas išmetamas atskiroje lentelėje, kurias sujungus galima gauti pilną loginių klaidų sąrašą.

Visos programos buvo bandomos su pradiniu duomenų rinkiniu, o vėliau patikrinta ar galutiniame tos klaidos buvo pastebėtos ir ištaisytos.

Dar kartą norėtuši priminti kodų agregavimo lygio taisyklę: „A“ lygio kodas lygus „B“ lygio kodų sumai. Ši tikrinimo programa sudaryta tokiu principu: aštuonių ženklų lygmenyje lyginamos atskiro kodo „A“ ir „B“ lygio gamybos, pardavimų ir pinigų suma.

```
I. /*A-B lygio tikrinimas */
```

```
proc sql;
create table mag.AB_lygis as
select * from(
    select distinct(prod_kod8)as prod_kod8, dalis, im_kod,
    sum(lygis_a)as lygis_A, sum(lygis_b)as lygis_B
    from(
        select substr(prod_kod,1,8)as prod_kod8, prod_mtv,
prod_lygis, dalis, im_kod,
        (pag_natura+pard_vis0+pard_vid+lt_vis0+lt_vid)as
lygis_A, 0 as lygis_B
        from mag.pirmas2
        where prod_lygis='A'
    union
        select substr(prod_kod,1,8)as prod_kod8, prod_mtv,
prod_lygis, dalis, im_kod, 0 as lygis_A,
        (pag_natura+pard_vis0+pard_vid+lt_vis0+lt_vid)as lygis_B
        from mag.pirmas2
        where prod_lygis='B'
    )group prod_kod8, dalis, im_kod
) where (lygis_A < lygis_B or lygis_A > lygis_B) and prod_kod8 not in (
```

```

/*išmetam kodus, kurių matavimo vienetai skirtingi*/
select distinct(prod_kod8) from(

        select distinct(substr(prod_kod,1,8))as prod_kod8, prod_mtv
        from mag.pirmas2
        where prod_lygis in('A','B')
)group prod_kod8 having count(distinct(prod_mtv))>1

)
;quit;

```

Rezultatas – įmonės su aštuonženkliais gaminių kodais, kurie netenkina pagrindinės kodų agregavimo lygio taisyklės.

	prod_kod8	DALIS	IM_KOD	lygis_A	lygis_B
1	13922499	1	1505089	0	3464
2	13922499	1	4509851	7000	0
3	13922499	1	5529267	3079	9834
4	13922499	1	8453695	943002	940870
5	31091230	1	6672778	4006	2003

9 pav. Tikrinimo programos rezultatas

Matome, kad tikrinant pradinį duomenų rinkinį, išmetamos 5 klaidos. O iš turimos informacijos galimos klaidos vietą galime nusakyti eilutės tikslumu. Apie pirmas dvi šio sąrašo klaidas jau iškart galima pasakyti, kad ataskaitos atsiųstos be vieno agreguotų kodų, t.y. pirmu atveju neįrašytas bendresnis „A“ lygio kodas, antru – „B“ lygio.

Programa išmeta žymiai mažiau klaidų, nei anksčiau buvo parodyta diagramoje, dėl dviejų priežasčių: pirmoji – ji neaptinka eilučių, kurios buvo ištaisytos TSV darbuotojų duomenų įvedimo metu (o sistema tai buvo užfiksavusi) ir ji neima tuščių nereikalingų eilučių (kai duomenų nėra, bet kodas ataskaitoje įvestas). Tuo tarpu sistemoje tai fiksuojama kaip klaida.

Taigi, iš 5 rastų klaidų antrinio redagavimo metu buvo aptikta ir ištaisyta tik viena iš jų, o tai reiškia rastos papildomos 4 klaidos.

Ieškant tokio pobūdžio klaidų, būtinai reikia išskirti tokius kodus, kurių skirtinguose agregavimo lygiuose nesutampa matavimo vienetai. Į juos nereikėtų kreipti dėmesio, ir geriausia išmesti iš sąrašo. Kas ir buvo padaryta prieš tai buvusioje programoje. Juk eilutės suma, pagal ką ir yra tikrinama, bus ne vienoda.

1 lentelė Gaminių kodo su skirtingais matavimo vienetais pavyzdys

14.20.10.30.00	Kailiniai drabužiai ir drabužių priedai (išskyrus skrybėles ir galvos apdangalus)	-	A
14.20.10.30.10	Paltai ir puspalčiai iš natūralaus kailio	vnt.	B
14.20.10.30.20	Apykaklės iš natūralaus kailio	vnt.	B
14.20.10.30.80	Kiti kailiniai drabužiai ir drabužių priedai	vnt.	B

```

/*Ši dalis išmeta aštuonženklus kodus ir skirtingus matavimo vienetus.
Taigi, jeigu vienas kodas pasikartoja kelis kartus, tai reiškia, kad tam
pačiam kodui priklauso keli skirtingi matavimo vienetai*/

```

```

proc sql;
create table mag.AB_mtv as
select distinct(prod_kod8)/*,count(distinct(prod_mtv))as mtv*/
from(
        select distinct(substr(prod_kod,1,8))as prod_kod8, prod_mtv
        from mag.pirmas2
        where prod_lygis in('A','B')
)group prod_kod8 having count(distinct(prod_mtv))>1
;quit;

```

Rezultatas – aštuonženklis gaminio kodas ir Prodcod klasifikatoriaus jam priskirtas matavimo vienetas (čia matavimo vienetai yra užkomentuoti).

	prod_kod8
1	14201030
2	16232000
3	20421630

10 pav. Tikrinimo programos rezultatas

Tikrinant agregavimo lygį būtent šitie trys kodai buvo ignoruojami.

II. /*Tikrinama, kad pardavimai vidaus rinkoje neviršytų pardavimų iš viso*/

```

proc sql create table mag.pard_pardv as
select im_kod, prod_kod, dalis, pard_viso, pard_vid /*,
round(((pard_vid/pard_viso)*100),1)as proc*/
from mag.pirmas2
where pard_viso<pard_vid and pard_viso<>0 /*and
round(((pard_vid/pard_viso)*100),1)>1000 or*/
/*round(((pard_vid/pard_viso)*100),1)<10 and
round(((pard_vid/pard_viso)*100),1)<>.**/
/*order by proc*/
;quit;

```

Rezultatas – įmonių sąrašas su gaminių kodais, kurių pardavimai vidaus rinkoje viršija bendrus pardavimus.

Tokių klaidų vasario mėnesį nepasitaikė.

Patikrinti ar niekas neįrašė nereikalingos gamybos ar pardavimo apimčių esant „-“ matavimo vienetais (kai matavimo vienetai „-“, reikia nurodyti pardavimų iš viso ir vidaus rinkoje tik pinigine išraiška), galima paleidus kitą programą. Dažniausiai tokius matavimo vienetus turi kodai, kuriais koduojamos įvairios detalės.

III. /*Matavimo vienetai „-“*/

```

proc sql;
create table mag.mtv as
select im_kod, prod_kod, prod_mtv, pag_natura, pard_viso, pard_vid, lt_viso,
lt_vid
from mag.pirmas2

```

```
where prod_mtv='- ' and (pag_natura<>0 or pard_viso<>0 or pard_vid<>0)
; quit;
```

Rezultatas – klaidų sąrašas pagal įmones ir atskirus kodus.

Tokio tipo klaidų nerasta.

IV. /*Tikrina pardavimus, pinigus. Jei vienas iš jų yra, turi būti ir kitas*/

```
proc sql;
create table mag.pard_pin as
select im_kod, prod_kod, prod_mtv, dalis, pag_natura, pard_viso, pard_vid,
lt_viso, lt_vid
from mag.pirmas2
where prod_mtv<>"-" and dalis=1 and (pard_viso=0 or pard_viso is null) and
prod_kod not in (
select prod_kod /*išmetamos eilutės, kurios yra visiškai tuščios
arba yra užpildyta tik gamyba*/
from mag.pirmas2
where prod_mtv<>"-" and dalis=1 and (pard_viso=0 or pard_viso is
null)
and (pard_vid=0 or pard_vid is null) and (lt_viso=0 or lt_viso is
null) and (lt_vid=0 or lt_vid is null)
)
union
select im_kod, prod_kod, prod_mtv, dalis, pag_natura, pard_viso, pard_vid,
lt_viso, lt_vid
from mag.pirmas2
where prod_mtv<>"-" and dalis=1 and (lt_viso=0 or lt_viso is null) and
prod_kod not in (
select prod_kod
from mag.pirmas2
where prod_mtv<>"-" and dalis=1 and (pard_viso=0 or pard_viso is
null)
and (pard_vid=0 or pard_vid is null) and (lt_viso=0 or lt_viso is
null) and (lt_vid=0 or lt_vid is null)
)
; quit;
```

Rezultatas – gaminių kodų sąrašas pagal įmones, kur trūksta pardavimų, nes yra nurodyti pinigai ir atvirksčiai.

	IM_KOD	PROD_KOD	PROD_MTV	DALIS	PAG_NATURA	pard_viso	pard_vid	lt_viso	lt_vid
1	1008072	2540130000	kg	1	0	0	0	28000	28000
2	1072534	1412124000	vnt.	1	2	2	2	0	0
3	1072534	1413245500	vnt.	1	1	1	1	0	0
4	2044568	3299594000	kg	1	0	0	0	10700	10700
5	2044568	3299594010	kg	1	0	0	0	10700	10700
6	2089425	1610390000	m ²	1	30	30	30	0	0

11 pav. Tikrinimo programos rezultatas

Matome, kad čia yra taisytinos 6 vietos, tik negalime pasakyti kaip. Tad tokiais atvejais pirmiausia pasitikslinama TSV darbuotojų ar tai nebuvo jų suvedimo klaida, o jau tada skambinama respondentui nurodytais telefonais ir aiškinamasi tokios klaidos priežastys. Džiugu pastebėti, kad ir ši programa atrado klaidas, kurios nebuvo atrastos ir ištaisytos net ir po antrinio redagavimo.

V. /*Jeigu gamyba vidaus rinkoje ir iš viso lygi, tai ir pinigai turi būti lygūs. Ir analogiškas priešingas variantas*/

```
proc sql;
create table mag.pard_pin_lygus as
select im_kod, prod_kod, prod_mtv, pag_natura, pard_viso, pard_vid, lt_viso,
lt_vid
from mag.pirmas2
where dalis=1 and prod_mtv<>'-' and pard_viso=pard_vid and lt_viso<>lt_vid
union
select im_kod, prod_kod, prod_mtv, pag_natura, pard_viso, pard_vid, lt_viso,
lt_vid
from mag.pirmas2
where dalis=1 and prod_mtv<>'-' and pard_viso<>pard_vid and lt_viso=lt_vid
; quit;
```

Rezultatas – sąrašas neteisingai užpildytų eilučių pagal įmones.

	IM_KOD	PROD_KOD	PROD_MTV	PAG_NATURA	pard_viso	pard_vid	lt_viso	lt_vid
1	5888876	1624113300	vnt.	40	40	0	0	0

12 pav. Tikrinimo programos rezultatas

/*gamybos 2 mėnesių palyginimas*/
Analogiškai šios programos pagalba galim palyginti dviejų mėnesių pardavimus bei pinigus.

```
proc sql;
create table mag.gam2men as
select * from (

select im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis,
sum(pagaminta_1) as gamyba_1men, sum(pagaminta_2) as gamyba_2men,
round((sum(pagaminta_2)/sum(pagaminta_1)*100),1)as proc

from (

select im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis,
sum(pag_natura) as pagaminta_1, 0 as pagaminta_2/*, pag_sav, pard_vid,
pinigai_viso, pinigai_vid*/
from mag.sausis
where prod_lygis<>'A'
group im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis
union
select im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis, 0 as
pagaminta_1, sum(pag_natura) as pagaminta_2/*, pag_sav, pard_vid,
pinigai_viso, pinigai_vid*/
from mag.pirmas2
where prod_lygis<>'A'
group im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis

)group im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis

) where proc<>. and proc<>0 and (proc>10000 or proc<1)
order by proc

;quit;
```

Rezultatas – įmonių sąrašas su tiksliais kodais, kurių gamyba nuo praeito mėnesio ženkliai skiriasi.

	IM_KOD	PROD_KOD	PROD_PAV	PROD_MTV	DALIS	PROD_LYGIS	gamyba_1men	gamyba_2men	proc
1	3480009	1413149000	Motenškos arba mergaičių kelinės, kombinezonai su antkrūtiniais ir petnešomis, bėdžiai ir šortai iš megztų arba netų tekstilės medžiagų	vnt.	1		50	5215	10430
2	7120576	2349125500	Kiti keraminiai gaminiai ir dirbiniai (paprastosios keramikos), įskaitant šildymo aparatus, ugniai neatsparias šonines briaunas, krosnių ir (arba) pakutų dalis, vazonus, rankenėles ir (arba) rutulines rankenas, parduotuvių ženklus, radiatorių drėkintuvus	kg	1		132	14000	10606
3	1066255	1419130000	Pirštuotos pirštinės, kumštinės pirštinės ir puspirštinės, megztos arba nertos	porų sk.	2		8	1010	12625
4	2615064	1512121000	Skrynios, lagaminai, lagaminėliai (skrynelės kosmetikai), lagaminėliai (diplomatai), portfeliai, mokyklinės kuprinės ir panašūs daiktai iš išdirbtos, kompozicinės arba lakinės odos, plastikų arba tekstilės medžiagų, aluminio arba kitų medžiagų	vnt.	1		2	279	13950
5	5592392	2511235000	Kitos statybinės konstrukcijos daugiausia iš lakštų: kitų	kg	1		40	5700	14250
6	6138336	1413245500	Vyriški arba berniukų kombinezonai su antkrūtiniais ir petnešomis (išskyrus megztus arba nertus, darbinus ar profesinius)	vnt.	2		11	1573	14300
7	5664478	0812119000	Statybinis smėlis: molingasis smėlis, kaolino smėlis, putnago (lauko špato) smėlis (išskyrus silicinį ir metalingąjį smėlį)	kg	1		18900	4382600	23188
8	8370946	1413147000	Motenškos arba mergaičių suknelės iš megztų arba netų tekstilės medžiagų	vnt.	1		10	2966	29660
9	3495854	1413347000	Motenškos arba mergaičių suknelės (išskyrus megztas ir nertas)	vnt.	2		1	357	35700
10	1066991	1812191000	Spausdinti visų rūšių kalendoriai, įskaitant bloknotų pavidalo kalendorius	vnt.	1		1	3000	300000
11	8510073	1812130000	Spausdinti laikraščiai, žurnalai ir periodiniai leidiniai, leidžiami rečiau negu keturis kartus per savaitę	vnt.	2		400000	1812130000	453033

13 pav. Tikrinimo programos rezultatas

Pažvelgus į sąrašą matome akivaizdžią klaidą 11 eilutėje. Pildant ataskaitą ataskaitos antroje dalyje vietoj gamybos apimtys buvo įrašytas gaminio kodas. Dar reiktų atkreipti dėmesį į tuos kodus, kurių matavimo vienetai yra kg, nes kaip taisyklė šių gaminių gamybos apimtys pokičiai tik išskirtiniais (vienetiniiais) atvejais gali padidėti arba sumažėti kelis kartus palyginus su praeito mėnesio gamybos apimtimi. Tačiau šitiems pataisymams vis tiek reikalinga papildoma informacija, pvz. visų praeitų metų 12 mėnesių dinamika. Geriausia būtų kreiptis į ataskaitą pildžiusį asmenį.

Likusių programų, tikrinančių kainų neatitikimus, atsiskaitymo pilnumą bei ryšį su kitomis ataskaitų formomis, kodai pateikiami 4 priede. Programos suprogramuotos ir jų veikimas patikrintas pasinaudojant Statistikos departamento turimomis duomenų bazėmis ir disponuojama informacija.

3. LOGISTINĖ REGRESIJA

Logistinės regresijos modelis taikomas, kai priklausomas kintamasis yra kategorinis. Šis modelis prognozuoja kategorinio kintamojo reikšmių tikimybes. Logistinė regresija tinka galiojant gana bendroms prielaidoms, t.y. kintamieji X_1, X_2, \dots, X_k nebūtinai turi būti normalieji, nereikalaujama normaliai pasiskirsčiusių paklaidų.

Y priklausomas kintamasis, įgyjantis dvi reikšmes (1 – įvykis, 0 – ne įvykis),

$\vec{x} = \{x_1, \dots, x_m\}$ nepriklausomų kintamųjų vektorius.

$P\{Y_i = 1|\vec{x}\} = p_i$ - tai įvykio tikimybė, o $P\{Y_i = 0|\vec{x}\} = 1 - p_i$ - ne įvykio tikimybė.

Logistinės regresijos modelis atrodo taip:

$$p_i = \frac{\exp\{z(\vec{x}_i)\}}{1 + \exp\{z(\vec{x}_i)\}}, \quad z(\vec{x}_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m,$$

Galimybė įvykti įvykiui $Y_i = 1$ gaunama $\frac{p_i}{1-p_i} = \exp\{z(\vec{x}_i)\}$.

```
proc logistic data=log.duom desc;  
class ivede lygis(ref=first) mtv(ref=last);  
model Y=ivede lygis mtv lt_viso/rsquare;  
run;
```

Čia *ivede* – nepriklausomas kintamasis įgyjantis tokias reikšmes: El.forma, KAU, KLP, PNV, SL, VLN arba Kitas. T.y. suvedamos ataskaitos buvo suskaidytos pagal penkias teritorines statistikos valdybas, išskirta elektroninė forma ir „kitas“ – tai bet kurio iš SD darbuotojo suvesta ataskaita.

lygis – tai kintamasis įgyjantis dvi reikšmes, 1 – gaminiai be agregavimo lygio ir 2 – gaminiai, kurių gaminio kodo agregavimo lygis yra „A“ arba „B“.

Mtv – įgyjamos reikšmės: 1, kai gaminys neturi matavimo vienetų, t.y. „-“, ir 2 kitais atvejais.

Lt_viso – kiekybinis kintamasis; pinigai už parduotą produkciją.

Atlikus logistinę analizę statistinio paketo SAS pagalba, gauti tokie rezultatai:

The LOGISTIC Procedure

Model Information

Data Set	LOG.DUOM
Response Variable	Y
Number of Response Levels	2
Number of Observations	8608
Model	binary logit

Response Profile

Ordered Value	Y	Total Frequency
1	1	177
2	0	8431

Probability modeled is Y=1.

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	63.6634	9	<.0001
Score	90.1555	9	<.0001
Wald	74.3449	9	<.0001

Norint išsiaiškinti ar kovariantė kaip nors įtakoja priklausomojo kintamojo reikšmių tikimybes, reikia tikrinti hipotezę, kad lygūs nuliui koeficientai prie visų pseudokintamųjų. Matome, kad mūsų atveju tikrinama hipotezė ($H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$) apie kokybinės kovariantės įtaką atmetama (Likelihood Ratio < 0.0001). T.y. bent vienas nepriklausomas kintamasis modelyje reikalingas ir jis įtakoja klaidos atsiradimą.

Type III Analysis of Effects

Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
ivede	6	39.1221	<.0001
lygis	1	5.0125	0.0252
mtv	1	31.1783	<.0001
lt_viso	1	0.7332	0.3919

Iš lentelės matome, kad klaidos atsiradimą įtakoja matavimo vienetai (būtent kai jie yra „-“), gaminio lygis (kai gaminiai yra A,B lygio) ir ataskaitos patekimo į duomenų bazę būdas. Būtent elektroninės formos ir Šiaulių TSV suvestos ataskaitos.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard		Wald		
		Estimate	Error	Chi-Square	Pr > ChiSq	
Intercept	1	-3.1851	0.1139	781.5903	<.0001	
ivede EI.forma	1	-0.6103	0.1561	15.2882	<.0001	
ivede KAU	1	-0.0592	0.1557	0.1447	0.7037	
ivede KLP	1	0.2248	0.2174	1.0690	0.3012	
ivede PNV	1	-0.3581	0.3017	1.4088	0.2353	
ivede SI	1	-0.6742	0.2653	6.4561	0.0111	
ivede VLN	1	-0.1991	0.1928	1.0671	0.3016	
lygis	2	1	0.1866	0.0833	5.0125	0.0252
mtv	1	1	0.5189	0.0929	31.1783	<.0001
lt_viso	1	-1.36E-7	1.584E-7	0.7332	0.3919	

Taigi, įsitikinome, kad elektroniniu būdu siunčiamoms ataskaitoms būtinai reikalinga papildoma duomenų kontrolė. Taip pat tikrinant statistinius duomenis didesnę dėmesį reikėtų kreipti ir į išsiskaidančius gaminio kodus bei gaminius, kurie neturi matavimo vienetų.

IŠVADOS

- Išanalizavus Įmonės pramoninių gaminių gamybos tyrimą buvo pastebėta, kad silpnosios šio tyrimo vietos - staistinių duomenų patikimumas. Daugybei atsitiktinių ir sisteminių klaidų "išgaudymui" neužtenka vien tik programuotojų sudarytų duomenų įvedimo programose kontrolių. Todėl duomenų kokybei pagerinti ir buvo nuspręsta sukurti papildomas antriniam redagavimui skirtas programas. Jos buvo kuriamas atsižvelgiant į dažniausiai pasikartojančias klaidas.
- Išanalizavus galimas paklaidų priežastis, pastebėta, kad dažniausiai nesąmoningai daromos klaidos dėl ataskaitos pildymo nesupratimo ir instrukcijų nesilaikymo. Sąmoningoms klaidoms įtaką gali turėti tokie faktoriai kaip nenoras skelbti tam tikro pobūdžio informaciją, taip pat kvalifikuotų darbuotojų trūkumas.
- Gaminų gamybos ataskaitose tarp dažniausiai pasitaikančių klaidų fiksuojamos loginės klaidos: kaip agregavimo lygio neatitikimas, gamybos apimčių pateikimas neatsižvelgus į matavimo vienetus ("-"), kai jos rodyti visai nereikia ir kt.
- Logistinės regresijos analizės metu buvo parodyta, kad statistiniuose duomenyse klaidos atsiradimą įtakoja ataskaitos perdavimo būdas, gaminio agregavimo lygis ir matavimo vienetai.
- Pasiūlytos antriniam redagavimui skirtos programos loginėms ir kainų / gamybos / pardavimų apimčių palyginimo klaidų nustatymui.

SANTRAUKA

Statistikos departamentas siekdamas užtikrinti statistinės informacijos kokybę vadovaujasi Europos statistikos sistemos kokybės deklaracijos ir Europos statistikos praktikos kodekso nuostatomis. Tuo tarpu Pramonės statistikos skyrius savo skelbiamos informacijos kokybę pagerinti gali padidinus pirminių statistinių duomenų kontrolių ir klaidų „išgaudymo“ skaičiumi. Taigi, vienas iš pagrindinių šio darbo uždavinių ir buvo – gerinti Pramonės statistikos skyriaus skelbiamų duomenų (apie pagrindinių gaminių gamybos ir pardavimo apimtis ir jų pokyčius) kokybę, t.y. automatizuoti įmonių pramoninių gaminių gamybos ataskaitose loginių klaidų paiešką. Loginių klaidų „išgaudymui“ buvo pasiūlytos programos, kurių veikimas buvo patikrintas su pirminių (duomenys iki antrinio redagavimo) duomenų rinkiniu, o gauti rezultatai – tai įmonių, padariusių vienokio ar kitokio tipo loginę klaidą, sąrašas. Atlikus logistinę regresiją, buvo gauta, kad klaidos atsiradimą įtakoja ataskaitos pateikimo būdas (t.y. elektroninės formos siuntimas), gaminio kodo agregavimo lygis (t.y. tie kodai, kurie išsiskaido į smulkesnius kodus) bei matavimo vienetai kai jie yra „-“.

REZIUMÈ

In order to warrant the quality of statistical information the Department of Statistics to the Government of the Republic of Lithuania follows Quality declaration of the European Statistical System and Code of practice for European statistics. The Division of statistics on industry, accordingly, in order to improve the quality of information it publishes may increase the number of primary statistical data controls and error eliminations. Thus, the primary aim of this thesis is to improve the quality of data (on the amount and changes of production and trade of major products) that is published by the Division of statistics on industry, i.e. to automate the search for the logical errors in the reports on the production of industry products at enterprises. A software application was developed for the elimination of logical errors. It was tested with the set of primary statistical data (the data before the secondary editing) and it provided as its result a list of companies that have made some kind of mistake. By means of logistic regression it was found that the appearance of a mistake is conditioned by the method of submission of the report (i.e., the uploading of an electronic form), by the level of aggregation of a product code (i.e., the codes that resolve into smaller codes) and by the measures when they are “-”.

LITERATŪROS SĄRAŠAS

1. Čekanavičius V., Murauskas G. (2000), Statistika ir jos taikymai II. Vilnius, TEV. 238 p.
2. D.Delwiche L., J.Slaughter S. (2003), The little SAS book. Cary NC, SAS Institute. 350 p.
3. Ekonominės veiklos rūšių klasifikatorius EVRK 2 red., Statistikos departamentas, p. 249.
4. <http://www.techonthenet.com/>
5. Krapavickaitė D., Plikusas A. Imčių teorijos pagrindai. Vilnius: Technika, 2005, p. 312
6. Levulienė R. (2005), Statistika su SAS I. Mokomoji priemonė. 184 p.
7. LR Statistikos įstatymo pakeitimo įstatymas 1999 m. gruodžio 23 d. Nr. VIII-1511 Vilnius (Žin., 1993, Nr.54-1048; 1995, Nr.55-1355; 1997, Nr.65-1542).
8. Martišius S.A., Kėdaitis V. Statistika 2: statistikos išvados ir sprendimai. Vilnius, Vilniaus universiteto leidykla, 2004, p. 341.
9. PRODCOM Classification, Eurostat, Version: June 15th, 2008, p.293
10. Recommended Practices for Editing and Imputation in Cross-Sectional Business Surveys, 2007, Eurostat (žiūr.: http://edimbus.istat.it/EDIMBUS1/document/RPM_EDIMBUS/RPM_EDIMBUS.pdf).
11. Rekomendacijos: Statistinės informacijos rengimo proceso vykdymo ir dokumentavimo aprašymas, Vilnius: Statistikos departamentas prie Lietuvos Respublikos Vyriausybės, 2005.
12. S. Martišius, V. Kėdaitis. Statistika:II dalis: statistikos išvados ir sprendimai. Vilnius:Vilniaus universiteto l-kla, 2004, 341 p.
13. Statistikos departamentas prie Lietuvos Respublikos Vyriausybės (www.stat.gov.lt)
14. Statistikos departamento Pramonės skyriaus pateikta metodinė medžiaga: „Įmonės pramoninių gaminių gamybos tyrimo metodika“ Patvirtinta Statistikos departamento prie Respublikos Vyriausybės generalinio direktoriaus 2006 m. įsakymu.
15. Statistikos departamento Pramonės skyriaus pateikta metodinė medžiaga: Produktų, gaminių ir paslaugų klasifikatoriaus (PGPK) įdiegimo analizė ir įvertinimas, 2005.
16. Vytautas Kazakevičius. Kokybinių duomenų analizė. Mokomoji priemonė, p. 70.
17. Vytautas Kėdaitis. Matavimai ir duomenų rinkimo metodai, Vilnius, VU, 2003, p. 80

PRIEDAI

1 Priedas

Pagrindinės sąvokos

Įmonė – statistinio stebėjimo vienetas, esantis juridiniu vienetu (arba juridinių vienetų junginiu), gaminantis prekes ar teikiantis paslaugas.

Statistinės ataskaitos formos užpildymo laikas – laikas, kurį įmonė skiria statistikai reikalingos informacijos parengimui bei statistinės ataskaitos formos užpildymui.

Statistinės atsakomybės našta – darbo laiko sąnaudos, kurias patiria įmonė, vykdydama teisės aktuose nustatytus privalomus informacinio pobūdžio įpareigojimus.

Atliktų pramoninių darbų vertė – tai pajamos, kurias įmonė gavo arba gaus už atliktą darbą (paslaugą), ir pagamintai produkcijai jau yra išrašyta sąskaita faktūra, kuri gali būti dar neapmokėta.

Pagaminta produkcija natūriniais vienetais – visa per atsiskaitomąjį laikotarpį pagaminta produkcija, nesvarbu ar ją pardavė kaip prekę, ar panaudojo tolesnei gamybai, ar liko sandėliuose.

Parduota produkcija natūriniais vienetais – įmonės per atsiskaitomąjį laikotarpį pagamintos produkcijos pardavimai natūriniais vienetais, kuriems yra išrašyta sąskaita faktūra, kuri gali būti dar neapmokėta.

Parduota produkcija natūriniais vienetais Lietuvos rinkoje - įmonės per atsiskaitomąjį laikotarpį pagamintos produkcijos pardavimai Lietuvos rinkoje, kuriems yra išrašyta sąskaita faktūra, kuri gali būti dar neapmokėta.

Parduota produkcija veikusiomis kainomis (be PVM ir akcizų) - įmonės per atsiskaitomąjį laikotarpį pagamintos produkcijos pardavimai vertine išraiška, kuriems yra išrašyta sąskaita faktūra, kuri gali būti dar neapmokėta.

Parduota produkcija veikusiomis kainomis (be PVM ir akcizų) Lietuvos rinkoje - įmonės per atsiskaitomąjį laikotarpį pagamintos produkcijos pardavimai Lietuvos rinkoje vertine išraiška, kuriems yra išrašyta sąskaita faktūra, kuri gali būti dar neapmokėta.

Produkcija, panaudota savo reikmėms – tai įmonėje pagaminta produkcija, kuri panaudojama įmonės reikmėms (tik mėnesinio tyrimo rodiklis).

Produkcija, suvartota tolesnei gamybai – tai įmonėje pagaminta produkcija iš kurios vėliau toje pačioje įmonėje pagaminama kitos rūšies produkcija (tik mėnesinio tyrimo rodiklis).

Įmonė, gaminanti produkciją iš nuosavų žaliavų ir medžiagų, pusfabrikačių (visas medžiagas įmonė nusipirko pati), pagamintos produkcijos kiekį ir pardavimo vertę nurodo ataskaitos P-12 pirmojoje dalyje (iš nuosavų medžiagų bei žaliavų, pusfabrikačių).

Įmonė, gaminanti tiek iš nuosavų žaliavų bei medžiagų ir pusfabrikačių, tiek iš užsakovo medžiagų bei žaliavų, pusfabrikačių (už užsakovo žaliavas bei medžiagas, pusfabrikačius įmonė nemoka), pildo I ir II ataskaitos dalis. I –joje dalyje įmonė nurodo pagamintos produkcijos apimtį ir parduotos produkcijos vertę tik tos dalies produkcijos, kuri pagaminta iš įmonės nuosavų žaliavų. II – joje dalyje – pagamintos produkcijos iš užsakovo žaliavų apimtį ir atliktų darbų vertę (pajamas, kurias įmonė gavo už atliktą darbą).

Įmonė, gaminanti produkciją tik iš užsakovų žaliavų ir medžiagų, pusfabrikačių pildo tik ataskaitos P – 12 antrąją dalį (iš užsakovų medžiagų ir žaliavų, pusfabrikačių). Nurodomas visas pagamintos produkcijos kiekis, o į parduotos produkcijos apimtį įskaičiuojama tik atliktų darbų vertė, t. y. pajamos, kurias įmonė gavo už atliktą darbą.

Įmonė, turinti medžiagas bei žaliavas ir pusfabrikačius, tačiau pati negaminanti, o turimas medžiagas bei žaliavas ir pusfabrikačius perduodanti gaminti produkciją sutartiniu pagrindu Lietuvos teritorijoje esančioms įmonėms, nurodo tik parduotos produkcijos apimtį ir vertę.

Vartojami sutrumpinimai

EEB - Europos Ekonominės Bendrijos

EVRK –Ekonominės veiklos rūšių klasifikatorius

IT- Informacinės technologijos

KN –Kombinuotoji nomenklatūra

PGPK - Produktų, gaminių ir paslaugų klasifikatorius

TSV- Teritorinė statistikos valdyba

2 Priedas

Klasifikatoriai

PRODCOM klasifikatorius

Tai tarptautinis Europos lygmens klasifikatorius, skirtas pramonės produkcijos tyrimams, jis sudaro sąlygas palyginti šalyje gaminamos produkcijos ir užsienio prekybos statistinius duomenis. PRODCOM - trečioji grandis klasifikatorių NACE – CPA – PRODCOM grandinėje. Šių klasifikatorių struktūra, klasifikavimo principai ir kodavimo sistema tie patys. Tačiau PRODCOM apima ne visas NACE ir CPA klases, o tik išgaunamosios pramonės produkciją, elektros, dujų ir vandens tiekimą ir pramonės gaminių gamybą.

Kurdami Prodcom, EUROSTAT darbuotojai vadovavosi šiais pagrindiniais reikalavimais:

- detalizacijos laipsnis turi atitikti valstybės ir privačių asmenų interesus;
- nauja sistema turi būti maksimaliai palyginama su užsienio prekybos statistika (poreikis skaičiuoti vidaus suvartojimą);
- teisinis pagrindas, kuris įpareigoja visas EB šalis vykdyti šią sistemą.

Lietuvoje prekių gamybos apimtys pradėtos rinkti pagal Prodcom sąrašą nuo 2002 metų.

Sąrašas į skyrius skirstomas pagal EVRK (nuo 2009 metų pagal NACE rev 2) skyrius. Kiekvieno sąrašo skyriaus dalos yra išdėstytos didėjančia tvarka pagal jų Prodcom kodus.

Pirmame Prodcom sąrašo stulpelyje pateikiami klasifikuojamų Prodcom gaminių kodai. Kiekviena dalį identifikuojama pagal aštuonių skaitmenų kodą: 13.20.11.10 (nacionalinė versija pagal 10). Pirmieji 6 skaitmenys (13.20.10) yra klasifikatoriaus CPA (Sajungos Produktų klasifikatorius pagal veiklos rūšį) kodai. Pirmieji 4 skaitmenys (13.20) atitinka EVRK klases. Prie Prodcom kodo gali būti raidės A ir B. Privalomosios dalos su raide A skaidosi į B dalis, į kurias duomenys įrašomi pasirinktinai. A lygio kodas yra B lygio kodų suma. Taip pat kiekvienas kodas turi jam priskirtą matavimo vienetą.

Sąsajos su kitais klasifikatoriais: Tarptautinis standartinis gamybinis visų ekonominės veiklos rūšių klasifikatorius (ISIC); Statistinis Europos Bendrijos ekonominės veiklos rūšių klasifikatorius (NACE); Ekonominės veiklos rūšių klasifikatorius (EVRK); Svarbiausias produktų klasifikatorius (CPC); Produktų, gaminių ir paslaugų klasifikatorius (PGPK); Produktų pagal veiklos rūšį klasifikatorius (CPA); Harmonizuota sistema (HS); Kombinuota nomenklatūra (CN).

Atnaujinimo periodiškumas - PRODCOM atnaujinamas kiekvienais metais.

PGPK klasifikatorius

Produktų, gaminių ir paslaugų klasifikatorius - tai nacionalinė tarptautinių klasifikatorių CPA ir PRODCOM versija. PGPK skiriasi nuo Prodcom sąrašo 10 ženklų kodais (Prodcom 8 ženklai) ir įtrauktais papildomais gaminių kodais. Šiuos gaminius sudaro:

- tik Lietuvoje gaminami gaminiai (pvz. obuolių sūriai, midaus trauktinės..)
- kai kurios Prodcom dalos buvo susmulkintos į kelis kodus (pvz. dešros į rūkytas ir virtas dešras....)
- išskirti svarbūs Lietuvai gaminiai (pvz. ruginiai miltai, ruginė duona....).

Produktų, gaminių ir paslaugų klasifikatorių sudaro dvi dalys:

Pirmoji dalis - tai nacionalinė versija, parengta pagal CPA 2002 klasifikatoriaus nepramoninio sektoriaus skyrius. Ji vartojama žemės ūkio, miškininkystės, medžioklės ir žuvininkystės produkcijos bei nepramoninių paslaugų ir šalyje gaminamos pramonės produkcijos apskaitai pagal jų gamybinę kilmę.

Antroji dalis - nacionalinė versija, parengta pagal CPA 2002 pramoninio sektoriaus dalis ir Europos Bendrijos pramonės gaminių sąrašą PRODCOM, kuris atnaujinamas kiekvienais metais. Šis klasifikatorius vartojamas šalyje gaminamos pramonės produkcijos apskaitai pagal gamybinę jos kilmę.

Sąsajos su kitais klasifikatoriais: Tarptautinis standartinis gamybinis visų ekonominės veiklos rūšių klasifikatorius (ISIC); Statistinis Europos Bendrijos ekonominės veiklos rūšių klasifikatorius (NACE), (PGPK keturų kodo ženklų lygmenyje struktūriškai atitinka NACE); Svarbiausias produktų klasifikatorius (CPC); Europos Bendrijos produkcijos sąrašas (PRODCOM); Harmonizuota sistema (HS); Kombinuota nomenklatūra (CN).

Atnaujinimo periodiškumas - PGPK pirmosios dalies atnaujinamo periodiškumas priklauso nuo CPA atnaujinimo. PGPK antroji dalis atnaujinama kiekvienais metais atitinkamai pagal tarptautinio klasifikatoriaus PRODCOM pakeitimus.

EVRK klasifikatorius

Ekonominės veiklos rūšių klasifikatorius - tai tarptautinio klasifikatoriaus nacionalinė versija, parengta pagal Europos Sąjungos NACE klasifikatorių. Jis skirtas duomenų, susijusių su veiklos vienetais, skirstymui į kategorijas pagal veiklos rūšis. Juo remiantis rengiami statistiniai duomenys apie produkciją, įvairių rūšių gamybos išlaidas (darbo, žaliavų, energijos ir kt.), kapitalo formavimą ir finansines operacijas statistinių vienetų veikloje.

Statistikos departamento generalinio direktoriaus 2007 m. spalio 31 d. įsakymu Nr. 226 (Žin., 2007, Nr. 119-4877) nuo 2008 m. sausio 1 d. Ekonominės veiklos rūšių klasifikatoriaus 1.1 redakcija (EVRK 1.1 red., parengtas remiantis statistiniu Europos Bendrijos ekonominės veiklos rūšių klasifikatoriumi (NACE Rev.1)) keičiama Ekonominės veiklos rūšių klasifikatoriaus 2 redakcija (EVRK 2 red. - remiantis NACE Rev.2).

Ekonominės veiklos rūšys turi prasmę, kai ištekliai, tokie kaip įrenginiai, darbo jėga, technologijos, informaciniai tinklai ar produktai, suderinami taip, kad juos naudojant sukuriama atitinkamos prekės ar paslaugos. Veikla yra apibūdinama gamybos (prekių ar paslaugų) sąnaudomis, gamybos procesu ir išleidžiamąja produkcija.

EVRK grindžiamas hierarchine kodavimo sistema:
pirmasis lygmuo – sekcijos, identifikuojamos vieno ženklo nuo A iki Q alfabetiniu kodu,
tarpinis lygmuo – posekcijai identifikuojami dviejų ženklų alfabetiniu kodu,
antrasis lygmuo – skyriai, identifikuojami dviejų ženklų skaitmeniniu kodu,
trečiasis lygmuo – grupės, identifikuojamos trijų ženklų skaitmeniniu kodu,
ketvirtasis lygmuo – klasės, identifikuojamos keturių ženklų skaitmeniniu kodu,
penktasis lygmuo (nacionalinis) – poklasiai, identifikuojami šešių ženklų skaitmeniniu kodu.
Alfabetiniai ženklai kodo pozicijos identifikavimui nenaudojami.

Sąsajos su kitais klasifikatoriais: Tarptautinis standartinis gamybinis visų ekonominės veiklos rūšių klasifikatorius (ISIC); Statistinis Europos Bendrijos ekonominės veiklos rūšių klasifikatorius (NACE), (tarp EVRK ir NACE yra griežta elementų sąsaja keturių kodo ženklų lygmenyje); Produktų pagal veiklos rūšį klasifikatorius (CPA), (CPA keturių kodo ženklų lygmenyje struktūriškai atitinka EVRK); Europos Bendrijos produkcijos sąrašas (PRODCOM); Produktų, gaminių ir paslaugų klasifikatorius (PGPK), (PGPK keturių kodo ženklų lygmenyje struktūriškai atitinka EVRK).

Klasifikatoriaus atnaujinimo periodiškumas tiesiogiai priklauso nuo tarptautinio klasifikatoriaus NACE atnaujinimo periodiškumo, kurio peržiūra buvo 2007 metais.

LR AVGVK klasifikatorius

LR AVGVK – tai LR Administracinių vienetų ir gyvenamųjų vietovių klasifikatorius. Klasifikatorius skirtas duomenis klasifikuoti administracinių vienetų ir gyvenamųjų vietovių pagrindu.

Klasifikavimo objektas Lietuvos Respublikos aukštesnieji administraciniai vienetai – apskritys, administraciniai vienetai – savivaldybės ir gyvenamosios vietovės. Pagal tipus gyvenamosios vietovės formaliai turi būti skirstomos į miestus, miestelius ir kaimus. Naudojami

tipai yra: didysis miestas, miestas, miestelis, kaimas, viensėdis, geležinkelio stotis, miesto dalis, kaimo dalis.

Kodo struktūra - 3 lygmenys: 1 lygmuo — apskritis; 2 lygmuo— savivaldybė; 3 lygmuo— gyvenamoji vietovė.

Atnaujinimo periodiškumas - įvykus pasikeitimams, tai yra Lietuvos Respublikos adresų registre, tvarkomame valstybės įmonės registrų centras, įregistruojant (ar išregistruojant) gyvenamąsias vietas, Lietuvos Respublikos Vyriausybei priėmus atitinkamus teisės aktus.

Sąsąjū su kitais klasifikatoriais nėra.

Apie šiuos ir kitus klasifikatorius naudingos informacijos galima rasti Statistikos departamento internetiniame puslapyje www.stat.gov.lt → KLASIFIKATORIAI → Centrinė klasifikatorių duomenų bazė.

3 Priedas

LR AVGVK klasifikatorius

Kodas	Apskritis	Kodas	Savivaldybė
1	Alytaus	11	Alytaus miesto
		15	Druskininkų
		33	Alytaus rajono
		38	Varenos rajono
		59	Lazdijų rajono
2	Kauno	12	Birštono
		19	Kauno miesto
		46	Jonavos rajono
		49	Kaišiadorių rajono
		52	Kauno rajono
		53	Kėdainių rajono
		69	Prienuų rajono
		72	Raseinių rajono
3	Klaipėdos	21	Klaipėdos miesto
		23	Neringos
		25	Palangos miesto
		55	Klaipėdos rajono
		56	Kretingos rajono
		75	Skuodo rajono
		88	Šilutės rajono
4	Marijampolės	18	Marijampolės
		39	Vilkaviškio rajono
		48	Kalvarijos
		58	Kazlų Rūdos
		84	Šakių rajono
5	Panevėžio	27	Panevėžio miesto
		36	Biržų rajono
		57	Kupiškio rajono
		66	Panevėžio rajono
		67	Pasvalio rajono
		73	Rokiškio rajono
6	Šiaulių	29	Šiaulių miesto
		32	Akmenės rajono
		47	Joniškio rajono
		54	Kelmės rajono
		65	Pakruojo rajono
		71	Radviliškio rajono
		91	Šiaulių rajono
7	Tauragės	63	Pagėgių
		77	Tauragės rajono
		87	Šilalės rajono
		94	Jurbarko rajono
8	Telšių	61	Mažeikių rajono
		68	Plungės rajono

		74	Rietavo
		78	Telšių rajono
9	Utenos	30	Visagino
		34	Anykščių rajono
		43	Zarasų rajono
		45	Ignalinos rajono
		62	Molėtų rajono
		82	Utenos rajono
10	Vilniaus	13	Vilniaus miesto
		41	Vilniaus rajono
		42	Elektrėnų
		79	Trakų rajono
		81	Ukmergės rajono
		85	Šalčininkų rajono
		86	Švenčionių rajono
		89	Širvintų rajono

4 Priedas

Programų kodai

3 mėnesių gamybos lyginimas atskiram kodui. Jungiamos dvi 2008 ir 2009 metų lentelės. Dėl pasikeitusios klasifikatoriaus versijos 2008 metų duomenys pirmiausia turi būti perkoduoti. Išmetami kodai, kurių gamyba su praėitu (ar dviem ankstesniais) mėnesiu skiriasi daugiau nei 10 kartų.

```
select im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis,
sum(pagaminta_1) as gamyba_1men, sum(pagaminta_2) as gamyba_2men,
sum(pagaminta_3) as gamyba_12men,
--round (decode( sum(pagaminta_1), 0, null,
(sum(pagaminta_2)/sum(pagaminta_1))*100), 0) as skirt1_proc,
-- round (decode( sum(pagaminta_2), 0, null,
(sum(pagaminta_3)/sum(pagaminta_2))*100), 0) as skirt2_proc
round (decode( sum(pagaminta_1), 0, null,
(sum(pagaminta_3)/sum(pagaminta_1))*100), 0) as skirt3_proc
from (
select im_kod,/*substr(prod_kod,1,8)as prod_kod8*/prod_kod, prod_pav,
prod_mtv, dalis, prod_lygis, sum(pag_natura) as pagaminta_1, 0 as
pagaminta_2, 0 as pagaminta_3/*, pag_sav, pard_vid, pinigai_viso,
pinigai_vid*/
from std_pramone.p12_excel_2009
where substr(periodas,6,2)=01
group by im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis
union
select im_kod,/*substr(prod_kod,1,8)as prod_kod8*/prod_kod, prod_pav,
prod_mtv, dalis, prod_lygis, 0 as pagaminta_1, sum(pag_natura) as
pagaminta_2, 0 as pagaminta_3/*, pag_sav, pard_vid, pinigai_viso,
pinigai_vid*/
from std_pramone.p12_excel_2009
where substr(periodas,6,2)=02
group by im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis
union
select b.im_kod,/*substr(prod_kod,1,8)as prod_kod8*/prod8 as prod_kod,
prod_pav, prod_mtv, dalis, prod_b_list, 0 as pagaminta_1, 0 as pagaminta_2,
sum(pag_natura) pagaminta_3/*, pag_sav, pard_vid, pinigai_viso, pinigai_vid*/
from std_pramone.p12_2008 a, std_pramone.klas_prodcom,
std_pramone.pr_imones b, pr_zana.r20083
where prod=prod_id and prod_metai=2007 and a.im_id=b.im_id and
metai=2008
and menuo=12 and prod_kod=prod7
group by prod_kod, prod_pav, prod_mtv, dalis, prod_b_list,
b.im_kod, prod8
-- select im_kod,/*substr(prod_kod,1,8)as prod_kod8*/prod_kod, prod_pav,
prod_mtv, dalis, prod_lygis, 0 as pagaminta_1, 0 as pagaminta_2,
sum(pag_natura) as pagaminta_3/*, pag_sav, pard_vid, pinigai_viso,
pinigai_vid*/
-- from std_pramone.p12_excel_2009
-- where substr(periodas,6,2)=03
-- group by im_kod, prod_kod, prod_pav, prod_mtv, dalis,
prod_lygis
)
--where dalis = 1
--where nvl(prod_b_list,'A') = 'A'
where substr(prod_kod,1,2) in (08,22,31,32) and (prod_lygis='B' or prod_lygis
is null)
group by im_kod, prod_kod, prod_pav, prod_mtv, dalis, prod_lygis
```



```

) --where (skirt2_proc>1000 or skirt2_proc<10) /*and not skirt2_proc='0' and
prod_mtv in ('vnt.')

```

Lyginamos vidaus ir eksporto kainos. Išmetamos įmonės, kurių vidaus ir eksporto kainos skiriasi daugiau nei 6 kartus.

```

select im_kod, im_pav, menuo, prod_kod, prod_pav, prod_mtv, /*prod_lygis,*/
/*dalis,*/ pag_natura, pard_viso, pard_vid, pard_exp, pinigai_viso,
pinigai_vid,
pinigai_exp, vnt_k, vnt_k_v, vnt_k_e, round(decode(vnt_k_v, 0, null,
(vnt_k_e/vnt_k_v)*100),0)as sk_proc
from(
select im_kod, im_pav, substr(periodas,6,2)as menuo, prod_kod, prod_pav,
prod_mtv, prod_lygis, dalis, pag_natura, pard_viso, pard_vid,
(pard_viso - pard_vid) AS pard_exp, pinigai_viso, pinigai_vid, (pinigai_viso
- pinigai_vid) AS pinigai_exp,
ROUND((DECODE(pard_viso,0, NULL,(pinigai_viso/pard_viso)*1000)),2)AS vnt_k,
ROUND((DECODE(pard_vid,0,NULL,(pinigai_vid/pard_vid)*1000)),2)AS vnt_k_v,
ROUND((DECODE((pard_viso-pard_vid),0,NULL,((pinigai_viso-
pinigai_vid)/(pard_viso-pard_vid))*1000)),2)AS vnt_k_e
from std_pramone.pl2_excel_2009
where substr(periodas,6,2)=01 and (prod_lygis='B' OR prod_lygis IS NULL) and
dalis=1 and substr(prod_kod,1,2) in (08,22,31,32)
) where round(decode(vnt_k_v, 0, null, (vnt_k_e/vnt_k_v)*100),0)>500 or
round(decode(vnt_k_v, 0, null, (vnt_k_e/vnt_k_v)*100),0)<20
order by sk_proc

```

4 mėnesių kainų palyginimas atskiram kodui ir pagal įmones.

```

select distinct(prod_kod), im_kod, prod_pav, prod_mtv, dalis, prod_b_list,
"12", "01", "02", "03",
/*round(decode("01", 0, null, ("02"/"01")*100),0)as proc1, round(decode("02",
0, null, ("03"/"02")*100),0)as proc2,
round(decode("01", 0, null, ("03"/"01")*100),0)as proc3,*/
round(decode("12", 0, null, ("01"/"12")*100),0)as proc4, round(decode("12",
0, null, ("02"/"12")*100),0)as proc5
from(
select distinct(prod_kod), im_kod, prod_pav, prod_mtv, dalis, prod_b_list,
sum("12")as "12",sum("01")as "01", sum("02")as "02", sum("03")as "03"
from(
select distinct(prod8)as prod_kod, im_kod, prod_pav, prod_mtv, dalis,
prod_b_list,
round(decode(pard_viso,0,null,(pinigai_viso/pard_viso)*1000),2)as "12",
0 as "01", 0 as "02", 0 as "03"
from std_pramone.pl2_2008, std_pramone.klas_prodcom, pr_zana.r20083
where prod=prod_id and prod_metai=2007 and menuo=12 and metai=2008 and
prod_kod=prod7 and prod_mtv not in ('-') and dalis=1
/*and substr(prod_kod,1,2) in (25,36,14)*/
union
select distinct(prod_kod), im_kod, prod_pav, prod_mtv, dalis, prod_lygis as
prod_b_list, 0 as "12",
round(decode(pard_viso,0,null,(pinigai_viso/pard_viso)*1000),2)as "01", 0 as
"02", 0 as "03"
from std_pramone.pl2_excel_2009

```

```

where dalis=1 /*and substr(prod_kod,1,2) in (08,22,31,32)*/ and
substr(periodas,6,2)=01 and prod_mtv not in ('-')
union
select distinct(prod_kod), im_kod, prod_pav, prod_mtv, dalis, prod_lygis as
prod_b_list, 0 as "12",
0 as "01", round(decode(pard_viso,0,null,(pinigai_viso/pard_viso)*1000),2)as
"02", 0 as "03"
from std_pramone.p12_excel_2009
where dalis=1 /*and substr(prod_kod,1,2) in (08,22,31,32)*/ and
substr(periodas,6,2)=02 and prod_mtv not in ('-')
union
select distinct(prod_kod), im_kod, prod_pav, prod_mtv, dalis, prod_lygis as
prod_b_list, 0 as "12",
0 as "01", 0 as "02",
round(decode(pard_viso,0,null,(pinigai_viso/pard_viso)*1000),2)as "03"
from std_pramone.p12_excel_2009
where dalis=1 /*and substr(prod_kod,1,2) in (08,22,31,32)*/ and
substr(periodas,6,2)=03 and prod_mtv not in ('-')
)where substr(prod_kod,1,2) in (08,22,31,32) /*and im_kod in(1080691)*/
group by prod_kod, im_kod, prod_pav, prod_mtv, dalis, prod_b_list
)where (prod_b_list='B' or prod_b_list is null)
and (round(decode("12", 0, null, ("01"/"12")*100),0)<20 or round(decode("12",
0, null, ("01"/"12")*100),0)>500)
and not(round(decode("12", 0, null, ("01"/"12")*100),0)=0)
--(round(decode("02", 0, null, ("03"/"02")*100),0)<15 or round(decode("02",
0, null, ("03"/"02")*100),0)>700)
--(round(decode("01", 0, null, ("03"/"01")*100),0)<50 or round(decode("01",
0, null, ("03"/"01")*100),0)>200)
--and round(decode("02", 0, null, ("03"/"02")*100),0) <> 0 and
not(prod_b_list = 'A') and prod_mtv='vnt.' and substr(prod_pav,1,4) not in
'Kiti'
order by proc4

```

Tikrinama ar nebuvo atsiskaityta litais vietoj tūkstančiai litų.

```

select distinct(a.im_kod), b.im_kod2,
substr(replace(replace(replace(replace(replace(replace(upper(a.im_PAV), 'UŽDAR
OJI', 'U'), 'AKCINĖ', 'A'), 'BENDROVĖ', 'B'), 'ŽEMĖS', 'Ž'), 'ŪKIO', 'Ū'), 'KOOPERATYVA
S', 'K'),1,50) pav,
b.im_veikla2, substr(a.periodas,7,1)as menuo, sum(a.pag_natura)as pagaminta,
sum(a.pard_viso)as parduota, sum(a.pard_vid)as parduota_vid,
sum(a.pinigai_viso) as tlt_viso, sum(a.pinigai_vid) as tlt_vid
from std_pramone.p12_excel_2009 a, std_pramone.pr_imones b
where im_metai=2009 and a.im_kod=b.im_kod and periodas in ('2009.02') and
im_frm_id=179 /*and substr(im_veikla2,1,1)='8'*/
and a.pinigai_viso>10000 and a.im_kod not in (select im_kod from
std_pramone.pr_imones where im_frm_id=179 and im_metai=2009 and
im_paj>120000000)
group by a.im_kod, b.im_kod2, a.im_PAV, b.im_veikla2, a.periodas
order by im_veikla2

```

Atskiros įmonės pinigai lyginami su P-11 ataskaitos duomenimis (jei ji atsiskaito P-11)

```

select imone,pavad,veikla,case WHEN apskr in (1,2,4) THEN 'Kaunas'
when apskr in (5,9) then 'Panevezys'
when apskr in (6,8) then 'Siauliai'
when apskr in (3,7) then 'Klaipeda'
else 'Vilnius'

```

```

END AS aps,p12,p11,skirt,men11 as bukle
from(
select a.imone, c.im_pav as pavad, c.im_veikla2 as veikla, c.im_apskr as
apskr, a.p12, a.p11, a.skirt, a.men11
from(
select imone, sum(p12)as p12, sum(p11)as p11, sum(p12)-sum(p11) as skirt,
sum(men11)as men11
from (
select b.im_kod as imone, sum(b.pinigai_viso) p12, 0 as p11, 0 as men11
from std_pramone.p12_excel_2009 b, std_pramone.klas_procom a
where a.prod_kod=b.prod_kod and substr(periodas,6,2)=01 and (PROD_lygis
in('A') or PROD_lygis is NULL)
group by b.im_kod
UNION
SELECT distinct(IM_KOD) as imone, 0 as p12, dl as p11, men1 as men11
FROM std_pramone.PRAM_IMONES, std_pramone.PRAM_DUOMENYS,
std_pramone.PRAM_RODIKLIAI/*, std_pramone.PRAM_ATASKAITOS*/
WHERE d_im_id=im_id and d_metai=im_metai and d_metai=2009 and im_frm_id=175
and rod_metai=d_metai and rod_id=d_rod_id and rod_frm_id=175 and
rod_kod='02'
)
having (sum(p12)-sum(p11))<>0 and sum(p11)<>0
group by imone
)a, std_pramone.pram_imes b, std_pramone.pr_imes c /*im_reg.imn_pram b*/
where a.imone=b.im_kod and b.im_metai=2009 and c.im_metai=2009 and
a.imone=c.im_kod and c.im_frm_id=179
/*a.imone=b.imone and abs(skirt)>10 ... imones in(...)*/
)
where /*imone in(select distinct(im_kod7) from std_pramone.p12_2006 )*/
imone in (select * from (select distinct(im_kod) from std_pramone.pr_imes
where im_frm_id=179 and im_metai=2009 and im_pozymis='-1'))
and substr(veikla,1,3) not in (105)
order by imone --veikla

```

Statistiniai duomenys pagal atskiras įmones. Spausdina visus duomenis apie norimą (prie įmonės kodo nurodytą) įmonę, atskirų gaminių pardavimus, gamybą, ir pinigus pagal mėnesius. Prijungiama ir 2008 metų lentelė tikslesniam „įmonės paveiklui“ nupiešti.

```

select distinct(a.im_kod), b.im_kod2,
substr(replace(replace(replace(replace(upper(a.im_PAV), 'UŽDAR
OJI', 'U'), 'AKCINĖ', 'A'), 'BENDROVĖ', 'B'), 'ŽEMĖS', 'Ž'), 'ŪKIO', 'Ū'), 'KOOPERATYVA
S', 'K'), 1, 50) pav,
b.im_bukle, b.im_apskr as apskritis, b.im_veikla2 as im_veikla, a.prod_kod,
a.prod_pav, a.prod_mtv as mtv, a.dalis, a.prod_lygis,
substr(a.periodas, 1, 4) as metai,
substr(a.periodas, 7, 1) as menuo,
a.pag_natura, a.pag_sav_reik, a.pard_viso, a.pard_vid, (a.pard_viso-
a.pard_vid) as pard_exp, a.pinigai_viso as tlt_viso,
a.pinigai_vid as tlt_vid, (a.pinigai_viso-a.pinigai_vid) as tlt_exp,
round((decode(a.pard_viso, 0, null, (a.pinigai_viso/a.pard_viso)*1000)), 2) as
vnt_k,
round((decode(a.pard_vid, 0, null, (a.pinigai_vid/a.pard_vid)*1000)), 2) as
vnt_k_vid,
round((decode((a.pard_viso-a.pard_vid), 0, null, ((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid)*1000)), 2) as vnt_k_exp,
round(((decode(a.pard_vid, 0, null, (a.pinigai_vid/a.pard_vid)*1000))/
(decode((a.pard_viso-a.pard_vid), 0, null, ((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid)*1000))))*100, 0) as proc
from std_pramone.p12_excel_2009 a, std_pramone.pr_imes b

```

```

where im_metai=2009 and a.im_kod=b.im_kod /*and periodas in ('2009.01')*/ and
im_frm_id=179 and a.im_kod in (4412014)
--and substr(im_veikla2,1,2)='22'
--and substr(prod_kod,1,8)=1061272
--order by prod_kod

union

select distinct(b.im_kod), '0' as im_kod2,
substr(replace(replace(replace(replace(replace(replace(upper(b.im_PAV), 'UŽDAR
OJI', 'U'), 'AKCINĖ', 'A'), 'BENDROVĖ', 'B'), 'ŽEMĖS', 'Ž'), 'ŪKIO', 'Ū'), 'KOOPERATYVA
S', 'K'),1,50) pav,
b.im_bukle, b.im_apskr as apskritis, im_veikla, prod_kod, prod_pav, prod_mtv
as mtv, a.dalis, prod_b_list as prod_lygis,
substr(a.metai,1,4)as metai, substr(a.menuo,1,2)as menuo,
a.pag_natura, a.pag_sav as pag_sav_reik, a.pard_viso,
a.pard_vid,(a.pard_viso-a.pard_vid)as pard_exp, a.pinigai_viso as tlt_viso,
a.pinigai_vid as tlt_vid,(a.pinigai_viso-a.pinigai_vid)as tlt_exp,
round((decode(a.pard_viso,0,null,(a.pinigai_viso/a.pard_viso)*1000)),2)as
vnt_k,
round((decode(a.pard_vid,0,null,(a.pinigai_vid/a.pard_vid)*1000)),2)as
vnt_k_vid,
round(decode((a.pard_viso-a.pard_vid),0,null,((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid))*1000),2)as vnt_k_exp,
round(((decode(a.pard_vid,0,null,(a.pinigai_vid/a.pard_vid)*1000))/
(decode((a.pard_viso-a.pard_vid),0,null,((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid))*1000)))*100,0) as proc
from std_pramone.pl2_2008 a, std_pramone.klas_prodcom, std_pramone.pr_imeses
b
where prod_id=a.prod and prod_metai=2007 and a.metai in (2008) and a.im_id =
b.im_id and b.im_frm_id = 179 and a.dalis in (1, 2)
and (a.menuo between 1 and 12) and b.im_kod in (4412014)
/*and substr(veikla,1,2)=25*/ /*and (prod_b_list='A' or prod_b_list is
null)*/
/*and substr(prod_kod,1,10)='3612129080' */
group by b.im_kod, b.im_bukle, b.im_pav, b.im_apskr, b.im_veikla, prod_kod,
prod_pav, prod_mtv, a.dalis, prod_b_list, a.metai, a.menuo, a.pag_natura,
a.pag_sav, a.pard_viso, a.pard_vid, a.pinigai_viso, a.pinigai_vid,
im_veikla2
--order by menuo-- metai, prod_kod, menuo
-- prijungiam sia dali, jeigu reikia lyginti 2008 su 2007 ar net su 2006
order by prod_kod, prod_pav --metai, menuo

```

Statistiniai duomenys pagal gaminio kodus. Skirtumas su praeita programa tame, kad pirmoje mes rinkomės pagal įmonę įrašę tikslų jos kodą, o čia nurodome gaminio kodą. Ir gauname visą sąrašą įmonių gaminančių šį gaminį.

```

select distinct(a.im_kod), b.im_kod2,
substr(replace(replace(replace(replace(replace(replace(upper(a.im_PAV), 'UŽDAR
OJI', 'U'), 'AKCINĖ', 'A'), 'BENDROVĖ', 'B'), 'ŽEMĖS', 'Ž'), 'ŪKIO', 'Ū'), 'KOOPERATYVA
S', 'K'),1,50) pav,
b.im_veikla2, a.prod_kod, a.prod_pav, a.prod_mtv as mtv, a.prod_lygis,
a.dalis, substr(a.periodas,1,4)as metai, substr(a.periodas,7,1)as menuo,
a.pag_natura, /*a.pag_sav_reik,*/ a.pard_viso, a.pard_vid,(a.pard_viso-
a.pard_vid)as pard_exp, a.pinigai_viso as tlt_viso,
a.pinigai_vid as tlt_vid,(a.pinigai_viso-a.pinigai_vid)as tlt_exp,
round((decode(a.pard_viso,0,null,(a.pinigai_viso/a.pard_viso)*1000)),2)as
vnt_k,

```

```

round((decode(a.pard_vid,0,null,(a.pinigai_vid/a.pard_vid)*1000)),2)as
vnt_k_vid,
round(decode((a.pard_viso-a.pard_vid),0,null,((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid))*1000),2)as vnt_k_exp,
round(((decode(a.pard_vid,0,null,(a.pinigai_vid/a.pard_vid)*1000))/
(decode((a.pard_viso-a.pard_vid),0,null,((a.pinigai_viso-
a.pinigai_vid)/(a.pard_viso-a.pard_vid))*1000)))*100,0) as proc
from std_pramone.p12_excel_2009 a, std_pramone.pr_imones b
where im_metai=2009 and a.im_kod=b.im_kod and periodas in
('2009.01','2009.02') and im_frm_id=179 /*and a.im_kod in (1011398)*/ /*and
substr(im_veikla2,1,2)='22' */
and substr(prod_kod,1,10) in (3299594000)

order by vnt_k_vid --prod_kod

union

select distinct(b.im_kod), b.im_kod2,
substr(replace(replace(replace(replace(upper(im_PAV),'UŽDAROJ
I','U'),'AKCINĖ','A'),'BENDROVĖ','B'),'ŽEMĖS','Ž'),'ŪKIO','Ū'),'KOOPERATYVAS'
,'K'),1,50) pav,
b.im_veikla2, prod_kod, prod_pav, prod_mtv as mtv, prod_b_list as
prod_lygis, dalis, substr(metai,1,4)as metai, substr(menuo,1,2)as menuo,
pag_natura, /*pag_sav,*/ pard_viso, pard_vid,(pard_viso-pard_vid)as pard_exp,
pinigai_viso as tlt_viso,
pinigai_vid as tlt_vid,(pinigai_viso-pinigai_vid)as tlt_exp,
round((decode(pard_viso,0,null,(pinigai_viso/pard_viso)*1000)),2)as vnt_k,
round((decode(pard_vid,0,null,(pinigai_vid/pard_vid)*1000)),2)as vnt_k_vid,
round(decode((pard_viso-pard_vid),0,null,((pinigai_viso-
pinigai_vid)/(pard_viso-pard_vid))*1000),2)as vnt_k_exp,
round(((decode(pard_vid,0,null,(pinigai_vid/pard_vid)*1000))/
(decode((pard_viso-pard_vid),0,null,((pinigai_viso-
pinigai_vid)/(pard_viso-pard_vid))*1000)))*100,0) as proc
from std_pramone.p12_2008 a, std_pramone.klas_prodcom, std_pramone.pr_imones
b
where prod=prod_id and a.im_id=b.im_id and prod_metai=2007 and metai in
(2008) and b.im_frm_id = 179
/*and substr(veikla,1,2)=25*/ /*and (prod_b_list='A' or prod_b_list is
null)*/ /*and menuo=9*/ /*and dalis=1*/
/*and b.im_kod='2514548'*/ and substr(prod_kod,1,10)in('2524285000') and
(menuo between 1 and 12) --order by im_kod/*menuo*/
--group by im_kod7, bukle, pavad, jur_vietove, veikla, prod_kod, prod_pav,
prod_mtv, dalis,
--prod_b_list, metai, menuo, pag_natura, pag_sav, pard_viso,
pard_vid,pinigai_viso, pinigai_vid
--order by vnt_kaina_vid/*prod_kod, menuo, pard_vid*/
--)group by menuo

```

Gamyba lyginama su pardavimais. Spausdina įmonę ir gaminio kodą, kurio gamyba nuo pardavimų skiriasi daugiau nei 10 kartų.

```

select distinct(prod_kod)as prod_kod, prod_pav, prod_lygis, prod_mtv, dalis,
im_kod, pav, metai, menuo, pagaminta, parduota,
round(decode(pagaminta,0,null,parduota/pagaminta*100),0) as procent
from(

select distinct(prod_kod)as prod_kod, prod_pav, prod_lygis, prod_mtv, dalis,
b.im_kod,

```

```

substr(replace(replace(replace(replace(replace(replace(upper(b.im_PAV), 'UŽDAR
OJI', 'U'), 'AKCINĖ', 'A'), 'BENDROVĖ', 'B'), 'ŽEMĖS', 'Ž'), 'ŪKIO', 'Ū'), 'KOOPERATYVA
S', 'K'), 1, 50) pav,
b.im_veikla2, substr(periodas, 1, 4) as metai, substr(periodas, 6, 2) as menuo,
sum(pag_natura) as pagaminta, sum(pard_viso) as parduota
from std_pramone.p12_excel_2009 a, std_pramone.pr_imones b
where im_metai=2009 and a.im_kod=b.im_kod and periodas in
('2009.01', '2009.02', '2009.03') and im_frm_id=179
and substr(prod_kod, 1, 2) in (31, 32) and (prod_lygis='B' or prod_lygis is
null)
group by prod_kod, b.im_kod, b.im_pav, prod_pav, a.dalis,
substr(periodas, 6, 2), prod_mtv, substr(periodas, 1, 4), prod_lygis,
b.im_veikla2
--order by im_kod, pinigai

) where round(decode(pagaminta, 0, null, parduota/pagaminta*100), 0) > 1000 or
round(decode(pagaminta, 0, null, parduota/pagaminta*100), 0) < 10
-- and not(parduota=0)
order by im_kod, menuo, procent, prod_kod

```

Gamyba nuo 2002 iki 2008 metų. Spausdina atskiram gaminio kodui visų čia išvardintų metų gamybą.

```

select distinct(prod_kod), /*prod_pav, prod_mtv,*/ sum(gamyba_2008) as
GAMYBA_2008, sum(gamyba_2007) as GAMYBA_2007, sum(gamyba_2006) as GAMYBA_2006,
sum(gamyba_2005) as GAMYBA_2005, sum(gamyba_2004) as GAMYBA_2004,
sum(gamyba_2003) as GAMYBA_2003, sum(gamyba_2002) as GAMYBA_2002
from(
/*2008*/
select prod_kod, prod_pav, prod_mtv, metai, sum(pag_natura) as gamyba_2008, 0
as gamyba_2007, 0 as gamyba_2006, 0 as gamyba_2005, 0 as gamyba_2004,
0 as gamyba_2003, 0 as gamyba_2002
from std_pramone.p12_2008, std_pramone.klas_prodcom
where prod=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/ and
prod_metai=2007 and metai=2008
and prod_kod in
('3611121000', '3613109000', '3613105000', '3614123000', '3614125000', '3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, metai/*, pard_viso, pinigai_viso*/

UNION
/*2007*/
select prod_kod, prod_pav, prod_mtv, metai, 0 gamyba_2008, sum(pag_natura) as
gamyba_2007, 0 as gamyba_2006, 0 as gamyba_2005, 0 as gamyba_2004,
0 as gamyba_2003, 0 as gamyba_2002
from std_pramone.p12_2006, std_pramone.klas_prodcom
where prod=prod_id and /*(prod_b_list='A' or prod_b_list is null) and*/
prod_metai=2005 and metai=2007
and prod_kod in
('3611121000', '3613109000', '3613105000', '3614123000', '3614125000', '3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, metai/*, pard_viso, pinigai_viso*/

UNION
/*2006*/
select prod_kod, prod_pav, prod_mtv, metai, 0 as gamyba_2008, 0 as
gamyba_2007, sum(pag_natura) as gamyba_2006, 0 as gamyba_2005, 0 as
gamyba_2004,
0 as gamyba_2003, 0 as gamyba_2002
from std_pramone.p12_2006, std_pramone.klas_prodcom

```

```

where prod=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/ and
prod_metai=2005 and metai=2006
and prod_kod in
('3611121000','3613109000','3613105000','3614123000','3614125000','3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, metai

UNION
/*2005*/
select prod_kod, prod_pav, prod_mtv, metai, 0 as gamyba_2008, 0 as
gamyba_2007, 0 as gamyba_2006, sum(pag_natura) as gamyba_2005, 0 as
gamyba_2004,
0 as gamyba_2003, 0 as gamyba_2002
from std_pramone.sum_prod_2005, std_pramone.klas_prodcom
where prod=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/ and
prod_metai=2003 and metai=2005
and prod_kod in
('3611121000','3613109000','3613105000','3614123000','3614125000','3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, metai

UNION
/*2004*/
select prod_kod, prod_pav, prod_mtv, metai, 0 as gamyba_2008, 0 as
gamyba_2007, 0 as gamyba_2006, 0 as gamyba_2005, sum(pag_natura) as
gamyba_2004,
0 as gamyba_2003, 0 as gamyba_2002
from std_pramone.sum_prod_2004, std_pramone.klas_prodcom
where prod=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/ and
prod_metai=2003 and metai=2004
and prod_kod in
('3611121000','3613109000','3613105000','3614123000','3614125000','3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, metai

UNION
/*2003*/
select prod_kod, prod_pav, prod_mtv, dp_metai as metai, 0 as gamyba_2008, 0
as gamyba_2007, 0 as gamyba_2006, 0 as gamyba_2005, 0 as gamyba_2004,
sum(dp1) as gamyba_2003, 0 as gamyba_2002
from std_pramone.duom_p02, std_pramone.klas_prodcom
where dp_prod_id=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/
and prod_metai=2003 and dp_metai=2003
and prod_kod in
('3611121000','3613109000','3613105000','3614123000','3614125000','3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, dp_metai

UNION
/*2002*/
select prod_kod, prod_pav, prod_mtv, dp_metai as metai, 0 as gamyba_2008, 0
as gamyba_2007, 0 as gamyba_2006, 0 as gamyba_2005, 0 as gamyba_2004,
0 as gamyba_2003, sum(dp1) as gamyba_2002
from std_pramone.duom_p02, std_pramone.klas_prodcom
where dp_prod_id=prod_id /*and (prod_b_list='A' or prod_b_list is null)*/
and prod_metai=2002 and dp_metai=2002
and prod_kod in
('3611121000','3613109000','3613105000','3614123000','3614125000','3663313000',
'3663630000')
group by prod_kod, prod_mtv, prod_pav, dp_metai

) group by prod_kod /*prod_mtv, prod_pav,*/

```