



HARIDUS- JA  
TEADUSMINISTEERIUM



# PIAAC uuringu andmeanalüüsikoolitus

28. augustil 2014 Tartus

Vivika Halapuu  
Haridus- ja Teadusministeerium

Andres Võrk  
Tartu Ülikool/Praxis

## Päevakava

10:00-11:30	Uuringust üldiselt ja selle metoodilised eripärad, ülevaade IEA IDB Analyzeri võimalustest
11:30-12:00	Milline on matemaatilise kirjaoskuse tase erineva õppevaldkonna lõpetanud kõrgharitute hulgas Eestis, Poolas ja Slovakkias?
12:00-12:45	Lõuna
12:45-13:30	Milline on tehnoloogiarikas keskkonnas probleemilahendusoskuse tase valitud riikides vanusegrupiti?
13:30-14:15	Milline on IKT oskuste kasutussagedus Eesti, Poola ja Slovakkia tööturul?
14:15-14:30	Paus
14:30-16:00	Palga seos kõrgeima omandatud haridustaseme ja matemaatilise kirjaoskuse tasemega

# Ülevaade uuringust

# PIAAC – rahvusvaheline täiskasvanute oskuste uuring

- Osales 24 riiki (peamiselt OECD riigid, v.a Küpros ja Venemaa) ja ca 160 000 inimest
  - Austraalia oma mikroandmeid ei avalikustanud
  - Venemaa andmetega probleemid
  - Prantsusmaa andmed hilinesid
- Lisandumas 9 riiki: Tšiili, Kreeka, Indoneesia (ainult pealinna piirkonnas), Israel, Leedu, Uus-Meremaa, Sloveenia, Singapur ja Türgi
- Eestis osales uuringus 7632 inimest

## Mida ja kuidas mõõdeti?


Uuringus mõõdeti järgmisi oskusi:

- Funktsionaalset lugemisoskust, sh baasilisi lugemisoskusi
- Matemaatilist kirjaoskust
- Probleemilahendusoskust tehnoloogiarikkas keskkonnas (v.a Prantsusmaa, Hispaania, Itaalia ja Küpros)

Nende oskuste ühisnimetajana kasutame läbivalt ka terminit **infotöötlusoskused**.

Mõõtmise nii arvuti- kui paberipõhiselt.

# Funktsionaalne lugemisoskus

Osa 1

Ülesanne 1 - Küsimus 1/3

Vaadake eelkooli reeglite loetelu. Allolevale küsimusele vastamiseks märkige ära loetelus olev informatsioon.

Mis kell hiljemalt peaksid lapsed eelkooli saabuma?

## Eelkooli reeglid

Tere tulemast meie eelkooli! Meid on ees ootamas aasta täis lusti, õppimist ja üksteisega tutvumist. Palun võtke mõni hetk meie eelkooli reeglitega tutvumiseks.


- Palun tooge oma laps kohale kella 9.00-ks.
- Uneajaks pange lapsele kaasa väike tekk või padi ja/või väike pelune mänguasi.
- Riietage oma laps mugavalt ja tooge kaasa vahetuseks mõeldud riided.
- Palume mitte kaasa lubada ehteid või maiustusi. Kui teie lapsel on sünnipäev, rääkige teiste laste kostitamisest oma lapse õpetajaga.
- Palun tooge oma laps täielikult rietatuna, mitte pidžaamas.
- Palun kirjutage registreerumisel oma täisallkiri. Seda nõuab protseduurieskiri. Täname Teid!
- Hommikusööki pakutakse kuni kella 7.30-ni.
- Ravimid peavad olema etikettidega varustatud originaalpakendites ning kantud igas klassis asuval ravinilehele.
- Küsimuste korral pöörduge oma klassi õpetaja või pr Männiaru või pr Tammiku poole.



... on oskus mõista, hinnata ja kasutada kirjalikke tekste selleks, et ühiskonnas edukalt toimida, saavutada oma eesmärged ning arendada oma teadmisi ja võimeid.

# Matemaatiline kirjaoskus

... oskus hankida, kasutada, tõlgendada ning edastada matemaatilist teavet ja matemaatilisi ideid (arvud, kogused, mõõtmed, tõenäosused, protsendid jne) selleks, et tulla toime elus ette tulevate matemaatilisi teadmisi nõudvate olukordadega.

 Osa 1


Ülesanne 11 - Küsimus 1/1

Lugege artiklit tuulegeneraatorite kohta. Sisestage oma vastus allolevale küsimusele, kasutades numbriklahve.

Mitut tuulegeneraatorit on vaja, et asendada tuumareaktori poolt toodetud elektrit?

## Tuulegeneraatorid


2005. aastal sulges Rootsi valitsus viimase tuumareaktori Barsebäcki elektrijaamas. Reaktor tootis igal aastal keskmiselt 3572 GWh elektrienergiat.



Rootsis jätkub töö suurte tuulegeneraatoreid kasutavate rannäärsete tuuleparkide rajamisel. Iga tuulegeneraator toodab umbes 6000 MWh elektrienergiat aastas.

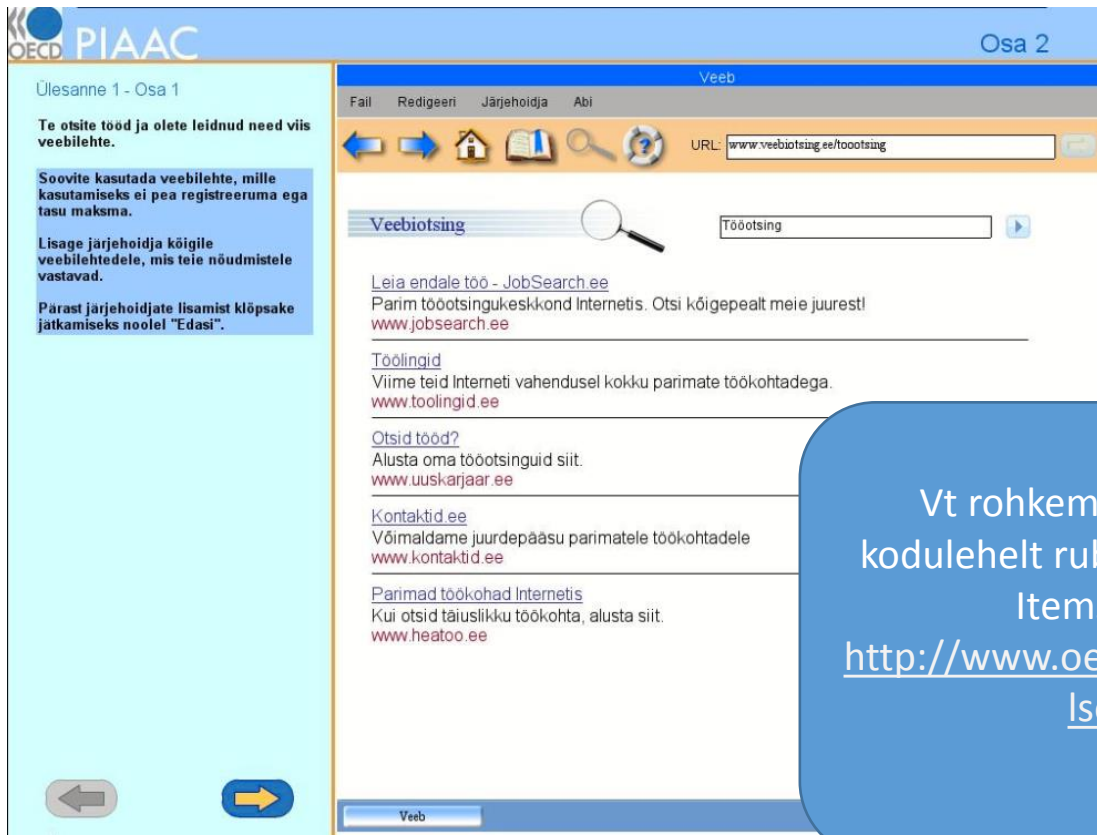
**Informatsiooniks:**  
Elektrienergiat mõõdetakse vatt-tundides(Wh)

1 kWh	= 1 kilovatt-tund	= 1000 Wh
1 MWh	= 1 megavatt-tund	= 1 000 000 Wh
1 GWh	= 1 gigavatt-tund	= 1 000 000 000 Wh



# Probleemilahendusoskus tehnoloogiarikkas keskkonnas

...on oskus kasutada digitaalset tehnoloogiat, kommunikatsioonivahendeid ja arvutivõrgustikke selleks, et hankida ja hinnata infot, suhelda ja täita praktilisi ülesandeid.



Vt rohkem näidisülesandeid OECD PIAACi kodulehelt rubriigist „Publications“ (fail „Sample Items“) ja uuringukeskkonnas:  
<http://www.oecd.org/site/piaac/educationandskillsonlineassessment.htm>  
(inglise keeles)



## Miks need oskused?

- Globaalne konkurents ja kasvav vajadus teha keerukamat tööd
- Vajalikud eri olukordades: tervisega seotud info, finantskirjaoskus, töö, edasi õppimine, e-riigis osalemine jne
- Neile baseerub mitmete teiste oskuste omandamine
- Õpitavad, mistõttu saab neid poliitika kujundamisega mõjutada
- Võrreldavalt mõõdetavad, kaalutud mõõta ka nt sotsiaalseid oskusi, loovust, ettevõtlikkust, ...

# Andmete kasutamise võimalused

## RAHVUSVAHELISED ANDMEBAASID:

1. PIAAC põhiuuringu rahvusvaheline mikroandmetega andmebaas (*PUF* failid)
2. Rahvusvaheline andmestik, mida on võimalik kasutada üle võrgu *International Data Exploreriga (IDEga)*
3. PIAAC Põhjamaade andmebaas

## RIIKLIK ANDMEBAAS:

4. PIAAC uuringu riiklik mikroandmetega andmebaas

# 1. PIAAC uuringu rahvusvaheline mikroandmetega andmebaas

- *International Public Use Data Files, PUF*
- Andmed alla laetavad aadressilt:  
<http://www.oecd.org/site/piaac/publicdataandanalysis.htm>
- Sisaldab kõigi uuringus osalenud riikide (v.a Austraalia ja Küpros) andmefaile, milles riigid said omal valikul teha andmete mahasurumisi.


# Andmete allalaadimine OECD kodulehelt (1)

## OECD Skills Surveys

**PIAAC**

- Survey of Adult Skills
- Skills assessed
- Public data and analysis**
- Policy decision-making
- Country-specific materials
- Take the test
- Join the Survey
- Publications
- Events and Multimedia
- Contact us
- En ESPAÑOL

[PIAAC](#)




### OECD Skills Surveys

As part of its Programme for the International Assessment of Adult Competencies (PIAAC), the OECD collects and analyses data that assist governments in assessing, monitoring and analysing the level and distribution of skills among their adult populations as well as the utilisation of skills in different contexts.


The Survey of Adult Skills, implemented in 24 countries, and the [Education and Skills Online Assessment](#) for individuals are part of the package of tools available to support countries develop, implement and evaluate policies that foster both the development of skills and the optimal use of existing skills.

### First results




The 2013 edition of [Skills Outlook](#) provides a summary of the first results and key findings of the Survey of Adults Skills (PIAAC). It enables countries to see themselves in the light of other countries' performance.

### Data related products



**Data Products** of the Survey of Adults Skills (PIAAC) allow users to explore the survey's data directly through a diverse set of products: Data Explorer, Data Analyser, Public Use Files and Web-packages.

### Country-specific material



[Country notes](#), [interactive data](#) and [multilingual summaries](#) regarding twenty-two OECD Members participating in the first round of the Survey of Adult Skills (PIAAC), two OECD Partners, as well as the European Union

» [Read about the Survey of Adult Skills in different languages:](#)

[em Português](#)

[العربية](#)

[на русском языке](#)

[中文](#)

## Andmete allalaadimine OECD kodulehelt (2)

### International Data Explorer

The International Data Explorer (IDE) allows you to create statistical tables and charts to explore adults' (16-65 years of age) performance in literacy, numeracy, and problem solving in technology-rich environments. Data can be analysed by country as well as in relation to demographic characteristics, social and linguistic background, education, use of skills and employment status.

### **International Public Use Data Files**

Public Use Files containing individual unit record data in SAS and SPSS format are available for downloading for each of the countries participating in the Survey of Adult Skills other than Australia. These contain both responses to the background questionnaire and the cognitive assessment. Full documentation of the contents of the files is provided. Please note that these files are primarily designed for use by researchers and data analysts. The files are also large and may take some time to download. To access the Australian Public Use File please write to: [microdata.access@abs.gov.au](mailto:microdata.access@abs.gov.au)

**Public Use Files (PUF)**. Please note that these files were updated on 7 November 2013. Information about the changes can be found [here](#).

**German PIAAC Scientific Use File** and its **User Guide**. Please note that access to some data is subject to prior authorisation. Please see 'data and documents' tab on the GESIS page cited previously.

**International codebook**

**Derived variables codebook**

**Background Questionnaire**

**PIAAC Data Analysis Manual (SAS®, STATA®)**

**SAS® Macro** (version 2.0). Please note that this file was updated on 18 December 2013; 27 March 2014 and 11 July 2014. (For any queries or to be informed of any update on this macro, please contact [edu.piaac@oecd.org](mailto:edu.piaac@oecd.org)).

**STATA® Macro** (version 1.2). Please note that this file was updated on 9 January 2014.

**Data compendia - Background variables**. Please note that these files were updated on 13 November 2013.

## Andmete allalaadimine OECD kodulehelt (3)

**vs-web-fs-1.oecd.org - /piaac/puf-data/**

[\[To Parent Directory\]](#)

12/18/2013	5:32 PM	<dir>	<a href="#">CSV</a>
10/7/2013	1:22 PM	<dir>	<a href="#">SAS</a>
11/7/2013	4:39 PM	<dir>	<a href="#">SPSS</a>
9/17/2013	1:44 PM	168	<a href="#">web.config</a>

- 22 riigi andmefailid
- Austraalia ja Küprose andmeid ei ole.
- Alla laadides paigutada kõik failid ühte kausta.
- Nimetusi mitte muuta.
  - prg – general response file
  - aut – Austria [riigi nimetus]
  - p1 – PIAACi 1. voor

**vs-web-fs-1.oecd.org - /piaac/puf-data/SPSS/**

[\[To Parent Directory\]](#)

11/7/2013	3:38 PM	17342801	<a href="#">prgautp1.sav</a>
11/7/2013	3:38 PM	18223489	<a href="#">prgbelp1.sav</a>
11/7/2013	3:38 PM	88828649	<a href="#">prgcanp1.sav</a>
11/7/2013	3:38 PM	20554361	<a href="#">prgczep1.sav</a>
11/7/2013	3:38 PM	18833481	<a href="#">prgdeup1.sav</a>
11/7/2013	3:39 PM	24970113	<a href="#">prgdnkp1.sav</a>
11/7/2013	3:39 PM	20201025	<a href="#">prgespp1.sav</a>
11/7/2013	3:39 PM	25275697	<a href="#">prgestp1.sav</a>
11/7/2013	3:39 PM	18842233	<a href="#">prgfinp1.sav</a>
11/7/2013	3:39 PM	23513209	<a href="#">prgfrap1.sav</a>
11/7/2013	3:39 PM	30109817	<a href="#">prggbrp1.sav</a>
11/7/2013	3:39 PM	19982089	<a href="#">prgirlp1.sav</a>
11/7/2013	3:39 PM	15429713	<a href="#">prgitap1.sav</a>
11/7/2013	3:39 PM	17505817	<a href="#">prgjpn1.sav</a>
11/7/2013	3:39 PM	22215481	<a href="#">prgkor1.sav</a>
11/7/2013	3:39 PM	18028401	<a href="#">prgnldp1.sav</a>
11/7/2013	3:39 PM	17722497	<a href="#">prgnorp1.sav</a>
11/7/2013	3:39 PM	30633017	<a href="#">prgpolp1.sav</a>
11/7/2013	3:39 PM	13378297	<a href="#">prgrusp1.sav</a>
11/7/2013	3:39 PM	18922497	<a href="#">prgsvkp1.sav</a>
11/7/2013	3:39 PM	15716497	<a href="#">prgswep1.sav</a>
11/7/2013	3:39 PM	17015249	<a href="#">prgusap1.sav</a>

## Rahvusvahelise mikroandmebaasi üldinfo

- Vastajaid igast riigist keskmiselt 5000 (min 4500 kui mõõtmine ühes keeles, ainult FL ja MK; min 5000, kui mõõtmine ühes keeles, FL, MK, PT.
- Võimalikud lisavalimid (Kanada, Taani, Austraalia) selles ei kajastu.
- Riigispetsiifilised küsimused ei kajastu.
- Erinevad mahasurumised.
- Avalikult kättesaadavas rahvusvahelises andmebaasis ligi 160000 vaatlust, üle 1300 tunnuse.



## 2. Rahvusvaheline andmestik, mida on võimalik kasutada üle võrgu *International Data Exploreriga (IDE)*

- Kättesaadav aadressil:  
<http://www.oecd.org/site/piaac/publicdataandanalysis.htm>.
- Online andmetöötlusvõimalus on ehitatud andmefaili peale, milles andmed võivad olla detailsemal kujul kui mikroandmetega andmefailis, kuid andmeid endid näha ei ole.
- Eelised:
  - andmetöötlus kiire ja online'is
  - kiire graafiline ülevaade tulemustest

## IDE analüüsivõimalused

- Korraga saab analüüsida ühte mõõdetud oskuste domeenidest või mõnda taustatunnust teiste taustatunnuste lõikes
- Nende tunnustena, mille lõikes nt lugemisoskust uuritakse, saab kasutada üksnes kategoorilisi tunnuseid.

## IDE suurimad puudused

- Ei ole võimalik näha, kui suure valimi pealt üks või teine analüüs tehti.
- Analüütilised võimalused väga piiratud (kirjeldav ülevaade erinevates lõigetes).
- Ei näidata nende arvutuste tulemusi, mille aluseks on vähem kui 60 juhtumit.
- Lisaks: joonistel ei saa muuta riikide järjestust.

### 3. PIAAC Põhjamaade andmebaas

- PIAACi Põhjamaade koostöös valmiv andmebaas, kus Soome, Rootsi, Norra, Taani ja Eesti rahvusvahelistele mikroandmetega andmefailidele lisatakse registriandmeid.
- Vaheversioon on valmis, kuid ei ole veel avalik.
- Andmestik asub füüsiliselt Taani Statistikaametis.
- Analüütikutel on võimalik taotleda juurdepääsu andmetele, esitades selleks [taotluse Taani Statistikaametisse](#).
- Ligipääs andmetele [kaugpöördusssüsteemi](#) kaudu: andmeid arvutisse alla laadida ei saa.

## Põhjamaade andmebaasi loomise eesmärk

- Suurendada veelgi niigi ulatusliku andmebaasi analüüsipotentsiaali
- Saada täpsemat infot tunnuste kohta, mis inimeste endi poolt raporteerituna võivad olla „mustad“. Seeläbi teha paremaid analüüse.
- Võrrelda Põhjamaid ülejäänud maailmaga (Põhjamaade PIAACi aruanne). Eesti jaoks oluline positsioneerida ennast Põhjamaade hulgas.

## Andmed Rahvastikuregistrist:

- Sünniriik
- Kodakondsus
- Immigratsioonistaatus
- Perekonnatüüp\*
- Abielustaatus
- Viimases abielustaatuses oldud aastate arv
- Laste arv (kokku, alla 18 aastased lapsed, alla 7 aastased lapsed)
- Regioon (NUTS3)
- Elukoha linnastumisaste
- Riiki elama asumise aeg

\* - Eesti kohta seda tunnust ei ole

## Andmed Eesti Hariduse Infosüsteemist:

- Kõrgeim omandatud haridustase (tase, valdkond)
- Ajaliselt kõige viimasena omandatud haridustase (tase, valdkond)
- Praegused õpingud (algusaeg, tase, valdkond)
- Osalus elukestvas õppes\*

\* - Eesti kohta seda tunnust ei ole

## Andmed Maksukohuslaste registrist/ Äriregistrist:

- Hõivestaatus\*\*
- Tegevusala
- Amet
- Ametialane staatus
- Sektor
- Töötajate arv töökohas\*
- Töötajate arv organisatsioonis/ettevõttes
- Töötajate sooline jaotus töökohas\*
- Töötajate hariduslik jaotus töökohas\*
- Töötajate arv organisatsioonis/ettevõttes\*
- Töötajate sooline jaotus organisatsioonis/ettevõttes\*
- Töötajate hariduslik jaotus organisatsioonis/ettevõttes\*
- Töine sissetulek
- Kasutatav sissetulek
- Majapidamise kasutatav sissetulek\*

\* - Eesti kohta seda tunnust ei ole

\*\* - Eesti kohta tuletati tunnus mitmetest registritest saadud info põhjal



## Andmed Töötuna ja tööotsijana arvel olevate isikute ning tööturuteenuste osutamise registrist ja Töötuskindlustuse andmekogust:

- Tööturutoetused (summa, kestus)
- Osalus tööturumeetmetes
- Osalus tööturumeetmetes – tööturukoolitus
- Osalus tööturumeetmetes – subsideeritud töö
- Osalus tööturumeetmetes – nõustamine
- Osalus tööturumeetmetes – rahaline vm toetus oma ettevõtte alustamiseks
- Tööotsimise kategooria

## Andmed Ravikindlustuse andmekogust, Sotsiaalteenuste ja –toetuste andmeregistrist ning Riiklikust pensionikindlustuse registrist

- Ajutise töövõimetuse hüvitised (summa, saamisaeg)
- Toimetulekutoetus (saamise fakt, kestus)
- Vanemahüvitise saamise kestus päevades
- Töövõimetushüvitised/puudetoetused (saamise fakt, saamise algusaeg)
- Kas inimene saab varajast pensioni või mitte

## 4. PIAAC uuringu riiklik mikroandmetega andmebaas

- Andmebaas sarnaneb rahvusvahelise mikroandmetega andmefailiga, kuid sisaldab detailsemaid tunnuseid nt hariduse ja sünniriigi kohta.
- Lisaks detailsem info osadest registritest.
- Ei ole avalikult kättesaadav.

# Valminud ja valmimas aruanded

# Valminud ja valmimas eesti keeles:

## Valminud:

- **Halapuu, V., Valk, A.** (2013). Täiskasvanute oskused Eestis ja maailmas: PIAAC uuringu esmased tulemused. Tartu: Haridus- ja Teadusministeerium
- **Anspal, S., Järve, J., Jürgenson, A., Masso, M., Seppo, I.** (2014). Oskuste kasulikkus tööturul: PIAAC uuringu temaatiline aruanne nr 1. Tartu: Haridus- ja Teadusministeerium
- **Saar, E., Unt, M., Lindemann, K., Reiska, E., Tamm, A.** (2014). Oskused ja elukestev õpe: kellelt ja mida on Eestil oskuste parandamiseks õppida? PIAAC uuringu temaatiline aruanne nr 2. Tartu: Haridus- ja Teadusministeerium
- Mõned lõputööd

## Valmimas:

- Oskuste tipud ja mahajääjad Eestis
- Oskuste roll palgalõhede seletamisel
- Oskused ja valmisolek kasutada IKTd
- Formaalaridus ja oskused
- Oskuste (mitte)vastavus Eesti tööturu vajadustele
- TLÜ RASI projekt
- Projekt „Counterfactual Impact Evaluation of Estonian Adult Vocational Training Activity“

## Valminud inglise jm keeltes

- **OECD** (2013). OECD Skills Outlook 2013: First Results from the Survey of Adult Skills
- **OECD** (2013). The Survey of Adult Skills – Reader's Companion
- **OECD** (2013). Technical Report of the Survey of Adult Skills (PIAAC)
- **Dinis da Costa, P., Rodrigues, M., Vera-Toscano, E., Weber, A.** (2014). Education, Adult Skills and Social Outcomes: Empirical evidence from the Survey on Adult Skills (PIAAC 2013). European Commission JRC Science and Policy Reports
- **Flisi, S., Goglio, V., Meroni, E., Rodrigues, M., Vera-Toscano, E.** (2014). Occupational mismatch in Europe: Understanding overeducation and overskilling for policy making. European Commission JRC Science and Policy Reports
- **Hanushek, E. A., Schwerdt, G. Wiederhold, S., Woessmann, L.** (2014). Returns to Skills Around the World. Evidence from PIAAC
- **Rosdahl, A.** (2014). Fra 15 til 27 år: PISA 2000-ELEVERNE I 2011/12. SFI – Det Nationale Forskningscenter for Velfærd

## Valmimas inglise keeles

- The use of skills in the workplace
- Skills mismatch
- Skills and labour market outcomes
- Digital literacy, problem solving in TRE, and ICT use
- Trends, ageing and the determinants of skills
- A closer look at the population with low levels of proficiency
- Põhjamaade PIAACi aruanne
- Skill acquisition, skill loss, and age. A comparative study of Cognitive Foundation Skills (CFS)

## Mõned ideed seni katmata teemadest

- Ettevõtlusega alustamine (eestispetsiifilised küsimused)
- Osa I küsimused: erinevad hoiakulised küsimused
- Inimressursijuhtimise tavad (eestispetsiifilised küsimused)
- Tööotsimise kohta käiv info (Eesti andmetega problemaatiline)
- Detailsem vaade probleemilahenduse domeenile
- Sektorivaated (problemaatiline)



# Taustaküsimustik

# Rahvusvaheline andmefail – struktuur

Osa A – Üldine info vastaja kohta

Osa B – Haridus ja koolitus

Osa C – Praegune tööalane staatus ja tööajalugu

Osa D – Praegune töökoht

Osa E – Viimane töökoht viimase 12 kuu jooksul

Osa F – Töökohal kasutatavad oskused

Osa G – Töökohal kasutatavad lugemis-, arvutamis- ja IKT oskused

Osa H – Igapäevaelus kasutatavad lugemis-, arvutamis- ja IKT oskused

Osa I – Mitmesugused küsimused vastaja kohta

Osa J – Taustainformatsioon

## Rahvusvaheline andmefail – struktuur

- Dispositsioonikoodid
- Üldiselt intervjuust
- Erinevat tüüpi ülesannete juurde suunavad tunnused
- Oskuste skoorid (PVLIT1..10; PVNUM1...10; PVPSL1...10)
- Valimi disaini iseloomustavad tunnused
- Kaalud, replikatsioonikaalud

# Tunnuste nimetused

## KÕIGE TAVAPÄRASEM

- J\_Q04a – tegu on J osa küsimusega, mille järjenumbriks on 04a.
- J\_Q04c2 – tegu on J osa küsimuse, mille järjenumbriks on 04c teise alaosaga.

## TRENDITUNNUSED

- J\_Q04c2\_T – tegu on sisuliselt sama tunnusega nagu J\_Q04c2, kuid tunnust on muudetud selliselt, et ta oleks võrreldav IALSi või ALL uuringu taustaküsimustikuga.

## TULETATUD KOONDTUNNUSED

- IMYRS – tegu on konsortsiumi poolt tuletatud koondtunnusega.

## NN COARSENEDE TUNNUSED

- J\_Q03b\_C – tegu on nn *coarsened* tunnusega, st tunnusega, mille vastuste skaalat on ühest otsast koondatud (nt kategooria „rohkem kui 4“). Erandiks ISCO08\_C, ISIC4\_C, kus C viitab sõnale *current*.

## TULETATUD TUNNUSED

- C\_D05 – tuletatud (*derived*) tunnus. Tuletatud taustaküsimustiku täitmise käigus.

**PVLIT1...10, PVNUM1...10, PVPSL1...10, SPWGT0, SPWGT1...80**

# Üksikvastuste teooria ja selle mõju andmete analüüsimisele

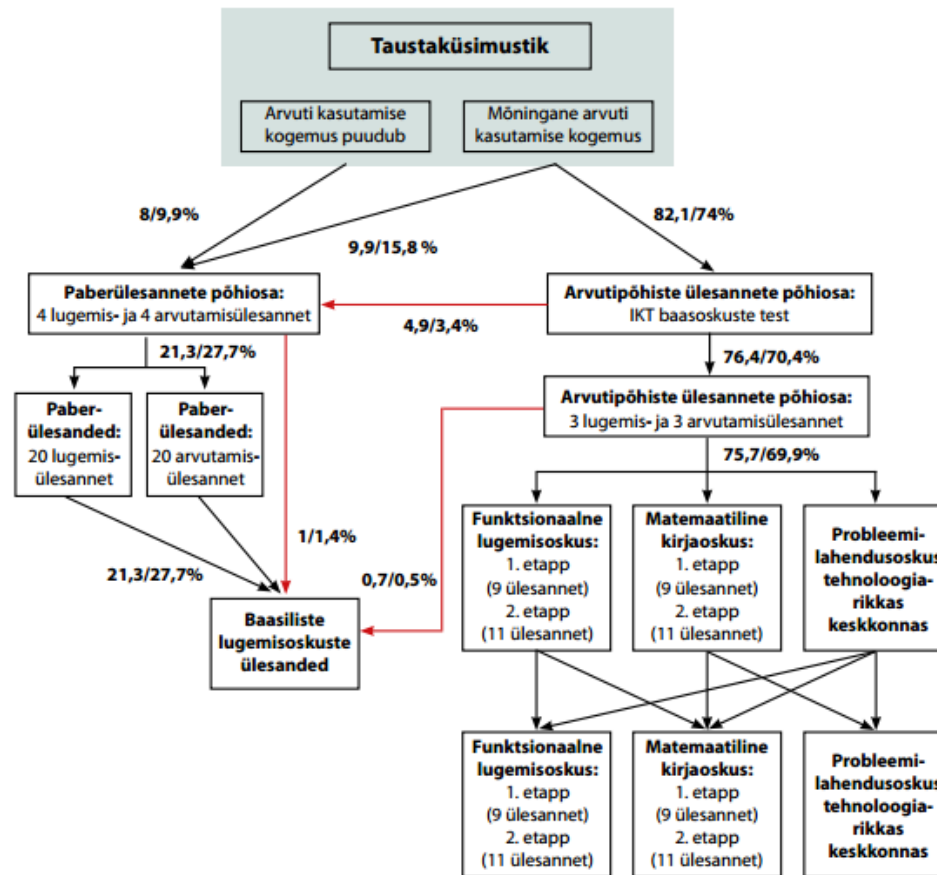
## Kõik ei vastanud kõigele

Nii taustaküsimustiku kui ülesannete lahendamise osas suunamised!

### Suunamised taustaküsimustikus:

- D blokile vastasid üksnes uuringu hetkel hõivatud
- E blokile need, kes olid olnud hõivatud viimase 5 aasta jooksul
- F ja G bloki küsimusi (töökohal kasutatavad oskused) küsiti ainult nendelt, kes olid olnud hõivatud 12 kuu jooksul enne uuringu toimumist

## Erinevad vastamisteed ülesannete lahendamise osas



## Miks kõik kõigele ei vasta?

- PIAAC uuringus mõõdeti kolme erinevat domeeni, lisaks baasilised lugemisoskused (58 FL ülesannet, 56 MK ja 14 PL ülesannet)

Keskmine ajakulu minutites:

	Funktsionaalne lugemisoskus	Matemaatiline kirjaoskus	Probleemilahenduoskus tehnoloogiarikas keskkonnas
Ülesanne	1	1,04	2,69
Moodul	19,9	20,8	18,8

- Lisaks mahukas taustaküsimustik
- Kõikidelt inimestelt kõige küsimine võtnuks liiga kaua aega



## PIAAC uuringu disain

- PIAACi disaini puhul on tegu mitmese maatriksdisainiga (*multiple matrix design*): populatsiooni esindavad teatud inimesed, kellest igaüks lahendab teatud osa ülesannete kogumist.
- Ülesanded, mida inimene lahendab, sõltuvad tema oskuste tasemest (aga ka vastustest teatud taustaküsimustiku küsimustele)

Kõik ei lahenda igat liiki ülesandeid. Küsimus, kuidas saada kõigile kõigi oskuste skoorid?

Vastus – IRT

## Üksikvastuste teooria (Item Response Theory – IRT)

IRT annab võimaluse hinnata mudeleid, mis seovad omavahel

- jälgitava vastamismustri
- mittejälgitava alustunnuse (võimekuse) ning
- ülesannete parameetrid (raskusaste, diskrimineerimisparameeter jne).

Selleks, et teooriat saaks rakendada, peab lahendatud ülesannetes esinema mingigi kattuvus, mis võimaldab infot „laenata“.

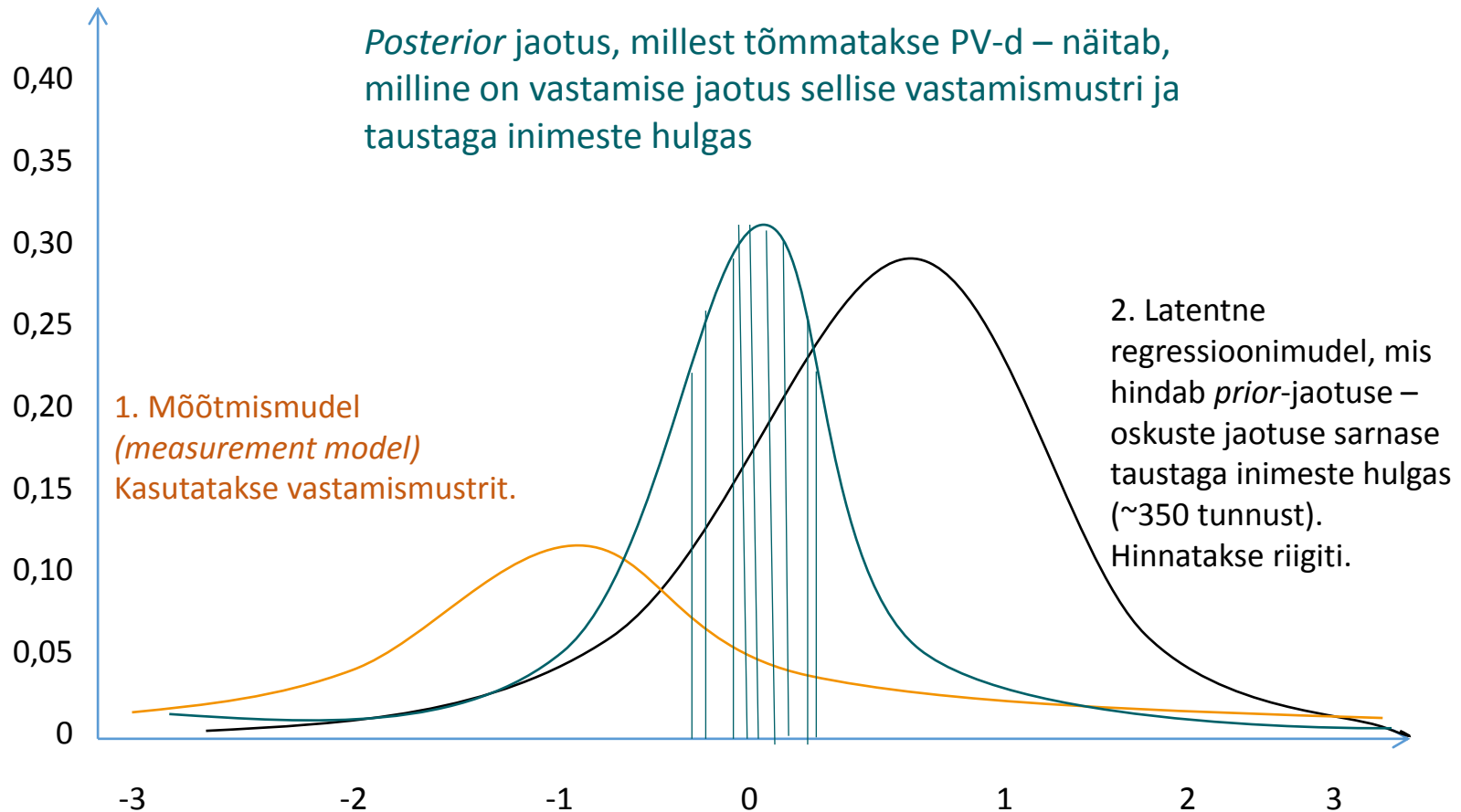
## Mida see kõik praktikas tähendab?

	Ül 1	Ül 2	Ül 3	Ül 4	Ül 5	Ül 6	Ül 7	Ül 8
Inimene A	1	1	0	1	1			
Inimene B	1	0	1	1	0			
Inimene C			1	1	1	1	0	
Inimene D			1	0	1	1	0	
Inimene E				1	1	1	1	1
Inimene F				1	1	0	1	1
Inimene G				0	1	0	1	0

Info õigete/valede vastuste kohta

Palju taustainformatsiooni inimeste endi kohta

## 2-tasandiline IRT mudel



## Tuletatud tunnused ehk *plausible value'd*

- Nii funktsionaalse lugemisoskuse, matemaatilise kirjaoskuse kui tehnoloogiarikas keskkonnas probleemilahendusoskuse tulemusi esitatakse skaalal 0...500
- Vastajad paigutuvad 500-punktilisele skaalale oma võimekuse, ülesanded oma raskusastme põhjal
- Iga domeeni kohta esitatud 10 tuletatud väärtust

*Plausible value'd* on mõeldud gruppide ja mitte indiviidide kohta järelduste tegemiseks!

## Oskuste tasemed

Funktsionaalne lugemisoskus/ matemaatiline kirjaoskus		Probleemilahendusoskus tehnoloogiarikas keskkonnas	
0-175	Alla 1. taset	0-240	Alla 1. taset
176-225	1. tase	241-290	1. tase
226-275	2. tase	291-340	2. tase
276-325	3. tase	341-500	3. tase
326-375	4. tase		
376-500	5. tase		

## Plausible value'de kasutamine arvutustes

- Korrektsete tulemuste saamiseks tuleks teha kõik arvutused kõigi replikatsioonikaalude ja kõigi PVdega ning võtta tulemustest keskmine.
- Kindlasti ei tohi kasutada PVde keskmist, kuna sel juhul luuakse ideaaljuht, mis ei ole reaalne.
- Lihtsustatud juhtude puhul kasutada pigem ühte PVd ja arvestada sellega, et standardvigu tõenäoliselt alahinnatakse.

## Plausible value'dest ja valimi üldisest iseloomust tulenev mõju andmeanalüüsile

- Kuna PV-d on juhuslikud suurused *posterior* jaotusest, käib nendega kaasas täiendava vea komponent.
- Valimi võtmisest tingitud vea arvutamiseks peame kasutama nii populatsioonikaalu kui replikatsioonikaale (kuna tegu pole juhuvalimiga, vaid kihistatud juhuvalimiga)
- **Viga = valimi võtmisest tingitud viga (valimiviga) + andmete imputeerimisest tingitud viga (mõõtmisviga)**



# Kas tavapärasel viisil andmete analüüsile lähenemine viiks valede järeldusteni?

	Matemaatilise kirjaoskuse keskmine tulemus	SRS	SMP+IMP
Mehed	276	0,78	0,86
Naised	270	0,64	0,80
Erinevus	-6	1,01	1,28
	Funktsionaalse lugemisoskuse keskmine tulemus	SRS	SMP+IMP
Mehed	275	0,83	1,09
Naised	277	0,68	0,81
Erinevus	-2	1,07	1,24

	Mean	Standard Error	
		SRS	SMP + IMP
Overall Mathematics Score	491	1.5	5.2
Mathematics Score for Girls	495	2.0	5.5
Mathematics Score for Boys	487	2.1	7.6
Difference Between Girls & Boys	8	2.9	8.3

Allikas: TIMSSi andmed, New Zealand (1999), E. Gonzaleze ettekandest

Replikatsioonikaalud. Tuletatud  
väärtuste ja replikatsioonikaalude  
kasutamine analüüsis

## *Plausible values* (tõepärane/usutav väärtus)

- Eesmärk hinnata pädevusi populatsiooni tasemel: keskmist, hajuvust, seoseid taustateguritega
- Valimisse sattunud eri inimeste käest küsitakse erinevaid asju!
  - Kõikide komponentide küsimine oleks nii küsitlejale kui vastajale liiga ajakulukas
  - Võimalusel kohandatakse küsimuse raskusastet vastavalt vastaja võimetele
- Eelduste ja statistiliste mudelitega pannakse inimesed hiljem samale skaalale ja üldistatakse tulemusi => iga inimene ei saa mitte ühte kindlat väärtust vaid tõenäosusjaotuse enda pädevuse jaotuvuse kohta
- Iga inimese tõenäosusjaotust iseloomustatakse sealt võetud juhuslike väärtustega => “plausible values”
  - Võiks anda andmefailis iga inimese jaotust iseloomustavad parameetrid, aga seda liiga keeruline (jaotus mitmedimensionaalne)

## Statistiku hindamine kui andmetes *plausible values*

- Leia soovitatav statistik  $\beta$  (keskmine, regressioonikordaja jmt) iga PV korral (olgu  $J$  tükki)
  - Ebaefektiivne on võtta statistik  $\beta$  ainult ühe PV korral (nt esimene *plausible value*), ei arvesta imputeerimisest tekkivat variatsiooni
  - Vale on võtta PV-de keskmine ja selle põhjal arvutada üks statistik
- Statistiku punkthinnang on keskmine kõikidest  $\beta$ -statistikutest
- Variatsioon = valimi variatsioon + imputeerimisvariatsioon (Rubin 1987)

$$V_{total} = V_{jrr}(\beta) + V_{imp}(\beta)$$

( “within-imputation variance” + “between imputation variance” )

Kui imputeerimisvariatsioon suur, siis parem

$$V_{total} = V_{jrr}(\beta) + (1 + 1/J) \times V_{imp}(\beta)$$

$$\beta = \frac{\sum_{j=1}^J \beta_j}{J}$$

## Imputeerimisest tulev variatsioon

- Leiame statistiku iga PV korral
- Leiame statistikute dispersiooni alljärgneva valemi alusel

$$V_{imp}(\beta) = \frac{\sum_{j=1}^J (\beta_j - \beta)^2}{J - 1}$$

- Erinevatest imputeerimistest tulenev variatsioon
- Kui kasutada ainult ühte PV väärtust, siis alahindame tegelikku variatsiooni

## Valimi variatsiooni hindamine (imputeerimisesisene variatsioon)

- Tegemist ei ole lihtsa juhuvalikuga, vaid keerulisema valikudisainiga uuringuga (kihid, klastrid, mittevastamine, kaalude kalibreerimine).
- Riikide vahel valikudisainid erinevad
  - Väga keerukas oleks kõikide riikide puhul seda välja uurida + andmekaitse probleemid
- Vaja universaalset meetodit statistiku standardvea leidmiseks, kui riike omavahel võrrelda
- Kasutame simulatsioonimeetodeid ja leiame iga *plausible value* korral dispersiooni. Lõplik dispersioon on nende keskmine

$$V_{jrr}(\beta) = \frac{\sum_{j=1}^J V_{j,jrr}(\beta)}{J}$$

- Praktikas kasutatud ka dispersioonihinnangut ainult esimese PV korral

## Variatsiooni hindamine PIAACis - *Replication sampling*

- Kasutame replikatsiooni valimeid - Süstemaatiliselt kustutatud valimist osa vaatlusi (kaal = 0)
  - Jackknife Repeated Replication (JRR)
    - Variant 1 (JK1) või variant 2 (JK2)
  - Balanced Repeated Replication (BRR)
    - Fay's variant – ei kasutata tegelikult
- Kohanda allesjäänud valimikaalused -> replikatsioonikaalud
- Leida punkthinnangud kasutades kogu valimit ja ka iga replikatsioonikaalu korral
- Leia replikatsioonihinnangute dispersiooni kogu valimi põhjal hinnatud hinnangute suhtes
- Statas: svy, brrweight või jkrweight

## Valimi variatsiooni hindamine

■ Dispersioon  $V_{j,jrr}(\beta) = f \sum_{r=1}^R (\beta_{j,r} - \beta_j)^2$

$R$  – replikatsioonide arv

$f$  – sõltub replikatsioonivalimi moodustamise viisist

■ Jackknife type 1  $f=(R-1)/R$

■ Jackknife type 2  $f=1$

■ BRR fay  $f=1/[R(1-fay)^2]$

- NB! Alternatiiv on kasutada dispersiooni arvutamisel mitte lähtekaaludega saadud suurust vaid replikatsioonitulemuste keskmist

$$V_{j,jrr}(\beta) = f \sum_{r=1}^R (\beta_{j,r} - \bar{\beta}_j)^2$$

$$\bar{\beta}_j = \frac{\sum_{r=1}^R \beta_{j,r}}{R}$$



## Replikatsioonikaalud PIAACis

Variable	Obs	Mean	Std. Dev.	Min	Max
spfw0	7632	117.4218	24.39007	60.58855	227.555
spfw1	7632	117.4218	27.58829	0	409.5915
spfw2	7632	117.4218	28.07782	0	386.5593
spfw3	7632	117.4218	28.22376	0	431.7424
spfw4	7632	117.4218	27.47511	0	427.6749
spfw5	7632	117.4218	27.91954	0	334.2408
spfw75	7632	117.4218	28.0138	0	342.7931
spfw76	7632	117.4218	27.51332	0	366.7626
spfw77	7632	117.4218	27.39523	0	333.0671
spfw78	7632	117.4218	27.58833	0	431.8828
spfw79	7632	117.4218	27.49441	0	403.1724
spfw80	7632	117.4218	27.60039	0	396.6776

# Replikatsioonikaalude meetodid PIAACis

	Meetod	Replikatsioonide arv
Austria	Jackknife 1	80
Belgium	Jackknife 2	80
Canada	Jackknife 1	80
Cyprus	Jackknife 2	80
Czech Republic	Jackknife 2	80
Denmark	Jackknife 1	80
Estonia	Jackknife 2	80
Finland	Jackknife 2	80
Germany	Jackknife 1	80
Ireland	Jackknife 2	80
Italy	Jackknife 2	80
Japan	Jackknife 2	80
Korea, Republic of	Jackknife 2	80
Netherlands	Jackknife 2	80
Norway	Jackknife 2	80
Poland	Jackknife 2	80
Slovak Republic	Jackknife 2	80
Spain	Jackknife 2	80
Sweden	Jackknife 2	80
United Kingdom	Jackknife 2	80
United States	Jackknife 2	45

# Replikatsioonikaalude meetodid

- JK1 - *delete-one jackknife ~ delete-a-group jackknife ~ random groups approach*
  - Valim jagatakse kihtideks, igast kihist visatakse juhuslikult välja üks vaatlus, teiste kaalu suurendatakse proportsionaalselt
- JK2 – *the paired jackknife, the stratified jackknife*
  - Valim jagatakse kihtideks (nt 80), igast kihist juhuslikult pooled vaatlused sisse, teise poole kaal korrutatakse kahega
  - Sobilik, kui valimis olevad üksused on kihistatud ja võetud süstemaatilise valiku teel sorteeritud nimekirjast
  - Tegelikult veel kaalude kalibreerimine

## Analüüsi võimalused (käsitsi)

1. Kasutades PV-de keskmist – vale dispersioon

2. Ignoreerides PV ja disainikaalu

Võta esimene PV, eelda, et LJV => **punkthinnang nihkega**, alahindab dispersiooni imputeerimise komponenti, teine komponent teadmata suunas nihkes

Statas: reg pvlit1 male

3. Ignoreerides PV ja replikatsioonikaale

Võta esimene PV ja valimikaal => nihketa punkthinnang, alahindab dispersiooni imputeerimise komponenti, teine komponent teadmata suunas nihkes

Statas: reg pvlit1 male [pw=spfw0]

4. Kasutades esimest PV ja replikatsioonikaale

svyset [pw=spfw0], jkrweight(spfw1-spfw80, multiplier(1)) vce(jackknife) mse  
svy: reg pvlit1 male

5. Kasutades kõiki PV – tsükkel üle PV ja käsitsi arvutused

6. Kasutades kõiki PV ja replikatsioonikaale - tsükkel üle PV, svy käsk, käsitsi arvutused

## Analüüsi võimalused kasutades PV ja replikatsioonikaale

1. Käsitsi ise kirjutades valemid lahti või kombineerides valmisolevate käskudega

### Spetsiaalsed käsud

2. **SPSSis** IDB analyzer – abiprogramm, mis annab ette vajalikud sisendparameetrid SPSSi makrodele
3. **Statas** – ado-failide pakett „piaactools“  
ssc install piaactools
4. **SASis** pakett PIAAC\_Tool.sas

## Näide SPSSi makrost JB\_RegPC.ieasps

```
!do !pv = 1 !to !pverr
```

```
...
```

```
!let !dvar = !concat(!rootpv,!pv,!tailpv)
```

```
weight by !wgt.
```

```
regression var = !dvar !xvar
```

```
...
```

```
!do !r = 1 !to !nrwgt.
```

```
weight by !concat(!wgtroot,!r).
```

```
...
```

```
regression var = !dvar !xvar
```

```
/ dependent = !dvar
```

```
/ method = enter
```

```
/ desc = all
```

# Näide Stata do-failist piaacreg.ado

```
program define _piaacregpv, rclass
```

```
syntax [varlist(numeric default=none)], [overpv(string) overval(string) pvdep(string)  
pvindep1(string) pvindep2(string) pvindep3(string)] weight(varlist max=1) rep(varlist) jk(integer)  
cmd(string) ...
```

```
forvalues ipv=1(1)10 {
```

```
....
```

```
if "`pvindep1'"!=" local pvindep="`pvindep1`ipv'"
```

```
if `pv'!=0 local outcome="`pvdep`ipv'
```

```
`cmd' `outcome' `pvindep' `varlist' [pw=`weight']
```

```
foreach replication of local rep {
```

```
    qui `cmd' `outcome' `pvindep' `varlist' [pw=`replication'] ...
```

```
...
```

```
    foreach var in `pvindep' `varlist' `stala' {
```

```
        local sq_`var' = (_b[`var'] - `fulstat_`var')^2
```

```
        local sumofsq_`var'=`sumofsq_`var'+ `sq_`var'"
```

```
..
```

```
}
```

## Veel kommentaare

- Riikidevaheline võrdlus – võib käsitleda riike kui sõltumatuid
  - NB! Kui võrdled ühte riiki kõigi riigi keskmisega, vaja kohandada riikide keskmise hinnangu standardviga (pole sõltumatu)
$$SE_{(\alpha-a)} = \sqrt{SE_{\alpha}^2 + \left( \frac{(C-1)^2 - 1}{C^2} \right) SE_a^2}$$
  - Riigi keskmine tavaliselt kaalumata
- Riigisisene gruppide võrdlus – pead arvestama, et grupid ei ole teineteisest sõltumatud, kuna võeti ühe valimidisainiga.
  - Peab gruppide võrdlust tegema nt regressioonimeetodiga
  - Välja arvatud juhul kui uuritav grupitunnus oli eelkihistamise aluseks (Eestis vanus ja sugu)



# *Hands-on* ülesanded

Milline on matemaatilise kirjaoskuse tase erineva õppevaldkonna lõpetanud kõrgharitute hulgas Eestis, Poolas ja Slovakkias?

Milliste valdkondade lõpetajate matemaatiline kirjaoskus on statistiliselt oluliselt kõrgem/madalam sotsiaalteaduste, majanduse ja õigusteaduse lõpetajate omast?

Kas Eesti õpetajate matemaatilise kirjaoskuse tase on statistiliselt oluliselt madalam kui kõigi vaatlusaluste postsotsialistlike riikide õpetajate oma?

## Meid huvitavad tunnused

- Kas inimesed on kõrgharidus või mitte? B\_Q01a, B\_Q01a\_T, B\_Q01a\_T1, B\_Q01a3, B\_Q01a3\_C, ISCED\_HF, ISCED\_HF\_C, EDCAT6, EDCAT7, EDCAT8, EDLEVEL3
- Mis valdkonnas on inimene kõrghariduse omandanud? B\_Q01b
- Inimeste matemaatilise kirjaoskuse skoorid PVNUM1...PVNUM10

# Tulemuste tõlgendamiseks

Valdkond 1 - Üldharidus

Valdkond 2 - Õpetajakoolitus ja kasvatusteadus

Valdkond 3 - Humanitaarteadused, keeled ja kunst

Valdkond 4 - Sotsiaalteadused, majandus ja õigusteadus

Valdkond 5 - Täppisteadused, matemaatika ja arvutiõpetus

Valdkond 6 - Tehnika, tootmine ja ehitus

Valdkond 7 - Põllumajandus ja veterinaaria

Valdkond 8 - Tervishoid ja hoolekanne

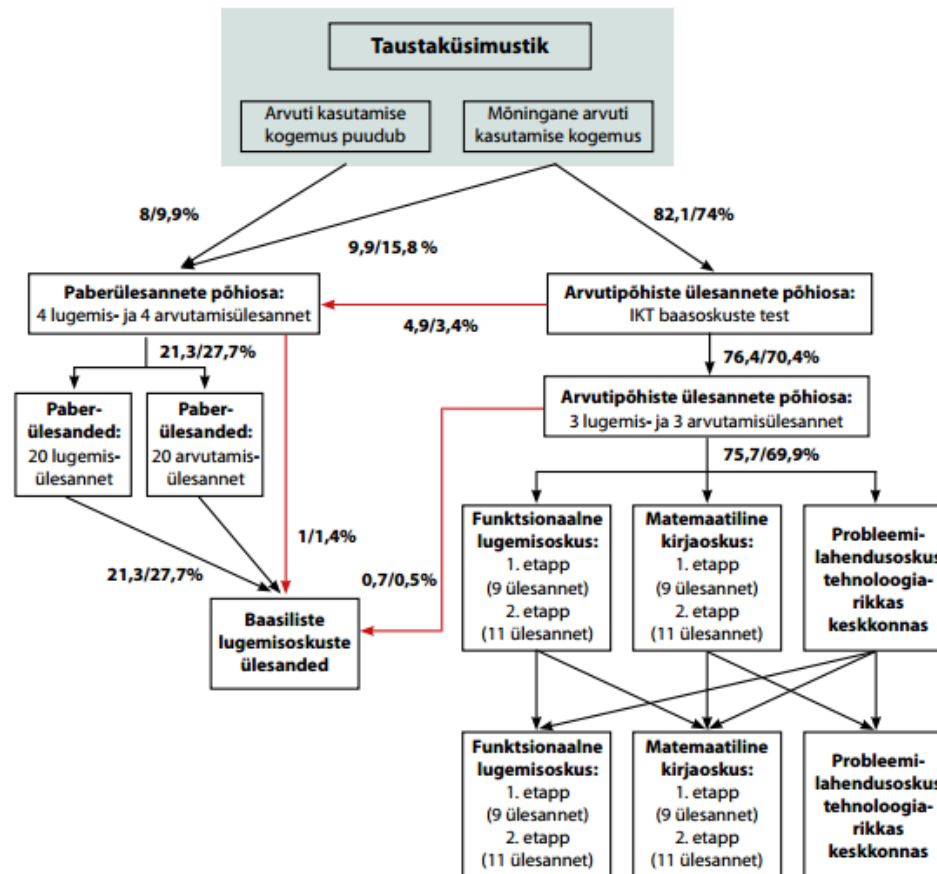
Valdkond 9 - Teenused

Milline on tehnoloogiarikas keskkonnas  
probleemilahendusoskuse tase Eestis, Poolas  
ja Slovakkias erinevates 10-aastastes  
vanusegruppides?

## Meid huvitavad tunnused

- Inimeste vanus AGE\_R, AGE5LFS, AGE10LFS
- Tehnoloogiarikkas keskkonnas probleemilahenduoskuse skoor PVPSL1...PVPSL10
- Midagi veel?

# Meeldetuletuseks



## Meid huvitavad tunnused

- Inimeste vanus AGE\_R, AGE5LFS, AGE10LFS
- Tehnoloogiarikkas keskkonnas probleemilahenduoskuse skoor PVPSL1...PVPSL10
- Midagi veel? Jah – PBROUTE!



## Oskuste kasutussageduse tunnuste kasutamine analüüsid:

Milline on IKT oskuste kasutussagedus erineva haridustasemega hõivatute hulgas Eesti, Poola ja Slovakkia tööturgudel?

## Meid huvitavad tunnused

- Hõivestaatus C\_D05
- IKT kasutussagedus tööol ICTWORK, ICTWORK\_WLE\_CA
- Midagi veel? Jah – G\_Q04 (Kas Te kasutate/kasutasite oma praeguses/viimases töös arvutit?)
- Haridustase – EDLEVEL3

## Palga seos kõrgeima omandatud haridustaseme ja matemaatilise kirjaoskuse tasemega

## Meid huvitavad tunnused

- Palk D\_Q16b, EARNHR, EARNHRDCL, EARNHRPPP, EARNHRBONUS, EARNHRBONUSDCL, EARNHRBONUSPPP, EARNMTH, EARNMTHPPP, EARNMTHBONUS, EARNMTHBONUSPPP, EARNMTHALL, EARNMTHALLDCL, EARNMTHALLPPP
- Kõrgeim omandatud haridustase
- Matemaatilise kirjaoskuse skoorid

## Palgatunnused PIAACis

	N	Min	Max	Keskmine	St. hälve
Current work - Earnings - Gross pay	3991	1,00	45000,0	769,6	1578,0
<b>Hourly earnings</b> <u>excluding</u> bonuses for wage and salary earners (derived)	4055	0,17	1621,8	5,8	27,6
<b>Hourly earnings</b> <u>excluding</u> bonuses for wage and salary earners, PPP corrected \$US (derived)	4055	0,32	2985,3	10,8	50,8
<b>Hourly earnings</b> <u>including</u> bonuses for wage and salary earners (derived)	4055	0,17	1621,8	6,0	27,7
<b>Hourly earnings</b> <u>including</u> bonuses for wage and salary earners, PPP corrected \$US (derived)	4055	0,32	2985,3	11,0	51,0
<b>Monthly earnings</b> <u>excluding</u> bonuses for wage and salary earners (derived)	4061	20,79	7222,2	885,2	687,3
<b>Monthly earnings</b> <u>excluding</u> bonuses for wage and salary earners, PPP corrected \$US (derived)	4061	38,27	13294,1	1629,4	1265,1
<b>Monthly earnings</b> <u>including</u> bonuses for wage and salary earners (derived)	4061	20,79	23469,6	909,9	793,0
<b>Monthly earnings</b> <u>including</u> bonuses for wage and salary earners, PPP corrected \$US (derived)	4061	38,27	43201,0	1675,0	1459,7
<b>Monthly earnings</b> for <b>self-employed</b> , PPP corrected \$US (derived)	355	0,00	138054,1	4169,5	12397,6
<b>Monthly earnings</b> <u>including</u> bonuses for <b>wage and salary earners and self-employed</b> (derived)	4416	0,00	75000,0	1018,9	2086,0
<b>Monthly earnings</b> <u>including</u> bonuses for <b>wage and salary earners and self-employed</b> , PPP corrected \$US (derived)	4416	0,00	138054,1	1875,5	3839,7

Aitäh!