



Funded by
the European Union

КЫРГЫЗ РЕСПУБЛИКАСЫНЫН
БИЛИМ БЕРҮҮ ЖАНА ИЛИМ
МИНИСТРЛИГИ



МИНИСТЕРСТВО
ОБРАЗОВАНИЯ И НАУКИ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ



THE WORLD BANK



Особенности обработки данных исследования PISA

Колачев Никита Игоревич

Иссык-Куль 2022

PISA



1. Источники специальной информации об исследовании PISA для психометриков.
2. Типы данных, предоставляемых исследованием PISA. Обзор баз данных.
3. Процент выполнения задания: теоретические аспекты, пример расчёта на основе данных компендиума, пример расчёта на основе базы данных, особенности расчёта для заданий адаптивного теста.
4. Работа с plausible values (PV): что это такое, как получаются, способы работы (вредные и хорошие советы)
5. Объединение разных баз данных.
6. Использование Data Explorer для расчёта описательных статистик.



PISA

1. Источники специальной информации об исследовании PISA для психометриков.

← ↻ 🏠 🔒 <https://www.oecd.org/pisa/> 🔍 ⚙️ 🌐 🛡️ ⚙️ ⚙️ 🗂️ 👤 ⋮



BETTER POLICIES FOR BETTER LIVES

Follow us







Programme for International Student Assessment



[Home](#) [About](#) [PISA Test](#) [Innovation](#) [Data](#) [Publications](#) [Webinars](#) [Join PISA](#) [FAQ](#)

[Français](#) [Deutsch](#) [Español](#)

What is PISA?

PISA is the OECD's Programme for International Student Assessment. PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges.



1. Источники специальной информации об исследовании PISA для психометриков.

Featured – PISA Results

PISA 2018

[What Students Know and Can Do \(Volume I\)](#)

[Where All Students Can Succeed \(Volume II\)](#)

[What School Life Means for Students' Lives \(Volume III\)](#)

[Executive Summaries of the Volumes I, II and III](#)

[PISA 2018: Insights and Interpretations](#)

[Are Students Smart about Money? \(Volume IV\)](#)

[Effective Policies, Successful Schools \(Volume V\)](#)

[Are Students ready to Thrive in an Interconnected World? \(Volume VI\)](#)

[Sky's the Limit: Growth mindset, students, and schools in PISA](#)

[21st-Century Readers: Developing Literacy Skills in a Digital World](#)

Previous PISA cycles

[2015](#) [2012](#) [2009](#) [2006](#) [2003](#) [2000](#)

Methodology & Manuals

Assessment and Analytical Frameworks

The theory underlying the PISA surveys – what they aim to achieve and how they are developed – including background questionnaires completed by students, school principals and parents.

[2025](#) [2022](#) [2018](#) [2015](#) [2012](#) [2009](#) [2006](#) [2003](#) [2000](#)

Technical Reports

Detailed descriptions of the implementation of the surveys.

[2018](#) [2015](#) [2012](#) [2009](#) [2006](#) [2003](#) [2000](#)

Survey Implementation Tools

Field trial materials, translation manuals and technical guidelines used to carry out the surveys.

[2022](#) [2018](#) [2015](#) [2012](#)

1. Источники специальной информации об исследовании PISA для психометриков.

Chapter 1: Programme for International Student Assessment - an Overview

Chapter 2: Test Design and Development (text), Tables and Figures

Chapter 3: Context Questionnaire Development

Chapter 4: Sample Design

Chapter 5: Translation

Chapter 6: Survey Operations

Chapter 7: Quality Control

Chapter 8: Survey Weighting (text), Tables

Chapter 9: Scaling PISA Data (text), Tables, Figures

Chapter 10: Data Management Procedures

Chapter 11: Sampling Outcomes (text), Tables

Chapter 12: Scaling Outcomes (text), Tables, Figures

Chapter 13: Coding Reliability

Chapter 14: Data Adjudication (text), Figure

Chapter 15: Proficiency Scale Construction

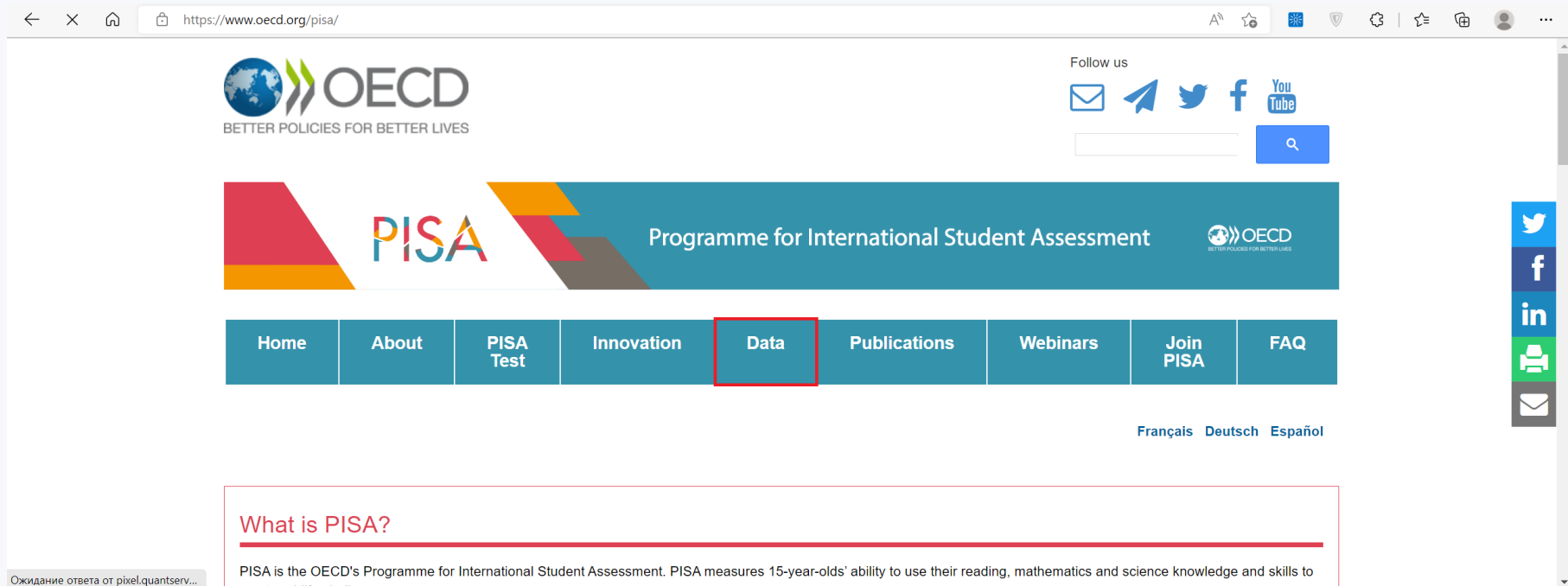
Chapter 16: Scaling procedures and construct validation of context questionnaire data (text), Tables

Chapter 17: Questionnaire Design and Computer-based Questionnaire Platform (text), Table

Chapter 18: Computer-based Platform

- Annex A: Item Pool Classification
- Annex B: Contrast Coding
- Annex C: Standard Errors of Means, Sample Sizes, School Variance Estimates and Other Sampling Outcomes
- Annex D: Mapping of ISCED Levels to Years
- Annex E: National Household Possessions
- Annex F: Final Distribution of RMSD Values Across Groups for Each Scale
- Annex G: Common and Unique Item Parameters in Each Domain, by Country and Language
- Annex H: New Procedures for PISA 2018 Population Modelling
- Annex I: PISA 2018 Technical Standards
- Annex J: The Development and Implementation of PISA: A Collaborative Effort
- Annex K: Uses and Reporting of Process Data
- Annex L: Equated P+
- Annex M: PISA 2018 Main Survey Design for Ultra-Orthodox Schools in Israel

1. Источники специальной информации об исследовании PISA для психометриков.



The screenshot shows the official PISA website (https://www.oecd.org/pisa/). The header features the OECD logo with the tagline "BETTER POLICIES FOR BETTER LIVES" and a "Follow us" section with icons for email, Telegram, Twitter, Facebook, and YouTube. Below this is a search bar. The main banner displays the PISA logo and the text "Programme for International Student Assessment" alongside the OECD logo. A navigation menu contains links for Home, About, PISA Test, Innovation, Data (highlighted with a red box), Publications, Webinars, Join PISA, and FAQ. Language options for Français, Deutsch, and Español are provided. The "What is PISA?" section is partially visible at the bottom, stating: "PISA is the OECD's Programme for International Student Assessment. PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real life challenges."



1. Источники специальной информации об исследовании PISA для психометриков.

PISA Database

The PISA database contains the full set of responses from individual students, school principals and parents. These files will be of use to statisticians and professional researchers who would like to undertake their own analysis of the PISA data. The files available on this page include background questionnaires, data files in ASCII format (from 2000 to 2012), codebooks, compendia and SAS™ and SPSS™ data files in order to process the data.

[2018](#) [2015](#) [2012](#) [2009](#) [2006](#) [2003](#) [2000](#)



1. Источники специальной информации об исследовании PISA для психометриков.

Questionnaires

Below are the master questionnaires administered to students who sat the computer-based assessment; each country translated these questionnaires into their own language(s). These master questionnaires are also available in [French](#). Students who sat the paper-based assessment also completed questionnaires that can be found via the following links: [English](#) and [French](#).

- [Student questionnaire](#) administered to students
- [Student questionnaire \(*une heure* version\)](#) administered to low-performing students
- [Well-being questionnaire](#) administered to students
- [Educational career questionnaire](#) administered to students
- [Financial literacy questionnaire](#) administered to students
- [ICT familiarity questionnaire](#) administered to students
- [School questionnaire](#) administered to school principals
- [Teacher questionnaire](#) administered to teachers of the test language
- [Teacher questionnaire](#) administered to other teachers
- [Parent questionnaire](#) administered to parents (via paper)

Codebook and Compendia

Code book

Compendia

- [Questionnaire items](#) (42.8 MB)
- [Questionnaire items](#) (Financial literacy - 12.8MB)
- [Cognitive items](#) (19.3 MB)
- [Cognitive items](#) (Financial literacy - 4.3 MB)

1. Источники специальной информации об исследовании PISA для психометриков.

SAS (TM) Data Files (compressed)

Databases including Global Competence variables were uploaded in October 2020.

- [Student questionnaire data files](#) (597 MB)
- [School questionnaire data file](#) (3.7 MB)
- [Teacher questionnaire data file](#) (17.3 MB)
- [Cognitive item data file](#) (584 MB)
- [Moscow City data file](#) (23.6 MB)
- [Questionnaire timing data files](#) (184 MB)
- [Additional data file for Viet Nam*](#) (PVs and cognitive) (2.1 MB)
- [Financial literacy data file](#) (249 MB)
- [Cognitive items total time/visits data file](#) (193 MB) *See PISA 2018 Technical Report Annex K for details*

SPSS (TM) Data Files (compressed)

Databases including Global Competence variables were uploaded in October 2020.

- [Student questionnaire data file](#) (489 MB)
- [School questionnaire data file](#) (3.1 MB)
- [Teacher questionnaire data file](#) (12.8 MB)
- [Cognitive item data file](#) (466 MB)
- [Moscow City data file](#) (20 MB)
- [Questionnaire timing data files](#) (188 MB)
- [Additional data file for Viet Nam*](#) (PVs and Cognitive) (1.1 MB)
- [Financial literacy data file](#) (206 MB)
- [Cognitive items total time/visits data file](#) (180 MB) *See PISA 2018 Technical Report Annex K for details*

* Due to issues identified during data adjudication, caution is required when analysing these data. For further information, see [Annexes A4 and A6](#) of [PISA 2018 Results \(Volume I\): What Students Know and Can Do](#) and Chapter 14 of [PISA 2018 Technical Report](#).

Compendia

- Предобработанная информация (описательного характера)

Data Files

- Самое интересное для психометрика!!!

Базы данных содержатся отдельно:

- Данные анкетирования учащихся (с PV)
- Данные анкетирования образовательной организации
- Данные анкетирования учителей
- Данные о выполнении заданий тестов
- Данные по отдельным выборкам исследований
- Мета-данные, лог-файлы и пр.

Процент выполнения задания: теоретические аспекты, пример расчёта на основе данных компендиума, пример расчёта на основе базы данных, особенности расчёта для заданий адаптивного теста.

- Для дихотомических заданий:

$$P_{jd} = \frac{n_j}{N}$$

где n_1 - число учащихся с оценкой 1 на задание j , N – общее число учащихся.

- Для политомических заданий:

$$P_{jp} = \frac{\sum_{m=1}^l m_j \cdot p_{jm}}{k_j}$$

где m_j - балл политомической шкалы оценивания, p_{jm} - вероятность получения m -го балла в j -ом задании, k_j - макс. балл по заданию.

Процент выполнения задания: теоретические аспекты, пример расчёта на основе данных компендиума, пример расчёта на основе базы данных, особенности расчёта для заданий адаптивного теста.

- В исследовании 2018 года для измерения читательской грамотности было применено адаптивное тестирование. Оно отличается от традиционного подхода тем, что подбирает наиболее оптимальные задания для учащихся разного уровня подготовленности.
- Таким образом, нарушается требование случайности предъявления заданий и репрезентативности усреднённых значений. Следовательно, традиционный подход к расчёту трудности задания не применим. Нужно пользоваться готовыми расчётами (P+Equated).

Более подробная информация: [Ch.09-Scaling-PISA-Data.pdf \(oecd.org\)](https://www.oecd.org/pisa/data/pisa2018-data/ch09-scaling-pisa-data.pdf)

Работа с plausible values (PV): что это такое, как получаются, способы работы (вредные и невредные советы).

- Правдоподобные значения (plausible values) - это специально созданные значения, которые похожи на индивидуальные результаты тестирования и имеют приблизительно такое же распределение, как и измеряемый латентный признак.
- Правдоподобные значения были разработаны в качестве вычислительного приближения для получения согласованных оценок характеристик популяции в ситуациях оценки, когда индивидам предлагается слишком мало заданий, чтобы можно было точно оценить их способности.
- Правдоподобные значения представляют собой случайные выборки из эмпирически полученного распределения результатов тестирования, которые зависят от наблюдаемых значений результатов выполнения заданий и контекстных переменных.



Работа с plausible values (PV): что это такое, как
получаются, способы работы (вредные и невредные
советы).

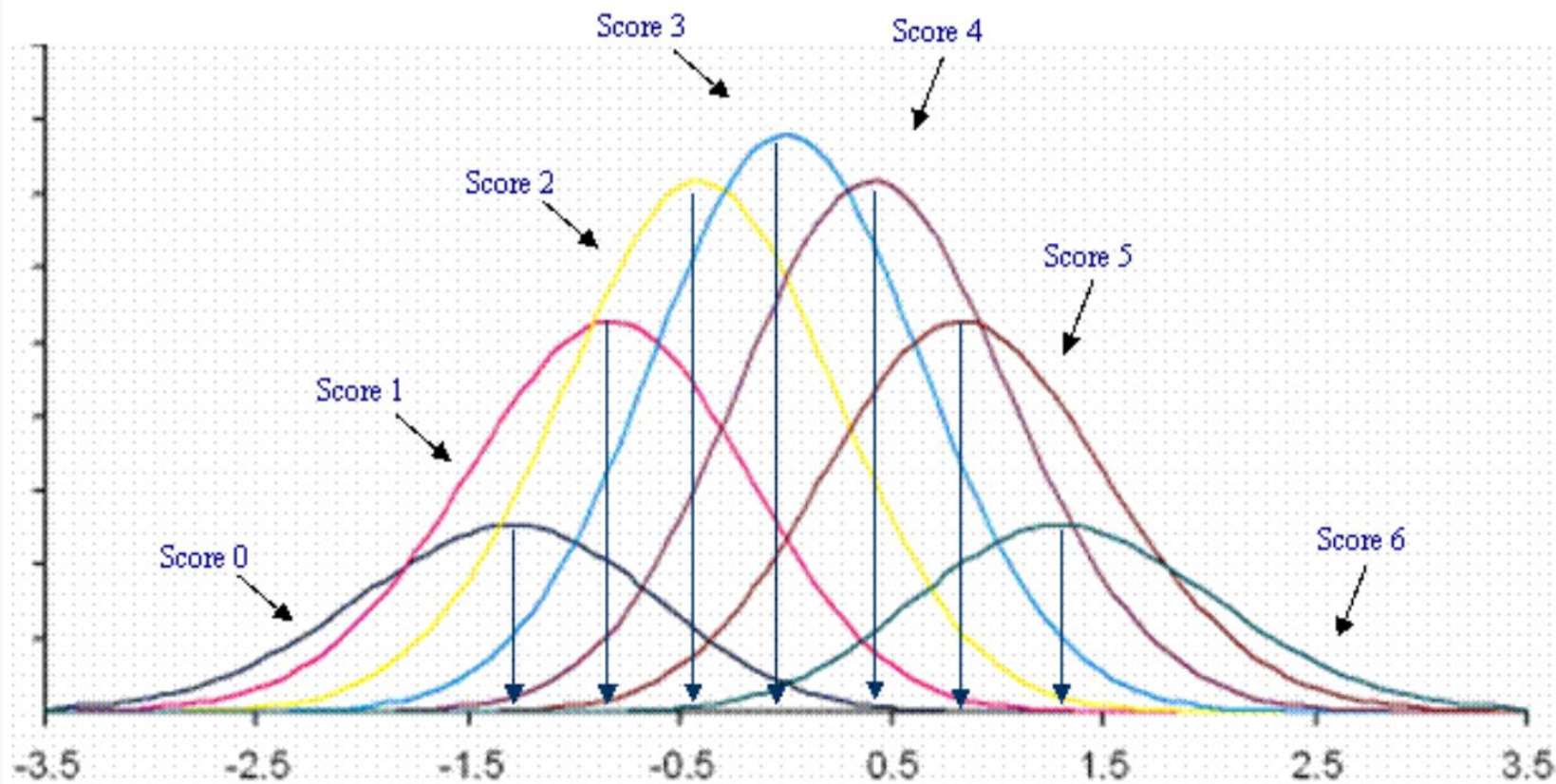


Figure: Proficiency on Logit Scale

Работа с plausible values (PV): что это такое, как
получаются, способы работы (вредные и невредные
советы).

Советы по работе с plausible values (PV).

Вредные (-):

1. Усреднить все PV в базе данных и работать с усреднённым показателем
2. Брать только одно PV из 10 и работать с ним

Хорошие (+):

1. Использовать методы множественной импутации (multiple imputation)
2. Использовать специальное ПО для работы с данными международных сравнительных исследований (Data Explorer, IDB Analyzer, специальные пакеты для программы R, Stata и др.)