



Математическая модель расчета индексов субъективного благополучия и их визуализация

Для оценки СБ использовалась «Операциональная модель субъективного благополучия подростков РФ», которая включает 12 субдоменов СБ: 1 – Удовлетворенность своими достижениями; 2 – Удовлетворенность собой; 3 – Удовлетворенность здоровьем и самочувствием; 4 – Удовлетворенность подростка собственной психологической и физической безопасностью; 5 – Удовлетворенность учетом мнения подростков при решении вопросов, затрагивающих его интересы; 6 – Удовлетворенность внутренней поддерживающей сетью; 7 – Удовлетворенность внешней сетью; 8 – Удовлетворенность «нормализацией жизни»; 9 – Удовлетворенность материальным положением своим и своей семьи; 10 – Удовлетворенность социальными контекстами; 11 – Удовлетворенность своим хронотопом (своим прошлым, настоящим и будущим), 12 – Удовлетворенность возможностью доверять себе и другому.

Данные по стрессам содержатся в субдомене «Стрессы» и рассматриваются отдельно по запросу.

В таблице 1 приводятся обозначения, определения, размерности числовых характеристик субдоменов СБ (субъективное благополучие детей).

Таблица 1. Числовые характеристики субъективного благополучия детей (СБ).

Субдомены субъективного благополучия		Числовые характеристики СубД		
Имена субдоменов	Сокращенные имена субдоменов	Средние по всей базе, μ	Средние по текущей выборке, η	Относительное благополучие $\delta\eta = (\eta - \mu) / \mu$
Удовлетворенность ↓	Удовлетворенность ... ↓	в баллах	в баллах	в % (процентах)
Своими достижениями	Достижениями	μ_1	η_1	$\delta\eta_1 = (\eta_1 - \mu_1) / \mu_1$
Собой	Собой	μ_2	η_2	$\delta\eta_2 = (\eta_2 - \mu_2) / \mu_2$
Здоровьем и самочувствием	Здоровьем и самочувствием	μ_3	η_3	$\delta\eta_3 = (\eta_3 - \mu_3) / \mu_3$
Подростка собственной психологической и физической безопасностью	Безопасностью	μ_4	η_4	$\delta\eta_4 = (\eta_4 - \mu_4) / \mu_4$



Учетом мнения подростков при решении вопросов, затрагивающих его интересы	Учетом мнения	μ_5	η_5	$\delta\eta_5 = (\eta_5 - \mu_5) / \mu_5$
Субдомены субъективного благополучия		Числовые характеристики СубД		
Внутренней поддерживающей сетью	Внутренней сетью	μ_6	η_6	$\delta\eta_6 = (\eta_6 - \mu_6) / \mu_6$
Внешней сетью	Внешней сетью	μ_7	η_7	$\delta\eta_7 = (\eta_7 - \mu_7) / \mu_7$
Нормализацией жизни	Нормализацией жизни	μ_8	η_8	$\delta\eta_8 = (\eta_8 - \mu_8) / \mu_8$
Материальным положением своим и своей семьи	Материальным положением	μ_9	η_9	$\delta\eta_9 = (\eta_9 - \mu_9) / \mu_9$
Социальными контекстами	Социальными контекстами	μ_{10}	η_{10}	$\delta\eta_{10} = (\eta_{10} - \mu_{10}) / \mu_{10}$
Своим хронотопом (своим прошлым, настоящим и будущим)	Хронотопом	μ_{11}	η_{11}	$\delta\eta_{11} = (\eta_{11} - \mu_{11}) / \mu_{11}$
Доверием	Доверием	μ_{12}	η_{12}	$\delta\eta_{12} = (\eta_{12} - \mu_{12}) / \mu_{12}$
Интегральные статистические величины				
Средние по субдоменам	$\langle \delta\eta \rangle = \sum_k \delta\eta_k / 12$ – сумма по всем 12-ти субдоменам (сумма по всем k)			
Дисперсия, $D\delta\eta$	$D_{\delta\eta} = \sum_k (\delta\eta_k - \langle \delta\eta \rangle)^2 / 12$ (сумма по всем k); $D_{\delta\eta} = \langle \delta\eta^2 \rangle - \langle \delta\eta \rangle^2$			
Среднеквадратичное (стандартное) отклонение, $\sigma\delta\eta$	$\sigma_{\delta\eta} = D_{\delta\eta}^{1/2}$ – корень квадратный из дисперсии			
Z-оценка, $Z\delta\eta$	$Z_k = (\delta\eta_k - \langle \delta\eta \rangle) / \sigma_{\delta\eta}$ – отклонения $\delta\eta_k$ от его среднего в единицах $\sigma_{\delta\eta}$			
Условия нормировки Z-оценок	$\langle Z \rangle = \sum_k Z_k / 12 = 0$; $\sigma_z = 1$			

Для оценки отклонений субдоменов СБ от среднего значения использовалась Z-оценка.

Описание информационной системы СБ

Первый этап – сбор и предварительная обработка информации. База данных является основой системы. Исходными данными для ее формирования являются ответы на вопросы тестов, опросников и разного вида психологических методик. Сами вопросы с выпадающими списками ответов или высылаются респондентам по электронной почте, или размещаются на сервисах для создания опросов (тестов) и сбора результатов («Анкетолог»). Ответы поступают аналогично – либо по электронной почте, либо выгружаются с платформы. В обоих случаях это многочисленные таблицы Excel в форматах .xls или .xslm. Все полученные данные заносятся в единую базу данных (далее – база).



Следующий этап – первоначальная проверка данных, требует долгой и кропотливой работы над полученной базой. Для ускорения этой работы была разработана специальная программа, указывающая на ошибки ввода, то есть на несоответствие ответов подростков ответам, заложенным в выпадающих списках. Как правило, отсеву подлежит до 5–7% анкет опрошенных.

Следующий этап полностью автоматизирован и решает задачу преобразования ответов в соответствующие им баллы (5–балльная система Лайкерта).

Таким образом, после указанных этапов получают две базы данных: текстовая и соответствующая ей числовая в баллах.

База данных ответов обрабатывается с целью получения статистики распределения респондентов по регионам, а также статистики каждого ответа на каждый вопрос опроса в процентах (доля общего числа респондентов, ответивших на данный вопрос конкретным ответом). На этой основе возможно построение многочисленного количества диаграмм.

Первоначальная обработка полученных данных заканчивается передачей базы данных баллов в систему для статистической обработки.

Система хранения и обработки данных домена СБ была создана в виде электронной таблицы Microsoft Excel 2021 по следующим причинам:

- доступность для любого пользователя, возможность перенесения базы данных с одного компьютера на другой;
- современный Excel позволяет создавать относительно большие таблицы (базы данных) размером до $2^{20} \sim 1.000.000$ строк (записей) и до $2^{14} \sim 16.000$ столбцов (полей);
- электронная таблица Excel имеет большое количество встроенных функций, в том числе статистических и позволяет строить графики любой степени сложности;
- система автофильтров Excel позволяет делать любые выборки из базы даже малоподготовленным пользователям.

Все результаты проходят проверку значимости на уровне надежности $P = 95\%$, поэтому их можно в дальнейшем обрабатывать в любых программах, например, SPSS или Jamovi.

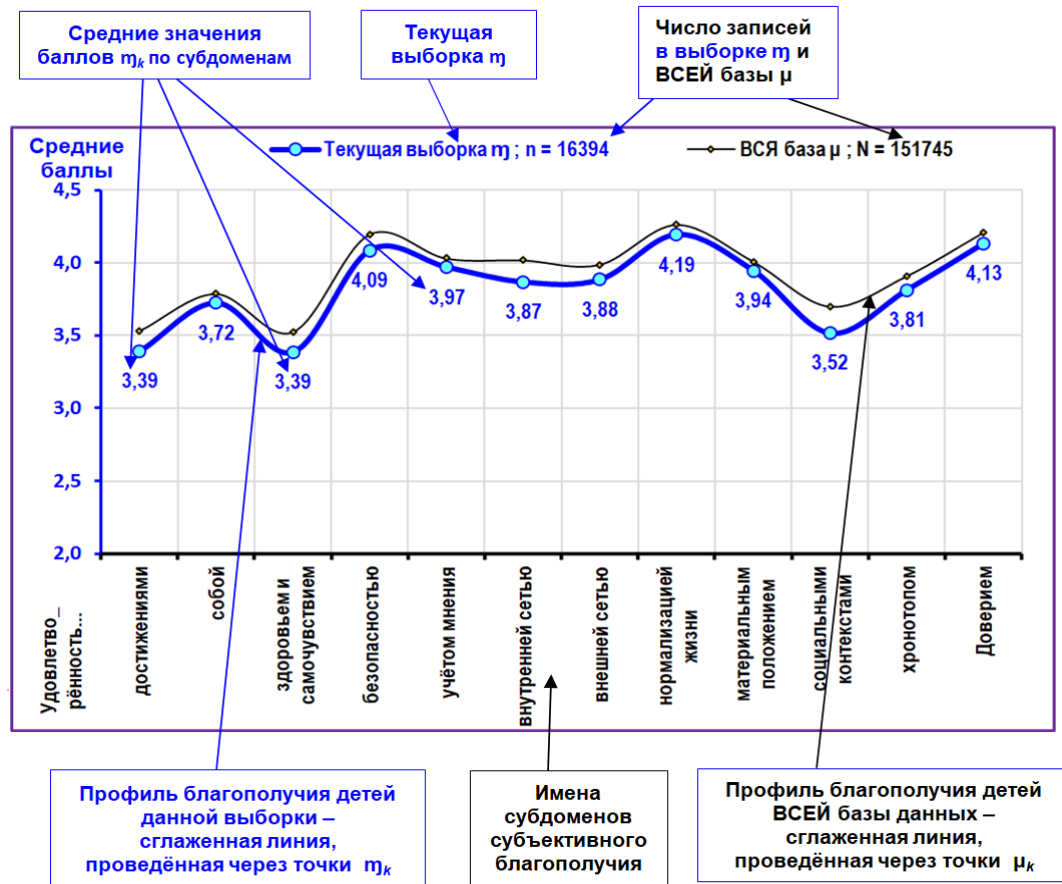
Созданная система хранения и обработки данных является минимально достаточным инструментом объективного измерения индексов субдоменов СБ детей.

Визуализация результатов расчетов представляется в виде различных профилей и Z- оценок.

На рисунке 1 приводится диаграмма профилей СБ всех детей в базе данных (черная линия) и профиль СБ для любой выбранной из базы категории детей, например, мальчики или дети-мигранты (синяя линия), которая автоматически получается после каждого включения любого фильтра.



Рисунок 1. Профиль благополучия в баллах (от 1 до 5) по субдоменам СБ



Зная точку отсчета (вся база), для каждого субдомена k может быть рассчитано относительное благополучие $\delta\eta_k = (\eta_k - \mu_k) / \mu_k$. Эти величины безразмерны и могут быть выражены в процентах. Они показывают, на сколько процентов они отклоняются от средних по всей базе (черная линия на рисунке 2).

Пример такого профиля приводится на рисунке 2, где средняя (красная) линия – это усредненное значение по всем 12-ти субдоменам $\langle \delta\eta \rangle = \sum \delta\eta_k / 12$. Также рассчитывается среднеквадратичное (стандартное) отклонение σ .

Эти две интегральные величины являются основными параметрами, с помощью которых строятся Z-оценки.

Этот параметр определяется как отношение разности величины x и ее средней $\langle x \rangle$ на σ : $Zx = s \cdot (x - \langle x \rangle) / \sigma$ (1)

где x – случайная величина,

$\langle x \rangle$ – среднее всех x_k ,

$\sigma = Dx^{1/2}$ – среднеквадратическое (стандартное) отклонение,

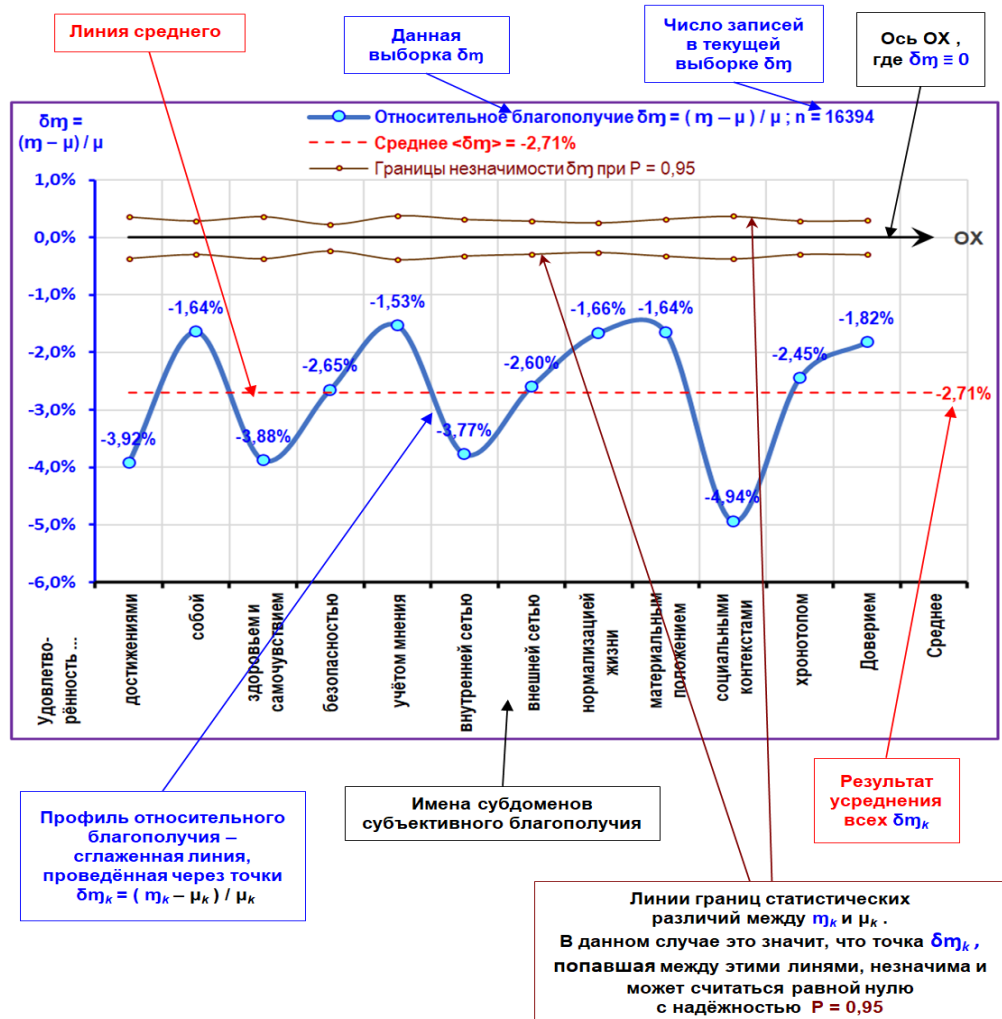
$Dx = \sum (x_k - \langle x \rangle)^2 / N = \langle x^2 \rangle - \langle x \rangle^2$ – дисперсия,

$\langle x^2 \rangle = \sum x_k^2 / N$ – средний квадрат x по всем измерениям k .

Из определения (1) следует, что величина Zx безразмерна и может интерпретироваться, как отклонение величины x от среднего в единицах σ .



Рисунок 2 – Профиль относительного субъективного благополучия $\delta\eta = (\eta - \mu) / \mu$ по субдоменам СБ



Для расчета Z-оценок по формуле (1) статистические данные x_k должны иметь, во-первых, одинаковые единицы измерения (например, в виде десятичных дробей, процентов, промилле и так далее) и, во-вторых, их масштаб должен допускать их простое сравнение друг с другом. На рисунке 3 даны пояснения к диаграмме Z-оценок по величинам параметров субдоменов, где каждому субдомену соответствует знак его величины (уровня): Н – высокий, М – средний и L – низкий.



Рисунок 3. Диаграмма Z-оценок по величинам параметров субдоменов СБ



Алгоритмы расчетов индексов СБ в разрезах регионов и субдоменов

Все изложенное выше имело отношение к визуализации числовых характеристик в разрезе субдоменов относительно одной из групп детей, например, детей-сирот, девушек, юношей, детей-инвалидов и так далее.

Рассмотрим более общий случай, когда таких групп несколько, например, детей, проживающих в различных регионах (были выбраны регионы, где проживает более тысячи детей из нашей базы данных (см. Табл. 2).

Как видно из рисунка 4, где приводится блок-схема алгоритмов получения итоговых индексов СБ в разрезе регионов и субдоменов, число регионов не ограничено, как не ограничено и число субдоменов.

Необходимо отметить, что вся система обработки работает автоматически при любых изменениях базы данных, например, включении фильтров по любому набору полей. Так переключение фильтров с одного региона на другой позволяет получить средние баллы субдоменов j по отфильтрованным регионам m_{ij} и заполнить столбцами чисел таблицу блока 3 (рис.4). При этом, если выключаются все фильтры базы данных, то получаем общие средние баллы μ_j по всей базе (блок 2 на рис.4), относительно которых затем рассчитывается относительное благополучие детей $\delta \eta_{ij} = (m_{ij} - \mu_i) / \mu_i$ в блоке 4 (рис.4).



Индекс детского благополучия. Математическая модель субъективного индекса.

Таблица 2. Относительное субъективное благополучие $\delta\eta_i$ по субдоменам i в региональном разрезе ($n > 1\,000$ чел.)

Относительное СБ Субдоменов в разрезе регионов	Республика Башкортостан	Чеченская Республика	Приморский край	Московская область	Новгородская область	Омская область	Рязанская область	Сахалинская область	Тюменская область	Ульяновская область	Москва	ДНР	Среднее, < $\delta\eta$ >	Сигма, $\sigma_{\delta\eta}$
Удовлетворенность ↓	54 234	1 117	16 394	4 157	2 629	6 343	1 084	1 993	25 428	6 725	16 229	13 341	← объемы выборок	
достижениями	0,8%	4,8%	-3,9%	-0,7%	1,3%	-2,0%	0,9%	-4,6%	1,1%	-1,1%	3,0%	-2,1%	-0,21%	2,62%
собой	0,2%	0,1%	-1,6%	-2,4%	1,3%	-1,6%	1,7%	-3,5%	1,9%	0,2%	0,7%	-1,6%	-0,38%	1,65%
здоровьем и самочувствием	2,6%	3,7%	-3,9%	-5,3%	1,0%	-3,8%	-0,4%	-7,1%	1,9%	1,3%	-4,1%	-1,2%	-1,28%	3,34%
безопасностью	1,1%	3,2%	-2,7%	-3,7%	1,5%	-0,2%	0,1%	-4,3%	1,7%	0,9%	-1,6%	-1,5%	-0,46%	2,22%
учетом мнения	-1,3%	-3,5%	-1,5%	-1,7%	2,5%	-1,9%	2,5%	-2,5%	2,7%	0,5%	3,8%	-0,8%	-0,10%	2,31%
внутренней сетью	1,9%	4,6%	-3,8%	-3,1%	1,0%	-2,1%	1,2%	-6,2%	0,7%	0,9%	-1,0%	-1,3%	-0,60%	2,80%
внешней сетью	0,0%	4,2%	-2,6%	-1,4%	2,0%	-0,2%	1,2%	-3,7%	1,5%	0,1%	1,8%	-1,5%	0,12%	2,10%
Нормализацией жизни	0,3%	-2,6%	-1,7%	-2,2%	1,2%	-2,4%	1,5%	-2,9%	1,2%	1,1%	0,8%	-0,5%	-0,52%	1,65%
Материальным положением	0,0%	-0,1%	-1,6%	-0,9%	0,1%	-1,8%	0,3%	-3,2%	1,1%	-0,1%	2,0%	-0,4%	-0,38%	1,31%
Социальными контекстами	1,2%	4,5%	-4,9%	-1,8%	3,5%	0,4%	0,5%	-6,4%	0,9%	0,6%	1,3%	-1,9%	-0,18%	3,01%
хронотопом	1,0%	2,0%	-2,5%	-1,6%	0,7%	-2,8%	1,2%	-4,4%	1,1%	0,3%	1,0%	-2,3%	-0,53%	1,98%
Доверием	0,2%	2,8%	-1,8%	-1,5%	0,9%	-1,2%	0,9%	-3,6%	1,7%	0,2%	0,9%	-1,6%	-0,18%	1,72%
Среднее, < $\delta\eta$ >	0,67%	1,98%	-2,71%	-2,19%	1,42%	-1,63%	0,97%	-4,37%	1,46%	0,41%	0,72%	-1,39%	-0,39%	2,23%
Сигма, $\sigma_{\delta\eta}$	0,97%	2,74%	1,10%	1,24%	0,86%	1,14%	0,74%	1,41%	0,53%	0,61%	2,03%	0,57%	0,33%	0,59%



Рисунок 4. Блок схема расчетов индексов СБ в разрезах регионов и субдоменов

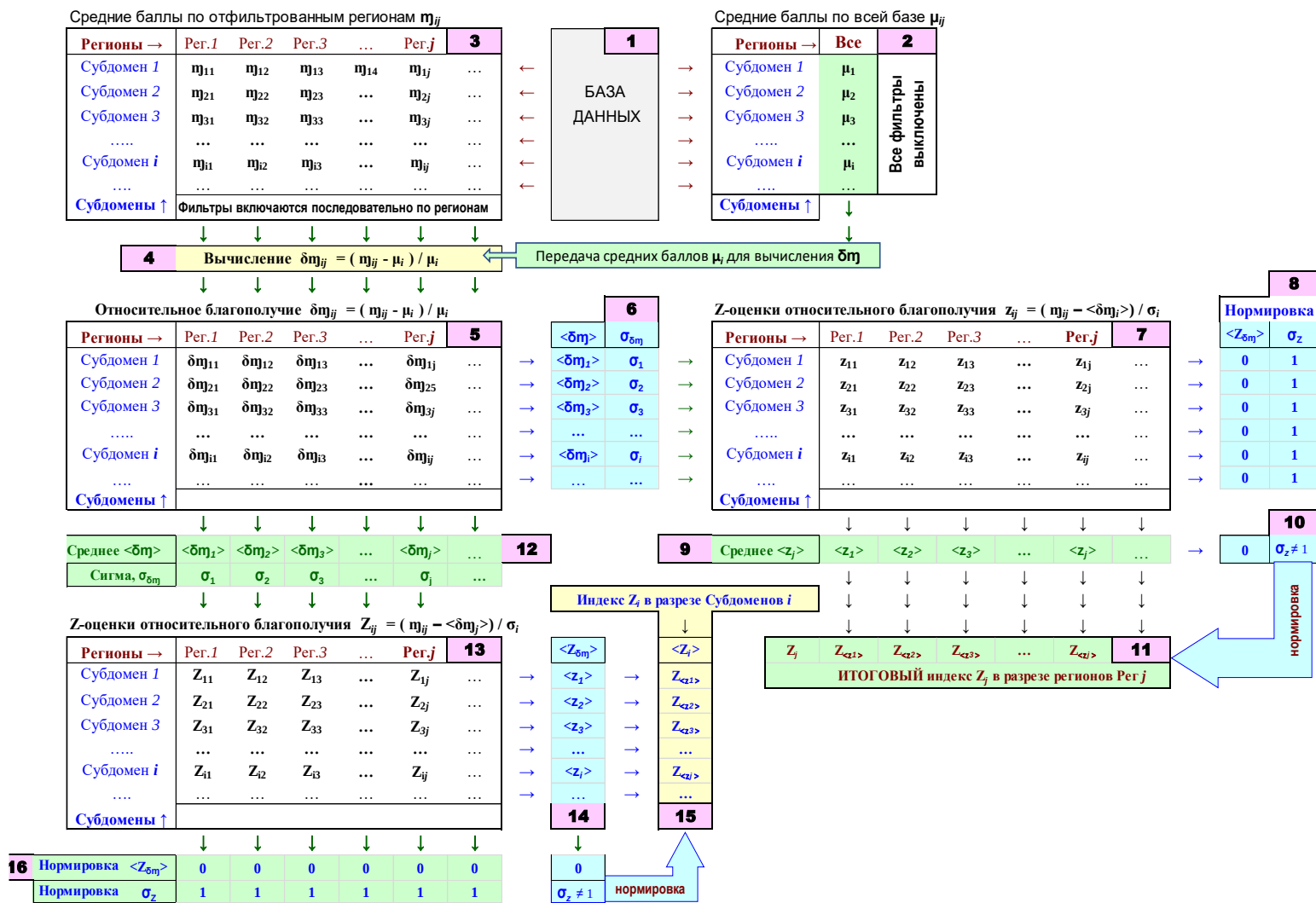




Таблица 3. Z- оценки относительного субъективного благополучия $Z_{\delta mi}$ по субдоменам i в региональном разрезе ($n > 1\,000$ чел.)

Z-оценки Относительных СБСубдоменов в разрезе регионов	Республика Башкортостан	Чеченская Республика	Приморский край	Московская область	Новгородская область	Омская область	Рязанская область	Сахалинская область	Тюменская область	Ульяновская область	Москва	ДНР	Средние, < Z >	Сигмы, σ_z
	54 234	1 117	16 394	4 157	2 629	6 343	1 084	1 993	25 428	6 725	16 229	13 341	← объемы выборок	
Достижениями	0,38	1,91	-1,41	-0,19	0,58	-0,68	0,42	-1,68	0,50	-0,34	1,22	-0,72	0,00	1,00
Собой	0,35	0,29	-0,74	-1,22	1,02	-0,74	1,26	-1,89	1,38	0,35	0,66	-0,74	0,00	1,00
Здоровьем и самочувствием	1,16	1,49	-0,79	-1,21	0,68	-0,76	0,26	-1,74	0,95	0,77	-0,85	0,02	0,00	1,00
Безопасностью	0,70	1,65	-1,01	-1,46	0,88	0,12	0,25	-1,73	0,97	0,61	-0,51	-0,47	0,00	1,00
Учетом мнения	-0,52	-1,47	-0,61	-0,69	1,13	-0,78	1,13	-1,04	1,21	0,26	1,69	-0,30	0,00	1,00
Внутренней сетью	0,89	1,86	-1,14	-0,89	0,57	-0,54	0,64	-2,00	0,47	0,54	-0,14	-0,25	0,00	1,00
Внешней сетью	-0,06	1,94	-1,29	-0,72	0,90	-0,15	0,51	-1,81	0,66	-0,01	0,80	-0,77	0,00	1,00
Нормализацией жизни	0,49	-1,26	-0,72	-1,02	1,04	-1,14	1,22	-1,44	1,04	0,98	0,80	0,01	0,00	1,00
Материальным положением	0,29	0,22	-0,93	-0,39	0,37	-1,08	0,52	-2,15	1,13	0,22	1,82	-0,01	0,00	1,00
Социальными контекстами	0,46	1,55	-1,57	-0,54	1,22	0,19	0,22	-2,07	0,36	0,26	0,49	-0,57	0,00	1,00
Хронотопом	0,77	1,27	-1,00	-0,54	0,62	-1,15	0,87	-1,95	0,82	0,42	0,77	-0,89	0,00	1,00
Доверием	0,22	1,73	-0,95	-0,77	0,63	-0,60	0,63	-2,00	1,09	0,22	0,63	-0,83	0,00	1,00
Средние, < Z >	0,43	0,93	-1,01	-0,80	0,80	-0,61	0,66	-1,79	0,88	0,36	0,61	-0,46	0,00	0,86
Итоговые Z-оценки (Индексы СБ в разрезе регионов)													Нормировки	
	0,50	1,08	-1,18	-0,93	0,93	-0,71	0,77	-2,08	1,02	0,41	0,71	-0,54	0,00	1,00

Результаты этих расчетов ложатся в таблицу блока 5 (рис.4) «Относительное благополучие $\delta\pi_{ij} = (\pi_{ij} - \mu_i) / \mu_i$ ».

Теперь возможно решение двух задач: 1 – Рассчитать индексы СБ в разрезе регионов; 2 – Рассчитать индексы СБ в разрезе субдоменов.

Первая задача – вычислим среднее относительное благополучие $\langle \delta\pi_i \rangle$ и среднеквадратичное отклонение σ_i для каждого субдомена, то есть усреднение идет по строкам таблицы блока 5 (табл. 2 и рис.4). Зная среднее $\langle \delta\pi_i \rangle$ и среднеквадратичное σ_i , считаем теперь z-оценки относительного благополучия $z_{ij} = (\pi_{ij} - \langle \delta\pi_i \rangle) / \sigma_i$, заполняя таблицу блока 7 (рис.4 и табл. 3).

Проверяем условия нормировки, считая построчно для каждого субдомена i среднее $\langle z_i \rangle$ и среднеквадратичное σ_{z_i} , как и должно, оказываются равными 0 и 1, соответственно (рис.4). Теперь, поскольку нас интересуют индексы СБ в разрезе регионов, а распределение их по субдоменам отсутствует, то усредним теперь z-оценки относительного благополучия z_{ij} по субдоменам j , то есть по столбцам сверху вниз, заполняя этими средними $\langle z_j \rangle$ строку блока (рис.4.). Проверая нормировки, выясняем, что среднее строки $\langle z_j \rangle$, как и должно, равно нулю, а среднеквадратичное $\sigma_{\langle z_j \rangle}$ не равно единице, что вызывает определенные проблемы (в конкретном численном примере Таблицы 3 $\sigma_{\langle z_j \rangle} \approx 0,86 \neq 1$).

Как показывает математический расчет, нормировка строки $\langle z_j \rangle$ элементарна – для этой нормировки достаточно каждый элемент строки $\langle z_j \rangle$ поделить $\sigma_{\langle z_j \rangle}$. Полученная строка и будет итоговой z-оценкой или Индексами СБ в разрезе регионов.

Полученные результаты приводятся в таблицах 4 и 5, и визуализируются диаграммой на рисунке 5.

Таблица 4. Общие Z_{tot} -оценки в разрезе регионов (Индексы СБ)

Регионы ↓	Z_{tot}
Республика Башкортостан	0,50
Чеченская Республика	1,08
Приморский край	-1,18
Московская область	-0,93
Новгородская область	0,93
Омская область	-0,71
Рязанская область	0,77
Сахалинская область	-2,08
Тюменская область	1,02
Ульяновская область	0,41
Москва	0,71
Донецкая Народная Республика	-0,54
Нормировка: Среднее, $\langle Z_{tot} \rangle$	0,00
Нормировка: Сигма, $\sigma_{Z_{tot}}$	1,00



Таблица 5. Общие Z_{tot} -оценки в разрезе регионов, сортированные по z-оценкам (Индексы СБ)

Регионы ↓	Z _{tot}
Чеченская Республика	1,08
Тюменская область	1,02
Новгородская область	0,93
Рязанская область	0,77
Москва	0,71
Республика Башкортостан	0,50
Ульяновская область	0,41
Донецкая Народная Республика	-0,54
Омская область	-0,71
Московская область	-0,93
Приморский край	-1,18
Сахалинская область	-2,08
Нормировка: Среднее, <Z _{tot} >	0,00
Нормировка: Сигма, σ _{Ztot}	1,00

Рисунок 5 – Z-оценки относительного субъективного благополучия в региональном разрезе ($n > 1\,000$ чел.)



Рассмотрим теперь кратко поблочный алгоритм получения Индексов СБ в разрезе субдоменов. Принципиально общая логика здесь мало отличается от той, что описана выше, за исключением характера усреднений при обработке исходного блока 5 (рис.4.) и далее по блокам 12, 13, 14, 15 (рис.4.). Там где раньше шла обработка по строкам (по горизонтали), будет обработка по столбцам, то есть по вертикали и наоборот. То же относится к числовому примеру в Таблице 6, где вычисления z-оценок идет по вертикали, то есть по каждому региону j отдельно, а затем расчет средних $\langle Z_i \rangle$ по горизонтали, то есть по каждому субдомену i . Как и



Индекс детского благополучия. Математическая модель субъективного индекса.

ожидалось, среднеквадратичное $\sigma_{\langle Z_i \rangle} = 0,28 \neq 1$, поэтому применяем стандартный метод нормирования, деля голубой столбец поэлементно на $\sigma_{\langle Z_i \rangle}$.

В результате получается столбец Z_{tot} , который затем преобразуется (Табл. 7., Табл. 8) для визуализации индексов СБ в разрезе субдоменов (рис. 6.).

Таблица 7. Общие Z-оценки в разрезе субдоменов (Индексы СБ)

Удовлетворенность ↓	Z_{tot}
Достижениями	-0,79
Собой	0,38
Здоровьем и самочувствием	-2,03
Безопасностью	-0,04
Учетом мнения	2,05
Внутренней сетью	-0,79
Внешней сетью	1,04
Нормализацией жизни	0,57
Материальным положением	-0,08
Социальными контекстами	0,18
Хронотопом	-0,85
Доверием	0,37
Нормировка: Среднее, $\langle Z_{tot} \rangle$	0,00
Нормировка: Сигма, $\sigma_{Z_{tot}}$	1,00

Таблица 8. Общие Z-оценки в разрезе субдоменов (Индексы СБ) сортированные по Z_i

Удовлетворенность ↓	Z_{tot}
Учетом мнения Н	2,05
Внешней сетью Н	1,04
Нормализацией жизни Н	0,57
Собой М	0,38
Доверием М	0,37
Социальными контекстами М	0,18
Безопасностью М	-0,04
Материальным положением М	-0,08
Достижениями L	-0,79
Внутренней сетью L	-0,79
Хронотопом L	-0,85
Здоровьем и самочувствием L	-2,03
Нормировка: Среднее, $\langle Z_{tot} \rangle$	0,00
Нормировка: Сигма, $\sigma_{Z_{tot}}$	1,00



Рисунок 6 – Z-оценки относительного субъективного благополучия в разрезе субдоменов



Существует еще один способ получения индексов СБ в разрезе субдоменов, который предполагает замену строк на столбцы не в стадии обработки, а в исходных данных. То есть предлагается переформатировать исходную таблицу 5, заменив в ней строки на столбцы и далее использовать первую цепочку алгоритмов. Такая операция называется транспонированием.

Таблица 6. Z-оценки относительного СБ δm_i по субдоменам i в разрезе субдоменов ($n > 1\,000$ чел.)

[illegible]



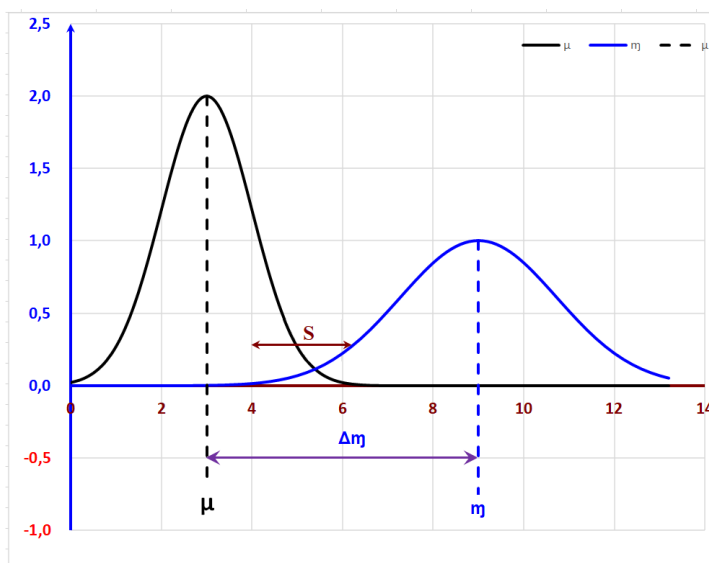
Расчет значимых величин (Т-тест Стьюдента)

Как видно из рисунков 1 и 2 различия между средними значениями η и μ оказываются весьма незначительными (порядка нескольких процентов). Это заставляет нас задаться вопросом — является ли разность этих величин значимой.

Для решения этой проблемы воспользуемся Т-тестом Стьюдента.

Идея этого теста иллюстрируется ниже следующими Рисунками 7 и 8.

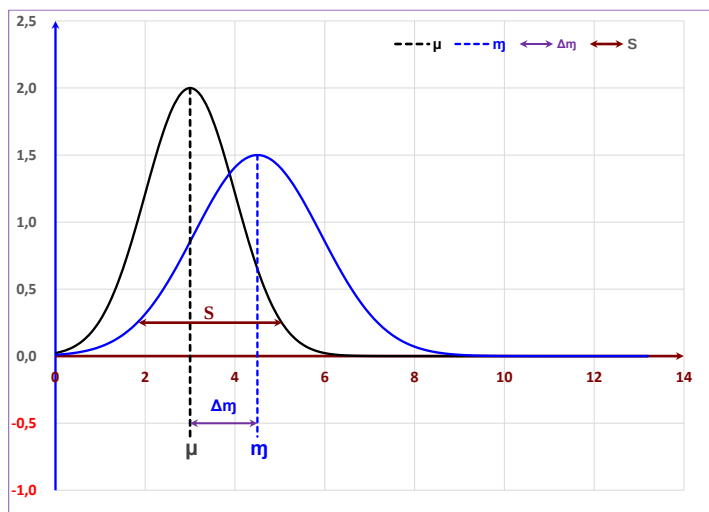
Рисунок — 7 Распределения случайных величин со средними μ и η когда $\Delta\eta = |\eta - \mu| \gg S$



Как видно из рисунка 7, эти распределения различны, как различны их средние, потому что размер области их «пересечения» S много меньше расстояния между средними $|\eta - \mu| \gg S$.

На рисунке 8 представлена обратная ситуация, где размер области «пересечения» распределений S больше расстояния между средними $|\eta - \mu| < S$, и распределения и их средние могут быть неразличимы.

Рисунок 8 — Распределения случайных величин с разностью средних $\Delta\eta = |\eta - \mu| < S$





Размер области **S** зависит от количества точек, попадающих в нее и ширины распределений, и потому размер области **S** рассчитывается как доверительный интервал точек **S_t**, попавших в область «пересечения» **S**, который рассчитывается по формулам:

$$\text{Доверительный интервал, } S_t = \langle \sigma \rangle \cdot t_{\text{крит}}, \quad (1)$$

где коэффициент $t_{\text{крит}} = \text{Stud}_{\text{обр}}(1-P; f)$ – обратное распределение Стьюдента с надежностью $P = 0,95$ и числом степеней свободы **f**:

$$t_{\text{крит}} = \text{Stud}_{\text{обр}}(1-P; f) \quad (2)$$

Число точек в области «пересечения» **f** считается по формуле:

$$f = \frac{1}{\frac{1}{(N-1) \left[1 + \frac{d \cdot N}{D \cdot n} \right]^2} + \frac{1}{(n-1) \left[1 + \frac{D \cdot n}{d \cdot N} \right]^2}} \quad (3)$$

Здесь **N** и **n** – число записей во всей базе данных и объем текущей выборки соответственно; **D** и **d** – дисперсия величин данного субдомена во всей базе данных и дисперсия в текущей выборке, соответственно.

Средний разброс, это среднеквадратичное отклонение точек в области **S**:

$$\langle \sigma \rangle = (d/n + D/N)^{1/2} \quad (4)$$

Рассчитанный по вышеприведенным формулам доверительный интервал **S_t**, сравнивается с разностью средних $|\bar{m} - \mu|$ и, если $|\bar{m} - \mu| > S_t$, то делается вывод что средние **m** и **μ** различимы и величина относительного благополучия $\delta m = (\bar{m} - \mu) / \mu$ значима с надежностью $P \geq 0,95$.

В противоположном случае, когда $|\bar{m} - \mu| < S_t$, поделив обе части этого неравенства на **μ**, получаем:

$$-S_t / \mu < \delta m = (\bar{m} - \mu) / \mu < S_t / \mu, \text{ или, переобозначив,}$$

$$G^- < \delta m < G^+, \text{ где } G^\pm = \pm S_t / \mu \quad (5)$$

Здесь **G[±]** имеет смысл границ области незначимости **δm** с надежностью $P = 0,95$

На всех рисунках, где представлены профили относительного благополучия **δm**, границы незначимости **G[±]** обозначены тонкими коричневыми линиями, симметричными относительно оси ОХ. Ширина области незначимости **δm** **G[±]** при $N \gg n \gg 1$ увеличивается с увеличением надежности **P** и с уменьшением **n** ($G^\pm \sim n^{-1/2}$)

Выяснение значимости относительного благополучия **δm** производится при обработке базы автоматически сразу после вычисления **δm** последовательно по формулам (4), (3), (2), (1), (5).

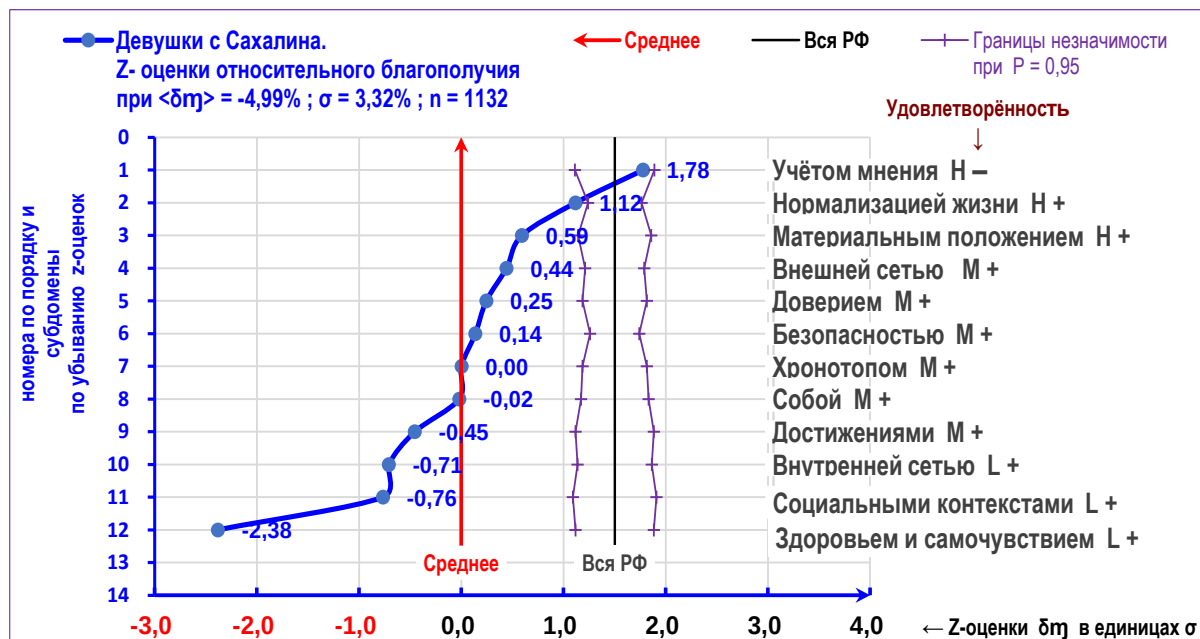
Относительное благополучие **δm**, попав в область незначимости (рис.8), можно считать равным нулю с высокой надежностью $P = 0,95$, что, в свою очередь, означает что **m** = **μ**, то есть данная ситуация, описываемая данным конкретным относительным благополучием **δm_k**, не отличается от средней ситуации по всей базе или по всей России с той же высокой надежностью $P = 0,95$.

На рисунке 9 показан расчетный лист «СУБДОМ» с вычислениями числовых характеристик в том числе областей незначимости субдоменов СБ при включенных



фильтрах «Девушка» и «Сахалинская область». По этим данным построена обобщающая диаграмма z-оценок субдоменов СБ.

Рисунок 9. Обобщенный график z-оценок и других числовых характеристик субдоменов СБ «Девушки, Сахалинская область»



На рисунке 10 показан расчетный лист Excel, где рассчитываются числовые характеристики субдоменов СБ и их значимости.



Рисунок 10. Лист Excel, где рассчитываются числовые характеристики

H8																	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1			Вся база μ ; $N = 151745$	Ко всем значениям прибавлена малая $\epsilon = 1E-13 = 10^{-13}$													
2		Вся	Дисперсия D_o	0,64	0,44	0,64	0,33	0,90	0,58	0,47	0,41	0,61	0,69	0,47	0,56		
3		База.	Средние μ	3,53	3,79	3,52	4,20	4,03	4,02	3,99	4,26	4,01	3,70	3,91	4,21		
4		Эталон	Объёмы выборок N	151 745	151 745	151 745	151 745	151 745	151 745	151 745	151 745	151 745	151 745	151 745	151 745		
5		Теку-	Дисперсия D	0,59	0,51	0,59	0,33	0,80	0,69	0,43	0,40	0,67	0,74	0,49	0,56		
6		шая	Средние μ	3,30	3,60	3,07	4,01	4,07	3,72	3,85	4,21	3,89	3,42	3,71	4,03		
7		выборка	Объёмы выборок n	1 132	1 132	1 132	1 132	1 132	1 132	1 132	1 132	1 132	1 132	1 132	1 132		
8			Надёжность $P =$	1,00	1,00	1,00	1,00	0,92	1,00	1,00	1,00	1,00	1,00	1,00	1,00		
9		0,033	$= \sigma_{\Delta\eta}$	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	-4,99%	
10		Ось	ОХ	0	0	0	0	0	0	0	0	0	0	0	0	0	
11			Текущая выборка n ; $n = 1132$														
12			Независимый двухвыборочный t-тест Стьюдента при надёжности $P = 0,95$	$\leftrightarrow$ 0,95													
13		Средний разброс, $\langle \sigma \rangle = (D_o/n_o + D_d/n_d)^{0,5} =$		0,0229	0,0212	0,0228	0,0172	0,0266	0,0247	0,0196	0,0189	0,0243	0,0256	0,0209	0,0224		
14		Число степеней свободы, $f =$		1 150	1 146	1 149	1 148	1 150	1 145	1 150	1 148	1 147	1 147	1 147	1 148		
15		Доверительный интервал, $\Delta_t = \langle \sigma \rangle \cdot t_{крит} =$		0,0449	0,0416	0,0448	0,0336	0,0522	0,0485	0,0384	0,0371	0,0478	0,0502	0,0409	0,0439		
16		Δ_t меньше или больше $\Delta\eta$		А	А	А	А	В	А	А	А	А	А	А	А		
17		$\Delta\eta = \mu_j - \mu =$		0,23	0,19	0,45	0,19	0,04	0,30	0,14	0,05	0,12	0,28	0,19	0,18		
18		Различия или значимости		+	+	+	+	-	+	+	+	+	+	+	+		
19		Относительное благополучие $\delta\eta = (\mu_j - \mu) / \mu$; $n = 1132$		-6,50%	-5,05%	-12,90%	-4,53%	0,92%	-7,35%	-3,52%	-1,27%	-3,02%	-7,53%	-4,98%	-4,17%	Среднее $\langle \delta\eta \rangle$	
20		Границы незначимости $\delta\eta$ при $P = 0,95$		1,27%	1,10%	1,27%	0,80%	1,30%	1,21%	0,96%	0,87%	1,19%	1,36%	1,05%	1,04%	-4,99%	
21		Нижние границы незначимости		-1,27%	-1,10%	-1,27%	-0,80%	-1,30%	-1,21%	-0,96%	-0,87%	-1,19%	-1,36%	-1,05%	-1,04%		
22				+	+	+	+	+	+	+	+	+	+	+	+		
23		Среднее $\langle \delta\eta \rangle = -4,99\%$															
24				$f = \frac{1}{\frac{1}{(N-1) \cdot (1 + d \cdot N / D \cdot n)^2} + \frac{1}{(n-1) \cdot (1 + D \cdot n / d \cdot N)^2}}$													
25																	
26																	
27																	
28																	
29			Z - оценки относительного благополучия $\delta\eta$ при $\langle \delta\eta \rangle = -0,05$ и $\sigma = 0,033$														
30		Уровень	средний	средний	низкий	средний	высокий	низкий	средний	высокий	высокий	низкий	средний	средний			
31		$Z_{\delta\eta} =$	-0,45	-0,02	-2,38	0,14	1,78	-0,71	0,44	1,12	0,59	-0,76	0,00	0,25			
32		Значимости $\delta\eta$	+	+	+	+	-	+	+	+	+	+	+	+	+	Среднее $\langle \delta\eta \rangle$	
33		Относительное благополучие, $\delta\eta = (\mu_j - \mu) / \mu =$		-6,50%	-5,05%	-12,90%	-4,53%	0,92%	-7,35%	-3,52%	-1,27%	-3,02%	-7,53%	-4,98%	-4,17%	-4,99%	
34																	
35																	
36																	
37																	
38		12	318	319	320	321	322	323	324	325	326	327	328	332	336		
39			ИНДЕКСЫ СУБД	Удовлетворённость	Достижениями	Собой	Здоровьем и самочувствием	Безопасностью	Учёбом мнения	Внутренней сетью	Внешней сетью	Нормализацией жизни	Материальным положением	Социальными контекстами	Хронотопом	Доверием	Среднее
40				1	2	3	4	5	6	7	8	9	10	14	15		1,82%
41																	2,77%
42																	-0,20%
		БАЛЛЫ	Лист2	СУБДОМ	Буллинг	5 Мнение	Безопасность	Стрессы	t - тест	РЕГИ	Тексты						
		Готово															

субдоменов СБ и их значимости