

Успешность футбольных сборных

Эта база данных посвящена успешности футбольных сборных разных стран мира и факторам, которые потенциально могут на неё влиять. База данных имеет панельный формат (то есть, для каждой страны в наличии есть набор показателей за несколько лет), что позволяет изучать как мгновенные эффекты (к примеру, влияние проведения или объявления о проведении в будущем Чемпионата Мира по футболу на показатели национальной футбольной сборной), так и долгосрочные, такие как отдача расходов на спорт на успехи футбольной сборной на горизонте нескольких лет.

База состоит из нескольких таблиц:

- 1) **MAIN_football.csv.** Файл содержит в себе следующие показатели по 170 странам за период с 2003 по 2021 годы:

Переменная	Объяснение	Единицы измерения
country	название страны	
iso	трёхбуквенный код страны (для удобства сопоставления с данными из других источников)	
unicef_reg	код региона по классификации ЮНИСЕФ	
un_reg	код региона по классификации ООН	
un_subreg	код субрегиона по классификации ООН	
inc_group	детальная классификация уровня дохода (Всемирный Банк): <ul style="list-style-type: none"> • Low Income — низкий уровень дохода • Lower Middle Income — доход ниже среднего • Upper Middle Income — доход выше среднего • High Income — высокий уровень дохода 	
inc_gen_group	обобщённая классификация уровня дохода (Всемирный Банк): <ul style="list-style-type: none"> • Low Income — низкий уровень дохода • Middle Income — средний уровень дохода • High Income — высокий уровень дохода 	

Переменная	Объяснение	Единицы измерения
capital	столица	
lat	широта столицы	градусы
long	долгота столицы	градусы
coords	координаты столицы	
timezone	часовой пояс столицы	
year	год	
gii	индекс гендерного неравенства*	
gdp	ВВП на душу населения по ППС	доллары США (по текущему курсу)
temp	среднегодовая температура в столице страны	градусы Цельсия
host_future	дамми-переменная на признание страны хозяйкой предстоящего Чемпионата мира по футболу среди мужчин (ЧМ): <ul style="list-style-type: none"> • 1 — страна объявлена хозяйкой ЧМ (начиная с года объявления* до года проведения ЧМ) • 0 — иначе <p>* Если решение было принято во второй половине года (с июля по декабрь), единица ставится со следующего за ним года</p>	
host_now	дамми-переменная на проведение ЧМ в стране в текущем году: <ul style="list-style-type: none"> • 1 — в стране проводился ЧМ в текущем году • 0 — иначе 	
host_past	дамми-переменная на проведение ЧМ в стране в прошлом*: <ul style="list-style-type: none"> • 1 — в стране проводился ЧМ в прошлом** • 0 — иначе <p>* У стран-хозяек ЧМ 2006-2018 единицы стоят после года проведения ЧМ; у остальных стран (где ЧМ проводился до 2003 г.) — во все годы</p> <p>** Так как Бразилия и Германия неоднократно проводили ЧМ (до 2003 г. и в период, отраженный в данных), у них в этой переменной единицы стоят во все годы</p>	

Переменная	Объяснение	Единицы измерения
men_rank	мужской рейтинг FIFA, усредненный по всем обновлениям за соответствующий год (принимает значение 0, если страна не была включена в рейтинг FIFA)	
women_rank	женский рейтинг FIFA, усредненный по всем обновлениям за соответствующий год (принимает значение 0, если страна не была включена в рейтинг FIFA)	

* Рассчитывается ООН на основе пяти показателей: Maternal mortality ratio; Adolescent birth rate; Share of seats in parliament; Population with at least some secondary education; Labour force participation rate. Источник: hdr.undp.org/data-center/documentation-and-downloads

2) В файлах **obesity_adults.csv**, **obesity_children.csv**, **thinness_adults.csv**, **thinness_children.csv** содержатся данные Всемирной организации здравоохранения по ожирению и недостатку веса у взрослых (файлы “_adults”) и детей (файлы “_children”) для разных стран в разные годы, они рассчитываются отдельно для мужчин и женщин, для детей — ещё и отдельно для возрастных групп от 5 до 9 и от 10 до 19.

- переменные **obesity_...** описывают долю людей (в процентах) с индексом массы тела (ИМТ) более 30 для взрослых и более +2 стандартных отклонений от медианного веса для детей
- переменные **thinness_...** описывают долю людей (в процентах) с индексом массы тела (ИМТ) менее 18.5 для взрослых и менее -2 стандартных отклонений от медианного веса для детей

Показатели могут быть сопоставлены с основным файлом данных по трёх-символьному коду страны **countryCode** и году **year**. Учитывайте, что в файле **thinness_children.csv** присутствуют данные не за все годы.

3) В файле **gdp_expenditure_shares.csv** содержатся данные Международного валютного фонда (полученные через ресурсы Всемирного Банка) по долям отдельных статей в ВВП странах мира за разные годы. Представлены показатели:

- **expenditure_sport_recreation** — доля расходов на спортивные и рекреационные услуги в ВВП
- **expenditure_health** — доля расходов на здравоохранение в ВВП
- **expenditure_education** — доля расходов на образование в ВВП

Стоит учитывать, что эти показатели содержат очень большое количество пропусков, при использовании их совместно с основным файлом данных может понадобиться сужение набора рассматриваемых стран. Это может быть критичным, к примеру, для исследований, фокусирующихся на бедных странах, и не быть критичным для исследований, посвящённых развитым странам.

4) В файле **travel_tourism.csv** содержатся данные Travel & Tourism Competitiveness Index Всемирного Экономического Форума (полученные через ресурсы всемирного банка). Представлены следующие показатели:

- **tt_safety** — доля людей, считающих, что гулять в одиночестве ночью в их городе безопасно (от 0 до 1)
- **tt_stadiums** — количество больших спортивных стадионов в стране
- **tt_neet** — доля NEET (не работающих и не обучающихся) среди всех молодых людей в возрасте от 15 до 24
- **tt_tourism_services** — оценка конкурентоспособности туристических сервисов по оценке от 1 до 7

Стоит учитывать, что эти данные доступны только за 2019 и 2021 годы, поэтому их использование потребует значительного сокращения рассматриваемого временного интервала

База составлена на основе данных из открытых источников:

1) Общая информация о странах и регионах

JME Regional Classification:

data.unicef.org/wp-content/uploads/2018/05/JME_Regional-Classifications.xlsx

2) Столицы, их координаты и часовые пояса

Python countryinfo: pypi.org/project/countryinfo

Python tzwhere: pypi.org/project/tzwhere

GPS Coordinates: gps-coordinates.net

3) Коэффициент гендерного неравенства (GII)

Latest Human Development composite indices tables:

hdr.undp.org/data-center/documentation-and-downloads

4) ВВП на душу населения

World Economic Outlook Database:

imf.org/en/Publications/WEO/weo-database/2021/October

GDP per capita PPP | World:

tradingeconomics.com/country-list/gdp-per-capita-ppp?continent=world

GDP per capita, PPP (current international \$):

data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD

5) Численность населения

World Economic Outlook Database:

imf.org/en/Publications/WEO/weo-database/2021/October

Population | World: tradingeconomics.com/country-list/population?continent=world

World Population by Year:

worldometers.info/world-population/world-population-by-year

6) Среднегодовая температура

Average Temperature by Country: tradingeconomics.com/country-list/temperature

Annual Surface Temperature Change:

climatedata.imf.org/datasets/4063314923d74187be9596f10d034914/explore

7) Решения о странах, принимающих ЧМ

FIFA World Cup Host Announcement Decision:

web.archive.org/web/20101214080245/http://de.fifa.com/mm/document/fifafacts/mencompwc/51/97/81/fs-201_13a_fwc-host-announcement.pdf

8) Мужской рейтинг FIFA

Men's Ranking: fifa.com/fifa-world-ranking/men?dateId=id13505

FIFA World Ranking 1992-2023: kaggle.com/datasets/cashncarry/fifaworldranking/data

9) Женский рейтинг FIFA

Women's Ranking: [fifa.com/fifa-world-ranking/women?dateId=ranking_20211210](https://www.fifa.com/fifa-world-ranking/women?dateId=ranking_20211210)

FIFA World Ranking Women 2003-2023:
[kaggle.com/datasets/cashncarry/fifa-world-ranking-women/data](https://www.kaggle.com/datasets/cashncarry/fifa-world-ranking-women/data)

10) Данные об ожирении и недостатке веса

Prevalence of obesity among adults, BMI ≥ 30 (crude estimate) (%)
[who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-adults-bmi--30-\(crude-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-adults-bmi--30-(crude-estimate)-(-))

Prevalence of underweight among adults, BMI < 18.5 (crude estimate) (%)
[who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-underweight-among-adults-bmi-18-\(crude-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-underweight-among-adults-bmi-18-(crude-estimate)-(-))

Prevalence of thinness among children and adolescents, BMI < -2 standard deviations below the median (crude estimate) (%)
[who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-thinness-among-children-and-adolescents-bmi--2-standard-deviations-below-the-median-\(crude-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-thinness-among-children-and-adolescents-bmi--2-standard-deviations-below-the-median-(crude-estimate)-(-))

Prevalence of obesity among children and adolescents, BMI $> +2$ standard deviations above the median (crude estimate) (%)
[who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-children-and-adolescents-bmi-2-standard-deviations-above-the-median-\(crude-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-children-and-adolescents-bmi-2-standard-deviations-above-the-median-(crude-estimate)-(-))

11) Статьи ВВП

Expenditure on recreational & sporting services, Domestic currency
prosperitydata360.worldbank.org/en/indicator/IMF+GFSCFOG+GERS_G14_XDC

12) Количество стадионов

prosperitydata360.worldbank.org/en/indicator/WEF+TTDI+STADIUM