極度に貧しい！？

2023-12-11

UNICEF [[Link](https://www.unicef.or.jp/kodomo/sdgs/17goals/1-poverty/)]

1. 世界では、六人に一人（3億5600万人）の子どもたちが「極度にまずしい」暮らしをしています。

どのようなデータから、このようなことがわかるのでしょうか。まずは、極度の貧困とは、どのように定義しているのでしょうか。

#### Fact Sheet: An Adjustment to Global Poverty Lines [[Link](https://www.worldbank.org/en/news/factsheet/2022/05/02/fact-sheet-an-adjustment-to-global-poverty-lines)]

The World Bank updated the global poverty lines in September 2022. The decision, announced in May, follows the release in 2020 of new purchasing power parities (PPPs)—the main data used to convert different currencies into a common, comparable unit and account for price differences across countries. The new extreme poverty line of $2.15 per person per day, which replaces the $1.90 poverty line, is based on 2017 PPPs. Here you find more information about this change and what it means for measuring global poverty.

世界銀行は2022年9月に世界の貧困ラインを更新した。5月に発表されたこの決定は、2020年に新購買力平価（PPP）が発表されたことを受けてのものである。新購買力平価（PPP）とは、さまざまな通貨を共通の比較可能な単位に変換し、貧困の範囲を説明するために使用される主要なデータである。国ごとの価格差。1.90ドルの貧困線に代わる、1人当たり1日当たり2.15ドルという新たな極度の貧困線は、2017年のPPPに基づいている。ここでは、この変化とそれが世界の貧困の測定に何を意味するかに関する詳細情報をご覧いただけます。

そこで、この購買力平価（PPP: purchasing power parities）をまずは、調べてみましょう。国や、地域ごとの、GDP, PPP （国内総生産（Gross Domestic Prodoct））を、人口で割った、一人当たりの、GDP PPP です。2017年を基準として、インフレ率を調整し、ドル換算されています。総人口の指標も使います。

* GDP per capita, PPP (constant 2017 international $): NY.GDP.PCAP.PP.KD

GDP per capita, PPP (constant 2017 international $) を、指標名（Indicator Name）といい、NY.GDP.PCAP.PP.KD を、指標コード（Indicator Code）と言います。後者を、WDI（世界開発指標、World Development Indicator）と呼ぶこともあります。

### 準備

Step 1. （R に機能を付け加える）パッケージのインストール（最初だけ）

install.packages("tidyverse")  
install.packages("WDI")  
install.packages('showtext')

Step 2. パッケージを使えるように読み込みます。

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.3 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ ggplot2 3.4.4 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.0  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(WDI)  
library(showtext)

## Loading required package: sysfonts  
## Loading required package: showtextdb

showtext\_auto()

Step 3. データを保存するための data という名前のディレクトリ（フォルダー）を作成します。（最初だけ）

dir.create("data")

## Warning in dir.create("data"): 'data' already exists

Step 4. エラーが生じた時に、調べやすいので、‘システム言語（System Language）’ を英語にしておきます。（最初だけ）

Sys.setenv(LANG = "en")

Step 5. データを読み込みます。WDI パッケージをつかうと、簡単に、データを読み込むことができます。わかりやすい名前（gdppcap, pop）をつけ、人口（pop）も読み込んでおきます。extra = TRUE としておくと、使い情報を一緒に読み込むことができます。

df\_pcap <- WDI(indicator = c(gdppcap = "NY.GDP.PCAP.PP.KD"), extra = TRUE)

Step 6. データをみてみます。

head(df\_pcap)

## country iso2c iso3c year gdppcap status lastupdated region capital  
## 1 Afghanistan AF AFG 1991 NA 2023-10-26 South Asia Kabul  
## 2 Afghanistan AF AFG 1987 NA 2023-10-26 South Asia Kabul  
## 3 Afghanistan AF AFG 1990 NA 2023-10-26 South Asia Kabul  
## 4 Afghanistan AF AFG 1989 NA 2023-10-26 South Asia Kabul  
## 5 Afghanistan AF AFG 1988 NA 2023-10-26 South Asia Kabul  
## 6 Afghanistan AF AFG 2019 2079.922 2023-10-26 South Asia Kabul  
## longitude latitude income lending  
## 1 69.1761 34.5228 Low income IDA  
## 2 69.1761 34.5228 Low income IDA  
## 3 69.1761 34.5228 Low income IDA  
## 4 69.1761 34.5228 Low income IDA  
## 5 69.1761 34.5228 Low income IDA  
## 6 69.1761 34.5228 Low income IDA

Step 7. データの列（columns、変数　variables）はどのようなものがあるかを表示します。

str(df\_pcap)

## 'data.frame': 16758 obs. of 13 variables:  
## $ country : chr "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...  
## $ iso2c : chr "AF" "AF" "AF" "AF" ...  
## $ iso3c : chr "AFG" "AFG" "AFG" "AFG" ...  
## $ year : int 1991 1987 1990 1989 1988 2019 2018 2017 2016 2015 ...  
## $ gdppcap : num NA NA NA NA NA ...  
## ..- attr(\*, "label")= chr "GDP per capita, PPP (constant 2017 international $)"  
## $ status : chr "" "" "" "" ...  
## $ lastupdated: chr "2023-10-26" "2023-10-26" "2023-10-26" "2023-10-26" ...  
## $ region : chr "South Asia" "South Asia" "South Asia" "South Asia" ...  
## $ capital : chr "Kabul" "Kabul" "Kabul" "Kabul" ...  
## $ longitude : chr "69.1761" "69.1761" "69.1761" "69.1761" ...  
## $ latitude : chr "34.5228" "34.5228" "34.5228" "34.5228" ...  
## $ income : chr "Low income" "Low income" "Low income" "Low income" ...  
## $ lending : chr "IDA" "IDA" "IDA" "IDA" ...

Step 8. 必要な列だけ取り出します。

df\_pcap <- df\_pcap |> select(c(1,2,4,5,8,12))  
head(df\_pcap)

## country iso2c year gdppcap region income  
## 1 Afghanistan AF 1991 NA South Asia Low income  
## 2 Afghanistan AF 1987 NA South Asia Low income  
## 3 Afghanistan AF 1990 NA South Asia Low income  
## 4 Afghanistan AF 1989 NA South Asia Low income  
## 5 Afghanistan AF 1988 NA South Asia Low income  
## 6 Afghanistan AF 2019 2079.922 South Asia Low income

Step 9. （ちょっと高度ですが）region, income, lending には、どのようなものがあるか、みてみます。

df\_pcap |> select(region, income) |> lapply(unique)

## $region  
## [1] "South Asia" "Aggregates"   
## [3] "Europe & Central Asia" "Middle East & North Africa"  
## [5] "East Asia & Pacific" "Sub-Saharan Africa"   
## [7] "Latin America & Caribbean" "North America"   
## [9] NA   
##   
## $income  
## [1] "Low income" "Aggregates" "Upper middle income"  
## [4] "Lower middle income" "High income" NA   
## [7] "Not classified"

### 視覚化（Visualization）

図 1. 世界の　GDP per Capita の推移（経年変化）を見てみます。

COUNTRY <- "World"  
df\_pcap |> filter(country == COUNTRY) |> drop\_na(gdppcap) |>  
 ggplot(aes(year, gdppcap)) + geom\_line() +  
 labs(title = "世界の一人当たりの GDP の平均の推移")

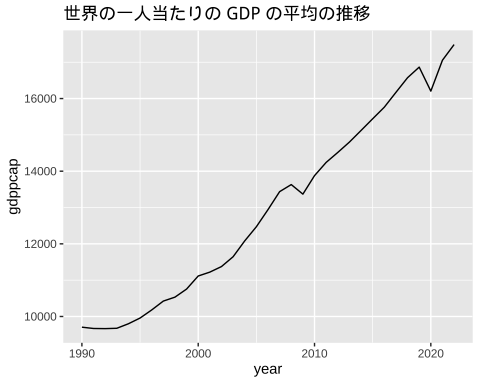


図 2. Low Income の 推移をみてみましょう。

COUNTRY <- "Low income"  
df\_pcap |> filter(country == COUNTRY) |> drop\_na(gdppcap) |>  
 ggplot(aes(year, gdppcap)) + geom\_line() +  
 labs(title = "低所得国の一人当たりの GDP の平均の推移")

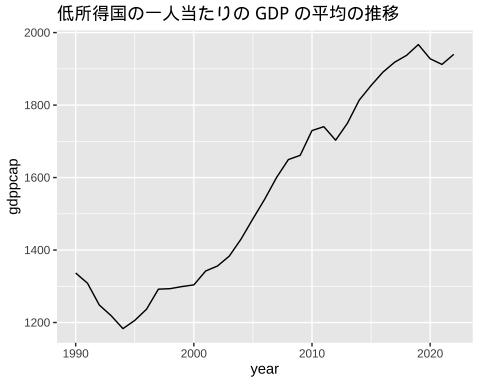


図 3. Low Income の 推移を、１日あたりに変えてみてみましょう。

COUNTRY <- "Low income"  
df\_pcap |> filter(country == COUNTRY) |> drop\_na(gdppcap) |>  
 ggplot(aes(year, gdppcap/365)) + geom\_line() +  
 labs(title = "世界の１日一人当たりの GDP の平均の推移")



図 4. 2022年の、Low Income Country について、GDP per Capita の少ないから順に並べてみましょう。

df\_pcap |> filter(year == 2022, region != "Aggregates") |>   
 filter(income == "Low income") |>  
 drop\_na(gdppcap) |> arrange(gdppcap)

図 5. 一人当たりのGDP の2022年の分布をみてみましょう。

df\_pcap |> filter(year == 2022, region != "Aggregates") |> drop\_na(gdppcap) |>   
 ggplot(aes(gdppcap)) + geom\_histogram(binwidth = 10000) +  
 labs(title = "2022年の 一人当たりの GDP の国の分布")

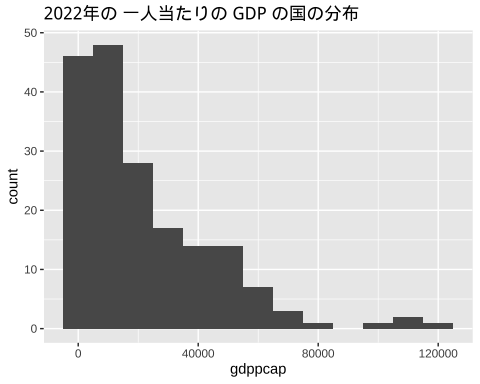


図 6. 地域ごとに色分けしてみてみましょう。

df\_pcap |> filter(year == 2022, region != "Aggregates") |>   
 drop\_na(gdppcap) |>   
 ggplot(aes(gdppcap, fill = region)) +   
 geom\_histogram(col = "black", linewidth = 0.2, binwidth = 10000) +  
 labs(title = "2022年の 一人当たりの GDP の国の分布",   
 subtitle = "地域情報付き")

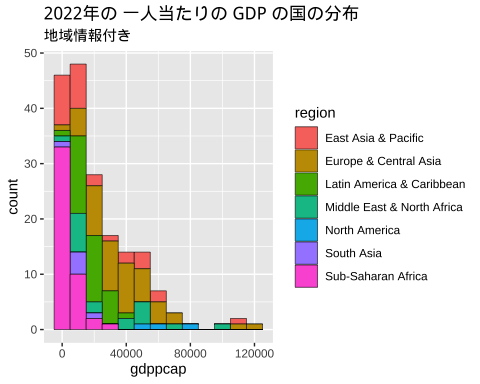
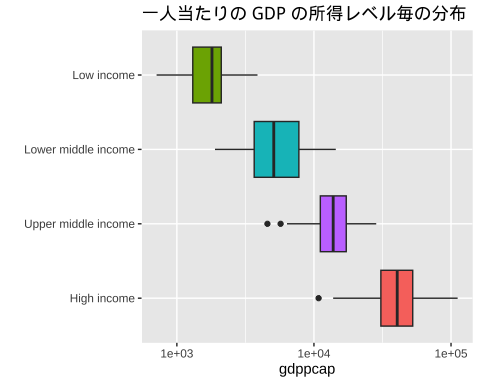


図 7. 箱ひげ図で見てみましょう。

df\_pcap |> filter(year == 2020) |> drop\_na(gdppcap) |>   
 filter(income != "Aggregates") |>   
 ggplot(aes(gdppcap, factor(income, levels = c("High income", "Upper middle income", "Lower middle income", "Low income")), fill = income)) + geom\_boxplot() + scale\_x\_log10() +  
 labs(y = "") +  
 theme(legend.position = "none") +  
 labs(title = "一人当たりの GDP の所得レベル毎の分布")



### 課題

* 図1から図7 それぞれからどんなことが分かりましたか。
* 図1から図7 でわからなかったことで、このデータから、どのようなことが知りたいですか。
* こんな図は、描けないかなどの提案はありますか。
* 極度の貧困の人について知りたいときには、あとどのような情報が必要でしょうか。