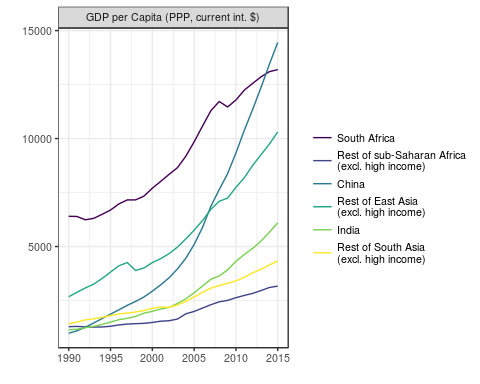
# Trends in Macro Indicators

library(WDI)  
library(data.table)  
library(ggplot2)  
library(stringr)  
library(viridis)  
  
setwd("~/Projects/2017-agra-aasr")  
load("./tmp/2017-agra-aasr\_WDI.RData")  
  
#####################################################################################  
# 2017.06.15 WDI Time Series for Peter  
#####################################################################################  
# We need:  
# - cereal yields in kg/ha for SSA, with/without ZAF, China, South Asia (not China) for 2000-2015  
# - GDP per capita for SSA, with/without ZAF  
# - GDP shares by sector (ag, services, industry) SSA with/without ZAF  
# WDI variables are:  
# - AG.YLD.CREL.KG  
# - NY.GDP.PCAP.PP.CD  
# - NV.AGR.TOTL.ZS  
# - NV.SRV.TETC.ZS  
# - NV.IND.TOTL.ZS  
  
# SSF ZG Sub-Saharan Africa (all income levels) Aggregates  
# SSA ZF Sub-Saharan Africa (developing only) Aggregates  
# ECS Z7 Europe & Central Asia (all income levels) Aggregates  
# ECA 7E Europe & Central Asia (developing only) Aggregates  
# EAS Z4 East Asia & Pacific (all income levels) Aggregates  
# EAP 4E East Asia & Pacific (developing only) Aggregates  
  
reg <- WDI\_data[[2]]  
reg <- data.table(reg)  
reg[region %like% "Sub-Saharan Africa", .N, keyby=country]  
  
# We need mean cereal yield weighted by cereal production I assume  
vars <- c("AG.YLD.CREL.KG", "NY.GDP.PCAP.PP.CD",  
 "NV.AGR.TOTL.ZS", "NV.SRV.TETC.ZS", "NV.IND.TOTL.ZS")  
  
dt = WDI(indicator=vars, country=c("ZG", "ZF", "Z4", "4E", "CN"), start=1990, end=2015)  
dt <- data.table(dt)  
dt <- melt(dt, id.vars=1:3)  
  
levels(dt$variable)  
dt[, varLabel := factor(variable, labels=c(  
 "Cereal Yield\n(kg/ha)",  
 "GDP per Capita\n(PPP, current int. $)",  
 "Agriculture\n(value added, % of GDP)",  
 "Services\n(valude added, % of GDP)",  
 "Industry\n(value added, % of GDP)"))]  
  
setnames(dt, "country", "Region")  
dt[, Region := factor(Region)]  
  
ggplot(dt[year<2015 & Region=="Sub-Saharan Africa"], aes(year, value, color=Region)) + geom\_line() +  
 xlab("") + ylab("") +  
 facet\_grid(varLabel ~ ., scales="free\_y")  
  
ggsave("./out/MB/2017-agra-aasr\_WDI\_SSA\_ts.png", width=7, heigh=9, units="in")  
  
ggplot(dt[year<2015 & iso2c=="ZF"], aes(year, value, color=Region)) + geom\_line() +  
 xlab("") + ylab("") +  
 facet\_grid(varLabel ~ ., scales="free\_y")  
  
ggsave("./out/MB/2017-agra-aasr\_WDI\_ZF\_ts.png", width=7, heigh=9, units="in")  
  
ggplot(dt[year<2015], aes(year, value, color=Region)) + geom\_line() +  
 xlab("") + ylab("") +  
 facet\_grid(varLabel ~ ., scales="free\_y")  
  
ggsave("./out/MB/2017-agra-aasr\_WDI\_ts.png", width=7, heigh=9, units="in")  
  
#####################################################################################  
# => noticed that only diff between region `SSA` and `SSA (excluded high income)` is  
# Equatorial Guinea as ZAF is not included in High Income.  
# So we need to regenerate regional aggregates by hand, using the correct formulas:  
# - yield = sum(yield\*production)/sum(production)  
# - GPD per capita = sum(GDP)/sum(population)  
# - value\_added\_% = 100\*sum(value\_added)/sum(GDP)  
# Also need to exclude `China` from `East Asia`  
  
# Generate new regional groupings  
reg[region %like% "Sub-Saharan Africa", unique(iso3c), by=income]  
reg[region %like% "East Asia", unique(country), by=.(income, iso3c)]  
reg[region %like% "East Asia" & !income %like% "High", unique(country), by=income]  
reg[region %like% "South Asia" & !income %like% "High", unique(country), by=.(income, iso3c)]  
  
reg[, group := NULL]  
reg[region %like% "Sub-Saharan Africa" & !income %like% "High" & iso3c != "ZAF",  
 group := "Rest of sub-Saharan Africa (excl. high income)"]  
reg[region %like% "East Asia" & !income %like% "High" & iso3c != "CHN",  
 group := "Rest of East Asia (excl. high income)"]  
reg[region %like% "South Asia" & !income %like% "High" & iso3c != "IND",  
 group := "Rest of South Asia (excl. high income)"]  
  
reg[iso3c=="ZAF", group := "South Africa"]  
reg[iso3c=="CHN", group := "China"]  
reg[iso3c=="IND", group := "India"]  
  
reg[!is.na(group), unique(country), by=group]  
  
# Re-download all needed WDI indicators  
vars <- c("AG.YLD.CREL.KG", "AG.PRD.CREL.MT",  
 #"NY.GDP.PCAP.PP.CD",  
 "NY.GDP.MKTP.PP.CD", "SP.POP.TOTL",  
 #"NV.AGR.TOTL.ZS", "NV.SRV.TETC.ZS", "NV.IND.TOTL.ZS", "NV.IND.MANF.ZS",  
 "NV.AGR.TOTL.KD", "NV.SRV.TETC.KD", "NV.IND.TOTL.KD", "NV.IND.MANF.KD", "NY.GDP.MKTP.KD",  
 #"EA.PRD.AGRI.KD",  
 "AG.LND.AGRI.K2", "SL.AGR.EMPL.ZS", "SL.TLF.TOTL.IN")  
  
WDIsearch(vars)  
dt = WDI(indicator=vars, country=reg[!is.na(group), unique(iso2c)], start=1990, end=2015)  
dt <- data.table(dt)  
  
# Merge in group names  
setkey(reg, iso2c)  
setkey(dt, iso2c)  
dt[reg, Region := group]  
  
# Re-compute regional aggregates  
# Make sure to only include full records (hence ifelse())  
dt.group <- dt[, .(  
 AG.YLD.CREL.KG = sum(AG.YLD.CREL.KG\*AG.PRD.CREL.MT, na.rm=T)/sum(ifelse(is.na(AG.YLD.CREL.KG), NA, AG.PRD.CREL.MT), na.rm=T),  
 NY.GDP.PCAP.PP.CD = sum(ifelse(is.na(SP.POP.TOTL), NA, NY.GDP.MKTP.PP.CD), na.rm=T)/sum(ifelse(is.na(NY.GDP.MKTP.PP.CD), NA, SP.POP.TOTL), na.rm=T),  
 NV.AGR.TOTL.ZS = 100\*sum(ifelse(is.na(NY.GDP.MKTP.KD), NA, NV.AGR.TOTL.KD), na.rm=T)/sum(ifelse(is.na(NV.AGR.TOTL.KD), NA, NY.GDP.MKTP.KD), na.rm=T),  
 NV.SRV.TETC.ZS = 100\*sum(ifelse(is.na(NY.GDP.MKTP.KD), NA, NV.SRV.TETC.KD), na.rm=T)/sum(ifelse(is.na(NV.SRV.TETC.KD), NA, NY.GDP.MKTP.KD), na.rm=T),  
 NV.IND.TOTL.ZS = 100\*sum(ifelse(is.na(NY.GDP.MKTP.KD), NA, NV.IND.TOTL.KD), na.rm=T)/sum(ifelse(is.na(NV.IND.TOTL.KD), NA, NY.GDP.MKTP.KD), na.rm=T),  
 NV.IND.MANF.ZS = 100\*sum(ifelse(is.na(NY.GDP.MKTP.KD), NA, NV.IND.MANF.KD), na.rm=T)/sum(ifelse(is.na(NV.IND.MANF.KD), NA, NY.GDP.MKTP.KD), na.rm=T),   
 lABOR.AG.PROD = sum(NV.AGR.TOTL.KD, na.rm=T)/sum(ifelse(is.na(NV.AGR.TOTL.KD), NA, (SL.AGR.EMPL.ZS/100)\*SL.TLF.TOTL.IN), na.rm=T),  
 lAND.AG.PROD = sum(NV.AGR.TOTL.KD, na.rm=T)/sum(ifelse(is.na(NV.AGR.TOTL.KD), NA, AG.LND.AGRI.K2), na.rm=T)  
), by=.(Region, year)]  
  
# Replace Inf with NA  
dt.group <- dt.group[, lapply(.SD, function(x) ifelse(x==Inf, NA, x))]  
  
# Impute missing by country, by indicator using linear interpolation  
setorder(dt.group, Region, year)  
dt.group.imp <- dt.group[, lapply(.SD, zoo::na.approx, na.rm=F), .SDcols=-c(1:2), by=Region]  
dt.group.imp <- cbind(year=dt.group$year, dt.group.imp)  
  
dt.group <- melt(dt.group.imp, id.vars=1:2)  
levels(dt.group$variable)  
dt.group[, varLabel := factor(variable, labels=c(  
 "Cereal Yield (kg/ha)",  
 "GDP per Capita (PPP, current int. $)",  
 "Agriculture (value added, % of GDP)",  
 "Services (value added, % of GDP)",  
 "Industry (value added, % of GDP)",  
 "Manufacturing (value added, % of GDP)",  
 "Agricultural Value Added per Worker",  
 "Agricultural Value Added per Hectare"))]  
  
dt.group[, Region := factor(Region, levels=c(  
 "South Africa", "Rest of sub-Saharan Africa (excl. high income)",  
 "China", "Rest of East Asia (excl. high income)",  
 "India", "Rest of South Asia (excl. high income)"))]  
  
# Ruttanograms  
dt.rutt <- dcast(dt.group[year %between% c(2000, 2014) & variable %like% ".AG.PROD"], Region+year~variable)  
setnames(dt.rutt, 3:4, c("labor", "land"))  
setorder(dt.rutt, Region, year)  
dt.rutt[, yearLabel := NULL]  
dt.rutt[!is.na(land) & !is.na(labor), yearLabel := ifelse(year==min(year) | year==max(year), year, NA), by=Region]  
  
# Save data  
fwrite(setnames(setcolorder(dt, c(1,2,16,3:15)), "Region", "region"), "./out/MB/2017-agra-aasr\_WDI\_ts (corrected).csv")  
  
# Save workspace  
save.image("./tmp/2017-agra-aasr\_WDI.RData")

Long-term time series of macro-economic indicators across regions. Note that these series are aggregated using WDI country-level indicators. Table @ref(tab:agg-regions) below lists all countries included in each region. The raw data is available at [2017-agra-aasr\_WDI\_ts (corrected).csv](https://github.com/mbacou/2017-agra-aasr/blob/master/out/MB/2017-agra-aasr_WDI_ts%20(corrected).csv).

ggplot(dt.group[varLabel %in% levels(varLabel)[2]],  
 aes(year, value, color=Region)) +  
 geom\_line() + xlab("") + ylab("") +  
 facet\_wrap(~varLabel, nrow=1, scales="free\_y") +  
 scale\_color\_manual(name="",  
 values=viridis\_pal()(6),  
 labels=str\_replace(levels(dt.group$Region), fixed(" ("), "\n(")) +  
 theme\_bw(base\_size=10) + theme(legend.position="right")

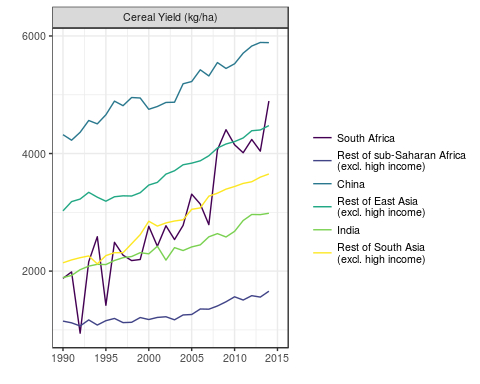


Regional Trends in GDP per Capita (1990-2015) Source: WDI.

ggsave("../out/MB/2017-agra-aasr\_WDI\_ts\_agg-gdp.png", width=6, height=4, units="in")

ggplot(dt.group[varLabel %in% levels(varLabel)[1]],  
 aes(year, value, color=Region)) +  
 geom\_line() + xlab("") + ylab("") +  
 facet\_wrap(~varLabel, nrow=1, scales="free\_y") +  
 scale\_color\_manual(name="",  
 values=viridis\_pal()(6),  
 labels=str\_replace(levels(dt.group$Region), fixed(" ("), "\n(")) +  
 theme\_bw(base\_size=10) + theme(legend.position="right")

## Warning: Removed 6 rows containing missing values (geom\_path).

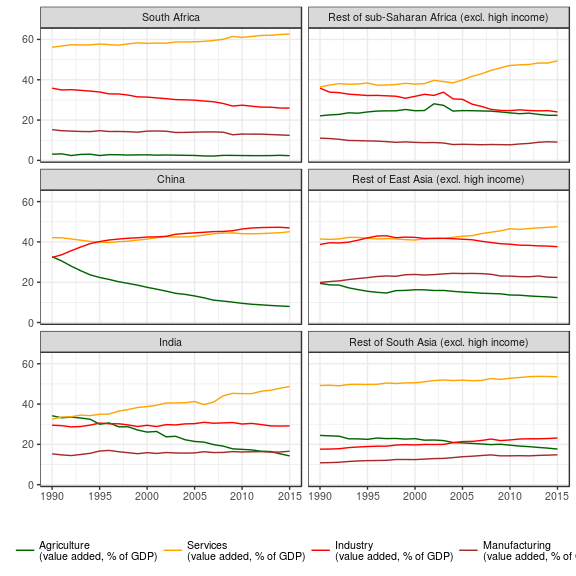


Regional Trends in Cereal Yield (1990-2015) Source: WDI.

ggsave("../out/MB/2017-agra-aasr\_WDI\_ts\_agg-yield.png", width=6, height=4, units="in")

## Warning: Removed 6 rows containing missing values (geom\_path).

ggplot(dt.group[varLabel %like% "value added"],  
 aes(year, value, color=varLabel)) +  
 geom\_line() + xlab("") + ylab("") +  
 facet\_wrap(~Region, ncol=2) +  
 scale\_color\_manual(name="",  
 values=c("darkgreen", "orange", "red", "brown"),  
 labels=str\_replace(levels(dt.group$varLabel)[3:6], fixed("("), "\n(")) +  
 theme\_bw(base\_size=10) + theme(legend.position="bottom")



Regional Trends in Sectoral Value Added (1990-2015) Source: WDI.

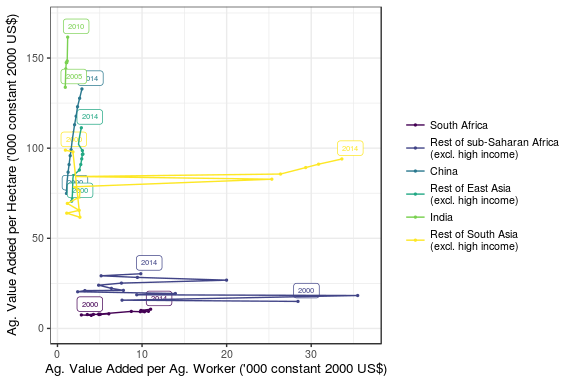
ggsave("../out/MB/2017-agra-aasr\_WDI\_ts\_gdp-sector.png", width=6, height=7, units="in")

Note that India is missing multiple years of agricultural employment statistics. Simple **linear interpolation** is used to impute these missing values.

ggplot(dt.rutt,  
 aes(labor/1000, land/1000, color=Region, group=Region, label=yearLabel)) +  
 geom\_point(size=.5) + geom\_label(size=2, nudge\_x=1, nudge\_y=6, show.legend=F) +  
 geom\_path(na.rm=T) + ylim(0,170) +  
 xlab("Ag. Value Added per Ag. Worker ('000 constant 2000 US$)") +  
 ylab("Ag. Value Added per Hectare ('000 constant 2000 US$)") +  
 scale\_color\_manual(name="",  
 labels=str\_replace(levels(dt.group$Region), fixed("("), "\n("),  
 values=viridis\_pal()(6)) +  
 theme\_bw(base\_size=10) + theme(legend.position="right")

## Warning: Removed 9 rows containing missing values (geom\_point).

## Warning: Removed 78 rows containing missing values (geom\_label).



Trends in Land and Labor Productivity (2000-2014). Source: WDI.

#ggsave("../out/MB/2017-agra-aasr\_RUTT\_ts (corrected).svg", width=6, heigh=4, units="in")

The regional aggregation used in the above graphs is provided here.

kable(reg[!is.na(group), .(`ISO2 Code`=iso2c, `ISO3 Code`=iso3c, `Income Level`=income), keyby=.(Region=group, Country=country)],  
 caption="Regional Aggregation")

Regional Aggregation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Region | Country | ISO2 Code | ISO3 Code | Income Level |
| China | China | CN | CHN | Upper middle income |
| India | India | IN | IND | Lower middle income |
| Rest of East Asia (excl. high income) | American Samoa | AS | ASM | Upper middle income |
| Rest of East Asia (excl. high income) | Cambodia | KH | KHM | Low income |
| Rest of East Asia (excl. high income) | Fiji | FJ | FJI | Lower middle income |
| Rest of East Asia (excl. high income) | Indonesia | ID | IDN | Lower middle income |
| Rest of East Asia (excl. high income) | Kiribati | KI | KIR | Lower middle income |
| Rest of East Asia (excl. high income) | Korea, Dem. Rep. | KP | PRK | Low income |
| Rest of East Asia (excl. high income) | Lao PDR | LA | LAO | Lower middle income |
| Rest of East Asia (excl. high income) | Malaysia | MY | MYS | Upper middle income |
| Rest of East Asia (excl. high income) | Marshall Islands | MH | MHL | Lower middle income |
| Rest of East Asia (excl. high income) | Micronesia, Fed. Sts. | FM | FSM | Lower middle income |
| Rest of East Asia (excl. high income) | Mongolia | MN | MNG | Lower middle income |
| Rest of East Asia (excl. high income) | Myanmar | MM | MMR | Low income |
| Rest of East Asia (excl. high income) | Palau | PW | PLW | Upper middle income |
| Rest of East Asia (excl. high income) | Papua New Guinea | PG | PNG | Lower middle income |
| Rest of East Asia (excl. high income) | Philippines | PH | PHL | Lower middle income |
| Rest of East Asia (excl. high income) | Samoa | WS | WSM | Lower middle income |
| Rest of East Asia (excl. high income) | Solomon Islands | SB | SLB | Lower middle income |
| Rest of East Asia (excl. high income) | Thailand | TH | THA | Upper middle income |
| Rest of East Asia (excl. high income) | Timor-Leste | TL | TLS | Lower middle income |
| Rest of East Asia (excl. high income) | Tonga | TO | TON | Lower middle income |
| Rest of East Asia (excl. high income) | Tuvalu | TV | TUV | Lower middle income |
| Rest of East Asia (excl. high income) | Vanuatu | VU | VUT | Lower middle income |
| Rest of East Asia (excl. high income) | Vietnam | VN | VNM | Lower middle income |
| Rest of South Asia (excl. high income) | Afghanistan | AF | AFG | Low income |
| Rest of South Asia (excl. high income) | Bangladesh | BD | BGD | Low income |
| Rest of South Asia (excl. high income) | Bhutan | BT | BTN | Lower middle income |
| Rest of South Asia (excl. high income) | Maldives | MV | MDV | Upper middle income |
| Rest of South Asia (excl. high income) | Nepal | NP | NPL | Low income |
| Rest of South Asia (excl. high income) | Pakistan | PK | PAK | Lower middle income |
| Rest of South Asia (excl. high income) | Sri Lanka | LK | LKA | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Angola | AO | AGO | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Benin | BJ | BEN | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Botswana | BW | BWA | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Burkina Faso | BF | BFA | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Burundi | BI | BDI | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Cameroon | CM | CMR | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Cape Verde | CV | CPV | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Central African Republic | CF | CAF | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Chad | TD | TCD | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Comoros | KM | COM | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Congo, Dem. Rep. | CD | COD | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Congo, Rep. | CG | COG | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Cote d'Ivoire | CI | CIV | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Eritrea | ER | ERI | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Ethiopia | ET | ETH | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Gabon | GA | GAB | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Gambia, The | GM | GMB | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Ghana | GH | GHA | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Guinea | GN | GIN | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Guinea-Bissau | GW | GNB | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Kenya | KE | KEN | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Lesotho | LS | LSO | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Liberia | LR | LBR | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Madagascar | MG | MDG | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Malawi | MW | MWI | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Mali | ML | MLI | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Mauritania | MR | MRT | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Mauritius | MU | MUS | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Mayotte | YT | MYT | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Mozambique | MZ | MOZ | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Namibia | NA | NAM | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Niger | NE | NER | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Nigeria | NG | NGA | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Rwanda | RW | RWA | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Sao Tome and Principe | ST | STP | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Senegal | SN | SEN | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Seychelles | SC | SYC | Upper middle income |
| Rest of sub-Saharan Africa (excl. high income) | Sierra Leone | SL | SLE | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Somalia | SO | SOM | Low income |
| Rest of sub-Saharan Africa (excl. high income) | South Sudan | SS | SSD | Not classified |
| Rest of sub-Saharan Africa (excl. high income) | Sudan | SD | SDN | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Swaziland | SZ | SWZ | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Tanzania | TZ | TZA | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Togo | TG | TGO | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Uganda | UG | UGA | Low income |
| Rest of sub-Saharan Africa (excl. high income) | Zambia | ZM | ZMB | Lower middle income |
| Rest of sub-Saharan Africa (excl. high income) | Zimbabwe | ZW | ZWE | Low income |
| South Africa | South Africa | ZA | ZAF | Upper middle income |