project1

February 17, 2021

[1]:
```
!pip install wbdata

import pandas as pd
import wbdata
import numpy as np
import re
import seaborn as sns
import matplotlib.pyplot as plt
from matplotlib import animation as ani
import matplotlib.patches as mpatches
```

```
Requirement already satisfied: wbdata in /opt/conda/lib/python3.8/site-packages
(0.3.0)
Requirement already satisfied: tabulate>=0.8.5 in /opt/conda/lib/python3.8/site-
packages (from wbdata) (0.8.7)
Requirement already satisfied: requests>=2.0 in /opt/conda/lib/python3.8/site-
packages (from wbdata) (2.25.1)
Requirement already satisfied: appdirs<2.0,>=1.4 in
/opt/conda/lib/python3.8/site-packages (from wbdata) (1.4.4)
Requirement already satisfied: decorator>=4.0 in /opt/conda/lib/python3.8/site-
packages (from wbdata) (4.4.2)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/opt/conda/lib/python3.8/site-packages (from requests>=2.0->wbdata) (1.25.7)
Requirement already satisfied: idna<3,>=2.5 in /opt/conda/lib/python3.8/site-
packages (from requests>=2.0->wbdata) (2.8)
Requirement already satisfied: certifi>=2017.4.17 in
/opt/conda/lib/python3.8/site-packages (from requests>=2.0->wbdata) (2019.11.28)
Requirement already satisfied: chardet<5,>=3.0.2 in
/opt/conda/lib/python3.8/site-packages (from requests>=2.0->wbdata) (3.0.4)
```

0.1 Loading the Data

[2]:
```
# Return list of all country/region codes:
wbdata.get_country()
```

[2]:
```
id    name
----
--------------------------------------------------------------------------------
```

ABW	Aruba
AFG	Afghanistan
AFR	Africa
AGO	Angola
ALB	Albania
AND	Andorra
ARB	Arab World
ARE	United Arab Emirates
ARG	Argentina
ARM	Armenia
ASM	American Samoa
ATG	Antigua and Barbuda
AUS	Australia
AUT	Austria
AZE	Azerbaijan
BDI	Burundi
BEA	East Asia & Pacific (IBRD-only countries)
BEC	Europe & Central Asia (IBRD-only countries)
BEL	Belgium
BEN	Benin
BFA	Burkina Faso
BGD	Bangladesh
BGR	Bulgaria
BHI	IBRD countries classified as high income
BHR	Bahrain
BHS	Bahamas, The
BIH	Bosnia and Herzegovina
BLA	Latin America & the Caribbean (IBRD-only countries)
BLR	Belarus
BLZ	Belize
BMN	Middle East & North Africa (IBRD-only countries)
BMU	Bermuda
BOL	Bolivia
BRA	Brazil
BRB	Barbados
BRN	Brunei Darussalam
BSS	Sub-Saharan Africa (IBRD-only countries)
BTN	Bhutan
BWA	Botswana
CAA	Sub-Saharan Africa (IFC classification)
CAF	Central African Republic
CAN	Canada
CEA	East Asia and the Pacific (IFC classification)
CEB	Central Europe and the Baltics
CEU	Europe and Central Asia (IFC classification)
CHE	Switzerland
CHI	Channel Islands

CHL Chile
CHN China
CIV Cote d'Ivoire
CLA Latin America and the Caribbean (IFC classification)
CME Middle East and North Africa (IFC classification)
CMR Cameroon
COD Congo, Dem. Rep.
COG Congo, Rep.
COL Colombia
COM Comoros
CPV Cabo Verde
CRI Costa Rica
CSA South Asia (IFC classification)
CSS Caribbean small states
CUB Cuba
CUW Curacao
CYM Cayman Islands
CYP Cyprus
CZE Czech Republic
DEA East Asia & Pacific (IDA-eligible countries)
DEC Europe & Central Asia (IDA-eligible countries)
DEU Germany
DFS IDA countries classified as Fragile Situations
DJI Djibouti
DLA Latin America & the Caribbean (IDA-eligible countries)
DMA Dominica
DMN Middle East & North Africa (IDA-eligible countries)
DNF IDA countries not classified as Fragile Situations
DNK Denmark
DNS IDA countries in Sub-Saharan Africa not classified as fragile situations
DOM Dominican Republic
DSA South Asia (IDA-eligible countries)
DSF IDA countries in Sub-Saharan Africa classified as fragile situations
DSS Sub-Saharan Africa (IDA-eligible countries)
DZA Algeria
EAP East Asia & Pacific (excluding high income)
EAR Early-demographic dividend
EAS East Asia & Pacific
ECA Europe & Central Asia (excluding high income)
ECS Europe & Central Asia
ECU Ecuador
EGY Egypt, Arab Rep.
EMU Euro area
ERI Eritrea
ESP Spain
EST Estonia
ETH Ethiopia

EUU European Union
FCS Fragile and conflict affected situations
FIN Finland
FJI Fiji
FRA France
FRO Faroe Islands
FSM Micronesia, Fed. Sts.
FXS IDA countries classified as fragile situations, excluding Sub-Saharan Africa
GAB Gabon
GBR United Kingdom
GEO Georgia
GHA Ghana
GIB Gibraltar
GIN Guinea
GMB Gambia, The
GNB Guinea-Bissau
GNQ Equatorial Guinea
GRC Greece
GRD Grenada
GRL Greenland
GTM Guatemala
GUM Guam
GUY Guyana
HIC High income
HKG Hong Kong SAR, China
HND Honduras
HPC Heavily indebted poor countries (HIPC)
HRV Croatia
HTI Haiti
HUN Hungary
IBB IBRD, including blend
IBD IBRD only
IBT IDA & IBRD total
IDA IDA total
IDB IDA blend
IDN Indonesia
IDX IDA only
IMN Isle of Man
IND India
INX Not classified
IRL Ireland
IRN Iran, Islamic Rep.
IRQ Iraq
ISL Iceland
ISR Israel
ITA Italy

JAM Jamaica
JOR Jordan
JPN Japan
KAZ Kazakhstan
KEN Kenya
KGZ Kyrgyz Republic
KHM Cambodia
KIR Kiribati
KNA St. Kitts and Nevis
KOR Korea, Rep.
KWT Kuwait
LAC Latin America & Caribbean (excluding high income)
LAO Lao PDR
LBN Lebanon
LBR Liberia
LBY Libya
LCA St. Lucia
LCN Latin America & Caribbean
LDC Least developed countries: UN classification
LIC Low income
LIE Liechtenstein
LKA Sri Lanka
LMC Lower middle income
LMY Low & middle income
LSO Lesotho
LTE Late-demographic dividend
LTU Lithuania
LUX Luxembourg
LVA Latvia
MAC Macao SAR, China
MAF St. Martin (French part)
MAR Morocco
MCO Monaco
MDA Moldova
MDE Middle East (developing only)
MDG Madagascar
MDV Maldives
MEA Middle East & North Africa
MEX Mexico
MHL Marshall Islands
MIC Middle income
MKD North Macedonia
MLI Mali
MLT Malta
MMR Myanmar
MNA Middle East & North Africa (excluding high income)
MNE Montenegro

MNG Mongolia
MNP Northern Mariana Islands
MOZ Mozambique
MRT Mauritania
MUS Mauritius
MWI Malawi
MYS Malaysia
NAC North America
NAF North Africa
NAM Namibia
NCL New Caledonia
NER Niger
NGA Nigeria
NIC Nicaragua
NLD Netherlands
NOR Norway
NPL Nepal
NRS Non-resource rich Sub-Saharan Africa countries
NRU Nauru
NXS IDA countries not classified as fragile situations, excluding Sub-Saharan Africa
NZL New Zealand
OED OECD members
OMN Oman
OSS Other small states
PAK Pakistan
PAN Panama
PER Peru
PHL Philippines
PLW Palau
PNG Papua New Guinea
POL Poland
PRE Pre-demographic dividend
PRI Puerto Rico
PRK Korea, Dem. People's Rep.
PRT Portugal
PRY Paraguay
PSE West Bank and Gaza
PSS Pacific island small states
PST Post-demographic dividend
PYF French Polynesia
QAT Qatar
ROU Romania
RRS Resource rich Sub-Saharan Africa countries
RUS Russian Federation
RWA Rwanda
SAS South Asia

SAU Saudi Arabia
SDN Sudan
SEN Senegal
SGP Singapore
SLB Solomon Islands
SLE Sierra Leone
SLV El Salvador
SMR San Marino
SOM Somalia
SRB Serbia
SSA Sub-Saharan Africa (excluding high income)
SSD South Sudan
SSF Sub-Saharan Africa
SST Small states
STP Sao Tome and Principe
SUR Suriname
SVK Slovak Republic
SVN Slovenia
SWE Sweden
SWZ Eswatini
SXM Sint Maarten (Dutch part)
SXZ Sub-Saharan Africa excluding South Africa
SYC Seychelles
SYR Syrian Arab Republic
TCA Turks and Caicos Islands
TCD Chad
TEA East Asia & Pacific (IDA & IBRD countries)
TEC Europe & Central Asia (IDA & IBRD countries)
TGO Togo
THA Thailand
TJK Tajikistan
TKM Turkmenistan
TLA Latin America & the Caribbean (IDA & IBRD countries)
TLS Timor-Leste
TMN Middle East & North Africa (IDA & IBRD countries)
TON Tonga
TSA South Asia (IDA & IBRD)
TSS Sub-Saharan Africa (IDA & IBRD countries)
TTO Trinidad and Tobago
TUN Tunisia
TUR Turkey
TUV Tuvalu
TWN Taiwan, China
TZA Tanzania
UGA Uganda
UKR Ukraine
UMC Upper middle income

```
URY   Uruguay
USA   United States
UZB   Uzbekistan
VCT   St. Vincent and the Grenadines
VEN   Venezuela, RB
VGB   British Virgin Islands
VIR   Virgin Islands (U.S.)
VNM   Vietnam
VUT   Vanuatu
WLD   World
WSM   Samoa
XKX   Kosovo
XZN   Sub-Saharan Africa excluding South Africa and Nigeria
YEM   Yemen, Rep.
ZAF   South Africa
ZMB   Zambia
ZWE   Zimbabwe
```

To see possible datasets we can access via the API, use `get_source()`

[3]:
```
wbdata.get_source()
```

[3]:
```
  id  name
----  ------------------------------------------------------------------------
   1  Doing Business
   2  World Development Indicators
   3  Worldwide Governance Indicators
   5  Subnational Malnutrition Database
   6  International Debt Statistics
  11  Africa Development Indicators
  12  Education Statistics
  13  Enterprise Surveys
  14  Gender Statistics
  15  Global Economic Monitor
  16  Health Nutrition and Population Statistics
  18  IDA Results Measurement System
  19  Millennium Development Goals
  20  Quarterly Public Sector Debt
  22  Quarterly External Debt Statistics SDDS
  23  Quarterly External Debt Statistics GDDS
  24  Poverty and Equity
  25  Jobs
  27  Global Economic Prospects
  28  Global Financial Inclusion
  29  The Atlas of Social Protection: Indicators of Resilience and Equity
  30  Exporter Dynamics Database - Indicators at Country-Year Level
  31  Country Policy and Institutional Assessment
```

```
32  Global Financial Development
33  G20 Financial Inclusion Indicators
34  Global Partnership for Education
35  Sustainable Energy for All
36  Statistical Capacity Indicators
37  LAC Equity Lab
38  Subnational Poverty
39  Health Nutrition and Population Statistics by Wealth Quintile
40  Population estimates and projections
41  Country Partnership Strategy for India (FY2013 - 17)
43  Adjusted Net Savings
44  Readiness for Investment in Sustainable Energy
45  Indonesia Database for Policy and Economic Research
46  Sustainable Development Goals
50  Subnational Population
54  Joint External Debt Hub
57  WDI Database Archives
58  Universal Health Coverage
59  Wealth Accounts
60  Economic Fitness
61  PPPs Regulatory Quality
62  International Comparison Program (ICP) 2011
63  Human Capital Index
64  Worldwide Bureaucracy Indicators
65  Health Equity and Financial Protection Indicators
66  Logistics Performance Index
67  PEFA 2011
68  PEFA 2016
69  Global Financial Inclusion and Consumer Protection Survey
70  Economic Fitness 2
71  International Comparison Program (ICP) 2005
72  PEFA_Test
73  Global Financial Inclusion and Consumer Protection Survey (Internal)
75  Environment, Social and Governance (ESG) Data
76  Remittance Prices Worldwide (Sending Countries)
77  Remittance Prices Worldwide (Receiving Countries)
78  ICP 2017
79  PEFA_GRPFM
80  Gender Disaggregated Labor Database (GDLD)
81  International Debt Statistics: DSSI
```

[4]:
```
SOURCE = 40 # "Population estimates and projections"

indicators_40 = wbdata.get_indicator(source=SOURCE)
indicators_40
```

[4]:
```
id                 name
-----------------
--------------------------------------------------------------------
SH.DTH.0509        Number of deaths ages 5-9 years
SH.DTH.1014        Number of deaths ages 10-14 years
SH.DTH.1519        Number of deaths ages 15-19 years
SH.DTH.2024        Number of deaths ages 20-24 years
SH.DTH.IMRT        Number of infant deaths
SH.DTH.IMRT.FE     Number of infant deaths, female
SH.DTH.IMRT.MA     Number of infant deaths, male
SH.DTH.MORT        Number of under-five deaths
SH.DTH.MORT.FE     Number of under-five deaths, female
SH.DTH.MORT.MA     Number of under-five deaths, male
SH.DTH.NMRT        Number of neonatal deaths
SH.DYN.0509        Probability of dying among children ages 5-9 years (per
1,000)
SH.DYN.1014        Probability of dying among adolescents ages 10-14 years (per
1,000)
SH.DYN.1519        Probability of dying among adolescents ages 15-19 years (per
1,000)
SH.DYN.2024        Probability of dying among youth ages 20-24 years (per 1,000)
SH.DYN.MORT        Mortality rate, under-5 (per 1,000 live births)
SH.DYN.MORT.FE     Mortality rate, under-5, female (per 1,000 live births)
SH.DYN.MORT.MA     Mortality rate, under-5, male (per 1,000 live births)
SH.DYN.NMRT        Mortality rate, neonatal (per 1,000 live births)
SM.POP.NETM        Net migration
SP.DYN.AMRT.FE     Mortality rate, adult, female (per 1,000 female adults)
SP.DYN.AMRT.MA     Mortality rate, adult, male (per 1,000 male adults)
SP.DYN.CBRT.IN     Birth rate, crude (per 1,000 people)
SP.DYN.CDRT.IN     Death rate, crude (per 1,000 people)
SP.DYN.IMRT.FE.IN  Mortality rate, infant, female (per 1,000 live births)
SP.DYN.IMRT.IN     Mortality rate, infant (per 1,000 live births)
SP.DYN.IMRT.MA.IN  Mortality rate, infant, male (per 1,000 live births)
SP.DYN.LE00.FE.IN  Life expectancy at birth, female (years)
SP.DYN.LE00.IN     Life expectancy at birth, total (years)
SP.DYN.LE00.MA.IN  Life expectancy at birth, male (years)
SP.DYN.TFRT.IN     Fertility rate, total (births per woman)
SP.POP.0004.FE     Population ages 00-04, female
SP.POP.0004.FE.5Y  Population ages 00-04, female (% of female population)
SP.POP.0004.MA     Population ages 00-04, male
SP.POP.0004.MA.5Y  Population ages 00-04, male (% of male population)
SP.POP.0014.FE.IN  Population ages 0-14, female
SP.POP.0014.FE.ZS  Population ages 0-14, female (% of female population)
SP.POP.0014.MA.IN  Population ages 0-14, male
SP.POP.0014.MA.ZS  Population ages 0-14, male (% of male population)
SP.POP.0014.TO     Population ages 0-14, total
SP.POP.0014.TO.ZS  Population ages 0-14 (% of total population)
```

```
SP.POP.0509.FE     Population ages 05-09, female
SP.POP.0509.FE.5Y  Population ages 05-09, female (% of female population)
SP.POP.0509.MA     Population ages 05-09, male
SP.POP.0509.MA.5Y  Population ages 05-09, male (% of male population)
SP.POP.1014.FE     Population ages 10-14, female
SP.POP.1014.FE.5Y  Population ages 10-14, female (% of female population)
SP.POP.1014.MA     Population ages 10-14, male
SP.POP.1014.MA.5Y  Population ages 10-14, male (% of male population)
SP.POP.1519.FE     Population ages 15-19, female
SP.POP.1519.FE.5Y  Population ages 15-19, female (% of female population)
SP.POP.1519.MA     Population ages 15-19, male
SP.POP.1519.MA.5Y  Population ages 15-19, male (% of male population)
SP.POP.1564.FE.IN  Population ages 15-64, female
SP.POP.1564.FE.ZS  Population ages 15-64, female (% of female population)
SP.POP.1564.MA.IN  Population ages 15-64, male
SP.POP.1564.MA.ZS  Population ages 15-64, male (% of male population)
SP.POP.1564.TO     Population ages 15-64, total
SP.POP.1564.TO.ZS  Population ages 15-64 (% of total population)
SP.POP.2024.FE     Population ages 20-24, female
SP.POP.2024.FE.5Y  Population ages 20-24, female (% of female population)
SP.POP.2024.MA     Population ages 20-24, male
SP.POP.2024.MA.5Y  Population ages 20-24, male (% of male population)
SP.POP.2529.FE     Population ages 25-29, female
SP.POP.2529.FE.5Y  Population ages 25-29, female (% of female population)
SP.POP.2529.MA     Population ages 25-29, male
SP.POP.2529.MA.5Y  Population ages 25-29, male (% of male population)
SP.POP.3034.FE     Population ages 30-34, female
SP.POP.3034.FE.5Y  Population ages 30-34, female (% of female population)
SP.POP.3034.MA     Population ages 30-34, male
SP.POP.3034.MA.5Y  Population ages 30-34, male (% of male population)
SP.POP.3539.FE     Population ages 35-39, female
SP.POP.3539.FE.5Y  Population ages 35-39, female (% of female population)
SP.POP.3539.MA     Population ages 35-39, male
SP.POP.3539.MA.5Y  Population ages 35-39, male (% of male population)
SP.POP.4044.FE     Population ages 40-44, female
SP.POP.4044.FE.5Y  Population ages 40-44, female (% of female population)
SP.POP.4044.MA     Population ages 40-44, male
SP.POP.4044.MA.5Y  Population ages 40-44, male (% of male population)
SP.POP.4549.FE     Population ages 45-49, female
SP.POP.4549.FE.5Y  Population ages 45-49, female (% of female population)
SP.POP.4549.MA     Population ages 45-49, male
SP.POP.4549.MA.5Y  Population ages 45-49, male (% of male population)
SP.POP.5054.FE     Population ages 50-54, female
SP.POP.5054.FE.5Y  Population ages 50-54, female (% of female population)
SP.POP.5054.MA     Population ages 50-54, male
SP.POP.5054.MA.5Y  Population ages 50-54, male (% of male population)
SP.POP.5559.FE     Population ages 55-59, female
```

SP.POP.5559.FE.5Y	Population ages 55-59, female (% of female population)
SP.POP.5559.MA	Population ages 55-59, male
SP.POP.5559.MA.5Y	Population ages 55-59, male (% of male population)
SP.POP.6064.FE	Population ages 60-64, female
SP.POP.6064.FE.5Y	Population ages 60-64, female (% of female population)
SP.POP.6064.MA	Population ages 60-64, male
SP.POP.6064.MA.5Y	Population ages 60-64, male (% of male population)
SP.POP.6569.FE	Population ages 65-69, female
SP.POP.6569.FE.5Y	Population ages 65-69, female (% of female population)
SP.POP.6569.MA	Population ages 65-69, male
SP.POP.6569.MA.5Y	Population ages 65-69, male (% of male population)
SP.POP.65UP.FE.IN	Population ages 65 and above, female
SP.POP.65UP.FE.ZS	Population ages 65 and above, female (% of female population)
SP.POP.65UP.MA.IN	Population ages 65 and above, male
SP.POP.65UP.MA.ZS	Population ages 65 and above, male (% of male population)
SP.POP.65UP.TO	Population ages 65 and above, total
SP.POP.65UP.TO.ZS	Population ages 65 and above (% of total population)
SP.POP.7074.FE	Population ages 70-74, female
SP.POP.7074.FE.5Y	Population ages 70-74, female (% of female population)
SP.POP.7074.MA	Population ages 70-74, male
SP.POP.7074.MA.5Y	Population ages 70-74, male (% of male population)
SP.POP.7579.FE	Population ages 75-79, female
SP.POP.7579.FE.5Y	Population ages 75-79, female (% of female population)
SP.POP.7579.MA	Population ages 75-79, male
SP.POP.7579.MA.5Y	Population ages 75-79, male (% of male population)
SP.POP.80UP.FE	Population ages 80 and above, female
SP.POP.80UP.FE.5Y	Population ages 80 and above, female (% of female population)
SP.POP.80UP.MA	Population ages 80 and above, male
SP.POP.80UP.MA.5Y	Population ages 80 and above, male (% of male population)
SP.POP.AG00.FE.IN	Age population, age 00, female, interpolated
SP.POP.AG00.MA.IN	Age population, age 00, male, interpolated
SP.POP.AG01.FE.IN	Age population, age 01, female, interpolated
SP.POP.AG01.MA.IN	Age population, age 01, male, interpolated
SP.POP.AG02.FE.IN	Age population, age 02, female, interpolated
SP.POP.AG02.MA.IN	Age population, age 02, male, interpolated
SP.POP.AG03.FE.IN	Age population, age 03, female, interpolated
SP.POP.AG03.MA.IN	Age population, age 03, male, interpolated
SP.POP.AG04.FE.IN	Age population, age 04, female, interpolated
SP.POP.AG04.MA.IN	Age population, age 04, male, interpolated
SP.POP.AG05.FE.IN	Age population, age 05, female, interpolated
SP.POP.AG05.MA.IN	Age population, age 05, male, interpolated
SP.POP.AG06.FE.IN	Age population, age 06, female, interpolated
SP.POP.AG06.MA.IN	Age population, age 06, male, interpolated
SP.POP.AG07.FE.IN	Age population, age 07, female, interpolated
SP.POP.AG07.MA.IN	Age population, age 07, male, interpolated
SP.POP.AG08.FE.IN	Age population, age 08, female, interpolated
SP.POP.AG08.MA.IN	Age population, age 08, male, interpolated

SP.POP.AG09.FE.IN	Age population, age 09, female, interpolated
SP.POP.AG09.MA.IN	Age population, age 09, male, interpolated
SP.POP.AG10.FE.IN	Age population, age 10, female, interpolated
SP.POP.AG10.MA.IN	Age population, age 10, male, interpolated
SP.POP.AG11.FE.IN	Age population, age 11, female, interpolated
SP.POP.AG11.MA.IN	Age population, age 11, male, interpolated
SP.POP.AG12.FE.IN	Age population, age 12, female, interpolated
SP.POP.AG12.MA.IN	Age population, age 12, male, interpolated
SP.POP.AG13.FE.IN	Age population, age 13, female, interpolated
SP.POP.AG13.MA.IN	Age population, age 13, male, interpolated
SP.POP.AG14.FE.IN	Age population, age 14, female, interpolated
SP.POP.AG14.MA.IN	Age population, age 14, male, interpolated
SP.POP.AG15.FE.IN	Age population, age 15, female, interpolated
SP.POP.AG15.MA.IN	Age population, age 15, male, interpolated
SP.POP.AG16.FE.IN	Age population, age 16, female, interpolated
SP.POP.AG16.MA.IN	Age population, age 16, male, interpolated
SP.POP.AG17.FE.IN	Age population, age 17, female, interpolated
SP.POP.AG17.MA.IN	Age population, age 17, male, interpolated
SP.POP.AG18.FE.IN	Age population, age 18, female, interpolated
SP.POP.AG18.MA.IN	Age population, age 18, male, interpolated
SP.POP.AG19.FE.IN	Age population, age 19, female, interpolated
SP.POP.AG19.MA.IN	Age population, age 19, male, interpolated
SP.POP.AG20.FE.IN	Age population, age 20, female, interpolated
SP.POP.AG20.MA.IN	Age population, age 20, male, interpolated
SP.POP.AG21.FE.IN	Age population, age 21, female, interpolated
SP.POP.AG21.MA.IN	Age population, age 21, male, interpolated
SP.POP.AG22.FE.IN	Age population, age 22, female, interpolated
SP.POP.AG22.MA.IN	Age population, age 22, male, interpolated
SP.POP.AG23.FE.IN	Age population, age 23, female, interpolated
SP.POP.AG23.MA.IN	Age population, age 23, male, interpolated
SP.POP.AG24.FE.IN	Age population, age 24, female, interpolated
SP.POP.AG24.MA.IN	Age population, age 24, male, interpolated
SP.POP.AG25.FE.IN	Age population, age 25, female, interpolated
SP.POP.AG25.MA.IN	Age population, age 25, male, interpolated
SP.POP.BRTH.MF	Sex ratio at birth (male births per female births)
SP.POP.DPND	Age dependency ratio (% of working-age population)
SP.POP.DPND.OL	Age dependency ratio, old (% of working-age population)
SP.POP.DPND.YG	Age dependency ratio, young (% of working-age population)
SP.POP.GROW	Population growth (annual %)
SP.POP.TOTL	Population, total
SP.POP.TOTL.FE.IN	Population, female
SP.POP.TOTL.FE.ZS	Population, female (% of total population)
SP.POP.TOTL.MA.IN	Population, male
SP.POP.TOTL.MA.ZS	Population, male (% of total population)
SP.RUR.TOTL	Rural population
SP.RUR.TOTL.ZG	Rural population growth (annual %)
SP.RUR.TOTL.ZS	Rural population (% of total population)

```
SP.URB.GROW        Urban population growth (annual %)
SP.URB.TOTL        Urban population
SP.URB.TOTL.IN.ZS  Urban population (% of total population)
```

0.2 Population function

[5]:
```
def population(sex, year, age, country):
    """Function that takes in a SEX ("Male", "Female"), ... , COUNTRY=, ..
    and returns a statistic for the given function arguments"""
    if sex == "Male":
        variable_labels = {"SP.POP." + str(age[0])+str(age[1]) + ".MA": sex}
    elif sex=="Female":
        variable_labels = {"SP.POP." + str(age[0])+str(age[1]) + ".FE": sex}
    pop_stats = wbdata.get_dataframe(variable_labels, country=country)
    pop_stats = pop_stats.filter(like=str(year), axis=0)
    return pop_stats[sex][0]
```

[6]:
```
population("Female", 2010, (15,19), "CHN")
```

[6]:
```
45907253.0
```

[7]:
```
def population_dataframes(indicators):
    """Returns a pandas DataFrame indexed by Region or Country and Year,
    with columns giving counts of people in different age-sex groups."""
    # Create a dictionary of all corresponding ID's and keys in SOURCE 40 from␣
↪WBDATA
    labels = {}
    for i in range(len(indicators)):
        col_id = indicators[i]['id']
        col_name = indicators[i]['name']
        labels[col_id] = col_name

    def find_labels(indicators):
        """Helper function that akes in an indicator object to filter through␣
↪variable
        keys and parse through necessary ID's to obtain relevant data. """
        # Filter through column ID strings to obtain relevant population data.
        r = re.compile("(SP.POP).[\d]{2}[A-ZO-9]{2}.[MAFE]{2}$")
        col_keys = list(filter(r.match, labels))

        # Add total population column at the end of COL_KEYS
        col_keys.append('SP.POP.TOTL')

        labels_filtered = {}
        for key, value in labels.items():
            if key in col_keys:
```

```
                labels_filtered[key] = value
        return labels_filtered

    df_labels = find_labels(indicators)
    return wbdata.get_dataframe(df_labels)
```

0.3 Population DataFrames

[8]:
```
pop_df = population_dataframes(indicators_40)
pop_df.head()
```

[8]:

country	date	Population ages 00-04, female	Population ages 00-04, male	\
Afghanistan	1960	760938.0	780471.0	
	1961	795378.0	808289.0	
	1962	818678.0	831163.0	
	1963	834934.0	851787.0	
	1964	850992.0	873022.0	

country	date	Population ages 05-09, female	Population ages 05-09, male	\
Afghanistan	1960	583953.0	598721.0	
	1961	594819.0	617341.0	
	1962	611717.0	637537.0	
	1963	632901.0	657620.0	
	1964	654635.0	676153.0	

country	date	Population ages 10-14, female	Population ages 10-14, male	\
Afghanistan	1960	544194.0	523122.0	
	1961	547964.0	528908.0	
	1962	550019.0	538026.0	
	1963	551891.0	550452.0	
	1964	556047.0	565257.0	

country	date	Population ages 15-19, female	Population ages 15-19, male	\
Afghanistan	1960	447872.0	478075.0	
	1961	460206.0	481616.0	
	1962	477029.0	485476.0	
	1963	495604.0	490025.0	
	1964	511923.0	496099.0	

country	date	Population ages 20-24, female	Population ages 20-24, male	\
Afghanistan	1960	382542.0	415970.0	
	1961	388240.0	421947.0	

```
            1962                      394896.0                      429297.0
            1963                      402961.0                      437345.0
            1964                      413203.0                      445246.0

                  …  Population ages 60-64, male  \
country     date  …
Afghanistan 1960  …                      93053.0
            1961  …                      93359.0
            1962  …                      94140.0
            1963  …                      95186.0
            1964  …                      96187.0

                  Population ages 65-69, female  Population ages 65-69, male  \
country     date
Afghanistan 1960                      56624.0                      67283.0
            1961                      57393.0                      67330.0
            1962                      58326.0                      67197.0
            1963                      59399.0                      67052.0
            1964                      60569.0                      67117.0

                  Population ages 70-74, female  Population ages 70-74, male  \
country     date
Afghanistan 1960                      34655.0                      40748.0
            1961                      35363.0                      41520.0
            1962                      36055.0                      42186.0
            1963                      36757.0                      42704.0
            1964                      37461.0                      42987.0

                  Population ages 75-79, female  Population ages 75-79, male  \
country     date
Afghanistan 1960                      16990.0                      19683.0
            1961                      17750.0                      20500.0
            1962                      18364.0                      21050.0
            1963                      18819.0                      21331.0
            1964                      19098.0                      21378.0

                  Population ages 80 and above, female  \
country     date
Afghanistan 1960                             7486.0
            1961                             8358.0
            1962                             9041.0
            1963                             9456.0
            1964                             9513.0

                  Population ages 80 and above, male  Population, total
country     date
Afghanistan 1960                            8294.0          8996973.0
```

```
                1961                                9283.0          9169410.0
                1962                               10013.0          9351441.0
                1963                               10376.0          9543205.0
                1964                               10269.0          9744781.0

[5 rows x 35 columns]
```

0.3.1 Cleaning the data

[9]:
```
# Change "date" index to type INT.
pop_df.index = pop_df.index.set_levels(pop_df.index.levels[1].astype(int),␣
 ↪level=1)

# Rename MultiIndex column "date" to "year"
pop_df = pop_df.rename_axis(index=['country', 'year'])

# Delete 2020 row for every country code
pop_df = pop_df[~pop_df.index.get_level_values('year').isin([2020])]

# Add population total breakdowns for male and female
cols_f = [col for col in pop_df.columns if 'female' in col]
cols_m = [col for col in pop_df.columns if col not in cols_f]
pop_df['Population, total female'] = pop_df.loc[:,cols_f].sum(axis=1)
pop_df['Population, total male'] = pop_df.loc[:, cols_m].sum(axis=1)
pop_df.head()
```

[9]:
```
                    Population ages 00-04, female  Population ages 00-04, male  \
country     year
Afghanistan 1960                         760938.0                     780471.0
            1961                         795378.0                     808289.0
            1962                         818678.0                     831163.0
            1963                         834934.0                     851787.0
            1964                         850992.0                     873022.0

                    Population ages 05-09, female  Population ages 05-09, male  \
country     year
Afghanistan 1960                         583953.0                     598721.0
            1961                         594819.0                     617341.0
            1962                         611717.0                     637537.0
            1963                         632901.0                     657620.0
            1964                         654635.0                     676153.0

                    Population ages 10-14, female  Population ages 10-14, male  \
country     year
Afghanistan 1960                         544194.0                     523122.0
            1961                         547964.0                     528908.0
            1962                         550019.0                     538026.0
```

```
            1963                       551891.0                       550452.0
            1964                       556047.0                       565257.0

                  Population ages 15-19, female  Population ages 15-19, male  \
country     year
Afghanistan 1960                       447872.0                       478075.0
            1961                       460206.0                       481616.0
            1962                       477029.0                       485476.0
            1963                       495604.0                       490025.0
            1964                       511923.0                       496099.0

                  Population ages 20-24, female  Population ages 20-24, male  \
country     year
Afghanistan 1960                       382542.0                       415970.0
            1961                       388240.0                       421947.0
            1962                       394896.0                       429297.0
            1963                       402961.0                       437345.0
            1964                       413203.0                       445246.0

                  …  Population ages 65-69, male  \
country     year  …
Afghanistan 1960  …                      67283.0
            1961  …                      67330.0
            1962  …                      67197.0
            1963  …                      67052.0
            1964  …                      67117.0

                  Population ages 70-74, female  Population ages 70-74, male  \
country     year
Afghanistan 1960                       34655.0                       40748.0
            1961                       35363.0                       41520.0
            1962                       36055.0                       42186.0
            1963                       36757.0                       42704.0
            1964                       37461.0                       42987.0

                  Population ages 75-79, female  Population ages 75-79, male  \
country     year
Afghanistan 1960                       16990.0                       19683.0
            1961                       17750.0                       20500.0
            1962                       18364.0                       21050.0
            1963                       18819.0                       21331.0
            1964                       19098.0                       21378.0

                  Population ages 80 and above, female  \
country     year
Afghanistan 1960                               7486.0
            1961                               8358.0
```

```
                1962                                    9041.0
                1963                                    9456.0
                1964                                    9513.0

                      Population ages 80 and above, male  Population, total  \
country     year
Afghanistan 1960                                  8294.0          8996973.0
            1961                                  9283.0          9169410.0
            1962                                 10013.0          9351441.0
            1963                                 10376.0          9543205.0
            1964                                 10269.0          9744781.0

                      Population, total female  Population, total male
country     year
Afghanistan 1960                    4347395.0                13646547.0
            1961                    4439156.0                13899660.0
            1962                    4535392.0                14167491.0
            1963                    4636170.0                14450236.0
            1964                    4741529.0                14748029.0

[5 rows x 37 columns]
```

0.4 Comparing Population Growth Rates of China and India

[10]:
```
# Rename DATE to YEAR
def growth_df(region):
    df=wbdata.get_dataframe({'SP.POP.GROW': 'Total Population growth (annual␣
 ↪%)',
                             'SP.RUR.TOTL.ZG': 'Rural population growth (annual␣
 ↪%)',
                             'SP.URB.GROW': 'Urban population growth (annual␣
 ↪%)'}, country=region)
    df.index = df.index.rename('year')
    df.index = df.index.astype(int)
    df = df.drop([2020])
    return df
```

[11]:
```
chn_gr = growth_df('CHN')
chn_gr.head()
```

[11]:
```
      Total Population growth (annual %)  Rural population growth (annual %)  \
year
2019                            0.357291                           -2.513529
2018                            0.455900                           -2.420469
2017                            0.559121                           -2.310812
2016                            0.541479                           -2.275352
```

```
2015                              0.508137                              -2.242449

      Urban population growth (annual %)
year
2019                              2.292727
2018                              2.491628
2017                              2.693540
2016                              2.744069
2015                              2.769551
```

[12]:
```
ind_gr = growth_df('IND')
ind_gr.head()
```

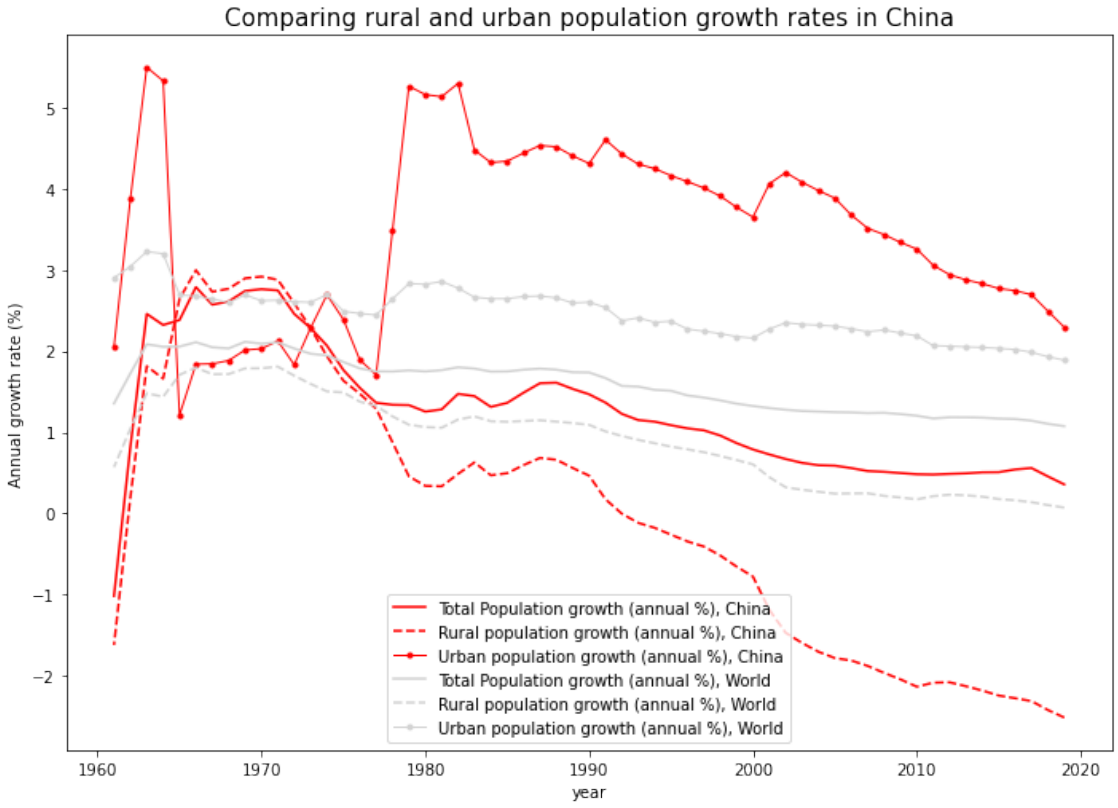
[12]:
```
      Total Population growth (annual %)  Rural population growth (annual %)  \
year
2019                              1.015106                              0.342850
2018                              1.037323                              0.387627
2017                              1.062597                              0.435052
2016                              1.089880                              0.485585
2015                              1.116633                              0.533714

      Urban population growth (annual %)
year
2019                              2.305597
2018                              2.308965
2017                              2.314448
2016                              2.317931
2015                              2.322891
```

[13]:
```
wld_gr = growth_df('WLD')
wld_gr.head()
```

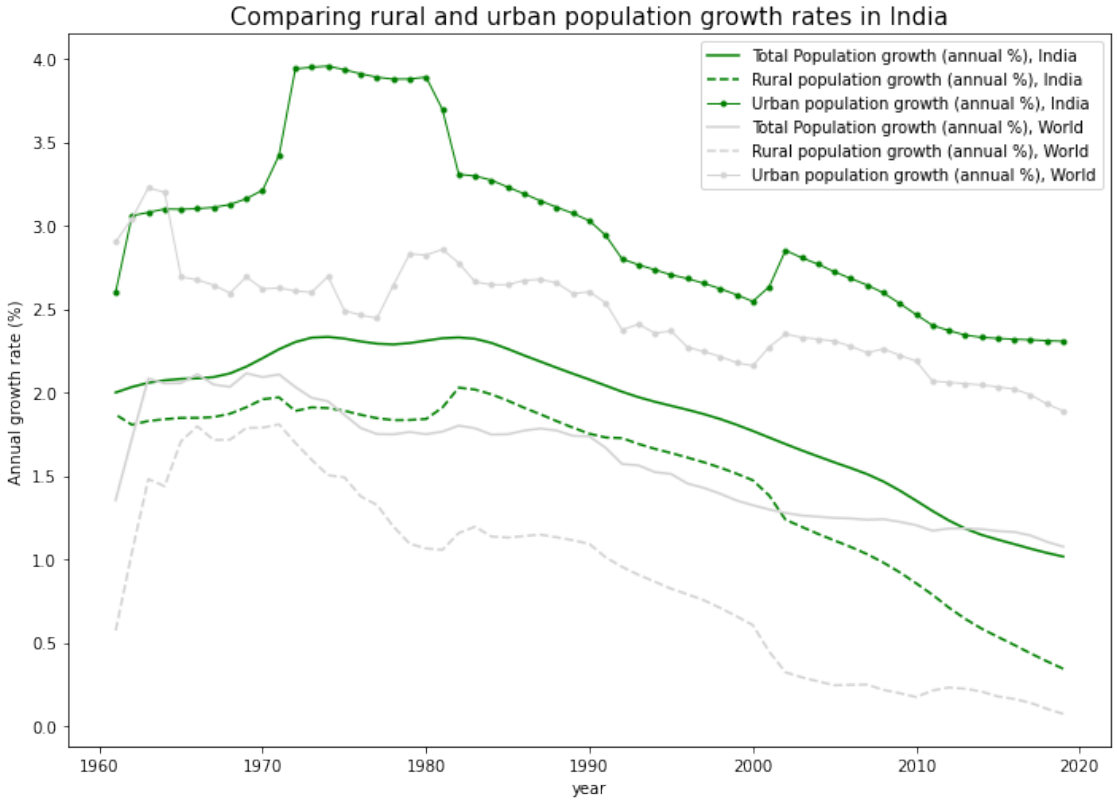
[13]:
```
      Total Population growth (annual %)  Rural population growth (annual %)  \
year
2019                              1.074675                              0.072327
2018                              1.103609                              0.102808
2017                              1.142040                              0.138527
2016                              1.162578                              0.162037
2015                              1.168120                              0.176929

      Urban population growth (annual %)
year
2019                              1.887461
2018                              1.929861
2017                              1.984537
2016                              2.018914
2015                              2.032497
```

[14]:
```
def urb_rur_plot(df, region_name, color):
    fig = plt.figure(figsize=(11.75, 8.25))
    ax = plt.axes()

    plt.plot(df.index, df.iloc[:, 0],  color=color)
    plt.plot(df.index, df.iloc[:, 1],  '--', color=color)
    plt.plot(df.index, df.iloc[:, 2],  '.-', color=color, linewidth=1)

    plt.plot(wld_gr.index, wld_gr.iloc[:, 0],  color='lightgray')
    plt.plot(wld_gr.index, wld_gr.iloc[:, 1],  '--', color='lightgray')
    plt.plot(wld_gr.index, wld_gr.iloc[:, 2],  '.-', color='lightgray',␣
↪linewidth=1);

    ax.legend((df.columns+", {}".format(region_name)).append(wld_gr.columns+",␣
↪World"))

    plt.title('Comparing rural and urban population growth rates in {}'.
↪format(region_name), {'fontsize': 15})
    plt.xlabel('year')
    plt.ylabel('Annual growth rate (%)')

    plt.show();
```

[15]:
```
urb_rur_plot(chn_gr, "China", 'red')
```



[16]:
```
urb_rur_plot(ind_gr, "India", 'green')
```



0.5 Population Pyramids

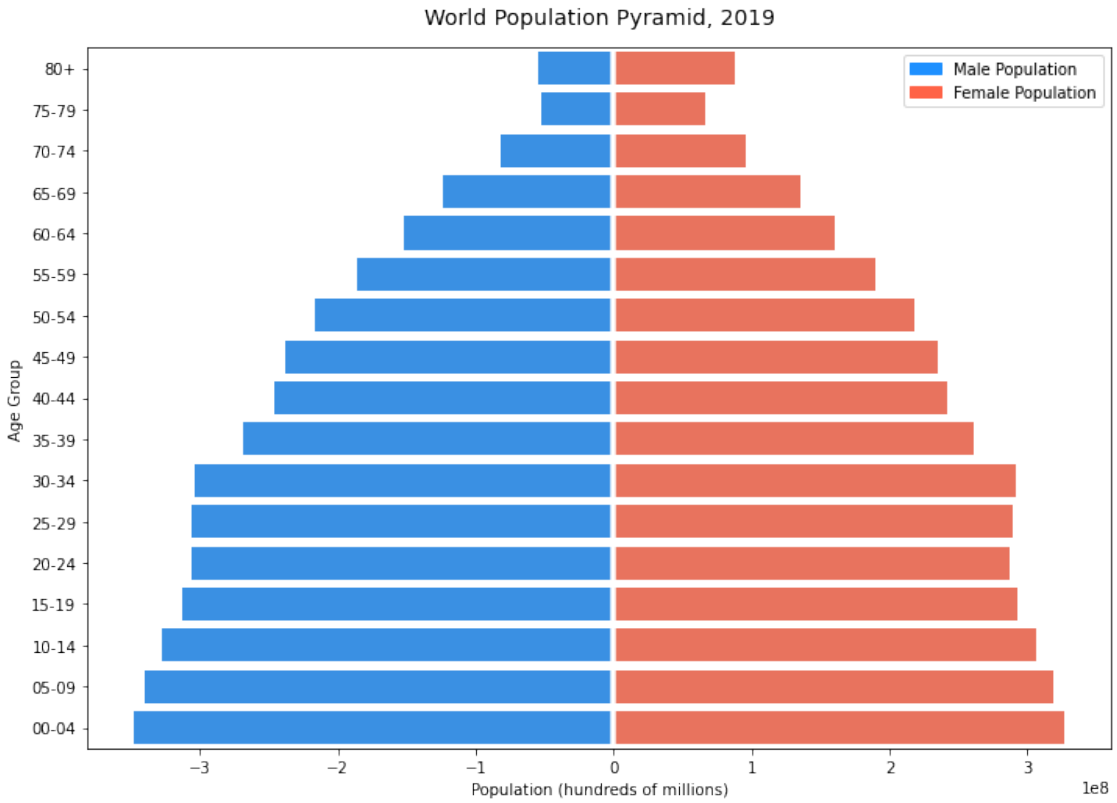
For the sake of simplicity, visualizations will use world data only.

[17]:
```
def generate_ppy_df(pop_df, country, year):
        """Takes in a population DataFrame and selects population
        data for a specific country/region in YEAR, then modifies the
        DataFrame in a way that can be visualized. """
        new_df = pop_df[pop_df.index.get_level_values('country').
↪isin([country])]
        country_df = new_df[new_df.index.get_level_values('year').isin([year])]
        ages = [col[16:21] for col in country_df.columns[::2]]
        ages = ages[:-3]
        ages.append('80+')
        f_pop = list(country_df.iloc[0])[::2][:-2]
        m_pop = list(country_df.iloc[0])[1::2][:-1]
        ppy_df = pd.DataFrame({'Age': ages,
```

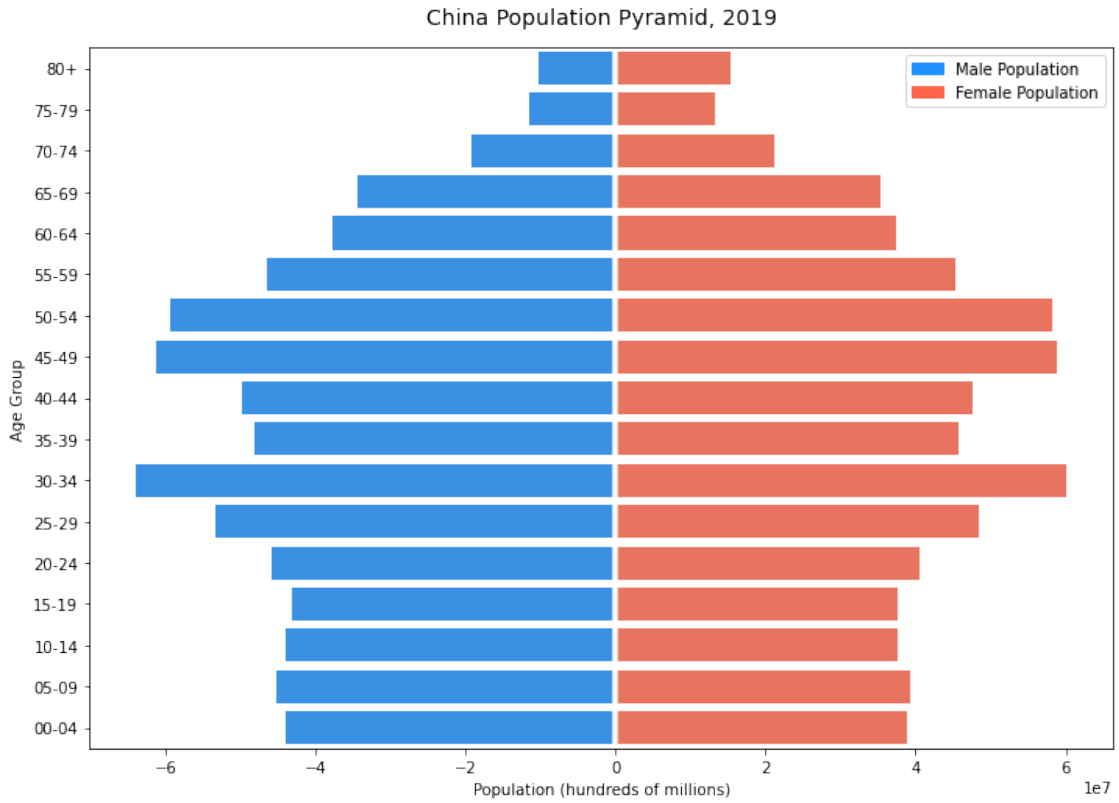
```
                        'Population, total female': f_pop,
                        'Population, total male': m_pop})
        # Convert male population data to negative values for displaying␣
↪purposes.
        ppy_df['Population, total male'] = ppy_df['Population, total male']*-1
        return ppy_df


def population_pyramid(region, year):
    """Takes as input a pandas DataFrame with columns providing counts
    of people by age-sex groups, and constructs a population pyramid graph
    for visualizing the data."""
    df=generate_ppy_df(pop_df, region, year)
    fig = plt.figure(figsize=(11.75,8.25))
    ax = plt.axes()

    f_patch = mpatches.Patch(color='tomato', label='Female Population')
    m_patch = mpatches.Patch(color='dodgerblue', label='Male Population')
    plt.legend(handles=[m_patch, f_patch],
               loc='upper right')

    ageAxis = df['Age'][::-1]
    fig = sns.barplot(x='Population, total female',
                            y='Age',
                            data=df,
                            order=ageAxis,
                            lw=0,
                            color='tomato')
    fig = sns.barplot(x='Population, total male',
                            y='Age',
                            data=df,
                            order=ageAxis,
                            lw=0,
                            color='dodgerblue')


    fig.axvline(lw=3.5, color='w')
    fig.set_title("{} Population Pyramid, {}".format(region, year),␣
↪fontdict={'fontsize':14,'fontweight':3}, pad=15)
    fig.set(xlabel="Population (hundreds of millions)", ylabel="Age Group");
```

Reorganized DataFrame to make suitable for graphing.

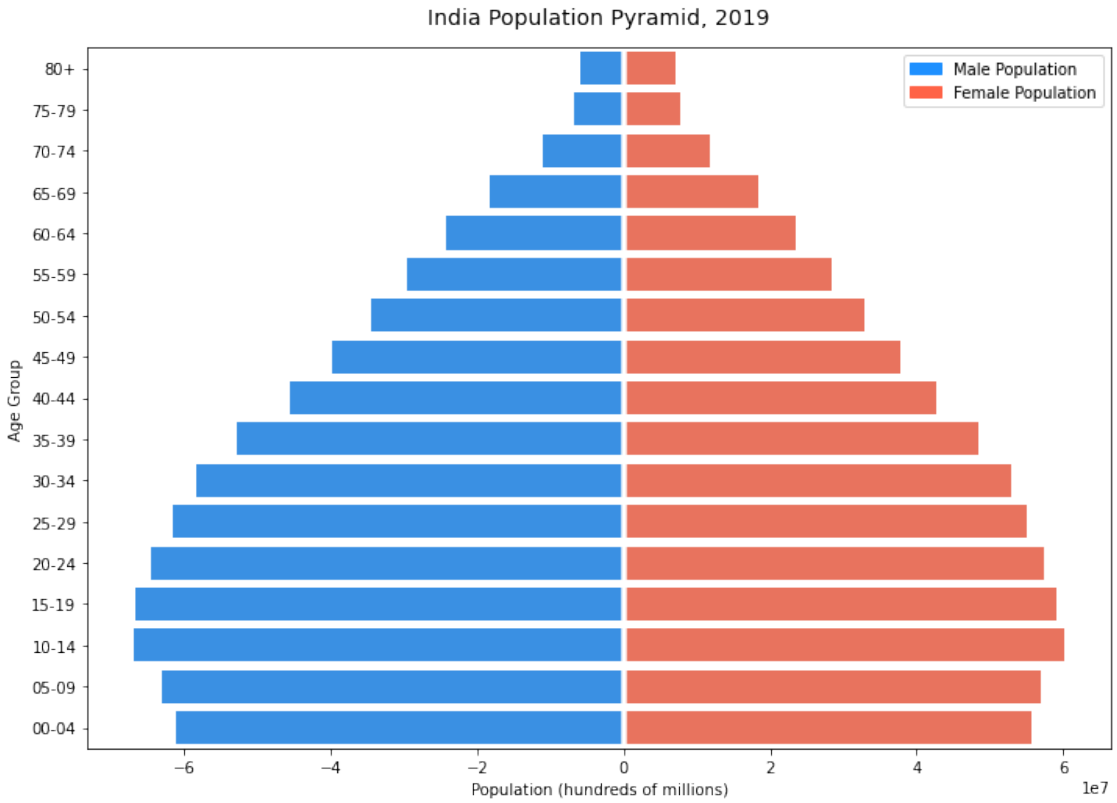
[18]:
```
population_pyramid("World", 2019)
```



[19]:
```
population_pyramid("China", 2019)
```



[20]:
```
population_pyramid("India", 2019)
```



0.6 Animated Population Pyramid

[27]:
```
%matplotlib notebook
from matplotlib import animation, rc
from IPython.display import HTML

Writer = animation.writers['ffmpeg']
writer = Writer(fps=60, metadata=dict(artist='Me'), bitrate=1800)

fig2 = plt.figure()
ax2 = plt.axes()
f_patch = mpatches.Patch(color='tomato', label='Female Population')
m_patch = mpatches.Patch(color='dodgerblue', label='Male Population')
plt.legend(handles=[m_patch, f_patch])

region = "World"       # Replace with desired region

def animate(i):
    df=generate_ppy_df(pop_df, region, 1960+i)
    ageAxis=df['Age'][::-1]
```

```
    p = sns.barplot(x='Population, total female', y='Age', data=df,␣
↪order=ageAxis, lw=0, color='tomato')
    p = sns.barplot(x='Population, total male', y='Age', data=df,␣
↪order=ageAxis, lw=0, color='dodgerblue')
    p.axvline(lw=3.5, color='w')
    p.set_title("{r} Population Pyramid, {y}".format(r=region, y=1960+i),␣
↪fontdict={'fontsize':14,'fontweight':3}, pad=15)
    plt.xlabel('Population (hundreds of millions)')
    plt.ylabel('Age Group')

anim=animation.
↪FuncAnimation(fig2,animate,blit=False,frames=60,interval=100,repeat=True)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[22]:
```
# anim.save('china_ppy.gif', writer=writer)     # Replace with appropriate␣
↪graph title
```

[ ]:
```

```