World Population

1. Introduction

The World Population Data dataset provides total population statistics for each country from 1960 to 2020, as well as some additional information by country, such as its region, income group, and special notes (if any).

The dataset can be found here

2. Setting up my environment

Load the 'tidyverse' package

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

Import the dataset

```
world_pop <- read_csv("world_pop_data.csv")
```

```
## Rows: 266 Columns: 64
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (3): Country Code, Indicator Name, Indicator Code
## dbl (61): 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
country <- read_csv("metadata_country.csv")
```

```
## Rows: 265 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (5): Country Code, Region, IncomeGroup, SpecialNotes, TableName
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

View the tables

```
head(world_pop)
```

```
## # A tibble: 6 x 64
##   `Country Code` `Indicator Name`   `Indicator Code` `1960` `1961` `1962` `1963`
##   <chr>          <chr>              <chr>             <dbl>  <dbl>  <dbl>  <dbl>
## 1 ABW            "\"Population, to~ SP.POP.TOTL      5.42e4 5.54e4 5.62e4 5.67e4
## 2 AFE            "\"Population, to~ SP.POP.TOTL      1.31e8 1.34e8 1.38e8 1.41e8
## 3 AFG            "\"Population, to~ SP.POP.TOTL      9.00e6 9.17e6 9.35e6 9.54e6
## 4 AFW            "\"Population, to~ SP.POP.TOTL      9.64e7 9.84e7 1.01e8 1.03e8
## 5 AGO            "\"Population, to~ SP.POP.TOTL      5.45e6 5.53e6 5.61e6 5.68e6
## 6 ALB            "\"Population, to~ SP.POP.TOTL      1.61e6 1.66e6 1.71e6 1.76e6
## # i 57 more variables: `1964` <dbl>, `1965` <dbl>, `1966` <dbl>, `1967` <dbl>,
## #   `1968` <dbl>, `1969` <dbl>, `1970` <dbl>, `1971` <dbl>, `1972` <dbl>,
## #   `1973` <dbl>, `1974` <dbl>, `1975` <dbl>, `1976` <dbl>, `1977` <dbl>,
## #   `1978` <dbl>, `1979` <dbl>, `1980` <dbl>, `1981` <dbl>, `1982` <dbl>,
## #   `1983` <dbl>, `1984` <dbl>, `1985` <dbl>, `1986` <dbl>, `1987` <dbl>,
## #   `1988` <dbl>, `1989` <dbl>, `1990` <dbl>, `1991` <dbl>, `1992` <dbl>,
## #   `1993` <dbl>, `1994` <dbl>, `1995` <dbl>, `1996` <dbl>, `1997` <dbl>, ...
```

```
head(country)
```

```
## # A tibble: 6 x 5
##   `Country Code` Region                    IncomeGroup    SpecialNotes TableName
##   <chr>          <chr>                     <chr>          <chr>        <chr>
## 1 ABW            Latin America & Caribbean High income    "null"       Aruba
## 2 AFE            null                      null           "\"26 count~ Africa E~
## 3 AFG            South Asia                Low income     "Fiscal yea~ Afghanis~
## 4 AFW            null                      null           "\"22 count~ Africa W~
## 5 AGO            Sub-Saharan Africa        Lower middle ~ "null"       Angola
## 6 ALB            Europe & Central Asia     Upper middle ~ "null"       Albania
```

Merge the two tables

```
pop_and_country_wide <- merge(x=world_pop, y=country, by="Country Code", all.x=TRUE)
head(pop_and_country_wide)
```

```
##   Country Code      Indicator Name Indicator Code      1960      1961      1962
## 1          ABW "Population, total"    SP.POP.TOTL     54208     55434     56234
## 2          AFE "Population, total"    SP.POP.TOTL 130836765 134159786 137614644
## 3          AFG "Population, total"    SP.POP.TOTL   8996967   9169406   9351442
## 4          AFW "Population, total"    SP.POP.TOTL  96396419  98407221 100506960
## 5          AGO "Population, total"    SP.POP.TOTL   5454938   5531451   5608499
## 6          ALB "Population, total"    SP.POP.TOTL   1608800   1659800   1711319
##        1963      1964      1965      1966      1967      1968      1969
## 1     56699     57029     57357     57702     58044     58377     58734
## 2 141202036 144920186 148769974 152752671 156876454 161156430 165611760
## 3   9543200   9744772   9956318  10174840  10399936  10637064  10893772
## 4 102691339 104953470 107289875 109701811 112195950 114781116 117468741
## 5   5679409   5734995   5770573   5781305   5774440   5771973   5803677
## 6   1762621   1814135   1864791   1914573   1965598   2022272   2081695
##        1970      1971      1972      1973      1974      1975      1976
## 1     59070     59442     59849     60236     60527     60653     60586
## 2 170257189 175100167 180141148 185376550 190800796 196409937 202205766
## 3  11173654  11475450  11791222  12108963  12412960  12689164  12943093
## 4 120269044 123184308 126218502 129384954 132699537 136173544 139813171
## 5   5890360   6041239   6248965   6497283   6761623   7023994   7279630
## 6   2135479   2187853   2243126   2296752   2350124   2404831   2458526
##        1977      1978      1979      1980      1981      1982      1983
```

```
## 1     60366     60102     59972     60097     60561     61341     62213
## 2 208193045 214368393 220740384 227305945 234058404 240999134 248146290
## 3  13171294  13341199  13411060  13356500  13171679  12882518  12537732
## 4 143615715 147571063 151663853 155882270 160223588 164689764 169279422
## 5   7533814   7790774   8058112   8341290   8640478   8952971   9278104
## 6   2513546   2566266   2617832   2671997   2726056   2784278   2843960
##        1984      1985      1986      1987      1988      1989      1990
## 1     62826     63024     62645     61838     61072     61033     62152
## 2 255530063 263161451 271050065 279184536 287524258 296024639 304648010
## 3  12204306  11938204  11736177  11604538  11618008  11868873  12412311
## 4 173991851 178826553 183785612 188868567 194070079 199382783 204803865
## 5   9614756   9961993  10320116  10689247  11068051  11454784  11848385
## 6   2904429   2964762   3022635   3083605   3142336   3227943   3286542
##        1991      1992      1993      1994      1995      1996      1997
## 1     64623     68240     72495     76705     80324     83211     85450
## 2 313394693 322270073 331265579 340379934 349605660 358953595 368440591
## 3  13299016  14485543  15816601  17075728  18110662  18853444  19357126
## 4 210332267 215976366 221754806 227692136 233807627 240114179 246613750
## 5  12248901  12657361  13075044  13503753  13945205  14400722  14871572
## 6   3266790   3247039   3227287   3207536   3187784   3168033   3148281
##        1998      1999      2000      2001      2002      2003      2004
## 1     87280     89009     90866     92892     94992     97016     98744
## 2 378098393 387977990 398113044 408522129 419223717 430246635 441630149
## 3  19737770  20170847  20779957  21606992  22600774  23680871  24726689
## 4 253302310 260170348 267214544 274433894 281842480 289469530 297353098
## 5  15359600  15866871  16395477  16945753  17519418  18121477  18758138
## 6   3128530   3108778   3089027   3060173   3051010   3039616   3026939
##        2005      2006      2007      2008      2009      2010      2011
## 1    100028    100830    101226    101362    101452    101665    102050
## 2 453404076 465581372 478166911 491173160 504604672 518468229 532760424
## 3  25654274  26433058  27100542  27722281  28394806  29185511  30117411
## 4 305520588 313985474 322741656 331772330 341050537 350556886 360285439
## 5  19433604  20149905  20905360  21695636  22514275  23356247  24220660
## 6   3011487   2992547   2970017   2947314   2927519   2913021   2905195
##        2012      2013      2014      2015      2016      2017      2018
## 1    102565    103165    103776    104339    104865    105361    105846
## 2 547482863 562601578 578075373 593871847 609978946 626392880 643090131
## 3  31161378  32269592  33370804  34413603  35383028  36296111  37171922
## 4 370243017 380437896 390882979 401586651 412551299 423769930 435229381
## 5  25107925  26015786  26941773  27884380  28842482  29816769  30809787
## 6   2900401   2895092   2889104   2880703   2876101   2873457   2866376
##        2019      2020                    Region         IncomeGroup
## 1    106310    106766 Latin America & Caribbean         High income
## 2 660046272 677243299                      null               null
## 3  38041757  38928341                South Asia          Low income
## 4 446911598 458803476                      null               null
## 5  31825299  32866268        Sub-Saharan Africa Lower middle income
## 6   2854191   2837743     Europe & Central Asia Upper middle income
##
## 1
## 2                                                                  "26 countries, stretching from th
## 3
## 4 "22 countries, stretching from the westernmost point of Africa, across the equator, and partly alo
## 5
```

```
## 6
##                     TableName
## 1                       Aruba
## 2 Africa Eastern and Southern
## 3                 Afghanistan
## 4  Africa Western and Central
## 5                      Angola
## 6                     Albania
```

```
colnames(pop_and_country_wide)
```

```
##  [1] "Country Code"   "Indicator Name" "Indicator Code" "1960"
##  [5] "1961"           "1962"           "1963"           "1964"
##  [9] "1965"           "1966"           "1967"           "1968"
## [13] "1969"           "1970"           "1971"           "1972"
## [17] "1973"           "1974"           "1975"           "1976"
## [21] "1977"           "1978"           "1979"           "1980"
## [25] "1981"           "1982"           "1983"           "1984"
## [29] "1985"           "1986"           "1987"           "1988"
## [33] "1989"           "1990"           "1991"           "1992"
## [37] "1993"           "1994"           "1995"           "1996"
## [41] "1997"           "1998"           "1999"           "2000"
## [45] "2001"           "2002"           "2003"           "2004"
## [49] "2005"           "2006"           "2007"           "2008"
## [53] "2009"           "2010"           "2011"           "2012"
## [57] "2013"           "2014"           "2015"           "2016"
## [61] "2017"           "2018"           "2019"           "2020"
## [65] "Region"         "IncomeGroup"    "SpecialNotes"   "TableName"
```

Reshape the data from wide to long

```
pop_and_country_long <- gather(pop_and_country_wide, Year, Population, `1960`:`2020`, factor_key = TRUE)
pop_and_country_long$Year=as.numeric(levels(pop_and_country_long$Year))[pop_and_country_long$Year]
head(pop_and_country_long)
```

```
##   Country Code      Indicator Name Indicator Code                    Region
## 1          ABW "Population, total"    SP.POP.TOTL Latin America & Caribbean
## 2          AFE "Population, total"    SP.POP.TOTL                      null
## 3          AFG "Population, total"    SP.POP.TOTL                South Asia
## 4          AFW "Population, total"    SP.POP.TOTL                      null
## 5          AGO "Population, total"    SP.POP.TOTL        Sub-Saharan Africa
## 6          ALB "Population, total"    SP.POP.TOTL     Europe & Central Asia
##           IncomeGroup
## 1         High income
## 2                null
## 3          Low income
## 4                null
## 5 Lower middle income
## 6 Upper middle income
##
## 1
## 2                                                                "26 countries, stretching from th
## 3
## 4 "22 countries, stretching from the westernmost point of Africa, across the equator, and partly alo
## 5
## 6
```

```
##                     TableName Year Population
## 1                       Aruba 1960      54208
## 2 Africa Eastern and Southern 1960  130836765
## 3                 Afghanistan 1960    8996967
## 4  Africa Western and Central 1960   96396419
## 5                      Angola 1960    5454938
## 6                     Albania 1960    1608800
```

3. Questions

3.1. How did the population of my country change over time?

Filter data

```
japan_pop <- pop_and_country_long %>%
  filter(TableName == "Japan")
head(japan_pop)
```

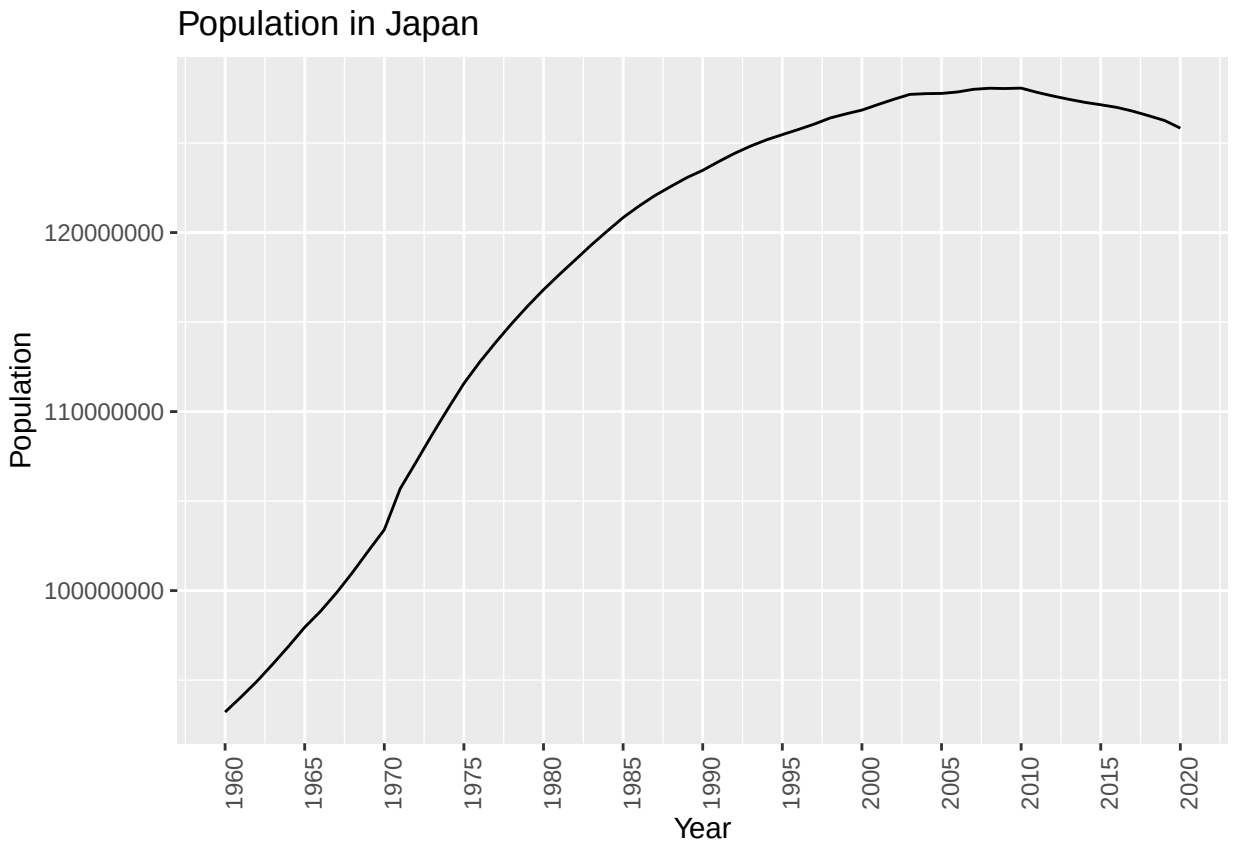
```
##   Country Code      Indicator Name Indicator Code              Region
## 1          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
## 2          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
## 3          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
## 4          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
## 5          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
## 6          JPN "Population, total"    SP.POP.TOTL East Asia & Pacific
##   IncomeGroup
## 1 High income
## 2 High income
## 3 High income
## 4 High income
## 5 High income
## 6 High income
##                                                                SpecialNotes
## 1 Fiscal year end: March 31; reporting period for national accounts data: CY.
## 2 Fiscal year end: March 31; reporting period for national accounts data: CY.
## 3 Fiscal year end: March 31; reporting period for national accounts data: CY.
## 4 Fiscal year end: March 31; reporting period for national accounts data: CY.
## 5 Fiscal year end: March 31; reporting period for national accounts data: CY.
## 6 Fiscal year end: March 31; reporting period for national accounts data: CY.
##   TableName Year Population
## 1     Japan 1960   93216000
## 2     Japan 1961   94055000
## 3     Japan 1962   94933000
## 4     Japan 1963   95900000
## 5     Japan 1964   96903000
## 6     Japan 1965   97952000
```

Create a plot

```
options(scipen = 999)

ggplot(japan_pop, aes(x=Year, y=Population)) +
  geom_line() +
  scale_x_continuous(breaks = seq(1960, 2020, 5)) +
  scale_y_continuous(breaks = seq(0, 130000000, 10000000)) +
  theme(axis.text.x = element_text(angle = 90)) +
```

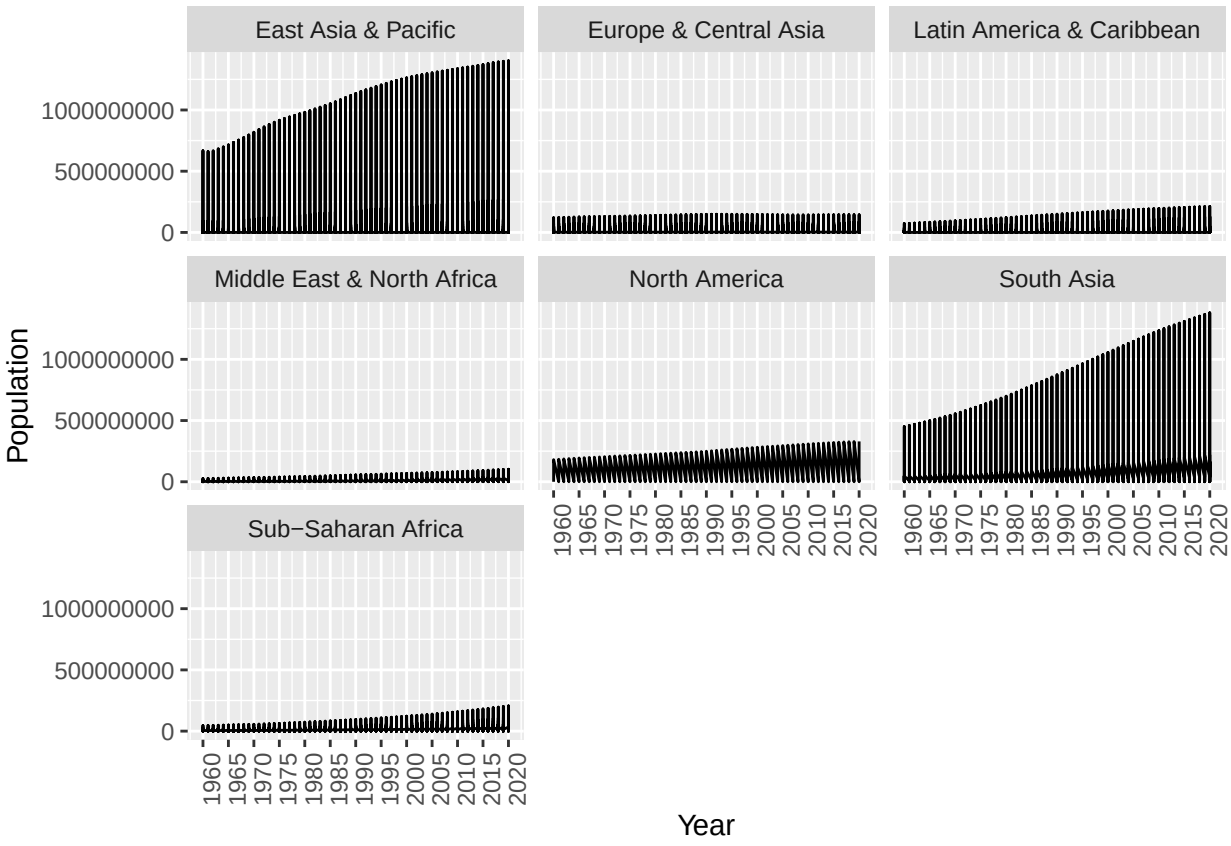
```
  ggtitle("Population in Japan")
```



3.2. How did the population in different parts of the world change over time?

Create a plot

```
pop_and_country_long %>%
  filter(Region != "null", Region != "NA") %>%
  ggplot(mapping = aes(x=Year, y=Population)) +
  geom_line() +
  scale_x_continuous(breaks = seq(1960, 2020, 5)) +
  theme(axis.text.x = element_text(angle = 90)) +
  facet_wrap(~ Region)
```



3.3. Which country or countries have experienced the highest increase/decrease in population over time?

Filter data

```
pop_1960_2020 <- select(pop_and_country_wide, "Country Code", "TableName", "Region", "IncomeGroup", "19
glimpse(pop_1960_2020)
```

```
## Rows: 266
## Columns: 6
## $ `Country Code` <chr> "ABW", "AFE", "AFG", "AFW", "AGO", "ALB", "AND", "ARB",~
## $ TableName      <chr> "Aruba", "Africa Eastern and Southern", "Afghanistan", ~
## $ Region         <chr> "Latin America & Caribbean", "null", "South Asia", "nul~
## $ IncomeGroup    <chr> "High income", "null", "Low income", "null", "Lower mid~
## $ `1960`         <dbl> 54208, 130836765, 8996967, 96396419, 5454938, 1608800, ~
## $ `2020`         <dbl> 106766, 677243299, 38928341, 458803476, 32866268, 28377~
```

Add new columns to compare population in 1960 and in 2020

```
pop_1960_2020 <- mutate(pop_1960_2020, divided = `2020` / `1960`, subtracted = `2020` - `1960`)
```

Highest population growth rate

```
div_highest_1960_2020 <- arrange(pop_1960_2020, desc(divided))
head(div_highest_1960_2020, 20)
```

```
##    Country Code                TableName                      Region
## 1           ARE United Arab Emirates Middle East & North Africa
## 2           QAT                Qatar Middle East & North Africa
## 3           KWT               Kuwait Middle East & North Africa
```

```
## 4           SXM Sint Maarten (Dutch part)  Latin America & Caribbean
## 5           DJI                  Djibouti Middle East & North Africa
## 6           JOR                    Jordan Middle East & North Africa
## 7           BHR                   Bahrain Middle East & North Africa
## 8           MAF  St. Martin (French part)  Latin America & Caribbean
## 9           OMN                      Oman Middle East & North Africa
## 10          SAU              Saudi Arabia Middle East & North Africa
## 11          CYM            Cayman Islands  Latin America & Caribbean
## 12          CIV             Côte d'Ivoire         Sub-Saharan Africa
## 13          NER                     Niger         Sub-Saharan Africa
## 14          UGA                    Uganda         Sub-Saharan Africa
## 15          TCA  Turks and Caicos Islands  Latin America & Caribbean
## 16          KEN                     Kenya         Sub-Saharan Africa
## 17          GMB             "Gambia, The"         Sub-Saharan Africa
## 18          AGO                    Angola         Sub-Saharan Africa
## 19          MDV                  Maldives                 South Asia
## 20          ZMB                    Zambia         Sub-Saharan Africa
##            IncomeGroup    1960     2020    divided subtracted
## 1          High income   92417  9890400 107.019271    9797983
## 2          High income   47383  2881060  60.803664    2833677
## 3          High income  269026  4270563  15.874165    4001537
## 4          High income    2833    40812  14.405930      37979
## 5  Lower middle income   83634   988002  11.813401     904368
## 6  Upper middle income  933102 10203140  10.934646    9270038
## 7          High income  162429  1701583  10.475857    1539154
## 8          High income    3898    38659   9.917650      34761
## 9          High income  551735  5106622   9.255570    4554887
## 10         High income 4086534 34813867   8.519167   30727333
## 11         High income    7870    65720   8.350699      57850
## 12 Lower middle income 3503559 26378275   7.528994   22874716
## 13          Low income 3388774 24206636   7.143184   20817862
## 14          Low income 6767092 45741000   6.759329   38973908
## 15         High income    5825    38718   6.646867      32893
## 16 Lower middle income 8120082 53771300   6.622014   45651218
## 17          Low income  365049  2416664   6.620109    2051615
## 18 Lower middle income 5454938 32866268   6.025049   27411330
## 19 Upper middle income   89873   540542   6.014509     450669
## 20 Lower middle income 3070780 18383956   5.986738   15313176
```

Highest population growth number

```
sub_highest_1960_2020 <- arrange(pop_1960_2020, desc(subtracted))
head(sub_highest_1960_2020, 20)
```

```
##    Country Code                                TableName     Region
## 1           WLD                                    World       null
## 2           IBT                          IDA & IBRD total       null
## 3           LMY                   Low & middle income       null
## 4           MIC                         Middle income       null
## 5           IBD                             IBRD only       null
## 6           EAR                 Early-demographic dividend       null
## 7           LMC                   Lower middle income       null
## 8           UMC                   Upper middle income       null
## 9           IDA                             IDA total       null
## 10          EAS                   East Asia & Pacific       null
```

```
## 11          SAS                                    South Asia       null
## 12          TSA                      South Asia (IDA & IBRD)       null
## 13          LTE                    Late-demographic dividend       null
## 14          EAP East Asia & Pacific (excluding high income)       null
## 15          TEA            East Asia & Pacific (IDA & IBRD)       null
## 16          IND                                       India South Asia
## 17          SSF                          Sub-Saharan Africa       null
## 18          TSS              Sub-Saharan Africa (IDA & IBRD)       null
## 19          SSA  Sub-Saharan Africa (excluding high income)       null
## 20          IDX                                     IDA only       null
##            IncomeGroup       1960       2020  divided subtracted
## 1                 null 3032156070 7752840547 2.556874 4720684477
## 2                 null 2299245319 6562212357 2.854072 4262967038
## 3                 null 2264230620 6509474374 2.874917 4245243754
## 4                 null 2126448264 5844325339 2.748398 3717877075
## 5                 null 1919643259 4853608684 2.528391 2933965425
## 6                 null  980003345 3332105361 3.400096 2352102016
## 7                 null  989984004 3330652547 3.364350 2340668543
## 8                 null 1136464260 2513672792 2.211836 1377208532
## 9                 null  379602060 1708603673 4.501039 1329001613
## 10                null 1041673567 2352037717 2.257941 1310364150
## 11                null  572839530 1856882402 3.241540 1284042872
## 12                null  572839530 1856882402 3.241540 1284042872
## 13                null 1096903448 2307887395 2.104002 1210983947
## 14                null  894875757 2105003391 2.352286 1210127634
## 15                null  883445587 2079198305 2.353510 1195752718
## 16 Lower middle income  450547675 1380004385 3.062949  929456710
## 17                null  227233184 1136046775 4.999476  908813591
## 18                null  227233184 1136046775 4.999476  908813591
## 19                null  227191484 1135948313 4.999960  908756829
## 20                null  259210418 1134444535 4.376539  875234117
```

Lowest population growth rate

```
div_lowest_1960_2020 <- arrange(pop_1960_2020, divided)
head(div_lowest_1960_2020, 20)
```

```
##    Country Code                      TableName                    Region
## 1           BGR                       Bulgaria     Europe & Central Asia
## 2           LVA                         Latvia     Europe & Central Asia
## 3           HUN                        Hungary     Europe & Central Asia
## 4           HRV                        Croatia     Europe & Central Asia
## 5           LTU                      Lithuania     Europe & Central Asia
## 6           BIH         Bosnia and Herzegovina     Europe & Central Asia
## 7           GEO                        Georgia     Europe & Central Asia
## 8           UKR                        Ukraine     Europe & Central Asia
## 9           KNA            St. Kitts and Nevis Latin America & Caribbean
## 10          SRB                         Serbia     Europe & Central Asia
## 11          ROU                        Romania     Europe & Central Asia
## 12          EST                        Estonia     Europe & Central Asia
## 13          CZE                 Czech Republic     Europe & Central Asia
## 14          CEB Central Europe and the Baltics                     null
## 15          DEU                        Germany     Europe & Central Asia
## 16          BLR                        Belarus     Europe & Central Asia
## 17          PRT                       Portugal     Europe & Central Asia
```

```
## 18          ITA                         Italy     Europe & Central Asia
## 19          DMA                      Dominica Latin America & Caribbean
## 20          RUS             Russian Federation     Europe & Central Asia
##             IncomeGroup      1960      2020   divided subtracted
## 1  Upper middle income   7867374   6927288 0.8805083    -940086
## 2          High income   2120979   1901548 0.8965426    -219431
## 3          High income   9983967   9749763 0.9765420    -234204
## 4          High income   4140181   4047200 0.9775418     -92981
## 5          High income   2778550   2794700 1.0058124      16150
## 6  Upper middle income   3225664   3280815 1.0170976      55151
## 7  Upper middle income   3645600   3714000 1.0187623      68400
## 8  Lower middle income  42664646  44134693 1.0344559    1470047
## 9          High income     51199     53192 1.0389265       1993
## 10 Upper middle income   6608000   6908224 1.0454334     300224
## 11 Upper middle income  18406905  19286123 1.0477657     879218
## 12         High income   1211537   1331057 1.0986515     119520
## 13         High income   9602006  10698896 1.1142355    1096890
## 14                null  91401764 102246330 1.1186472   10844566
## 15         High income  72814900  83240525 1.1431798   10425625
## 16 Upper middle income   8198000   9398861 1.1464822    1200861
## 17         High income   8857716  10305564 1.1634561    1447848
## 18         High income  50199700  59554023 1.1863422    9354323
## 19 Upper middle income     60020     71991 1.1994502      11971
## 20 Upper middle income 119897000 144104080 1.2018990   24207080
```

Lowest population growth number

```
sub_lowest_1960_2020 <- arrange(pop_1960_2020, subtracted)
head(sub_lowest_1960_2020, 20)
```

```
##    Country Code                      TableName                    Region
## 1           BGR                       Bulgaria     Europe & Central Asia
## 2           HUN                        Hungary     Europe & Central Asia
## 3           LVA                         Latvia     Europe & Central Asia
## 4           HRV                        Croatia     Europe & Central Asia
## 5           KNA           St. Kitts and Nevis Latin America & Caribbean
## 6           NRU                          Nauru       East Asia & Pacific
## 7           TUV                         Tuvalu       East Asia & Pacific
## 8           PLW                          Palau       East Asia & Pacific
## 9           GIB                      Gibraltar     Europe & Central Asia
## 10          DMA                       Dominica Latin America & Caribbean
## 11          FRO                 Faroe Islands     Europe & Central Asia
## 12          LTU                     Lithuania     Europe & Central Asia
## 13          MCO                        Monaco     Europe & Central Asia
## 14          SMR                    San Marino     Europe & Central Asia
## 15          BMU                       Bermuda             North America
## 16          LIE                  Liechtenstein     Europe & Central Asia
## 17          VGB         British Virgin Islands Latin America & Caribbean
## 18          GRD                       Grenada Latin America & Caribbean
## 19          GRL                     Greenland     Europe & Central Asia
## 20          VCT St. Vincent and the Grenadines Latin America & Caribbean
##             IncomeGroup    1960    2020   divided subtracted
## 1  Upper middle income 7867374 6927288 0.8805083    -940086
## 2          High income 9983967 9749763 0.9765420    -234204
## 3          High income 2120979 1901548 0.8965426    -219431
```

```
## 4          High income 4140181 4047200 0.9775418     -92981
## 5          High income   51199   53192 1.0389265       1993
## 6          High income    4377   10834 2.4752113       6457
## 7  Upper middle income    5321   11792 2.2161248       6471
## 8          High income    9769   18092 1.8519808       8323
## 9          High income   23420   33691 1.4385568      10271
## 10 Upper middle income   60020   71991 1.1994502      11971
## 11         High income   34624   48865 1.4113043      14241
## 12         High income 2778550 2794700 1.0058124      16150
## 13         High income   22461   39244 1.7472063      16783
## 14         High income   15440   33938 2.1980570      18498
## 15         High income   44400   63903 1.4392568      19503
## 16         High income   16501   38137 2.3111933      21636
## 17         High income    8053   30237 3.7547498      22184
## 18 Upper middle income   89927  112519 1.2512260      22592
## 19         High income   32500   56367 1.7343692      23867
## 20 Upper middle income   80970  110947 1.3702235      29977
```

3.4. Which country or countries have been experiencing the highest increase/decrease in population in the last ten years?

Filter data

```
pop_2010_2020 <- select(pop_and_country_wide, "Country Code", "TableName", "Region", "IncomeGroup", "20
glimpse(pop_2010_2020)
```

```
## Rows: 266
## Columns: 6
## $ `Country Code` <chr> "ABW", "AFE", "AFG", "AFW", "AGO", "ALB", "AND", "ARB",~
## $ TableName      <chr> "Aruba", "Africa Eastern and Southern", "Afghanistan", ~
## $ Region         <chr> "Latin America & Caribbean", "null", "South Asia", "nul~
## $ IncomeGroup    <chr> "High income", "null", "Low income", "null", "Lower mid~
## $ `2010`         <dbl> 101665, 518468229, 29185511, 350556886, 23356247, 29130~
## $ `2020`         <dbl> 106766, 677243299, 38928341, 458803476, 32866268, 28377~
```

Add new columns to compare population in 2010 and in 2020

```
pop_2010_2020 <- mutate(pop_2010_2020, divided = `2020` / `2010`, subtracted = `2020` - `2010`)
```

Highest population growth rate

```
div_highest_2010_2020 <- arrange(pop_2010_2020, desc(divided))
head(div_highest_2010_2020, 20)
```

```
##    Country Code          TableName                      Region
## 1           OMN               Oman Middle East & North Africa
## 2           QAT              Qatar Middle East & North Africa
## 3           GNQ  Equatorial Guinea         Sub-Saharan Africa
## 4           MDV           Maldives                 South Asia
## 5           NER              Niger         Sub-Saharan Africa
## 6           KWT             Kuwait Middle East & North Africa
## 7           UGA             Uganda         Sub-Saharan Africa
## 8           AGO             Angola         Sub-Saharan Africa
## 9           JOR             Jordan Middle East & North Africa
## 10          COD "Congo, Dem. Rep."         Sub-Saharan Africa
## 11          LBN            Lebanon Middle East & North Africa
## 12          TCD               Chad         Sub-Saharan Africa
```

```
## 13          BHR            Bahrain Middle East & North Africa
## 14          BDI            Burundi         Sub-Saharan Africa
## 15          GAB              Gabon         Sub-Saharan Africa
## 16          IRQ               Iraq Middle East & North Africa
## 17          ZMB             Zambia         Sub-Saharan Africa
## 18          GMB      "Gambia, The"         Sub-Saharan Africa
## 19          TZA           Tanzania         Sub-Saharan Africa
## 20          MLI               Mali         Sub-Saharan Africa
##            IncomeGroup     2010     2020  divided subtracted
## 1          High income  3041435  5106622 1.679017    2065187
## 2          High income  1856329  2881060 1.552020    1024731
## 3  Upper middle income   943640  1402985 1.486780     459345
## 4  Upper middle income   365730   540542 1.477981     174812
## 5           Low income 16464025 24206636 1.470274    7742611
## 6          High income  2991884  4270563 1.427383    1278679
## 7           Low income 32428164 45741000 1.410533   13312836
## 8  Lower middle income 23356247 32866268 1.407172    9510021
## 9  Upper middle income  7261541 10203140 1.405093    2941599
## 10          Low income 64563853 89561404 1.387176   24997551
## 11 Upper middle income  4953064  6825442 1.378024    1872378
## 12          Low income 11952134 16425859 1.374303    4473725
## 13         High income  1240864  1701583 1.371289     460719
## 14          Low income  8675606 11890781 1.370599    3215175
## 15 Upper middle income  1624146  2225728 1.370399     601582
## 16 Upper middle income 29741977 40222503 1.352382   10480526
## 17 Lower middle income 13605986 18383956 1.351167    4777970
## 18          Low income  1793199  2416664 1.347683     623465
## 19 Lower middle income 44346532 59734213 1.346987   15387681
## 20          Low income 15049352 20250834 1.345628    5201482
```

Highest population growth number

```
sub_highest_2010_2020 <- arrange(pop_2010_2020, desc(subtracted))
head(sub_highest_2010_2020, 20)
```

```
##    Country Code                                      TableName Region IncomeGroup
## 1           WLD                                          World   null        null
## 2           LMY                           Low & middle income   null        null
## 3           IBT                              IDA & IBRD total   null        null
## 4           MIC                                 Middle income   null        null
## 5           LMC                           Lower middle income   null        null
## 6           IBD                                     IBRD only   null        null
## 7           EAR                  Early-demographic dividend   null        null
## 8           IDA                                     IDA total   null        null
## 9           SSF                           Sub-Saharan Africa   null        null
## 10          TSS             Sub-Saharan Africa (IDA & IBRD)   null        null
## 11          SSA  Sub-Saharan Africa (excluding high income)   null        null
## 12          PRE                   Pre-demographic dividend   null        null
## 13          IDX                                      IDA only   null        null
## 14          LDC Least developed countries: UN classification   null        null
## 15          SAS                                   South Asia   null        null
## 16          TSA                       South Asia (IDA & IBRD)   null        null
## 17          HPC      Heavily indebted poor countries (HIPC)   null        null
## 18          FCS    Fragile and conflict affected situations   null        null
## 19          UMC                           Upper middle income   null        null
```

```
## 20          AFE                  Africa Eastern and Southern   null        null
##          2010       2020  divided subtracted
## 1  6921877071 7752840547 1.120049  830963476
## 2  5739242600 6509474374 1.134204  770231774
## 3  5792408567 6562212357 1.132899  769803790
## 4  5225459282 5844325339 1.118433  618866057
## 5  2878875404 3330652547 1.156928  451777143
## 6  4427061950 4853608684 1.096350  426546734
## 7  2907916947 3332105361 1.145874  424188414
## 8  1365346617 1708603673 1.251407  343257056
## 9   869025115 1136046775 1.307266  267021660
## 10  869025115 1136046775 1.307266  267021660
## 11  868935345 1135948313 1.307287  267012968
## 12  731868780  970795671 1.326461  238926891
## 13  909297040 1134444535 1.247606  225147495
## 14  836614841 1057438163 1.263949  220823322
## 15 1638792927 1856882402 1.133079  218089475
## 16 1638792927 1856882402 1.133079  218089475
## 17  624219296  823480038 1.319216  199260742
## 18  735448194  930006546 1.264544  194558352
## 19 2346583878 2513672792 1.071205  167088914
## 20  518468229  677243299 1.306239  158775070
```

Lowest population growth rate

```
div_lowest_2010_2020 <- arrange(pop_2010_2020, divided)
head(div_lowest_2010_2020, 20)
```

```
##    Country Code                      TableName                      Region
## 1           SYR           Syrian Arab Republic Middle East & North Africa
## 2           PRI                    Puerto Rico  Latin America & Caribbean
## 3           BIH         Bosnia and Herzegovina      Europe & Central Asia
## 4           LTU                      Lithuania      Europe & Central Asia
## 5           LVA                         Latvia      Europe & Central Asia
## 6           MDA                        Moldova      Europe & Central Asia
## 7           AND                        Andorra      Europe & Central Asia
## 8           BGR                       Bulgaria      Europe & Central Asia
## 9           HRV                        Croatia      Europe & Central Asia
## 10          SRB                         Serbia      Europe & Central Asia
## 11          ROU                        Romania      Europe & Central Asia
## 12          UKR                        Ukraine      Europe & Central Asia
## 13          GRC                         Greece      Europe & Central Asia
## 14          ALB                        Albania      Europe & Central Asia
## 15          PRT                       Portugal      Europe & Central Asia
## 16          HUN                        Hungary      Europe & Central Asia
## 17          CEB Central Europe and the Baltics                       null
## 18          GEO                        Georgia      Europe & Central Asia
## 19          VIR          Virgin Islands (U.S.)  Latin America & Caribbean
## 20          BMU                        Bermuda              North America
##            IncomeGroup      2010      2020   divided subtracted
## 1           Low income  21362541  17500657 0.8192217   -3861884
## 2          High income   3721525   3194034 0.8582595    -527491
## 3  Upper middle income   3705478   3280815 0.8853959    -424663
## 4          High income   3097282   2794700 0.9023072    -302582
## 5          High income   2097555   1901548 0.9065545    -196007
```

```
## 6  Upper middle income   2861487   2617820 0.9148460    -243667
## 7           High income     84454     77265 0.9148767      -7189
## 8  Upper middle income   7395599   6927288 0.9366771    -468311
## 9           High income   4295427   4047200 0.9422113    -248227
## 10 Upper middle income   7291436   6908224 0.9474435    -383212
## 11 Upper middle income  20246871  19286123 0.9525483    -960748
## 12 Lower middle income  45870741  44134693 0.9621535   -1736048
## 13          High income  11121341  10715549 0.9635123    -405792
## 14 Upper middle income   2913021   2837743 0.9741581     -75278
## 15          High income  10573100  10305564 0.9746965    -267536
## 16          High income  10000023   9749763 0.9749741    -250260
## 17                null 104421447 102246330 0.9791698   -2175117
## 18 Upper middle income   3786695   3714000 0.9808025     -72695
## 19          High income    108357    106290 0.9809242      -2067
## 20          High income     65124     63903 0.9812512      -1221
```

Lowest population growth number

```
sub_lowest_2010_2020 <- arrange(pop_2010_2020, subtracted)
head(sub_lowest_2010_2020, 20)
```

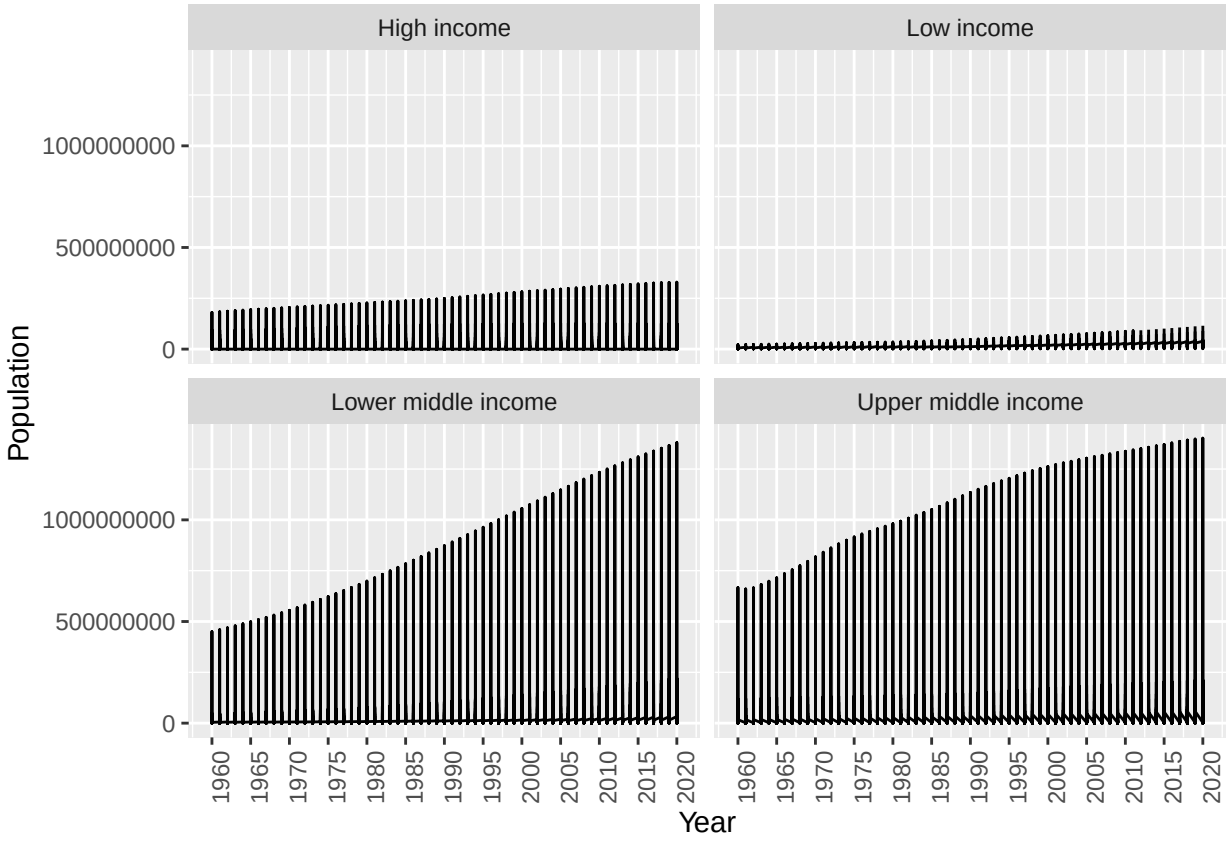
```
##    Country Code                      TableName                      Region
## 1           SYR           Syrian Arab Republic Middle East & North Africa
## 2           JPN                          Japan        East Asia & Pacific
## 3           CEB Central Europe and the Baltics                        null
## 4           UKR                        Ukraine      Europe & Central Asia
## 5           ROU                        Romania      Europe & Central Asia
## 6           PRI                    Puerto Rico  Latin America & Caribbean
## 7           BGR                       Bulgaria      Europe & Central Asia
## 8           BIH         Bosnia and Herzegovina      Europe & Central Asia
## 9           GRC                         Greece      Europe & Central Asia
## 10          SRB                         Serbia      Europe & Central Asia
## 11          LTU                      Lithuania      Europe & Central Asia
## 12          PRT                       Portugal      Europe & Central Asia
## 13          HUN                        Hungary      Europe & Central Asia
## 14          HRV                        Croatia      Europe & Central Asia
## 15          MDA                        Moldova      Europe & Central Asia
## 16          LVA                         Latvia      Europe & Central Asia
## 17          POL                         Poland      Europe & Central Asia
## 18          BLR                        Belarus      Europe & Central Asia
## 19          ALB                        Albania      Europe & Central Asia
## 20          GEO                        Georgia      Europe & Central Asia
##            IncomeGroup      2010      2020   divided subtracted
## 1           Low income  21362541  17500657 0.8192217   -3861884
## 2          High income 128070000 125836021 0.9825566   -2233979
## 3                 null 104421447 102246330 0.9791698   -2175117
## 4  Lower middle income  45870741  44134693 0.9621535   -1736048
## 5  Upper middle income  20246871  19286123 0.9525483    -960748
## 6           High income   3721525   3194034 0.8582595    -527491
## 7  Upper middle income   7395599   6927288 0.9366771    -468311
## 8  Upper middle income   3705478   3280815 0.8853959    -424663
## 9           High income  11121341  10715549 0.9635123    -405792
## 10 Upper middle income   7291436   6908224 0.9474435    -383212
## 11          High income   3097282   2794700 0.9023072    -302582
## 12          High income  10573100  10305564 0.9746965    -267536
```

```
## 13         High income  10000023   9749763 0.9749741    -250260
## 14         High income   4295427   4047200 0.9422113    -248227
## 15 Upper middle income   2861487   2617820 0.9148460    -243667
## 16         High income   2097555   1901548 0.9065545    -196007
## 17         High income  38042794  37950802 0.9975819     -91992
## 18 Upper middle income   9490583   9398861 0.9903355     -91722
## 19 Upper middle income   2913021   2837743 0.9741581     -75278
## 20 Upper middle income   3786695   3714000 0.9808025     -72695
```

3.5. How does income group affect a country's population growth?

Create a plot

```
pop_and_country_long %>%
  filter(IncomeGroup != "null", IncomeGroup != "NA") %>%
  ggplot(mapping = aes(x=Year, y=Population)) +
  geom_line() +
  scale_x_continuous(breaks = seq(1960, 2020, 5)) +
  theme(axis.text.x = element_text(angle = 90)) +
  facet_wrap(~ IncomeGroup)
```



4. Findings

- The population of Japan peaked in 2010, and then started decreasing. Japan ranks the second lowest population growth number from 2010 to 2020, decreased by 2,233,979.
- South Asia sees the highest population growth, followed by East Asia & Pacific. The population of Europe & Central Asia has not changed much since 1960.
- From 1960 to 2020, 8 out of top 20 countries with the highest population growth rate are in Middle

East & North Africa. 17 and 11 countries in Europe & Central Asia are in the top 20 countries with the lowest population growth rate and the lowest population growth number, respectively.

- From 2010 to 2020, 12 countries in Sub-Saharan Africa experienced high growth in population. Concerning the lowest growth rate and number, Europe & Central Asia accounts for most of the countries.
- Lower middle income countries have the highest population growth rate.