Examining Gender Inequality in Global Health: An Empirical Study at the Country and Region Level

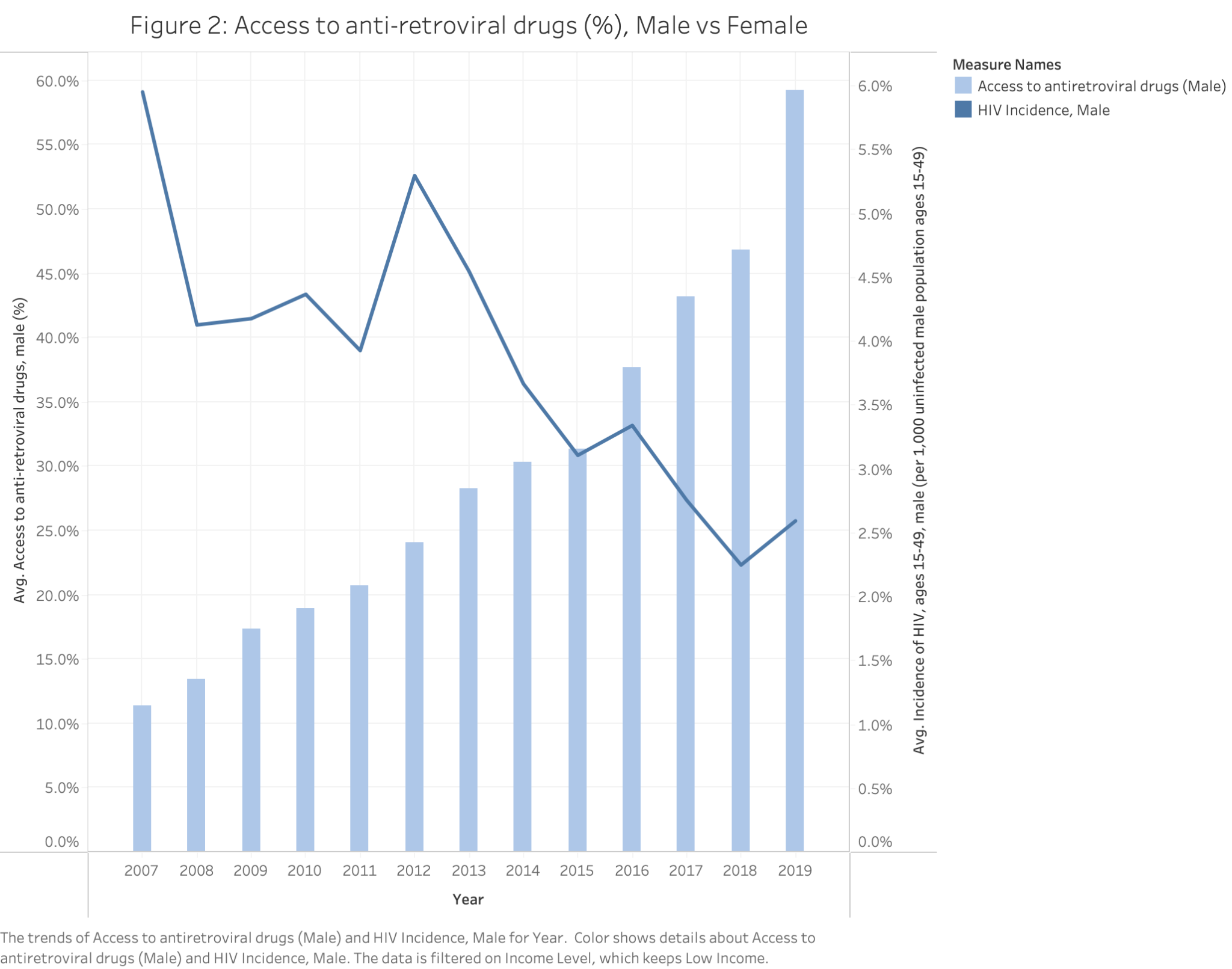
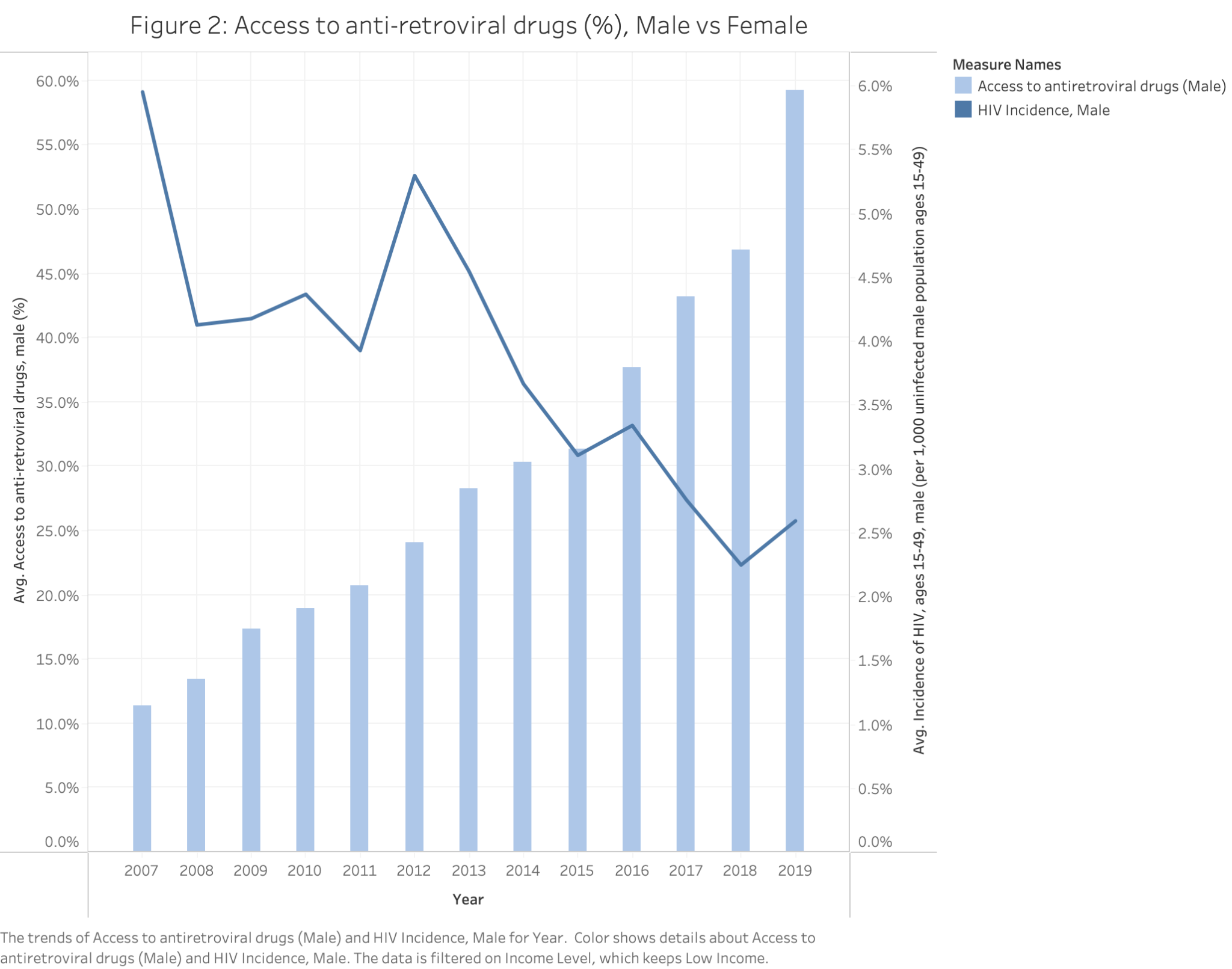
Team Ghost Rider:

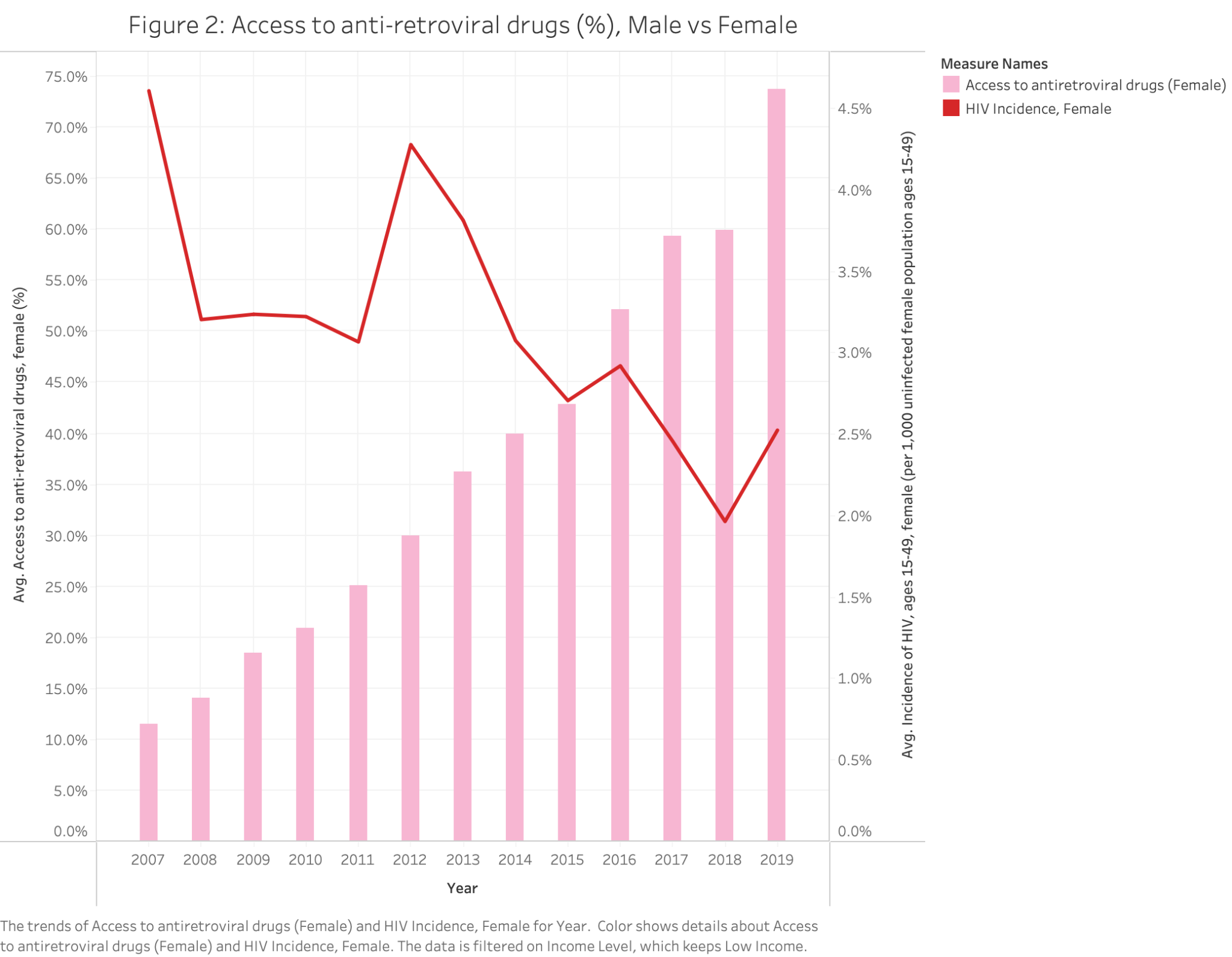
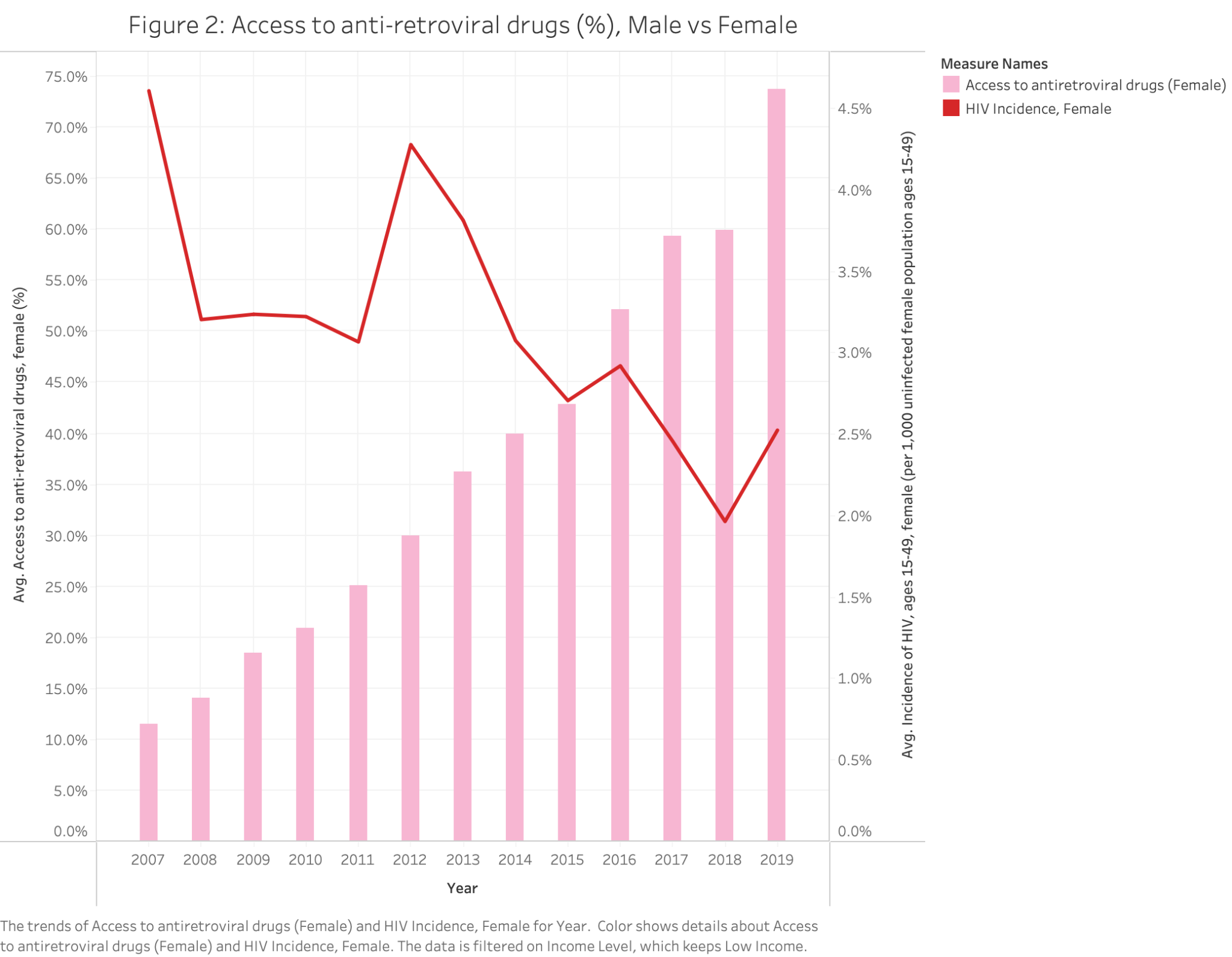
Hayley Demetres, Tu Ho, Paul Kim

Matrix:

| **Serial #** | **Hypothesis** | **Question** | **Analytic Type** | **Variables** | **Chart** | **Conclusion** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | There is a negative correlation between access to anti-retroviral drugs and incidence of HIV among females and males over time. | Does access to anti-retroviral drugs affect the incidence of HIV among females and males? | Predictive | Access to anti-retroviral drugs, female (%) and male (%), Incidence of HIV, ages 15-49, female and male (per 1,000 uninfected population ages 15-49) | Bar + Line chart | Higher access to anti-retroviral drugs is correlated with lower incidence of HIV among females and males. |
| 2 | There is a correlation between the proportion of women in national parliaments and reduced fertility rate over time. | How does the proportion of women in national parliaments correlate with fertility rate? | Descriptive | Proportion of seats held by women in national parliaments (%), Fertility rate, total (births per woman) | Tree map | Higher representation of women in national parliaments is correlated with lower fertility rate. |
| 3 | There is a positive correlation between immunization against measle and life expectancy among females and males. | How does immunization against measles affect life expectancy among females and males? | Predictive | Immunization, measles (% of children ages 12-23 months), Life expectancy at birth, female (years), male (years) | Multiple Line Chart | Higher immunization rate against measles is correlated with higher life exp. among females and males. |
| 4 | There is a relationship between a country's economic situation and incidence of HIV among females and males. | Is there a relationship between a country's economic situation and incidence of HIV among females and males? | Descriptive | GDP per capita, PPP (current international $), Incidence of HIV, ages 15-49, female and male (per 1,000 uninfected population ages 15-49) | Pie Chart | Countries with better economic situations typically have lower HIV rates for both genders. |
| 5 | There is an association between gender parity in school enrollment and fertility rate over time. | Is gender parity in school enrollment at primary and secondary level associated with fertility rate? | Predictive | School enrollment, primary and secondary (gross), gender parity index (GPI), Fertility rate, total (births per woman) | Bubble chart | Higher gender parity index (GPI) in school enrollment at primary and secondary levels is associated with lower fertility rate. |
| 6 | There is a positive correlation between the proportion of women in national parliaments and life expectancy among females. | How does the proportion of women in national parliaments correlate with life expectancy among females? | Descriptive | Proportion of seats held by women in national parliaments (%), Life expectancy at birth, female (years) | Bar + Line Chart | Representation of women in national parliaments is positively correlated with life expectancy among females. |
| 7 | There is an association between gender parity in school enrollment and life expectancy among females and males. | Is gender parity in school enrollment affecting life expectancy among females and males? | Descriptive | School enrollment, primary and secondary (gross), gender parity index (GPI), Life expectancy at birth, female (years), Life expectancy at birth, male (years) | Scatter Plot with Regression Line | Higher GPI in primary & secondary school enrollment is correlated with higher life exp. for both genders. |
| 8 | There is a negative correlation between unemployment and mortality from CVD, cancer, diabetes, or CRD among females and males. | How does unemployment rate affect mortality from CVD, cancer, diabetes, or CRD for both genders? | Predictive | unemployed, female (% of female employment) (modeled ILO estimate), unemployed, male (% of male employment), Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, female (%), Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, male (%) | Bubble Chart | Higher rates of unemployment correlate with higher mortality from CVD, cancer, diabetes, or CRD for both genders. |
| 9 | There is a positive correlation between unemployment rate among females and fertility rate over time. | Is the unemployment rate among females affecting fertility rate over time? | Descriptive | Unemployment, female (% of female labor force) (modeled ILO estimate), Fertility rate, total (births per woman) | Bar + Line chart | Higher unemployment rates among females is correlated with higher fertility rates over time. |
| 10 | There is a positive correlation between self-employment and life expectancy among females and males . | How does self-employment correlate with life expectancy among both genders? | Predictive | Self-employed, female (% of female employment) (modeled ILO estimate), Life expectancy at birth, female (years) | Area chart | Higher rate of self-employment is correlated with higher life exp. for both genders. |
| 11 | There is a relationship between a country's economic situation and mortality from CVD, cancer, diabetes, or CRD among females and males. | Is there a relationship between a country's economic situation and mortality from CVD, cancer, diabetes, or CRD among females and males? | Predictive | Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, female (%), male (%) | Multiple line charts | Higher-income countries are more likely to exhibit lower mortality from CVD, cancer, diabetes or CRD for both genders. |
| 12 | There is a positive correlation between access to anti-retroviral drugs and life expectancy among both genders. | Does access to anti-retroviral drugs positively correlate with life expectancy among both genders? | Descriptive | Access to anti-retroviral drugs, female (%) and male (%), Life expectancy at birth, female (years), Life expectancy at birth, male (years) | Tree maps | Access to anti-retroviral drugs is correlated with higher life exp. among both genders. |
| 13 | There is an association between regions with high gender parity indexes and mortality from CVD, cancer, diabetes, or CRD among females and males. | Is there an association between region's with higher gender parity index and mortality rate from CVD, cancer, diabetes, or CRD among males? | Descriptive | Population, male, Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, male (%), Region, School Enrollment | Map | Higher GPI in school enrollment is correlated with a reduced mortality rate from CVD, cancer, diabetes, or CRD among females and males. |
| 14 | There is a negative correlation between the proportion of seats held by women in national parliaments and incidence of HIV among females. | How does the proportion of seats held by women in national parliaments correlate with incidence of HIV in females? | Descriptive | Proportion of seats held by women in national parliaments (%), Incidence of HIV, females | Area Chart | There is an inverse relationship between women in national parliaments and HIV incidence in females. |
| 15 | There is a positive correlation between unemployment rates and incidence of HIV among females and males. | Is unemployment rates correlated with incidence of HIV among females and males? | Descriptive | Unemployment, female (% of female labor force) (modeled ILO estimate), Unemployment, male (% of male labor force) (modeled ILO estimate), Incidence of HIV, ages 15-49, female (per 1,000 uninfected female population ages 15-49), Incidence of HIV, ages 15-49, male (per 1,000 uninfected male population ages 15-49), Year | Bar Graph | Higher unemployment rates are correlated with higher incidence of HIV among females and males. |

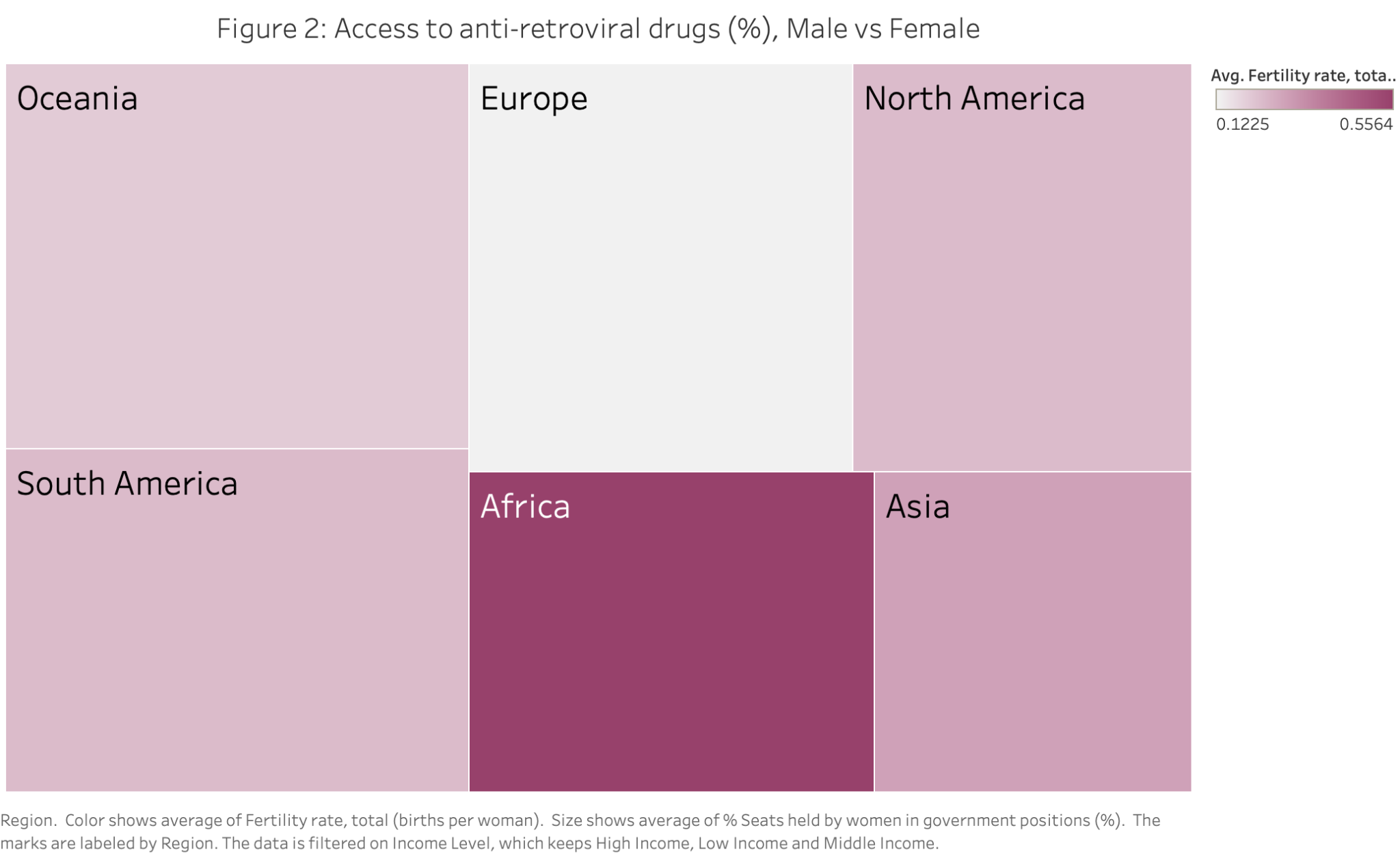
Figure #1- Access to anti-retroviral drugs vs. Incidence of HIV amongst males and females





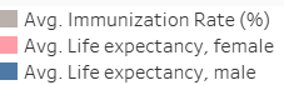
* Line and bar chart shows the relationship between access to antiretroviral drugs and HIV incidents for males and females.
* Over the years, access to antiretroviral drugs has increased for both genders, and incidents of HIV have sequentially decreased.
* While the relationship between the variables is the same for both genders, Males have gained access to the drugs at a faster rate compared to females.

Figure #2- Women in Parliament vs Fertility Rates



* This heat map demonstrates the relationship between women in parliament and fertility rates by region. Size indicates women in parliament, color represents fertility rate.
* Africa has the highest average fertility rate, which is an average representation of women in parliament. Europe has a slightly higher representation of women in the national parliament, which has a very low fertility rate.
* Based on our definition of fertility rate, women in Africa are forecasted to be able to have the most children, while women in Europe are less fertile. There does not seem to be a relationship between representation in government positions and fertility rates.

Figure #3 - Measles Immunization Rates vs Life Expectancy in Males and Females

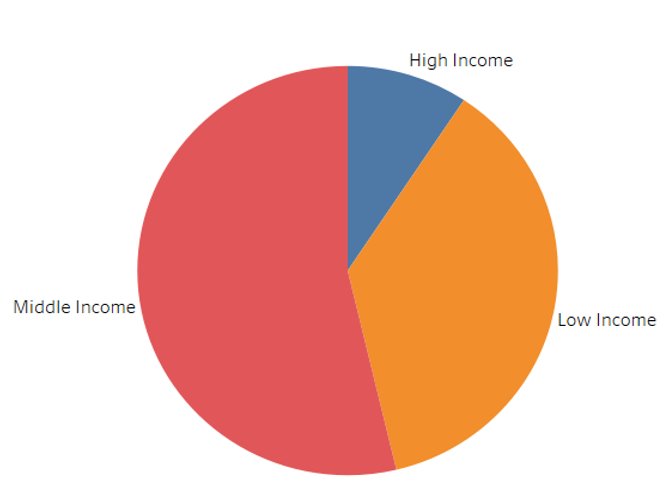




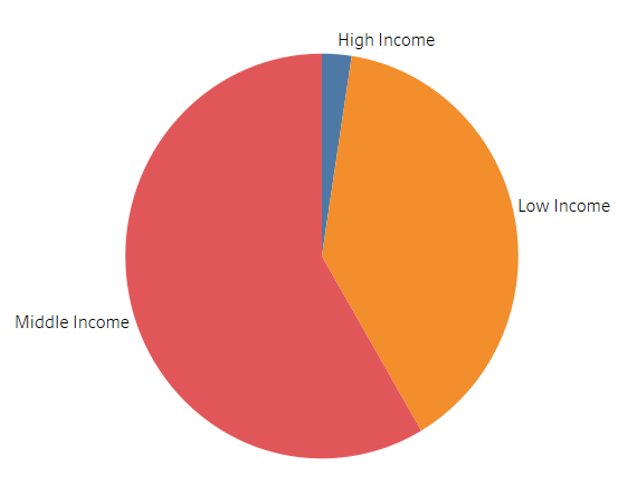
* Line graph showing measles immunization rate and life expectancy for males and females over time in low-income countries​
* While life expectancy for females remained below that of males, the gap has reduced in recent years.​
* Increased immunization rates have contributed to longer life expectancy and may help reduce gender inequity in health.

Figure #4- Economic Levels vs HIV Incidences in Males and Females

Incidence of HIV- Females

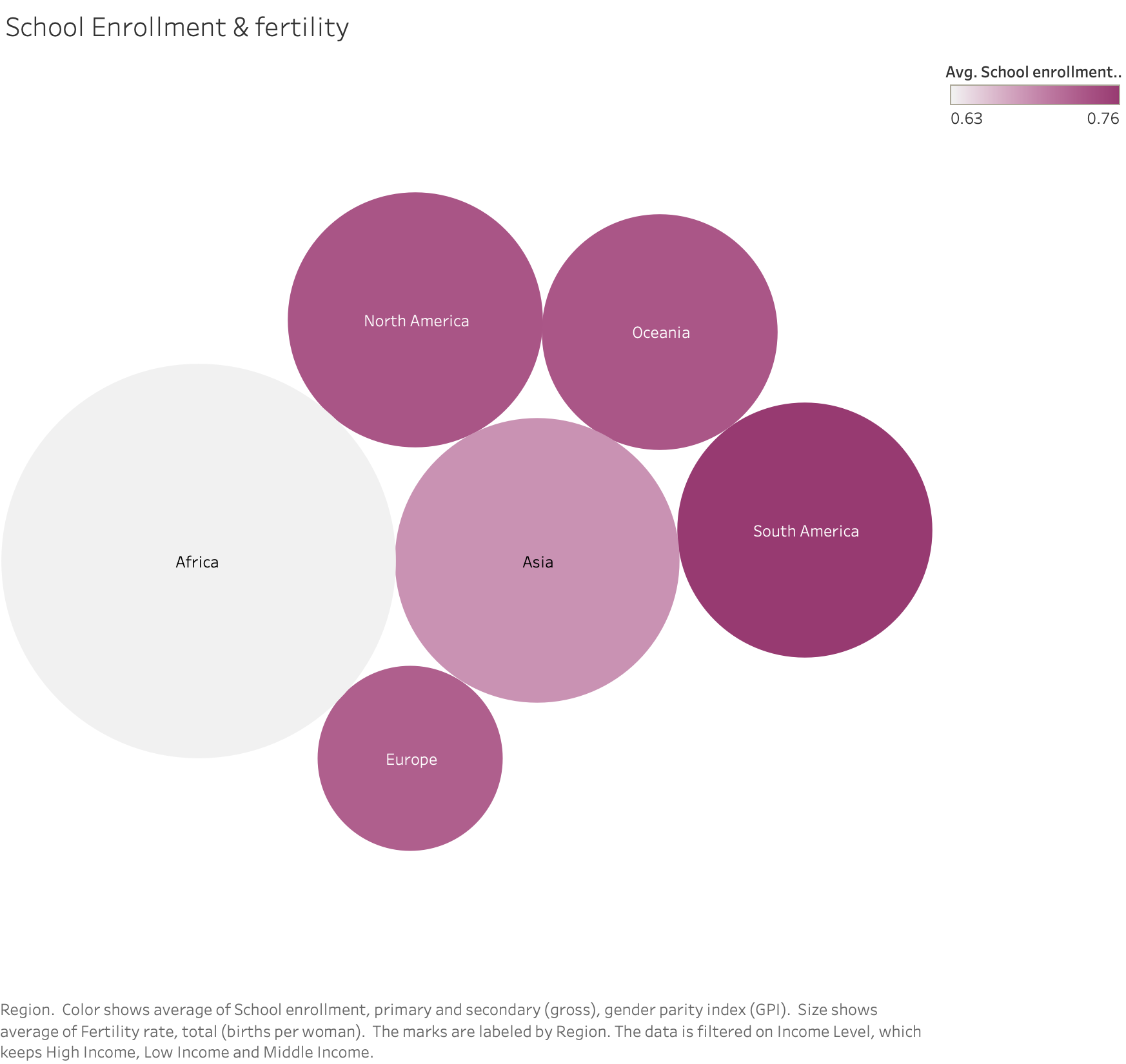
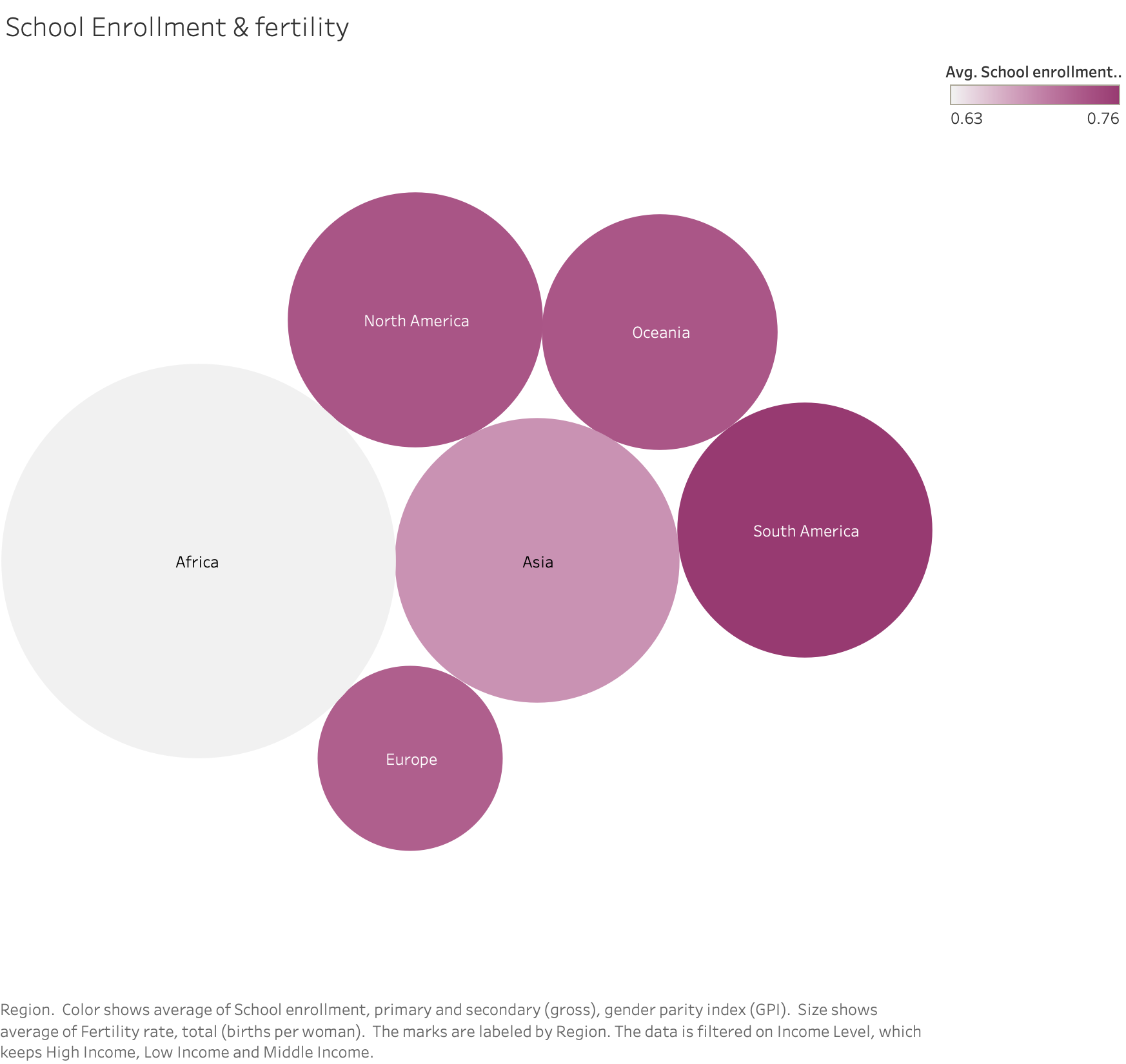


Incidence of HIV- Males



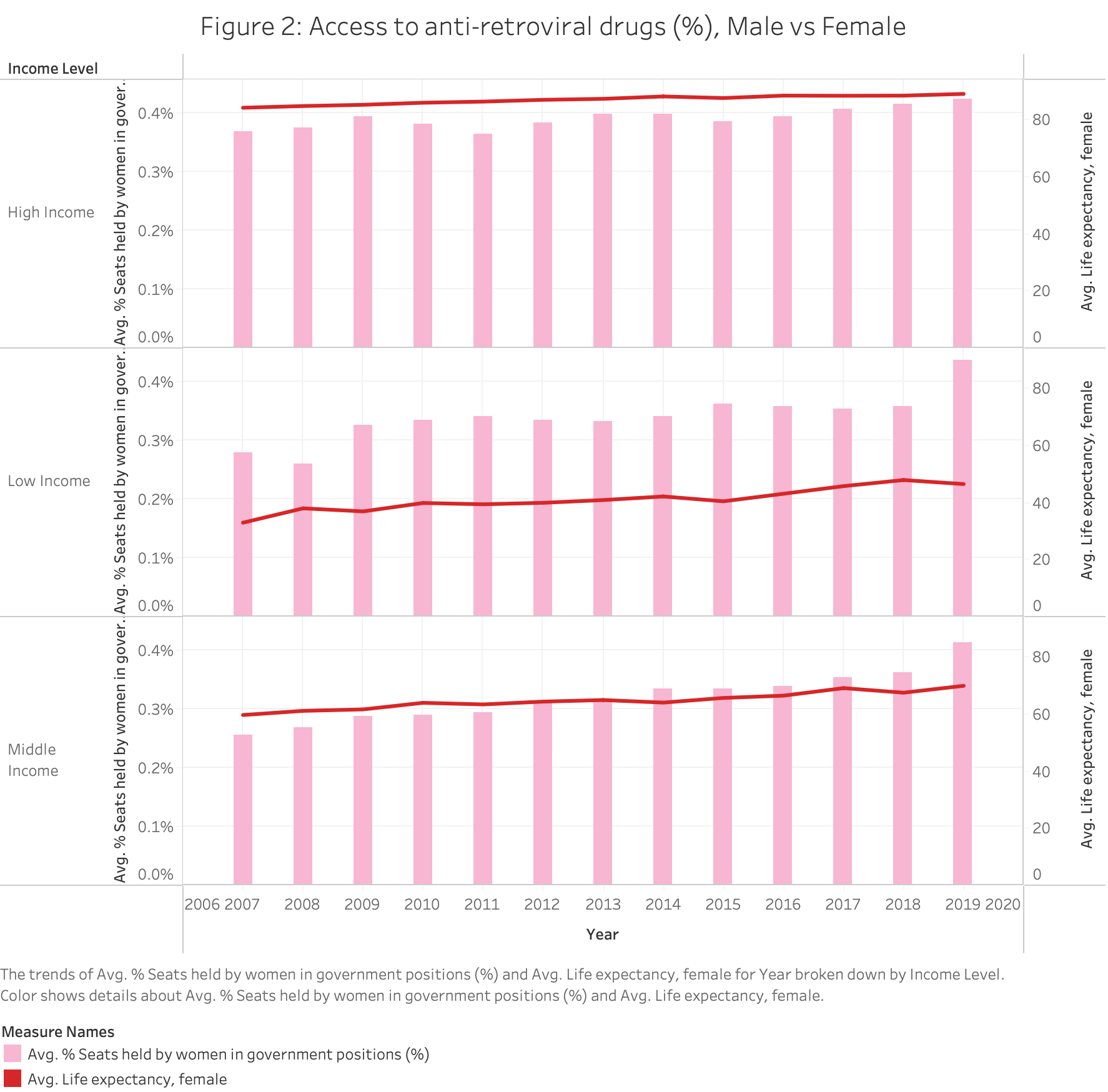
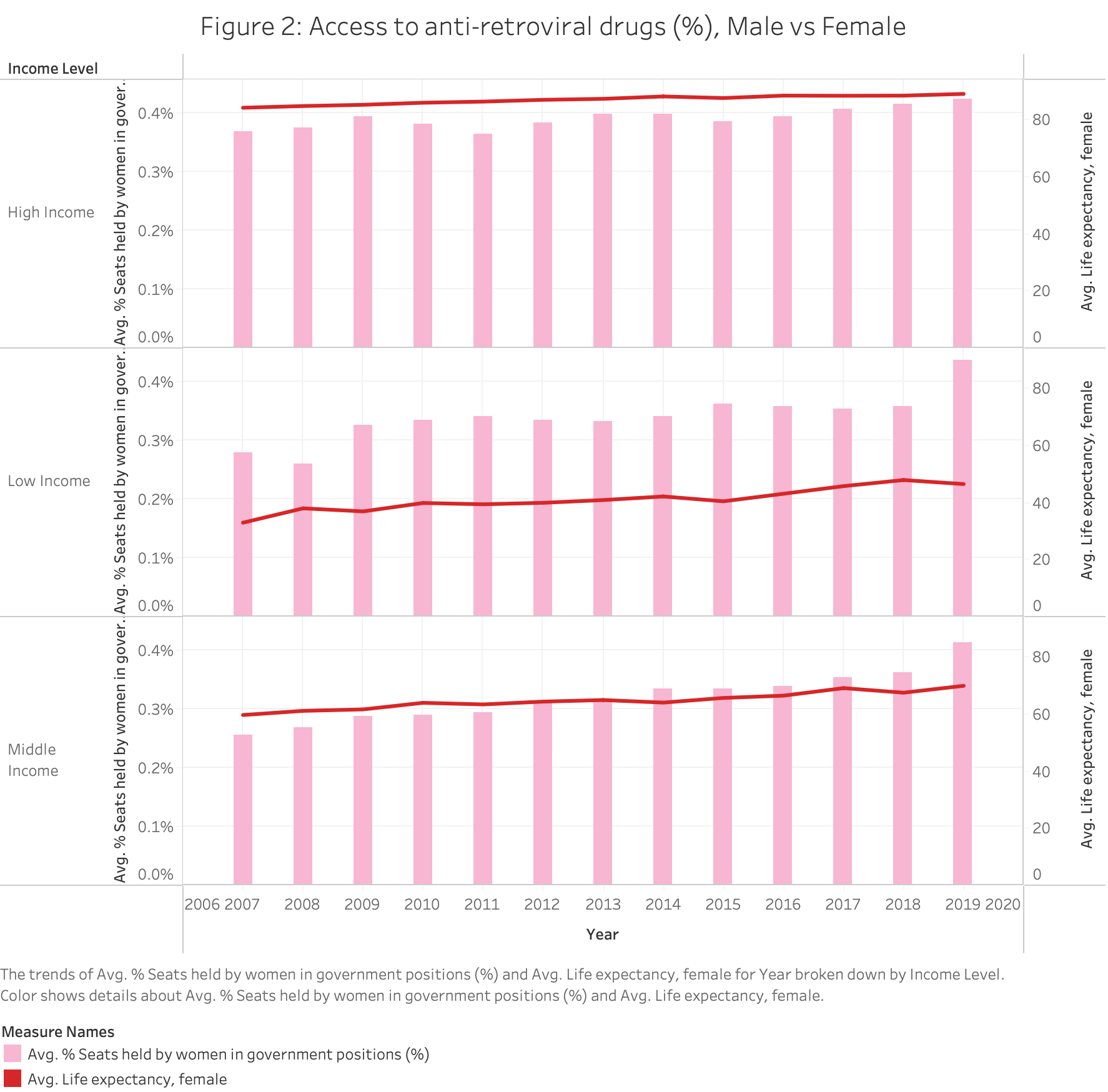
* Pie Charts showing the proportion of recorded HIV incidents in different income levels by gender.
* ​HIV incidence in females was about the same as males in low-income countries, but notably higher in middle-income nations.​
* This implies negligence in the education & prevention of HIV for women in middle-income countries.​

Figure #5- School Enrollment and Fertility Rates



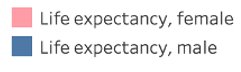
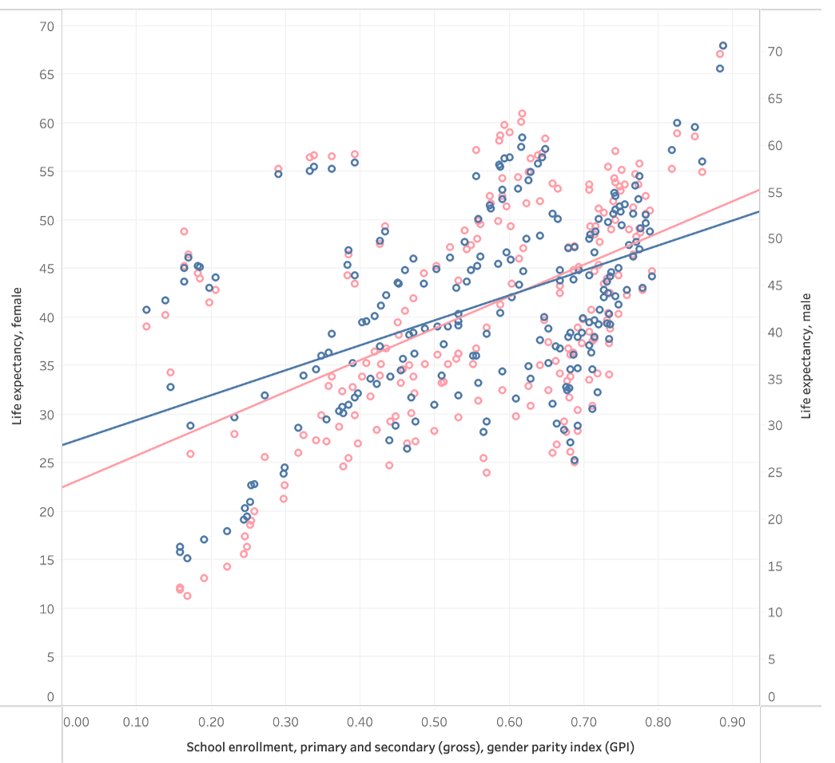
* Bubble chart shows relationship between school enrollment and fertility rates by Region. The bigger the circle, the higher the fertility rate, the darker the circle the higher the school enrollment for females.
* Africa is the region with the lowest female enrollment rate, but highest fertility rate for women.
* Increased female enrollment in schools is correlated with lower fertility rates.

Figure #6- Proportion of Women in National Parliament and Life Expectancy among females



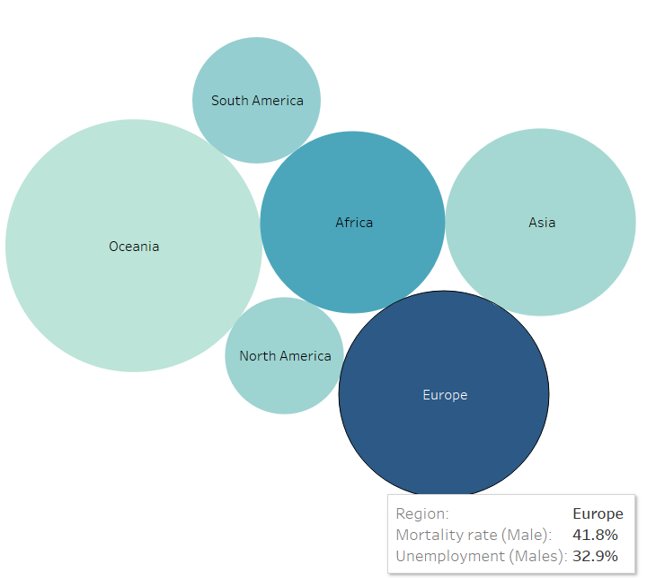
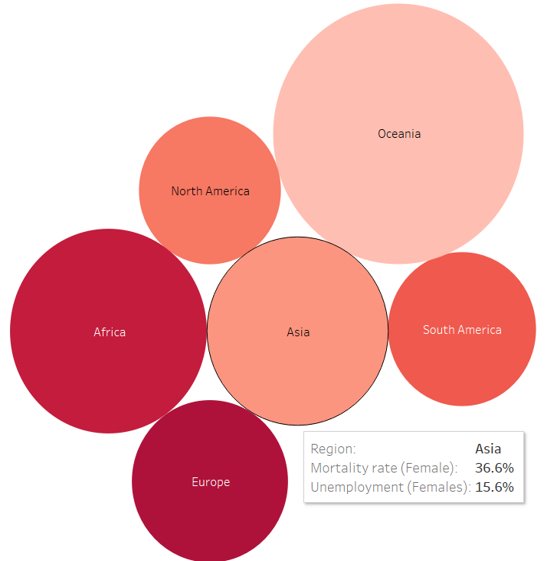
* Bar graph and line chart combo shows the percentage of seats held by women in parliament and female life expectancy by varying income level.
* Female representation in parliament has been increasing in countries across all three income levels, while average life expectancy for females has also increased over time, with low income countries growing at the slowest rate of the three groups.
* Low income countries lack the same motivation to promote women to government and decision-making roles, which may be a contributing factor to why life expectancy of females is lagging behind higher income countries.

Figure #7- GPI in School and Life Expectancy by Gender



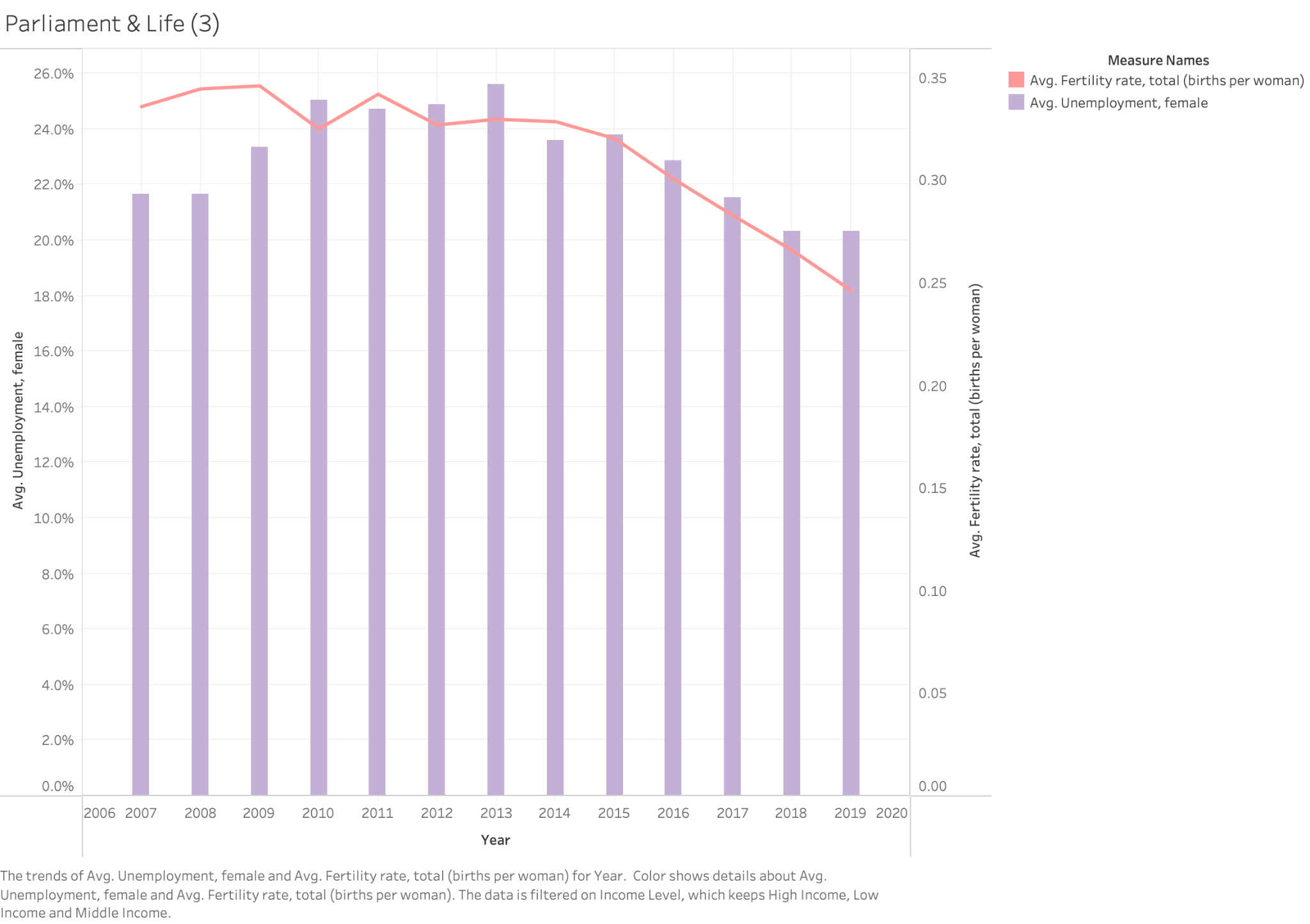
* Scatter plot with regression lines displaying gender parity in school enrollment & life expectancy for males and females.​
* Life expectancy is positively correlated with school enrollment GPI, but female life expectancy exceeded male life expectancy as parity increased.​
* This suggests increasing equity in school enrollment may have reduced overall gender inequity in society and improved health for women.

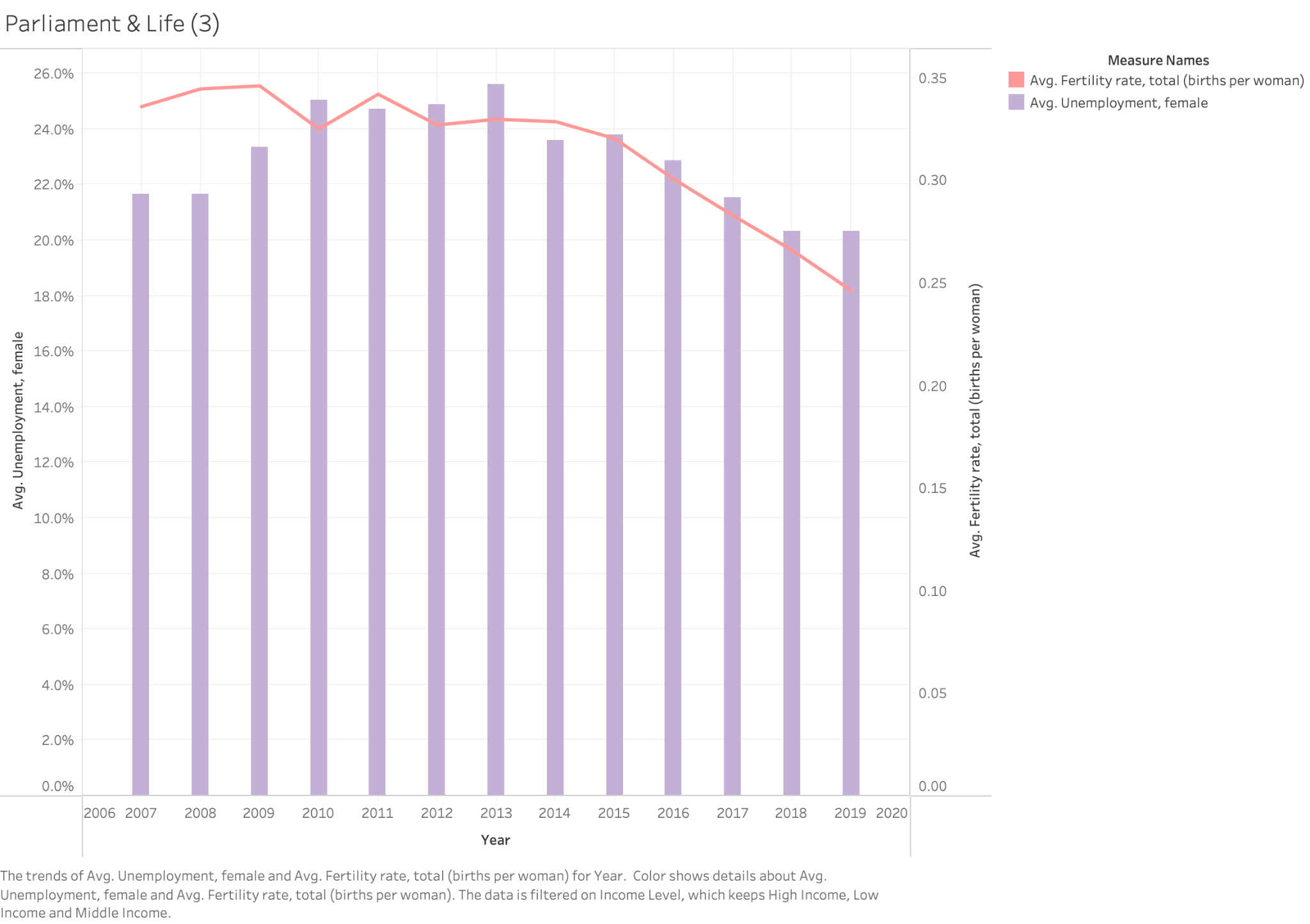
Figure #8- UnemploymentRate vs Mortality from CVD, cancer, diabetes, or CRD by Gender



* Bubble Charts showing unemployment and mortality rate by Region for both males and females. Size indicates mortality rate and color indicates unemployment rate. ​
* Oceania countries had low unemployment rates but high mortality rates for both genders. ​
* Higher unemployment is correlated with higher mortality rates.

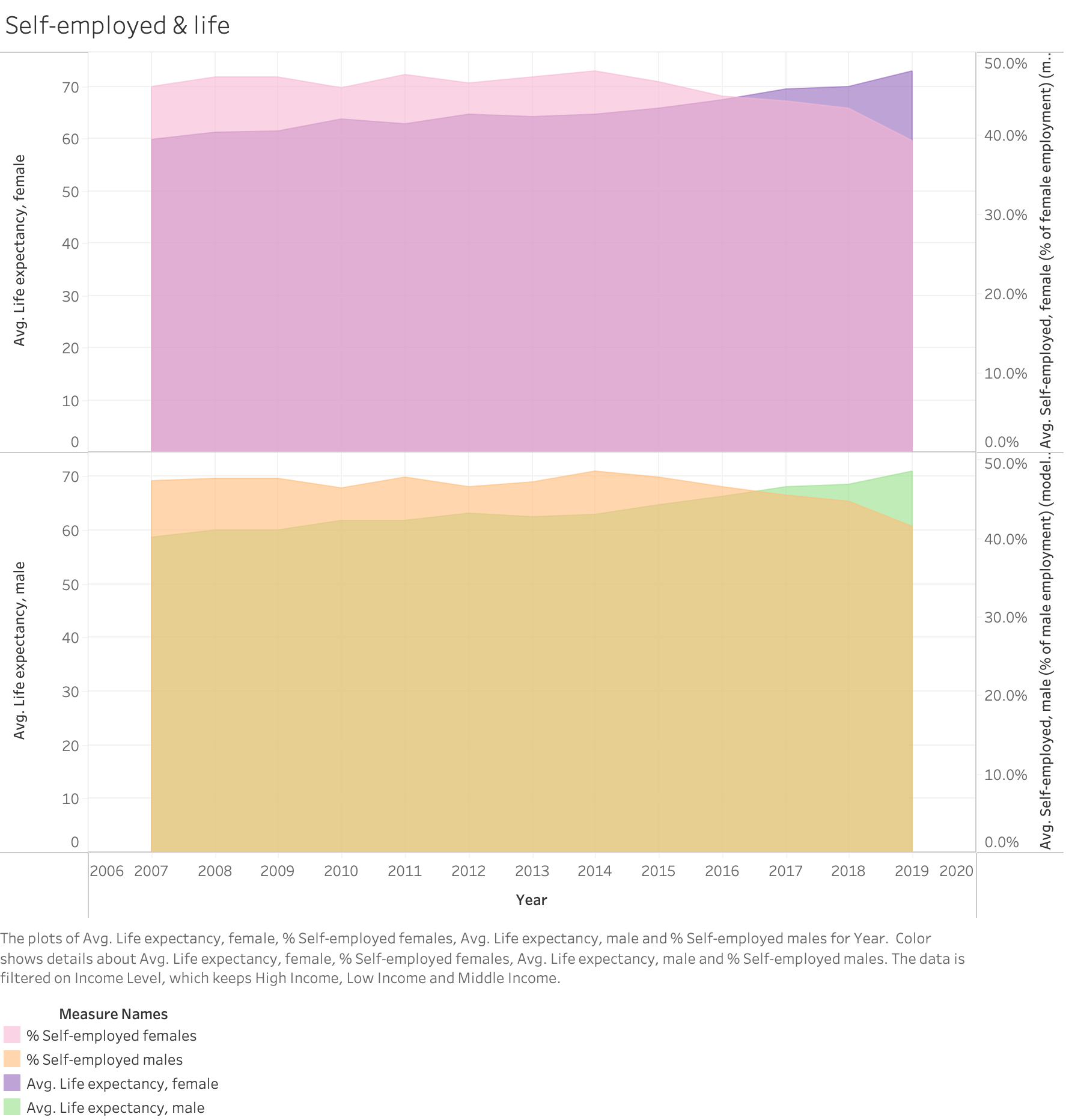
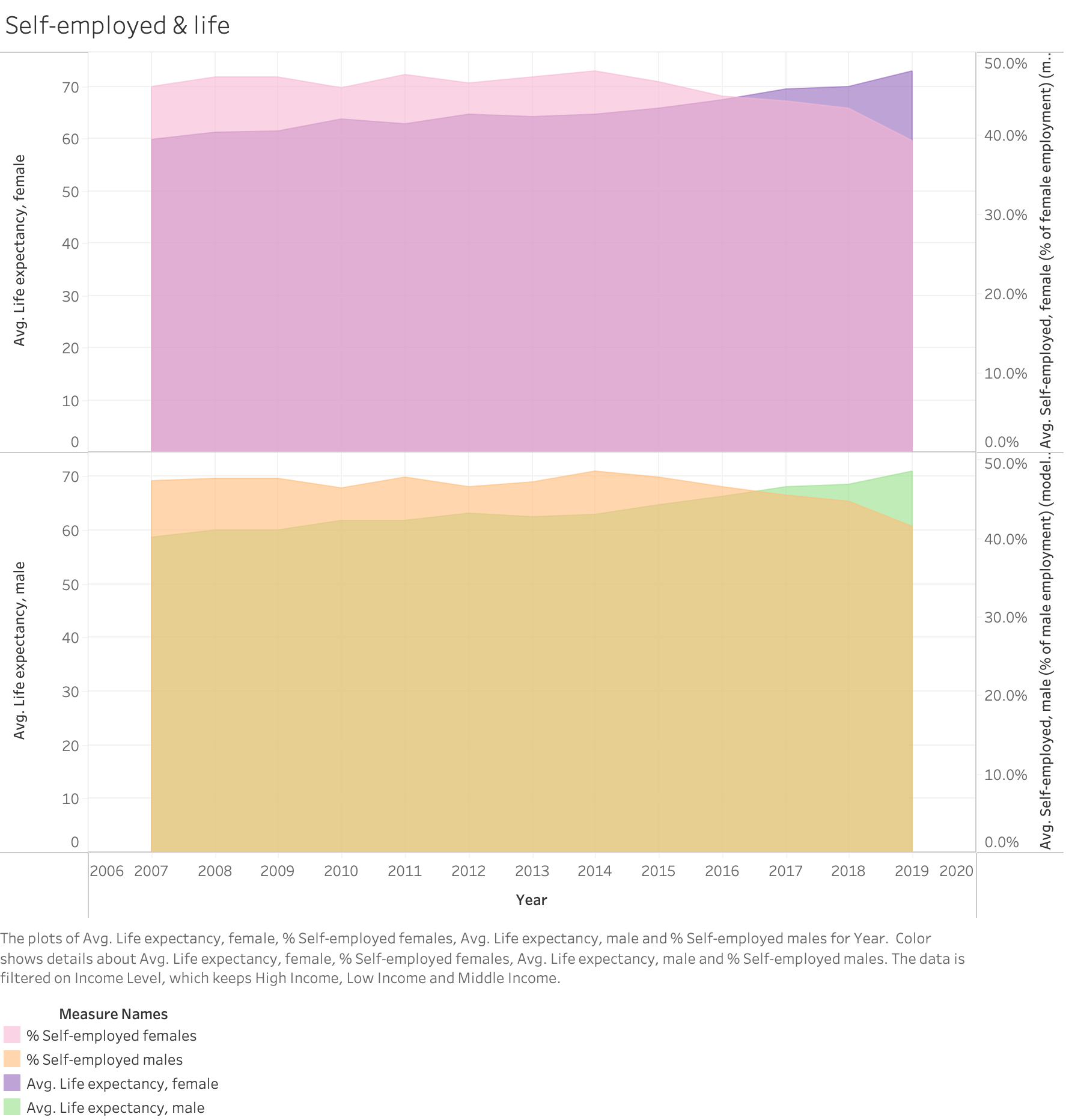
Figure #9- Female Unemployment and Fertility





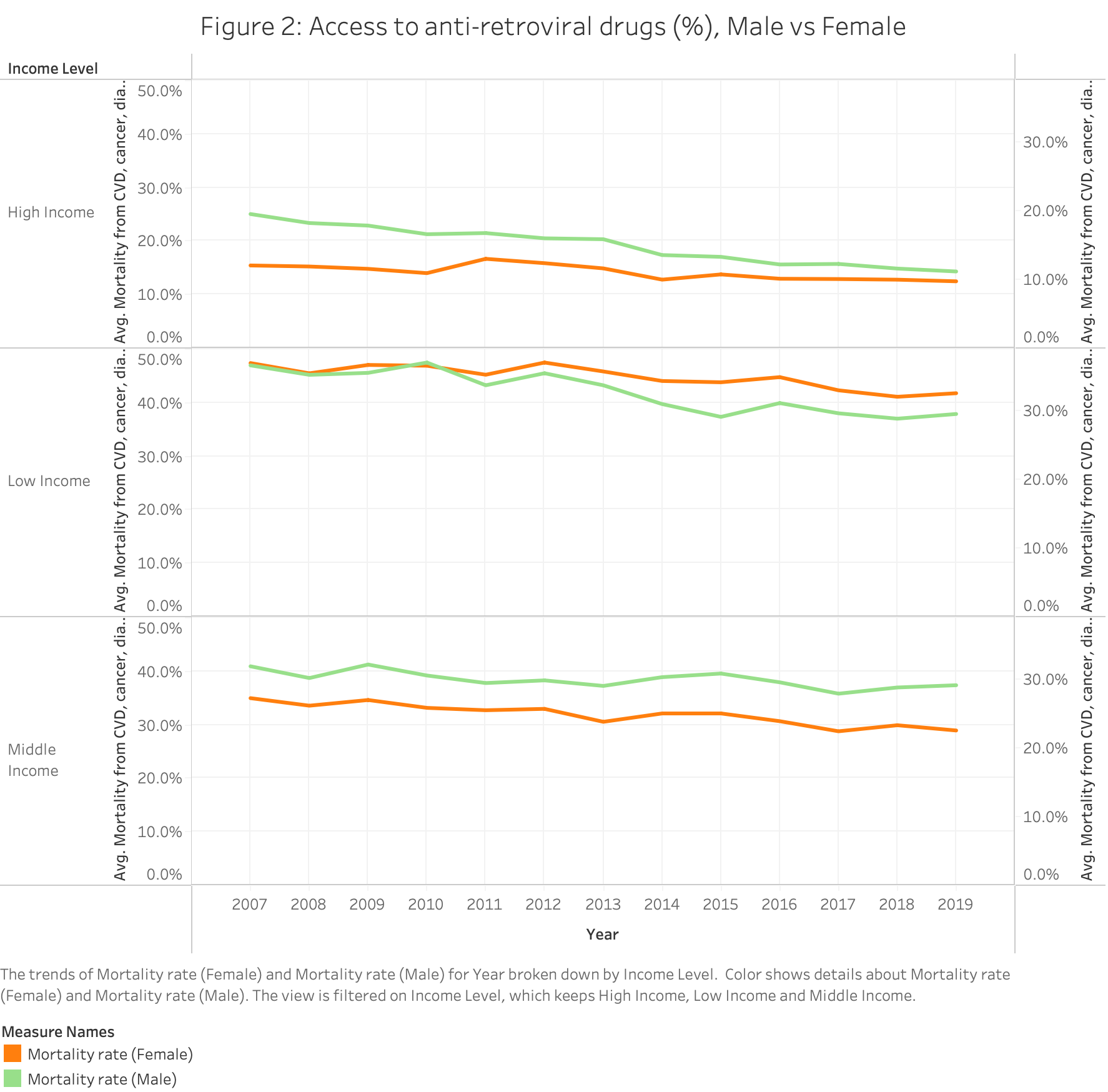
* Combo chart of bar and line graph shows the yearly trend of unemployment and fertility rates for females.
* Average fertility rates have been declining globally since 2006, while unemployment rate has fluctuated, hitting its peak around 2010-2013.
* This chart suggests that there is no clear relationship between women’s fertility and unemployment rates.

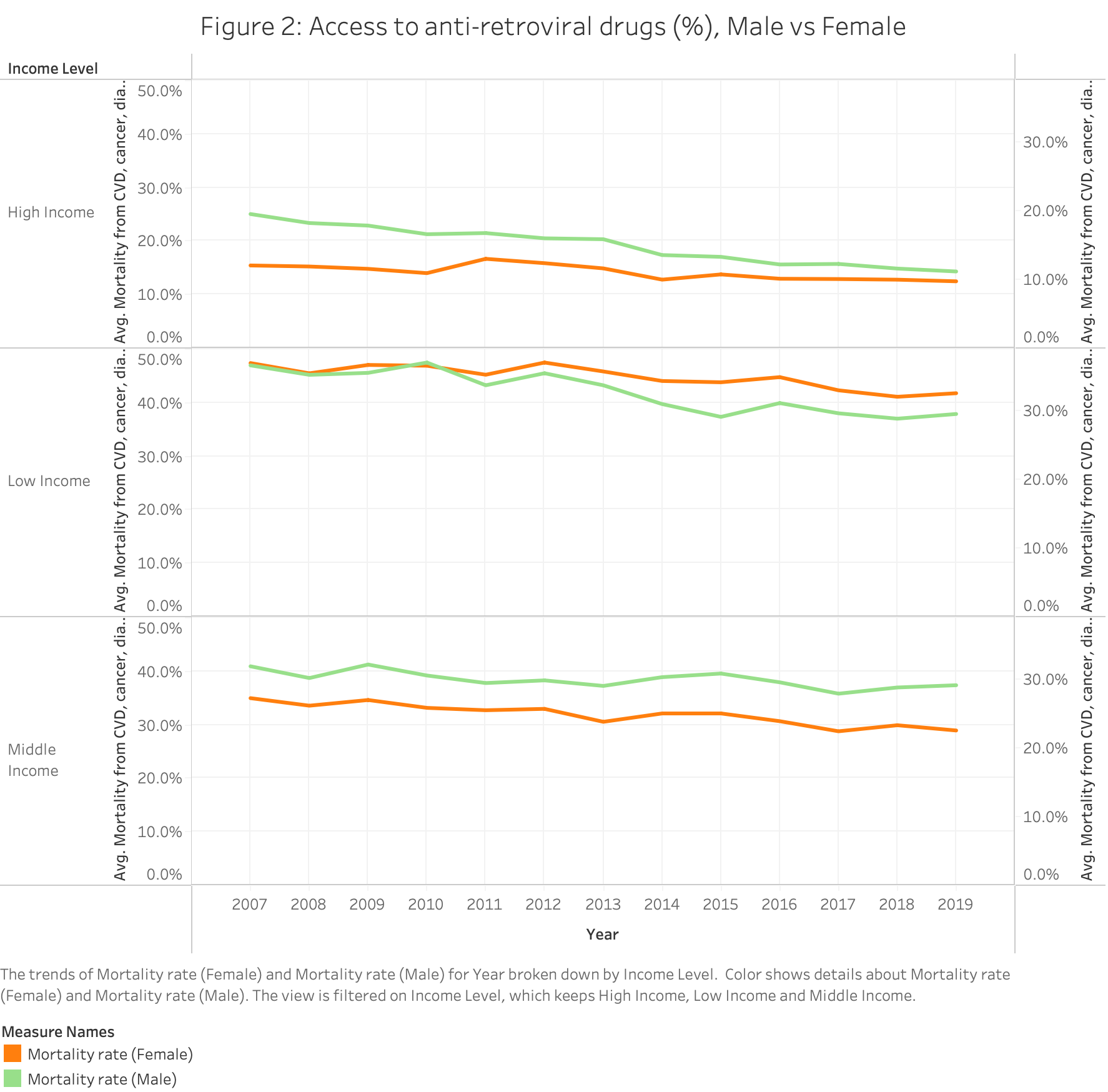
Figure #10- Correlation of Self-Employment and Life Expectancy



* These two side-by-side area charts explore the relationship between self-employed individuals and life expectancy for each gender.
* Both graphs show a similar relationship, with self-employed work decreasing over the years (which is described as household work), while life expectancy has increased.
* There seems to be no difference between males and females who are self-employed and as it pertains to life expectancy.

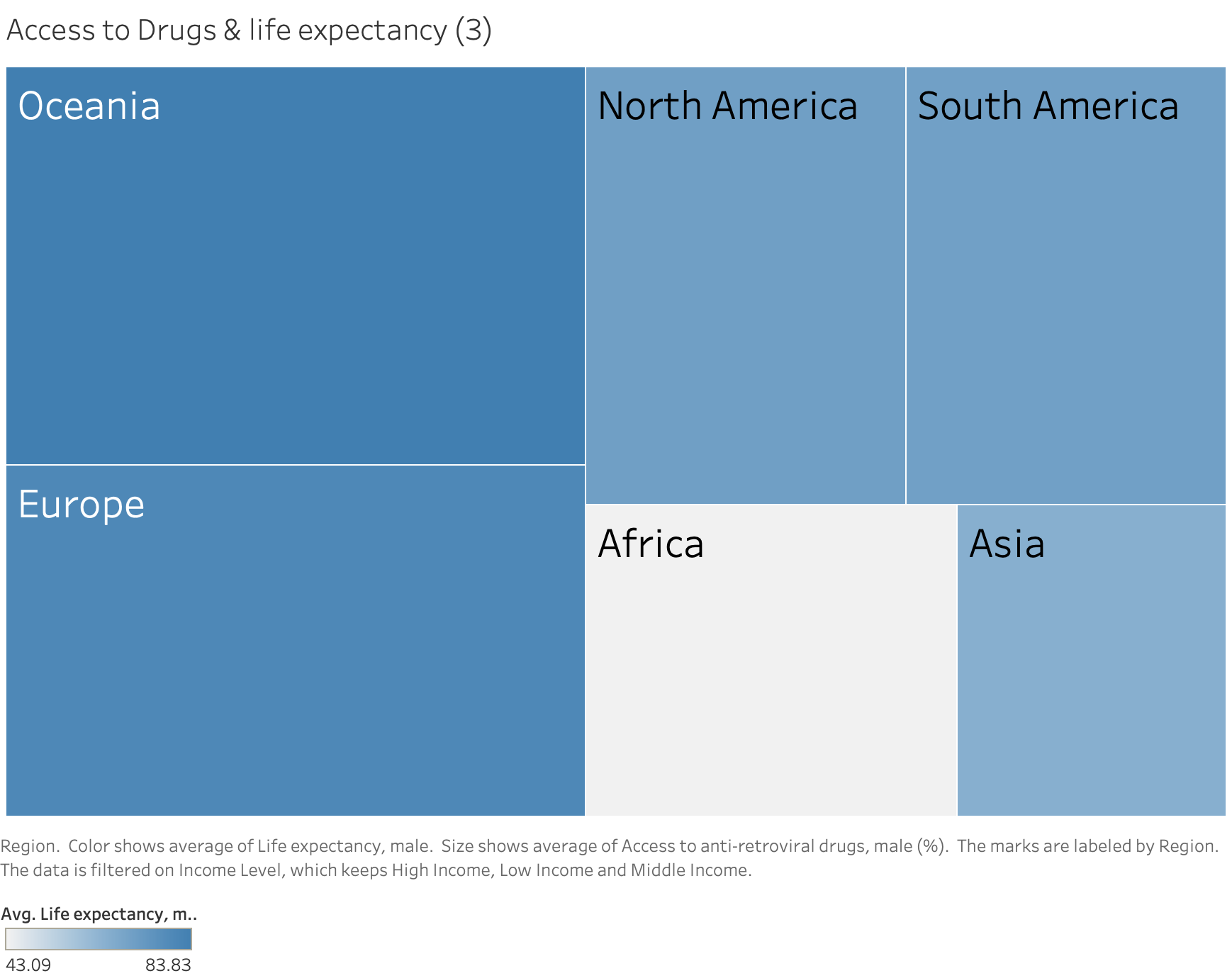
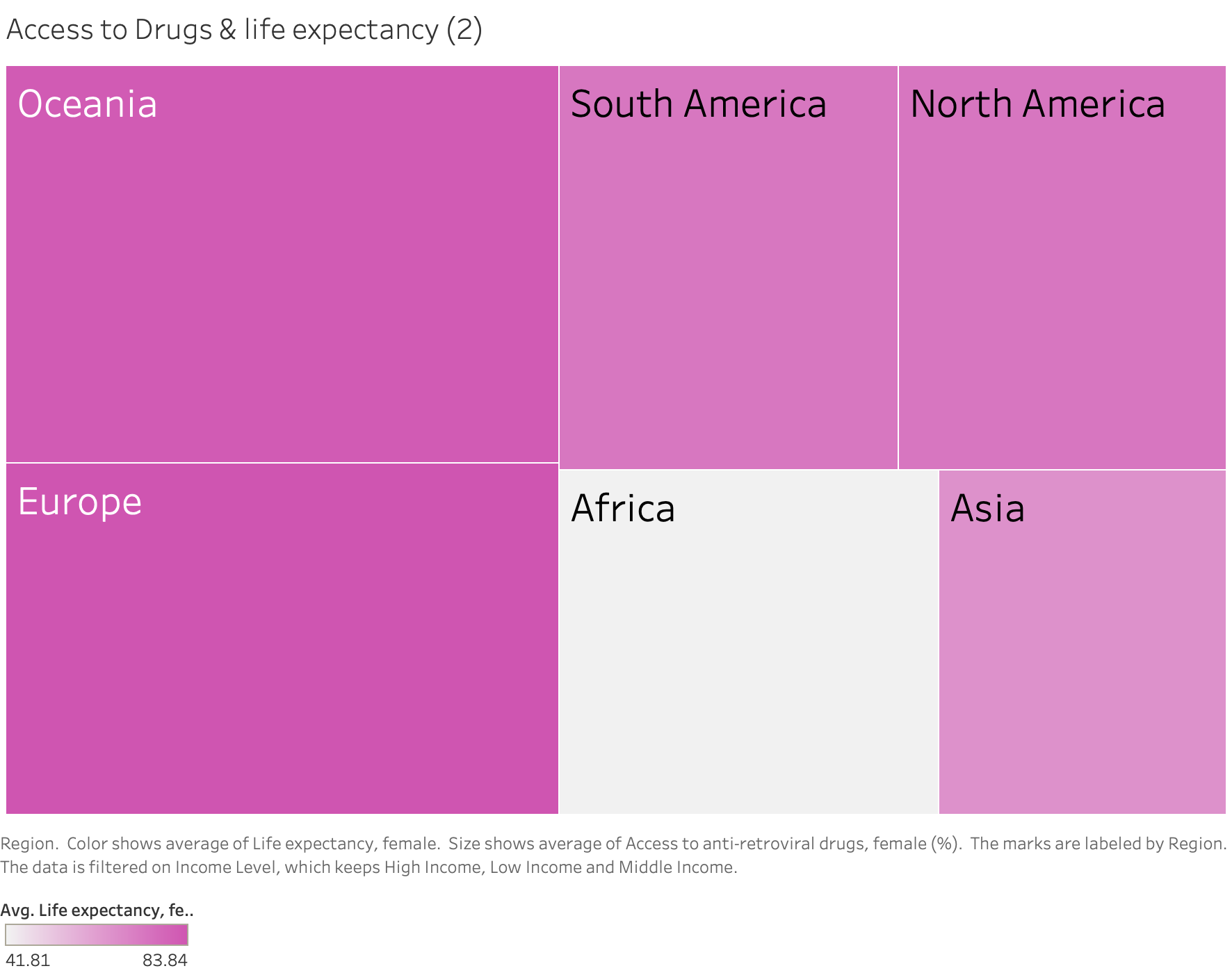
Figure #11- Correlation between Economic Conditions by Country and Mortality Rates

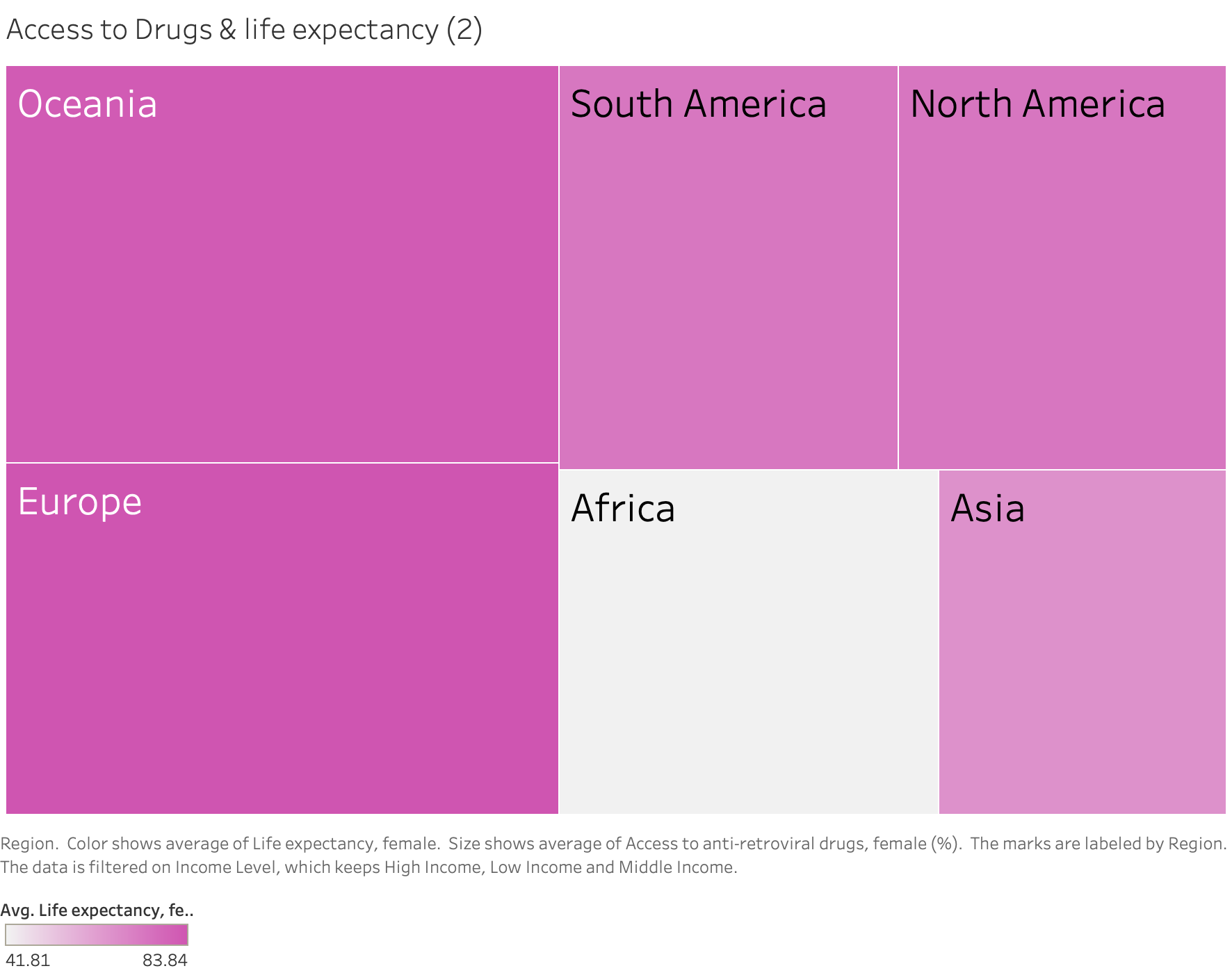
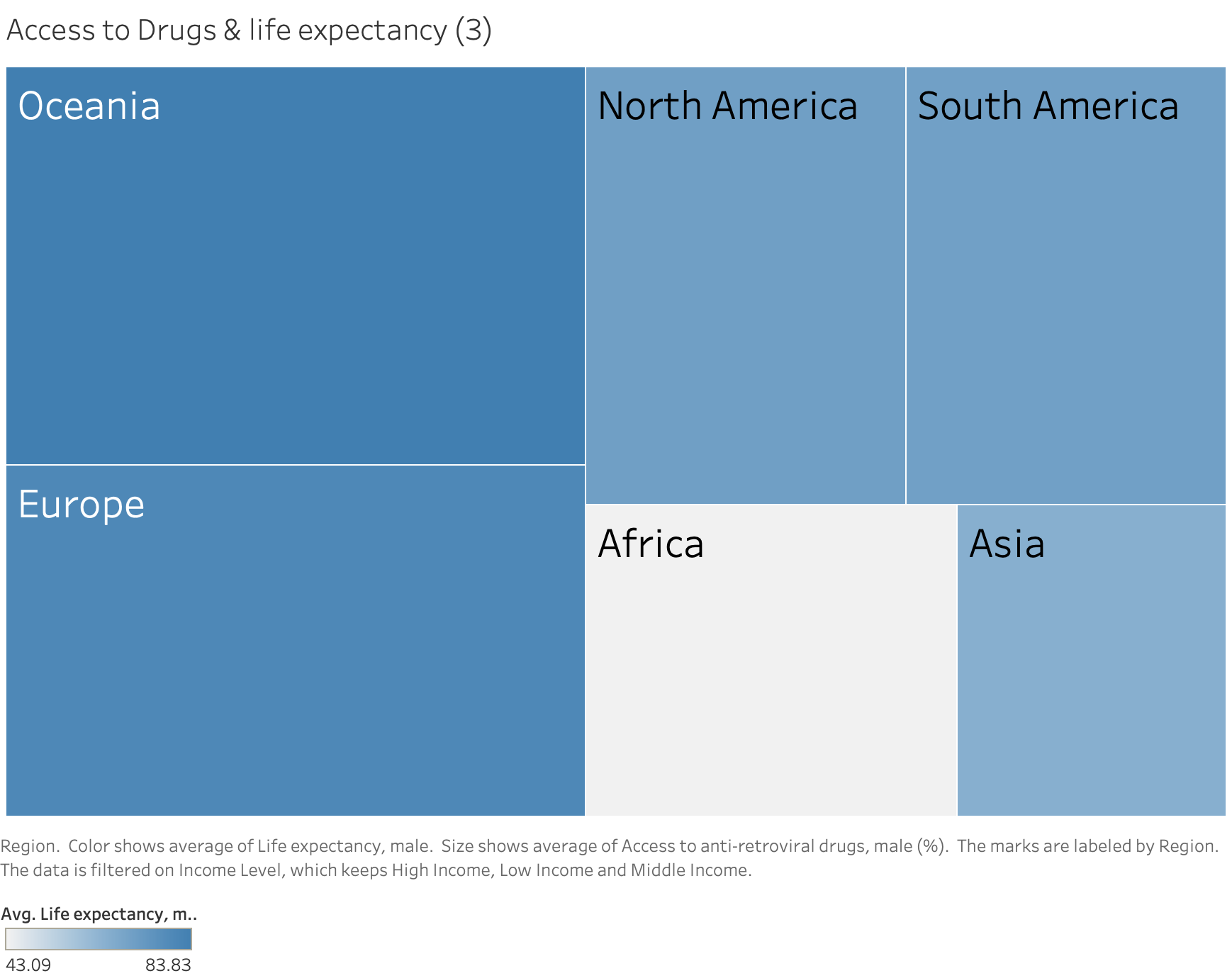




* Three side-by-side line graphs show the trend of mortality rates for males and females based on income level.
* Female mortality rates are lower than males in middle income and high income countries, but higher than males in low income countries.
* This suggests that in low income countries, less resources and attention is geared toward women's health, leading to increased mortality for females.

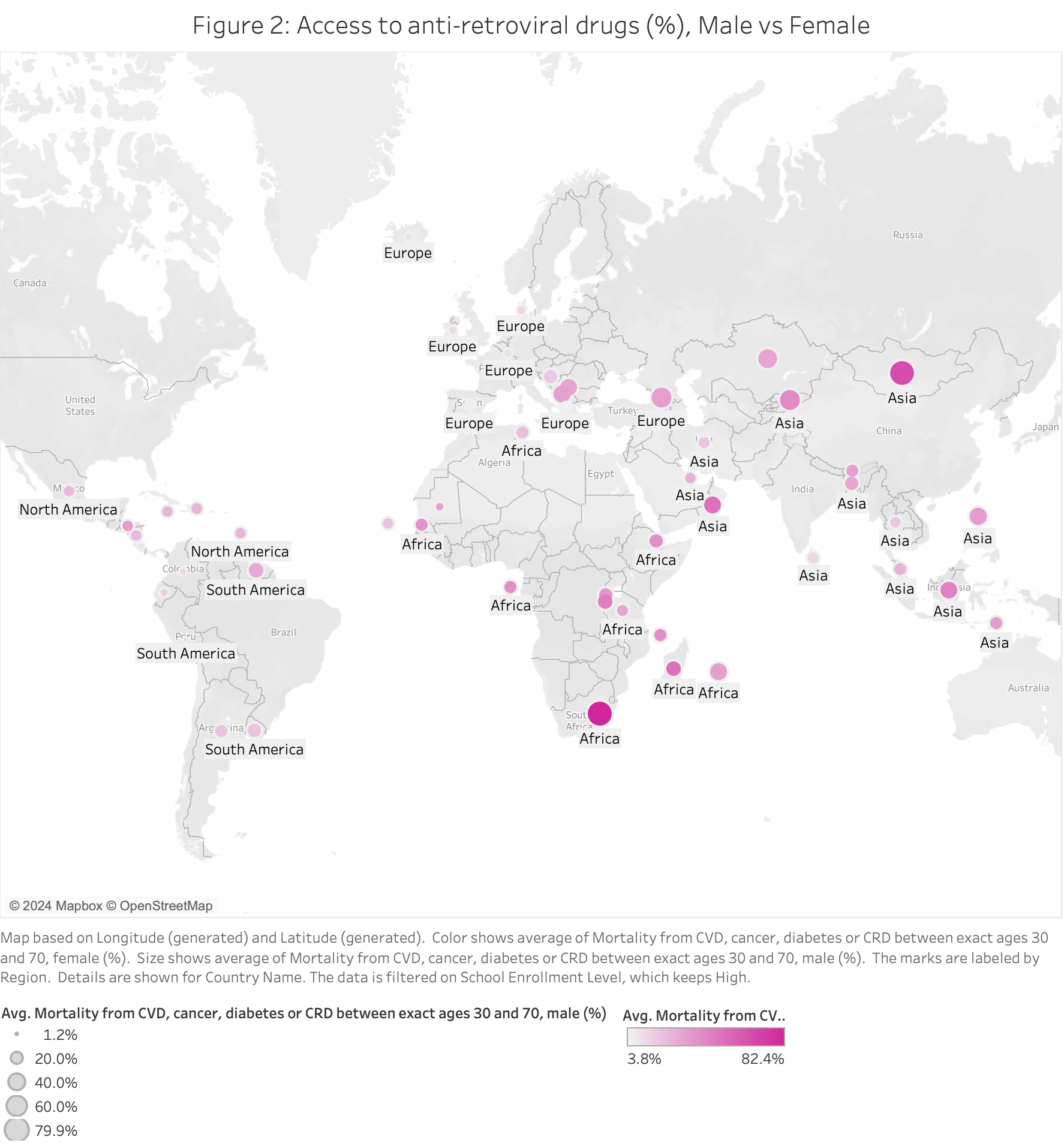
Figure #12- Access to Anti-Retroviral Drugs and Life Expectancy by Gender

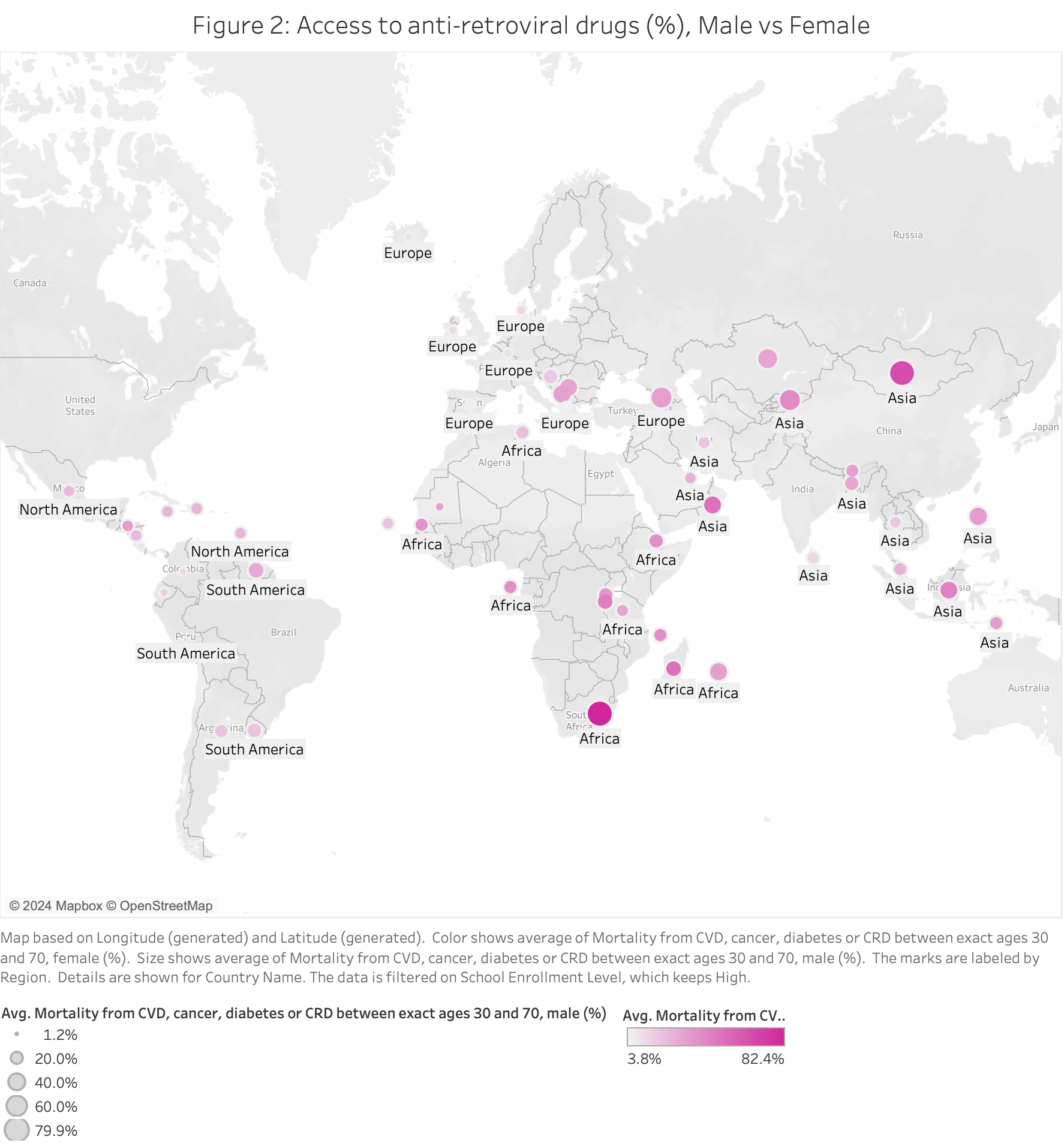


* The two tree maps show access to antiretroviral drugs and life expectancy by region.
* For both genders, Africa has the lowest life expectancy and the second lowest access to anti-retroviral drugs.
* As Africa is predominantly where the HIV Incidence reports are documented, this chart suggests a socioeconomic issue deeper than gender - access to HIV drugs is lowest in the countries that need it most, impacting both genders.

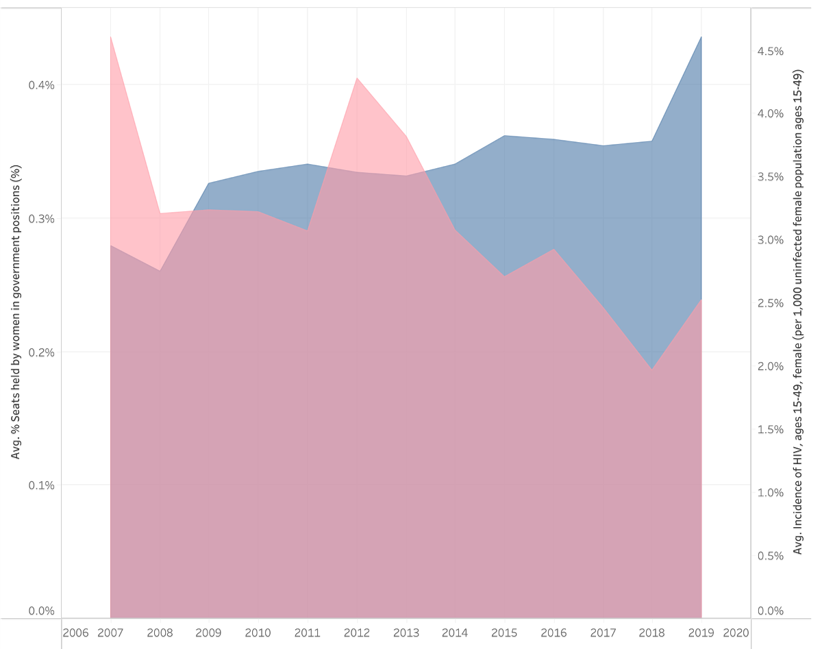
Figure #13- School Enrollment GPI and Mortality Rates



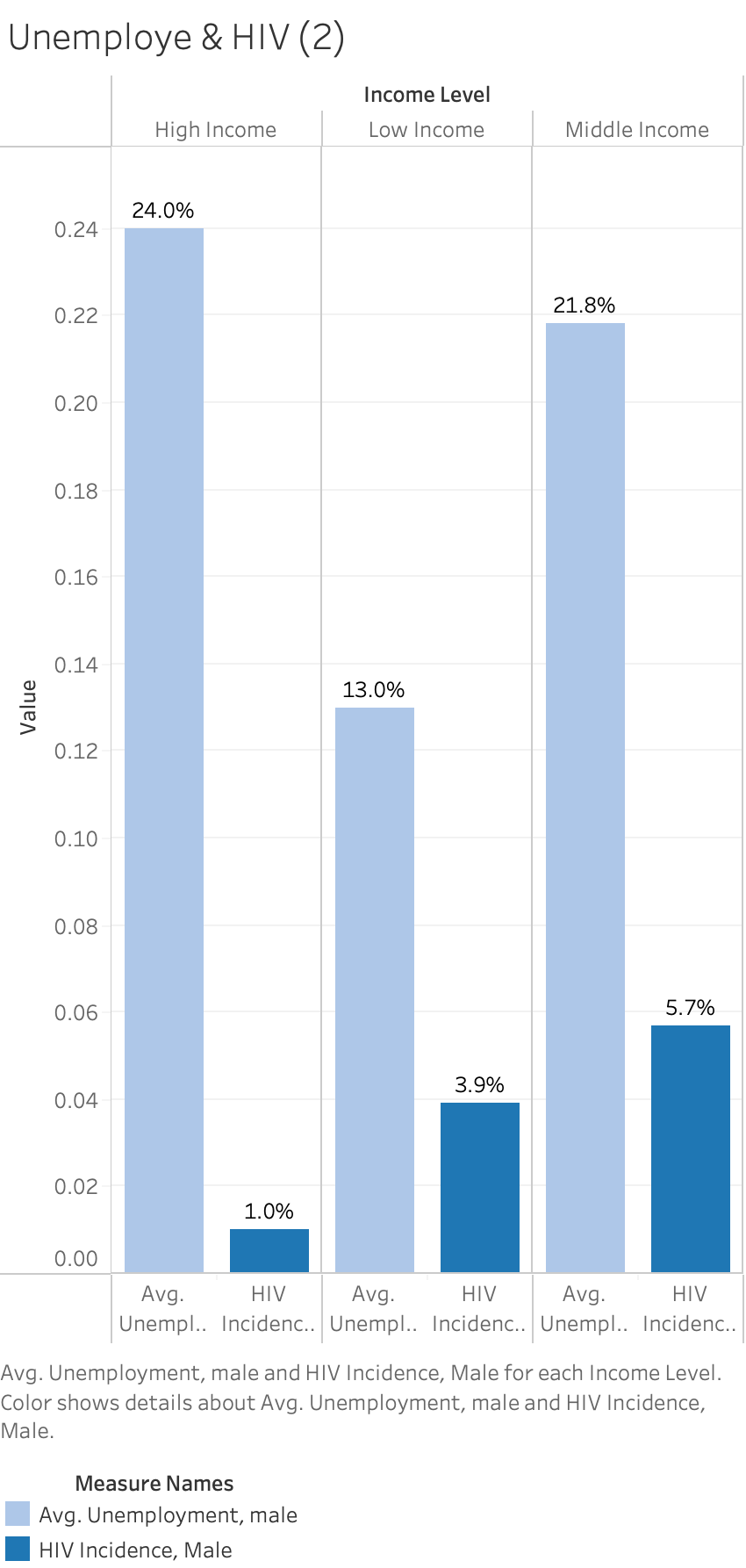
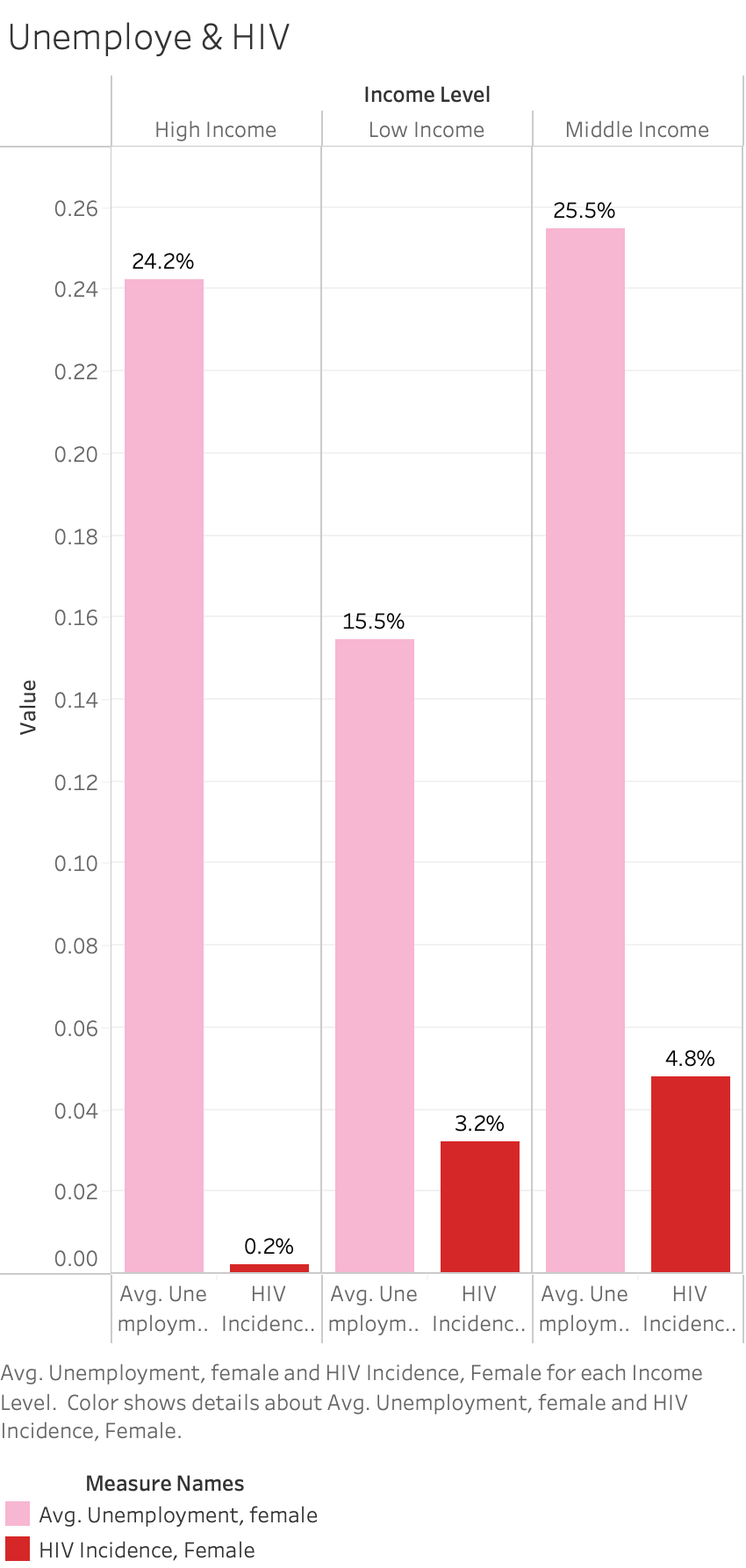


* Map represents the average mortality for males and females by country, filtered for high school enrollment countries. Size of circle indicates measure for male, the deepness of the color, for females.
* Asia and Africa represents the country with the highest mortality rate amongst males and females, out of the high level of school enrollment.
* Higher increased enrollment of women in school overall leads to lower mortality rates from CVS, cancer, diabetes, or CRD, with the exception of Africa and Asia, which could be due to cultural differences.

Figure #14- Seats Held by Women in Parliament and HIV Incidents

* Area chart illustrating the relationship between the % of government positions held by women & incidence of HIV among females from 2007 to 2019​
* There was an increase in the proportion of political positions occupied by women, but HIV incidence among females fluctuated during this time. ​
* No correlation is observed between the proportion of government positions held by women and incidence of HIV among women.

Figure #15- Unemployment and HIV Incidences Across Genders

* Side by side bar charts represent the relationship between unemployment and HIV, pink and red bars for females, and blue bars for males.
* In middle income countries, unemployment is higher for women than for men, however HIV incidents are more frequent for men than women.
* In middle income countries, lower unemployment in men is correlated with higher incidence of HIV compared to women.

Figure #16a - Descriptive Statistics of Dataset

| **Variable** | **Count** | **Mean** | **Std** | **Min** | **25%** | **50%** | **75%** | **Max** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Access to anti-retroviral drugs, female (%) | 1043 | 41.91371 | 24.78453 | 1 | 21 | 41 | 62 | 98 |
| Access to anti-retroviral drugs, male (%) | 1043 | 36.6069 | 22.81682 | 1 | 18 | 32 | 54 | 97 |
| Fertility rate, total (births per woman) | 1043 | 2.915721 | 1.538994 | 1.14 | 1.722 | 2.345 | 3.986 | 7.559 |
| GDP per capita, PPP (current international $) | 1043 | 16604.06 | 18348.25 | 558.9002 | 3869.955 | 10861.36 | 21604.6 | 119652.8 |
| Immunization, measles (% of children ages 12-23 months) | 1043 | 87.64621 | 12.79022 | 27 | 85 | 92 | 96 | 99 |
| Incidence of HIV, ages 15-49, female | 1043 | 1.411342 | 5.11357 | 0.01 | 0.05 | 0.15 | 0.47 | 42.22 |
| Incidence of HIV, ages 15-49, male | 1043 | 1.025369 | 2.977849 | 0.01 | 0.14 | 0.31 | 0.57 | 24.34 |
| Life expectancy at birth, female (years) | 1043 | 73.10283 | 9.123131 | 45.163 | 66.6205 | 76.086 | 80.0115 | 86.7 |
| Life expectancy at birth, male (years) | 1043 | 67.93706 | 8.379194 | 40.811 | 62.3155 | 68.949 | 73.8 | 81.7 |
| Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, female (%) | 1043 | 17.08955 | 6.884573 | 5.7 | 11.5 | 16.9 | 21.6 | 42.9 |
| Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, male (%) | 1043 | 23.53231 | 9.068081 | 9.9 | 16.7 | 22.2 | 29.3 | 55.9 |
| Population, female | 1043 | 12191227 | 19219011 | 85124 | 1725685 | 5409472 | 14471225 | 1.33E+08 |
| Population, male | 1043 | 12026971 | 19260244 | 84721 | 1668532 | 5203501 | 14601447 | 1.35E+08 |
| Proportion of seats held by women in national parliaments (%) | 1043 | 21.87426 | 11.53467 | 0.332226 | 12.76596 | 20.50633 | 28.66667 | 63.75 |
| School enrollment, primary and secondary (gross), gender parity index (GPI) | 1043 | 0.977951 | 0.077974 | 0.60961 | 0.972415 | 0.99833 | 1.018335 | 1.15494 |
| Self-employed, female (% of female employment) | 1043 | 45.63352 | 30.67842 | 2.094409 | 16.19129 | 39.61991 | 73.95222 | 98.81054 |
| Self-employed, male (% of male employment) | 1043 | 43.22483 | 23.27762 | 2.458484 | 24.02953 | 40.86448 | 59.45428 | 92.37865 |
| Unemployment, female (% of female labor force) | 1043 | 8.876871 | 6.851447 | 0.218 | 3.8705 | 6.424 | 12.1755 | 36.326 |
| Unemployment, male (% of male labor force) | 1043 | 7.161685 | 5.340937 | 0.259 | 3.273 | 5.364 | 9.9325 | 25.599 |

Figure 16b: Correlation Matrix

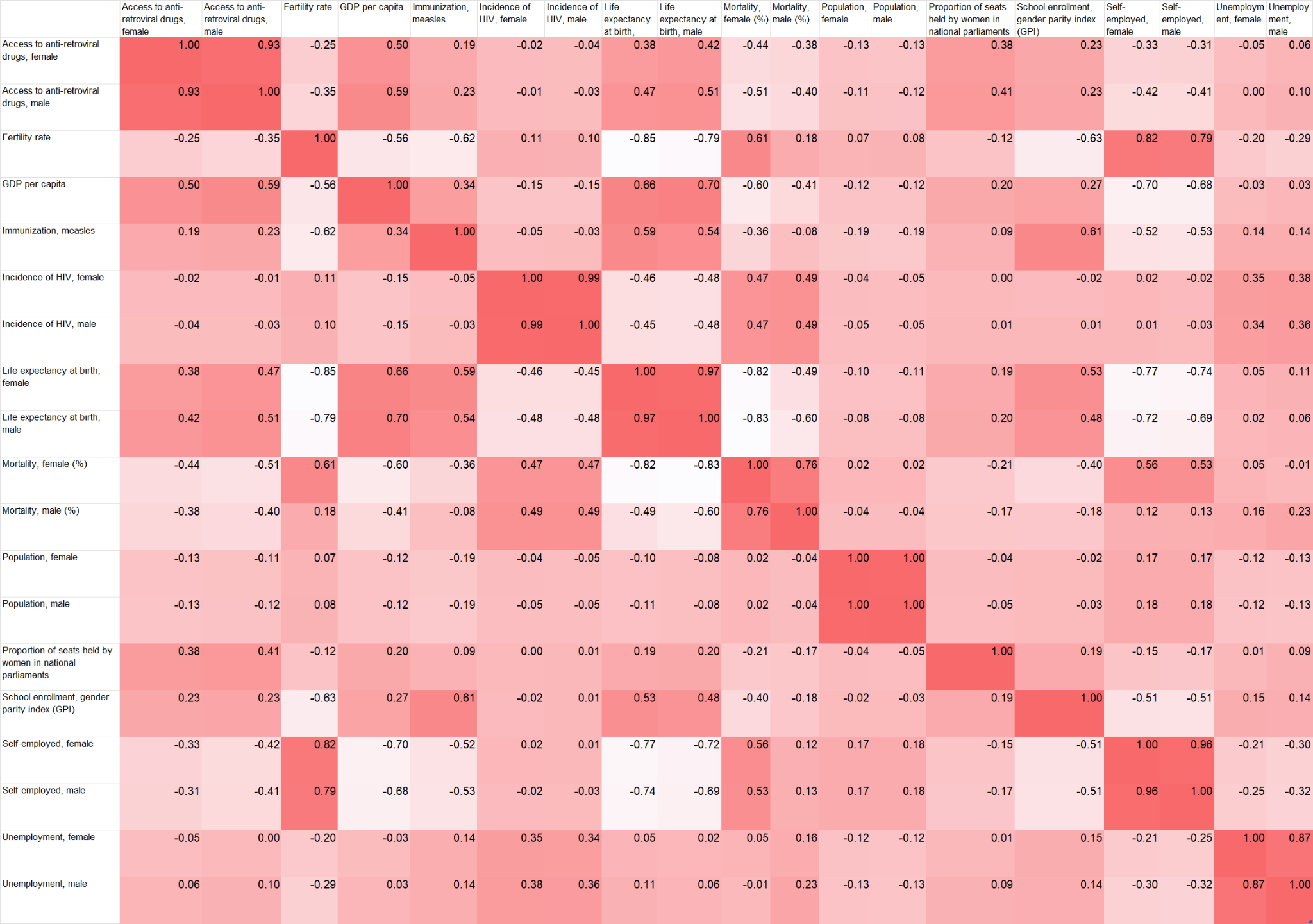
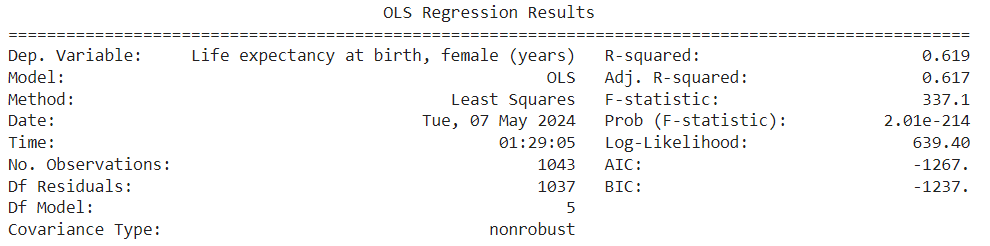
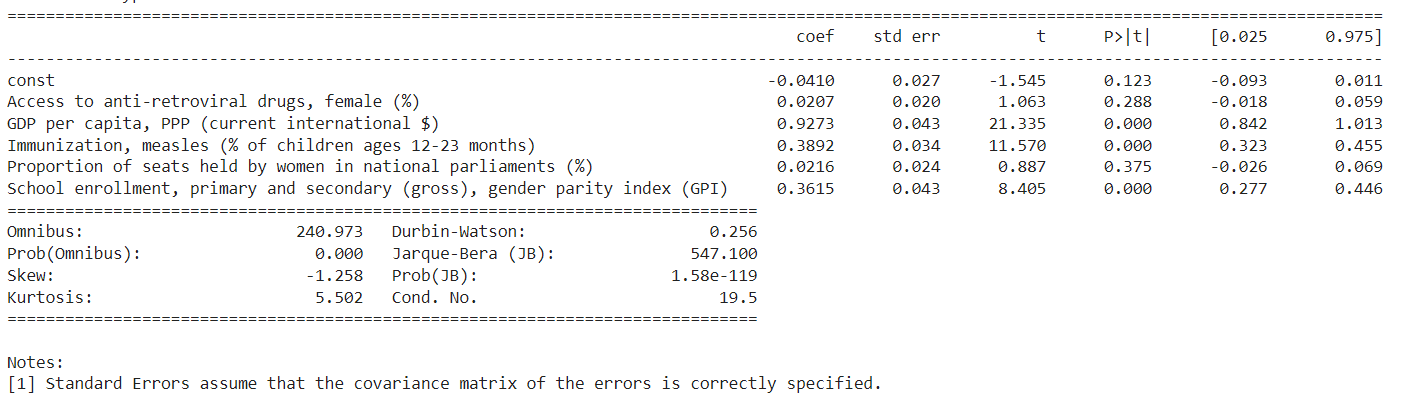


Figure #16c - Multiple Regression





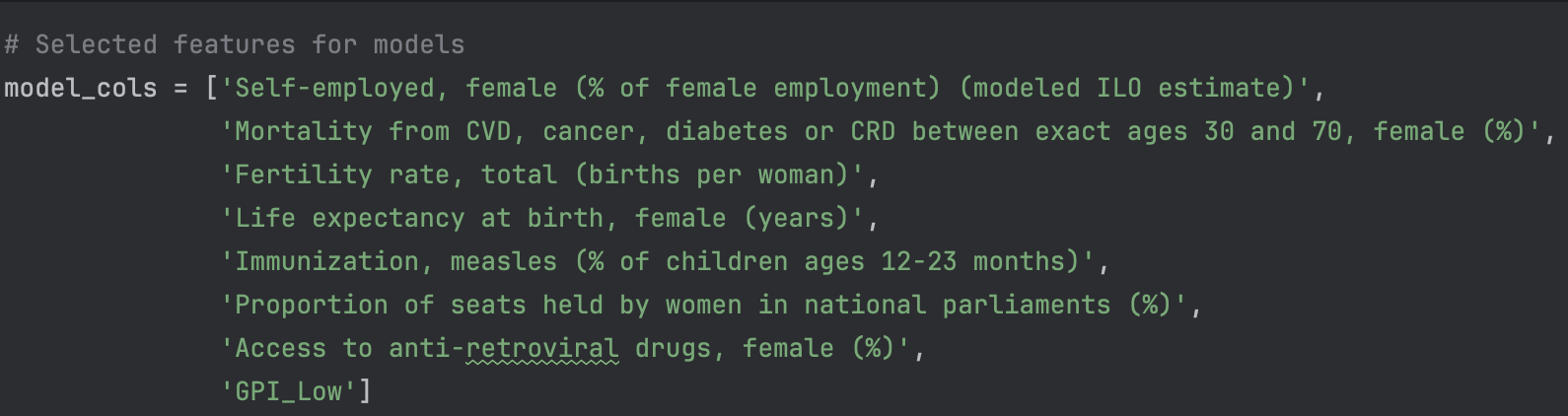
* Multiple linear regression to predict life expectancy in females using the variables, ‘Access to anti-retroviral drugs, female’, ‘GDP per capita’, ‘Immunization, measles’, ‘Proportion of seats held by women in national parliaments’, and ‘School Enrollment (GPI)’.
* The GDP per capita has a highly significant positive effect on female life expectancy, with each unit increase in GDP per capita associated with an increase of approximately 0.9273 years in female life expectancy.
* Policy measures aimed at improving economic conditions and increasing GDP per capita could potentially lead to improvements in female life expectancy.

Figure #17- Machine Learning

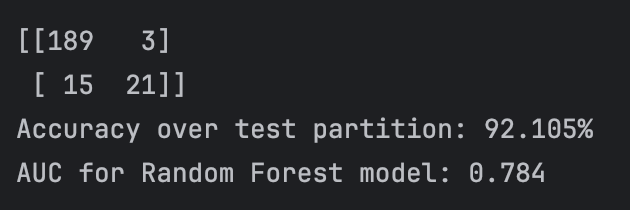
Using school enrollment Gender Parity Index (GPI) as a proxy for overall gender parity, machine learning was used to predict whether a given country could be classified as having a low GPI. According to the World Bank, a GPI of less than 1 suggests females are more disadvantaged than males. Countries were classified as ‘Low GPI’ if their school enrollment GPI was below the 25th percentile in the dataset, which was 0.6815.

Although the classification was created based on the non-normalized output data, the normalized dataset was used to train and test the models. All records 2016 and older were used as the training partition, while the data from 2017-2019 were used to test.

The final features selected were a mix of economic, social, and health-related variables, with emphasis on the variables for females; male predictors were removed to account for potential multicollinearity between features bifurcated by gender (i.e. ‘self-employed, female’ & ‘self-employed, male’ as seen above in Fig. 16b):



While several model candidates were considered, the Random Forest classifier performed the best, with an outstanding 92% accuracy over the testing data:



However, the moderately high AUC signifies that while the selected features were strong predictors, further refinement is needed in order to optimize the model and make more precise predictions. Additional features and/or more robust feature engineering may help improve the performance of the model and elucidate the socioeconomic factors that drive gender parity.