CMSE201_ExampleProject_1

November 10, 2019

1 *Is there an Urban Paradox? A look at Urban Advantage and Possible Issues*

Background: In the last several years, the world has become increasingly urban, with large cities appearing all over the world. Regions see vast increases in their urban population, such as Asia which sees an additional 48 million people move into cities annually, while Africa grows its own urban population by four percent a year ("Advantage or Paradox?", 2018). Along with this occurence, people now live closer to everything they need, allowing for great advances in public health and economic growth. It is seen that on average, urban households experience higher income, improved infrastructure, better education and access to services ("Advantage or Paradox?", 2018). It has come to be known as the "urban advantage".

Recent studies from UNICEF, however, reveal that "not all urban children are benefitting equally"("Advantage or Paradox?", 2018). According to their findings, while on average there does seem to be an urban advantage, a closer look reveals that in some regions the poorest children in urban regions actually fare worse than their rural counterparts, in a phenomenon referred to as the "urban paradox" ("Advantage or Paradox?", 2018). It suggests that wealth disparity and poverty in becoming an increasing problem in urban areas and that large sections of the urban population might be falling behind.

It is important to understand and research this concept as it is necessary to see where the issue might occur and thus work to focus on and fix these problems before they develop even further. As the world grows increasingly urban, the urban paradox will continue to become more of an issue and so it is crucial to look at now to see how it might affect the future as well as how people might prevent its spread. This study will look to evaluate urban advantage across ten indicators of health, to show where it exists when the advantage is greatest. It will then look to see how wealth impacts urban advantage and where there might be the greatest wealth disparity, and finally whether there is evidence of the urban paradox in any indicators and if so, where it exists. It will look to see how well these results match up with the findings of UNICEF.

Methodology: This study will require large amounts of data analysis, largely stemming from the data collected and made public by UNICEF. As a first step, each file was imported using the pandas library, the first containing general data on population and a value for each indicator for each country. The ten indicators served as a measure of health for the people of the group, including access to drinking water, access to sanitation services, the prescence of a skilled attendant during childbirth, birth registration prevelance, DTP3 immunization coverage, completion of primary education, comprehensive knowledge about HIV/AIDS for both genders, the rate at which children are stunted, and the under five mortality rate. A high percentage is desired for each indicator except for in the case of stunting and mortality, as both are signs of poor healthcare and malnutrition.

The second dataset revealed similar data to the first except it included data split up between all wealth quintiles of the countries, so that differences in health with wealth could be analyzed. Once the data was imported, the .head() function was also used to see the first few rows of data for each case.

[1]:
```python
import pandas #importing the pandas library
# These were two datasets containing data from all observed countries,␣
↪including percentage values for each indicator both urban and rural as well␣
↪as total
urban_paradox = pandas.read_csv('UrbanParadox_data_new.csv') # includes general␣
↪data for urban and rural populations, with values for each indicator both␣
↪urban and rural as well as urban, rural, and total populations for each␣
↪country
wealth = pandas.read_csv('UrbanParadox_data_wealth.csv') # includes similar␣
↪data to previous file but shows data for all wealth quintiles for both rural␣
↪and urban populations
```

[2]:
```python
#showing the first few lines of data to see column names and check work
urban_paradox.head()
```

[2]:
```
     Area name Iso3code  Region                            Indicator name  \
0  Afghanistan      AFG    Asia  At least basic drinking water services
1  Afghanistan      AFG    Asia  At least basic drinking water services
2  Afghanistan      AFG    Asia  At least basic drinking water services
3       Angola      AGO  Africa  At least basic drinking water services
4       Angola      AGO  Africa  At least basic drinking water services

  Survey type  Year Residence  Indicator value  Unweighted population  \
0         DHS  2015     Total         0.630880               201833.0
1         DHS  2015     Urban         0.864950                52023.0
2         DHS  2015     Rural         0.554475               149810.0
3         DHS  2015     Total         0.488314                73914.0
4         DHS  2015     Urban         0.656363                42719.0

   Weighted population  Small Sample
0         196048.30060           0.0
1          48245.90109           0.0
2         147802.39950           0.0
3          77465.13155           0.0
4          49803.87501           0.0
```

[3]:
```python
wealth.head()
```

[3]:
```
     Area name Iso3code Region                            Indicator name  \
0  Afghanistan      AFG   Asia  At least basic drinking water services
1  Afghanistan      AFG   Asia  At least basic drinking water services
2  Afghanistan      AFG   Asia  At least basic drinking water services
3  Afghanistan      AFG   Asia  At least basic drinking water services
4  Afghanistan      AFG   Asia  At least basic drinking water services
```

```
  Survey type  Year Residence Wealth quintile  Indicator value  \
0         DHS  2015     Urban              Q1         0.712094
1         DHS  2015     Urban              Q2         0.822749
2         DHS  2015     Urban              Q3         0.905302
3         DHS  2015     Urban              Q4         0.941093
4         DHS  2015     Urban              Q5         0.942910

   Unweighted population  Weighted population  Small Sample
0                18763.0          9928.450283           0.0
1                10164.0          9232.637160           0.0
2                 8160.0          9107.088417           0.0
3                 7574.0          9735.917840           0.0
4                 7362.0         10241.807390           0.0
```

As the original data was imported as strings, the next section of code sought to convert any columns containing numbers into floating point numbers. It was then further consolidated by splitting up data into rural and urban sections for both datasets, and also creating some data separated by the wealthiest (Q5) and poorest (Q1) quintiles. This was then even further split into data for each indicator.

[4]:
```
# again converting strings into floating point numbers for two additional␣
 ↪datasets
urban_paradox[['Indicator value', 'Unweighted population', 'Weighted␣
 ↪population', 'Small Sample']] = urban_paradox[['Indicator value',␣
 ↪'Unweighted population', 'Weighted population', 'Small Sample']].
 ↪apply(pandas.to_numeric)
wealth[['Indicator value', 'Unweighted population', 'Weighted population',␣
 ↪'Small Sample']] = wealth[['Indicator value', 'Unweighted population',␣
 ↪'Weighted population', 'Small Sample']].apply(pandas.to_numeric)
```

[5]:
```
#splitting up urban_paradox dataset
paradox_urban = urban_paradox['Residence'] == 'Urban' #masking dataset to find␣
 ↪all urban values
urban_all = urban_paradox[paradox_urban] # creating a new dataset with just␣
 ↪urban values

#breaking up by indicator
u_water_vals = urban_all['Indicator name'] == 'At least basic drinking water␣
 ↪services'
u_water = urban_all[u_water_vals]

u_san_vals = urban_all['Indicator name'] == 'At least basic sanitation services'
u_san = urban_all[u_san_vals]

u_birth_vals = urban_all['Indicator name'] == 'Birth Registration Rate'
u_birth = urban_all[u_birth_vals]

u_dpt3_vals = urban_all['Indicator name'] == 'DPT3 coverage'
```

```
u_dpt3 = urban_all[u_dpt3_vals]

u_hivm_vals = urban_all['Indicator name'] == 'HIV Knowledge Young Men'
u_hivm = urban_all[u_hivm_vals]

u_hivw_vals = urban_all['Indicator name'] == 'HIV Knowledge Young Women'
u_hivw = urban_all[u_hivw_vals]

u_edu_vals = urban_all['Indicator name'] == 'Primary Education Completion Rate'
u_edu = urban_all[u_edu_vals]

u_sba_vals = urban_all['Indicator name'] == 'Skilled Birth Attendants'
u_sba = urban_all[u_sba_vals]

u_stunt_vals = urban_all['Indicator name'] == 'Stunting Rate'
u_stunt = urban_all[u_stunt_vals]

u_u5mr_vals = urban_all['Indicator name'] == 'Under 5 Mortality Rate'
u_u5mr = urban_all[u_u5mr_vals]


#same as above for rural values
paradox_rural = urban_paradox['Residence'] == 'Rural'
rural_all = urban_paradox[paradox_rural]


#breaking up by indicator
r_water_vals = rural_all['Indicator name'] == 'At least basic drinking water␣
 ↪services'
r_water = rural_all[r_water_vals]

r_san_vals = rural_all['Indicator name'] == 'At least basic sanitation services'
r_san = rural_all[r_san_vals]

r_birth_vals = rural_all['Indicator name'] == 'Birth Registration Rate'
r_birth = rural_all[r_birth_vals]

r_dpt3_vals = rural_all['Indicator name'] == 'DPT3 coverage'
r_dpt3 = rural_all[r_dpt3_vals]

r_hivm_vals = rural_all['Indicator name'] == 'HIV Knowledge Young Men'
r_hivm = rural_all[r_hivm_vals]

r_hivw_vals = rural_all['Indicator name'] == 'HIV Knowledge Young Women'
r_hivw = rural_all[r_hivw_vals]

r_edu_vals = rural_all['Indicator name'] == 'Primary Education Completion Rate'
```

```
r_edu = rural_all[r_edu_vals]

r_sba_vals = rural_all['Indicator name'] == 'Skilled Birth Attendants'
r_sba = rural_all[r_sba_vals]

r_stunt_vals = rural_all['Indicator name'] == 'Stunting Rate'
r_stunt = rural_all[r_stunt_vals]

r_u5mr_vals = rural_all['Indicator name'] == 'Under 5 Mortality Rate'
r_u5mr = rural_all[r_u5mr_vals]



#splitting up wealth dataset
wealth_urban = wealth['Residence'] == 'Urban' #making an all urban dataset
u_wealth = wealth[wealth_urban]

uquin5 = u_wealth['Wealth quintile'] == 'Q5' # Quintile 5 urban dataset
uq5 = u_wealth[uquin5]

#breaking up by indicator
u5_water_vals = uq5['Indicator name'] == 'At least basic drinking water␣
↪services'
u5_water = uq5[u5_water_vals]

u5_san_vals = uq5['Indicator name'] == 'At least basic sanitation services'
u5_san = uq5[u5_san_vals]

u5_birth_vals = uq5['Indicator name'] == 'Birth Registration Rate'
u5_birth = uq5[u5_birth_vals]

u5_dpt3_vals = uq5['Indicator name'] == 'DPT3 coverage'
u5_dpt3 = uq5[u5_dpt3_vals]

u5_hivm_vals = uq5['Indicator name'] == 'HIV Knowledge Young Men'
u5_hivm = uq5[u5_hivm_vals]

u5_hivw_vals = uq5['Indicator name'] == 'HIV Knowledge Young Women'
u5_hivw = uq5[u5_hivw_vals]

u5_edu_vals = uq5['Indicator name'] == 'Primary Education Completion Rate'
u5_edu = uq5[u5_edu_vals]

u5_sba_vals = uq5['Indicator name'] == 'Skilled Birth Attendants'
u5_sba = uq5[u5_sba_vals]

u5_stunt_vals = uq5['Indicator name'] == 'Stunting Rate'
```

```
u5_stunt = uq5[u5_stunt_vals]

u5_u5mr_vals = uq5['Indicator name'] == 'Under 5 Mortality Rate'
u5_u5mr = uq5[u5_u5mr_vals]




uquin1 = u_wealth['Wealth quintile'] == 'Q1' # Quintile 1 urban dataset
uq1 = u_wealth[uquin1]

#breaking up by indicator
u1_water_vals = uq1['Indicator name'] == 'At least basic drinking water␣
 ↪services'
u1_water = uq1[u1_water_vals]

u1_san_vals = uq1['Indicator name'] == 'At least basic sanitation services'
u1_san = uq1[u1_san_vals]

u1_birth_vals = uq1['Indicator name'] == 'Birth Registration Rate'
u1_birth = uq1[u1_birth_vals]

u1_dpt3_vals = uq1['Indicator name'] == 'DPT3 coverage'
u1_dpt3 = uq1[u1_dpt3_vals]

u1_hivm_vals = uq1['Indicator name'] == 'HIV Knowledge Young Men'
u1_hivm = uq1[u1_hivm_vals]

u1_hivw_vals = uq1['Indicator name'] == 'HIV Knowledge Young Women'
u1_hivw = uq1[u1_hivw_vals]

u1_edu_vals = uq1['Indicator name'] == 'Primary Education Completion Rate'
u1_edu = uq1[u1_edu_vals]

u1_sba_vals = uq1['Indicator name'] == 'Skilled Birth Attendants'
u1_sba = uq1[u1_sba_vals]

u1_stunt_vals = uq1['Indicator name'] == 'Stunting Rate'
u1_stunt = uq1[u1_stunt_vals]

u1_u5mr_vals = uq1['Indicator name'] == 'Under 5 Mortality Rate'
u1_u5mr = uq1[u1_u5mr_vals]




wealth_rural = wealth['Residence'] == 'Rural' #making an all rural dataset
r_wealth = wealth[wealth_rural]
```

```
rquin5 = r_wealth['Wealth quintile'] == 'Q5' # Quintile 5 rural dataset
rq5 = r_wealth[rquin5]

#breaking up by indicator
r5_water_vals = rq5['Indicator name'] == 'At least basic drinking water␣
↪services'
r5_water = rq5[r5_water_vals]

r5_san_vals = rq5['Indicator name'] == 'At least basic sanitation services'
r5_san = rq5[r5_san_vals]

r5_birth_vals = rq5['Indicator name'] == 'Birth Registration Rate'
r5_birth = rq5[r5_birth_vals]

r5_dpt3_vals = rq5['Indicator name'] == 'DPT3 coverage'
r5_dpt3 = rq5[r5_dpt3_vals]

r5_hivm_vals = rq5['Indicator name'] == 'HIV Knowledge Young Men'
r5_hivm = rq5[r5_hivm_vals]

r5_hivw_vals = rq5['Indicator name'] == 'HIV Knowledge Young Women'
r5_hivw = rq5[r5_hivw_vals]

r5_edu_vals = rq5['Indicator name'] == 'Primary Education Completion Rate'
r5_edu = rq5[r5_edu_vals]

r5_sba_vals = rq5['Indicator name'] == 'Skilled Birth Attendants'
r5_sba = rq5[r5_sba_vals]

r5_stunt_vals = rq5['Indicator name'] == 'Stunting Rate'
r5_stunt = rq5[r5_stunt_vals]

r5_u5mr_vals = rq5['Indicator name'] == 'Under 5 Mortality Rate'
r5_u5mr = rq5[r5_u5mr_vals]



rquin1 = r_wealth['Wealth quintile'] == 'Q1' # Quintile 1 rural dataset
rq1 = r_wealth[rquin1]

#breaking up by indicator
r1_water_vals = rq1['Indicator name'] == 'At least basic drinking water␣
↪services'
r1_water = rq1[r1_water_vals]

r1_san_vals = rq1['Indicator name'] == 'At least basic sanitation services'
```

```
r1_san = rq1[r1_san_vals]

r1_birth_vals = rq1['Indicator name'] == 'Birth Registration Rate'
r1_birth = rq1[r1_birth_vals]

r1_dpt3_vals = rq1['Indicator name'] == 'DPT3 coverage'
r1_dpt3 = rq1[r1_dpt3_vals]

r1_hivm_vals = rq1['Indicator name'] == 'HIV Knowledge Young Men'
r1_hivm = rq1[r1_hivm_vals]

r1_hivw_vals = rq1['Indicator name'] == 'HIV Knowledge Young Women'
r1_hivw = rq1[r1_hivw_vals]

r1_edu_vals = rq1['Indicator name'] == 'Primary Education Completion Rate'
r1_edu = rq1[r1_edu_vals]

r1_sba_vals = rq1['Indicator name'] == 'Skilled Birth Attendants'
r1_sba = rq1[r1_sba_vals]

r1_stunt_vals = rq1['Indicator name'] == 'Stunting Rate'
r1_stunt = rq1[r1_stunt_vals]

r1_u5mr_vals = rq1['Indicator name'] == 'Under 5 Mortality Rate'
r1_u5mr = rq1[r1_u5mr_vals]
```

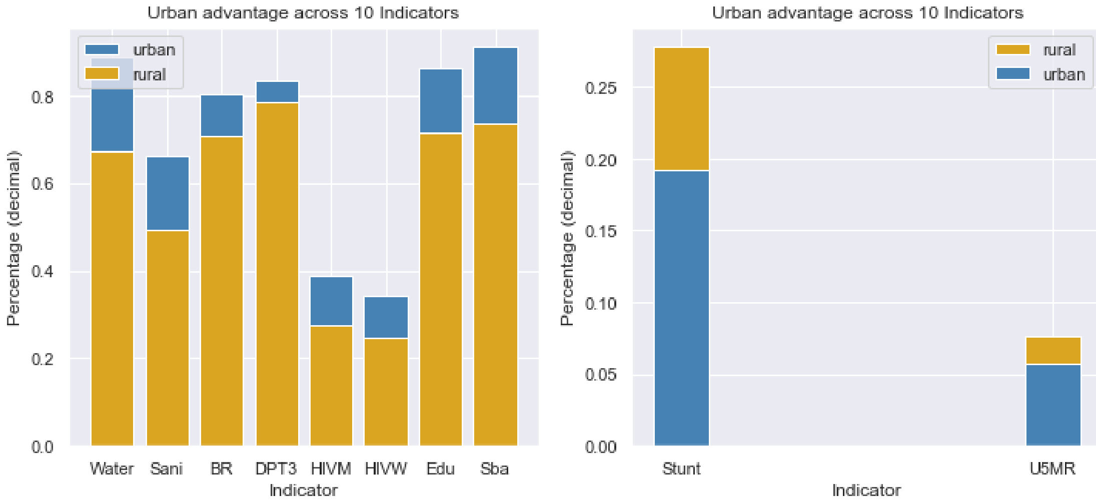
With all data sorted, it was then analyzed through plotting using Python's matplotlib library with the aid of seaborn. In the first chart, the average urban and rural indicator values for all countries were found and then plotted so as to see the general urban advantage. The ten health indicators have shortened names on the x-axis.

[6]:
```
# importing necessary libraries
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns

# initializing lists for plotting
x_axis = ['Water', 'Sani', 'BR', 'DPT3', 'HIVM', "HIVW", 'Edu', 'Sba'] # order␣
↪of indicators to be plotted
urban_average = []
rural_average = []

# making separate lists (and a separate plot) for stunting and under 5␣
↪mortality as high values are not desired in these categories
x_axis_2 = ['Stunt', 'U5MR']
urban_average_2 = []
rural_average_2 = []
```

```
# finding the mean value for all countries for each indicator to plot the␣
↪average urban advantage among all countries
ans1 = u_water['Indicator value'].mean()
urban_average.append(ans1) # appending these values to the lists to be plotted

ans2 = u_san['Indicator value'].mean()
urban_average.append(ans2)

ans3 = u_birth['Indicator value'].mean()
urban_average.append(ans3)

ans4 = u_dpt3['Indicator value'].mean()
urban_average.append(ans4)

ans5 = u_hivm['Indicator value'].mean()
urban_average.append(ans5)

ans6 = u_hivw['Indicator value'].mean()
urban_average.append(ans6)

ans7 = u_edu['Indicator value'].mean()
urban_average.append(ans7)

ans8 = u_sba['Indicator value'].mean()
urban_average.append(ans8)

ans9 = u_stunt['Indicator value'].mean()
urban_average_2.append(ans9)

ans10 = u_u5mr['Indicator value'].mean() / 1000 # dividing by 1000 as previous␣
↪unit was deaths per 1000 live births (making it a percentage)
urban_average_2.append(ans10)

# same process as above but for all rural values appended to rural list
ans11 = r_water['Indicator value'].mean()
rural_average.append(ans11)

ans12 = r_san['Indicator value'].mean()
rural_average.append(ans12)

ans13 = r_birth['Indicator value'].mean()
rural_average.append(ans13)

ans14 = r_dpt3['Indicator value'].mean()
rural_average.append(ans14)

ans15 = r_hivm['Indicator value'].mean()
```

```
rural_average.append(ans15)

ans16 = r_hivw['Indicator value'].mean()
rural_average.append(ans16)

ans17 = r_edu['Indicator value'].mean()
rural_average.append(ans17)

ans18 = r_sba['Indicator value'].mean()
rural_average.append(ans18)

ans19 = r_stunt['Indicator value'].mean()
rural_average_2.append(ans19)

ans20 = r_u5mr['Indicator value'].mean() /1000
rural_average_2.append(ans20)

# plotting average urban and rural values for all indicators to show that urban␣
 ↪advantage exists among all indicators when not considering wealth
plt.figure(figsize= (12,5)) #setting figure size
sns.set() #using seaborn library for better visualization

plt.subplot(1,2,1) # using subplots so that mortality and stunting can be␣
 ↪separated to show that urban has lower values for these which is desired for␣
 ↪those indicators
plt.bar(x_axis, urban_average, label = 'urban', align = 'center', color =␣
 ↪'steelblue')
plt.bar(x_axis, rural_average, label = 'rural', align = 'center', color =␣
 ↪'goldenrod')
plt.legend()
plt.title('Urban advantage across 10 Indicators')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(1,2,2)
plt.bar(x_axis_2, rural_average_2, label = 'rural', align = 'center', width = 0.
 ↪15, color = 'goldenrod') #setting width to be similar to first subplot for␣
 ↪better visualization
plt.bar(x_axis_2, urban_average_2, label = 'urban', align = 'center', width = 0.
 ↪15, color = 'steelblue')
plt.legend()
plt.title('Urban advantage across 10 Indicators') #same titles
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
```

[6]:
```
Text(0, 0.5, 'Percentage (decimal)')
```



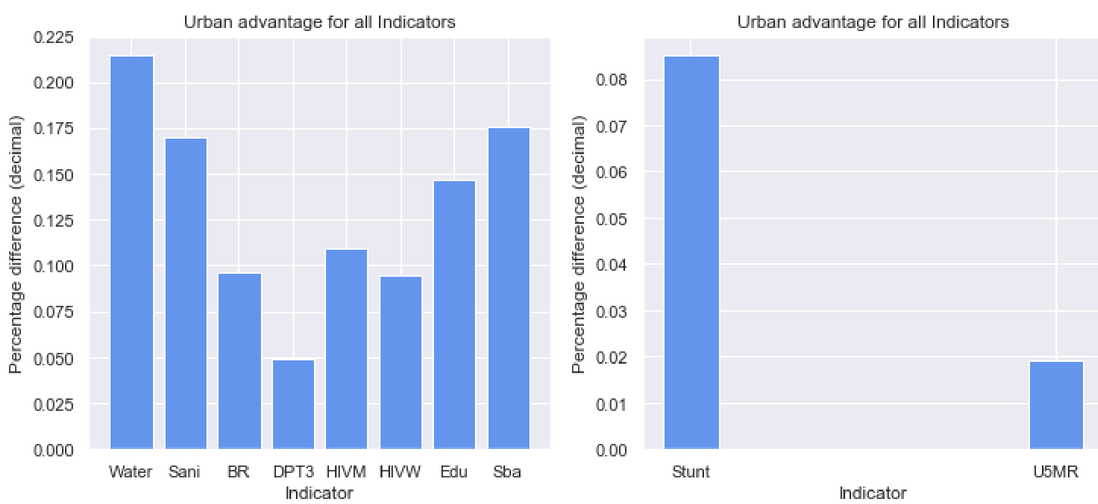
As can be seen, there is a clear urban advantage across all indicators while not scanning for wealth or other factors and just using average values. The rural bar appears to be higher for both stunting and under 2 child mortality, which is as expected as lower rates of these are desired to indicate better health. All other indicators show the blue urban bar rising above the rural example, showing the advantage.

The advantage is not equal across all indicators however, as can be seen on the plot. For some indicators the advantage was greater than for others. To see the average advantage for all indicators, the difference between the two columns was plotted to find the average advantage. For stunting and mortality, the urban value was subtracted from the rural value, while for all other indicators the rural was subtracted from the urban.

[7]:
```
differences = [] # initializing list of differences for plotting
differences_2 = [] # separate list for stunt rate and mortality as these have␣
 ↪urban subtracted from rural

water_diff = ans1-ans11 # finding difference
differences.append(water_diff) # adding to list

sani_diff = ans2-ans12
differences.append(sani_diff)

br_diff = ans3-ans13
differences.append(br_diff)

dpt3_diff = ans4-ans14
differences.append(dpt3_diff)

hivm_diff = ans5-ans15
differences.append(hivm_diff)

hivw_diff = ans6-ans16
```

```
differences.append(hivw_diff)

edu_diff = ans7-ans17
differences.append(edu_diff)

sba_diff = ans8-ans18
differences.append(sba_diff)

stunt_diff = ans19-ans9
differences_2.append(stunt_diff)

mort_diff = ans20-ans10
differences_2.append(mort_diff)

#plotting differences
plt.figure(figsize= (12,5))

plt.subplot(1,2,1)
plt.bar(x_axis, differences, color = 'cornflowerblue')
plt.title('Urban advantage for all Indicators')
plt.xlabel('Indicator')
plt.ylabel('Percentage difference (decimal)')

plt.subplot(1,2,2)
plt.bar(x_axis_2, differences_2, width = 0.15, color = 'cornflowerblue')
plt.title('Urban advantage for all Indicators')
plt.xlabel('Indicator')
plt.ylabel('Percentage difference (decimal)')
```

[7]: Text(0, 0.5, 'Percentage difference (decimal)')



The smallest advatage seems to appear with DPT3 coverage, while the largest is with access to basic drinking water services. UNICEF claims that "much of the urban advantage disappears if we control for wealth" ("Advantage or Paradox?", 2018). The second dataset included indicator values separated by wealth quintile. For the next section, the poorest (Q1) quintile was plotted with the wealthiest (Q5) quintile for both urban and rural areas to see the impact of wealth on urban advantage. The means for each quitile for each indicator amongst all countries was again found and then added to a list before it was then plotted.

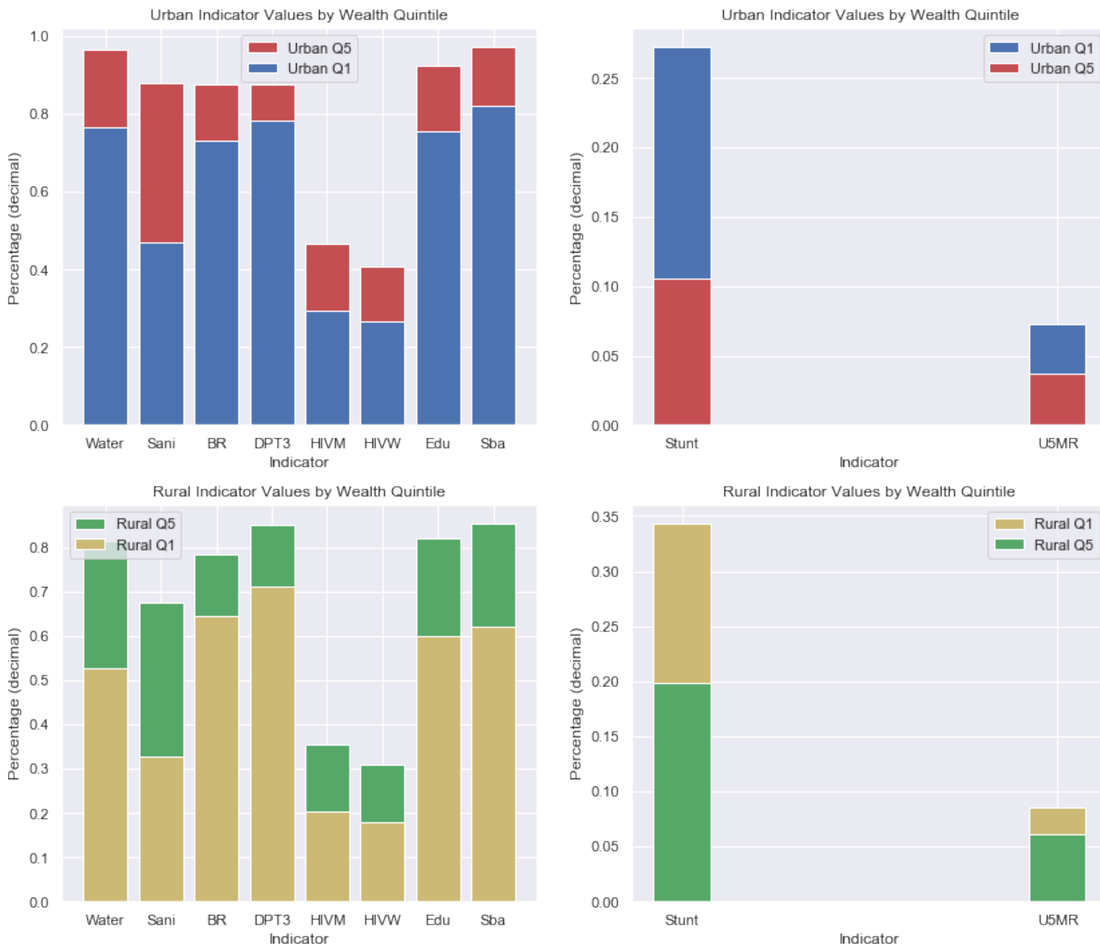
[8]:
```
# initializing lists for plotting
urban_average_Q5 = []
rural_average_Q5 = []
urban_average_Q1 = []
rural_average_Q1 = []

# making separate lists for stunting and under 5 mortality
urban_average_2_Q5 = []
rural_average_2_Q5 = []
urban_average_2_Q1 = []
rural_average_2_Q1 = []

# finding the mean value for each quitile for each indicator amongst all␣
↪countries
uq5w = u5_water['Indicator value'].mean()
urban_average_Q5.append(uq5w) # appending these values to the lists
uq1w = u1_water['Indicator value'].mean()
urban_average_Q1.append(uq1w)

uq5san = u5_san['Indicator value'].mean()
urban_average_Q5.append(uq5san)
uq1san = u1_san['Indicator value'].mean()
urban_average_Q1.append(uq1san)

uq5birth = u5_birth['Indicator value'].mean()
urban_average_Q5.append(uq5birth)
uq1birth = u1_birth['Indicator value'].mean()
urban_average_Q1.append(uq1birth)

uq5dpt3 = u5_dpt3['Indicator value'].mean()
urban_average_Q5.append(uq5dpt3)
uq1dpt3 = u1_dpt3['Indicator value'].mean()
urban_average_Q1.append(uq1dpt3)

uq5hivm = u5_hivm['Indicator value'].mean()
urban_average_Q5.append(uq5hivm)
uq1hivm = u1_hivm['Indicator value'].mean()
urban_average_Q1.append(uq1hivm)

uq5hivw = u5_hivw['Indicator value'].mean()
```

```
urban_average_Q5.append(uq5hivw)
uq1hivw = u1_hivw['Indicator value'].mean()
urban_average_Q1.append(uq1hivw)

uq5edu = u5_edu['Indicator value'].mean()
urban_average_Q5.append(uq5edu)
uq1edu = u1_edu['Indicator value'].mean()
urban_average_Q1.append(uq1edu)

uq5sba = u5_sba['Indicator value'].mean()
urban_average_Q5.append(uq5sba)
uq1sba = u1_sba['Indicator value'].mean()
urban_average_Q1.append(uq1sba)

uq5stunt = u5_stunt['Indicator value'].mean()
urban_average_2_Q5.append(uq5stunt)
uq1stunt = u1_stunt['Indicator value'].mean()
urban_average_2_Q1.append(uq1stunt)

uq5u5mr = u5_u5mr['Indicator value'].mean() / 1000
urban_average_2_Q5.append(uq5u5mr)
uq1u5mr = u1_u5mr['Indicator value'].mean() / 1000
urban_average_2_Q1.append(uq1u5mr)

# same process as above but for all rural values
rq5w = r5_water['Indicator value'].mean()
rural_average_Q5.append(rq5w)
rq1w = r1_water['Indicator value'].mean()
rural_average_Q1.append(rq1w)

rq5san = r5_san['Indicator value'].mean()
rural_average_Q5.append(rq5san)
rq1san = r1_san['Indicator value'].mean()
rural_average_Q1.append(rq1san)

rq5birth = r5_birth['Indicator value'].mean()
rural_average_Q5.append(rq5birth)
rq1birth = r1_birth['Indicator value'].mean()
rural_average_Q1.append(rq1birth)

rq5dpt3 = r5_dpt3['Indicator value'].mean()
rural_average_Q5.append(rq5dpt3)
rq1dpt3 = r1_dpt3['Indicator value'].mean()
rural_average_Q1.append(rq1dpt3)

rq5hivm = r5_hivm['Indicator value'].mean()
rural_average_Q5.append(rq5hivm)
```

```
rq1hivm = r1_hivm['Indicator value'].mean()
rural_average_Q1.append(rq1hivm)

rq5hivw = r5_hivw['Indicator value'].mean()
rural_average_Q5.append(rq5hivw)
rq1hivw = r1_hivw['Indicator value'].mean()
rural_average_Q1.append(rq1hivw)

rq5edu = r5_edu['Indicator value'].mean()
rural_average_Q5.append(rq5edu)
rq1edu = r1_edu['Indicator value'].mean()
rural_average_Q1.append(rq1edu)

rq5sba = r5_sba['Indicator value'].mean()
rural_average_Q5.append(rq5sba)
rq1sba = r1_sba['Indicator value'].mean()
rural_average_Q1.append(rq1sba)

rq5stunt = r5_stunt['Indicator value'].mean()
rural_average_2_Q5.append(rq5stunt)
rq1stunt = r1_stunt['Indicator value'].mean()
rural_average_2_Q1.append(rq1stunt)

rq5u5mr = r5_u5mr['Indicator value'].mean() / 1000
rural_average_2_Q5.append(rq5u5mr)
rq1u5mr = r1_u5mr['Indicator value'].mean() / 1000
rural_average_2_Q1.append(rq1u5mr)

# plotting average urban and rural values for all indicators for both␣
 ↪wealthiest (Q5) and poorest (Q1) quiltile
plt.figure(figsize= (14,12))

plt.subplot(2,2,1)
plt.bar(x_axis, urban_average_Q5, label = 'Urban Q5', align = 'center', color =␣
 ↪'r')
plt.bar(x_axis, urban_average_Q1, label = 'Urban Q1', align = 'center', color =␣
 ↪'b')
plt.legend()
plt.title('Urban Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,2)
plt.bar(x_axis_2, urban_average_2_Q1, label = 'Urban Q1', align = 'center',␣
 ↪color = 'b', width = 0.15)
plt.bar(x_axis_2, urban_average_2_Q5, label = 'Urban Q5', align = 'center',␣
 ↪color = 'r', width = 0.15)
```

```
plt.legend()
plt.title('Urban Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,3)
plt.bar(x_axis, rural_average_Q5, label = 'Rural Q5', align = 'center', color =␣
 ↪'g')
plt.bar(x_axis, rural_average_Q1, label = 'Rural Q1', align = 'center', color =␣
 ↪'y')
plt.legend()
plt.title('Rural Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,4)
plt.bar(x_axis_2, rural_average_2_Q1, label = 'Rural Q1', align = 'center',␣
 ↪color = 'y', width = 0.15)
plt.bar(x_axis_2, rural_average_2_Q5, label = 'Rural Q5', align = 'center',␣
 ↪color = 'g', width = 0.15)
plt.legend()
plt.title('Rural Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
```

[8]:
```
Text(0, 0.5, 'Percentage (decimal)')
```



Here it can be seen that wealth does make quite the impact on the indicator values, as Q5 always shows a better average than Q1. From this it is difficult to see whether urban or rural wealth disparities are greater, and so the average difference between the Q5 and Q1 values was calculated and printed below to see which group, on average, experienced greater wealth disparity.

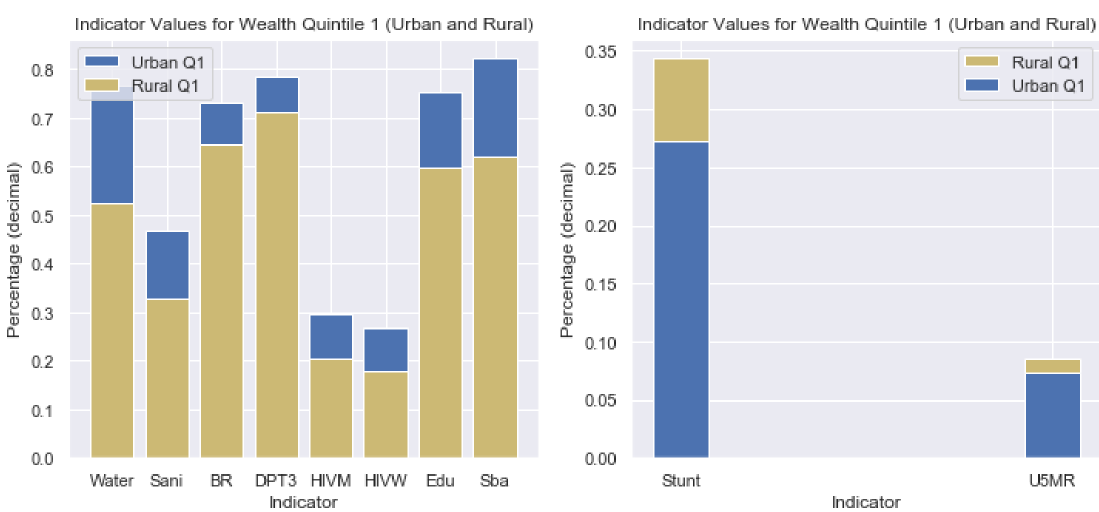
[52]:
```
print('The average urban disparity across all indicators is', (((uq5w - uq1w) *100)+((uq5san - uq1san) * 100)+((uq5birth - uq1birth) * 100)+((uq5dpt3 - uq1dpt3) * 100)+((uq5hivm - uq1hivm) * 100)+((uq5hivw - uq1hivw) *100)+((uq5edu - uq1edu) * 100)+((uq5sba - uq1sba) * 100)+((uq1stunt - uq5stunt) * 100)+( (uq1u5mr - uq5u5mr) * 100))/10, 'percent.')
print('The average rural disparity across all indicators is', (((rq5w - rq1w) *100)+((rq5san - rq1san) * 100)+((rq5birth - rq1birth) * 100)+((rq5dpt3 - rq1dpt3) * 100)+((rq5hivm - rq1hivm) * 100)+((rq5hivw - rq1hivw) *100)+((rq5edu - rq1edu) * 100)+((rq5sba - rq1sba) * 100)+((rq1stunt - rq5stunt) * 100)+( (rq1u5mr - rq5u5mr) * 100))/10, 'percent.')
```

```
The average urban disparity across all indicators is 16.792352587975834 percent.
The average rural disparity across all indicators is 18.220969514531934 percent.
```

It is seen that the average wealth disparity between Q1 and Q5 is similar in both urban and rural areas, although this data shows the average rural disparity to be approximately two percent higher. As UNICEF found that the poorest urban children in some regions were actually worse of than their rural counterparts, the next stage sought to compare urban and rural Q1 indicator values, using the same data from the previous chart simply placing Q1 data together ("Advantage or Paradox?", 2018).

[10]:
```
plt.figure(figsize= (12,5))

plt.subplot(1,2,1) # plotting Q1 values for both urban and rural areas averaged␣
↪across countries
plt.bar(x_axis, urban_average_Q1, label = 'Urban Q1', align = 'center', color =␣
↪'b')
plt.bar(x_axis, rural_average_Q1, label = 'Rural Q1', align = 'center', color =␣
↪'y')
plt.title('Indicator Values for Wealth Quintile 1 (Urban and Rural)')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
plt.legend()

plt.subplot(1,2,2)
plt.bar(x_axis_2, rural_average_2_Q1, label = 'Rural Q1', align = 'center',␣
↪color = 'y', width = 0.15)
plt.bar(x_axis_2, urban_average_2_Q1, label = 'Urban Q1', align = 'center',␣
↪color = 'b', width = 0.15)
plt.title('Indicator Values for Wealth Quintile 1 (Urban and Rural)')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
plt.legend()
```

[10]:
```
<matplotlib.legend.Legend at 0x215c0b28e80>
```



From this, it appears that there is still a slight urban advantage across all indicators, which would seem to oppose the findings of UNICEF. To look more closely and see where the urban paradox might occur, urban and rural Q1 values were plotted for each country in charts separated by indicator. This employed the use again of the matplotlib library. Where the rural bar passes the urban, or, in the case of stunting and under 5 mortality, where the urban bar passes the rural, there is urban paradox. Some charts had fewer countries on the x-axis, due to lack of data. The number of countries for which there was an urban paradox, or where the rural indicator value was better than the urban, was also calculated and printed. It was also turned into a percentage to put numbers into better perspective, and the country regions where there was paradox were found and represented in a pie chart to see if there was a specific region with more paradox.

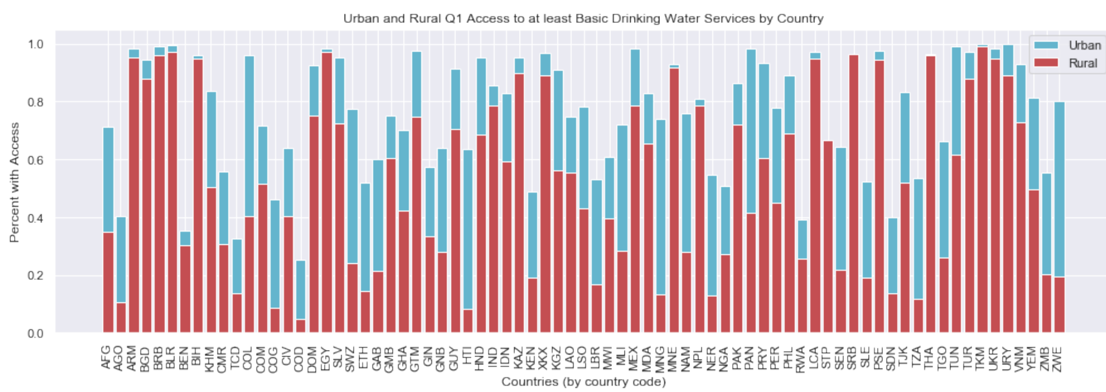
[11]:
```
import numpy as np #importing numpy library to help with function

def urban_paradox(urban_array, rural_array): # function to find number of␣
↪countries where rural indicator value is greater than urban indicator value
    count = 0 # initializing count
    x = np.array(rural_array) > np.array(urban_array) # finding places where␣
↪rural value is higher than urban
    for i in x:
        if i == True:
            count += 1 # adding the number of countries where condition is true␣
↪to the count
    return(count) # returning the count
```

[12]:
```
def find_region(array): # function to create list of countries with paradox in␣
↪each region so that a pie chart can be made
    Africa = 0 #initializing counts
    Asia = 0
    Latin_America_Caribbean = 0
    Europe = 0

    for i in array: # adding to count
        if i == 'Africa':
            Africa += 1
        if i == 'Asia':
            Asia += 1
        if i == 'Latin America and the Caribbean':
            Latin_America_Caribbean += 1
        if i == 'Europe':
            Europe += 1
    return[Africa, Asia, Latin_America_Caribbean, Europe] #giving back list
```

[13]:
```
plt.figure(figsize= (17,5)) # using a larger width for all figures due to large␣
↪amount of countries

plt.bar(u1_water['Iso3code'], u1_water['Indicator value'], label = 'Urban',␣
↪color = 'c') # plotting indicator value against country code
```

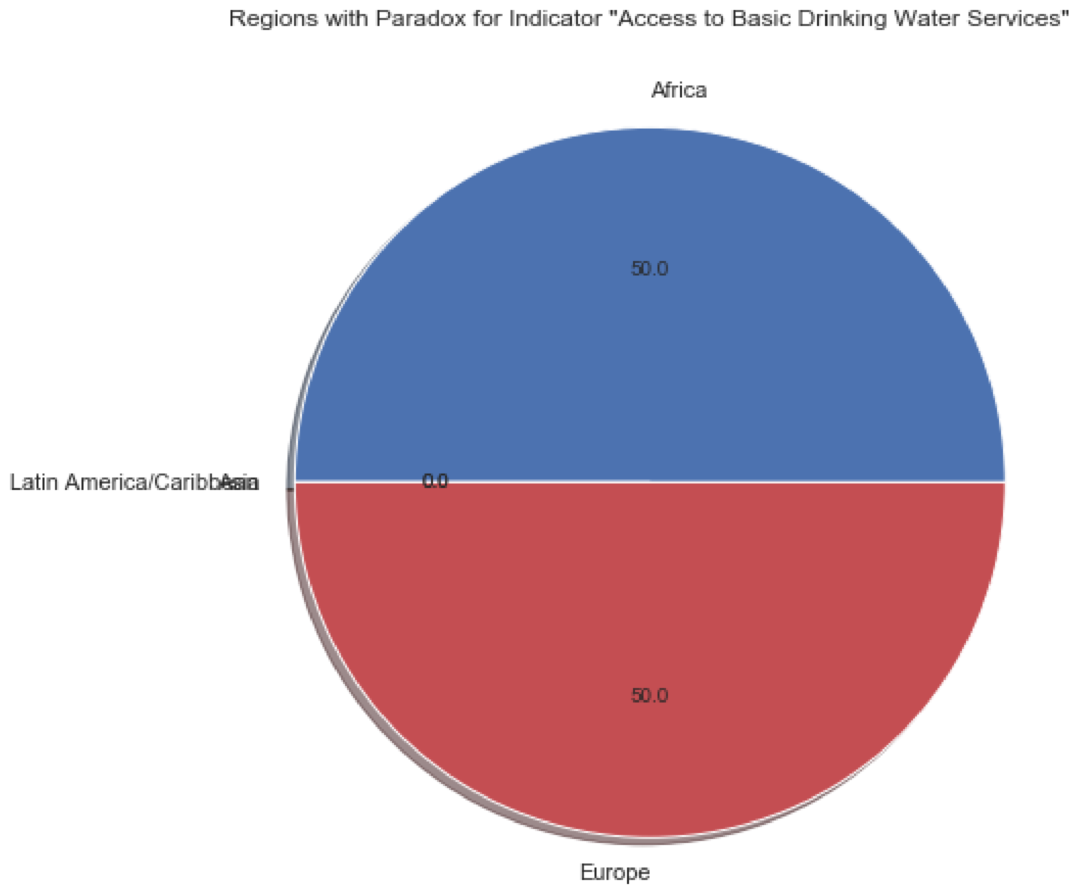
```
plt.bar(r1_water['Iso3code'], r1_water['Indicator value'], label = 'Rural',␣
 ↪color = 'r')
plt.xticks(rotation=90) # rotating x tick labels as they would not fit␣
 ↪otherwise
plt.legend()
plt.title('Urban and Rural Q1 Access to at least Basic Drinking Water Services␣
 ↪by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with Access')
# much the same process for all other indicators

print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_water['Indicator value'], r1_water['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox(u1_water['Indicator value'], r1_water['Indicator value'])/
 ↪len(u1_water['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 2
The percentage of countries experiencing an urban paradox is 2.666666666666667
percent.
```



[14]:
```
x = np.array(r1_water['Indicator value']) > np.array(u1_water['Indicator␣
 ↪value']) # finding countries where there is paradox
# masking
region = u1_water['Region']
nations_water = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8)) # plotting using pie chart to see what regions have␣
 ↪paradox
plt.pie(nations_water, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Access to Basic Drinking Water␣
 ↪Services"')
```

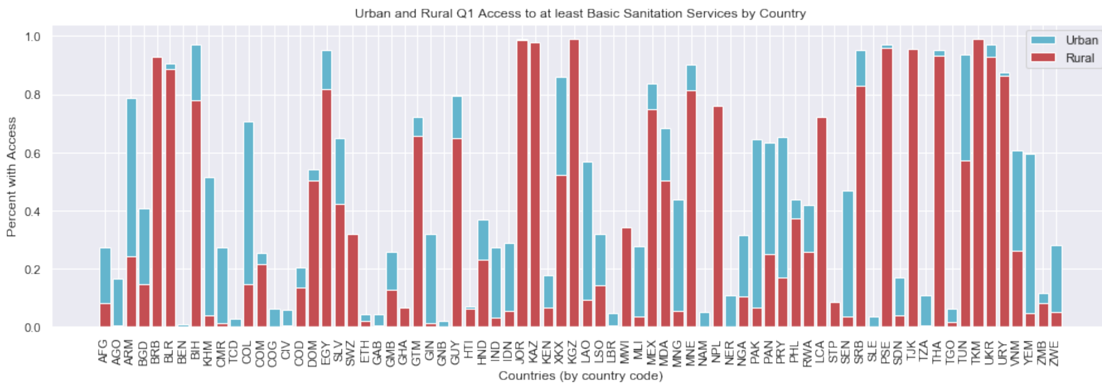
[14]: Text(0.5, 1.0, 'Regions with Paradox for Indicator "Access to Basic Drinking Water Services"')



[15]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_san['Iso3code'], u1_san['Indicator value'], label = 'Urban', color =␣
 ↪'c')
plt.bar(r1_san['Iso3code'], r1_san['Indicator value'], label = 'Rural', color =␣
 ↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Access to at least Basic Sanitation Services by␣
 ↪Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with Access')

print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_san['Indicator value'], r1_san['Indicator value']))
```

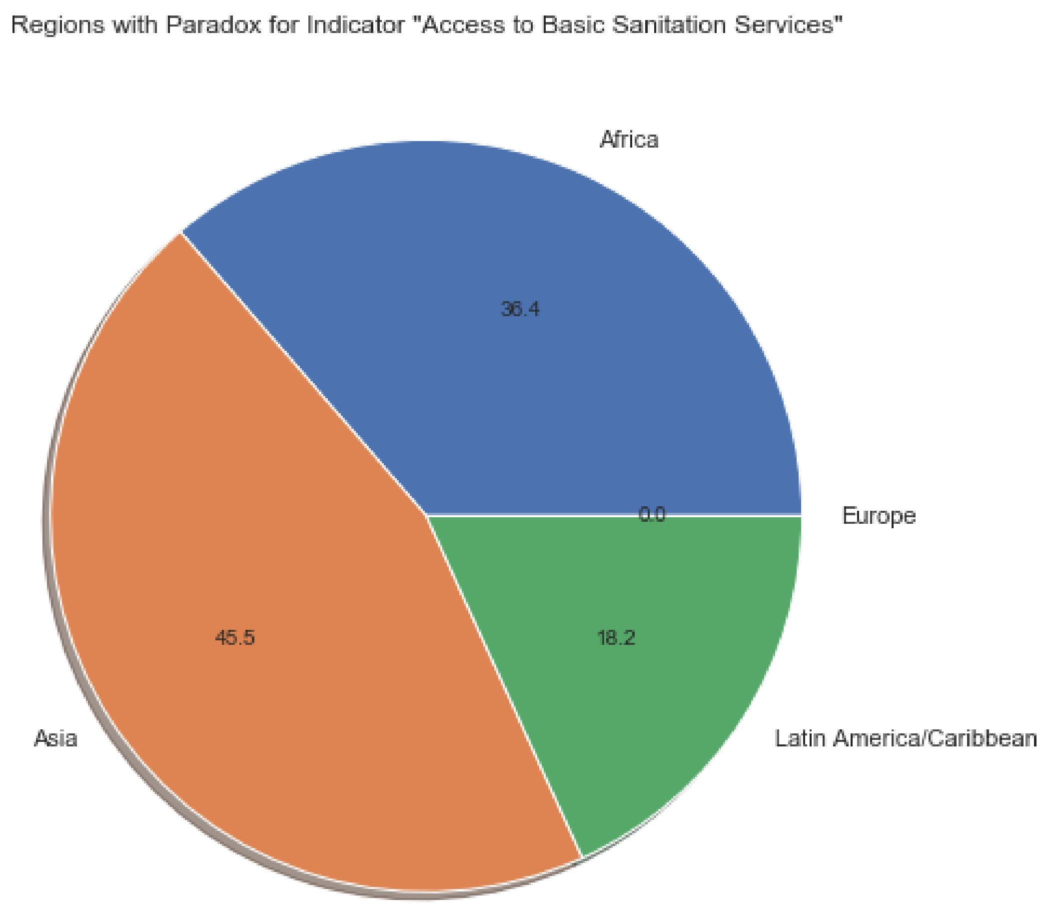
```
print('The percentage of countries experiencing an urban paradox is',␣
↪(urban_paradox(u1_san['Indicator value'], r1_san['Indicator value'])/
↪len(u1_san['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 11
The percentage of countries experiencing an urban paradox is 14.864864864864865
percent.
```



[36]:
```
# similar process to last pie chart
x = np.array(r1_san['Indicator value']) > np.array(u1_san['Indicator value'])
region = u1_san['Region']

nations_san = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_san, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Access to Basic Sanitation␣
↪Services"')
```

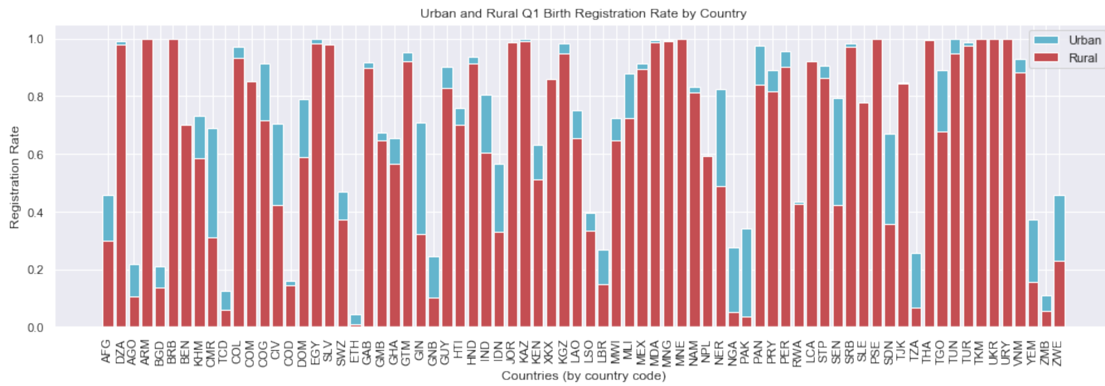
[36]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Access to Basic Sanitation
Services"')
```



[17]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_birth['Iso3code'], u1_birth['Indicator value'], label = 'Urban',␣
 ↪color = 'c')
plt.bar(r1_birth['Iso3code'], r1_birth['Indicator value'], label = 'Rural',␣
 ↪color = 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Birth Registration Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Registration Rate')

print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_birth['Indicator value'], r1_birth['Indicator value']))
```

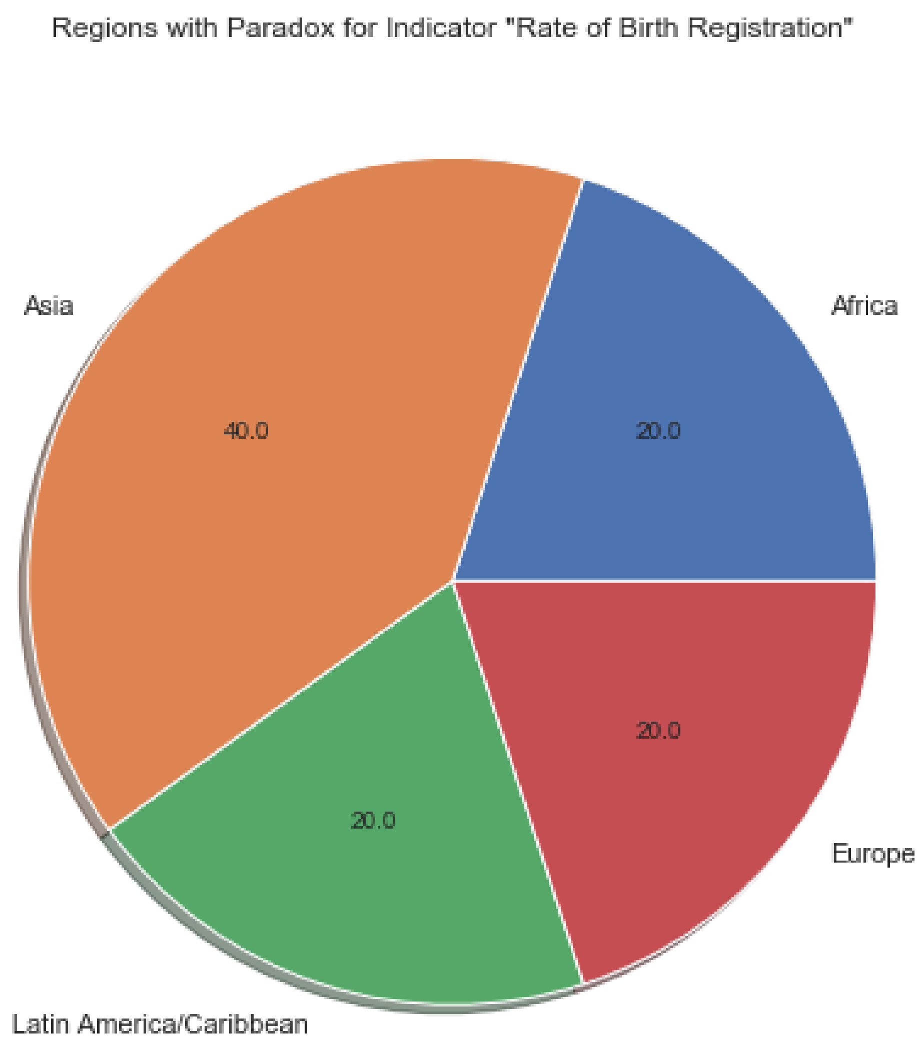
```
print('The percentage of countries experiencing an urban paradox is',␣
↪(urban_paradox(u1_birth['Indicator value'], r1_birth['Indicator value'])/
↪len(u1_birth['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 15
The percentage of countries experiencing an urban paradox is 20.27027027027027
percent.
```



[37]:
```
x = np.array(r1_birth['Indicator value']) > np.array(u1_birth['Indicator␣
↪value'])
region = u1_birth['Region']

nations_birth = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_birth, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Rate of Birth Registration"')
```

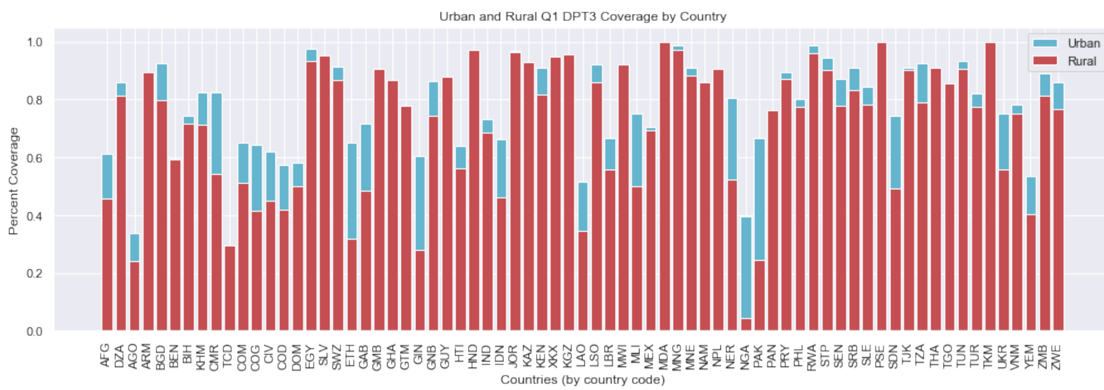
[37]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Rate of Birth
Registration"')
```



[19]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_dpt3['Iso3code'], u1_dpt3['Indicator value'], label = 'Urban', color␣
↪= 'c')
plt.bar(r1_dpt3['Iso3code'], r1_dpt3['Indicator value'], label = 'Rural', color␣
↪= 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 DPT3 Coverage by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent Coverage')
```

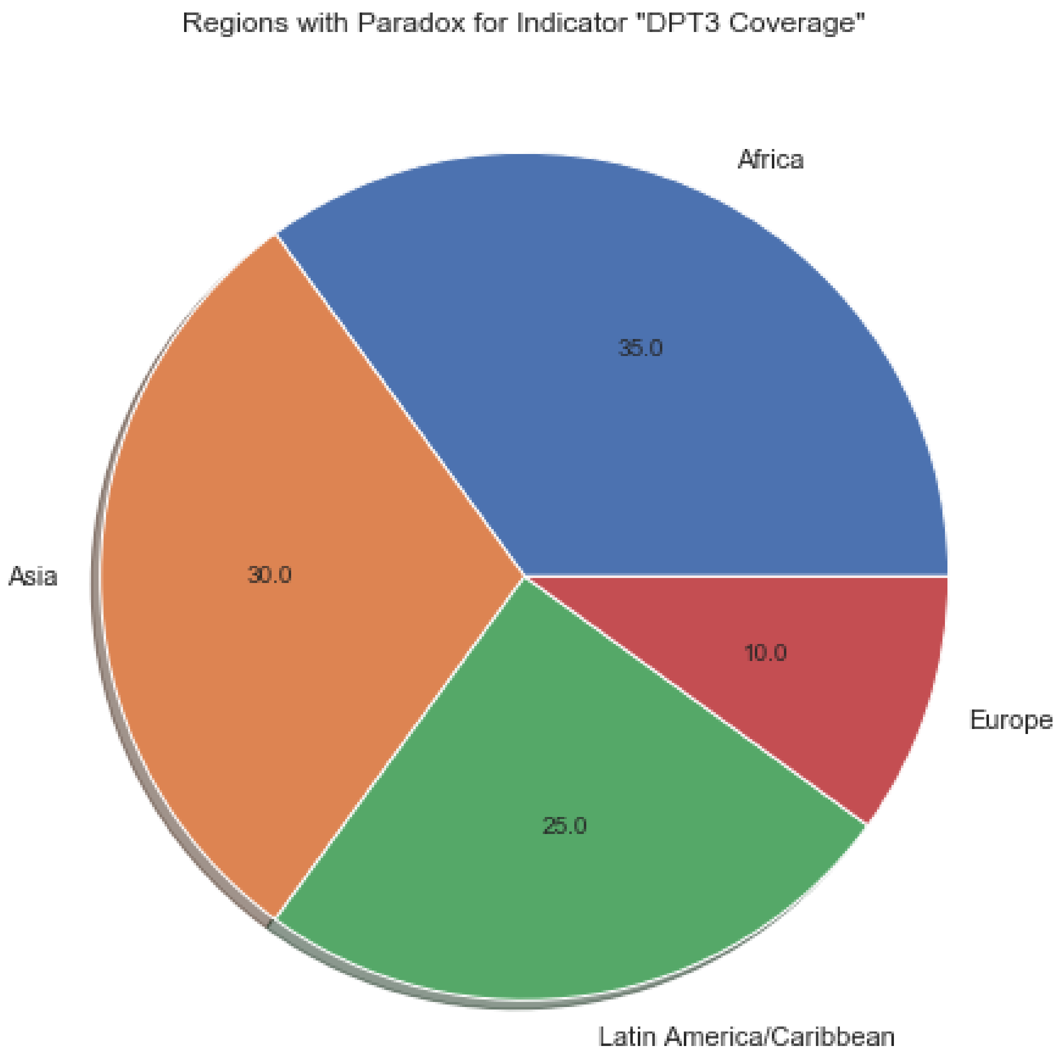
```
print('The number of countries with an urban paradox is',␣
↪urban_paradox(u1_dpt3['Indicator value'], r1_dpt3['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
↪(urban_paradox(u1_dpt3['Indicator value'], r1_dpt3['Indicator value'])/
↪len(u1_dpt3['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 20
The percentage of countries experiencing an urban paradox is 28.169014084507044
percent.
```



[38]:
```
x = np.array(r1_dpt3['Indicator value']) > np.array(u1_dpt3['Indicator value'])
region = u1_dpt3['Region']

nations_dpt3 = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_dpt3, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "DPT3 Coverage"')
```

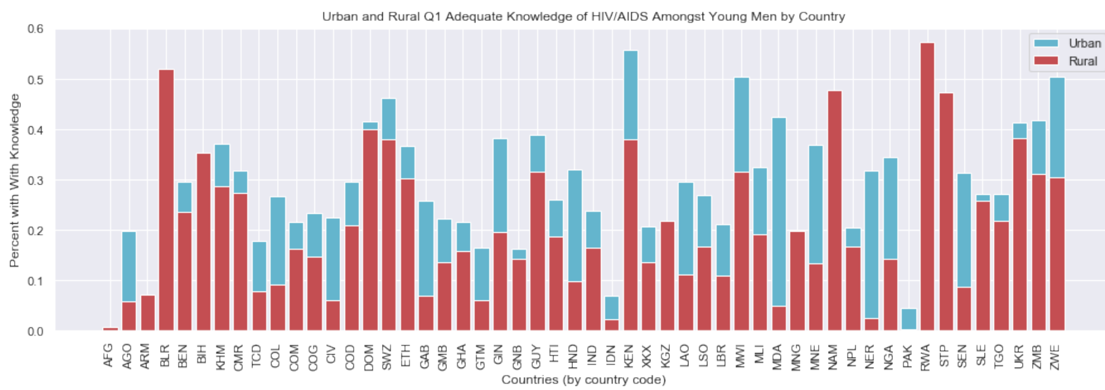
[38]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "DPT3 Coverage"')
```



[21]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_hivm['Iso3code'], u1_hivm['Indicator value'], label = 'Urban', color␣
 ↪= 'c')
plt.bar(r1_hivm['Iso3code'], r1_hivm['Indicator value'], label = 'Rural', color␣
 ↪= 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young Men␣
 ↪by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with With Knowledge')
```

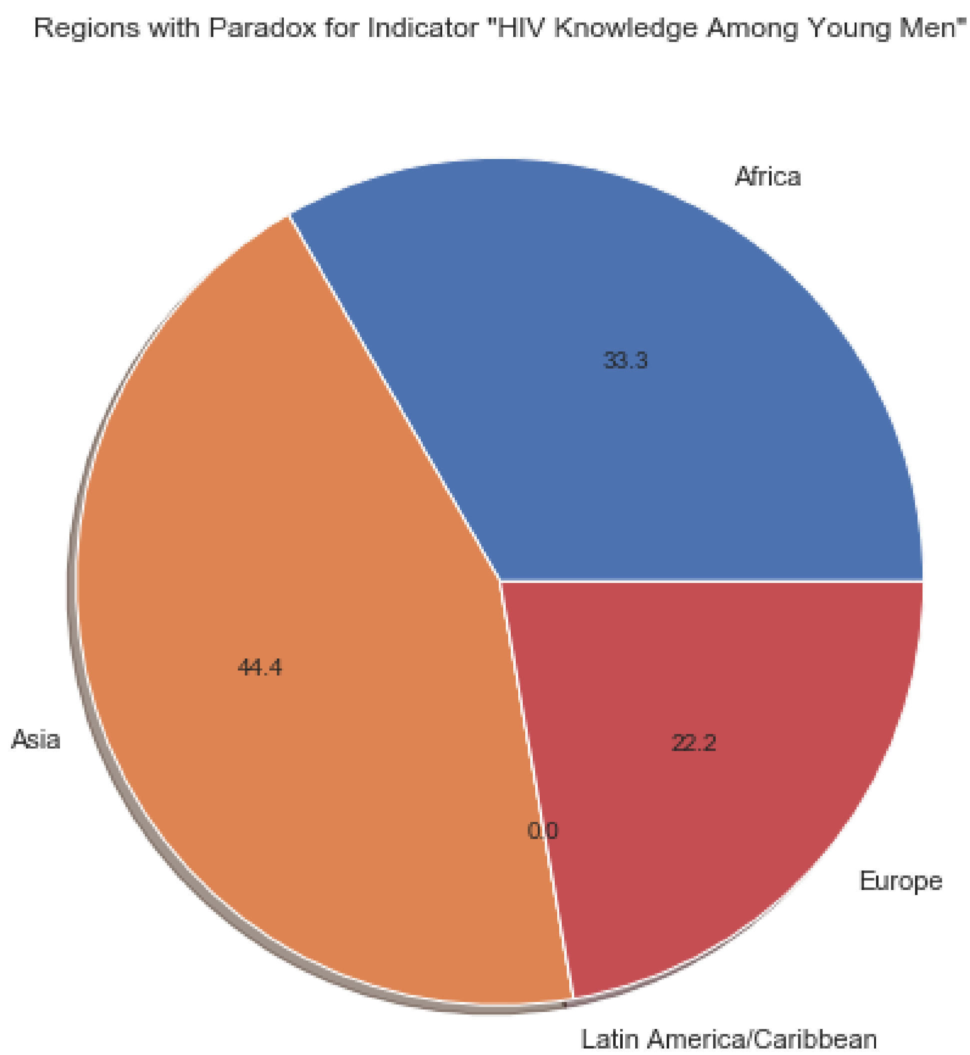
```
print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_hivm['Indicator value'], r1_hivm['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox(u1_hivm['Indicator value'], r1_hivm['Indicator value'])/
 ↪len(u1_hivm['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 9
The percentage of countries experiencing an urban paradox is 17.307692307692307
percent.
```



[39]:
```
x = np.array(r1_hivm['Indicator value']) > np.array(u1_hivm['Indicator value'])
region = u1_hivm['Region']

nations_hivm = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_hivm, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "HIV Knowledge Among Young Men"')
```

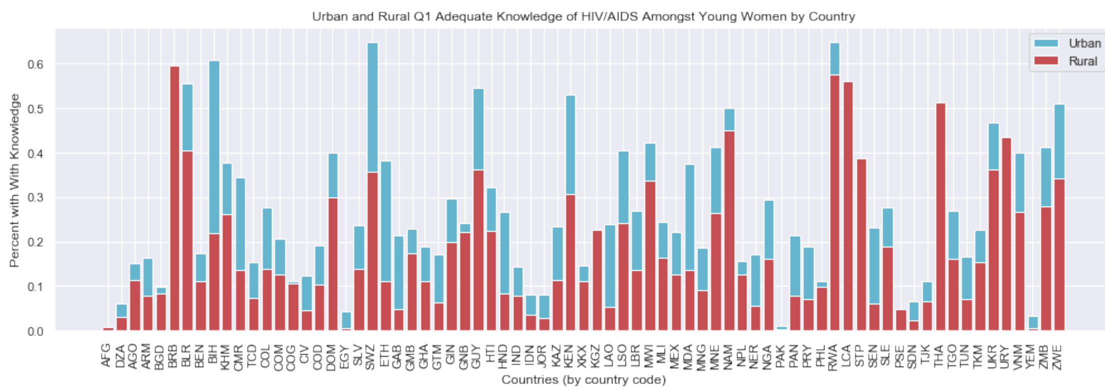
[39]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "HIV Knowledge Among Young
Men"')
```



[23]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_hivw['Iso3code'], u1_hivw['Indicator value'], label = 'Urban', color␣
 ↪= 'c')
plt.bar(r1_hivw['Iso3code'], r1_hivw['Indicator value'], label = 'Rural', color␣
 ↪= 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young␣
 ↪Women by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with With Knowledge')
```

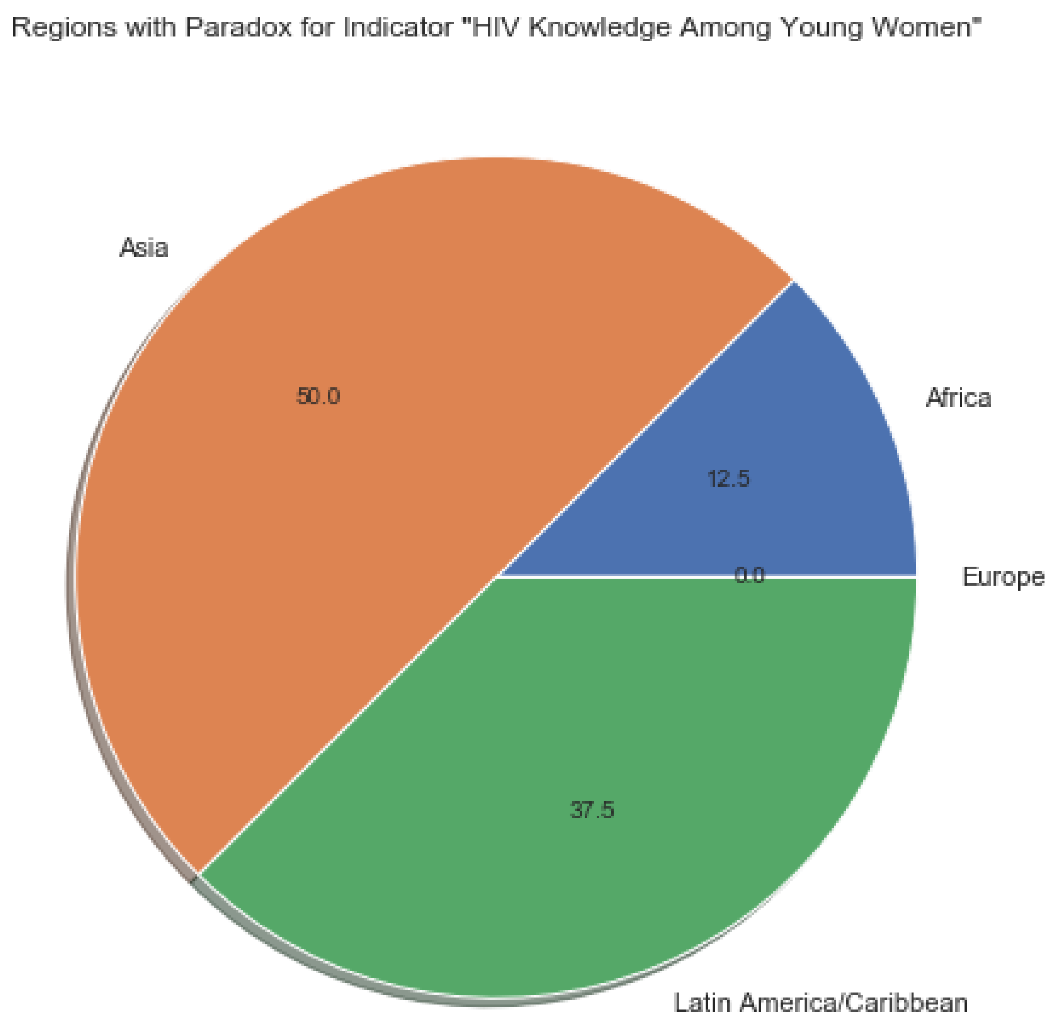
```
print('The number of countries with an urban paradox is',␣
    ↪urban_paradox(u1_hivw['Indicator value'], r1_hivw['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
    ↪(urban_paradox(u1_hivw['Indicator value'], r1_hivw['Indicator value'])/
    ↪len(u1_hivw['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 8
The percentage of countries experiencing an urban paradox is 10.95890410958904
percent.
```



[40]:
```
x = np.array(r1_hivw['Indicator value']) > np.array(u1_hivw['Indicator value'])
region = u1_hivw['Region']

nations_hivw = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_hivw, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "HIV Knowledge Among Young␣
    ↪Women"')
```

[40]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "HIV Knowledge Among Young
Women"')
```

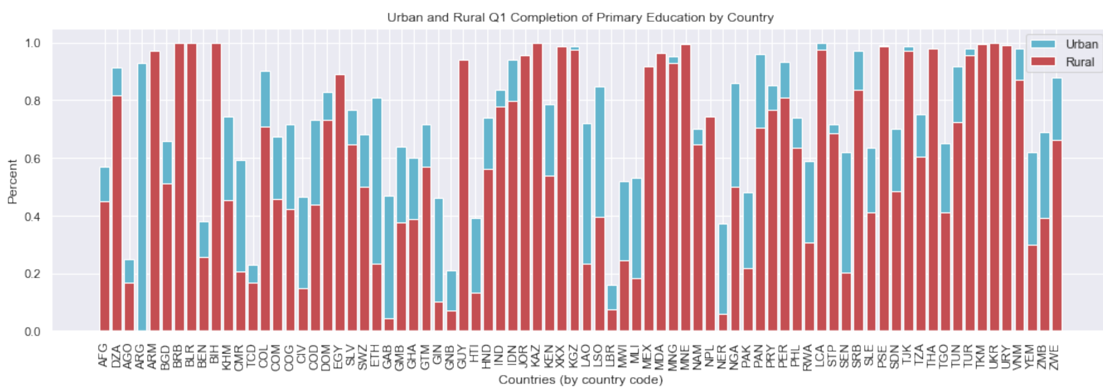


[25]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_edu['Iso3code'], u1_edu['Indicator value'], label = 'Urban', color =␣
 ↪'c')
plt.bar(r1_edu['Iso3code'], r1_edu['Indicator value'], label = 'Rural', color =␣
 ↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Completion of Primary Education by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent')
```

```
print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_edu['Indicator value'], r1_edu['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox(u1_edu['Indicator value'], r1_edu['Indicator value'])/
 ↪len(u1_edu['Indicator value']))*100, 'percent.')
```

```
C:\Users\Annalise\Anaconda3\lib\site-packages\ipykernel_launcher.py:5:
RuntimeWarning: invalid value encountered in greater
  """
```

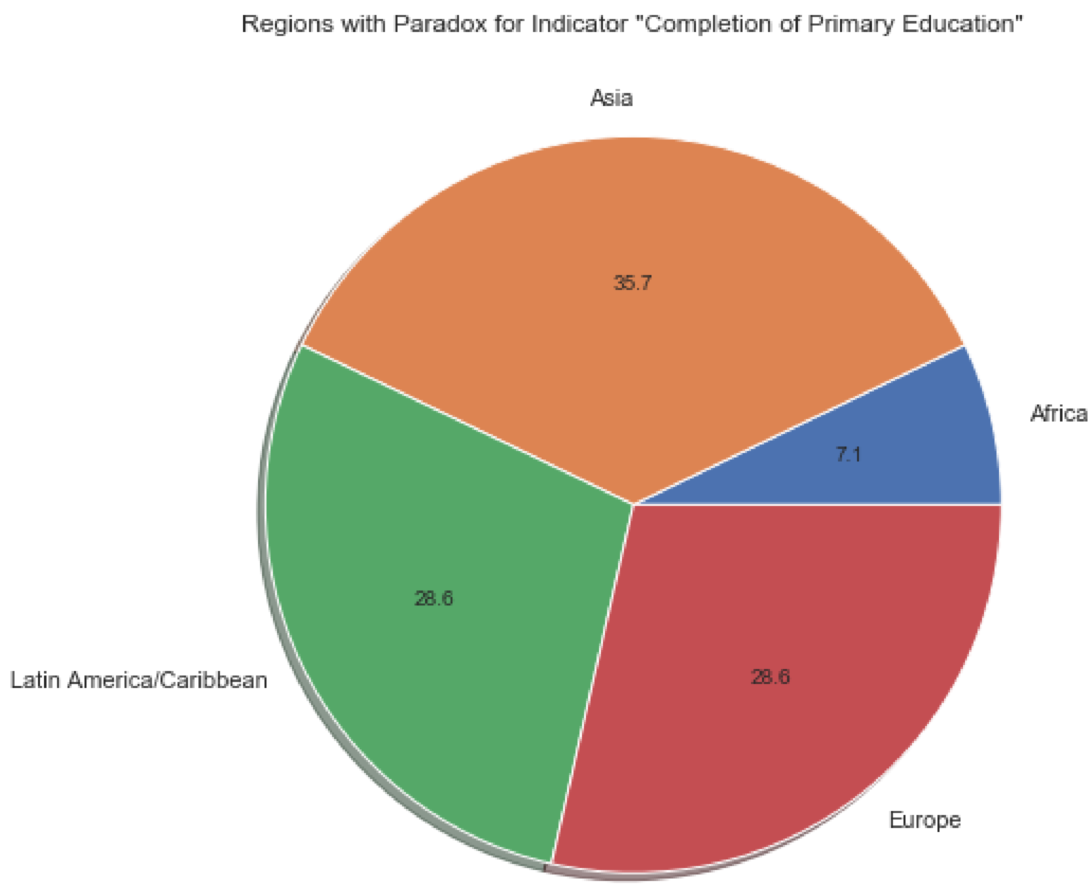
```
The number of countries with an urban paradox is 14
The percentage of countries experiencing an urban paradox is 17.94871794871795
percent.
```



[41]:
```
x = np.array(r1_edu['Indicator value']) > np.array(u1_edu['Indicator value'])
region = u1_edu['Region']

nations_edu = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_edu, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Completion of Primary␣
 ↪Education"')
```

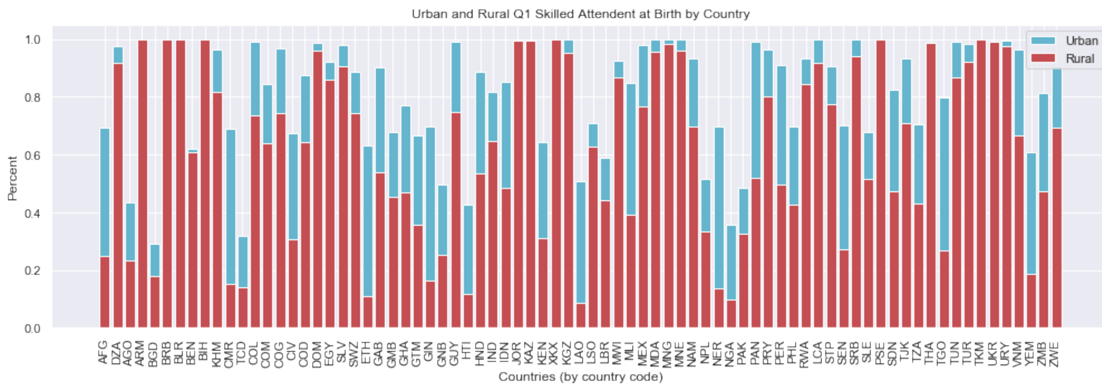
```
C:\Users\Annalise\Anaconda3\lib\site-packages\ipykernel_launcher.py:1:
RuntimeWarning: invalid value encountered in greater
  """Entry point for launching an IPython kernel.
```

[41]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Completion of Primary
Education"')
```



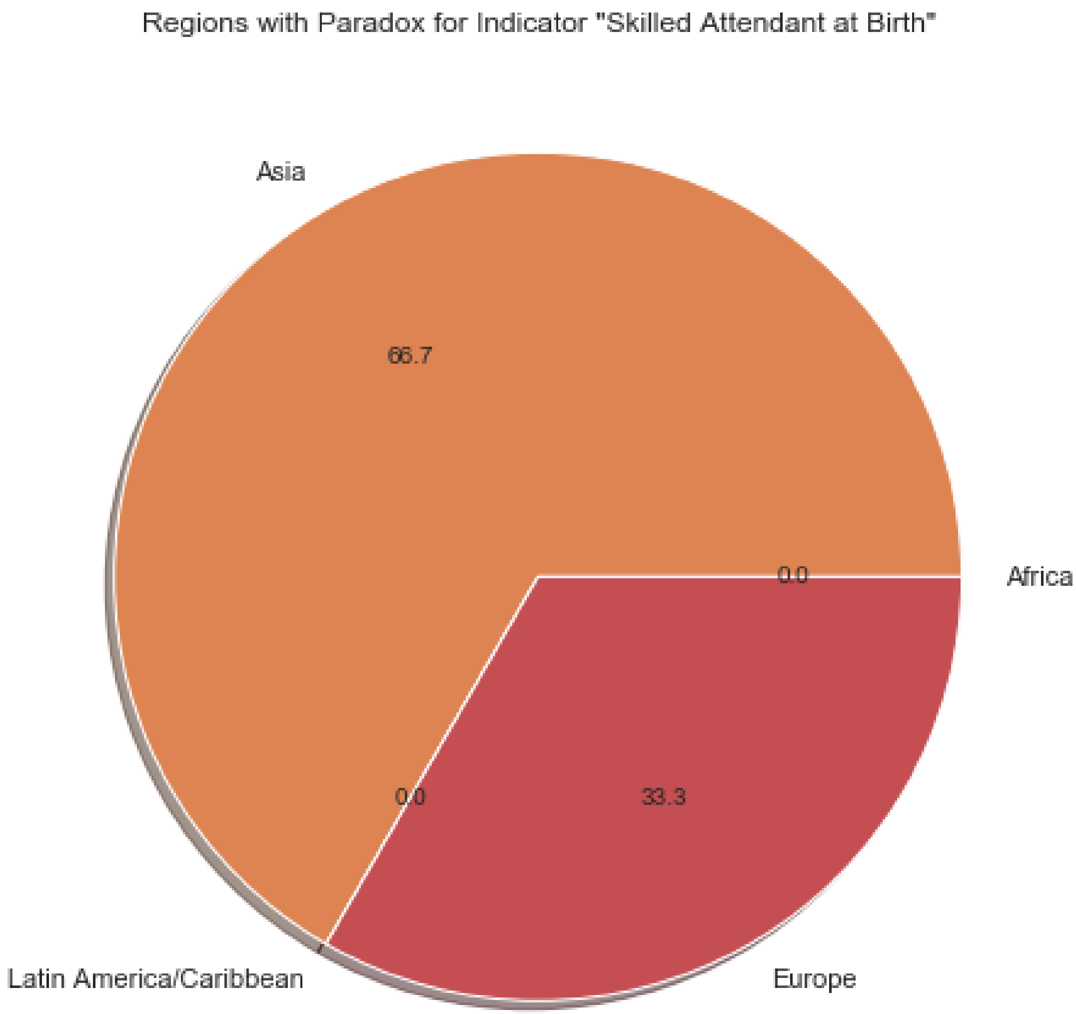
[27]:
```
plt.figure(figsize= (17,5))

plt.bar(u1_sba['Iso3code'], u1_sba['Indicator value'], label = 'Urban', color =␣
 ↪'c')
plt.bar(r1_sba['Iso3code'], r1_sba['Indicator value'], label = 'Rural', color =␣
 ↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Skilled Attendent at Birth by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent')

print('The number of countries with an urban paradox is',␣
 ↪urban_paradox(u1_sba['Indicator value'], r1_sba['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox(u1_sba['Indicator value'], r1_sba['Indicator value'])/
 ↪len(u1_sba['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 6
The percentage of countries experiencing an urban paradox is 7.792207792207792
percent.
```



[42]:
```python
x = np.array(r1_sba['Indicator value']) > np.array(u1_sba['Indicator value'])
region = u1_sba['Region']

nations_sba = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_sba, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Skilled Attendant at Birth"')
```

[42]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Skilled Attendant at
Birth"')
```

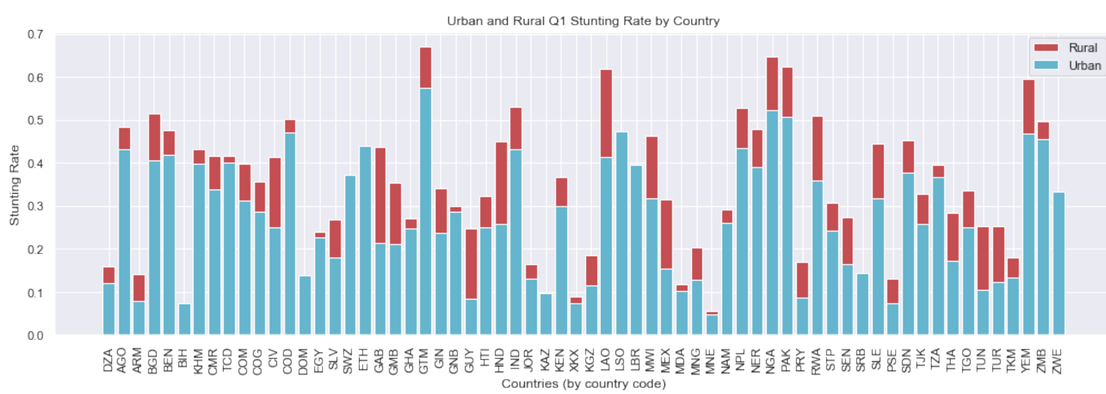


[29]:
```
def urban_paradox_2(urban_array, rural_array): # making another function as␣
↪lower values are desired for stunting and mortality
    count = 0
    x = np.array(rural_array) < np.array(urban_array) # finding places where␣
↪rural value is lower than urban
    for i in x:
        if i == True:
            count += 1 # adding the number of countries where condition is true␣
↪to the count
    return(count) # returning the count
```

[30]:
```
plt.figure(figsize= (17,5))
```

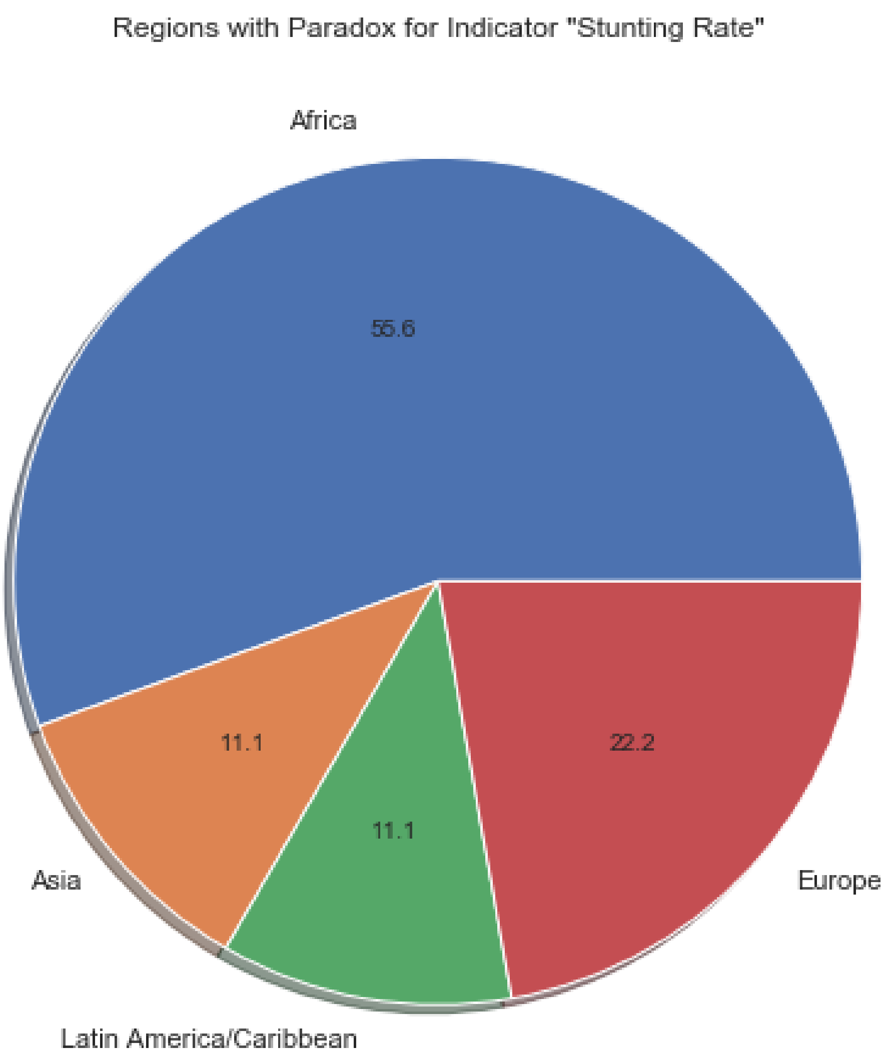
```
plt.bar(r1_stunt['Iso3code'], r1_stunt['Indicator value'], label = 'Rural',␣
 ↪color = 'r')
plt.bar(u1_stunt['Iso3code'], u1_stunt['Indicator value'], label = 'Urban',␣
 ↪color = 'c')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Stunting Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Stunting Rate')

print('The number of countries with an urban paradox is',␣
 ↪urban_paradox_2(u1_stunt['Indicator value'], r1_stunt['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox_2(u1_stunt['Indicator value'], r1_stunt['Indicator value'])/
 ↪len(u1_stunt['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 9
The percentage of countries experiencing an urban paradox is 14.0625 percent.
```



[43]:
```
x = np.array(r1_stunt['Indicator value']) < np.array(u1_stunt['Indicator␣
 ↪value'])
region = u1_stunt['Region']

nations_stunt = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_stunt, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Stunting Rate"')
```

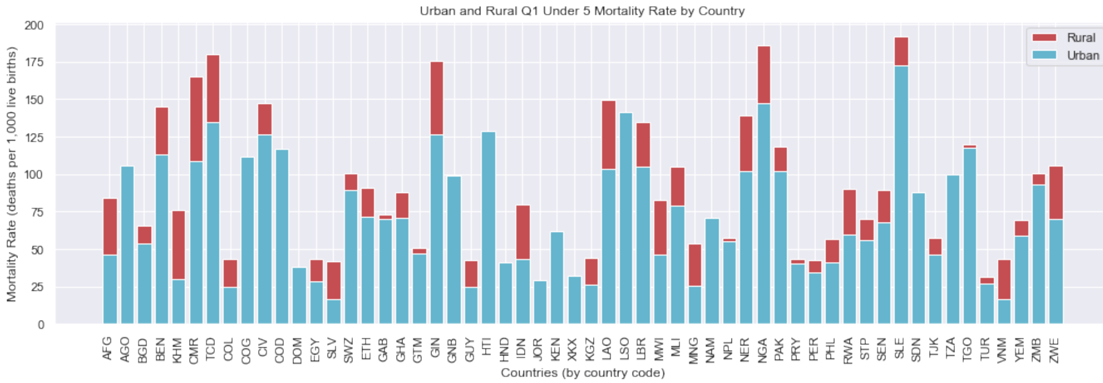
[43]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Stunting Rate"')
```



[32]:
```
plt.figure(figsize= (17,5))

plt.bar(r1_u5mr['Iso3code'], r1_u5mr['Indicator value'], label = 'Rural', color␣
 ↪= 'r')
plt.bar(u1_u5mr['Iso3code'], u1_u5mr['Indicator value'], label = 'Urban', color␣
 ↪= 'c')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Under 5 Mortality Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Mortality Rate (deaths per 1,000 live births)')
```

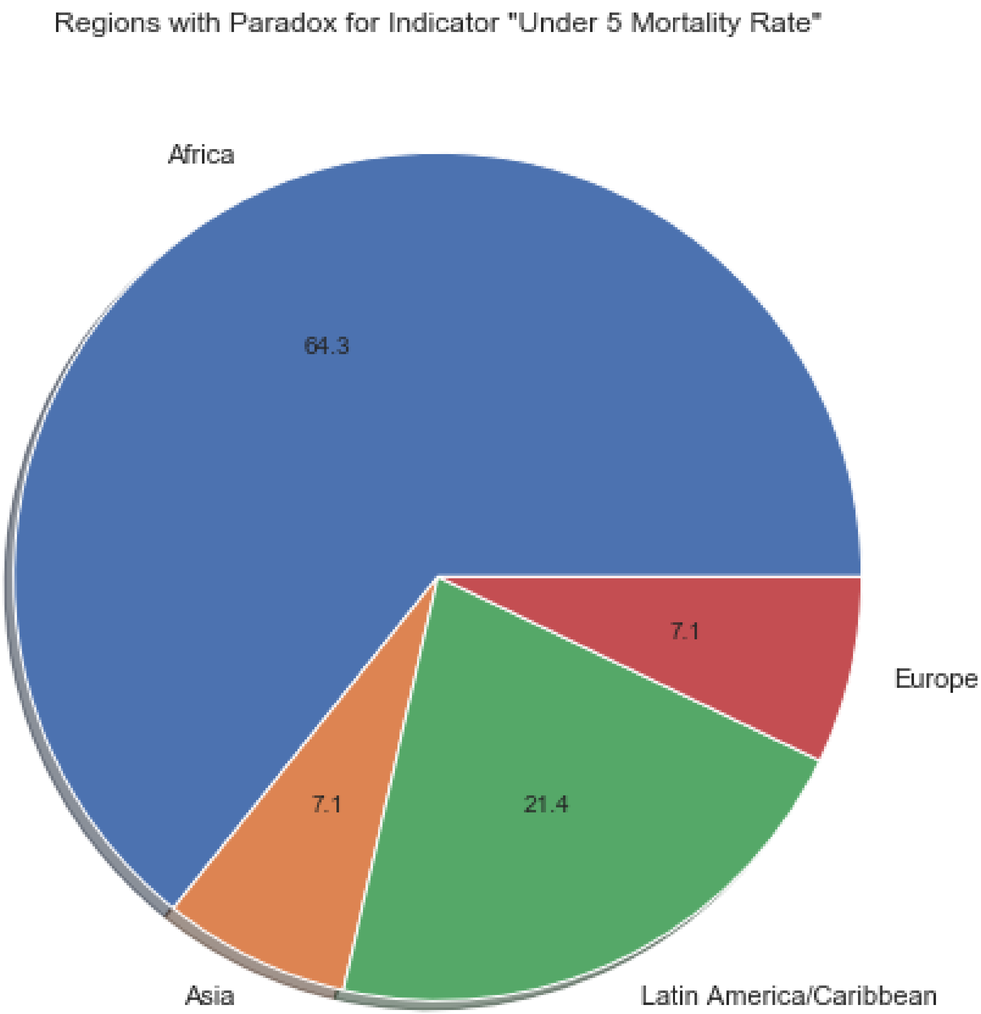
```
print('The number of countries with an urban paradox is',␣
 ↪urban_paradox_2(u1_u5mr['Indicator value'], r1_u5mr['Indicator value']))
print('The percentage of countries experiencing an urban paradox is',␣
 ↪(urban_paradox_2(u1_u5mr['Indicator value'], r1_u5mr['Indicator value'])/
 ↪len(u1_u5mr['Indicator value']))*100, 'percent.')
```

```
The number of countries with an urban paradox is 14
The percentage of countries experiencing an urban paradox is 25.0 percent.
```



[44]:
```
x = np.array(r1_u5mr['Indicator value']) < np.array(u1_u5mr['Indicator value'])
region = u1_u5mr['Region']

nations_u5mr = find_region(region[x])
labels = ['Africa', 'Asia', 'Latin America/Caribbean', 'Europe']

plt.figure(figsize= (8,8))
plt.pie(nations_u5mr, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Under 5 Mortality Rate"')
```

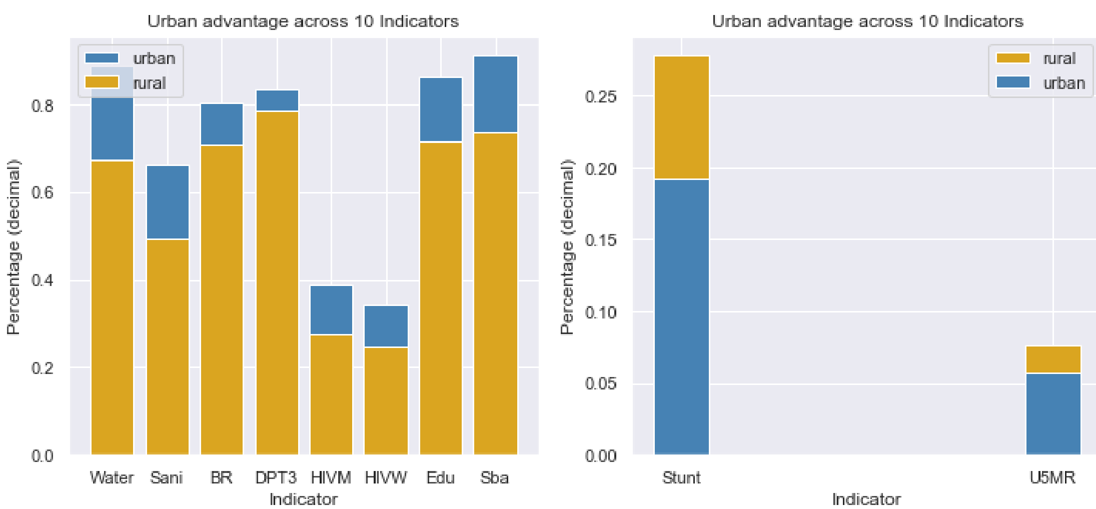
[44]:
```
Text(0.5, 1.0, 'Regions with Paradox for Indicator "Under 5 Mortality Rate"')
```



These charts show that there is indeed some urban paradox, as urban Q1 indicator values are found to be worse in some countries. This better matches with the results from UNICEF's study. Furthermore, in the indicator areas which were previously found to have the least urban advantage, there is the greatest urban paradox. The pie charts show a great variation in where there appears to be paradox, although it does appear to have some more prevalence in Africa and Asia.

Results: It is clear that there is an urban advantage across all indicators looking at average statistics. When urban indicator values are plotted with rural indicator values, there is alower stunting and mortality rate for urban values, and higher urban numbers for all other indicators. This follows the findings of UNICEF, which also saw a clear urban advantage across all indicators ("Advantage or Paradox?", 2018). The size of the advantage varied across indicators, with the lowest advantage found in DPT3 coverage and the highest in basic access to drinking water services.

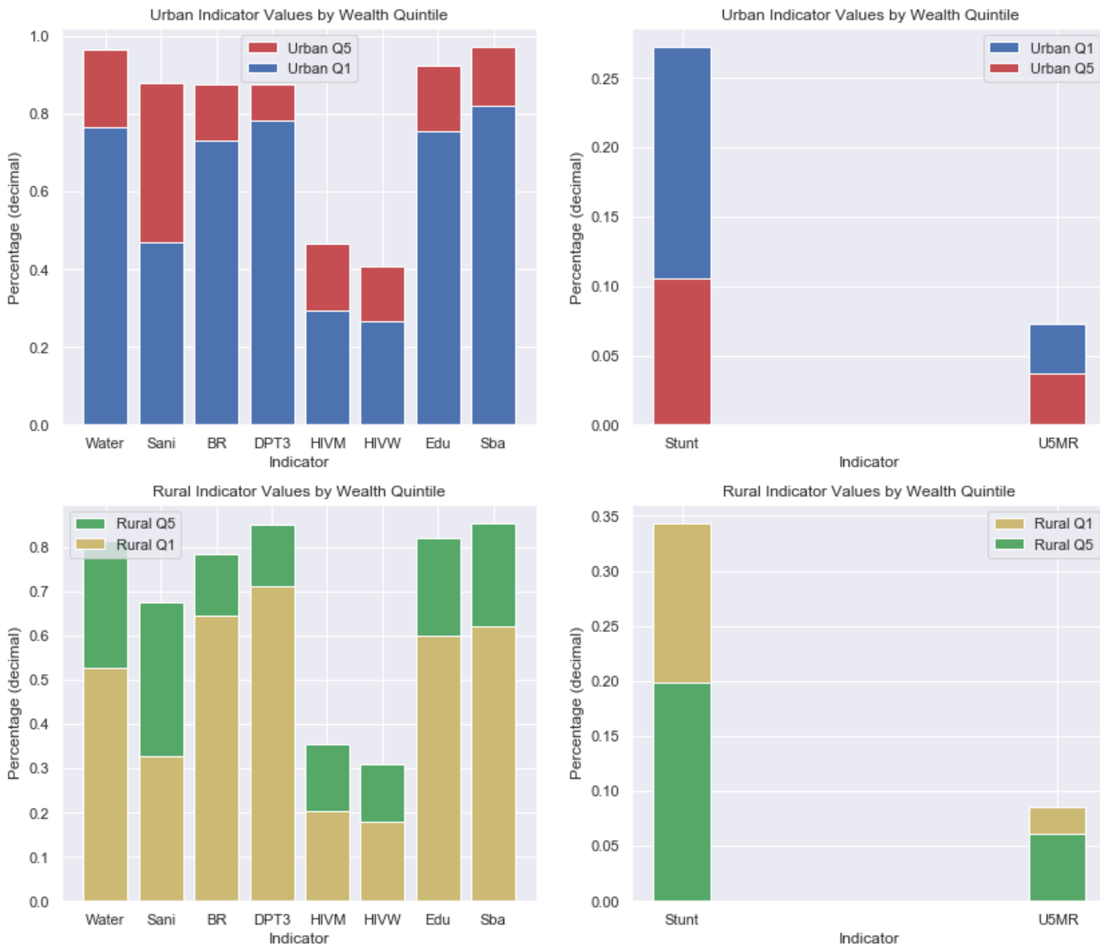
[51]:
```
plt.figure(figsize= (12,5))
sns.set()

plt.subplot(1,2,1)
plt.bar(x_axis, urban_average, label = 'urban', align = 'center', color =␣
↪'steelblue')
plt.bar(x_axis, rural_average, label = 'rural', align = 'center', color =␣
↪'goldenrod')
plt.legend()
plt.title('Urban advantage across 10 Indicators')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(1,2,2)
plt.bar(x_axis_2, rural_average_2, label = 'rural', align = 'center', width = 0.
↪15, color = 'goldenrod')
plt.bar(x_axis_2, urban_average_2, label = 'urban', align = 'center', width = 0.
↪15, color = 'steelblue')
plt.legend()
plt.title('Urban advantage across 10 Indicators')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
plt.savefig('Urban Advantage Across 10 Indicators') # saving figure to use in␣
↪google slides presentation
```



When indicator data is plotted for both the richest and poorest quintiles, there is another clear advantage that comes with wealth. There are better indicator values for quintile five in both urban and rural cases. The wealth disparity seems to be slightly higher in rural as to urban situations, with an average difference around 18 and 16 percent disparity, respectively. The data reveals that averages may mask inequities within both urban and rural areas.

[53]:

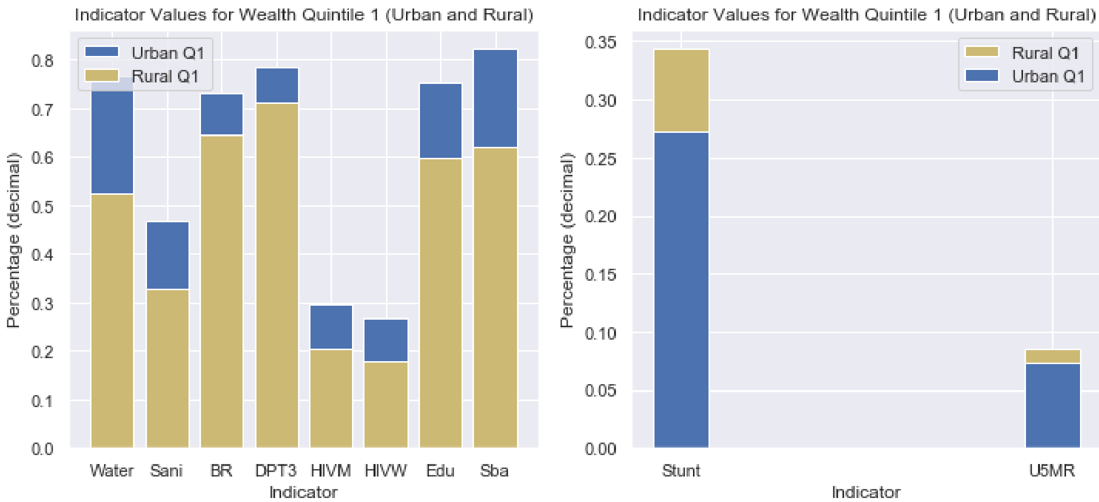
```
plt.figure(figsize= (14,12))

plt.subplot(2,2,1)
plt.bar(x_axis, urban_average_Q5, label = 'Urban Q5', align = 'center', color =␣
 ↪'r')
plt.bar(x_axis, urban_average_Q1, label = 'Urban Q1', align = 'center', color =␣
 ↪'b')
plt.legend()
plt.title('Urban Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,2)
plt.bar(x_axis_2, urban_average_2_Q1, label = 'Urban Q1', align = 'center',␣
 ↪color = 'b', width = 0.15)
plt.bar(x_axis_2, urban_average_2_Q5, label = 'Urban Q5', align = 'center',␣
 ↪color = 'r', width = 0.15)
plt.legend()
plt.title('Urban Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,3)
plt.bar(x_axis, rural_average_Q5, label = 'Rural Q5', align = 'center', color =␣
 ↪'g')
plt.bar(x_axis, rural_average_Q1, label = 'Rural Q1', align = 'center', color =␣
 ↪'y')
plt.legend()
plt.title('Rural Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.subplot(2,2,4)
plt.bar(x_axis_2, rural_average_2_Q1, label = 'Rural Q1', align = 'center',␣
 ↪color = 'y', width = 0.15)
plt.bar(x_axis_2, rural_average_2_Q5, label = 'Rural Q5', align = 'center',␣
 ↪color = 'g', width = 0.15)
plt.legend()
plt.title('Rural Indicator Values by Wealth Quintile')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')

plt.savefig('Indicators by Wealth Quintiles')
```



A main topic of this study is to look at urban paradox, referring to locations where the poorest children in rural areas are actually fairing better than their urban counterparts, going against the concept of urban advantage. When plotting the averages for Q1 data in both rural and urban situations, there still seemed to be an urban advantage, contrary to the suggestions of UNICEF ("Advantage or Paradox?", 2018).

[54]:
```
plt.figure(figsize= (12,5))

plt.subplot(1,2,1)
plt.bar(x_axis, urban_average_Q1, label = 'Urban Q1', align = 'center', color = 'b')
plt.bar(x_axis, rural_average_Q1, label = 'Rural Q1', align = 'center', color = 'y')
plt.title('Indicator Values for Wealth Quintile 1 (Urban and Rural)')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
plt.legend()

plt.subplot(1,2,2)
```

```
plt.bar(x_axis_2, rural_average_2_Q1, label = 'Rural Q1', align = 'center',␣
 ↪color = 'y', width = 0.15)
plt.bar(x_axis_2, urban_average_2_Q1, label = 'Urban Q1', align = 'center',␣
 ↪color = 'b', width = 0.15)
plt.title('Indicator Values for Wealth Quintile 1 (Urban and Rural)')
plt.xlabel('Indicator')
plt.ylabel('Percentage (decimal)')
plt.legend()

plt.savefig('Average Indicator Values for Q1')
```



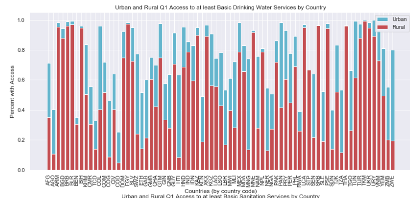
As was observed when shifting from urban and rural averages to controlling for wealth, using the mean value can disguise lots of disparity between values. When quintile one indicator values were plotted by country, the urban paradox became evident across all indicators to varying degrees. When the rural bars exceeded those of urban across all indicators, exempt from stunting rate and mortality (in which case rural would have fallen behind), there was an sign of paradox, as it reveals that urban children are faring worse than rural examples. Interestingly, the indicators which saw the least urban advantage also displayed the greatest urban paradox, with 28 percent of countries showing paradox in the case of DPT3 coverage, which had the least advantage. This proved that the urban paradox does exist in varying levels across indicators.

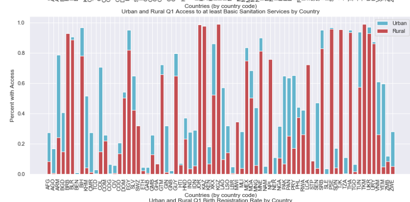
[55]:
```
plt.figure(figsize= (14,70))

plt.subplot(10,1,1)
plt.bar(u1_water['Iso3code'], u1_water['Indicator value'], label = 'Urban',␣
 ↪color = 'c')
plt.bar(r1_water['Iso3code'], r1_water['Indicator value'], label = 'Rural',␣
 ↪color = 'r')
plt.xticks(rotation=90)
plt.legend()
```

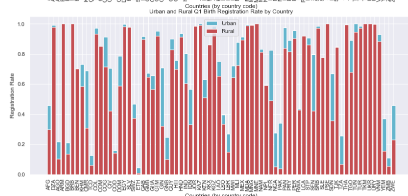
```
plt.title('Urban and Rural Q1 Access to at least Basic Drinking Water Services␣
↪by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with Access')

plt.subplot(10,1,2)
plt.bar(u1_san['Iso3code'], u1_san['Indicator value'], label = 'Urban', color =␣
↪'c')
plt.bar(r1_san['Iso3code'], r1_san['Indicator value'], label = 'Rural', color =␣
↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Access to at least Basic Sanitation Services by␣
↪Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with Access')

plt.subplot(10,1,3)
plt.bar(u1_birth['Iso3code'], u1_birth['Indicator value'], label = 'Urban',␣
↪color = 'c')
plt.bar(r1_birth['Iso3code'], r1_birth['Indicator value'], label = 'Rural',␣
↪color = 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Birth Registration Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Registration Rate')

plt.subplot(10,1,4)
plt.bar(u1_dpt3['Iso3code'], u1_dpt3['Indicator value'], label = 'Urban', color␣
↪= 'c')
plt.bar(r1_dpt3['Iso3code'], r1_dpt3['Indicator value'], label = 'Rural', color␣
↪= 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 DPT3 Coverage by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent Coverage')

plt.subplot(10,1,5)
plt.bar(u1_hivm['Iso3code'], u1_hivm['Indicator value'], label = 'Urban', color␣
↪= 'c')
plt.bar(r1_hivm['Iso3code'], r1_hivm['Indicator value'], label = 'Rural', color␣
↪= 'r')
plt.xticks(rotation=90)
plt.legend()
```

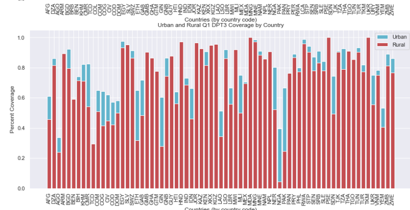
```
plt.title('Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young Men␣
 ↪by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with With Knowledge')

plt.subplot(10,1,6)
plt.bar(u1_hivw['Iso3code'], u1_hivw['Indicator value'], label = 'Urban', color␣
 ↪= 'c')
plt.bar(r1_hivw['Iso3code'], r1_hivw['Indicator value'], label = 'Rural', color␣
 ↪= 'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young␣
 ↪Women by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent with With Knowledge')

plt.subplot(10,1,7)
plt.bar(u1_edu['Iso3code'], u1_edu['Indicator value'], label = 'Urban', color =␣
 ↪'c')
plt.bar(r1_edu['Iso3code'], r1_edu['Indicator value'], label = 'Rural', color =␣
 ↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Completion of Primary Education by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent')

plt.subplot(10,1,8)
plt.bar(u1_sba['Iso3code'], u1_sba['Indicator value'], label = 'Urban', color =␣
 ↪'c')
plt.bar(r1_sba['Iso3code'], r1_sba['Indicator value'], label = 'Rural', color =␣
 ↪'r')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Skilled Attendent at Birth by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Percent')

plt.subplot(10,1,9)
plt.bar(r1_stunt['Iso3code'], r1_stunt['Indicator value'], label = 'Rural',␣
 ↪color = 'r')
plt.bar(u1_stunt['Iso3code'], u1_stunt['Indicator value'], label = 'Urban',␣
 ↪color = 'c')
plt.xticks(rotation=90)
plt.legend()
```

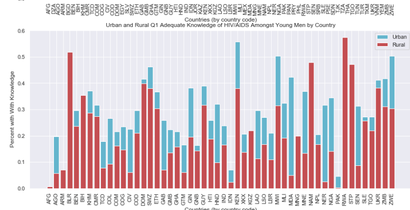
```
plt.title('Urban and Rural Q1 Stunting Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Stunting Rate')

plt.subplot(10,1,10)
plt.bar(r1_u5mr['Iso3code'], r1_u5mr['Indicator value'], label = 'Rural', color␣
 ↪= 'r')
plt.bar(u1_u5mr['Iso3code'], u1_u5mr['Indicator value'], label = 'Urban', color␣
 ↪= 'c')
plt.xticks(rotation=90)
plt.legend()
plt.title('Urban and Rural Q1 Under 5 Mortality Rate by Country')
plt.xlabel('Countries (by country code)')
plt.ylabel('Mortality Rate (deaths per 1,000 live births)')


plt.savefig('Indicator Q1 values by Country')
```

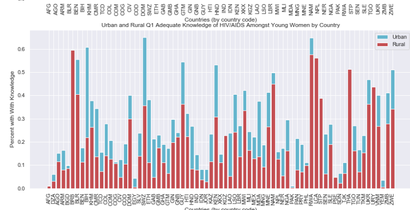

Urban and Rural Q1 Access to at least Basic Drinking Water Services by Country
Percent with Access
Countries (by country code)
Urban
Rural

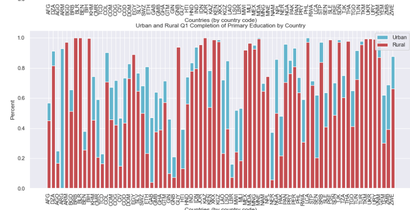

Urban and Rural Q1 Access to at least Basic Sanitation Services by Country
Percent with Access
Countries (by country code)
Urban
Rural

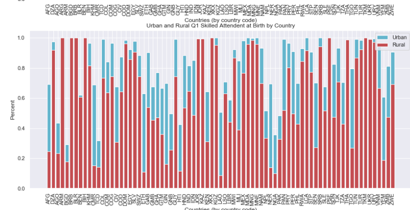

Urban and Rural Q1 Birth Registration Rate by Country
Registration Rate
Countries (by country code)
Urban
Rural

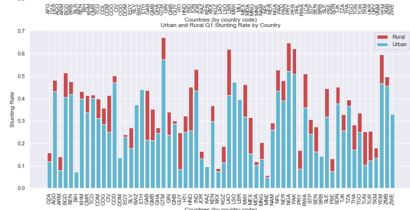

Urban and Rural Q1 DPT3 Coverage by Country
Percent Coverage
Countries (by country code)
Urban
Rural

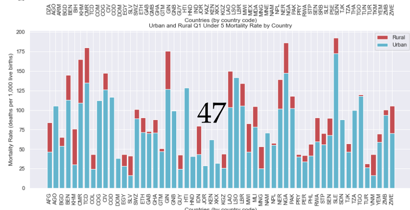

Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young Men by Country
Percent with With Knowledge
Countries (by country code)
Urban
Rural


Urban and Rural Q1 Adequate Knowledge of HIV/AIDS Amongst Young Women by Country
Percent with With Knowledge
Countries (by country code)
Urban
Rural


Urban and Rural Q1 Completion of Primary Education by Country
Percent
Countries (by country code)
Urban
Rural


Urban and Rural Q1 Skilled Attendant at Birth by Country
Percent
Countries (by country code)
Urban
Rural

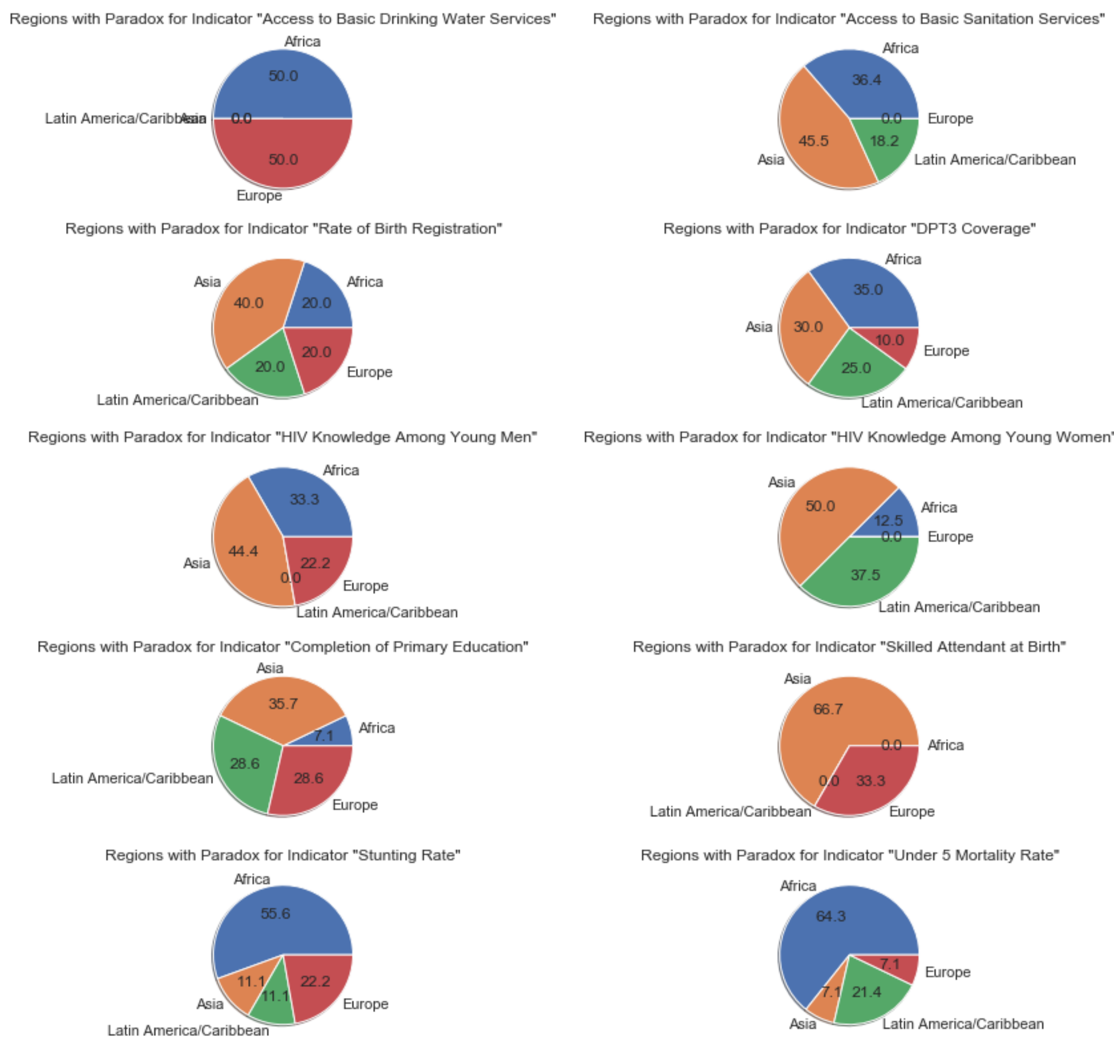

Urban and Rural Q1 Stunting Rate by Country
Stunting Rate
Countries (by country code)
Rural
Urban


Urban and Rural Q1 Under 5 Mortality Rate by Country
Mortality Rate (deaths per 1,000 live births)
Countries (by country code)
Rural
Urban

It was also desired to see if there are specific regions in which cases of urban paradox seem to be highest. The pie charts for the prevalence of urban paradox for each indicator in each region revealed great variance, although there was an increased amount of incidence in both African and Asian regions.

[56]:
```
plt.figure(figsize= (14,14))

plt.subplot(5,2,1)
plt.pie(nations_water, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Access to Basic Drinking Water␣
↪Services"')

plt.subplot(5,2,2)
plt.pie(nations_san, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Access to Basic Sanitation␣
↪Services"')

plt.subplot(5,2,3)
plt.pie(nations_birth, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Rate of Birth Registration"')

plt.subplot(5,2,4)
plt.pie(nations_dpt3, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "DPT3 Coverage"')

plt.subplot(5,2,5)
plt.pie(nations_hivm, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "HIV Knowledge Among Young Men"')

plt.subplot(5,2,6)
plt.pie(nations_hivw, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "HIV Knowledge Among Young␣
↪Women"')

plt.subplot(5,2,7)
plt.pie(nations_edu, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Completion of Primary␣
↪Education"')

plt.subplot(5,2,8)
plt.pie(nations_sba, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Skilled Attendant at Birth"')

plt.subplot(5,2,9)
plt.pie(nations_stunt, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Stunting Rate"')
```

```python
plt.subplot(5,2,10)
plt.pie(nations_u5mr, labels = labels, autopct = '%.1f', shadow = True)
plt.title('Regions with Paradox for Indicator "Under 5 Mortality Rate"')


plt.savefig('Regions with Paradox Pie Charts')
```



Synthesis/Discussion: From this study, it can be seen that the urban paradox is indeed an issue which should be further looked into. Although on average there appears to be an urban advantage, when controlling for wealth one can see that much of this advantage disappears and in some cases it becomes a disadvantage. There is great wealth disparity in both urban and rural populations which impact the ten health indicator values for the worse in the case of the poorest people. The urban paradox seems to focus on Africa and Asia, both of which include several developing nations, which might be a part of the issue. If the study were to be done again, it would be beneficial to find a way to show the actual values of the chunks on the pie charts, as certain indicators had more pradox that others, which does not show in the charts. It would also be ideal to possibly look at some of these indicator values over time, as it might help to see how urban paradox grows or falls over the years.

Works Cited:

Advantage or Paradox: The challenge for children and young people of growing up urban. (2018, November). Retrieved April 16, 2019, from https://data.unicef.org/resources/urban-paradox-report/