rural_population_growth

May 1, 2022

Program to load population growth data, clean up and export as cleaned up CSV file.

[7]:
```
import pandas as pd
import numpy as np

df_pop_growth = pd.read_csv ('../Data/Rural_Pop_growth_WorldBank/
 ↪rural_pop_growth_data.csv')

df_pop_growth.head()

#pd.unique(df_pop_growth['Series Code'])
```

[7]:

	Series Name	Series Code	Country Name	Country Code	1960 [YR1960]	1961 [YR1961]	1962 [YR1962]	1963 [YR1963]	1964 [YR1964]	1965 [YR1965]	...	2012 [YR2012]	2013 [YR2013]
0	Rural population (% of total population)	SP.RUR.TOTL.ZS	Afghanistan	AFG	91.599	91.316	91.024	90.724	90.414	90.096	...	75.84	75.627
1	Rural population (% of total population)	SP.RUR.TOTL.ZS	Africa Eastern and Southern	AFE	85.2953119729611	85.0555405636455	84.8143917635466	84.5555861095586	84.2809240839644	83.9980078257727	...	66.9283128426341	66.4924520296653
2	Rural population (% of total population)	SP.RUR.TOTL.ZS	Africa Western and Central	AFW	85.3296712092282	84.9464227530417	84.550718370091	84.1359792709003	83.7050160428521	83.2586392998873	...	57.0989645749267	56.4692956838611
3	Rural population (% of total population)	SP.RUR.TOTL.ZS	Albania	ALB	69.295	69.057	68.985	68.914	68.842				
4	Rural population (% of total population)	SP.RUR.TOTL.ZS	Algeria	DZA	69.49	68.203	66.786	65.338	63.859				

```
3             68.77  ...             45.67            44.613
4            62.357  ...            31.085            30.424

      2014 [YR2014]     2015 [YR2015]     2016 [YR2016]     2017 [YR2017]  \
0            75.413            75.197             74.98             74.75
1  66.0475756414397  65.5949160938201  65.1348990608866   64.667627399344
2  55.8428690387667  55.2186269982415  54.5981908883918  53.9806664573227
3            43.577            42.566            41.579            40.617
4            29.779            29.152            28.541            27.948

      2018 [YR2018]     2019 [YR2019]     2020 [YR2020] 2021 [YR2021]
0            74.505            74.246            73.974            ..
1  64.1922298807226  63.7086779320966  63.2166938509081            ..
2  53.3668617220008  52.7571471216999  52.1513745647494            ..
3            39.681            38.771            37.888            ..
4            27.371            26.811            26.267            ..

[5 rows x 66 columns]
```

[8]:
```
df_pop_growth_transpose = df_pop_growth.melt(id_vars=["Series Name", "Series␣
↪Code", "Country Name","Country Code"],
        var_name="Year",
        value_name="Rur_perc")
```

[9]:
```
df_pop_growth_transpose.head(20)
```

[9]:
```
                                Series Name     Series Code  \
0   Rural population (% of total population)  SP.RUR.TOTL.ZS
1   Rural population (% of total population)  SP.RUR.TOTL.ZS
2   Rural population (% of total population)  SP.RUR.TOTL.ZS
3   Rural population (% of total population)  SP.RUR.TOTL.ZS
4   Rural population (% of total population)  SP.RUR.TOTL.ZS
5   Rural population (% of total population)  SP.RUR.TOTL.ZS
6   Rural population (% of total population)  SP.RUR.TOTL.ZS
7   Rural population (% of total population)  SP.RUR.TOTL.ZS
8   Rural population (% of total population)  SP.RUR.TOTL.ZS
9   Rural population (% of total population)  SP.RUR.TOTL.ZS
10  Rural population (% of total population)  SP.RUR.TOTL.ZS
11  Rural population (% of total population)  SP.RUR.TOTL.ZS
12  Rural population (% of total population)  SP.RUR.TOTL.ZS
13  Rural population (% of total population)  SP.RUR.TOTL.ZS
14  Rural population (% of total population)  SP.RUR.TOTL.ZS
15  Rural population (% of total population)  SP.RUR.TOTL.ZS
16  Rural population (% of total population)  SP.RUR.TOTL.ZS
17  Rural population (% of total population)  SP.RUR.TOTL.ZS
18  Rural population (% of total population)  SP.RUR.TOTL.ZS
19  Rural population (% of total population)  SP.RUR.TOTL.ZS
```

	Country Name	Country Code	Year	Rur_perc
0	Afghanistan	AFG	1960 [YR1960]	91.599
1	Africa Eastern and Southern	AFE	1960 [YR1960]	85.2953119729611
2	Africa Western and Central	AFW	1960 [YR1960]	85.3296712092282
3	Albania	ALB	1960 [YR1960]	69.295
4	Algeria	DZA	1960 [YR1960]	69.49
5	American Samoa	ASM	1960 [YR1960]	33.789
6	Andorra	AND	1960 [YR1960]	41.55
7	Angola	AGO	1960 [YR1960]	89.565
8	Antigua and Barbuda	ATG	1960 [YR1960]	60.344
9	Arab World	ARB	1960 [YR1960]	68.7658583608498
10	Argentina	ARG	1960 [YR1960]	26.389
11	Armenia	ARM	1960 [YR1960]	48.725
12	Aruba	ABW	1960 [YR1960]	49.224
13	Australia	AUS	1960 [YR1960]	18.471
14	Austria	AUT	1960 [YR1960]	35.28
15	Azerbaijan	AZE	1960 [YR1960]	47.337
16	Bahamas, The	BHS	1960 [YR1960]	40.288
17	Bahrain	BHR	1960 [YR1960]	17.68
18	Bangladesh	BGD	1960 [YR1960]	94.865
19	Barbados	BRB	1960 [YR1960]	63.223

[10]:
```
#Change year to 4 digits only.
df_pop_growth_transpose['Year'] = df_pop_growth_transpose['Year'].str[:4]

#Separate % of total and % growth into 2 DFs
df_rural_pop = df_pop_growth_transpose[df_pop_growth_transpose["Series Name"]␣
 ↪== 'Rural population (% of total population)']
df_rural_pop_growth = df_pop_growth_transpose[df_pop_growth_transpose["Series␣
 ↪Name"] == 'Rural population growth (annual %)']
```

[15]:
```
#If Rur_perc is ".." convert to NA
df_rural_pop = df_rural_pop.replace('..',np.NaN)
df_rural_pop_growth = df_rural_pop_growth.replace('..',np.NaN)
df_rural_pop_growth.rename(columns = {'Rur_perc':'Rur_pop_growth'}, inplace =␣
 ↪True)
```

[16]:
```
df_rural_pop.to_csv("../Data/Rural_Pop_growth_WorldBank/rural_pop_cleaned.csv")
df_rural_pop_growth.to_csv("../Data/Rural_Pop_growth_WorldBank/
 ↪rural_pop_growth_cleaned.csv")
```

[ ]:
```

```