IMPORTING ALL ESSENTIAL LIBRARIES

In [90]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

FUNCTION WHICH WILL BE USED IN DATA PREPROCESSING

In [91]:
```
def removing_comma(dataframe,stri):
    li=[]
    for i in range(len(dataframe)):
        s=""
        for j in dataframe[stri][i]:
            if j!=',':
                s=s+j
        li.append(int(s))
    return li

def removing_percent(dataframe,stri):
    gr=[]
    for i in range(len(dataframe)):
        s=""
        for j in dataframe[stri][i]:
            if j!='%':
                s=s+j
        gr.append(float(s))
```

analyzing 1.1 dataset

In [9]:

In [10]:

Out[10]:

	Unnamed: 0	year	popu1ation_growth_rate	Population_Density	growth_rate
0	0	2024	1,441,719,852	438.58	0.92%
1	1	2023	1,428,627,663	434.60	0.81%
2	2	2022	1,417,173,173	431.11	0.68%
3	3	2021	1,407,563,842	428.19	0.80%
4	4	2020	1,396,387,127	424.79	0.96%

In [11]:

Out[11]:
```
Unnamed: 0                0
year                      0
popu1ation_growth_rate    0
Population_Density        0
growth_rate               0
dtype: int64
```

In [12]:

Out[12]: (75, 5)

In [13]:

Out[13]:
```
Unnamed: 0                 int64
year                       int64
popu1ation_growth_rate    object
Population_Density       float64
growth_rate               object
dtype: object
```

In [14]:

In [15]:

Out[15]:

	Unnamed: 0	year	total_population	Population_Density	population_growth_rate
0	0	2024	1,441,719,852	438.58	0.92%
1	1	2023	1,428,627,663	434.60	0.81%
2	2	2022	1,417,173,173	431.11	0.68%
3	3	2021	1,407,563,842	428.19	0.80%
4	4	2020	1,396,387,127	424.79	0.96%

In [16]:

In [17]:

In [18]:

In [19]:

In [20]:

In [21]:

Out[21]:

	year	total_population	Population_Density	population_growth_rate
0	2024	1441719852	438.58	0.92
1	2023	1428627663	434.60	0.81
2	2022	1417173173	431.11	0.68
3	2021	1407563842	428.19	0.80
4	2020	1396387127	424.79	0.96

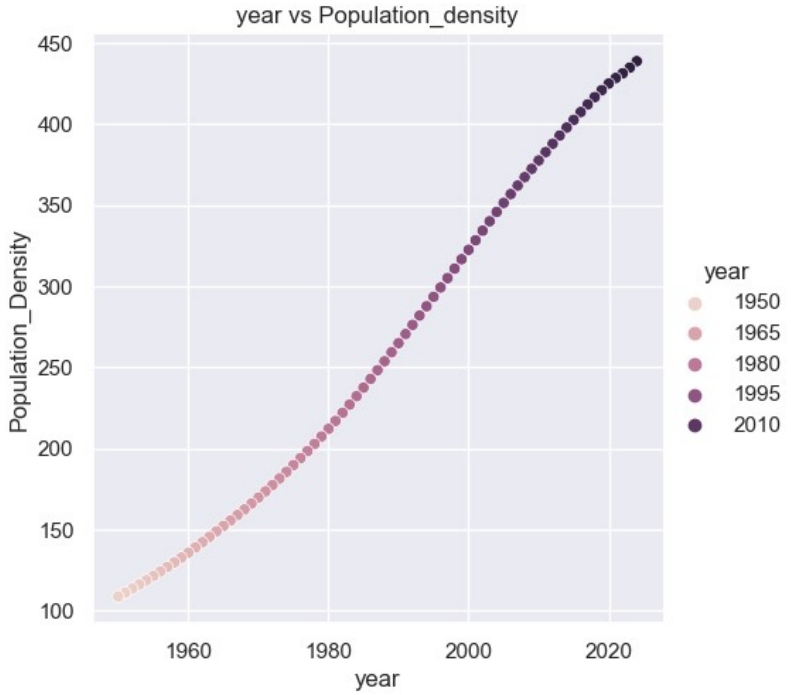
In [22]:

year vs population density

In [23]:
```
sbn.relplot(x="year",y="Population_Density",data=data1,kind="scatter",hue="
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[23]: Text(0.5, 1.0, 'year vs Population_density')

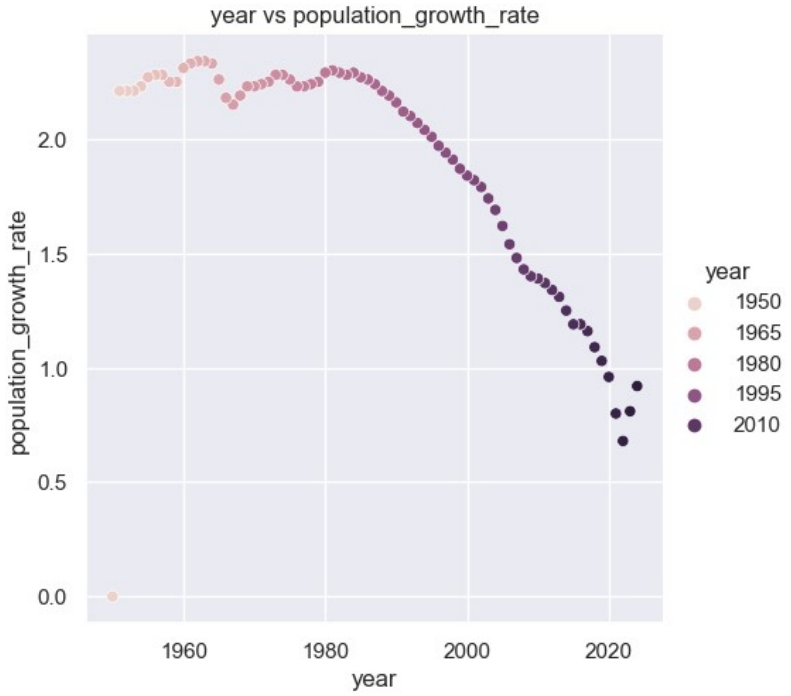


year vs population_growth_rate

In [24]:
```
sbn.relplot(x="year",y="population_growth_rate",data=data1,kind="scatter",hu
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[24]: Text(0.5, 1.0, 'year vs population_growth_rate')

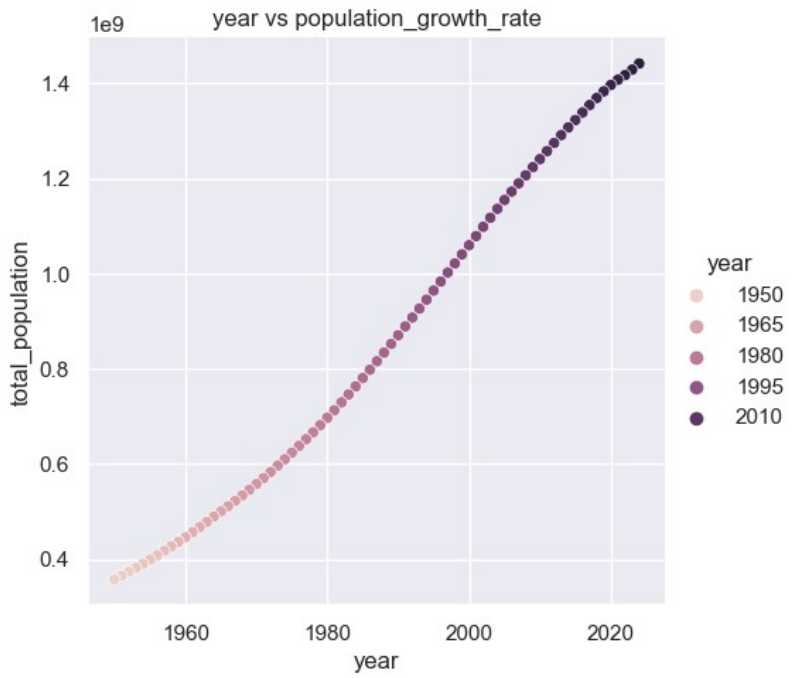


year vs population

In [25]:
```
sbn.relplot(x="year",y="total_population",data=data1,kind="scatter",hue="ye
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[25]: Text(0.5, 1.0, 'year vs population_growth_rate')

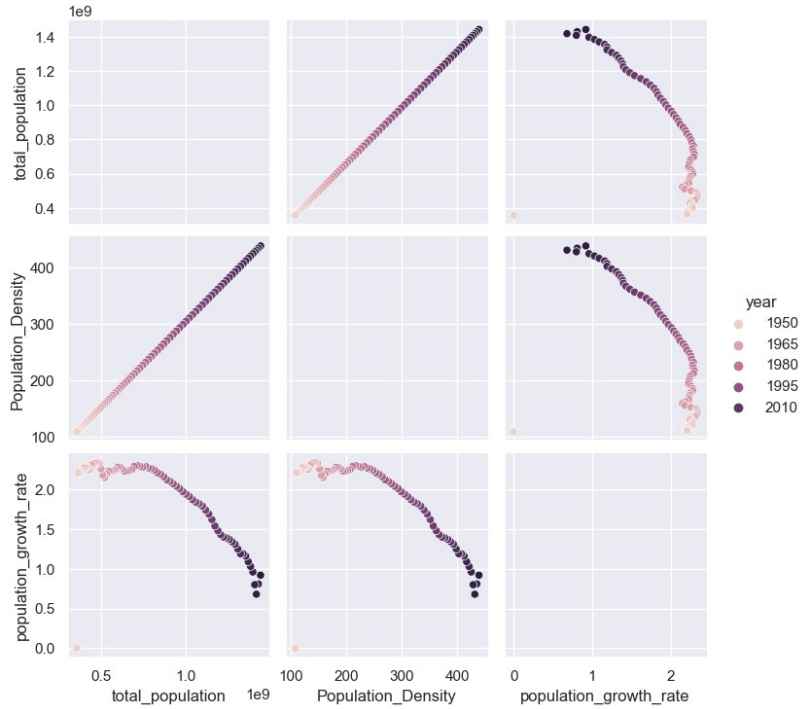


pair plot for out data1

In [26]:
```
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[26]: <seaborn.axisgrid.PairGrid at 0x2452c151e50>



analysing 1.4 dataset(rural population historical data

In [27]:
```
```

In [28]:
```
```

Out[28]:

	Unnamed: 0	year	Population	percent_of_total	Change
0	0	2022	908,804,812	64.13	-0.06%
1	1	2021	909,384,771	64.61	0.08%
2	2	2020	908,684,959	65.07	0.26%
3	3	2019	906,325,664	65.53	0.35%
4	4	2018	903,131,481	65.97	0.44%

In [29]:

Out[29]: (63, 5)

In [30]:

Out[30]:
```
Unnamed: 0          0
year                0
Population          0
percent_of_total    0
Change              0
dtype: int64
```

In [31]:

In [32]:

In [33]:

In [34]:

In [35]:

In [36]:

Out[36]:

	year	Population	percent_of_total	Change
0	2022	908804812	64.13	-0.06
1	2021	909384771	64.61	0.08
2	2020	908684959	65.07	0.26
3	2019	906325664	65.53	0.35
4	2018	903131481	65.97	0.44

In [37]:

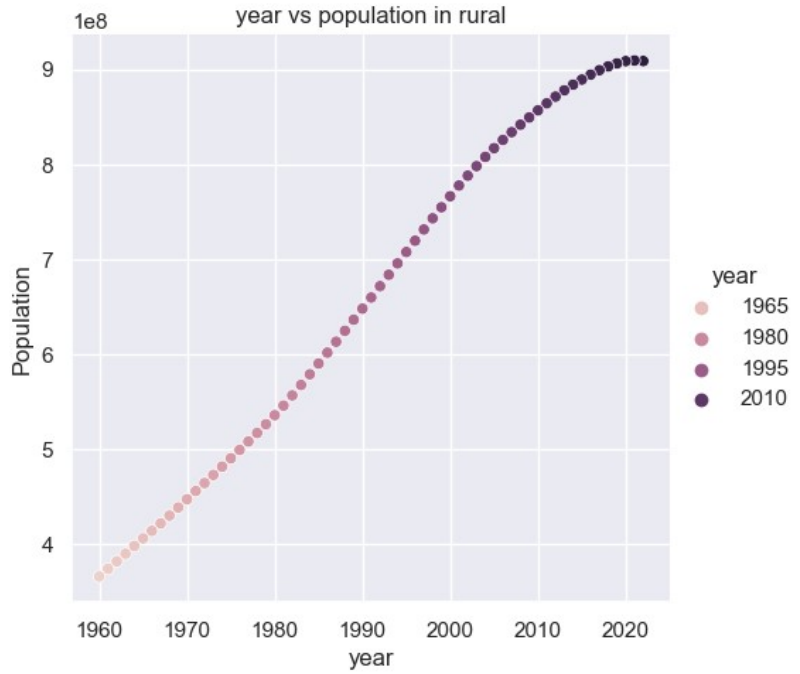
Out[37]:
```
year                int64
Population          int64
percent_of_total    float64
Change              float64
dtype: object
```

YEAR VS POPULATION IN RURAL

In [38]:
```
sbn.relplot(x="year",y="Population",kind="scatter",hue="year",data=data2)
plt.title("year vs population in rural")
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[38]: Text(0.5, 1.0, 'year vs population in rural')

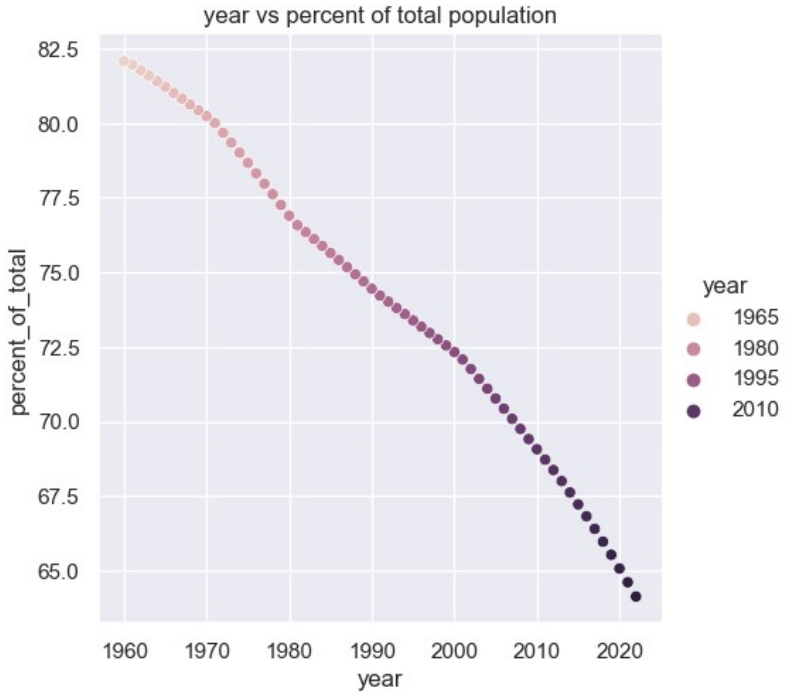


YEAR VS PERCENT OF TOTAL POPULATION

In [39]:
```
sbn.relplot(x="year",y="percent_of_total",kind="scatter",hue="year",data=da
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[39]:
```
Text(0.5, 1.0, 'year vs percent of total population')
```

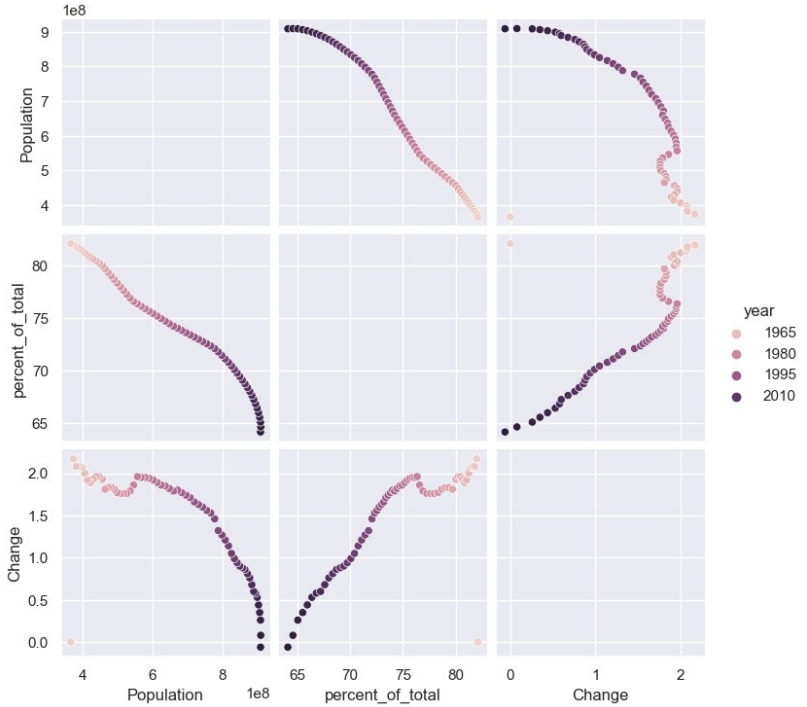


pairplot for data2

In [40]:
```
sbn.pairplot(data2,hue="year")
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[40]:
```
<seaborn.axisgrid.PairGrid at 0x2452cd29910>
```



analyzing 1.5 dataset(Urban population historical data)

In [41]:

In [42]:

Out[42]:

	Unnamed: 0	year	Population	percent_of_total	Change
0	0	2022	508,368,361	35.87	2.02%
1	1	2021	498,179,071	35.39	2.13%
2	2	2020	487,702,168	34.93	2.26%
3	3	2019	476,786,386	34.47	2.32%
4	4	2018	465,871,825	34.03	2.36%

In [43]:

Out[43]: (63, 5)

In [44]:

Out[44]:
```
Unnamed: 0          0
year                0
Population          0
percent_of_total    0
Change              0
dtype: int64
```

In [45]:

In [46]:
```
li=removing_comma(data3,"Population")
```

In [47]:
```
ch=removing_percent(data3,'Change')
```

In [48]:

Out[48]:

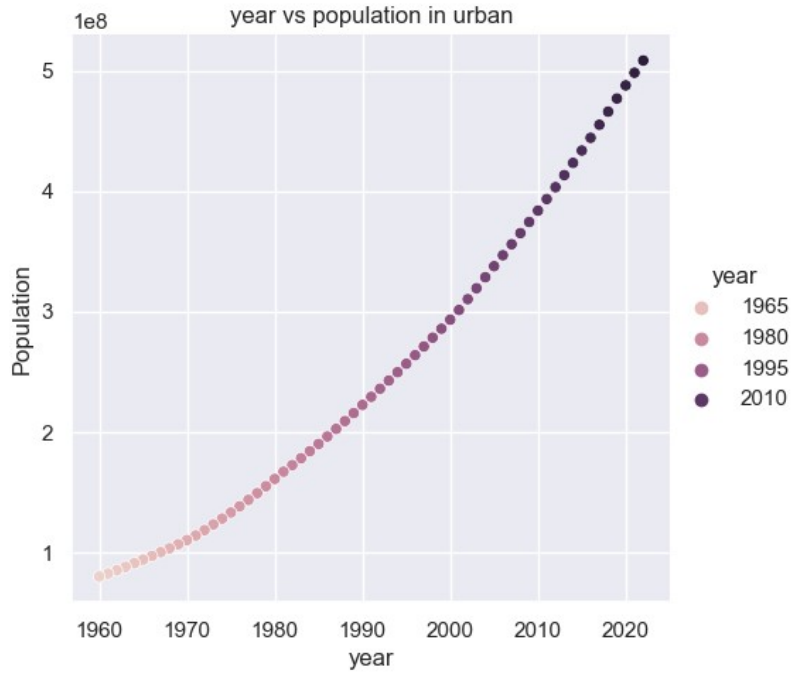
	year	Population	percent_of_total	Change
0	2022	508368361	35.87	2.02
1	2021	498179071	35.39	2.13
2	2020	487702168	34.93	2.26
3	2019	476786386	34.47	2.32
4	2018	465871825	34.03	2.36

YEAR VS POPULATION IN URBAN

In [49]:
```
sbn.relplot(x="year",y="Population",kind="scatter",hue="year",data=data3)
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[49]: Text(0.5, 1.0, 'year vs population in urban')

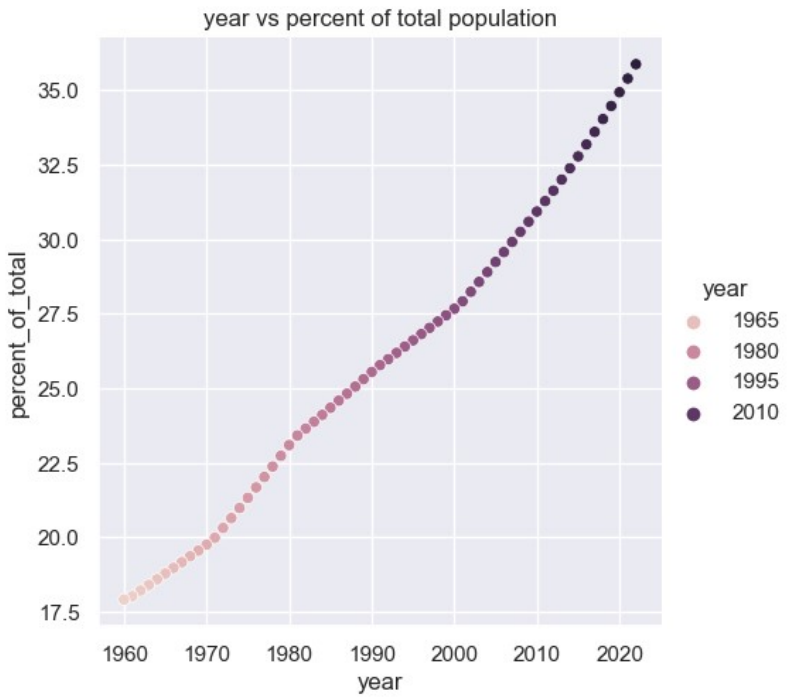


YEAR VS PERCENT OF TOTAL POPULATION

In [50]:
```
sbn.relplot(x="year",y="percent_of_total",kind="scatter",hue="year",data=da
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[50]:
```
Text(0.5, 1.0, 'year vs percent of total population')
```



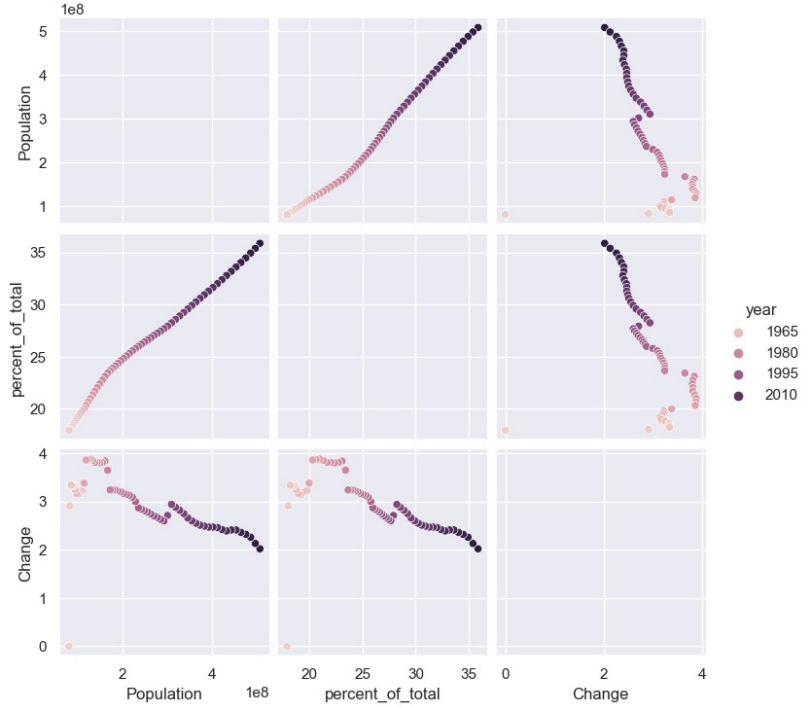
PAIRPLOT FOR DATA3

In [51]:
```

```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

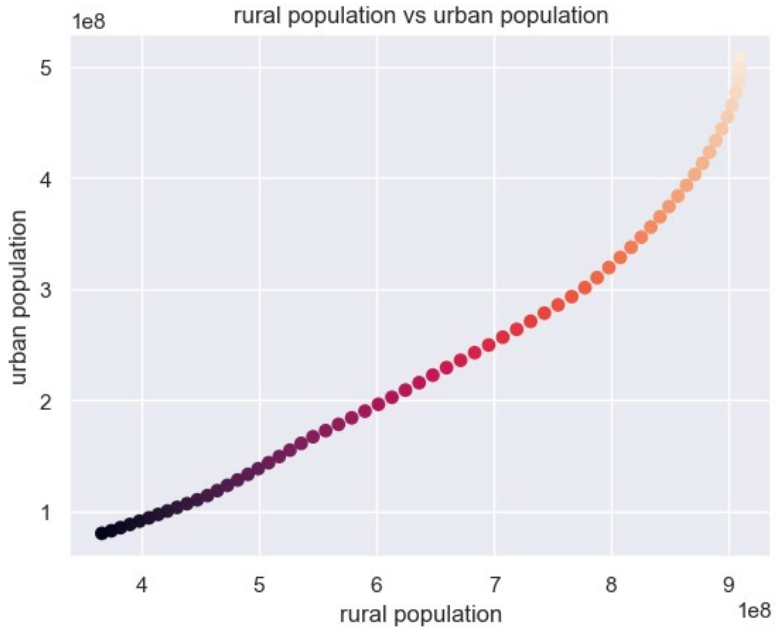
Out[51]:
```
<seaborn.axisgrid.PairGrid at 0x2452e968850>
```



rural vs urban (pop vs pop)

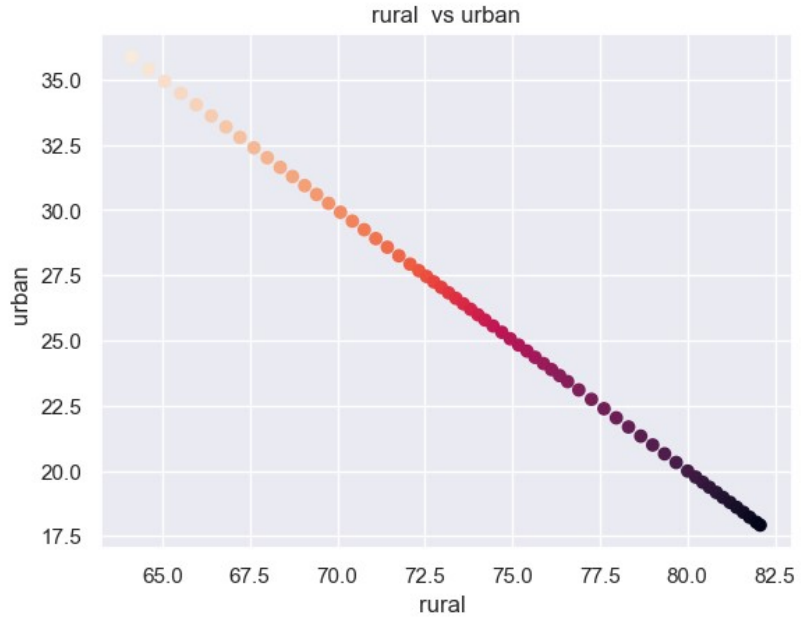
In [52]:
```
plt.title("rural population vs urban population")
plt.scatter(data2["Population"],data3["Population"],c=data2["year"])
plt.xlabel("rural population")
```

Out[52]: Text(0, 0.5, 'urban population')



In [53]:
```
plt.title("rural  vs urban")
plt.scatter(data2["percent_of_total"],data3["percent_of_total"],c=data2["ye
plt.xlabel("rural ")
```

Out[53]: Text(0, 0.5, 'urban ')



PERCENT OF TOTAL POPULATION

analysis of DATASET 1.6-2.0

In [54]:

In [55]:
```
data4.head()
```

Out[55]:

	Unnamed: 0	year	Birth_Rate	Growth_Rate
0	0	2024	16.750	-1.170%
1	1	2023	16.949	-1.250%
2	2	2022	17.163	-1.230%
3	3	2021	17.377	-1.220%
4	4	2020	17.592	-1.200%

In [56]:
```
li=removing_percent(data4,"Growth_Rate")
```

In [57]:

In [58]:

In [59]:

In [60]:

Out[60]:

	Unnamed: 0	year	Death_Rate	Growth_Rate
0	0	2024	7.473	0.770%
1	1	2023	7.416	0.490%
2	2	2022	7.380	0.490%
3	3	2021	7.344	0.480%
4	4	2020	7.309	0.490%

In [61]:
```
li=removing_percent(data5,"Growth_Rate")
```

In [62]:

In [63]:

In [64]:

In [65]:

Out[65]:

	Unnamed: 0	year	Fertility_Rate	Growth_Rate
0	0	2024	2.122	-0.790%
1	1	2023	2.139	-0.930%
2	2	2022	2.159	-0.920%
3	3	2021	2.179	-0.950%
4	4	2020	2.200	-0.900%

In [66]:

In [67]:
```
li=removing_percent(data6,"Growth_Rate")
```

In [68]:

In [69]:

In [70]:

Out[70]:

	Unnamed: 0	year	Infant_Mortality_Rate	Growth_Rate
0	0	2024	25.799	-3.080%
1	1	2023	26.619	-3.890%
2	2	2022	27.695	-3.740%
3	3	2021	28.771	-3.610%
4	4	2020	29.848	-3.480%

In [71]:

In [72]:
```
li=removing_percent(data7,"Growth_Rate")
```

In [73]:

In [74]:

In [75]:

Out[75]:

	Unnamed: 0	year	Life_Expectancy
0	0	2024	70.62
1	1	2023	70.42
2	2	2022	70.19
3	3	2021	69.96
4	4	2020	69.73

In [76]:

In [77]:
```
data_ultimate=pd.merge(data4,data5,on="year")
data_ultimate=pd.merge(data_ultimate,data6,on="year")
data_ultimate=pd.merge(data_ultimate,data7,on="year")
```

In [78]:

Out[78]:

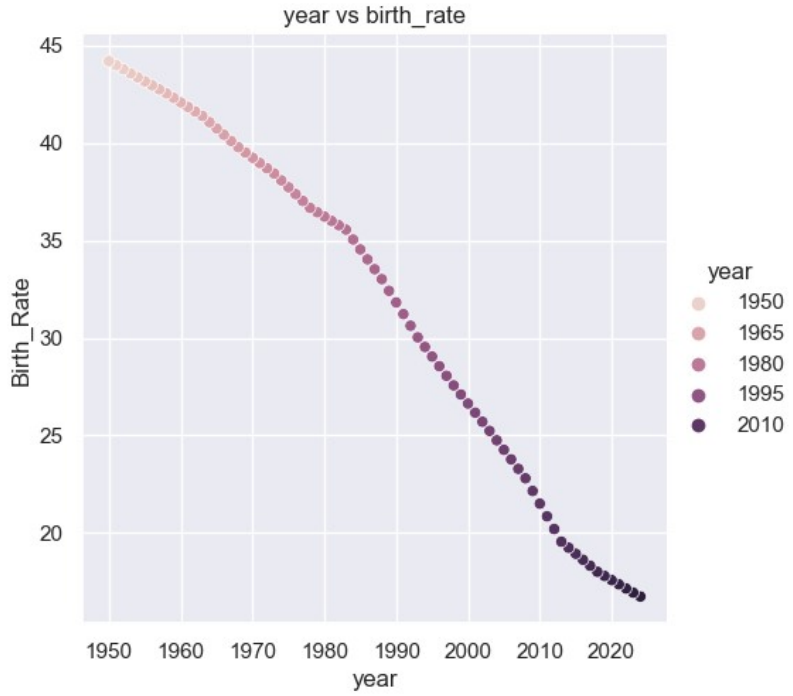
	year	Birth_Rate	birth_rate_growth_rate	Death_Rate	death_rate_growth_rate	Fertility_Rate
0	2024	16.750	-1.17	7.473	0.77	2.122
1	2023	16.949	-1.25	7.416	0.49	2.139
2	2022	17.163	-1.23	7.380	0.49	2.159
3	2021	17.377	-1.22	7.344	0.48	2.179
4	2020	17.592	-1.20	7.309	0.49	2.200

YEAR VS BIRTH RATE

```
In [79]: sbn.relplot(x="year",y="Birth_Rate",kind="scatter",hue="year",data=data_ult
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

```
Out[79]: Text(0.5, 1.0, 'year vs birth_rate')
```

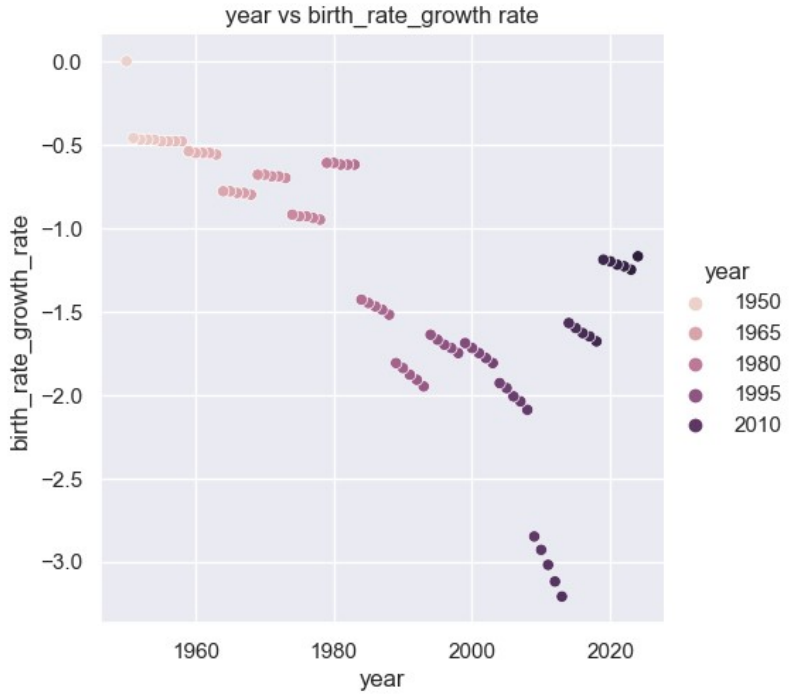


YEAR VS BIRTH RATE GROWTH RATE

```
In [80]: sbn.relplot(x="year",y="birth_rate_growth_rate",kind="scatter",hue="year",da
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

```
Out[80]: Text(0.5, 1.0, 'year vs birth_rate_growth rate')
```

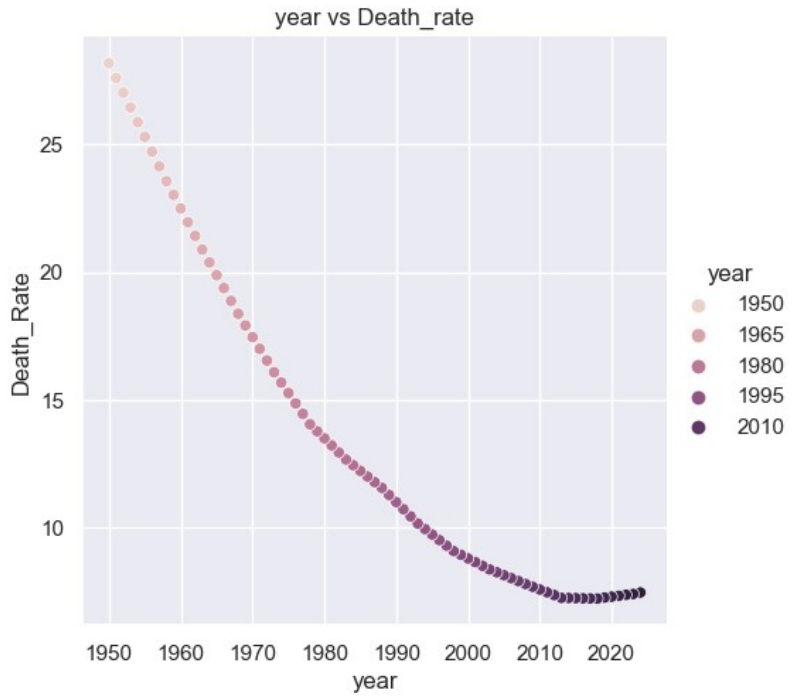


YEAR VS DEATH RATE

In [81]:
```
sbn.relplot(x="year",y="Death_Rate",kind="scatter",hue="year",data=data_ult
plt.title("year vs Death_rate")
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[81]: Text(0.5, 1.0, 'year vs Death_rate')

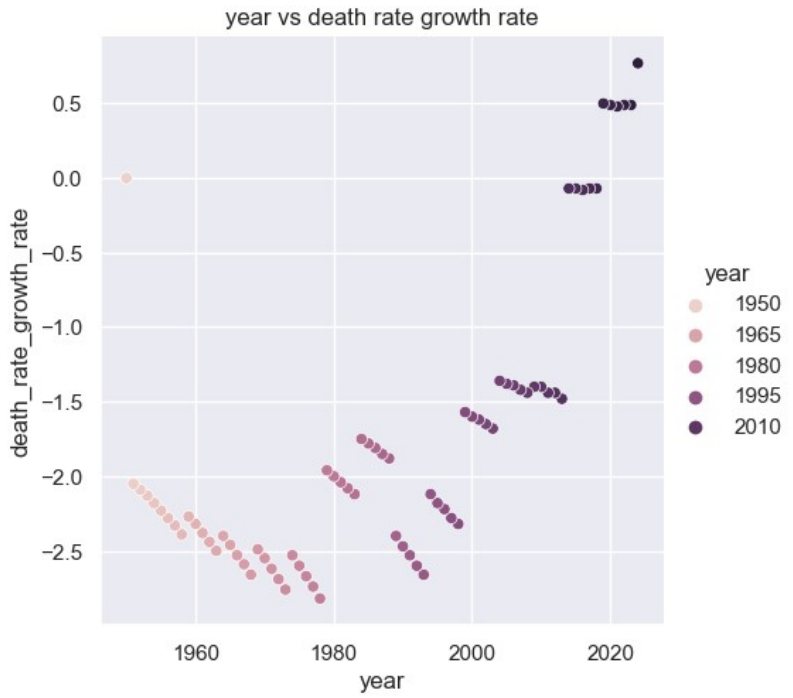


YEAR VS death_rate_growth_rate

In [82]:
```
sbn.relplot(x="year",y="death_rate_growth_rate",kind="scatter",hue="year",da
plt.title("year vs death rate growth rate")
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[82]: Text(0.5, 1.0, 'year vs death rate growth rate')

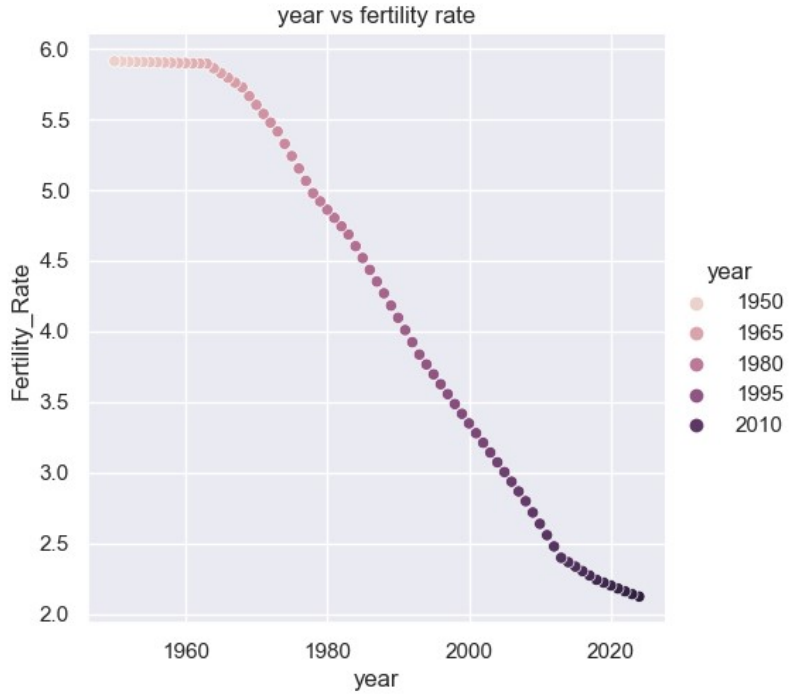


YEAR VS FERTILITY_RATE

In [83]:
```
sbn.relplot(x="year",y="Fertility_Rate",kind="scatter",hue="year",data=data
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[83]: Text(0.5, 1.0, 'year vs fertility rate')



In [84]:
```

```

Out[84]:

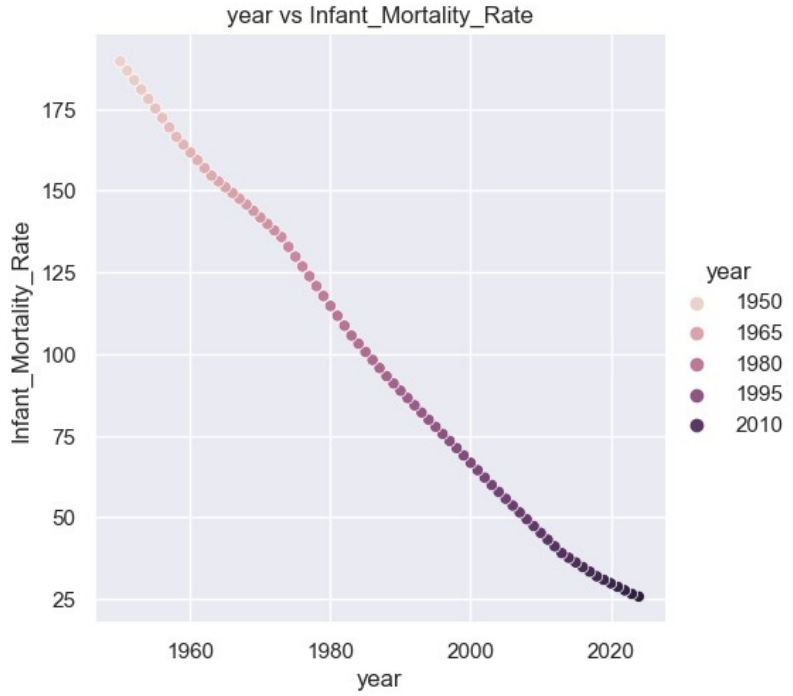
	year	Birth_Rate	birth_rate_growth_rate	Death_Rate	death_rate_growth_rate	Fertility_Rate
0	2024	16.750	-1.17	7.473	0.77	2.122
1	2023	16.949	-1.25	7.416	0.49	2.139
2	2022	17.163	-1.23	7.380	0.49	2.159
3	2021	17.377	-1.22	7.344	0.48	2.179
4	2020	17.592	-1.20	7.309	0.49	2.200

YEAR VS INFANT MORTALITY RATE

In [85]:
```
sbn.relplot(x="year",y="Infant_Mortality_Rate",kind="scatter",hue="year",da
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[85]: Text(0.5, 1.0, 'year vs Infant_Mortality_Rate')

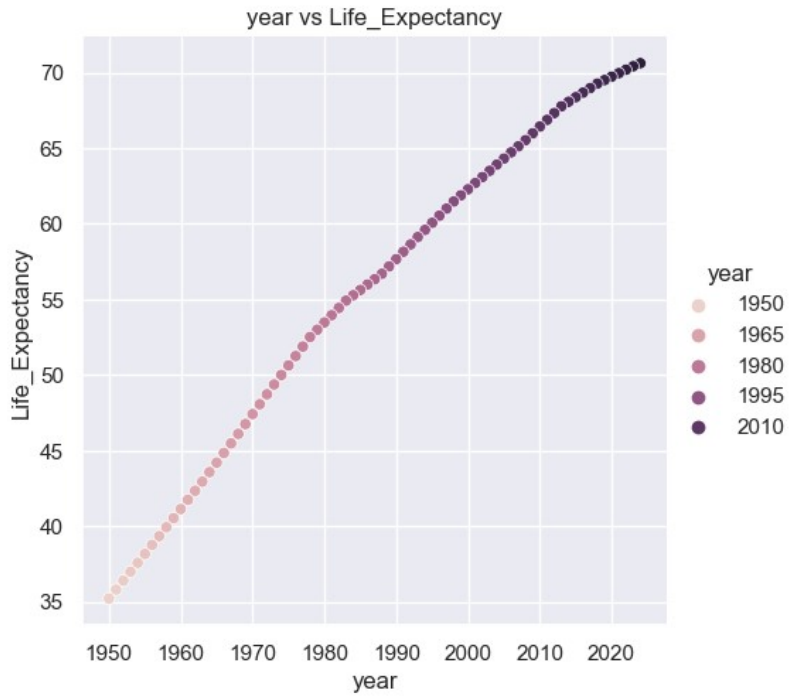


YEAR VS Life_Expectancy

In [86]:
```
sbn.relplot(x="year",y="Life_Expectancy",kind="scatter",hue="year",data=dat
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[86]:
```
Text(0.5, 1.0, 'year vs Life_Expectancy')
```

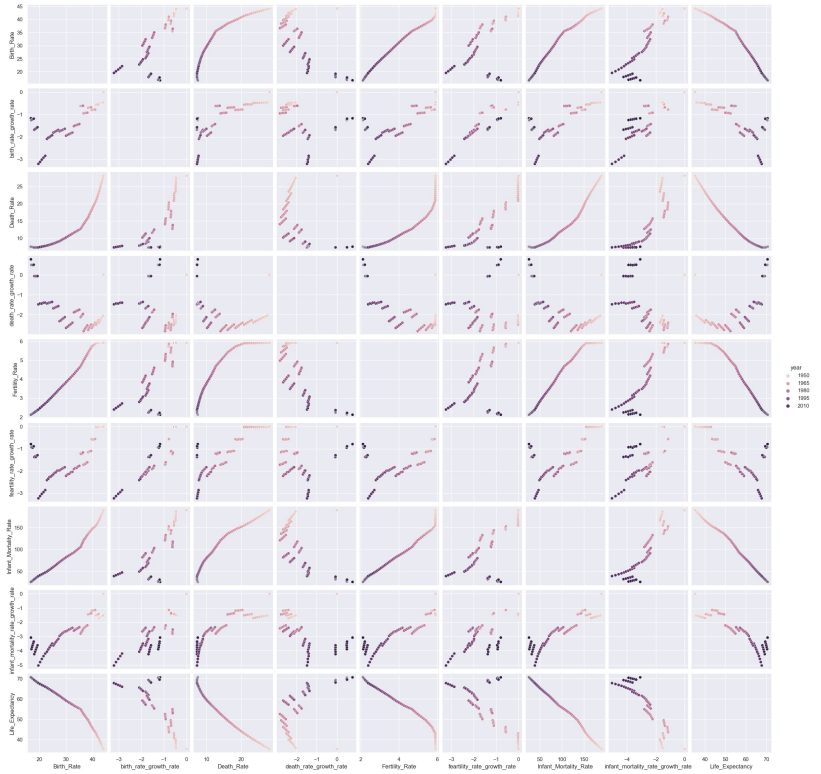


PAIR PLOT FOR ULTIMATE DATA

In [87]:
```
```

```
C:\Users\win10\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWa
rning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[87]:
```
<seaborn.axisgrid.PairGrid at 0x24530268e90>
```



In [ ]:
```
```