▼ Data Visualization

Project 1

Rhichard Koh

```
import pandas as pd
import plotly.express as px

income = pd.read_csv("gdp_pcap.csv")
babiesPerWoman = pd.read_csv("children_per_woman_total_fertility.csv")
population = pd.read_csv("pop.csv")
literacyRate = pd.read_csv("literacy_rate_adult.csv")
lifeExpectancy = pd.read_csv("lex.csv")
fertilityRate = pd.read_csv("sp_dyn_tfrt_in.csv")
```

income

	country	1800	1801	1802	1803	1804	1805	1806	1807	1808	 2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
0	Afghanistan	583	583	583	583	583	583	583	583	583	 3570	3700	3720	3830	3930	4030	4130	3900	4070	4030
1	Angola	452	454	456	458	460	462	464	466	468	 20.6k	21.1k	21.7k	21.9k	22.3k	22.5k	23k	23.5k	24k	24.6k
2	Albania	569	571	572	574	576	577	579	581	582	 37.6k	38.5k	39.4k	39.9k	40.9k	41.9k	41.4k	42.4k	43.7k	45.7k
3	Andorra	1660	1660	1670	1670	1670	1680	1680	1680	1690	 59.2k	60k	60.3k	61k	61.7k	62.7k	61.9k	62.5k	62.9k	62.1k
4	UAE	1390	1390	1390	1400	1410	1410	1420	1420	1430	 86.2k	85.4k	84.1k	84.3k	84.6k	83.9k	84.2k	84k	84.1k	84.3k
190	Samoa	1940	1940	1940	1940	1940	1940	1940	1950	1950	 12.6k	12.9k	13.3k	13.6k	13.9k	14.3k	14.7k	14.7k	15.6k	16k
191	Yemen	1220	1230	1230	1230	1240	1240	1240	1250	1250	 4680	4620	4680	4880	5060	5260	5370	5490	5610	5730
192	South Africa	2120	2100	2070	2050	2030	1920	1920	2080	1820	 36.5k	37.3k	38.1k	39k	39.6k	40.2k	40.9k	41.6k	42.1k	42.7k
193	Zambia	646	648	650	651	653	654	656	658	659	 10.6k	10.9k	10.8k	11.1k	11.4k	11.7k	12k	12.3k	12.5k	12.8k
194	Zimbabwe	1170	1170	1170	1170	1170	1170	1180	1180	1180	 5970	6190	6360	6570	6600	6750	6890	7080	7200	7380

195 rows × 302 columns

babiesPerWoman

	country	1800	1801	1802	1803	1804	1805	1806	1807	1808	 2091	2092	2093	2094	2095	2096	2097	2098	2099	21(
0	Aruba	5.64	5.64	5.64	5.64	5.64	5.64	5.64	5.64	5.64	 1.82	1.82	1.82	1.82	1.82	1.82	1.82	1.82	1.83	1.8
1	Afghanistan	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	 1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.7
2	Angola	6.93	6.93	6.93	6.93	6.93	6.93	6.93	6.94	6.94	 2.54	2.52	2.50	2.48	2.47	2.45	2.43	2.42	2.40	2.4
3	Albania	4.60	4.60	4.60	4.60	4.60	4.60	4.60	4.60	4.60	 1.78	1.78	1.78	1.79	1.79	1.79	1.79	1.79	1.79	1.7
4	Netherlands Antilles	5.80	5.80	5.80	5.80	5.80	5.80	5.80	5.80	5.80	 2.00	2.00	2.01	2.01	2.01	2.01	2.01	2.02	2.02	2.0
197	Samoa	6.98	6.98	6.98	6.98	6.98	6.98	6.98	6.98	6.98	 2.10	2.09	2.08	2.07	2.06	2.05	2.04	2.03	2.02	2.0
198	Yemen	6.88	6.88	6.88	6.88	6.88	6.88	6.88	6.88	6.88	 1.68	1.68	1.69	1.69	1.69	1.69	1.70	1.70	1.70	1.7
199	South Africa	6.47	6.47	6.47	6.47	6.47	6.47	6.47	6.47	6.47	 1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.8
200	Zambia	6.71	6.71	6.71	6.71	6.71	6.71	6.71	6.71	6.71	 2.59	2.57	2.56	2.54	2.53	2.51	2.50	2.48	2.46	2.4
201	Zimbabwe	6.75	6.75	6.75	6.75	6.75	6.75	6.75	6.75	6.75	 1.85	1.85	1.85	1.84	1.84	1.84	1.83	1.83	1.83	1.8

202 rows × 302 columns

population

	country	1800	1801	1802	1803	1804	1805	1806	1807	1808	 2091	2092	2093	2094	2095	2096	2097	209
0	Afghanistan	3.28M	 108M	108M	109M	109M	109M	110M	110M	110								
1	Angola	1.57M	 125M	126M	127M	128M	129M	130M	131M	131								
2	Albania	400k	402k	404k	405k	407k	409k	411k	413k	414k	 1.35M	1.32M	1.29M	1.26M	1.23M	1.21M	1.18M	1.15
3	Andorra	2650	2650	2650	2650	2650	2650	2650	2650	2650	 62.5k	62.2k	61.9k	61.7k	61.4k	61.2k	60.9k	60.
4	UAE	40.2k	 13.5M	13.5M	13.6M	13.7M	13.8M	13.8M	13.9M	14								
192	Samoa	47.3k	47.2k	47.2k	 432k	434k	435k	436k	437k	438k	439k	441						
193	Yemen	2.59M	 73.2M	73.3M	73.5M	73.6M	73.7M	73.8M	73.9M	74								
194	South Africa	1.45M	1.45M	1.46M	1.46M	1.47M	1.47M	1.48M	1.49M	1.49M	 75.6M	75.5M	75.4M	75.3M	75.2M	75.1M	75M	74.8
195	Zambia	747k	758k	770k	782k	794k	806k	818k	831k	843k	 60.3M	60.7M	61.1M	61.5M	61.9M	62.2M	62.6M	62.9
400	i i	4 0014	4 0014	4 0014	4 0014	4 0014	4 0014	4 0014	4 0014	4 0014	05 484	05 014	05 014	05 444	05 484	05 584	^	^r 3

literacyRate

	country	1975	1976	1977	1978	1979	1980	1981	1982	1983	 2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Ē
0	Aruba	NaN	 NaN	NaN	96.8	NaN	ī														
1	Afghanistan	NaN	NaN	NaN	NaN	18.2	NaN	NaN	NaN	NaN	 NaN	NaN	39.0								
2	Angola	NaN	 NaN	NaN	70.4																
3	Anguilla	NaN	 NaN	NaN																	
4	Albania	NaN	 NaN	NaN	NaN	NaN	NaN	NaN	95.9	NaN	NaN	96.8									
151	Samoa	NaN	 NaN	NaN	98.5	NaN	NaN	NaN	NaN	NaN	NaN	98.8									
152	Yemen	NaN	 NaN	NaN	54.8	NaN	NaN	NaN	NaN	NaN	NaN	65.3									
153	South Africa	NaN	NaN	NaN	NaN	NaN	76.2	NaN	NaN	NaN	 NaN	NaN	NaN	NaN	NaN	88.7	NaN	NaN	NaN	93.0	
154	Zambia	NaN	 69.2	NaN	NaN	NaN	NaN	61.4	NaN	NaN	NaN	NaN									
155	Zimbabwe	NaN	77.8	NaN	 NaN	NaN	83.6														

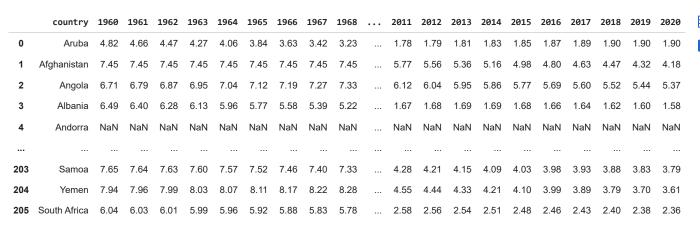
lifeExpectancy

156 rows × 38 columns

	country	1000	1001	1002	1002	1004	1005	1000	1007	1000		2001	2002	2002	2004	2005	2006	2007	2000	2099	2100	F
	country	1900	1991	1002	1803	1004	1902	1900	1007	1000	•••	2091	2092	2093	2094	2095	2090	2097	2098	2099	2100	н
0	Afghanistan	28.2	28.2	28.2	28.2	28.2	28.2	28.1	28.1	28.1		75.5	75.7	75.8	76.0	76.1	76.2	76.4	76.5	76.6	76.8	1
1	Angola	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0	27.0		78.8	79.0	79.1	79.2	79.3	79.5	79.6	79.7	79.9	80.0	
2	Albania	35.4	35.4	35.4	35.4	35.4	35.4	35.4	35.4	35.4		87.4	87.5	87.6	87.7	87.8	87.9	88.0	88.2	88.3	88.4	
3	Andorra	NaN		NaN																		
4	UAE	30.7	30.7	30.7	30.7	30.7	30.7	30.7	30.7	30.7		82.4	82.5	82.6	82.7	82.8	82.9	83.0	83.1	83.2	83.3	
190	Samoa	25.4	25.4	25.4	25.4	25.4	25.4	25.4	25.4	25.4		79.8	79.9	80.0	80.1	80.3	80.4	80.5	80.6	80.7	80.8	
191	Yemen	23.4	23.4	23.4	23.4	23.4	23.4	23.4	23.4	23.4		76.9	77.0	77.1	77.3	77.4	77.5	77.6	77.8	77.9	78.0	
192	South Africa	33.5	33.5	33.5	33.5	33.5	33.5	33.5	33.5	33.5		76.4	76.5	76.7	76.8	77.0	77.1	77.3	77.4	77.5	77.7	
193	Zambia	32.6	32.6	32.6	32.6	32.6	32.6	32.6	32.6	32.6		75.8	76.0	76.1	76.3	76.4	76.5	76.7	76.8	77.0	77.1	
194	Zimbabwe	33.7	33.7	33.7	33.7	33.7	33.7	33.7	33.7	33.7		73.3	73.4	73.5	73.7	73.8	73.9	74.0	74.2	74.3	74.4	

195 rows × 302 columns

fertilityRate



income = income.melt(id_vars=['country'], var_name='Year', value_name='income')
babiesPerWoman = babiesPerWoman.melt(id_vars=['country'], var_name='Year', value_name='babiesPerWoman')
population = population.melt(id_vars=['country'], var_name='Year', value_name='population')
literacyRate = literacyRate.melt(id_vars=['country'], var_name='Year', value_name='literacyRate')
lifeExpectancy = lifeExpectancy.melt(id_vars=['country'], var_name='Year', value_name='lifeExpectancy')
fertilityRate = fertilityRate.melt(id_vars=['country'], var_name='Year', value_name='fertilityRate')

income = income.sort_values(by=['country','Year']).reset_index(drop=True)
babiesPerWoman = babiesPerWoman.sort_values(by=['country','Year']).reset_index(drop=True)
population = population.sort_values(by=['country','Year']).reset_index(drop=True)
literacyRate = literacyRate.sort_values(by=['country','Year']).reset_index(drop=True)
lifeExpectancy = lifeExpectancy.sort_values(by=['country','Year']).reset_index(drop=True)
fertilityRate = fertilityRate.sort_values(by=['country','Year']).reset_index(drop=True)

income

	country	Year	income	
0	Afghanistan	1800	583	ıl.
1	Afghanistan	1801	583	
2	Afghanistan	1802	583	
3	Afghanistan	1803	583	
4	Afghanistan	1804	583	
58690	Zimbabwe	2096	6750	
58691	Zimbabwe	2097	6890	
58692	Zimbabwe	2098	7080	
58693	Zimbabwe	2099	7200	
58694	Zimbabwe	2100	7380	

58695 rows × 3 columns

babiesPerWoman



population

	country	Year	population	-
0	Afghanistan	1800	3.28M	ıl.
1	Afghanistan	1801	3.28M	
2	Afghanistan	1802	3.28M	
3	Afghanistan	1803	3.28M	
4	Afghanistan	1804	3.28M	
59292	Zimbabwe	2096	35.5M	
59293	Zimbabwe	2097	35.5M	
59294	Zimbabwe	2098	35.6M	
59295	Zimbabwe	2099	35.6M	
59296	Zimbabwe	2100	35.6M	

literacyRate

59297 rows × 3 columns

	country	Year	literacyRate	
0	Afghanistan	1975	NaN	ılı
1	Afghanistan	1976	NaN	
2	Afghanistan	1977	NaN	
3	Afghanistan	1978	NaN	
4	Afghanistan	1979	18.2	
5767	Zimbabwe	2007	NaN	
5768	Zimbabwe	2008	NaN	
5769	Zimbabwe	2009	NaN	
5770	Zimbabwe	2010	NaN	
5771	Zimbabwe	2011	83.6	
5772 rc	ws × 3 column	าร		

lifeExpectancy

		country	Year	lifeExpectancy	⊞
	٥	Afahanistan	1200	၁ 2 ၁	-
ferti	lityRat	e			
		country	Year	fertilityRate	
	0	Afghanistan	1960	7.45	11.
	1	Afghanistan	1961	7.45	_
	2	Afghanistan	1962	7.45	
	3	Afghanistan	1963	7.45	
	4	Afghanistan	1964	7.45	
	12683	Zimbabwe	2016	3.80	
	12684	Zimbabwe	2017	3.71	
	12685	Zimbabwe	2018	3.62	
	12686	Zimbabwe	2019	3.53	
	12687	Zimbabwe	2020	3.46	
	12688 rd	ows × 3 columi	าร		
				esPerWoman, on=[lation, on=['cou	'country','Year'])
merge	d = pd.	merge(merge	d,life	Expectancy, on=['country','Year'])
					<pre>country','Year']) ountry','Year'])</pre>
J	·	0 . 0		,	
		ontinent =			
	-	stan': 'Asi ': 'Europe'			
	Algeria	': 'Africa'			
	-	: 'Africa', and Barbuda	a': 'N	orth America',	
'	Argenti	na': 'South			
		ı': 'Asia', .jan': 'Asia	٠,		
'	Bahrain	': 'Asia',			
	_	lesh': 'Asia ': 'Europe'			
'	Belize'	: 'North Ame		,	
		'Africa', : 'Asia',			
		': 'South A	merica	١,	
		and Herzegova': 'Africa		'Europe',	
		: 'South Ame	•	,	
		: 'Asia', a': 'Europe			
	_	Faso': 'Af	-		
		': 'Africa' a': 'Asia',	,		
		n': 'Africa	٠,		
		rde': 'Afri African Re	-	': 'Africa',	
		'Africa',	Judiic	. All Ica ,	
		'South Ame	rica',		
		a': 'South A	Americ	a',	
		': 'Africa' Dem. Rep.':		ca'	
		ica': 'Nort			
		Ivoire": 'A		,	
		'North Amer			
		: 'Europe',	' · 'No	rth America',	
		': 'South A		-	
		'Africa', ador': 'Nor	th Ame	rica'.	
				,	

```
'Equatorial Guinea': 'Africa',
'Eritrea': 'Africa',
'Estonia': 'Europe',
'Eswatini': 'Africa',
'Ethiopia': 'Africa',
'Gabon': 'Africa',
'Gambia': 'Africa',
'Georgia': 'Asia',
'Ghana': 'Africa',
'Greece': 'Europe',
'Guatemala': 'North America',
'Guinea': 'Africa',
'Guinea-Bissau': 'Africa',
'Guyana': 'South America',
'Haiti': 'North America',
'Honduras': 'North America',
'Hungary': 'Europe',
'India': 'Asia',
'Indonesia': 'Asia',
'Iran': 'Asia',
'Iraq': 'Asia',
'Israel': 'Asia',
'Italy': 'Europe',
'Jamaica': 'North America',
'Jordan': 'Asia',
'Kazakhstan': 'Asia',
'Kenya': 'Africa',
'Kuwait': 'Asia',
'Kyrgyz Republic': 'Asia',
'Lao': 'Asia',
'Latvia': 'Europe',
'Lebanon': 'Asia',
'Lesotho': 'Africa',
'Liberia': 'Africa',
'Libya': 'Africa',
'Lithuania': 'Europe',
'Madagascar': 'Africa',
'Malawi': 'Africa',
'Malaysia': 'Asia',
'Maldives': 'Asia',
'Mali': 'Africa',
'Malta': 'Europe',
'Mauritania': 'Africa',
'Mauritius': 'Africa',
'Mexico': 'North America',
'Moldova': 'Europe',
'Mongolia': 'Asia',
'Montenegro': 'Europe',
'Morocco': 'Africa',
'Mozambique': 'Africa',
'Myanmar': 'Asia',
'Namibia': 'Africa',
'Nepal': 'Asia',
'Nicaragua': 'North America',
'Niger': 'Africa',
'Nigeria': 'Africa',
'North Korea': 'Asia',
'North Macedonia': 'Europe',
'Oman': 'Asia',
'Pakistan': 'Asia',
'Palestine': 'Asia',
'Panama': 'North America',
'Papua New Guinea': 'Oceania',
'Paraguay': 'South America',
'Peru': 'South America',
'Philippines': 'Asia',
'Poland': 'Europe',
'Portugal': 'Europe',
'Oatar': 'Asia',
'Romania': 'Europe',
'Russia': 'Europe',
'Rwanda': 'Africa',
'Samoa': 'Oceania',
'Sao Tome and Principe': 'Africa',
'Saudi Arabia': 'Asia',
'Senegal': 'Africa',
'Serbia': 'Europe',
```

```
'Seychelles': 'Africa',
    'Sierra Leone': 'Africa',
    'Singapore': 'Asia',
    'Slovenia': 'Europe',
    'South Africa': 'Africa',
    'Spain': 'Europe',
    'Sri Lanka': 'Asia',
    'Sudan': 'Africa',
    'Suriname': 'South America',
    'Syria': 'Asia',
    'Tajikistan': 'Asia',
    'Tanzania': 'Africa',
    'Thailand': 'Asia',
    'Timor-Leste': 'Asia'.
    'Togo': 'Africa',
    'Tonga': 'Oceania',
    'Trinidad and Tobago': 'North America',
    'Tunisia': 'Africa',
    'Turkey': 'Asia',
    'Turkmenistan': 'Asia',
    'UAE': 'Asia',
    'Uganda': 'Africa',
    'Ukraine': 'Europe',
    'Uruguay': 'South America',
    'Uzbekistan': 'Asia',
    'Vanuatu': 'Oceania',
    'Venezuela': 'South America',
    'Vietnam': 'Asia',
    'Yemen': 'Asia',
    'Zambia': 'Africa',
    'Zimbabwe': 'Africa',
}
merged['continent'] = merged['country'].map(country_to_continent)
def convertMk(string):
 if string[-1] == 'M':
   string = string.replace("M", "")
   string = float(string)
 string = string * 1000000
elif string[-1] == 'k':
   string = string.replace("k", "")
   string = float(string)
 string = string * 1000
elif string[-1] == 'B':
   string = string.replace("B", "")
   string = float(string)
   string = string * 1000000000
 return float(string)
merged['population'] = merged['population'].astype(str).apply(convertMk)
merged
```

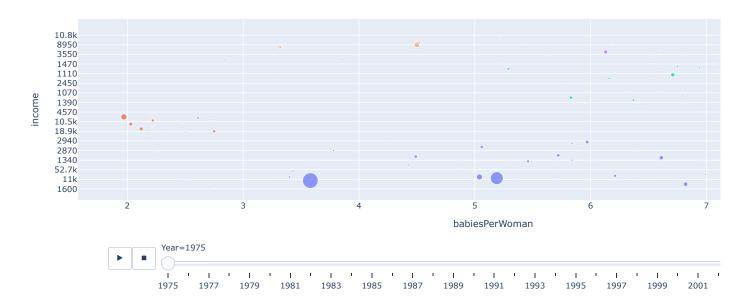
https://colab.research.google.com/drive/1kUmMPUwZFNVqiiGJC9JOHA0tzf9BIK3I#scrollTo=8NqDZe0r5ILY&printMode=true

country Year income babiesPerWoman population lifeExpectancy fertilityRate literacyRate continent

title="Decrease of Babies per Woman when Income Increases")

fig1.write_html("Decrease of Babies per Woman when Income Increases.html")
fig1.show()

Decrease of Babies per Woman when Income Increases



Decrease of Babies per Woman when Life Expectancy Increases

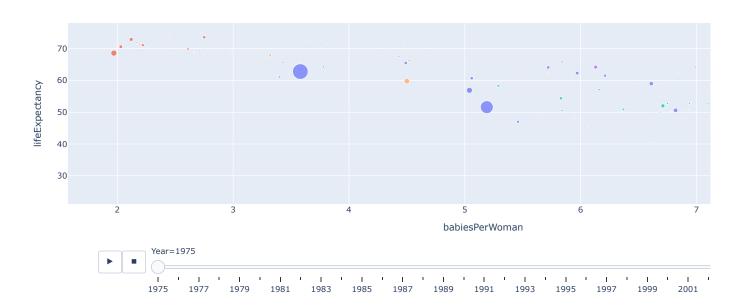
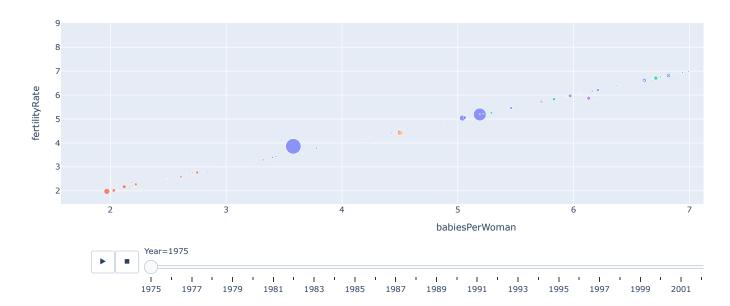
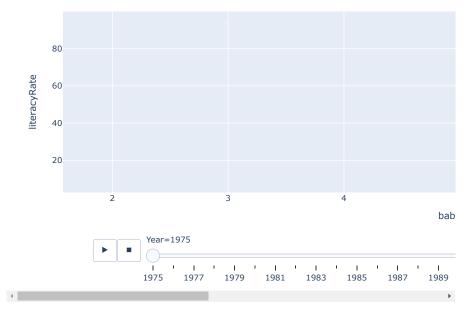


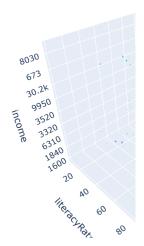
fig3.write_html("Positive Correlation of Babies per Woman and Fertility Rate.html") fig3.show()

Positive Correlation of Babies per Woman and Fertility Rate

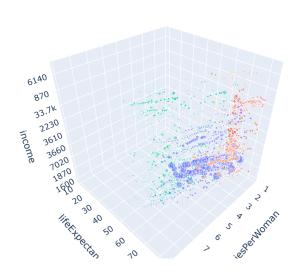


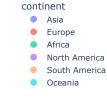
Decrease of Babies per Woman when Literacy Rate Increases





We can observe that literacy rate influences the number of babies per woman when compared to income.





We can see that the higher the life expectancy and income the less babies are created per woman.