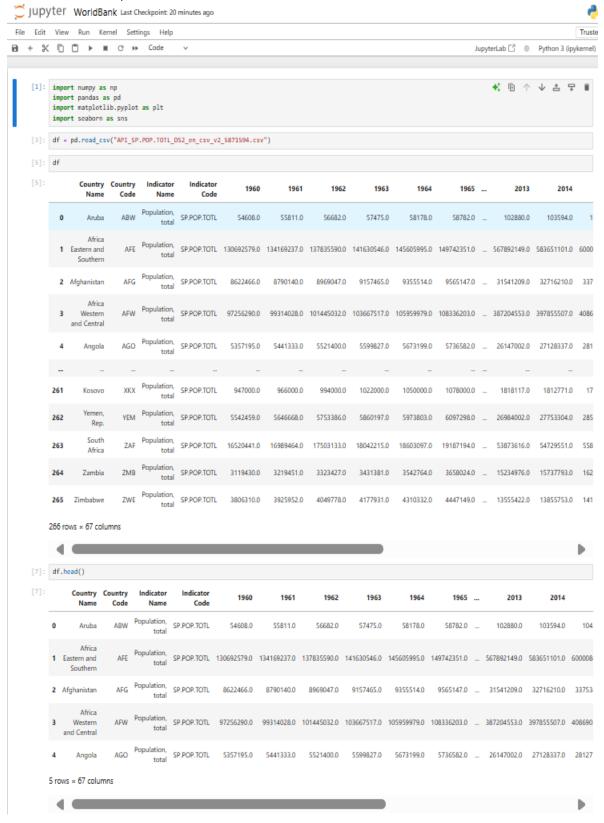
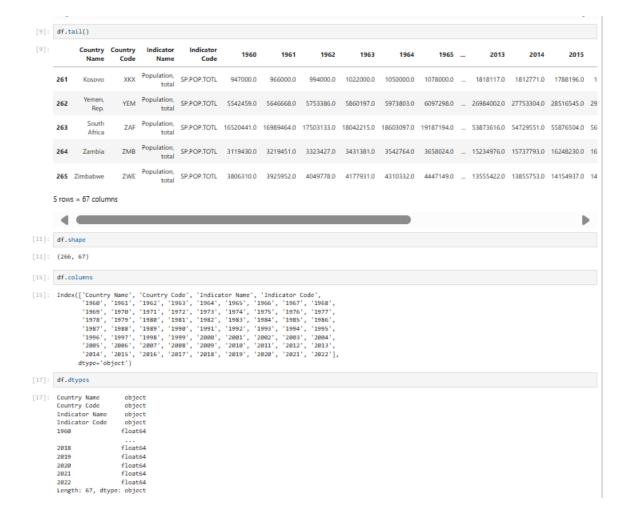
PRODIGY INFOTECH

PRODIGY_DS_01

ANKIT BHARDWAJ | DATA SCIENCE INTERN





[19]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 266 entries, 0 to 265
Data columns (total 67 columns):
Columns

	eindex: 266 entr		
	columns (total (67 columns): Non-Null Count	Dtype
0	Country Name	266 non-null	object
1		266 non-null	object
2	Indicator Name	266 non-null	object
3		266 non-null	object
4	1960	264 non-null	float64
5	1961	264 non-null	float64
6	1962	264 non-null	float64
7	1963	264 non-null	float64
8	1964	264 non-null	float64
9	1965	264 non-null	float64
	1966	264 non-null	float64
11	1967	264 non-null	float64
	1968	264 non-null	float64
	1969	264 non-null	float64
	1970	264 non-null	float64
15		264 non-null	float64
	1972	264 non-null	float64
	1973	264 non-null	float64
18	1974	264 non-null	float64
	1975	264 non-null	float64
	1976	264 non-null	float64
21	1977	264 non-null	float64
22	1977 1978	264 non-null	float64
23	1979	264 non-null	float64
24	1980	264 non-null	float64
25	1981	264 non-null	float64
26	1982	264 non-null	float64
27	1983	264 non-null	float64
28	1984	264 non-null	float64
29	1985	264 non-null	float64
38	1986	264 non-null	float64
31	1987	264 non-null	float64
32	1988	264 non-null	float64
33	1989	264 non-null	float64
34	1998	265 non-null 265 non-null	float64 float64
35 36	1991 1992	265 non-null	float64
37	1993	265 non-null	float64
38	1994	265 non-null	float64
39		265 non-null	float64
48	1996	265 non-null	float64
41	1997	265 non-null	float64
	1998	265 non-null	float64
	1999	265 non-null	float64
44	2000	265 non-null	float64
	2001	265 non-null	float64
46	2882	265 non-null	float64
47	2883	265 non-null	float64
48	2884	265 non-null	float64
49	2885	265 non-null	float64
	2886	265 non-null	float64
51	2007	265 non-null	float64
52	2008	265 non-null	float64
53		265 non-null	float64
54	2010	265 non-null	float64
55 56	2011 2012	265 non-null 265 non-null	float64 float64
56	2012	265 non-null	float64 float64
58	2013	265 non-null	float64
59	2014	265 non-null	float64
68	2015	265 non-null	float64
61	2016	265 non-null	float64
62	2017	265 non-null	float64
63	2019	265 non-null	float64
64	2020	265 non-null	float64
65	2021	265 non-null	float64
66	2022	265 non-null	float64
dtyp	es: float64(63), ry usage: 139.4+	object(4)	
memo	ry usage: 139.4+	KB	

[21]: df.describe() 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 ... 20 $\textbf{count} \quad 2.640000e + 02 \quad 2.640000e$ mean 1.172712e+08 1.188807e+08 1.210511e+08 1.237333e+08 1.264378e+08 1.291813e+08 1.320404e+08 1.348980e+08 1.378358e+08 1.408789e+08 ... 2.927787e+ std 3.695439e+08 3.740897e+08 3.808061e+08 3.895039e+08 3.982439e+08 4.071153e+08 4.164504e+08 4.257424e+08 4.353218e+08 4.452927e+08 ... 9.186849e+ min 2.646000e+03 2.888000e+03 3.171000e+03 3.481000e+03 3.811000e+03 4.161000e+03 4.531000e+03 4.930000e+03 5.354000e+03 5.354000e+03 ... 1.069400e+ 25% 5.132212e+05 5.231345e+05 5.337595e+05 5.449288e+05 5.566630e+05 5.651150e+05 5.691470e+05 5.773872e+05 5.832700e+05 5.875942e+05 ... 1.697753e+ 50% 3.757486e+06 3.887144e+06 4.023896e+06 4.139356e+06 4.224612e+06 4.277636e+06 4.331825e+06 4.385700e+06 4.450934e+06 4.530800e+06 ... 1.014958e+ 75% 2.670606e+07 2.748694e+07 2.830289e+07 2.914708e+07 3.001684e+07 3.084892e+07 3.163010e+07 3.29247e+07 3.249927e+07 3.247149e+07 ... 6.023395e+ max 3.031474e+09 3.072422e+09 3.126850e+09 3.193429e+09 3.260442e+09 3.328209e+09 3.398480e+09 3.468371e+09 3.540164e+09 3.614573e+09 ... 7.229732e+ 8 rows × 63 columns 1 [23]: df.duplicated().sum() [23]: 0 [25]: df.isna().sum().any() [25]: True [27]: df = df.fillna(method = "ffill") df.head() C:\Users\Ankit Bhardwaj\AppData\Local\Temp\ipykernel_35556\3568191538.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead. df - df.fillna(method - "ffill") Country Country Indicator Name Code Name 1960 1961 1962 1963 1964 1965 ... 2013 2014 ABW Population, SP.POP.TOTL Aruba 54608.0 55811.0 56682.0 57475.0 58178.0 58782.0 ... 102880.0 103594.0 Africa AFE Population, total SP.POP.TOTL 130692579.0 134169237.0 137835590.0 141630546.0 145605995.0 149742351.0 ... 567892149.0 583651101.0 600008 1 Eastern and Southern AFG Population, sp.POP.TOTL 8622466.0 8790140.0 8969047.0 9157465.0 9355514.0 9565147.0 ... 31541209.0 32716210.0 33753 2 Afghanistan Africa AFW Population, total SP.POP.TOTL 97256290.0 99314028.0 101445032.0 103667517.0 105959979.0 108336203.0 ... 387204553.0 397855507.0 408690 and Central AGO Population, total SP.POP.TOTL 5357195.0 5441333.0 5521400.0 5599827.0 5673199.0 5736582.0 .. 26147002.0 27128337.0 28127 Angola 5 rows × 67 columns

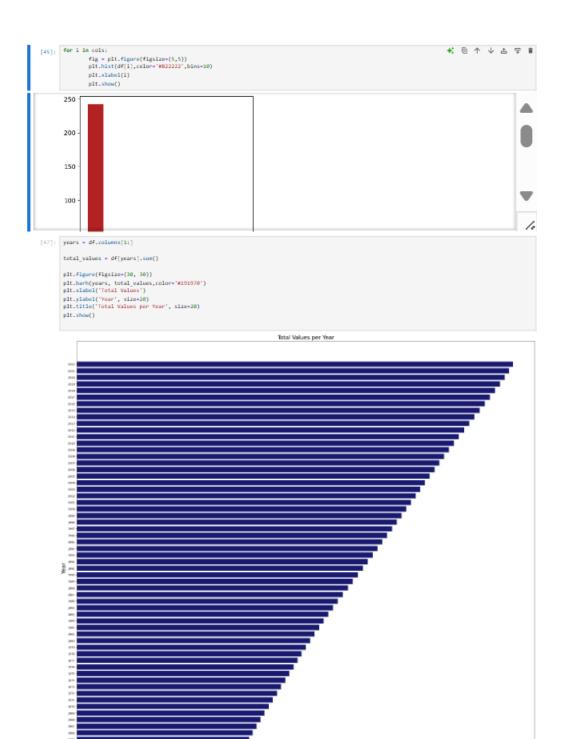
```
[29]: df.isna().sum().any()
```

```
[31]: df['Country Name'].unique()
```

```
[31]: df['Country Name'].unique()

[31]: array(['Aruba', 'Africa Eastern and Southern', 'Afghanistan', 'Africa Western and Central', 'Angola', 'Albania', 'Andorra', 'Arab World', 'United Arab Emirates', 'Argentina', 'Arrenia', 'Arrenia', 'American Samoa', 'Antigua and Barbuda', 'Australia', 'Australia', 'Australia', 'Australia', 'Arcebaijan', 'Bulgaria', 'Bahsanin', 'Bahamas, The', 'Bosnia and Herzegovina', 'Belarus', 'Belize', 'Bermuda', 'Bolivia', 'Brazil', 'Barbados', 'Brunei Darussalam', 'Bhutan', 'Botswana', 'Central African Republic', 'Cameroon', 'Congo, Dem. Rep.', 'Congo, Rep.', 'Colobbia', 'Comeros', 'Compo, Dem. Rep.', 'Congo, Rep.', 'Colobbia', 'Generos', 'Cabe Vorde', 'Costa Rica', 'Caribbean small states', 'Cuba', 'Curacao', 'Cayman Islands', 'Cyprus', 'Czechia', 'Germany', 'Djibouti', 'Dominican Republic', 'Algeria', 'Esta Asia & Pacific', 'Europe & Central Asia (excluding high income)', 'Early-demographic dividend', 'Esta Asia & Pacific', 'Europe & Central Asia', 'Ecuador', 'Egypt, Arab Rep.', 'Europe & Central Asia', 'Ecuador', 'Egypt, Arab Rep.', 'Europe area', 'Eritre', 'Spani', 'Estonia', 'Ethiopia', 'European', 'France', 'Faroe Islands', 'Micronesia, Fed. Stz.', 'Gabon', 'United Kingdom', 'Georgia', 'Ghana', 'Gibraltar', 'Guinca', 'Gambia, The', 'Guinca-Bissau', 'Equatorial Guinca', 'Greecel And', 'Guinca-Bissau', 'Equatorial Guinca', 'Greecel, 'Greenda', 'Greenland', 'Guatemala', 'Guam', 'Guyana', 'Hagh income', 'Hong Kong SAR, China', 'Honduras', 'Heavily indobted poor countries (HIPC)', 'Croatia', 'Haiti', 'Hungary', 'IBRO only', 'IDA & ElsBo total', 'JDA kotal', 'Hati', 'Hungary', 'IBRO only', 'IDA & ElsBo total', 'JDA kotal', 'Hati', 'Hungary', 'JBRO only', 'JDA & ElsBo total', 'JDA kotal', 'Hangary', 'JBRO only', 'JDA & ElsBo total', 'JDA kotal',
                                                                                                                                                                                                                                                         'Guam', 'Guyana', 'High income', 'Hong Kong SAR, China', 'Hondursi', 'Heavily indebted poor countries (HFC)', 'Croatia', 'Haiti', 'Hungary', 'IBRD only', 'IDA & IBRD total', 'IDA total', 'IDA blend', 'Indonesia', 'IDA only', 'Isle of Man', 'India', 'Not classified', 'Ireland', 'Iran, Islamic Rep.', 'Iraq', 'Iccland', 'Israel', 'Italy', 'Jamaica', 'Jordan', 'Japan', 'Kazakhstan', 'Kenya', 'Kyrgyz Republic', 'Cambodia', 'Kiribati', 'St. Kitts and Nevis', 'Korea, Rep.', 'Kuwait', 'Latin America & Caribbean (cxcluding high income)', 'Lao PDR', 'Ichard', 'Liberia', 'Librati', 'Libr
                                                                                                                                                                                                                                                              'Lebanon', 'Liberia', 'Libya', 'St. Lucia',
'Latin America & Caribbean',
'Least developed countries: UN classification', 'Low income',
'Liechtenstein', 'Sri Lanka', 'Lower middle income',
                                                                                                                                                                                                                                                Least developed countries: UN classification', 'Low income',
'liechtenstein', 'Sri Lanka', 'Lower middle income',
'Low & middle income', 'Lesotho', 'Late-demographic dividend',
'Lithuania', 'Luxembourg', 'Latvia', 'Macao SAR, China',
'St. Martin (French part)', 'Morocco', 'Monacoo', 'Moldova',
'Madagascar', 'Maldives', 'Middle East & North Africa', 'Mexico',
'Marshall Islands', 'Middle income', 'North Macedonia', 'Mali',
'Malta', 'Myanmar',
'Middle East & North Africa (cxcluding high income)', 'Montenegro',
'Mongolia', 'Northern Mariana Islands', 'Mozambique', 'Mauritania',
'Mow Caledonia', 'Miger', 'Nigeria', 'Nicaragua', 'Netherlands',
'Now Caledonia', 'Niger', 'Nigeria', 'Nicaragua', 'Netherlands',
'Norway', 'Nopal', 'Mauru', 'New Caland', 'GeCD members', 'Oman',
'Other small states', 'Pakistan', 'Panama', 'Peru', 'Philippines',
'Palau', 'Papua New Guinea', 'Feland', 'Pre-demographic dividend',
'Post-demographic dividend', 'French Polynessia', 'Qatar',
'Rost demographic dividend', 'French Polynessia', 'Qatar',
'Romania', 'Russian Federation', 'Rwanda', 'South Asia',
'Saudi Arabia', 'Sudan', 'Senegal', 'Singapore', 'Solomen Islands',
'Sub-Saharan Africa (cxcluding high income)', 'South Sudan',
'Sub-Saharan Africa', 'Small states', 'Sao Tome and Principe',
'Suriname', 'Slovak Republic', 'Slovenia', 'Sweden', 'Eswatini',
'Sith Maarten (Dutch part)', 'Soychelles', 'Syrian Arab Republic',
'Turks and Caicos Islands', 'Chad',
'East Asia & Pacific (IDA & IBRD countries)', 'Togo', 'Thailand',
'Tajikistan', 'Turkmenistan', 'Turkenistan', 'Turkmenistan', 'Turkenistan', 'Turkeni
                                                                                                                                                                                                                                                                  'Europe & Central Asia (IDA & IBRD countries)', 'Togo', 'Thailand',
'Tajikistan', 'Turkmenistan',
'Latin America & the Caribbean (IDA & IBRD countries)',
                                                                                                                                                                                                                                                     'Latin America & the Caribboan (IDA & IBRD countries)',
'Timor-Leste', 'Middle East & North Africa (IDA & IBRD countries)',
'Tonga', 'South Asia (IDA & IBRD'),
'Sub-Saharan Africa (IDA & IBRD'),
'Sub-Saharan Africa (IDA & IBRD countries)', 'Trinidad and Tobago',
'Tunisia', 'Turkiye', 'Tuvalu', 'Tanzania', 'Uganda', 'Ukraine',
'Upper middle income', 'Uruguay', 'United States', 'Uzbekistan',
'St. Vincent and the Grenadines', 'Venezuela, RB',
'Stritish Virgin Islands', 'Virgin Islands (U.S.)', 'Vietnam',
'Vanuatu', 'World', 'Samoa', 'Kosovo', 'Yemen, Rep.',
'South Africa', 'Zambia', 'Zimbabwe'], dtype-object)
```

```
[33]: df['Country Code'].unique()
  [33]: array(['ABW', 'AFE', 'AFG', 'AFW', 'AGO', 'ARG', 'ARM', 'ASM', 'ATG', 'AUS', 'BEN', 'BEN', 'BGD', 'BGR', 'BHN', 'BNU', 'BOL', 'BRA', 'BRB', 'BNN',
                                                                                                                                                                                              'ALB', 'AND', 'ARB', 'ARE',
'AUT', 'AZE', 'BDI', 'BEL',
'BHS', 'BIH', 'BLR', 'BLZ',
'BTN', 'BWA', 'CAF', 'CAN',
                                                                                                              'CHI', 'CHL', 'CHN', 'CPV', 'CRI', 'CSS',
                                                                                                                                                                                                                          'CMR',
'CUM',
'DZA',
                                                                                                                                                                                                                                                                                 'COG',
'CYP',
'EAR',
                                                         'CEB',
                                                                                                                                                                                                'CIV'
                                                                                                                                                                                                                                                      'C00'
                                                         'COL', 'COM',
'CZE', 'DEU',
                                                                                                              'DDI', 'DMA',
                                                                                                                                                                     "DNK",
                                                                                                                                                                                                  'DOM',
                                                                                                                                                                                                                                                        'EAP'
                                                         'EAS', 'ECA',
'ETH', 'EUU',
'GBR', 'GEO',
                                                                                                             'ECS', 'ECU',
'FCS', 'FIN',
'GHA', 'GIB',
                                                                                                                                                                                                'EMU', 'ERI',
'FRA', 'FRO',
'GMB', 'GNB',
                                                                                                                                                                    'EGY'.
                                                                                                                                                                                                                                                        'ESP'
                                                                                                                                                                                                                                                                                 'EST'.
                                                                                                                                                                     'FJI',
                                                                                                                                                                                                                                                        'GNQ'
                                                                                                                                                                                                'HIC', 'HKG',
'IDA', 'IDB',
'IRQ', 'ISL',
                                                          'GRD', 'GRL',
'HRV', 'HTI',
                                                                                                              'GTM', 'GUM', 'GUY',
                                                                                                              'HUN'
                                                                                                                                        'IRD', 'IRT',
                                                                                                                                                                                                                                                       "IDN"
                                                        'KNA',
                                                                                                                                                                                                                                                                                 'LCN'.
                                                                                                                                                                                                                                                                                 'LTU',
'MDV',
'MNA',
                                                                                                                                                                                                                                                                               'NAC'
                                                                                                                                                                                                                                                                               'PLW',
                                                       'NZI', 'OED', 'OMN', 'OSS', 'PAK', 'PAN', 'PER', 'PHL', 'PLL', 'PPG', 'PDG', 'PRE', 'PRI', 'PRK', 'PRT', 'PSF', 'PSS', 'PSS', 'PST', 'PYF', 'QAT', 'ROU', 'RUS', 'RNA', 'SAS, 'SAU', 'SDN', 'SEN', 'SGR', 'SLB', 'SLE', 'SLV', 'SMR', 'SOM', 'SRE', 'SSA', 'SSS', 'SSF', 'STP', 'SUR', 'SVK', 'SVW', 'SWE', 'SKZ', 'SSM', 'SYK', 'STP', 'SUR', 'SVK', 'SVW', 'SWE', 'SKZ', 'SXM', 'SYC', 'SYR', 'TCA', 'TCD', 'TEA', 'TEC', 'TGO', 'THA', 'TJK', 'TKM', 'TLA', 'TLA', 'TCD', 'TEA', 'TSA', 'TSS', 'TTO', 'TUN', 'TUR', 'TUR', 'TUT', 'TUA', 'TCA', 'UKR', 'UKR', 'UKC', 'UKR', 'USA', 'USB', 'VCT', 'VEN', 'VGB', 'VIR', 'VMM', 'VUT', 'WLD', 'NSM', 'XKK', 'YEM', 'ZAF', 'ZMB', 'ZME'], dtype-object)
 [35]: df['Indicator Name'].unique()
  [35]: array(['Population, total'], dtype-object)
 [37]: df['Indicator Code'].unique()
 [37]: array(['SP.POP.TOTL'], dtype=object)
 [39]: df.drop(['Indicator Name','Indicator Code','Country Code'],axis = 1, inplace = True)
  [41]: df.columns
  [41]: Index(['Country Name', '1968', '1961', '1962', '1963', '1964', '1965', '1966',
                                                       [Country Name', 1968', 1961', 1962', 1963', 1964', 1965', 1966', 1966', 1966', 1966', 1971', 1972', 1973', 1974', 1975', 1976', 1977', 1978', 1979', 1988', 1981', 1982', 1983', 1984', 1985', 1986', 1987', 1988', 1989', 1991', 1992', 1991', 1994', 1995', 1996', 1991', 1998', 1999', 2808', 2001', 2002', 2003', 2004', 2005', 2006', 2007', 2008', 2009', 2011', 2011', 2012', 2011', 2014', 2015', 2016', 2017', 2018', 2019', 2022', 2021', 2022'],
                                                  dtype='object')
1985', 1986', 1987', 1988', 1989', 1999', 1991', 1992', 1993', 1994', 1995', 1996', 1997', 1998', 1999', 12000', 12001', 12002', 12003', 12004', 12005', 12006', 12007', 12008', 12009', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011', 12011
```



V

1

40000 value

30000

20000