world-happiness-report-2024

October 3, 2024

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings("ignore", category=FutureWarning)
```

[3]:
```
df = pd.read_csv('World Happiness Report 2024.csv')
```

0.1 About Dataset

- Country name: The name of the country for which the data is reported.
- Year: The year the data was collected.
- Life Ladder: A scale measure of life satisfaction.
- Log GDP per capita: Logarithm of GDP per capita, indicating economic prosperity.
- Social support: Measure of perceived social support.
- Healthy life expectancy at birth: Expected number of years in good health from birth.
- Freedom to make life choices: Measure of individual freedom in making life decisions.
- Generosity: Measure of generosity or charitable giving.
- Perceptions of corruption: Measure of perceived government corruption.
- Positive affect: Level of positive emotions experienced.
- Negative affect: Level of negative emotions experienced.

[5]:
```
df.head()
```

[5]:
```
  Country name  year  Life Ladder  Log GDP per capita  Social support  \
0  Afghanistan  2008     3.723590            7.350416        0.450662
1  Afghanistan  2009     4.401778            7.508646        0.552308
2  Afghanistan  2010     4.758381            7.613900        0.539075
3  Afghanistan  2011     3.831719            7.581259        0.521104
4  Afghanistan  2012     3.782938            7.660506        0.520637

   Healthy life expectancy at birth  Freedom to make life choices  Generosity  \
0                         50.500000                      0.718114    0.164055
1                         50.799999                      0.678896    0.187297
2                         51.099998                      0.600127    0.117861
3                         51.400002                      0.495901    0.160098
```

```
4                  51.700001                      0.530935    0.234157

   Perceptions of corruption  Positive affect  Negative affect
0                   0.881686         0.414297         0.258195
1                   0.850035         0.481421         0.237092
2                   0.706766         0.516907         0.275324
3                   0.731109         0.479835         0.267175
4                   0.775620         0.613513         0.267919
```

```
[7]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2363 entries, 0 to 2362
Data columns (total 11 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Country name                      2363 non-null   object
 1   year                              2363 non-null   int64
 2   Life Ladder                       2363 non-null   float64
 3   Log GDP per capita                2335 non-null   float64
 4   Social support                    2350 non-null   float64
 5   Healthy life expectancy at birth  2300 non-null   float64
 6   Freedom to make life choices      2327 non-null   float64
 7   Generosity                        2282 non-null   float64
 8   Perceptions of corruption         2238 non-null   float64
 9   Positive affect                   2339 non-null   float64
 10  Negative affect                   2347 non-null   float64
dtypes: float64(9), int64(1), object(1)
memory usage: 203.2+ KB
```

- Statistical Analysis

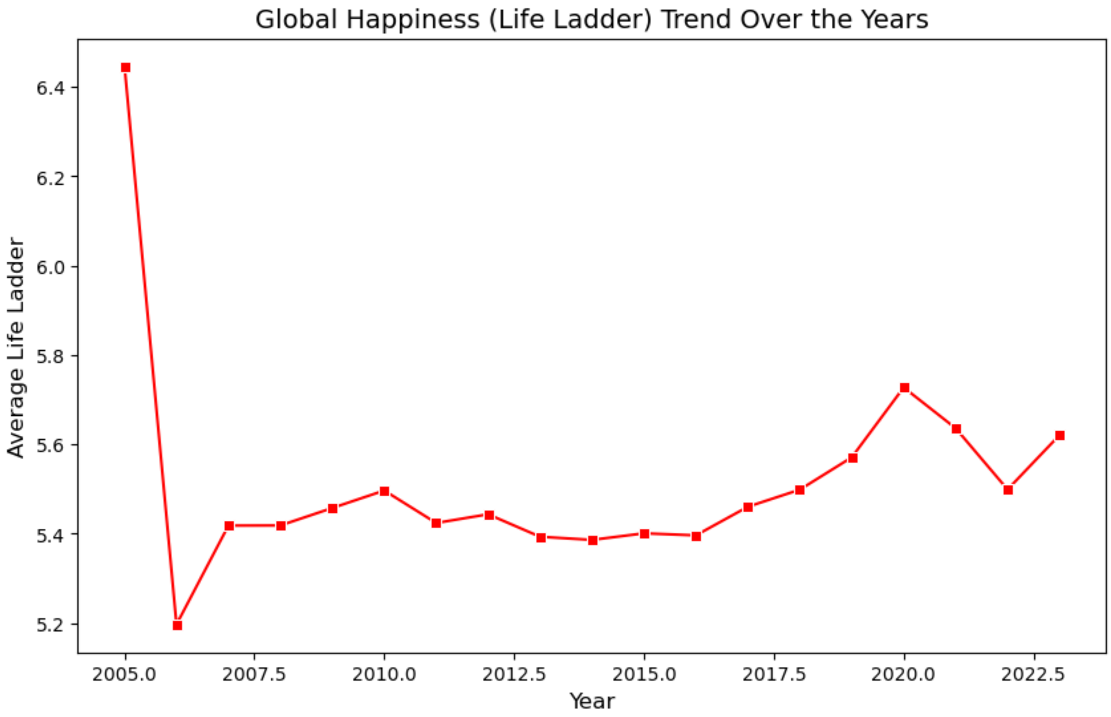
```
[10]: df.describe()
```

```
[10]:
              year  Life Ladder  Log GDP per capita  Social support  \
count  2363.000000  2363.000000         2335.000000     2350.000000
mean   2014.763860     5.483562            9.399673        0.809369
std       5.059436     1.125523            1.152062        0.121211
min    2005.000000     1.281271            5.526723        0.228217
25%    2011.000000     4.646750            8.506165        0.743811
50%    2015.000000     5.448725            9.502946        0.834395
75%    2019.000000     6.323592           10.392974        0.903779
max    2023.000000     8.018934           11.675588        0.987343

       Healthy life expectancy at birth  Freedom to make life choices  \
count                       2300.000000                   2327.000000
mean                          63.401828                      0.750290
std                            6.842644                      0.139366
```

```
min                            6.720000                      0.228301
25%                           59.195001                      0.660706
50%                           65.099998                      0.771122
75%                           68.552502                      0.861740
max                           74.599998                      0.985178

        Generosity  Perceptions of corruption  Positive affect  \
count  2282.000000                2238.000000      2339.000000
mean      0.000096                   0.743959         0.651878
std       0.161386                   0.184871         0.106249
min      -0.339547                   0.035198         0.178886
25%      -0.111940                   0.686784         0.571977
50%      -0.021613                   0.798497         0.663420
75%       0.093569                   0.867557         0.737262
max       0.699570                   0.983276         0.883586

       Negative affect
count      2347.000000
mean          0.273160
std           0.087133
min           0.082737
25%           0.208559
50%           0.262175
75%           0.326211
max           0.704590
```

- How has the global happiness score (Life Ladder) trended over the years?

[13]:
```
plt.figure(figsize=(10,6))
sns.lineplot(data=df, x='year', y='Life Ladder', marker='s', color='red', ci =␣
 ↪None)

plt.title('Global Happiness (Life Ladder) Trend Over the Years', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Average Life Ladder', fontsize=12)
plt.show()
```



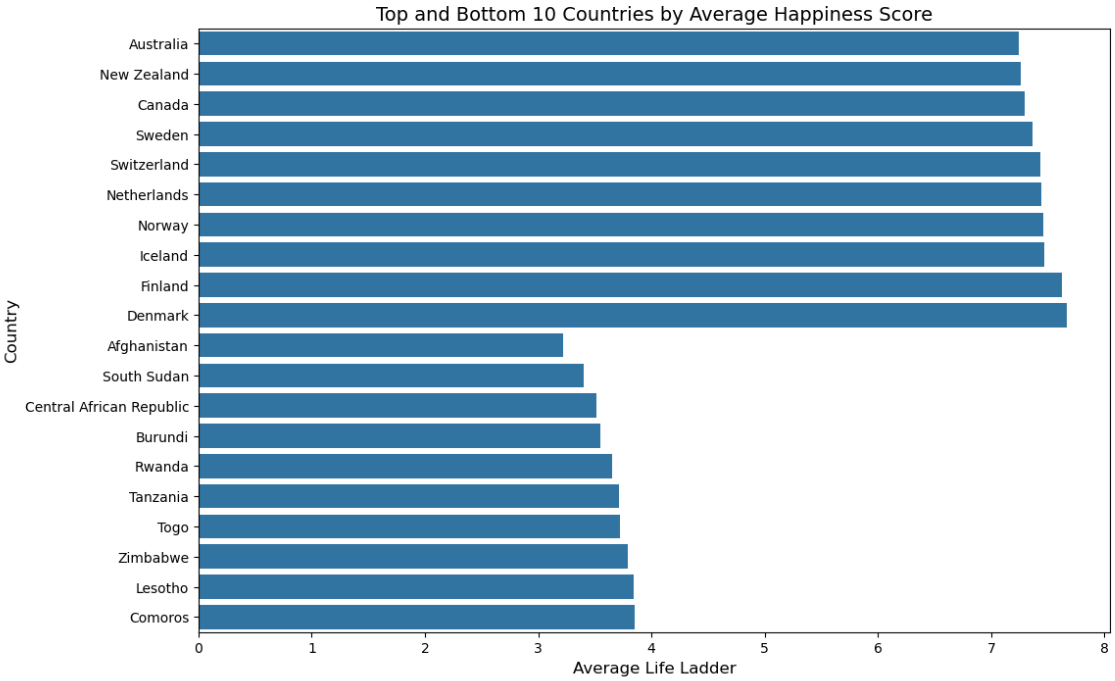
- Which countries have the highest and lowest average happiness scores ?

[15]:
```
country_ladder_avg = df.groupby('Country name')['Life Ladder'].mean().
 ↪sort_values()
country_ladder_avg
```

[15]:
```
Country name
Afghanistan                 3.219917
South Sudan                 3.401875
Central African Republic    3.514954
Burundi                     3.548124
Rwanda                      3.654473
                              …
Netherlands                 7.440393
Norway                      7.463886
Iceland                     7.467971
Finland                     7.624132
Denmark                     7.664026
Name: Life Ladder, Length: 165, dtype: float64
```

[17]:
```
country_ladder_avg = df.groupby('Country name')['Life Ladder'].mean().
 ↪sort_values()

top_10_countries = country_ladder_avg.tail(10)
```

```
bottom_10_countries = country_ladder_avg.head(10)

top_bottom_countries = pd.concat([top_10_countries, bottom_10_countries])

plt.figure(figsize=(12, 8))
sns.barplot(x=top_bottom_countries.values, y=top_bottom_countries.index)
sns.set_palette("Set3")


plt.title('Top and Bottom 10 Countries by Average Happiness Score', fontsize=14)
plt.xlabel('Average Life Ladder', fontsize=12)
plt.ylabel('Country', fontsize=12)
plt.show()
```



- How have Log GDP per capita, Social support, and Healthy life expectancy at birth trended over the years?

[19]:
```
df.head()
```

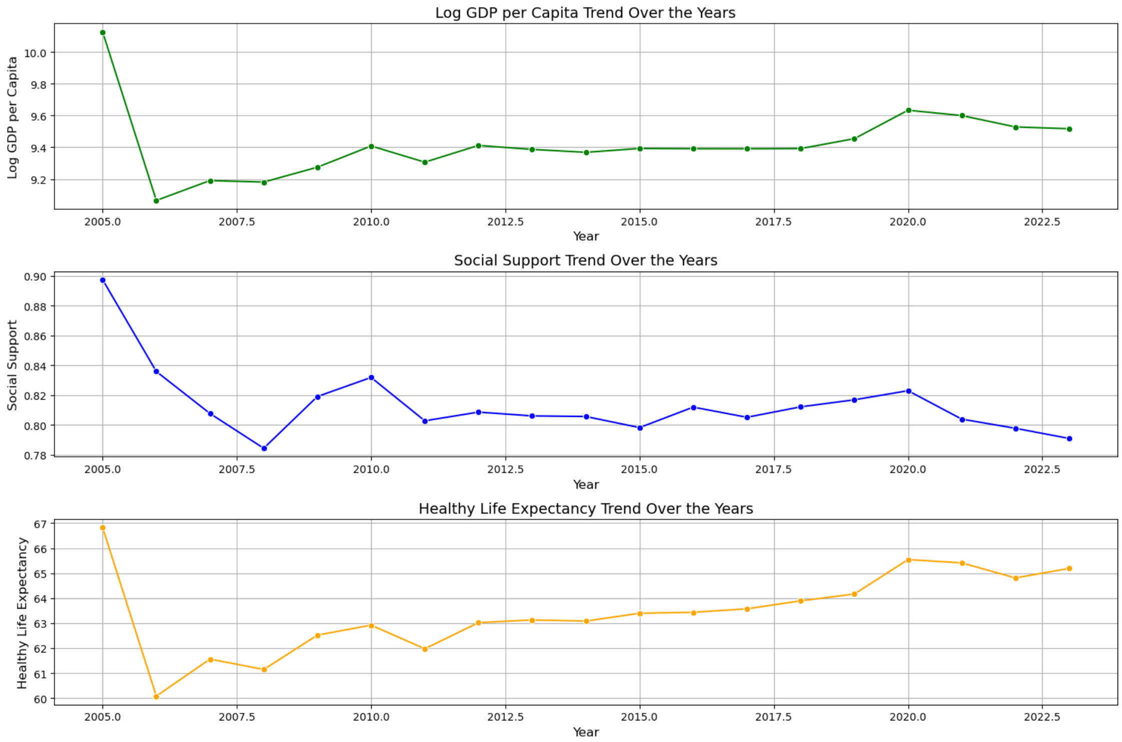
[19]:
```
  Country name  year  Life Ladder  Log GDP per capita  Social support  \
0  Afghanistan  2008     3.723590            7.350416        0.450662
1  Afghanistan  2009     4.401778            7.508646        0.552308
2  Afghanistan  2010     4.758381            7.613900        0.539075
3  Afghanistan  2011     3.831719            7.581259        0.521104
4  Afghanistan  2012     3.782938            7.660506        0.520637
```

```
   Healthy life expectancy at birth  Freedom to make life choices  Generosity  \
0                         50.500000                      0.718114    0.164055   
1                         50.799999                      0.678896    0.187297   
2                         51.099998                      0.600127    0.117861   
3                         51.400002                      0.495901    0.160098   
4                         51.700001                      0.530935    0.234157   

   Perceptions of corruption  Positive affect  Negative affect  
0                   0.881686         0.414297         0.258195  
1                   0.850035         0.481421         0.237092  
2                   0.706766         0.516907         0.275324  
3                   0.731109         0.479835         0.267175  
4                   0.775620         0.613513         0.267919  
```

[21]:
```python
df['year'].value_counts()
```

[21]:
```
year
2017    147
2011    146
2014    144
2019    143
2015    142
2012    141
2016    141
2018    141
2022    140
2023    138
2013    136
2010    124
2021    122
2020    116
2009    114
2008    110
2007    102
2006     89
2005     27
Name: count, dtype: int64
```

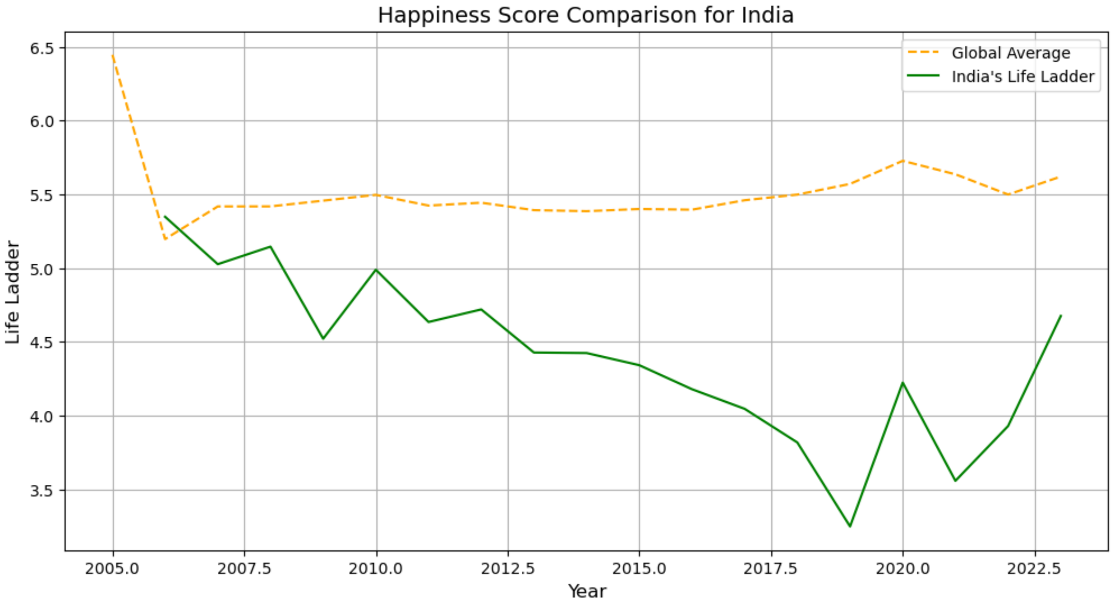
[23]:
```python
plt.figure(figsize=(15, 10))
plt.subplot(3, 1, 1)
sns.lineplot(data=df, x='year', y='Log GDP per capita', marker='o',␣
 ↪color='green', ci = None)
plt.title('Log GDP per Capita Trend Over the Years', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Log GDP per Capita', fontsize=12)
plt.grid(True)
plt.subplot(3, 1, 2)
sns.lineplot(data=df, x='year', y='Social support',  marker='o', color='blue',␣
 ↪ci = None)
plt.title('Social Support Trend Over the Years', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Social Support', fontsize=12)
plt.grid(True)

plt.subplot(3, 1, 3)
sns.lineplot(data=df, x='year', y='Healthy life expectancy at birth',␣
 ↪marker='o', color='orange', ci = None)
plt.title('Healthy Life Expectancy Trend Over the Years', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Healthy Life Expectancy', fontsize=12)
plt.grid(True)

plt.tight_layout()
plt.show()
```



- To compare the happiness score of a india to the global average

[27]:

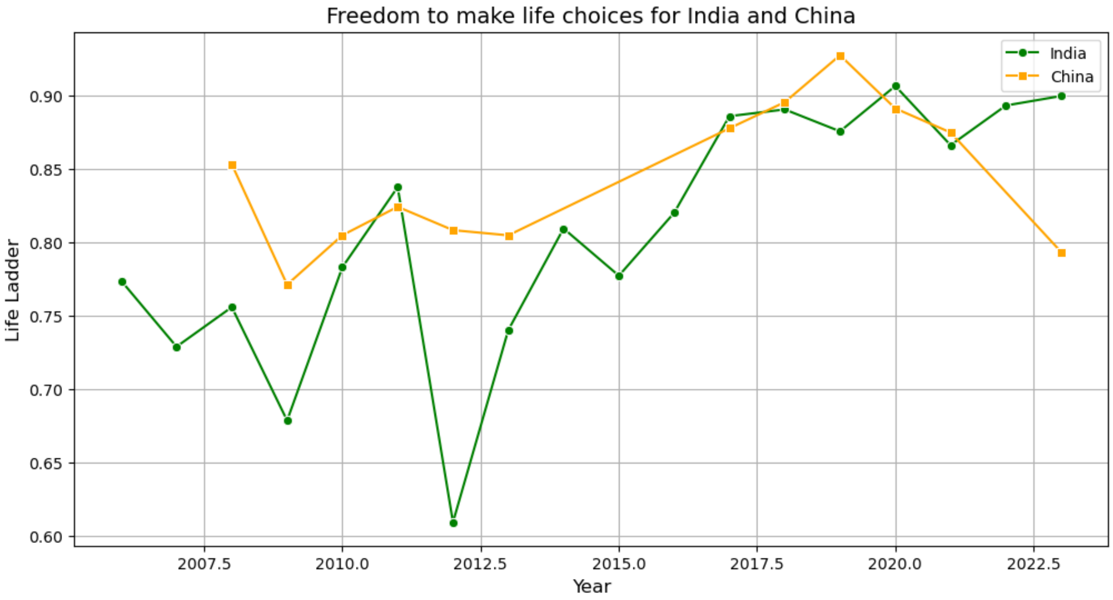
```
df_country = df[df['Country name'] == 'India']

global_avg_life_ladder = df['Life Ladder'].mean()

plt.figure(figsize=(12, 6))
sns.lineplot(data=df, x='year', y='Life Ladder', label='Global Average',␣
 ↪color='orange', linestyle='--', ci = None)
sns.lineplot(data=df_country, x='year', y='Life Ladder', label="India's Life␣
 ↪Ladder", color='green')
plt.title(f'Happiness Score Comparison for India', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Life Ladder', fontsize=12)
plt.legend()
plt.grid(True)
plt.show()
```



- To compare the Freedom to make life choices for India and China

[29]:

```
india = df[df['Country name'] == 'India']
china = df[df['Country name'] == 'China']

plt.figure(figsize=(12, 6))
sns.lineplot(data=india, x='year', y='Freedom to make life choices',␣
 ↪label='India', color='green', marker = 'o', ci = None)
sns.lineplot(data=china, x='year', y='Freedom to make life choices',␣
 ↪label="China", color='orange', marker = 's')
plt.title('Freedom to make life choices for India and China', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Life Ladder', fontsize=12)
plt.legend()
plt.grid(True)
plt.show()
```



[ ]:

Thank You