In [2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [7]:
```
import pandas as pd

# Load datasets
df_2015 = pd.read_csv("2015.csv")
df_2019 = pd.read_csv("2019.csv")

# Rename columns
df_2015 = df_2015.rename(columns={
    'Happiness Score': 'score_15',
    'Economy (GDP per Capita)': 'gdp_15',
    'Family': 'family_15',
    'Health (Life Expectancy)': 'health_15',
    'Freedom': 'freedom_15',
    'Trust (Government Corruption)': 'trust_15',
    'Generosity': 'generosity_15',
    'Country': 'country',
    'Happiness Rank': 'rank_15'
})

df_2019 = df_2019.rename(columns={
    'Score': 'score_19',
    'GDP per capita': 'gdp_19',
    'Social support': 'family_19',
    'Healthy life expectancy': 'health_19',
    'Freedom to make life choices': 'freedom_19',
    'Perceptions of corruption': 'trust_19',
    'Generosity': 'generosity_19',
    'Country or region': 'country',
    'Overall rank': 'rank_19'
})

# Merge on 'country'
df = pd.merge(df_2015, df_2019, on='country')

# Create change columns
df['score_change'] = df['score_19'] - df['score_15']
df['gdp_change'] = df['gdp_19'] - df['gdp_15']
df['family_change'] = df['family_19'] - df['family_15']
df['health_change'] = df['health_19'] - df['health_15']
df['freedom_change'] = df['freedom_19'] - df['freedom_15']
df['generosity_change'] = df['generosity_19'] - df['generosity_15']
df['trust_change'] = df['trust_19'] - df['trust_15']

# Preview
df.head()
```

Out[7]:

	country	Region	rank_15	score_15	Standard Error	gdp_15	family_15	health_1
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.9414
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.9478
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.8746
3	Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.8852
4	Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.9056

5 rows × 27 columns

In [9]:
```
#Insight #1: Top 10 Happiest Countries (2015 vs 2019)
```

In [8]:
```
#What was done:
#We sorted the dataset by score_15 and score_19 in descending order to co
#happiest countries in 2015 and 2019. This helps us understand which coun
#high and whether any country rose or fell in the rankings over time.
```

In [10]:
```
rank15= df.sort_values('score_15', ascending=False)[['country','score_15'
rank19 =df.sort_values('score_19', ascending=False)[['country','score_19'
```

In [12]:
```
rank15
```

Out[12]:

	country	score_15
0	Switzerland	7.587
1	Iceland	7.561
2	Denmark	7.527
3	Norway	7.522
4	Canada	7.427
5	Finland	7.406
6	Netherlands	7.378
7	Sweden	7.364
8	New Zealand	7.286
9	Australia	7.284

In [13]:
```
rank19
```

Out[13]:

	country	score_19
5	Finland	7.769
2	Denmark	7.600
3	Norway	7.554
1	Iceland	7.494
6	Netherlands	7.488
0	Switzerland	7.480
7	Sweden	7.343
8	New Zealand	7.307
4	Canada	7.278
12	Austria	7.246

In [15]:
```
#Insight #2: Bottom 10 Happiest Countries (2015 vs 2019)
#What was done:
#We sorted the dataset by happiness scores in ascending order to find
#the 10 countries with the lowest happiness scores in 2015 and 2019. This
#identify regions with persistent or worsening well-being issues.
```

In [17]:
```
low15 = df.sort_values('score_15', ascending=True)[['country','score_15']
low19 = df.sort_values('score_19', ascending=True)[['country','score_19']
```

In [19]:
```
low15
```

Out[19]:

	country	score_15
148	Togo	2.839
147	Burundi	2.905
146	Syria	3.006
145	Benin	3.340
144	Rwanda	3.465
143	Afghanistan	3.575
142	Burkina Faso	3.587
141	Ivory Coast	3.655
140	Guinea	3.656
139	Chad	3.667

In [20]:
```
low19
```

Out[20]:

	country	score_19
138	Central African Republic	3.083
143	Afghanistan	3.203
136	Tanzania	3.231
144	Rwanda	3.334
127	Yemen	3.380
122	Malawi	3.410
146	Syria	3.462
119	Botswana	3.488
111	Haiti	3.597
108	Zimbabwe	3.663

In [21]:
```
# #Insight #3: Average Global Happiness Score (2015 vs 2019)
# #We calculated the mean happiness score for each year using the score_1
# #score_19 columns. This helps us understand whether global well-being i
# or declined over the 4 years.
```

In [22]:
```
mean15 = df['score_15'].mean()
mean19 = df['score_19'].mean()
print(mean15)
print(mean19)
```

```
5.378288590604026
5.433872483221476
```

In [23]:
```
# The global average happiness score slightly increased from 2015 to 2019
# While improvement is modest, it shows that some countries did get happi
# even if a few declined.
# The gap between happiest and least happy countries remains wide, sugges
# uneven growth in well-being.
```

In [24]:
```
# 🔷 Insight #4: Countries with the Most Increase and Decrease in Happine
# 📌 What was done:
# We used the score_change column (calculated as score_19 - score_15) to

# Top 10 countries where happiness increased the most

# Top 10 countries where happiness decreased the most
```

In [27]:
```
top_improved = df.sort_values('score_change', ascending=False)[['country'
```

In [28]:
```
top_declined = df.sort_values('score_change', ascending=True)[['country',
```

In [29]:
```
top_improved
```

Out[29]:

	country	score_15	score_19	score_change
145	Benin	3.340	4.883	1.543
141	Ivory Coast	3.655	4.944	1.289
148	Togo	2.839	4.085	1.246
98	Honduras	4.788	5.860	1.072
142	Burkina Faso	3.587	4.587	1.000
97	Hungary	4.800	5.758	0.958
81	Romania	5.124	6.070	0.946
133	Gabon	3.896	4.799	0.903
135	Cambodia	3.819	4.700	0.881
140	Guinea	3.656	4.534	0.878

In [31]:
```
top_declined
```

Out[31]:

	country	score_15	score_19	score_change
21	Venezuela	6.810	4.707	-2.103
90	Lesotho	4.898	3.802	-1.096
80	Zambia	5.129	4.107	-1.022
108	Zimbabwe	4.610	3.663	-0.947
111	Haiti	4.518	3.597	-0.921
122	Malawi	4.292	3.410	-0.882
119	Botswana	4.332	3.488	-0.844
127	Yemen	4.077	3.380	-0.697
15	Brazil	6.983	6.300	-0.683
94	Swaziland	4.867	4.212	-0.655

In [32]:
```
# Benin, Ivory Coast, and Togo showed the most improvement — all African
# happiness is increasing in some lower-income regions.

# Countries facing economic collapse (Venezuela), conflict (Yemen, Zimbab
# issues (Lesotho, Haiti) saw large declines.

# Even some middle-income nations like Brazil saw happiness drop — a sign
# alone doesn't guarantee well-being.
```

In [ ]:
```
# 🔷 Insight #5: Which Features Increased or Decreased from 2015 to 2019
# 📌 What was done:
# We computed the average value of each key feature in 2015 and 2019 (e.g
# healthier, freer, richer, or more trusting over time.
```

In [53]:
```
# Features to compare
features = ['score','gdp', 'family', 'health', 'freedom', 'generosity', '
```

```
# Create comparison table
for feat in features :
    avg_15 = df[f'{feat}_15'].mean()
    avg_19 = df[f'{feat}_19'].mean()
    change = avg_19-avg_15
    print(f"📊 {feat.capitalize()}: 2015 = {avg_15:.3f}, 2019 = {avg_19:.
```

```
📊 Score: 2015 = 5.378, 2019 = 5.434, Change = +0.056
📊 Gdp: 2015 = 0.846, 2019 = 0.914, Change = +0.068
📊 Family: 2015 = 0.992, 2019 = 1.215, Change = +0.223
📊 Health: 2015 = 0.635, 2019 = 0.732, Change = +0.097
📊 Freedom: 2015 = 0.429, 2019 = 0.394, Change = -0.036
📊 Generosity: 2015 = 0.237, 2019 = 0.185, Change = -0.052
📊 Trust: 2015 = 0.142, 2019 = 0.110, Change = -0.031
```

In [37]:
```
# ✅ GDP, Family support, and Health all increased globally between 2015
# life expectancy, and perceived social support.

# ❌ Freedom, Generosity, and Trust in government all declined, suggestin
# and institutional confidence.
```

In [39]:
```
# --------UNIVARIATE ANALYSIS---------. LET'S SEE AND ANALYSE THE SINGLE
```

In [41]:
```
df.columns
```

Out[41]:
```
Index(['country', 'Region', 'rank_15', 'score_15', 'Standard Error', 'gd
p_15',
       'family_15', 'health_15', 'freedom_15', 'trust_15', 'generosity_1
5',
       'Dystopia Residual', 'rank_19', 'score_19', 'gdp_19', 'family_1
9',
       'health_19', 'freedom_19', 'generosity_19', 'trust_19', 'score_ch
ange',
       'gdp_change', 'family_change', 'health_change', 'freedom_change',
       'generosity_change', 'trust_change'],
      dtype='object')
```

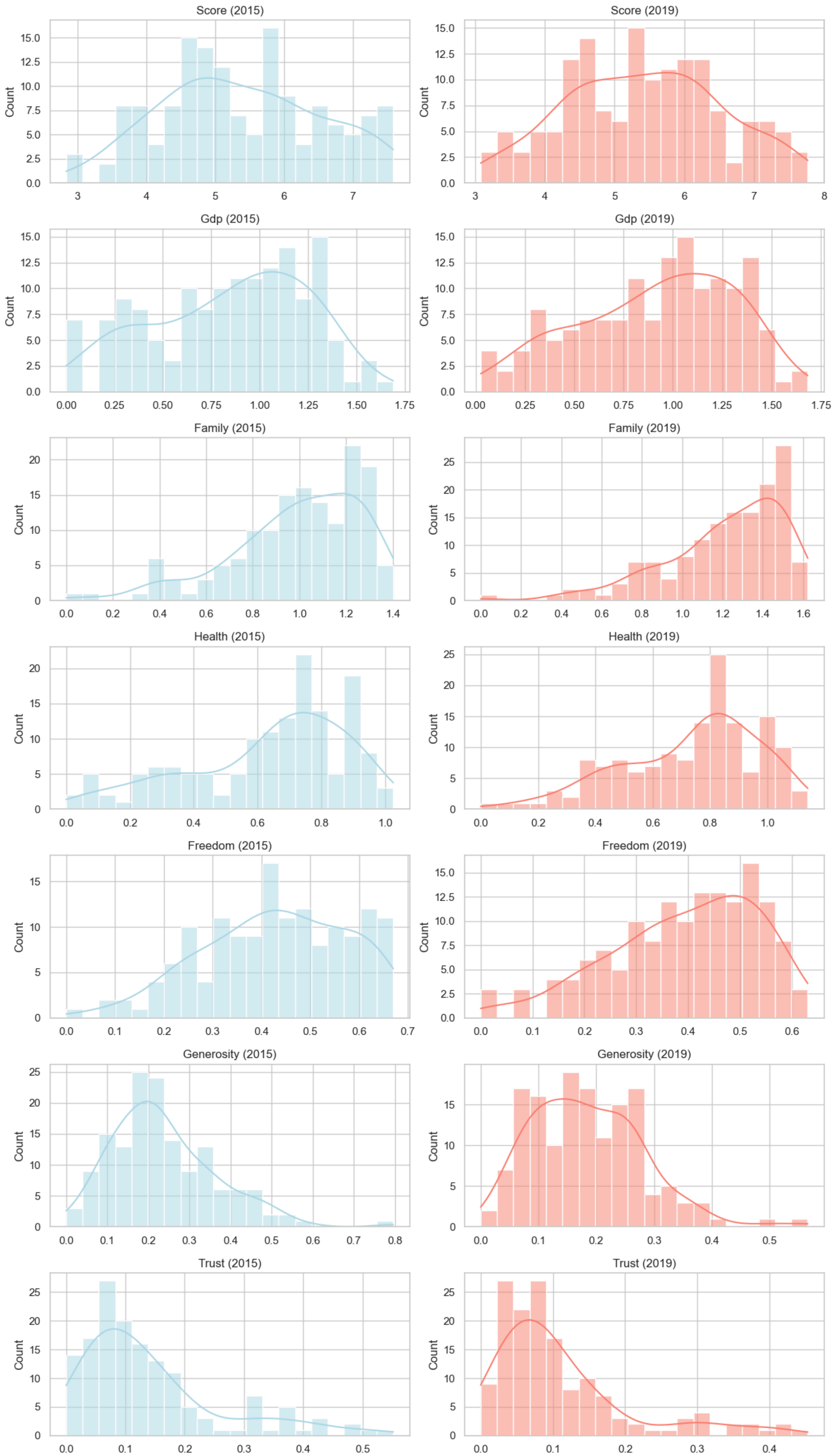
In [50]:
```
df=df.drop(columns=['Region'])
```

In [51]:
```
df.columns
```

Out[51]:
```
Index(['country', 'rank_15', 'score_15', 'gdp_15', 'family_15', 'health_
15',
       'freedom_15', 'trust_15', 'generosity_15', 'rank_19', 'score_19',
       'gdp_19', 'family_19', 'health_19', 'freedom_19', 'generosity_1
9',
       'trust_19', 'score_change', 'gdp_change', 'family_change',
       'health_change', 'freedom_change', 'generosity_change', 'trust_ch
ange'],
      dtype='object')
```

In [52]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Features to plot
features = ['score', 'gdp', 'family', 'health', 'freedom', 'generosity',

# Set the layout: one row per feature, two columns (2015 and 2019)
plt.figure(figsize=(12, len(features)*3))
sns.set(style='whitegrid')
```

```
for i, feat in enumerate(features):
    # Plot for 2015
    plt.subplot(len(features), 2, 2*i + 1)
    sns.histplot(df[f'{feat}_15'], kde=True, bins=20, color='lightblue')
    plt.title(f"{feat.capitalize()} (2015)")
    plt.xlabel('')

    # Plot for 2019
    plt.subplot(len(features), 2, 2*i + 2)
    sns.histplot(df[f'{feat}_19'], kde=True, bins=20, color='salmon')
    plt.title(f"{feat.capitalize()} (2019)")
    plt.xlabel('')

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


Score (2015)
Score (2019)
Gdp (2015)
Gdp (2019)
Family (2015)
Family (2019)
Health (2015)
Health (2019)
Freedom (2015)
Freedom (2019)
Generosity (2015)
Generosity (2019)
Trust (2015)
Trust (2019)
Count

In [54]:
```
df.describe()
```

Out[54]:

	rank_15	score_15	gdp_15	family_15	health_15	freedom_15
count	149.000000	149.000000	149.000000	149.000000	149.000000	149.000000
mean	79.369128	5.378289	0.846230	0.992332	0.634892	0.429376
std	46.125175	1.157782	0.404755	0.277737	0.247618	0.148194
min	1.000000	2.839000	0.000000	0.000000	0.000000	0.000000
25%	39.000000	4.518000	0.546490	0.851880	0.467210	0.328180
50%	79.000000	5.253000	0.901980	1.035160	0.698050	0.434500
75%	119.000000	6.295000	1.154060	1.223930	0.813250	0.546040
max	158.000000	7.587000	1.690420	1.402230	1.025250	0.669730

8 rows × 23 columns

In [ ]:
```
# 🔷 Insight #6: Which Features Correlate Most with Happiness?
# (aka: What drives happiness in 2019?)

# 🎯 Goal:
# To find which features (GDP, health, trust, etc.) are most closely rela
```

In [60]:
```
corr_15=df_2015.corr(numeric_only=True)['score_15'].sort_values(ascending
corr_15
```

Out[60]:
```
score_15             1.000000
gdp_15               0.780966
family_15            0.740605
health_15            0.724200
freedom_15           0.568211
Dystopia Residual    0.530474
trust_15             0.395199
generosity_15        0.180319
Standard Error      -0.177254
rank_15             -0.992105
Name: score_15, dtype: float64
```

In [61]:
```
corr_19=df_2019.corr(numeric_only=True)['score_19'].sort_values(ascending
corr_19
```

Out[61]:
```
score_19         1.000000
gdp_19           0.793883
health_19        0.779883
family_19        0.777058
freedom_19       0.566742
trust_19         0.385613
generosity_19    0.075824
rank_19         -0.989096
Name: score_19, dtype: float64
```

In [62]:
```
# lets create a heatmap to visualize
```

In [65]:
```
df[df['score_19'] > df['score_15']]
```

Out[65]:

	country	rank_15	score_15	gdp_15	family_15	health_15	freedom_15	tı
2	Denmark	3	7.527	1.32548	1.36058	0.87464	0.64938	(
3	Norway	4	7.522	1.45900	1.33095	0.88521	0.66973	(
5	Finland	6	7.406	1.29025	1.31826	0.88911	0.64169	(
6	Netherlands	7	7.378	1.32944	1.28017	0.89284	0.61576	(
8	New Zealand	9	7.286	1.25018	1.31967	0.90837	0.63938	(
...	...	...	...	...	...	...	...	
142	Burkina Faso	152	3.587	0.25812	0.85188	0.27125	0.39493	(
145	Benin	155	3.340	0.28665	0.35386	0.31910	0.48450	(
146	Syria	156	3.006	0.66320	0.47489	0.72193	0.15684	(
147	Burundi	157	2.905	0.01530	0.41587	0.22396	0.11850	(
148	Togo	158	2.839	0.20868	0.13995	0.28443	0.36453	

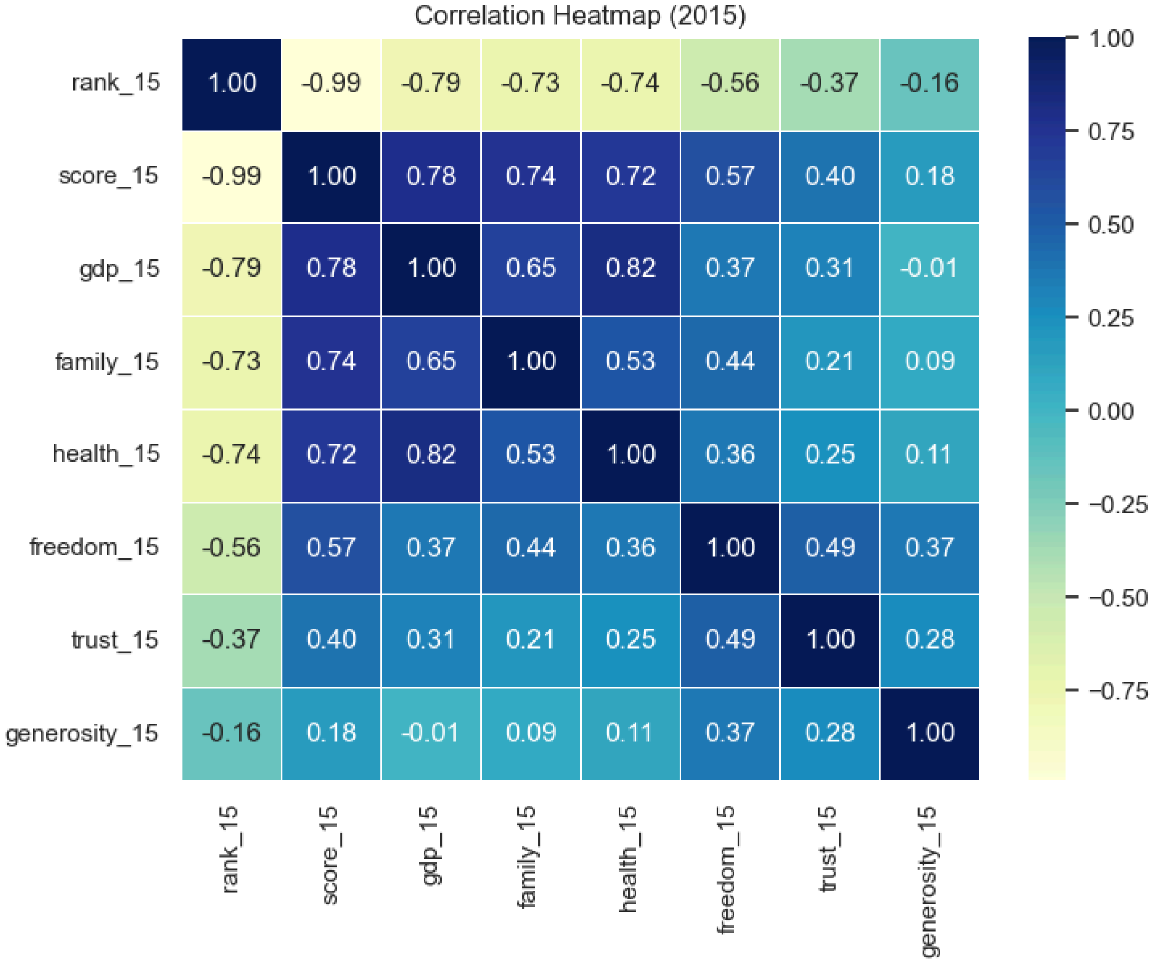
77 rows × 24 columns

In [75]:

```
df_2015.drop(columns=['Dystopia Residual','Standard Error'], inplace=True
```

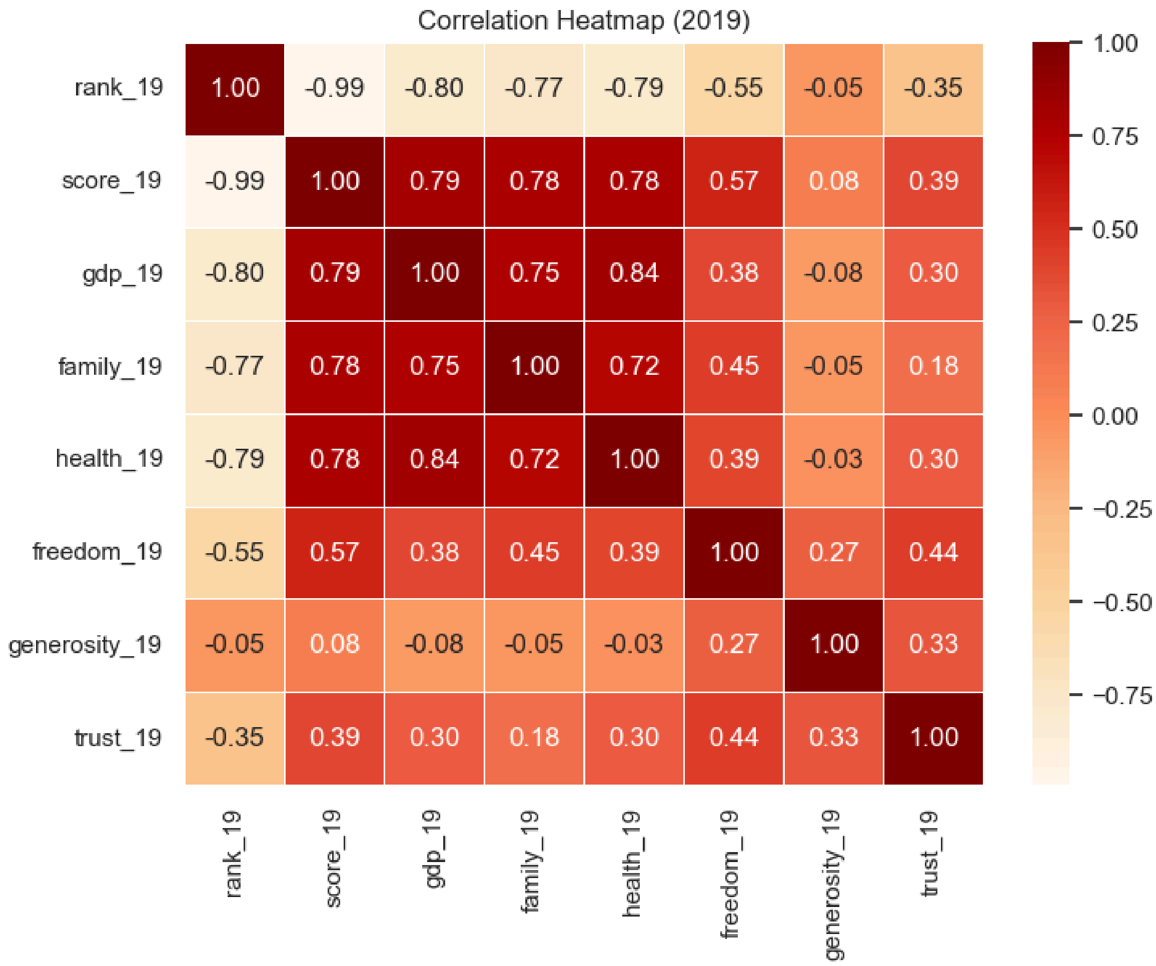
In [76]:

```
import seaborn as sns
import matplotlib.pyplot as plt

# Correlation matrix for 2015
corr_2015 = df_2015.corr(numeric_only=True)

plt.figure(figsize=(8, 6))
sns.heatmap(corr_2015, annot=True, fmt=".2f", cmap="YlGnBu", linewidths=0
plt.title("Correlation Heatmap (2015)")
plt.show()
```



In [77]:

```
# Correlation matrix for 2019
corr_2019 = df_2019.corr(numeric_only=True)

plt.figure(figsize=(8, 6))
sns.heatmap(corr_2019, annot=True, fmt=".2f", cmap="OrRd", linewidths=0.5
plt.title("Correlation Heatmap (2019)")
plt.show()
```



In [80]:
```
df.head()
```

Out[80]:

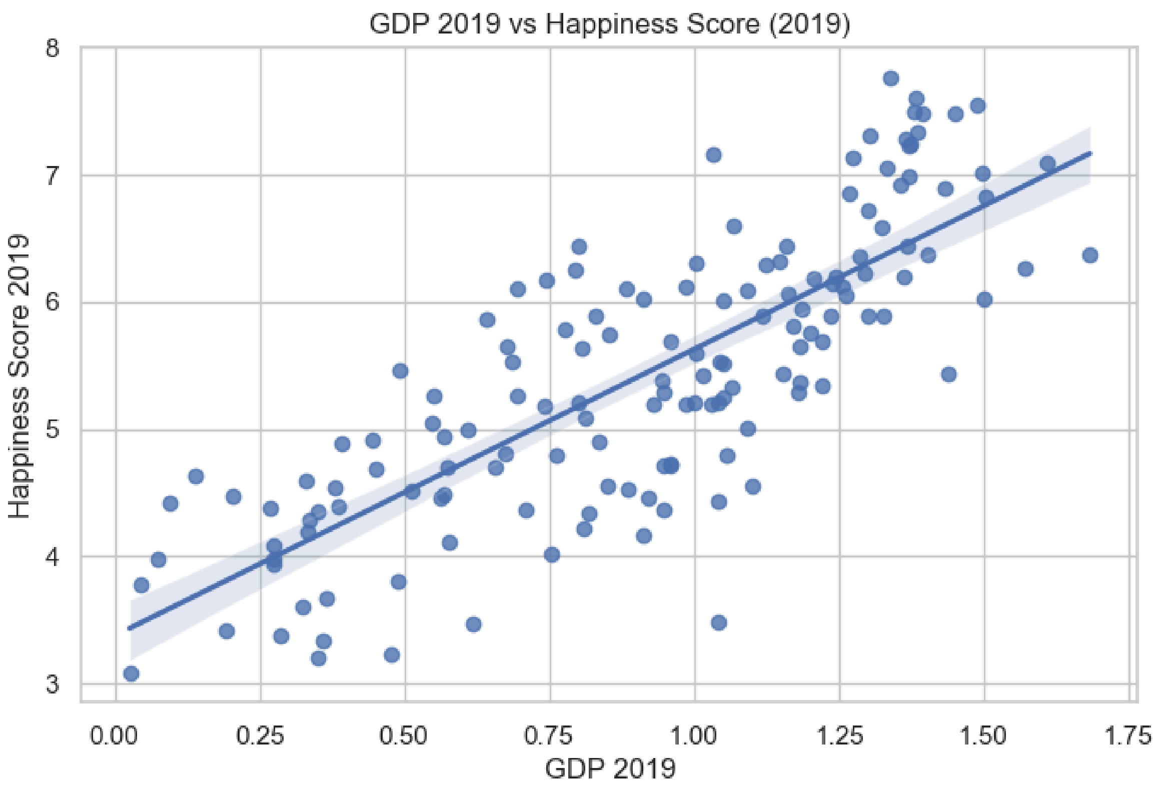
	country	rank_15	score_15	gdp_15	family_15	health_15	freedom_15	trus
0	Switzerland	1	7.587	1.39651	1.34951	0.94143	0.66557	0.41
1	Iceland	2	7.561	1.30232	1.40223	0.94784	0.62877	0.14
2	Denmark	3	7.527	1.32548	1.36058	0.87464	0.64938	0.48
3	Norway	4	7.522	1.45900	1.33095	0.88521	0.66973	0.36
4	Canada	5	7.427	1.32629	1.32261	0.90563	0.63297	0.32

5 rows × 24 columns

In [ ]:
```
#----BIVARIATE ANALYSIS----
# 📌 Why is this important?

# While univariate analysis showed how individual factors changed over ti
# bivariate analysis helps us understand which features actually drive ha
# how strongly.

# 📌 How we'll do it:

# Use sns.regplot() to visualize the relationship between each feature an

# Calculate correlation (.corr()) to quantify strength

# Interpret each plot and draw insights
```

In [81]:
```
#A. score vs gdp
```

In [90]:
```
plt.figure(figsize=(8,5))
sns.regplot(data=df, x='gdp_19', y='score_19', )
plt.title('GDP 2019 vs Happiness Score (2019)')
plt.xlabel('GDP 2019')
plt.ylabel('Happiness Score 2019')
plt.show()

print(df['gdp_19'].corr(df['score_19']))
```



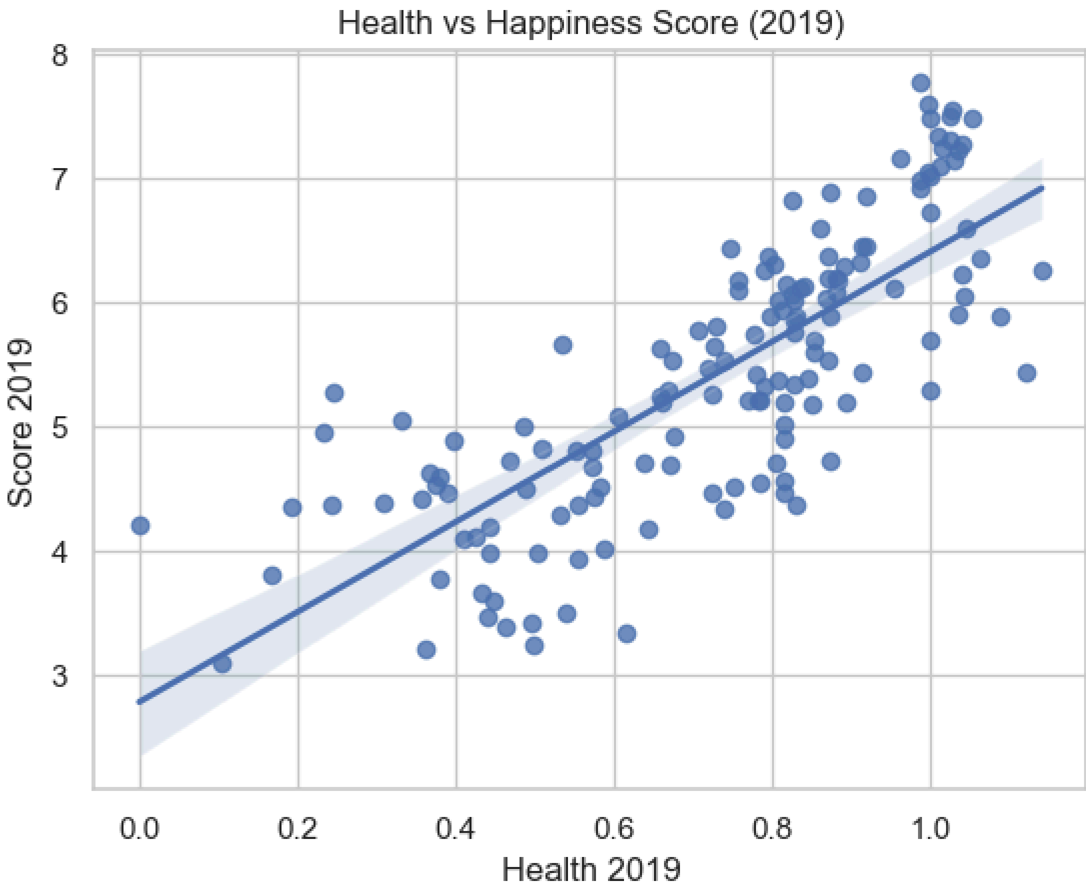

```
0.7964065217934378
```

In [88]:
```
# Bivariate Analysis: GDP vs Happiness Score

# We analyzed how GDP per capita in 2019 is related to the happiness scor
# A regression plot and correlation coefficient show a strong positive re
# (r ≈ 0.75). This confirms that economic prosperity remains a major driv
# national happiness.
```

In [89]:
```
#B. Score vs Health
```

In [93]:
```
sns.regplot(data=df, x='health_19', y='score_19')
plt.title('Health vs Happiness Score (2019)')
plt.xlabel('Health 2019')
plt.ylabel('Score 2019')

print (df['score_19'].corr(df['health_19']).round(3))
```

```
0.78
```

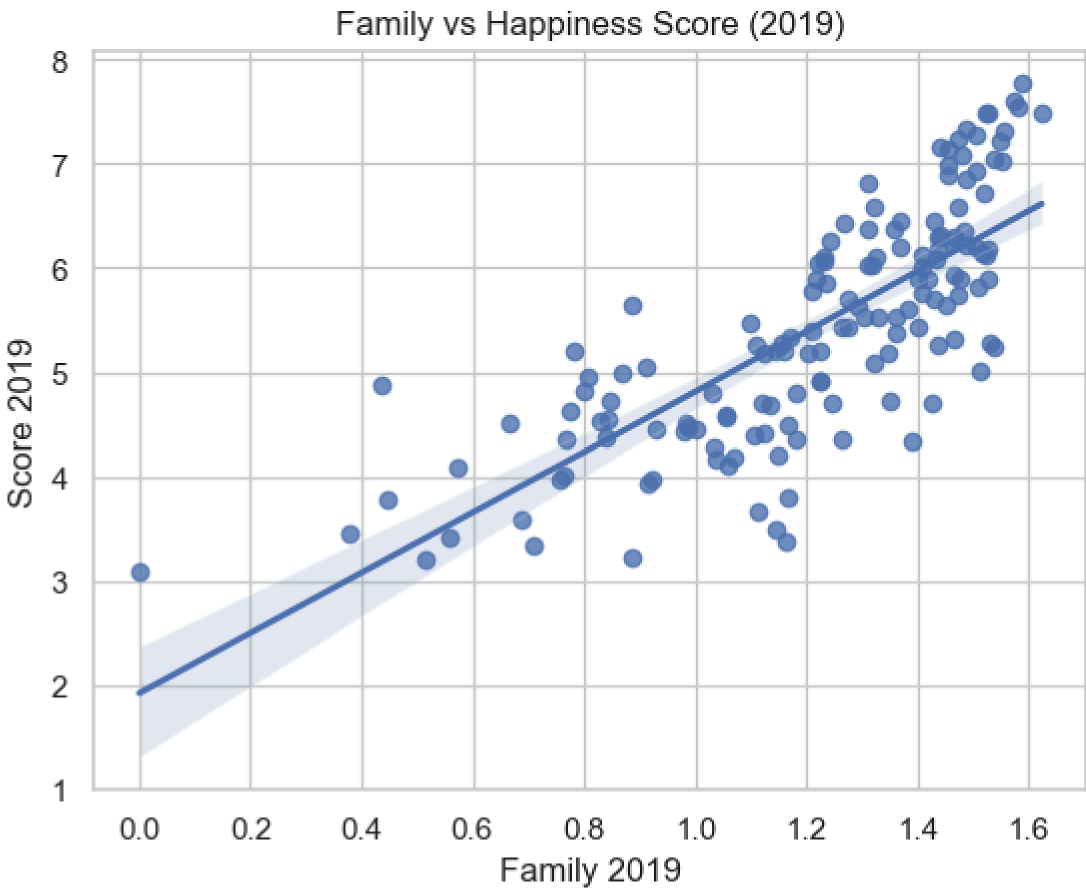


In [94]:
```
# Health vs Happiness Score (2019):
# The regression plot and correlation coefficient reveal a strong positiv
# relationship between health (measured as life expectancy) and happiness
# The correlation value is +0.779, even stronger than GDP.
# This suggests that longer, healthier lives are highly associated with
# higher happiness levels among countries in 2019.
```

In [95]:
```
#C. Score vs Family
```

In [96]:
```
sns.regplot(data=df, x='family_19', y='score_19')
plt.title('Family vs Happiness Score (2019)')
plt.xlabel('Family 2019')
plt.ylabel('Score 2019')

print (df['score_19'].corr(df['family_19']).round(3))
```

```
0.773
```



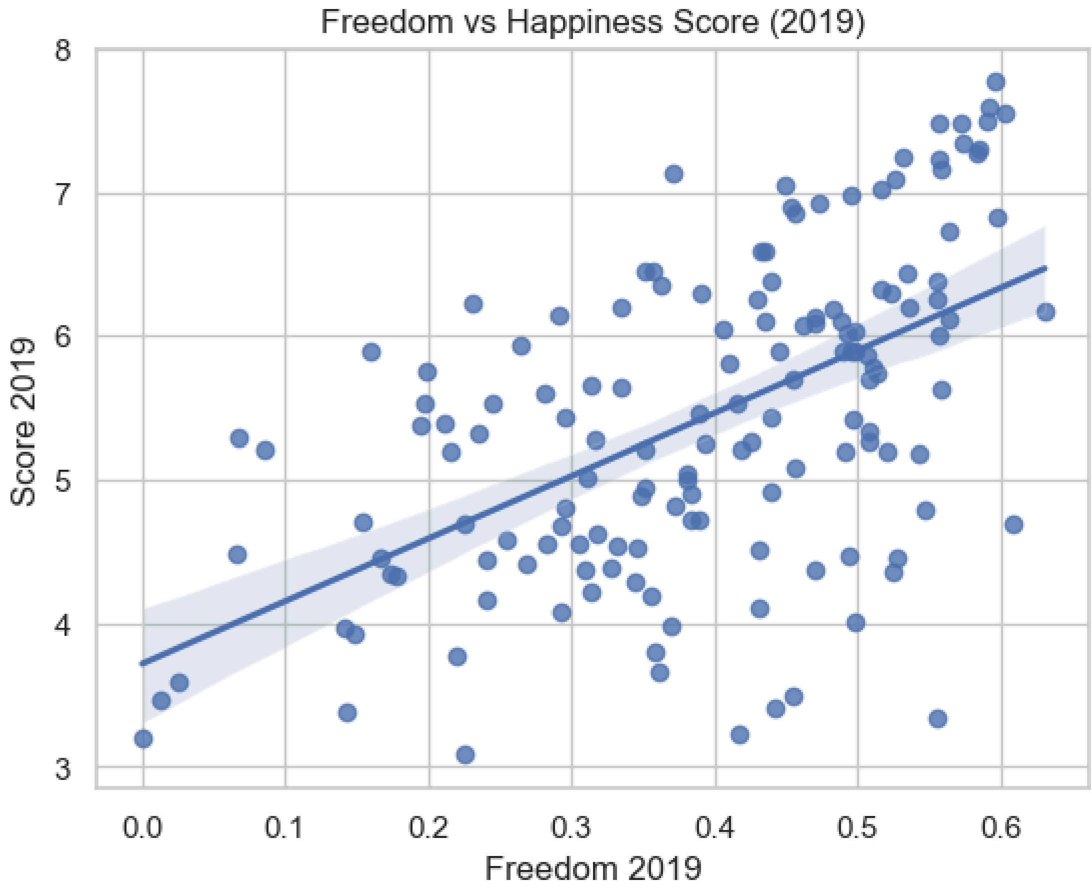
In [98]:

```
# Family Support vs Happiness Score (2019):
# There is a strong positive correlation (r = 0.773) between perceived fa
# support and national happiness.
# The regression plot shows that countries where people report having soc
# support in times of need tend to be significantly happier.
# This confirms that relationships and community play a vital role in wel
```

In [97]:

```
sns.regplot(data=df, x='freedom_19', y='score_19')
plt.title('Freedom vs Happiness Score (2019)')
plt.xlabel('Freedom 2019')
plt.ylabel('Score 2019')

print (df['score_19'].corr(df['freedom_19']).round(3))
```

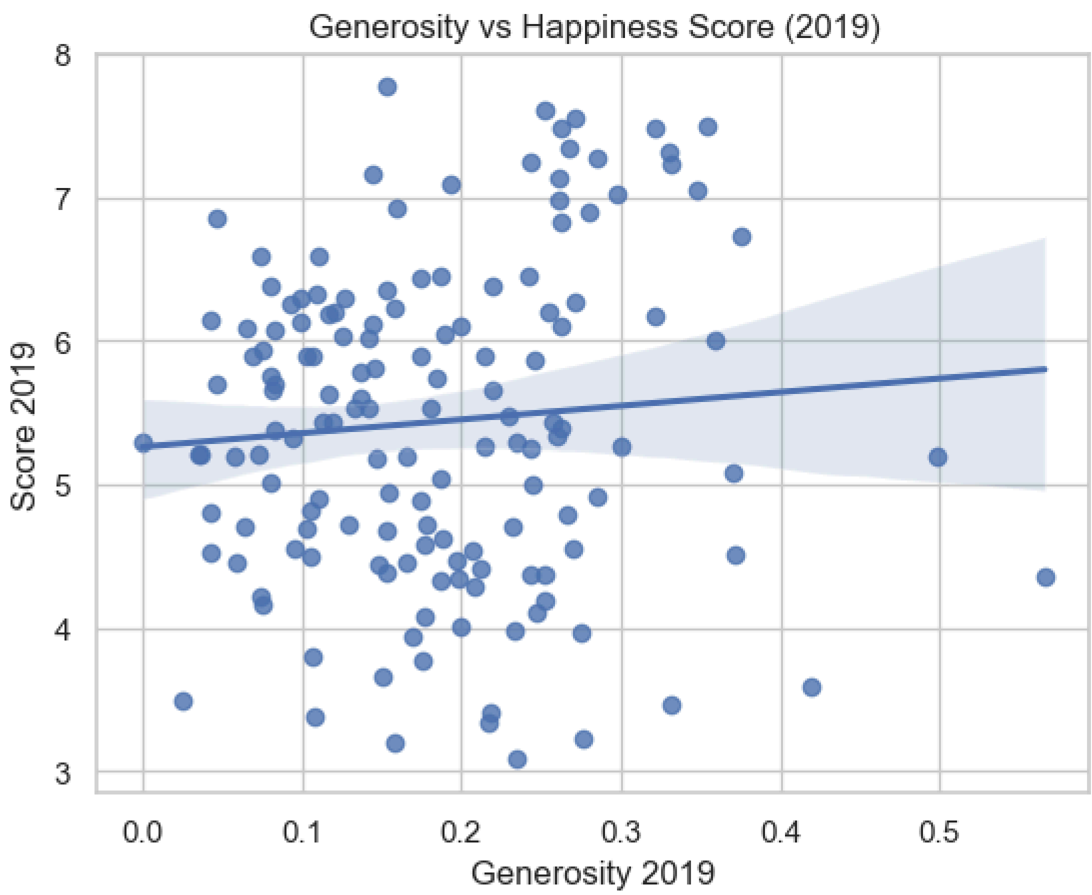
```
0.558
```



In [99]:
```
# Freedom vs Happiness Score (2019):
# The correlation between freedom and happiness score is moderate, at 0.5
# The plot suggests that greater freedom to make life choices is generall
# with higher happiness, although the relationship is not as strong as GD
# or family.
# Still, freedom remains an important contributing factor to overall well
```

In [100…
```
sns.regplot(data=df, x='generosity_19', y='score_19')
plt.title('Generosity vs Happiness Score (2019)')
plt.xlabel('Generosity 2019')
plt.ylabel('Score 2019')

print (df['score_19'].corr(df['generosity_19']).round(3))
```

```
0.083
```



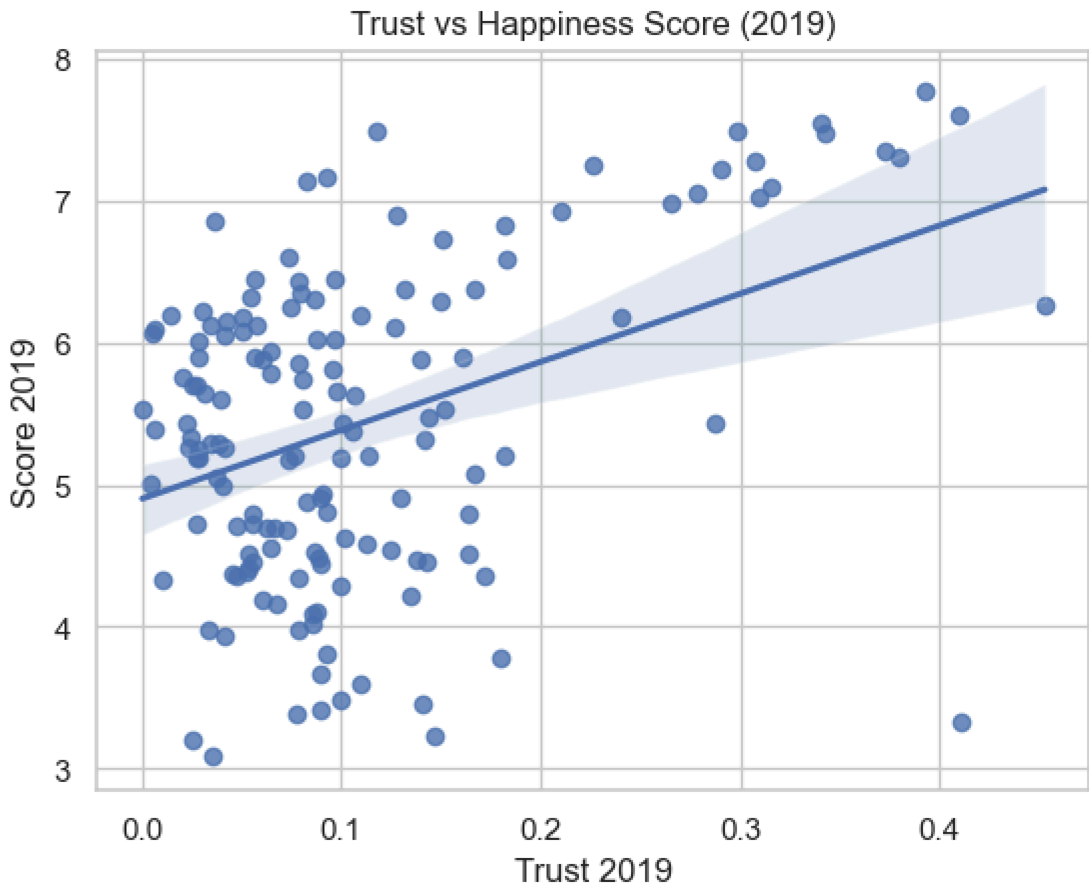
In [101…

```
# The relationship between generosity and happiness score is very weak, w
# correlation of just 0.083.
# Despite a slight upward trend in the plot, the low correlation suggests
# generosity alone does not significantly influence national happiness sc
# This could be due to variability in how generosity is expressed or meas
# across different cultures.
```

In [102…

```
sns.regplot(data=df, x='trust_19', y='score_19')
plt.title('Trust vs Happiness Score (2019)')
plt.xlabel('Trust 2019')
plt.ylabel('Score 2019')

print (df['score_19'].corr(df['trust_19']).round(3))
```

```
0.411
```



In [104…

```
# Trust vs Happiness Score (2019):
# There is a moderate positive correlation between trust in government an
# happiness score, with a correlation coefficient of 0.411.
# This suggests that in countries where citizens have greater confidence
# in their government, happiness scores tend to be higher.
# While not as strong as GDP or health, trust is still a meaningful contr
# to well-being.
```

SUMMARY OF BIVARIATE ANALYSIS The bivariate analysis reveals that the strongest positive relationships with happiness in 2019 were observed in:

Health (Life Expectancy) (r = 0.779)

Family Support (r = 0.773)

GDP per Capita (r = 0.749) These features have a clear, linear influence on happiness, as shown in the regression plots.

Freedom (r = 0.558) and Trust in Government (r = 0.411) showed moderate associations, suggesting they do play a role in happiness, but not as dominantly.

Surprisingly, Generosity had a very weak correlation (r = 0.083), indicating that individual altruistic behavior may not translate into higher national happiness scores, or may vary widely by region and reporting.

In [106…

```
%pip install plotly
```

```
Collecting plotly
  Downloading plotly-6.1.2-py3-none-any.whl.metadata (6.9 kB)
Collecting narwhals>=1.15.1 (from plotly)
  Downloading narwhals-1.44.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: packaging in /Library/Frameworks/Python.fra
mework/Versions/3.12/lib/python3.12/site-packages (from plotly) (25.0)
Downloading plotly-6.1.2-py3-none-any.whl (16.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 16.3/16.3 MB 4.0 MB/s eta 0:0
0:00a 0:00:01
Downloading narwhals-1.44.0-py3-none-any.whl (365 kB)
Installing collected packages: narwhals, plotly
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 2/2 [plotly]2m1/2 [plotly]
Successfully installed narwhals-1.44.0 plotly-6.1.2
Note: you may need to restart the kernel to use updated packages.
```

In [107…

```python
import plotly.express as px
```

---- Visual Dashboard ---- A. World Map of Happiness Score

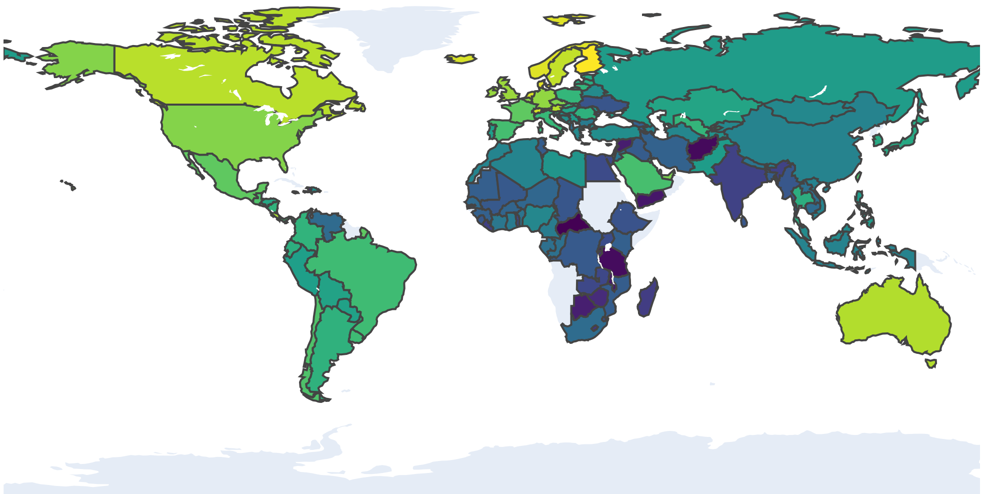
In [112…

```python
#This choropleth map displays the distribution of happiness scores across
# year 2019. Each country is shaded based on its average happiness score,
# representing higher levels of happiness.
```

In [111…

```python
fig = px.choropleth(
    df,
    locations = 'country',
    locationmode = 'country names',
    color = 'score_19',
    hover_name = 'country',
    color_continuous_scale = 'viridis',
    title = 'World Happiness Score (2019)'
)
fig.update_layout(
    geo=dict(showframe=  False, showcoastlines = False),
    coloraxis_colorbar = dict(title='Happiness Score')
)
fig.show(renderer="notebook")
```

World Happiness Score (2019)



Nordic countries (such as Finland, Denmark, and Norway) appear the happiest, shown by the lightest shades.

Countries in Sub-Saharan Africa, South Asia, and conflict-affected regions tend to have lower happiness scores, indicated by darker shades.

This map visually reinforces global inequalities in well-being, showing that geographic region, development status, and governance quality play a significant role in national happiness.

The map provides a high-level overview before diving into deeper statistical relationships explored later in the report.

B. Visually compare the top gainers and top decliners in happiness from 2015 to 2019.

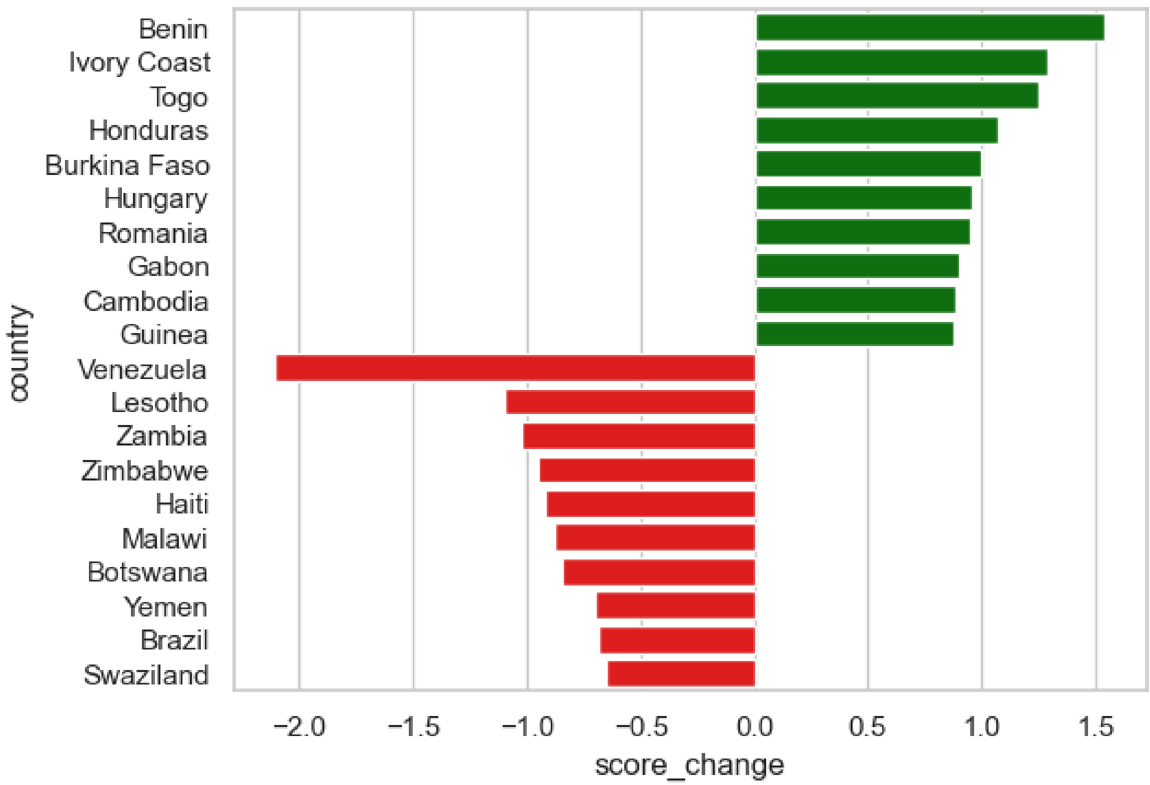
In [121…

```
top_increase=df.sort_values(by='score_change', ascending=False).head(10)
top_decrease=df.sort_values(by='score_change', ascending=True).head(10)
change_df = pd.concat([top_increase,top_decrease])

sns.barplot(data=change_df, x='score_change', y='country',
            palette= ['green' if x>0 else 'red' for x in change_df['score_
```

```
/var/folders/mc/346138y553185_ymspptqg200000gn/T/ipykernel_98104/355556863
6.py:5: FutureWarning:



Passing `palette` without assigning `hue` is deprecated and will be remove
d in v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for
the same effect.
```

Out[121…

```
<Axes: xlabel='score_change', ylabel='country'>
```



This horizontal bar chart highlights the top 10 countries that improved the most in happiness score and the top 10 that declined the most from 2015 to 2019.

✅ Top gainers include countries like Benin, Ivory Coast, and Togo, which saw dramatic improvements in well-being. ❌ Top decliners include Venezuela, Lesotho, and Zambia, reflecting political instability, economic crisis, or social issues.

The plot uses green bars for improvements and red bars for declines. The dashed vertical line at 0 represents no change.

C. Comparison of Global Averages (2015 vs 2019)

In [122…

```
df.head()
```

Out[122...

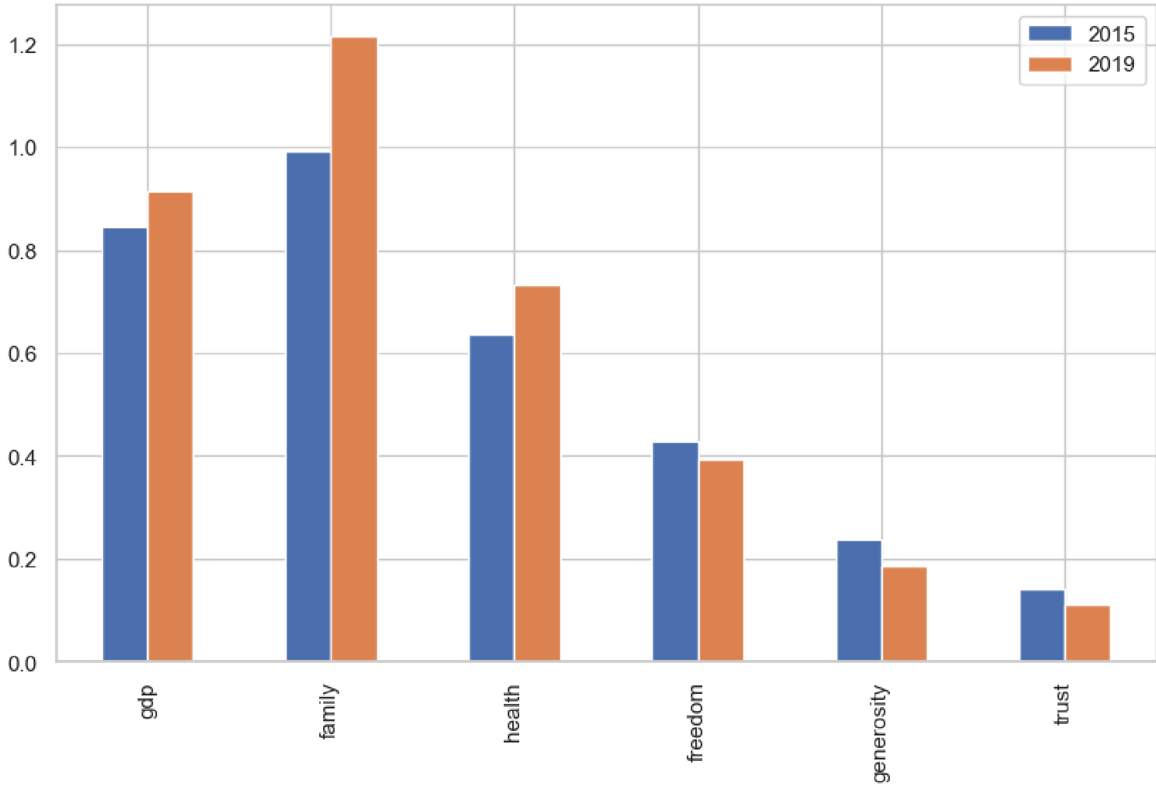
	country	rank_15	score_15	gdp_15	family_15	health_15	freedom_15	trus
0	Switzerland	1	7.587	1.39651	1.34951	0.94143	0.66557	0.41
1	Iceland	2	7.561	1.30232	1.40223	0.94784	0.62877	0.14
2	Denmark	3	7.527	1.32548	1.36058	0.87464	0.64938	0.48
3	Norway	4	7.522	1.45900	1.33095	0.88521	0.66973	0.36
4	Canada	5	7.427	1.32629	1.32261	0.90563	0.63297	0.32

5 rows × 24 columns

In [138...

```
features = ['gdp', 'family', 'health', 'freedom', 'generosity', 'trust']
data = {
    '2015': [df[f'{f}_15'].mean() for f in features],
    '2019': [df[f'{f}_19'].mean() for f in features]
}

mean_df = pd.DataFrame(data, index=features)

mean_df.plot(kind='bar', figsize=(10, 6))
```

Out[138... <Axes: >



In [131...

```
# This grouped bar chart compares the global average values of key happin
# 2015 and 2019.

# GDP, Health, and Family Support showed noticeable increases, suggesting
# social development.

# However, Freedom, Generosity, and Trust in Government slightly declined
# in perceived autonomy, charitable giving, and institutional trust.
```

```
# This plot highlights that while material and social well-being improved
# and civic values may have weakened globally between 2015 and 2019.
```

📝 Final Summary This project explored global patterns in happiness using the World Happiness Reports from 2015 and 2019, combining data analysis and visualization to understand the key drivers of well-being across countries.

We conducted a full Exploratory Data Analysis (EDA) that included:

Data Cleaning & Feature Engineering: Combined and aligned 2015 and 2019 datasets, created change variables.

Univariate Analysis: Studied distribution of happiness scores and contributing factors like GDP, Family, Health, Freedom, Trust, and Generosity for both years.

Bivariate Analysis: Investigated relationships between features and happiness scores using correlation and regression plots.

Feature Impact: Found that Health, Family, and GDP had the strongest positive correlation with happiness; Trust and Generosity had minimal impact.

Score Change Analysis: Identified countries with the highest increases (e.g., Benin, Togo) and decreases (e.g., Venezuela, Zambia) in happiness.

Dashboards:

🌍 World Map of 2019 Happiness Scores

🔼🔽 Top Countries by Score Change

📊 Comparison of Global Feature Averages (2015 vs 2019)

🎯 Conclusion This analysis shows that:

Material well-being (GDP) and social support (Family, Health) are major contributors to happiness.

Trust in government, Freedom, and Generosity saw global declines from 2015 to 2019.

Several developing countries made major progress, while some wealthier nations saw a decline, indicating that happiness isn't purely tied to income.

Overall, this project provides a meaningful look at what shapes happiness globally, and how priorities might shift for policymakers, development programs, and individuals.

Data Source: World Happiness Reports (2015, 2019) via Kaggle.

In [ ]: