IEG World Bank Project Performance Ratings:

Evaluating What Performance Results Mean for Regions and Project Output

Luciano Colla Hadad

Overview

This report provides an analysis of the World Bank Project Performance Ratings dataset, focusing on various aspects of project success and performance. The aim of this report is to highlight key trends within the data while offering potential recommendations to improve the efficiency and/or quality of projects that the World Bank decides to pursue.

Purpose

The World Bank is a global financial institution that provides loans and grants to the governments of low and middle-income countries with the aim to reduce poverty and support development. Given the nature of such an institution, it is important to assess the successfulness of projects of which much resources are being invested in.

Objectives

The primary objective of this analysis is to look at how successful projects are across different regions, the impact the type of global practice has to their successfulness, the correlation between bank performance and monitoring quality, the patterns between the completion year and project sustainability, as well as general trends from the different country lending groups.

Questions

To accomplish these objectives, we asked 5 questions.

1. Which regions and countries have the highest and lowest project success rates?

2. How does the 'Global Practice' impact the likelihood of project success across different regions?

3. Does the 'M&E Quality' (Monitoring and Evaluation Quality) correlate with the overall 'Bank Performance'?

4. What patterns emerge from the 'Final Closing FY' in relation to project success and sustainability?

5. Are there specific 'Country Lending Groups' that show consistent trends in project performance?

Importing the Data

```python
# Import python packages
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import geopandas as gpd
from tabulate import tabulate
from scipy.stats import f_oneway

# Load CSV file into dataframe
file = 'IEG World Bank Project Performance Ratings.csv'
df = pd.read_csv(file)
```

Data Preparation

```python
[2]  # Rename 'WB Region' column to 'Region'
     df.rename(columns = {'WB Region' : 'Region'}, inplace = True)

     # Extract unique regions and countries
     region_country_mapping = df[['Region', 'Country']].drop_duplicates()

     # Create a dictionary of regions and corresponding countries
     region_country_dict = region_country_mapping.groupby('Region')['Country'].apply(list).to_dict()

     # Define numeric scales for performance ratings
     satisfaction_scale = {'Highly Satisfactory': 6,
                           'Satisfactory': 5,
                           'Moderately Satisfactory': 4,
                           'Moderately Unsatisfactory': 3,
                           'Unsatisfactory': 2,
                           'Highly Unsatisfactory': 1}

     monitor_eval_scale = {'High': 4,
                           'Substantial': 3,
                           'Modest': 2,
                           'Negligible': 1}

     # Add numeric performance columns
     df['Numeric Outcome'] = df['Outcome'].replace(satisfaction_scale)
     df['Numeric Entry'] = df['Quality at Entry'].replace(satisfaction_scale)
     df['Numeric Supervision'] = df['Quality of Supervision'].replace(satisfaction_scale)
     df['Numeric BP'] = df['Bank Performance'].replace(satisfaction_scale)
     df['Numeric M&E'] = df['M&E Quality'].replace(monitor_eval_scale)

     # Set color palette
     colors = sns.color_palette('Paired')
```

} ChatGPT Assisted

Q1. Project Success Rate

```python
# Create a new DataFrame for outcome satisfaction, filtering out missing values
outcome_satisfaction = df.dropna(subset=['Outcome']).copy()

# Group by 'Region' and calculate the average numeric outcome
outcome_region_success = outcome_satisfaction.groupby('Region').agg(
    Rate=('Numeric Outcome', 'mean'),
    Projects=('Country', 'count'),
    Country=('Country', 'nunique')
)
outcome_region_success.rename(columns={'Rate': 'Success Rate'}, inplace=True)

# Add count and % total of successes to region_sucess DataFrame
outcome_region_success['Count'] = outcome_satisfaction[outcome_satisfaction['Numeric Outcome'] > 3].groupby('Region').size()
outcome_region_success['% of Total'] = (outcome_region_success['Count'] / outcome_region_success['Projects']).round(4) * 100


# Sort by success rate in descending order
outcome_region_success = outcome_region_success.sort_values(by='Success Rate').reset_index()
```

```python
# Group by 'Region' and calculate the average numeric outcome
outcome_region_success = outcome_satisfaction.groupby('Region').agg(
    Rate=('Numeric Outcome', 'mean'),
    Projects=('Country', 'count'),
    Country=('Country', 'nunique')
)
outcome_region_success.rename(columns={'Rate': 'Success Rate'}, inplace=True)
```


Average Project Success Rate by Region

The **success rate** is calculated based on the 'Numeric Outcome' of which is aggregated by the average.
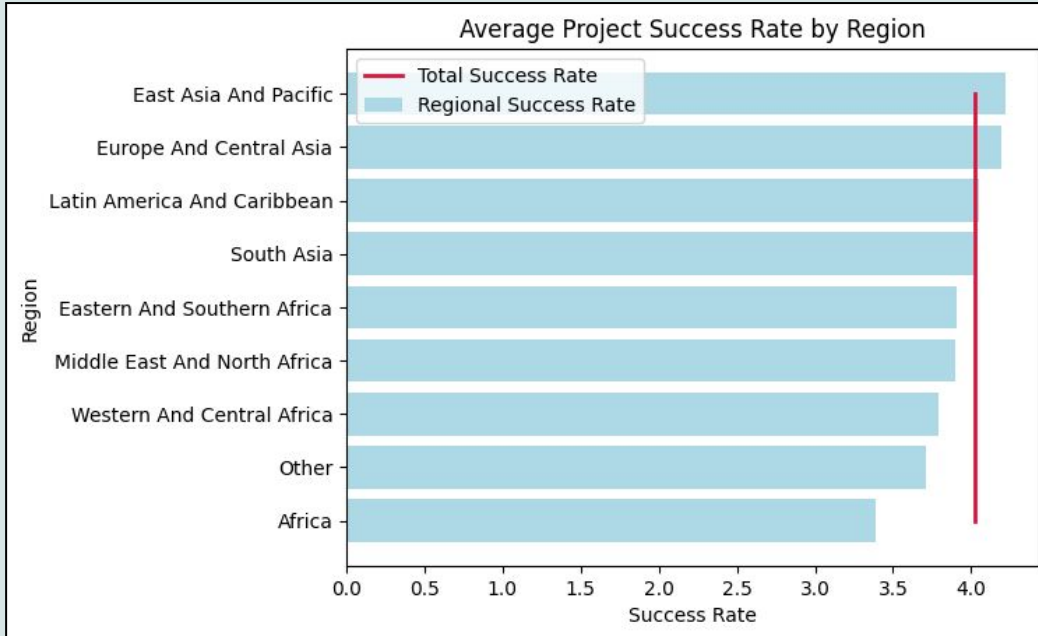
Q1. Success Rate Across Regions

Average Success Rate: 4.02

Africa – 3.39
Moderately Unsatisfactory

Other (World) – 3.71
Moderately Unsatisfactory

Europe and Central Asia – 4.20
Moderately Satisfactory

East Asia and Pacific – 4.22
Moderately Satisfactory

Region	Success Rate
Africa	3.39130
Other	3.71429
Western And Central Africa	3.79040
Middle East And North Africa	3.90118
Eastern And Southern Africa	3.91162
South Asia	4.01389
Latin America And Caribbean	4.05052
Europe And Central Asia	4.19624
East Asia And Pacific	4.22403

Q1. Pie Chart Code

```python
# --------------------------------- Pie Chart ---------------------------------

# Count the number of projects in Africa and Non-Africa regions
africa_regions = outcome_region_success['Projects'][outcome_region_success['Region'].str.contains('Africa')].sum()
non_africa_regions = outcome_region_success['Projects'][~outcome_region_success['Region'].str.contains('Africa')].sum()

# Create a DataFrame to hold the counts
africa_counts = pd.DataFrame({'Region': ['Africa Regions', 'Non-Africa Regions'], 'Projects': [africa_regions, non_africa_

# Plot
plt.figure(figsize=(12, 3))
plt.pie(africa_counts['Projects'], labels=africa_counts['Region'], autopct='%1.1f%%', startangle=140, colors=colors, textp
plt.title('Proportion of Projects in Africa vs. Non-Africa Regions', size=14)
plt.axis('equal')  # Equal aspect ratio ensures that pie is drawn as a circle.
plt.show()
```

Separates Regions

Region	Success Rate	Projects	Country	Count	% of Total
Africa	3.39130	46	1	24	52.17
Other	3.71429	7	1	3	42.86
Western And Central Africa	3.79040	854	23	566	66.28
Middle East And North Africa	3.90118	425	13	305	71.76
Eastern And Southern Africa	3.91162	792	27	571	72.10
South Asia	4.01389	648	9	484	74.69
Latin America And Caribbean	4.05052	1049	36	809	77.12
Europe And Central Asia	4.19624	958	33	769	80.27
East Asia And Pacific	4.22403	924	25	746	80.74

Proportion of Projects in Africa vs. Non-Africa Regions

Non-Africa Regions 62.9%
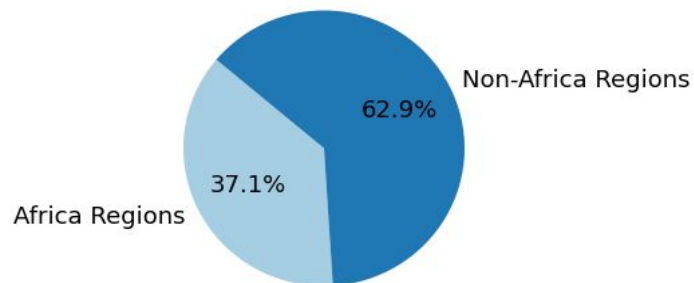
Africa Regions 37.1%

Q1. Recommendations

Given the lower success rates of projects in African regions but their high project count, it is recommended that the World Bank focuses on regional projects within Africa rather than continental projects.

This approach may improve project success rates by addressing specific regional needs.

Q2. Global Practice Impact

We analyzed the impact of Global Practice on project success across different regions to identify high-performing practices and regions, and to provide tailored recommendations for improvement.

We prepared the data and created 3 visualizations:

- Stacked Bar of Average Numeric Outcome of Global Practice Per Region
- Heat Map of Average Numeric Outcome per Region
- Summary Table of Numeric Outcome per Region

Q2. Global Practice Impact Base Code

```python
# Group by 'Global Practice' and 'Region', then calculate mean numeric outcome
grouped_data = outcome_satisfaction.groupby(['Global Practice', 'Region'])['Numeric Outcome'].mean().reset_index()

# Pivot the grouped data for easier plotting
pivot_data = grouped_data.pivot(index='Global Practice', columns='Region', values='Numeric Outcome')

# Sort the pivot data by the total average numeric outcome in descending order
pivot_data['Total'] = pivot_data.sum(axis=1)

# Calculate total for sorting, then sort and drop the total column
pivot_data = pivot_data.sort_values(by='Total', ascending=False).drop(columns='Total')
```

Q2. Numeric Outcome by Global Practice and Region

```python
pivot_data.plot(kind='barh', stacked=True, figsize=(14, 8), width=0.8, color=colors)
plt.title('Average Numeric Outcome by Global Practice (Stacked by Region)')
plt.xlabel('Average Numeric Outcome')
plt.ylabel('Global Practice')
plt.legend(title='Region', bbox_to_anchor=(1, 1), loc='upper right')
```


Average Numeric Outcome by Global Practice (Stacked by Region)
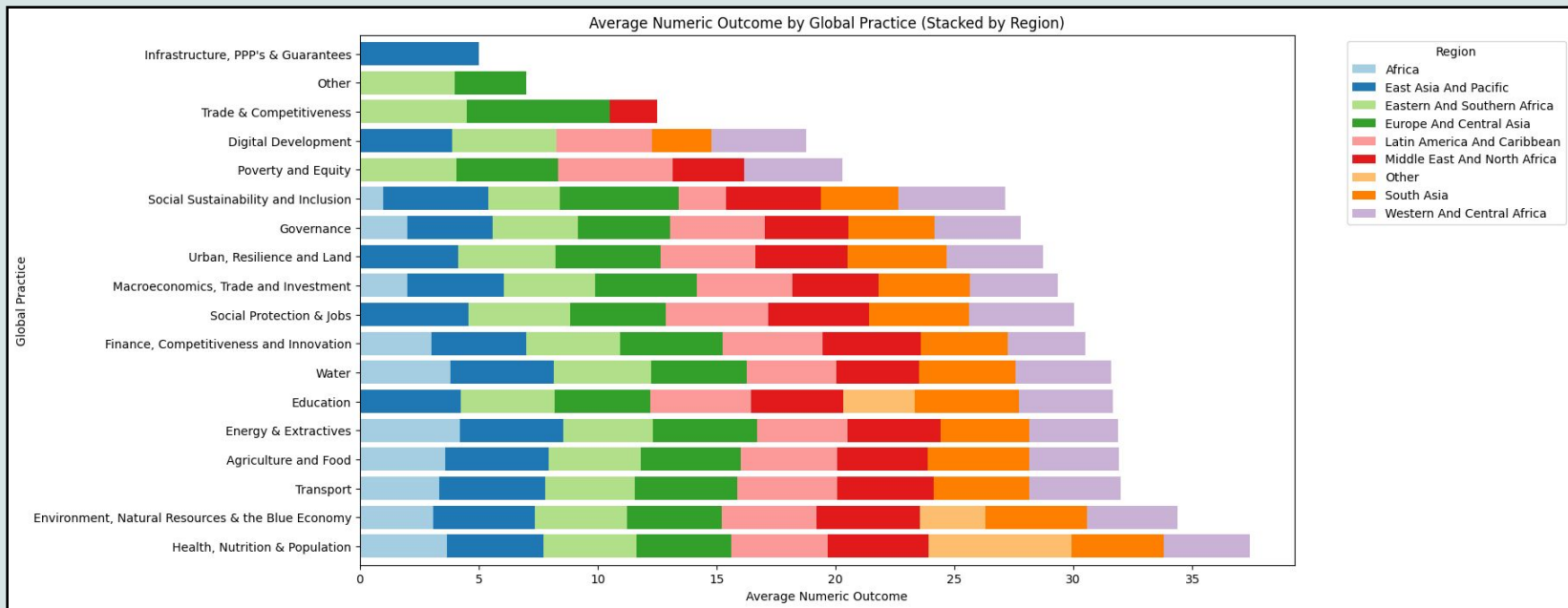
Q2. AvG Numeric Outcome by Region Heat Map

```python
# Load the world map shapefile
world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))

# Define region mapping based on continents/sub-regions
region_mapping = {
    'Africa': ['Africa'],
    'East Asia And Pacific': ['Asia', 'Oceania'],
    'Eastern And Southern Africa': ['Africa'],
    'Europe And Central Asia': ['Europe', 'Asia'],
    'Latin America And Caribbean': ['South America', 'North America'],
    'Middle East And North Africa': ['Africa', 'Asia'],
    'South Asia': ['Asia'],
    'Western And Central Africa': ['Africa']
}

# Group by 'Region' and calculate the average numeric outcome
region_avg_outcome = outcome_satisfaction.groupby('Region')['Numeric Outcome'].mean().reset_index()

# Expand the region_avg_outcome dataframe to map multiple continents to each region
expanded_data = []
for region, continents in region_mapping.items():
    if region in region_avg_outcome['Region'].values:
        average_outcome = region_avg_outcome.loc[region_avg_outcome['Region'] == region, 'Numeric Outcome'].values[0]
        for continent in continents:
            expanded_data.append([continent, average_outcome])

# Create a new dataframe with the expanded data
expanded_region_avg = pd.DataFrame(expanded_data, columns=['Continent', 'Numeric Outcome'])

# Group by continent to get the average numeric outcome
continent_avg_outcome = expanded_region_avg.groupby('Continent')['Numeric Outcome'].mean().reset_index()

# Merge the continental averages with the world map
world['Continent'] = world['continent'].replace({'Seven seas (open ocean)': 'Other'})
merged = world.set_index('Continent').join(continent_avg_outcome.set_index('Continent'), how='left')

# Plot
fig, ax = plt.subplots(1, 1, figsize=(15, 12))
world.boundary.plot(ax=ax, linewidth=1)
merged.plot(column='Numeric Outcome', ax=ax, legend=True,
            legend_kwds={'label': "Average Numeric Outcome by Continent",
                         'orientation': "horizontal"},
            missing_kwds={"color": "lightgrey", "label": "No Data"})

plt.title('Average Project Performance Outcomes - Heatmap')

# Adjust axis limits to make the map rectangular
ax.set_xlim(-180, 180)
ax.set_ylim(-60, 90)
ax.set_aspect(1.5)

plt.show()
```
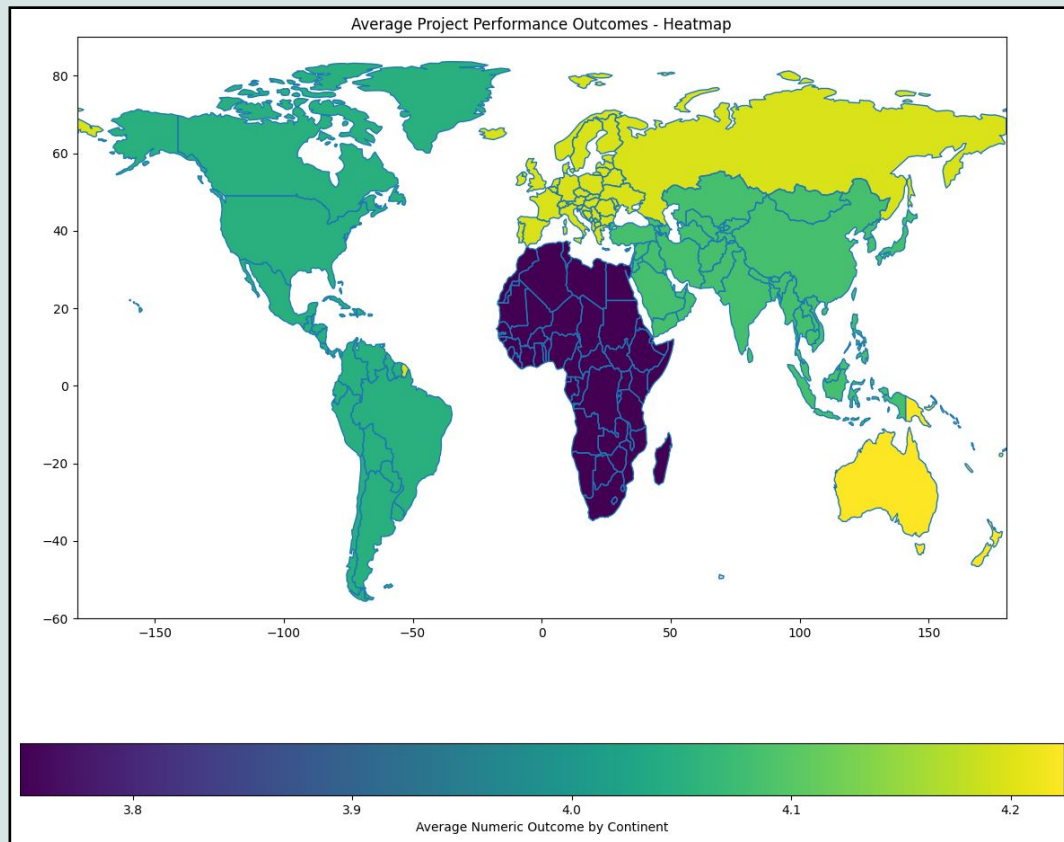

Average Project Performance Outcomes - Heatmap

Q2. Analysis Summary

Summary Table:

	Continent	Numeric Outcome
0	Africa	3.74862
1	Asia	4.08383
2	Europe	4.19624
3	North America	4.05052
4	Oceania	4.22403
5	South America	4.05052

- **Higher Success Rates**: East Asia and Pacific and Europe and Central Asia in:
  - Infrastructure, PPPs & Guarantees
  - Trade & Competitiveness
  - Finance, Competitiveness and Innovation

- **Lower Success Rates**: African Regions in:
  - Agriculture and Food
  - Education

Q2. Recommendations

'Agriculture and Food' and 'Education' Global Practices in African regions require tailored strategies and additional support to enhance their effectiveness. High-performing regions can serve as models for best practices.

'Infrastructure, PPPs & Guarantees' and 'Finance, and Competitiveness and Innovation' Global Practices in regions such as 'East Asia and Pacific' and 'Europe and Central Asia' can be good models

Focus should be made on region-specific interventions to enhance project effectiveness, especially in Africa and other lower-performing areas.

Q3. M&E Quality and Bank Performance

Analyzing the relationship between M&E Quality and Bank Performance to determine the impact of high-quality monitoring and evaluation practices.

Methodology:

- **Correlation Analysis:**
  - Calculated the correlation coefficient between M&E Quality and Bank Performance.
- **Visualization:**
  - Created a box plot to visually represent the relationship.
  - Highlighted the average bank performance across different M&E Quality ranks.

Q3. Summary Table and Correlation Coefficient

	M&E Quality	BP Average	BP Count
0	Negligible	2.99786	468
1	Modest	3.81049	2174
2	Substantial	4.49578	1658
3	High	5.08621	116

Correlation Coefficient: 0.57

```python
# -------------------------------- Summary Table --------------------------------

# Calculate the average numeric BP for each M&E quality
me_avg_bp = me_quality_performance.groupby('M&E Quality').agg(
    Average=('Numeric BP', 'mean'),
    Count=('Numeric BP', 'count')
).reset_index()
me_avg_bp.rename(columns={'Average': 'BP Average', 'Count': 'BP Count'}, inplace=True)
me_avg_bp.sort_values(by='BP Average', inplace=True)
me_avg_bp.reset_index(drop=True, inplace=True)

# Print the formatted table
print(tabulate(me_avg_bp, headers='keys', tablefmt='psql'))

# Calculate the correlation coefficient
bp_correlation = me_quality_performance[['Numeric M&E', 'Numeric BP']].corr().iloc[0, 1]

# Present the correlation coefficient
print(f"Correlation coefficient: {bp_correlation:.2f}\n")
```

Q3. Box Plot Code

```python
# --------------------------------- Box Plot ---------------------------------

# Create the box plot
plt.figure(figsize=(7, 5))
sns.boxplot(x='M&E Quality', y='Numeric BP', data=me_quality_performance, color='lightblue')

# Overlay a point plot for mean values
sns.pointplot(x='M&E Quality', y='Numeric BP', data=me_quality_performance, estimator=np.mean, color='crimson')

# Add labels and title
plt.xlabel('M&E Quality')
plt.ylabel('Bank Performance')
plt.title('Box Plot of Bank Performance by M&E Quality')

# Set custom tick labels for y-axis
plt.yticks(ticks=list(satisfaction_scale.values()), labels=list(satisfaction_scale.keys()))

# Show the plot
plt.show()
```

(parameter) data: Any | None

Q3. Box Plot Visualization
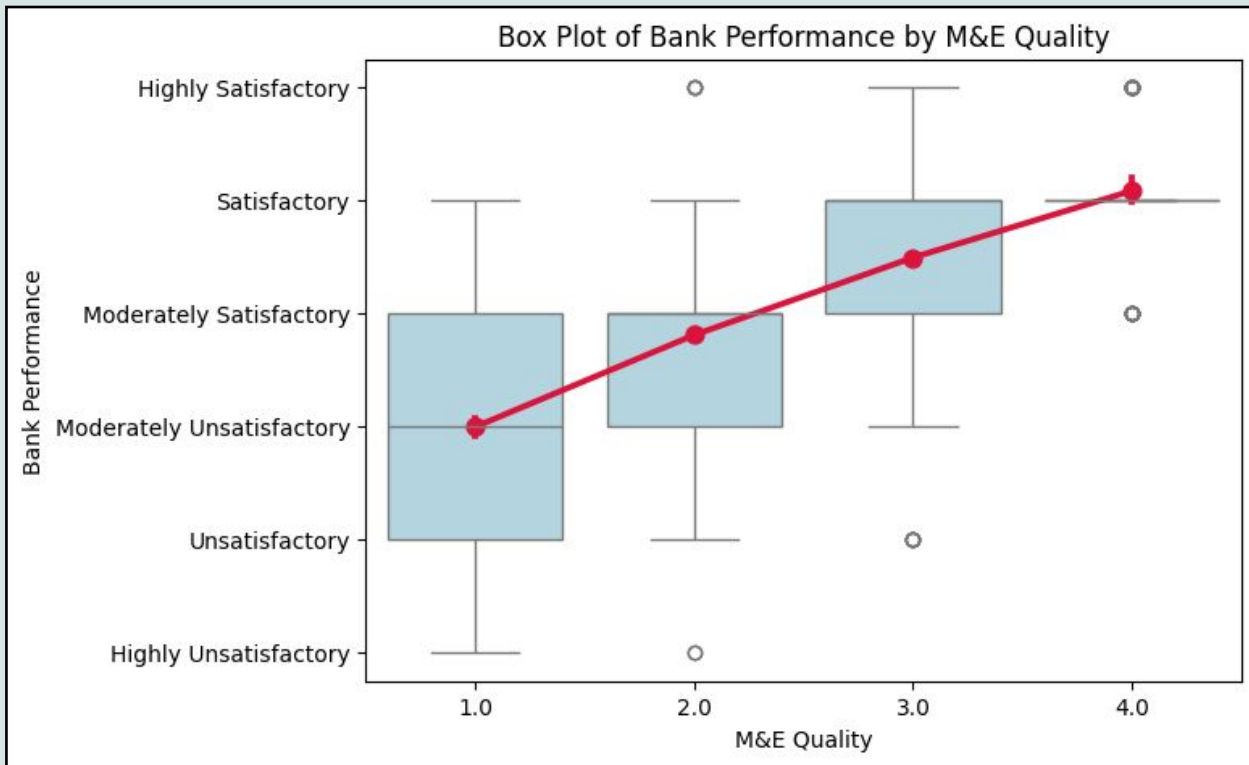
Q3. Analysis Summary

Key Findings:

- Correlation Coefficient: **0.57** (Moderate Positive Correlation)
- High M&E Quality projects achieved the highest mean bank performance.
- Negligible M&E Quality projects had the lowest performance metrics.

Conclusion:

This analysis shows the importance of investing in high-quality monitoring and evaluation practices to enhance overall bank performance.

Q3. Recommendations

There is a clear moderate positive correlation between high-quality Monitoring and Evaluation (M&E) practices and better bank performance.

It is recommended to invest in robust M&E frameworks to ensure projects remain on track, issues are addressed promptly, and intended outcomes are achieved, leading to better overall project success.

Q4. Final Closing Fiscal Year

We analyzed the relationship between the final closing fiscal year of projects and their success and sustainability, as it can provide valuable insights for project planning and management.

To analyze and find trends over time, we created a timeline graph showing how performance and sustainability scores changed over the years.

Q4. Performance and Sustainability Trends across Final Closing Fiscal Year

```python
# Visualization: Line plot for performance trends over time
plt.figure(figsize=(14, 8))
sns.lineplot(x='Final Closing FY', y='M&E Quality Numeric', data=closing_year_performance, marker='o', label='Average M&E Quality', markers=True)
sns.lineplot(x='Final Closing FY', y='Success', data=closing_year_performance, marker='o', label='Success Rate', markers=True)
plt.title('Project Performance and Sustainability Trends by Final Closing FY')
plt.xlabel('Final Closing Fiscal Year')
plt.ylabel('Average Rating / Success Rate')
plt.legend()
plt.show()
```
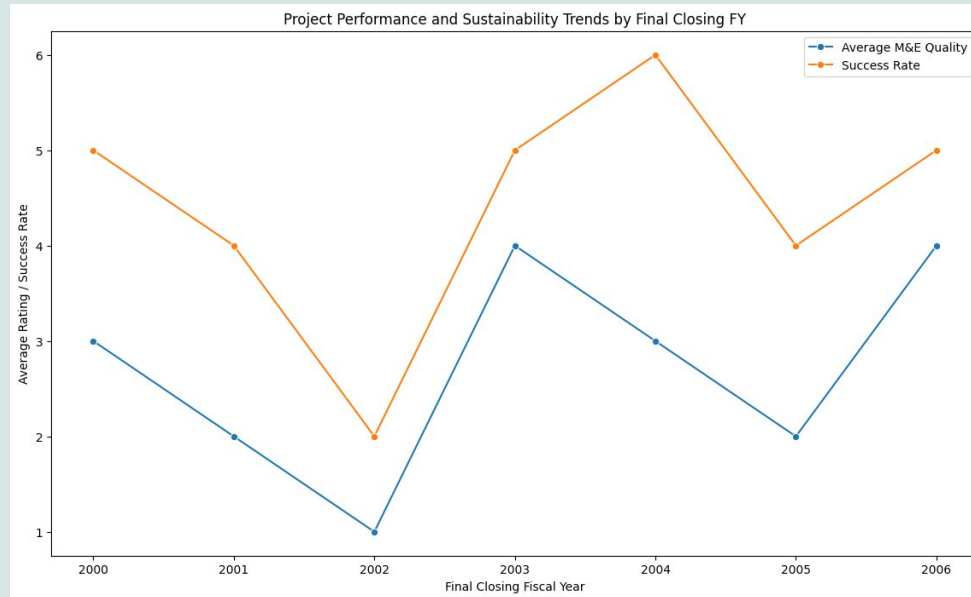
Q4. Analysis Summary

Initially, both M&E quality and success rates started relatively high in 2000 but then declined sharply until 2002, reaching their lowest points.

After 2002, there is a noticeable improvement, with both metrics peaking around 2003.

There is a strong correlation between Average M&E Quality and Success Rate.

Q4. Conclusion

Our analysis reveals a clear correlation between rigorous monitoring and evaluation (M&E) practices and the success rates of World Bank projects.

The data shows that projects with higher M&E quality tend to have higher success rates, highlighting the importance of comprehensive monitoring throughout the project lifecycle.

Q4. Recommendations

The fluctuations in M&E quality and success rates over different fiscal years suggest a need for ongoing efforts to stabilize and enhance project performance.

This includes adopting more hands-on and proactive M&E practices throughout the project lifecycle to maintain high standards and improve project outcomes.
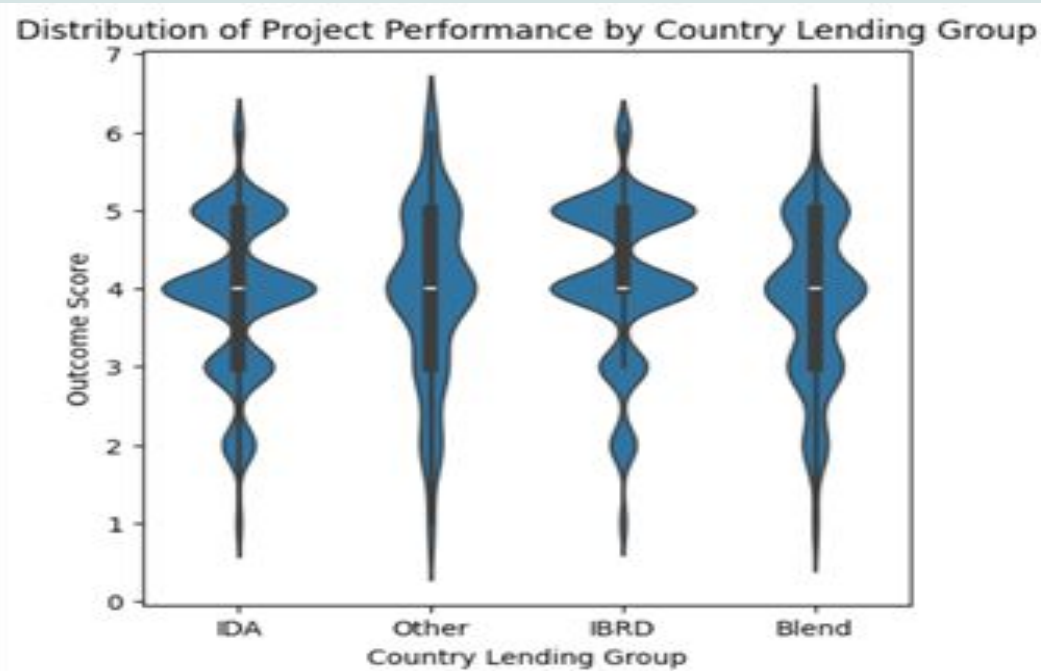
Q5. Country Lending Groups

Our study examined the trends in project performance across different 'Country Lending Groups'.

Our objective is to achieve valuable Insights for financial strategists and country program managers in order to influence decisions on engaging lending groups for future projects

We accomplished the goal by:

- Analyzing the outcomes categorized
- Identifying patterns that suggest consistent performance levels.
- Adapting statistical correlation to visual representations

```python
# --------------------------------- Violin Plot ---------------------------------

plt.figure(figsize=(5, 5))
sns.violinplot(x='Country Lending Group', y='Numeric Outcome', data=lending_satisfaction)
plt.title('Distribution of Project Performance by Country Lending Group')
plt.ylabel('Outcome Score')
plt.xlabel('Country Lending Group')
plt.yticks(ticks=list(satisfaction_scale.values()), labels=list(satisfaction_scale.keys()))
plt.show()
```


Distribution of Project Performance by Country Lending Group

```
# Create bar plot for average numeric outcome by country lending group
sns.barplot(x='Country Lending Group', y='Outcome Avg', data=lending_group_success, color='lightblue')

# Create standard error
std_error = lending_group_success['Outcome Stdv'] / np.sqrt(lending_group_success['Outcome Stdv'].count())

# Add error bars using outcome stdv
plt.errorbar(lending_group_success['Country Lending Group'], lending_group_success['Outcome Avg'], yerr=std_error,
```


Mean Project Performance by Country Lending Group
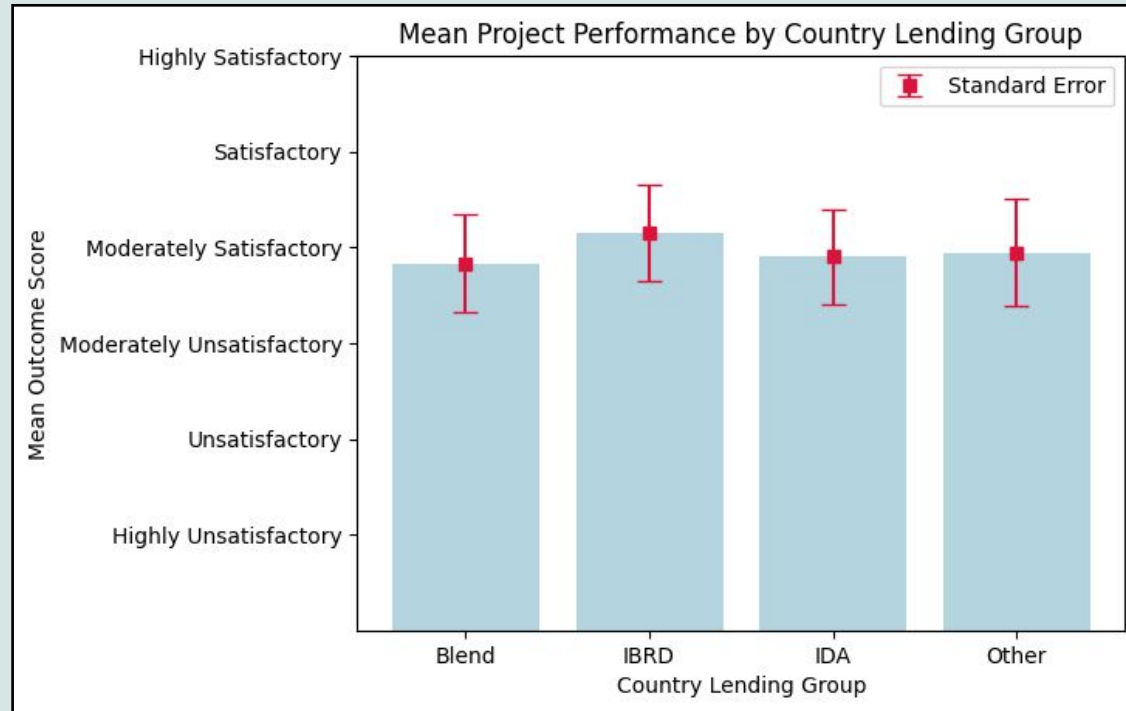
Q5. Analysis Summary

```python
# Print the formatted table
print(tabulate(lending_group_success, headers='keys', tablefmt='psql'))

# Calculate total success rate
lending_avg = lending_satisfaction['Numeric BP'].mean()
print(f"Average Bank Performance: {lending_avg:.2f}")
```

	Country Lending Group	Outcome Avg	Outcome Stdv
0	Blend	3.83200	1.023
1	IBRD	4.14900	1.005
2	IDA	3.90400	0.997
3	Other	3.94700	1.12

Consistency and Trends:

- IDA: Consistent performance with lower standard deviation.
- IBRD: Higher variability in outcomes.

Graph Analysis:

- IDA projects: Higher median performance scores.
- IBRD projects: More variability and spread in outcomes.

Q5. Conclusion

By examining the mean and standard deviation of 'Outcome' scores and visualizing the distribution of performance outcomes, it is possible to identify which groups are associated with higher or more consistent performance.

Analysis indicates specific consistent trends in 'IDA' group having the most stable outcomes, making it favorable for future projects.

Q5. Recommendations

The 'IDA' lending group shows consistent performance with lower variability and higher median performance scores.

Future engagements should prioritize the 'IDA' group, while factors causing variability in the 'IBRD' group should be investigated.

Enhanced monitoring practices should be implemented to optimize project performance and support strategic planning.

Closing Remarks

We implemented Python and Business Analytics to identify trends and patterns in the World Bank Project Performance Ratings dataset

By implementing our recommendations, the World Bank can:

- Enhance project efficiency and effectiveness

- Better support global development and

- Achieve its mission of reducing poverty.

Continuous assessment and adaptation are crucial to ensuring that resources are invested in the most impactful projects.

References

- Sullivan, D. (2015). *NoSQL for Mere Mortals*. Addison-Wesley Professional.

- The World Bank. (n.d.). https://www.worldbank.org/en/about/what-we-do.prin

- The World Bank. (2023, October 5). *Africa Overview*. https://www.worldbank.org/en/region/afr/overview

- Pandas Development Team. (2008). *pandas: Python Data Analysis Library*. https://pandas.pydata.org/

- Oliphant, T. (2006). *NumPy*. https://numpy.org/

- Hunter, J. D. (2003). *Pyplot API*. https://matplotlib.org/3.5.3/api/_as_gen/matplotlib.pyplot.html

- Waskom, M. (2014). *Seaborn*. https://seaborn.pydata.org/

References

- GeoPandas Development Team. (2013). *GeoPandas*. https://geopandas.org/en/stable/

- SciPy Community. (2001). *SciPy Stats*. https://docs.scipy.org/doc/scipy/reference/stats.html

- Lopuhin, K. (2010). *Tabulate*. https://pypi.org/project/tabulate/

- Pandas Development Team. (2008). *DataFrame.replace*.

  https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.replace.html

- Hunter, J. D. (2003). *Named colors*. https://matplotlib.org/stable/gallery/color/named_colors.html

- Practical Python for Data Science Team. (n.d.). *Seaborn Palette*.

  https://www.practicalpythonfordatascience.com/ap_seaborn_palette

- Stack Overflow Community. (2011, August 16). *How to set the label's size on a pie chart*.

  https://stackoverflow.com/questions/7082345/how-to-set-the-labels-size-on-a-pie-chart

THANK YOU!