ai_impact_analysis

February 8, 2026

1 Global AI Content Impact Dataset Analysis

This notebook provides a comprehensive analysis of the Global AI Content Impact Dataset, examining how artificial intelligence adoption affects various industries and countries.

[113]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from datetime import datetime
%matplotlib inline
```

[114]:
```
# Load the dataset
df = pd.read_csv('Global_AI_Content_Impact_Dataset.csv')
```

[115]:
```
# Display basic information about the dataset
print("Dataset Shape:", df.shape)
print("\nColumn Names:")
print(df.columns.tolist())
print("\nDataset Info:")
print(df.info())
```

```
Dataset Shape: (200, 12)

Column Names:
['Country', 'Year', 'Industry', 'AI Adoption Rate (%)', 'AI-Generated Content
Volume (TBs per year)', 'Job Loss Due to AI (%)', 'Revenue Increase Due to AI
(%)', 'Human-AI Collaboration Rate (%)', 'Top AI Tools Used', 'Regulation
Status', 'Consumer Trust in AI (%)', 'Market Share of AI Companies (%)']

Dataset Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 12 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   Country                                 200 non-null    object
 1   Year                                    200 non-null    int64
 2   Industry                                200 non-null    object
```

```
 3   AI Adoption Rate (%)                          200 non-null    float64
 4   AI-Generated Content Volume (TBs per year)    200 non-null    float64
 5   Job Loss Due to AI (%)                        200 non-null    float64
 6   Revenue Increase Due to AI (%)                200 non-null    float64
 7   Human-AI Collaboration Rate (%)               200 non-null    float64
 8   Top AI Tools Used                             200 non-null    object
 9   Regulation Status                             200 non-null    object
 10  Consumer Trust in AI (%)                      200 non-null    float64
 11  Market Share of AI Companies (%)              200 non-null    float64
dtypes: float64(7), int64(1), object(4)
memory usage: 18.9+ KB
None
```

[116]:
```
# Display first few rows
df.head()
```

[116]:

	Country	Year	Industry	AI Adoption Rate (%)	AI-Generated Content Volume (TBs per year)	Job Loss Due to AI (%)	Revenue Increase Due to AI (%)	Human-AI Collaboration Rate (%)	Top AI Tools Used	Regulation Status	Consumer Trust in AI (%)	Market Share of AI Companies (%)
0	South Korea	2022	Media	44.29	33.09	16.77	46.12	74.79	Bard	Strict	40.77	18.73
1	China	2025	Legal	34.75	66.74	46.89	52.46	26.17	DALL-E	Strict	35.67	35.02
2	USA	2022	Automotive	81.06	96.13	10.66	45.60	39.66	Stable Diffusion	Moderate	54.47	
3	France	2021	Legal	85.24	93.76	27.70	78.24	29.45	Claude	Moderate	51.84	
4	France	2021	Gaming	78.95	45.62	17.45	1.05	21.70	Midjourney	Strict	41.77	

```
2                              22.76
3                               1.93
4                              21.41
```

[117]:
```
# Summary statistics
df.describe()
```

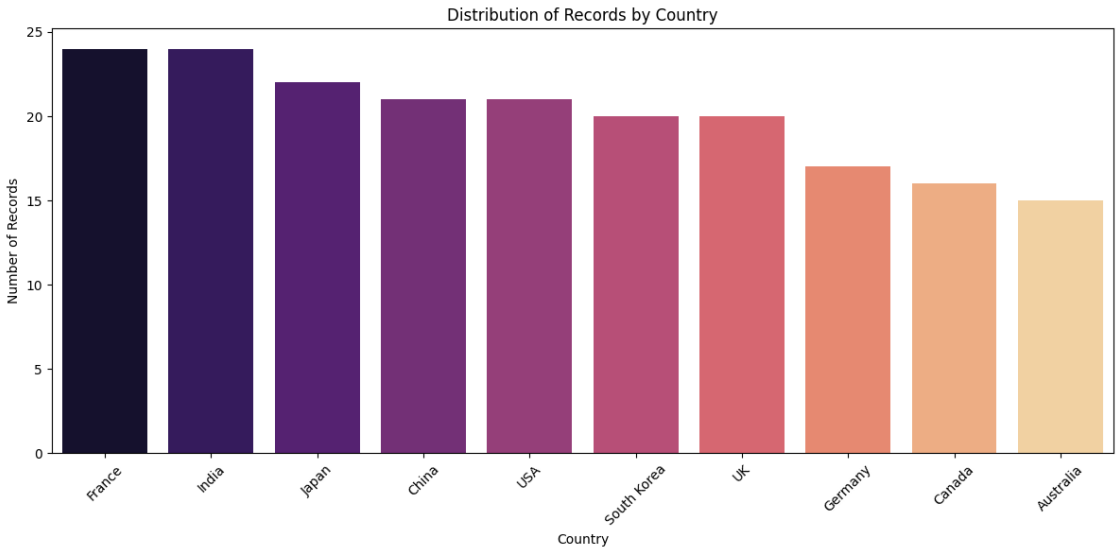
[117]:
```
              Year  AI Adoption Rate (%)  \
count   200.000000            200.000000
mean   2022.315000             54.265850
std       1.825496             24.218067
min    2020.000000             10.530000
25%    2021.000000             33.222500
50%    2022.000000             53.310000
75%    2024.000000             76.220000
max    2025.000000             94.760000

       AI-Generated Content Volume (TBs per year)  Job Loss Due to AI (%)  \
count                                   200.00000              200.000000
mean                                     46.07260               25.788250
std                                      29.16122               13.901105
min                                       1.04000                0.090000
25%                                      20.32250               14.995000
50%                                      44.32000               25.735000
75%                                      71.62000               37.417500
max                                      99.06000               49.710000

       Revenue Increase Due to AI (%)  Human-AI Collaboration Rate (%)  \
count                      200.000000                       200.000000
mean                        39.719450                        54.102150
std                         23.829545                        19.247079
min                          0.140000                        20.210000
25%                         17.907500                        37.770000
50%                         42.100000                        54.515000
75%                         58.697500                        69.402500
max                         79.550000                        88.290000

       Consumer Trust in AI (%)  Market Share of AI Companies (%)
count                200.000000                        200.000000
mean                  59.425150                         26.569550
std                   17.319668                         14.023729
min                   30.120000                          1.180000
25%                   44.755000                         14.052500
50%                   59.215000                         27.390000
75%                   74.885000                         38.432500
max                   89.880000                         49.280000
```

[118]:
```
# Check for missing values
df.isnull().sum()
```

[118]:
```
Country                                    0
Year                                       0
Industry                                   0
AI Adoption Rate (%)                       0
AI-Generated Content Volume (TBs per year)    0
Job Loss Due to AI (%)                     0
Revenue Increase Due to AI (%)             0
Human-AI Collaboration Rate (%)            0
Top AI Tools Used                          0
Regulation Status                          0
Consumer Trust in AI (%)                   0
Market Share of AI Companies (%)           0
dtype: int64
```

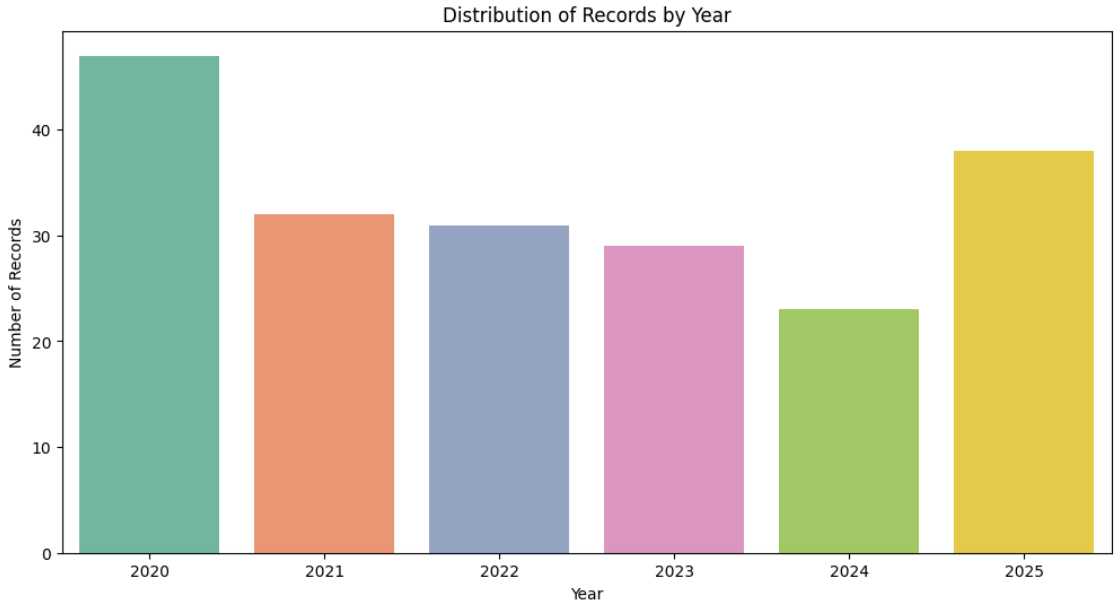
[119]:
```
# Distribution of countries
plt.figure(figsize=(12, 6))
country_counts = df['Country'].value_counts()
sns.barplot(x=country_counts.index, y=country_counts.values,palette='magma')
plt.title('Distribution of Records by Country')
plt.xticks(rotation=45)
plt.ylabel('Number of Records')
plt.tight_layout()
plt.show()
```

```
/tmp/ipykernel_692781/632759182.py:4: FutureWarning:

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

  sns.barplot(x=country_counts.index, y=country_counts.values,palette='magma')
```

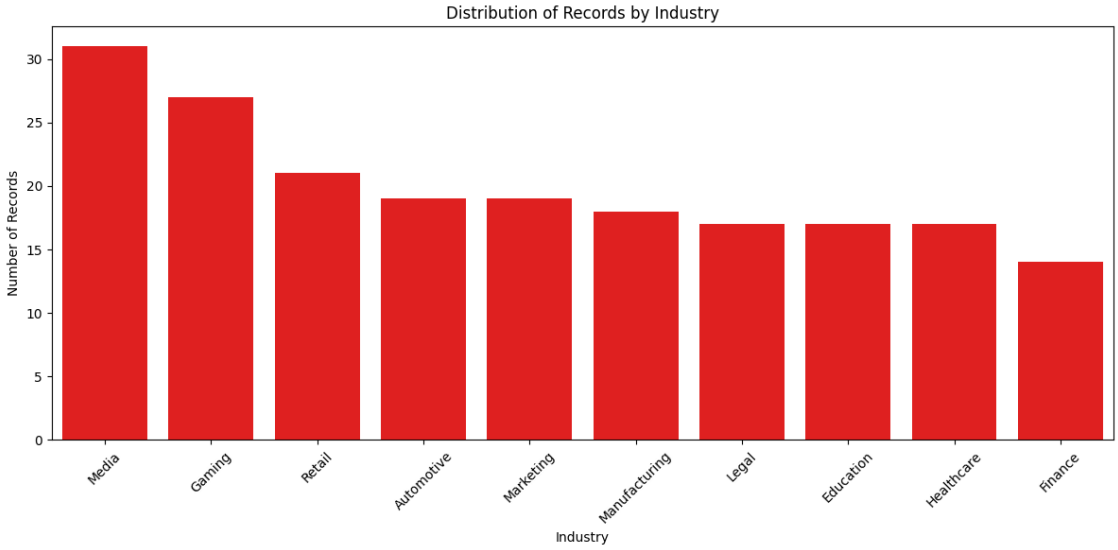


[136]:
```
# Distribution of years
plt.figure(figsize=(12, 6))
year_counts = df['Year'].value_counts().sort_index()
sns.barplot(x=year_counts.index, y=year_counts.values,palette='Set2')
plt.title('Distribution of Records by Year')
plt.xlabel('Year')
plt.ylabel('Number of Records')
plt.show()
```

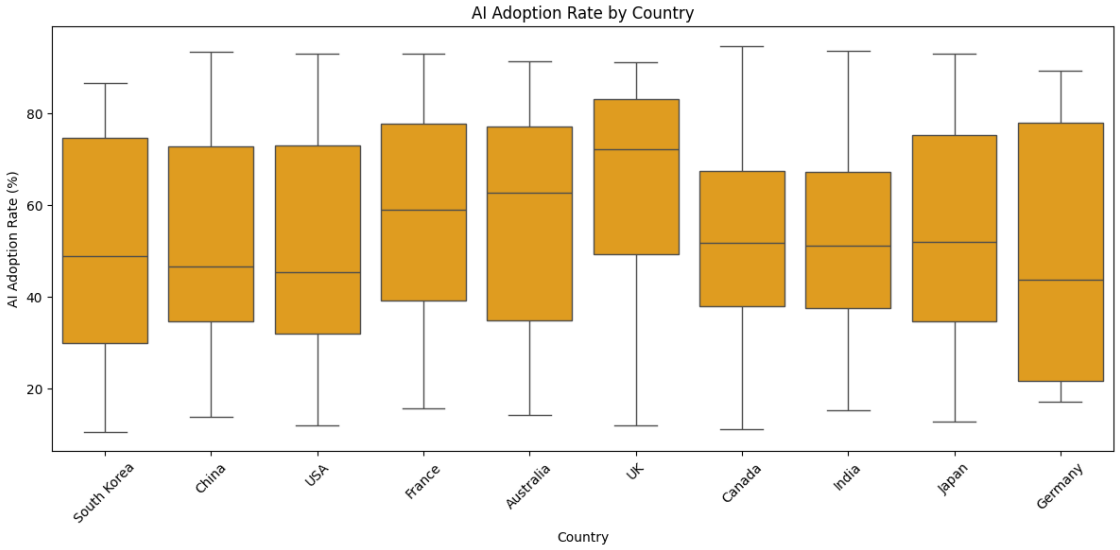
```
/tmp/ipykernel_692781/321646476.py:4: FutureWarning:

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

  sns.barplot(x=year_counts.index, y=year_counts.values,palette='Set2')
```



[155]:
```
# Distribution of industries
plt.figure(figsize=(12, 6))
industry_counts = df['Industry'].value_counts()
sns.barplot(x=industry_counts.index, y=industry_counts.values,color='red')
plt.title('Distribution of Records by Industry')
plt.xticks(rotation=45)
plt.ylabel('Number of Records')
plt.tight_layout()
plt.show()
```

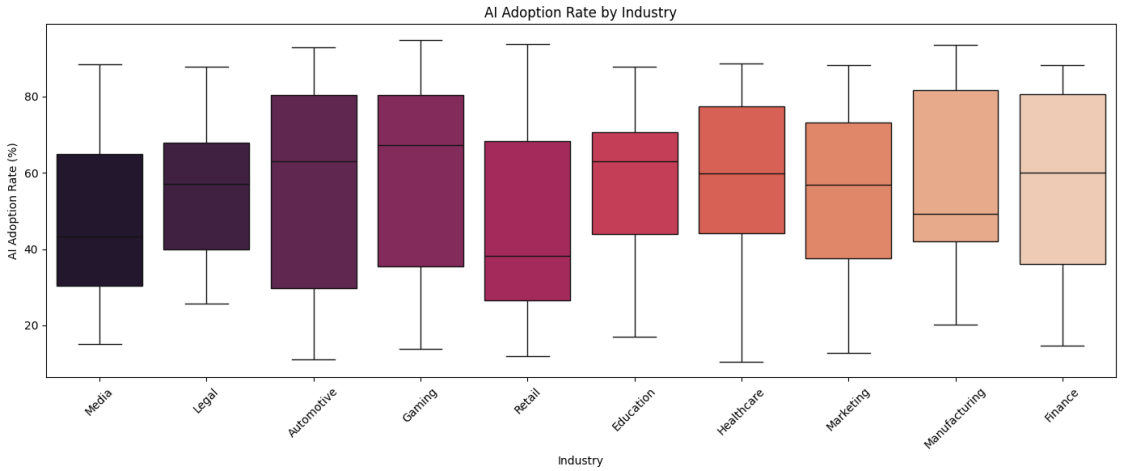


[140]:
```
# AI Adoption Rate by Country
plt.figure(figsize=(12, 6))
sns.boxplot(data=df, x='Country', y='AI Adoption Rate (%)',color='orange')
plt.title('AI Adoption Rate by Country')
plt.xticks(rotation=45)
plt.ylabel('AI Adoption Rate (%)')
plt.tight_layout()
plt.show()
```



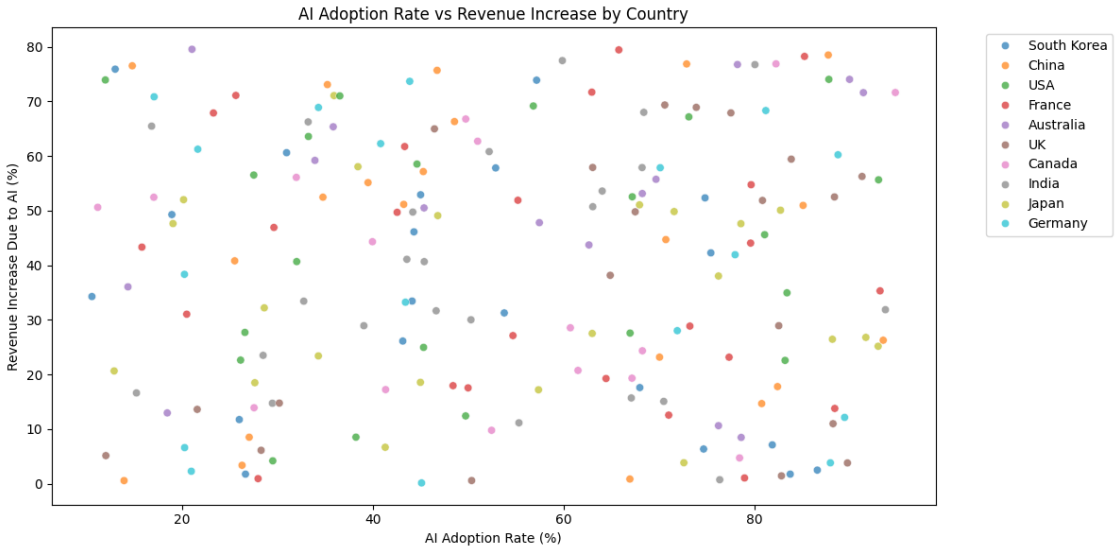
[151]:
```
# AI Adoption Rate by Industry
plt.figure(figsize=(14, 6))
sns.boxplot(data=df, x='Industry', y='AI Adoption Rate (%)',palette='rocket')
plt.title('AI Adoption Rate by Industry')
plt.xticks(rotation=45)
plt.ylabel('AI Adoption Rate (%)')
plt.tight_layout()
plt.show()
```

```
/tmp/ipykernel_692781/3762641706.py:3: FutureWarning:

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

  sns.boxplot(data=df, x='Industry', y='AI Adoption Rate (%)',palette='rocket')
```

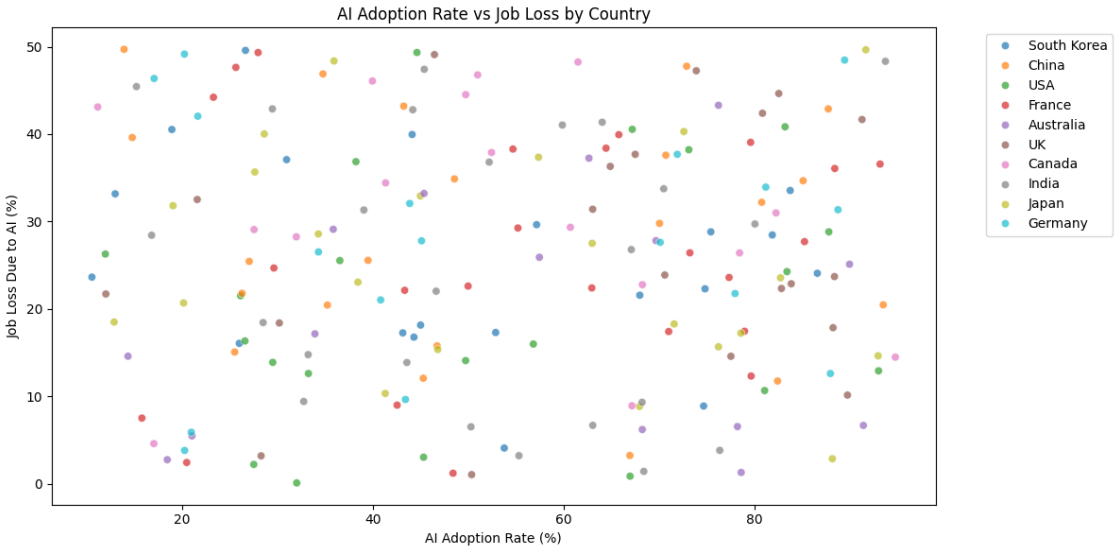


[143]:
```
# Correlation matrix for numerical variables
numerical_cols = [
    'AI Adoption Rate (%)',
    'AI-Generated Content Volume (TBs per year)',
    'Job Loss Due to AI (%)',
    'Revenue Increase Due to AI (%)',
    'Human-AI Collaboration Rate (%)',
    'Consumer Trust in AI (%)',
    'Market Share of AI Companies (%)'
]

correlation_matrix = df[numerical_cols].corr()
plt.figure(figsize=(12, 10))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', center=0,
            square=True, fmt='.2f')
plt.title('Correlation Matrix of Numerical Variables')
plt.tight_layout()
plt.show()
```

[144]:
```
# Relationship between AI Adoption Rate and Revenue Increase
plt.figure(figsize=(12, 6))
sns.scatterplot(data=df, x='AI Adoption Rate (%)', y='Revenue Increase Due to␣
↪AI (%)',
                hue='Country', alpha=0.7)
plt.title('AI Adoption Rate vs Revenue Increase by Country')
plt.xlabel('AI Adoption Rate (%)')
plt.ylabel('Revenue Increase Due to AI (%)')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```

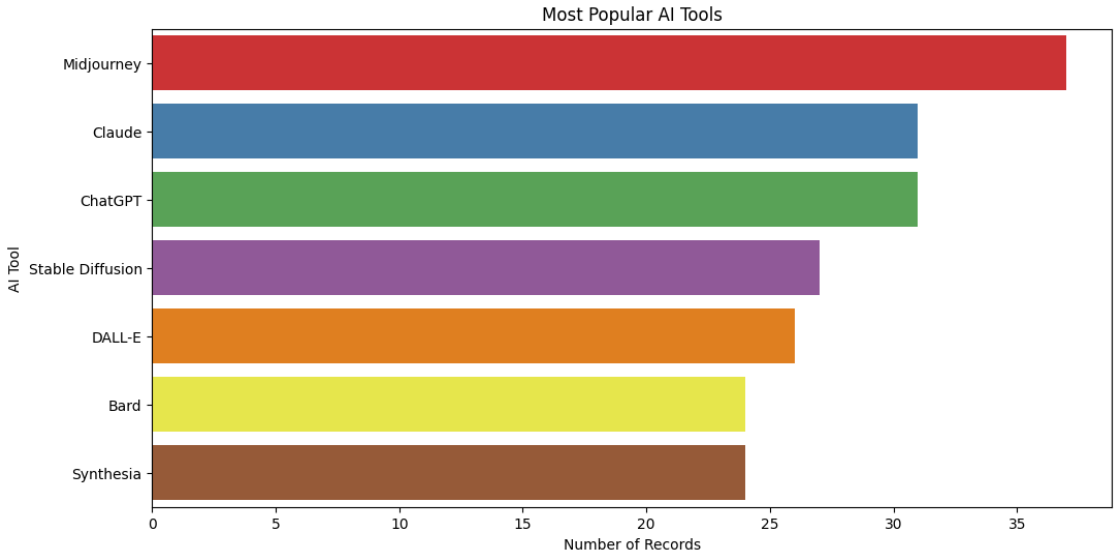


[145]:
```
# Relationship between Job Loss and AI Adoption Rate
plt.figure(figsize=(12, 6))
sns.scatterplot(data=df, x='AI Adoption Rate (%)', y='Job Loss Due to AI (%)',
                hue='Country', alpha=0.7)
plt.title('AI Adoption Rate vs Job Loss by Country')
plt.xlabel('AI Adoption Rate (%)')
plt.ylabel('Job Loss Due to AI (%)')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```

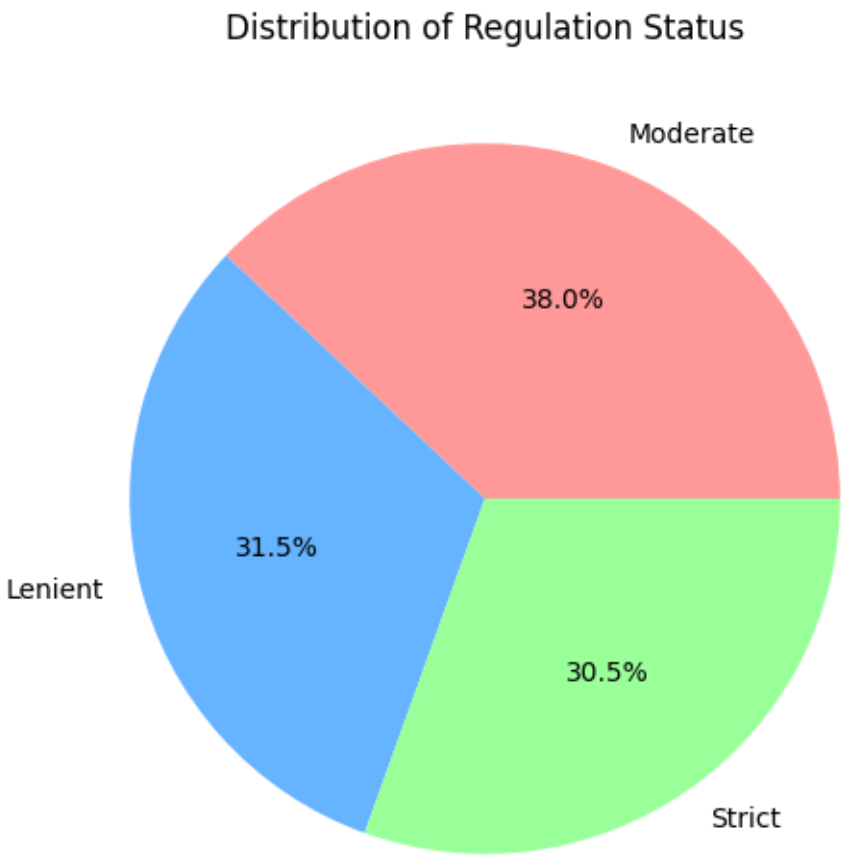


[146]:
```
# Top AI tools used
plt.figure(figsize=(12, 6))
tool_counts = df['Top AI Tools Used'].value_counts()
sns.barplot(x=tool_counts.values, y=tool_counts.index,palette='Set1')
plt.title('Most Popular AI Tools')
plt.xlabel('Number of Records')
plt.ylabel('AI Tool')
plt.show()
```

```
/tmp/ipykernel_692781/3483395137.py:4: FutureWarning:

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

  sns.barplot(x=tool_counts.values, y=tool_counts.index,palette='Set1')
```

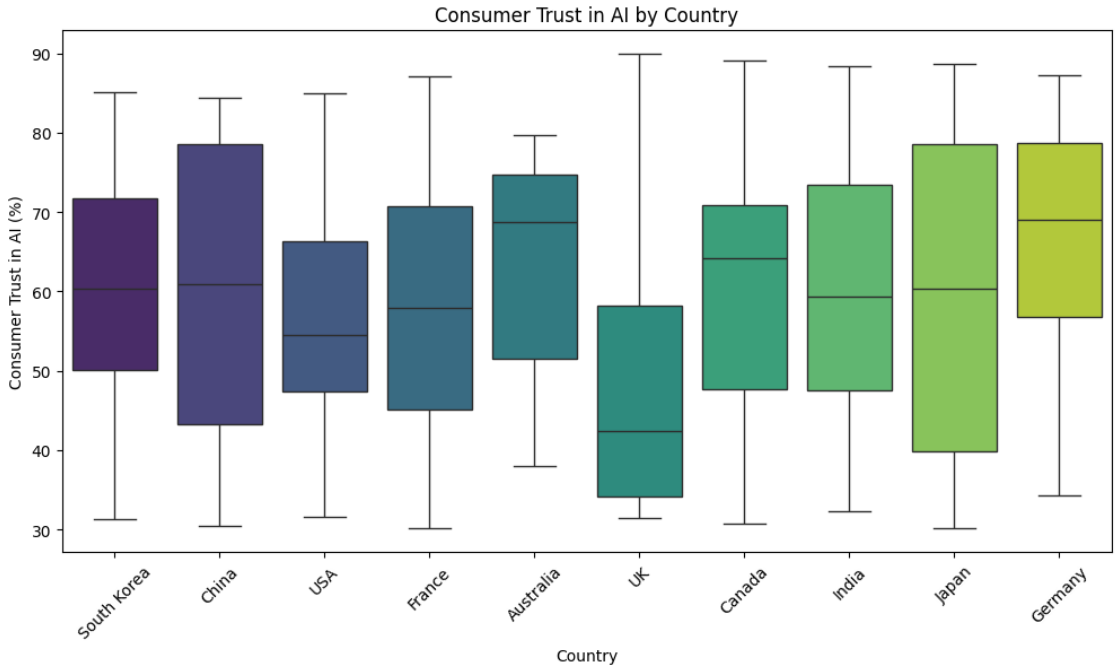


[147]:
```
# Regulation Status distribution
plt.figure(figsize=(12, 6))
regulation_counts = df['Regulation Status'].value_counts()
colors = ['#ff9999', '#66b3ff', '#99ff99']
plt.pie(regulation_counts.values, labels=regulation_counts.index, autopct='%1.
  ↪1f%%', colors=colors)
plt.title('Distribution of Regulation Status')
plt.show()
```



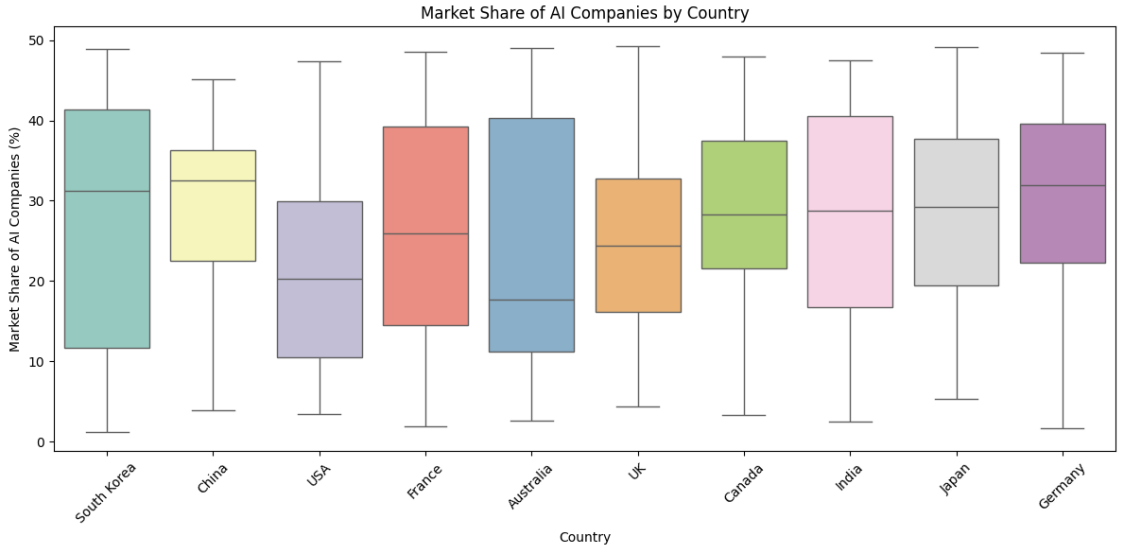
[156]:
```
# Consumer Trust by Country
plt.figure(figsize=(12, 6))
sns.boxplot(data=df, x='Country', y='Consumer Trust in AI␣
 ↪(%)',palette='viridis')
plt.title('Consumer Trust in AI by Country')
plt.xticks(rotation=45)
plt.ylabel('Consumer Trust in AI (%)')
plt.show()
```

```
/tmp/ipykernel_692781/4098806775.py:3: FutureWarning:

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

  sns.boxplot(data=df, x='Country', y='Consumer Trust in AI
(%)',palette='viridis')
```



[150]:
```
# Market Share of AI Companies by Country
plt.figure(figsize=(12, 6))
sns.boxplot(data=df, x='Country', y='Market Share of AI Companies␣
↪(%)',palette='Set3')
plt.title('Market Share of AI Companies by Country')
plt.xticks(rotation=45)
plt.ylabel('Market Share of AI Companies (%)')
plt.tight_layout()
plt.show()
```

```
/tmp/ipykernel_692781/1605442204.py:3: FutureWarning:

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

  sns.boxplot(data=df, x='Country', y='Market Share of AI Companies
(%)',palette='Set3')
```



[131]:
```
# Average metrics by country
country_metrics = df.groupby('Country').agg({
    'AI Adoption Rate (%)': 'mean',
    'Revenue Increase Due to AI (%)': 'mean',
    'Job Loss Due to AI (%)': 'mean',
    'Consumer Trust in AI (%)': 'mean',
    'Market Share of AI Companies (%)': 'mean'
}).round(2)

print("Average Metrics by Country:")
country_metrics
```

```
Average Metrics by Country:
```

[131]:

Country	AI Adoption Rate (%)	Revenue Increase Due to AI (%)	\
Australia	56.08	49.70	
Canada	52.19	38.75	
China	52.89	42.60	
France	56.52	39.55	
Germany	51.46	40.58	
India	51.81	40.08	
Japan	54.21	34.61	
South Korea	50.56	34.25	
UK	64.69	36.12	
USA	52.08	43.52	

Country	Job Loss Due to AI (%)	Consumer Trust in AI (%)	\

```
Australia                      18.82                      63.58
Canada                         30.99                      62.15
China                          29.09                      60.32
France                         26.48                      58.44
Germany                        28.10                      64.95
India                          25.23                      59.92
Japan                          25.51                      61.00
South Korea                    25.54                      60.02
UK                             27.13                      50.14
USA                            20.71                      56.20

              Market Share of AI Companies (%)
Country
Australia                                    23.69
Canada                                       28.28
China                                        27.06
France                                       27.19
Germany                                      30.46
India                                        28.34
Japan                                        27.15
South Korea                                  27.33
UK                                           25.14
USA                                          20.98
```

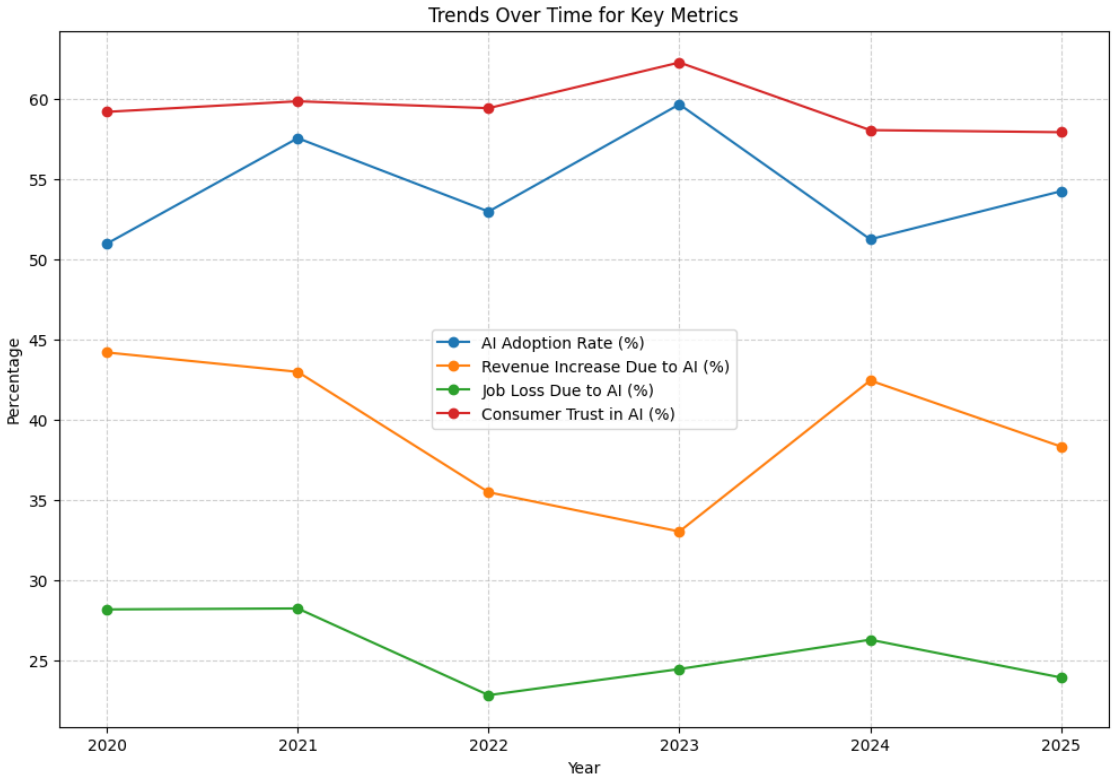
[132]:
```python
# Average metrics by industry
industry_metrics = df.groupby('Industry').agg({
    'AI Adoption Rate (%)': 'mean',
    'Revenue Increase Due to AI (%)': 'mean',
    'Job Loss Due to AI (%)': 'mean',
    'Consumer Trust in AI (%)': 'mean',
    'Market Share of AI Companies (%)': 'mean'
}).round(2)

print("Average Metrics by Industry:")
industry_metrics
```

```
Average Metrics by Industry:
```

[132]:
```
              AI Adoption Rate (%)  Revenue Increase Due to AI (%)  \
Industry
Automotive                   54.89                           46.48
Education                    57.03                           39.54
Finance                      55.76                           36.26
Gaming                       60.42                           33.23
Healthcare                   55.73                           38.59
Legal                        56.08                           41.82
Manufacturing                57.01                           42.81
Marketing                    54.24                           36.81
```

```
Media                          47.26                          43.72
Retail                         47.91                          37.68

               Job Loss Due to AI (%)  Consumer Trust in AI (%)  \
Industry
Automotive                      28.92                     61.24
Education                       26.14                     51.65
Finance                         27.79                     56.18
Gaming                          27.20                     61.34
Healthcare                      25.58                     60.16
Legal                           28.23                     57.79
Manufacturing                   32.75                     58.27
Marketing                       19.58                     58.67
Media                           22.75                     61.76
Retail                          21.85                     62.73

               Market Share of AI Companies (%)
Industry
Automotive                                29.59
Education                                 22.84
Finance                                   29.00
Gaming                                    24.06
Healthcare                                29.64
Legal                                     20.24
Manufacturing                             25.96
Marketing                                 28.68
Media                                     30.03
Retail                                    24.60
```

[133]:
```python
# Trends over time for key metrics
time_trends = df.groupby('Year').agg({
    'AI Adoption Rate (%)': 'mean',
    'Revenue Increase Due to AI (%)': 'mean',
    'Job Loss Due to AI (%)': 'mean',
    'Consumer Trust in AI (%)': 'mean'
}).round(2)

plt.figure(figsize=(12, 8))
for col in time_trends.columns:
    plt.plot(time_trends.index, time_trends[col], label=col, marker='o')

plt.title('Trends Over Time for Key Metrics')
plt.xlabel('Year')
plt.ylabel('Percentage')
plt.legend()
plt.grid(True, linestyle='--', alpha=0.6)
plt.show()
```



[134]:
```
# Summary of findings
print("\nSUMMARY OF KEY FINDINGS:")
print(f"1. Dataset contains {df.shape[0]} records across {df['Country'].
 ↪nunique()} countries, "
      f"{df['Industry'].nunique()} industries, and {df['Year'].nunique()} years␣
 ↪({df['Year'].min()}-{df['Year'].max()})")
print(f"2. Most popular AI tools: {', '.join(df['Top AI Tools Used'].
 ↪value_counts().head(3).index)}")
print(f"3. Average AI adoption rate: {df['AI Adoption Rate (%)'].mean():.2f}%")
print(f"4. Average revenue increase due to AI: {df['Revenue Increase Due to AI␣
 ↪(%)'].mean():.2f}%")
print(f"5. Average job loss due to AI: {df['Job Loss Due to AI (%)'].mean():.
 ↪2f}%")
print(f"6. Average consumer trust in AI: {df['Consumer Trust in AI (%)'].mean():
 ↪.2f}%")
print(f"7. Countries with highest average AI adoption: {', '.
 ↪join(country_metrics.nlargest(3, 'AI Adoption Rate (%)').index)}")
print(f"8. Industries with highest average revenue increase from AI: {', '.
 ↪join(industry_metrics.nlargest(3, 'Revenue Increase Due to AI (%)').index)}")
```

```
SUMMARY OF KEY FINDINGS:
```

1. Dataset contains 200 records across 10 countries, 10 industries, and 6 years (2020-2025)
2. Most popular AI tools: Midjourney, Claude, ChatGPT
3. Average AI adoption rate: 54.27%
4. Average revenue increase due to AI: 39.72%
5. Average job loss due to AI: 25.79%
6. Average consumer trust in AI: 59.43%
7. Countries with highest average AI adoption: UK, France, Australia
8. Industries with highest average revenue increase from AI: Automotive, Media, Manufacturing