

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
df = pd.read_csv('global_llm_benchmark_dataset_models.csv')
```

```
print(df.isnull().sum())
```

```
model_name          0
provider            0
release_date        0
context_window      0
api_input_cost_per_1k_usd  0
api_output_cost_per_1k_usd  0
open_source         0
mmlu_score          0
gsm8k_score         0
humaneval_score     0
dtype: int64
```

```
df['api_input_cost_per_1k_usd'] = df['api_input_cost_per_1k_usd'].fillna(0)
df['api_output_cost_per_1k_usd'] = df['api_output_cost_per_1k_usd'].fillna(0)
df['mmlu_score'] = df['mmlu_score'].fillna(df['mmlu_score'].mean())
```

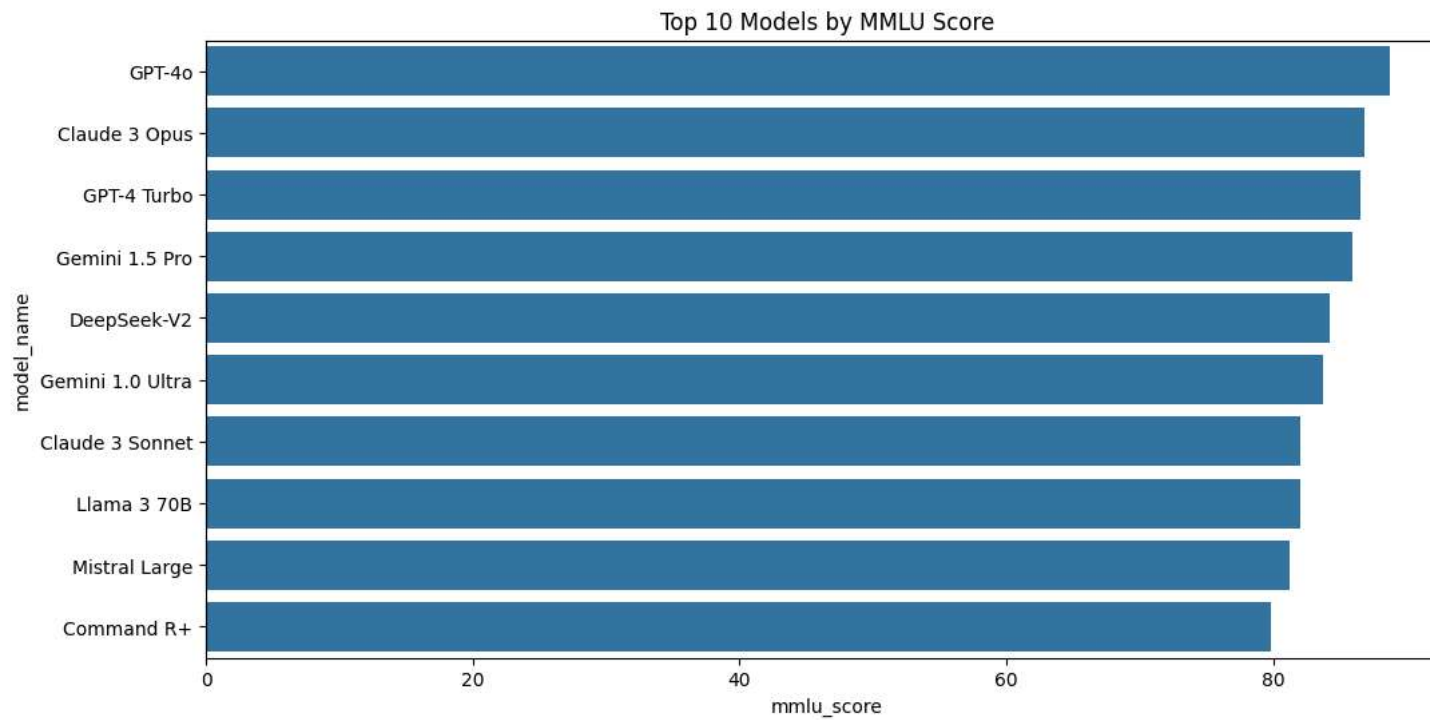
```
df.dropna(subset=['model_name'], inplace=True)
```

```
Q1 = df['context_window'].quantile(0.25)
Q3 = df['context_window'].quantile(0.75)
IQR = Q3 - Q1
```

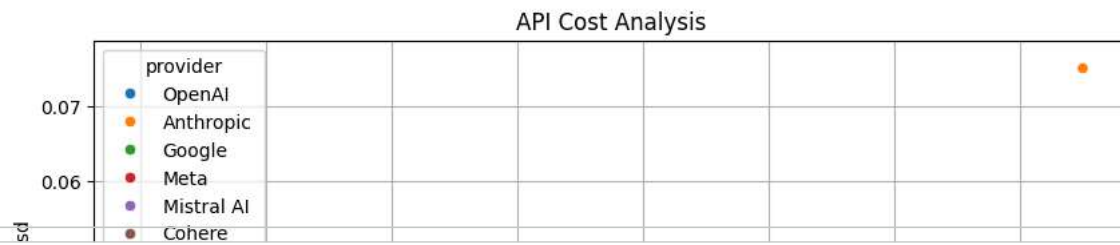
```
lower_limit = Q1 - 1.5 * IQR
upper_limit = Q3 + 1.5 * IQR
```

```
df_cleaned = df[(df['context_window'] >= lower_limit) & (df['context_window'] <= upper_limit)]
```

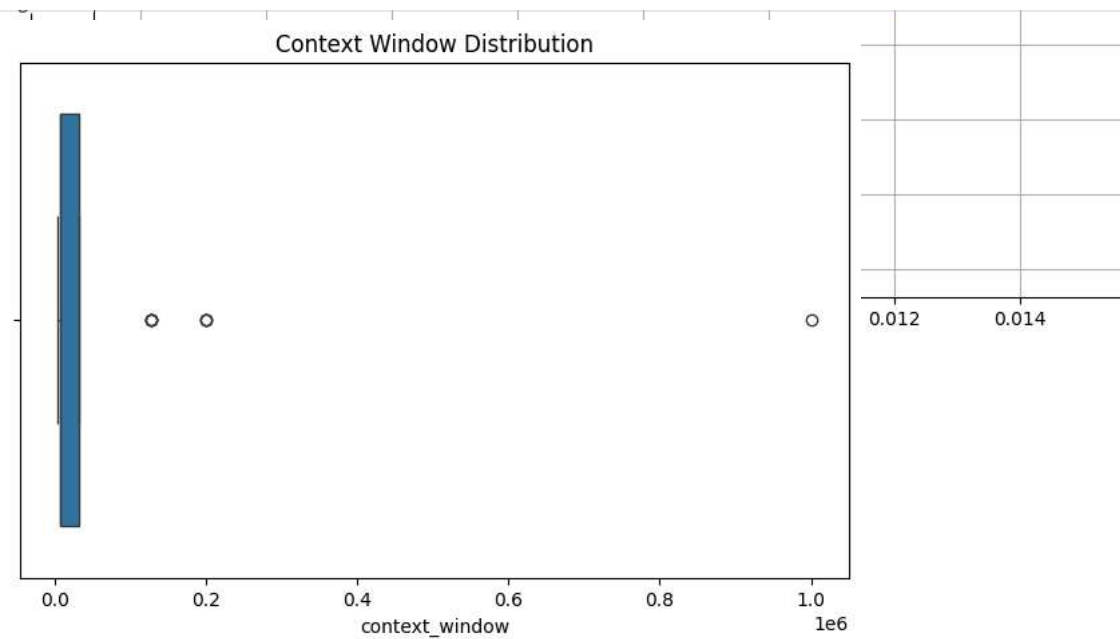
```
plt.figure(figsize=(12, 6))
sns.barplot(x='mmlu_score', y='model_name', data=df.sort_values('mmlu_score', ascending=False).head(10))
plt.title('Top 10 Models by MMLU Score')
plt.show()
```



```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='api_input_cost_per_1k_usd', y='api_output_cost_per_1k_usd', hue='provider', data=df)
plt.title('API Cost Analysis')
plt.grid(True)
plt.show()
```



```
plt.figure(figsize=(8, 5))
sns.boxplot(x=df['context_window'])
plt.title('Context Window Distribution')
plt.show()
```



model_name	provider	release_date	context_window	api_input_cost_per_1k_usd	api_output_cost_per_1k_usd	open_source	mmlu_score	gsm8k_score	humaneval_score
GPT-4o	OpenAI	2024-05-13	128000	0.005	0.015	False	88.7	92.0	90.2
GPT-4 Turbo	OpenAI	2023-11-06	128000	0.01	0.03	False	86.5	89.0	85.4
GPT-3.5 Turbo	OpenAI	2023-03-01	16000	0.001	0.002	False	70.0	57.1	48.1
Claude 3 Opus	Anthropic	2024-03-04	200000	0.015	0.075	False	86.8	95.0	84.9
Claude 3 Sonnet	Anthropic	2024-03-04	200000	0.003	0.015	False	82.0	92.3	78.5
Claude 3 Haiku	Anthropic	2024-03-04	200000	0.00025	0.00125	False	75.2	80.0	63.0
Gemini 1.5 Pro	Google	2024-02-15	1000000	0.007	0.021	False	85.9	94.4	74.4
Gemini 1.0 Ultra	Google	2023-12-06	32000	0.012	0.036	False	83.7	90.0	67.0
Llama 3 70B	Meta	2024-04-18	8000	0.0	0.0	True	82.0	89.0	81.7
Llama 3 8B	Meta	2024-04-18	8000	0.0	0.0	True	68.4	74.6	62.2
Llama 2 70B	Meta	2023-07-18	4096	0.0	0.0	True	69.8	54.4	29.9
Mistral Large	Mistral AI	2024-02-26	32000	0.008	0.024	False	81.2	86.5	76.8
Mixtral 8x7B	Mistral AI	2023-12-11	32000	0.0	0.0	True	74.5	83.7	40.2
Command R+	Cohere	2024-03-13	128000	0.003	0.015	False	79.8	88.0	66.0
Grok-1	xAI	2023-11-04	8192	0.0	0.0	False	73.0	63.2	63.0
DeepSeek-V2	DeepSeek	2024-05-01	128000	0.002	0.006	True	84.2	90.1	82.3
Qwen1.5-72B	Alibaba	2024-02-01	32000	0.0	0.0	True	77.3	84.5	53.0
OpenModel-1B	Various	2024-01-01	8192	0.0	0.0	True	60.5	50.6	40.7
OpenModel-2B	Various	2024-01-01	8192	0.0	0.0	True	61.0	51.2	41.4
OpenModel-3B	Various	2024-01-01	8192	0.0	0.0	True	61.5	51.8	42.1
OpenModel-4B	Various	2024-01-01	8192	0.0	0.0	True	62.0	52.4	42.8
OpenModel-5B	Various	2024-01-01	8192	0.0	0.0	True	62.5	53.0	43.5
OpenModel-6B	Various	2024-01-01	8192	0.0	0.0	True	63.0	53.6	44.2
OpenModel-7B	Various	2024-01-01	8192	0.0	0.0	True	63.5	54.2	44.9
OpenModel-8B	Various	2024-01-01	8192	0.0	0.0	True	64.0	54.8	45.6
OpenModel-9B	Various	2024-01-01	8192	0.0	0.0	True	64.5	55.4	46.3
OpenModel-10B	Various	2024-01-01	8192	0.0	0.0	True	65.0	56.0	47.0
OpenModel-11B	Various	2024-01-01	8192	0.0	0.0	True	65.5	56.6	47.7
OpenModel-12B	Various	2024-01-01	8192	0.0	0.0	True	66.0	57.2	48.4
OpenModel-13B	Various	2024-01-01	8192	0.0	0.0	True	66.5	57.8	49.1
OpenModel-14B	Various	2024-01-01	8192	0.0	0.0	True	67.0	58.4	49.8
OpenModel-15B	Various	2024-01-01	8192	0.0	0.0	True	67.5	59.0	50.5
OpenModel-16B	Various	2024-01-01	8192	0.0	0.0	True	68.0	59.6	51.2
OpenModel-17B	Various	2024-01-01	8192	0.0	0.0	True	68.5	60.2	51.9
OpenModel-18B	Various	2024-01-01	8192	0.0	0.0	True	69.0	60.8	52.6
OpenModel-19B	Various	2024-01-01	8192	0.0	0.0	True	69.5	61.4	53.3
OpenModel-20B	Various	2024-01-01	8192	0.0	0.0	True	70.0	62.0	54.0