

ChatGPT Vs Gemini SQL Case Study Analysis



VS



Presented By Rupal Hiwarkar

CHALLENGE



OpenAI

Vs

Gemini

The challenge is to craft SQL queries to extract insights from the Gemini Vs ChatGPT database.

Table Of Contents



OpenAI

VS

Gemini

- ☐ Capabilities
- ☐ Models
- ☐ Benchmarks

1)What are the average scores for each capability on both the Gemini Ultra and GPT-4 models?

```
3
4  -- 1)What are the average scores for each capability on both the Gemini Ultra and GPT-4 models?
5  • SELECT c.CapabilityName,
6         AVG(b.ScoreGemini) AS AvgScoreGeminiUltra,
7         AVG(b.ScoreGPT4) AS AvgScoreGPT4
8  FROM Capabilities c
9  LEFT JOIN Benchmarks b ON c.CapabilityID = b.CapabilityID
10 GROUP BY c.CapabilityName;
11
12
13
14
15
```

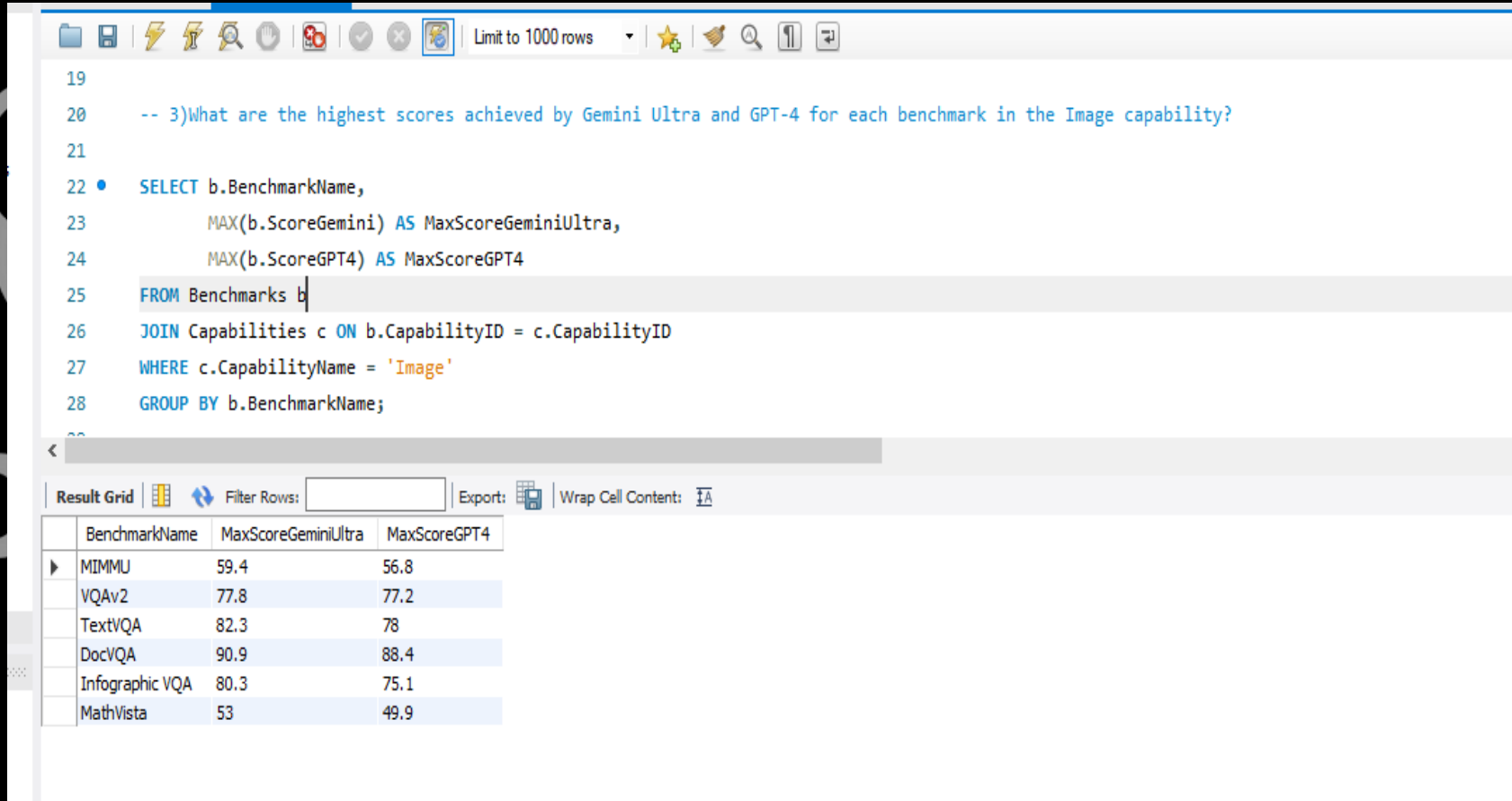
CapabilityName	AvgScoreGeminiUltra	AvgScoreGPT4
General	88.20000076293945	86.4000015258789
Reasoning	85.51666768391927	86.43333435058594
Math	73.12500095367432	72.45000076293945
Code	72.55000114440918	70.45000076293945
Image	73.9500020345052	70.89999961853027
Video	58.70000076293945	51.14999961853027
Audio	23.84999918937683	23.350000381469727

2) Which benchmarks does Gemini Ultra outperform GPT-4 in terms of scores?

```
12
13 -- 2)Which benchmarks does Gemini Ultra outperform GPT-4 in terms of scores?
14
15 • SELECT BenchmarkName
16 FROM Benchmarks
17 WHERE ScoreGemini > ScoreGPT4;
18
19
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
BenchmarkName			
▶ MMLU			
Big-Bench Hard			
DROP			
GSM8K			
MATH			
HumanEval			
NaturalCode			
MIMMU			
VQAv2			
TextVQA			
DocVQA			
Infographic VQA			
MathVista			
VATEX			
Perception Test...			
CoVOST 2			

3)What are the highest scores achieved by Gemini Ultra and GPT-4 for each benchmark in the Image capability?



```
19
20 -- 3)What are the highest scores achieved by Gemini Ultra and GPT-4 for each benchmark in the Image capability?
21
22 • SELECT b.BenchmarkName,
23         MAX(b.ScoreGemini) AS MaxScoreGeminiUltra,
24         MAX(b.ScoreGPT4) AS MaxScoreGPT4
25 FROM Benchmarks b
26 JOIN Capabilities c ON b.CapabilityID = c.CapabilityID
27 WHERE c.CapabilityName = 'Image'
28 GROUP BY b.BenchmarkName;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	BenchmarkName	MaxScoreGeminiUltra	MaxScoreGPT4
▶	MIMMU	59.4	56.8
	VQAv2	77.8	77.2
	TextVQA	82.3	78
	DocVQA	90.9	88.4
	Infographic VQA	80.3	75.1
	MathVista	53	49.9

4) Calculate the percentage improvement of Gemini Ultra over GPT-4 for each benchmark?

```
31  -- 4)Calculate the percentage improvement of Gemini Ultra over GPT-4 for each benchmark?
32
33  •  SELECT BenchmarkName,
34         ((ScoreGemini - ScoreGPT4) / ScoreGPT4) * 100 AS PercentageImprovement
35  FROM Benchmarks
36  WHERE ScoreGPT4 IS NOT NULL;
37
```

Result Grid |  Filter Rows: | Export:  | Wrap Cell Content: 

	BenchmarkName	PercentageImprovement
▶	MMLU	4.166664827017169
	Big-Bench Hard	0.601684728256294
	DROP	1.8541408797380168
	HellaSwag	-7.869884323011742
	GSM8K	2.608697310737942
	MATH	0.5671062918850502
	HumanEval	11.04477839683419
	Natura12Code	1.3531799449960928
	MIMMU	4.577468879826893
	VQAv2	0.7772100093707153
	TextVQA	5.512824425330528
	DocVQA	2.8280542498272814
	Infographic VQA	6.924107434474319
	MathVista	6.212421601857841
	VATEX	11.964287076677595
	Perception Test...	18.142552190703228
	CoVOST 2	37.80068023523442
	FLEURS	-56.81818329598289

5) Retrieve the benchmarks where both models scored above the average for their respective models.

```
37
38  -- 5)Retrieve the benchmarks where both models scored above the average for their respective models?
39
40  WITH AvgScores AS (
41      SELECT AVG(ScoreGemini) AS AvgGemini, AVG(ScoreGPT4) AS AvgGPT4
42      FROM Benchmarks
43  )
44  SELECT BenchmarkName
45  FROM Benchmarks, AvgScores
46  WHERE ScoreGemini > (SELECT AvgGemini FROM AvgScores)
47      AND ScoreGPT4 > (SELECT AvgGPT4 FROM AvgScores);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	BenchmarkName
▶	MMLU
	Big-Bench Hard
	DROP
	HellaSwag
	GSM8K
	HumanEval
	Natural12Code
	VQAv2
	TextVQA
	DocVQA
	Infographic VQA

6) Which benchmarks show that Gemini Ultra is expected to outperform GPT-4 based on the next score?

```
48
49 -- 6) Which benchmarks show that Gemini Ultra is expected to outperform GPT-4 based on the next score?
50
51 • SELECT BenchmarkName
52 FROM Benchmarks
53 WHERE ScoreGemini > (SELECT AVG(ScoreGemini) FROM Benchmarks)
54 AND ScoreGPT4 < (SELECT AVG(ScoreGPT4) FROM Benchmarks);
55
56
57
58
--
```

Result Grid |  Filter Rows: | Export:  | Wrap Cell Content: 

BenchmarkName

7) Classify benchmarks into performance categories based on score ranges?

```
55
56  -- 7)Classify benchmarks into performance categories based on score ranges?
57
58  •  SELECT BenchmarkName,
59         CASE
60             WHEN ScoreGemini >= 90 AND ScoreGPT4 >= 90 THEN 'High Performance'
61             WHEN ScoreGemini >= 80 AND ScoreGPT4 >= 80 THEN 'Good Performance'
62             WHEN ScoreGemini >= 70 AND ScoreGPT4 >= 70 THEN 'Fair Performance'
63             ELSE 'Low Performance'
64         END AS PerformanceCategory
65  FROM Benchmarks;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

BenchmarkName	PerformanceCategory
MMLU	Good Performance
MMLU	Low Performance
Big-Bench Hard	Good Performance
Big-Bench Hard	Low Performance
DROP	Good Performance
DROP	Low Performance
HellaSwag	Good Performance
HellaSwag	Low Performance
GSM8K	High Performance
GSM8K	Low Performance
MATH	Low Performance
MATH	Low Performance
HumanEval	Low Performance
HumanEval	Low Performance

8) Retrieve the rankings for each capability based on Gemini Ultra scores?

```
66
67  -- 8) Retrieve the rankings for each capability based on Gemini Ultra scores?
68
69  WITH GeminiRankings AS (
70      SELECT c.CapabilityName, b.BenchmarkName, b.ScoreGemini,
71             DENSE_RANK() OVER (PARTITION BY c.CapabilityID ORDER BY b.ScoreGemini DESC) AS GeminiRank
72      FROM Capabilities c
73      JOIN Benchmarks b ON c.CapabilityID = b.CapabilityID
74      WHERE b.ScoreGemini IS NOT NULL
75  )
76  SELECT CapabilityName, BenchmarkName, ScoreGemini, GeminiRank
77  FROM GeminiRankings
78  WHERE GeminiRank = 1;
79
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	CapabilityName	BenchmarkName	ScoreGemini	GeminiRank
►	General	MMLU	90	1
	Reasoning	HellaSwag	95.3	1
	Math	GSM8K	94.4	1
	Code	Natura12Code	74.9	1
	Image	DocVQA	90.9	1
	Video	VATEX	62.7	1
	Audio	CoVOST 2	40.1	1

9) Convert the Capability and Benchmark names to uppercase.

```
80
81 -- 9) Convert the Capability and Benchmark names to uppercase?
82
83 • SELECT UPPER(CapabilityName) AS CapabilityName, UPPER(BenchmarkName) AS BenchmarkName
84 FROM Capabilities c
85 JOIN Benchmarks b ON c.CapabilityID = b.CapabilityID;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

CapabilityName	BenchmarkName
GENERAL	MMLU
GENERAL	MMLU
REASONING	BIG-BENCH HARD
REASONING	BIG-BENCH HARD
REASONING	DROP
REASONING	DROP
REASONING	HELLASWAG
REASONING	HELLASWAG
MATH	GSM8K
MATH	GSM8K
MATH	MATH
MATH	MATH
CODE	HUMANEVAL
CODE	HUMANEVAL
CODE	NATURA12CODE
CODE	NATURA12CODE
IMAGE	MIMMU
IMAGE	VQAV2
IMAGE	TEXTVQA
IMAGE	DOCVQA

[illegible]

Open

THANK YOU



vs Gemini

NEXT-GEN AI SHOWDOWN