

“SQL CASE STUDY”



Presented by Korrapati Jaswanth

CHALLENGE

Challenge is to craft SQL queries to extract insights from the Gemini Vs Chatgpt database.

Table Of Contents

- 1 Models
- 2 Capabilities
- 3 Benchmarks

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

```
92
93 #1)What are the average scores for each capability on both the Gemini Ultra and GPT-4 models?
94 • select capabilityname ,round( avg(scoregemini) )as avg_scoregemini , round(avg(scoregpt4)) as avg_scoregpt from capabilities
95 inner join Benchmarks
96 on
97 Capabilities.Capabilityid=Benchmarks.Capabilityid
98 group by capabilityname;
99
```

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

capabilityname	avg_scoregemini	avg_scoregpt
General	88	86
Reasoning	86	86
Math	73	72
Code	73	70
Image	74	71
Video	59	51
Audio	24	23

Result Grid
Form Editor
Field Types

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

```
101
102
103 #2)Which benchmarks does Gemini Ultra outperform GPT-4 in terms of scores?
104 • select benchmarkname from Benchmarks
105 where scoregemini > scoregpt4;
106
107
108
109
```

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

benchmarkname
MMLU
Big-Bench Hard
DROP
GSM8K
MATH
HumanEval
Natura12Code
MIMMU
VQAv2
TextVQA

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

```
110
111      #3)What are the highest scores achieved by Gemini Ultra and GPT-4 for each benchmark in the Image capability?
112 •  select  benchmarkname , max(scoregemini) as max_gemini , max(scoregpt4) as max_gpt4
113   from  Benchmarks
114   where  capabilityid=5
115   group  by benchmarkname;
116
117
118
```

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

benchmarkname	max_gemini	max_gpt4
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?

```
120
121 #4)Calculate the percentage improvement of Gemini Ultra over GPT-4 for each benchmark?
122 • select benchmarkname , round(((scoregemini-scoregpt4)/scoregpt4*100),0) as percentage_improvement from Benchmarks
123 where scoregpt4 is not null
124 order by percentage_improvement ;
125
126
127
```

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

	benchmarkname	percentage_improvement
▶	FLEURS	-57
	HellaSwag	-8
	Big-Bench Hard	1
	MATH	1
	Natura12Code	1
	VQAv2	1
	DROP	2
	GSM8K	3
	DocVQA	3
	MMLU	4

Result 60 ×

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

```
127
128      #5)Retrieve the benchmarks where both models scored above the average for their respective models?
129 •  select benchmarkname ,scoregemini,scoregpt4  from Benchmarks
130      where scoregemini >(select avg(scoregemini)  from Benchmarks)
131      and  scoregpt4>(select avg(scoregpt4)  from Benchmarks);
132
133
134
135
```

	benchmarkname	scoregemini	scoregpt4
▶	MMLU	90	86.4
	Big-Bench Hard	83.6	83.1
	DROP	82.4	80.9
	HellaSwag	87.8	95.3
	GSM8K	94.4	92
	HumanEval	74.4	67
	Natura12Code	74.9	73.9
	VQAv2	77.8	77.2
	TextVQA	82.3	78
	DocVQA	90.9	88.4

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

```
131
132
133      #8) Retrieve the rankings for each capability based on Gemini Ultra scores?
134 •  select capabilityname , scoregemini , DENSE_RANK() OVER(ORDER BY scoregemini DESC) as rnk from Capabilities
135      inner join Benchmarks
136      on
137          Capabilities.capabilityid=Benchmarks.capabilityid;
138
139
```

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

	capabilityname	scoregemini	rnk
▶	Reasoning	95.3	1
	Math	94.4	2
	Math	92	3
	Image	90.9	4
	General	90	5
	Reasoning	87.8	6
	General	86.4	7
	Reasoning	83.6	8
	Reasoning	83.1	9
	Reasoning	82.4	10

Result 62 ×

7) Convert the Capability and Benchmark names to uppercase?

```
140
141
142 #9)Convert the Capability and Benchmark names to uppercase?
143 • select upper(capabilityname) as capability_name , upper(benchmarkname) as benchmark_name from Capabilities
144 inner join Benchmarks
145 on
146 Capabilities.capabilityid=Benchmarks.capabilityid;
147
```

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

	capability_name	benchmark_name
▶	GENERAL	MMLU
	GENERAL	MMLU
	REASONING	BIG-BENCH HARD
	REASONING	BIG-BENCH HARD
	REASONING	DROP
	REASONING	DROP
	REASONING	HELLASWAG
	REASONING	HELLASWAG
	MATH	GSM8K
	MATH	GSM8K

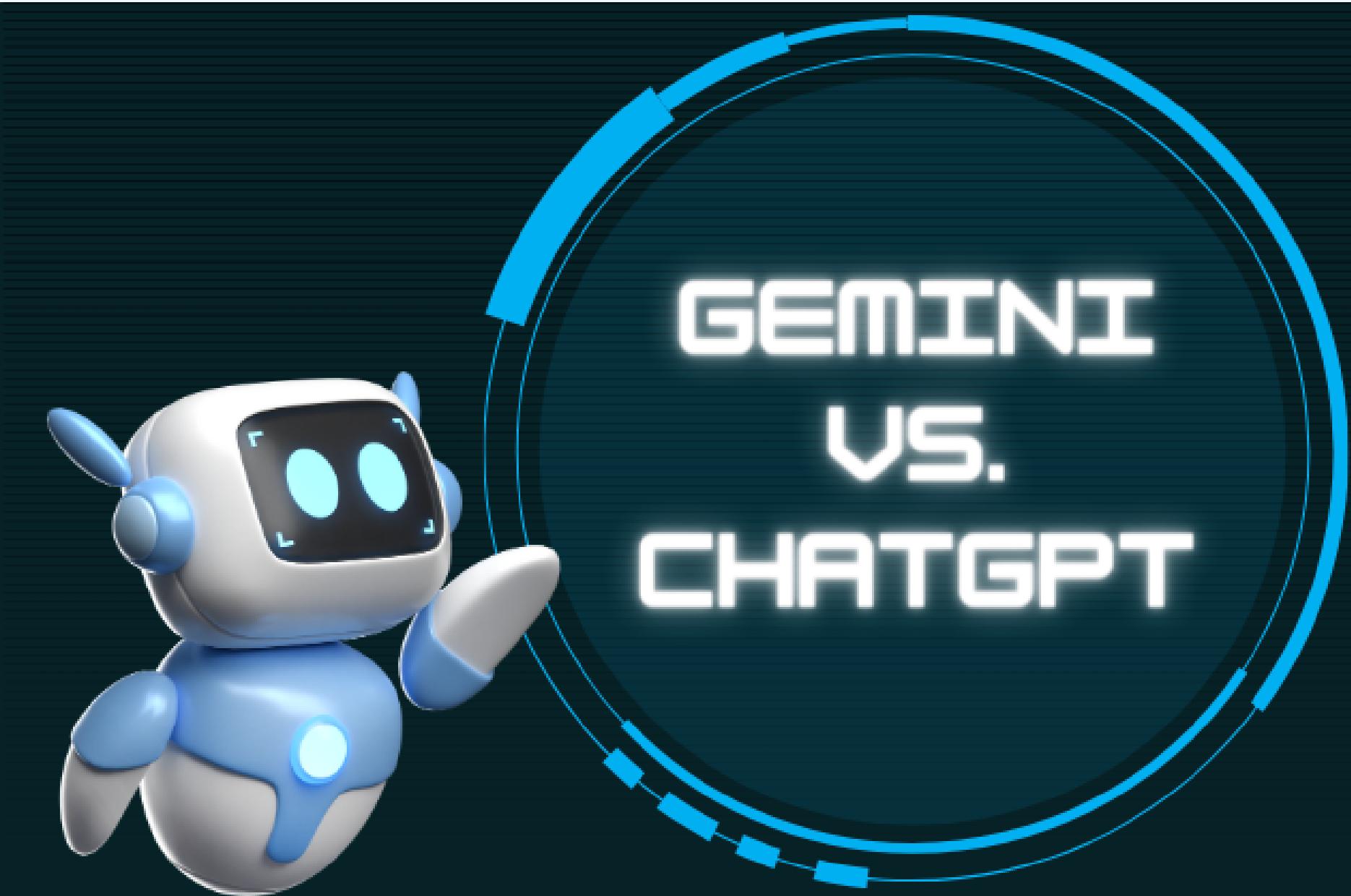
8) Can you provide the benchmarks along with their descriptions in a concatenated format?

```
149
150 #10) Can you provide the benchmarks along with their descriptions in a concatenated format?
151 • select concat(benchmarkname , " " , description ) as benchmarkname
152 from benchmarks;
153
154
155
156
157
```

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

benchmarkname
MMLU Representation of questions in 57 subjects
MMLU Representation of questions in 57 subjects
Big-Bench Hard Diverse set of challenging tasks ...
Big-Bench Hard Diverse set of challenging tasks ...
DROP Reading comprehension (Fl Score)
DROP Reading comprehension (Fl Score)
HellaSwag Commonsense reasoning for everyd...
HellaSwag Commonsense reasoning for everyd...
GSM8K Basic arithmetic manipulations, incl. Grad...
GSM8K Basic arithmetic manipulations, incl. Grad...

Result 64 ×



**“Process is More
Important Than
the Results”**



A background featuring a light blue and white marbled or wash effect at the top, transitioning to a white area with scattered gold leaf pieces (circles and ovals) and two large, semi-transparent gold geometric shapes (hexagons and triangles) on the left and right sides. The overall aesthetic is elegant and modern.

THANK YOU