Table 30: **Benchmark scores for language: Korean (ko)**. The highest and second-best scores are shown in **bold** and underlined, respectively.

	Model	MLogiQA	INCLUDE	MMMLU	MT-AIME24	PolyMath	Average
Thinking Mode	Gemini2.5-Pro	<u>75.0</u>	**88.0**	**85.9**	<u>76.7</u>	<u>50.0</u>	**75.1**
	QwQ-32B	**76.2**	72.0	81.8	60.0	40.0	66.0
	Qwen3-235B-A22B	71.2	<u>80.0</u>	<u>84.7</u>	**80.0**	**55.7**	<u>74.3</u>
	Qwen3-32B	71.2	74.0	79.2	**80.0**	48.5	70.6
	Qwen3-30B-A3B	68.8	72.0	78.6	<u>76.7</u>	46.6	68.5
	Qwen3-14B	67.5	74.0	79.6	<u>76.7</u>	46.0	68.8
	Qwen3-8B	60.0	<u>80.0</u>	74.7	<u>76.7</u>	42.3	66.7
	Qwen3-4B	66.2	74.0	68.8	70.0	40.6	63.9
	Qwen3-1.7B	53.8	66.0	57.8	43.3	25.2	49.2
	Qwen3-0.6B	33.8	52.0	41.5	13.3	11.8	30.5
Non-thinking Mode	GPT-4o-2024-1120	**63.7**	**80.0**	**80.5**	13.3	12.9	50.1
	Gemma-3-27b-IT	58.8	<u>76.0</u>	75.9	20.0	18.3	49.8
	Qwen2.5-72B-Instruct	58.8	68.0	76.7	6.7	17.7	45.6
	Qwen3-235B-A22B	**63.7**	<u>76.0</u>	<u>79.8</u>	**33.3**	**27.9**	**56.1**
	Qwen3-32B	<u>60.0</u>	74.0	77.2	<u>26.7</u>	<u>21.2</u>	<u>51.8</u>
	Qwen3-30B-A3B	52.5	72.0	72.5	16.7	20.7	46.9
	Qwen3-14B	52.5	68.0	73.3	20.0	18.7	46.5
	Qwen3-8B	52.5	<u>76.0</u>	66.5	23.3	16.3	46.9
	Qwen3-4B	46.2	74.0	59.9	13.3	16.6	42.0
	Qwen3-1.7B	48.8	58.0	46.0	6.7	9.0	33.7
	Qwen3-0.6B	40.0	52.0	36.9	0.0	5.5	26.9

Table 31: **Benchmark scores for language: Indonesian (id)**. The highest and second-best scores are shown in **bold** and underlined, respectively.

	Model	INCLUDE	MMMLU	MT-AIME24	PolyMath	Average
Thinking Mode	Gemini2.5-Pro	<u>80.0</u>	<u>86.3</u>	**83.3**	<u>51.3</u>	**75.2**
	QwQ-32B	76.4	83.7	73.3	47.3	70.2
	Qwen3-235B-A22B	<u>80.0</u>	**87.2**	<u>80.0</u>	**53.5**	**75.2**
	Qwen3-32B	<u>80.0</u>	82.0	76.7	45.6	71.1
	Qwen3-30B-A3B	**81.8**	80.4	<u>80.0</u>	44.9	<u>71.8</u>
	Qwen3-14B	78.2	79.6	70.0	45.3	68.3
	Qwen3-8B	72.7	77.7	70.0	43.8	66.0
	Qwen3-4B	70.9	72.3	66.7	41.2	62.8
	Qwen3-1.7B	63.6	61.2	36.7	26.8	47.1
	Qwen3-0.6B	36.4	46.6	10.0	12.6	26.4
Non-thinking Mode	GPT-4o-2024-1120	<u>80.0</u>	<u>81.1</u>	10.0	14.7	46.4
	Gemma-3-27b-IT	76.4	75.9	13.3	22.6	47.0
	Qwen2.5-72B-Instruct	74.5	78.8	10.0	16.6	45.0
	Qwen3-235B-A22B	**81.8**	**81.9**	**33.3**	**27.5**	**56.1**
	Qwen3-32B	**81.8**	77.2	23.3	24.3	<u>51.6</u>
	Qwen3-30B-A3B	70.9	76.4	<u>30.0</u>	<u>25.9</u>	50.8
	Qwen3-14B	70.9	74.1	26.7	24.6	49.1
	Qwen3-8B	78.2	69.6	20.0	21.6	47.4
	Qwen3-4B	67.3	66.5	13.3	19.0	41.5
	Qwen3-1.7B	52.7	49.0	3.3	10.8	29.0
	Qwen3-0.6B	52.7	40.0	3.3	5.1	25.3