Table 5: **Comparison among Qwen3-14B-Base, Qwen3-30B-A3B-Base, and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.**

	Gemma-3-12B Base	**Qwen2.5-14B Base**	**Qwen2.5-32B Base**	**Qwen2.5-Turbo Base**	**Qwen3-14B Base**	**Qwen3-30B-A3B Base**
Architecture	Dense	Dense	Dense	MoE	Dense	MoE
# Total Params	12B	14B	32B	42B	14B	30B
# Activated Params	12B	14B	32B	6B	14B	3B
General Tasks						
MMLU	73.87	79.66	**83.32**	79.50	81.05	81.38
MMLU-Redux	70.70	76.64	**81.97**	77.11	79.88	81.17
MMLU-Pro	44.91	51.16	55.10	55.60	61.03	**61.49**
SuperGPQA	24.61	30.68	33.55	31.19	34.27	**35.72**
BBH	74.28	78.18	**84.48**	76.10	81.07	81.54
Math & STEM Tasks						
GPQA	31.31	32.83	**47.97**	41.41	39.90	43.94
GSM8K	78.01	90.22	**92.87**	88.32	92.49	91.81
MATH	44.43	55.64	57.70	55.60	**62.02**	59.04
Coding Tasks						
EvalPlus	52.65	60.70	66.25	61.23	**72.23**	71.45
MultiPL-E	43.03	54.79	58.30	53.24	61.69	**66.53**
MBPP	60.60	69.00	73.60	67.60	73.40	**74.40**
CRUX-O	52.00	61.10	67.80	60.20	**68.60**	67.20
Multilingual Tasks						
MGSM	64.35	74.68	78.12	70.45	**79.20**	79.11
MMMLU	72.50	78.34	**82.40**	79.76	79.69	81.46
INCLUDE	63.34	60.26	64.35	59.25	64.55	**67.00**

Table 6: **Comparison among Qwen8B-Base and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.**

	Llama-3-8B Base	**Qwen2.5-7B Base**	**Qwen2.5-14B Base**	**Qwen3-8B Base**
Architecture	Dense	Dense	Dense	Dense
# Total Params	8B	7B	14B	8B
# Activated Params	8B	7B	14B	8B
General Tasks				
MMLU	66.60	74.16	**79.66**	76.89
MMLU-Redux	61.59	71.06	**76.64**	76.17
MMLU-Pro	35.36	45.00	51.16	**56.73**
SuperGPQA	20.54	26.34	30.68	**31.64**
BBH	57.70	70.40	78.18	**78.40**
Math & STEM Tasks				
GPQA	25.80	36.36	32.83	**44.44**
GSM8K	55.30	85.36	**90.22**	89.84
MATH	20.50	49.80	55.64	**60.80**
Coding Tasks				
EvalPlus	44.13	62.18	60.70	**67.65**
MultiPL-E	31.45	50.73	54.79	**58.75**
MBPP	48.40	63.40	69.00	**69.80**
CRUX-O	36.80	48.50	61.10	**62.00**
Multilingual Tasks				
MGSM	38.92	63.60	74.68	**76.02**
MMMLU	59.65	71.34	**78.34**	75.72
IINCLUDE	44.94	53.98	**60.26**	59.40