## F. Statistics of Datasets

LongBench is a meticulously designed benchmark suite that evaluates the capabilities of language models in handling extended documents and complex information sequences. This benchmark was created for multi-task evaluation of long-context inputs and includes 17 datasets covering tasks such as single-document QA (Kočiský et al., 2018; Dasigi et al., 2021), multi-document QA (Yang et al., 2018; Ho et al., 2020; Trivedi et al., 2022; He et al., 2018), summarization (Huang et al., 2021; Zhong et al., 2021; Fabbri et al., 2019; Wu et al., 2023), few-shot learning (Li & Roth, 2002; Gliwa et al., 2019; Joshi et al., 2017), synthetic tasks and code generation (Guo et al., 2023; Liu et al., 2024d). The datasets feature an average input length ranging from $1K$ to $18K$ tokens, requiring substantial memory for KV cache management.

Table 19 shows the statistics of the datasets that we used in our experiments.

| DATASET | # TRAIN | # TEST |
|---|---|---|
| GSM8K (COBBE ET AL., 2021) | 7,473 | 1,319 |
| LONGBENCH (BAI ET AL., 2024) | - | 4,750 |
| NIAH* (KAMRADT, 2023) | - | 800 |

Table 19: Dataset Statistics. # TRAIN and # TEST represent the number of training and test samples, respectively. *: The size of the NIAH test set varies based on the context length and step size, typically around 800 samples per evaluation.

## G. Prompt

Table 20 shows the prompt for the Figure 1

| The prompt for demonstration |
|---|
| . . . . . .<br>. . . . . .<br>The purple-crested turaco (Gallirex porphyreolophus) or, in South Africa, the purple-crested loerie, (Khurukhuru in the Luvenḓa (Venḓa) language) is a species of bird in the clade Turaco with an unresolved phylogenetic placement. Initial analyses placed the purple-crested turaco in the family Musophagidae, but studies have indicated that these birds do not belong to this family and have been placed in the clade of Turacos with an unresolved phylogeny. It is the National Bird of the Kingdom of Eswatini, and the crimson flight feathers of this and related turaco species are important in the ceremonial regalia of the Swazi royal family. This bird has a purple-coloured crest above a green head, a red ring around their eyes, and a black bill. The neck and chest are green and brown. The rest of the body is purple, with red flight feathers. Purple-crested turacos are often seen near water sources, where they can be observed drinking and bathing, which helps them maintain their vibrant plumage. Purple-crested turacos are considered to be large frugivores that are known to carry cycad seeds from various plant species long distances from feeding to nesting sites. After fruit consumption, they regurgitate the seeds intact where they can germinate nearby. G. porphyreolophus primarily consumes fruits whole like many other large frugivores which are suggested to be necessary for effective ecosystem functioning. Among similar turacos, the purple-crested turaco have faster minimum transit times when consuming smaller seed diets than larger seed diets, and G. porphyreolophus has been shown to have significantly faster pulp (seedless fruit masses) transit time than another closely related Turaco when fed only the pulp of larger-seeding fruits than smaller-seeding fruits. In addition to their frugivorous diet, these birds are occasionally seen foraging for other food items such as nuts and leaves, which provide essential nutrients. They are also known to coexist with various other animals, including those that might enjoy strawberries and other similar fruits. The purple-crested turaco's role in seed dispersal is crucial, and their interaction with different elements of their habitat, including water and diverse plant materials, highlights their importance in maintaining ecological balance.<br>. . . . . .<br>. . . . . . |

Table 20: The prompt for demonstration

Here we provide the CoT prompt exemplars for GSM8K which is used in section 4.1.