DATASET STATISTICS

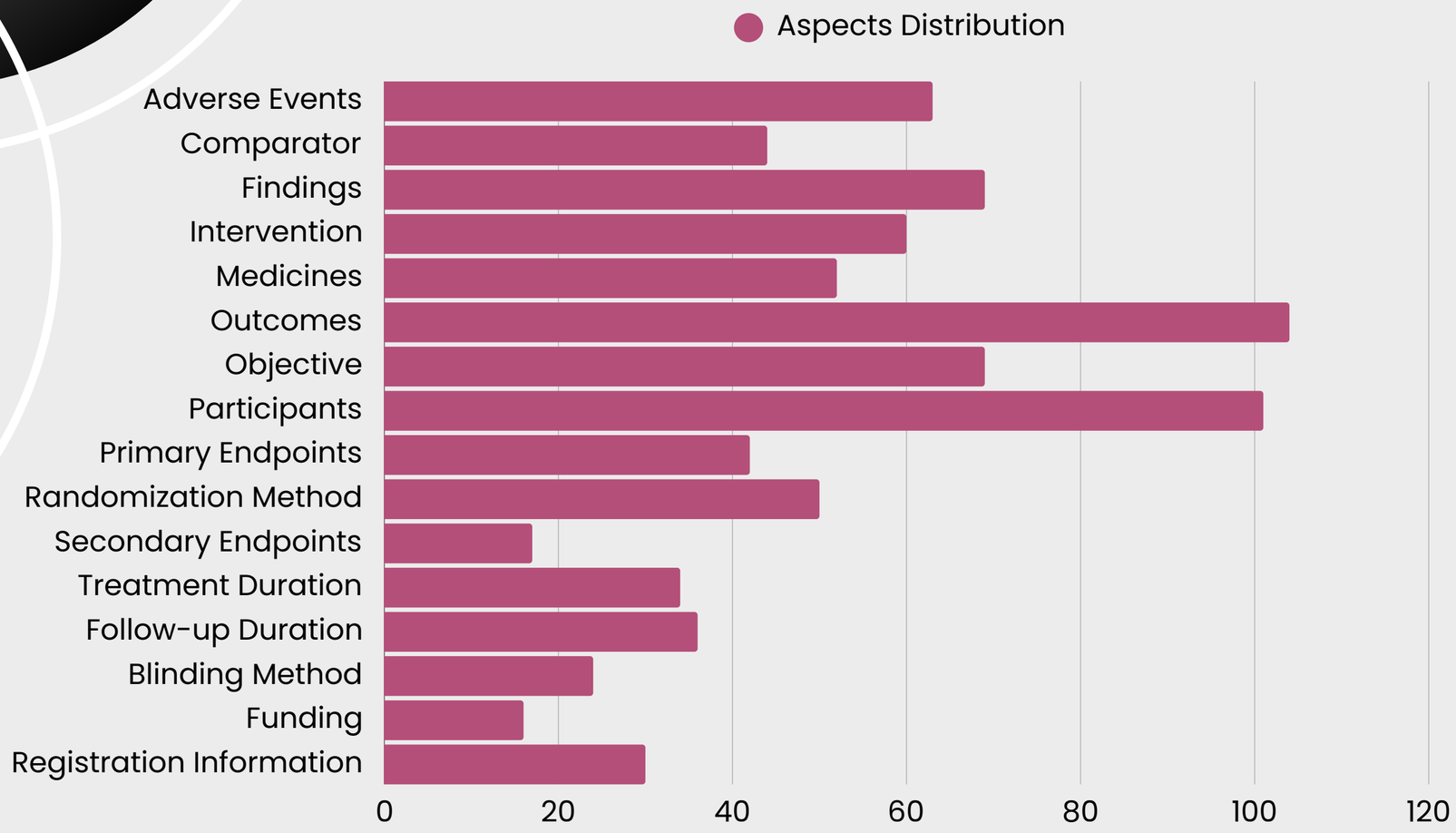
Number of total pairs:

811

Number of unique sentences:

525

```
(pp2025) root@inception:~/praxis#  python dataset_statistics.py

Dataset Statistics
==================
Total pairs: 811

Examples:
Pair 1
Sentence: Since the 5-year analysis, no new safety signals were observed.
Aspect: ae
Keyphrases: [{'sentence': 9, 'kps': ['no new safety signals']}]
----------------------------------------
Pair 2
Sentence: PURPOSE: In the phase III CheckMate 067 trial, durable clinical benefit was demonstrated previously with nivolum
ab plus ipilimumab and nivolumab alone versus ipilimumab.
Aspect: c
Keyphrases: [{'sentence': 0, 'kps': ['ipilimumab']}]
----------------------------------------
Pair 3
Sentence: PATIENTS AND METHODS: Patients with previously untreated unresectable stage III or stage IV melanoma were random
ly assigned 1:1:1 to receive nivolumab 1 mg/kg plus ipilimumab 3 mg/kg once every 3 weeks (four doses) followed by nivolum
ab 3 mg/kg once every 2 weeks (n = 314), nivolumab 3 mg/kg once every 2 weeks (n = 316), or ipilimumab 3 mg/kg once every
3 weeks (four doses; n = 315).
Aspect: c
Keyphrases: [{'sentence': 2, 'kps': ['ipilimumab 3 mg/kg once every 3 weeks']}]
----------------------------------------
Pair 4
Sentence: CONCLUSION: These 6.5-year CheckMate 067 results, which include the longest median OS in a phase III melanoma tr
ial reported to date and the first report of MSS, showed durable, improved clinical outcomes with nivolumab plus ipilimuma
b or nivolumab versus ipilimumab in patients with advanced melanoma and, in descriptive analyses, with the combination ove
r nivolumab monotherapy.
Aspect: f
Keyphrases: [{'sentence': 10, 'kps': ['durable , improved clinical outcomes']}]
----------------------------------------
Pair 5
Sentence: PATIENTS AND METHODS: Patients with previously untreated unresectable stage III or stage IV melanoma were random
ly assigned 1:1:1 to receive nivolumab 1 mg/kg plus ipilimumab 3 mg/kg once every 3 weeks (four doses) followed by nivolum
ab 3 mg/kg once every 2 weeks (n = 314), nivolumab 3 mg/kg once every 2 weeks (n = 316), or ipilimumab 3 mg/kg once every
3 weeks (four doses; n = 315).
Aspect: i
Keyphrases: [{'sentence': 2, 'kps': ['untreated unresectable stage III or stage IV melanoma', 'nivolumab 1 mg/kg plus ipil
imumab 3 mg/kg once every 3 weeks ( four doses )', 'nivolumab 3 mg/kg once every 2 weeks', 'nivolumab 3 mg/kg once every 2
 weeks', 'ipilimumab 3 mg/kg once every 3 weeks', 'four doses']}]
----------------------------------------
Aspects distribution: {'ae': 63, 'c': 44, 'f': 69, 'i': 60, 'm': 52, 'o': 104, 'ob': 69, 'p': 101, 'pe': 42, 'r': 50, 'se'
: 17, 'td': 34, 'fd': 36, 'b': 24, 'fu': 16, 'rf': 30}
Number of unique aspects: 16
Number of unique sentences: 525
```



Number of unique aspects:

16

MODIFICATIONS

The errors you're encountering are related to activating functionality that requires working with bf16 (bfloat16 blocks) and with caching.

- **Model layers:** The error (where it can't find `encoder` or `decoder`) arises because the `unsloth` model uses a different layer structure than classic models (like those from HuggingFace). This is because the `unsloth/llama-3-8b-bnb-4bit` model has a custom architecture. When we started checking the layers for metrics, the program tried to access layers that don't exist in this model.
- **Caching:** Caching and layers are interconnected because caching implies that the model uses a buffer to store intermediate states (which speeds up computation). Disabling the cache could have affected the model's internal operations, as the model expects the cache to be enabled. Disabling caching directly led to other errors.

The specific error:

- The error you're getting says that the model and tensors are on different devices (GPU and CPU at the same time). This isn't a library issue, but rather how your code is managing the devices for the model and tensors.

Copy

```
RuntimeError: Expected all tensors to be on the same device, but found at least two devices, cuda:1 and cuda:0!
```

💡 This means that parts of the model or your inputs/labels ended up on different GPUs. It looks like your code is accidentally placing them on different devices.

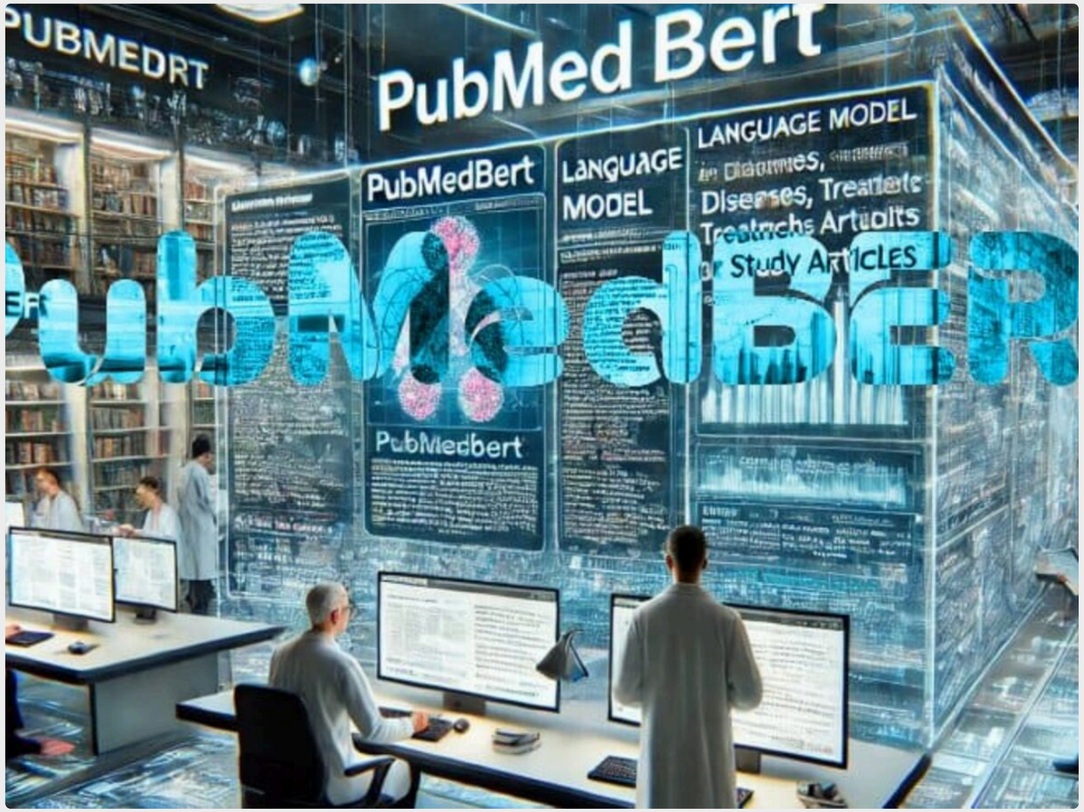
To do:

```python
def extract_key_phrases(self, text):
    """
    Extracts a list of key phrases (kps) from a string.
    """
    try:
        # Find the block "**Key Phrases**: {...}"
        pattern = r"\*\*Key Phrases\*\*:\s*\{.*?\}"
        match = re.search(pattern, text)
        if not match:
            return []

        # Extract only the dictionary as a string
        dict_str = match.group(0).split(":", 1)[1].strip()

        # Replace single quotes with double quotes (for valid JSON)
        dict_str = dict_str.replace("'", '"')

        # Load as dictionary
        kp_dict = json.loads(dict_str)

        # Return key phrases
        return kp_dict.get("kps", [])

    except Exception as e:
        print(f"[extract_key_phrases] Parsing error: {e}")
        return []
```

```
ValueError: Found input variables with inconsistent numbers of samples: [426, 0]
wandb:
```

NEW MODEL

PubMedBERT


PUBMEDRT
PubMed Bert
PubMedBert
LANGUAGE MODEL
LANGUAGE MODEL
Study Articles
PubMedbert