

# PIQA-PT

Nicholas Kluge Corrêa  
Universität Bonn  
Center for Science and Thought  
Konrad-Zuse-Platz 1-3 53227 Bonn  
kluge@uni-bonn.de

## Abstract

We introduce a small-scale dataset in Brazilian Portuguese, consisting of 100 instances of physical commonsense reasoning questions, inspired by the PIQA benchmark (Physical Interaction: Question Answering) (Bisk et al., 2019). This 100-sample dataset balances prototypical and atypical uses of objects characteristic of Brazilian culture, e.g., the string of a Berimbau is made of "metal" instead of "nylon". By adapting PIQA into Brazilian Portuguese, the dataset supports research and development of natural language understanding systems in a low-resource language setting.

## 1 Dataset Description

- **Dataset Name:** PIQA Portuguese.
- **Dataset Creator:** Nicholas Kluge Corrêa.
- **Annotations:** Human-annotated by the creator
- **Source:** Manually created using regional themes, objects, and contextual characteristics of Brazilian culture, particularly those of southern Brazilian culture.
- **Language:** Portuguese (pt-BR)
- **License:** Any open license the lead's for the "MRL 2025 Shared Task" care to adopt (as long as it is open).

## 2 Dataset Summary

This dataset is a Portuguese adaptation of the Physical Interaction Question Answering (PIQA) (Bisk et al., 2019) task, explicitly tailored to Brazilian cultural contexts. It consists of multiple-choice commonsense reasoning tasks where a prompt describes an action or situation, and the model must select the more plausible completion/solution. The dataset contains 100 examples where each prompt

has two possible solutions, and the label indicates which one is correct. The examples encompass food, traditions, regional objects, daily activities, and environmental contexts that are common to Brazil.

## 3 Dataset Structure

Every sample of this dataset has the following structure:

- `prompt` (string): The question or description of the scenario (e.g. "Para fazer tapioca, é melhor usar frigideira uma ...").
- `solution0` (string): Candidate solution 0 (e.g. "frigideira antiaderente pois ela gruda menos").
- `solution1` (string): Candidate solution 1 (e.g. "panela de pressão pois ela gruda menos").
- `label` (int): The correct solution index (0 or 1).

The dataset currently contains 100 samples.

## 4 Curation Rationale

The dataset was created to evaluate how multilingual and culturally grounded commonsense reasoning models perform when exposed to culturally specific contexts in Brazilian Portuguese. The creator manually annotated the dataset. Each prompt was paired with two possible solutions, with one labeled as the correct option.

## 5 Considerations for Using the Dataset

This dataset helps assess and improve multilingual reasoning models in culturally specific contexts, promoting the inclusion of underrepresented cultural and linguistic data in commonsense reasoning benchmarks. However, it has some limitations: it is

small (100 examples), focused on southern Brazilian culture, and may not generalize to all Brazilian or Portuguese contexts, and only provides binary choices (`solution0` vs. `solution1`). From an ethical perspective, the dataset contains no sensitive or personal data, though cultural representations are simplified and may not capture the full complexity of Brazilian diversity.

## References

Yonatan Bisk, Rowan Zellers, Ronan Le Bras, Jianfeng Gao, and Yejin Choi. 2019. [Piqa: Reasoning about physical commonsense in natural language](#).