

# S3\_Contents

September 1, 2025

```
[1]: import boto3, os
import pandas as pd
from dotenv import load_dotenv
import psycopg2
import psycopg2.extras as pgx
import json
import sys
import praw
import time, datetime as dt

[2]: load_dotenv("/home/ubuntu//deds2025b_proj/opt/reddit_pipeline/.env")
bucket_name = os.environ["LAKE_BUCKET"]
s3 = boto3.client("s3")

[3]: def print_s3_tree(bucket_name, prefix="", indent=""):
    s3 = boto3.client("s3")
    paginator = s3.get_paginator("list_objects_v2")

    # Use delimiter to simulate directories
    for page in paginator.paginate(Bucket=bucket_name, Prefix=prefix,
    ↪ Delimiter="/"):
        # Print subfolders
        if "CommonPrefixes" in page:
            for sub in page["CommonPrefixes"]:
                folder = sub["Prefix"].split("/")[-2] # get last folder name
                print(f"{indent} {folder}/")
                print_s3_tree(bucket_name, prefix=sub["Prefix"], indent=indent,
    ↪+ " ")

        # Print files
        if "Contents" in page:
            for obj in page["Contents"]:
                if obj["Key"].endswith("/"):
                    continue # skip "folder marker" objects
                file_name = obj["Key"].split("/")[-1]
                print(f"{indent} {file_name}")
```

```
[4]: print(f"S3 Bucket: {bucket_name}")
      print_s3_tree(bucket_name)
```

```
S3 Bucket: deds2025b-proj-oncolens
bronze/
  reddit/
    dt=2025-08-24/
      subreddit=AskPH/
        run_id=20250824T041150Z/
          comments.jsonl.gz
          posts.jsonl.gz
      subreddit=OffMyChestPH/
        run_id=20250824T041150Z/
          comments.jsonl.gz
          posts.jsonl.gz
      subreddit=adulthoodph/
        run_id=20250824T041150Z/
          comments.jsonl.gz
          posts.jsonl.gz
    dt=2025-08-29/
      subreddit=BRCA/
        run_id=20250829T223859Z/
          comments.jsonl.gz
          posts.jsonl.gz
      subreddit=IndustrialPharmacy/
        run_id=20250829T223859Z/
          comments.jsonl.gz
          posts.jsonl.gz
      subreddit=LivingWithMBC/
        run_id=20250829T111752Z/
          comments.jsonl.gz
          posts.jsonl.gz
        run_id=20250829T111757Z/
          comments.jsonl.gz
          posts.jsonl.gz
        run_id=20250829T112257Z/
          comments.jsonl.gz
          posts.jsonl.gz
        run_id=20250829T112302Z/
          comments.jsonl.gz
          posts.jsonl.gz
        run_id=20250829T112446Z/
          comments.jsonl.gz
          posts.jsonl.gz
        run_id=20250829T223859Z/
          comments.jsonl.gz
          posts.jsonl.gz
      subreddit=Philippines/
```

```

        run_id=20250829T084342Z/
            comments.jsonl.gz
            posts.jsonl.gz
        run_id=20250829T085448Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=breastcancer/
        run_id=20250829T223859Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=cancer/
        run_id=20250829T223859Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=ovariancancer_new/
        run_id=20250829T223859Z/
            comments.jsonl.gz
            posts.jsonl.gz
dt=2025-09-01/
    subreddit=BRCA/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=IndustrialPharmacy/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=LivingWithMBC/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=breastcancer/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=cancer/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
    subreddit=ovariancancer_new/
        run_id=20250901T141054Z/
            comments.jsonl.gz
            posts.jsonl.gz
gold/
    reddit/
        comment_enriched/
            dt=2025-08-29/
                subreddit=BRCA/

```

```

        part-0.parquet
        subreddit=IndustrialPharmacy/
        part-0.parquet
        subreddit=LivingWithMBC/
        part-0.parquet
        subreddit=Philippines/
        part-0.parquet
        subreddit=breastcancer/
        part-0.parquet
        subreddit=cancer/
        part-0.parquet
        subreddit=ovariancancer_new/
        part-0.parquet
    post_enriched/
        dt=2025-08-29/
        subreddit=BRCA/
        part-0.parquet
        subreddit=IndustrialPharmacy/
        part-0.parquet
        subreddit=LivingWithMBC/
        part-0.parquet
        subreddit=Philippines/
        part-0.parquet
        subreddit=breastcancer/
        part-0.parquet
        subreddit=cancer/
        part-0.parquet
        subreddit=ovariancancer_new/
        part-0.parquet
    marts/
        reddit/
            comments_daily/
                dt=2024-09-24/
                part_0001_part_00.parquet
                dt=2024-09-25/
                part_0000_part_00.parquet
                dt=2024-09-26/
                part_0000_part_00.parquet
                dt=2024-09-27/
                part_0001_part_00.parquet
                dt=2024-09-29/
                part_0001_part_00.parquet
                dt=2024-09-30/
                part_0000_part_00.parquet
                dt=2024-10-01/
                part_0000_part_00.parquet
                dt=2024-10-02/
                part_0000_part_00.parquet

```

dt=2024-10-03/  
part\_0000\_part\_00.parquet  
dt=2024-10-04/  
part\_0001\_part\_00.parquet  
dt=2024-10-05/  
part\_0000\_part\_00.parquet  
dt=2024-10-06/  
part\_0000\_part\_00.parquet  
dt=2024-10-07/  
part\_0000\_part\_00.parquet  
dt=2024-10-08/  
part\_0000\_part\_00.parquet  
dt=2024-10-09/  
part\_0001\_part\_00.parquet  
dt=2024-10-10/  
part\_0001\_part\_00.parquet  
dt=2024-10-12/  
part\_0001\_part\_00.parquet  
dt=2024-10-13/  
part\_0000\_part\_00.parquet  
dt=2024-10-15/  
part\_0000\_part\_00.parquet  
dt=2024-10-16/  
part\_0001\_part\_00.parquet  
dt=2024-10-17/  
part\_0001\_part\_00.parquet  
dt=2024-10-18/  
part\_0001\_part\_00.parquet  
dt=2024-10-19/  
part\_0000\_part\_00.parquet  
dt=2024-10-20/  
part\_0000\_part\_00.parquet  
dt=2024-10-21/  
part\_0000\_part\_00.parquet  
dt=2024-10-22/  
part\_0001\_part\_00.parquet  
dt=2024-10-23/  
part\_0000\_part\_00.parquet  
dt=2024-10-24/  
part\_0000\_part\_00.parquet  
dt=2024-10-25/  
part\_0001\_part\_00.parquet  
dt=2024-10-29/  
part\_0000\_part\_00.parquet  
dt=2024-10-30/  
part\_0001\_part\_00.parquet  
dt=2024-10-31/  
part\_0001\_part\_00.parquet

dt=2024-11-01/  
part\_0000\_part\_00.parquet  
dt=2024-11-02/  
part\_0001\_part\_00.parquet  
dt=2024-11-03/  
part\_0001\_part\_00.parquet  
dt=2024-11-05/  
part\_0001\_part\_00.parquet  
dt=2024-11-06/  
part\_0000\_part\_00.parquet  
dt=2024-11-07/  
part\_0000\_part\_00.parquet  
dt=2024-11-08/  
part\_0001\_part\_00.parquet  
dt=2024-11-09/  
part\_0000\_part\_00.parquet  
dt=2024-11-10/  
part\_0000\_part\_00.parquet  
dt=2024-11-12/  
part\_0001\_part\_00.parquet  
dt=2024-11-13/  
part\_0000\_part\_00.parquet  
dt=2024-11-14/  
part\_0000\_part\_00.parquet  
dt=2024-11-15/  
part\_0001\_part\_00.parquet  
dt=2024-11-16/  
part\_0001\_part\_00.parquet  
dt=2024-11-18/  
part\_0001\_part\_00.parquet  
dt=2024-11-19/  
part\_0001\_part\_00.parquet  
dt=2024-11-20/  
part\_0000\_part\_00.parquet  
dt=2024-11-21/  
part\_0001\_part\_00.parquet  
dt=2024-11-22/  
part\_0001\_part\_00.parquet  
dt=2024-11-23/  
part\_0001\_part\_00.parquet  
dt=2024-11-24/  
part\_0000\_part\_00.parquet  
dt=2024-11-25/  
part\_0001\_part\_00.parquet  
dt=2024-11-26/  
part\_0000\_part\_00.parquet  
dt=2024-11-27/  
part\_0001\_part\_00.parquet

dt=2024-11-28/  
part\_0000\_part\_00.parquet  
dt=2024-11-29/  
part\_0000\_part\_00.parquet  
dt=2024-11-30/  
part\_0000\_part\_00.parquet  
dt=2024-12-01/  
part\_0001\_part\_00.parquet  
dt=2024-12-02/  
part\_0000\_part\_00.parquet  
dt=2024-12-03/  
part\_0000\_part\_00.parquet  
dt=2024-12-04/  
part\_0001\_part\_00.parquet  
dt=2024-12-05/  
part\_0000\_part\_00.parquet  
dt=2024-12-06/  
part\_0001\_part\_00.parquet  
dt=2024-12-07/  
part\_0000\_part\_00.parquet  
dt=2024-12-08/  
part\_0001\_part\_00.parquet  
dt=2024-12-09/  
part\_0001\_part\_00.parquet  
dt=2024-12-10/  
part\_0001\_part\_00.parquet  
dt=2024-12-11/  
part\_0000\_part\_00.parquet  
dt=2024-12-12/  
part\_0000\_part\_00.parquet  
dt=2024-12-17/  
part\_0000\_part\_00.parquet  
dt=2024-12-18/  
part\_0000\_part\_00.parquet  
dt=2024-12-19/  
part\_0001\_part\_00.parquet  
dt=2024-12-20/  
part\_0001\_part\_00.parquet  
dt=2024-12-21/  
part\_0001\_part\_00.parquet  
dt=2024-12-22/  
part\_0001\_part\_00.parquet  
dt=2024-12-23/  
part\_0001\_part\_00.parquet  
dt=2024-12-24/  
part\_0001\_part\_00.parquet  
dt=2024-12-25/  
part\_0000\_part\_00.parquet

dt=2024-12-30/  
part\_0001\_part\_00.parquet  
dt=2024-12-31/  
part\_0001\_part\_00.parquet  
dt=2025-01-01/  
part\_0001\_part\_00.parquet  
dt=2025-01-02/  
part\_0001\_part\_00.parquet  
dt=2025-01-03/  
part\_0001\_part\_00.parquet  
dt=2025-01-04/  
part\_0000\_part\_00.parquet  
dt=2025-01-06/  
part\_0001\_part\_00.parquet  
dt=2025-01-07/  
part\_0000\_part\_00.parquet  
dt=2025-01-08/  
part\_0000\_part\_00.parquet  
dt=2025-01-09/  
part\_0001\_part\_00.parquet  
dt=2025-01-10/  
part\_0000\_part\_00.parquet  
dt=2025-01-11/  
part\_0001\_part\_00.parquet  
dt=2025-01-12/  
part\_0000\_part\_00.parquet  
dt=2025-01-13/  
part\_0000\_part\_00.parquet  
dt=2025-01-14/  
part\_0001\_part\_00.parquet  
dt=2025-01-15/  
part\_0001\_part\_00.parquet  
dt=2025-01-16/  
part\_0000\_part\_00.parquet  
dt=2025-01-18/  
part\_0001\_part\_00.parquet  
dt=2025-01-21/  
part\_0001\_part\_00.parquet  
dt=2025-01-26/  
part\_0001\_part\_00.parquet  
dt=2025-01-27/  
part\_0000\_part\_00.parquet  
dt=2025-01-29/  
part\_0001\_part\_00.parquet  
dt=2025-01-31/  
part\_0001\_part\_00.parquet  
dt=2025-02-03/  
part\_0000\_part\_00.parquet



dt=2025-02-04/  
part\_0000\_part\_00.parquet  
dt=2025-02-05/  
part\_0001\_part\_00.parquet  
dt=2025-02-06/  
part\_0001\_part\_00.parquet  
dt=2025-02-07/  
part\_0001\_part\_00.parquet  
dt=2025-02-09/  
part\_0000\_part\_00.parquet  
dt=2025-02-12/  
part\_0001\_part\_00.parquet  
dt=2025-02-13/  
part\_0000\_part\_00.parquet  
dt=2025-02-15/  
part\_0000\_part\_00.parquet  
dt=2025-02-16/  
part\_0001\_part\_00.parquet  
dt=2025-02-18/  
part\_0001\_part\_00.parquet  
dt=2025-02-19/  
part\_0000\_part\_00.parquet  
dt=2025-02-20/  
part\_0001\_part\_00.parquet  
dt=2025-02-22/  
part\_0000\_part\_00.parquet  
dt=2025-02-23/  
part\_0000\_part\_00.parquet  
dt=2025-02-24/  
part\_0001\_part\_00.parquet  
dt=2025-02-25/  
part\_0001\_part\_00.parquet  
dt=2025-02-26/  
part\_0001\_part\_00.parquet  
dt=2025-02-27/  
part\_0001\_part\_00.parquet  
dt=2025-02-28/  
part\_0001\_part\_00.parquet  
dt=2025-03-02/  
part\_0001\_part\_00.parquet  
dt=2025-03-05/  
part\_0000\_part\_00.parquet  
dt=2025-03-06/  
part\_0001\_part\_00.parquet  
dt=2025-03-07/  
part\_0000\_part\_00.parquet  
dt=2025-03-08/  
part\_0000\_part\_00.parquet

dt=2025-03-10/  
part\_0001\_part\_00.parquet  
dt=2025-03-11/  
part\_0001\_part\_00.parquet  
dt=2025-03-12/  
part\_0001\_part\_00.parquet  
dt=2025-03-13/  
part\_0001\_part\_00.parquet  
dt=2025-03-16/  
part\_0001\_part\_00.parquet  
dt=2025-03-17/  
part\_0001\_part\_00.parquet  
dt=2025-03-18/  
part\_0001\_part\_00.parquet  
dt=2025-03-19/  
part\_0000\_part\_00.parquet  
dt=2025-03-20/  
part\_0001\_part\_00.parquet  
dt=2025-03-21/  
part\_0001\_part\_00.parquet  
dt=2025-03-27/  
part\_0000\_part\_00.parquet  
dt=2025-03-28/  
part\_0001\_part\_00.parquet  
dt=2025-03-29/  
part\_0000\_part\_00.parquet  
dt=2025-04-02/  
part\_0001\_part\_00.parquet  
dt=2025-04-03/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-04/  
part\_0001\_part\_00.parquet  
dt=2025-04-05/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-07/  
part\_0001\_part\_00.parquet  
dt=2025-04-08/  
part\_0001\_part\_00.parquet  
dt=2025-04-09/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-10/  
part\_0001\_part\_00.parquet  
dt=2025-04-11/  
part\_0001\_part\_00.parquet  
dt=2025-04-12/

part\_0000\_part\_00.parquet  
dt=2025-04-13/  
part\_0000\_part\_00.parquet  
dt=2025-04-14/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-16/  
part\_0001\_part\_00.parquet  
dt=2025-04-17/  
part\_0000\_part\_00.parquet  
dt=2025-04-18/  
part\_0000\_part\_00.parquet  
dt=2025-04-19/  
part\_0001\_part\_00.parquet  
dt=2025-04-20/  
part\_0001\_part\_00.parquet  
dt=2025-04-21/  
part\_0000\_part\_00.parquet  
dt=2025-04-22/  
part\_0001\_part\_00.parquet  
dt=2025-04-23/  
part\_0001\_part\_00.parquet  
dt=2025-04-24/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-25/  
part\_0000\_part\_00.parquet  
dt=2025-04-26/  
part\_0001\_part\_00.parquet  
dt=2025-04-27/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-28/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-29/  
part\_0001\_part\_00.parquet  
dt=2025-04-30/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-01/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-02/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-03/  
part\_0001\_part\_00.parquet

dt=2025-05-04/  
part\_0001\_part\_00.parquet  
dt=2025-05-05/  
part\_0000\_part\_00.parquet  
dt=2025-05-06/  
part\_0001\_part\_00.parquet  
dt=2025-05-07/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-08/  
part\_0001\_part\_00.parquet  
dt=2025-05-09/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-10/  
part\_0001\_part\_00.parquet  
dt=2025-05-11/  
part\_0001\_part\_00.parquet  
dt=2025-05-12/  
part\_0001\_part\_00.parquet  
dt=2025-05-13/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-14/  
part\_0001\_part\_00.parquet  
dt=2025-05-15/  
part\_0001\_part\_00.parquet  
dt=2025-05-16/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-17/  
part\_0000\_part\_00.parquet  
dt=2025-05-18/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-19/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-20/  
part\_0001\_part\_00.parquet  
dt=2025-05-21/  
part\_0001\_part\_00.parquet  
dt=2025-05-22/  
part\_0000\_part\_00.parquet  
dt=2025-05-23/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-24/

```

    part_0001_part_00.parquet
dt=2025-05-25/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-05-26/
    part_0001_part_00.parquet
dt=2025-05-27/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-05-28/
    part_0000_part_00.parquet
dt=2025-05-29/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-05-30/
    part_0000_part_00.parquet
dt=2025-05-31/
    part_0001_part_00.parquet
dt=2025-06-01/
    part_0000_part_00.parquet
dt=2025-06-02/
    part_0001_part_00.parquet
dt=2025-06-03/
    part_0001_part_00.parquet
dt=2025-06-04/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-06-05/
    part_0000_part_00.parquet
dt=2025-06-06/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-06-08/
    part_0000_part_00.parquet
dt=2025-06-09/
    part_0000_part_00.parquet
dt=2025-06-10/
    part_0001_part_00.parquet
dt=2025-06-11/
    part_0001_part_00.parquet
dt=2025-06-12/
    part_0000_part_00.parquet
dt=2025-06-13/
    part_0001_part_00.parquet
dt=2025-06-14/
    part_0001_part_00.parquet
dt=2025-06-15/
    part_0000_part_00.parquet

```

dt=2025-06-16/  
part\_0001\_part\_00.parquet  
dt=2025-06-17/  
part\_0000\_part\_00.parquet  
dt=2025-06-18/  
part\_0000\_part\_00.parquet  
dt=2025-06-19/  
part\_0001\_part\_00.parquet  
dt=2025-06-20/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-06-21/  
part\_0001\_part\_00.parquet  
dt=2025-06-22/  
part\_0000\_part\_00.parquet  
dt=2025-06-23/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-06-24/  
part\_0001\_part\_00.parquet  
dt=2025-06-25/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-06-26/  
part\_0001\_part\_00.parquet  
dt=2025-06-27/  
part\_0000\_part\_00.parquet  
dt=2025-06-28/  
part\_0000\_part\_00.parquet  
dt=2025-06-29/  
part\_0000\_part\_00.parquet  
dt=2025-06-30/  
part\_0000\_part\_00.parquet  
dt=2025-07-01/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-02/  
part\_0001\_part\_00.parquet  
dt=2025-07-03/  
part\_0001\_part\_00.parquet  
dt=2025-07-04/  
part\_0001\_part\_00.parquet  
dt=2025-07-05/  
part\_0000\_part\_00.parquet  
dt=2025-07-06/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-07/

part\_0001\_part\_00.parquet  
dt=2025-07-08/  
part\_0001\_part\_00.parquet  
dt=2025-07-09/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-10/  
part\_0000\_part\_00.parquet  
dt=2025-07-11/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-12/  
part\_0001\_part\_00.parquet  
dt=2025-07-13/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-14/  
part\_0001\_part\_00.parquet  
dt=2025-07-15/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-16/  
part\_0001\_part\_00.parquet  
dt=2025-07-17/  
part\_0000\_part\_00.parquet  
dt=2025-07-18/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-19/  
part\_0000\_part\_00.parquet  
dt=2025-07-20/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-21/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-22/  
part\_0000\_part\_00.parquet  
dt=2025-07-23/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-24/  
part\_0001\_part\_00.parquet  
dt=2025-07-25/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-26/  
part\_0000\_part\_00.parquet

```

    part_0001_part_00.parquet
dt=2025-07-27/
    part_0000_part_00.parquet
dt=2025-07-28/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-07-29/
    part_0001_part_00.parquet
dt=2025-07-30/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-07-31/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-01/
    part_0000_part_00.parquet
dt=2025-08-02/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-03/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-04/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-05/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-06/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-07/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-08/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-09/
    part_0001_part_00.parquet
dt=2025-08-10/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-11/
    part_0000_part_00.parquet
    part_0001_part_00.parquet
dt=2025-08-12/
    part_0000_part_00.parquet
    part_0001_part_00.parquet

```



dt=2025-08-13/  
part\_0000\_part\_00.parquet  
dt=2025-08-14/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-15/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-16/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-17/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-18/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-19/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-20/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-21/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-22/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-23/  
part\_0001\_part\_00.parquet  
dt=2025-08-24/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-25/  
part\_0001\_part\_00.parquet  
dt=2025-08-26/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-27/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-28/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-29/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet

```
subreddit_daily/  
  dt=2024-09-24/  
    part_0000_part_00.parquet  
  dt=2024-09-26/  
    part_0001_part_00.parquet  
  dt=2024-09-29/  
    part_0000_part_00.parquet  
  dt=2024-10-01/  
    part_0000_part_00.parquet  
  dt=2024-10-02/  
    part_0001_part_00.parquet  
  dt=2024-10-05/  
    part_0001_part_00.parquet  
  dt=2024-10-06/  
    part_0000_part_00.parquet  
  dt=2024-10-08/  
    part_0000_part_00.parquet  
  dt=2024-10-09/  
    part_0001_part_00.parquet  
  dt=2024-10-10/  
    part_0001_part_00.parquet  
  dt=2024-10-15/  
    part_0001_part_00.parquet  
  dt=2024-10-16/  
    part_0000_part_00.parquet  
  dt=2024-10-20/  
    part_0000_part_00.parquet  
  dt=2024-10-22/  
    part_0000_part_00.parquet  
  dt=2024-10-23/  
    part_0001_part_00.parquet  
  dt=2024-10-31/  
    part_0001_part_00.parquet  
  dt=2024-11-01/  
    part_0000_part_00.parquet  
  dt=2024-11-05/  
    part_0000_part_00.parquet  
  dt=2024-11-07/  
    part_0001_part_00.parquet  
  dt=2024-11-09/  
    part_0000_part_00.parquet  
  dt=2024-11-11/  
    part_0001_part_00.parquet  
  dt=2024-11-14/  
    part_0000_part_00.parquet  
  dt=2024-11-18/  
    part_0001_part_00.parquet  
  dt=2024-11-19/
```

part\_0000\_part\_00.parquet  
dt=2024-11-21/  
part\_0000\_part\_00.parquet  
dt=2024-11-22/  
part\_0001\_part\_00.parquet  
dt=2024-11-23/  
part\_0001\_part\_00.parquet  
dt=2024-11-26/  
part\_0000\_part\_00.parquet  
dt=2024-11-28/  
part\_0001\_part\_00.parquet  
dt=2024-11-29/  
part\_0001\_part\_00.parquet  
dt=2024-11-30/  
part\_0000\_part\_00.parquet  
dt=2024-12-05/  
part\_0000\_part\_00.parquet  
dt=2024-12-06/  
part\_0001\_part\_00.parquet  
dt=2024-12-10/  
part\_0001\_part\_00.parquet  
dt=2024-12-11/  
part\_0001\_part\_00.parquet  
dt=2024-12-18/  
part\_0001\_part\_00.parquet  
dt=2024-12-19/  
part\_0001\_part\_00.parquet  
dt=2024-12-20/  
part\_0000\_part\_00.parquet  
dt=2024-12-23/  
part\_0001\_part\_00.parquet  
dt=2024-12-30/  
part\_0001\_part\_00.parquet  
dt=2024-12-31/  
part\_0000\_part\_00.parquet  
dt=2025-01-03/  
part\_0000\_part\_00.parquet  
dt=2025-01-07/  
part\_0000\_part\_00.parquet  
dt=2025-01-09/  
part\_0000\_part\_00.parquet  
dt=2025-01-10/  
part\_0000\_part\_00.parquet  
dt=2025-01-11/  
part\_0001\_part\_00.parquet  
dt=2025-01-12/  
part\_0000\_part\_00.parquet  
dt=2025-01-13/

part\_0000\_part\_00.parquet  
dt=2025-01-14/  
part\_0000\_part\_00.parquet  
dt=2025-01-26/  
part\_0000\_part\_00.parquet  
dt=2025-01-31/  
part\_0000\_part\_00.parquet  
dt=2025-02-02/  
part\_0000\_part\_00.parquet  
dt=2025-02-03/  
part\_0000\_part\_00.parquet  
dt=2025-02-06/  
part\_0000\_part\_00.parquet  
dt=2025-02-12/  
part\_0000\_part\_00.parquet  
dt=2025-02-15/  
part\_0001\_part\_00.parquet  
dt=2025-02-18/  
part\_0001\_part\_00.parquet  
dt=2025-02-20/  
part\_0000\_part\_00.parquet  
dt=2025-02-22/  
part\_0001\_part\_00.parquet  
dt=2025-02-25/  
part\_0000\_part\_00.parquet  
dt=2025-02-26/  
part\_0000\_part\_00.parquet  
dt=2025-03-05/  
part\_0001\_part\_00.parquet  
dt=2025-03-10/  
part\_0000\_part\_00.parquet  
dt=2025-03-15/  
part\_0001\_part\_00.parquet  
dt=2025-03-17/  
part\_0000\_part\_00.parquet  
dt=2025-03-27/  
part\_0000\_part\_00.parquet  
dt=2025-04-02/  
part\_0000\_part\_00.parquet  
dt=2025-04-03/  
part\_0001\_part\_00.parquet  
dt=2025-04-05/  
part\_0001\_part\_00.parquet  
dt=2025-04-07/  
part\_0001\_part\_00.parquet  
dt=2025-04-08/  
part\_0000\_part\_00.parquet  
dt=2025-04-09/

part\_0000\_part\_00.parquet  
dt=2025-04-11/  
part\_0001\_part\_00.parquet  
dt=2025-04-13/  
part\_0001\_part\_00.parquet  
dt=2025-04-16/  
part\_0001\_part\_00.parquet  
dt=2025-04-17/  
part\_0001\_part\_00.parquet  
dt=2025-04-18/  
part\_0001\_part\_00.parquet  
dt=2025-04-19/  
part\_0000\_part\_00.parquet  
dt=2025-04-20/  
part\_0000\_part\_00.parquet  
dt=2025-04-22/  
part\_0001\_part\_00.parquet  
dt=2025-04-23/  
part\_0000\_part\_00.parquet  
dt=2025-04-24/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-25/  
part\_0000\_part\_00.parquet  
dt=2025-04-26/  
part\_0000\_part\_00.parquet  
dt=2025-04-27/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-28/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-04-29/  
part\_0001\_part\_00.parquet  
dt=2025-04-30/  
part\_0001\_part\_00.parquet  
dt=2025-05-01/  
part\_0000\_part\_00.parquet  
dt=2025-05-02/  
part\_0000\_part\_00.parquet  
dt=2025-05-03/  
part\_0000\_part\_00.parquet  
dt=2025-05-04/  
part\_0000\_part\_00.parquet  
dt=2025-05-05/  
part\_0000\_part\_00.parquet  
dt=2025-05-06/  
part\_0001\_part\_00.parquet

dt=2025-05-07/  
part\_0001\_part\_00.parquet  
dt=2025-05-08/  
part\_0001\_part\_00.parquet  
dt=2025-05-09/  
part\_0000\_part\_00.parquet  
dt=2025-05-14/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-15/  
part\_0001\_part\_00.parquet  
dt=2025-05-16/  
part\_0000\_part\_00.parquet  
dt=2025-05-17/  
part\_0001\_part\_00.parquet  
dt=2025-05-18/  
part\_0001\_part\_00.parquet  
dt=2025-05-19/  
part\_0001\_part\_00.parquet  
dt=2025-05-21/  
part\_0000\_part\_00.parquet  
dt=2025-05-22/  
part\_0000\_part\_00.parquet  
dt=2025-05-23/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-24/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-05-27/  
part\_0000\_part\_00.parquet  
dt=2025-05-28/  
part\_0000\_part\_00.parquet  
dt=2025-05-29/  
part\_0001\_part\_00.parquet  
dt=2025-05-30/  
part\_0000\_part\_00.parquet  
dt=2025-05-31/  
part\_0000\_part\_00.parquet  
dt=2025-06-04/  
part\_0000\_part\_00.parquet  
dt=2025-06-08/  
part\_0000\_part\_00.parquet  
dt=2025-06-09/  
part\_0000\_part\_00.parquet  
dt=2025-06-11/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet

dt=2025-06-13/  
part\_0001\_part\_00.parquet  
dt=2025-06-14/  
part\_0001\_part\_00.parquet  
dt=2025-06-15/  
part\_0001\_part\_00.parquet  
dt=2025-06-17/  
part\_0000\_part\_00.parquet  
dt=2025-06-18/  
part\_0001\_part\_00.parquet  
dt=2025-06-20/  
part\_0001\_part\_00.parquet  
dt=2025-06-21/  
part\_0000\_part\_00.parquet  
dt=2025-06-22/  
part\_0000\_part\_00.parquet  
dt=2025-06-23/  
part\_0000\_part\_00.parquet  
dt=2025-06-24/  
part\_0001\_part\_00.parquet  
dt=2025-06-25/  
part\_0000\_part\_00.parquet  
dt=2025-06-26/  
part\_0001\_part\_00.parquet  
dt=2025-06-27/  
part\_0001\_part\_00.parquet  
dt=2025-06-28/  
part\_0001\_part\_00.parquet  
dt=2025-06-29/  
part\_0001\_part\_00.parquet  
dt=2025-07-01/  
part\_0001\_part\_00.parquet  
dt=2025-07-02/  
part\_0001\_part\_00.parquet  
dt=2025-07-03/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-05/  
part\_0000\_part\_00.parquet  
dt=2025-07-06/  
part\_0000\_part\_00.parquet  
dt=2025-07-07/  
part\_0000\_part\_00.parquet  
dt=2025-07-08/  
part\_0000\_part\_00.parquet  
dt=2025-07-09/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet

dt=2025-07-10/  
part\_0000\_part\_00.parquet  
dt=2025-07-11/  
part\_0000\_part\_00.parquet  
dt=2025-07-12/  
part\_0000\_part\_00.parquet  
dt=2025-07-13/  
part\_0000\_part\_00.parquet  
dt=2025-07-16/  
part\_0001\_part\_00.parquet  
dt=2025-07-18/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-19/  
part\_0001\_part\_00.parquet  
dt=2025-07-20/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-21/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-22/  
part\_0001\_part\_00.parquet  
dt=2025-07-23/  
part\_0001\_part\_00.parquet  
dt=2025-07-24/  
part\_0001\_part\_00.parquet  
dt=2025-07-25/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-26/  
part\_0001\_part\_00.parquet  
dt=2025-07-27/  
part\_0001\_part\_00.parquet  
dt=2025-07-28/  
part\_0000\_part\_00.parquet  
dt=2025-07-29/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-30/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-07-31/  
part\_0001\_part\_00.parquet  
dt=2025-08-01/  
part\_0001\_part\_00.parquet  
dt=2025-08-02/  
part\_0000\_part\_00.parquet



part\_0001\_part\_00.parquet  
dt=2025-08-03/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-04/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-05/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-06/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-07/  
part\_0001\_part\_00.parquet  
dt=2025-08-08/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-09/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-10/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-11/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-12/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-13/  
part\_0000\_part\_00.parquet  
dt=2025-08-14/  
part\_0001\_part\_00.parquet  
dt=2025-08-15/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-16/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-17/  
part\_0000\_part\_00.parquet  
dt=2025-08-18/  
part\_0000\_part\_00.parquet  
dt=2025-08-19/  
part\_0000\_part\_00.parquet  
part\_0001\_part\_00.parquet  
dt=2025-08-20/

```

        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-21/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-22/
        part_0001_part_00.parquet
dt=2025-08-23/
        part_0001_part_00.parquet
dt=2025-08-24/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-25/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-26/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-27/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-28/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
dt=2025-08-29/
        part_0000_part_00.parquet
        part_0001_part_00.parquet
tags_daily/
dt=2024-09-24/
        part_0001_part_00.parquet
dt=2024-09-25/
        part_0000_part_00.parquet
dt=2024-09-26/
        part_0001_part_00.parquet
dt=2024-09-27/
        part_0001_part_00.parquet
dt=2024-09-29/
        part_0000_part_00.parquet
dt=2024-09-30/
        part_0001_part_00.parquet
dt=2024-10-01/
        part_0001_part_00.parquet
dt=2024-10-02/
        part_0000_part_00.parquet
dt=2024-10-03/
        part_0000_part_00.parquet
dt=2024-10-04/
        part_0001_part_00.parquet

```

dt=2024-10-05/  
part\_0001\_part\_00.parquet  
dt=2024-10-06/  
part\_0001\_part\_00.parquet  
dt=2024-10-07/  
part\_0001\_part\_00.parquet  
dt=2024-10-08/  
part\_0000\_part\_00.parquet  
dt=2024-10-09/  
part\_0001\_part\_00.parquet  
dt=2024-10-10/  
part\_0001\_part\_00.parquet  
dt=2024-10-13/  
part\_0000\_part\_00.parquet  
dt=2024-10-15/  
part\_0000\_part\_00.parquet  
dt=2024-10-16/  
part\_0001\_part\_00.parquet  
dt=2024-10-17/  
part\_0000\_part\_00.parquet  
dt=2024-10-18/  
part\_0000\_part\_00.parquet  
dt=2024-10-20/  
part\_0000\_part\_00.parquet  
dt=2024-10-21/  
part\_0000\_part\_00.parquet  
dt=2024-10-22/  
part\_0000\_part\_00.parquet  
dt=2024-10-23/  
part\_0001\_part\_00.parquet  
dt=2024-10-25/  
part\_0000\_part\_00.parquet  
dt=2024-10-30/  
part\_0000\_part\_00.parquet  
dt=2024-10-31/  
part\_0001\_part\_00.parquet  
dt=2024-11-01/  
part\_0000\_part\_00.parquet  
dt=2024-11-02/  
part\_0000\_part\_00.parquet  
dt=2024-11-03/  
part\_0001\_part\_00.parquet  
dt=2024-11-05/  
part\_0001\_part\_00.parquet  
dt=2024-11-07/  
part\_0000\_part\_00.parquet  
dt=2024-11-08/  
part\_0000\_part\_00.parquet

dt=2024-11-09/  
part\_0001\_part\_00.parquet  
dt=2024-11-10/  
part\_0000\_part\_00.parquet  
dt=2024-11-11/  
part\_0000\_part\_00.parquet  
dt=2024-11-12/  
part\_0000\_part\_00.parquet  
dt=2024-11-14/  
part\_0000\_part\_00.parquet  
dt=2024-11-15/  
part\_0000\_part\_00.parquet  
dt=2024-11-16/  
part\_0001\_part\_00.parquet  
dt=2024-11-18/  
part\_0000\_part\_00.parquet  
dt=2024-11-19/  
part\_0001\_part\_00.parquet  
dt=2024-11-20/  
part\_0001\_part\_00.parquet  
dt=2024-11-21/  
part\_0001\_part\_00.parquet  
dt=2024-11-22/  
part\_0001\_part\_00.parquet  
dt=2024-11-23/  
part\_0001\_part\_00.parquet  
dt=2024-11-24/  
part\_0001\_part\_00.parquet  
dt=2024-11-25/  
part\_0001\_part\_00.parquet  
dt=2024-11-26/  
part\_0001\_part\_00.parquet  
dt=2024-11-27/  
part\_0000\_part\_00.parquet  
dt=2024-11-28/  
part\_0000\_part\_00.parquet  
dt=2024-11-29/  
part\_0001\_part\_00.parquet  
dt=2024-11-30/  
part\_0000\_part\_00.parquet  
dt=2024-12-01/  
part\_0001\_part\_00.parquet  
dt=2024-12-02/  
part\_0001\_part\_00.parquet  
dt=2024-12-03/  
part\_0001\_part\_00.parquet  
dt=2024-12-05/  
part\_0000\_part\_00.parquet

dt=2024-12-06/  
part\_0000\_part\_00.parquet  
dt=2024-12-07/  
part\_0001\_part\_00.parquet  
dt=2024-12-08/  
part\_0001\_part\_00.parquet  
dt=2024-12-09/  
part\_0000\_part\_00.parquet  
dt=2024-12-11/  
part\_0001\_part\_00.parquet  
dt=2024-12-17/  
part\_0000\_part\_00.parquet  
dt=2024-12-19/  
part\_0001\_part\_00.parquet  
dt=2024-12-20/  
part\_0000\_part\_00.parquet  
dt=2024-12-23/  
part\_0001\_part\_00.parquet  
dt=2024-12-30/  
part\_0000\_part\_00.parquet  
dt=2024-12-31/  
part\_0000\_part\_00.parquet  
dt=2025-01-03/  
part\_0000\_part\_00.parquet  
dt=2025-01-04/  
part\_0001\_part\_00.parquet  
dt=2025-01-07/  
part\_0001\_part\_00.parquet  
dt=2025-01-08/  
part\_0000\_part\_00.parquet  
dt=2025-01-09/  
part\_0000\_part\_00.parquet  
dt=2025-01-10/  
part\_0001\_part\_00.parquet  
dt=2025-01-11/  
part\_0000\_part\_00.parquet  
dt=2025-01-12/  
part\_0001\_part\_00.parquet  
dt=2025-01-13/  
part\_0000\_part\_00.parquet  
dt=2025-01-14/  
part\_0001\_part\_00.parquet  
dt=2025-01-16/  
part\_0001\_part\_00.parquet  
dt=2025-01-18/  
part\_0001\_part\_00.parquet  
dt=2025-01-21/  
part\_0000\_part\_00.parquet

dt=2025-01-26/  
part\_0001\_part\_00.parquet  
dt=2025-01-27/  
part\_0001\_part\_00.parquet  
dt=2025-01-29/  
part\_0001\_part\_00.parquet  
dt=2025-01-31/  
part\_0000\_part\_00.parquet  
dt=2025-02-02/  
part\_0001\_part\_00.parquet  
dt=2025-02-03/  
part\_0001\_part\_00.parquet  
dt=2025-02-04/  
part\_0000\_part\_00.parquet  
dt=2025-02-05/  
part\_0001\_part\_00.parquet  
dt=2025-02-06/  
part\_0000\_part\_00.parquet  
dt=2025-02-07/  
part\_0000\_part\_00.parquet  
dt=2025-02-09/  
part\_0001\_part\_00.parquet  
dt=2025-02-12/  
part\_0001\_part\_00.parquet  
dt=2025-02-13/  
part\_0000\_part\_00.parquet  
dt=2025-02-15/  
part\_0000\_part\_00.parquet  
dt=2025-02-16/  
part\_0000\_part\_00.parquet  
dt=2025-02-18/  
part\_0001\_part\_00.parquet  
dt=2025-02-19/  
part\_0000\_part\_00.parquet  
dt=2025-02-20/  
part\_0000\_part\_00.parquet  
dt=2025-02-22/  
part\_0001\_part\_00.parquet  
dt=2025-02-23/  
part\_0000\_part\_00.parquet  
dt=2025-02-24/  
part\_0001\_part\_00.parquet  
dt=2025-02-25/  
part\_0001\_part\_00.parquet  
dt=2025-02-26/  
part\_0000\_part\_00.parquet  
dt=2025-02-27/  
part\_0001\_part\_00.parquet

dt=2025-02-28/  
part\_0001\_part\_00.parquet  
dt=2025-03-02/  
part\_0001\_part\_00.parquet  
dt=2025-03-05/  
part\_0001\_part\_00.parquet  
dt=2025-03-06/  
part\_0001\_part\_00.parquet  
dt=2025-03-07/  
part\_0000\_part\_00.parquet  
dt=2025-03-10/  
part\_0000\_part\_00.parquet  
dt=2025-03-11/  
part\_0001\_part\_00.parquet  
dt=2025-03-12/  
part\_0000\_part\_00.parquet  
dt=2025-03-13/  
part\_0001\_part\_00.parquet  
dt=2025-03-15/  
part\_0001\_part\_00.parquet  
dt=2025-03-16/  
part\_0000\_part\_00.parquet  
dt=2025-03-17/  
part\_0000\_part\_00.parquet  
dt=2025-03-18/  
part\_0001\_part\_00.parquet  
dt=2025-03-19/  
part\_0001\_part\_00.parquet  
dt=2025-03-20/  
part\_0001\_part\_00.parquet  
dt=2025-03-27/  
part\_0000\_part\_00.parquet  
dt=2025-03-28/  
part\_0001\_part\_00.parquet  
dt=2025-03-29/  
part\_0001\_part\_00.parquet  
dt=2025-04-02/  
part\_0000\_part\_00.parquet  
dt=2025-04-03/  
part\_0000\_part\_00.parquet  
dt=2025-04-04/  
part\_0001\_part\_00.parquet  
dt=2025-04-05/  
part\_0000\_part\_00.parquet  
dt=2025-04-07/  
part\_0000\_part\_00.parquet  
dt=2025-04-08/  
part\_0001\_part\_00.parquet

dt=2025-04-09/  
part\_0001\_part\_00.parquet  
dt=2025-04-11/  
part\_0001\_part\_00.parquet  
dt=2025-04-12/  
part\_0001\_part\_00.parquet  
dt=2025-04-13/  
part\_0000\_part\_00.parquet  
dt=2025-04-14/  
part\_0001\_part\_00.parquet  
dt=2025-04-16/  
part\_0000\_part\_00.parquet  
dt=2025-04-17/  
part\_0001\_part\_00.parquet  
dt=2025-04-18/  
part\_0001\_part\_00.parquet  
dt=2025-04-19/  
part\_0000\_part\_00.parquet  
dt=2025-04-20/  
part\_0000\_part\_00.parquet  
dt=2025-04-21/  
part\_0001\_part\_00.parquet  
dt=2025-04-22/  
part\_0001\_part\_00.parquet  
dt=2025-04-23/  
part\_0000\_part\_00.parquet  
dt=2025-04-24/  
part\_0001\_part\_00.parquet  
dt=2025-04-25/  
part\_0001\_part\_00.parquet  
dt=2025-04-26/  
part\_0000\_part\_00.parquet  
dt=2025-04-27/  
part\_0000\_part\_00.parquet  
dt=2025-04-28/  
part\_0001\_part\_00.parquet  
dt=2025-04-29/  
part\_0001\_part\_00.parquet  
dt=2025-04-30/  
part\_0000\_part\_00.parquet  
dt=2025-05-01/  
part\_0000\_part\_00.parquet  
dt=2025-05-02/  
part\_0001\_part\_00.parquet  
dt=2025-05-03/  
part\_0000\_part\_00.parquet  
dt=2025-05-04/  
part\_0001\_part\_00.parquet



dt=2025-05-05/  
part\_0001\_part\_00.parquet  
dt=2025-05-06/  
part\_0000\_part\_00.parquet  
dt=2025-05-07/  
part\_0000\_part\_00.parquet  
dt=2025-05-08/  
part\_0001\_part\_00.parquet  
dt=2025-05-09/  
part\_0001\_part\_00.parquet  
dt=2025-05-10/  
part\_0000\_part\_00.parquet  
dt=2025-05-11/  
part\_0000\_part\_00.parquet  
dt=2025-05-12/  
part\_0000\_part\_00.parquet  
dt=2025-05-13/  
part\_0001\_part\_00.parquet  
dt=2025-05-14/  
part\_0001\_part\_00.parquet  
dt=2025-05-15/  
part\_0000\_part\_00.parquet  
dt=2025-05-16/  
part\_0001\_part\_00.parquet  
dt=2025-05-17/  
part\_0001\_part\_00.parquet  
dt=2025-05-18/  
part\_0000\_part\_00.parquet  
dt=2025-05-19/  
part\_0001\_part\_00.parquet  
dt=2025-05-20/  
part\_0001\_part\_00.parquet  
dt=2025-05-21/  
part\_0001\_part\_00.parquet  
dt=2025-05-22/  
part\_0000\_part\_00.parquet  
dt=2025-05-23/  
part\_0001\_part\_00.parquet  
dt=2025-05-24/  
part\_0000\_part\_00.parquet  
dt=2025-05-25/  
part\_0001\_part\_00.parquet  
dt=2025-05-26/  
part\_0001\_part\_00.parquet  
dt=2025-05-27/  
part\_0001\_part\_00.parquet  
dt=2025-05-28/  
part\_0001\_part\_00.parquet

dt=2025-05-29/  
part\_0000\_part\_00.parquet  
dt=2025-05-30/  
part\_0001\_part\_00.parquet  
dt=2025-05-31/  
part\_0000\_part\_00.parquet  
dt=2025-06-01/  
part\_0001\_part\_00.parquet  
dt=2025-06-02/  
part\_0001\_part\_00.parquet  
dt=2025-06-03/  
part\_0000\_part\_00.parquet  
dt=2025-06-04/  
part\_0001\_part\_00.parquet  
dt=2025-06-05/  
part\_0001\_part\_00.parquet  
dt=2025-06-06/  
part\_0001\_part\_00.parquet  
dt=2025-06-08/  
part\_0000\_part\_00.parquet  
dt=2025-06-09/  
part\_0001\_part\_00.parquet  
dt=2025-06-10/  
part\_0000\_part\_00.parquet  
dt=2025-06-11/  
part\_0000\_part\_00.parquet  
dt=2025-06-12/  
part\_0000\_part\_00.parquet  
dt=2025-06-13/  
part\_0000\_part\_00.parquet  
dt=2025-06-14/  
part\_0000\_part\_00.parquet  
dt=2025-06-15/  
part\_0001\_part\_00.parquet  
dt=2025-06-16/  
part\_0001\_part\_00.parquet  
dt=2025-06-17/  
part\_0001\_part\_00.parquet  
dt=2025-06-18/  
part\_0001\_part\_00.parquet  
dt=2025-06-19/  
part\_0001\_part\_00.parquet  
dt=2025-06-20/  
part\_0001\_part\_00.parquet  
dt=2025-06-21/  
part\_0001\_part\_00.parquet  
dt=2025-06-22/  
part\_0001\_part\_00.parquet

dt=2025-06-23/  
part\_0001\_part\_00.parquet  
dt=2025-06-24/  
part\_0000\_part\_00.parquet  
dt=2025-06-25/  
part\_0000\_part\_00.parquet  
dt=2025-06-26/  
part\_0001\_part\_00.parquet  
dt=2025-06-27/  
part\_0001\_part\_00.parquet  
dt=2025-06-28/  
part\_0001\_part\_00.parquet  
dt=2025-06-29/  
part\_0001\_part\_00.parquet  
dt=2025-06-30/  
part\_0000\_part\_00.parquet  
dt=2025-07-01/  
part\_0000\_part\_00.parquet  
dt=2025-07-02/  
part\_0001\_part\_00.parquet  
dt=2025-07-03/  
part\_0001\_part\_00.parquet  
dt=2025-07-04/  
part\_0001\_part\_00.parquet  
dt=2025-07-05/  
part\_0001\_part\_00.parquet  
dt=2025-07-06/  
part\_0001\_part\_00.parquet  
dt=2025-07-07/  
part\_0001\_part\_00.parquet  
dt=2025-07-08/  
part\_0001\_part\_00.parquet  
dt=2025-07-09/  
part\_0001\_part\_00.parquet  
dt=2025-07-10/  
part\_0000\_part\_00.parquet  
dt=2025-07-11/  
part\_0000\_part\_00.parquet  
dt=2025-07-12/  
part\_0000\_part\_00.parquet  
dt=2025-07-13/  
part\_0000\_part\_00.parquet  
dt=2025-07-14/  
part\_0001\_part\_00.parquet  
dt=2025-07-15/  
part\_0000\_part\_00.parquet  
dt=2025-07-16/  
part\_0000\_part\_00.parquet

dt=2025-07-17/  
part\_0001\_part\_00.parquet  
dt=2025-07-18/  
part\_0000\_part\_00.parquet  
dt=2025-07-19/  
part\_0001\_part\_00.parquet  
dt=2025-07-20/  
part\_0000\_part\_00.parquet  
dt=2025-07-21/  
part\_0001\_part\_00.parquet  
dt=2025-07-22/  
part\_0001\_part\_00.parquet  
dt=2025-07-23/  
part\_0001\_part\_00.parquet  
dt=2025-07-24/  
part\_0000\_part\_00.parquet  
dt=2025-07-25/  
part\_0001\_part\_00.parquet  
dt=2025-07-26/  
part\_0001\_part\_00.parquet  
dt=2025-07-27/  
part\_0000\_part\_00.parquet  
dt=2025-07-28/  
part\_0000\_part\_00.parquet  
dt=2025-07-29/  
part\_0001\_part\_00.parquet  
dt=2025-07-30/  
part\_0001\_part\_00.parquet  
dt=2025-07-31/  
part\_0001\_part\_00.parquet  
dt=2025-08-01/  
part\_0000\_part\_00.parquet  
dt=2025-08-02/  
part\_0001\_part\_00.parquet  
dt=2025-08-03/  
part\_0001\_part\_00.parquet  
dt=2025-08-04/  
part\_0001\_part\_00.parquet  
dt=2025-08-05/  
part\_0000\_part\_00.parquet  
dt=2025-08-06/  
part\_0000\_part\_00.parquet  
dt=2025-08-07/  
part\_0000\_part\_00.parquet  
dt=2025-08-08/  
part\_0001\_part\_00.parquet  
dt=2025-08-09/  
part\_0000\_part\_00.parquet

```

dt=2025-08-10/
    part_0000_part_00.parquet
dt=2025-08-11/
    part_0000_part_00.parquet
dt=2025-08-12/
    part_0000_part_00.parquet
dt=2025-08-13/
    part_0000_part_00.parquet
dt=2025-08-14/
    part_0000_part_00.parquet
dt=2025-08-15/
    part_0001_part_00.parquet
dt=2025-08-16/
    part_0001_part_00.parquet
dt=2025-08-17/
    part_0000_part_00.parquet
dt=2025-08-18/
    part_0001_part_00.parquet
dt=2025-08-19/
    part_0000_part_00.parquet
dt=2025-08-20/
    part_0000_part_00.parquet
dt=2025-08-21/
    part_0000_part_00.parquet
dt=2025-08-22/
    part_0001_part_00.parquet
dt=2025-08-23/
    part_0000_part_00.parquet
dt=2025-08-24/
    part_0000_part_00.parquet
dt=2025-08-25/
    part_0001_part_00.parquet
dt=2025-08-26/
    part_0000_part_00.parquet
dt=2025-08-27/
    part_0000_part_00.parquet
dt=2025-08-28/
    part_0000_part_00.parquet
dt=2025-08-29/
    part_0001_part_00.parquet
silver/
    reddit/
        authors/
            dt=2025-08-29/
                part-0.parquet
            dt=2025-09-01/
                part-0.parquet
        comments/

```

```

dt=2025-08-29/
  subreddit=BRCA/
    part-0.parquet
  subreddit=IndustrialPharmacy/
    part-0.parquet
  subreddit=LivingWithMBC/
    part-0.parquet
  subreddit=Philippines/
    part-0.parquet
  subreddit=breastcancer/
    part-0.parquet
  subreddit=cancer/
    part-0.parquet
  subreddit=ovariancancer_new/
    part-0.parquet
dt=2025-09-01/
  subreddit=BRCA/
    part-0.parquet
  subreddit=IndustrialPharmacy/
    part-0.parquet
  subreddit=LivingWithMBC/
    part-0.parquet
  subreddit=breastcancer/
    part-0.parquet
  subreddit=cancer/
    part-0.parquet
  subreddit=ovariancancer_new/
    part-0.parquet
posts/
  dt=2025-08-29/
    subreddit=BRCA/
      part-0.parquet
    subreddit=IndustrialPharmacy/
      part-0.parquet
    subreddit=LivingWithMBC/
      part-0.parquet
    subreddit=Philippines/
      part-0.parquet
    subreddit=breastcancer/
      part-0.parquet
    subreddit=cancer/
      part-0.parquet
    subreddit=ovariancancer_new/
      part-0.parquet
  dt=2025-09-01/
    subreddit=BRCA/
      part-0.parquet
    subreddit=IndustrialPharmacy/

```

```

        part-0.parquet
    subreddit=LivingWithMBC/
        part-0.parquet
    subreddit=breastcancer/
        part-0.parquet
    subreddit=cancer/
        part-0.parquet
    subreddit=ovariancancer_new/
        part-0.parquet
subreddits/
    dt=2025-08-29/
        part-0.parquet
    dt=2025-09-01/
        part-0.parquet
tmp/
    redshift/
        comment_enriched/
            dt=2025-08-29/
                part-ace70697dd5b42a18c3c6ff7c6830ff4.parquet
        post_enriched/
            dt=2025-08-29/
                part-b9afbb831c634aa5af610dea98c05047.parquet
        stg_tag_rows/
            dt=2025-08-29/
                part-7cd9cd4e771544809f6d8bd37f14962a.parquet

```