


```
In [7]: import os
import json
import time
from collections import namedtuple
import heapq
import uuid
import pandas as pd
import s3fs
import pyarrow.parquet as pq

endpoint_url='https://storage.budsc.midwest-datascience.com'
# s3 = s3fs.S3FileSystem(
#     anon=True,
#     client_kwargs={
#         'endpoint_url': endpoint_url
#     }
# )

acceleration_columns = [
    'offset',
    'id',
    'ride_id',
    'uuid',
    'x',
    'y',
    'z',
    # 't'
]
Acceleration = namedtuple('Acceleration', acceleration_columns)
def read_accelerations():
    df = pq.ParquetDataset(
        'bdd/accelerations'
    ).read_pandas().to_pandas()

    df = df[acceleration_columns].sort_values(by=['offset'])

    records = [Acceleration(*record) for record in df.to_records(index=False)]

    return records
accelerations = read_accelerations()

location_columns = [
    'offset',
    'id',
    'ride_id',
    'uuid',
    'course',
    'latitude',
    'longitude',
    'geohash',
    'speed',
    'accuracy',
    # 't'
]
Location = namedtuple('Location', location_columns)
def read_locations():
```

```

df = pq.ParquetDataset(
    'bdd/locations').read_pandas().to_pandas()

df = df[location_columns].sort_values(by=['offset'])

records = [Location(*record) for record in df.to_records(index=False)]

return records

locations = read_locations()

```

```

In [3]: import os
import json
from pathlib import Path
data_dir = Path('bdd/accelerations/t=000.0')
full_df = pd.concat(
    pd.read_parquet(parquet_file)
    for parquet_file in data_dir.glob('*.parquet')
)
#full_df.to_csv('csv_file.csv')

```

```

In [5]: accelerations

```

```

4865a7995', x=-1.167, y=0.043000000000000003, z=-0.156),
Acceleration(offset=26.696912529556645, id='85c61911b7fe2ced1000c33c9e93270
6', ride_id='6760ffa3f41908695d1405b776c3e8d5', uuid='dad7eae44e784b549c8c5a3
aa051a8c7', x=-1.086, y=0.002, z=-0.01),
Acceleration(offset=26.70506088652284, id='58682c5d48cad9d9e103431d773615b
f', ride_id='c9a2b46c9aa515b632eddc45c4868482', uuid='19b9aa10588646b3bf22c9b
4865a7995', x=-0.761, y=0.14300000000000002, z=-0.027),
Acceleration(offset=26.716912529556648, id='85c61911b7fe2ced1000c33c9e93270
6', ride_id='6760ffa3f41908695d1405b776c3e8d5', uuid='dad7eae44e784b549c8c5a3
aa051a8c7', x=-1.0050000000000001, y=0.033, z=0.013000000000000001),
Acceleration(offset=26.72506088652284, id='58682c5d48cad9d9e103431d773615b
f', ride_id='c9a2b46c9aa515b632eddc45c4868482', uuid='19b9aa10588646b3bf22c9b
4865a7995', x=-0.9530000000000001, y=0.042, z=-0.151),
Acceleration(offset=26.736912529556648, id='85c61911b7fe2ced1000c33c9e93270
6', ride_id='6760ffa3f41908695d1405b776c3e8d5', uuid='dad7eae44e784b549c8c5a3
aa051a8c7', x=-1.045, y=-0.021, z=0.0),
Acceleration(offset=26.74506088652284, id='58682c5d48cad9d9e103431d773615b
f', ride_id='c9a2b46c9aa515b632eddc45c4868482', uuid='19b9aa10588646b3bf22c9b
4865a7995', x=-1.077, y=0.085, z=-0.017),
Acceleration(offset=26.755912529556646, id='85c61911b7fe2ced1000c33c9e93270

```