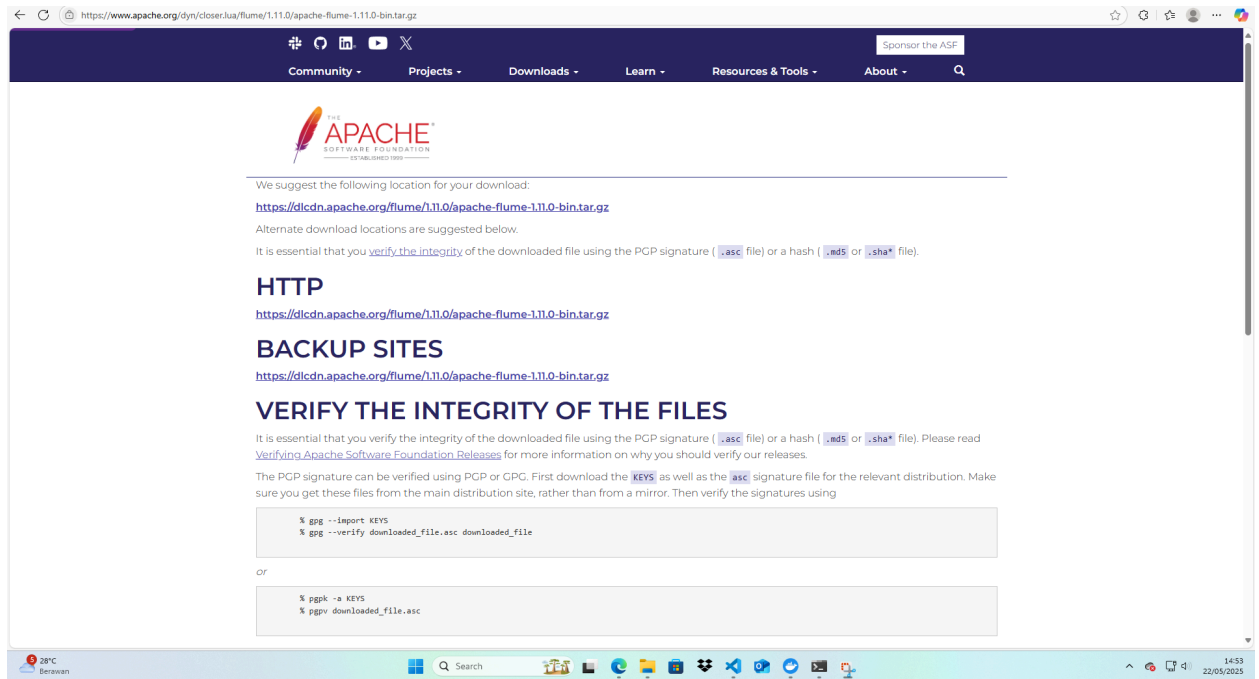


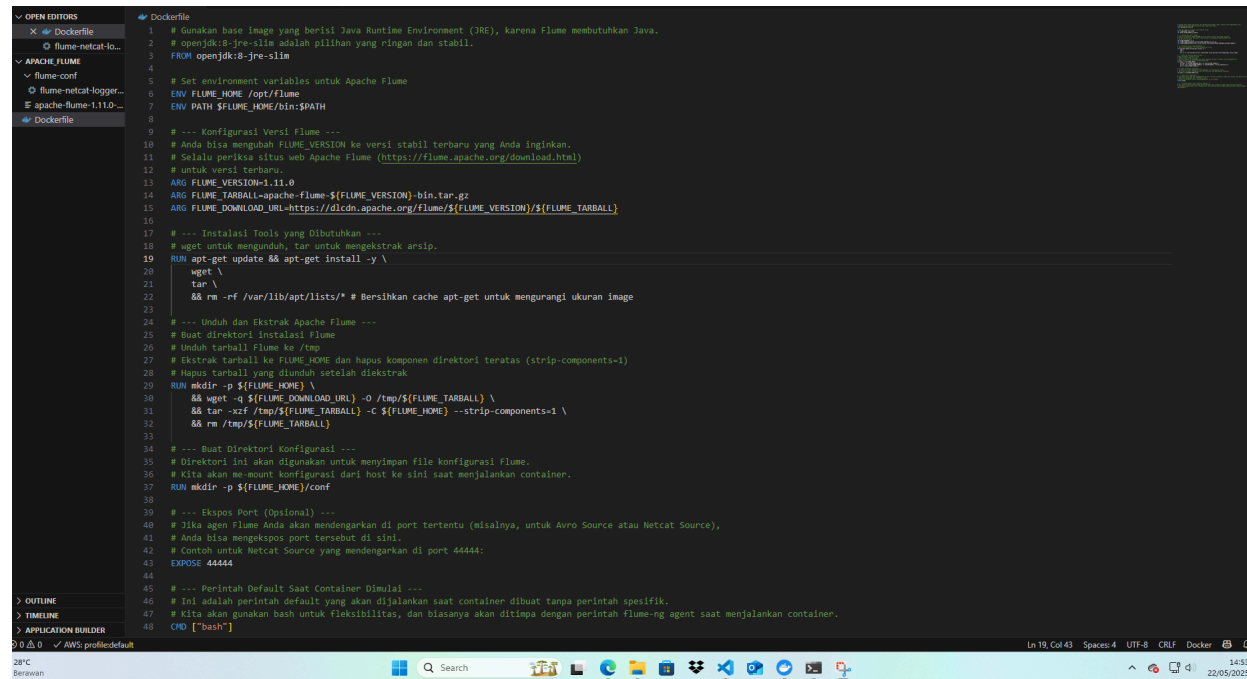
Instalasi dan konfigurasi apache flume

1. Download apache flume.tar

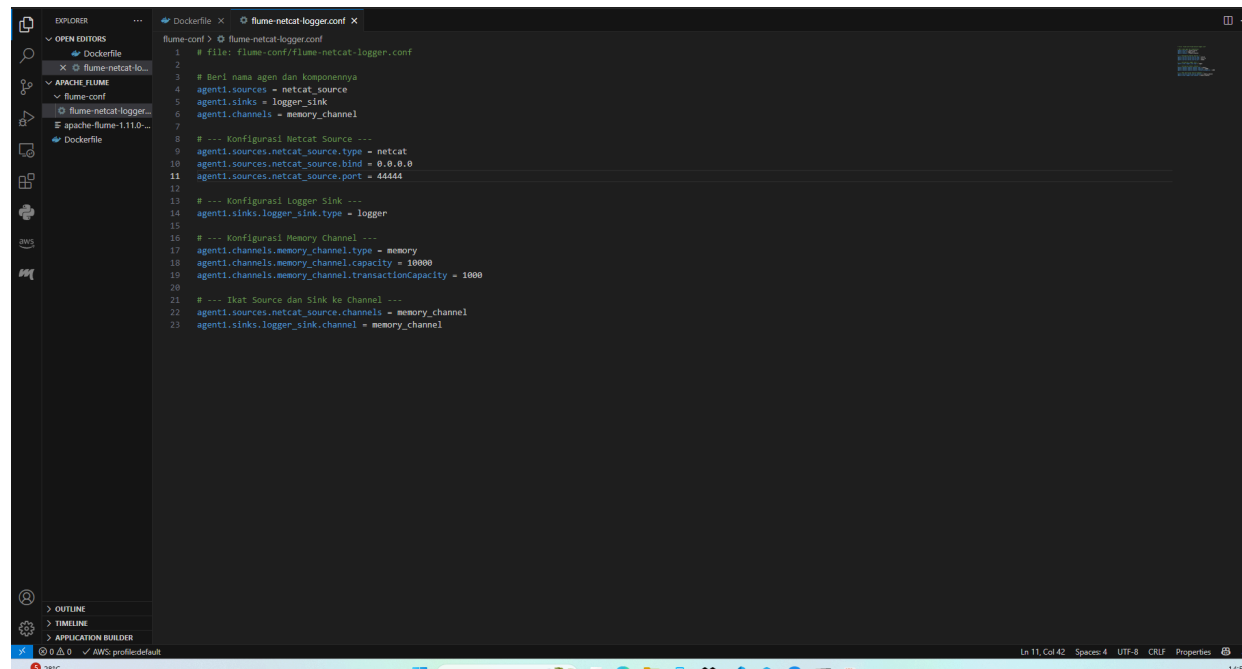


2. Buat file konfigurasi untuk apache flume untuk dijalankan di docker

a. Buat file Dockerfile

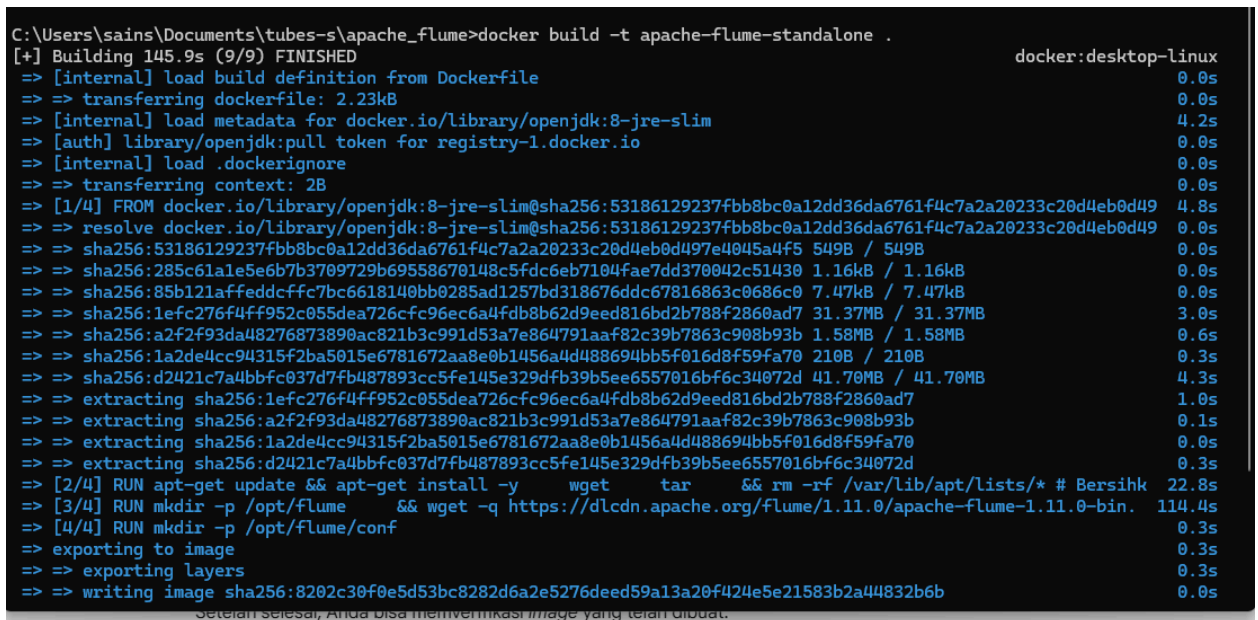


b. Buat file konfigurasi flume



```
1 # flume: flume-conf/flume-netcat-logger.conf
2
3 # Beri nama agen dan komponennya
4 agent1.sources = netcat_source
5 agent1.sinks = logger_sink
6 agent1.channels = memory_channel
7
8 # --- Konfigurasi Netcat Source ---
9 agent1.sources.netcat_source.type = netcat
10 agent1.sources.netcat_source.bind = 0.0.0.0
11 agent1.sources.netcat_source.port = 44444
12
13 # --- Konfigurasi logger Sink ---
14 agent1.sinks.logger_sink.type = logger
15
16 # --- Konfigurasi Memory Channel ---
17 agent1.channels.memory_channel.type = memory
18 agent1.channels.memory_channel.capacity = 10000
19 agent1.channels.memory_channel.transactionCapacity = 1000
20
21 # --- Ikat Source dan Sink ke Channel ---
22 agent1.sources.netcat_source.channels = memory_channel
23 agent1.sinks.logger_sink.channel = memory_channel
```

3. Build docker image untuk apache flume



```
C:\Users\sains\Documents\tubes-s\apache_flume>docker build -t apache-flume-standalone .
[+] Building 145.9s (9/9) FINISHED
=> [internal] load build definition from Dockerfile
=> => transferring dockerfile: 2.23kB
=> [internal] load metadata for docker.io/library/openjdk:8-jre-slim
=> [auth] library/openjdk:pull token for registry-1.docker.io
=> [internal] load .dockerignore
=> => transferring context: 2B
=> [1/4] FROM docker.io/library/openjdk:8-jre-slim@sha256:53186129237fbb8bc0a12dd36da6761f4c7a2a20233c20d4eb0d49
=> => resolve docker.io/library/openjdk:8-jre-slim@sha256:53186129237fbb8bc0a12dd36da6761f4c7a2a20233c20d4eb0d49
=> => sha256:53186129237fbb8bc0a12dd36da6761f4c7a2a20233c20d4eb0d497e4045a4f5 549B / 549B
=> => sha256:285c61a1e5e6b7b3709729b69558670148c5fdc6eb7104fae7dd370042c51430 1.16kB / 1.16kB
=> => sha256:85b121affeddcffc7bc6618140bb0285ad1257bd318676ddc67816863c0686c0 7.47kB / 7.47kB
=> => sha256:1efc276f4ff952c055dea726cfc96ec6a4fdb8b62d9eed816bd2b788f2860ad7 31.37MB / 31.37MB
=> => sha256:a2f2f93da48276873890ac821b3c991d53a7e864791aaf82c39b7863c908b93b 1.58MB / 1.58MB
=> => sha256:1a2de4cc94315f2ba5015e6781672aa8e0b1456a4d488694bb5f016d8f59fa70 210B / 210B
=> => sha256:d2421c7a4bbfc037d7fb487893cc5fe145e329dfb39b5ee6557016bf6c34072d 41.70MB / 41.70MB
=> => extracting sha256:1efc276f4ff952c055dea726cfc96ec6a4fdb8b62d9eed816bd2b788f2860ad7 1.0s
=> => extracting sha256:a2f2f93da48276873890ac821b3c991d53a7e864791aaf82c39b7863c908b93b 0.1s
=> => extracting sha256:1a2de4cc94315f2ba5015e6781672aa8e0b1456a4d488694bb5f016d8f59fa70 0.0s
=> => extracting sha256:d2421c7a4bbfc037d7fb487893cc5fe145e329dfb39b5ee6557016bf6c34072d 0.3s
=> [2/4] RUN apt-get update && apt-get install -y wget tar && rm -rf /var/lib/apt/lists/* # Bersihk 22.8s
=> [3/4] RUN mkdir -p /opt/flume && wget -q https://dlcdn.apache.org/flume/1.11.0/apache-flume-1.11.0-bin. 114.4s
=> [4/4] RUN mkdir -p /opt/flume/conf
=> exporting to image
=> exporting layers
=> writing image sha256:8202c30f0e5d53bc8282d6a2e5276deed59a13a20f424e5e21583b2a44832b6b 0.0s
```

Setelah selesai, Anda bisa memverifikasi image yang telah dibuat.

4. Jalankan container

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker run -d ^
More? -p 444444:44444 ^
More? --name flume_agent_01 ^
More? -v "C:/Users/sains/Documents/tubes-s/apache_flume/flume-conf/flume-netcat-logger.conf":/opt/flume/conf/flume.conf ^
More? apache-flume-standalone ^
More? /opt/flume/bin/flume-ng agent ^
More? --conf /opt/flume/conf ^
More? --conf-file /opt/flume/conf/flume.conf ^
More? --name agent1 ^
More? -Dflume.root.logger=INFO,console
9eab3fdbf8eb513edfa6671ca9f43fc97439f3cef812b49af86f509d6016d6d6
```

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker ps
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS
```

5. Masukkan data ke folder direktori apache flume

Name	Date modified	Type	Size
data-input	22/05/2025 15:13	File folder	
flume-conf	22/05/2025 15:13	File folder	
apache-flume-1.11.0-bin.tar	22/05/2025 14:49	Compressed Archi...	85.333 KB
Dockerfile	22/05/2025 14:49	File	3 KB
data_gempa	22/05/2025 14:43	Microsoft Excel C...	116 KB

6. Buat file yg berisi konfigurasi untuk memasukkan data csv ke apache flume yang akan diteruskan ke hdfs

```
flume-conf > flume-csv-hdfs.conf
1 # Nama agen dan komponennya
2 a1.sources = spooled_dir_source
3 a1.sinks = hdfs_sink
4 a1.channels = memory_channel
5
6 # --- Konfigurasi Spooling Directory Source ---
7 # Source ini akan memantau direktori /flume-data-input di dalam container
8 a1.sources.spooled_dir_source.type = spooldir
9 a1.sources.spooled_dir_source.spoolDir = /flume-data-input
10 a1.sources.spooled_dir_source.fileHeader = true
11 # Deserializer LINE membaca setiap baris sebagai satu event.
12 a1.sources.spooled_dir_source.deserializer = LINE
13 # Menandakan file yang sudah diproses ke direktori 'done'
14 a1.sources.spooled_dir_source.completedSuffix = .DONE
15 a1.sources.spooled_dir_source.basenameHeader = true # Menambahkan nama file ke header event
16
17 # --- Konfigurasi Memory Channel ---
18 a1.channels.memory_channel.type = memory
19 a1.channels.memory_channel.capacity = 10000
20 a1.channels.memory_channel.transactionCapacity = 1000
21
22 # --- Konfigurasi HDFS Sink ---
23 a1.sinks.hdfs_sink.type = hdfs
24 a1.sinks.hdfs_sink.channel = memory_channel
25 # URL HDFS Anda. Ganti 'hadoop-master' dengan hostname kontainer Hadoop Anda
26 # dan '9000' dengan port NameNode Anda.
27 # Kita akan menggunakan kontainer Hadoop dari contoh sebelumnya, 'bigdata-hadoop' adalah hostname-nya.
28 a1.sinks.hdfs_sink.hdfs.path = hdfs://bigdata-hadoop:9000/user/flume/csv_data/%Y-%m-%d
29 a1.sinks.hdfs_sink.hdfs.filePrefix = csv_data_
30 a1.sinks.hdfs_sink.hdfs.fileSuffix = .csv
31 a1.sinks.hdfs_sink.hdfs.rollInterval = 0 # Tanpa roll berdasarkan waktu
32 a1.sinks.hdfs_sink.hdfs.rollSize = 104857600 # roll file setiap 100M (100 * 1024 * 1024 bytes)
33 a1.sinks.hdfs_sink.hdfs.rollCount = 0 # Tanpa roll berdasarkan jumlah event
34 a1.sinks.hdfs_sink.hdfs.useLocalTimeStamp = true # Gunakan timestamp dari Flume JVM
35 a1.sinks.hdfs_sink.hdfs.filetype = DataStream # Simpan sebagai file teks
36 a1.sinks.hdfs_sink.hdfs.writeFormat = Text # Format penulisan: teks
37 a1.sinks.hdfs_sink.hdfs.batchSize = 1000 # Jumlah event per batch ke HDFS
38 a1.sinks.hdfs_sink.hdfs.maxOpenFiles = 5000 # Batas file terbuka di HDFS
```

7. Stop dan hapus container sebelumnya agar tidak ada konflik
Selanjutnya jalankan lagi container yg baru yg sudah ada file datanya

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker stop flume_agent_01
flume_agent_01

C:\Users\sains\Documents\tubes-s\apache_flume>docker rm flume_agent_01
flume_agent_01

C:\Users\sains\Documents\tubes-s\apache_flume>docker run -d ^
More? --name flume_csv_agent ^
More? -v "C:/Users/sains/Documents/tubes-s/apache_flume/flume-conf/flume-csv-hdfs.conf":/opt/flume/conf/flume.conf ^
More? -v "C:/Users/sains/Documents/tubes-s/apache_flume/data-input":/flume-data-input ^
More? apache-flume-standalone ^
More? /opt/flume/bin/flume-ng agent ^
More? --conf /opt/flume/conf ^
More? --conf-file /opt/flume/conf/Flume.conf ^
More? --name al ^
More? -Dflume.root.logger=INFO,console
c3ce85943a59ba45a408a560c043a721ecbfd5059b8477804953d741c47c5c89

C:\Users\sains\Documents\tubes-s\apache_flume>
```

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker ps

CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS
c3ce85943a59   apache-flume-standalone             "/opt/flume/bin/flum..." About a minute ago Up About a minute 44444/tcp

flume_csv_agent
e519c20be8f5   bigdata-hadoop                      "/bin/bash -c '/boot..." About an hour ago Up About an hour 0.0.0.0:3306->3306/tcp, 0.0.0.0:8000->8000/tcp, 0.0.0.0:8088->8088/tcp, 0.0.0.0:9000->9000/tcp, 0.0.0.0:9866->9866/tcp, 0.0.0.0:9870->9870/tcp, 0.0.0.0:10000->10000/tcp, 0.0.0.0:10002->10002/tcp, 0.0.0.0:50030->50030/tcp, 0.0.0.0:32768->1004/tcp, 0.0.0.0:32769->1006/tcp, 0.0.0.0:32770->1180/tcp, 0.0.0.0:32771->8020/tcp, 0.0.0.0:32772->8030/tcp, 0.0.0.0:32773->8031/tcp, 0.0.0.0:32774->8032/tcp, 0.0.0.0:32775->8033/tcp, 0.0.0.0:32776->8040/tcp, 0.0.0.0:32777->8041/tcp, 0.0.0.0:32778->8042/tcp, 0.0.0.0:32779->9864/tcp, 0.0.0.0:32780->9867/tcp, 0.0.0.0:32781->10020/tcp, 0.0.0.0:32782->19888/tcp, 0.0.0.0:32783->50070/tcp, 0.0.0.0:32784->50470/tcp bigdata-hadoop

C:\Users\sains\Documents\tubes-s\apache_flume>
```

8. Menjalankan perintah untuk meneruskan data dari flume ke hdfs

9. Melakukan verifikasi apakah file sudah masuk ke hdfs atau belum

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker exec -it bigdata-hadoop /hadoop/bin/hdfs dfs -ls /gempa/input/
Found 1 items
-rw-r--r-- 1 root supergroup 118425 2025-05-22 08:39 /gempa/input/data_gempa.csv

C:\Users\sains\Documents\tubes-s\apache_flume>docker exec -it bigdata-hadoop /hadoop/bin/hdfs dfs -ls /gempa/input/
Found 1 items
-rw-r--r-- 1 root supergroup 118425 2025-05-22 08:39 /gempa/input/data_gempa.csv

C:\Users\sains\Documents\tubes-s\apache_flume>
```

10. Datanya sudah berhasil masuk ke hdfs dengan format CSV

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker exec -it bigdata-hadoop /hadoop/bin/hdfs dfs -cat /gempa/input/data_gempa.csv
Event ID,Date time,Latitude,Longitude,Magnitude,Mag Type,Depth (km),Phase Count,Azimuth Gap,Location,Agency
bmkg2024ccsf,2024-01-30T22:32:20.889489Z,1.491829395,99.6288899,2.887956762,MLv,165,35,118.1091671,"Northern Sumatra, Indonesia",BMKG
bmkg2024cctb,2024-01-30T09:49:36.13532,4.366219997,96.84862958,3.96293717,MLv,5,58,46.14782476,"Northern Sumatra, Indonesia",BMKG
bmkg2024cbbw,2024-01-30T01:38:49.267026Z,-3.534729988,102.44801474,3.582914497,MLv,10,38,77.04136963,"Southern Sumatra, Indonesia",BMKG
bmkg2024cbab,2024-01-30T08:15:42.92897Z,0.283133921,97.01425317,4.123478801,MLv,26,41,202.6468864,"Northern Sumatra, Indonesia",BMKG
bmkg2024calx,2024-01-29T16:34:19.290276Z,0.4767362177,98.34382629,3.226258213,MLv,4,32,93.5999729,"Northern Sumatra, Indonesia",BMKG
bmkg2024bzgl,2024-01-29T01:11:58.854074Z,-6.327270031,103.6217957,3.870640216,MLv,13,43,221.3026276,"Southwest of Sumatra, Indonesia",BMKG
bmkg2024byus,2024-01-28T19:15:42.127995Z,-0.842182636,102.8365173,2.873935574,MLv,31,36,210.5345993,"Southern Sumatra, Indonesia",BMKG
bmkg2024byra,2024-01-28T17:23:08.673881Z,4.514056683,96.4898624,2.328394463,MLv,10,18,84.06783529,"Northern Sumatra, Indonesia",BMKG
bmkg2024byvm,2024-01-28T15:18:16.29369Z,-6.278318882,103.6838521,3.369606575,MLv,13,30,226.5658864,"Southwest of Sumatra, Indonesia",BMKG
bmkg2024bydp,2024-01-28T10:37:25.988456Z,-2.566192627,102.2488632,3.891933208,MLv,2,37,55.69261265,"Southern Sumatra, Indonesia",BMKG
bmkg2024bxzd,2024-01-28T07:21:19.982621Z,4.873150163,95.99866486,3.836825487,MLv,11,11,130.5513458,"Northern Sumatra, Indonesia",BMKG
bmkg2024bxnd,2024-01-28T06:51:07.583526Z,1.988808831,98.36688995,2.761889372,MLv,55,41,73.85723572,"Northern Sumatra, Indonesia",BMKG
bmkg2024bxne,2024-01-28T01:48:16.582463Z,-0.1162820192,100.3288847,2.431973444,MLv,10,17,122.0183394,"Southern Sumatra, Indonesia",BMKG
bmkg2024bxwx,2024-01-27T19:36:48.270761Z,-8.626338959,105.978981,3.231705134,MLv,10,37,270.0153503,"South of Java, Indonesia",BMKG
bmkg2024xbwb,2024-01-27T17:40:42.077041Z,-3.839478731,99.71772766,3.841358116,MLv,12,58,239.6386948,"Southwest of Sumatra, Indonesia",BMKG
bmkg2024xbwl,2024-01-27T16:50:58.016875Z,-0.0975858485,100.5130345,4.616822398,MLv,3,84,32.38021851,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwy,2024-01-27T16:35:40.983329Z,0.2177918656,97.0996317,4.067109784,MLv,26,78,201.4032326,"Northern Sumatra, Indonesia",BMKG
bmkg2024xbwh,2024-01-27T08:45:43.411692Z,-0.5941882133,100.5099182,3.221390735,MLv,8,41,38.47946167,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwi,2024-01-27T08:42:28.37332Z,-0.607478559,100.4877472,2.792772331,MLv,10,37,61.46913147,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwh,2024-01-27T08:41:17.684993Z,-0.607855277,100.5148849,2.383285139,MLv,10,21,94.48898315,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwd,2024-01-27T08:15:03.51365Z,-0.6866630859,100.5018997,3.884417177,MLv,8,39,56.71483375,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwdh,2024-01-27T08:11:32.565295Z,-0.6197630134,100.5136640,2.289294425,MLv,10,24,96.42337836,"Southern Sumatra, Indonesia",BMKG
bmkg2024xbwcd,2024-01-27T07:36:55.299719Z,0.1292646825,100.0288914,3.28468976,MLv,10,36,104.9448853,"Northern Sumatra, Indonesia",BMKG
bmkg2024bvnw,2024-01-27T08:12:14.618952Z,0.2605628967,96.77553558,3.597632847,MLv,31,34,210.9622498,Off West Coast of Northern Sumatra, BMKG
bmkg2024buzr,2024-01-26T17:22:58.245514Z,-4.668157681,103.6215134,2.986724382,MLv,99,29,79.62498845,"Southern Sumatra, Indonesia",BMKG
bmkg2024buzl,2024-01-26T17:06:01.428478Z,-3.294524097,101.5526607,2.11163185,MLv,31,23,158.6545269,"Southern Sumatra, Indonesia",BMKG
bmkg2024buzp,2024-01-26T16:51:45.778052Z,2.588378532,96.62932587,2.763146247,MLv,5,32,97.85159099,"Northern Sumatra, Indonesia",BMKG
bmkg2024buhk,2024-01-26T07:59:23.83742Z,-1.582182537,99.37985992,3.09283376,MLv,10,31,88.82979886,"Southern Sumatra, Indonesia",BMKG
bmkg2024bubs,2024-01-26T05:06:51.070213Z,-5.875364384,105.932845,2.80160752,MLv,190,32,174.0245285,"Sunda Strait, Indonesia",BMKG
bmkg2024bstr,2024-01-25T23:51:12.708983Z,3.043182564,95.8158493,3.877874051,MLv,16,32,165.5041687,Off West Coast of Northern Sumatra, BMKG
bmkg2024bstw,2024-01-25T15:02:53.528479Z,-6.859526634,105.2411346,3.696589077,MLv,26,43,211.1286697,"Sunda Strait, Indonesia",BMKG
bmkg2024bsdl,2024-01-25T03:43:04.662211Z,3.241672993,96.66138829,3.811488809,MLv,10,20,88.262146,"Northern Sumatra, Indonesia",BMKG
bmkg2024bsdg,2024-01-24T19:02:46.949457Z,-3.698475838,101.5774612,3.389378225,MLv,20,33,182.3262863,"Southern Sumatra, Indonesia",BMKG
bmkg2024bqpx,2024-01-24T07:44:35.024851Z,2.668890156,95.72974396,3.734058251,MLv,15,25,227.5066986,Off West Coast of Northern Sumatra, BMKG
bmkg2024bqpl,2024-01-24T07:31:22.655583Z,-0.2118413462,100.3476944,2.66220229,MLv,10,15,113.874115,"Southern Sumatra, Indonesia",BMKG
bmkg2024bqbl,2024-01-24T03:59:02.52656Z,-2.982845545,100.6426163,3.473447218,MLv,14,23,154.586792,"Southern Sumatra, Indonesia",BMKG
bmkg2024bqbd,2024-01-24T01:15:15.928113Z,4.856840955,95.98483276,2.583009643,MLv,10,20,118.5791168,"Northern Sumatra, Indonesia",BMKG
```