

```
In [2]: import pandas as pd
import seaborn as sns
import numpy as np
import datetime as dt
import matplotlib.pyplot as plt
import re
```

```
In [8]: df = pd.read_csv("C:/Users/PCC/Downloads/HealthApp_2k.log_structured.csv")
```

```
In [9]: templates = pd.read_csv("C:/Users/PCC/Downloads/HealthApp_2k.log_templates.csv")
```

```
In [12]: df
```

	Lineld	Time	Component	Pid	Content	EventId	EventTemplate
	0	1 20171223-22:15:29:606	Step_LSC	30002312	onStandStepChanged 3579	E42	onStandStepChanged <*>
	1	2 20171223-22:15:29:615	Step_LSC	30002312	onExtend:1514038530000 14 0 4	E39	onExtend:<*> <*> <*> <*>
	2	3 20171223-22:15:29:633	Step_StandReportReceiver	30002312	onReceive action: android.intent.action.SCREEN_ON	E41	onReceive action: android.intent.action.SCREEN_ON
	3	4 20171223-22:15:29:635	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E43	processHandleBroadcastAction action:android.in...
	4	5 20171223-22:15:29:635	Step_StandStepCounter	30002312	flush sensor data	E12	flush sensor data
	...	...	...	...	...	...	...
	1995	1996 20171224-0:58:53:985	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
	1996	1997 20171224-0:59:7:581	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
	1997	1998 20171224-1:0:0:794	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
	1998	1999 20171224-1:1:0:935	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
	1999	2000 20171224-1:2:35:789	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...

2000 rows × 7 columns

```
In [13]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 7 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   LineId           2000 non-null  int64  
1   Time             2000 non-null  object  
2   Component        2000 non-null  object  
3   Pid              2000 non-null  int64  
4   Content          2000 non-null  object  
5   EventId          2000 non-null  object  
6   EventTemplate    2000 non-null  object  
dtypes: int64(2), object(5)
memory usage: 109.5+ KB
```

```
In [58]: # Check for missing values and decide how to handle them. You can either remove rows with missing values or impute
# them with appropriate values (mean, median, mode, etc.). To check for missing values:

df.isnull()
```

	Lineld	Time	Component	Content	EventId	EventTemplate
	0	False	False	False	False	False
	1	False	False	False	False	False
	2	False	False	False	False	False
	3	False	False	False	False	False
	4	False	False	False	False	False
	...	...	...	...	...	...
	1995	False	False	False	False	False
	1996	False	False	False	False	False
	1997	False	False	False	False	False
	1998	False	False	False	False	False
	1999	False	False	False	False	False

In [54]: df.notnull()

Out[54]:

	Lineld	Time	Component	Content	EventId	EventTemplate
0	True	True	True	True	True	True
1	True	True	True	True	True	True
2	True	True	True	True	True	True
3	True	True	True	True	True	True
4	True	True	True	True	True	True
...	...	...	...	...	...	...
1995	True	True	True	True	True	True
1996	True	True	True	True	True	True
1997	True	True	True	True	True	True
1998	True	True	True	True	True	True
1999	True	True	True	True	True	True

2000 rows × 6 columns

In [56]: df.shape

Out[56]: (2000, 6)

In [15]: df.head(5)

Out[15]:

	Lineld	Time	Component	Pid	Content	EventId	EventTemplate
0	1	20171223-22:15:29:606	Step_LSC	30002312	onStandStepChanged 3579	E42	onStandStepChanged <*>
1	2	20171223-22:15:29:615	Step_LSC	30002312	onExtend:1514038530000 14 0 4	E39	onExtend:<*> <*> <*> <*>
2	3	20171223-22:15:29:633	Step_StandReportReceiver	30002312	onReceive action: android.intent.action.SCREEN_ON	E41	onReceive action: android.intent.action.SCREEN_ON
3	4	20171223-22:15:29:635	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E43	processHandleBroadcastAction action:android.in...
4	5	20171223-22:15:29:635	Step_StandStepCounter	30002312	flush sensor data	E12	flush sensor data

In [57]: df.tail

Out[57]:

	Lineld	Time	Component	Content	EventId	EventTemplate
1995	1996	20171224-0:58:53:985	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1996	1997	20171224-0:59:7:581	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1997	1998	20171224-1:0:0:794	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1998	1999	20171224-1:1:0:935	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1999	2000	20171224-1:2:35:789	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...

In [17]: df.describe()

Out[17]:

	Lineld	Pid
count	2000.000000	2000.0
mean	1000.500000	30002312.0
std	577.494589	0.0
min	1.000000	30002312.0
25%	500.750000	30002312.0
50%	1000.500000	30002312.0
75%	1500.250000	30002312.0
max	2000.000000	30002312.0

In [25]: df.count()

Out[25]:

LineId	2000
Time	2000
Component	2000
Pid	2000
Content	2000
EventId	2000
EventTemplate	2000
dtype: int64	

In [32]: df.min()

Out[32]: LineId 1  
Time 20171223-22:15:29:606  
Component HiH\_  
Pid 30002312  
Content getDiffTotalSteps= 1513958400215##0  
EventId E1  
EventTemplate Alarm uploadStaticsToDB totalSteps=<\*>:<\*>:<\*>...  
dtype: object

In [34]: df.head(5)

	Lineld	Time	Component	Pid	Content	EventId	EventTemplate
0	1	20171223-22:15:29:606	Step_LSC	30002312	onStandStepChanged 3579	E42	onStandStepChanged <*>
1	2	20171223-22:15:29:615	Step_LSC	30002312	onExtend:1514038530000 14 0 4	E39	onExtend:<*> <*> <*> <*>
2	3	20171223-22:15:29:633	Step_StandReportReceiver	30002312	onReceive action: android.intent.action.SCREEN_ON	E41	onReceive action: android.intent.action.SCREEN_ON
3	4	20171223-22:15:29:635	Step_LSC	30002312	processHandleBroadcastAction action:android.in...	E43	processHandleBroadcastAction action:android.in...
4	5	20171223-22:15:29:635	Step_StandStepCounter	30002312	flush sensor data	E12	flush sensor data

In [36]: df = df.drop('Pid', axis=1)

In [37]: df

	Lineld	Time	Component	Content	EventId	EventTemplate
0	1	20171223-22:15:29:606	Step_LSC	onStandStepChanged 3579	E42	onStandStepChanged <*>
1	2	20171223-22:15:29:615	Step_LSC	onExtend:1514038530000 14 0 4	E39	onExtend:<*> <*> <*> <*>
2	3	20171223-22:15:29:633	Step_StandReportReceiver	onReceive action: android.intent.action.SCREEN_ON	E41	onReceive action: android.intent.action.SCREEN_ON
3	4	20171223-22:15:29:635	Step_LSC	processHandleBroadcastAction action:android.in...	E43	processHandleBroadcastAction action:android.in...
4	5	20171223-22:15:29:635	Step_StandStepCounter	flush sensor data	E12	flush sensor data
...	...	...	...	...	...	...
1995	1996	20171224-0:58:53:985	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1996	1997	20171224-0:59:7:581	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...

In [39]: dfs23 = df[df['EventId'] == 'E22']  
dfs23['Time'] = pd.to\_datetime(dfs23['Time'], format='%Y%m%d-%H:%M:%S:%f')  
dfs23['Steps'] = dfs23['Content'].str.extract(r'##(\d+)##')  
dfs23['Steps'] = dfs23['Steps'].astype(int)  
plt.figure(figsize=(12, 6))  
plt.plot(dfs23['Time'], dfs23['Steps'], marker='o', linestyle='--')  
plt.title('Time vs. Steps for Event E22')  
plt.xlabel('Time')  
plt.ylabel('Steps')  
plt.grid(True)  
plt.xticks(rotation=45)  
plt.tight\_layout()  
plt.show()

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\2267965645.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead  
  
See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Time'] = pd.to\_datetime(dfs23['Time'], format='%Y%m%d-%H:%M:%S:%f')  
C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\2267965645.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead  
  
See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Steps'] = dfs23['Content'].str.extract(r'##(\d+)##')  
C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\2267965645.py:4: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead  
  
See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Steps'] = dfs23['Steps'].astype(int)

Time vs. Steps for Event E22

Time	Steps
23 22:15	7000
23 22:15.5	7100
23 22:16	7200
23 22:16.5	7200
23 22:17	7200
23 22:17.5	7200
23 22:18	7200
23 22:18.5	7200
23 22:19	7200
23 22:19.5	7200
23 22:20	7200
23 22:20.5	7200
23 22:21	7200
23 22:21.5	7200
23 22:22	7200
23 22:22.5	7200
23 22:23	7200
23 22:23.5	7200
23 22:24	7200
23 22:24.5	7200
23 22:25	7200
23 22:25.5	7200
23 22:26	7200
23 22:26.5	7200
23 22:27	7200
23 22:27.5	7200
23 22:28	7200
23 22:28.5	7200
23 22:29	7200
23 22:29.5	7200
23 22:30	7200
23 22:30.5	7200
23 22:31	7200
23 22:31.5	7200
23 22:32	7200
23 22:32.5	7200
23 22:33	7200
23 22:33.5	7200
23 22:34	7200
23 22:34.5	7200
23 22:35	7200
23 22:35.5	7200
23 22:36	7200
23 22:36.5	7200
23 22:37	7200
23 22:37.5	7200
23 22:38	7200
23 22:38.5	7200
23 22:39	7200
23 22:39.5	7200
23 22:40	7200
23 22:40.5	7200
23 22:41	7200
23 22:41.5	7200
23 22:42	7200
23 22:42.5	7200
23 22:43	7200
23 22:43.5	7200
23 22:44	7200
23 22:44.5	7200
23 22:45	7200
23 22:45.5	7200
23 22:46	7200
23 22:46.5	7200
23 22:47	7200
23 22:47.5	7200
23 22:48	7200
23 22:48.5	7200
23 22:49	7200
23 22:49.5	7200
23 22:50	7200
23 22:50.5	7200
23 22:51	7200
23 22:51.5	7200
23 22:52	7200
23 22:52.5	7200
23 22:53	7200
23 22:53.5	7200
23 22:54	7200
23 22:54.5	7200
23 22:55	7200
23 22:55.5	7200
23 22:56	7200
23 22:56.5	7200
23 22:57	7200
23 22:57.5	7200
23 22:58	7200
23 22:58.5	7200
23 22:59	7200
23 22:59.5	7200
23 23:00	7200
23 23:00.5	7200
23 23:01	7200
23 23:01.5	7200
23 23:02	7200
23 23:02.5	7200
23 23:03	7200
23 23:03.5	7200
23 23:04	7200
23 23:04.5	7200
23 23:05	7200
23 23:05.5	7200
23 23:06	7200
23 23:06.5	7200
23 23:07	7200
23 23:07.5	7200
23 23:08	7200
23 23:08.5	7200
23 23:09	7200
23 23:09.5	7200
23 23:10	7200
23 23:10.5	7200
23 23:11	7200
23 23:11.5	7200
23 23:12	7200
23 23:12.5	7200
23 23:13	7200
23 23:13.5	7200
23 23:14	7200
23 23:14.5	7200
23 23:15	7200
23 23:15.5	7200
23 23:16	7200
23 23:16.5	7200
23 23:17	7200
23 23:17.5	7200
23 23:18	7200
23 23:18.5	7200
23 23:19	7200
23 23:19.5	7200
23 23:20	7200
23 23:20.5	7200
23 23:21	7200
23 23:21.5	7200
23 23:22	7200
23 23:22.5	7200
23 23:23	7200
23 23:23.5	7200
23 23:24	7200
23 23:24.5	7200
23 23:25	7200
23 23:25.5	7200
23 23:26	7200
23 23:26.5	7200
23 23:27	7200
23 23:27.5	7200
23 23:28	7200
23 23:28.5	7200
23 23:29	7200
23 23:29.5	7200
23 23:30	7200
23 23:30.5	7200
23 23:31	7200
23 23:31.5	7200
23 23:32	7200
23 23:32.5	7200
23 23:33	7200
23 23:33.5	7200
23 23:34	7200
23 23:34.5	7200
23 23:35	7200
23 23:35.5	7200
23 23:36	7200
23 23:36.5	7200
23 23:37	7200
23 23:37.5	7200
23 23:38	7200
23 23:38.5	7200
23 23:39	7200
23 23:39.5	7200
23 23:40	7200
23 23:40.5	7200
23 23:41	7200
23 23:41.5	7200
23 23:42	7200
23 23:42.5	7200
23 23:43	7200
23 23:43.5	7200
23 23:44	7200
23 23:44.5	7200
23 23:45	7200
23 23:45.5	7200
23 23:46	7200
23 23:46.5	7200
23 23:47	7200
23 23:47.5	7200
23 23:48	7200
23 23:48.5	7200
23 23:49	7200
23 23:49.5	7200
23 23:50	7200
23 23:50.5	7200
23 23:51	7200
23 23:51.5	7200
23 23:52	7200
23 23:52.5	7200
23 23:53	7200
23 23:53.5	7200
23 23:54	7200
23 23:54.5	7200
23 23:55	7200
23 23:55.5	7200
23 23:56	7200
23 23:56.5	7200
23 23:57	7200
23 23:57.5	7200
23 23:58	7200
23 23:58.5	7200
23 23:59	7200
23 23:59.5	7200
24 00:00	0
24 00:00.5	0
24 00:01	0
24 00:01.5	0
24 00:02	0
24 00:02.5	0
24 00:03	0
24 00:03.5	0
24 00:04	0
24 00:04.5	0
24 00:05	0
24 00:05.5	0
24 00:06	0
24 00:06.5	0
24 00:07	0
24 00:07.5	0
24 00:08	0
24 00:08.5	0
24 00:09	0
24 00:09.5	0
24 00:10	0
24 00:10.5	0
24 00:11	0
24 00:11.5	0
24 00:12	0
24 00:12.5	0
24 00:13	0
24 00:13.5	0
24 00:14	0
24 00:14.5	0
24 00:15	0
24 00:15.5	0
24 00:16	0
24 00:16.5	0
24 00:17	0
24 00:17.5	0
24 00:18	0
24 00:18.5	0
24 00:19	0
24 00:19.5	0
24 00:20	0
24 00:20.5	0
24 00:21	0
24 00:21.5	0
24 00:22	0
24 00:22.5	0
24 00:23	0
24 00:23.5	0
24 00:24	0
24 00:24.5	0
24 00:25	0
24 00:25.5	0
24 00:26	0
24 00:26.5	0
24 00:27	0
24 00:27.5	0
24 00:28	0
24 00:28.5	0
24 00:29	0
24 00:29.5	0
24 00:30	0

```
In [41]: dfs23['Time'] = pd.to_datetime(dfs23['Time'], format='%Y%m%d-%H:%M:%S.%f')

dfs23['Date'] = dfs23['Time'].dt.date

print(dfs23[['Date', 'Steps']])

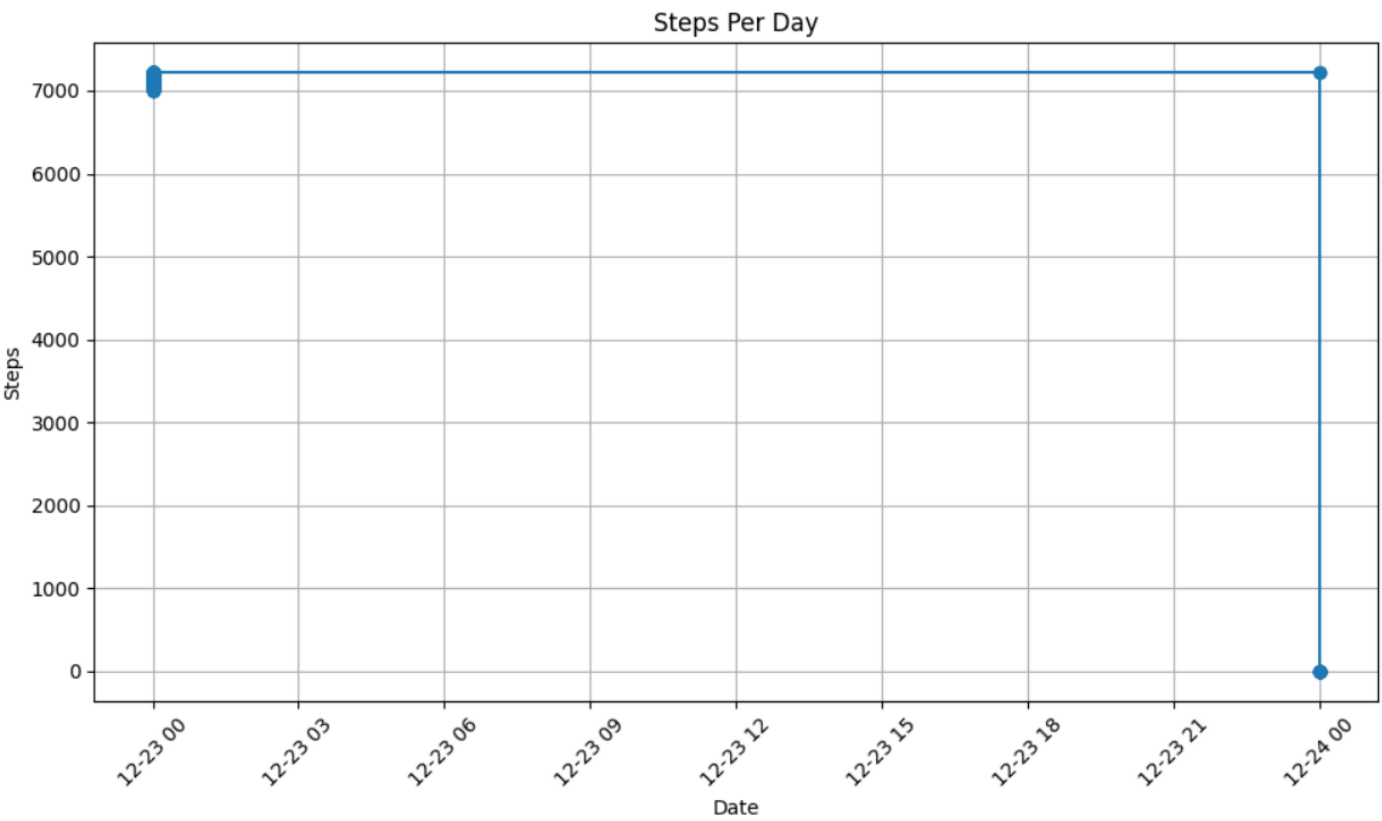
plt.figure(figsize=(10, 6))
plt.plot(dfs23['Date'], dfs23['Steps'], marker='o')
plt.title('Steps Per Day')
plt.xlabel('Date')
plt.ylabel('Steps')
plt.xticks(rotation=45)
plt.grid(True)
plt.tight_layout()
plt.show()
```

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\2098986252.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead  
  
See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Time'] = pd.to\_datetime(dfs23['Time'], format='%Y%m%d-%H:%M:%S.%f')  
C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\2098986252.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead  
  
See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Date'] = dfs23['Time'].dt.date

	Date	Steps
5	2017-12-23	6993
15	2017-12-23	7007
22	2017-12-23	7008
29	2017-12-23	7009
36	2017-12-23	7011
...	...	...
1925	2017-12-24	0
1935	2017-12-24	0
1951	2017-12-24	0
1957	2017-12-24	0
1966	2017-12-24	0

[242 rows x 2 columns]

[242 rows x 2 columns]



dfs23

	Lineld	Time	Component	Content	EventId	EventTemplate	Steps	Date
5	6	2017-12-23 22:15:29.635	Step_SPUtills	getTodayTotalDetailSteps = 1514038440000##699...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	6993	2017-12-23
15	16	2017-12-23 22:15:29.950	Step_SPUtills	getTodayTotalDetailSteps = 1514038440000##700...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	7007	2017-12-23
22	23	2017-12-23 22:15:30.632	Step_SPUtills	getTodayTotalDetailSteps = 1514038440000##700...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	7008	2017-12-23
29	30	2017-12-23 22:15:31.142	Step_SPUtills	getTodayTotalDetailSteps = 1514038440000##700...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	7009	2017-12-23
36	37	2017-12-23 22:15:32.145	Step_SPUtills	getTodayTotalDetailSteps = 1514038440000##701...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	7011	2017-12-23
...	...	...	...	...	...	...	...	...
1925	1926	2017-12-24 00:11:57.442	Step_SPUtills	getTodayTotalDetailSteps = 1514045400000##0##...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	0	2017-12-24
1935	1936	2017-12-24 00:15:53.285	Step_SPUtills	getTodayTotalDetailSteps = 1514045400000##0##...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	0	2017-12-24
1951	1952	2017-12-24 00:25:16.438	Step_SPUtills	getTodayTotalDetailSteps = 1514045640000##0##...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	0	2017-12-24
1957	1958	2017-12-24 00:25:17.438	Step_SPUtills	getTodayTotalDetailSteps = 1514046240000##0##...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	0	2017-12-24
1966	1967	2017-12-24 00:28:25.453	Step_SPUtills	getTodayTotalDetailSteps = 1514046240000##0##...	E22	getTodayTotalDetailSteps = <*>##<*>##<*>##<*>#...	0	2017-12-24

242 rows × 8 columns

242 rows × 8 columns

```
dfs23['Date'] = dfs23['Time'].dt.date
dfs23['Hour'] = dfs23['Time'].dt.hour
dfs23['Minute'] = dfs23['Time'].dt.minute

steps_increase_threshold = 1000
min_duration_minutes = 30

workout_sessions = []
current_session = None

for index, row in dfs23.iterrows():
    if row['Steps'] > steps_increase_threshold:
        if current_session is None:
            current_session = {'start_time': row['Time']}
        else:
            current_session['end_time'] = row['Time']
    else:
        if current_session is not None:
            duration = (current_session['end_time'] - current_session['start_time']).total_seconds() / 60
            if duration >= min_duration_minutes:
                workout_sessions.append(current_session)
            current_session = None

if workout_sessions:
    print("Workout sessions:")
    for session in workout_sessions:
        print(f"Start: {session['start_time']} - End: {session['end_time']}")
```

Workout sessions:  
Start: 2017-12-23 22:15:29.635000 - End: 2017-12-24 00:00:00.234000

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\4098350854.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Date'] = dfs23['Time'].dt.date

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\4098350854.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfs23['Hour'] = dfs23['Time'].dt.hour

```
: dfsd24 = df[df['EventId'] == 'E4']
dfsd24['Calories'] = dfsd24['Content'].str.extract(r'(\d+)')
dfsd24['Calories'] = dfsd24['Calories'].astype(int)
dfsd24.head()
```

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\3852311565.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfsd24['Calories'] = dfsd24['Content'].str.extract(r'(\d+)')

C:\Users\PCC\AppData\Local\Temp\ipykernel\_14684\3852311565.py:3: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
dfsd24['Calories'] = dfsd24['Calories'].astype(int)

	Lineld	Time	Component	Content	EventId	EventTemplate	Calories
	8	9	20171223-22:15:29:645	Step_ExtSDM	calculateCaloriesWithCache totalCalories=126775	E4 calculateCaloriesWithCache totalCalories=<*>	126775
	17	18	20171223-22:15:29:959	Step_ExtSDM	calculateCaloriesWithCache totalCalories=126797	E4 calculateCaloriesWithCache totalCalories=<*>	126797
	24	25	20171223-22:15:30:639	Step_ExtSDM	calculateCaloriesWithCache totalCalories=126818	E4 calculateCaloriesWithCache totalCalories=<*>	126818
	31	32	20171223-22:15:31:157	Step_ExtSDM	calculateCaloriesWithCache totalCalories=126861	E4 calculateCaloriesWithCache totalCalories=<*>	126861
	38	39	20171223-22:15:32:156	Step_ExtSDM	calculateCaloriesWithCache totalCalories=126882	E4 calculateCaloriesWithCache totalCalories=<*>	126882

```
]: dfsd24['Time'] = pd.to_datetime(dfsd24['Time'], format='%Y%m%d-%H:%M:%S:%f')

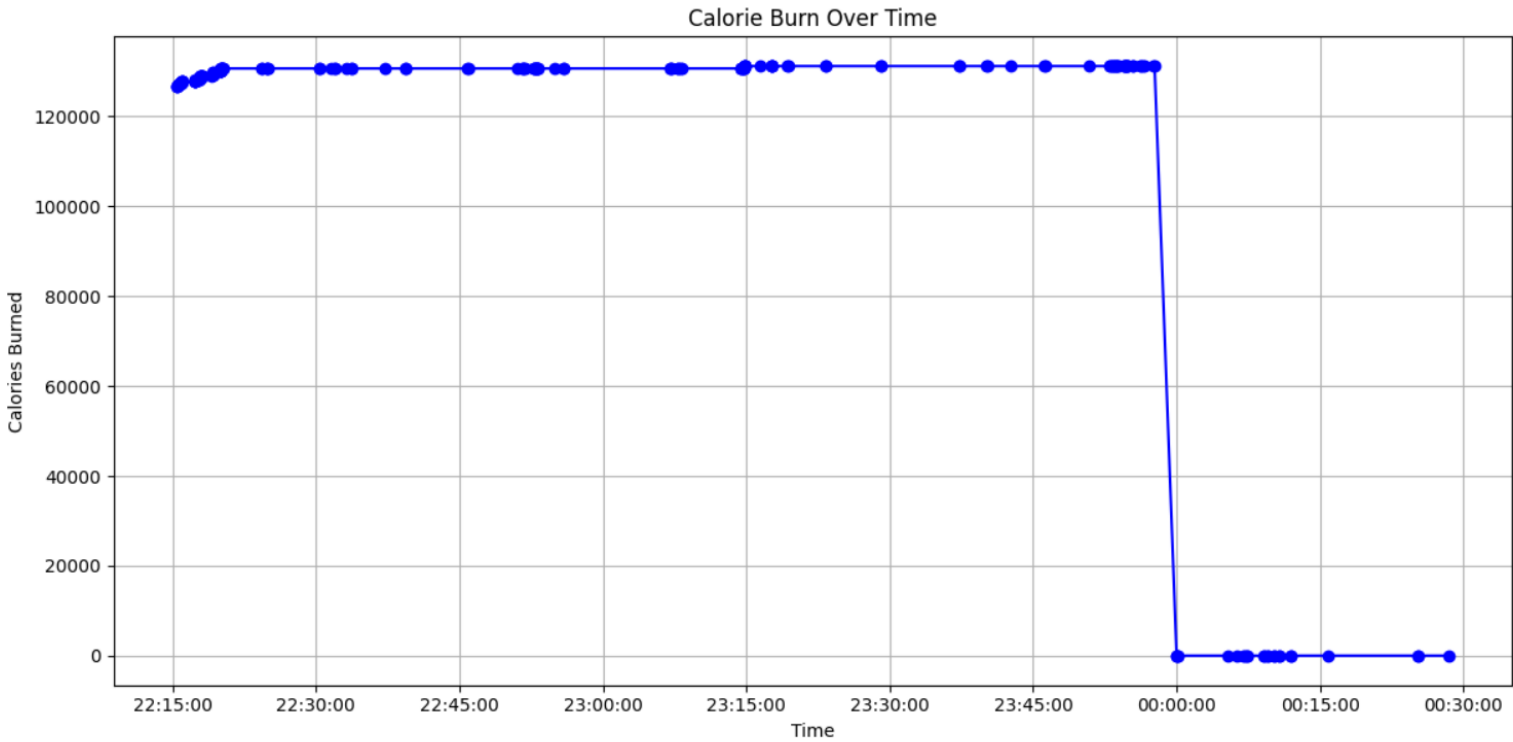
plt.figure(figsize=(12, 6))
plt.plot(dfsd24['Time'], dfsd24['Calories'], markers='o', linestyle='--', color='b')
plt.title('Calorie Burn Over Time')
plt.xlabel('Time')
plt.ylabel('Calories Burned')
plt.grid(True)
plt.tight_layout()

plt.gca().xaxis.set_major_formatter(plt.matplotlib.dates.DateFormatter('%H:%M:%S'))

plt.show()

C:\Users\PCC\AppData\Local\Temp\ipykernel_14684\2321828199.py:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
dfsd24['Time'] = pd.to_datetime(dfsd24['Time'], format='%Y%m%d-%H:%M:%S:%f')
```



Time

```
.]: # To remove rows with missing values:
df.dropna()
```

	Lineld	Time	Component	Content	EventId	EventTemplate
0	1	20171223-22:15:29:606	Step_LSC	onStandStepChanged 3579	E42	onStandStepChanged <*>
1	2	20171223-22:15:29:615	Step_LSC	onExtend:1514038530000 14 0 4	E39	onExtend:<*> <*> <*> <*>
2	3	20171223-22:15:29:633	Step_StandReportReceiver	onReceive action: android.intent.action.SCREEN_ON	E41	onReceive action: android.intent.action.SCREEN_ON
3	4	20171223-22:15:29:635	Step_LSC	processHandleBroadcastAction action:android.in...	E43	processHandleBroadcastAction action:android.in...
4	5	20171223-22:15:29:635	Step_StandStepCounter	flush sensor data	E12	flush sensor data
...	...	...	...	...	...	...
1995	1996	20171224-0:58:53:985	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1996	1997	20171224-0:59:7:581	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1997	1998	20171224-1:0:0:794	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1998	1999	20171224-1:1:0:935	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...
1999	2000	20171224-1:2:35:789	Step_LSC	processHandleBroadcastAction action:android.in...	E44	processHandleBroadcastAction action:android.in...

2000 rows × 6 columns

```
!]: # Check for null values
df.isnull().values.any()
```

!]: False

Value: 10220

## Conclusions

# From the given observations app loads detailed steps data when event E22 occurs. On further analysis (refer to cell 3) the following can be concluded: on 23-12-2017(from 10:15PM to 23:57PM) person's total step count = 7214 on 24-12-2017(from 00:00AM to 00:29AM) person's total step count = 0 Person's average number of steps per day = 23898 (approx) Workout sessions (Refer to cell 4): Start: 2017-12-23 22:15:29.635000 - End: 2017-12-24 00:00:00.234000 On 23-12-2017 after 11:14PM there is no change in steps or there isn't any movements so the person might be resting or put their phone on rest

## Calorie Counter

From the given observations app loads detailed calorie data when event E4 occurs. On further analysis (refer to cell 7, 8) the following can be concluded: on 23-12-2017(from 10:15PM to 23:57PM) person's total(cumulative) calories = 131208 on 24-12-2017(from 00:00AM to 00:29AM) person's total calories = 0

## Confirmations

On 23-12-2017 there is no any major change in steps and calories from 11:14 PM, this confirms that person is resting or put thier phone on rest