AI on Hi-ASAP 2020: Q&A

**Contact information:**

Speaker:

Chun-Hu: [lch0909@gate.sinica.edu.tw](mailto:lch0909@gate.sinica.edu.tw)

Vincent Wang: [phdzen@gate.sinica.edu.tw](mailto:phdzen@gate.sinica.edu.tw)

Joanne: [joannehu@gate.sinica.edu.tw](mailto:joannehu@gate.sinica.edu.tw)

Mark Tsou: [marktsou09@gate.sinica.edu.tw](mailto:marktsou09@gate.sinica.edu.tw)

1. Get calibration factor

**Question 1:** From Ha Chi

I got the .csv output for date 4/10/2020 only (in attachment). But I copied all csv and xlsx files from the "ASLung\_Calibration\_Factor\_RawData" folder you provided into the same name folder of pycharm\_project. Why couldn't I get the calibration factors results of date 04/08 and 04/09 even when I have modified the data time format of time column to be consistent with date 04/10?

**Answer:**

The fully data comparison period is from 4/8 to 4/10, we set the last date of 4/10 as “Start\_date”.

The aslung device save data day by day, you will have, so there are three data files in a fully comparison period.

**Question 2:** From Ha Chi

Could you please explain more clearly the slice No.4 about "Modify the data time format of time column"?

**Answer:**

The data time format should be “yyyy/mm/dd hh:mm:ss”, “yyyy-mm-dd hh:mm:ss” or "mm/dd/yyyy hh:mm:ss", you can modified the data time format in the excel file.

Ex:

yyyy/mm/dd hh:mm:ss=>2020/10/06 09:46:00

yyyy-mm-dd hh:mm:ss=>2020-10-06 09:46:00

mm/dd/yyyy hh:mm:ss=>10/06/2020 09:46:00

**Question 3:** From Thiha

When I run the maximum concentration of the PM, I am getting these error messages. How can I fix the error? Thank you.

Please kindly see the attached file.

Error Message:

AS-Lung ID :

AL-0339

C:\Users\hp\PycharmProjects\ASLung\_AI\venv\lib\site-packages\pandas\core\frame.py:4163: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

return super().drop(

as\_id: AL-0339

………………….(Omit)

packages\numpy\core\fromnumeric.py", line 87, in \_wrapreduction

return ufunc.reduce(obj, axis, dtype, out, \*\*passkwargs)

ValueError: zero-size array to reduction operation minimum which has no identity

**Answer:**

The data set is not correct. Thiha copied data cleaning data set to get calibration factor folder.

Please check the data set is correct.

**Question 4:** From Thiha

I get another error message. How can I fix it?

**Error messages**

Maximum concentration of standard PM: 200

AS-Lung ID :

Traceback (most recent call last):

File "ASLungCalFractor\_2Segment\_Regesssion\_AI.py", line 49, in <module>

std\_df\_list = std\_df(Standard\_file\_path[0])

IndexError: list index out of range

AL-0400

Process finished with exit code 0

**Answer:**

Did you copy the reference PM to the folder of “ASLung\_Calibration\_Factor\_RawData”

**std\_df\_list = std\_df(Standard\_file\_path[0])**

**IndexError: list index out of range**

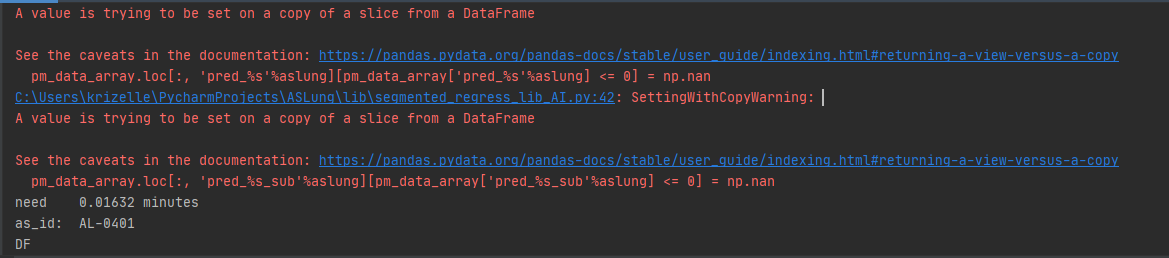
This error code mean that “run.py” cannot find reference PM

Please check that “standard\_2020\_Lab\_Grimm026\_0410.xls” is in “ASLung\_Calibration\_Factor\_RawData” folder.

**Question 5:** From Krizelle

I was able to run the code. However, a consistent “error” was seen. I also have a different result compared to the one from the slides.

Error message



**Answer:**

This is a warning, please ignore it.

**By the way, when you copy the calibration factor to the google sheet, please adjust the width of the column to show the start date.**

**If not, you will get “#####” when you past then to the google sheet.**

1. Data cleaning

**Question 1:** From Ha Chi

I followed the guideline in the presentation but there is no any outputs if I use the google link provided in GoogleSheetAndAPI.xlsx (I had copied the output of calibration factors to this link before copy to run.py)

I tested with the original run.py, it run well. So I believe the step of getting calibration factors might go wrong somewhere. Could you please help to solve this problem?

**Answer:**

The data set of get calibration factor and calibrate AS-Lung data are different.

There is no calibration factor to fit the test data set of AS-Lung in yours google sheet, so you cannot see any output in the folder.

If you want to try how to calibrate AS-Lung data, please do not change the calibration factor API or copy the factors to yours google sheet(GoogleSheetAndAPI.xlsx). I have modified the out in the running window, it will show “**There is no calibration factor at the sampling time of XXXXX !**”

The factors of calibrate AS-Lung are in the google sheet (<https://docs.google.com/spreadsheets/d/1yuvjPvsr1sEzm_pXpZWpnMxEC3WbsKn65V-apohp86E/edit#gid=0>), the link also show in the PDF file(page 10)

BTW, I found a little bug when write aslung id into the calibration factor file, it has been fixed.

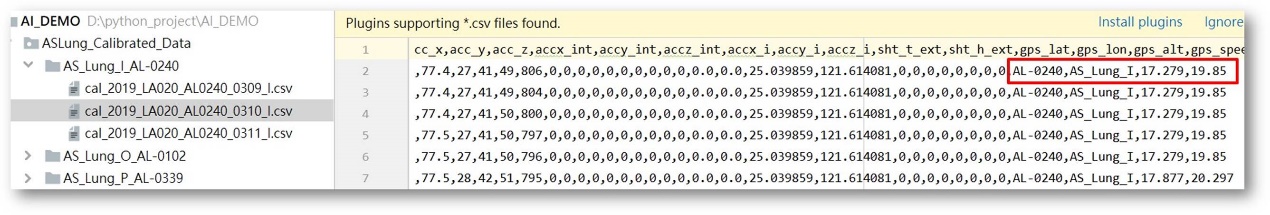
If you download the source code before 10/6, please update it.

**Question 2:** From Krizelle

I was also able to run the code and I saw the 3 columns. However, I do not know if the values are correct. The data here and the ones from the slides are different.

**Answer:**

I have check the data, the data is the same with mine.



If you want to check another data, you can calculate it by yourself.

Here are the full calibration factors for DEMO.

<https://docs.google.com/spreadsheets/d/1yuvjPvsr1sEzm_pXpZWpnMxEC3WbsKn65V-apohp86E/edit?usp=sharing>

**Question 3:**

Error messages:

Data file: D:\python\_project\AI\_DEMO\ASLung\_RawData\LA002\AS\_Lung\_O\2018\_LA002\_AL-0136\_1024\_O.csv

Folder: LA002\_AS\_Lung\_O

Calculate data file of : 2018\_LA002\_AL-0136\_1024\_O.csv

data error: Error tokenizing data. C error: Expected 31 fields in line 514, saw 53

**Answer:**

Open the data with Notepad++, go to line 514, the data format is incorrect.

Modified the data format and run again.

