# Taguette highlights

So the example ... event day, it was in February and the following columns have the variables, environmental factors.

The temperature and the other variables like precipitation, solar radiation and many empty data.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

So it is uploaded but I haven't seen it yet.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

No I don't have to pick any of the linkage purpose options I suppose.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

I don't understand this paragraph.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

So, it depends on the purpose I can choose if the events are for my specific example or season-alike if I want to share it with others.

I suppose this is that.

Okay, now I select the one for me.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay, to choose from an area within a certain distance I suppose I do not have enough details at the moment to select that.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

I will choose within an area.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Region levels, NUTS ... Okay if I click here and choose Switzerland I will see the different NUTS.

Ah okay, Switzerland is here, okay.

Okay, so now we will choose CH05 for St. Gallen. So I can come back and select NUTS-2.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Temporal units month and aggregation method the sum.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

So now I can link the data, Oh something is going wrong.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

I don't know if step 2 is ongoing or finished.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

MODERATOR: In the first step you have uploaded the example table and not your table because you clicked the upload data table button. You will have to upload it again.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Some important features unclear

MODERATOR: You also missed to select two options from the linkage options.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Some important features unclear

So drag and drop, I can select the file here. I hope I have uploaded the correct file but it does not say anything.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

MODERATOR: You have a message right below it.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Some important features unclear

Oh okay, with this message I can go to the next step.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Oh but I did not prepare the metadata file for this.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

MODERATOR: No problem, you can use the example.

Document: P1\_GMT20220711.odt Tags: Moderator guidance

Yeah, I have completed this ... NUTS-2 ... month ... sum ...

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay now it is working.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Within an area CH05, I will get all the data. It will not be only St. Gallen I suppose.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay, I can see it is also loading step 3.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay, now I have to export but I already clicked on the link. Do I have to do it again?

Document: P1\_GMT20220711.odt Tags: Technical web issues

MODERATOR: I think it did not work. Could you please show me the table you imported?

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Technical web issues

Sure, I will reshare my screen. Can you see it now?

Document: P1\_GMT20220711.odt Tags: Technical web issues

MODERATOR: Oh, I see the length of your time window is 30days and you have selected months. So you get a single data value or none. Do you mind resharing your screen and then I will resume the timer?

Document: P1\_GMT20220711.odt Tags: Moderator guidance

MODERATOR: Oh, I see the length of your time window is 30days and you have selected months. So you get a single data value or none.

Document: P1\_GMT20220711.odt Tags: Some important features unclear

Do you mind resharing your screen and then I will resume the timer?

Document: P1\_GMT20220711.odt Tags: Technical web issues

Sure, okay now if it is working. Export output. But I do not get any output.

Document: P1\_GMT20220711.odt Tags: Technical web issues

MODERATOR: Did you get any pop-up window after clicking the button because I could not see it due to your Zoom settings.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Technical web issues

I clicked on export output but nothing happened. I am checking downloads and I do not see anything.

Document: P1\_GMT20220711.odt Tags: Technical web issues

MODERATOR: Okay let's try to refresh, link and export again because it should have popped up an export window. The last query seemed to work because I can see it here but maybe the SUM is not working, I am not sure I will have to check.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Technical web issues

Sure, yes of course. I have selected the same options.

Okay, now it is working.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

I can see a zip file here.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Oh I guess I can input my event file now. I will check if it is a .csv file because I am not sure.

I will upload the correct one ... upload .

Document: P1\_GMT20220711.odt Tags: Some important features unclear

MODERATOR: If you click the data table button you upload the data table in the example not yours. You will have to refresh the website again. Apologies I need to include a message there mentioning that you don't need to click the button if you had already input your own table.

Document: P1\_GMT20220711.odt Tags: Moderator guidance, Some important features unclear

So now I have my csv file uploaded, and I will be using the metadata file from the example because I don't have one.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay, within an area ... NUTS-2 ... day ... mean.

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

Okay, it's good it works!

Document: P1\_GMT20220711.odt Tags: Data linkage process clear

It seems like you have a lot of variables, no missing data, the events here with the dates.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

I have 3 patients but in the output I have 4 patients, why is that?

Did it take into account my CSV file? I think it was only 3 events/patients when I uploaded my data.

Document: P1\_GMT20220711.odt Tags: Output data unclear

Let me check that. Oh no, I had four. It is fine.

I asked for 30 days so that is why I have November, December and so on.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

Yeah, okay it seems pretty good, that is one file.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

MODERATOR: Do you mind going back to the tasks?

Document: P1\_GMT20220711.odt Tags: Moderator guidance

A, B, C and D and then you have the dates, the lag, which I suppose is 30 days ... 30 rows for A.

For B, the same but I never get the full 30 days, it is 29 or 28 ... Does the month influence the number of days?

The columns are fine and the colours are explained here.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

I have to put A to search for event A, oh I see I have only the A now.

Document: P1\_GMT20220711.odt Tags: Report helpful

So here, I have them in blue.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

Okay, each line is a variable and I asked for 30 days and that is why I have from 0 to 30 days.

Document: P1\_GMT20220711.odt Tags: Visualization of plots clear

If I want to select one variable in the plot, let me see. Okay, I understand now.

Document: P1\_GMT20220711.odt Tags: Visualization of plots clear

The columns correspond to the weather changes.

Document: P1\_GMT20220711.odt Tags: Visualization of plots clear

Z-scores, yes okay.

Document: P1\_GMT20220711.odt Tags: Visualization of plots clear

It has changed a lot from the previous version.

Document: P1\_GMT20220711.odt Tags: Improved tool usability and output data

Okay, here I can select the individual events: A, B, C and D.

Document: P1\_GMT20220711.odt Tags: Visualization of plots clear

For D, I do not have any line, I don't understand.

Document: P1\_GMT20220711.odt Tags: Visualization of plots unclear

Yes, I am confident.

Document: P1\_GMT20220711.odt Tags: Improved tool usability and output data

I suppose I need to read the report with more time to read the details so I can be more confident in using the data.

Document: P1\_GMT20220711.odt Tags: Improved tool usability and output data

No, it is really complete.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

You have all the csv files, the interpretation in the html file, so for now I do not have any other query.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

I suppose if we are going to use other variables we have to estimate the other variables with the available data. But it would be done later on the analysis.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

My impression is that it is really complete.

Document: P1\_GMT20220711.odt Tags: Output data ready for analysis

It is usable.

Document: P1\_GMT20220711.odt Tags: Improved tool usability and output data

I would like to use it in real conditions to see if there is something lacking, but the first impression is that I have everything available for the analysis.

Document: P1\_GMT20220711.odt Tags: Real use case applicable

You provided a description of event here, just a comment but putting an asterisk

which is almost like a footnote or something. because at the moment I am developing terminology around principles.

It is good to have a clear link between this and this so it is not out of sync but it has a purpose.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Not sure if the wording of that first sentence ...

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I like the temporal and spatial aspects but it is probably more scientific in language but it may be space there.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

This is kind of the location and timestamp ...

Just notice the word POINT comes up but I note that the location point is a different term that the one previously used

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I would like to click to see the example. Okay, it is a spreadsheet, tabular form data ...

I don't know if this is external or part of the platform but some of these do not have units.

This needs to be clear as the temperature could be Kelvin or Celsius which could change everything.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Other than that it seems fine, in terms of preview and download different datasets that are available.

I am guessing you can download it as CSV or other options but I am not going to go there.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

The language is the same which is good to see the consistency between the interface and the output.

Document: P2\_GMT20220714.odt Tags: Positive comments on text

The only thing I would think of is that the metrics are missing there.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

So if I remember correctly this is like what it does. Yeah, okay I am thinking if there is a way to put coision into this like putting 1, 2 and 3 here.

Just to show that when you come down here that is what Iread exactly

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Okay, so step 1, step 2, step 3 ... very nice flow to it.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Okay so you can export the sample which is very good so you don't have to start from scratch.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I suppose that this is why this export is permitted, very good.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Yes, the naming of the file is not an issue when exporting the table.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

When uploading the table I guess you can import any file but I won't test that just yet because the scenarios might have to do with that.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

For example, the format Iam importing the data.

There is no reference for the type of file ... Does it have to be CSV, EXCEL or another format that is accepted.

Old versions of excel like version 3 which sometimes people save files as. You can choose from conversion options.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

So linkage seems to me the word that is used but I would probably like to see "event" and "environmental" featured here.

Because for me I had to go back and remind myself of it.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

As a first time user Ineed to be very transparent, but as a fifth time user, this is second nature.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear, Improved tool usability and output data

Environmental is not featured here at all so you might want to add "linking event and environmental data".

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Something like this around step 2 terminology.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Yeah, these are the ethical aspects going on. I assume that the data the person has is personal data so it makes sense having it there.

Document: P2\_GMT20220714.odt Tags: Data protection importance

I don't know if the last sentence serves any purpose though.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I know that when you are processing personal data you need to be very transparent and comply with the ethics.

Document: P2\_GMT20220714.odt Tags: Data protection importance

Are these purposes the only possible ones or examples? Are these the only options you can undertake with this platform?

So it is either to conduct analysis or prepare it for an open data repository.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Metadata input there are recommended and minimum options. That is okay.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I would have to figure out what this section is in terms of the minimum.

Okay, here it is.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Season-alike dates is a term Iam not familiar with. Imight need a footnote or tooltip with a definition or clarification.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Unless it is seasonal, but if there is a difference it definitely needs clarification from a researcher point of view.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Oh yeah it is here, so it is seasonal I think. But it has a spatial component.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Oky, this is something more around the risk of re-identification of an individual. Somehow the way you are producing the data ensures anonymity.

That is my interpretation but this sentence is quite a sentence.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear, Improvement to text or feature

I am also trying to think about this so, this step 2 is about linkage options but this is almost a warning, caveat or constraint.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I feel it is a constraint. It could improve the way it is written by using bullet points with specific titles at the start of each sentence with an explanation.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Like a constraint almost.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I would be curious how much this would affect the result. I will choose as input because I would be trying to conduct data analysis.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

But then I assume that if I was trying to deposit the data in an open data repository the season-alike would be the preferred output.

That is my interpretation of what I am getting.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I don't know if there is almost a need for step 2a and 2b.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Where 1 is about is about conducting data analysis and you keep it nice and clean.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear, Positive comments on text

I don't think they would like to do season-alike for data analysis.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Then, asking if you want to conduct data analysis or prepare data to deposit in an open data repository. Just a viewpoint.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Within an area or distance ... it depends if it is an urban or rural environment ...

Document: P2\_GMT20220714.odt Tags: Real use case applicable

I think it is getting from A to B ... Distance would be probably more appropriate but I would have to think about that.

Document: P2\_GMT20220714.odt Tags: Real use case applicable

This is the maximum resolution but I think this is due to the data availability.

Document: P2\_GMT20220714.odt Tags: Data availability comment

I want to start clicking on these but I can only select one. I don't want to mess with the system just yet.

I will play with that later.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Okay, it takes from seconds to minutes, that is very good.

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

The outputs are pretty straight forward. I am very familiar with FAIR.

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

I am curious about whether the graph files are ready to be shared because I can conduct data analysis and would have to meet these criteria already.

Unless anonymity still exists in the end, so I am curious about this.

Document: P2\_GMT20220714.odt Tags: Data protection importance

It is not necessary for data analysis. I am trying to interpret where data analysis begins and where linking the data.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

I don't know if Ifollow into the HDR definition but I at the same time I would find this useful.

Document: P2\_GMT20220714.odt Tags: Real use case applicable, Researcher's expertise

I would find it useful actually because I would come from the environmental data side.

Document: P2\_GMT20220714.odt Tags: Real use case applicable

I would add "environmental data background" in the scenario.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Looking at the platform, again I am trying to imagine what it would change if coming from an individual or population stamp point, in terms of data.

Especially anonymity etc.

Document: P2\_GMT20220714.odt Tags: Data protection importance

At this point I would probably like to upload the dataset that has been provided, even though it might be the same.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

One thing that I don't have is the description of some of these but I don't know how important that would be.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

There are no descriptors of the column names.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

The length of the event is 14 days and a lag of 7 days, I am not sure about that completely.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

I don't know if other formats would work but I trust the developers.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

it says uploaded so this would be another way of uploading the file.

Document: P2\_GMT20220714.odt Tags: Output data unclear

When I dragged and dropped the success message was different so I am not sure if I have repeated the process twice.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

I probably don't have time to do both of these but I will choose data analysis.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

The warning message is very helpful here.

Document: P2\_GMT20220714.odt Tags: Positive comments on text

Okay it is very good and I change between the purposes so that is great.

Document: P2\_GMT20220714.odt Tags: Positive comments on text

The terminology is different here.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I would prefer recommended metadata ... Okay, where did this come from?

Unlike the previous step where the table was there this was not available to me.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Hold on, so task 1 is complete actually? No? I might have to ask what I have missed.

Document: P2\_GMT20220714.odt Tags: Experimental methodology

So what I end up doing is exporting the metadata file and importing it if that is what is required.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Okay, so I could have dragged the table I had just exported.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

When I clicked on recommended ... Yes, this is a bit tricky ... If I would have explored the platform already ...

Document: P2\_GMT20220714.odt Tags: Some important features unclear

There is a lot going in step 2 ... You need to upload a file. Then, go through all of this, which is quite intense or hopefully with not many changes.

Then, come back to the platform. So this could be in the middle if I wanted to do a quick analysis but it serves the function.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

I suppose options yeah, what are these options? Are they to do with case and season or others?

I am going to go as input because I am not sure what is season-alike dates are.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

I am going to go with distance as a preference. Okay I got an option for distance in km.

I am going to choose a 10km radius unless this can cause problems.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I would usually say long/lat instead of lan/lon.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

I don't know if this small dataset will take long. I always like to start small and then, meaning is always good.

Document: P2\_GMT20220714.odt Tags: Experimental methodology

Ah! I can only choose one aggregation method. I wish there was a (s) so I could select more than one.

That way Ican actually, I suppose I want to do data analysis so the more outputs I can get, the easier my life is.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

And so, the link button is key and ... technically the task is completed once the spinner stops.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Unfortunately my button is not working

Document: P2\_GMT20220714.odt Tags: Technical web issues

MODERATOR: It happened to me this morning, could you click the button again?

Document: P2\_GMT20220714.odt Tags: Moderator guidance

Sure. Okay, now it works. It must be a glitch.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

But the export button does not work ... nothing is downloading

Document: P2\_GMT20220714.odt Tags: Technical web issues

MODERATOR: Maybe you need to refresh the web there as you linked twice the data. Then, you could fill again all the options quickly

Document: P2\_GMT20220714.odt Tags: Moderator guidance

Okay, I will refresh. Okay upload the table, the metadata that is already there.

Document: P2\_GMT20220714.odt Tags: Technical web issues

If the metadata is already there by default, this is a suggestion that you have to upload.

Okay I will select the same linkage options ... and link

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I don't think the link success message was there when linked the first time.

So when the message does not appear the link was not done correctly.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Okay extract all, zip file.

I had an interrupted action about the length of the file being too long to export.

Document: P2\_GMT20220714.odt Tags: Output data unclear

I might be wrong but the compressed name and the name of the files, all of those have lots of characters. They are rather lengthy.

Document: P2\_GMT20220714.odt Tags: Output data unclear

I have 4 excel files and 2 ttl and 1 e-report which I think is the html. I have a lot of options here.

It basically has several different excel files which I will open all briefly, individually.

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

I don't know if it is possible to have them all in one excel file in different tabs.

Document: P2\_GMT20220714.odt Tags: CSV output with missing data

This is a lot of the air quality, this is more the raw data with the events A, B, C and D.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

This was the daily conditions for A in January and February, August and July.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

For some reason, the months are backwards by default. For A, February comes first then January, not sure why but it does.

Some months have missing data unfortunately. But there is a range of different information.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

Okay, so this file is the data related to what happened that day ... I guess In the daily resolution the dataset is very small for me to do an analysis beyond this.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

And also it says January and February, even though Iask for daily I am getting monthly values.

Document: P2\_GMT20220714.odt Tags: Output data unclear

Okay, in the journey file, there are the 10km buffers but the journey file does not serve too much.

Document: P2\_GMT20220714.odt Tags: Output data unclear

Okay, this one has the normals, ah yeah okay ... normals, values, sdev .... normals as normalised?

Document: P2\_GMT20220714.odt Tags: Output data unclear

Okay, the lag, okay so this is based on the query information and the initial dataset when Iasked for 7 days of data.

So the dates reflect that.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

Okay, so we could analyse this, but I need to think on the analysis approach to analyse this data.

But that would be for another thing. It is clear that there are some data gaps.

Document: P2\_GMT20220714.odt Tags: Real use case applicable

Okay, all the files seem very clear and I will close the csv files.

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis

I will look at the graph files but I don't know how to open them.

Let's try this ... Oh, it did not work but I am not overly concerned.

I can open them with firefox ... there you go, okay.

Document: P2\_GMT20220714.odt Tags: Output data unclear

Okay, the report is very nice and helpful.

Document: P2\_GMT20220714.odt Tags: Report helpful

A lot of explanations and descriptions about the data.

This is very nice.

Document: P2\_GMT20220714.odt Tags: Report helpful

Okay, this is very nice and helpful.

Document: P2\_GMT20220714.odt Tags: Report helpful

So interestingly, the order on which the files are listed in the UI. I feel this one should be at the top, to look at them in this order.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

Because the other files are not definitely as useful as the interactive report, which is by far the most valuable file.

Document: P2\_GMT20220714.odt Tags: Report helpful

The other files are almost as supplementary information. So you need to instruct people to look at the interactive report first.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature, Report helpful

One thing I would query is, because the interactive report is an html file ... could this be linked here?

Because at the end of the day you have exported the file but maybe you could get people to view it here because you are essentially exporting a link of the file.

But a link is a link that you could produce but you cannot hyperlink something that does not exist.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

Definitely the interactive report is the priority file to view. It is very very ... There is a lot going on there that deserves attention.

Document: P2\_GMT20220714.odt Tags: Report helpful

So is that task 2 completed?

Document: P2\_GMT20220714.odt Tags: Experimental methodology

Okay so I'm going to go through the report.

There is temporal and spatial information, there are lag considerations ...

We have a lot more going on in terms of ... This is an interactive report, fair ...

You can do lots of things with it and it is quite helpful.

Document: P2\_GMT20220714.odt Tags: Report helpful

The plot information buttons are good explanations. This is probably why you need to play around once or twice as the actual tool.

Document: P2\_GMT20220714.odt Tags: Report helpful

This is very well presented.

Document: P2\_GMT20220714.odt Tags: Report helpful

Yes, I would need more time to read the details but it seems really complete

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

I suppose my research focus might be on analysing the air quality with the health impacts. So the lag would be extremely valuable to me.

It is probably the trigger period as opposed to the duration.

Document: P2\_GMT20220714.odt Tags: Real use case applicable

One of the things here is that this is more just the dataset, not necessary the data analysis.

But it prepares the data very well for transforming it to something that can be used in SPSS or other programs.

so just trying to events, so the events are the individuals ... That is perfectly ...

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data, Real use case applicable

The only thing is that for my analysis I might want to exclude certain columns and pieces of information but I can easily do that with excel.

I guess I have no choice but to take everything in the export, that is the only opportunity for me to actually remove the variable I do not need.

So from my research point I want to be more specific about data that Ineed.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

But some researchers about Big Data that don't understand the problem fully might want to include all.

So I guess it depends on if your audience is not someone who wants to narrow it to certain pollutants that works.

Document: P2\_GMT20220714.odt Tags: Real use case applicable

Okay this is individual events A, B, C, D ... very different plots.

I can understand why you would like to see one, maybe see them side by side now that they are 4.

But if it starts getting to large datasets then it is not feasible.

Document: P2\_GMT20220714.odt Tags: Report helpful

So in terms of confidence, I would be highly confident because everything is available to me.

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

Probably I would like to narrow it down, but that is coming from my approach.

Document: P2\_GMT20220714.odt Tags: Real use case applicable, Suggestions for additional features

However, it is better to have everything available and you can select that after retrospectively not having regrets of not having included something.

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data, Real use case applicable

All the csv files, the interpretation in the html file so for now I do not have any other query

Document: P2\_GMT20220714.odt Tags: Improved tool usability and output data

Other variables can be estimated from the available variables

Document: P2\_GMT20220714.odt Tags: Real use case applicable

I mentioned being able to compare visually, again how appropriate that point would be for this ... I don't know.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

One thing is to compare 2 events. But again this report has to have some limits and constraints so I am not too concerned.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

I am just wondering if I have to go back to the user interface. What could I modify to re-create a new report based on having looked at the first report?

That is my approach, not everyone's but I can remove some of the data to simplify the report?

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

Like a second chance almost because I recognised that some of the pages of the report are relevant but others not.

Document: P2\_GMT20220714.odt Tags: Complex report content

But just going back to the UI and yeah. Removing some of the environmental variables while the event stays put, but the environmental data I don't have any control over that.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

The other feature would be perhaps in this case ... in step 2, the metadata table only appeared when I clicked those items.

I was surprised and it came about that. Oh there is a file I need to import here, while in step 1 I always was clear from the beginning that Ineeded to upload a file.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

In being prepared for the process, I was prepared for 1 but not for 2, which I could have been.

I think that is quite important as this could be an intensive file to prepare.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Maybe as an event as step 0 that you will need an event file ...

Step 1, 2 and 3 are more about the process and step 0 is more about having the information ready for the process itself.

This did cause by default which would slow me down in the process.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Lastly, the aggregation method, could I have more than 1? But again, I don't know if that would be highly beneficial.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

So it is more about the temporal unit and aggregation method so you could choose more than one. But again it could create chaos with files and the report could become extensive itself.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

Lastly, the word event if you can ... serdif report you have this fantastic information with icons ... but the event seems like the event was something that needed to be explained.

Anything that could be interlinked or information boxes can be beneficial.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

It is really complete, usable and I will have to use in real conditions but the first impression is that I have everything

Document: P2\_GMT20220714.odt Tags: Output data ready for analysis, Improved tool usability and output data

I mean going through step 1 was very clear and very simple perhaps. It could be more complex but once I understand the length and the lag it is very clear.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Again, step 2 may need two sections one for conducting analysis and another for preparing the data for an open data repository because they are 2 very different tasks.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Also this appearance of the very important metadata table that I find very important, but just pops out as a requirement.

Document: P2\_GMT20220714.odt Tags: Some important features unclear

Continuing with step 2, I had to choose some spatial and temporal factors and that was fine.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

I would like to have the option to select multiple options but it could get overly complicated in the report, the data outputs and the duration for running.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

Step 3: the interactive report should get the attention first and the other data files become supplementary.

The interactive report is directed to reviewing that and then you know what you have in your hands in terms of excel files.

Document: P2\_GMT20220714.odt Tags: Improvement to text or feature

The actual report itself ... it is very well linked in terms of explanations and terminology, etc., which I think is good.

Document: P2\_GMT20220714.odt Tags: Report helpful

There is little I would be suggesting. There is a lot going on to be honest, but in fairness the clarity and detail going into the explanations is very much existing.

Document: P2\_GMT20220714.odt Tags: Report helpful

A lot of information is clearly linked to the metadata files and I am guessing it comes back to the metadata file being properly prepared.

Document: P2\_GMT20220714.odt Tags: Report helpful

I don't know if there is a way to check if it works after messing with the metadata file with spaces and back slash.

Document: P2\_GMT20220714.odt Tags: Suggestions for additional features

That processing steps ... All of that is very clear.

Document: P2\_GMT20220714.odt Tags: Data linkage process clear

Yes, so this is the air pollution information which is important.

Okay, yeah the units are also here.

Two standard deviations is the threshold for the z-scores ...

I suppose again as the report the one thing I am curious about is how does this ... Can you export this report?

Because it is currently an html file, I don't know how long this html file survives. Oh no, it is a downloaded html file.

It is not something linked. This is one of the things I would be curious about is the export ... Oh yes, you can export these figures.

Document: P2\_GMT20220714.odt Tags: Report helpful

There is a lot going on here as well ... Yes, so there is a lot in this report that is extremely valuable.

Document: P2\_GMT20220714.odt Tags: Report helpful

It is not necessary for data analysis but very useful for the methods chapter of a paper for example.

Document: P2\_GMT20220714.odt Tags: Real use case applicable, Report helpful

It is the input and output data, so it is useful. Importing and downloading is very useful indeed.

Document: P2\_GMT20220714.odt Tags: Report helpful

There is a lot going on when you hover over the data but I don't know how it helps but it is there, which I suppose is better than not being there.

Document: P2\_GMT20220714.odt Tags: Report helpful

Drag and drop ... okay ... It says there the event data was successfully uploaded.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Oh wow, nice! You can see the data here. That is really nice.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

You can click upload the data table. Same message appears again about the data being successfully uploaded.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Yeah, likange options ... That is very interesting.

My purpose is to conduct data analysis ... Okay, it explains that here.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Okay, so the recommended metadata is more extensive than the minimum. I will go for recommended metadata then.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Okay, I think I remember these ones. I will go for “distance” ... Let's say 10km is good Monthly temporal units and mean

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

MODERATOR: The example you submitted only has a few days so if you select “month” you will get a single month.

Document: P3\_GMT20220719.odt Tags: Moderator guidance, Some important features unclear

Yeah , this is nice actually. Easy to follow with three steps.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Then, the information about the metadata input is nicely summarised.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

It is better to spot them in this table format.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

I like that when you click the upload button for the tables.

It gives you a confirmation message that it has been successfully uploaded, and that you can proceed to the next option, which is a good thing.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

The spatial option "within an area" ... The NUTS-3 areas ... I used to remember those things but I don't think I remember what different NUTS mean.

I think if you are working with climate data it should be easy or you can always google it.

Document: P3\_GMT20220719.odt Tags: Some important features unclear

Distance in km seems like an easy option.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Oh nice, export.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Can you still see my screen? Okay, can you see it now? Okay ...

Document: P3\_GMT20220719.odt Tags: Technical web issues

I will start with the report.

Oh nice, I am just going to click here, oh wow! Amazing!

Document: P3\_GMT20220719.odt Tags: Report helpful

I am going to browse quickly first.

Wow, this is nice hahaha

Document: P3\_GMT20220719.odt Tags: Report helpful

Oh, nice!

Document: P3\_GMT20220719.odt Tags: Report helpful

Oh wow this is nice. This is nice.

Document: P3\_GMT20220719.odt Tags: Report helpful

Oh, interesting.

Document: P3\_GMT20220719.odt Tags: Report helpful

Oh, nice wow, thi is really ...

Document: P3\_GMT20220719.odt Tags: Report helpful

Let me see, so we got event A and all the dates up to 12 days.

For event B ... Can I ask a question? The lag number here does it refer to the days prior to the event? Oh okay here it is.

Then, the lag numbers correspond to the number of lag that you put in the input with your dataset. Oh, okay.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

The dataset I put in it says, event, lat, lon, length and lag ... So the length referes to the number of lags generated okay.

So for event B, we start from 7 all the way to 14 ... okay.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

This is really nice.

Document: P3\_GMT20220719.odt Tags: Report helpful

Oh wow, nice! From the event going back 25 days for all the air pollution variables. Oh wow, that is amazing.

The weather background is the colouring on the right.

Document: P3\_GMT20220719.odt Tags: Visualization of plots clear

For example, the air pollution was higher at the time of the events and then decreased for some days. That is interesting.

Okay, so I assume this is the average. Oh really nice.

Document: P3\_GMT20220719.odt Tags: Visualization of plots clear

Okay, so we can change the colorbar axis ... nice.

For example, SO2 is above this lag.

Document: P3\_GMT20220719.odt Tags: Visualization of plots clear

Oh amazing, wow! This is really nice.

I can check the concentrations for each event ... normal and high concentrations here.

Document: P3\_GMT20220719.odt Tags: Visualization of plots clear

I can change the axis and it changes the background pattern ... Wow interesting, this is really nice.

Pollution data against wind speed or temperature.

Document: P3\_GMT20220719.odt Tags: Visualization of plots clear

Oh nice. A lot of data.

I can see the events A, B, C and D, the month.

Okay, that is nice.

These are month aggregates ... You have them in different formats.

Oh wow, this is

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

Yeah, this is really quite more than what Ihad expected.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

This gives you more data, more options, the whole website is easy to follow with three steps.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

We don't get lost in the information about metadata.

It guides you through defining what format you need to upload.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

If you need to upload the metadata yourself, it is something to think about because Iam assuming that before you conduct the data analysis you would have gone through the application process.

Document: P3\_GMT20220719.odt Tags: Some important features unclear

Some of the metadata stuff is just copying and pasting. So if you have to do or share that at least you have that example there.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Then, within like seconds you get your results. Sometimes when I do it, it can take days.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data, Real use case applicable

This is amazing because yeah, it gives you different options about time and space.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

The one thing ... Okay then you can an aggregate of let's say pollution readings within 10 or 20km you can check how much it changes.

This would be an extra step that you would not really need within this step here.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

I noticed that for some of the pollution data there is no measurement reading, maybe this is due to the fact that there are no pollution stations nearby.

Then, you can always replace that with some other distance like 10-25km which is what I have read in a manuscript.

Document: P3\_GMT20220719.odt Tags: Suggestions for additional features

I know that for temperature that is always well done with good modelling and there is good temperature data in most areas.

Document: P3\_GMT20220719.odt Tags: Real use case applicable

It is nice that you can visualise things. The average between let's say your sample and then you can link it to individuals but it does not tell you much information unless you are interested in the outliers.

Document: P3\_GMT20220719.odt Tags: Report helpful

Yeah, it is really user friendly . It is like, yeah somehow the human brain likes the easy way out hahaha into doing things.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

It helps you understand what you are doing and then, put it together. You save us a lot of time.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data, Real use case applicable

You would save a researcher like months of having to learn how to do the linkage.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

In my case, I had to link half a million people and my computer just freaked out.

Of course these are just 4 people, but I just want to put more people, click and go have lunch and hopefully then in a few hours it comes out.

Document: P3\_GMT20220719.odt Tags: Real use case applicable

Yeah, I guess in some cases when you are doing a few cases if you had the support from different countries, maybe an extra variable could help.

But you would know which country they belong to so you don't need to ...

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

It is nice that you have standardised, in the UK they use Celsius, but I don't think it matters here.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

From a country perspective you don't need to label who comes from which country

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Again thinking about data protection.

Document: P3\_GMT20220719.odt Tags: Data protection importance

I like that it gives you a nice summary at the beginning.

Then you have an option to look at the example output, which is something else.

Document: P3\_GMT20220719.odt Tags: Positive comments on text

Okay, yeah I mean you get so much information at the beginning to have an idea of what you need.

Document: P3\_GMT20220719.odt Tags: Positive comments on text

The learning curve is like my first time doing this ... It would take me less than an hour. I mean it took me longer here but if you are working in the field you would know what you are looking for.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

Then, you have three steps: upload health events, choos linkage options and export; and that is it.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

The results give you some sort of zip that you can download and check a summary of the data without using an external software like R.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

I mean this could be a one week work. Imean within one week you would be able to do your project, you know.

Document: P3\_GMT20220719.odt Tags: Real use case applicable

I really can't think of more features. I think less is more. Having what you have now is just enough.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

Most people are trying to understand how environmental factors like air pollution influence a person having a health event.

It has everything. Yeah, so I would recommend no additions to the features.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

Less is more. This is easy to follow and to get an idea of what you are doing.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

It provides definitions of everything.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

Now if I need to put in a dataset I can organise it this way and then.

Because most health datasets would have lat/lon, data and pseudonymised data of hospital admission, time they were diagnosed, time they experienced a relapse.

Document: P3\_GMT20220719.odt Tags: Real use case applicable

Having it in this format is ... And the lags is something that you ... So the length time to gather data, so it is something that you input yourself like two weeks before the event happened starting at lag 0.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

From 0 to 14, from 7 to 14 ... This is the one thing that might not be clear. The lag is clear to me but the length is something a bit unclear.

Document: P3\_GMT20220719.odt Tags: Some important features unclear

The rest is just ... I especially like the metadata bit because it is a lot of information which most people probably did not care to put together before uploading the data.

The information would be sitting somewhere in a DPIA.

Document: P3\_GMT20220719.odt Tags: Data protection importance

Then, this is nicely explained and within an area, the NUTS thing I cannot remember but it is about blocks but you can google that.

Document: P3\_GMT20220719.odt Tags: Some important features unclear

Temporal unit and aggregation method seems straight forward.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Oh I did not realise that, there you actually explain the NUTS ... Okay, okay. Then, that is just my bad for not reading.

You explain here what NUTS-1 -2 -3 mean. Okay, that is perfect, at a regional level.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear

Oh that is nice, that is a nice option especially if you want to map an average at a regional level.

Then, if you have 2 or 3 people you can get the average to address reidentification.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear, Data protection importance

For purposes of data analysis at least in my case you want to get as close to the individual as possible regarding location.

Having the regions is a nice second option to aggregate at the regional level to do some geographical comparisons.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear, Improved tool usability and output data

Oh, user friendly, easy to use, time saving hahaha like headache saving.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

Having multiple options, easy, playing a bit with it helps, having the data in the right format is important.

Document: P3\_GMT20220719.odt Tags: Data linkage process clear, Improved tool usability and output data

The event lat/lon location and date are almost always given in most hospital or medical records.

The lag and length I did not get. In my case, I just want to know from the month of diagnosis going 6 or 12 months back.

Document: P3\_GMT20220719.odt Tags: Real use case applicable

It is good that you get an average for the lag.

Document: P3\_GMT20220719.odt Tags: Report helpful

Overall I think it is a great experience, three steps, less is more, you have the information there.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

Everything is explained, you have links to places where you can check if you need more information.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

Only essential information is shown which is important.

Document: P3\_GMT20220719.odt Tags: Positive comments on text

Sometimes having extra information could send people through a rabbit hole trying to understand .

Document: P3\_GMT20220719.odt Tags: Improvement to text or feature

The website is just so simple and nicely designed. I like the three data formats you get the data with the specific lags that you wanted.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

Also the average for a particular month and then you have like data for multiple variables.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

In my case I could only do the linkage of a specific pollutant or climate variable at a time, and here you get everything in one go.

Document: P3\_GMT20220719.odt Tags: Output data ready for analysis

It is almost like magic. If you don't understand computers it is almost like magic, happening in one go hahaha Definitely, like from a user perspective it is user friendly and time saving.

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

I think it is also cost saving because once you save time in doing the linkage you can use this time to do other stuff.

Nice work, this is really nice!

Document: P3\_GMT20220719.odt Tags: Improved tool usability and output data

Okay that is nice. Can I go to the next part or I can complete task 1 in this window?

Document: P4\_GMT20220720.odt Tags: Experimental methodology

Okay, let me see. I can understand this part for now.

Document: P4\_GMT20220720.odt Tags: Positive comments on text

Okay, so I think this is the file you shared with me.

Oh okay, scrolling down I can find the rest of stuff.

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

Oh, so how can I upload the file? Oh okay, it is here.

So, upload data table and then I go to the linkage options

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

Okay, so I choose the minimum. Again do I need to import the data? Oh okay ...

I can export this table.

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

(The participant reads again the metadata upload instructions as they are not clear)

Document: P4\_GMT20220720.odt Tags: Some important features unclear

Oh okay, we have information here.

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

I can't understand the meaning of as input or season-alike, I cannot understand that completely.

Document: P4\_GMT20220720.odt Tags: Some important features unclear

I am not sure if it is related to the problem ... season-alike or as input ... Oh okay I see it now.

Interval of dates around the input date.

For input, I cannot select a special date. Oh I understand now, a date within the interval selected.

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

Okay, so distance in km from event point. Mmmm within a distance or area? I will go with "area".

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

Region ... I don't know about NUTS but Ican understand what is explained here.

Document: P4\_GMT20220720.odt Tags: Data linkage process clear

I want to choose days and ... mean, max ... Oh just one? Why is it not including more than one aggregation method?

A range would be useful as a data summary.

Document: P4\_GMT20220720.odt Tags: Suggestions for additional features

I hope I did not miss anything ... Okay, I will go up, this? ... Oh, right I forgot that ...

I must have missed something else ... No ...

Document: P4\_GMT20220720.odt Tags: Some important features unclear

MODERATOR: You have selected everything but it is not working. Would you mind refreshing the page and selecting the options quickly to see if it works?

Document: P4\_GMT20220720.odt Tags: Moderator guidance, Technical web issues

Yes, sure ...

It still does not work ... Mmmm ...

Document: P4\_GMT20220720.odt Tags: Technical web issues

MODERATOR: Let's try to choose different options to see if it works but in theory it should work. I am checking it on my end and it is working.

Document: P4\_GMT20220720.odt Tags: Moderator guidance

Oh, it could be related to my Internet connection settings ... Okay, let me see ... My internet is awful these days ... It takes a lot of time to do anything.

Document: P4\_GMT20220720.odt Tags: Technical web issues

Let me refresh the website again and let's hope it works. No ...

Document: P4\_GMT20220720.odt Tags: Technical web issues

MODERATOR: I think it is because of your firewall and Internet settings. Let's try another browser then.

Document: P4\_GMT20220720.odt Tags: Moderator guidance, Technical web issues

Okay, I have Firefox but I don't think that's it ... No it is not working with Firefox either ... Let me restart my Internet and computer.

Document: P4\_GMT20220720.odt Tags: Technical web issues

MODERATOR: Okay, let me send you the export that you would be getting and we can resume from task 2.

Document: P4\_GMT20220720.odt Tags: Moderator guidance, Technical web issues

I think I have finished task 1 and 2 then.

Document: P4\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: Oh, please read task 2 again.

Document: P4\_GMT20220720.odt Tags: Moderator guidance

Okay, all that great.

Document: P4\_GMT20220720.odt Tags: Report helpful

I do not remember what the "AVG" function was ...

Document: P4\_GMT20220720.odt Tags: Improvement to text or feature

Okay, I can click this and it is great.

Document: P4\_GMT20220720.odt Tags: Report helpful

Okay, great to have the sources and the process here. Nice.

Oh, nice it explains about the area that I have chosen in the linkage options.

Document: P4\_GMT20220720.odt Tags: Report helpful

Oh, nice result, I like this! Oh, that is so nice, I love this map. I love visualisations like this.

Document: P4\_GMT20220720.odt Tags: Report helpful

All the steps here are completely clear, great.

Document: P4\_GMT20220720.odt Tags: Report helpful

Okay, it is nice to see how this is working.

Document: P4\_GMT20220720.odt Tags: Report helpful

Nice graph to see the time windows for each event. I can select and deselect, which is also nice.

I can even zoom in and out ... Oh nice, I can save the pictures. That is nice!

I thought I could change the shape of this graph axis but that is not necessary.

Document: P4\_GMT20220720.odt Tags: Report helpful

This is just for events here, we cannot have other dates here not events ... So no visit times.

Document: P4\_GMT20220720.odt Tags: Output data unclear

The environmental variables seem good but I am not very familiar with all the options available for weather data.

I was wondering if we can have the information for every date not just for events, for every visit date.

Document: P4\_GMT20220720.odt Tags: Output data unclear

Okay, this is fine. It is complete.

I am curious about what a data cube is ... Oh, you have an explanation here which is good to have.

Document: P4\_GMT20220720.odt Tags: Positive comments on text

Your events do not mean exactly what I would call events. Let's try to go to the next part.

I think I missed something because I am missing the definition of event.

For me, an event means the time that was recorded when a flare happened.

Document: P4\_GMT20220720.odt Tags: Output data unclear

Okay I will filter the table for "B". Okay, nice.

Document: P4\_GMT20220720.odt Tags: Report helpful

Can I filter according to the variables of the table or only the event column?

It would be nice to filter according to variables as well, but this is great.

Document: P4\_GMT20220720.odt Tags: Suggestions for additional features

It is nice to have this explanation. It is nice to have a definition of everything here. It is okay.

Document: P4\_GMT20220720.odt Tags: Positive comments on text

It is nice that the plot is interactive. I remember the interactive plot from the previous version that was also nice.

That is nice.

Document: P4\_GMT20220720.odt Tags: Visualization of plots clear

Nice, so nice and beautiful.

Document: P4\_GMT20220720.odt Tags: Visualization of plots clear

Okay, nice ... Great, I think I have completed task 2.

Document: P4\_GMT20220720.odt Tags: Visualization of plots clear

What do you mean about discuss?

Document: P4\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: Just to talk about it.

Document: P4\_GMT20220720.odt Tags: Moderator guidance

It was very nice, I liked it. Especially the fact that the plot shows me a general view about the data.

Document: P4\_GMT20220720.odt Tags: Visualization of plots clear

This descriptive view allows me to have an idea of how the data looks like, and it gives me information about its background.

Document: P4\_GMT20220720.odt Tags: Report helpful

That is nice. I had a lot of problems preprocessing the data lately so by using this, the preprocessing is solved.

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

The problem is just that the meaning of events is not clear ... I don't know if it could be any time recorded event or only when a flare happened.

Document: P4\_GMT20220720.odt Tags: Output data unclear

It would be great if you could include all the clinical features I need for my analysis.

Document: P4\_GMT20220720.odt Tags: Real use case applicable, Data availability comment

That is because I need to consider both the environmental and clinical datasets, but I think it is fine.

Document: P4\_GMT20220720.odt Tags: Real use case applicable

It really made me more confident because linking the data is really hard and sometimes I am not sure if Ilink them correctly or not.

So it is nice that it does the linkage for you.

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

As I have explained already in task 3, the most important part is that I can't have the patient information linked to environmental variables for every visit time that I have on the clinical records.

This would be what I need. I would need biomarkers, blood tests, biopsys or others.

Document: P4\_GMT20220720.odt Tags: Data availability comment

That is fantastic, I really love it! That is really great!

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

If I ignore the fact that there is no clinical information or biomarker data, I can say that is fantastic.

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

Specially the plots, I love to visualise the data. Even though the data cannot be published sometimes, it makes it easy to understand.

Document: P4\_GMT20220720.odt Tags: Report helpful

It saves time and it is publishable.

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

I love this interface and congratulations!

Document: P4\_GMT20220720.odt Tags: Improved tool usability and output data

I have not really prepared anything so I will have to use the example.

Document: P5\_GMT20220720.odt Tags: Experimental methodology

I can rewrite the table I guess. I mean I don't know if I need to rewrite it or if I can use it as it is.

Document: P5\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: This is up to you, no problem, whatever works best for you.

Document: P5\_GMT20220720.odt Tags: Moderator guidance

MODERATOR: Would you mind thinking aloud while you complete the tasks?

Document: P5\_GMT20220720.odt Tags: Moderator guidance

I am actually reading the instructions. I can understand that you can select to conduct data analysis or deposit in an open data repository.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

If this is the case, you need explicit permission as per GDPR rules.

Document: P5\_GMT20220720.odt Tags: Data protection importance

Then you can have recommended or minimum metadata input, which is like just some data or larger file with recommended.

You can edit the example here using the metadata provided

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

When you click the export button you can export the metadata table. It has been downloaded to my computer.

I presume you can upload the metadata table with more examples here.

I am just opening the CSV file that has been downloaded to see how it looks like

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

Yes so now my impression would be to click upload metadata.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

Then, you can select the following options below. So, you can select the event dates as input or season-alike and I will go and read the explanation.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

I will go for the input option.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

Yeah, I like the distance option here where you can choose the diameter of the circle from the point.

Distance is km ...

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

Then you can select the temporal unit to be day, month or year.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

The aggregation method might be either mean, max, min ...

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

It seems like you can only have one of those formats per attempt (i.e. query).

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

I am like the clicking type of person so I would just click on "Link" here and I hope it is fine.

Let's try it, I guess.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

Yeah, so now the linkage process has been done to my understanding.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

I can export the output, which I did and I have it as a ZIP file in my computer.

Document: P5\_GMT20220720.odt Tags: Data linkage process clear

So, I am going back to the file. I am now seeing the SERDIF report but I'm not sure I am sharing the right screen.

Document: P5\_GMT20220720.odt Tags: Technical web issues

MODERATOR: Yes, all good Ican see the report.

Document: P5\_GMT20220720.odt Tags: Moderator guidance

Yes, that is something in the input field to search the data table based on my input.

When you type "A" you get event A, the same for event B.

Document: P5\_GMT20220720.odt Tags: Report helpful

I am not sure if I understood it correctly but it seems you cannot browse the whole table, only the events.

For example if I set temperature you will somehow filter that column but I am not sure.

Document: P5\_GMT20220720.odt Tags: Suggestions for additional features

So I clicked on the air parcel information box to check what it said.

It includes the environmental factors included here like weather and pollution ... particulate matter ... Then, also the sample variability.

Document: P5\_GMT20220720.odt Tags: Positive comments on text

So you can choose let's say solar radiation and it sets in the background.

I can just leave the pollutants I am interested in but Ican make them appear again. This makes sense.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

I am a bit confused with the boxes in the legend ... So I think I should click in here ... Okay, I like that.

Document: P5\_GMT20220720.odt Tags: Visualization of plots unclear

Yeah then, I understand. I was confused with the word box at first but now it seems straight forward.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

Then we have this other overview figure which is Z-scores.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

I would probably welcome something that could hide all the variables by one click and one click.

But I guess you are interested in all the data or you can use the source data but it is something that comes to my mind.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

I like the boxes now, it makes sense.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

You can also choose the background displayed in the event tab.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

If I select event C, I see wind speed by default but the background says "background-temperature", it may be by default.

This took me some time to get the idea that when I change the background it changes in the plot as well but by default it does not correlate.

Because you know, I got maximum temperature here but then I see wind speed there.

Document: P5\_GMT20220720.odt Tags: Visualization of plots unclear

It may be the case that the data has already been given but, does this mean that the data is from 12 days prior to the event?

That is how I would interpret that, but I'm not sure this is right. 0 would be the day of the event.

Document: P5\_GMT20220720.odt Tags: Output data unclear

Yeah, so I mean like my main impression is that you got the background like "maximum temperature" displayed on a colour scale like dark blue.

You have the concentration of whatever you are interested in and years, you could pose a research question on that.

I mean it is yes, nice.

Document: P5\_GMT20220720.odt Tags: Visualization of plots clear

I mean I can really imagine myself using this in research practice when you have the environmental data you are interested in.

Document: P5\_GMT20220720.odt Tags: Improved tool usability and output data, Real use case applicable

You could link it with the day of onset in vasculitis in our case.

I can imagine that you could select "months" even though I selected days.

Document: P5\_GMT20220720.odt Tags: Real use case applicable

I can imagine myself working with this type of data.

Document: P5\_GMT20220720.odt Tags: Improved tool usability and output data

Yes, I mean it is pretty straight forward for a clinician without any technical knowledge, as I am.

It should be fine.

Document: P5\_GMT20220720.odt Tags: Improved tool usability and output data

I am not really sure how much discussion I need to provide ...

Document: P5\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: It is up to you, if you feel you have shared what you wanted to share you can move to the next task. Unless you want to share more of your thoughts for this task.

Document: P5\_GMT20220720.odt Tags: Moderator guidance

Yeah, it took me some time to get the right impression from the whole plot but I am not sure if I understand all 100% but I'm getting there.

Document: P5\_GMT20220720.odt Tags: Visualization of plots unclear

t would be nice to play with real data because it is actually nice.

Document: P5\_GMT20220720.odt Tags: Real use case applicable

Yes so ... It has been interesting for me.

I can imagine myself linking real data with environmental exposures, for example in my country, with the disease event I am interested in.

Even selecting years would make sense.

Document: P5\_GMT20220720.odt Tags: Improved tool usability and output data, Real use case applicable

MODERATOR: Yes, please.

Document: P6\_GMT20220720.odt Tags: Moderator guidance

I think the example is clear and I will click the example output data.

Document: P6\_GMT20220720.odt Tags: Positive comments on text

Is this already a paper? Okay, so I will go back.

Document: P6\_GMT20220720.odt Tags: Output data unclear

These are only links for the websites from where the data is taken from.

Document: P6\_GMT20220720.odt Tags: Positive comments on text

Can I upload here the CSV that you sent me?

I will upload the CSV file. Okay, you can use the example directly here too.

So the next step is linkage options.

Document: P6\_GMT20220720.odt Tags: Data linkage process clear

What does recommended and minimum metadata mean? Ah okay it is here.

Document: P6\_GMT20220720.odt Tags: Positive comments on text

Maybe here, when you have information in a button it is more intuitive if you add a "(i)" symbol. Otherwise you might skip that.

I think it would be useful to add the "(i)" symbol in the buttons with tooltips.

Document: P6\_GMT20220720.odt Tags: Improvement to text or feature

I don't know what Ihave to do here because I only have one CSV file, right? So ... I might click upload metadata.

Document: P6\_GMT20220720.odt Tags: Some important features unclear

I understand the first step but the second one confuses me a bit.

I don't understand if I can select the different tables and upload them individually or ...

Document: P6\_GMT20220720.odt Tags: Some important features unclear

Do all the tables that appear here get uploaded when I click the upload button?

I don't know what kind of ... If I upload a file I don't know which ... Like the different fields I need to fill in.

Different tables have different values.

Document: P6\_GMT20220720.odt Tags: Some important features unclear

I prefer to choose distance because even though I had read about NUTS, but I don't know the exact meaning.

Document: P6\_GMT20220720.odt Tags: Some important features unclear

Okay, so multiples of 10km ...

Document: P6\_GMT20220720.odt Tags: Data linkage process clear

Temporal unit: month ... then, the mean.

Document: P6\_GMT20220720.odt Tags: Data linkage process clear

MODERATOR: The file that you uploaded in step 1 only has a few days, so if you select months you will only get a single value.

Document: P6\_GMT20220720.odt Tags: Moderator guidance, Some important features unclear

Oh, okay so days then, that is true, thanks.

Document: P6\_GMT20220720.odt Tags: Some important features unclear

Yeah, it would be great that when you upload the file you could see like the head of the table here.

I don't know if it is like this or not now, but maybe it is this one, the one that you sent.

Document: P6\_GMT20220720.odt Tags: Suggestions for additional features

I will wait for the output ...

(The participant reads the step 3 linkage success banner)

Document: P6\_GMT20220720.odt Tags: Data linkage process clear

The interactive report ... Okay, so export to downloads ... I opened the ZIP file ... I got an error in my computer ...

Document: P6\_GMT20220720.odt Tags: Output data unclear

MODERATOR: Oh, you can click on skip since you already have everything there.

Document: P6\_GMT20220720.odt Tags: Moderator guidance

I guess this is like a summary of all the files included in the ZIP file.

Document: P6\_GMT20220720.odt Tags: Output data ready for analysis

Is it only to explore this report? Okay ...

Can I go on to task 3?

Document: P6\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: Maybe you could explore the report in more detail. For example, by reading a couple of paragraphs and interacting with it.

Document: P6\_GMT20220720.odt Tags: Moderator guidance

You could add an index in the beginning to go directly to the section that you need.

This feature could be useful for more experienced people that might want to skip the introduction and go directly to the data exploration and resulting figures.

Document: P6\_GMT20220720.odt Tags: Improvement to text or feature

I think I need to figure out the difference between events A, B, C and D.

Maybe I need to read everything to understand it ... I'm not sure.

Document: P6\_GMT20220720.odt Tags: Output data unclear

Oh here, this is the thing I meant in the beginning with the index to navigate to specific sections.

Document: P6\_GMT20220720.odt Tags: Report helpful

Okay I will try to search in the table, this is interesting in case the categories are spelled wrong.

Document: P6\_GMT20220720.odt Tags: Report helpful

Okay, here you have the index to go directly to what you want.

Document: P6\_GMT20220720.odt Tags: Report helpful

Why is it different when you click 2 times the legend? Is there something wrong here? Maybe not ...

I am trying to understand this ... Okay, some boxes appear or disappear ... It would be good that you can see all at the same time when they are activated.

Because if not I do not know what I'm looking at when I have all of them active.

Document: P6\_GMT20220720.odt Tags: Visualization of plots unclear

Maybe, you can show the first figure from the beginning and then you click on the other figures to change them.

Document: P6\_GMT20220720.odt Tags: Improvement to text or feature

Yes, I think that I am going to look at the CSV files in the ZIP file.

Yes, Ithink it would be very useful if this tool can do the linkage for you and you don't have to go to each webpage and get the data that you need.

Document: P6\_GMT20220720.odt Tags: Improved tool usability and output data

This can do it for you. If the data sources are from Copernicus and EEA I think they are very reliable sources so Ithink I would be confident in using this tool.

Document: P6\_GMT20220720.odt Tags: Improved tool usability and output data

For me step 2 is difficult to understand ...

Document: P6\_GMT20220720.odt Tags: Some important features unclear

I thought that you could link your environmental data with respiratory or cardiovascular diseases.

So do you have access to health data or do we have to upload it? That is what I am thinking.

Would this interface include the epidemiological records for different diseases in different countries?

Document: P6\_GMT20220720.odt Tags: Data availability comment

So then maybe if you need to upload the health data you need to explain how the input events file should look like.

Maybe the information is already there but I am not sure where to find it.

Document: P6\_GMT20220720.odt Tags: Data availability comment

Besides step 2, the other steps are clear.

Document: P6\_GMT20220720.odt Tags: Data linkage process clear, Some important features unclear

I do not have a lot of expertise in this but I think it is quite easy to use and user-friendly.

Document: P6\_GMT20220720.odt Tags: Improved tool usability and output data

So I think everybody can do it. If you read all the information in detail I think you can understand everything but I did not have a lot of time today.

Document: P6\_GMT20220720.odt Tags: Improved tool usability and output data

We have some nice little graphics here with a cross-section of the country with the time series of yeah some environmental data. That is cool.

Document: P7\_GMT20220715.odt Tags: Positive comments on text

Okay that is fine.

Document: P7\_GMT20220715.odt Tags: Positive comments on text

Okay so event A, these are two different events. These could be like the flare events or whatever health outcome.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

Oh sorry, it was the same flare event ... Oh yes it is the same flare event but at different lags. Then, these are all the environmental data.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

Okay, so we can upload our own data here or we can edit the table below by clicking on the cell and so on.

Let me see ... by clicking ... Okay ... So event A, longitude, latitude ... Let's just keep it as it is as I am not 100% sure what to change.

So Iunderstand what this is ... the name of the event, longitude, latitude ...

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

What does this export button do? Let me see ... Okay, yeah.

This is how the data needs to look like if we were to input the data ourselves.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Okay, that is fine, the upload was successful.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Okay, let me see ... This is a specific longitude and latitude point, what if you just wanted the square region here or something like that?

I will keep going and try to understand it later.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

Okay so it is just how you get the metadata.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Sorry, let me check what I was supposed to do again. Oh, okay.

Document: P7\_GMT20220715.odt Tags: Experimental methodology

This is just metadata stuff.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Let's just leave it as it is with input.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Let's just select distance ... Oh, okay okay, I see.

So we have this coordinate in the input and ...

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Let's try to click on the "area".

I have no idea what NUTS means but let's check the information link here.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

I guess it lets you pick ... Okay, NUTS-2 for all these countries.

I am not 100% sure what this is but it would probably not be that bad if I spend some time looking at it.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Let's just go with the circle. Radius of 10km ... Let's make that 20km.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Temporal unit let's say day and mean for the aggregation method.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

I take it that it draws a circle around the events at this date and then I ask for the temporal unit.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

Going through this in order now, the thing that confused me is that the lat/lon point was a very specific coordinate.

Maybe it is because I read it too fast but step 1 left me thinking, if I want a specific region what do I do?

I found out later that you could by drawing a circle around that region. That makes sense.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

Anyways, let's just link ... Oh wait, what am I missing?

I forgot to upload the metadata there, okay.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

Now I can link it, okay. It makes sense.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

I think so ... If I was doing it a second time, now that I have looked at the whole thing, it makes a lot more sense now and it would be faster and easier.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

It takes a few seconds and then I can export the output.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

I am not sure what a TTL file is but let's just open these files up.

Let's make this a bit bigger. I am not sure what TTL file is ...

Document: P7\_GMT20220715.odt Tags: Output data unclear

These are the events with all the environmental stuff.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

That is interesting. That is fine, it makes sense that different regions have different data but these two events should be in the same region.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

There is a little bit of inconsistency there because they are in different months.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

Okay, so this CSV file is marked as sdev so it must be the standard deviation of the data I guess.

Document: P7\_GMT20220715.odt Tags: Output data unclear

I don't know what "normals" means from the top of my head but it probably explains it somewhere.

Document: P7\_GMT20220715.odt Tags: Output data unclear

Let's have a look at the data. Okay, so this is the real data, that is fine.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

The CSV file marked as sdev is the standard deviation of the data I guess.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

The CSV file marked as normals could be the z-statistic but I am not sure.

Document: P7\_GMT20220715.odt Tags: Output data unclear

Okay, so this is the interactive report ... What do I need to do? Let me go back to the tasks.

Document: P7\_GMT20220715.odt Tags: Experimental methodology

So we have some metadata here with different tabs I can explore.

Okay, this is just a dictionary of the data and the input data.

Document: P7\_GMT20220715.odt Tags: Report helpful

Processing steps ... Okay, so this is to check if I have done everything right and I can also check the input again.

Document: P7\_GMT20220715.odt Tags: Report helpful

This is the step by step linkage process which is fine. I will skip ahead.

Document: P7\_GMT20220715.odt Tags: Report helpful

Okay, so this explains what the files I was looking at.

It might be better if you call them "mean" because "normal" can mean a few things.

Okay, so this CSV file is about the means.

Document: P7\_GMT20220715.odt Tags: Output data unclear

Maybe I am misunderstanding this ... So I don't see "AsInPM10" in the raw data, it is only in the sdev and normal files.

There is PM10 here ... Okay ... There seems to be some inconsistency in the columns unless I'm misunderstanding something.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

That could be a glitch or something I'm misunderstanding. I will assume these are correct.

Event B and C ... There is a value for C here but not for B ... Is this the same? Let me check the values.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

I have July and August for event B monthly value but there is no daily data ... Again I am not sure if there is a mistake there or if I am misunderstanding something.

Because my understanding is that this is the aggregated data I asked for and the normals and sdev should be a further aggregated dataset, which is telling me the average should be ...

I am not sure about this, this would be my understanding of it. I am not sure.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

Maybe the explanation about the normals is what I am missing here, and then about the area mentioned.

Document: P7\_GMT20220715.odt Tags: Improvement to text or feature

The variables are all links. Let's check the links out of interest.

Okay they link to Copernicus, which is good but for now I am not going to read all of that.

It is good the links have some information there.

Document: P7\_GMT20220715.odt Tags: Positive comments on text

Data sharing ... Okay, this is all fine.

Document: P7\_GMT20220715.odt Tags: Report helpful

Okay, so here we have the data table ... With the events ID: A, B, C and D ... The dates and lags ... One, two, three days before the event ...

The other columns are the covariates, the Z-scores for each ...

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

Okay, the data table has the raw values but it is colour coded by whether it is above or below two standard deviations.

That is pretty handy, it is fine.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

You might be able to see some trends for some outlier events ... I don't see trends here though, at least related to statistically significant values ...

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

Let's type B in the filter input ... It does not do anything ... Maybe I need to put it into inverted commas ... Oh, wait.

So I thought it would just isolate the B events. Oh for its column values ... If I type CO what happens ... What if I type A?

Document: P7\_GMT20220715.odt Tags: Suggestions for additional features

Oh yeah it works, maybe I did not notice that before. That makes sense, this is filtering out whatever values I input.

Document: P7\_GMT20220715.odt Tags: Report helpful

I want to see if Ican type more than 2 ... Oh, it does not work, it is fine as it is only a simple way to isolate the events, that's fine.

Document: P7\_GMT20220715.odt Tags: Suggestions for additional features

Filtering is handy.

Document: P7\_GMT20220715.odt Tags: Report helpful

Okay, if I wanted I could click these buttons to get more information about the background of the data but I am not going to read through all of it.

Oh, actually I might get extra information about the data as I go along.

Document: P7\_GMT20220715.odt Tags: Report helpful

So the background button, this is kind of handy. Let's look at the plot first.

The plot presents the concentration of pollutants related to these events.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

Maybe I can't see it ... Oh these are the boxes ... This button changes the weather variable in the background.

The colour bar tells us the scale of the colours for the weather variables.

We can remove the pollutants as we see or need to.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

Okay, let me have a look again ... Okay, so these lines are the mean values per lag and the boxplots show us the variability.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

I did not see that at first.

Document: P7\_GMT20220715.odt Tags: Visualization of plots unclear

For example here, we only have a line meaning that there is probably only a single value that corresponds to these means.

Then we see boxes that show the variability, this is pretty handy.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

Let's say I want to look at PM2.5 or PM10 ... Then, I have PM10. Okay that's pretty cool, great.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

The z-score figure has two horizontal lines which are sort of like confidence intervals, actually +2 and -2 sdev for a specific variable.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

The plot works the same way ... I can click on the boxes to display the boxplots as before.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

Okay, this is for each event, let's try A.

Okay, B ... Then of course these events are ...

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

We asked for a 0 day lag here but there we have the 8 lag so these values must have been specified from the start.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

I think I did not pay much attention at the start. Yeah, okay we started with a 7 day lag, okay.

That is why it is only displaying starting from the 7th day lag.

Document: P7\_GMT20220715.odt Tags: Visualization of plots clear

It all seems pretty cool.

Document: P7\_GMT20220715.odt Tags: Report helpful

The data table seems very handy in particular, I like this.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

It seems pretty easy to use and very intuitive, I think.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

For my purposes as a statistician I would just say that I would probably draw my own plots with the dataset instead of using the plots provided.

I would probably be more comfortable doing that.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

But this can be very good for medical researchers who do not have experience drawing plots. So that is good.

Document: P7\_GMT20220715.odt Tags: Report helpful

I would think so but let's have a look at the data.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

The normals and sdev variables might not be that useful to me because I can compute them myself.

Document: P7\_GMT20220715.odt Tags: Output data unclear

The raw table is fine. It is in wide format which is the format I would like it in.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

There are a lot of missing values but I think this is not something you can fix. This is just how the data is. which is unfortunate.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

In terms of the structure of the data, it all seems pretty good.

I don't think I have any issues here.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

Again, I would not use the sdev and normals anyway if I was going to do an analysis.

Document: P7\_GMT20220715.odt Tags: Output data unclear

I am still not sure what it actually means as there seems to be some inconsistency with column headers.

It seems there are some inconsistencies of what is available and what not.

There were no values for that variable but we still have it in the aggregates.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

For example, there are no C6H6 values here but we have them in the normals.

I am not sure what is going on there or if there is a glitch.

Document: P7\_GMT20220715.odt Tags: CSV output with missing data

Otherwise, the actual data is good though. This is the main thing at least for my purposes so I am happy with this.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

For my purposes, I don't think so. It is one of those things that if I actually sat down to do an analysis, I might think of things as I come to it.

I think this is good, it is all I need.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

The plots are a nice quick way to get a quick look at the data. It is nice to have it here anyway to quickly look at things.

Document: P7\_GMT20220715.odt Tags: Report helpful

Even though I would draw my own plots.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

I am trying to think if there would be another plot I would do ... No, this is fine to be honest

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

I think this is really good. This is all very helpful and it seems very complete to me.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis, Report helpful

You have all this information here and the links for further information if you wanted.

It would be a big thing to sit through but it is good that it is there.

Document: P7\_GMT20220715.odt Tags: Output data ready for analysis

You have all the proper warnings and all.

Everything seems very transparent and organised, which I like a lot.

Document: P7\_GMT20220715.odt Tags: Report helpful

The data is in the format I need.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

The figures are nice to have just as a quick look at the data.

Document: P7\_GMT20220715.odt Tags: Report helpful

Contact information is great to have in case you have any questions.

Document: P7\_GMT20220715.odt Tags: Report helpful

I do not have anything against this. It all seems very good.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

Linking it was ... I struggled a little bit because I didn't quite understand the process at first.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

What am Isupposed to upload here? Because it seemed very specific at first.

I don't want specific coordinates, I want a region or a country.

Document: P7\_GMT20220715.odt Tags: Some important features unclear

Obviously, when I got to that point I realised I could draw a circle around that event and pick this country with the NUTS code.

Document: P7\_GMT20220715.odt Tags: Data linkage process clear

If I was to use this after the first time, Ithink this would be very easy to use.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

It is just that I stumbeled over the first time I used it but it seems very good.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data

Other than the normals and sdev inconsistency with the raw data, which I still cannot understand, the rest seems fine.

Document: P7\_GMT20220715.odt Tags: Improved tool usability and output data, CSV output with missing data

There is the file you sent me earlier in the email which I have to do. I will save it on the desktop.

Document: P8\_GMT20220720.odt Tags: Experimental methodology

Sorry, this was me not being prepared as I should have been.

Document: P8\_GMT20220720.odt Tags: Experimental methodology

MODERATOR: You could use the example in step 1 if it is easier for you. The data is already there.

Document: P8\_GMT20220720.odt Tags: Moderator guidance

Okay, I am happy to use that one. That makes more sense.

Okay, 4 events then.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

Okay, we can do data analysis.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

I am going to have recommended metadata.

I am a bit lost here, okay so we have an event ... I do not understand terribly well.

I will upload metadata anyway, why not? Okay ...

Document: P8\_GMT20220720.odt Tags: Some important features unclear

I will go the distance in km and 10km seems fine to me.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

I am interested in day to day environmental data.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

I am going to go with max for now ... Okay.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

Now we are going to export it as CSV and TTL ... Okay that is the end of task 1.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

Okay, so I have to go to the ZIP file and extract the files, hopefully this is the intelligent thing to do.

Document: P8\_GMT20220720.odt Tags: Data linkage process clear

We don't have to worry about personal data here because it is fake data, which is cool.

Document: P8\_GMT20220720.odt Tags: Data protection importance

This is the metadata about my data ... I think last time I got to choose if I wanted air pollution or weather ... Oh wait this is everything.

That seems reasonable.

It has a licence ... temporal resolution ... That all seems perfectly reasonable.

Document: P8\_GMT20220720.odt Tags: Report helpful

Let me go back to the tasks ...

Document: P8\_GMT20220720.odt Tags: Experimental methodology

So data which is probably more important ... Okay ...

There is a table and also some charts.

Event data and lag ...

Event A, we have all these environmental variables.

Now, interestingly I selected "max" as my aggregating method ...

Document: P8\_GMT20220720.odt Tags: Output data ready for analysis

The red are high values and the blue normally low values.

I am slightly confused because this site does not seem to be smaller than the other one ...

It could be because I am not a statistician.

Red values are big if I understood them correctly so ... But that is not the case.

Okay, this looks like a small value and not a big one.

There is something I do not totally understand there.

Document: P8\_GMT20220720.odt Tags: Output data unclear

I do not remember setting the lag or length, perhaps I had the opportunity but I went for defaults.

Interestingly, these are days going backwards from my event date with a lag 0.

Document: P8\_GMT20220720.odt Tags: Output data unclear

Some values don't exist prior to 13 days back from the event date and even more they don't exist for other periods.

However, we have all these values for other variables like CO and others.

Document: P8\_GMT20220720.odt Tags: Output data unclear

The whole period of interest is there.

Document: P8\_GMT20220720.odt Tags: Output data ready for analysis

Okay, so we have a filter functionality here for the table. B ... Mmmm that may refer to the column, I will try it and see what happens.

Okay, but there are no letters in any column so that does not work terribly well.

Document: P8\_GMT20220720.odt Tags: Suggestions for additional features

Search the table for its column values ... Let's try NO, nothing comes up because there are no columns with the value NO in it.

It does not seem very interesting to type in numbers and search for specific numbers.

Document: P8\_GMT20220720.odt Tags: Suggestions for additional features

But I guess we could look for events ... Ah, I have it now! I understand what is going on here.

Document: P8\_GMT20220720.odt Tags: Report helpful

I did not notice multiple events before, I understand it better now.

Document: P8\_GMT20220720.odt Tags: Output data ready for analysis

You have data for 12 days prior to event A but there is no data for event B for these variables because it was a different time around August 2011.

Even less data for events C and D. This is the only explanation I can think of for the blank values there.

Document: P8\_GMT20220720.odt Tags: CSV output with missing data

I now understand what this "B" thing is for.

Document: P8\_GMT20220720.odt Tags: Report helpful

I will go with event A as it is the most interesting to me. I have already explained that I don't understand the blue and the red terribly well.

Document: P8\_GMT20220720.odt Tags: Output data unclear

I could look at this to understand if there is something going on. At the end they are just numbers so who cares about numbers, let's move on.

Document: P8\_GMT20220720.odt Tags: Output data unclear

I am still interested in what happened with my "max" selection but I can worry about that later.

Document: P8\_GMT20220720.odt Tags: Some important features unclear

Okay, so you have averaged the result across the 4 events A, B, C and D.

Except of course that some of these things do not exist for some events.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

Some of these purple ones might be better ... I wonder if you can turn the variables on and off ...

I am tempted to click but ... Okay, so I thought the NO variables were the most interesting because they appear for all 4.

I am going to leave all the ones with nitrogen ... Oh, I did not do that right I don't think.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

I am looking at different nitrogen pollutants. I don't know what the background means, let's try to change that.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

Mmmm ... I don't know what that means, let me read the information here ...

Document: P8\_GMT20220720.odt Tags: Visualization of plots unclear

It looks to me that what is underneath is ... Okay I will try to change it ... Oh yeah, it is the background colour.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

Let me change to something that means something to me like solar radiation.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

I am wondering why I would want to pick one of those and put it behind but I am not an environmental researcher.

It could be that I want that or more than one. I am not sure why you have done it this way.

It seems that you have preferred pollution. I have a lot of control over the pollution and I can do lots of stuff for them.

But it seems that I can only have one weather related variable at a time for whatever reason.

Document: P8\_GMT20220720.odt Tags: Researcher's expertise

Apparently solar radiation is high for this particular period here.

The average solar radiation across all 4 events at each lag, if I understood correctly.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

I chose the NO because I had a reason for that. I am going to go here and choose the same again.

I could have included O3 also so I can prove I understand again.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

Okay, now I can look at this plot and say if there is something interesting happening in relation to these variables at a time that might be relevant to my event.

If I was doing that I would probably ... This variable is the grey one ... There are medians ... My statistics are not grey so I don't know what could be interesting here.

If I have understood correctly .... Okay, so this particular value of NO is the average of the 4 events 21 days prior to the lag prior to the event which is 0 in this case.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

Why don't I have some of these plots with data going back in time?

There is probably a reason for that but I cannot think of what it is.

Oh, okay I have solar radiation going back there for longer lags.

Document: P8\_GMT20220720.odt Tags: Visualization of plots unclear

Now, we are completely outside my comfort zone because I am not used to working with z-scores believe it or not.

I could think about it for a long time and get it but let's go on.

Document: P8\_GMT20220720.odt Tags: Researcher's expertise

It is interesting that the solar radiation is 2 sdev and before we had the values.

Z-scores are about how comfortable we are in the variability of the data.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

If I cared, Icare but not really.

Document: P8\_GMT20220720.odt Tags: Researcher's expertise

This is a single event and I am looking at event A which was the one with more data.

We have all the data. NO looks the most volatile.

I presume you can change them on and off if I wanted.

Yeah, okay.

Document: P8\_GMT20220720.odt Tags: Visualization of plots clear

AS Imade clear, Iam not an environmental researcher so I would not know what to do with this data really.

Document: P8\_GMT20220720.odt Tags: Researcher's expertise

How would I work out if there is something interesting about this data ... I think I would have to export the data.

I would use the data for some type of statistical analysis asking some of my colleagues.

Document: P8\_GMT20220720.odt Tags: Real use case applicable

All of the data is available, which is a way of looking at it.

The data exists in CSV files and everything else which I can import to the analysis.

Document: P8\_GMT20220720.odt Tags: Improved tool usability and output data

I don't think I would. All the data is there but I am not a statistician so Icannot judge that aspect.

Document: P8\_GMT20220720.odt Tags: Output data ready for analysis, Researcher's expertise

we went faster as you would probably have expected.

Document: P8\_GMT20220720.odt Tags: Improved tool usability and output data

What did I miss? I still don't understand what happens when I pick "maximum". I don't know how this affected my data.

Document: P8\_GMT20220720.odt Tags: Some important features unclear

I am still confused by the red and the blue. I am probably not getting that.

Document: P8\_GMT20220720.odt Tags: Output data unclear

I could do it all as I went along.

Document: P8\_GMT20220720.odt Tags: Improved tool usability and output data

It was definitely simpler than the last time.

Document: P8\_GMT20220720.odt Tags: Improved tool usability and output data

I definitely like the diagram because that confused me last time significantly.

Document: P8\_GMT20220720.odt Tags: Positive comments on text

When it came to linking, I started to get a bit confused ... Metadata is fine ... I have all these options.

Document: P8\_GMT20220720.odt Tags: Some important features unclear

I don't know what happened when I selected max ...

Oh maybe the aggregation method ...

Document: P8\_GMT20220720.odt Tags: Some important features unclear

Maybe the aggregation would be at the temporal scale like for months.

Document: P8\_GMT20220720.odt Tags: Some important features unclear

Oh, where do I upload the file? Oh step 1, it seems here.

So … “upload data table”, alright.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I want to conduct data analysis and I will select minimum metadata.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

Upload metadata ... I will do that.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I am not not sure what the NUTS areas mean. I will go for NUTS-3.

Document: P9\_GMT20220721.odt Tags: Some important features unclear

I could look it up here and check what they are. Okay, very good.

NUTS-3 are all of these regions, I might change it to NUTS-2 then.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I will select a year for the temporal unit and I am going to go for the minimum for the aggregation method.

Document: P9\_GMT20220721.odt Tags: Some important features unclear

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "year" as the temporal unit you will only get a single data point for each event.

Document: P9\_GMT20220721.odt Tags: Moderator guidance, Some important features unclear

Okay, I will go with days then.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

So far I'm managing to get through. I like the layout the way it is.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I can go from part to part and I can get instructions of what I am doing as I go.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I am going just to take a look at this. Okay, this is the data ... I have chosen what I wanted ...

I am wondering why I got all of that data but I guess it is because I just want to do a data analysis.

Well, I linked it anyway which was my first task to link it all together.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear, Output data ready for analysis

I got the raw data file but I got this nice html file that has been put all together for me.

Document: P9\_GMT20220721.odt Tags: Report helpful

I won't know what data I am looking at in this table but I can see that if I knew what the data meant I could go and get it.

Document: P9\_GMT20220721.odt Tags: Output data unclear

The thing I like about the report is that it was very quick for me to get this report and it looks like a very comprehensive page that I am getting.

Document: P9\_GMT20220721.odt Tags: Report helpful

I can access the data table and the exploration etc.

I presume that if I go through these parts in the index, it would bring me to the specific section so that I can get to different parts.

Document: P9\_GMT20220721.odt Tags: Report helpful

Okay, I want to push some of those things here ... The information in the boxes here tells me more about air parcels.

A huge amount of data has been created for me that I can make sense of all of the data that is being shown in the charts.

Document: P9\_GMT20220721.odt Tags: Visualization of plots clear

What is the difference between all of these plots ... I am wondering what these could be telling me.

Should I be able to see all of them together or maybe that is what "A" is?

Why do I need to see them all separately? That is what is coming into my mind.

Document: P9\_GMT20220721.odt Tags: Visualization of plots unclear

I can also check the contact information and send an email if I have any difficulties.

Document: P9\_GMT20220721.odt Tags: Report helpful

So do you need me to explore anymore? There is lots of information here of what Ican do and see. I can see the licence and title here.

Document: P9\_GMT20220721.odt Tags: Experimental methodology

MODERATOR: If you don't mind, you could explore a bit more in detail in the metadata section of the report.

Document: P9\_GMT20220721.odt Tags: Moderator guidance

Oh, I did not spot that I could go in there to see this information.

Document: P9\_GMT20220721.odt Tags: Complex report content

Input data tells me what you did, data structure, data sharing ... Okay.

Document: P9\_GMT20220721.odt Tags: Report helpful

The processing steps ... Oh very good, there is lots of data in here that I would have missed.

Why did I not see that?

Document: P9\_GMT20220721.odt Tags: Report helpful, Complex report content

I am wondering about each category ... When I click metadata in the index it brings me here ...

I am wondering if there should be tabs instead of the buttons in a menu.

Document: P9\_GMT20220721.odt Tags: Improvement to text or feature

When I go down to data ... I see the tables here.

When Igot down here I got the same mechanism here.

Document: P9\_GMT20220721.odt Tags: Improvement to text or feature

I have spotted that I can search within the data table.

I will search for "2.5", It allows me to get information there.

Document: P9\_GMT20220721.odt Tags: Report helpful

My overall statement is that there is a lot of data, all very compactly and quickly created.

Document: P9\_GMT20220721.odt Tags: Output data ready for analysis

I also got the raw data in the spreadsheets if I wanted to take a look myself.

Document: P9\_GMT20220721.odt Tags: Output data ready for analysis

All the data is in the zip file which I could go through if I wanted.

Document: P9\_GMT20220721.odt Tags: Output data ready for analysis

Yes, I think my view on that would be that I have now got two tools:

(1) that is allowing me to make my criteria for my selection known and a very usable interface for me to allow me to do that,

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

(2) I have now got a report that is very interactive that I can drill down on to see the result of those criteria I put in for my data.

Document: P9\_GMT20220721.odt Tags: Report helpful

I can search for it. I can understand the context. There is a lot of information about the context, the licence and the anonymisation, the raw data to back that up.

Document: P9\_GMT20220721.odt Tags: Report helpful

I could go back and change the constraints if I wanted to. Maybe I created something odd, maybe I did not have the constraints right but it was quite quick to change the constraints and redo that.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

What I also could notice and I hope this is true, if I want to run another analysis I presume I get another webpage report.

I could start comparing different analyses with different constraints.

Document: P9\_GMT20220721.odt Tags: Report helpful

What I really like and I would be confident about is that if I was to write a paper or something about this.

I actually have that zip file to go back and check the data.

Document: P9\_GMT20220721.odt Tags: Output data ready for analysis

You could imagine that for each publication that you would do you could use that as your artefact because some journals would ask for the data to supply with it.

Maybe it is in an anonymised state or a state that we might be able to publish if that is appropriate.

Document: P9\_GMT20220721.odt Tags: Data protection importance

I think yes, I would be confident that I can use it for my research.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

The overall linkage process diagram looks like buttons to me.

Document: P9\_GMT20220721.odt Tags: Improvement to text or feature

When I first came in here first, I thought I might have to do these things first.

I did not ally notice that this was my step 1 as it was a bit far down the page.

The diagram does not look like a process, maybe an arrow or make it look more like a diagram.

Document: P9\_GMT20220721.odt Tags: Improvement to text or feature

I think these three steps are actually quite straightforward.

Document: P9\_GMT20220721.odt Tags: Data linkage process clear

I presume that the people that would be doing the analysis with this data would be able to choose the more appropriate one.

I was just picking one but you have the description there.

If it was my own data I would be able to choose which one to do.

I think this part of the website, once I got started, I was quite able to get going.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data, Real use case applicable

I can't think of any additional features and it is not like me.

You must have covered a lot of things.

Document: P9\_GMT20220721.odt Tags: Output data ready for analysis

My overall experience is that it is very smooth, clear, usable and it is really useful.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

I think that is the most important part of it.

I can see the output being very useful.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

Even if the interface became more complicated with more options because it is so useful I would be willing to put the extra time on this to make those choices, taking my time.

I think that's well done. It is really really smooth and useful.

Document: P9\_GMT20220721.odt Tags: Improved tool usability and output data

Okay, so ... Drag and drop is not working ... I will have to select it manually.

I am not sure if the manual thing works either. This could be a security setting in my browser perhaps or ...

Document: P10\_GMT20220802.odt Tags: Some important features unclear

MODERATOR: We could try to refresh the website and try again. It could also be a security setting from your browser because it is working for me on my laptop.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Okay, I will try to change that now. I will try it on a different browser.

I might try it on Chrome and see if it works there. Chrome is not working either I'm afraid.

I will try Edge but it is not working either.

Document: P10\_GMT20220802.odt Tags: Technical web issues

MODERATOR: I have Firefox working here for me, I don't know what could be happening.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Yeah, it is not working.

Document: P10\_GMT20220802.odt Tags: Technical web issues

MODERATOR: I would advise to go ahead with the example data and see if we can make it until the end.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Okay, I will use the example data then.

There, okay, I have added the events manually.

Document: P10\_GMT20220802.odt Tags: Data linkage process clear

Nice pop up there.

Document: P10\_GMT20220802.odt Tags: Positive comments on text

I am not sure where I will be uploading metadata but okay

Document: P10\_GMT20220802.odt Tags: Some important features unclear

Okay this is the type of metadata that you require.

Document: P10\_GMT20220802.odt Tags: Data linkage process clear

The recommended metadata is very extensive.

It is hard to imagine researchers trying to satisfy that.

For example, right class and right comment, maybe we would need some explanation about what they are.

Some are not accessible like dpv:[...].

Okay, that is pretty hardcore.

Document: P10\_GMT20220802.odt Tags: Some important features unclear

I don't know what season-alike events mean. Okay.

Document: P10\_GMT20220802.odt Tags: Some important features unclear

The sinta of this sentence is not great. It needs to be corrected.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

I don't think it is possible to assign a single aggregation method for all variables.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

We will go within an area region level NUTS-3. I will check here what it means.

Document: P10\_GMT20220802.odt Tags: Some important features unclear

Eurostat website ... Maybe you could add a couple of lines about what NUTS are.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

Excellent ... NUTS-0 ... I don't know if you can see my screen now.

For each NUTS level, you are getting increased granularity.

They got these at the national level and then 1 2 3.

They have assigned a random ... Okay we will go with NUTS-3.

Document: P10\_GMT20220802.odt Tags: Data linkage process clear

Let's go with “month”.

Document: P10\_GMT20220802.odt Tags: Some important features unclear

Aggregation method you should include min to max.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

MODERATOR: It is not working as you are not getting a spinning sign. The dates need to be without the slash but with the dash between year month and day.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

MODERATOR: You would have to refresh and copy paste the event data again.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Okay, like that.

Document: P10\_GMT20220802.odt Tags: Technical web issues

I think to be usable you need to add a checker to the events input to get an error message.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

We are spinning here but I don't know how long it takes to spin.

I had chosen up to 2020 as the data available was up until 2020.

This is still spinning ... What are the parameters that determine the performance?

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

MODERATOR: No it did not work ... Let's try to link again.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

It is a small dataset with only 12 events, NUTS-3 is as small as possible ...

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "year" as the temporal unit you will only get a single data point for each event.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

I will try days then.

Document: P10\_GMT20220802.odt Tags: Some important features unclear

It could be because I am using my own data.

Document: P10\_GMT20220802.odt Tags: Technical web issues

MODERATOR: Still not working with these settings.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

I will just refresh the page and use the example data.

Document: P10\_GMT20220802.odt Tags: Technical web issues

Could you provide me with an example of when you would use the sum?

Document: P10\_GMT20220802.odt Tags: Some important features unclear

MODERATOR: An example would be if you are interested in cumulative pollution, if you want to study rain or particularities of a time series that the mean would smooth.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Yeah, that works so the issue is that custom data does not work.

I could send you the link to the dataset and you can try later.

Document: P10\_GMT20220802.odt Tags: Technical web issues

Okay, so I can export it now.

Document: P10\_GMT20220802.odt Tags: Data linkage process clear

I now have a ZIP file which did not like the unzip.

You should try to name the ZIP file shorter because in Windows computers you have difficulties with long file names.

I will shorten the file name and extract again. Let me take a screenshot of the error.

Document: P10\_GMT20220802.odt Tags: Output data unclear

MODERATOR: The error is related to having a colon in the name. You can remove the timestamp from the file.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

MODERATOR: Let me do that for you and send you the zip file again.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

Okay, you can click on each of these variables.

Here you got the EEA data dictionary for the pollutants.

Document: P10\_GMT20220802.odt Tags: Report helpful

“Indices” is spelled wrong.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

The search input just filters but it is not entirely clear what it is filtering by.

If you put 7.0, it does search for all the 7.0 in the table.

It is not very clear if it is what it does. Yeah ...

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

So red is greater than 2 sdev and blue lower than 2 sdev.

I am surprised to see that there are no blue fields ... There are more red than blue.

Document: P10\_GMT20220802.odt Tags: Output data ready for analysis

Event D has only these three parameters the rest are missing.

I guess that event D is in a location where there is no weather data.

Document: P10\_GMT20220802.odt Tags: CSV output with missing data

I am surprised there is no weather data associated with D as the source of weather data is Copernicus, which has a pretty complete coverage.

PM2.5 is only available for some of the days. It does not link it all for event B.

Document: P10\_GMT20220802.odt Tags: CSV output with missing data

I would use "prior" not "previous" and "colour" is in the US spelling.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

I will select the maximum temperature there.

Okay, this is to show in a colour change the evolution in the time series, only for those available.

Document: P10\_GMT20220802.odt Tags: Visualization of plots clear

Oh sorry, these are bars. Okay they are bars as opposed to lines.

It is not that easy to see the box plots, unless you remove them one by one.

It would be nice to remove all lines as a whole or have an option to have them all off by default.

Document: P10\_GMT20220802.odt Tags: Visualization of plots unclear

There might be an issue with temperature values there for the climate linkage.

The precipitation is in mm I guess, this value is very unusual for Ireland.

I am assuming this is Irish data but ... Oh it is not.

Document: P10\_GMT20220802.odt Tags: Output data unclear

If we don't have full coverage we do not see a line.

PM10 seems to correlate with air pressure.

Document: P10\_GMT20220802.odt Tags: Visualization of plots clear

I will unclick some of the pollutants.

Okay so these are the z-scores for the climate variables.

I don't think you can see the z-scores for the pollution data or I can't see it anyway.

They are nice plots, alright.

Document: P10\_GMT20220802.odt Tags: Visualization of plots clear

The z-scores only apply to weather, not pollution.

Document: P10\_GMT20220802.odt Tags: Visualization of plots unclear

It does not like PM2.5 which is our primary target.

Document: P10\_GMT20220802.odt Tags: Data availability comment

What do these plots represent then? Okay, the plots are the average of the 4 events.

Document: P10\_GMT20220802.odt Tags: Visualization of plots clear

I would probably summarise. It is really cool.

Document: P10\_GMT20220802.odt Tags: Report helpful

It is very slick, has a nice interface, and is very easy to follow.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data

It has tremendous detail on the data protection which might be a bit of overkill.

Perhaps we could trimmer that down a bit but it is important.

Document: P10\_GMT20220802.odt Tags: Complex report content

The interface is really nice but a bit buggy.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data, Technical web issues

So you need to be very specific in the data that you put in or it does not work at all.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

If you had to work on something I would suggest the importing aspect by introducing more tolerance or by being more explicit about the format of the CSV file.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

I had an issue with importing the CSV file. I don't know what that is but there was a bug there.

Document: P10\_GMT20220802.odt Tags: Suggestions for additional features

My usual type of comment is that not using Copernicus for pollution is a missed opportunity which dramatically reduces the usability.

Document: P10\_GMT20220802.odt Tags: Data availability comment

In many cases the pollution data is not available, even in the selected dataset that you have selected towards maximum coverage or randomly.

Once you start trying with multiple input datasets you might find out that the coverage is .

Document: P10\_GMT20220802.odt Tags: Data availability comment

I would be concerned about how the outputs are coming to us. For example, we had a value of -3ºC for temperature which is not linked correctly.

Understanding where that comes from going back to the raw data might be important.

Document: P10\_GMT20220802.odt Tags: Output data unclear

The plots are very slick. It seems to me that you can only set weather but maybe there is a button there to change that. I could not see that easily.

Document: P10\_GMT20220802.odt Tags: Report helpful

The plots are very slick but I would probably change the default to only showing one of the pollutants.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

It is very busy when you get there but it is easy to remove them with the legend. Maybe you could add a note about that.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

We have 4 event summaries here so you could maybe add a title that says that it is a summary graph and you can split them afterwards.

Document: P10\_GMT20220802.odt Tags: Improvement to text or feature

It is really fantastic, well done.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data

Tremendously detailed, well designed with attention to detail but you know thinking about your user, will they care about all this GDPR stuff?

Wouldn't they care more about the coverage of their area? If there is no measuring station data ...

I think that is where the balance lies.

Document: P10\_GMT20220802.odt Tags: Output data ready for analysis, Data availability comment

I would definitely add Copernicus as a pollution source.

Document: P10\_GMT20220802.odt Tags: Data availability comment

I think I have done all my tasks, haven't I?

Document: P10\_GMT20220802.odt Tags: Experimental methodology

MODERATOR: You have skipped tasks 3 and 4 and went directly to task 5.

Document: P10\_GMT20220802.odt Tags: Moderator guidance

I am confident in the tool, which is very easy to use,

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data

but I am not confident in the data that is coming out accurately.

Document: P10\_GMT20220802.odt Tags: CSV output with missing data

It was quite easy to find areas where it was not accurate. I could certainly be uneasy with that.

Especially if you need additional features at the start of the linkage process.

Document: P10\_GMT20220802.odt Tags: CSV output with missing data

Add Copernicus for pollution data.

Document: P10\_GMT20220802.odt Tags: Data availability comment

The rest is covered in great detail, perhaps too much.

Document: P10\_GMT20220802.odt Tags: Output data ready for analysis

The website is quite verbose with lots of text but I think you need it all.

Document: P10\_GMT20220802.odt Tags: Complex report content

I think the interface is lovely.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data

It has come out a lot from the previous versions.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data

It has great potential for being an important tool for the global research community.

Document: P10\_GMT20220802.odt Tags: Improved tool usability and output data, Real use case applicable

MODERATOR: Would you mind thinking aloud as you complete the task?

Document: P12\_GMT20220902.odt Tags: Moderator guidance

Oh yeah sorry. This dataset contains many pollution and weather variables like temperature, PM2.5, PM10, humidity ... Wow, lots of variables, cool.

Document: P12\_GMT20220902.odt Tags: Output data ready for analysis

I will upload the example data. Okay, this is the example data.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

What is the meaning of “metadata report”? Metadata report is quite confusing for me.

Document: P12\_GMT20220902.odt Tags: Improvement to text or feature

Okay, I will conduct data analysis.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

Okay, cool. I will use the recommended metadata.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

I am not sure what season-alike events are but I will follow the steps to understand it better.

Document: P12\_GMT20220902.odt Tags: Some important features unclear

I will try season-alike. What is season-alike? Oh these are intervals, a measure of the aggregate length.

This is like to make it default ...

Document: P12\_GMT20220902.odt Tags: Some important features unclear

What's this of “within an area”?

Document: P12\_GMT20220902.odt Tags: Some important features unclear

I am going to use distance, 15 km then.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

I will choose days and I will choose the mean as this is what I would normally use.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

Link, I can click “link now”. Okay, this is the last step of task 1.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

I will just export the output then. I am going to save the file in my local machine and then unzip it.

Document: P12\_GMT20220902.odt Tags: Data linkage process clear

Okay, should I switch the window from the website to my desktop?

Document: P12\_GMT20220902.odt Tags: Technical web issues

MODERATOR: As you prefer, you can also use the one you are using now.

Document: P12\_GMT20220902.odt Tags: Moderator guidance

It is nice to see the previous steps I have done.

Document: P12\_GMT20220902.odt Tags: Report helpful

Wow, oh yes. The green points are the datasets selected. Great visualisation, nice.

Document: P12\_GMT20220902.odt Tags: Report helpful

Events A, B, C and D ... This event has two weeks of data for January ...

Yes, exactly I choose the mean to aggregate the data.

Document: P12\_GMT20220902.odt Tags: Report helpful

Okay, if I click on the variables I can see the documentation, cool.

Document: P12\_GMT20220902.odt Tags: Report helpful

Okay, this is cool. I can see the units by hovering over the variable name.

Document: P12\_GMT20220902.odt Tags: Report helpful

I want to search for the lag ... Okay, so I will type 2012 or January 4. Okay, efficient, cool.

Document: P12\_GMT20220902.odt Tags: Report helpful

Event A or event B, nice. I will search for events and dates.

Document: P12\_GMT20220902.odt Tags: Report helpful

Can I compare the values if I use mathematical symbols?

Document: P12\_GMT20220902.odt Tags: Suggestions for additional features

Okay, this is the interesting part for me.

Okay, these are the weather variables to construct the visualisation.

Okay, I can click them to make them appear and disappear, cool.

These are the boxplots.

Okay, I will try some variables here.

Document: P12\_GMT20220902.odt Tags: Visualization of plots clear

So there is no way to show the value of the wind speed by hovering over the plot, is that correct?

Document: P12\_GMT20220902.odt Tags: Visualization of plots unclear

Okay ... I will show some more box plots.

Document: P12\_GMT20220902.odt Tags: Visualization of plots clear

Okay, I am going to click event A. Ah, great! It is very easy to make individual plots by the direct output of this UI.

Document: P12\_GMT20220902.odt Tags: Visualization of plots clear, Improved tool usability and output data

Very nice report.

Document: P12\_GMT20220902.odt Tags: Report helpful

I think I am confident in getting some reports out of the UI.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

The one thing that is a bit difficult for me is that Ihad to read lots of text on the website to understand what I am going to do.

Document: P12\_GMT20220902.odt Tags: Complex report content

I had to know the meaning of the labels to understand the interactions.

Document: P12\_GMT20220902.odt Tags: Improvement to text or feature

Generally, I think it is a great tool.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

I think it is an easy tool to link the data.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

After several iterations, for example, for the first time you need half an hour to understand everything.

If I was familiar with the tool it may take me 10 min to finish the task.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

The learning curve is a bit high for new learners but after that I think it is a great tool towards data analysis.

Representing the time series and all the things I needed to get more understanding of the health events.

combining the weather and the according air pollution data.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

I wanted to say that environmental data may not be able to represent all that we want.

It could be interesting to link water or soil data.

Document: P12\_GMT20220902.odt Tags: Data availability comment

The interaction is very good for me. There are a few small expectations on the data.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

Yes, I said before in task 3, I would want more data to be added like sensor data about remote sensing, well remote sensing is already included in the data sources.

I would add water and soil data.

Document: P12\_GMT20220902.odt Tags: Data availability comment

The environment is a big area and more data could be beneficial but then if we have too much data it would be harder to analyse.

If we just focus on air quality and climate Ithink it is enough to conduct an analysis.

Document: P12\_GMT20220902.odt Tags: Data availability comment

Okay, so the interaction is very good. There are a lot of components on the website and it is easy to interact with the data.

Document: P12\_GMT20220902.odt Tags: Output data ready for analysis, Improved tool usability and output data

For example, I could follow the instructions very well.

I can read them one by one and get some knowledge from the instructions so Ican do the interaction more smoothly.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

The visualisations are good looking and efficient.

Document: P12\_GMT20220902.odt Tags: Report helpful

I don't know if it is because I selected too much data but the linkage took a few/several minutes.

Document: P12\_GMT20220902.odt Tags: Improvement to text or feature

I don't know if there would be an option to optimise and get a quicker query response from the website for the data linkage process.

Maybe it is hard, I know.

Document: P12\_GMT20220902.odt Tags: Improvement to text or feature

It is good to use and just to make it ... I think if you put in more effort it is not going to be more perfected.

I think it is a very good starting point and I am happy with it.

Document: P12\_GMT20220902.odt Tags: Improved tool usability and output data

Okay, so I have to connect data and health events.

So these are the health events and I have an option to upload.

I think you have shared with me an example of events data.

Let me save it over here. That's alright.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

I do not understand a lot of these countries ... Maybe they are related to the climate data but I am aware of these two data sources.

Document: P13\_GMT20220908.odt Tags: Data availability comment

I do not understand why there are only a few countries or if it is just an example.

I think that should not be an issue.

Document: P13\_GMT20220908.odt Tags: Data availability comment

I can see clearly that there are three options, details, it is very clear I have three steps.

The icons seem very clear to me.

Document: P13\_GMT20220908.odt Tags: Positive comments on text

I have the event data over here, that is fine. The data was successfully uploaded.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

I thought there was a chance for me to drag and drop or select a file.

Maybe you could highlight the box in blue or something to understand that this is a clickable option.

I did not realise at first this was a clickable option but now I understand it and I can import the event data using that.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

I can see that it has been successfully uploaded. Should I click again to upload data?

I think now the data has been properly uploaded.

Let me just check it ... Okay, so the data here is the data Ihave uploaded.

Document: P13\_GMT20220908.odt Tags: Some important features unclear

I would like to go with the minimum metadata as Ido not require much.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

I think this is another option to upload the metadata.

I don't have the metadata at this moment but I can use the sample that is shown over here.

I clicked on upload and it has been uploaded successfully.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

I would like to have ... Let me read that to understand what "as input" and "season-alike" mean.

I do not understand the difference between these two event dates.

Document: P13\_GMT20220908.odt Tags: Some important features unclear

Maybe a definition would be useful to have here. I can't see the definition.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

So I can see that season-alike is mentioned over here.

What about "as input"? I do not understand this.

Document: P13\_GMT20220908.odt Tags: Some important features unclear

You could add a definition of both of them that would make this part easier.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

I would like to have the season-alike 8 days before and after the event.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

Maybe I should go with a 5km distance radius as a circle. Maybe 10km is too much.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

Temporal unit I will go with year and aggregation method the mean or average.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

MODERATOR: Apologies, but the 5km distance is not a valid input as you can only select multiples of 10km.

Document: P13\_GMT20220908.odt Tags: Moderator guidance

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "year" as the temporal unit you will only get a single data point for each event.

Document: P13\_GMT20220908.odt Tags: Moderator guidance, Some important features unclear

Oh, okay multiples of 10. Sure sure, let me see if I can make this calculation again by making it days.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

I don't know when it will be back to the top. Oh, okay it is calculating right now.

Document: P13\_GMT20220908.odt Tags: Technical web issues

Perhaps if I could see a processing time over here together with the spinning sign it would be nicer.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

Then, I could have an idea of how long it takes to link.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

It clearly states that the linkage was done successfully and that I can go to the next step.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

The files are all in the FAIR format data which I am aware of and it should be good.

Document: P13\_GMT20220908.odt Tags: Data linkage process clear

LEt me just export this over here and see what Iget after unzipping the file.

Perfect, I can see all of them in a nice CSV format.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

I can understand that these are all the pollutants and the weather data that are linked from.

I have an idea about the air pollution and Copernicus data so Ithink they have been linked correctly.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

This seems to be a very good result that I can see with the CSV and TTL.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

Okay, I see the html file with a summary of all the details here.

I can see them properly and I have a lot of analytics being generated, that is really impressive.

Document: P13\_GMT20220908.odt Tags: Report helpful

I think that I'll be using this in my case to generate these plots in order to perform several experiences.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear, Report helpful

I think it is a wonderful tool. It is very clear.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

It gives a description of the data but if I was interested in some of it I would know where to go.

Extra information is provided in the form of hyperlinks so that is also very nice.

Document: P13\_GMT20220908.odt Tags: Report helpful

I can see the lead author's name so I can contact them if I need to.

This is very good.

Document: P13\_GMT20220908.odt Tags: Report helpful

I am very happy with the results.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

I think this seems to be very clear and systematic.

Document: P13\_GMT20220908.odt Tags: Report helpful

The SERDIF report is talking about the metadata, what data it is using and how the data is being used and logged.

Document: P13\_GMT20220908.odt Tags: Report helpful

Let me hover over the column names. I can see the units and variable descriptions.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

The user interface is very clear and by hovering you can get the information that you need without googling or searching it on the Internet.

Document: P13\_GMT20220908.odt Tags: Report helpful

I can see some of the background values are highlighted if the z-scores are more than 2 sdev.

I think it is a very clear way of identifying the particular pollutants for these particular events.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

I can see pollutants for these particular dates that are over 2 sdev.

I think this is a clear way of understanding this.

It is showing for other events as well so I am pretty happy about this.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

Overall it talks about several aspects of data exploration. I can see a dropdown menu here.

I am pretty sure if I click here I will be able to select from a list of different parameters.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

Alright, this is really nice if Ijust hover around and I can also see the concentrations of several pollutants. That is very impressive.

Let me check if I can do it for the precipitation amount ... Yeah, okay.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

If I do it for the precipitation amount, the secondary Y axis changes with the amount of precipitation.

Then, I can see the concentration in [um] for the other parameters. That is very impressive.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

I would like to do it for another one to see how it looks.

All of them are mentioned in bands. Essentially, for all of these particular days I think the solar radiation is of this particular value.

This is mentioned over here.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

It is slightly difficult for me to understand first how this band looks but now I can clearly get it.

Document: P13\_GMT20220908.odt Tags: Visualization of plots unclear

Maybe it is a minor comment but if I wanted to get the particular value of this band that would be good.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

This is very interesting as I click on here to see all the pollutant concentrations.

But I would also be interested in seeing the weather values for example for solar irradiance.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

I am pretty sure this would be high but I cannot see the exact value.

Perhaps if I could click on top of the bands and get the value over here that would be very impressive.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

Otherwise, it is absolutely very good.

I can see another weather variable as well.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

Before we were talking about concentration and now about z-scores.

That is very nice.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

Even for a particular event I could see the plots.

This is impressive, the plots are nice and clearly done.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

I can see some options to zoom in and out. This is amazing as I can download the plot as PNG.

The plot buttons are very intuitive.

I can see the generated PNG and it is amazing.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear

This is very nice and clear and I can use this on my dashboard and in the presentations properly.

Document: P13\_GMT20220908.odt Tags: Visualization of plots clear, Real use case applicable

MODERATOR: Before doing that, could you explore the metadata categories in more detail?

Document: P13\_GMT20220908.odt Tags: Moderator guidance

MODERATOR: You have some categories below dataset descriptors.

Document: P13\_GMT20220908.odt Tags: Moderator guidance

These are the different metadata categories. I can click on the input data which talks about what I have inputted.

Processing steps are the processing steps I did in the previous example when I was clicking on.

Here we have a distance of 10km radius.

This shows the different steps I took when generating the report.

This is also very useful.

Document: P13\_GMT20220908.odt Tags: Report helpful

I am curious about what this is. If I click on "Step 2: Results" it shows me on a geographic map the specific location of the events and datasets.

If I click on here it talks about the dataset of type air quality and EPA dataset.

Document: P13\_GMT20220908.odt Tags: Report helpful

I don't quite understand what the colour area "index 0" means.

Perhaps someone else working in this area would be able to understand this.

Document: P13\_GMT20220908.odt Tags: Improvement to text or feature

Let's see, this is also very nice. I can see the pan options and I can drag the different events. This is very interesting.

I can zoom in/out. The events are in Switzerland. There are 4 events each of them linked properly.

Document: P13\_GMT20220908.odt Tags: Report helpful

So that is brilliant

Document: P13\_GMT20220908.odt Tags: Report helpful

let me go to dataset structure.

I am aware of the structure but this is quite easy for me to understand.

There are two datasets, the climate from Copernicus and the air pollution from the EPA.

It may be mentioned here whether it was mentioned in the SERDIF data.

Okay, climate data is from Copernicus and air pollution from EEA.

Okay, this is very much in line.

Document: P13\_GMT20220908.odt Tags: Report helpful

Data sharing ... I need to know about it because it may contain some user specific data or confidential data.

I need to know who are the responsibilities and data controllers before I share the data with third parties.

I think this is clear and it is inline with the strict requirements of GDPR which are absolutely important to make sure we do not risk using the data in a different format.

Document: P13\_GMT20220908.odt Tags: Data protection importance, Report helpful

The metadata seems to be perfectly clear.

Document: P13\_GMT20220908.odt Tags: Report helpful

I really enjoy the overall feel of the webpage. I like things in a light background and the letters are in balck, not the dark mode.

I like this because it is very clear to me.

Document: P13\_GMT20220908.odt Tags: Positive comments on text

Yes, indeed I am very confident in using the linked data and Ithink this is one of the key things we need to do.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

Not only as a health data researcher but also for other applications as well.

Document: P13\_GMT20220908.odt Tags: Real use case applicable

I guess the SERDIF platform was fairly straightforward and I did not need any additional details or features to start the analysis of the linked data.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis, Improved tool usability and output data

After these exercises, I feel interested and I would like to know more details about this particular SERDIF platform and the related publications that come out of it.

Even though I can access the HELICALproject and ADAPT centre, it would be good to have a list of related publications for the community to track and explore.

I understand it may be a bit challenging to maintain a community of publications that come out of this but perhaps related papers from SERDIF would be a very good thing.

That is because I am keen to understand more about SERDIF.

Document: P13\_GMT20220908.odt Tags: Suggestions for additional features

So this is the only thing I would like to see, so related publications coming out of it.

Document: P13\_GMT20220908.odt Tags: Suggestions for additional features

Otherwise the rest of the features are quite straightforward.

Document: P13\_GMT20220908.odt Tags: Output data ready for analysis

My overall experience has been absolutely amazing.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

I think this is one of the key tools that any health data researcher would benefit from.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

I don't think this is restricted to health as such a user interface for any particular application would be absolutely wonderful.

Document: P13\_GMT20220908.odt Tags: Real use case applicable

I am really excited to work on this particular sERDIF platform as it is clear, systematic, straightforward and very user intuitive.

Document: P13\_GMT20220908.odt Tags: Improved tool usability and output data

I am looking forward to working with the SERDIF tool in the future.

Document: P13\_GMT20220908.odt Tags: Real use case applicable

So, first I have to upload health event data.

Drop or select a file ... What I do need here is event data.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

Okay so I have to upload the data here. Do I have to upload the data or not?

Okay, done.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

I want to do data analysis for example.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

Within an area or distance? Within an area ... What does NUTS mean?

Document: P14\_GMT20220909.odt Tags: Some important features unclear

Okay so there are different NUTS for different countries.

What is the difference between NUTS-2 and NUTS-3? They seem quite similar to me.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

EUROSTAT is something that you are not developing. It is external.

Whatever, I will choose NUTS-3.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

I want for example the year distribution.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

Aggregation method for example “mean”.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

Why can I not link? It does not allow me to link.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

I don't understand this metadata input. I already downloaded the data file that is needed.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "year" as the temporal unit you will only get a single data point for each event.

Document: P14\_GMT20220909.odt Tags: Moderator guidance, Some important features unclear

Okay, I see. Ineed to select "days" then.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

This data is already included in the excel that I have uploaded. Should I upload the dataset another time?

Document: P14\_GMT20220909.odt Tags: Some important features unclear

MODERATOR: This is another one about metadata.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

But I do not have another one.

Document: P14\_GMT20220909.odt Tags: Some important features unclear

MODERATOR: You can use the sample that is provided there.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

I have to upload the metadata. Oh, okay I see.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

Event date as input. Okay, so I want to link it.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

I think another important point is what you mentioned in the email that the data needs to be prepared in this way.

So this is important for the platform.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

Okay, export output. Task completed. I will move on to task 2.

Document: P14\_GMT20220909.odt Tags: Data linkage process clear

Can you see my screen? I think you can't.

Document: P14\_GMT20220909.odt Tags: Technical web issues

MODERATOR: Yeah, I can see it now.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Red ... Does it mean that they are positive?

Document: P14\_GMT20220909.odt Tags: Output data unclear

Lag ... Standard deviation ... Pfff ...

Document: P14\_GMT20220909.odt Tags: Output data unclear

So it means that here at the event-A, not I don't get it.

At lag 0, we have all this data. I don't understand this day 0.

What is the message here?

Document: P14\_GMT20220909.odt Tags: Output data unclear

MODERATOR: Maybe checking the column can help you, the second column.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Why has it changed? Why does it start from 12? Oh, no it starts with 1.

I understand the number of 1, 2, 3, 4, 5 and so on as the number of samples that you were collecting.

I don't understand the table.

Document: P14\_GMT20220909.odt Tags: Output data unclear

Then, lag ... day 1 ...

The z-scores are everything that is presented here, right?

Document: P14\_GMT20220909.odt Tags: Output data unclear

MODERATOR: The z-scores are only in the background as stated in the paragraph above.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

MODERATOR: Maybe if you go on with the report the plots might help you understand the table.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Oh, data exploration you mean.

For me it is hard to understand the table.

Document: P14\_GMT20220909.odt Tags: Output data unclear

Okay, so dates are backwards and on the Y axis we have concentration. Concentration of what?

Oh okay, concentration of pollutants.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

So per day this is the concentration of particulate matter. Okay.

These are concentrations of pollutants and these are the weather ones.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

What is the plot telling me then? Higher wind speed means there should be less particles.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

Why does the z-score plot tell me?

The plot is complicated to understand.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

The plots represent the tendencies so the peak here means that the maximum temperature around 5ªC is related with the maximum concentration of PM10.

How would that be connected with the days? Okay, that all happened 20 days before today.

Higher wind speeds, lower the concentration of the particles.

When the wind is slow or stable more particles can be detected.

More temperatures, more particles ... There seems to be a connection there.

Higher humidity, more particles, meaning that there is a positive correlation.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

The plots seem to be very similar.

I am confused about the z-scores. The z-scores tell us if the changes are significant or not.

Okay, so ... Okay, so everything within this window I imagine it can be significant, right?

Wait no, it should not.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

The standard deviations ... Yeah usually I would have box plots or bars to understand the values.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: The box plots are available if you click the boxes on the legend.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Oh, okay. This is not intuitive.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

Okay, so I see now.

For example, I just want to see PM2.5 ... I don't have them. Okay, so there are no values for PM2.5, right?

If I just stay with PM10, okay. I will check with the wind speed.

Okay, beautiful. 5 days ago there was very low wind and a higher concentration of PM10.

Whether this is significant or not, I can check with the z-scores. This is what z-scores tell me right?

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

So the z-score is when you have the bar plot it would be the deviations there, right?

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: You could try to plot the box plot by clicking the legend in the plot as you did before.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Ah, yes okay. There it is. If I click it here ... What does this line here mean?

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: That there is only a single value for this variable across all the events that you have linked.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

No, no this dashed line what does it mean.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: Above the red line is a high value, 2 sdev above the mean, and below the blue line, 2 sdev below the mean.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

So, this means that you cannot trust this data then.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: No. Values within the two dashed lines are normal values that are not considered abnormal values.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

This is abnormal then ...

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

MODERATOR: No. It is telling you how far a value is from the mean.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Okay, so this happened and this is it. This is what z-scores are trying to tell you, okay.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

MODERATOR: Would you mind checking the metadata section before moving to task 3? You can click on the categories on the left.

Document: P14\_GMT20220909.odt Tags: Moderator guidance, Complex report content

Ah so these are the geographical point of where the events happened.

Document: P14\_GMT20220909.odt Tags: Report helpful

I do not understand the lag thing ...

Document: P14\_GMT20220909.odt Tags: Output data unclear

MODERATOR: There is a diagram in the introduction of the SERDIF interface.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Lag ... Pfff ... before event

No, I don't understand this either.

Document: P14\_GMT20220909.odt Tags: Output data unclear

MODERATOR: Don't worry, you can move on to the next metadata sections.

Document: P14\_GMT20220909.odt Tags: Moderator guidance

Probably in the future we will have to use this data.

Document: P14\_GMT20220909.odt Tags: Real use case applicable

I am not very confident though with this kind of data.

Document: P14\_GMT20220909.odt Tags: Output data unclear

I will need a lot of additional information.

Document: P14\_GMT20220909.odt Tags: Improvement to text or feature

Starting from the basics, but I guess this would be the basics but I do not understand the events.

Document: P14\_GMT20220909.odt Tags: Output data unclear

The plots are pretty clear but I don't know the significance of those events.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear, Output data unclear

How can we build a correlation to understand if the variables are significant or not?

I would like to make sure the correlations are significant to trust this is what is happening in the data, so the confidence in this data.

I would like to visualise it in this kind of plot.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

The z-scores are only telling you if there are abnormal values.

I am not sure whether it is makeable, but probably it would be easier to merge these two plots.

Like putting the star sign we have in p-values, this would make it easier to understand.

It would mean that this is an abnormality and that you would not expect that to happen.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

So those event figures are the same figures but for particular events.

So they are separated data.

Okay, so there are good points.

Document: P14\_GMT20220909.odt Tags: Visualization of plots clear

This is what I am missing here.

So the overall figures are the mean or sum of all the available events.

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

So what I think it would be relevant, I don't know how, but to say that the significance of this.

Whether it is similar when we compare multiple days, how confident are we with this data?

Document: P14\_GMT20220909.odt Tags: Visualization of plots unclear

It is nice to have this type of platform.

Document: P14\_GMT20220909.odt Tags: Improved tool usability and output data

If I look at this table Iam not very confident in understanding what it is telling me.

Document: P14\_GMT20220909.odt Tags: Output data unclear

It is one day before ... Well with lag I am not clear at all ...

So these variables are the mean for 1 day for a single event.

Here the PM what does it mean? Okay, the colour here is above the threshold.

Document: P14\_GMT20220909.odt Tags: Output data unclear

I think you should add more explanation about what the colours mean.

Document: P14\_GMT20220909.odt Tags: Improvement to text or feature

The units? I do not see any units here? Where are the units? Is it true that there are no units?

Ah, okay here. Oh no there are no units.

Document: P14\_GMT20220909.odt Tags: Improvement to text or feature

When you click on data it changes very wierd the sorting of the rows.

Document: P14\_GMT20220909.odt Tags: Output data unclear

Okay, what do those numbers in the first column mean? Remove this column as it does not bring any information.

Okay, alphabetical order or descending order when I click on this.

Document: P14\_GMT20220909.odt Tags: Improvement to text or feature

Do I have to give you another example? Oh, no okay, I have to explain ...

Document: P15\_GMT20220914.odt Tags: Experimental methodology

MODERATOR: You can just scroll down and keep going, no need to provide another example.

Document: P15\_GMT20220914.odt Tags: Moderator guidance

I can select 1 event for instance the A event and drag it to the box.

No, this is not working.

Document: P15\_GMT20220914.odt Tags: Some important features unclear

I understand that I have to put all the data in this step but I don't understand why I cannot drag the events from the table to the box.

Document: P15\_GMT20220914.odt Tags: Some important features unclear

Okay, so first I downloaded the data. Then, I selected the file from my laptop and uploaded it.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

I have to choose the metadata I consider important for my research.

For instance, open data.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

I choose that but I am not completely sure it will ...

Document: P15\_GMT20220914.odt Tags: Some important features unclear

I will choose the recommended metadata.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

Is this the same? I have uploaded the data ...

Document: P15\_GMT20220914.odt Tags: Some important features unclear

Season-alike event ... interval around the event ... I will leave it as the default.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

Okay, distance. I don't know what NUTS mean.

Document: P15\_GMT20220914.odt Tags: Some important features unclear

The temporal unit is days because the interval is in days.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

The aggregation method ... The “mean”.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

I can export the output.

Document: P15\_GMT20220914.odt Tags: Data linkage process clear

I have several files: metadata, datatable, ttl which is a graphic and a html is the interactive report.

Document: P15\_GMT20220914.odt Tags: Output data ready for analysis

MODERATOR: Do you want to read task 2 again?

Document: P15\_GMT20220914.odt Tags: Moderator guidance

MODERATOR: You have already downloaded the output. You have it in the downloads folder on your computer.

Document: P15\_GMT20220914.odt Tags: Moderator guidance

MODERATOR: Then, you can open the html report to complete this task. Html files can be opened with your browser.

Document: P15\_GMT20220914.odt Tags: Moderator guidance

Like with Google Chrome? Oh, yeah sorry, I am not a very tech person.

Document: P15\_GMT20220914.odt Tags: Improvement to text or feature

Okay, there is a data table with three different indices: event, day and lag.

Document: P15\_GMT20220914.odt Tags: Output data ready for analysis

I do not see the z-scores ... I don't know where the z-scores are.

Document: P15\_GMT20220914.odt Tags: Output data unclear

Okay, it is general and it is not represented as a column.

Document: P15\_GMT20220914.odt Tags: Output data ready for analysis

If I want to see event C, okay then the table is filtered only for event C

Document: P15\_GMT20220914.odt Tags: Report helpful

The plot represents the concentration of pollutants for all the lags.

The lag in this case is days as I requested days.

Wind speed is in the background.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

Okay, here we have information about the air parcel weather and pollution.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

I am not sure if particulate matter is the only pollution variable but it is here.

Document: P15\_GMT20220914.odt Tags: Visualization of plots unclear

The sample variability is related to the length of each box.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

Okay, the pollution part of the air parcel is what is represented on the plot by the different lines and the concentration is related to all of these parameters of pollution.

Wind speed is represented on the right part only, which I can change for other weather parameters.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

In this figure, I see the z-scores of the weather and air pollution data.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

I can also swap the wind speed for temperature as I did before.

These are the events ... The number of observations that are above the mean, the higher the more different from the mean.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

I can also see the boxes with different lengths.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

I am not sure about the colours of the boxes now ...

Document: P15\_GMT20220914.odt Tags: Visualization of plots unclear

Here, I can see more clearly the plots per each event.

Document: P15\_GMT20220914.odt Tags: Visualization of plots clear

In my opinion I have some problems with this kind of data because I am not used to working with this kind of data.

Document: P15\_GMT20220914.odt Tags: Researcher's expertise

I think I would need more practice with the interface just to be more confident.

Document: P15\_GMT20220914.odt Tags: Output data unclear

At first sight Isee that the data is more tidy and it is useful to see the data in these interactive tables.

The tables allow me to show a single event or for specific dates for all the pollutants I need to see.

Document: P15\_GMT20220914.odt Tags: Report helpful

The plots are useful for me to visualise the data in another way that is more easy to do an interpretation for me than in a table.

Document: P15\_GMT20220914.odt Tags: Report helpful

I think it is a good tool for this purpose to explore the health data linked with the environmental events.

Document: P15\_GMT20220914.odt Tags: Improved tool usability and output data

I think there is too much information which can be confusing for me.

Document: P15\_GMT20220914.odt Tags: Complex report content

In my opinion, for me it is easier to follow less instructions or graph/diagram instructions.

Otherwise I become a bit overwhelmed with all the information but Iconsider that this is my case.

Document: P15\_GMT20220914.odt Tags: Complex report content

The information is enough and I would not need any extra information.

Document: P15\_GMT20220914.odt Tags: Output data ready for analysis

My overall experience is positive.

Document: P15\_GMT20220914.odt Tags: Improved tool usability and output data

It is a good tool to explore and organise the data in terms of visualising the different events.

Document: P15\_GMT20220914.odt Tags: Report helpful

It is easy to relate them with environmental data.

Document: P15\_GMT20220914.odt Tags: Improved tool usability and output data

It is also useful to change the background scenarios that affect your data and it is easy to follow the evolution of the events.

The z-score is a valuable tool to understand the extreme or abnormal events and how they are separated from the mean values.

It is easy to detect them on the plot.

Document: P15\_GMT20220914.odt Tags: Report helpful

It is also good to see all the events individually in case you have lots of data as it allows you to work easily with each event.

Document: P15\_GMT20220914.odt Tags: Report helpful

As I said before, my experience is positive and I only have to practice more.

Document: P15\_GMT20220914.odt Tags: Improved tool usability and output data

Event, day, lag ... Okay, that is fine.

Document: P16\_GMT20220902.odt Tags: Output data ready for analysis

Can I use your example data instead of my own?

Is the file already in there? Oh okay.

Document: P16\_GMT20220902.odt Tags: Experimental methodology, Some important features unclear

I want to do data analysis.

Document: P16\_GMT20220902.odt Tags: Data linkage process clear

We will go with season-alike events 10 days before and after the event.

Document: P16\_GMT20220902.odt Tags: Data linkage process clear

We will go with “distance” ... Oh no within an area ... This is subjective right? I can select one of them randomly.

Document: P16\_GMT20220902.odt Tags: Data linkage process clear

Why is it not working? Is it supposed to work?

Document: P16\_GMT20220902.odt Tags: Some important features unclear

MODERATOR: You are missing something in step 1.

Document: P16\_GMT20220902.odt Tags: Moderator guidance, Some important features unclear

Why is it not working?

Document: P16\_GMT20220902.odt Tags: Some important features unclear

MODERATOR: Try yourself first and if not I will answer.

Document: P16\_GMT20220902.odt Tags: Moderator guidance

I should have probably picked a smaller area so it is quicker but anyway.

Document: P16\_GMT20220902.odt Tags: Improvement to text or feature

Now I get a zip file and Ican use the data for analysis.

Oh cool, very good.

Document: P16\_GMT20220902.odt Tags: Data linkage process clear

Cool, is that it? Am I finished?

Document: P16\_GMT20220902.odt Tags: Experimental methodology

Wow, that was quick. It must be better than last time, simpler and easier to understand.

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

MODERATOR: You have finished task 1, but you have 4 more tasks to go.

Document: P16\_GMT20220902.odt Tags: Moderator guidance

Oh, okay I thought I was off the hook.

Where are the tasks?

Document: P16\_GMT20220902.odt Tags: Experimental methodology

MODERATOR: The tasks are in the informed consent.

Document: P16\_GMT20220902.odt Tags: Moderator guidance

MODERATOR: Apologies but you would be in task 2.

Document: P16\_GMT20220902.odt Tags: Moderator guidance

Oh okay ... Export the data linkage.

Where is my file now ... Can you see this?

Document: P16\_GMT20220902.odt Tags: Technical web issues

Hovering does not seem to work to get the variable units, oh wait it does.

Document: P16\_GMT20220902.odt Tags: Report helpful

Concentration of the particles ...

Document: P16\_GMT20220902.odt Tags: Visualization of plots clear

Okay the z-scores ...

Document: P16\_GMT20220902.odt Tags: Visualization of plots clear

Okay, the events have different lags here.

Okay that looks pretty cool.

Document: P16\_GMT20220902.odt Tags: Visualization of plots clear

MODERATOR: Do you mind exploring the metadata section of the report?

Document: P16\_GMT20220902.odt Tags: Moderator guidance, Complex report content

Yeah, they are all what we have agreed before.

Processing steps ... distance datasets ... When we have options you could choose the pollution data source also.

Lag of the event and interval ... okay.

Document: P16\_GMT20220902.odt Tags: Report helpful

Data structure ... They are all the variables here.

Document: P16\_GMT20220902.odt Tags: Report helpful

That all looks very good. It is definitely much better than the initial interface

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

I am confident in using the linked data for further analysis in R

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

Let me just have another look at the raw data ... Event, day ...

I presume event could be the identifier of a patient.

Yes, I think this is the format I have given to you before.

Document: P16\_GMT20220902.odt Tags: Output data ready for analysis

It is very straightforward and easy to understand.

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

The output is what I would expect.

Document: P16\_GMT20220902.odt Tags: Output data ready for analysis

It is good to get all the information on the metadata as well.

Document: P16\_GMT20220902.odt Tags: Report helpful

It is nice that it is very specific about the date of the event, the time lag and the duration.

This is the same as my export.

Document: P16\_GMT20220902.odt Tags: Output data ready for analysis

The plots are nice.

Document: P16\_GMT20220902.odt Tags: Report helpful

I think it looks good. I don't know how you would improve on that further.

Document: P16\_GMT20220902.odt Tags: Output data ready for analysis

I think it is much quicker as well. The last time I was trying to use it for a while but this one is more explanatory.

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

This is essentially what Iget feedback on the last time.

You essentially made what I had on my head.

Document: P16\_GMT20220902.odt Tags: Improved tool usability and output data

Okay, I can see the events and the different variables at a moment and place. Perfect.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

So, these are the data sources available.

I will have to choose events within this window and countries, I guess. Perfect.

Document: P17\_GMT20220921.odt Tags: Positive comments on text

Wait, I need to download the data that was in the email you sent. Here, perfect.

Document: P17\_GMT20220921.odt Tags: Experimental methodology

Before dragging the file let me open and check the data.

4 events with the dates that they fall within the time interval.

So, what I am going to do is to drag it here. Perfect.

I uploaded the data table.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

This is privacy stuff ... Okay.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

Spatial linkage ... not fully sure of what differences mean.

Document: P17\_GMT20220921.odt Tags: Some important features unclear

"within and area" or distance? ... Oh the administrative levels and the distance I can choose the values, okay.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

Even though this is arbitrary, I am going with geographical distance and not political boundaries.

10 km is fine even though it depends on the event I am looking for.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

I will choose "days" and the "mean" ... Okay.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

This is not working because I probably missed something.

Document: P17\_GMT20220921.odt Tags: Some important features unclear

Okay, now it works perfectly. I tend to forget about this privacy/metadata studd even though I know it is important.

Document: P17\_GMT20220921.odt Tags: Data protection importance

Okay, I can link this now and my task was to link the data.

So, when this is ready, the task will be completed.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

Okay, I have to export and explore the interactive report.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

Can I open the ttl with my linux? ... I don't think so.

Document: P17\_GMT20220921.odt Tags: Output data unclear

Okay, this is the input data and the data we generated as CSV files.

Okay, I can see the query here but I'm not sure what it is doing here.

Document: P17\_GMT20220921.odt Tags: Output data unclear

Okay, this values file is what I was expecting.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

For some of them there are no values which could be an indication that we are out of range.

These files include the summaries I think.

Document: P17\_GMT20220921.odt Tags: CSV output with missing data

Okay, all good. This is the input, exactly.

Document: P17\_GMT20220921.odt Tags: Report helpful

The processing steps represent the steps and choices I made in the interface.

Oh, nice map. We have events in Zurich. That is cool.

Document: P17\_GMT20220921.odt Tags: Report helpful

This plot is about the time of the events, okay.

Document: P17\_GMT20220921.odt Tags: Report helpful

Okay, yes one of the CSV was about the sdev of the values. Here I have thresholds, very good.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

For event C, there were two data sources but the Copernicus one is not working here.

We only have pollution but no problem.

Document: P17\_GMT20220921.odt Tags: CSV output with missing data

Here we have unusually high values.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

Okay, the plots are interactive, very nice.

Oh yeah, the legend is interactive too.

Document: P17\_GMT20220921.odt Tags: Visualization of plots clear

Ah the background button is related to the colour ... Okay, that is nice.

Document: P17\_GMT20220921.odt Tags: Visualization of plots clear

My only tip for this would be to use red colours for temperature, at least for maximum temperature.

Document: P17\_GMT20220921.odt Tags: Improvement to text or feature

The plot is the average throughout all of the events. Okay, it is cool.

Okay, this is to change from the mean to the deviation from all of the events ... Very good.

Document: P17\_GMT20220921.odt Tags: Visualization of plots clear

Okay, this plots the anomalies. Okay, here we are above the threshold.

It is interesting that it happened a few days before a peak.

Document: P17\_GMT20220921.odt Tags: Visualization of plots clear

Okay, these are the plots for each event separated.

That is why now we do not have the boxes.

We have some data here but not here for example.

Document: P17\_GMT20220921.odt Tags: Visualization of plots clear

MODERATOR: Would you mind checking the categories you haven't explored from the metadata section?

Document: P17\_GMT20220921.odt Tags: Moderator guidance, Complex report content

Ah, yes the deviation between events, exactly.

Okay, from 1 event area based on the 10km that I put.

Document: P17\_GMT20220921.odt Tags: Report helpful

Okay, data sharing ... Okay, so I think that's it.

Oh, very good to have the ORCID there.

Document: P17\_GMT20220921.odt Tags: Report helpful

Confident ... In what terms? Do you mean in understanding what it is?

Document: P17\_GMT20220921.odt Tags: Experimental methodology

MODERATOR: Yes understanding what it is and if you would be able to use it for your analysis as it is together with the information provided.

Document: P17\_GMT20220921.odt Tags: Moderator guidance

I think so. Here is an example with 4 events but it is important to access the information of all the events individually, not only the final summaries.

This allows us to narrow down and check there is nothing weird going on at some specific point in time for an event.

I think this is nice and Icould use this.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data, Report helpful

The interactive report I was exploring just now is very useful.

Document: P17\_GMT20220921.odt Tags: Report helpful

Always in the end I would be using the CSV raw data if I had to do something more.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

But I think to get a first impression on the data this is perfect.

Yeah, that is great.

Document: P17\_GMT20220921.odt Tags: Report helpful

The only problem I saw is that it did not say why some data was missing.

For example, the Copernicus data was lacking for some of the events when this is supposed to be exactly the same period of coverage.

But this is okay.

Document: P17\_GMT20220921.odt Tags: CSV output with missing data

No that I can think of right now.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

At the beginning I had a bit of trouble with the spatial linkage but I understood that it is to aggregate the data within these political categories.

Document: P17\_GMT20220921.odt Tags: Some important features unclear

Oh, I can check the NUTS on the external website, very nice.

Document: P17\_GMT20220921.odt Tags: Data linkage process clear

Yeah, I do not miss anything really.

Document: P17\_GMT20220921.odt Tags: Output data ready for analysis

For me, the main word to describe this is "easy".

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

It is like super easy to use. I mean you just put the data input as a CSV and in a matter of 1 minute you get all the linked data.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

This is great. It is super easy and I did not find tools like this out there on the Web.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

It is easy and intuitive so everybody could use this.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

People without a Computer Science background would be able to use this.

I think it is great.

Yes, "easy" is the main word.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

Obviously I was doing it fast but it is the first time I do this so I would have expected to have more doubts.

If I was doing this repeatedly I would even be more sure of what I am doing and more experienced with it.

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

I think it is great really. What I would like is to have the temporal and spatial coverage to be greater if you want me to say anything.

Document: P17\_GMT20220921.odt Tags: Data availability comment

But I know that this data might not be available for all the places in the world.

But it is hard especially for air pollution data, Iguess I don't know.

Document: P17\_GMT20220921.odt Tags: Data availability comment

I know that this is still underdevelopment so this is easily solvable and what is hard here is the infrastructure which here works very well.

So, that's my overall experience. Task completed

Document: P17\_GMT20220921.odt Tags: Improved tool usability and output data

The first thing that I would say is that there is too much text.

Document: P18\_GMT20220905.odt Tags: Complex report content

So I don't know where to start or sort of end.

I mean I guess where to start but I mean ideally I would like to do this directly.

Document: P18\_GMT20220905.odt Tags: Complex report content

Upload an excel here, set the linkage options and export the zip.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

When you say environmental data, do you mean anything or for this particular task I should select between Copernicus climate data or air pollution data?

Any of these ... Okay.

Document: P18\_GMT20220905.odt Tags: Data availability comment

Let me check the example quickly.

This is the format I need to have, which is exactly the one you sent in the CSV.

Event, date, latitude, longitude, length and lag to link them ...

Yeah, ideally I would like to see this more as the first thing I see is where to upload it.

Document: P18\_GMT20220905.odt Tags: Complex report content

So this is the format and I can upload the data table. No, that's wrong.

First I need to drag and drop and then ... Let me store the excel file on my desktop.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

The date needs to be in another format.

Okay, so give me one moment as I need to change the format of the date.

Can you see my screen? I am having some issues with formatting the dates because it says it is not in the right format.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: Yes, the date should be in the YYYY-MM-DD format so you need to change the slashes for dashes.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay, but my excel is not letting me do that ...

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: If you want to use the same data as in the example, you can use the table that is already there in the UI and click "upload data table" there.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Oh, okay. So maybe I don't know ... These sorts of conversions, it would be good to do them automatically.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Potentially to convert latitude and longitude automatically.

Maybe someone wants to give you the Swedish or British coordinates.

So if they know the CRS string then they can use it and you can do the conversions automatically to the coordinates you need to link the environmental data.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

So, the link again. I think there is a lot of text here.

Document: P18\_GMT20220905.odt Tags: Complex report content

I would like to conduct data analysis.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

Yeah, I would have thought that there might be potential issues with data confidentiality.

So this is something you need to think about the future.

The app could be used offline and then upload the data and link it with whatever.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Again, the other thing I was thinking is that I don't think it would make much sense when working offline to provide the data types available.

For now it is interesting to get whatever you want from Copernicus but if you wanted to make it for a private network or whatever.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

The environmental and health data should be inputted by the user so they can choose what data they want and then you can do the linkage offline with the app.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Metadata input ... I don't know what that means actually but I would use the “recommended”.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Okay, so this is the metadata. Why do we need this?

I don't understand why these are important to make the link actually.

I would have assumed that based on your interface I would get other columns ... No, you want ...

Document: P18\_GMT20220905.odt Tags: Some important features unclear

If the data table format is this one, I would have guessed that you would add another column with the exposure of interest.

You could have different confounders and select more than one.

I don't see why you need this metadata but let's use the one that you have there.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Is there an option to skip uploading metadata?

Document: P18\_GMT20220905.odt Tags: Some important features unclear

I don't understand what "season-alike" dates mean.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Yeah, I mean Isee what you are trying to do here.

I am thinking how important that would be in terms of data governance.

Anyhow, if you are sure this is fine in data governance, this is fine.

Document: P18\_GMT20220905.odt Tags: Data protection importance

Ah okay, you mean here that I am using the date that I have inputted in the first step and then link it based on days, months or years.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

Spatial linkage can be "within an area" or "distance".

I mean this is a bit ... I don't understand it well as it depends on the outcome and environmental exposure.

Document: P18\_GMT20220905.odt Tags: Real use case applicable

What I mean here is ... Now you have a lat/lon location which could be the centroid of something or the coordinates of something else or whatever.

So I don't know what this is and I know the data sources are grids. I don't understand if both of the data sources are grids, what does this mean?

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Does "distance" mean "closest distance"? Yeah, again I am not sure what this means because you cannot be generic to cover all the options.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Mmmm ... "within an area" I can see the European NUTS regions ... "distance" ...

Okay, to me it was not clear that if I clicked on that I would get other options.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

So basically you use the point and draw a distance ... Okay ... What do you do if several exposure points fall within this distance?

Do you compute the mean here? This is also something that should be mentioned here to explain how you calculate that.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

Okay, it is down here.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

I think it would be nice to have population weights here because in epidemiology we always want to have population weights.

Document: P18\_GMT20220905.odt Tags: Data availability comment

You are saying that within 10km from the event point and within one month or whatever we take the daily and spatial mean.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "month" as the temporal unit you will only get a single data point for each event.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

MODERATOR: The length in the input table refers to the length of the time series that will be associated to each event, in this case it would be of 14 days.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Ah, okay. So it calculates ... I would have assumed that it ... Wait, so why does it calculate the month across all dates and not ...

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: I would advise to check the diagram on the top in the introduction. You will see that you are getting data for a specific event. It gives you a granular view per event so you can do the aggregate afterwards.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

MODERATOR: Okay, we have some points within a region which are datasets and the event is within this region. What we are doing is to aggregate all the environmental datasets per event for a certain time window.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

But here we have 4 events in 4 different rows, regardless. I would assume that you compute the mean of the month of all A.

The temporal aggregation in my mind, it should have been location specific.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: Exactly, in this case because you are only selecting 14 days if you compute the mean you would get a single point for each event because you don't have more than 30 days.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay, I see. So will get the same dataset in a way.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

MODERATOR: So you would get one data point for all the variables that are available, so yeah more or less you are right. It would be more or less what you have there but with environmental data as columns.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay, so yes it would take the same table with some variables on the right.

But if I do "daily" I would get 14 rows for each event with all the data.

Document: P18\_GMT20220905.odt Tags: Output data ready for analysis

Then the lag is 0 ... So it means ... In epidemiology lag means a different thing, the lag effect.

Yeah this could be confusing but I see what you are saying.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

From the date of the event and from the 14 days after the event we are going to get environmental data and lag 0 means that we are not going to get environmental data longer than this.

But in epidemiology the lag is more about the days before the event occurs, right.

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: The time window is the days before the event and the lag is the delay of the time window. For example, if there is a lag of one week, the time window can start 7 days before the event.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

But in the diagram ... Oh, okay the event is here, sorry I did not see it.

The event is here, the lag is the previous days and then length is how many days I want to get data from.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

But why do we need the length here?

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: The length will depend on each use case as different studies might define different exposure windows as this could be more relevant for the health event. For example, researchers might want to study one month before the event, the cumulative exposure of one year or the acute response after one week. So it brings flexibility.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

So the length defines the length of the exposure. Then, the lag makes no sense, what does 3 days lag mean in a year?

Document: P18\_GMT20220905.odt Tags: Some important features unclear

MODERATOR: In this particular example maybe the lag would need to be 365 days to make sense. For example, if in the hospital they did not record a clear date, you might want to take a safe range and say 5 days before the event the person was already having some symptoms. It is to add flexibility for researchers as a parameter. If you are not going to use it you can always write 0.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay, I see. So what we want here is days just to get more data.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

A progress bar would be nice to have here to check how much time it takes

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

Okay, so this is the interactive report generated.

Document: P18\_GMT20220905.odt Tags: Data linkage process clear

This is the metadata that we were discussing.

The input data is what we inputted at the start.

Document: P18\_GMT20220905.odt Tags: Report helpful

The processing steps have the spatial linkage as 10km.

So this is like the verbose of the linkage.

Document: P18\_GMT20220905.odt Tags: Report helpful

MODERATOR: You can also click on the results to check what happened on each step.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Switzerland ... These are the events and this is the area for each of them ... Okay.

We have 4 events.

I am assuming that the time window plot would be very ugly with 400 events.

One event in Bern, Zurich ... Okay, I see.

Document: P18\_GMT20220905.odt Tags: Report helpful

You choose the average method ... I guess here you mean the "mean".

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

Z-scores ... The mean of what is observed or measured ... The mean of what? The mean of your datasets, the exposure dataset ... I don't know.

But I guess you get an idea about the extremes in a way.

Document: P18\_GMT20220905.odt Tags: Output data unclear

So, red is very high and blue is very low.

Okay, so let me type event A ... Okay, so we have the event and all the variables.

Document: P18\_GMT20220905.odt Tags: Output data ready for analysis

What would be also interesting in epidemiology is to have more information about the spatial resolution of this kind of thing.

Potentially is from 25km if it is from Copernicus but maybe this is on the metadata.

I would have liked to see how precise the linkage is.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

It is very hard to see the different colours.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

The shades are ... Concentration is [um] so we are talking about air pollutants.

We have wind speed as shades because you cannot put it on the Y axis.

This is for ... Is it for all events or for one event?

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

The mean of each variable across all of the events per lag ... Yeah, again what we are saying in publications is that you should be able to look at the plot and be able to know what it is.

So you need to add a title saying that, then the background buttons, lag and concentrations or whatever.

Because this really throws someone's attention, good.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

What is a bit puzzling is if the z-scores are calculated based on the event data or on the exposure data.

If the sdec deviations are calculated from the events or from the actual sdev of the exposure data.

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

I would assume that it would be for the events because you are interested in the extremes with respect to exposure, right?

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

MODERATOR: The definition of normals and sdev are in the metadata section of the report. The metrics have been calculated for each event for all the historical data that is available for a specific region.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay so it is based on the exposure data and not on the event data.

Both of them have merit. In one we can check if it is extreme in normal climate but in the other if it is extreme in my dataset.

If you have cases and controls it could be useful here. If the cases have constantly higher values then it could indicate an association.

Document: P18\_GMT20220905.odt Tags: Complex report content

MODERATOR: Previously I had it like this but then I changed it to historical extremes as it was more useful for researchers.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

I mean both are useful and make sense. Imean again, as a statistician z-value is not the best way to do this but whatever, you would use this.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

MODERATOR: What would you advise instead of z-scores for example?

Document: P18\_GMT20220905.odt Tags: Moderator guidance

I don't know why on the linkage stage we are interested in picking up the extremes ... It is a little bit going towards the analysis step.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

In my view and I have done several time the analysis of environmental data, yes you are correct the first is the linkage and then I want to have some sanity checks.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Some plots that show me whatever, the exposure plots where my outcome is and how it is exactly extracted on the point.

The first thing I do is to conduct some sanity check and then I start doing exploratory analysis like you have here with temporal plots and so on.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

But computing z-values is no longer exploratory analysis in my eyes but more like I am going towards analysis so I am observing that.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

We have some extreme values ... I don't know what it tells you exactly ...

So we have an extreme value 11 days before the event exposure ... I don't know what that means.

We see that for example PM might have some sort of ... No this is the mean across the events. So what does it tell you?

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

MODERATOR: So just the mean could be a bit misleading but if you click on the boxes on the legend. Underneath the legend, for example can you see the boxes there? You can click one of them to visualise them together with the legend.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

MODERATOR: Narrow boxes can mean that the events have similar exposures. If they are wide, the values across the events are not similar.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

What does it mean if the value of this is 2 sdev above the mean?

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

MODERATOR: It means that you had an anomaly in the climate and air pollution data worth taking a look at. It is to provide context to the data. This would be telling you if a value is high or low compared to the normals. For example, if a value at lag 7 for all the events is high then there might be something in the air that might be related with the event.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

MODERATOR: It was trying to give you a bit more exploratory analysis or information about the context of the data before your analysis. Each researcher might conduct their own analysis so I can see why this can be a bit misleading there.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Yeah, again it is the mean of the events so you are flagging if the mean of the events is extreme not if a particular event has extreme values.

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

MODERATOR: Yeah, it was more of an overview and then on the next step you have them individually.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

Okay, I think this also can lead a lot of people to biassed results.

What I am saying is that, if you select your lags based on what you see on the data then of course you are going to report something significant.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

The goal here and in general is that you initially have to set the lag period (e.g. three days) regardless of the data.

I do not think this is something you should see on your exploratory analysis because it can bias your decisions.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

But again, it is information and some people may need it but others not ... I have never used a plot like this.

But again, it is my take.

Document: P18\_GMT20220905.odt Tags: Improvement to text or feature

MODERATOR: We can discuss it after.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

So what you are saying is that for A we have different concentrations ... This can be very massive for many events.

Why don't we have lines for this variable here? Maybe it is incomplete.

Document: P18\_GMT20220905.odt Tags: Visualization of plots unclear

As I told you, I would like to see some sanity checks to make sure I am using the correct thing.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

So basically I would like to select Copernicus for instance to get an idea this is the correct thing I am using.

For example, by plotting the temperature or whatever. Then, if I am sure this is the data I want, I would go through all of this procedure.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

And then some sanity checks between the linkage of the events and exposure.

You had it somewhere I think but maybe some more ... Like a map of temperature and then put the point on the map and see interactively that this is the right thing.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Again, this is one date but just for sanity checks.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Now I am thinking, why do we have 5 points? Everything should be from the same source here.

Document: P18\_GMT20220905.odt Tags: CSV output with missing data

MODERATOR: I can explain it to you later.

Document: P18\_GMT20220905.odt Tags: Moderator guidance

So again, what I would have ideally liked to be more confident in using the linked data is to have sanity checks to see some values of the exposure and the outcome and see that what it has been done in the background this what I wanted to do.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

I would have liked more information with respect to selecting the exposure but this is on me.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

What if I don't want maximum temperature and only wind speed for instance?

How do I select that?

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

Then you have Copernicus for only a few countries which is a bit restrictive.

Document: P18\_GMT20220905.odt Tags: Data availability comment

I mean I would have liked to go to Copernicus instead on this page ... I mean Copernicus ... It should be relatively easy to implement a variable selection from the variables on your website.

So people are more flexible on it.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

I don't know. I am just saying that it may be interesting to select the metrics that you want because now we have 4 points but the analysis I am conducting on my research has 1.5 Million points.

I cannot do all the variables for Copernicus as it would be a waste of memory and time.

Document: P18\_GMT20220905.odt Tags: Real use case applicable, Suggestions for additional features

I would have liked some sanity checks.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

I think it was a pleasant experience and it was relatively easy.

Document: P18\_GMT20220905.odt Tags: Improved tool usability and output data

I think that too much text is a bit daunting but at the same time you need to have the information there.

So I guess it is hard to select how much text and information you should have there.

Document: P18\_GMT20220905.odt Tags: Complex report content

It was easy.

Document: P18\_GMT20220905.odt Tags: Improved tool usability and output data

I would say that in my experience it is nice that we have these integrated datasets inside but maybe what if I want to upload my datasets to do the linkage.

If I can input my event dataset on my own why I cannot add my environmental dataset too?

Document: P18\_GMT20220905.odt Tags: Data availability comment

But I think it was a pleasant experience.

I really enjoyed the html with the summary. It is nice to see some plots directly.

Document: P18\_GMT20220905.odt Tags: Report helpful

Yeah, anyhow it is nice to see the plots directly but I need more sanity checks here so Ican check that the linkage has been done correctly.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

For temperature, you could have boxplots over the year to see the seasonality or other things for people to get some sanity checks so people are confident on the data.

Document: P18\_GMT20220905.odt Tags: Suggestions for additional features

The last thing if you asked me if I could use it, I would have said no.

That is because I am using the health data in a private network without internet access so even if I did have internet access.

I would not have been able to upload the postal codes of people or the corresponding geocodes of people that have been hospitalised daily.

It is something against patient confidentiality so I would not be able to perform this task.

But for another dataset that I can upload online it is fine.

Document: P18\_GMT20220905.odt Tags: Data protection importance, Suggestions for additional features

I will read all of this stuff aloud as there is quite a lot of text.

Document: P19\_GMT20220912.odt Tags: Complex report content

I can see there are 3 steps and the overall flow of the website looks like there is introductory information at the start and then the steps.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

The introductory information looks like the background information I would need to know to use the resource.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

I suppose you mean a geographical area as opposed to a disease area.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

I think although linking environmental data with clinical data is not my exact current area of work I am interested in.

Document: P19\_GMT20220912.odt Tags: Researcher's expertise

It sounds like this is a time period that is being linked to a geographical area, which is fairly clear.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

Maybe having the words "geography" or "in time" or similar words would make it more clear it would be helpful.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

The result of a semantic query, although I know that it is what is happening here, that made the sentence harder to understand.

I would add another sentence to clarify "semantic technology".

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

The diagram probably can be something similar as we have a map of Ireland and data points in a geographical area.

There we have datasets and data points and now we have the same data in a graph form.

This represents an individual that had an event.

These are the different environmental variables within the data like UV radiation, PM, temperature and so on.

This is the exposure data and this is the amount of exposure this person had.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

Is this exposure data likely to coincide with the event?

A lag is in the context of trying to work out the cause and effect or relationship mechanism.

The lag could be the time for the disease to develop.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

The example here makes sense.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

I think A-flare-01 looks quite code-like, perhaps that could be expanded to something that could explain a bit more about being a flare for a particular patient.

It is a good code but it could take longer for an intermediate reader to understand that.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

Is it Czechia? Maybe for an English speaking person it would be easier Czech Republic.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

Sorry, I might be a bit too penequity but I will click to the data sources out of interest.

It looks like the sites of Copernicus even though I am not quite familiar with it.

The pollution data is from EEA.

Document: P19\_GMT20220912.odt Tags: Data availability comment

It sounds like a really clear three steps, three is the magic number.

I always like to summarise things in threes, which is a nice way to do it.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

It looks like Ican drag and drop or select a file.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Oh, so I can click on that. Maybe you can add some additional markup of that text because I assumed that when I clicked on it I wouldn't be able to click on it.

Also, when I hovered over a hand or a change in the pointer did not happen which you would expect for an interactive part.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

So, what I will maybe do is to click on the fields and then export them, which is quite nice to play around with.

I will change some of these items out of interest.

I will make the length all 20.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Okay, when Ifill that in it works but for whatever reason it looks like ... Yeah, so when I enter a value but I click out of the table it does not seem to record the value.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

Lag ... So this might be due to my lack of familiarity with environmental data but I don't know why I would select a different lag period for a clinical event.

I don't know when I would have different lags but for purposes of using this I will add just a row.

Document: P19\_GMT20220912.odt Tags: Researcher's expertise

Event E ... I don't know exactly where this lat/lon is but I am sure if I were a researcher using this that wouldn't be a problem.

Document: P19\_GMT20220912.odt Tags: Real use case applicable

It is overall very nice and clear in a clean interface to use.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

I will just now upload the table, okay that was very slick.

I thought I had to export it and then upload it again but the "upload table" button is here with a clear message here that it has worked.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

I did not realise that immediately from the text above that.

It wasn't until I got down there that I realised that I could upload the data there.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

So all of the steps seem to be that they will be recorded and that my interaction with it will be recorded.

If that is the case, if I were to use it with real data I would be given a login to trace what the user is doing or maybe I miss understood that.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

Oh wait it is more just recording my methods of how it is done so Ican check what I did.

I might read that again as I think I might be slow in reading it.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

Oh, okay. If I use this interface, aspects of the personal data might end up in the metadata report.

If I need to share that or reuse it within another context then it would require the appropriate data protection permissions in order to do that.

That makes sense I think.

Document: P19\_GMT20220912.odt Tags: Data protection importance

This sounds to conduct data analysis within a team and the other to publish it on the web.

There might be an intermediate step where you might want to publish some of the results of the data analysis but not all the data.

For some research user there might be some inbetween stage between those two.

They sound like maybe the two extremes on the spectrum on how you might use the output for.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

Okay, cool. There was a nice clear cross to remove the message when I wanted too.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

I will select recommended metadata. Again, I can add and edit the data as I want.

These are probably some example ones that you can use.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Maybe someone that is conducting this type of research might know which type of metadata they would want to attach to their output.

By the nature of working on the field, I might not know what metadata to attach at first.

Document: P19\_GMT20220912.odt Tags: Researcher's expertise, Some important features unclear

Unless somebody has me like a form like this with an explainer of each row.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

Some of this looks a bit complicated, I do not necessarily understand all of it.

It looks like it has to do with permissions on the organisation that is carrying out the research which may want to be attached to this output.

I will just upload the example one.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

So I suppose health data is always personal data unless it is irrevocably anonymised.

Then, due to the nature of the data like rare disease data, the linked data might still be personal data.

Document: P19\_GMT20220912.odt Tags: Data protection importance

Ah, okay that sounds cool and potentially very useful if you want to share some data making it more anonymous.

It sounds like it changes the time but it uses times with similar environmental exposures, which sounds very cool.

It is not something I came across or I have used.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Okay, I will try season-alike as it sounds cool. I will increase the time interval but I don't know if this is going to work quite well.

If you extend the interval long enough you might reach different seasons which are likely to have different weather conditions.

I will go with that for now but this is more of a research question as opposed to an issue to do with the interface.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear, Real use case applicable

I will change that to 5km ... Oh, I need to enter a valid value.

It indicated to me after I made the error that I needed to add multiples of 10.

I will do 20 or 30 then. It tells me in a slightly indirect way that I need to select.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Maybe changing that to an input box with a dropdown list where you can only select the valid options.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

The aggregation methods ... I am not sure what this exactly means.

I will go with the mean to see what happens.

Document: P19\_GMT20220912.odt Tags: Some important features unclear

Oh, maybe this is a way to summarise the environmental data for each particular day or a particular time point that you are using.

So, I will go with the mean now.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

I guess it is good for researchers to have this text to read while they wait for their data to be linked.

I am looking forward to seeing the interactive report, it sounds nice.

Sounds good, so I will try to export it now.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

Can you still see my screen here?

Document: P19\_GMT20220912.odt Tags: Technical web issues

I am sure we can figure out what is going on here by checking the html report.

Document: P19\_GMT20220912.odt Tags: Report helpful

I am not sure what these datasets are at the moment, it looks like a TTL dataset.

I was expecting to have a different dataset as a TTL.

There is also the metadata as a TTL.

Document: P19\_GMT20220912.odt Tags: Output data unclear

I am not immediately sure about the differences between these CSV files from their names.

But I am sure that by having a look at the data one could work it out.

Document: P19\_GMT20220912.odt Tags: Output data unclear

I will look at the report then.

I always love this automated output, it is nice.

There is quite a lot to explore here which is very cool.

Document: P19\_GMT20220912.odt Tags: Report helpful

I assume I will go through and have a look at the output and describe what I am seeing ?

Document: P19\_GMT20220912.odt Tags: Experimental methodology

MODERATOR: Yes, you can go section by section and it is up to you how much you want to explore.

Document: P19\_GMT20220912.odt Tags: Moderator guidance

This is very cool. It is quite satisfying to immediately get some output to look at before you analyse the data yourself.

Document: P19\_GMT20220912.odt Tags: Report helpful

The way it is presented here may be a way that the researcher might not immediately have thought of using the data.

I think that is really cool.

Okay, if I click on that Ijust go down to the section of the document as a hyperlink.

Document: P19\_GMT20220912.odt Tags: Report helpful

I really like this way of exploring it. I think if it was just all the data in one page that might be a bit too much text.

It is really nice to have everything laid out like this.

I won't go through all these links at the moment but it is useful to have them there.

Document: P19\_GMT20220912.odt Tags: Report helpful

Nice way of describing it.

Document: P19\_GMT20220912.odt Tags: Report helpful

So I got the data table structured in slices ... Let's see if I could see where those slices might be ...

So we got event, date, lag, so these might be the slices.

I think slice means this. Most data use data would think of rows unless event-A ... Looks like it is all event-A until event-B.

It looks like there is additional data per event.

This is because I set a definite number of lags and lengths.

Document: P19\_GMT20220912.odt Tags: Data linkage process clear

I think it went up to a 100.

Maybe that is why it took a bit longer because it was a lot of data perhaps?

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

I am not sure about the slices but it might just be how they are described.

Document: P19\_GMT20220912.odt Tags: Output data unclear

“Indices” is spelled wrong as the "e" is supposed to be an "i".

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

It sounds that these are the couple of datasets I saw already so I want to check if these represent that.

Okay, I presume that the z-scores are in the table here.

+/- 2 sdev sounds like a sensible z-score to use but I am not sure this is something I can adapt at later stages in the process.

Document: P19\_GMT20220912.odt Tags: Output data ready for analysis

I will try to type something in the search, cool.

That is really cool.

I presume I can search for numbers as well, maybe, okay.

I can probably search for dates as well, yeah.

That is really cool. It looks like there is probably a lot that you can do with that.

Document: P19\_GMT20220912.odt Tags: Report helpful

I am sure that this search bar has lots of functionalities or things I could do so maybe some examples what you might want to do it would be good.

Maybe a bit more elaboration on how you would like to search.

This looks like a quite powerful picture.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature, Suggestions for additional features

Then I can hover over the column names to get the units, which is very useful.

Document: P19\_GMT20220912.odt Tags: Report helpful

I am not sure if I had selected too many things there or if it takes some time to show up.

Cool, so that looks very nice. You can represent them either as box plots of line charts.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

I can hover over that and Ican get information about the different time bars as opposed to the box plots.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

I have used Plotly before. You can do some handy things like zoom in/out if that is useful.

I think because you have used Plotly before I know that Ican download the plot as a PNG.

I know that may not be that clear to the user but they need to go to this sidebar.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

I will try to change the weather variables in the background, that is cool.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

It is very minor but maybe the colour could change based on the weather variable because people could associate certain colours to certain variables.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

It is obviously very clear what the background variable is.

Yeah, I can select the variables I need. I don't know if it is maybe my browser but having NO2 line up to each other can help individuals realise that there is two options for NO2.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

Okay, so here we have it for z-scores and the other one was for raw values.

This plot has been standardised.

Okay, now we have a different colour scale because it can go both sides of the mean.

That makes sense, it looks good.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

Wow, that is nice. I always like a pretty plot.

Document: P19\_GMT20220912.odt Tags: Visualization of plots clear

If I have any questions I can always contact the developer.

Here we have the acknowledgements.

Document: P19\_GMT20220912.odt Tags: Report helpful

I might just go back to here ... I want to have a look at these files.

I really do not like numbers here, I hope that CSV can change that to default.

I may just be thinking that this looks like ... Maybe you need a bit more background on what the files contain.

Document: P19\_GMT20220912.odt Tags: Output data unclear

I am not immediately sure if the values have been standardised or they are absolute values.

I guess it must be absolute numbers for this but Iwas expecting a bit more data here.

Maybe it is my own lack of familiarity with the data.

Document: P19\_GMT20220912.odt Tags: Output data unclear

Okay, so it looks like data for each event over a period that Iselected in the dataset.

This time might be giving me standardised values.

Maybe it is something that I have selected and Iam being forgetful.

It may just need to have a better understanding from my side to know what to expect.

Document: P19\_GMT20220912.odt Tags: Output data unclear

Turtle file may not be very human readable.

Document: P19\_GMT20220912.odt Tags: Output data unclear

This may be some metadata that describes the things I did along the way.

Document: P19\_GMT20220912.odt Tags: Report helpful

I will answer this kind of hypothetically as I am not analysing the environment at the moment.

Although it is something I might be interested in the not so distant future.

Document: P19\_GMT20220912.odt Tags: Researcher's expertise

I will probably help one of my peers with a bit of her work as well.

Document: P19\_GMT20220912.odt Tags: Real use case applicable

If I was able to familiarise myself with the interface and the sources of data as well, I think I would become confident pretty quickly using this.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

It may require a better familiarisation with the underlying resources: Copernicus and EEA.

Document: P19\_GMT20220912.odt Tags: Data availability comment

I think that I could become confident pretty quickly.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

I might need a bit of advice from someone that has used the interface before or I might need to read the documentation closer.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

But I think it was pretty easy to follow the steps to select different options and it was all pretty clear to obtain the data.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

It is probably a totally different question and out of scope of this but then, using the data and how to do your analysis I don't know if I am necessarily confident with that but this would be a separate question in respect to this interface.

Document: P19\_GMT20220912.odt Tags: Real use case applicable, Researcher's expertise

I think, and again a caveat of not being an expert on this, I think there's lots of data there, lots of background data and sources to the underlying data which I am sure that hey will contain tons of useful data.

It looks like you can get the data over different formats and time periods.

So it seems there is a lot of flexibility on how you could use all of the potential available data.

Document: P19\_GMT20220912.odt Tags: Real use case applicable, Researcher's expertise

The units of the different types of measurements were available in a hover over manner but it may well be available in some of the actual datasets.

Probably it was from memory or not everywhere.

Maybe adding some metadata attached to it having the units in some sort of format besides hover over but I am sure they are probably somewhere.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

I cannot think of any other features before starting the analysis.

Document: P19\_GMT20220912.odt Tags: Output data ready for analysis

I think it was a really nice user experience and it has a nice, clean and satisfying interface.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

There was a good amount of text and explanations of what was happening as you went along but it was not overwhelming.

Document: P19\_GMT20220912.odt Tags: Positive comments on text

There was good use of either messages that showed up or hovering over items to get additional information that didn't have to stay there the whole time.

The user can keep the interface without too much information at the same time.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

As I mentioned previously, I like the three step process. It always seems manageable to a user to go through three steps.

It is quite nice to break it out into three steps. Obviously within each step there are maybe a few steps as well but it seems much more manageable, user-friendly and fun to have three nice and clear steps to go through.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data

The output html looks great. It is very useful.

It is a nice tool to play around with the data and to start exploring it.

Document: P19\_GMT20220912.odt Tags: Report helpful

Something that I wasn't totally clear about is exactly what all the other files represented at the end.

How I might use them or what data contained each of them.

Document: P19\_GMT20220912.odt Tags: Output data unclear

Again, that could be due to my lack of familiarity within the area.

Document: P19\_GMT20220912.odt Tags: Researcher's expertise

Maybe having something explaining that at some stage.

Document: P19\_GMT20220912.odt Tags: Improvement to text or feature

So overall experience ... Yeah, really nice interface and it looks like you could do a lot of useful work with it for sure.

Document: P19\_GMT20220912.odt Tags: Improved tool usability and output data, Real use case applicable

It is not immediately clear to me what point(x,y) means here.

I am guessing that it is a latitude and longitude coordinate.

But one has to make that assumption.

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

So if that was the event at that time with a lag that at that time.

Then you get the spill of environmental variables that you have produced.

Document: P20\_GMT20220906.odt Tags: Output data ready for analysis

As you probably know, I am quite interested in SO2 after one of my peers did that analysis on the biobank so I will focus on that for the purposes of this experiment.

Document: P20\_GMT20220906.odt Tags: Real use case applicable

Then, you've got all these files and I am not an informatician so I don't know what to get from here so I am going to ignore that.

Document: P20\_GMT20220906.odt Tags: Output data unclear

This is just describing what I have to do or what you guys do to link the data.

I guess it is what I have to do.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

So then I need to go on excel and create my own or can I just use this?

Document: P20\_GMT20220906.odt Tags: Experimental methodology

MODERATOR: Yes, you can just use this and upload the data table.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

I want to conduct some data analysis.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

It is quite nice that you have these little bubbles with explanations that are quite user-friendly. I like that.

Document: P20\_GMT20220906.odt Tags: Positive comments on text

Okay I will name this GPA vasculitis. I am lost here but I assume the ADAPT centre is fine here given that you are conducting all of these studies.

Document: P20\_GMT20220906.odt Tags: Some important features unclear

This is describing the dataset and where it sits from an informatics perspective.

I don't understand half of this stuff so I will assume that it is fine.

I am going to press the button to see what happens.

Document: P20\_GMT20220906.odt Tags: Some important features unclear

It is helpful to always remind researchers of government aspects. It is important.

Document: P20\_GMT20220906.odt Tags: Data protection importance

My intention is not to share at this stage, it is for my personal use, so I will go as input.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

I am not quite clear if I need to click all these ... I do not know what NUTS area is but I hope this link will tell me.

Document: P20\_GMT20220906.odt Tags: Some important features unclear

I am interested in the UK so that is NUTS-2. I am still not clear what NUTS are but I hope it provides some spatial resolution for my analysis.

Okay, so I decided I want NUTS-2.

Document: P20\_GMT20220906.odt Tags: Some important features unclear

So pragmatically speaking I think 25km seems okay.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

MODERATOR: Yeah, I need to fix that but the distance needs to be a multiple of 10.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

I think the temporal would be of "months" ...

Document: P20\_GMT20220906.odt Tags: Some important features unclear

MODERATOR: In step 1, you have selected only a few days for each event, so if you select "month" as the temporal unit you will only get a single data point for each event.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

So, I need to go for "days", thank you.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

Okay, I will export the zip file.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

The descriptions are generally very good.

Document: P20\_GMT20220906.odt Tags: Report helpful

Very clear.

Document: P20\_GMT20220906.odt Tags: Report helpful

I don't know what AVG function is but I guess it is the average function or something.

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

Again, I know what you are doing this experiment because the language of informaticians is not always the same of other human beings haha

Document: P20\_GMT20220906.odt Tags: Researcher's expertise

MODERATOR: Could you explore the rest of metadata categories there on the left?

Document: P20\_GMT20220906.odt Tags: Moderator guidance, Complex report content

Data input ... So, it is a good glossary there, very helpful.

Document: P20\_GMT20220906.odt Tags: Report helpful

Processing steps ... I guess this summarises my decisions.

Document: P20\_GMT20220906.odt Tags: Report helpful

Data structure ... The interest of my outcomes

Document: P20\_GMT20220906.odt Tags: Report helpful

Data sharing ... It tells me what I can and cannot do.

Okay, that is very helpful.

Document: P20\_GMT20220906.odt Tags: Report helpful

It is quite a nice description as well.

Document: P20\_GMT20220906.odt Tags: Positive comments on text

"indices" has a typo there.

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

Again, wide format might be a statistical notation so you might want to explain this.

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

Z-scores ... Nice explanation.

Document: P20\_GMT20220906.odt Tags: Positive comments on text

Sdev ... I had to work that one myself.

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

It seems sensible, helpful.

Document: P20\_GMT20220906.odt Tags: Report helpful

So what I need clarification on is what the event means. Do the events A, B, C and D represent a single patient?

I think the event here is a relapse and a single patient can have more than one relapse.

That is why you have got more than one C.

Document: P20\_GMT20220906.odt Tags: Experimental methodology

MODERATOR: An event is a date that can be associated with a single visit to the hospital. Then for example, you could code an event like patient1-visit1 and so on. It could be anything in this case.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

MODERATOR: Anything that has a date and a location. It could be relapse1, relapse2 which could be from the same or different patients.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

What stage has been defined?

Document: P20\_GMT20220906.odt Tags: Experimental methodology

MODERATOR: This would refer to the input data on the user interface that you have inputted in step 1.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

But for now, what is the commonality for an event coded A and all the ones coded A, and the rest that have C for example.

Document: P20\_GMT20220906.odt Tags: Experimental methodology

MODERATOR: Here it is just an example so they are random events but for the purposes of the experiment you can imagine that A and B are different patients.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

So for the databases that are available, you did not have SO2 for these particular data resources.

Document: P20\_GMT20220906.odt Tags: CSV output with missing data

It is good that the lag is built into this as a common challenge in all of this environmental analysis.

Document: P20\_GMT20220906.odt Tags: Real use case applicable

Let's try with solar radiation. So here we have all the different pollutants.

It is a nice visualisation.

Document: P20\_GMT20220906.odt Tags: Visualization of plots clear

It is exactly what it says there as it allows you to explore the data in a very efficient way.

It is very nice, I like that.

Document: P20\_GMT20220906.odt Tags: Visualization of plots clear

MODERATOR: You could also interact with the legend by clicking on the boxes there to get box plots instead of just the mean.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

Sorry, which box do I click?

Document: P20\_GMT20220906.odt Tags: Visualization of plots unclear

MODERATOR: For example, if you click the orange box for NO2 you will display the box plot for NO2. The boxes are below the X axis, underneath the lag.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

Okay, these guys.

It comes and goes. I get that but ...

Document: P20\_GMT20220906.odt Tags: Visualization of plots unclear

MODERATOR: So you can click the box just below that, the squared box. If there are no values you will get a line but if there are values you will get all the raw values.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

Oh okay, got it. Interesting.

Document: P20\_GMT20220906.odt Tags: Visualization of plots clear

The SO2, I am assuming that there is no data for the previous lags and this is a limitation of the dataset.

So how come that SO2 is here but I cannot see it here?

Document: P20\_GMT20220906.odt Tags: CSV output with missing data

MODERATOR: Some events have it and some others don't.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

Okay, so just these that I am looking at don't have it.

I know I should be asking questions but if these are from the datasets ... Just because in this person's geography there was not a local dataset.

Document: P20\_GMT20220906.odt Tags: Data availability comment

MODERATOR: You could increase the circle or look for another time period.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

The other background variables which you can contextualise with.

Yeah this is useful. Let's make some boxes appear.

Document: P20\_GMT20220906.odt Tags: Visualization of plots clear

I hope you let my colleague have a go at this, he would love this.

Good, okay.

Document: P20\_GMT20220906.odt Tags: Report helpful

How would I get the actual raw data to pull into a statistical program and go above and beyond descriptives of the datasets?

Document: P20\_GMT20220906.odt Tags: Output data unclear

MODERATOR: That would be in one of the files that you have downloaded.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

Okay, I will have a quick look at them then.

Okay, very good.

Document: P20\_GMT20220906.odt Tags: Output data ready for analysis

Is that everything?

Document: P20\_GMT20220906.odt Tags: Experimental methodology

MODERATOR: You can move to task 3 now.

Document: P20\_GMT20220906.odt Tags: Moderator guidance

As I said upfront, my role is not to get into the weeds but it is a helpful exercise to understand what interfaces my team uses.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data, Researcher's expertise

Actually, it was a lot easier than I anticipated.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

If I had a few hours of my life to devote to this, I think I would pick it up pretty quickly, which is saying something for somebody like me.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

So yeah, that does make things as easy as possible.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

Hopefully I have identified some areas that require a bit more clarity but you clearly made a strong effort to make this as accessible to all sorts of researchers and not focused informaticians or statisticians.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

Even someone like me, after playing with this for an hour I could get quite comfortable with this Ithink.

So well done.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

I think my only confusion was on those events as discussed when defining them.

Document: P20\_GMT20220906.odt Tags: Experimental methodology

That could be made a little bit easier but otherwise, you being able to see that spreadsheet makes me comfortable and I can immediately pull that into a statistical program that I would be comfortable with, so that is okay.

Document: P20\_GMT20220906.odt Tags: Data linkage process clear

You might want to have at the end of the report "Here is your data!".

I know it is sort of there but just to make it even easier.

There are a few spreadsheets there so sign posting directly after exploring the data so "Here is your spreadsheet! Now go away and have some fun"

Document: P20\_GMT20220906.odt Tags: Improvement to text or feature

It is the easiest interface I have come across in this particular space.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data

Again, the caveat is that I'm not in the weeds of this, so it is not as I have done this myself in another context.

But having supervised many students and seeing the troubles that they go through in other troubles of data linkage, this is as easy as I have encountered.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data, Real use case applicable, Researcher's expertise

That's a great step forward and I’m sure that the framework that you have put together here could be applied to many other different types of data so congratulations.

Document: P20\_GMT20220906.odt Tags: Improved tool usability and output data, Real use case applicable

I really like the plot/graph by the way.

Document: P21\_GMT20220906.odt Tags: Improvement to text or feature

Okay, I'm just going to go and use this table because I do not have a table prepared exactly with these kind of data.

Document: P21\_GMT20220906.odt Tags: Experimental methodology

I will add some more events.

I am going to hope this new one does not fall into event C.

One second because Iam trying to find a nice ...

So the events are in Switzerland, we can make an extra event in Lucerne.

Document: P21\_GMT20220906.odt Tags: Data linkage process clear

Just one note here, one comment, sometimes people have the coordinates in degree, minute format, so it would be nice to provide a way to automatically convert this.

If someone is not really experienced they might struggle with that.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

Any easy step that you can add might help, any step like that might be useful.

I had to do it not so long ago for my data so I thought it might be useful for someone else.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

It is nice to get this information.

Document: P21\_GMT20220906.odt Tags: Positive comments on text

We don't want “minimum”, let's choose “recommended”.

I am just going to assume this is the metadata for the data including the row that I added.

So I am going to upload it.

Document: P21\_GMT20220906.odt Tags: Data linkage process clear

Within an area ... Oh nice! Okay.

Document: P21\_GMT20220906.odt Tags: Data linkage process clear

I think this step concludes task 1, which I think I achieved successfully.

It is quite easy since I used the example template but whatever.

Document: P21\_GMT20220906.odt Tags: Data linkage process clear

Now I am going to export the output which is going to be task 2.

It comes into a zip file and then I guess I can just extract it.

Document: P21\_GMT20220906.odt Tags: Data linkage process clear

It is really nice both aesthetically and functionally wise. I think this is very nice.

Document: P21\_GMT20220906.odt Tags: Report helpful

Processing steps ... This is really nice because you can trace back everything.

Document: P21\_GMT20220906.odt Tags: Report helpful

MODERATOR: You can also explore the results for each processing step if you want.

Document: P21\_GMT20220906.odt Tags: Moderator guidance

Oh wow, here are the events and stations used to generate the data that is really nice.

Document: P21\_GMT20220906.odt Tags: Report helpful

Maybe you could use some rounding for the coordinates but still.

Just because I am very neat and picky but I think this looks great and you really get an idea of what happens.

Document: P21\_GMT20220906.odt Tags: Improvement to text or feature, Report helpful

These are all the areas that we have used in Switzerland, great.

Document: P21\_GMT20220906.odt Tags: Report helpful

It works very well. It is also nice that it is a local html because then you don't have problems like server response or anything.

It is very smooth compared to usual websites, depending on the servers of course, but you would have to pay for a lot of server usage.

In this case, this will always be usable, which is nice.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data, Report helpful

The time window is also very nicely shown.

You can export it if you want.

I think this is really great because you can see when the event happened.

I wonder how it would look for a lot of events but still you could get more or less an idea.

I really like it.

If you didn't know what you were selecting this would make it very clear what you are doing.

Document: P21\_GMT20220906.odt Tags: Report helpful

Okay so the maximum values should be ...

Okay, really nice.

Really nice and intuitively explained with smooth transitions.

Document: P21\_GMT20220906.odt Tags: Output data ready for analysis

Okay, so as I understand.

Is this for a single event or ... I assume this is the mean for each of the lags.

So basically grouping by lag and then doing the mean of each of the variables during that lag.

I think that is how you should look at it.

Document: P21\_GMT20220906.odt Tags: Visualization of plots clear

I wonder why we have 25 lag but I guess it is because we have different lags.

It can become a bit annoying when we have selected different lags for different days but I guess it is exactly what makes sense.

This was my instruction so usually I would not select different lags unless I have a very specific reason and Ithink that makes sense.

Really nice actually, you can select the background like temperature for instance.

Document: P21\_GMT20220906.odt Tags: Visualization of plots clear

Then the same but with z-scores, right?

Document: P21\_GMT20220906.odt Tags: Visualization of plots clear

I mean all these variables are kind of normal but if you were using a kind of non normal distributed variables, z-scores don't work that well.

Well, they don't work at all.

But I think these variables are close enough to a normal distribution where you can use z-scores in this sense.

Document: P21\_GMT20220906.odt Tags: Improvement to text or feature

A z-score really gives you the capability to see strong anomalies happening to one variable.

So we have them for each of them, nice.

Document: P21\_GMT20220906.odt Tags: Visualization of plots clear

This one is not working for some reason, maybe because the variable is not working for some reason.

Okay, if that is the case I think it should be somehow communicated because someone might think there is a bug if you get no feedback.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

I would say that feedback when some of the variables are not available at all.

For example, event E, which is the one I selected, we don't have data at all it seems. Maybe I added a date that was not included in the records or is it in a location without data.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

We cannot see which event is it ... Oh, yes. As we see the NUTS-3 area has no stations meaning no data.

So what we could do is to go upscale and do NUTS-2 and probably we would get data.

It would be nice to know this but I don't know.

I guess you could also give a message for this.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

Like if some of the events when you are located given an area ... But I guess I could have selected a circle and then it tries to find the closest one.

Actually this station is super close and probably we would have been able to use that station.

But yeah. Some kind of feedback on that if you are not getting data for an event for whatever reason, it would help to easily pinpoint.

I know it is not so easy to pinpoint. I know it is not that easy to do so that's also nice I guess.

I think that was a bit of a coincidence that I selected that place but it was good to pinpoint that.

You can see it here but just to know that it is not ...

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

I think I would be very confident using this.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

I think not only because Itrust the person who did it but because you can trace back where all the data is coming from.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data, Report helpful

You not only have the metadata information but you have really specific plots that show you specifically where your data is coming from.

Both in the temporal and the spatial scale. I think it is great to see that we are aggregating over this area and using the data from these stations to aggregate it.

Document: P21\_GMT20220906.odt Tags: Report helpful

Showing the temporal scale is also great because the user can really get an idea of what this means.

Depending on how used they are to working with lags and windows, people might get lost with that.

Seeing it graphically it really helps to visualise the data that is being fed and I think this is very important.

Document: P21\_GMT20220906.odt Tags: Report helpful

So I would be very confident using this data.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

Sometimes Ithink Iwould like to have a bit more feedback when data cannot be ... I mean, of course it depends ...

We could assume that people that are going to use this are quite capable of getting data.

But I assume well, we did not get data for these, could we get it?

Maybe just giving a hint like: "We could not get data for your specified level of resolution for these specific events, maybe try to lower the resolution to get more data."

Some small hints like that and providing feedback on why there is no data for specific values that would be completely useful.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

I guess you can download this table ... Can you export his table as a CSV?

Yes, I think you get the export as a document itself. Yes, you get the values. I did not say anything.

Document: P21\_GMT20220906.odt Tags: Output data ready for analysis

Basically more feedback on certain things that did not work.

I know that it is not easy but providing feedback when things don't work, but not when things don't work because there is an error that you could not have thought of, but these errors that you can expect might happen.

Well it is not an error it is that there is no data for that specific location and time just say it or give something so the user knows why to discard the data point or actually get data that might not be available in the initial resolution that they might have wanted.

Document: P21\_GMT20220906.odt Tags: Suggestions for additional features

I would say well I don't remember when I tried last time but this looks really well.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

I think it looks really scientific in a way, the explanations are very nice, but it looks both scientific and useful.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

Like it does not look hard and you get kind of walked along each step which is nice.

There are information nodes. You get the sources for everything.

You get explanations for everything and the small things.

I think it looks great.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

As I said, usually I would be more prone to using a web-based service but I think if everybody's computer can display and represent an html file like this, it is not an expensive process anyway.

If in this way you don't really need to maintain the server side of this and you can just get the report as an html is up to the user's ability to use the page.

Most users won't even realise they are not on the Internet and that they are doing it on their computer anyway.

Document: P21\_GMT20220906.odt Tags: Report helpful

I think the whole way this works is great and I also like it from an aesthetic point of view. I like how it works.

I think it is a polished work and it really works.

Document: P21\_GMT20220906.odt Tags: Improved tool usability and output data

Okay, so I took your example and I have added a few rows to it.

So, I can just go in to get the file. Okay, successfully uploaded going to step 2.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

I am unsure what exactly this means but I will click the minimum here.

It looks like here Ihave to upload the metadata.

What I will actually do is export that file ... Oh, wait, let's cancel that and I will use the example that is here.

Document: P22\_GMT20220908.odt Tags: Some important features unclear

Okay, it is telling me that I can have the dates as input or season alike.

So the example that I'm uploading it has specific dates on it so I will take the dates as input.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

I think I am going to go with "within an area" and I am going to trust that the lat/lon falls within ...

Wait no, what I will do is ... I will do a circle and I will do distance in km from the lat/lon point.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

I think that what the circle is going to do is to map to the closest Copernicus or other location within 10km of the lat/lon coordinate of the event.

If there is more than one, I guess it averages between those but maybe I can ask you about it afterwards.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

I will take the temporal unit as a day and select the maximum over a day.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

What this is going to do is apply the maximum operator over all those climate and air pollution variables in the same day.

I am assuming that it is down to daily resolution so maybe that won't be an issue right.

So I will get one value for a day.

Then, I will link the data.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

Okay, obviously where I was going wrong is the metadata so that is clear to me now.

Document: P22\_GMT20220908.odt Tags: Some important features unclear

I have about seven events in that event data spreadsheet, well 6 events.

It is taking a bit longer than I might have expected for 6 events hahaha

Document: P22\_GMT20220908.odt Tags: Improvement to text or feature

So it is just exporting as a zip file.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear

Now I am going to take a look at the zip file.

CSV and a graph ... You can't see this now but Iam looking inside the zip file.

Right, so what I'm getting out is a number of files here.

So I am getting what looks like a TTL file, a normalised and sdev CSV files and a CSV file with raw values.

Within that CSV file with the ... I think this looks like the raw values of the CSV file.

Document: P22\_GMT20220908.odt Tags: Output data ready for analysis

MODERATOR: Would you mind reading task 2?

Document: P22\_GMT20220908.odt Tags: Moderator guidance

The metadata is just a standard example with rare disease A.

Document: P22\_GMT20220908.odt Tags: Complex report content

The data is telling us the atmospheric surface, linked data ...

Document: P22\_GMT20220908.odt Tags: Output data ready for analysis

MODERATOR: Would you mind exploring the rest of metadata categories there on the left?

Document: P22\_GMT20220908.odt Tags: Moderator guidance

Okay, right the input data ... Okay so this gives us the head of that import ile.

Now, what question that comes out is what happens when we leave out one of these fileds here like length, lag ...

What does it default to?

For most modelling parts, I think the length could be set easily and the lag could be 0 anyway.

Document: P22\_GMT20220908.odt Tags: Suggestions for additional features

The time window is on the horizontal axis. One thing that I would say here is that if you have a lot of events this will become really crowded here.

But the other thing is that if you have recurring patients, let's say that instead of an event you might have a patient id or number, and that may be repeated in the vasculitis case if we have relapses.

So I am wondering how would it cope if instead I had a patient ID and I have a repeated index. Would it handle this okay?

That is why I tried to use A-0 but I should have probably tried with A.

Document: P22\_GMT20220908.odt Tags: Improvement to text or feature

Okay, and we are looking at the max.

We can't have everything but it is going to apply the maximum operator for everything.

Document: P22\_GMT20220908.odt Tags: Report helpful

Can we just get out the raw data as well without applying any kind of aggregation?

Document: P22\_GMT20220908.odt Tags: Data availability comment

Then, it is giving z-scores ... So here, it may be a little bit ambiguous.

It is not clear how the mean is actually calculated, right?

Is it over a particular area or circle of radius 10km?

Over what length of period of time, so may not be initially clear.

Document: P22\_GMT20220908.odt Tags: Output data unclear

Yeah, okay this is very helpful because the colours highlight straight away values that are 2 sdev away from the mean as normalised values.

Actually what we are seeing here is that the normalisation is done in the background and the table is showing the raw values.

If the raw value is colored red it would be +2sdev away from the normalised mean.

Document: P22\_GMT20220908.odt Tags: Output data ready for analysis

So that is actually helpful. Looking at this the other byproduct of this is that it gives the different pollutants and what kind of a normal value is.

This normal value is specific to an area so how this normalised mean value is computed in the first place and over what kind of time scales are we thinking.

Document: P22\_GMT20220908.odt Tags: Report helpful

Okay, so what I'm looking at here is an overlaid time series plot with each of the pollutants here.

Okay, so if I click this for PM2.5 ... I have asked for daily resolution, which is giving me a single line which represents the level of PM2.5.

Let me unclick this but click the NOx, it is giving me box plots.

Document: P22\_GMT20220908.odt Tags: Visualization of plots clear

Okay ... Right if we want to look at ... Sorry but the Zoom conversation box was covering the second axis here hahaha.

Document: P22\_GMT20220908.odt Tags: Technical web issues

I have to say that I am really not a fan of plots with 2 vertical axes, this is incredibly confusing.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

The purpose of this plot is to judge a correlation really between all these weather variables that are listed here.

For example, sea level pressure and one of the pollution variables.

Document: P22\_GMT20220908.odt Tags: Visualization of plots clear

I am not sure what the purpose of this is.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

Again, it is the same sort of idea but we are looking at z-scores instead with the background variable as well.

Maybe we are looking at correlations between variables.

Document: P22\_GMT20220908.odt Tags: Visualization of plots clear

So these are the different pollutants for the different events so we can drill down to the personal level and look at individuals.

Document: P22\_GMT20220908.odt Tags: Visualization of plots clear

Event E we are not getting anything so I am assuming maybe the lat/lon is off.

I am just pulling out my spreadsheet ...

Document: P22\_GMT20220908.odt Tags: CSV output with missing data

MODERATOR: You could use the map in the processing steps category of the metadata. You can press the zoom out button to see everything to get the appropriate width.

Document: P22\_GMT20220908.odt Tags: Moderator guidance

Okay, this is event A0, A, B, C, D and E ... Okay, so I remember what I did now.

I copied the lat/lon of A into E so I have 2 events happening at the same lat/lon but let me confirm that though.

No A and A0 happened at the same lat/lon coordinates.

I am not getting any values for event E ...

Document: P22\_GMT20220908.odt Tags: Some important features unclear

MODERATOR: You could check the countries available in the interface.

Document: P22\_GMT20220908.odt Tags: Moderator guidance

Yeah, that must be the reason why because the event is located in Germany.

I will go back and check that earlier that I got everything okay.

So I might have to change the regions in order to get that one and everything else in Switzerland.

Document: P22\_GMT20220908.odt Tags: Report helpful

So A and A0 but the actual plot looks different ... This is line and line+point so I might wonder why that is the case.

Does this signify something as a user or it is just probably nothing? Okay.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

So what I am going to do is to have a look at the ZIP file I have downloaded.

I can pull out ... I would say that the names on the CSV and everything that comes in the ZIP folder are kind of long winded.

Document: P22\_GMT20220908.odt Tags: Output data unclear

Would there be a way to put a string there in any way? It is fine obviously as you can rename them yourself.

But just for convenience.

Document: P22\_GMT20220908.odt Tags: Output data unclear

Yeah, I am pretty sure I understand what this data actually means and I would be confident in using it.

Document: P22\_GMT20220908.odt Tags: Improved tool usability and output data

Just one moment. I am looking at what happened to event E in my spreadsheet or to event C. We have missing values for some of the components.

This obviously has to do with the values not being available for the satellite or something else.

There is not much you can do about that.

Document: P22\_GMT20220908.odt Tags: Data availability comment, CSV output with missing data

Would it be useful? If I am uploading a spreadsheet with a ton of dates and events and I just downloaded this, would it be useful to output a message on the main screen that there are missing values or some event locations do not have complete records?

Would it be useful to do that for a user?

Document: P22\_GMT20220908.odt Tags: CSV output with missing data

But I would be confident in using the spreadsheet.

Document: P22\_GMT20220908.odt Tags: Improved tool usability and output data

I think just to be clear in naming the files in the zip folder. It would be helpful, right.

You have a -normals, -values, -sdev and another CSV.

Document: P22\_GMT20220908.odt Tags: Output data unclear

MODERATOR: Would you mind making a suggestion on how the filename should look like?

Document: P22\_GMT20220908.odt Tags: Moderator guidance

Yeah okay, the -values just raw values ... -normals, let me have a look at that, the -normals are the sample means over whatever aggregation you’re using to normalise it.

So you could maybe name the -normals and -sdev files as something like normalisation-means and normalisation-sdev to make it clear these files are to do with calculating normalised values.

Document: P22\_GMT20220908.odt Tags: Output data unclear

I think you can add more information about how this normalisation is calculated because what Iselected was an area of 10km so I am not sure how the normalisation happens if I increase it to 20 km.

I am not sure over what time scale the normalisation is.

Is it over the study period or dates that I have presented or is it over a much longer time scale like the entire Copernicus back catalogue?

Document: P22\_GMT20220908.odt Tags: Improvement to text or feature

I think how the table is presented is very helpful as it gives a very much sense to the researcher of the scale of values one might expect.

Easier for cross-referencing within different regions as well.

Document: P22\_GMT20220908.odt Tags: Report helpful

I think that the experience using this dashboard with the exception of the two vertical axis plots, is a subjective thing, because I always find them very confusing.

So I mean you just don't take that as a harsh criticism and it is personal preference.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

The colour has to do with the bands on the background which were not immediately clear to me.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

I don't know if looking at them I do get a sense of the correlation in some way but it would take me a while to get comfortable with.

I mean, is this something that a researcher would immediately catch on and use it in this context.

I am not sure, maybe, maybe not.

Document: P22\_GMT20220908.odt Tags: Visualization of plots unclear

I think that the whole kind of lay off and the fact that you can upload event dates is much better.

Document: P22\_GMT20220908.odt Tags: Data linkage process clear, Improved tool usability and output data

From the statistician side what we would be doing is uploading that spreadsheet of days putting a lag of 0 and a length of maybe 90 days or something.

You have probably seen my colleague's work. Then, fitting this kind of a weighting function or inferring the weights on 90 days.

Then, trying to figure out which of these variables might be predictive of the flare case.

Document: P22\_GMT20220908.odt Tags: Real use case applicable

It does solve a problem for the interoperability of this data that we had because processing this prior to this with dates and messing around with the spreadsheet in the background.

Document: P22\_GMT20220908.odt Tags: Improved tool usability and output data, Real use case applicable

It does take a long time to run, but what is the computational complexity of those queries?

Document: P22\_GMT20220908.odt Tags: Improvement to text or feature

So here we have step 1, and I have got some data that I received in my email with this structure here.

I guess I am just going to upload that data table.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

So I will just make sure I can find it on my email.

I am just going to save it to my desktop and then I will go back.

Document: P23\_GMT20220909.odt Tags: Experimental methodology

Okay, there we go. My gut feeling here was that the drag and drop into this section makes sense but it was not obvious to me that it was a clickable button that I could select and get the option to see my files.

Document: P23\_GMT20220909.odt Tags: Improvement to text or feature

Okay, sure thing, there we go.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

Maybe I have uploaded the file two times, maybe I double did that ... But okay.

Document: P23\_GMT20220909.odt Tags: Some important features unclear

I presume this does not seem to be personal data per se so I'm going to conduct data analysis.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

I am not a GDPR expert here but it should be fine.

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

If I hover over this I get additional information. Okay so recommended.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

Okay, since I do not have this information readily available I guess I will say that this is the metadata for the sake of this experiment and upload this metadata.

Document: P23\_GMT20220909.odt Tags: Some important features unclear

I am not sure what the purpose of the metadata linkage option is here but it may become obvious to me later.

Document: P23\_GMT20220909.odt Tags: Some important features unclear

Okay, so here I kind of get an option to decrease the granularity of the data. It is my feeling here.

I am interested to do it as the input, not these season-alike dates.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

I guess within an area ... The NUTS level ... I am going to open the link.

Okay, I am just going to check Sweden even if the data is not for Sweden.

Here we have the area for example.

This is I guess an area of the event where I would live.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

Since I don't know where my data is located but I could guess it is somewhere in central Europe from the coordinates.

Document: P23\_GMT20220909.odt Tags: Experimental methodology

I might just change and select the distance within a circle from the event and choose 10km as a good distance. But I couldn't really tell from that.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

The temporal units I presume are "days".

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

The aggregation method I am not sure what it means at all.

I will test with "mean" and see what you get out.

Document: P23\_GMT20220909.odt Tags: Some important features unclear

I have uploaded a small dataset and I think this is the same dataset.

I have selected what kind of analysis I want to do.

I have uploaded metadata to be included and I have selected how I want my data linkage to happen.

I am going to try and link it and I hope it works out.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

I am just going to export the output and open the zip.

We have the data and in a way I am done with task 1.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

I am opening the html but I don't think you see that.

Document: P23\_GMT20220909.odt Tags: Technical web issues

MODERATOR: Maybe, you could open it in the browser window that you are sharing.

Document: P23\_GMT20220909.odt Tags: Moderator guidance

Here we have the disclaimer again that I should not break any GDPR laws.

Document: P23\_GMT20220909.odt Tags: Data protection importance

Okay, so here there is a description of the dataset, version, publisher, licence ...

Now, I get that going back I could have filled this in with something more personal.

Document: P23\_GMT20220909.odt Tags: Experimental methodology

MODERATOR: Would you mind exploring the rest of metadata categories there on the left?

Document: P23\_GMT20220909.odt Tags: Moderator guidance

Oh yeah, here we only have the descriptors.

Document: P23\_GMT20220909.odt Tags: Complex report content

Here we have the input data with the CSV that you provided.

It is interesting because here you need to know what you are doing I guess.

Document: P23\_GMT20220909.odt Tags: Report helpful

This would be in terms of uploading the data and the disclaimer that this is considered personal data.

It is hard to know to what extent this is pseudonymised or if it is anonymised.

Because you are kind of providing a lot of information.

Say that this is a rare disease, in my case it would be vasculitis, and you know that you have an event lat/lon and a date it is hard to make things fully anonymised.

Document: P23\_GMT20220909.odt Tags: Real use case applicable, Data protection importance

Processing steps ... This is information about what I filled out before with a circle of 10 km for each event.

This is fascinating.

Document: P23\_GMT20220909.odt Tags: Report helpful

So here we have the results ...

There we go ... Here we have the events in red, and what do we mean with the datasets-used?

Are these the datasets that provide environmental data?

Document: P23\_GMT20220909.odt Tags: Visualization of plots unclear

Here there is the more complex part, I am not really into environmental research myself making it harder to understand the data.

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

Okay, here you get the information with Event A that happened at this location.

Here you have different information and you got air quality and weather data.

I was wondering if I could click on that but I can't.

Document: P23\_GMT20220909.odt Tags: Report helpful

Event A here ... I can remove or add the events.

So the first event here happened early 2011 ...

This is how my time window relates to the actual event.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

If it happens in ... I am confused about what the lag means here.

I am just going to go back to the data to see what it is.

Document: P23\_GMT20220909.odt Tags: Visualization of plots unclear

Okay so here I have the lag and the length for each event ... The time interval to gather data from ... Okay, sure.

It kind of makes sense. So we are gathering data from 14 days before the event and in some cases it is 7 days before the event.

Here I am not sure about the Y axis. Oh okay it is just the events.

So this might be a task for me?

Oh okay I can select them here ... Here I selected a distance within 10km ...

Are the steps connected somehow? Do I need to do the task here?

Document: P23\_GMT20220909.odt Tags: Improvement to text or feature

MODERATOR: This is simply telling you what happened in the background. You don't have to do anything.

Document: P23\_GMT20220909.odt Tags: Moderator guidance

Okay, now I am moving out of my comfort zone as a medical researcher going to this.

Weather and air information ...

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

It kind of provides me with information about what stands out for each variable (z-scores).

Document: P23\_GMT20220909.odt Tags: Output data ready for analysis

Okay, so either I am scrolling or I could search.

Okay, I could get the B event if I type B in the search bar.

Document: P23\_GMT20220909.odt Tags: Report helpful

We have the 14 dates related to event B. It happened on the 12 and Iget the variables that stood out during this time like an unusually high level of CO a few days before the event.

Lower temperature than usual a few days before ... PM10 stands out 12-13 days before ... It rained an unusual amount these days and we had higher air pressure three days before ...

This is really fascinating actually.

Document: P23\_GMT20220909.odt Tags: Report helpful

What are the standards in relation to? I mean I understand but, what is the normal value here? Is it about my dataset here or doing some kind of ... What is in relation to?

I think I can get this information from the metadata section Iguess.

Document: P23\_GMT20220909.odt Tags: Report helpful

I had jumped these parts from the metadata section before.

Document: P23\_GMT20220909.odt Tags: Complex report content

Here we have the data structure ... I am not sure what a data cube is, oh just a vocabulary. It is useful to have the link there.

Okay, the normal values are the monthly means and where the normal values are found.

They are the monthly mean over a time period of 10 years specified here and the same for the sdev. Okay, fair enough.

Here Iget information about the variables so I am interested in clicking the link about PM.

I get some information about that I guess. I get information about the units, notation, measurement equipment ...

Document: P23\_GMT20220909.odt Tags: Report helpful

I am looking at the data sharing now ... Here we get information about my dataset that I did not read.

Here it is what I said about the dataset being pseudonymised.

Maybe I broke the law here by uploading it hahaha Oh well.

Document: P23\_GMT20220909.odt Tags: Report helpful

Out of curiosity I will look at the other events.

Yeah, event B is in another time period. B looks like it was in the summer.

Here we have the temperatures ... Cool, fascinating.

Document: P23\_GMT20220909.odt Tags: Report helpful

Here we have a lack of data for C and Ican go back to the map and see what kind of data we had available.

We have air quality and weather data but for some reason I do not get temperature data for the C.

I guess this data is missing for some reason. Same for D here. I am going to look at all the data.

Okay so it is just missing data.

Document: P23\_GMT20220909.odt Tags: CSV output with missing data

It is very information heavy. It is also good but overwhelming I think.

Document: P23\_GMT20220909.odt Tags: Complex report content

Okay, here you get some kind of extra information about how I can interpret the data in a way.

What are the effects of the air variables or how they are related to each other.

I will just look at the pollution.

Document: P23\_GMT20220909.odt Tags: Report helpful

It depends on what kind of event we are dealing with here. If we are thinking about vasculitis the kind of incubation period is not really well known.

You will have to know how stationary your event was because it will affect this PM since it is combustion sources that would differ immensely in a few metres.

If I am on the highway it will be much higher than if I am a few metres from it I presume.

I will have to take that into account when interpreting the data.

Document: P23\_GMT20220909.odt Tags: Real use case applicable

Okay we will see, let's look at the plot here.

Here we have the concentration of a number of particles.

In the background Ihave the wind speed. So here I guess I can remove the ones I am not interested in.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

For example I could look at CO.

I am bad at understanding graphs ... This is the lag in days ... What event I'm looking at here?

What am I actually looking at here?

Is this going to change based on what I am selecting here ... No, I don't think so.

Document: P23\_GMT20220909.odt Tags: Visualization of plots unclear

This is just a visual in relation to an aggregate of all events.

Okay so this is the mean temperature.

Solar radiation would be interesting in vasculitis in a way.

We have higher solar radiation here ... Okay.

Here we have the boxes that we talked about earlier.

I can see we get some kind of medians ...

Okay, so this is an aggregate of all the events.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

13 days before the events we have an unusual narrow distribution.

Maybe this is just a low concentration of NOx.

Fair enough, the purpose of this is not for me to interpret the data, I am just curious.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

Okay, so this is Ozone and solar radiation which should go together somehow.

Document: P23\_GMT20220909.odt Tags: Real use case applicable

I should be better at statistics ... I am not sure about sample variability and z-scores.

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

Okay, here we have things that are standing out beyond a standard deviation of 2.

This line here ... NO2 is higher there and at those times the wind speed is also high.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

So here we can look at one event after another.

Okay for that event in August I get information about the particles and how it relates to the wind speed, temperature and other weather information.

How much it has rained ...

Okay, cool. Fair enough.

Document: P23\_GMT20220909.odt Tags: Visualization of plots clear

It depends ... Disclaimer here: I am not conducting environmental data as such or dealing with relapsing events or using environmental data to predict events.

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

I am maybe not the best kind of researcher to answer that but given time and everything and if I knew more about environmental data and I had a real hypothesis or something that I wanted to explore in terms of the particular particle that I would be interested in or weather variable, I think I would be quite confident in using this.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data, Researcher's expertise

Then I would be quite interested in knowing if I could export the raw linked data, which I guess I might have got in the zip file.

Yes, that is the raw data in the CSV.

Document: P23\_GMT20220909.odt Tags: Output data ready for analysis

My general feeling is that the html gives a good overview, like a shiny app in R, where you can explore the data in a way.

It is almost like this could be used for hypothesis building.

Document: P23\_GMT20220909.odt Tags: Report helpful

It gives you a good overview and I think I could use the data from the CSVs for research.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data

I am going to say that I have finished discussing how confident I am.

I would say that I am fairly confident.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data

Again, I am not really doing this kind of research so it is hard for me to know.

Document: P23\_GMT20220909.odt Tags: Researcher's expertise

I don't think I would need extra features or information.

Document: P23\_GMT20220909.odt Tags: Output data ready for analysis

To me it seems that I get a lot of information from the metadata about where the data Iwork with is coming from so I could always go back to the sources if I wanted to know more about the process of how data has been collected and has been linked.

Document: P23\_GMT20220909.odt Tags: Report helpful

So I don't think I would need anymore data.

Yeah, I don't think so.

Document: P23\_GMT20220909.odt Tags: Output data ready for analysis

I think when I get to the SERDIF UI I get a lot of information like an example use case which Ithink is good.

I am not sure to what extent this figure makes sense to me, I think it does.

I get information from what areas we have data to link.

I get an overview of the linkage process ... Sure thing.

I think this is interesting, it makes sense.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data

You have the event and you need a standardised CSV, which is fairly straightforward, with variables that make sense.

You kind of know the format of the data and what you need.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

In step 2, I am curious about the purpose here ... Fair enough Ican fill that up in here, it makes sense.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

If I was doing this for real with my subject level data I would be interested in the linkage process to understand what happens with my data when it is uploaded here.

I think that is good information to know here but it is probably somewhere here.

I know from experience that all of this is kind of shaky to use real data.

Document: P23\_GMT20220909.odt Tags: Real use case applicable, Suggestions for additional features

This way I could test it with different levels of granularity.

I am thinking about vasculitis. I would probably do it at the regional level and not with distance circles and probably increase the temporal units.

Document: P23\_GMT20220909.odt Tags: Real use case applicable

I am not really sure what the aggregation method is here and how it will impact the data that I get back.

I could always test it if I wanted to.

Document: P23\_GMT20220909.odt Tags: Real use case applicable

I get confirmation of the linkage process and it is easy to export.

Document: P23\_GMT20220909.odt Tags: Data linkage process clear

I am going to the html without looking at the CSVs, and I think this makes sense.

Document: P23\_GMT20220909.odt Tags: Report helpful

You had to tell me verbally to look at the metadata subheadings and click those but I might have been able to figure that out.

Document: P23\_GMT20220909.odt Tags: Complex report content

Then this gives me some information about how I should interpret the data.

This gives a good visual overview.

Document: P23\_GMT20220909.odt Tags: Report helpful

I think my overall experience doing it for the first time in a video was fairly simple and intuitive.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data

If I had to do it again it would be even simpler.

Document: P23\_GMT20220909.odt Tags: Improved tool usability and output data

So you have the date, temperature and some other environmental data.

Document: P24\_GMT20220923.odt Tags: Output data ready for analysis

I will try to edit one of these cells.

Just thinking that when you can only see the number after typing and you cannot visualise while you are typing.

Oh, it is gone now. If you click on any of these things you cannot see the numbers.

Document: P24\_GMT20220923.odt Tags: Technical web issues

I do not have any to import here, well it is my understanding.

So I am just going to use what you attached to my email, is that right?

Document: P24\_GMT20220923.odt Tags: Experimental methodology

MODERATOR: You could either use the one in the email or use this one here, up to you.

Document: P24\_GMT20220923.odt Tags: Moderator guidance

Okay, I will use this one here but I will add a row.

Let's say with event E, let's make up a number for lat/lon.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

This one did not save my number when I typed it in.

If I click out of the box it does not save the number the first few times. I will try it again to understand why this happens.

If you click out of the box it does not save if but if I click enter it saves the value.

Document: P24\_GMT20220923.odt Tags: Technical web issues

I would have thought that if I clicked out of the box it would save the value.

Again, it did not save that one when I clicked out of the box.

Document: P24\_GMT20220923.odt Tags: Technical web issues

I was going to say that you should include the format of the date to type but it is there when you hold over the column heading.

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

So that is okay, maybe it would be nice to see the format that is behind date.

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

Again, if you click outside of it it does not save the number.

Document: P24\_GMT20220923.odt Tags: Technical web issues

That is interesting, if you put a wrong date only with month and year, it still allows you to proceed.

I wonder what happens if I put in a date that does not exist. Even non-real dates exist.

You should probably add some limits there.

Document: P24\_GMT20220923.odt Tags: Suggestions for additional features

And again, I need to click return or enter to save the value there.

Document: P24\_GMT20220923.odt Tags: Technical web issues

Length ... What is this? ... Okay ...

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

There is a lot of reading here, a lot of bits to read.

Document: P24\_GMT20220923.odt Tags: Complex report content

Spatial distance ... Within an area or distance ... I am not sure what this refers to.

Document: P24\_GMT20220923.odt Tags: Some important features unclear

Distance in km from the event point.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

Temporal units, I will go for "days".

Aggregation method, okay.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

Okay, "upload datatable" I had to do that. Okay.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

So now I want to link it but Icannot click it.

I did upload it ... but yeah, here we go I forgot these options.

Document: P24\_GMT20220923.odt Tags: Some important features unclear

Oh, it says here that you need to click enter.

Did it say that on the previous table? Oh, yes, it did. I did not read that before.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

Perfect ... subject with a rare disease ...

Upload metadata

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

Okay, export please. That works. I will go to my downloads there.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

I have awful lots of documents now but I will open the report maybe.

Document: P24\_GMT20220923.odt Tags: Output data unclear

Yes, I opened the report kind of naturally.

So that is done ...

Let's see what this report says.

Document: P24\_GMT20220923.odt Tags: Report helpful

There is a nice data protection piece

Document: P24\_GMT20220923.odt Tags: Data protection importance

There is a typo there "intent" not "intend".

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

So there is a lot of environmental data here. That is fine.

Then we have a data table.

So just looking at this, I have some kind of event and the environmental data which I can understand easily enough.

Document: P24\_GMT20220923.odt Tags: Output data ready for analysis

Okay, so this is a person in Dublin for example, but I don't see the lat/lon here ... And these are the environmental characteristics of that place at that time period.

Document: P24\_GMT20220923.odt Tags: Output data unclear

Okay ... I am not sure what the ... Okay, all of the environmental data presented in one chart.

Double click on the legend ... There was a hint for me there on the right-top corner that now has disappeared.

Mmmm ... Not sure why it is gone.

Okay, so Ican include or exclude various data points.

Let's look at Nitrogen and Iam going to unclick all those features.

So this is the Nitrogen level over time.

Document: P24\_GMT20220923.odt Tags: Visualization of plots clear

I'm not sure why there are different bars of colour.

Document: P24\_GMT20220923.odt Tags: Visualization of plots unclear

I'd suppose that is the average wind speed on a day. I am not sure what this grey refers to ... Background wind speed and it says 20 ...

Document: P24\_GMT20220923.odt Tags: Visualization of plots unclear

I'd suppose I do not have a lot of experience with the environmental data per se but anyhow if I click on these charts ...

Document: P24\_GMT20220923.odt Tags: Researcher's expertise

Okay, so I will look at NOx, there it is. Okay, so I guess the box plots just show the error bars, the variation in concentration throughout the day.

Document: P24\_GMT20220923.odt Tags: Visualization of plots unclear

I am just looking back here to see what the z-score is again.

That is funny there, those ones that only have little pieces of information there.

It is the PM2.5. I suppose there is missing data.

Document: P24\_GMT20220923.odt Tags: CSV output with missing data

Okay, so I suppose A, B, C and D here are the different lat/lon points that we have chosen.

There is no legend ... Oh there it is if you go a bit further down.

Oh look the little hint appeared there again saying that I can double click the legend to make it appear or disappear.

I am going to stick with NOx for the time being.

It is quite clear from the greying out which is being included and which not, that is quite intuitive.

Document: P24\_GMT20220923.odt Tags: Visualization of plots clear

So what I am interpreting of these background colours is that on day 16 there was a wind speed of 2m/s.

Document: P24\_GMT20220923.odt Tags: Visualization of plots clear

The colours are very similar. I actually don't think that the colours on the background match the colours in the gradient here.

That is because the background colours are transparent and I don't think there is any patient with this royal blue in the background.

Otherwise you would not be able to see the lines in front

Document: P24\_GMT20220923.odt Tags: Visualization of plots unclear

So maybe the colour in the legend on the right needs to match the actual colour on the background of the plot.

Particularly if they are matching colours.

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

We have linked ... I am trying to work out what I have actually linked.

We have linked longitudinal data, like location based data. Let me just go back to the home page.

Okay, lat/lon and date based data and Ihave linked what it seems like environmental things like the air quality or precipitation.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

This process was relatively easy to do for my dataset.

Document: P24\_GMT20220923.odt Tags: Improved tool usability and output data

I have the environmental data.

Yeah, I thought that was relatively straight forward.

Document: P24\_GMT20220923.odt Tags: Improved tool usability and output data

I mean there is a lot of text on these pages.

Document: P24\_GMT20220923.odt Tags: Complex report content

I am just wondering if there are good explanations for what each part does and you can hover over them, so that is good.

All of my buttons worked, at least the ones Ihave used.

Document: P24\_GMT20220923.odt Tags: Improved tool usability and output data

NUTS, I have not heard of that before ... Okay, so I will click on Ireland. Iwonder what this NUTS thing is.

Document: P24\_GMT20220923.odt Tags: Some important features unclear

Nomenclature Units of Statistics ... Okay. I will go back to my page now.

So I am NUTS-2 because I am in Ireland.

Document: P24\_GMT20220923.odt Tags: Some important features unclear

So I have a date and I have a lag, and the lag has to do with the number of days.

But, why don't I have the lat/lon recorded in this excel file?

Document: P24\_GMT20220923.odt Tags: Output data unclear

It looks like I can edit that as well so that is fine.

What else do I have here? Okay, so we have lat/lon, date, length, lag, point, datause, query ... It is getting very academic there now, alright

Document: P24\_GMT20220923.odt Tags: Output data unclear

MODERATOR: Now you have completed task 2. Would you mind completing the rest of the tasks by answering aloud the questions posed on the tasks?

Document: P24\_GMT20220923.odt Tags: Moderator guidance

I am looking at the report with a nice navigation bar at the top ... I am going to try the search box here.

I will search for "health" ... Oh, let me try that again ...

So I have searched for the word "health" here but I have not actually done anything for me that Ican see.

How about I go pollution? Okay, Iam not sure that the search function is working.

When I click enter or return it does not search for anything there.

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

To be honest, I would need to know exactly what this data is that Iam looking at.

But I would have thought that on this data table you would have included both the environmental information and the lat/lon information.

Document: P24\_GMT20220923.odt Tags: Output data unclear

Okay, hovering over the column names gives me the units, that is good.

It is certainly presented in a nice way.

Document: P24\_GMT20220923.odt Tags: Report helpful

Yeah I think that if it had the lat/lon information there would be important.

Just even just to compare the old dataset with the combined/linked dataset.

Document: P24\_GMT20220923.odt Tags: CSV output with missing data

It is relatively intuitive.

I have no idea what I was doing when I opened this up or what I had to do.

Document: P24\_GMT20220923.odt Tags: Improved tool usability and output data

I briefly read the instructions and I certainly did not read all of the instructions on these pages because It is very detailed.

It is relatively intuitive.

Document: P24\_GMT20220923.odt Tags: Improved tool usability and output data

I uploaded the health event data, I could do that.

I added a row and it was no problem.

Document: P24\_GMT20220923.odt Tags: Data linkage process clear

Drag and drop this ... I cannot drag and drop this but I am sure I can drag my own document there.

I will try to drag and drop something there.

I am not sure if that worked there. It does not seem to be refreshing when I upload anything.

Document: P24\_GMT20220923.odt Tags: Technical web issues

The data table was successfully uploaded but it does not cover the second time.

It does not seem to have to work for the new one.

Document: P24\_GMT20220923.odt Tags: Technical web issues

When I clicked "link" there it brought me to the top, almost like I missed something.

Document: P24\_GMT20220923.odt Tags: Improvement to text or feature

I am not sure what is going on there but I might need to refresh the page if I wanted to submit a new query.

I will see if this works now ... Upload metadata ... Link, okay I see that Ineeded to refresh the page.

Document: P24\_GMT20220923.odt Tags: Technical web issues