# Taguette highlights

Okay, I read this and I think the text is clear.

Document: P2\_GMT20211119.odt Tags: Home tab clear

I am not exactly sure what all the images are for or whether they are there to help me.

Document: P2\_GMT20211119.odt Tags: Wording not clear

I think that's everything I get that I get to be held to support to do my queries.

Document: P2\_GMT20211119.odt Tags: Home tab clear

Now, I am going to log in with the credentials provided.

Okay, I am in.

Document: P2\_GMT20211119.odt Tags: Log in easy

I am going to select definite events for people in Ireland.

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

I am going to go for Dublin.

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

Length of days, I am going to put in ... Time window length, so I am going to go for a seven day period.

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

Time window lag, so I think this means either side of that.

Document: P2\_GMT20211119.odt Tags: Query process complex

(The participant reads the temporal unit tool tip)

I don't know what that is. Oh sorry, maybe that is at those temporal windows there.

Document: P2\_GMT20211119.odt Tags: Query process complex

I am going to try to submit that but I am not doing something right.

Document: P2\_GMT20211119.odt Tags: Query process complex

No events, okay. I think I have ... Maybe if I change this.

Document: P2\_GMT20211119.odt Tags: Query process clear

Let's see what I did there. Oh sorry, now I am getting this.

Temporal units, I am gonna go by the day

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

So let me get back to my tasks.

(The participant reads the task again)

Document: P2\_GMT20211119.odt Tags: Task sequence complex

I was not sure whether I was inputting a query.

Document: P2\_GMT20211119.odt Tags: Query process complex

Maybe just something like enter a query, even though you have query inputs up there.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion, Additional explanation

So this is my query and I was looking for definite events and here I have environmental information.

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful

(The participant reads the column heading descriptions from the data table)

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful, Task sequence complex

I am not sure what NOx means.

Document: P2\_GMT20211119.odt Tags: Additional explanation

(The participant reads the task again)

Document: P2\_GMT20211119.odt Tags: Task sequence complex

When I was reading it, I was expecting that the data provenance, lineage, use and limitations would have a button for each. So that surprised me a bit.

I want to explore them now.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Each of the three buttons but I got five buttons.

Document: P2\_GMT20211119.odt Tags: Task sequence complex

I am seeing example events Ireland, Wexford and 4 returns, so that is the provenance.

Document: P2\_GMT20211119.odt Tags: Metadata content clear and useful

I am getting this information here, which I am not sure what it means.

Document: P2\_GMT20211119.odt Tags: Metadata content not useful or confusing

Okay, here. So it is telling me where all the metadata is coming from I think.

Document: P2\_GMT20211119.odt Tags: Metadata content not useful or confusing

FAIR metadata export, I think I am not getting how this is all changing.

Let me just go back.

Document: P2\_GMT20211119.odt Tags: Navigation complex

So that is the provenance, so when I close that it goes away. The same for the lineage and full metadata.

Document: P2\_GMT20211119.odt Tags: Navigation complex

I would have expected not to have all of them appear at the same time, but when I click another one the previous would disappear.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Maybe also a button to show them all together.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

FAIR metadata export, I am getting a file coming down here.

Document: P2\_GMT20211119.odt Tags: Simple export

I still have this text area showing.

Let's see ...

Document: P2\_GMT20211119.odt Tags: Navigation complex

Let's see ... (The participant reads the sub task again)

Document: P2\_GMT20211119.odt Tags: Task sequence complex

I think there was only one. Oh, no I better go back and take a look.

How many data sets were used? Three data sets.

Document: P2\_GMT20211119.odt Tags: Metadata content clear and useful

Okay, the data license is there and it is a creative commons license.

Document: P2\_GMT20211119.odt Tags: Search feature simple

The data use is just rdfs:comment. It is an rdfs about the data use but not telling me anything about it

Document: P2\_GMT20211119.odt Tags: Metadata content not useful or confusing

I have to ask about the identification, so this should be an interesting one. Where is my identification risk?

Okay, so the information seems to be here but it may be better to have it summarized in a better particular way so that you can expand it in more detail. If I want to go further.

Document: P2\_GMT20211119.odt Tags: Search feature simple

Okay, so the information seems to be here but it may be better to have it summarized in a better particular way so that you can expand it in more detail. If I want to go further.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

Okay, so I can't find the risk.

MODERATOR: Do you want to try to use Control+F?

Oh, you told me that.

I am using Control+F but it is not doing anything.

MODERATOR: Control+F opens a little box on the top and you can type what you want to look for?

Okay, so risk.

Document: P2\_GMT20211119.odt Tags: Search information process complex

MODERATOR: Do you want to try to use Control+F?

Document: P2\_GMT20211119.odt Tags: Moderator comment

MODERATOR: Control+F opens a little box on the top and you can type what you want to look for?

Document: P2\_GMT20211119.odt Tags: Moderator comment

So this has risk ... Data sets storage ... I am not getting an understanding of what the risk is.

MODERATOR: Maybe you could use the arrows in the Control+F search box at the top right corner of your screen.

Document: P2\_GMT20211119.odt Tags: Search information process complex

MODERATOR: Maybe you could use the arrows in the Control+F search box at the top right corner of your screen.

Document: P2\_GMT20211119.odt Tags: Moderator comment

Okay, now I got some description here now. It took me a while to find that. There is something there that is telling me.

So my answer is that it could be just clearer or I could be faster.

Document: P2\_GMT20211119.odt Tags: Search information process complex

Now I am going to click on these extra buttons and I am going to get rid of that one, ups.

I don't know why that happened.

Document: P2\_GMT20211119.odt Tags: Navigation complex

So do I understand the columns?

(The participant reads the column heading descriptions)

So yes, they are clearly the ones I got there. If I am an expert in all of those things, I would know.

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful

Maybe if I am an expert I would like to know more information, maybe ranges or something if that is relevant.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

So yes, I guess I can find this information.

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful

So I am not told. Instinctively, I would say that the red ones are warnings or that they are higher of a particular range of something. The blue ones might be lower then.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Maybe the range would have been nice to have up here if that is the case.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion, Additional explanation

So we can hide them. I am not quite sure where the toggle columns button is but I can use the little eye to do that.

Document: P2\_GMT20211119.odt Tags: Navigation complex

I suppose ... but wait, how do I get the hidden columns back? I don't know if I have them all displayed now because I got rid of some of them.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Maybe I go and I do this. I am not sure how I am going to get them back or maybe they are all there because I don't remember.

Oh no because I don't have the event column.

Document: P2\_GMT20211119.odt Tags: Navigation complex

MODERATOR: Maybe try the toggle columns button.

Oh, very good. Now I found it. Okay, so I am making them disappear, not making them appear. Okay I got all the columns now.

Okay, so that is obvious to me now and that is a handy feature.

Document: P2\_GMT20211119.odt Tags: Data table features useful, Navigation complex

MODERATOR: Maybe try the toggle columns button.

Document: P2\_GMT20211119.odt Tags: Moderator comment

I didn't notice these up here. I don't remember seeing that before

Document: P2\_GMT20211119.odt Tags: Wording not clear

(The participant reads the data background paragraph).

Okay, so the red are high and the blue are low, so what I expected them to be.

Document: P2\_GMT20211119.odt Tags: Data table features useful

So now, Can I make sense now? is my question.

So I have an event with a particular name and I have all of this data here. I am understanding that on this particular day this event occurred and these are all the environmental factors that happened there. These are now all linked together from different data sets.

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful

I did have three data sets but I am not sure which environmental variable comes from which data sets but I know that they are linked.

Document: P2\_GMT20211119.odt Tags: Metadata content not useful or confusing

So, maybe the data set that it came from might be a useful thing to know. I don't know how you can mark them but that would be something useful.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

Maybe, these ones being red, I don't know about the significance of them being red but maybe that is up to the user to interpret that. I can see that these are red for the data set for that column and that event. But what does it mean for the event? is the whole row affected or is there a way of signaling that row instead of just the cell?

Document: P2\_GMT20211119.odt Tags: Navigation complex

This data table with all of these events that occurred on the previous slide. Now I am hearing relative time from the event (day).

Relative for the event (day) ... Relative time from the event day -2 and event-530 and the range is there.

Document: P2\_GMT20211119.odt Tags: Visualization useful

So again, if there was a high and maybe this would come across, but it would be nice to see why I am getting these events or maybe it is all that it exists.

Document: P2\_GMT20211119.odt Tags: Event concept and approach not clear

I am going to pick rain. Okay so I just got four events here and I had four events there.

Then, I have four events there. So there are four events.

Document: P2\_GMT20211119.odt Tags: Visualization useful

So the events are in different orders in all of the screens.

Document: P2\_GMT20211119.odt Tags: Navigation complex

I am not sure what these types of things are used for.

Document: P2\_GMT20211119.odt Tags: Plot design complex

I am not sure if the events are supposed to be associated with this or this are just for the time range that I am in.

Document: P2\_GMT20211119.odt Tags: Plot design complex

It is loading up and I am going to read the tasks again.

(The participant reads the task again)

Document: P2\_GMT20211119.odt Tags: Task sequence complex

I wonder if it would be nice that you could look at a collective of the three of them together maybe for 1 event if that was useful.

For example, if you were to choose event-58, you could get three plots together for that event.

I think I would like to add on this as well what the data sets were that contributed to this.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

So I think that the data sets would be interesting to see and maybe something about the color coding on what the different things mean.

I would say you would get used to it quite quickly with all the different symbols once you started using it.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion, Additional explanation

Back to it now, so I am used to this by now.

Document: P2\_GMT20211119.odt Tags: Overall positive experience

I will input a new query for Ireland but this time I am going to look at no events.

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

I will go for 2 days and I got that ...

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

Okay, so I have 23 data sets here ... Oh sorry, no data available.

Document: P2\_GMT20211119.odt Tags: Query process complex

So I am going to add Mayo in the locations.

I am not getting any data ... Oh because I already got definite up there.

Document: P2\_GMT20211119.odt Tags: Query process complex

Okay so the events I am looking for are no events in all of these places ... Why am I not getting to input the temporal unit?

Document: P2\_GMT20211119.odt Tags: Query process complex

Oh sorry, that is just telling me there. I was getting confused between two things here and the two things there.

Document: P2\_GMT20211119.odt Tags: Query process complex

This is just giving me information I think.

So I have 20 and I don't need all of them, let's remove ...

Okay, it is still saying that.

I am going to submit this ... Mmmm.

MODERATOR: Maybe try clicking data availability.

Oh yes, that is right, very good. So Now I get the temporal unit and the aggregation method.

Document: P2\_GMT20211119.odt Tags: Query process complex

MODERATOR: Maybe try clicking data availability.

Document: P2\_GMT20211119.odt Tags: Moderator comment

So I am going for days and mean again.

Document: P2\_GMT20211119.odt Tags: Query inputs and elements useful

So the last one I got 3 data sets and I have 3 data sets now.

The data sets are all the same for all of the events. It looks like there are 3 data sets again for all of those events.

Document: P2\_GMT20211119.odt Tags: Metadata content clear and useful

So the last one was created commons. So I need to get rid of data provenance and then go to this button.

I should then have a window here to search for the license.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Ups ... How do you spell license?

Okay I am struggling now where this license is but I am determined.

Document: P2\_GMT20211119.odt Tags: Search information process complex

MODERATOR: Maybe you can read one of the previous tasks to find out were the license was the first time.

Document: P2\_GMT20211119.odt Tags: Moderator comment

Maybe I just say that I cannot find it.

Document: P2\_GMT20211119.odt Tags: Search information process complex

Okay, so that was in the full metadata exploration.

Document: P2\_GMT20211119.odt Tags: Navigation complex

I can see that this is creative commons. So I did not read the other one properly but I do know it was creative commons but I can't remember the type of creative commons license because I didn't pay attention there. I only read the creative commons piece.

Document: P2\_GMT20211119.odt Tags: Search feature simple

I think Q2 might have more but again I did not pay attention to that but I think I can find that out easily.

Document: P2\_GMT20211119.odt Tags: Data table features useful

It might be interesting to see about those types of things, like the license, to have a little icon to say things like what is the license, the identification use. You know like a little symbol on it. That would be better than having to go in and look.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

I am going to look at the heat map and select rain again.

So I have got more events in this than I had in the other one. I am not really sure of what the differences are in terms of the data. I am not thinking of the data when trying to compare as it is not my area of information.

I did not pick up a pattern in the last one.

I can compare here.

Document: P2\_GMT20211119.odt Tags: Visualization useful

I wonder because you are asking me about comparing queries, I wonder if it would have been nice to see more than one query on at the same time.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

If that is something that is ... I suppose that then you would have to have many information about the different queries even though we are here. We would probably have to show them on the right hand side as well.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

I think I did that already but I can go back to the first query and have a look at it.

Document: P2\_GMT20211119.odt Tags: Task sequence complex

Okay so I did that.

Document: P2\_GMT20211119.odt Tags: Simple export

How do I open this now? Okay, so I can see it there.

Document: P2\_GMT20211119.odt Tags: Export name and format unknown

MODERATOR: No problem, you can go ahead, no need to explore that.

Document: P2\_GMT20211119.odt Tags: Moderator comment

I think I also did that.

Document: P2\_GMT20211119.odt Tags: Simple export

I am not sure what I am getting when downloading it all. Is it just for this particular query or it is just for the two queries I did?

So maybe add something that tells me which query I am downloading.

Document: P2\_GMT20211119.odt Tags: Additional explanation

Yes, I think the way it is formatted with the export I can see that it is in a format it would be useful to input. There are a lot of details in there that would be useful in there that would be useful to input into another research.

Document: P2\_GMT20211119.odt Tags: Export (meta)data useful

Okay, so I don't think I got stuck too badly.

Document: P2\_GMT20211119.odt Tags: Overall positive experience

I am just looking at the screen here and I didn't pick up that the Q1 and Q2 were there. I could have flicked between the two queries. I wasn't aware that this was available to me.

Document: P2\_GMT20211119.odt Tags: Navigation complex

Overall, I don't think I got stuck too badly on the first go with content I wouldn't be extremely familiar with. I mean I know about the concept but I am not used to the data or type of plots that would be use. I think this dashboard would be a type of user interface I could get used to.

Document: P2\_GMT20211119.odt Tags: Overall positive experience

It was quite quick and I was able to ... Once I found the buttons, I got used to using the buttons and I think I will be able to navigate it quite quickly.

Document: P2\_GMT20211119.odt Tags: Overall positive experience

The only comment then would be in terms of icons on tables or maybe doing a combination of things but that is up to the researchers what they actually would find useful in combining.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

Maybe more visibility around the characteristics of each plot or the data sets in terms of putting icons around the provenance or identification risk or something like that, as well as being able to find it. Just some summary, so as long as the summary was able to capture all of the nuances of the information behind. You don't want to give somebody a false sense of security if something is of low risk if it isn't. I know all that has to be taken into account.

Document: P2\_GMT20211119.odt Tags: Additional feature suggestion

Overall, I think the right hand side was very useful, even though I missed some things, which I did not know why I missed them.

Document: P2\_GMT20211119.odt Tags: Overall positive experience

The left hand side I think I was getting a bit confused on the order to do the things in. Whether I had to check this data availability or when I could submit etcetera. That one maybe would be the one you would not get used to. Maybe there is a sequence that is out of place there.

Document: P2\_GMT20211119.odt Tags: Query process complex

I am going to check what a knowledge graph is. I don't know if you can see my screen but I have to log in here as well. I think with the same credentials as before. I will check that later.

Document: P3\_GMT20211125.odt Tags: Additional source exploration

Okay, this all makes sense to me.

Document: P3\_GMT20211125.odt Tags: Home tab clear

I think I could understand all that is here

Document: P3\_GMT20211125.odt Tags: Home tab clear

I am not going to say ask completed because I want to look at these links.

Knowledge graph, this is graph db ... It takes a while ... loading loading.

Document: P3\_GMT20211125.odt Tags: Additional source exploration

It takes a while ... loading loading.

Document: P3\_GMT20211125.odt Tags: System issue

I will check the methodology, ah the GitHub okay.

Document: P3\_GMT20211125.odt Tags: Additional source exploration

I don't know if the graph db page is working but it is taking a while.

Document: P3\_GMT20211125.odt Tags: Additional source exploration

I guess what I was expecting when clicking on here was a description of what a knowledge graph is exactly. I am not sure I will find it here for what it says here.

Document: P3\_GMT20211125.odt Tags: Wording not clear, Additional explanation

Yeah, this is still loading and we will get stuck here for a while.

Document: P3\_GMT20211125.odt Tags: System issue

I will try to log in into graph db again otherwise I will skip.

Document: P3\_GMT20211125.odt Tags: Additional source exploration

MODERATOR: It is weird because for me it is working, feel free to go ahead with the tasks.

Document: P3\_GMT20211125.odt Tags: Moderator comment

Okay, it was loading for a while.

Document: P3\_GMT20211125.odt Tags: System issue

I think the information was enough especially to understand how this can be useful to people you want to, you know, get hold of this data, explore it somehow and search for that.

Document: P3\_GMT20211125.odt Tags: Home tab clear

Not much, I had a bit of a doubt what a knowledge graph is exactly and that is why I clicked there. Maybe adding an explanation or a link to a place, I mean you don't need to write it here.

Document: P3\_GMT20211125.odt Tags: Wording not clear, Additional explanation

Okay, this is what I did before on another website.

Okay, super fast.

Document: P3\_GMT20211125.odt Tags: Log in easy

Mmmm ... Interesting ... I guess the definite events are confirmed cases of something, example events, like a disease let's say. These events are not biologically confirmed but suspected somehow. These events are discarded. I am a weird person and I am going to choose the no, those that probably had a clinical form similar to what you would expect for this disease but then it was a no and it is something else.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

I am not very familiar with Ireland so I am going to go with Cork in the south.

Document: P3\_GMT20211125.odt Tags: Query process complex

Time window length (days) ... I guess it is the window of days that I want the data.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

I guess the lag is the lag of the data with respect to the clinical data ... I am not going to put a lot of lag.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

Temporal unit ... I have to first check if the data is available.

Document: P3\_GMT20211125.odt Tags: Query process clear

Ah okay, now I have to choose this temporal unit.

(The participant reads the temporal unit tool tip)

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

Okay, so this is an aggregation process that goes internally.

Document: P3\_GMT20211125.odt Tags: Query process clear

The aggregation method, this is hard to let the user choose this, maybe. Depends on the variable I guess. Interesting. I am going to put ... For me it is a bit confusing as I would like to know what sort of variables I am going to be selecting because I don't know.

Document: P3\_GMT20211125.odt Tags: Query process complex

I am going to select the mean as it makes general sense.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

Okay, the Q1 ... Where is the Q1 tab? I think I submitted this.

Okay, it is taking a while.

MODERATOR: I got a critical error from the hosting platform. Could you try to refresh the dashboard and re-submit a query? Could you also try a different location?

Okay, I will try for Dublin.

Okay, so 20 days with a 5 days lag.

Data is available.

Ah okay, it is fine.

Document: P3\_GMT20211125.odt Tags: System issue

MODERATOR: I got a critical error from the hosting platform. Could you try to refresh the dashboard and re-submit a query? Could you also try a different location?

Document: P3\_GMT20211125.odt Tags: Moderator comment

Yeah this is the Q1 tab. I wonder what Q1 means. Is it like query 1? Okay.

Document: P3\_GMT20211125.odt Tags: Wording not clear

I guess if I do queries, I will have other tabs here.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

Okay, so if I go here ... Event type no, yeah as I have chosen the no. This is the query I guess.

The data sets used are here, which should have the same dates and the lag I guess.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

I put the lag of 5 days, but it says 2021-10-12 and it goes from ... Ah no it is the same file. It changes only this EEA-28A ... Okay.

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

(The participant reads the sub tasks again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

This is the data provenance and these are the data sets where information comes from, but I do not have a clear idea of what they mean.

They all have the same date which is a bit weird to me.

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

Data lineage ... I guess it is this here. This was the provenance and the data lineage is like code. Probably, it is what is going behind it, like the exact ...

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

This is called data lineage, maybe the name is a bit confusing as I don't know what the lineage is exactly.

Document: P3\_GMT20211125.odt Tags: Wording not clear

Okay, this is what is behind it.

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

For me it is not useful to look at because it is a bit ... I guess for a clinician it is even less useful because they cannot read this code.

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

Are these buttons to export things? Yeah, I can download stuff okay.

Document: P3\_GMT20211125.odt Tags: Simple export

(The participant reads the sub task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

Well I guess it is the number of data sets here I guess 25 or 30.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

(The participant reads the sub task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

I guess I would do a Control+F and search for the words and let's see if it works.

I am not sure if these things work here ... Oh, yeah.

Identification risk ... (The participant reads the identification risk comment) I don't know if this is the paragraph they are asking me to find but okay.

Document: P3\_GMT20211125.odt Tags: Search feature simple

The license must be here at some place. Maybe it is in the upper or lower part.

Document: P3\_GMT20211125.odt Tags: Search information process complex

A this is the data set, ah no the data use. Yeah it is here ... (The participant reads the data use comment)

Document: P3\_GMT20211125.odt Tags: Search feature simple

I am missing the license, which I did not find. Ah, I am writing it wrong, yeah.

Document: P3\_GMT20211125.odt Tags: Search information process complex

Well here there is a license at the end, this is a creative commons. There are these types of creative commons ... okay ... this means, oh yeah I knew that. Yeah, yeah, one of these types.

Document: P3\_GMT20211125.odt Tags: Search feature simple

The dates are days, so this is the aggregation I choose. I guess the aggregation was before and this is data displayed by dates.

I think the event is the type of the event in the format of the database.

Here I have data about pollution, so this is all air pollution data.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

(The participant reads the column heading descriptions)

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

Here if I hover over it does not say what it is. In some it says for example particulate matter but in others it does not say.

Document: P3\_GMT20211125.odt Tags: Additional explanation

There are some values here in red and some in blue. Intuitively, I would say that red values are bad values or something. Maybe values of air pollution over a threshold. I think I am right because this value is the highest in the column ... Oh, actually, this is the highest value in all the columns ... Oh no, I think it is values over a threshold by column, I don't know. Maybe WHO recommendation threshold or these things, I am not sure.

Document: P3\_GMT20211125.odt Tags: Wording not clear, Navigation complex

(The participant reads the sub task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

I was going ahead with the task. Actually here there is the precise meaning of what they are.

(The participant reads the data background paragraph)

Values higher values of 2 ... Well ... The Z-score, okay, extreme values of the distribution, okay.

High are red and low are blue.

Document: P3\_GMT20211125.odt Tags: Data table features useful

Okay, in the beginning maybe I was expecting a climatologically sound threshold like coming from another thing but I see it is about the distribution of the data cutting some things, well marking some of the edges. Yeah, like the 95% interval for the data.

Document: P3\_GMT20211125.odt Tags: Data table features useful, Wording not clear

If I go like this, they go away. I can end up with a super small data set.

Document: P3\_GMT20211125.odt Tags: Data table features useful

The toggle is this, okay, I can say which ones I want and which ones I don't.

Document: P3\_GMT20211125.odt Tags: Data table features useful

(The participant reads the sub task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

My only doubt is ... The events here, how many events per day are there? I guess there are days where there are more events than others. The data table is data per day but the events are just the same thing in all these columns. I guess these days are 0, 0, 0, 0, 1, 10. I don't know.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear

For me this is very complete regarding the environmental data and the pollution and climate data, that is super useful.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

But I am not sure how I will use that. I would need more data about the disease or these events. I guess I cannot see it because this is private information or something. I would like to have a column that has the number of events per day or something like that and so on.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear, Additional explanation

I guess this data is lagged as this is what I have asked for in the end. I guess this event here and the data for this time, the dew point temperature is the lagged one for 5 days. I am not sure. Maybe I would like to see this here somehow. This is data that is lagged 5 days or something like this.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

I mean I think to explore the relationship with the event and some of this data with the table I cannot do it. If it was myself I would download it and do it myself. Just plot some variables or do some correlation analysis first step thing so see the connection to the events or something like that. I think there was a dashboard here somehow, they said before. I can do it myself by downloading it or this web page can do it also a little bit. Some plots of the data and so on.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful, Visualization useful

If it was myself I would download it and do it myself. Just plot some variables or do some correlation analysis first step thing so see the connection to the events or something like that. I think there was a dashboard here somehow, they said before. I can do it myself by downloading it or this web page can do it also a little bit. Some plots of the data and so on.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

I think the information is ready to assess this connection, which is good. Most of the time what is more consuming is to get this table because there are these variables and all these sources. I mean now it is very easy as I can just go here and put this and retrieve the data, which is great.

Document: P3\_GMT20211125.odt Tags: Simple export, Export (meta)data useful

I mean now it is very easy as I can just go here and put this and retrieve the data, which is great.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

For me this is more than useful, as a user I would download it and I will look at things myself. But I guess maybe some users that are not so comfortable with programming maybe they would like to see it here as an interactive dashboard and so on. This depends on the preference of the user I guess.

Document: P3\_GMT20211125.odt Tags: Simple export, Export (meta)data useful, Visualization useful

I can download it as I can see here from the export button, so all good.

Document: P3\_GMT20211125.odt Tags: Simple export

No, in the table no.

Document: P3\_GMT20211125.odt Tags: Data table features useful

Heat map, heat map ... Where is the heat map?

Do I have to select one variable? No, the arrow in the data table would order things differently.

Here is a heat map ...

Document: P3\_GMT20211125.odt Tags: Navigation complex

Maybe now I would like to close the metadata windows. If I click again they go away, okay, nice.

Document: P3\_GMT20211125.odt Tags: Navigation complex

If I select temperature or O3 is not working ... PM10 ...

MODERATOR: It should be working.

I can refresh if you want.

MODERATOR: It is weird because I am not getting an error. Usually, I have an error regarding garbage collection but this time it is not that. Would you mind refreshing the page?

Yeah, I will refresh.

Document: P3\_GMT20211125.odt Tags: System issue

MODERATOR: It should be working.

Document: P3\_GMT20211125.odt Tags: Moderator comment

MODERATOR: It is weird because I am not getting an error. Usually, I have an error regarding garbage collection but this time it is not that. Would you mind refreshing the page?

Document: P3\_GMT20211125.odt Tags: Moderator comment

Okay, I can choose more than one event type ...

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

Oh, no events now. Ah, it says there are only two events here. This is hard to relate with the ... Oh there are just a few events ... Mmmm ... Interesting.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear

Can I increase this? No, events. This is a very uncommon disease I guess. Now there is just one event. If there is just one event, it's a bit hard to link.

Document: P3\_GMT20211125.odt Tags: Query process complex

Okay, ah. Now I understand what was happening before.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

Okay, so we can go to the heat map.

It is not working again.

MODERATOR: Can you try the box plot?

Let me try another browser, I will try Chrome.

MODERATOR: Nice thought, you are my first safari user. It is good that you point that out because the previous participants used Firefox or Chrome.

Yes, that could be the issue. I am going to go with Chrome.

MODERATOR: That could be the reason why the knowledge graph link was not working either.

Chrome is always a bit more stable.

Document: P3\_GMT20211125.odt Tags: System issue

MODERATOR: Can you try the box plot?

Document: P3\_GMT20211125.odt Tags: Moderator comment

MODERATOR: Nice thought, you are my first safari user. It is good that you point that out because the previous participants used Firefox or Chrome.

Document: P3\_GMT20211125.odt Tags: Moderator comment

MODERATOR: That could be the reason why the knowledge graph link was not working either.

Document: P3\_GMT20211125.odt Tags: Moderator comment

MODERATOR: There is a critical problem with Safari then.

Document: P3\_GMT20211125.odt Tags: System issue

(The participant reads the task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

So every color is the temperature, quite cold because this is ... My only doubt is that you don't allow me to choose a time frame. I don't know when this event took place, the day and the month. Oh, yes I know it is up there somewhere. Oh I cannot find it.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear

This here is the lag, the relative time from an event (day -1).

This is giving me a picture of the 30 days before the event if I am not mistaken. Because there is just one event.

Document: P3\_GMT20211125.odt Tags: Visualization useful

What would happen if there were 4 events?

Would we have 4 graphs like this or they would be overlapping? The four part conditions of the 4 events ... Relative time from the event.

Document: P3\_GMT20211125.odt Tags: Plot design complex

I am not sure the difference between the heat map and the bottom plot because the colors are the same. It is a bit redundant.

Document: P3\_GMT20211125.odt Tags: Plot design complex

Okay, if I choose some other variable, okay. Maybe I am a bit like this but for temperature I would prefer red or warmer colors, which are more intuitive. For rain, maybe more blues.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

Rain is much more chaotic, that is true. It did not rain here but is raining a lot there.

Document: P3\_GMT20211125.odt Tags: Visualization useful

The plot represents the interval of values that we say in the heat map. The median is 10.6, oh nice, the box plot descriptors are available when hovering over the box plot.

Document: P3\_GMT20211125.odt Tags: Visualization useful

Wow, nice. These are the different directions. This is temperature but here the color scale changes from the heat amp. This is more temperature-like.

I am not sure of what this is because this is a map north, south, east, west with wind speed. I don't get this wind speed.

Document: P3\_GMT20211125.odt Tags: Plot design complex

The number of the circles are not clear, are these for distance?

Document: P3\_GMT20211125.odt Tags: Plot design complex

Is this temperature for different locations?

Document: P3\_GMT20211125.odt Tags: Plot design complex

Polar plot ... This is not super clear to me what I am looking at right now. If in the center is supposed to be the event or something like this.

It looks like the temperature was recorded in different places.

Do I have temperature for different places then?

Document: P3\_GMT20211125.odt Tags: Plot design complex

Okay let's try O3. the places are the same. These are maybe stations where you have data from.

Maybe yes, these are stations where you measure the variables.

Document: P3\_GMT20211125.odt Tags: Plot design complex

But yeah, maybe I would add some explanation of what this is exactly or what this is.

Document: P3\_GMT20211125.odt Tags: Additional explanation

Let's go for definite like this. Uff one event in a 100 days.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear

This is something that I do not get. It's like there were no definite events in Dublin, that is what the availability is telling me.

Document: P3\_GMT20211125.odt Tags: Event concept and approach not clear, Query process clear

It takes more time because I have chosen many things here.

Okay, but wait. This is the same query ... I would like to do another one.

MODERATOR: There was a memory exceed problem.

Ah, my 300 was a bit too much.

Document: P3\_GMT20211125.odt Tags: System issue

MODERATOR: There was a memory exceed problem.

Document: P3\_GMT20211125.odt Tags: Moderator comment

Okay, now I have Q1 and Q3, and Q2 is dead, right?

Okay, amazing.

Document: P3\_GMT20211125.odt Tags: Query inputs and elements useful

This one is using more data sets by far. I guess that I put in more days but here we have two events (Q3) of high probability in Dublin.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

Okay, here I have the event date in 2016, quite old.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

Q3 uses 18 data sets and Q1 was using ... Ah, no I think in Q1 there are 21 data sets, so I guess Q1 wins.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

I guess but I will check. So the license here was in the data lineage ... Oh no, in the full metadata exploration.

Document: P3\_GMT20211125.odt Tags: Metadata content not useful or confusing

This has the by-SA creative commons license, which I don't remember what it means. The other query has the same license.

Document: P3\_GMT20211125.odt Tags: Search feature simple

Wait, I am going to close this otherwise it is in the middle.

Document: P3\_GMT20211125.odt Tags: Navigation complex

Q3 has some red values ... Oh, there is a problem here with the O3.

There tend to be more red values than blue values. I think there are more extreme values in Q3.

But the variables are different, toluene was not here before, which I think is a chemical compound or a metal.

Document: P3\_GMT20211125.odt Tags: Data table features useful, Export (meta)data useful

MODERATOR: We are having again the same problem with the memory. So we will have to refresh and submit the same query that you submitted before. The last query to check the plots would be fine.

Document: P3\_GMT20211125.odt Tags: Moderator comment

MODERATOR: We are having again the same problem with the memory. So we will have to refresh and submit the same query that you submitted before. The last query to check the plots would be fine.

Ah, my queries are too big.

I would like to try one with more events. Do you have any advice to check that?

MODERATOR: Wexford has many events to check.

Okay nice. I hope this does not crash the machine.

Document: P3\_GMT20211125.odt Tags: System issue

MODERATOR: Wexford has many events to check.

Document: P3\_GMT20211125.odt Tags: Moderator comment

Okay, here I have the events with different times. The events are super sparse: 2013, 2020, 2016, 2017 ... This one only uses three data sets.

Document: P3\_GMT20211125.odt Tags: Metadata content clear and useful

There is just one graph. Now I understand. Each event is here as a row in the heat map. The event-258 was basically in winter I guess, and the other ones in summer I guess. There is quite a difference.

Document: P3\_GMT20211125.odt Tags: Visualization useful

What was the task asking me to do?

(The participant reads the sub task again)

Document: P3\_GMT20211125.odt Tags: Task sequence complex

We can compare within the events.

Document: P3\_GMT20211125.odt Tags: Visualization useful

I am not sure because it is a bit complicated in the sense that events are very scarce. If there were more events I would suggest a time series plot but these events are super super scarce. You cannot put the time series of the events there for temperature and see it there. I don't know.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

I think for these types of events it is okay, it is nice.

Document: P3\_GMT20211125.odt Tags: Visualization useful

If I want the FAIR metadata, it downloads directly.

Document: P3\_GMT20211125.odt Tags: Simple export

I have another file with a weird extension. It is a ttl file and I don't know how to open this but I will figure that out I guess.

Document: P3\_GMT20211125.odt Tags: Export name and format unknown

So I can export the data basically. This is my CSV file, I don't know if you can see it but I have downloaded it.

Document: P3\_GMT20211125.odt Tags: Simple export

This is easy., just go here and it downloads when clicking export.

Document: P3\_GMT20211125.odt Tags: Simple export

Yeah, I think that would be useful. I think that it does what sometimes is very dull and awful to do, which is to go to different data sources and try to match all the days you are interested in there with the different variables from weather and pollution. Here you just click and get the data and I trust the data is correctly loaded.

Document: P3\_GMT20211125.odt Tags: Simple export, Export (meta)data useful

The only thing I would maybe add is an automatic analysis whether these events make sense to say that are linked to some of the variables. I am going to send you a paper that is trying to see if zoonotic spillovers or emergencies of diseases were linked, which is something very punctual like your data here. There is just one super random event, which was linked to the period of el Niño or something like that and they were doing some analysis. It was not super interesting but it is something that maybe I don't know. Comparing for the whole period how was the temperature when these events were taking place or with some lag periods. I don't know. I am going to send it to you afterwards if I find it.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

So, yeah I think it is very intuitive. If you don't use Safari, it is better.

Document: P3\_GMT20211125.odt Tags: Overall positive experience

It works smoothly.

Document: P3\_GMT20211125.odt Tags: Overall positive experience

Yeah, I don't know. Maybe for diseases where there are more events maybe I would group them. I don't know if this is thought for diseases that have lots of events. Then, I would group them by day instead of saying the temperature for the last 20 days. There will be an overlap if there are lots of events.

I don't know. I guess this is more the data that I use. I do not normally use such scarce data but yeah.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

I think for this purpose this is great, it works fine and it is useful.

Document: P3\_GMT20211125.odt Tags: Overall positive experience

What I would do personally is to download the data and then I would try to do something myself.

Document: P3\_GMT20211125.odt Tags: Export (meta)data useful

Sometimes maybe what you like also because if you want to compare if some variables can trigger this event you would need information about this variable during the whole period. Like during the last 30 years because you want to compare the values of temperature the last 10 days before the event but you don't have here super available the rest of temperature data if you want to compare it with the general values of temperature to check for a period with an anomaly or something like that. Maybe that is the only thing I would add. I don't know how though.

Document: P3\_GMT20211125.odt Tags: Additional feature suggestion

I believe at this point the information is well explained and enough to actually go on with the tasks.

Document: P4\_GMT20211125.odt Tags: Home tab clear

I have them here, so the task is completed as I believe I signed in.

Document: P4\_GMT20211125.odt Tags: Log in easy

Is the task the whole thing or are the sub tasks a task that I should say completed? Okay, okay, alright.

Document: P4\_GMT20211125.odt Tags: Task sequence complex

I am interested in the events that we are certain that are clear, so the definite event. I assume that the possible are those with certain uncertainties. At this point I am only interested in definite ones.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

I am interested in let's see ... Limerick because it is a town I like and one of the ones I know in Ireland.

Document: P4\_GMT20211125.odt Tags: Query process complex

That being said I see that you can also add Dublin for instance, which is cool.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

Why would I add extra? Just as a shortcut to add all of them together at the same time because you are interested in all the events in Ireland. I don't want to have to click every single one of the counties by themselves

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

I am going to select a time window of one month so I will select 30 days with a lag of up to 7 days.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

Why? Because I believe these are events for ANCA vasculitis, where the prodrome is not really known but 30 days seems to be a decent amount of exposure time and a week is probably not enough but in just case I am just interested in that.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

First of all, let's check if there is some data. Oh, no events or envo data available.

Document: P4\_GMT20211125.odt Tags: Query process clear

Probably I need to select less days ... Why is there no data available? Let's see if I just say Dublin. Is this a problem with my settings of time window length?

Document: P4\_GMT20211125.odt Tags: Query process complex

I understand that the time window length, I would specify this better. Like with some information in a question mark.

Document: P4\_GMT20211125.odt Tags: Additional explanation, Query process complex

I assume that basically I am going to get an aggregated value of all the time window length with the lag in respect of the event for the days that I selected.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful, Query process clear

I mean I know this because I am kind of used to this type of research and I think we had discussed this. Even though I am not sure. I think this should be specified more clearly. I still don't know what it really means.

Document: P4\_GMT20211125.odt Tags: Additional explanation, Query process complex

Okay, let's see the high probability. So we don't have data sets for Dublin that have definite events I am assuming. That is a pity but I will go on anyway with high probability events. We are no 100% sure but we are highly sure that these are the events we are interested in studying.

Document: P4\_GMT20211125.odt Tags: Query process clear

There is some data. I want to use daily data and the sum, not the mean.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

That being said I would like this option to be variable specific so we could input a dictionary. For example, if it was precipitation I would like the sum but if it is for temperature I would like the mean as the sum of the temperatures doesn't make a lot of sense. Since this is a variable specific factor I would like to have them like this in default but having the possibility to go into more advanced settings. If you want to open more advanced settings to specify for each of the variables.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

At this point I don't know the variables I am getting, which I think would also make sense before I decide on the aggregation method. How can I know how to aggregate the data if I don't really know the nature of the data at this point. So that is something I would change.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

I don't think I am going to remember all these tasks, so I am going to go step by step

Document: P4\_GMT20211125.odt Tags: Task sequence complex

We have here two events that have come from all these different data sets.

I mean I can count the data sets visually but I am a bit lazy to do that. I am not really going to count them but you could count them.

Document: P4\_GMT20211125.odt Tags: Metadata content clear and useful

Alright, then here I assume we have the daily values for each of the events. We have the max values in red I assume ... No, wait a second.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful, Task sequence complex

(The participant reads again the task)

Document: P4\_GMT20211125.odt Tags: Task sequence complex

There are not three buttons, there are five buttons. Well, except the export I think.

Document: P4\_GMT20211125.odt Tags: Task sequence complex, Navigation complex

Okay, yeah. This is the document showing exactly what was the query, which is nice in case you wanted to export this somehow or that you want to prove exactly how your data was aggregated, how it was formed and where it came from. I guess here you have everything which is quite valuable itself.

Document: P4\_GMT20211125.odt Tags: Metadata content clear and useful

It adds even more of that. It is not supposed to be readable for humans I assume but machine readable. Maybe it is a bit readable by humans depending on the kind of human.

Document: P4\_GMT20211125.odt Tags: Metadata content not useful or confusing

Let's say it would be nice to add a copy of one of those with a button that you copy the whole paragraph because I seriously doubt many people will read this. But they might want to copy and store it somehow or copy and paste it somewhere else.

I guess you can also download it like clicking the FAIR metadata export and this would do it. Exactly. Then you can download the data which can download it as an RDF.

Document: P4\_GMT20211125.odt Tags: Simple export, Export (meta)data useful

I don't know how many data sets there are so ...

Document: P4\_GMT20211125.odt Tags: Metadata content not useful or confusing

In the full metadata exploration text, I can try to do that.

(The participant reads the data use comment)

Seems I am a researcher in one of these projects, so I can use it.

Document: P4\_GMT20211125.odt Tags: Search feature simple

Identification risk, this is a long comment.

(The participant reads the identification risk comment)

Understood, so we understand that there are some risks here so we need to take them into account.

Document: P4\_GMT20211125.odt Tags: Search feature simple

Let's look at what is the license of the data set we are checking.

(The participant reads the data sets description comment)

This is a very verbose description but I guess that is good and you can know exactly what kind of aggregations were done. Basically, by getting this comment, that is a very nice description.

Document: P4\_GMT20211125.odt Tags: Search feature simple, Metadata content clear and useful

Alright, these are the data columns I assume, oh the data table I mean.

(The participant reads the column heading descriptions)

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

This one does not have a tooltip for some reason. I think maybe it is because of the width of the column. I could probably indirectly find that out but we don't want to do that I guess. It is NOx as NO2.

Document: P4\_GMT20211125.odt Tags: Additional explanation

Wet bulb air temperature, I don't know what exactly it is but I imagine that if you search it this is the real name from the ERA5.

Document: P4\_GMT20211125.odt Tags: Query inputs and elements useful

I think they are quite descriptive and it is pretty nice to have a short name and a proper description once you hover over. So that is cool.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

(The participant reads the data background paragraph)

Okay, I did not understand this before but after reading the data background paragraph, a Z-score is computed using all the values in the same column. Then, we get red values as high, basically over 95% over 2 in Z-score; and those that are blue, although I don't see many blue ones, are basically those that have a Z-score less than -2. This is my assumption based on the description I see here.

So I would say that: yes, I understand unless I don't. But I think I understand.

Document: P4\_GMT20211125.odt Tags: Data table features useful

Okay, I did not understand this before but after reading the data background paragraph, a Z-score is computed using all the values in the same column.

Document: P4\_GMT20211125.odt Tags: Wording not clear

If I hide this it gets hidden and with the toggle button ... Okay, it is fine, I like this ... Yes, this is completely usable and that is nice.

Document: P4\_GMT20211125.odt Tags: Data table features useful

Well, yes it is but it is not of course a trivial understanding. Of course you would need further analysis of this data to get a proper answer to whatever questions you are studying. But it is indeed a useful table to start to understand how the environmental variables might be linked to your events

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

Mmmm ... To the table itself ... You can export as CSV I assume, yes.

Document: P4\_GMT20211125.odt Tags: Simple export, Export (meta)data useful

No, I mean I could imagine some nice things like adding a gradient if you want instead of only the extremes but it would be a wasted effort.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

So far I think this is very nice.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

I am more interested in checking ... For instance I don't know what it is. It is probably a missing value, right? It is taking it as a zero I assume. The same as this. That is weird. Maybe since missing values could exist, maybe a proper category for these values. Same here, missing values should not be included in the distributions and get Z-scores. Either add a different background for missing values, a gray one or something, or at least do not compute them as anything. I don't know if they are taken as zeros but they shouldn't. It should be clear that whenever any aggregation is done later from this, they should not be taken into account.

Document: P4\_GMT20211125.odt Tags: Navigation complex

Let's make a heat map

I will try PM2.5, okay I see here. It can be a bit buggy.

Document: P4\_GMT20211125.odt Tags: System issue

So a heat map should be two variables I assume ... I am not sure.

Well, wait a second. Okay, let me see what happens.

Basically, the heat map allows us to compare per event. The X-axis shows time (days), relative time from the event, and the Y-axis the different events. So this allows us to compare and observe common patterns, time common patterns let's say.

Document: P4\_GMT20211125.odt Tags: Visualization useful

That being said, I don't know why it works for visibility but it does not work for others.

Okay, it works for temperature, SO2 but for my original choice, which was PM2.5 ... Okay, I think it is Heroku being a bit buggy together with Dash sometimes. I think it works.

Document: P4\_GMT20211125.odt Tags: System issue

We can do a more clear comparison of the distributions. We don't care about time anymore but we see a comparison between the events. We are able to see if the particulate matter in one event is lower or higher in general among the days that we selected compared to other events.

Document: P4\_GMT20211125.odt Tags: Visualization useful

This is a very cool plot.

Document: P4\_GMT20211125.odt Tags: Visualization useful

Okay, the polar plot. This one is a bit more complicated in general but I am making assumptions here. I think the polar plot should have some kind of description.

Document: P4\_GMT20211125.odt Tags: Additional explanation, Plot design complex

I already know what a polar plot is because I have discussed it with SERDIF creator and maintainer but I don't think most medical researchers will have any kind of clue what a polar plot is, although it looks pretty. This probably needs to be communicated.

Document: P4\_GMT20211125.odt Tags: Visualization useful, Additional explanation, Plot design complex

Other than that, this is showing the coordinates and I assume these are different values but I don't know how to connect this data from the table. I guess each circle is a day, and then depending on where the wind was coming that day and the seep it had ... No? So it is the mean of the values which are in certain coordinates compared to the event and then you can basically see ... I mean that this is what I would do in real life.

I know what a polar plot is but ... Anyway, not going to waste more time

Document: P4\_GMT20211125.odt Tags: Plot design complex

I think the polar plot is nice but it needs further explanation.

Document: P4\_GMT20211125.odt Tags: Visualization useful, Additional explanation

So I would say aloud that the polar plot needs a proper explanation.

Document: P4\_GMT20211125.odt Tags: Visualization useful, Additional explanation

I would also say that instead of the meaning ... Okay, I think here it says mean because I selected the mean ... Okay no, but we could do that though.

Document: P4\_GMT20211125.odt Tags: Plot design complex

I think scatter plots might be interesting, which you can color by the event or even adding extra categories.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

In this case, I have only used high probability events but I may want to compare high probability events with other categories like those that are not. Then, we can show a scatter plot for those ones that are not. So we can show the plot not by event or the box plot not by event but by category. I think that is the most useful visualization you can have. It is interesting to see the distribution of the events of the same condition or category like those that were certain. But what you are interested is comparing controls against treatment, or healthy against disease events. So usually that is what you would want to do. That is what I would miss.

That being said this can be done afterwards. So if I had to ask for something I would ask for that.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

How are we with time? Do you think I am being too slow? Maybe a bit.

Document: P4\_GMT20211125.odt Tags: Task sequence complex

So what if I select no events in Dublin, right?

Now we are going to compare the high probability with the no events. Unless there is no data available. Ah there is data available and the number of data sets is this 20.

Document: P4\_GMT20211125.odt Tags: Query process clear

So there is one event, which is not really that much to extract conclusions, but we have one event of someone that did not have any probability, well not only any probability but null probability of disease I assume. We have 20 environmental data sets.

Document: P4\_GMT20211125.odt Tags: Query process clear

We are going to use the same aggregation method as I want them to be comparable so I am only going to change the event category.

Document: P4\_GMT20211125.odt Tags: Query process clear

(The participant starts getting so pop outs to download the data sets)

MODERATOR: Apologies, Heroku glitch.

Document: P4\_GMT20211125.odt Tags: System issue

MODERATOR: Apologies, Heroku glitch.

Document: P4\_GMT20211125.odt Tags: Moderator comment

(The participant reads the sub task again)

Document: P4\_GMT20211125.odt Tags: Task sequence complex

Well, so I would have to count them now.

We have 20 here (Q2) and the other one also has 20 (Q1). I think because it is using the same location so I could probably use a different location to get different data sets. But yeah, they are exactly the same data sets.

Document: P4\_GMT20211125.odt Tags: Metadata content clear and useful

Good question, let's go to the full metadata exploration text. I think so as this is a creative commons license for Q1 and for Q2 we have the same license, the creative commons. So yes we have the same one.

Document: P4\_GMT20211125.odt Tags: Search feature simple

Well this is Q2, and doing a very quick exploration I would say ... If we don't count the missing values I would say that Q2 has more, but more or less, I should quantify it better.

Document: P4\_GMT20211125.odt Tags: Data table features useful, Export (meta)data useful

I think the most important one here would be the box plot as we will be able to compare the 3 events actually.

I will go to the box plot and compare the PM2.5 between the two events in Q1. Then, if we go to Q2 and do the box plot again, of course we only have one event.

We see that the median for event-no has a median of 9.475 and the median for the events that were with high probability were 10.13 and 10, which is a clear indication that PM2.5 is causing the high probability events so we should do something about it (joking).

Document: P4\_GMT20211125.odt Tags: Visualization useful

Yes, I would somehow allow comparison between different Qs. Q1 against Q2 with the same access and allow a side by side comparison. Now I have to remember this but the axis are not in the same range because this goes up to 50 and Q2 to 25. So visually ... I know how to read a box plot but to allow for a proper comparison the different queries should be merged into one data table so they can be combined.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

Okay, I will do it but I think I already did that.

Document: P4\_GMT20211125.odt Tags: Simple export

I have saved the FAIR data and metadata. I assume that this is the same but the FAIR ... Oh no, no, this is the FAIR metadata and this is the FAIR data.

Document: P4\_GMT20211125.odt Tags: Simple export, Export name and format unknown

So now I go to data provenance, do you want me to copy and paste this? Oh no, it was here the export button. Saved.

Document: P4\_GMT20211125.odt Tags: Simple export, Navigation complex

I know how to download the data table as you can see here in the previous downloads.

Document: P4\_GMT20211125.odt Tags: Simple export

But data provenance is this? Okay, now it is saved as a CSV.

Document: P4\_GMT20211125.odt Tags: Simple export

I will save and open it to check what's inside.

Document: P4\_GMT20211125.odt Tags: Simple export

Oh, it is a zip file with all of our queries inside, which is really nice, so we can make as many queries as we want. Then, merge the tables from the queries to do our analysis, which actually is really nice to do in our software of choice.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

I hadn't even read the full question and I can already answer because yes, it would be perfect for further exploration for an environmental researcher such as me to then load it in R or Python and then do my own analysis. This would be definitely useful.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

Well, in general the tasks were easy to do even though I made some comments in steps that could be more clarified. But in general, I think the job is very well done.

Document: P4\_GMT20211125.odt Tags: Overall positive experience

I do think that visualizations, since I am very peaky as I know how to do this analysis by myself

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

I am more interested in being able to fetch the data tables as we did in the previous task and then load them into my software of choice and do it myself. Running the query and getting the data that you specifically want, it works really well.

Document: P4\_GMT20211125.odt Tags: Export (meta)data useful

I would just add some more advanced settings for some complex scenarios where we want the average for some variables and the maximum for others, or we want a multiple of them. This should be doable.

Document: P4\_GMT20211125.odt Tags: Additional feature suggestion

But other than that I think it was a very smooth and easy experience to complete the tasks.

Document: P4\_GMT20211125.odt Tags: Overall positive experience

(The participant reads the text in the home tab) Events with locations and time and environmental data, then it goes through SERDIF and then it comes out and you get the data you need.

Document: P5\_GMT20211125.odt Tags: Home tab clear

Okay so that seems fair to me.

Document: P5\_GMT20211125.odt Tags: Home tab clear

So yeah, I think I got an overall understanding of the SERDIF framework.

Document: P5\_GMT20211125.odt Tags: Home tab clear

I don't see anything, nothing stands out as obviously missing here.

Document: P5\_GMT20211125.odt Tags: Home tab clear

I guess I had some sort of an idea of what was going on. The fact that it was quite high level was a nice thing.

Document: P5\_GMT20211125.odt Tags: Home tab clear

I don't necessarily need to see more information.

I can see that there are links. So you have links for further information, which I have completely ignored at first, so let me just ... Okay, the links work.

Document: P5\_GMT20211125.odt Tags: Additional source exploration

I guess this is the link to a Python dashboard or something, and that you can run something and some software. This is just something to do with the graph database, okay. But I am not sure I want to pursue this further anyway.

Document: P5\_GMT20211125.odt Tags: Additional source exploration

I am happy I know what is going on, so I think yeah. I am very happy with that.

Document: P5\_GMT20211125.odt Tags: Home tab clear

Let's try to log in in task two. I have them here so I will copy and paste that.

That seems to go fine.

Document: P5\_GMT20211125.odt Tags: Log in easy

Yeah, I will read that and go one at a time.

Document: P5\_GMT20211125.odt Tags: Task sequence complex

(The participant explores the event type pop out) So yeah, this is something I remember to talk about with other researchers, that definition is clear to me.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I think the events are pretty widespread.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I am surprised by the order of the four events, which is fine. I can see what's going on, so it is not a problem. But I would have thought that you would have gone for something like no, possible, high probability and definite.

Document: P5\_GMT20211125.odt Tags: Navigation complex

But anyway. The information there is pretty clear.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

We can look at different counties ... (The participant reads some of the Irish counties) Again, it seems like it is in a random order.

Document: P5\_GMT20211125.odt Tags: Navigation complex

I guess I can put it in the search box here just because I am being noisy and it is a county I know well.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I am just picking them arbitrarily but I guess I could pick a couple so actually it could be more interesting. Why don't I pick some that are close to each other so I can see?

I will do Wicklow, Dublin, Kildare and Meath or something like that because they are all sort of commuting to Dublin.

Okay, I guess it is pretty arbitrary for me.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I would say ... Time window length and lag ... I feel like you want a kind of broadened window. I feel like a month will be an interesting window to look at.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

Then, we got time window lag. So for this illness, I guess we are talking about vasculitis here, but I mean this is pretty broad here but I am thinking with vasculitis in my mind while we are dealing with this. I am trying to think about what the window is, I guess but I am far from an expert but I feel like the window is almost like a three month interval. Maybe we will just put 60 days as an average there.

So I guess these are the types of time frames I am interested in.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

So we have 14 event and 22 environmental data sets, okay so now I can move to the next steps.

Document: P5\_GMT20211125.odt Tags: Query process clear

So temporal units: hour, day, month, year ... Aggregation method: mean, sum, min and max ... I guess I am just playing around with this but can I click multiple of these or is it an either/or.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

So let me think now, I think a year is probably too broad for the temporal unit. I don't think we have to look at something annually. I don't know whether a day or month would be better. At the top of my head I don't know what would be better, probably a bit embarrassing to say so but I will pick days, so it is a bit more granular.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

Mean, sum, min, max I can remember now I think I guess I would pick the ... I guess it doesn't matter too much. I think I remember some researcher talking about the max being sometimes more important but we will see.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I mean these things were pretty clear.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I guess yeah I mean I kind of understood what was going on here.

Document: P5\_GMT20211125.odt Tags: Query process clear

Yeah I think that was all fine.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I think I understood everything.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful, Query process clear

The only thing that jumped on to me while going through this was to maybe you might want to look at a few of them at the same time for the temporal unit. I don't know if this answers your question or if this is a great idea. The same for the aggregation method. I mean I guess you can repeat the query a couple of times so it is not the end of the world either way.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

I am just trying to remember the applications that we worked on, usually there was a bit of uncertainty around this. You might want to experiment a bit to try to find out what works best. So from that point, you know, you might have to have a look at them at the same time to see. That is the thing that jumps out to me if you know that's it.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

There is the export button ... I want to make sure I am doing the right task ... Okay yeah I will pick a button and go through.

Document: P5\_GMT20211125.odt Tags: Task sequence complex

I guess we got this data set ... (The participant reads the column headings of the data provenance table) So I guess , yeah the event type was high probability that was specified and the case study would probably be the same for all of them as it was the only one that was there.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

Looking at this it is not very clear to me. It seems that the eg: is quite meaningful but I don't know what it means. Normally, that stands for example given as an expression or there is a Latin expression that I don't know. Is it just because maybe I don't know the data that we are using? It is just synthetic data or something like that and it is just example data to work with. That is how it is all recorded, I don't know. Maybe it is an acronym that is important but if it is I don't know what it means. But maybe that is not important.

Document: P5\_GMT20211125.odt Tags: Wording not clear

A record of the different counties and then the different event dates. We have like a timestamp 2014 ten past three in the morning.

We have a set of data sets that relate to Dublin for this event type and event. I am seeing multiple data sets used here. Whereas for the other entries it seems like less data sets are used ... Oh no that is not true there are more data sets here.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

Yeah, I will keep going but I am not entirely sure what I am looking at here. sure you can clarify pretty quicker but at the same time I should be able to look at this and understand it myself directly. Then, I wouldn't be hugely confident that I could say exactly what is going on. So I guess the thing that is a little bit confusing is that there are multiple data sets here. I guess I am sort of expecting, what I am expecting to see as default from my own experience is one data set used entry per row entry in this data table but we got multiple data entries. I guess this is just redundant and it is the same information over time. But typically that is what you would see in a standard spreadsheet or what I am used to seeing anyways.

Document: P5\_GMT20211125.odt Tags: Query process complex, Metadata content not useful or confusing

So anyway, just again. I am not being critical. I am sort of saying that you know this is you know. Okay, that makes me kind of wonder if this relates to a specific event, which meets certain criteria that was specified, why are there multiple data sets and how do they relate to each other. That would not be obvious to me.

I can see that it is blank here as well for this entry and I am not entirely sure for some entries then it is not included so it is hard to. It is not entirely clear to be what it is going on but I guess I understand the basic format here but it is a bit elusive to me.

Document: P5\_GMT20211125.odt Tags: Wording not clear, Query process complex, Metadata content not useful or confusing

Okay, so I can see underneath ... This is a data table with some kind of color coding here so the values that are more than 2 standard deviations from the mean are read and less that 2 is a blue color.

Document: P5\_GMT20211125.odt Tags: Data table features useful, Task sequence complex

I do not see any colors on this but we have references to an event. I can see how we can relate the data provenance table to the data table. This is obviously an event in Meath. There are other ones as well like event-526 on different days. So I guess there was some kind of lag that I was asking for. Maybe this is related to that.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful, Metadata content clear and useful

This is a week's worth of data. Oh yeah, this is great So the variable names are explained as you go over them. I guess you can hide them if you want but I am not going to click on that.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful, Task sequence complex

MODERATOR: If you don't mind going back to completing the previous task as this is the following one?

Document: P5\_GMT20211125.odt Tags: Moderator comment

Let me just double check. Data lineage ... Is this the same information as before?

So data lineage, I am not sure what I am looking at now. Sorry I got confused or something.

Data provenance and if I click data lineage the bottom box just closed.

So I think this is the data provenance tab and I have discussed what I think is going on here.

Document: P5\_GMT20211125.odt Tags: Navigation complex

(The participants reads the text in the data lineage window) I can see here it says a query to construct the data set, so I think it is SPARQL query code that is explaining what is going on.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

So I don't know anything about SPARQL but it looks like I can kind of see how these query languages work with the bindings, select and these. So I can kind of speculate about what is going on here. But I don't think this is something that I would give particularly scrutiny as a non expert. Especially, if you have some concern that the query that you put in on the table does not correspond to what it's coming out in the output. You are familiar with the language then you might go through and check what is going on. For me I can't say much more than that.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

Sorry, just scrolling down. I wonder if I should close the others, yeah okay, that is nice.

Document: P5\_GMT20211125.odt Tags: Navigation complex

So I mean what is going on here is that you are sort of a big file that is explaining where you are pulling all the information from I guess. The things you are talking about here like the themes with URL references. These are kinds of definitions for the different data sets. I guess you got these kinds of prefixes here that link with those references so I guess that is where it comes from.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

So again, I guess the data comes out of a magic box so for me I would be happy that this information is all correct. If there was any concern I guess you got to a certain point to search for it and understand what is going on. Something like these geometries that are the coordinates of the different regions to figure out where the events happen. I guess you can use that to try to figure out if everything makes sense. But I mean at first glance, it wouldn't be the first thing that you would be looking at but its is good to know that it is there as a background resource.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

So I think I am looking in the data provenance table so I am just looking through this. I can see that we got data sets used here and that you can look through that to see what is going on.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

My question is then this, possibly I had said that before in another task. The thing that I am wondering is if the data sets are unique or you would use the same data set twice or something like that.

I can see some data sets have an EEA and some that have MET so I guess it is querying Met Éireann or something like that. So presumably those are two different data sets anyway. If I go over the different rows, you wouldn't have the same data set twice, which I don't know if this makes sense.

I guess I am not sure since you go from entry to entry. Well I can see in fact that these two are the same.

Document: P5\_GMT20211125.odt Tags: Metadata content not useful or confusing

(The participant checks other data sets used to see if they are repeated through the events) Let me so ... 20 data sets for this event, which I guess that they are different. I am not going to do that again but I have lost count but I was around 21.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

MODERATOR: No problem, you can go ahead.

Document: P5\_GMT20211125.odt Tags: Moderator comment

I can say I don't know if I answered the question or if I am looking at this in the right way.

I don't know what you can take from that but the answer could be no. Maybe that is because of the query that I set up.

Document: P5\_GMT20211125.odt Tags: Query process complex

I think this is the data set use comment but I don't see much here. Let me see if there is anything else. There are only a couple of matches.

(The participant reads the data use comment) Let me now try the identification risk. Probably I should have copied and pasted that but anyway. (The participant reads the identification risk comment) I think that is fine, I am happy with that. So pretty clear about that.

Document: P5\_GMT20211125.odt Tags: Search feature simple

So let me look through this now and I will close that one.

Document: P5\_GMT20211125.odt Tags: Navigation complex

(The participant reads the sub task again)

Document: P5\_GMT20211125.odt Tags: Task sequence complex

(The participant reads the column heading descriptions)

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

I don't know if you would like me to go over all of these.

Document: P5\_GMT20211125.odt Tags: Task sequence complex

Oh this one does not have a description

Document: P5\_GMT20211125.odt Tags: Additional explanation

but for the most part these are very clear so I am happy with that.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

So this is explained, the principle for sure, but I don't see any values here but I wonder if I change the entries here. So I can see some red values here, the red are the high values. Okay I can see some entries here but these were missing before but it seems like a low value.

I guess compared to the other ones 17.5 seems high. So yeah I am very happy that I understand that, that is very clear.

Document: P5\_GMT20211125.odt Tags: Data table features useful

Yeah that is fine although it is not clear to me how I can get them back. That was one of the questions, oh by using the toggle columns button. Okay, okay that is fine and really easy to use.

Document: P5\_GMT20211125.odt Tags: Data table features useful

Yeah that is fine although it is not clear to me how I can get them back.

Document: P5\_GMT20211125.odt Tags: Navigation complex

I guess the big thing is that it is being divided into smaller regions. I would just make the table bigger, stretching in an out,that would be easier to look at more data at once.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

There are going to be issues with missing data depending on the type of data that we have. So here we would only have a small amount. I don't know how easy this can be but toggling in and out of the missing data. Maybe you could pick the columns and pick the missing ones and get a much more complete picture.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

I think the other features that you have there are very helpful to give you a scan of what is going on. Especially before you export the data you want to make sure that everything looks okay. So that would be the only thing I would say.

Document: P5\_GMT20211125.odt Tags: Data table features useful, Export (meta)data useful

Let me look at the temperature. I am probably not clever enough to know what is going on with the other environmental data anyway.

We have 6 events and I guess I asked for them daily. I can make sense of this data. This is pretty clear. Event-147 was very warm as we were seeing blue the whole time and that other one the opposite being very cold.

Document: P5\_GMT20211125.odt Tags: Visualization useful

This box below I don't think what is going on to be honest.

Document: P5\_GMT20211125.odt Tags: Plot design complex

Relative time from the event ... I don't know what that means.

Document: P5\_GMT20211125.odt Tags: Plot design complex

I know that sometimes it is useful to have a couple of complementary maps but I would need some content to explain what is going on here. I can't figure it out based on that information.

At a glance it looks like it has the same or similar heat map twice so I am not sure.

Yeah, I am not sure what is going on here.

Document: P5\_GMT20211125.odt Tags: Additional explanation, Plot design complex

I know what is going on here as this is very clear.

Document: P5\_GMT20211125.odt Tags: Visualization useful

Oh yeah, these are very pretty. It would be nice if the events would be ordered in the same way between plots. It would be interesting to compare, is it possible to do that? I don't think it has happened here.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

what I would take again from it is that I asked for a 30 day window length not lag, isn't it? What you can kind of see is that the data is skewed. Event-265 had moderate temperatures but there was a lot of rain around that event. So you can see the mean overall is higher but for this it was kind of skewed but you know the rainfall didn't seem to change except for event-256. Is the fact that there is no data here or the fact that there is no data. It looks like it is just missing. That wasn't a particularly warm day. so I don't think that was an issue or something like that.

Document: P5\_GMT20211125.odt Tags: Visualization useful

For event-147 the rain was actually, oh no I am looking at the wrong day. For event-526, it is just missing that information but the temperature is missing too. So it is not appearing for the heat map. That is a bit surprising for me that, well that is maybe an internal setting but that is something you can fix fairly easily. Maybe it just needs to be a check for that.

Document: P5\_GMT20211125.odt Tags: Plot design complex

Certainly the plots are pretty easy to read and if you hover over you could see what the entries are along the way.

Document: P5\_GMT20211125.odt Tags: Visualization useful

So I think that looks good to me.

Document: P5\_GMT20211125.odt Tags: Visualization useful

I think I am being picky because I have taken too long to talk about this stuff.

Document: P5\_GMT20211125.odt Tags: Task sequence complex

So it says here wind direction and speed. Is that what is going on here? It is telling me what the temperature was like in relation to the wind direction and speed that is associated with this. I was not entirely sure what I was going to see here. I guess that I would need this to be explained to me.

Document: P5\_GMT20211125.odt Tags: Additional explanation, Plot design complex

That it was being used in relation to another variable. I am not sure what I was expecting to be honest with you for this. But it seems like you got cold events associated with strong northerly winds and warmer events associated with south-westerly winds. I wouldn't be 100% confident that I interpreted that right but it is what it seemed to me.

Document: P5\_GMT20211125.odt Tags: Additional explanation, Plot design complex

Okay, right. I will keep going because I think I am taking too long.

Document: P5\_GMT20211125.odt Tags: Task sequence complex

There are a lot of things that you can do. You would have to think about a histogram to look at the distribution even though the box plots kind of cover that.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

Maybe a scatter plot so you can start to look at more than one variable at the same time.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

There are so many to choose from so it is worth discussing it. I don't have an obvious one from the top of my head I guess. Maybe if you wanted to look at a couple of different relationships that would be helpful. A time series plot for all of the events at the same time that might be useful to see. Those are the things from the top of my head.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

(The participant starts reading task 8)

Document: P5\_GMT20211125.odt Tags: Task sequence complex

MODERATOR: You just completed task 6 so you should proceed with task 7.

Document: P5\_GMT20211125.odt Tags: Moderator comment, Task sequence complex

the locations from another part of the country. I am going to be quite broad again, but maybe that is a bad idea.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I am quite reluctant to change the time window length by very much but maybe I can reduce the lag. Maybe I will make the time window length a bit smaller. I don't want to play around too much with that.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I think that is running.

MODERATOR: It is actually not running, you need to check if data is available first.

Oh sorry, I forgot to press that button. Okay, perfect.

(The participant reads the task again) MODERATOR: I got a critical error from the memory. So that is why it is not working.

What do you think I should do?

MODERATOR: Maybe just choose one location and choose for example 5 days for the time window length so we can get through. If not we will have to refresh the page.

Okay.

MODERATOR: So you got Q3, so we will have to do with that. You won't have Q2 anymore.

So query 3 okay.

Document: P5\_GMT20211125.odt Tags: System issue

MODERATOR: It is actually not running, you need to check if data is available first.

Document: P5\_GMT20211125.odt Tags: Moderator comment

MODERATOR: I got a critical error from the memory. So that is why it is not working.

Document: P5\_GMT20211125.odt Tags: Moderator comment

MODERATOR: Maybe just choose one location and choose for example 5 days for the time window length so we can get through. If not we will have to refresh the page.

Document: P5\_GMT20211125.odt Tags: Moderator comment

MODERATOR: So you got Q3, so we will have to do with that. You won't have Q2 anymore.

Document: P5\_GMT20211125.odt Tags: Moderator comment

So this is pretty consistent in how many data sets it uses. That event for Q1 uses more data sets and that gets to the point of the question.

Document: P5\_GMT20211125.odt Tags: Metadata content clear and useful

Right, where do I find that out? It wasn't clear to me if it was in the data lineage or in the full metadata exploration. Okay, dct:license, which I forgot to check on the last one.

Document: P5\_GMT20211125.odt Tags: Metadata content not useful or confusing, Task sequence complex

So creative commons, yeah creative commons by-sa/4.0.

Document: P5\_GMT20211125.odt Tags: Search feature simple

So let me try that for query 1. Oh, I cannot find it for this one ... maybe if I try it again. Why can't I find it for one but I can for this other? That seems wrong. Oh maybe I clicked on the wrong thing, that was the issue.

Document: P5\_GMT20211125.odt Tags: Navigation complex

Okay, yeah no that is fine. It looks like it is the same one creative commons by-sa/4.0. Okay that is the same one. I certainly found it in both places. I was a bit lazy in terms of checking it.

Document: P5\_GMT20211125.odt Tags: Search feature simple

So this case record seems to be a bit more complete for the variables I am looking at here. Certainly I can see more blues and reds in general.

So it is hard to compare as this query has some issues regarding the completeness of the data. But I think I am seeing more colors for the second data set for query 3. I can see some but it does not seem to be as many.

That would suggest that for whatever reason, it is a bit tricky to say but I think query 3 has more.

Document: P5\_GMT20211125.odt Tags: Data table features useful

Okay so let me pick where I am actually confident in what is going one. But equally I am more tempted to do the polar plot to see if I can make more sense of it.

Oh, no.

MODERATOR: I think you won't be able to do this part as one of the queries broke which affected the index of the plots. So you cannot do this part because it broke before.

Oh sorry, that's a pity. I guess I will have to move on.

Document: P5\_GMT20211125.odt Tags: System issue

But equally I am more tempted to do the polar plot to see if I can make more sense of it.

Document: P5\_GMT20211125.odt Tags: Plot design complex

MODERATOR: I think you won't be able to do this part as one of the queries broke which affected the index of the plots. So you cannot do this part because it broke before.

Document: P5\_GMT20211125.odt Tags: Moderator comment

I guess some raw information would be useful like how many data points my query resulted in and how many missing values per column that would be kind of useful information to know.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

Then, if you could have some way to combine the queries but I would need to think about that one more. I think it could be interesting to say whether I have more data in Q1 than Q3 or you could combine line plots in the same part to see if the pictures are similar or not. Like you know if you had two queries and the pictures are separated you know if something is going on or if thy are broadly similar. So something like that I guess.

Document: P5\_GMT20211125.odt Tags: Additional feature suggestion

I guessed I clicked the export and left it in the downloads folder.

Document: P5\_GMT20211125.odt Tags: Simple export

Oh there is a FAIR metadata export button here and that is like a text file or something. I haven't seen a .ttl file before.

Document: P5\_GMT20211125.odt Tags: Simple export, Export name and format unknown

Yes, sorry, the data table and the data provenance ... Okay, I can export that as a CSV.

Document: P5\_GMT20211125.odt Tags: Simple export

So that is a big zip file. Do you need me to check it or is it fine like this? Okay, I think it is fine.

Document: P5\_GMT20211125.odt Tags: Simple export

For me personally, the way I would use the data as I don't have experience in using RDF myself, the thing I would be looking for is the data table.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

Then, the other information, frankly, I would like to talk to you guys about what is considered best practices but potentially it will be helpful to have it saved somewhere, in case I need it for my records or something like that. But is it something that I would directly check myself and say oh I need to check this, that is not something I would use. But that is not to say that is not important. I appreciate that it could be quite important. Especially in a collaborative project you know. In this kind of EU project we participate to keep track of what we are doing so I can appreciate the importance from that point of view. But I don't see how I would use it directly.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

I think the short answer is that it is a pretty positive experience.

Document: P5\_GMT20211125.odt Tags: Overall positive experience

Certainly logging in was very straightforward.

Document: P5\_GMT20211125.odt Tags: Log in easy

Establishing the queries was, you know, a little bit unstable right now but I don't know if it is because somebody is using it at the same time.

Document: P5\_GMT20211125.odt Tags: System issue

I think the information was quite clear.

Document: P5\_GMT20211125.odt Tags: Overall positive experience

I like the fact that you can use multiple tabs for multiple queries.

Document: P5\_GMT20211125.odt Tags: Query inputs and elements useful

I feel as if the emphasis of the dashboard is in getting the data you are interested in but also to provide information about metadata and RDF. Probably, this is kind of important because it is the kind of thing that I would for me, how you actually resource that data I don't really know about that. So the fact that there is some background information that lets you think about what is involved is quite helpful to check what is going on there. So that could be useful.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

I think the plots that are there are useful

Document: P5\_GMT20211125.odt Tags: Visualization useful

but I don't know how much further you want to go with them as I can download the data and do it myself.

Document: P5\_GMT20211125.odt Tags: Export (meta)data useful

But it is useful to have a look before you get all excited about all this data, and you say hang on, I need to resubmit the query. It is useful to have a little bit to see what is going on. I don't think you need to go in depth too much there.

Document: P5\_GMT20211125.odt Tags: Visualization useful

I am putting my username and password to sign in, okay.

Document: P6\_GMT20211126.odt Tags: Log in easy

So events with locations and time, environmental data, individual environmental records, okay.

Document: P6\_GMT20211126.odt Tags: Home tab clear

Not now, I need to read it again I suppose.

Document: P6\_GMT20211126.odt Tags: Wording not clear

Maybe an example, a pure example of how it works.

Document: P6\_GMT20211126.odt Tags: Additional feature suggestion

Okay, I did it

Document: P6\_GMT20211126.odt Tags: Log in easy

So as a case study I have to select an example of Ireland events. Eh ... possible, definite, no, high probability. High probability.

Location: Cork. Time window from day 1 to eh 15 days.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Click to check data availability. Envo data available, I don’t know what envo is.

Document: P6\_GMT20211126.odt Tags: Wording not clear

Well, I think I used 15 days because for a rare disease event we need at least 1 week of exposure, that is why I used 15 days.

Document: P6\_GMT20211126.odt Tags: Query process clear

Yes, well, The high probability I do not exactly understand what it means. If the event is high probable, I suppose that's the meaning.

Document: P6\_GMT20211126.odt Tags: Event concept and approach not clear

Okay I did it, Oh no I did not. I need to fill a temporal unit: hour, day, month or year. So I will click on the day because I chose 15 days.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Aggregation method: mean, sum, min or max. I will use the mean as I suppose is better and then I will click on submit.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Ehm so the introductory paragraph, we have a data table ...

(The participant reads the data table text below.)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

So I have the first column with I suppose is the 15 days, yeah.

Event, the event-540, and the high probability events anonymised. Then, the exposures of particles of 1um? Yeah I can read it here: particulate matter <10um.

Then, the participant reads the rest of the column names ...

Document: P6\_GMT20211126.odt Tags: Export (meta)data useful

So for this event, I have this ... I have to understand the table ... So for this event, this was the mean temperature of the 15 days ...

Yeah, now I understand. So for this event, there was this temperature, pressure, mean sea level pressure precipitation on the 15 preceding days to the event.

For this event, etcetera etcetera.

Document: P6\_GMT20211126.odt Tags: Export (meta)data useful

How many events were there? Can I see that? Ah no you have it there, many more events. Do I know how many events? Because I have 2 tables but I cannot count the number of events. So okay, I will go back to reading the tasks.

Document: P6\_GMT20211126.odt Tags: Task sequence complex

Is Q1 this one?

Document: P6\_GMT20211126.odt Tags: Navigation complex

MODERATOR: I think it is because of your zoom setting in the browser so if you scroll up or decrease the zoom you will be able to see the heading of the interface.

Document: P6\_GMT20211126.odt Tags: Moderator comment

Three buttons below? Which are the three buttons below? Is it here? Because I have 4.

Document: P6\_GMT20211126.odt Tags: Task sequence complex, Navigation complex

(The participant reads the tasks again)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

Data comes from events in Ireland. The events are here as examples in the table, event-118. So I can go back to the table and check if it corresponds ... but how can I get back? Ah it's here.

Document: P6\_GMT20211126.odt Tags: Metadata content clear and useful, Navigation complex

For example event-118, I want to see the data table ... 118 where is it? Oh it's here. So this one is a high probability event from Cork, the county of Cork. It was on this event date.

Document: P6\_GMT20211126.odt Tags: Metadata content clear and useful

The data sets used ... I do not understand this data set ...

Document: P6\_GMT20211126.odt Tags: Metadata content not useful or confusing

The other events you put an example of where they come from, that's what I understand

Document: P6\_GMT20211126.odt Tags: Metadata content clear and useful

So data lineage, what's the difference with data provenance? Sorry I ... Oh ah, it's different.

Document: P6\_GMT20211126.odt Tags: Navigation complex

So, yeah that I cannot understand ... Data lineage, I suppose, is the script you used to get the data.

Document: P6\_GMT20211126.odt Tags: Metadata content not useful or confusing

Eh ...

(The participant does not explore this text correctly)

Document: P6\_GMT20211126.odt Tags: Metadata content not useful or confusing

(the participant reads a few times the task)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

So to check how many data sets were used I need to count here: 1, 2, 3, 4, 5, 6, 7, 8, 9 ... so you used 9 data sets, is that right?

Document: P6\_GMT20211126.odt Tags: Metadata content clear and useful

(The participant reads the task again)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

I am searching for identification risk ... I cannot find it.

So I have to go to full metadata text and here I have to identify data use here.

Identification risk, no I do not know where I need to go to find the identification risk.

Document: P6\_GMT20211126.odt Tags: Search information process complex

MODERATOR: Do you want to read the introductory paragraph again?

Document: P6\_GMT20211126.odt Tags: Moderator comment

(The participant reads the introductory paragraph again).

So I have to search for the risk in the full metadata text. Oh okay.

Control + F, and ... I need to click on control + F

Document: P6\_GMT20211126.odt Tags: Search information process complex

MODERATOR: it may be command + F in your computer

Document: P6\_GMT20211126.odt Tags: Moderator comment

Oh yeah, now yes.

I have to search identification risk ... let me see

(the participant reads again the introductory paragraph)

So control + F identification ... Ah here, there's the information.

Document: P6\_GMT20211126.odt Tags: Search information process complex

(the participant identifies the elements from the task)

Yeah, you put the information about the identification risk I'm not sure if I understood ...

(The participant reads the task again)

Document: P6\_GMT20211126.odt Tags: Search information process complex

Yeah, I can write in license ...

(the participant identifies the elements from the task)

Document: P6\_GMT20211126.odt Tags: Search feature simple

That is what I did before ... Case study ... Oh no it was here

Document: P6\_GMT20211126.odt Tags: Navigation complex

Yes

Document: P6\_GMT20211126.odt Tags: Data table features useful

Yeah, I click here and it's okay

Document: P6\_GMT20211126.odt Tags: Data table features useful

So again, I have these events and I can know if these were exposed to particulate matter for 15 days before the event, that’s what I understand.

What was the mean temperature 15 days before the event. The different variables 15 days before the event, etcetera.

That is what I understand.

Document: P6\_GMT20211126.odt Tags: Export (meta)data useful

(The participant reads the task again)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

For now I do not have any ideas

Document: P6\_GMT20211126.odt Tags: Data table features useful

when I click on PM10 I might need to see elsewhere the results MODERATOR: maybe try another one

Oh Okay maybe there's no data for it

Oh yeah it works for rain.

Document: P6\_GMT20211126.odt Tags: System issue

MODERATOR: maybe try another one

Document: P6\_GMT20211126.odt Tags: Moderator comment

15 days before these events, the precipitation amount 15 days before ...

Relative time from the event -2 ... event ... here there are the number of the events Here relative time from the event -1 ... it's from the Z-score I suppose ... the results 15 days before quantified in a Z-score for all the events Relative time from the event

I don't understand that so, relative time -2 days ... the event would be in 0, you have -0.5.

Document: P6\_GMT20211126.odt Tags: Plot design complex

Again PM10, okay.

So for this event, the PM10 was here and the event was exposed to more than 15ug/m3.

Okay, it is the mean exposure of PM10 15 days before the event. For all the events I think now I understand the box plots for each event.

Document: P6\_GMT20211126.odt Tags: Visualization useful

I can choose PM10 ... Mmmm ... and where are the events in this polar plot?

Document: P6\_GMT20211126.odt Tags: System issue

MODERATOR: maybe try another variable

Document: P6\_GMT20211126.odt Tags: Moderator comment

Okay, this is the mean for ozone, but where are the events in this figure?

I am trying to understand this figure: east, west, north, south ... The mean ozone in this time window in the area it was chosen.

The mean of ozone 15 preceding the events but which events ... or just in the area chosen ... the whole ozone concentration in a geographic manner I suppose that's the result.

Document: P6\_GMT20211126.odt Tags: Plot design complex

Okay, so now I have to go back. I was in Ireland and now I will choose another country.

Oh, I cannot choose another country but I can use definite instead of high probability.

Then I can keep the one before ... location I will change from Cork to Dublin.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Well, I will keep them both.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Time window, I will choose more time. Maybe here in the time window if you want to be larger than days maybe you can add something more to be quicker than adding days one by one. Oh no you can do like this and write the input yourself.

Document: P6\_GMT20211126.odt Tags: Query inputs and elements useful

Okay so Q2 is because Q1 was the first thing I searched and then Q2 is the second query, yeah.

Document: P6\_GMT20211126.odt Tags: Query process clear

Okay, so to know how many data sets were used for Q1 I need to go to the data provenance. For example here, I have the number of data sets for this query: 1, 2, 3, 4, 5, 6, 7, 8, 9 ... and for this other query ... data provenance 1, 2, 3, 4, 5, 6, 7, 8, 9. Both have the same number of data sets ad that is because I did not change the query much.

I do not know if we have the same events though. I will search for event-102 ... I do not have the same events.

I think you have the same numbers of data sets for Q1 and Q2

Document: P6\_GMT20211126.odt Tags: Metadata content clear and useful

For Q1 the license, where was it? I have to go to Full metadata exploration. Then, I have to go and check for the license again ...

I cannot find the license ...

It does not work when.

Document: P6\_GMT20211126.odt Tags: Search information process complex

MODERATOR: try maybe closing the Control + F window on the top, opening it again and click on the arrows

Document: P6\_GMT20211126.odt Tags: Moderator comment

I try again but there is more information than when I clicked before.

Document: P6\_GMT20211126.odt Tags: Navigation complex

Okay, now it works.

I will close and open the window again.

Document: P6\_GMT20211126.odt Tags: Navigation complex

Here there's the license for Q1, creative commons.

For Q2, yeah it's here. I do not remember exactly the end of the creative commons link but I suppose they are the same

Document: P6\_GMT20211126.odt Tags: Search feature simple

I did not check for all the columns in the table. I cannot seem to go further, maybe it's the end of the table.

Document: P6\_GMT20211126.odt Tags: Navigation complex

There are a lot of colors in Q2 and in the Q1 again, yeah there are maybe less colors.

Document: P6\_GMT20211126.odt Tags: Data table features useful

Oh no I have to check page 1 and page 2 of the table.

Document: P6\_GMT20211126.odt Tags: Navigation complex

No I was wrong, I think both have similar number of colored cells.

But yeah, I suppose it is about the same because one has less numbers but I would have to check it correctly.

In a first view, I suppose they are the same.

Document: P6\_GMT20211126.odt Tags: Data table features useful

Better box plot ... in Q1 I have a box plot with 8 events that were exposed to less than 25 ug/m3 of PM10.

So that was for the first query, and for the second query ...

I have to select PM10 again, yeah. So I have 8 events and the PM10 exposure distribution is less skewed compared to Q1.

Oh no, because the mean is 18 and here 25 son the distribution is similar.

Document: P6\_GMT20211126.odt Tags: Visualization useful

So I go to the first query and I have to export all ...

Document: P6\_GMT20211126.odt Tags: Task sequence complex

(The participant reads the task again)

Document: P6\_GMT20211126.odt Tags: Task sequence complex

Okay, so here I can download the FAIR metadata export and here the FAIR data export

Document: P6\_GMT20211126.odt Tags: Simple export

Data provenance ... I want to export it as CSV ... so how can I do it?

Sorry, I am looking for where I can download this information.

Document: P6\_GMT20211126.odt Tags: Navigation complex

MODERATOR: maybe click again the button data provenance

Document: P6\_GMT20211126.odt Tags: Moderator comment

Oh yeah, export ... csv

Document: P6\_GMT20211126.odt Tags: Navigation complex

Okay

Document: P6\_GMT20211126.odt Tags: Simple export

Yeah, I think it is usable and reusable, because you can search again for the same information.

Document: P6\_GMT20211126.odt Tags: Export (meta)data useful

I think you have to be concentrated in all the information you provide while thinking aloud ... maybe if there is less information it would be more understandable but at the same time you need all this information at one point. So maybe you can click on a button to click the whole information and at the first stage maybe you do not need all the information at once. That is my idea at this point.

Document: P6\_GMT20211126.odt Tags: Additional feature suggestion

But the experience was okay.

Document: P6\_GMT20211126.odt Tags: Overall positive experience

Okay, this is very clear I think.

Document: P7\_GMT20211126.odt Tags: Home tab clear

Yes, I think it is pretty clear. You describe the things you can do with the framework and the different options that you have

Document: P7\_GMT20211126.odt Tags: Home tab clear

Okay, so here I have the credentials from the email.

I'm going to log in. Okay I am already logged in.

Document: P7\_GMT20211126.odt Tags: Log in easy

The events ... I'm going to choose possible events

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

I'm going to select Dublin because I think it can be the most interesting.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

Time window in days ... let's try 4 months maybe or 20 days, and the lag 1 week.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

Let's check, maybe if I change this because there are no events.

Document: P7\_GMT20211126.odt Tags: Query process clear

Oh okay now I have 2 events and 20 data sets.

Document: P7\_GMT20211126.odt Tags: Query process clear

Temporal units about days

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

the aggregation method the mean

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

MODERATOR: Could you please select less days? Because it is an example data base and it would take a long time Okay, 30 days is okay?

MODERATOR: Sure, no problem

Document: P7\_GMT20211126.odt Tags: System issue

MODERATOR: Could you please select less days?

Document: P7\_GMT20211126.odt Tags: Moderator comment

MODERATOR: Sure, no problem

Document: P7\_GMT20211126.odt Tags: Moderator comment

Okay, I think everything here is very clear.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

I would not add any extra information because the interface here to get the data is quite clear and the message is very clear.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful, Query process clear

Okay, let's go by parts.

Document: P7\_GMT20211126.odt Tags: Task sequence complex

Okay, the information provided here is quite clear.

Document: P7\_GMT20211126.odt Tags: Metadata content clear and useful

Okay, the case study is this, the events are high probability, the county is Dublin, and the last column I assume is the list of the 20 data sets used in the query.

Document: P7\_GMT20211126.odt Tags: Metadata content clear and useful

Here we just have more information

Document: P7\_GMT20211126.odt Tags: Metadata content not useful or confusing

(The participant does not explore this information)

Document: P7\_GMT20211126.odt Tags: Metadata content not useful or confusing

Yes, around 20 for each one.

Document: P7\_GMT20211126.odt Tags: Metadata content clear and useful

Okay, so I can use control+F to search the information in the metadata text.

Document: P7\_GMT20211126.odt Tags: Search feature simple

Data use comment - Clear information

Document: P7\_GMT20211126.odt Tags: Search feature simple

Identification risk comment - Let's see, I think here there's information regarding on how the data set was made in terms of confidential information and so on. I think it is also clear.

Document: P7\_GMT20211126.odt Tags: Search feature simple

License - The license is creative commons, okay. If you need more information about creative commons you can go directly to the web page of the license type.

Document: P7\_GMT20211126.odt Tags: Search feature simple

Here if you put the cursor on top of the column heading you can see a brief description of the particular column ...

(The participant reads the data table column heading names and descriptions.)

Document: P7\_GMT20211126.odt Tags: Export (meta)data useful

Yeah, I can understand the colors because here it says the meaning of each color. The high values are read and the low values are blue.

Document: P7\_GMT20211126.odt Tags: Data table features useful

Okay, let's see. Yes, it can be hidden, perfect.

Document: P7\_GMT20211126.odt Tags: Data table features useful

The information provided is useful for a particular study, you have a good temporal resolution, description and it is quite easy to see the values.

Document: P7\_GMT20211126.odt Tags: Export (meta)data useful

I would not add any kind of extra information in the description and the colors are clear.

Document: P7\_GMT20211126.odt Tags: Data table features useful, Export (meta)data useful

This feature of hiding the columns with the eye icon is also easy to use.

Document: P7\_GMT20211126.odt Tags: Data table features useful

I think it is quite clear as you can see all the information in the box plot (median, max, min ...).

It is very intuitive.

Document: P7\_GMT20211126.odt Tags: Visualization useful

I think the map is very clear as it provides information when you put your cursor over the plots.

We could try with another variable but I assume it's the same. Let's try rain, we can see the precipitation amount per day.

Document: P7\_GMT20211126.odt Tags: Visualization useful

The plots are fine

Document: P7\_GMT20211126.odt Tags: Visualization useful

They are useful and pretty but also provide the numerical information with just hovering the cursor over the graph.

Document: P7\_GMT20211126.odt Tags: Visualization useful

Okay so let's first input another query.

Let's try with the same one but definite in the same place.

Let's try this.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

Oh so okay, two events.

Document: P7\_GMT20211126.odt Tags: Query process clear

Now we have a Q2 and the previous work was saved as Q1.

Document: P7\_GMT20211126.odt Tags: Query process clear

Now we have a Q2 and the previous work was saved as Q1.

Document: P7\_GMT20211126.odt Tags: Query inputs and elements useful

(The participant reads the task again)

Document: P7\_GMT20211126.odt Tags: Task sequence complex

Okay, so the first one because it has 20 data sets this one is just 2 data sets.

This is clear.

Document: P7\_GMT20211126.odt Tags: Metadata content clear and useful

The last one was creative commons and let's see this one ... Yeah same license.

Document: P7\_GMT20211126.odt Tags: Search feature simple

I think this one, oh no the first one. Let's bring back the columns I hid before ... I do not know what I did but ... oh okay I see.

Document: P7\_GMT20211126.odt Tags: Navigation complex

I think the first data set has more extreme values because it has more colored backgrounds.

Document: P7\_GMT20211126.odt Tags: Data table features useful

Q1 box plot rain and Q2 box plot the same, rain.

Document: P7\_GMT20211126.odt Tags: Visualization useful

Let's see ... Which was the question I can't remember?

Document: P7\_GMT20211126.odt Tags: Task sequence complex

(The participants reads the task again)

Document: P7\_GMT20211126.odt Tags: Task sequence complex

I would not add anything because the results are quite clear because when you select a new case study it automatically generates a new Q here that you can compare by clicking one tab and another. I think it is very transparent so it is fine for me.

Document: P7\_GMT20211126.odt Tags: Data table features useful, Visualization useful, Query inputs and elements useful, Query process clear

Okay, I think I have downloaded that.

Document: P7\_GMT20211126.odt Tags: Simple export

MODERATOR: it is fine if you don't want to save it. It is just to see if you can download the file.

Document: P7\_GMT20211126.odt Tags: Moderator comment

Okay I could save data and metadata

Document: P7\_GMT20211126.odt Tags: Simple export

Okay, okay ... I will click here and save this again ...

Document: P7\_GMT20211126.odt Tags: Simple export

MODERATOR: Do you mind doing task 8.b please?

Document: P7\_GMT20211126.odt Tags: Simple export

So Data provenance and Data table as CSV files Oh yeah yeah, I had forgotten the export button here. Okay, now I have the CSV here.

Document: P7\_GMT20211126.odt Tags: Navigation complex

Oh I couldn't see that ...

Oh, when you download the data you get a zip file with the data table for the two queries.

Okay, I think this is very useful. Thanks for reminding me of that.

Document: P7\_GMT20211126.odt Tags: Simple export, Export (meta)data useful

I think the raw data from the CSV is useful by itself. I mean if you are searching for some environmental data for experiments the normal file format you find is CSV because it is useful for many programming languages.

Document: P7\_GMT20211126.odt Tags: Export (meta)data useful

The metadata is also useful because you have a brief description for the data you are downloading about the data itself, the legal things and so on.

Document: P7\_GMT20211126.odt Tags: Export (meta)data useful

So yeah, I think all of the files are very useful.

Document: P7\_GMT20211126.odt Tags: Export (meta)data useful

MODERATOR: Would you mind doing task 8.c before moving on to task 9?

Document: P7\_GMT20211126.odt Tags: Moderator comment

Yeah, in general terms, I think the framework is very useful and it is very easy to use.

Document: P7\_GMT20211126.odt Tags: Overall positive experience

The information provided in the interface is quite clear.

Document: P7\_GMT20211126.odt Tags: Overall positive experience

Downloading the data seems not to be a problem for a person that has used a computer once in their life.

Document: P7\_GMT20211126.odt Tags: Simple export

So I think my overall experience is very good. I have no complaints and it is quite clear.

Document: P7\_GMT20211126.odt Tags: Overall positive experience

It is great work.

Document: P7\_GMT20211126.odt Tags: Overall positive experience

Alright, so I have logged in. Let me get the other tasks open.

Document: P1\_GMT20211118.odt Tags: Task sequence complex

(The participants starts reading task 3)

Document: P1\_GMT20211118.odt Tags: Task sequence complex

MODERATOR: Could you please start in task 1?

Document: P1\_GMT20211118.odt Tags: Moderator comment

Okay, yeah that's cool.

Yeah, I think so yeah, It seems pretty sensible just to get location based data for each patient along with their environmental data.

Document: P1\_GMT20211118.odt Tags: Home tab clear

I mean it seems fine.

Document: P1\_GMT20211118.odt Tags: Home tab clear

The one thing I would change is that you keep using the word scientific data, which is kind of broad, so something more specific like environmental data or what the environmental data actually is.

Document: P1\_GMT20211118.odt Tags: Wording not clear

I would just change instead of scientific data just say environmental data or something a bit more specific.

Document: P1\_GMT20211118.odt Tags: Wording not clear

I have already done that.

Document: P1\_GMT20211118.odt Tags: Log in easy, Task sequence complex

Case study, okay so example events in Ireland so I guess this is just test data or something.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Right these are the adjudicated events so let's say that we wanted the high probability events.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Oh we can select others, as many as we want. That is good.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Maybe one thing I would say is that I would put the event categories in order. So no, possible, high probability and definite. In order of probability if you get me or maybe in reverse. Just to make it a bit more clear.

Document: P1\_GMT20211118.odt Tags: Navigation complex

Let's just say for a moment I want the high probability and definite events. I guess the numbers besides it just represent the number of people it applies to, 154 high probability events 149 definite events.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Location, this is self explanatory. Let's say I want Dublin.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Again, I can input as many as I want.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

It may be better to have them in alphabetical order.

Document: P1\_GMT20211118.odt Tags: Navigation complex

Time window lag, not sure what this is ... Not sure what the lag would mean. I guess it is from the start of the time window.

Document: P1\_GMT20211118.odt Tags: Query process complex

The time window tells you how many days you want to go back and the lag tells you how many days back you want the start time window at.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

If I say ten, I guess this means that the time window starts 10 days before the event.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

I am guessing. That is what I would guess but I am not 100% sure.

Document: P1\_GMT20211118.odt Tags: Query process complex

Okay, let me read the tasks again.

Document: P1\_GMT20211118.odt Tags: Task sequence complex

The input options are all pretty clear,

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

except the lag, which is the only one that would lead to misunderstanding. I don't know if my interpretation is right but yeah.

Document: P1\_GMT20211118.odt Tags: Query process complex

Again, maybe just making more clear time window lag. It could be obvious to other people but yeah.

Document: P1\_GMT20211118.odt Tags: Query process complex

Oh sorry I did not see the stuff underneath.

Document: P1\_GMT20211118.odt Tags: Query process complex

Okay, so I clicked if there was data available ... There are 2 events and 20 environmental data sets.

Document: P1\_GMT20211118.odt Tags: Query process clear

Envo data sets, I am not sure what it means.

Document: P1\_GMT20211118.odt Tags: Wording not clear

Okay temporal unit, you can choose from year, month, days or hours. I guess I will go with days.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Oh okay, I have a little tooltip there.

(The participant reads the tool tips from the labels)

These little tooltips are nice,

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

maybe you could add one for time window length and lag, just a very brief explanation of what they are. This would be nice.

Document: P1\_GMT20211118.odt Tags: Additional explanation, Query process complex

I think these are the options I want. Let's say the mean.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

So I guess this is running the query.

Document: P1\_GMT20211118.odt Tags: Query process clear

(The participant reads the column headings of the data provenance table) Met data okay.

Document: P1\_GMT20211118.odt Tags: Metadata content clear and useful

Let's have a look at the data lineage.

Ups ... Okay because there are only two events.

I am a little bit lost here.

If I click this .. Okay, so this is the data provenance and if I click that, okay.

Sorry, the data lineage, okay.

Document: P1\_GMT20211118.odt Tags: Navigation complex

I am not sure what data lineage means.

Document: P1\_GMT20211118.odt Tags: Wording not clear

I don't know what this is saying.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

It might not be useful for me as a statistician but maybe other people might find it useful whatever it is.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

Again, I am not really sure what this is.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

I am guessing this is stuff I would not be super interested in.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

FAIR metadata export that opens up a text, this file here I am guessing.

FAIR data export ... okay so we can export these things.

This is good, okay.

Document: P1\_GMT20211118.odt Tags: Simple export, Export (meta)data useful, Task sequence complex

I got a bit lost there but I figured that out in the end.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

I guess if you would have a lot of events the data provenance table would be a lot of information. It tells you the date, county ... There were no definite events there for Dublin, which is fine. Then, the data sets used ... Example data set yeah, okay, cool. To generate this data that is cool.

Document: P1\_GMT20211118.odt Tags: Data table features useful, Export (meta)data useful

Again, I am not sure what the data lineage text is saying.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

I am not sure what data lineage is.

Document: P1\_GMT20211118.odt Tags: Wording not clear, Metadata content not useful or confusing

Okay, and the full metadata text is just double Dutch to me.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

I don't know if this is a flaw with the UI. This is not just the thing I would know about, so that's okay.

Document: P1\_GMT20211118.odt Tags: Metadata content not useful or confusing

(The participant reads the sub task again)

Document: P1\_GMT20211118.odt Tags: Task sequence complex

So all the data sets are here and I suppose I can count them out. But I am looking for the number somewhere that I am missing. I don't think so let's count them ... 20 data sets unless I miscounted. So there seems to be 20 in there.

Document: P1\_GMT20211118.odt Tags: Metadata content clear and useful

(The participant reads the sub task again)

Document: P1\_GMT20211118.odt Tags: Task sequence complex

Okay, I can look for that ... Ups sorry ... Data use ... The data use comment just seems to be rdfs:about.

Document: P1\_GMT20211118.odt Tags: Search information process complex

Oh there is a comment here ... (The participant reads the data use comment) So this is the comment you refer to.

Document: P1\_GMT20211118.odt Tags: Search feature simple

(The participant reads the identification risk comment)

Document: P1\_GMT20211118.odt Tags: Search feature simple

I think this could be written a bit more clearly. There was a typo in: it could be linked.

Document: P1\_GMT20211118.odt Tags: Wording not clear

Okay, let's go for the last one, the dct:license.

The license is creative commons.

Document: P1\_GMT20211118.odt Tags: Search feature simple

(The participant reads the column headings descriptions)

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

So if I click on it, ups, wow. Alright okay, so I accidentally got rid of a couple of columns there.

Document: P1\_GMT20211118.odt Tags: Navigation complex

I think this is nice, the tool tips are good.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

So this is just pollution data I guess.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

this one does not have a tooltip. That is why I clicked on the other one. It was obviously the date column but the tool tip did not appear, so I thought I didn't do that.

Document: P1\_GMT20211118.odt Tags: Additional explanation

So it is just weather and pollution data.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

Okay, toggle columns .... So I can get them back here, that is good.

Document: P1\_GMT20211118.odt Tags: Data table features useful

So if you click on them they disappear. Okay cool.

Document: P1\_GMT20211118.odt Tags: Data table features useful

So obviously this is the event data here and this is the data and time.

Okay midnight ... Yes, it is aggregated over days.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

I am guessing this are outliers. Red would be a high outlier and blue a low outlier, that's what I am guessing.

Document: P1\_GMT20211118.odt Tags: Data table features useful

I am assuming you are using the rule of thumb where if it is more than 2 standard deviations away or maybe more than 3 standard deviations away. I am guessing that's what it means.

Document: P1\_GMT20211118.odt Tags: Data table features useful

This red value looks pretty high compared to the other values in SO2.

Then, the blue seems to have low values. Yeah, that I guess are the outliers.

Document: P1\_GMT20211118.odt Tags: Data table features useful

I have already done that by mistake.

Document: P1\_GMT20211118.odt Tags: Task sequence complex

If I click on the eye, it disappears and you can add them back with the toggle columns.

That seems nice. No problems there.

Document: P1\_GMT20211118.odt Tags: Data table features useful

Is it useful? Yeah, I mean it is hard to link information. As far as you can link from a table of numbers it seems fine.

Document: P1\_GMT20211118.odt Tags: Data table features useful, Export (meta)data useful

I mean visualizations are always better to understand the data instead of a bunch of numbers.

Document: P1\_GMT20211118.odt Tags: Visualization useful

I mean as far as you can display data, as informative as it can be, I think it is a good idea to highlight the outliers.

Document: P1\_GMT20211118.odt Tags: Data table features useful, Export (meta)data useful

We do see some missing values here.

Okay, so this is a bit odd. There is a blue box here for a number ... So, yeah, that is a little bit odd here. I noticed that for O3 too. There are a lot of missing values here and these values are highlighted for some reason which is a little bit odd. I guess what it might be happening is that for these missing values it is substituting them by zero or something. Then, it comes out as an outlier. So, I think it is fine that these are empty but if they are missing these probably should not be included in the outlier calculation. That's one thing you can fix.

Document: P1\_GMT20211118.odt Tags: Navigation complex

I think this is fine, just highlighting the outliers that is a good idea.

Document: P1\_GMT20211118.odt Tags: Data table features useful

Also, having this tooltips is a good idea.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

It might be a good idea to have the tooltips for all of them because yeah this one is missing for example.

Document: P1\_GMT20211118.odt Tags: Additional explanation

There is also an arrow button to order the columns when clicked. It is sorting the rows by increasing values of this variable. I can also order them in descending order.

Document: P1\_GMT20211118.odt Tags: Data table features useful

The arrows could be a bit bigger because it is easy to accidentally click on the eye icon.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

This is a good idea and you can order them back.

Document: P1\_GMT20211118.odt Tags: Data table features useful

It might be nice if these arrows would be a bit bigger.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Maybe adding tooltips for some of them because they are missing, even if it is obvious for the sake of completeness I suppose.

Document: P1\_GMT20211118.odt Tags: Additional explanation

Highlighting the outliers is a good idea

Document: P1\_GMT20211118.odt Tags: Data table features useful

just to make sure you are not highlighting missing values.

Document: P1\_GMT20211118.odt Tags: Navigation complex

Other than that I think yeah, good idea.

Document: P1\_GMT20211118.odt Tags: Data table features useful, Export (meta)data useful

Yeah, like I said. Make these arrows bigger and complete the tool tips. I am not sure if there is anything else I would add.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Maybe this is a little bit pedantic, but if I have only asked for days, it may be a good idea to format the date appropriately. You are giving me the time here, which corresponds to midnight. If I just want every hour, it makes sense. If I want days, you can maybe format the dates to the day without the actual time, just getting the literal date.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Other than that I think it is pretty good.

Document: P1\_GMT20211118.odt Tags: Data table features useful, Export (meta)data useful

(The participant reads the task again)

Document: P1\_GMT20211118.odt Tags: Task sequence complex

I have selected PM10 and there is the heat map. We have a legend here with a gradient of colors. Lower values of PM10 are yellow and high values are dark blue. Then, you have a row for each event.

This is the time window I have asked for.

Document: P1\_GMT20211118.odt Tags: Visualization useful

The graphs are interactive so that is good.

Document: P1\_GMT20211118.odt Tags: Visualization useful

If I hover over this box here ... Relative time from the event day -19 for event-147 and then it gives me the exact value 18.8, which is around the middle and it makes sense.

That is nice.

Document: P1\_GMT20211118.odt Tags: Visualization useful

This smaller heat map underneath it just seems the same heat map as above. I am not sure the usefulness of this or what it is for.

Document: P1\_GMT20211118.odt Tags: Plot design complex

I will select a different one.

Okay, yeah. This is nice.

Document: P1\_GMT20211118.odt Tags: Visualization useful

We have some very extreme values that are throwing off the gradient here. Along here it is hard to tell the differences because we have a pretty massive event compared to the other. I wouldn't know from the top of my head how you can fix that, you could try and put a maximum color gradient for the outliers. This is kind of awkward to deal with sometimes when you have outliers. You might want to think about the heat map in the presence of outliers. You have this behavior where the colors kind of blend together. Even though by its own scale this should be more distinct, because we have this massive value it is throwing off all the gradient.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

I just don't really know the purpose of the bottom plot as it seems like the same heat map but compressed.

Document: P1\_GMT20211118.odt Tags: Plot design complex

I may get the usefulness of this if I had more events. If I had 20 events, I would have this huge yolk, it is just a more compressed view of it. It would be more useful if we had more events, it is like a smaller heat map. Maybe this is okay.

This is fine.

Document: P1\_GMT20211118.odt Tags: Plot design complex

You can look into resolving the outlier issue.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Okay yeah. Again it is nice that the graph is interactive. It is telling us where the quantiles are and the exact values.

Document: P1\_GMT20211118.odt Tags: Visualization useful

So that is nice for event-147 and event-143.

I guess these are the quantiles within the time window.

Document: P1\_GMT20211118.odt Tags: Visualization useful

We have these massive outliers. Well a massive outlier, that is nice.

Document: P1\_GMT20211118.odt Tags: Visualization useful

I think personally, you know this isn't like universal advice. I personally don't like the data points down here and if you have the outliers just put them in the graph. My issue here is that you have this point here , which is very hard to see, very small points. I would personally get rid of the data points. I would just have the box plots and really large points for the outliers. But again some statisticians like to have the data points so you don't have to take this advice as the court law.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Anyway, I like box plots. It is a good idea.

Document: P1\_GMT20211118.odt Tags: Visualization useful

Yeah, we saw this in the last experiment. I remember I said I had never seen these type of plots before. This looks at wind speed in each direction and then it gives you the mean PM10, the average PM10 for that wind speed. It is a good idea.

Document: P1\_GMT20211118.odt Tags: Visualization useful

So the color tells you ...

When we had a north-eastern wind of 10 miles/hour values tend to be higher. We can see that south-eastern winds the value of PM10 was much lower.

This is actually quite nice. It is really nice. I had never seen this before myself. It is a really nice way of visualizing data.

Document: P1\_GMT20211118.odt Tags: Visualization useful

It is this kind of plot that you have to look at them for a second to understand what is going on.

I think that is fine. So yeah.

Document: P1\_GMT20211118.odt Tags: Visualization useful

Let's look at SO2. So the outlier is missing from this graph it seems ... Anyway, we do see a pretty clear correlation between PM10 and the wind speed which maybe it is expected but I don't know much about environmental stuff. Maybe it is obvious but I don't know.

Yeah you would expect that with rain, so that is cool.

Document: P1\_GMT20211118.odt Tags: Visualization useful

Maybe a histogram. It is nice to give them alongside with box plots since it tells you if there are any peaks in the data. A box plot can't really tell you that but a histogram will so it is nice to have those as well.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Yeah just, I guess the only other thing is the time series plot. Just the basic time series plot. Just a line graph of the variable over the time window you know. That would be nice as well. A basic time series plot of the variable.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Okay, let's say I just want to see the none events in Dublin.

Well not in Dublin, in Wicklow.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Let's pick a big number of 100 days with no lag.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

So there are 2 events and 1 environmental data set.

Document: P1\_GMT20211118.odt Tags: Query process clear

Let's say over 100 days, so that is kind of a long period ...

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Yeah let's select the minimum in the aggregation method.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

I don't know if this is a glitch but this stuff popped up. I am guessing that is not supposed to happen.

Document: P1\_GMT20211118.odt Tags: System issue

So Q2 only uses 1 data set and the Q1 used 20 data sets

Document: P1\_GMT20211118.odt Tags: Metadata content clear and useful

So that was dct:license, so I think so. I remember creative commons. This is right by-sa/4.0 I can't remember the exact stuff here but it was creative commons.

Document: P1\_GMT20211118.odt Tags: Search feature simple

Let me get rid of the full metadata text.

Document: P1\_GMT20211118.odt Tags: Navigation complex

We only have one environmental data set so obviously Q1, as Q2 has only O3.

Document: P1\_GMT20211118.odt Tags: Data table features useful

Is there a way to directly compare. Oh I see I see. I did not see the query tabs up there before.

Document: P1\_GMT20211118.odt Tags: Navigation complex

It opens new tabs so you can directly compare them. So yeah obviously. I did not need to check for this particular.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

Query ... hold on let me just check something.

So the data provenance table was open again so I don't know if this is a glitch or something I did.

Document: P1\_GMT20211118.odt Tags: Navigation complex

I can see that the tabs have opened up so that is nice to compare them.

Document: P1\_GMT20211118.odt Tags: Visualization useful

(The participant reads the sub task again)

Document: P1\_GMT20211118.odt Tags: Task sequence complex

Okay yeah this might be better if they were in an alphabetical order. I am not sure about this ordering but it is better alphabetical.

Document: P1\_GMT20211118.odt Tags: Navigation complex

For Q1 only one of the events has an associated O3 value, so this is a high probability event where O3 values are between 25 and 47. The O3 values for none events are very different but it is also another part of the country. This is Wicklow and the other one is Dublin. It is kind of interesting that for the high probability events data is much more spread out than it is here in Q2. These two events in Q2 might have lived in different parts of Wicklow as well.

Document: P1\_GMT20211118.odt Tags: Visualization useful

It is probably a lot of work to make it nice and easy but maybe if you could add more grouped box plots. If there is a way to compare O3 and ... Maybe not actually, hold on. Maybe not. I don't know how useful it would be. I think it would be a lot of work for something that wouldn't be so useful.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

This is fine, I would like to see a histogram and a time series plot. These would be the main things really.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

This is the data and this the metadata.

Document: P1\_GMT20211118.odt Tags: Simple export

It is saved as a ttl file, which I have no idea what is either the rdf file for the data.

Document: P1\_GMT20211118.odt Tags: Export name and format unknown

Okay, export as CSV files. That's the data table so export.

Document: P1\_GMT20211118.odt Tags: Simple export

I have just noticed that the data provenance table and the data table are automatically saved with the same name as default. Maybe use more informative names like data\_prov and data\_table, which would be more clear.

Document: P1\_GMT20211118.odt Tags: Export name and format unknown

Yeah, so the files are the data tables for each query. Okay, that is useful.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

That seems fine.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

Yeah, I mean I think so. I wouldn't see why not. This would be very useful for me because this is pretty much the exact data I need.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

So for me I would want all of the events for all of the locations. So I guess one thing you can do ... If I was to use this and I wanted to export I would want all of these for my particular research. Using the drop down to add in all the locations and events is pretty time consuming so if there was a way to add them all at once. Maybe there is already an option but it would be nice to add all of them somehow.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

I mean this is the whole thing. This is exactly what we want to select by the time window.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

Very useful.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

Yeah, I think this is pretty nice. It has most of the things I think I would want anyway.

Document: P1\_GMT20211118.odt Tags: Overall positive experience

I think this stuff is pretty clear.

Document: P1\_GMT20211118.odt Tags: Home tab clear

Maybe this could be written a little bit clearer, again scientific data could have been worded more specific.

Document: P1\_GMT20211118.odt Tags: Wording not clear

If you had an option to just input all at once here.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

This is nice as well, I already know what the events are but I did not see this information button here that tells you the interpretation of the events based on the probabilities.

Document: P1\_GMT20211118.odt Tags: Query inputs and elements useful

You just need a way to get all the data to specify all the data for a certain time window.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

Just some little hiccups here and there. Again, not every column has a tool tip and the outliers, sorry the missing values as the outliers. The missing values should not have been included in the first place in the outlier calculation. That is a big mistake so make sure that is fixed.

Document: P1\_GMT20211118.odt Tags: Additional explanation, Navigation complex

All of the metadata stuff, I don't know how useful it would be for me but it is good that it's here.

Document: P1\_GMT20211118.odt Tags: Metadata content clear and useful

This confused me as well, when I clicked on data lineage I was confused because I thought these two were the same. Maybe it was a mistake on my part. Maybe it would be better if you only could have one metadata box open at the time as it is not immediately obvious that another box would appear underneath.

Document: P1\_GMT20211118.odt Tags: Navigation complex

As I was saying, I don't know if these are particularly useful for me but this is open science stuff. Having all this information here might be useful for somebody, so that is good.

Document: P1\_GMT20211118.odt Tags: Metadata content clear and useful

Having ways of downloading/exporting the data is good as well.

Document: P1\_GMT20211118.odt Tags: Export (meta)data useful

Might be nice to have a way to throw away a query. Well I don't know if it would be important to have it. Maybe just a small X besides it to throw away the query. If I have a bunch of queries but I only wanted to download three of those in the download all, it might be a bit handy.

Document: P1\_GMT20211118.odt Tags: Additional feature suggestion

It is good, yeah, a few little hiccups here and there but it is nice.

Document: P1\_GMT20211118.odt Tags: Overall positive experience

Okay, just a question: should I answer these questions aloud?

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: yes, please do.

Document: P8\_GMT20211129.odt Tags: Moderator comment

Yes, I think the info is easy

Document: P8\_GMT20211129.odt Tags: Home tab clear

but there are some words that I do not fully understand.

Document: P8\_GMT20211129.odt Tags: Wording not clear

MODERATOR: Which ones?

Document: P8\_GMT20211129.odt Tags: Moderator comment

Mmmm for example I don't know what SPARQL is, I don't know what this is but I also don't know if this is relevant or not.

Document: P8\_GMT20211129.odt Tags: Wording not clear

Also RDF, I don't know what this acronym means but I know what the FAIR principles are.

Document: P8\_GMT20211129.odt Tags: Wording not clear

I think the information is enough

Document: P8\_GMT20211129.odt Tags: Home tab clear

Okay, that's easy I hope.

Document: P8\_GMT20211129.odt Tags: Log in easy

Then an event is possible, definite, no, high probability. So I understand this is the probability of an event to happen.

So I will select high probability of the event to happen.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

Then, location Dublin.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

One week and a lag of two days.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

I think a definition or the meaning of lag could be included but I suppose the people that are going to use this dashboard will be familiar with it. Maybe it should be included in it.

Document: P8\_GMT20211129.odt Tags: Additional explanation

Events and envo data available ...

Document: P8\_GMT20211129.odt Tags: Query process clear

Temporal units are clear because there is a definition here.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

Aggregation method ... (reads the tool tip) ... okay I will select the mean.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

The aggregation methods I do not know if they are in reference to the environmental data or to the event ... or both?

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear, Query process complex

I do not know if in the final version of the dashboard the events will make reference to an actual disease because now it is called an example event ...

So the names are the number of cases per day ... or hour ...

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

MODERATOR: Would you mind following the task order please? The data table is for task 5.

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: Apologies but I cannot hear you.

Document: P8\_GMT20211129.odt Tags: Moderator comment

Yeah, I am looking at the data provenance ... case study ... high probability events ...

So I think this is giving me the data sets that are being used for the query.

Document: P8\_GMT20211129.odt Tags: Metadata content clear and useful

I don't know if this will give useful information but we'll see, okay.

Document: P8\_GMT20211129.odt Tags: Metadata content not useful or confusing

Ah no, first I need ... oh I do not know if I am in the correct place ...

I see the same table, I don't know if I am doing this correctly, because when I click data provenance I see a table but then when I click data lineage I see the same table. Sometimes it appears ... I don't know.

It's a bit confusing.

Document: P8\_GMT20211129.odt Tags: Navigation complex

Maybe it is because I am not familiar or maybe if I look into this in detail ... but I do not see any difference, not now.

So I will probably not use these buttons.

Document: P8\_GMT20211129.odt Tags: Navigation complex

Can I move to task 5?

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: could you please answer the questions from task 4? (4 'd' and 'e')

Document: P8\_GMT20211129.odt Tags: Moderator comment

20 data sets I think ...

Document: P8\_GMT20211129.odt Tags: Metadata content clear and useful

I do not understand this question or maybe it is too simple.

Document: P8\_GMT20211129.odt Tags: Task sequence complex

I don't know if you are asking me to say that the identification risk is high probability and license ...

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Do you want to read the question aloud and then I try to help you out?

Document: P8\_GMT20211129.odt Tags: Moderator comment

Yeah, I don't understand the question

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: So first, we need to make sure we display the Full Metadata Exploration text box and we hide the other so we know what we are looking at. It could also help to read the introductory paragraph again at the top.

Document: P8\_GMT20211129.odt Tags: Moderator comment

Oh okay I understand now. When I click one of the buttons the others are still displayed? This is confusing because I don't know which button corresponds to each window or how to close the others just to look for other metadata information.

Document: P8\_GMT20211129.odt Tags: Navigation complex

So the question ... I don't know the answer to this question ...

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Maybe if you want to read the introductory paragraph again in the tab, here in Q1, the text you have at the top, it might help you.

Document: P8\_GMT20211129.odt Tags: Moderator comment

Okay, the first time I read this part I did not understand this last part.

Document: P8\_GMT20211129.odt Tags: Wording not clear

What should I read aloud? I am now looking for these comments but I do not know what to say.

MODERATOR: Maybe try using the arrows in the Control + F

So I have found this, but I don't know ...

MODERATOR: The question is only asking if you can identify the comment and that's it.

If I can find it you mean?

MODERATOR: yes

Oh okay, yes I can find them.

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Maybe try using the arrows in the Control + F

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: The question is only asking if you can identify the comment and that's it.

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: yes

Document: P8\_GMT20211129.odt Tags: Moderator comment

Oh okay, yes I can find them.

Document: P8\_GMT20211129.odt Tags: Search feature simple

Also I don't know what data provenance, lineage mean ... but I guess that's another question.

Document: P8\_GMT20211129.odt Tags: Wording not clear

Okay the first column is the date, the second would be the count of events, PM10, toluene ... well there is a definition for them.

There are many more and I think it is fine because there is a definition with the units also there.

There is no problem in this tab.

Document: P8\_GMT20211129.odt Tags: Export (meta)data useful

So I see the explanation here, so I think they are outliers below 0.05. But I don't understand if this is like the original value or the computed one.

I think it is the one that is computed but it should be there that this is not the original value but the computed one.

Document: P8\_GMT20211129.odt Tags: Navigation complex

Okay, I can hide the columns with the eye icon. The toggle column button is more useful because if you click on the eye icon you won't know if there was a column between these two. Do you know what I mean? Maybe there could be another column to indicate which column has been hidden.

Document: P8\_GMT20211129.odt Tags: Data table features useful

Maybe there could be another column to indicate which column has been hidden.

Document: P8\_GMT20211129.odt Tags: Additional feature suggestion

I think there is a lot of info on many variables ... there are a lot but maybe I am missing the origin of this data ... from where this data comes from ...

Document: P8\_GMT20211129.odt Tags: Metadata content not useful or confusing

MODERATOR: That would be in the data provenance

Document: P8\_GMT20211129.odt Tags: Moderator comment

Yeah but I don't understand it here. I don't know if this is observational or satellite data, I don't know.

Because from the data sets I don't have this info, there is only the name of the file.

Document: P8\_GMT20211129.odt Tags: Metadata content not useful or confusing

Enough features. There are a lot of parameters but I would like what I have mentioned before regarding the info about the hidden columns.

Indeed a lot of variables.

Document: P8\_GMT20211129.odt Tags: Additional feature suggestion, Export (meta)data useful

I cannot see any plot.

Document: P8\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Do you want to read the task again?

Document: P8\_GMT20211129.odt Tags: Moderator comment

Well I don't know the name in English but I think this is like a 'Rosa dels vents' or something like this Maybe it would be useful to indicate which variables have been used to compute the plot. I can see the wind direction ... but ...

Document: P8\_GMT20211129.odt Tags: Plot design complex

So I think all the data I have selected appears in this box plot for the range I have selected, the time window.

Document: P8\_GMT20211129.odt Tags: Visualization useful

So the different events would be different diseases?

MODERATOR: What do you think about the plot though? Do you think it is useful?

Yeah it is useful but ...

Different events are different diseases right?

MODERATOR: An event could be for example a maxima in Kawasaki Disease, so that would be one event. A specific date of importance would be an event.

So the box plot would be for the whole week I have selected, right?

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

MODERATOR: What do you think about the plot though? Do you think it is useful?

Document: P8\_GMT20211129.odt Tags: Moderator comment

Different events are different diseases right?

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

MODERATOR: An event could be for example a maxima in Kawasaki Disease, so that would be one event. A specific date of importance would be an event.

Document: P8\_GMT20211129.odt Tags: Moderator comment

I think a temporal series would be useful to see the evolution in PM10 or any other parameters.

Document: P8\_GMT20211129.odt Tags: Additional feature suggestion

The plots are not very clear ...

Document: P8\_GMT20211129.odt Tags: Visualization useful

Well this would be like the time series I have just asked for.

Maybe to include the value inside the square ... oh okay I can see it now.

Document: P8\_GMT20211129.odt Tags: Visualization useful

For example, if you want to use this plot ... can you take this plot for a presentation? Well you would lose the hovering info.

I think maybe you can include it on this website.

Yeah because maybe the relative time is the day in the event which won't be necessary but the value like in PM10 should appear in the square box.

It is just my opinion.

Document: P8\_GMT20211129.odt Tags: Additional feature suggestion

Okay maybe it was not clear to me because I have started with the last plot. So if I go by order it is more clear to me now.

Document: P8\_GMT20211129.odt Tags: Task sequence complex

What I still do not understand is the polar plot ... the center is Dublin or are there different stations where the data has been taken? I don't know.

I had selected 20 data sets which I don't know if this corresponds to one cycle in the polar plot ... I'm not sure.

Document: P8\_GMT20211129.odt Tags: Plot design complex

Oh here I do not understand the events ... they are possible, high probability and no but definite I don't know what it means.

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

So I will select now possible ... Oh, I can select two of them.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

So I see now that there is an info button explaining the different options but I still do not understand the definite events.

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

And maybe you can indicate that you can select one or more at the same time because I've just discovered this now.

Document: P8\_GMT20211129.odt Tags: Additional explanation

Oh and you can also select more than one county ... country or county? ... well all is Ireland I'm going to select now two events high probability and possible for the city of Cork.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful, Query process complex

Two weeks and a lag of one.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

Temporal units per hour.

Document: P8\_GMT20211129.odt Tags: Query inputs and elements useful

MODERATOR: Apologies as we have exceeded the memory capacity because there are 38 event of this category Should I remove a category from the events? The high probability. Should I change the query?

MODERATOR: If you click again click to see data availability. Maybe try to query less events.

Document: P8\_GMT20211129.odt Tags: System issue

MODERATOR: Apologies as we have exceeded the memory capacity because there are 38 event of this category

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: If you click again click to see data availability. Maybe try to query less events.

Document: P8\_GMT20211129.odt Tags: Moderator comment

I also do not know the meaning of the number besides the event categories.

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear

Is it working now?

MODERATOR: well we now have Q3 and Q2 is not with us anymore, so please continue

Maybe it should open automatically then.

MODERATOR: it does but since one query was broken it does not recognize the Q3 as the query tab to show.

Document: P8\_GMT20211129.odt Tags: System issue

MODERATOR: well we now have Q3 and Q2 is not with us anymore, so please continue

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: it does but since one query was broken it does not recognize the Q3 as the query tab to show.

Document: P8\_GMT20211129.odt Tags: Moderator comment

My second query has more events than the first one ...

Document: P8\_GMT20211129.odt Tags: Event concept and approach not clear, Metadata content not useful or confusing

I will go to the full metadata exploration to look for the license.

I can switch to Q1 which has creative commons and the second one ... the full metadata exploration is not working.

Oh wait it's down here. They use the same license.

Document: P8\_GMT20211129.odt Tags: Search feature simple

the full metadata exploration is not working.

Oh wait it's down here.

Document: P8\_GMT20211129.odt Tags: Navigation complex

MODERATOR: maybe since there is one broken query it won't work. You could try to select another variable like temperature.

Okay, so not working.

MODERATOR: Apologies, but since one query was broken you cannot complete this part but no problem.

Ah okay.

Document: P8\_GMT20211129.odt Tags: Moderator comment, System issue

MODERATOR: maybe since there is one broken query it won't work. You could try to select another variable like temperature.

Document: P8\_GMT20211129.odt Tags: Moderator comment

MODERATOR: Apologies, but since one query was broken you cannot complete this part but no problem.

Document: P8\_GMT20211129.odt Tags: Moderator comment

I do not recognize the extension of this ... ttl file

Document: P8\_GMT20211129.odt Tags: Export name and format unknown

MODERATOR: not problem the exercise is just to download the data

Document: P8\_GMT20211129.odt Tags: Moderator comment

Okay, I have already downloaded it.

Document: P8\_GMT20211129.odt Tags: Simple export

I think that's easy.

Document: P8\_GMT20211129.odt Tags: Simple export

I think this is easy and it is easy to find with the export buttons.

Document: P8\_GMT20211129.odt Tags: System issue

So I guess this downloads the data tables for all the queries.

Well I only get query 1 but it is fine I understand.

Document: P8\_GMT20211129.odt Tags: Simple export

I think this data would be very useful for our research because I am always looking for environmental data to download from different countries and stations.

Document: P8\_GMT20211129.odt Tags: Export (meta)data useful

It is still not clear to me if it uses satellite or observational data, or maybe both. So this should be included somewhere.

Document: P8\_GMT20211129.odt Tags: Additional explanation, Metadata content not useful or confusing

Also the source, if they are official stations or I don't know. Just to know the reliability of the data.

Document: P8\_GMT20211129.odt Tags: Metadata content not useful or confusing

Then, I think it is quite clear the dashboard,

Document: P8\_GMT20211129.odt Tags: Overall positive experience

maybe some minor changes that maybe would make the experience easier into understanding what you can obtain from each plot, table but most importantly from the metadata buttons on the top.

Document: P8\_GMT20211129.odt Tags: Additional feature suggestion, Additional explanation

So overall I think it is a very useful tool.

Document: P8\_GMT20211129.odt Tags: Overall positive experience

Just thinking maybe I need some extra time to digest this to think if there could be extra options that can be included.

Document: P8\_GMT20211129.odt Tags: Query process complex

I don't know what that means. I'm here just selecting things.

I am not really sure what I am doing.

Log in the credentials.

(The participant starts reading task 3)

MODERATOR: Do you mind starting from task 1? Read the home tab ...

Yeah I don't know what that means.

MODERATOR: the home tab that you see in the dashboard. The right panel has a tab named 'home'.

Okay, okay.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

Log in the credentials.

Document: P9\_GMT20211129.odt Tags: Log in easy

MODERATOR: Do you mind starting from task 1? Read the home tab ...

Document: P9\_GMT20211129.odt Tags: Moderator comment

MODERATOR: the home tab that you see in the dashboard. The right panel has a tab named 'home'.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Yeah, I guess it's efficient ... yeah.

Document: P9\_GMT20211129.odt Tags: Home tab clear

What is SPARQL? I cannot see where this is.

Document: P9\_GMT20211129.odt Tags: Wording not clear

I guess just these acronyms RDF and SPARQL, maybe just don't assume that people would know what it means. That's all.

Document: P9\_GMT20211129.odt Tags: Wording not clear

But I mean it's good to have less text, the less text the better. So I understand you don't want to include too much text but just still don't assume that people know what it means.

Document: P9\_GMT20211129.odt Tags: Wording not clear

Well no ... I mean I know nothing about events.

Document: P9\_GMT20211129.odt Tags: Event concept and approach not clear

I don't know if I was supposed to read about it or something but like right now choosing case study or events and location ...

Document: P9\_GMT20211129.odt Tags: Task sequence complex

These don't mean anything to me. I can select anything from this.

Document: P9\_GMT20211129.odt Tags: Event concept and approach not clear

Yeah I don't know what this is about. I don't know what this means.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

You can see what I have selected: possible, Dublin ... yeah I am choosing the first options because I don't know what I am selecting.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

Yeah sorry, I mean I don't know what I am doing or what I am supposed to be doing here.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Okay, no problem. These are just some example events so a specific date with a location that you can get environmental data from it.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Okay, so it just happens that Dublin does not have data or that I am choosing the wrong time window?

Document: P9\_GMT20211129.odt Tags: Query process complex

What is the time window length? Should I suggest?

Document: P9\_GMT20211129.odt Tags: Query process complex

MODERATOR: The time window would mean the series of days before the event.

Document: P9\_GMT20211129.odt Tags: Moderator comment

I mean I literally have no idea of what I am selecting here.

I have no idea of what we know in any of this.

I guess, maybe I did not read something I was supposed to read?

I am really confused right now. I don't even know how to submit a query because I don't know what would be the credible numbers here.

Document: P9\_GMT20211129.odt Tags: Query process complex, Task sequence complex

MODERATOR: Don't worry, it's fine. It is just to get an overall experience. Maybe you can try ...

Document: P9\_GMT20211129.odt Tags: Moderator comment

MODERATOR: For Dublin there are no possible events, could try another location like Wexford that has some events for example.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Okay, I will select months and mean.

Document: P9\_GMT20211129.odt Tags: Query inputs and elements useful

Okay now I see some meteorological data.

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

Then, what am I supposed to do?

Document: P9\_GMT20211129.odt Tags: Task sequence complex

(The participant does not specifically answer or comment tasks 3 a, b, c and d but due to the initial confusion the moderator did not notify the participant about skipping those tasks)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

To me, what is slightly frustrating is that I still don't know what data I got or what ... yeah.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Maybe this was explained in the informed consent.

Document: P9\_GMT20211129.odt Tags: Moderator comment

(The participant starts reading the informed consent)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

So what are we trying to get?

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: We are trying to get environmental data, like weather or pollution, that is related to a particular event.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Does this come from measurements where ... Who is the source of this information? Just because I'm a climate scientist/meteorologist and I am used to know the source.

Document: P9\_GMT20211129.odt Tags: Metadata content not useful or confusing

MODERATOR: This is one of the tasks that you have there and ...

Document: P9\_GMT20211129.odt Tags: Moderator comment

(The participant starts reading the tasks again)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Do you want to for the next task as everything comes easier?

Document: P9\_GMT20211129.odt Tags: Moderator comment

So I am a little bit confused now ... There is a data table and then there is a heat map, I don't see any map.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Do you mind going back to the tasks and complete them in a sequential manner?

Document: P9\_GMT20211129.odt Tags: Moderator comment

So data provenance, data lineage and full metadata exploration. Am I supposed to tell you how useful I think these columns were?

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Yes, and also answer the questions in 4 d and 4 e.

Document: P9\_GMT20211129.odt Tags: Moderator comment

It looks like three, possibly yeah.

Document: P9\_GMT20211129.odt Tags: Metadata content not useful or confusing

What?

Document: P9\_GMT20211129.odt Tags: Task sequence complex

(The participant reads the task again)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

Okay yeah, the data can only be used to ... (the participant reads the data use comment) Okay so now identification risk ... (the participant reads the identification risk comment) Yeah, so I can find them.

Document: P9\_GMT20211129.odt Tags: Search feature simple

I don't know, normally I just download the data set, open the file with whatever, use a dump command that downloads a NetCDF file, and I tend to read there the information I need. But also I do not use medical data, I use climate model data so I understand that here you have to put more ... Yeah I never had to put in the identification risk in my data ...

Document: P9\_GMT20211129.odt Tags: Metadata content not useful or confusing

This is a bit different from what I usually do.

Document: P9\_GMT20211129.odt Tags: Metadata content not useful or confusing

Okay, so creative commons license.

Okay, it was possible for me to find this.

Document: P9\_GMT20211129.odt Tags: Search feature simple

(The participant starts reading all the column heading descriptions) Yeah, that's fine

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

Do you mean like a green, that there is no data?

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: in your case you didn't get many background colors, so don't worry about this.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Yeah, I can.

Document: P9\_GMT20211129.odt Tags: Data table features useful

What did I do now? Oh yeah, I changed the order.

Document: P9\_GMT20211129.odt Tags: Navigation complex

Yeah, the table yeah. The table is fine I think.

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

The arrows let you sort the columns, yeah?

Document: P9\_GMT20211129.odt Tags: Navigation complex

I think the table is fine.

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

I don't know if for the previous one (refers to full metadata exploration) you needed all that text. But for the metadata I don't know if there was some ... if you think we have to search for those things or whether there is one document that has them would have been more simple. But anyway, I don't know because I don't use these things in my data, so it is completely new to me.

Document: P9\_GMT20211129.odt Tags: Additional feature suggestion

When it comes to climate data and the table it was perfectly fine.

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

No, the table is fine

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful

Let's select the wet bulb temperature.

How am I going to select it?

Do I select only one field or what?

Maybe I need your help here.

Document: P9\_GMT20211129.odt Tags: Plot design complex, Navigation complex

(The participant reads the task again)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

MODERATOR: Next to the data table tab, there are the plots. You can also hide the full metadata text exploration by clicking the button so it does not bother you there.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Sorry, I did not get it.

Document: P9\_GMT20211129.odt Tags: Navigation complex

MODERATOR: If you click again the full metadata text exploration button, you will get a cleaner view of the dashboard. Then, you can click on the heat map and go ahead.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Oh sorry, select like that. I thought I would be selecting it first from the table.

Document: P9\_GMT20211129.odt Tags: Plot design complex, Navigation complex

Okay, it didn't say in your text to go first to the heat map tab ... it says explore the heat map ... so maybe you can edit that.

Document: P9\_GMT20211129.odt Tags: Task sequence complex

I don't know what it would be ... Yeah, this measure of relative time from the event ...

Document: P9\_GMT20211129.odt Tags: Plot design complex

MODERATOR: The issue is that you only have one day per event, so it is just one value.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Oh, so maybe it doesn't seem too sensible.

Document: P9\_GMT20211129.odt Tags: Plot design complex

What was this meant to be if I had more events?

Document: P9\_GMT20211129.odt Tags: Plot design complex

Is this -1 year or what's the time unit here?

MODERATOR: I just saw that it is not displayed correctly, so thanks for spotting this out. I need to change that. Also, if you have more days, you would get more values. Now you have only a value because you selected 7 days and the time unit of months, so it collapses everything to one value.

Okay so let's change this to day as it will make more sense.

(the participant submits another query)

Okay this is better.

Document: P9\_GMT20211129.odt Tags: System issue

MODERATOR: I just saw that it is not displayed correctly, so thanks for spotting this out. I need to change that. Also, if you have more days, you would get more values. Now you have only a value because you selected 7 days and the time unit of months, so it collapses everything to one value.

Document: P9\_GMT20211129.odt Tags: Moderator comment

I still have some doubts. I don't understand what this plot is showing ... relative time from the event with the units of minus, so I don't get it.

Document: P9\_GMT20211129.odt Tags: Plot design complex

When did this event-130 happen?

Document: P9\_GMT20211129.odt Tags: Event concept and approach not clear

MODERATOR: you can check that in the data provenance table

Document: P9\_GMT20211129.odt Tags: Moderator comment

Okay so event-130, there are 8 inputs for event-130 so each of my ... oh seven inputs, wait 8 ... I don't know why there are 8 inputs for one event.

Document: P9\_GMT20211129.odt Tags: Plot design complex

MODERATOR: No problem, let's move on and I will note down that it is not clear.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Then, I have no idea what this 'relative time from the event' means.

Document: P9\_GMT20211129.odt Tags: Wording not clear, Plot design complex

Okay, so box plot ... okay.

Where the 8 points are 8 measurements, okay.

Document: P9\_GMT20211129.odt Tags: Visualization useful

Man, what the heck is this?

This is just something I have never seen before. This is probably ... this is somehow spatial and something ... what the heck?

Why is there wind speed if I selected dew point temperature?

Document: P9\_GMT20211129.odt Tags: Plot design complex

MODERATOR: Apologies, what did you say?

Document: P9\_GMT20211129.odt Tags: Moderator comment

Why does this figure display wind speed in the middle?

Document: P9\_GMT20211129.odt Tags: Plot design complex

MODERATOR: I will explain that later. If it is not clear just mention that.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Maybe it is just something done in a community ...

Document: P9\_GMT20211129.odt Tags: Plot design complex

Okay, so ... The heat maps are clear as I have seen them before. Except for the fact that to me it was confusing the time in the x axis in negative values. Perhaps it is normal in some communities but to me it was not normal.

Document: P9\_GMT20211129.odt Tags: Visualization useful, Plot design complex

The box plot was totally obvious, yeah. The box plot is fine.

Document: P9\_GMT20211129.odt Tags: Visualization useful

The polar plot is not something I have seen before, so I don't know what the meaning is ... the wind speed and these circles ... I don't know what I am seeing here.

Document: P9\_GMT20211129.odt Tags: Plot design complex

Maybe not adding more visualizations, just more explanations: what this plot presents, what's the exact x and y axis and yeah.

Document: P9\_GMT20211129.odt Tags: Additional explanation

My Q1 was a bad choice because I have selected 7 days in a monthly output, so that was quite a mess.

Document: P9\_GMT20211129.odt Tags: Query process clear

How did you say I can close this?

Document: P9\_GMT20211129.odt Tags: Navigation complex

MODERATOR: clicking again, yeah.

Document: P9\_GMT20211129.odt Tags: Moderator comment

(The participant reads the task again)

Document: P9\_GMT20211129.odt Tags: Task sequence complex

They all seem to use three.

Document: P9\_GMT20211129.odt Tags: Metadata content clear and useful

Creative commons ... and this one ... yeah both have the same license

Document: P9\_GMT20211129.odt Tags: Search feature simple

I did not find them in my case ...

Document: P9\_GMT20211129.odt Tags: Navigation complex

How do I know which window is this one? I would like to close that window

Document: P9\_GMT20211129.odt Tags: Navigation complex

MODERATOR: This is the second one, data lineage.

Document: P9\_GMT20211129.odt Tags: Moderator comment

Maybe that would help. To know which one is which because I get confused on how to close them.

Document: P9\_GMT20211129.odt Tags: Additional explanation

Oh yeah the extreme events ... Q1 in my case has nothing so Q2 looks like it has more extreme events for example for rain.

Document: P9\_GMT20211129.odt Tags: Data table features useful

I mean because my ... Let's see how Q1 looks like, can it even make a box plot?

Yeah okay, this was not a sensible choice.

Document: P9\_GMT20211129.odt Tags: Query process clear

I don't know if it would be easy for you. Because you need to select the aggregation method separately, you either choose mean or maximum and minimum. Sometimes it is interesting to be able to plot mean maximum and minimum together, but I don't know how simple this is ... not every suggestion is always useful if it requires too much work.

So yeah, maybe something like that. But I mean, it's okay how this is done, you can always change if you want minimum or maximum.

Document: P9\_GMT20211129.odt Tags: Additional feature suggestion

Okay so let's see, I am supposed to export FAIR data and metadata ... okay.

Document: P9\_GMT20211129.odt Tags: Simple export

Okay, export. That looks nice.

Document: P9\_GMT20211129.odt Tags: Simple export, Export (meta)data useful

Also data table, export. Oh yeah, that looks very pretty.

Document: P9\_GMT20211129.odt Tags: Simple export, Export (meta)data useful

Oh I don't see the download all ... oh here.

Document: P9\_GMT20211129.odt Tags: Navigation complex

Okay download, perfect. Okay well.

Document: P9\_GMT20211129.odt Tags: Simple export, Export (meta)data useful

Yeah the tables are very pretty and the graphs are very nice.

Document: P9\_GMT20211129.odt Tags: Export (meta)data useful, Visualization useful

Okay, well you know, I think this is growing more in me as we went through, yeah.

Document: P9\_GMT20211129.odt Tags: Overall positive experience

My biggest obstacle is that I don't know what I was doing at the start and then I was expecting climate data that I am used to.

Document: P9\_GMT20211129.odt Tags: Event concept and approach not clear

It took me a while to comprehend what we were doing. Now that I kind of understand the origin of your data and purpose.

Document: P9\_GMT20211129.odt Tags: Event concept and approach not clear

Yeah I don't know. Maybe if you selected 8 measurements I would write something like the measurements or number of ensemble members. Write what they really are but like it is actually quite neat.

Document: P9\_GMT20211129.odt Tags: Additional feature suggestion

You can already create a really nice figure. Are the people using this dashboard allowed to use this figure in their publication? Yeah that's really neat.

Document: P9\_GMT20211129.odt Tags: Visualization useful

Yeah and then it is fine. Okay I've been getting this information from your sheet. I guess yeah they will be able to find data use and identification risk because it is written up here, above the metadata options.

Document: P9\_GMT20211129.odt Tags: Search feature simple

So yeah well done, I think you did well.

Document: P9\_GMT20211129.odt Tags: Overall positive experience

Home tab information, that's where I start yeah?

Document: P10\_GMT20211130.odt Tags: Task sequence complex

MODERATOR: Yes

Document: P10\_GMT20211130.odt Tags: Moderator comment

Okay, yeah. I already knew what I was doing but it tells me again quite well what's doing.

Document: P10\_GMT20211130.odt Tags: Home tab clear

Okay, what was the other question?

Document: P10\_GMT20211130.odt Tags: Task sequence complex

I can't think of what else needs to be added but then again, I am coming from a background of knowing about this stuff.

Document: P10\_GMT20211130.odt Tags: Home tab clear

Sign in ... I think I typed that correctly ... okay

Document: P10\_GMT20211130.odt Tags: Log in easy

Let's go with definite events in Galway.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Let's go with a shortish time window like 4 days. Although if this was vasculitis we would want a longer one, but that's alright.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Time window lag, so this presumably means that I can go back to let's say.

Document: P10\_GMT20211130.odt Tags: Query process complex

Okay wait, time window length ... I might have a window length of a week and this presumably allows me to go back and say starting two weeks before the event.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Now, what's not clear for me is whether the time window starts 14 days ago or starts 14+7 days ago. I will go with something in the middle here so now it will be three around three weeks ago but that's not entirely clear.

Document: P10\_GMT20211130.odt Tags: Query process complex

Okay, I might as well click for data availability I think.

Okay, 6 events and 4 data sets, that seems enough to me.

Document: P10\_GMT20211130.odt Tags: Query process clear

You can tell me if I should have more events but that seems okay to me.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

And temporal units ... I think I will go with days just because I am not sure that the things that happen every hour matter much to this.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

And aggregation ... Mmmm ... we will go with mean.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

I told you about the one that wasn't.

Document: P10\_GMT20211130.odt Tags: Query process complex

Apart from that I think we are okay.

Document: P10\_GMT20211130.odt Tags: Additional explanation, Query process complex

I don't remember choosing if I wanted weather or pollution but that's fine I didn't miss anything important.

Document: P10\_GMT20211130.odt Tags: Metadata content not useful or confusing

So that's okay, you have given me some weather and some pollution information ... humidity ...

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

So the bit now I am wondering about is ... Oh yeah I see the events now: event-259, event-114 ... So these are 4 different people that had events happening to them. I am seeing three there.

Ups what happened there.

I am guessing that the other person is here ... Ups ...

So why did I think there were 4 events, I must have missed that ... there is event-490 ... 2, 3, 4, 5, 6 (the participant is counting the events) ...

Oh maybe I saw 6 events or maybe it was 6 events and 4 data sets.

Okay, I am misremembering, my memory is poor.

Okay I can understand what's going on here.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

For each of the days we have ...

Document: P10\_GMT20211130.odt Tags: Task sequence complex

MODERATOR: If you don't mind this would be the next task, could you do task 4 first please?

Document: P10\_GMT20211130.odt Tags: Moderator comment

Data provenance, so where did this information come from in relation to 6 events.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

We have these data sets with big large numbers that don't mean much to me, but I guess they relate to Met Eireann or environmental agencies of some sort.

Document: P10\_GMT20211130.odt Tags: Metadata content not useful or confusing

Okay that seems reasonable.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

The information there is pretty much identical, is it?

Document: P10\_GMT20211130.odt Tags: Navigation complex

Sorry, I will go back to the tasks for a second to make sure I am not doing anything stupid.

(The participant reads the task again)

Document: P10\_GMT20211130.odt Tags: Task sequence complex

To me they look the same, but okay ...

Document: P10\_GMT20211130.odt Tags: Navigation complex

Okay, so again it seems that the information is the same.

Document: P10\_GMT20211130.odt Tags: Navigation complex

As far as I can see, these three things are giving me the same information but I might be missing something.

Oh something happened there ...

Okay I thought the first time when I was looking at data provenance it looked different ... Oh yes it did.

There is something further down here that I am not seeing.

Document: P10\_GMT20211130.odt Tags: Navigation complex

Okay, so this is the information about ... Well I don't know what this information is about ... Okay, so this information is about lots of more information about the data sets.

Okay, I'm sure if you cared that would be interesting.

The stuff on the bottom has different information.

Document: P10\_GMT20211130.odt Tags: Metadata content not useful or confusing

Scrolling around these little screens is not so easy and ... is there anything else?

Document: P10\_GMT20211130.odt Tags: Navigation complex

I'm not sure what data provenance and lineage

Document: P10\_GMT20211130.odt Tags: Wording not clear

I'm not sure what data provenance and lineage is but there is a lot of information definitely in there, which seems to relate to locations.

Document: P10\_GMT20211130.odt Tags: Metadata content not useful or confusing

I am not sure if I need to make some sense of it except it looks like location data.

Okay I will go back here and I will assume I need to go further down.

Okay, there is some other stuff regarding metadata of what everything came from.

Document: P10\_GMT20211130.odt Tags: Metadata content not useful or confusing

For each event, previously you said that there were 4 in total.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

(The participant reads the task again)

Document: P10\_GMT20211130.odt Tags: Task sequence complex

Go back, go back ... The provenance ... I do not know why this is not looking the same each time.

Document: P10\_GMT20211130.odt Tags: Navigation complex

1, 2, 3, 4 ... (The participant counts the data sets used for all the events) Okay, I was a bit slow there.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

Okay so that presumably means I have to go through this stuff.

I would love to search for it but I am not sure how this will work.

Document: P10\_GMT20211130.odt Tags: Search information process complex

MODERATOR: Do you mind reading the first paragraph again in the tab?

Document: P10\_GMT20211130.odt Tags: Moderator comment

The first paragraph in the tab or the first here?

Document: P10\_GMT20211130.odt Tags: Task sequence complex

(The participant reads the task again)

Document: P10\_GMT20211130.odt Tags: Task sequence complex

Oh yeah, I was about to do that.

So I put this here ... It is not finding it.

Okay, whatever is going on here it is not looking in this window for what I am asking for.

So the only copy of this that I am finding is up there. So the simple answer is no, I cannot find the data use tab.

Document: P10\_GMT20211130.odt Tags: Search information process complex

MODERATOR: try again in the third text box

Document: P10\_GMT20211130.odt Tags: Moderator comment

So the one I am in now, right?

I think is probably a Firefox problem

Document: P10\_GMT20211130.odt Tags: Navigation complex

MODERATOR: No, you have three matches of your search

Document: P10\_GMT20211130.odt Tags: Moderator comment

Yeah, oh yeah. I tried that earlier and it did not work but it did work this time.

This is because I am bad at browsing.

Document: P10\_GMT20211130.odt Tags: Search information process complex

Data use okay this is the answer ... (The participant reads the data use comment) ... Okay that's the answer to that one, sorry very slow.

Document: P10\_GMT20211130.odt Tags: Search feature simple

And the second thing, the identification risk comment ... Okay, I will be way quicker this time. Here it is ... (The participant reads the identification risk comment) ... I have been very slow but I am getting better.

Document: P10\_GMT20211130.odt Tags: Search feature simple

The license I guess is doing the same job here ... we will find out very soon ... creative commons license.

Document: P10\_GMT20211130.odt Tags: Search feature simple

The date of the events, this one has no hovering but I worked out that it is the identification for the event.

Document: P10\_GMT20211130.odt Tags: Additional explanation

(The participant reads the column name descriptions)

I was able to do that.

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

I am going to assume that the read ones are out of the ordinary but let me have another look.

It must say it somewhere ... Red ones are high and blue ones are low it says there ...

I guess it is something to do with standard deviations or something.

Yeah okay, I did understand that.

Document: P10\_GMT20211130.odt Tags: Data table features useful

Yes, I guess I can do that. I really don't care about any pollution because there is nothing there so it's not that exciting.

Document: P10\_GMT20211130.odt Tags: Data table features useful

I presumably can change the ordering of that, yeah.

So now I have a weird order so I will go back to ordering it by event.

Document: P10\_GMT20211130.odt Tags: Data table features useful

The question is if I do understand what I am seeing but again I do because I am coming to this with a lot of advantages knowing what we are trying to do. Like what the weather was for people who have some kind of event in the days leading to that event.

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

I am assuming that for each individual here I am getting ... (The participant counts the number of values per event) ... Okay so these are 8 days because I said I wanted days prior to this person's event. What I can tell then is that a person had an event on the 25th or 24th of July.

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

I had long discussions about, well it is not relevant to your study but, whether we should be providing this information in this form, but I guess if I know nothing else except that event and I don't know anything else about the person. Then we are not giving away any secrets.

Document: P10\_GMT20211130.odt Tags: Additional feature suggestion

No, that's the task completed.

Document: P10\_GMT20211130.odt Tags: Data table features useful

Okay let's look at this ...

There is probably a way of getting rid of all of that up there but it's okay. Let me see if I can do that. No, I cannot deal with all of that.

Once I've dealt with all of this it would be quite nice to get rid off all of this because it is a bit in my way.

But okay, yeah okay, so ...

Document: P10\_GMT20211130.odt Tags: Navigation complex

So I am going to do a heat map of the rain

Okay so here are my six events and my 8 days running up to the time of the event.

Yeah I get that, these are the days and these are the events.

Document: P10\_GMT20211130.odt Tags: Visualization useful

This seems to be a squashed version of this essentially, which would be good if there would be a lot of events.

Document: P10\_GMT20211130.odt Tags: Visualization useful

I am afraid it does not tell me much of the use of the impact of the rain for these events, probably because the time length is too short.

But anyway, that seems perfectly acceptable.

Document: P10\_GMT20211130.odt Tags: Visualization useful

This is telling me that presumably in the 8 days leading up to the event-418 there was a very cold temperature ... but there is quite a variation in temperatures, which is typical of Ireland, isn't it?

Well for the event-418 for the person event-418, whereas for the event-490 happens after a very narrow temperature variation.

I don't think I need to say much more about that one.

Document: P10\_GMT20211130.odt Tags: Visualization useful

Uh lovely. Polar plots are good for wind direction and speed and I did not pick wind speed or wind direction so let me go back in a second.

Document: P10\_GMT20211130.odt Tags: Visualization useful

Now that I remember what a polar plot is, okay. I don't see wind speed or direction there and that is probably because you have just done it.

Document: P10\_GMT20211130.odt Tags: Plot design complex

So, what is this doing? I guess this is the total, 8 times 6 so 48 data wind speed and directions for the 8 days leading to the events.

If you were not a statistician you might say that there is some pattern going on here but I guess that we are in Ireland and most of the weather comes from the west and south-west. The strongest winds come from the south and south-west and that is why none of these people had an event after a north-eastern wind because that is very rare.

Yeah, helpful information if there was more data that would be useful, I guess.

Document: P10\_GMT20211130.odt Tags: Visualization useful

Slightly relevant that the polar plot is only relevant if you have wind direction and speed, and I did not pick that up after I had picked up dew point temperature. But that's okay.

Document: P10\_GMT20211130.odt Tags: Plot design complex

Maybe trends but no, because we think that the environmental impact on weather is probably happening at the time, it could be because there was a change. So changes on these variables over a period of time but I am not an expert in this so I don't know how you would visualize that

Document: P10\_GMT20211130.odt Tags: Additional feature suggestion

Okay let's go for a different county

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

let's have a longer time window

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Well no, let's go for another time unit like hours to have a bit of variety.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Okay, I thought it said earlier that I could choose another time window here ... Oh I need to check data availability first.

Document: P10\_GMT20211130.odt Tags: Query process clear

Oh lots more data sets! Because it was Dublin, oh was planning to take Galway out but that might be the answer. So maybe I leave Galway out, so let's check for data availability.

No data for Dublin, that's surprising.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

I will change this one as well and select more event types.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

It looks like I can add counties here which I haven't anticipated so let's just do Cork and Galway.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

MODERATOR: If you could reduce the events or the locations as it is too much for the memory

Document: P10\_GMT20211130.odt Tags: Moderator comment

So as far as I can make out all of my events used 4 data sets so I will have to check on the second one.

Okay so the answer to the first one is that his one is using more data sets. Oh no you wanted the event that was using the most but I suspect all of the events will be using the same amount of data sets.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

I was going to change this to hours but I probably end up regretting that. Let me go back and change this.

Okay now I have 24 hours per one.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

Okay so hours and mean.

Document: P10\_GMT20211130.odt Tags: Query inputs and elements useful

So I am going to read back the tasks in a second while this loads.

(The participant reads the tasks again)

Document: P10\_GMT20211130.odt Tags: Task sequence complex

Okay, so here there's the data and I need to find the full metadata exploration.

MODERATOR: We got an exceeding memory issue, let's refresh the dashboard and quickly submit two queries again.

Okay sure, oh query 1 is still up there I did not notice that. That was not immediately obvious to me so there is something to remember.

I am going to refresh and submit the two queries then.

Document: P10\_GMT20211130.odt Tags: System issue

MODERATOR: We got an exceeding memory issue, let's refresh the dashboard and quickly submit two queries again.

Document: P10\_GMT20211130.odt Tags: Moderator comment

Okay sure, oh query 1 is still up there I did not notice that. That was not immediately obvious to me so there is something to remember.

Document: P10\_GMT20211130.odt Tags: Navigation complex

So we have fewer data sets related to Wexford than Galway, which I do not know why this happens but we can come back at it another day.

Document: P10\_GMT20211130.odt Tags: Query process clear

Okay so now we have Q1 and Q2. I would have missed this bit that there were multiple queries across the top here.

Document: P10\_GMT20211130.odt Tags: Navigation complex

The data provenance suggests that there are 3 data sets here in Q2 while there were 4 for the other ones.

Document: P10\_GMT20211130.odt Tags: Metadata content clear and useful

So if I can remember it was creative commons.

Okay so now I'm going to search in the full metadata exploration for the word license.

Okay so it is creative commons. This one is the same license as the previous one.

Document: P10\_GMT20211130.odt Tags: Search feature simple

The extreme values, I have to look at the values and they are down here somewhere.

My sense is that there are slightly more red values up here than in the other. Oh well, there are more blues here and a similar amount of reds ... So this one here has more extreme values, which is Q1.

Document: P10\_GMT20211130.odt Tags: Data table features useful

I quite like the polar plot so let's go to Q2 and load the same plot.

Document: P10\_GMT20211130.odt Tags: Visualization useful

I think it doesn't matter what I pick, which is a bit confusing.

Document: P10\_GMT20211130.odt Tags: Plot design complex

Surprise, surprise it is not so windy in Wexford, which is what I would expect as it is not in the west coast as Galway.

One of the days prior to somebody getting their event there was a strong north-eastern wind, which is unusual.

But yeah, I think I have done some comparisons there. Okay so I did a bit of that.

Document: P10\_GMT20211130.odt Tags: Visualization useful

No, I can't think of anything else other than maybe seeing trends. I suppose the plots on the other side are like the box plots, you would know but I don't know how you would do. There is no temporal aspect to it ... So you would not know when the interesting winds happen but if you had enough data you might see a pattern.

Document: P10\_GMT20211130.odt Tags: Additional feature suggestion

Oh no, something struck me suddenly, I was assuming that the values out here were strong but now I am thinking they have something else. The colored tile is not stronger the more wind there is. Ah this is telling how much rain there was! So we have the rain, wind speed and direction.

Document: P10\_GMT20211130.odt Tags: Visualization useful

Okay I did not catch this the first time.

Document: P10\_GMT20211130.odt Tags: Plot design complex

This is a very strong dry north-eastern windy day and this is a very strong rainy south day. I would not expect that all the rain would come from over here but I think it's fine. And here it was a calm, not very windy or rainy day.

Document: P10\_GMT20211130.odt Tags: Visualization useful

## Okay I think that happened

Document: P10\_GMT20211130.odt Tags: Simple export

The metadata file looks different but you probably cannot see my downloads.

It is a ttl file for the metadata and a RDF file for the data, fair enough.

Document: P10\_GMT20211130.odt Tags: Simple export

Ah yes I missed this one.

(The participant reads the task again)

Document: P10\_GMT20211130.odt Tags: Task sequence complex

I am going to try this here. There's an export button so I might just click that.

Okay so I got a CSV file.

Document: P10\_GMT20211130.odt Tags: Simple export

The data might have an export button too. Oh here I missed that the first time.

Document: P10\_GMT20211130.odt Tags: Navigation complex

This one is taking a bit longer. So I have completed this now even though my computer is taking a bit to download the file.

Document: P10\_GMT20211130.odt Tags: Simple export

Yes, I know the answer to this one. The data is obviously why we are doing this and the metadata I could open and have a look to see what's in there.

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

I don't know what happens if you click on a ttl file, it says what do you want to do with this and I don't know what I want to do with this.

I will open it with the browser and see what it is. This is going beyond the task here but anyway. It is not offering me a browser.

Document: P10\_GMT20211130.odt Tags: Export name and format unknown

Anyway, I believe this would be useful.

Document: P10\_GMT20211130.odt Tags: Export (meta)data useful

Okay, so my first impression is that way better than the last time. I can't fully remember but I remember being confused and I found it difficult to connect what we were trying to do with what was going on in the dashboard. But I don't know why now but you probably remember.

Document: P10\_GMT20211130.odt Tags: Overall positive experience

It is certainly better, not perfect but ... Could somebody come into this that did not have my background information, which is considerable, make sense of it?

Document: P10\_GMT20211130.odt Tags: Event concept and approach not clear

So going back to a couple of things, I think that there's a little bit of confusion here how the starting of the time window and the end of the time window and how they relate to each other. It is slightly confusing that you can't see these things until you have checked for the data availability, and so kind of confused. Maybe you can add a note there, well there is a note that kind of implies it. But that was a little bit confusing to me.

Document: P10\_GMT20211130.odt Tags: Navigation complex

I didn't know that the queries were going to appear and still be there so I thought I was going to remember that. But of course like this it makes more sense but still it was a bit confusing.

Document: P10\_GMT20211130.odt Tags: Navigation complex

The searching thing that was not your fault it was just me being stupid.

Document: P10\_GMT20211130.odt Tags: Search feature simple

They were the main things that caused me some trouble but otherwise I think it is great.

Document: P10\_GMT20211130.odt Tags: Overall positive experience

I am sure there is a way to get rid of these things (referring to the metadata boxes) but I cannot see it at the moment. It will be quite nice because you need to keep remembering to scroll.

Document: P10\_GMT20211130.odt Tags: Navigation complex

The left panel has the login information, which I have already done.

Document: P11\_GMT20211201.odt Tags: Log in easy, Task sequence complex

The home tab looks nice.

Document: P11\_GMT20211201.odt Tags: Home tab clear

I noticed that the methodology link takes you to the GitHub page.

Document: P11\_GMT20211201.odt Tags: Additional source exploration

I will just click on Knowledge Graphs, it takes you to another web page, which is probably a technical page.

Document: P11\_GMT20211201.odt Tags: Additional source exploration

The user probably does not need the details in this understanding.

Document: P11\_GMT20211201.odt Tags: Additional source exploration

This first question refers to the home page I'm looking at at the moment right?

Document: P11\_GMT20211201.odt Tags: Task sequence complex

My only comment in this one would be for the link on the knowledge graph and methodology to be like a Wikipedia page., a simple page that describes semantic data integration framework with a simple definition of a knowledge graph. I guess just a simple definition.

Document: P11\_GMT20211201.odt Tags: Additional explanation

But yeah it is really nice laid out, you can access, explore and export the data, which is quite clear.

Document: P11\_GMT20211201.odt Tags: Home tab clear

The visual is also quite helpful. You got an event with location and time and environmental data which is linked and then you have environmental data for a patient.

Document: P11\_GMT20211201.odt Tags: Home tab clear

I have already done that.

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Here in the query tab we got example events for Ireland.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Then in the events option we have possible, definite, no and high probability. I guess this refers to a disease event so yeah.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I will go for definite but it looks like you can select more than one option so I am going to select also high probability.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

The next option is the location which is like a county, so I will choose Cork.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

It looks again that I can choose more than one option like Cork, Dublin or Limerick.

I will start with those.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

MODERATOR: For this particular case, could you please choose less locations as we currently have a memory exceeding issue. A single location would be enough

Document: P11\_GMT20211201.odt Tags: Moderator comment

Oh yeah, sorry. I got excited there, okay.

I will go for Dublin.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Time window length I will go for 4 days.

So I am looking at events that happened in Dublin for a certain period and I want to see the lag.

I will go for let's say a few days.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

But before you get to that, it gives you an option to click if data is available.

Two events and 20 environmental data sets available.

Document: P11\_GMT20211201.odt Tags: Query process clear

Okay then you can go for the temporal unit you want.

I will go for days okay.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I am provided with a case study of events in Ireland and I am interested in definite and high probable cases of events of this particular disease.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I will be focusing on Dublin because this is sort of the major city, which is more likely to have a fair amount of cases.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Time window dates length, I chose a month because it is a rare disease so I think 30 days is more likely to help in these few cases.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

The lag period I am interested in happens in the weather characteristics and environmental conditions at least two weeks before the events.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I got the information that there are two events and 20 data sets.

Document: P11\_GMT20211201.odt Tags: Query process clear

Then, I was given the option to choose the temporal unit, which I'm interested in days;

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

and then I want the mean aggregation.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I would give some little extra information on what is envo data so a definition for that.

Document: P11\_GMT20211201.odt Tags: Additional explanation

Then the time window length, so that it is fairly clear that you are interested in data for events for 30 days but does not show you what year or period the events are happening. And the time window lag seems clear. For each event that is aggregated, I am interested in the average of the last 15 days, and then of course the mean. So it is just the time window length, maybe you could even add information for these two (referring to length and lag).

Document: P11\_GMT20211201.odt Tags: Additional explanation

Then, events not sure I have clicked before. Information about the events that's clear with a definition for each of the events.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

The rest is straight forward.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

Okay I see a data set with dates, events ... (The participant reads the rest of the data table column headings)

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

Okay I see a data set with dates, events ... (The participant reads the rest of the data table column headings) Some of the values show up as being red

Document: P11\_GMT20211201.odt Tags: Task sequence complex

MODERATOR: Do you mind doing task 4 please? As this is task 5 which will be the next one.

Document: P11\_GMT20211201.odt Tags: Moderator comment

I assumed that the introductory paragraph was this one.

Document: P11\_GMT20211201.odt Tags: Task sequence complex

When you say the three buttons below what do you mean?

Document: P11\_GMT20211201.odt Tags: Task sequence complex

MODERATOR: The three buttons in task 4.c are the ones you need to explore. Apologies as I should have written it more clearly

Document: P11\_GMT20211201.odt Tags: Moderator comment

(The participant reads the task again)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

So I have an export button, a case study type, county ... (The participant reads the rest of the data provenance table column headings) So it gives me a summary of what I have selected; and then, it gives the event, which is I think a total number of events. So it says eg:event-147 and then event type the one I have selected, high probability. Then, this is the county where it happened, Dublin. The event date shows you the period 2014 December the 10th and I assume these are the times in hours.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

It also shows you where the data is coming from. It is from EEA which I think is the European Union wide data set.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

Then the second row shows another event-347 ... (The participant reads the rest of the event information).

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

I am assuming that the data set names contain the number of when they were incorporated within the framework, which is 2021 October 12th.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

Got the same thing when I click.

Okay, I started with data lineage sorry for that.

Okay, so now I am going to data provenance.

Document: P11\_GMT20211201.odt Tags: Navigation complex

Okay, I see so the data provenance provides you the information I was reading but below it gives you a summary of what has been written there. It shows the link of the data sets ... (The participant reads the information in the data lineage text) ... Okay it contains some code which I assume is hard.

Document: P11\_GMT20211201.odt Tags: Metadata content not useful or confusing

So data lineage shows you some information about the technical links

Document: P11\_GMT20211201.odt Tags: Metadata content not useful or confusing

It shows you the same table ...

Document: P11\_GMT20211201.odt Tags: Navigation complex

Okay, it describes the data sets themes ... (The participant reads the themes in the metadata report) ... It gives you a summary of what this definition means divided into different categories ... (The participant reads with interest the different sections in the metadata report).

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

MODERATOR: You don't have to read it all because it is very extensive. I would advise to focus on the tasks and then you can read the file at your own time.

Document: P11\_GMT20211201.odt Tags: Moderator comment

So in the data provenance, I will just quickly look up for the data sets that were used. There's air pollution and climate data associated with multiple events.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

(The participant reads the full metadata exploration text instead of the data provenance table) I assume these are the three data sets, one for the air pollution, one for the weather and one for the particular health events.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

MODERATOR: You have this information in the data provenance table under the column 'data sets used'

Document: P11\_GMT20211201.odt Tags: Navigation complex

Oh okay, yeah. This assumes EEA and meteorological data, one for air pollution and one for meteorological data, so two data sets.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

Would this ... okay.

Would the data use comment be in this text?

Document: P11\_GMT20211201.odt Tags: Task sequence complex

MODERATOR: Yes. If you want to read the first paragraph in this tab maybe it will help you. The introductory paragraph at the very top.

Document: P11\_GMT20211201.odt Tags: Moderator comment

(The participant reads the data use comment)

Document: P11\_GMT20211201.odt Tags: Search feature simple

(The participant reads the identification risk comment)

Document: P11\_GMT20211201.odt Tags: Search feature simple

So the license is creative commons.

Document: P11\_GMT20211201.odt Tags: Search feature simple

Okay, that is clear. Sorry I took long, as I was trying to explore but I ended up going down the rabbit hole of looking at all these values.

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Okay so this is the data.

(The participant read the data table column heading descriptions) Okay that is clear.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

The date is clear and it is in ascending order. So from the events that I selected these are the number of events that had happened in those 30 days.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

Event-147 is not clear what that is.

Document: P11\_GMT20211201.odt Tags: Wording not clear

The pollutants are shown with the unit, which is clear.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

Oh ...

Document: P11\_GMT20211201.odt Tags: System issue

MODERATOR: Could you re-submit the query?

Document: P11\_GMT20211201.odt Tags: Moderator comment

Sure, at least this query submission thing is quite quick once you have familiarized with the platform.

Document: P11\_GMT20211201.odt Tags: Query process clear

MODERATOR: If you could then please continue with task 5 as you were close to completing it.

Document: P11\_GMT20211201.odt Tags: Moderator comment

I would maybe give a ... yeah ... I guess the advantage of this is a general event for a particular disease or condition interested in. It looks fine to me.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

So the read are towards the higher end of the distribution, let me see here.

So yes, I do.

Document: P11\_GMT20211201.odt Tags: Data table features useful

(The participant forgot to complete this because of the system issue while completing this task)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Yes, very useful.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

Oh wow, these are nice plots.

So this is a PM10 heat map where the low values are yellow going up to blue, which is high.

Relative time from the event would be the lagged period. so you can see for the events: event-147, that around day 45 you would have high values of PM10; and for the event-347 it is towards days 20, 25 and 28 you see how high values of PM10.

Okay this is very cool.

Document: P11\_GMT20211201.odt Tags: Visualization useful

The X-axis represents the concentration of the PM10 and the Y-axis the events.

We got the median, the top and bottom as maximum and minimum. This has slightly higher values of PM10 compared to that one, that is cool.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Oh nice.

So we will look at wind speed and PM10 and wind direction. You can see that PM10 was higher when coming from the north-east and south-east. So generally towards the east there are high levels.

The polar plot shows the wind speed at different ranges.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Okay, that's cool. I get some sense of geographical pattern in PM10 exposure.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Yeah this is great. I like this one and the heat map specially because it shows you from the time of the event the pattern in pollution levels.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Yeah, so the heat map is really useful it shows the pattern, averaged for the time period selected and lagged days. The events that happen during this time you can see that there is a sort of lagged, increased levels of air pollution, prior to the events, which is quite informative itself.

Document: P11\_GMT20211201.odt Tags: Visualization useful

The polar plot gives you the patter I guess by direction, which is quite useful. I really like this plot.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Still Ireland, I might go for no events in Cork. I will choose about 60 days, is this a bit much?

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

MODERATOR: Yes

Document: P11\_GMT20211201.odt Tags: Moderator comment

OK, I will go for 30 days with a lag period of 30 days.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

OK, I will check the data available.

Document: P11\_GMT20211201.odt Tags: Query process clear

MODERATOR: I think this is going to be too much as there are many events for these input options. If we could choose another time window, that would be great because there are currently 25 events.

Document: P11\_GMT20211201.odt Tags: Moderator comment

MODERATOR: I think this is going to be too much as there are many events for these input options. If we could choose another time window, that would be great because there are currently 25 events.

I can go let's say for this one, and reduce to 20 days the length and the lag.

MODERATOR: Okay, great thanks. Very few environmental data sets but for this experiment it's fine.

Mmmm maybe I choose something else, any advice?

MODERATOR: Try Wexford as it has few events but enough environmental data.

Okay, yeah this is fine.

Document: P11\_GMT20211201.odt Tags: System issue

can go let's say for this one, and reduce to 20 days the length and the lag.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

MODERATOR: Okay, great thanks. Very few environmental data sets but for this experiment it's fine.

Document: P11\_GMT20211201.odt Tags: Moderator comment

MODERATOR: Try Wexford as it has few events but enough environmental data.

Document: P11\_GMT20211201.odt Tags: Moderator comment

Should I choose days or months? I will stick to days because it is easier to compare.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Oh nice, okay.

Same thing but this time we have Q1 and Q2, which are query 1 and query 2.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

(The participant reads the data provenance table for Q2)

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

(The participant reads the data table for Q2)

Now I have some data for NO2.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

(The participant does not complete the sub task)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

(The participant does not complete the sub task)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

(The participant does not complete the sub task)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Let's check the heat map for PM2.5 ... Oh it is not working.

Document: P11\_GMT20211201.odt Tags: System issue

MODERATOR: Maybe try another variable while I check if there is an issue. Try temperature please Oh yeah that's nice, don't worry. I can also compare the temperature from the previous query.

Document: P11\_GMT20211201.odt Tags: Moderator comment

So we got from the time of the event going back 20 days for that lag. These are events with a period of lower temperatures just the day before the diagnosis date but for the event. The 10th and 15th days are also periods with low temperatures, that's the main takeaway.

Document: P11\_GMT20211201.odt Tags: Visualization useful

I will check the box plot for temperature.

Lower temperatures for these two and higher temperatures for those others.

Document: P11\_GMT20211201.odt Tags: Visualization useful

Polar plot. More events coming from the south-east okay.

Document: P11\_GMT20211201.odt Tags: Visualization useful

So if I compare this to the previous one ... Okay you can definitely see some differences between Wexford and Dublin for the selected events.

Document: P11\_GMT20211201.odt Tags: Visualization useful

(The participant reads the task again)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Yeah, I think this does the job. I would not add anything. Maybe ... I'm just thinking, a summary table ... Yeah actually the heat map and box plots do show you from an exploratory point of view. I don't know if this would be an extra or too much of an extra feature if you were to get a median or mean value at the time of the event. Yeah just median or mean at the time of the event with some of the summary of the polar plot.

Document: P11\_GMT20211201.odt Tags: Additional feature suggestion

So that is just my comment. You get an overall raw data, then you can explore a summary, and explore the summary through a heat map and box plot, and then yeah. I don't know if an extra summary table is necessary but I think this is good enough. It gives you an idea of the data you queried.

Document: P11\_GMT20211201.odt Tags: Additional feature suggestion

Okay so where is the export.

Document: P11\_GMT20211201.odt Tags: Navigation complex

(The participant reads the task again)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Yeah I can download the FAIR metadata and data, this is very cool.

Document: P11\_GMT20211201.odt Tags: Simple export

Okay.

Document: P11\_GMT20211201.odt Tags: Simple export

(The participant reads the sub task again)

Document: P11\_GMT20211201.odt Tags: Task sequence complex

Oh okay in the left panel.

Document: P11\_GMT20211201.odt Tags: Navigation complex

Can I quickly download all files?

Document: P11\_GMT20211201.odt Tags: Simple export

MODERATOR: Yes, absolutely.

Document: P11\_GMT20211201.odt Tags: Moderator comment

Oh that is great. Nice.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

Okay ... I think I am done. I think when you download the data sets.

Document: P11\_GMT20211201.odt Tags: Simple export

MODERATOR: As you want.

Document: P11\_GMT20211201.odt Tags: Moderator comment

I'm sorry I took a while, as I was trying to familiarize myself much with it but I noticed ...

Document: P11\_GMT20211201.odt Tags: Task sequence complex

So for the first part, the home page is perfectly fine giving you a brief summary of how the platform works, a link of the methodology of which I think should be something about how semantic data integration framework just some basics. It could be a link to GitHub that gives a quick summary or a Wikipedia page.

Then, the knowledge graph is just a simple definition. The rest is quite straightforward.

Document: P11\_GMT20211201.odt Tags: Home tab clear, Additional explanation

The query part is really, quite self explanatory, with the case study.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

So I don't know if events ... Maybe under each case study there should be a description of the kind of things describing what kind of events are related to. For example, this is an example for an event in Ireland from this registry, and some link describing the study or how the events were recorded. But yeah, just a basic link.

Document: P11\_GMT20211201.odt Tags: Additional explanation

The description of the events was quite useful as you know what you are clicking into.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Then, the location is quite straightforward.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

The time window, as I said, is difficult but yeah if you can split it for a year or months, but this could complicate things.

The time window lag and time unit are also pretty straightforward.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

Once you have the query, which is quite nice. Then it gives you a summary of what you have asked for.

Document: P11\_GMT20211201.odt Tags: Metadata content clear and useful

The one thing about the event was that I know get it but it was given a random number that is almost a random date that you got and it shows you the data sets used.

Document: P11\_GMT20211201.odt Tags: Query inputs and elements useful

I mean, I guess for some people it would have been helpful to have a simple let's say European data set or Copernicus or something (referring to the filename). It took me a while to understand the data sets that were being used, without going down to all these tabs.

Document: P11\_GMT20211201.odt Tags: Metadata content not useful or confusing

Data provenance, data lineage ... so I don't know if it was me that I did not read the instructions properly, but it took me a while to kind of figure out the first thing you want to know.

Document: P11\_GMT20211201.odt Tags: Task sequence complex

So the identification risk and where it comes from, if it is licensed. I wonder if there could be a simplified summary of this information because the curiosity is reading what you are interested in without getting lost in it. Some would apply for the other tabs.

Document: P11\_GMT20211201.odt Tags: Additional feature suggestion

The rest is really great. Once you download the data, which is the main goal for somebody like me, when you already know the variables or exposures you are interested in environmental data.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful, Overall positive experience

Then, you want to be able to say I am interested in events that happened in this period and I would like to you know include people from Dublin, Cork and all these places. Then, I would like to know about people that were diagnosed during this period and what was the air pollution exposure from the time of the event going back. Which is really great, you have already achieved this

Document: P11\_GMT20211201.odt Tags: Overall positive experience, Query inputs and elements useful, Query process clear

specially with the summary in the heat map side of things. This is a really great summary on the type of data that you are going to be getting.

Document: P11\_GMT20211201.odt Tags: Visualization useful

The box plots are sort of general, you can always generate them yourself with the downloaded data.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful, Visualization useful

But the polar plot was something I thought was interesting, which you will be able to plot but it also gives you a visual summary.

Document: P11\_GMT20211201.odt Tags: Visualization useful

I found the heat map and polar plot useful and the box plot was okay.

Document: P11\_GMT20211201.odt Tags: Visualization useful, Overall positive experience

Once you have exported the data, I like the format, which is that you have an event which is referring to a particular person.

Document: P11\_GMT20211201.odt Tags: Export (meta)data useful

For the date time, I would only keep the date and not the time because it takes away from wanting to pick up the information as easily as possible.

Document: P11\_GMT20211201.odt Tags: Additional feature suggestion

What is not clear is that when you download the data is maybe a column saying ... Well I am looking a t Q2 that I have downloaded that has the event-32 and it was the first date that is the 19-10-2014 and then the dates go forward. Yeah so they go from October until November but you are not sure if those are lagged. So you are not sure when the event had happened. In my case, I was expecting something like October going back to September. But in this case it is probably that the event happened in November going back to October. So maybe a column with the lag that corresponds to the date as in the heat map. I think this would complete the lay out of the export.

Document: P11\_GMT20211201.odt Tags: Additional feature suggestion

This is really something I appreciate, like somebody like me who has spent some months figuring out the data linkage in R and all these things. let's say I was able to get my data into the platform, this would save people so much time. I had to do weather and pollution linkage separately. Then, I limited myself to a certain number of air pollutants, which in this case you can expand the number of pollutants.

Document: P11\_GMT20211201.odt Tags: Overall positive experience

So besides what I am saying about small formatting issues, which you can never get perfectly; this is really quite useful for researchers in the field to be able to have some automated process that just allows you. Like if we had all the money in the world and time and space this would be gold!

Document: P11\_GMT20211201.odt Tags: Overall positive experience

So that is my feedback, this is a cool platform.

Document: P11\_GMT20211201.odt Tags: Overall positive experience

Okay, yeah. I thought the information in the home tab is fairly well summarized.

Document: P12\_GMT20211206.odt Tags: Home tab clear

Okay, so I need to go and get the credentials.

Mmmm are these the correct credentials?

Document: P12\_GMT20211206.odt Tags: Task sequence complex

MODERATOR: You are missing the last character for the password.

Document: P12\_GMT20211206.odt Tags: Moderator comment

Oh, yeah that's okay.

Document: P12\_GMT20211206.odt Tags: Log in easy

Okay so the case study here is: example events in Ireland. I am assuming that's flare events there.

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

Then, we get a choice of the different flare level classifications or what I would call flare categories: possible, definite, no or high probability.

So if we wanted to look at let's say high probability flare events, we can then look in a county.

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

So I take one where it is going to be more like Dublin.

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

The time window ... Mmmm ... okay there is no help here to give information on what the time window things mean. So I am not sure what is the difference between time window length and lag.

Document: P12\_GMT20211206.odt Tags: Additional explanation, Query process complex

Okay, so with that time window length and lag I am told there are events and envo data available. So 2 events and 20 environmental data sets.

Document: P12\_GMT20211206.odt Tags: Query process clear

Time window length and lag I am a bit confused on what this is because I have no reference date to begin with. So I am not sure at what point in time the length and lag refers to.

Document: P12\_GMT20211206.odt Tags: Event concept and approach not clear, Query process complex

The temporal unit is the unit of how the environmental data comes in. We could go for days

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

we could take the average over the day.

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

I would say that I am a bit confused about the time window length and lag.

Document: P12\_GMT20211206.odt Tags: Query inputs and elements useful

Looking down at the data table now ... Looks like the two high probability events are event-147 and event-347 and I am looking back ...

Maybe now it is clear to me, when event-147 happened I look back a time window of 7 days before that and this is the data that I am getting here. This would not have been so clear to begin with but now that I am seeing the table, it is.

Document: P12\_GMT20211206.odt Tags: Export (meta)data useful

So we will take a look at the data provenance.

Okay, so I guess this is an overview of where the data actually came from.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful

The example events in Ireland are the flare cases.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful

The event ID, which I guess means something internally in the knowledge graph, does not necessarily mean anything to me as there is no patient ID attached.

Document: P12\_GMT20211206.odt Tags: Wording not clear

The event type, county and event date. So if I look down to the event day 2014-12-10, this corroborates my interpretation for the time window length and lag when compared with the data table below.

These are the two windows outlined in the data provenance.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful

The data sets used again, these are sort of references from the knowledge graph, so they don't mean anything to me.

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

I guess in terms of the traceability, which is what data provenance is about, they could be helpful if I was trying to figure out where the data is coming from.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful

Hold on now, wait ... data provenance is this one and data lineage is ... Just trying to see the differences between these two windows here.

Document: P12\_GMT20211206.odt Tags: Navigation complex

Okay, obviously there we have the snippet of code as well for the data provenance that was used to run.

Okay so ... I am assuming this data table is just staying here at the end. So I am not fully sure what the data lineage is doing, it might be very similar to the data provenance.

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

Okay, well I am missing something this morning because these are very similar or maybe it is the example that I ran.

Document: P12\_GMT20211206.odt Tags: Navigation complex

Okay, so I would say in this case two data sets since after the first dash everything else is a date and ID identifier, and the same for the second event.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful, Metadata content not useful or confusing

Mmmm ... (The participant navigates for a bit between the windows and text).

Document: P12\_GMT20211206.odt Tags: Navigation complex

MODERATOR: Maybe it would help to read the introductory paragraph again in this tab.

Document: P12\_GMT20211206.odt Tags: Moderator comment

(The participant reads the introductory paragraph again)

Mmmm ... This task is not clear to me because I don't know if I am supposed to be looking at the code or whether I am supposed to be looking somewhere else.

Document: P12\_GMT20211206.odt Tags: Task sequence complex

MODERATOR: Maybe try using the arrows that you get when using Control+F

Document: P12\_GMT20211206.odt Tags: Moderator comment

Okay, but what am I doing Control+F on?

Document: P12\_GMT20211206.odt Tags: Search information process complex

MODERATOR: The whole page.

Document: P12\_GMT20211206.odt Tags: Moderator comment

Okay, so the data use ... (The participant reads the data use comment)

Document: P12\_GMT20211206.odt Tags: Search feature simple

Okay, right I am not sure of the actual purpose of this part. If you were a user, why would you be wanting to look through what looks like it is the code it is run in the system?

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

It does not make sense for a user interface better if this was displayed summarized.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

(The participant reads the identification risk comment)

Document: P12\_GMT20211206.odt Tags: Search feature simple

And the dct:license ... Okay ... I am struggling with that one and I think I will move on.

Document: P12\_GMT20211206.odt Tags: Search information process complex

MODERATOR: The license word is not spelled correctly as it is using the US spelling.

Document: P12\_GMT20211206.odt Tags: Moderator comment

Oh license okay, the license is in US spelling which is not very helpful.

Document: P12\_GMT20211206.odt Tags: Search information process complex

Okay so the license is creative commons.

Document: P12\_GMT20211206.odt Tags: Search feature simple

So understand the column headings ... Okay.

The column headings are fairly clear and you have the tool tip that appears underneath.

Document: P12\_GMT20211206.odt Tags: Export (meta)data useful

Now the cell background colors ... Mmmm ... What I am seeing here is that the cell background colors are red and blue, so obviously there's a Z-score calculated based on the 7 days of data here that uses I am assuming a standard deviation, which I am hoping is robust.

Document: P12\_GMT20211206.odt Tags: Data table features useful

But what I am noticing, which looks a bit strange to me, is that the blue squares which indicate low scores have no data. They appear to be empty. I was just wondering what could be the reason for that. Is it that the missing data is just treated as a low number?

Document: P12\_GMT20211206.odt Tags: Navigation complex

The red squares look reasonable enough but the blue squares I don't really understand. Certainly some of them are empty; then if we go over here, other ones have values in there. For example, all three there are all empty, that is not entirely clear to me.

Document: P12\_GMT20211206.odt Tags: Navigation complex

Okay, so we can hide SO2 as well.

Document: P12\_GMT20211206.odt Tags: Data table features useful

Yeah, so I think what would be

Document: P12\_GMT20211206.odt Tags: Export (meta)data useful

If I was looking at this as a modeler researcher modeling these events, what I would say is that you can highlight the red and the blue, which seem to be only calculated based on the data in the data table. So whether that is a good representation of your now very high values, that is not clear. So for example, if I am only looking for 3 weeks back and I look for high PM10 relative to those three weeks or should I be looking for high PM10 relative to a specific window back in time from the current point.

So the standardization is useful to get the red and blue squares but I question whether the standardization should be done over a kind of a longer period. The idea of standardizing is to kind of use a pool of data that you think is reliable to standardize against. So I wonder about that I wonder as well about the empty, the empty blue squares and whether these values are missing or how they are actually treated right? For example, you can see an empty square here for Toluene, but that is not blue.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

I think the whole thing of looking for the comments in these using Control+F was more of an exercise to ... I don't know if anyone would want to do this in general but if so I am not sure for a user interface whether this kind of display code by default is actually quite necessary or helpful.

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

I mean an option to print the log out to file might be useful for someone that wanted to keep track of what they have searched for.

Document: P12\_GMT20211206.odt Tags: Export (meta)data useful

Okay so this is the box plot for PM10.

This gives us an impression of the raw values for PM10.

Document: P12\_GMT20211206.odt Tags: Visualization useful

Okay ... Mmmm ... Right, I am not 100% sure of what this polar plot should be telling me or how it is helpful. I think maybe including a text box here on the side or somewhere explaining the kind of rule of thumb cases for this. I am assuming that if you see something very far out in one direction, it is going to be meaningful in some way. It is not entirely clear to me immediately.

Document: P12\_GMT20211206.odt Tags: Additional explanation, Plot design complex

Okay so I am looking at PM10 ... (The participant counts the negative numbers in the X-axis) ... Relative time from the event day. I did notice that earlier we have two times, the time window length and the time window lag. Which of these plots refer to? Is this relative from the event the actual date?

Document: P12\_GMT20211206.odt Tags: System issue, Plot design complex

MODERATOR: This is a mistake from my side, the -1 should be a -7 based on your query as I forgot to add the lag.

Document: P12\_GMT20211206.odt Tags: Moderator comment

So this is actually from the start after the lag, okay right.

Document: P12\_GMT20211206.odt Tags: Visualization useful

I can see the usefulness of this. Again we are only looking at a week here but probably more useful if we were looking further than that.

Document: P12\_GMT20211206.odt Tags: Visualization useful

I would imagine that those heat map colors are standardized based on only the things that are in the data table. It might be interesting to think about standardizing against a fixed amount of time, like standardizing it against all the days in the previous year or something like that. I know this is a challenging one to pick but if you were to take a week of PM2.5 for the winter and summer, something that flags as high in the week in the winter might be a lot lower that something that flags as high in the summer. Using the standardization of what is in the table.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

Again, I wonder about standardization; it might make it a more useful plot if it was standardized.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

I know that you got rid of the time series that you had in previous ones but ... yeah this one could be useful.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

I am going to pick something with a ... I will pick a period of 14 days ... Okay.

(The participant selects different input options)

Now, what I am going to do is check for the more extreme values, I am going to take the max for each day.

Okay, it is thinking ...

MODERATOR: Yeah, so we have a memory exceed problem.

Oh yeah, I thought it would happen, apologies.

MODERATOR: The problem is that Cork has 19 events and that is why you don't have Q2 appearing. If you could run another location with less events, like for example Wexford it should be fine. You won't have Q2 as this one is broken, but you will have Q3 on the top.

Okay, yeah I see. Let's take a look. It is fine.

Document: P12\_GMT20211206.odt Tags: System issue

MODERATOR: Yeah, so we have a memory exceed problem.

Document: P12\_GMT20211206.odt Tags: Moderator comment

MODERATOR: The problem is that Cork has 19 events and that is why you don't have Q2 appearing. If you could run another location with less events, like for example Wexford it should be fine. You won't have Q2 as this one is broken, but you will have Q3 on the top.

Document: P12\_GMT20211206.odt Tags: Moderator comment

So I am looking at the data provenance, well it looks like I have 5 events and it is using two data sets.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful

Again, I am not sure of what these identifiers at the end of the filename mean.

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

Okay, creative commons license again.

Document: P12\_GMT20211206.odt Tags: Search feature simple

Then, I am looking at extreme values ... Okay ... Mmmm ... Interestingly there are less extreme values in this one (Q2) and I wonder why it would be.

Actually, I have a different ... Oh, no maybe less pollution data here ... So I can't see PM10 and PM2.5 here, but I can see NO2.

Document: P12\_GMT20211206.odt Tags: Data table features useful

Okay again, the max values the Z-scores we are looking at a large maximum based on what is on the table.

Document: P12\_GMT20211206.odt Tags: Additional explanation

MODERATOR: Apologies, but since Q2 was broken the plots won't have the right index and you won't be able to compare the plots so feel free to skip this task.

Document: P12\_GMT20211206.odt Tags: Moderator comment

MODERATOR: Apologies, but since Q2 was broken the plots won't have the right index and you won't be able to compare the plots so feel free to skip this task.

Okay, that is okay.

Document: P12\_GMT20211206.odt Tags: System issue

I think maybe a little bit more detail on how the Z-scores are evaluated here just to how the cells are colored. I think that would be quite useful.

Because I get the reason for using them but just yeah, I am not sure about just using the ones that are in the table.

Document: P12\_GMT20211206.odt Tags: Additional explanation

Okay, Q1 ... (The participant exports the metadata and data files)

Document: P12\_GMT20211206.odt Tags: Simple export

I am looking at the FAIR data which is a RDF file and the metadata is a ttl file.

Document: P12\_GMT20211206.odt Tags: Export name and format unknown

MODERATOR: Could you please confirm that you got a pop out window?

Document: P12\_GMT20211206.odt Tags: Moderator comment

Yeah, I can confirm that I got a pop out window to download the files.

Document: P12\_GMT20211206.odt Tags: Simple export

You won't be able to see this probably, but I am going to check the data table and that everything is in there.

Document: P12\_GMT20211206.odt Tags: Simple export

So it comes as a CSV file called Q1.csv and it has 17 rows of data and the various columns with dates involved.

I am going to share this in a moment because I wanted to talk about the csv file I got from the download all option.

Document: P12\_GMT20211206.odt Tags: Export (meta)data useful

Yeah, so I would say it is good but there are some things that are not necessarily clear straight away. So again, maybe more sort of help in the tabs to help explain things, small things like the time window length and lag. These things might be clear for the use case, for people that are actually using these tools.

Document: P12\_GMT20211206.odt Tags: Additional explanation

I mean one thing would be in the home menu, to add an extra tab called you know general directions, where you could just have a diagram showing the time window length and lag with a rough schematic to tell people of what they should be doing.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

The Z-scores can concede the intent of their usefulness but I would wonder about the general usefulness if it is only based on the data that is on the data table. That is because, if we could imagine, you know PM2.5 is probably much higher in the summer than in the winter. If we were to take a period of two weeks over the summer, then a high PM2.5 value over those two weeks would be very different from a high value of PM2.5 during two weeks in the winter; but both would be colored red with a high Z-score. They are not really as comparable as we might think initially, so better find a better data set to compute Z-scores against.

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

Yeah, I did not fully grasp the data provenance and lineage bit, but I do not know much about SQL but I imagine that these are all very useful for actually keeping a hard record of the data search that was carried out and so on.

Document: P12\_GMT20211206.odt Tags: Metadata content clear and useful, Metadata content not useful or confusing

I think the heat map plot was very nice to show a general profile of the pollutant over the time since the event.

Document: P12\_GMT20211206.odt Tags: Visualization useful

I guess the RAM is not going to be an issue when this is up and running properly so I won't say much about that.

Document: P12\_GMT20211206.odt Tags: System issue

So yeah, I think ... I suppose you can choose to view those things or not as I don't find them very useful myself, but you can kind of click on and off which is good (referring to the metadata displays).

Document: P12\_GMT20211206.odt Tags: Metadata content not useful or confusing

So one thing that I mentioned previously is to be able to look at the patient level, to look for somebody that has two events for something, a time series of events. I think this is a bit challenging for this set up for how things are but ...

Document: P12\_GMT20211206.odt Tags: Additional feature suggestion

It seems good but I just cannot understand ... (The participant reads the step that was not clear again) ...

Document: P13\_GMT20211203.odt Tags: Wording not clear

Oh yeah I understand now.

That seems good, seems okay I think.

Document: P13\_GMT20211203.odt Tags: Home tab clear

Are these links part of the second task or can I check them here? ... Okay I will check them here.

So, okay methodology ... Steps to run the dashboard ... Oh the code you wrote for this, okay.

Knowledge graph, let's check that ... Oh log in, I need a username and password ...

Document: P13\_GMT20211203.odt Tags: Additional source exploration

Okay it seems good, as a starting is good.

Document: P13\_GMT20211203.odt Tags: Home tab clear

Okay, go here.

Document: P13\_GMT20211203.odt Tags: Log in easy

Case study, example events Ireland. I can't understand the meaning of the case ... Oh okay I understand.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

Example events of Ireland, it means you can add different case studies in this, which is good, great.

Document: P13\_GMT20211203.odt Tags: Query process clear

Possible, definite, no or high probability, so we can select them separated but, can I select more than one? Oh good good.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

I will just select one of them as I am curious about for example definite events, which have 149 records in this.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

It is good Location, for example just in Dublin as I just know Dublin.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

Time window length in days ... Does this mean the length of the follow up? Oh I don't know. It is a little weird for me this part, the time window length. It could have different meanings for different researchers, you know. For example, I think this is the follow up time for every patient. For me that would be the case and I will select 4 months.

Document: P13\_GMT20211203.odt Tags: Query process complex

MODERATOR: Could you please reduce the value to a few days as I currently have a problem with the memory for this dashboard?

Document: P13\_GMT20211203.odt Tags: Moderator comment

Yes, yes, no problem. I just wanted to try. I will select 5 days. No problem at all.

Document: P13\_GMT20211203.odt Tags: System issue

Time window lag ... I suppose it is about the history of the patient so I will select 5 days. I am not sure about this one. For me, it refers to the history of the patient before the diagnosis.

Document: P13\_GMT20211203.odt Tags: Query process complex

(The participant checks for available data but none is available) I think it is about these options ... Maybe I check again selecting for possible, definite, no or high probability.

Document: P13\_GMT20211203.odt Tags: Query process clear

Okay, 3 events and 20 envo data sets available ...

Document: P13\_GMT20211203.odt Tags: Query process clear

What? ... Oh environmental data sets.

Document: P13\_GMT20211203.odt Tags: Wording not clear

Temporal unit, I don't have anything to say about this, it is clear. I will choose daily.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

Aggregation method, I will select the mean.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

Then, I submit.

Document: P13\_GMT20211203.odt Tags: Query process clear

I am excited to see what I get.

Document: P13\_GMT20211203.odt Tags: Overall positive experience

Okay ... (The participant reads the task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

Well, this part seems good. It's good, it seems good and enough I think.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

Okay ... (The participant reads the column headings of the data provenance table).

Okay, all the events are high probability in county Dublin.

Event date ... that means ... Okay, the information is summarized according to the event date.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

I can't understand the meaning of data sets used. I can't understand this part.

Document: P13\_GMT20211203.odt Tags: Metadata content not useful or confusing

This is telling me that the information was extracted from these data sets.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

Does the event name contain the number of observations? Yeah, event-1 ... Is this the number of events or the number of cases? We have longitudinal data so these are different as we have more than one event per person, so that is different.

Document: P13\_GMT20211203.odt Tags: Wording not clear

The time is the time for the event data, which is great for survival analysis as we need the exact time.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

Again, this is the number of cases maybe. That would be great to show, the number of events for which case, the number of visits. In my language that would be the observation number for every case. Something like this would be good.

Document: P13\_GMT20211203.odt Tags: Wording not clear

Okay that is fine.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

I am not completely sure about the data sets but I think two groups of data are used there. I hope so.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

(The participant reads the sub task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

I cannot understand the task here. I do not have any idea about the last question 4.e.

Document: P13\_GMT20211203.odt Tags: Task sequence complex

MODERATOR: As it says in the task, you need to go to the full metadata exploration.

Document: P13\_GMT20211203.odt Tags: Moderator comment

Oh, yes, the next part.

I will go to data lineage but this is like the last part exactly. No there is something here ... What is this?

Document: P13\_GMT20211203.odt Tags: Navigation complex

This is about the information about the linkage of the data, which is not useful to be but it probably is for a computer scientist ... and maybe for me, because I need to link different data sets. Yes, this is good because you open all of your information. That is great if I understand it.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

(The participant reads the sub task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

It's ... I am not sure if this is in the metadata exploration.

Document: P13\_GMT20211203.odt Tags: Navigation complex

MODERATOR: Do you want to read the introductory paragraph again at the top of the tab?

Document: P13\_GMT20211203.odt Tags: Moderator comment

(The participant reads the introductory paragraph again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

I need to search with Control+F. Oh, yeah I forgot about that. I am like a fish, a red fish, just have 3 min of memory (joking).

Okay so I need to search about that.

Data use ... okay, there is another part here ... (The participant reads the data use comment) ... Oh fantastic! Good.

Document: P13\_GMT20211203.odt Tags: Search feature simple

Explaining this is good, great ... So you put keywords here ...

Identification risk ... Okay ... (The participant reads the identification risk comment) ... Do I need to read all of them? That's good to keep you informed about it.

Document: P13\_GMT20211203.odt Tags: Search feature simple

License ... So licenses are for rules? ... (The participant reads the license) I think this is good for research so you can open your data.

Document: P13\_GMT20211203.odt Tags: Search feature simple

That means ... Are these the metadata buttons ... Ups ... I downloaded something, which I am not sure why. I think it is the full metadata.

Document: P13\_GMT20211203.odt Tags: Navigation complex

MODERATOR: Do you want to read the task again?

Document: P13\_GMT20211203.odt Tags: Moderator comment

(The participant reads the task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

I think I cannot understand the meaning of 'underneath'?

Document: P13\_GMT20211203.odt Tags: Wording not clear

MODERATOR: It means below, at the bottom.

Document: P13\_GMT20211203.odt Tags: Moderator comment

(The participant reads the column heading descriptions)

For me it is clear, great, fine.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

Can I filter or something like that? Could you include more options for that?

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

Does this order by the maximum values? (referring to the ordering of the columns)

Document: P13\_GMT20211203.odt Tags: Navigation complex

Oh, how can I see PM10 again? I don't know how to find it.

Document: P13\_GMT20211203.odt Tags: Task sequence complex

MODERATOR: Do you want to read the task again?

Document: P13\_GMT20211203.odt Tags: Moderator comment

(The participant reads the sub task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

Okay, that is great. It is like filter columns. That is great.

Document: P13\_GMT20211203.odt Tags: Data table features useful

I would like for example to select PM10 in a special range.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

For me this is useful data to analyze not to understand something. As my background and need I would need a special view about min/max view for different events.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

It is good overall data for me as I could learn some basic descriptors of the data set.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

It is good, really tidy and that is good. I like this.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

Is there any demographic information? It would be good to include clinical, demographics and environmental information at the same time, not separately, at least for me.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

Oh, what happened? Oh okay ...

Document: P13\_GMT20211203.odt Tags: Navigation complex

Oh that's the heat map. It has more information about it.

It is a great map. I like it.

The X-axis is the relative time from the event ... Okay, that's good.

Document: P13\_GMT20211203.odt Tags: Visualization useful

I will select SO2 again. That is good.

Can we see the number of the outliers? Oh yeah, great informative plots! I really like this.

It is easier like this and good.

Document: P13\_GMT20211203.odt Tags: Visualization useful

I like your plots, they are great.

Document: P13\_GMT20211203.odt Tags: Visualization useful

Do we have a total for all the events for example? Oh no, just the events separately.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

I am not sure of what a polar plot is exactly ... I will try SO2.

Is it a type of circular plot? Oh, okay, does the polar plot mean something like geography? I can see something more in the south here.

I don't know about polar plots.

Document: P13\_GMT20211203.odt Tags: Plot design complex

It would be good if you can show more than one. You know about R so you can use something like ggplot. Plots like that if you understand me.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

It was great. I like this part so much.

Document: P13\_GMT20211203.odt Tags: Visualization useful

(The participant inputs different input options in the query panel)

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

What have I just downloaded? Well, I will remove this.

Document: P13\_GMT20211203.odt Tags: System issue

Okay, so Q1 and Q2.

Document: P13\_GMT20211203.odt Tags: Query inputs and elements useful

I need to go to data provenance ... okay ... there are the same.

Is this good? Okay. These two are the same, this part is clear.

Document: P13\_GMT20211203.odt Tags: Metadata content clear and useful

I will check their license ... I need to go to full metadata, I remember.

Is this the license for Q2? Then, the same license. Oh great I found it.

Document: P13\_GMT20211203.odt Tags: Search feature simple

Probably the first one because it had more cases but let me check it. Yes, the first one because it had more values overall I guess.

Document: P13\_GMT20211203.odt Tags: Data table features useful

Yes, going to the plots is my favorite part.

Document: P13\_GMT20211203.odt Tags: Overall positive experience

I will select SO2 in the heat map.

That means I have only selected two values per event as it says in the relative time from the event.

Document: P13\_GMT20211203.odt Tags: Visualization useful

I selected some different measures for both of these. Can I see both of them on the same page?

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

Relative time from the event ... I am not sure about how to interpret heat maps, I just have this and this which is different in the second one.

Okay, I have compared them.

Document: P13\_GMT20211203.odt Tags: Visualization useful

I am not sure what the aggregation method means now. Does this mean that for every variable you use the sum? Oh, okay I understand now. This is about an aggregation method, maybe it is because there is a different definition in my field.

Document: P13\_GMT20211203.odt Tags: Wording not clear, Query process complex

We can also download them so I think it is fine.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

Maybe include a summary of the data table with the aggregation method and the rest. But if we can download it, I think it is fine.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

MODERATOR: Do you want to read the task again?

Document: P13\_GMT20211203.odt Tags: Moderator comment

(The participant reads the task again)

Document: P13\_GMT20211203.odt Tags: Task sequence complex

Is it not working when I click?

Document: P13\_GMT20211203.odt Tags: Task sequence complex

MODERATOR: You have previously blocked the downloads from this site so you need to unblock them there in the top of the browser.

Document: P13\_GMT20211203.odt Tags: Moderator comment

Okay, FAIR data and metadata.

Document: P13\_GMT20211203.odt Tags: Simple export

Oh, yeah. I have downloaded the file.

Document: P13\_GMT20211203.odt Tags: Simple export

I will open the csv file and ... Okay this is about the data.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful

What is the data table? Where is this? ... There is no data table ...

Oh, it is at the end of the page ... I think I clicked something that I cannot remember.

Document: P13\_GMT20211203.odt Tags: Navigation complex

MODERATOR: You can use the Control+F if it helps you find the data table.

Document: P13\_GMT20211203.odt Tags: Moderator comment

Download all data tables ... Oh my god! It was right there! right in front of me (joking).

Document: P13\_GMT20211203.odt Tags: Navigation complex

MODERATOR: Can you do 8.c also?

Document: P13\_GMT20211203.odt Tags: Moderator comment

Okay, this is a zip file, good.

Document: P13\_GMT20211203.odt Tags: Simple export

Can I link both of the queries with each other? No, it is not so important as it will make it messy.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

Definitely, all of this information in this shape/format is important.

Document: P13\_GMT20211203.odt Tags: Export (meta)data useful, Overall positive experience

But for example, the information is linked with events and date time but for me it is important to have the number of observations for every case and the exact time for every case. I also need additional demographic information like age and gender together with environmental information, which will be very great. Age and gender is usually demographic information so if you can add these demographics at least it would be great. The clinical information would be harder but I think demographics would be enough.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

Generally, if I want to summarize my idea about the dashboard I would say that it is great.

Document: P13\_GMT20211203.odt Tags: Overall positive experience

It would be better to add the information I just said together with a range in the number of observations for every case. For example, not just the time.

Document: P13\_GMT20211203.odt Tags: Additional feature suggestion

It is great, well done.

Document: P13\_GMT20211203.odt Tags: Overall positive experience

MODERATOR: Would you mind answering the questions in the task aloud please?

Document: P14\_GMT20220124.odt Tags: Moderator comment

So, I got an overall understanding.

Document: P14\_GMT20220124.odt Tags: Home tab clear

There were a few syntax errors and typos in there.

Document: P14\_GMT20220124.odt Tags: Wording not clear

I don't think I would add anything besides giving an example of a potential use case. Moving from the abstract to the specific.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion, Additional explanation

Okay, so I've logged in already.

Document: P14\_GMT20220124.odt Tags: Log in easy

It is not clear what those events are.

Document: P14\_GMT20220124.odt Tags: Event concept and approach not clear

I am looking at these possible, definite, no or high probability, which implies we are talking about flare events I suspect.

Event classified as definite by an agent.

Document: P14\_GMT20220124.odt Tags: Query inputs and elements useful

It is hard to get rid of that (referring to the event pop up info box). Oh, I need to click that again.

Document: P14\_GMT20220124.odt Tags: Navigation complex

I can see that these are possible, definite, no or high probability flare events. From a researcher perspective it does not make a lot of sense to separate these out but irrespective I will choose definite and high probability. Generally, when we do research we combine these two groups.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

So we are going to look at the county now.

Document: P14\_GMT20220124.odt Tags: Query inputs and elements useful

Time window length ... I presume that means the interval to query.

Document: P14\_GMT20220124.odt Tags: Query process complex

I think this should be months rather than days. At least in my use case, it does not make sense to look at a day interval. But we can increase this to 44 days.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

MODERATOR: Would you mind selecting only a few days due to a memory problem I have with the current deployment of the dashboard?

Document: P14\_GMT20220124.odt Tags: Moderator comment

Time window lag ... This is presumably looking at introducing a lag period of X amount of days. For a flare event, we might want to look at 10 days.

Document: P14\_GMT20220124.odt Tags: Query inputs and elements useful

Events and envo data sets, envo I guess it means environmental. This is not exactly clear, I would use the term environmental there.

Document: P14\_GMT20220124.odt Tags: Wording not clear

So there are 19 events and 4 data sets being buried.

Document: P14\_GMT20220124.odt Tags: Query process clear

(The participant reads the temporal unit tool tip)

I will pick days.

Document: P14\_GMT20220124.odt Tags: Query inputs and elements useful

We are asked to choose an aggregation method but of course the aggregation method depends on the variable under discussion. It will differ if we are talking about rainfall or whatever. I don't think you can apply an aggregation method to the overall data set.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

For the purposes of this we will choose min/max, this should be min and max. There is no rationale to choose min or max separately. These two should be combined.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

Particularly, I would use the term environmental instead of envo.

Document: P14\_GMT20220124.odt Tags: Wording not clear

I would add what is an event and what the event types mean. These were not very clear.

Document: P14\_GMT20220124.odt Tags: Event concept and approach not clear, Additional explanation

Comprehend is a poor word there. Yeah, so this sentence needs to be rewritten as I don't really know what it means.

Document: P14\_GMT20220124.odt Tags: Wording not clear

Maybe you can put a little information box to explain the meaning of data provenance and lineage. Perhaps, just put one beside each of these buttons to provide a bit more information. Particularly, these two on the left.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

So data table, I would separate those terms into data and table, not data table.

Document: P14\_GMT20220124.odt Tags: Wording not clear

So how do we know if data is normally distributed? Are we supposed to know this before we start?

The assumption is that we know whether it is normally distributed. You might need something there to advise people about that.

Document: P14\_GMT20220124.odt Tags: Additional explanation

DateT I don't know what the T means ...

Document: P14\_GMT20220124.odt Tags: Wording not clear

MODERATOR: Before exploring the data table, do you mind completing task 4 please?

Document: P14\_GMT20220124.odt Tags: Moderator comment

So, again for usability all of these are in variable names in a database. Perhaps, these can have a more readable data label.

Document: P14\_GMT20220124.odt Tags: Wording not clear

I don't know what eg: means here, it implies you are providing an example of something. I presume each row represents an event, so if you would take out the eg: it would make sense.

Document: P14\_GMT20220124.odt Tags: Wording not clear

The numbers in the events are not clear what they represent ... Are they use case IDs? ... It is not clear what the number means but maybe it is not important.

Document: P14\_GMT20220124.odt Tags: Metadata content not useful or confusing

So we have 6 events here, it says that we have 19 here though ... Oh sorry, you can scroll down to the other events.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

These are the event dates. Given that this case study refers to patients, it would be important to know whether these events refer to separate patients or whether there are multiple events per individual. That is important from a statistical point of view.

Document: P14\_GMT20220124.odt Tags: Wording not clear

Again, in the data sets used the eg: does not make sense here. Data set EEA, I guess an expert user might understand what the names mean but I guess these are the data sets that describe the environmental data. So I guess it means something if you understand them.

Document: P14\_GMT20220124.odt Tags: Metadata content not useful or confusing

So that on first inspection seems to be the same. The data lineage and provenance seem to be the same ...

Document: P14\_GMT20220124.odt Tags: Navigation complex

MODERATOR: Every time you click a button a new area is displayed below.

Document: P14\_GMT20220124.odt Tags: Moderator comment

Oh okay, so this needs to be improved with a little information button.

Document: P14\_GMT20220124.odt Tags: Additional explanation

What we are looking at here is like SPARQL code perhaps in RDF. I can see the query, so I would say it is the code underpinning the query.

This would be useful for a technical user who wants to understand exactly how the data was generated.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

So interesting that the agents are represented as an ORCID. It will not be a bad wait to do it assuming that person has an ORCID. It might be useful.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

(The participant reads the full metadata text in detail showing interest in the content) Storing the data like this in RDF is nice.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

What we are missing here is the RDF for the event, how is that represented and linked to a patient. What the event and patients are in a standardized ontology.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

4 data sets.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

(The participant reads the data use comment)

Document: P14\_GMT20220124.odt Tags: Search feature simple

(The participant reads the identification risk comment)

Document: P14\_GMT20220124.odt Tags: Search feature simple

(The participant reads another description that i not the license)

Document: P14\_GMT20220124.odt Tags: Search feature simple

MODERATOR: Would you mind reading right next to the text that has been highlighted in your text) Oh, okay that is the license: creative commons. This is good for reusability.

Document: P14\_GMT20220124.odt Tags: Moderator comment

I see we have a series of environmental pollutants, pollution data is empty apart from PM10. Otherwise, we would only have weather data.

Document: P14\_GMT20220124.odt Tags: Export (meta)data useful

There are occasionally some pollution values. That may be because there is no daily data, which highlights the importance of using interpolated data.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

(The participant reads the column heading descriptions)

That is very nice.

Document: P14\_GMT20220124.odt Tags: Export (meta)data useful

Some of the descriptions are missing or cut off so I am not sure if this is the end or not.

Document: P14\_GMT20220124.odt Tags: Additional explanation

So red means high values and blue low values based on a low and high Z-scores and assuming the data is normally distributed as advised above.

Document: P14\_GMT20220124.odt Tags: Data table features useful

That seems enough to do that. The eye icon is not really like an eye but I could guess it.

Document: P14\_GMT20220124.odt Tags: Data table features useful

Yes, it is very useful assuming that the events that we are studying and the environmental variables displayed are relevant to the experimental question.

Document: P14\_GMT20220124.odt Tags: Export (meta)data useful

You are also listing here the events but it does not seem that all the events are listed here. Only a few of the events are listed, there were 19 events but only a few are displayed. It could be that there was no environmental data for it. I think that needs to be made clear.

Document: P14\_GMT20220124.odt Tags: Navigation complex

I think you need to add a filtering based on the events in question. The rest I have discussed the features as I went along really.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

Okay so we can select from a list of variables here. I will choose PM2.5.

I am going to get rid of this metadata thing first.

MODERATOR: Could you please try another variable while I explore why PM2.5 is not working?

It does not seem to be working?

MODERATOR: What browser are you using?

Firefox

MODERATOR: Then, it should be fine.

The box plot is working though

MODERATOR: Let me restart one second. Could you refresh the website and submit another query please?

Okay so I will login and submit another query.

MODERATOR: I would advise to select a region with fewer data like Wexford.

Sure, I would sort the locations alphabetically.

Document: P14\_GMT20220124.odt Tags: System issue

I am going to get rid of this metadata thing first.

Document: P14\_GMT20220124.odt Tags: Navigation complex

MODERATOR: Could you please try another variable while I explore why PM2.5 is not working?

Document: P14\_GMT20220124.odt Tags: Moderator comment

MODERATOR: What browser are you using?

Document: P14\_GMT20220124.odt Tags: Moderator comment

MODERATOR: Then, it should be fine.

Document: P14\_GMT20220124.odt Tags: Moderator comment

MODERATOR: Let me restart one second. Could you refresh the website and submit another query please?

Document: P14\_GMT20220124.odt Tags: Moderator comment

MODERATOR: I would advise to select a region with fewer data like Wexford.

Document: P14\_GMT20220124.odt Tags: Moderator comment

Sure, I would sort the locations alphabetically.

Document: P14\_GMT20220124.odt Tags: Navigation complex

Data seems to be dependent on where you are looking.

Document: P14\_GMT20220124.odt Tags: Query process clear

MODERATOR: Can you try temperature to see if the problem is the variable name selected?

Document: P14\_GMT20220124.odt Tags: Moderator comment

So we have relative time from the event -10, okay.

Now, I have all the events lined-up here. That is nice, you should be able to tell if there is a common environmental factor there within those ten days.

Then, perhaps it would line-up and you can visualize that.

This is the mean temperature reported.

Document: P14\_GMT20220124.odt Tags: Visualization useful

We could include some controls in another query.

Document: P14\_GMT20220124.odt Tags: Query inputs and elements useful

Did I break the query again?

MODERATOR: It does not say anything in the log for now. There is another task, which is to run another query, so you won't have to do it again. I believe that if we stay with this amount of data we should be fine for the experiment.

Document: P14\_GMT20220124.odt Tags: System issue

MODERATOR: It does not say anything in the log for now. There is another task, which is to run another query, so you won't have to do it again. I believe that if we stay with this amount of data we should be fine for the experiment.

Document: P14\_GMT20220124.odt Tags: Moderator comment

I think it would be very helpful in this heat map to define the category for this events, which are no and definite. To see if there is an effect in the occurrence of the event.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

That is the rain, seems to be working better now.

Document: P14\_GMT20220124.odt Tags: Visualization useful

Let's try PM2.5 ... Well it does not like PM2.5.

Let's try NO2, again not working. So it must be the pollution ones.

Document: P14\_GMT20220124.odt Tags: System issue

There are some rainy days here and there.

Document: P14\_GMT20220124.odt Tags: Visualization useful

You are displaying a lot of pop ups there which are not very clear. Oh yes, the median, min, max and the quartiles.

Document: P14\_GMT20220124.odt Tags: Visualization useful

Okay, although we submitted a query with the mean you are displaying the data as a box plot. This allows you to infer the median.

Okay, so this is the precipitation amount. However, it is not clear what it represents. Is it the amount over a period or a specific date?

I am assuming it represents the median value for the whole observation period.

Document: P14\_GMT20220124.odt Tags: Visualization useful

It also seems that there are more box plots that there are events, so one represents the mean and the other the median? It is not clear why there are more box plots. I think there are events in the Y-axis that are not displayed there. I think this is an issue with the size of the plot in the Y-axis.

Document: P14\_GMT20220124.odt Tags: Plot design complex

Very pretty but I am not sure what it means though.

Document: P14\_GMT20220124.odt Tags: Visualization useful, Plot design complex

We have north, east, south and west.

We have presumably high and low rainfall but I don't see how this reflects or relates to the events that are occurring.

It looks like it is reflecting wind speed also.

Document: P14\_GMT20220124.odt Tags: Plot design complex

If we go for temperature, it does not work but it works for PM2.5.

Document: P14\_GMT20220124.odt Tags: System issue

So it shows an association between the wind direction and speed with the variable of interest.

Document: P14\_GMT20220124.odt Tags: Visualization useful

It is hard to understand and it definitely needs an explanation of what it means.

Document: P14\_GMT20220124.odt Tags: Additional explanation

I think I would like a scatter plot as it is more transparent on the data you know.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

The box plot is useful though to show the min, median and max.

Document: P14\_GMT20220124.odt Tags: Visualization useful

I already did that. Easy to do.

Document: P14\_GMT20220124.odt Tags: Query process clear

The first one used 4 data sets and this one uses 3. I can tell from the different variables presented to me that they are using different data sets, which is a big issue when you are comparing people across different areas.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

Right, okay. It is just to go in here.

This one has a creative commons license, which is the same as the other one.

Document: P14\_GMT20220124.odt Tags: Search feature simple

It is hard for me to say at the moment. I think to do that I would have to look at the data table.

Document: P14\_GMT20220124.odt Tags: Navigation complex

MODERATOR: It is just more or less.

Document: P14\_GMT20220124.odt Tags: Moderator comment

Again you are kind of assuming that you can compare them side by side it was hard to remember what the other query was.

Document: P14\_GMT20220124.odt Tags: Navigation complex

Let's look at rain and if I change this, does this automatically do it?

Document: P14\_GMT20220124.odt Tags: Navigation complex

MODERATOR: At the very top there are Q1 and Q2

Document: P14\_GMT20220124.odt Tags: Moderator comment

I see okay, they are easy to compare there, I did not see Q1 and Q2. You can even call that query 1 and query 2 to make it more obvious.

Document: P14\_GMT20220124.odt Tags: Wording not clear, Navigation complex

It is very clear. With a few tweaks it is possible to explore the data quite nicely.

Document: P14\_GMT20220124.odt Tags: Overall positive experience

It always comes back to what the question is and for me it is not very relevant in vasculitis but it might be relevant to another environmental researcher.

Document: P14\_GMT20220124.odt Tags: Export (meta)data useful

Export the metadata and data, there it exports.

Document: P14\_GMT20220124.odt Tags: Simple export

The FAIR data is saved as an XML file, which probably would be more accessible if you had a couple of options there like CSV.

I will try anyway to see what comes out of this, yes it is an RDF file. So I presume I can open that with the browser.

Okay, so I can explore that now.

Document: P14\_GMT20220124.odt Tags: Export name and format unknown

MODERATOR: Could you please complete the other sub tasks before moving on to the next task?

Document: P14\_GMT20220124.odt Tags: Moderator comment

(The participant reads the sub tasks again)

Document: P14\_GMT20220124.odt Tags: Task sequence complex

This comes down as a file named data, which would be better if it is called data provenance as default even though I can change it.

Document: P14\_GMT20220124.odt Tags: Export name and format unknown

That is very straightforward.

Document: P14\_GMT20220124.odt Tags: Simple export

The data table itself. The CSV file which is again easily explorable which is easy to incorporate into an R script or something like that.

Document: P14\_GMT20220124.odt Tags: Simple export, Export (meta)data useful

This is saved as a zip file, good. That is easy.

Document: P14\_GMT20220124.odt Tags: Simple export

Yes, it would with the caveats I have described before, which is that a researcher will want to define their research question. They won't be interested in just exploring data. They want to be able to define the study and the events.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion, Export (meta)data useful

Presumably, from the location there should be an option to select all locations. In my case, I would like to get the data for all Ireland.

Document: P14\_GMT20220124.odt Tags: Additional feature suggestion

I think there are a few little things that can be tweaked and a few little bugs here and there but overall it is a very intuitive and easy to follow interface.

Document: P14\_GMT20220124.odt Tags: Overall positive experience

So it is very nice, well done.

Document: P14\_GMT20220124.odt Tags: Overall positive experience

The visualizations are nice and the data tables are very clear.

Document: P14\_GMT20220124.odt Tags: Export (meta)data useful, Visualization useful

You can also explore the data provenance very nicely.

Document: P14\_GMT20220124.odt Tags: Metadata content clear and useful

So that is a highly FAIR interface.

Document: P14\_GMT20220124.odt Tags: Overall positive experience

(The participant reads the sub tasks again)

Document: P15\_GMT20211201.odt Tags: Task sequence complex

I don't know what RDF and SPARQL are.

Document: P15\_GMT20211201.odt Tags: Wording not clear

Yes, I think so. It is well explained and I think the information is enough as an introduction.

Document: P15\_GMT20211201.odt Tags: Home tab clear

So I go to the email and my credentials are these ones, which I will copy to the SERDIF dashboard ... Okay.

Document: P15\_GMT20211201.odt Tags: Log in easy

The event ... wait a second ... Okay ... I suppose I have to select an option for all of these query inputs because I was not sure.

Document: P15\_GMT20211201.odt Tags: Event concept and approach not clear

Okay, so for instance I will choose high probability

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

the location of county Cork.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

The time window length in days ... I am not quite sure but I will choose seven days, which is a week.

Document: P15\_GMT20211201.odt Tags: Query process complex

The time window lag, I will choose one day as it is a part of the week.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

Then, I click to check data availability ... It seems that is working as there are 19 events, which is good I think, and envo data sets.

Document: P15\_GMT20211201.odt Tags: Query process clear

Is the 3.a sub task done or I have to continue here until the end? Okay, I need to continue.

Document: P15\_GMT20211201.odt Tags: Task sequence complex

The temporal unit is day

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

the aggregation method the mean for instance, I am not sure about that but I hope it works.

Document: P15\_GMT20211201.odt Tags: Query process complex

Yes, they are clear and it is easy to compute the data query.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

No, I think it is enough to complete a query to get the data that you would like for your work.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

What do you mean with three buttons below I can see five?

Document: P15\_GMT20211201.odt Tags: Task sequence complex

MODERATOR: The buttons in task 4.c not in the dashboard, apologies.

Document: P15\_GMT20211201.odt Tags: Moderator comment

Okay, I was looking for three buttons but I was seeing more than three and I could not understand that. Okay, thank you.

Document: P15\_GMT20211201.odt Tags: Task sequence complex, Navigation complex

I can see a table with 19 events, which correspond to what I queried before. I think this is okay.

Document: P15\_GMT20211201.odt Tags: Metadata content clear and useful

Which seems to be the same as before ...

Document: P15\_GMT20211201.odt Tags: Navigation complex

Ah, so you have the code here. So that's the difference between the first and second button.

Document: P15\_GMT20211201.odt Tags: Metadata content not useful or confusing

This includes the code and something else that I don't know but seems like extra information about the provenance of the data.

Document: P15\_GMT20211201.odt Tags: Metadata content not useful or confusing

So, I have 19 events and the data sets are the other part ... Okay, they are here ... So 9 data sets for each event. So that is correct.

Document: P15\_GMT20211201.odt Tags: Metadata content clear and useful

Okay, so the data use comment I think it is this part ... Oh no no ...

Document: P15\_GMT20211201.odt Tags: Search information process complex, Navigation complex

MODERATOR: Do you want to read the introductory paragraph at the top of the tab again?

Document: P15\_GMT20211201.odt Tags: Moderator comment

This part, oh right the Control+F part, thanks.

Document: P15\_GMT20211201.odt Tags: Search information process complex

I can see the identification risk here but I don't know if I had to read it aloud. Okay, the task is just to identify it.

Document: P15\_GMT20211201.odt Tags: Search feature simple, Task sequence complex

I will check the license ... The license is the creative commons.

Document: P15\_GMT20211201.odt Tags: Search feature simple

The data use ... I think I was typing it wrong so I could not find it but I have it here now on the screen.

Document: P15\_GMT20211201.odt Tags: Search feature simple

I was able to identify these three parts in the metadata.

Document: P15\_GMT20211201.odt Tags: Search feature simple

What does hovering mean?

Document: P15\_GMT20211201.odt Tags: Wording not clear

MODERATOR: To go over something with your mouse.

Document: P15\_GMT20211201.odt Tags: Moderator comment

The date has no tooltip description.

Document: P15\_GMT20211201.odt Tags: Additional explanation

(The participant reads the column heading descriptions)

Document: P15\_GMT20211201.odt Tags: Export (meta)data useful

This is my field more than yours, I think. (joking) So, I understand all the parameters shown.

Document: P15\_GMT20211201.odt Tags: Export (meta)data useful

The cell background colors ... Mmmm ... I see the unmarked cells, the blue ones which are lower than the permitted limit and the red cells that I understand are above the higher limit, the upper limit for this.

The red ones are higher than two times the standard deviation of the distribution, more or less.

Document: P15\_GMT20211201.odt Tags: Data table features useful

For instance I checked the O3 and rain and they disappeared from the table.

I am not sure about the eye button. It also works with the eye.

Document: P15\_GMT20211201.odt Tags: Navigation complex

(The participant reads the sub task again)

Document: P15\_GMT20211201.odt Tags: Task sequence complex

I am not sure about the meaning of the events ... Are they related to meteorological events or are there other information that is not shown which is related to the events? For example, I can think of health information ... Okay.

Document: P15\_GMT20211201.odt Tags: Event concept and approach not clear

These are the basic environmental parameters including the main pollutants and I think they are enough and consistent enough to link with health data. We used to use this data to link with health data, so I think the parameters chosen are correct in this sense.

Document: P15\_GMT20211201.odt Tags: Export (meta)data useful

Let me check it ... Well here the NOx are not shown, which could be a problem in the data collection, but I think tit would help to complete the amount of environmental data that is necessary to link with health data.

Document: P15\_GMT20211201.odt Tags: Additional feature suggestion

Metals ... Sometimes the metal content, not one metal in particular for example the Nickel. Sometimes the metals are shown as a cluster, which is interesting to show.

Document: P15\_GMT20211201.odt Tags: Additional feature suggestion

I will check PM10 but I am not sure if it is working.

Document: P15\_GMT20211201.odt Tags: System issue

MODERATOR: Could you please try another variable while I check why it is not working?

Document: P15\_GMT20211201.odt Tags: Moderator comment

I will try rain or temperature but I cannot see anything.

MODERATOR: Could you scroll down a bit?

Oh yeah, it is here.

Document: P15\_GMT20211201.odt Tags: Plot design complex

MODERATOR: Could you scroll down a bit?

Document: P15\_GMT20211201.odt Tags: Moderator comment

(The participant reads the sub task again)

Document: P15\_GMT20211201.odt Tags: Task sequence complex

The graph shows me ... I am not sure about the relative time from the event.

The graphic shows me the difference ... Okay, this shows the days before the events.

Document: P15\_GMT20211201.odt Tags: Plot design complex

It shows me the air temperature for a period before the event which includes 8 days before the event.

The dark blue colors from the scale represent the high temperatures and the low temperatures are in a pale color.

Document: P15\_GMT20211201.odt Tags: Visualization useful

I will check another one, yeah. Let's see rain for instance. So it shows me the same graphic type which is the rain amount detected days before the event for each of the selected events. The plot has the same color scale for high and low values. I can understand the graphic very well.

Document: P15\_GMT20211201.odt Tags: Visualization useful

I think there are a lot of graphics shown here and I think it is enough information. I would not add any other parameters as I commented previously.

Document: P15\_GMT20211201.odt Tags: Visualization useful

MODERATOR: Could you please make sure you have explored all the plots?

Document: P15\_GMT20211201.odt Tags: Moderator comment

Oh yeah I need to check the box and polar plots sorry.

Document: P15\_GMT20211201.odt Tags: Task sequence complex

This graph shows the amplitude of temperature for each event including the variance for each event.

You can check if the event's temperatures overlay and for rain I guess it would be the same. Yeah, there are events with more dispersed data and other ones with less dispersion.

Document: P15\_GMT20211201.odt Tags: Visualization useful

I will check the same two parameters ... Oh this is nice.

Document: P15\_GMT20211201.odt Tags: Visualization useful

Here we have a graph that shows the mean temperature of each event but it adds a new parameter which is the wind speed and direction recorded for each event. It gives us a new perspective with more information, not only the temperature but it adds the wind speed and direction. The data is displayed in a wind rose, which is useful to detect the aggregated events, which is useful to detect if the events had a similar wind direction or speed. The scale of colors is used to represent the temperature of the air masses.

Document: P15\_GMT20211201.odt Tags: Visualization useful

For rain I think it is the same, well the dispersion of the rain is different than temperature but it aggregates the events in one mask of points. There is only an isolated point which is a bit different from the rest.

Document: P15\_GMT20211201.odt Tags: Visualization useful

Okay, the location ... I was looking for a big city like Dublin so it would have more data.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

Oh no sorry, this could be the same.

Okay, submit.

Okay, good.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful, Query process clear

(The participant reads the sub task again two times)

Document: P15\_GMT20211201.odt Tags: Task sequence complex

Here I have ... Okay, Q1 has more data sets used ... Okay the event that uses more data sets is this one in Q2.

Document: P15\_GMT20211201.odt Tags: Metadata content clear and useful

I have to check in the full metadata exploration for Q1 and ... Q1 has this license here and ... If we compare it with Q2 ... It seems to be the same data set license.

Document: P15\_GMT20211201.odt Tags: Search feature simple

Approximately the Q1 has more extreme events (referring to the data values).

Document: P15\_GMT20211201.odt Tags: Data table features useful

One only, okay, so let's choose the heat map to compare the temperature from Q1 and Q2.

Okay, here. I will do the same for the Q2 heat map.

Oh, okay. The main difference is that there is a single event and in Q1 there are 19 events, which gives a different aspect to the graph. The rest of the representation is the same showing the mean temperature for the days before the case of study events. So the main difference is the number of events.

Document: P15\_GMT20211201.odt Tags: Visualization useful

Do you mean only with the plots or in general?

Document: P15\_GMT20211201.odt Tags: Task sequence complex

For me there are enough features to comprehend this data set.

Document: P15\_GMT20211201.odt Tags: Data table features useful, Export (meta)data useful

Download Q1 ... Okay, so FAIR metadata export and FAIR data export.

I have downloaded the data without a problem.

Document: P15\_GMT20211201.odt Tags: Simple export

Okay, export the data provenance and for the data table ... Where is that?

Document: P15\_GMT20211201.odt Tags: Navigation complex

Data table okay I have this ... Oh here data table export.

I have done that.

Document: P15\_GMT20211201.odt Tags: Simple export

Download all, okay without a problem.

Document: P15\_GMT20211201.odt Tags: Simple export

Do you mean in general?

Document: P15\_GMT20211201.odt Tags: Task sequence complex

MODERATOR: Yes, in general.

Document: P15\_GMT20211201.odt Tags: Moderator comment

Well, I think it is a good set of data, which is easy to work with. You can download the data so you can work with this data in different formats and programming languages like R or Python, which is very important.

Document: P15\_GMT20211201.odt Tags: Export (meta)data useful

I think the general display of the data is very easy to understand. For me it is a good way to represent and show the data.

Document: P15\_GMT20211201.odt Tags: Visualization useful, Overall positive experience

I think it is very reusable for this kind of link between health and environmental data because you can select an individual event and it represents it in a good way ... How can I say that? ... When you have a very big data set, it is interesting to just pick one event that you want.

Document: P15\_GMT20211201.odt Tags: Export (meta)data useful, Overall positive experience

The app environment is very user friendly.

Document: P15\_GMT20211201.odt Tags: Overall positive experience

I am not a normal user of these environments but the text in the dashboard gave me more information, which was useful for me to complete the tasks. If I did not have these guides, it would have been more difficult for me to follow the instructions.

Document: P15\_GMT20211201.odt Tags: Query inputs and elements useful

I think the second time that I have selected the events and data it was easier for me. You get used to the functioning of the application very easily. You get more confidence with the data distributions and buttons which is useful for me.

Document: P15\_GMT20211201.odt Tags: Overall positive experience

It is quite clear in terms of its purpose.

Document: P16\_GMT20211215.odt Tags: Home tab clear

The language is very clear

Document: P16\_GMT20211215.odt Tags: Home tab clear

but I suppose there are words like SPARQL, that I have no idea what they mean. Is it a necessary piece of information to have there really?

Document: P16\_GMT20211215.odt Tags: Wording not clear

I know about RDF and FAIR but I am curious about the level of knowledge of these acronyms and it confuses and overwhelms a person at this stage.

Those aspects would be the factors.

Document: P16\_GMT20211215.odt Tags: Wording not clear

I don't think I would add anything really. There is no need to add anything necessarily.

Document: P16\_GMT20211215.odt Tags: Home tab clear

I am looking at the figure to see how it adds value in many ways. There is a lot going on in this figure ... Mmmm ... Again there is R2RML that appears in the figure that has no link that I can see with the text, but RDF, SPARQL and FAIR are featured in the text.

Document: P16\_GMT20211215.odt Tags: Wording not clear

So that would be kind of the considerations for me, this figure needs to be as simple as possible. I think it is overly complex for its needs.

Document: P16\_GMT20211215.odt Tags: Wording not clear

I haven't clicked on methodology or knowledge graph but I would probably do so. I must admit that I probably do not tend to.

Document: P16\_GMT20211215.odt Tags: Additional source exploration

Okay, it is taking me to a GitHub with lots of information. It is not necessarily a text page that I am going to, which explains the methodology, but it is the actual methodology.

Document: P16\_GMT20211215.odt Tags: Additional source exploration

The knowledge graph is something I have to log in to. SO I can't actually go back from the knowledge graph so let me just close that and reopen the link from my email.

Document: P16\_GMT20211215.odt Tags: Additional source exploration

Okay, log in. We are in.

Document: P16\_GMT20211215.odt Tags: Log in easy

Oh, this is quite funny. I logged in and I almost did not realize that I logged in because only the menu on the left had changed. It is almost a surprise. There is nothing to be said other than that but it is a bit of a surprise. I would almost expect something different on the main page to indicate I came through that.

I don't think it is a huge deal though.

Document: P16\_GMT20211215.odt Tags: Log in easy

Events ... possible, definite, no and high probability ... I am not quite sure of what these mean or what they are in relation to.

Document: P16\_GMT20211215.odt Tags: Event concept and approach not clear

I would need clarity of what this means at this point but maybe it will be clearer in the final version.

Document: P16\_GMT20211215.odt Tags: Event concept and approach not clear

I will pick Dublin because I want to focus on understanding the problem.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

I can see I can pick multiple counties without any problem.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

The time window length in terms of days ... Okay so 14 and 7 days for the length and lag.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

The events, I suppose I should click on the tab. Event is classified by an agent. Maybe I missed that one before. I suppose an agent, I am not quite sure what an agent refers to in terms of user language for me. So that's one thing I would need clarification on.

Document: P16\_GMT20211215.odt Tags: Event concept and approach not clear, Additional explanation

I have to double click twice to close it. If I click somewhere else it does not seem to automatically close it. That is something I would avoid as usually when I click somewhere else this kind of information tabs disappear.

Document: P16\_GMT20211215.odt Tags: Navigation complex

So I am choosing a time window for 14 days.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

Temporal unit the text is a little bit small here for the information, different style to the previous information.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Envo data sets, I guess that might be environmental data sets but not sure.

Document: P16\_GMT20211215.odt Tags: Wording not clear

Okay, aggregation methods spatial and temporal aggregation.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

I am going to check for data availability. Okay so no events or data available.

The question is, why not? I will try different inputs.

Document: P16\_GMT20211215.odt Tags: Query process clear

Possible events as a larger data set ... Mmmm ... I am going to try some other counties to add variety and understand why that data set was limited.

I am going to reduce this to see if it was a factor. I am changing a lot of variables to see if it would work this time.

Document: P16\_GMT20211215.odt Tags: Query process complex

21 events and envo ... I am not sure what envo means.

Document: P16\_GMT20211215.odt Tags: Wording not clear

MODERATOR: Would you mind reducing the number of county locations due to a memory exceeding problem that the dashboard currently has?

Document: P16\_GMT20211215.odt Tags: Moderator comment

No problem, I will remove that before submitting.

Document: P16\_GMT20211215.odt Tags: System issue

Yeah, and again I am not sure what envo means?

Document: P16\_GMT20211215.odt Tags: Wording not clear

MODERATOR: Environmental.

Document: P16\_GMT20211215.odt Tags: Moderator comment

I don't think it is a word that can be shortened without leading to a person being unsure.

Document: P16\_GMT20211215.odt Tags: Wording not clear

Okay, so. I don't know if Dublin was still available. So Dublin clearly does not have data.

Document: P16\_GMT20211215.odt Tags: Query process complex

The question is then when I select possible, why is Dublin still an option if it does not have any data? So if anything, if there is a way to filter based on the previous classification that would be appreciated. Otherwise this iteration might be kind of more complicated.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

So Mayo, 10 Events 6 data sets. Is that too much? I think I will try that.

Document: P16\_GMT20211215.odt Tags: Query process clear

MODERATOR: Yes, this should be fine.

Document: P16\_GMT20211215.odt Tags: Moderator comment

So 10 events 6 environmental data sets, at this point I don't know what the environmental data sets are. This leads me to want to select those. I want a bit of control over what environmental data sets I am putting together. Have some logic. I mean that I can see from the picture that temperature, rainfall and pollution might be there but I am not 100% in control of what I can choose. Especially, I know my location ... And there is a link to the event. The event is obviously a health event.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion, Metadata content not useful or confusing

Okay, temporal unit again, temporal unit for me is important. I need to understand the environmental factors to be able to decide on that. Certain environmental factors only have data in hourly or daily resolution so it varies. I am going to select daily as a relatively good metric but hourly concentrations or peak concentrations might be the factor that dominates.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful, Metadata content not useful or confusing

Then, max because I am choosing daily. Actually, I would say daily rather than day. No actually no. But maximum daily. If I was doing hourly I would choose mean hourly and that would capture it so.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

I have been using spatiotemporal as terminology if you are talking about space and time in environmental data. I noticed in the tabs that you are using both terms and I would try to harmonize the language across the dashboard, just say spatiotemporal data.

Document: P16\_GMT20211215.odt Tags: Wording not clear

I have to admit that provenance is not a term that I use too often. So I wonder if there is a more appropriate word or you need to explain it. the same for lineage. Those terms I would need to have an understanding. Maybe when I click on the tabs I will have some text that explains them.

Document: P16\_GMT20211215.odt Tags: Wording not clear, Additional explanation

I don't understand the FAIR metadata for publishing either.

Document: P16\_GMT20211215.odt Tags: Wording not clear

Control+F, I am relying on something I would use in word or excel.

Document: P16\_GMT20211215.odt Tags: Search information process complex

I am not sure of what some of these mean but that is for another time (referring to the Control+F suggestions).

Document: P16\_GMT20211215.odt Tags: Wording not clear

Three buttons below? For me there are five buttons below.

Document: P16\_GMT20211215.odt Tags: Task sequence complex, Navigation complex

MODERATOR: Apologies, I refer to the three buttons in the tasks not in the dashboard.

Document: P16\_GMT20211215.odt Tags: Moderator comment

So as step B I am not clicking them just exploring ...

Document: P16\_GMT20211215.odt Tags: Task sequence complex

MODERATOR: You can do them together.

Document: P16\_GMT20211215.odt Tags: Moderator comment

(The participants reads the column headings of the data provenance table) I see six data sets used which represent the environmental data sets.

Document: P16\_GMT20211215.odt Tags: Metadata content clear and useful

I don't see this information really as exporting it, how useful it is to me. I am not quite sure if this is relevant there, I am not sure if this is a relevant piece of information in reference to the events. Other than the name of the event and data sets used and the timestamp as well. So I am not sure if I would want to export it.

Document: P16\_GMT20211215.odt Tags: Metadata content not useful or confusing

I do not see any difference with data provenance.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Then, let me click on data lineage again ... Oh way hold on a second something is going on here .. So it is opening multiple tabs here, it is an open-close. It does not clearly indicate that is not an open-close as it opens below when I click.

Document: P16\_GMT20211215.odt Tags: Navigation complex

There is quite a lot of staff going on here, which I am not sure about.

Document: P16\_GMT20211215.odt Tags: Metadata content not useful or confusing

That is the code, okay. Again I don't really understand this I would need an explanation of this tabs, maybe when I hover over them.

Document: P16\_GMT20211215.odt Tags: Additional explanation, Metadata content not useful or confusing

The three buttons I am not sure why I would export them. I guess if I want to run them in some sort of system like Matlab or something perhaps. I am not sure as I haven't notice any reference to Matlab so I don't know.

Document: P16\_GMT20211215.odt Tags: Metadata content not useful or confusing

(The participant reads the sub task again)

Document: P16\_GMT20211215.odt Tags: Task sequence complex

So there are 6 data sets used. I assume they are the same data sets for each event based on the name coding.

Document: P16\_GMT20211215.odt Tags: Metadata content clear and useful

So I suppose I can use the Control+F in theory.

I just tested if it is case sensitive, which is not.

I can see it appearing here.

Comment ... Not sure ... Not so clear on data set use.

Document: P16\_GMT20211215.odt Tags: Search information process complex

Let's try identification risk, there seems to be some text that explains that

Document: P16\_GMT20211215.odt Tags: Search feature simple

which I did not see so clearly with data use.

Document: P16\_GMT20211215.odt Tags: Search information process complex

Here there seems to be an explanation about the identification risk.

Document: P16\_GMT20211215.odt Tags: Search feature simple

Okay, so license ... Some comments here ...I think that's it. So this looks like the comment after it.

Document: P16\_GMT20211215.odt Tags: Search information process complex

I am going to search for the comment ... I don't know. I am not completely clear what is going on.

Document: P16\_GMT20211215.odt Tags: Search information process complex

MODERATOR: It is fine.

Document: P16\_GMT20211215.odt Tags: Moderator comment

So I assume it is not this table here. Oh I have been asked to open a file. So maybe it is actually.

Document: P16\_GMT20211215.odt Tags: Task sequence complex, Navigation complex

(The participant reads the task again)

Document: P16\_GMT20211215.odt Tags: Task sequence complex

There are four tabs, one with a data table.

Document: P16\_GMT20211215.odt Tags: Navigation complex

So here I can see the environmental factors, there are a lot of them.

Document: P16\_GMT20211215.odt Tags: Export (meta)data useful

I am not sure about these last columns ...

Document: P16\_GMT20211215.odt Tags: Wording not clear

Oh, okay. So when you hover there is an explanation which is good.

Document: P16\_GMT20211215.odt Tags: Export (meta)data useful

Okay, so the order of these variables is something I would be curious about. Some do not go below to where they are, so they are not in the group they should be into.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Hovering provides clarity but the short text is pretty standard in general.

Document: P16\_GMT20211215.odt Tags: Export (meta)data useful

Sea level pressure is more about pressure than sea level, so the M there does not indicate pressure. Actually pressure is the actual measure.

Document: P16\_GMT20211215.odt Tags: Wording not clear

I would change the wet bulb to temperature as well but it is generally okay.

Document: P16\_GMT20211215.odt Tags: Wording not clear

So there is some text here which I would have normally skipped or gotten past.

Document: P16\_GMT20211215.odt Tags: Wording not clear

(The participant read the data background paragraph) I would almost have guessed that without reading that.

It shows a variety of highs and lows.

Document: P16\_GMT20211215.odt Tags: Data table features useful

I noticed for the first time that there are 5 pages. One of the things I fear is that the limitation of 20 rows to avoid scrolling too much, I am surprised you won't have the 100 rows in the same page. If you are allowing it to do this, you should be allowed to do more than 20. I don't see the purpose of this. There is no point to have one of these scrolls within a page.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion, Navigation complex

Ah yes, I hadn't really noticed that before it was mentioned in the tasks, which isn't necessarily a good sign of being noted.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Oh, yeah ... (The participant hides some columns) I noticed that I can hide them but, how do I get them back?

Let me see ... Toggle columns. So there are two ways of hiding them and one way of bringing them back.

Document: P16\_GMT20211215.odt Tags: Navigation complex

It sounds quite silly but instead of toggle, hide/unhide columns might be a better way of referring to that for the less literate. Hide/unhide columns would make it more universal, much more clear.

Document: P16\_GMT20211215.odt Tags: Wording not clear

I don't think the eye ... I suppose it is a way of doing it faster, but if nobody notices it, it is just a bonus of being there until you have it explained somewhere. If it was explained my apologies as I did not see it.

Document: P16\_GMT20211215.odt Tags: Navigation complex

The data table ... All I really learned from that is what information is in the table.

Document: P16\_GMT20211215.odt Tags: Export (meta)data useful

I guess I can see the Z-scores but it hasn't led me to any form of conclusions, so I wouldn't be hugely benefited from the distribution, the highs and lows, but it provides more understanding than a plain table.

Document: P16\_GMT20211215.odt Tags: Data table features useful

There is some benefit but I am curious now about the subsequent plots as more indicative pieces of information.

Document: P16\_GMT20211215.odt Tags: Visualization useful

I am not seeing the value of this so much other than understanding the range of the information. The question is then: is this type of table the best way to present the data or this tab should be just a summary of metadata? That might be just as simple.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

I think an explanation, questioning this data table. Just making sure everyone understands what is in the metadata file in terms of different criteria.

Document: P16\_GMT20211215.odt Tags: Additional explanation

Yeah the high lows. If this did not have the -2 and +2 this would be just a table. So I am more curious to get to the next plot.

Document: P16\_GMT20211215.odt Tags: Data table features useful

There is no need in my opinion to have the variables as short text because you have a long white tab here. There is no need to shorthand.

Document: P16\_GMT20211215.odt Tags: Wording not clear, Plot design complex

I'm going to go in the wind direction as there was not much data for the pollutants.

Maybe for this event wind direction is a factor that would be interesting. I am changing the variable to O3.

Document: P16\_GMT20211215.odt Tags: Visualization useful

There are four events where O3 data was available. I thought there were ten events so only four of them are available.

I find that quite strange really. It is taking off the six events where O3 data was not available, but it does not come across here.

Document: P16\_GMT20211215.odt Tags: Plot design complex

That is straightforward in terms of understanding the events.

The horizontal axis represents the number of days considering a lag in there as well. I am not sure how the lag affects it.

Concentrations up to a 200.

The heat map gives me something ... I am not sure how useful it is yet.

Document: P16\_GMT20211215.odt Tags: Visualization useful

I am going for O3 too to add on the clarity for the variable.

So here in the box plot I see all 10 events, which I appreciate more. I would appreciate that in the heat map as it provides a homogeneous, being able to click on and see.

Document: P16\_GMT20211215.odt Tags: Visualization useful

I include all of the events and if they are blank, leave them blank.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

The 200 here appears again, there must be a reading that was extremely high for one of these days. There we go, there is a max that seems like a real outlier, bizarrely outlier.

Document: P16\_GMT20211215.odt Tags: Visualization useful

On my screen the blue is very hard to see until I hover over it but the hovering is quite good.

Document: P16\_GMT20211215.odt Tags: Visualization useful

The hovering over does add some interesting value to that and the dots are spread across. Interestingly, the dots do not seem to align up the center of the box plot, they seem to lie below. It is definitely more informative than the previous one.

Document: P16\_GMT20211215.odt Tags: Visualization useful

A little bit slower to load ... Okay.

Document: P16\_GMT20211215.odt Tags: System issue

This is a polar plot for O3. This I am not clear ... O3 concentration is linked to wind speed and direction., but I don't know the spatial or temporal units.

Is it all 10 days of my data? I don't have any control over the day. Is it an event? There is no real breakdown or further breakdown here, which I think is necessary.

This just provides me with a more holistic overview of Mayo for the whole thing.

Document: P16\_GMT20211215.odt Tags: Plot design complex

Yeah, the box plot and the heat map provided more value than this one.

Document: P16\_GMT20211215.odt Tags: Plot design complex

I am going to check rainfall. Yeah, not quite sure.

Document: P16\_GMT20211215.odt Tags: Plot design complex

There are no metrics here just to be clear. It could add more value if there were more drop downs or sub menus to break it down further.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

In terms of visualizations, there are improvements that can be made for them by having harmonized 10 events or clarity whether you can separate the events.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

I am going back to the heat map and ideally I would like to choose a pollutant which I don't think is available.

I will choose pressure, oh sorry, relative humidity.

I am not sure if it provides ...

Document: P16\_GMT20211215.odt Tags: Task sequence complex

MODERATOR: Task 7 refers to submitting another query on the left panel.

Document: P16\_GMT20211215.odt Tags: Moderator comment

So query, I am going to choose high probability and another location.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

I will check availability ... 12 events and 3 data sets, fingers crossed.

Document: P16\_GMT20211215.odt Tags: Query process clear

I will keep it daily ... Actually, hold on. Let's change to possible events to see if we could go hourly.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

MODERATOR: If you would like hourly, I would advise Wexford as it has less data and events.

Document: P16\_GMT20211215.odt Tags: Moderator comment

Okay, so Wexford. Let's check availability, 5 and 3.

Document: P16\_GMT20211215.odt Tags: Query process clear

I am doing the mean.

Document: P16\_GMT20211215.odt Tags: Query inputs and elements useful

One thing I forgot in my assessment is that I was dealing with the maximum pressure and the maximum temperatures ... Rather than the mean, there was no flexibility to have mean and max there in the table.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

MODERATOR: We got almost to the max memory permitted but I think it will work, please proceed to the next tasks.

Document: P16\_GMT20211215.odt Tags: Moderator comment

So we have Q1 with daily max values and Q2 with hourly mean values.

Document: P16\_GMT20211215.odt Tags: Query process clear

Mean as opposed to ... Mmmm ... Wind direction is not something that can be understood in a normal distribution just thinking about it. Again, this is another reason to understand if the normal distribution is not useful for some of these variables. I did not think about that before.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

Know is an unusual metric you might want to convert that to m/s or km/h. I don't know why I'm only noticing this now.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

The heat map is not working for the pollutants. I will try another variable.

Document: P16\_GMT20211215.odt Tags: System issue

There is a lot more going on here as I have hourly values for 14 days.

Document: P16\_GMT20211215.odt Tags: Visualization useful

I will go back to the questions before I continue here.

Document: P16\_GMT20211215.odt Tags: Task sequence complex

Q2 has more data as I went for possible events with hourly resolution.

Document: P16\_GMT20211215.odt Tags: Task sequence complex

MODERATOR: It refers to data sets used.

Document: P16\_GMT20211215.odt Tags: Moderator comment

Ah, excuse me, data sets. There are much less data sets available for Q2.

Data sets as opposed to data, which is unfortunate in many ways. But the higher resolution adds value.

Document: P16\_GMT20211215.odt Tags: Metadata content clear and useful

I am not clear about what license refers to.

Document: P16\_GMT20211215.odt Tags: Wording not clear

Am I going back to data provenance? No, I am going to search in the data lineage.

I honestly don't know but I think I would have to search for the meaning of that.

Document: P16\_GMT20211215.odt Tags: Search information process complex

Okay, creative commons license.

Document: P16\_GMT20211215.odt Tags: Search feature simple

My question is, why do I need to know that as a user? More than anything ...

Document: P16\_GMT20211215.odt Tags: Metadata content not useful or confusing

In Q2 there are 70 pages, which I am not going to scroll across because that would take too much time. So I can give a guarantee of the information.

Document: P16\_GMT20211215.odt Tags: Navigation complex

But based on the common wind directions not a fair representation. With hourly data you would assume that you would have more extreme values. Oh no because it is the mean excuse me, no. So the max actually would have the higher values. The finer resolution would give more extreme values but the daily max would give higher values. Hold on, in the sense that the extreme values are considered out of the 95% of the distribution then Q2. Oh yeah, complicated.

Document: P16\_GMT20211215.odt Tags: Data table features useful

I will choose the box plot because I found it the most informative for this one and compare the temperature.

Document: P16\_GMT20211215.odt Tags: Visualization useful

I will go back to Q1 and also select the temperature for the box plot.

So 5 events in this one vs 10 in that other one.

Comparing is fine but I am clicking between tabs, which makes it a little bit tricky. One thing to notice is that the mean max is different so comparing tab to tab is not the most effective because the magnitude scales are different. Mentally I need to convert them to the same scale to effectively compare. There is much more variety so I don't know if it is a fair comparison.

Document: P16\_GMT20211215.odt Tags: Navigation complex

This is max daily Q1 so I would expect a higher maximum and more variety more distributed, while this one which is Q2 that represents the mean hourly data with more harmonized results. I think I changed too many variables to make a fair comparison.

Document: P16\_GMT20211215.odt Tags: Visualization useful

MODERATOR: Don't worry it is fine.

Document: P16\_GMT20211215.odt Tags: Moderator comment

Again, I changed a lot of variables but going back to the queries that stayed. I would probably not be as extreme in my changes. I made too many changes. Therefore, after an iteration I would have a better way of comparing.

Document: P16\_GMT20211215.odt Tags: Query process clear

However, things that have 0-20 and 0-25 scale would be good to convert it to the same scale to avoid any sort of misinterpretation.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

But this is an informative step in the analysis and you would be anyway downloading the data to assess it.

Document: P16\_GMT20211215.odt Tags: Export (meta)data useful

If you could have scales or be able to show the two side by side, split the screen or the window to allow that. That would be extremely helpful.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

Maybe rather than hovering over to get the x-value in the box plot, to have the value displayed somehow for the box plot metrics. The information when hovering over offers a great value but really I am just trying to get the actual number or get rid of the y-axis information in either or both.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

I go to the button FAIR metadata export.

Document: P16\_GMT20211215.odt Tags: Simple export

MODERATOR: No need to save it if you don't want to.

Document: P16\_GMT20211215.odt Tags: Moderator comment

Okay, and then I click on FAIR data export.

So by clicking on it, it automatically allows me to export.

Document: P16\_GMT20211215.odt Tags: Simple export

So I got a table ... What is the data table? Is that the data table?

Document: P16\_GMT20211215.odt Tags: Task sequence complex

MODERATOR: The actual data table that you see there.

Document: P16\_GMT20211215.odt Tags: Moderator comment

Ah this one, okay. I am just jumping down here.

Document: P16\_GMT20211215.odt Tags: Navigation complex

There is an export there.

It downloads the data as a CSV.

Document: P16\_GMT20211215.odt Tags: Simple export

Going back to the FAIR data, it is a .ttl file. So the FAIR metadata is a RDF file. So there is a wide range of formats that we are dealing with here.

Document: P16\_GMT20211215.odt Tags: Export name and format unknown

So one thing I would include here is that you can tick on the types of files that you want.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

What we are talking about here is that you have used the word download while previously the word export seems to be the common language on the other two tabs and buttons. I know they are the same thing but again, harmonized language and having to go to multiple locations ...

Document: P16\_GMT20211215.odt Tags: Wording not clear

Do you really need to go to different locations to download the files?

Maybe this tab is like downloadable information and all the buttons can be inside with selects.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

For me the two buttons for the FAIR (meta)data exports are two buttons that I would not have here. So the other three are fine but those too are just dropped here. If anything is just for export. I just don't feel that they fit.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Then, the data table and the data provenance ... In the data provenance table, only when I click there I can find the export button.

Document: P16\_GMT20211215.odt Tags: Navigation complex

Data table has it here. Harmonizing this information would be useful, maybe in this download area you could have a clear link to download information.

If you open a tab Q1 and Q2 inside the download area, where you can have an expandable/contracting menu for the queries to download the files. So I don't know.

Document: P16\_GMT20211215.odt Tags: Additional feature suggestion

Download query should download data tables because you are only downloading the data tables. Maybe it's my language.

Document: P16\_GMT20211215.odt Tags: Wording not clear

Obviously, reckoning the constraints is feasible for the present system, which allows/restricts some way of querying. I suppose that for me I was overly adventurous changing files and then, considering that I was going to compare Q1 and Q2. They were not as easily comparable as I would have liked.

Document: P16\_GMT20211215.odt Tags: Query process complex

The possible description again needs clarification for the events, I know that there is a button but again, I am not sure what an agent means.

From my perspective, definite, possible and high probability mean in a statistical sense so like you know greater than 90%. Again, maybe because I am coming at it not from a rare disease angle. I would like clarification on what that is rather than saying by an agent.

Document: P16\_GMT20211215.odt Tags: Event concept and approach not clear, Additional explanation

Trying to reduce the heading and subheading, narrowing it down. So yeah possible could be in brackets (50%) so when I am selecting these options I am clear. I can actually put in both, I did not realize this before. I don't know if there is anything else.

Document: P16\_GMT20211215.odt Tags: Event concept and approach not clear, Additional explanation

I suppose it is more about clarifying what an acronym is. From a user perspective, the thing that FAIR was in the front page and that it was open access, etcetera. I don't feel some of this information is necessary at that stage.

Document: P16\_GMT20211215.odt Tags: Wording not clear

The Control+F is a strange one for me, I don't know its value. I don't know why a user would go into this data lineage or the full metadata exploration to search for these items. I am not 100% sure why they would do that but that's me.

Document: P16\_GMT20211215.odt Tags: Search information process complex, Metadata content not useful or confusing

I essentially see information provided as enough.

Document: P17\_GMT20211216.odt Tags: Home tab clear

Maybe here in events with location and time I would put the word health. So that the user knows that despite the symbols indicating. It would be nice to see the word health.

Document: P17\_GMT20211216.odt Tags: Additional explanation

Okay so ...

Document: P17\_GMT20211216.odt Tags: Log in easy

MODERATOR: Would you mind speaking aloud while you are completing the tasks.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, I am now reading the definition of the events. There are four categories. Event classified as definite by an agent. Maybe this is not fully clear.

Document: P17\_GMT20211216.odt Tags: Event concept and approach not clear

High probability, possible ... Okay so, I think I understand these.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

I don't know why the event description does not go away.

Document: P17\_GMT20211216.odt Tags: Navigation complex

I understand these other classified events.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

Maybe it would be nice to have a map because not everyone knows where these counties are, for example me.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion, Query process complex

I am going to choose Cork, I know where Cork is.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

Time window length in days, maybe here you can put in an (i) info saying what this length refers to.

Document: P17\_GMT20211216.odt Tags: Additional explanation

I understand the lag well, as it is the lag between the environmental data and the outcome,

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

The time window length in days is not that clear.

Document: P17\_GMT20211216.odt Tags: Query process complex

I check data availability. I can see this says: 19 events and 9 environmental data sets.

Document: P17\_GMT20211216.odt Tags: Query process clear

Okay ... Temporal unit let's select days.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

Aggregation method, it is not that relevant but let's go with the mean

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

I can see that the SERDIF is computing and calculating what I have asked.

Document: P17\_GMT20211216.odt Tags: Query process clear

I don't know if the choices make sense or not.

Document: P17\_GMT20211216.odt Tags: Query process complex

Now, I see the query one linking the health event with environmental forcing through location and time.

Document: P17\_GMT20211216.odt Tags: Query process clear

Okay, if I go down I see that there are 4 tabs: data table, heat map, box plot and polar plot.

(The participant reads the data table background paragraph)

Document: P17\_GMT20211216.odt Tags: Task sequence complex

This data is not normally distributed so I don't know if there is any change to this non normality, but I will ask this afterwards.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion

Data normalized, I can see Z-scores.

There are some events classified.

Document: P17\_GMT20211216.odt Tags: Data table features useful

MODERATOR: Would you mind going back to the current task as this is task 5?

Document: P17\_GMT20211216.odt Tags: Moderator comment

I can see here that I can export the data provenance table as a CSV file.

Document: P17\_GMT20211216.odt Tags: Simple export

Okay, if I open the CSV file what I have is ... I did not understand this quite well but it could be the way my computer is opening the CSV file.

Document: P17\_GMT20211216.odt Tags: Wording not clear

I don't understand what data lineage is

Document: P17\_GMT20211216.odt Tags: Wording not clear, Metadata content not useful or confusing

but I think these are the events that the selector has selected.

Document: P17\_GMT20211216.odt Tags: Query process clear

Mmmm ... I see here there is the whole set of events, okay.

Document: P17\_GMT20211216.odt Tags: Metadata content not useful or confusing

Okay, let's see ... I don't know if I can say that ... Am I looking in the right place?

Document: P17\_GMT20211216.odt Tags: Navigation complex

MODERATOR: No, you should be in the data provenance window.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Oh, but I had clicked that before. Okay I see it now.

Document: P17\_GMT20211216.odt Tags: Navigation complex

Okay, so I see here listed the number of events right here. The number of events I understand. They are all listed here.

Document: P17\_GMT20211216.odt Tags: Metadata content clear and useful

(The participant reads the sub task again)

Document: P17\_GMT20211216.odt Tags: Task sequence complex

Okay, so this is the number of data sets listed here, so 9 data sets, and 9, and 9 ...

(The participant keeps counting the number of datasets for all the events)

Document: P17\_GMT20211216.odt Tags: Metadata content clear and useful

MODERATOR: Yeah it is fine.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Data use comment ... Mmmm ... No, I cannot.

Document: P17\_GMT20211216.odt Tags: Search information process complex

MODERATOR: Would you mind reading the introductory paragraph at the top of the tab?

Document: P17\_GMT20211216.odt Tags: Moderator comment

(The participant reads the introductory paragraph again) Oh I see, Control+F ... I had to look for ...

Is it this one?

Document: P17\_GMT20211216.odt Tags: Search information process complex

MODERATOR: Maybe use the arrows in the Control+F.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, so yeah I can see there are two.

Document: P17\_GMT20211216.odt Tags: Search information process complex

MODERATOR: Can you then identify the comment please?

Document: P17\_GMT20211216.odt Tags: Moderator comment

What do you mean by identify?

Document: P17\_GMT20211216.odt Tags: Wording not clear, Task sequence complex

MODERATOR: Just read the comment.

Document: P17\_GMT20211216.odt Tags: Moderator comment

(The participant reads the data use comment)

Document: P17\_GMT20211216.odt Tags: Search feature simple

(The participant skips the rest of the identifications)

Document: P17\_GMT20211216.odt Tags: Task sequence complex

(The participant reads the data table column heading descriptions)

Document: P17\_GMT20211216.odt Tags: Export (meta)data useful

The cell background colors ... Well, I understand that the red colors are outliers or are identified events. That's what I would assume.

Document: P17\_GMT20211216.odt Tags: Data table features useful

I don't know ... What is the eye icon? ... Okay yeah.

Document: P17\_GMT20211216.odt Tags: Navigation complex

Yes, I can do it.

Document: P17\_GMT20211216.odt Tags: Data table features useful

Well, yes I think this is useful for a user that is exposed to this kind of data set from let's say the health community, I think is quite intuitive so they can select the types of environmental covariates they are interested in.

Yeah I think it's useful.

Document: P17\_GMT20211216.odt Tags: Export (meta)data useful

Well, so I see here that there are some plots. Can I explore them now?

Document: P17\_GMT20211216.odt Tags: Task sequence complex

MODERATOR: This would be the next task.

Document: P17\_GMT20211216.odt Tags: Moderator comment

I was going to say some visual display, maybe of the time series. I don't know if that may be something useful or not, highlighting the showcased events that are the output of the query. Maybe it is not relevant.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion

MODERATOR: Hiding the metadata windows could be helpful in your case because I can see you have a small screen. They will hide if you click them again.

Document: P17\_GMT20211216.odt Tags: Moderator comment, Navigation complex

So here I can select PM10, but I cannot see the heat map. Should I submit that?

Document: P17\_GMT20211216.odt Tags: System issue

MODERATOR: Could you please try another variable while I check why it is not working?

Document: P17\_GMT20211216.odt Tags: Moderator comment

Sure, I will submit another query. Oh no, sorry! I refreshed the page by accident.

Document: P17\_GMT20211216.odt Tags: System issue

MODERATOR: No worries, you can submit another query and resume where you left.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, so I had: high probability, Cork, 7 and 4, daily and mean.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

MODERATOR: No worries, let's try another variable now.

Document: P17\_GMT20211216.odt Tags: Moderator comment

(The participant reads the task again)

Document: P17\_GMT20211216.odt Tags: Task sequence complex

Here there are the events and the relative time from the event, which I understand as the lagged time.

I am not fully convinced though.

Document: P17\_GMT20211216.odt Tags: Plot design complex

MODERATOR: This is a mistake from my side. The -1 should be -5 as I forgot to add the lag there.

Document: P17\_GMT20211216.odt Tags: Moderator comment

I see okay.

The relative time from the event that means the lag you selected in the query, right? I understand that it should be this.

Document: P17\_GMT20211216.odt Tags: Visualization useful

I don't understand what is the plot underneath the heat map though.

Document: P17\_GMT20211216.odt Tags: Plot design complex

Let's try another variable like rain. Okay, so millimeters I see ...

Document: P17\_GMT20211216.odt Tags: Visualization useful

I understand that the plot underneath is a subset of the database, that is what is denoted by these bars because the patterns tend to represent this.

Document: P17\_GMT20211216.odt Tags: Plot design complex

Relative time from the event locates the event in time and the rainfall.

Document: P17\_GMT20211216.odt Tags: Visualization useful

Just to suggest more work, I would like a way to assess the significance of the pattern. Not the significance of the individual values but of the overall selection of events like here with regard to the entire database. I don't know if this makes sense or not. Just to make an assessment of how relevant the query is but maybe this is not useful.

Document: P17\_GMT20211216.odt Tags: Additional source exploration

Okay, nice. I see this is very useful.

Document: P17\_GMT20211216.odt Tags: Visualization useful

The events might be in different seasons and we know that there is a seasonality in temperature. I don't know if you removed the seasonality in temperature but in any case you can showcase in a different color the different seasons. Maybe these values are anomalies and the seasonality is already removed. In this case, information of whether these are anomalies or individual values would be useful but they seem individual values.

I think yeah an indication of the season would be useful.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion

Nice plot. Okay that is a typical polar plot I think. This is relevant.

It displays all the events, which is nice.

Document: P17\_GMT20211216.odt Tags: Visualization useful

I don't know if it makes any sense from the query I made but yeah it can be informative.

Document: P17\_GMT20211216.odt Tags: Visualization useful

Well, you know I am keen on the time series data. I think that time series represents well whether your points are anomalies. If you could make a time series representation and highlight the anomalies with a different symbol or color. Then, we would have an image of how far these values are from normal range values, these outliers.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion

What I mean is that in the type of representation that you have with each section of the heat map and here at the bottom plot of the heat map you have the representation of the database. If I am correct, because I don't know if I am. Here you have a visual understanding of the comparison between the selection of events and the overall dimension or parameter space of this variable. You don't have this on the box plot because you cannot compare how relevant it is in the overall behavior and variability, and the same for the polar plot.

Document: P17\_GMT20211216.odt Tags: Additional feature suggestion

Okay, one more query. So should I submit another query with a different location?

Document: P17\_GMT20211216.odt Tags: Task sequence complex

MODERATOR: You can submit another query with different query input options, change whatever you like up to you.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, I want to compare other types of events leaving the rest of the inputs the same as before.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

Okay, so I need to check for the data availability again.

I see that this is the same number of events for this category. I don't know if they are the same events or if it is by chance.

Document: P17\_GMT20211216.odt Tags: Query process clear

MODERATOR: It is by chance.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Oh wow! I see that now the dashboard is computing.

I imagine I will have to displace, maybe not.

Document: P17\_GMT20211216.odt Tags: System issue

MODERATOR: The system has a problem regarding the memory being exceeded so I would like to ask you if you could refresh the dashboard and submit another two queries please.

Document: P17\_GMT20211216.odt Tags: Moderator comment

MODERATOR: Could we select a location with fewer events so we avoid the memory problem?

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, okay. So I will change the location for the second query. The former query is kept in the memory, okay.

Document: P17\_GMT20211216.odt Tags: Query inputs and elements useful

So, do I want to compare the two queries then?

Document: P17\_GMT20211216.odt Tags: Task sequence complex

MODERATOR: Yes, please proceed where you left in the task list.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, this is Q1 ... How can I switch between queries? ... Oh, yeah I understand.

Document: P17\_GMT20211216.odt Tags: Navigation complex

The first one uses more data sets.

MODERATOR: How did you know that?

I saw that there were more events in the first one.

MODERATOR: The task is asking about the number of data sets used not the events available.

Document: P17\_GMT20211216.odt Tags: Task sequence complex

MODERATOR: How did you know that?

Document: P17\_GMT20211216.odt Tags: Moderator comment

MODERATOR: The task is asking about the number of data sets used not the events available.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, so I imagine that in the data provenance table there are more data sets used for this query. Is that correct?

Document: P17\_GMT20211216.odt Tags: Metadata content clear and useful

I don't know ... Where is the license?

I don't know where the license is.

Document: P17\_GMT20211216.odt Tags: Search information process complex

MODERATOR: Maybe reading the introductory paragraph again would help.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, I thought that is what I did ... Okay, so Control+F and I look for a license.

Creative commons 4.0. So now I should go to the other query to check.

Document: P17\_GMT20211216.odt Tags: Simple export

But do I need to run the query again?

Document: P17\_GMT20211216.odt Tags: Navigation complex

MODERATOR: You can find that at the very top.

Document: P17\_GMT20211216.odt Tags: Moderator comment

Okay, I can see now ... Was that here? Oh no, it was in full metadata exploration.

Document: P17\_GMT20211216.odt Tags: Navigation complex

Yes, they have the same license.

Document: P17\_GMT20211216.odt Tags: Search feature simple

I can see that Q2 has more.

Document: P17\_GMT20211216.odt Tags: Data table features useful

I think no. What I understand for now is fine.

Document: P17\_GMT20211216.odt Tags: Data table features useful

Okay, I already did that as you can see from the downloads I have on the bottom of my browser.

Document: P17\_GMT20211216.odt Tags: Simple export, Task sequence complex

Okay, I can export the data table.

Document: P17\_GMT20211216.odt Tags: Simple export

So, is that it?

Document: P17\_GMT20211216.odt Tags: Task sequence complex

MODERATOR: Yeah, you are only missing 8.c.

Document: P17\_GMT20211216.odt Tags: Moderator comment

(The participant reads the sub task again)

Document: P17\_GMT20211216.odt Tags: Task sequence complex

Okay, so download all. I have here the zip file.

Document: P17\_GMT20211216.odt Tags: Simple export

Yes, I think they would.

Document: P17\_GMT20211216.odt Tags: Export (meta)data useful

I would say it is interesting.

Document: P17\_GMT20211216.odt Tags: Overall positive experience

I think it is nice that you have first the numerical description of the events you are searching and that then, you can display the data in different kinds of plots for a preliminary assessment of the relevance for the covariates for the specific events.

Document: P17\_GMT20211216.odt Tags: Overall positive experience

It is easy to operate and it is also easy to derive all the data sets to work more in depth afterwards.

Document: P17\_GMT20211216.odt Tags: Export (meta)data useful, Overall positive experience

I would say it is a very interesting tool, yeah.

Document: P17\_GMT20211216.odt Tags: Overall positive experience