(SLIDE 1)

Greetings Everyone. My name is Ian Hambleton – I'm Professor of Statistics at UWI. I'll be leading the project from UWI.

So I'm a University Researcher based in the Caribbean, and for over 25 years now I've been collecting and using Caribbean data to generate evidence for the region. And this project in many ways builds on the experiences I've had – and I hope and suspect those experiences will resonate with you all as well.

So I'm going to spend a few minutes expanding on WHAT we're planning, and WHY we think this project can be great for your organisations, and also for the region.

And these are the 4 themes running through our project.

- Infrastructure.
- Training.
- Analytics & Communication.
- And Collaborative working

(SLIDE 2)

Broadly, using existing UWI infrastructure as our foundation - we'll be building a set of resources to enable and champion Caribbean data sharing and data re-use

And I'm sticking to the same four themes:

- NEXT. We'll be setting up an online infrastructure to enable easier data sharing and re-use,
- NEXT. We'll be developing training programs and longer term mentoring to help increase your capacity in data handling and data analytics – and we want the training content to be guided by your needs.
- NEXT. We'll be performing centralized analytics along with a communication program to showcase existing data, champion data re-use, and to actively communicate the stories around the data we all produce
- NEXT. And through workshops and conferencing, we want to start building more collaborations that could lead to more data use across disciplines – more data linkages.

Of course, each of you will have particular needs, and one of our first jobs will be to reach out to better understand your work, your challenges, and what support would be most important for you.

(SLIDE 3)

Why do we think our project can be important for the region.

So the data industry is moving fast.

Most of us can remember when handling of data used to be a simple affair. Collect it, clean it, use it, store it.

The move towards handling 'big data' – large volumes and varieties of data, collected quickly – is one reason why data handling is becoming more complicated. Data collection is now mostly online. It is sometimes automated, and the associated data cleaning requires us to understand different subject matters and specialist software.

Just as important, data handling infrastructures must now conform to a fast-changing set of regulations, and these define for example minimum standards of security and participant confidentiality.

And countries across the Caribbean are updating their Data Protection Laws.

So one goal of our project is to help Caribbean data professionals keep up with best practice data handling.

(SLIDE 4)

I'm sure that each of you at some point will have received datasets that are less than perfect.

And I think it's not too dramatic to say that data prepared without best-practice data handling is regularly hard to use, time-consuming to correct, and sometimes unusable.

For data to be useful in the long term it must be accurate, and the data must be well documented and easily available.

Data that don't meet these basic criteria is stored at institutions across the world, creating what we might call 'data graveyards'. And these graveyards use resources that might be better spent elsewhere.

So another goal of our project is to reduce data wastage across the Caribbean. Again – this is all about making data accessible.

(SLIDE 5)

But data availability and accessibility in the Caribbean remains limited

There are lots of examples of this problem, and I want to jump on one example using the SDGs.

The World Bank tracks the availability of data to monitor the SDGs.

(SLIDE 6)

And here is a graphic of SDG data availability. Each row is a world subregion, and the dots represent the percentage of the SDGs indictors for which data was reported in 2019 as a dark dot and in 2010 as a light dot.

SDG data availability was at 34% for the Caribbean in 2019, down from 38% in 2010. This compares to 57 percent globally in 2019, up from 51% in 2010.

So the Caribbean is struggling a bit to report on the SDGs, but these scores do not necessarily mean that data does not exist.

So a goal of our project is to increase data accessibility across the Caribbean.

(SLIDE 7)

Across the Caribbean, international collaborations are common, and the data that come from these collaborations are often collected and stored using overseas data infrastructures. This is great if it creates high quality data and stores that data securely.

Often though, we hear of examples where data are hosted overseas without a full understanding of long-term data access or the implications for looking after that data. And we see this as a very important risk in the creation of long-term data resources for the region.

And so a goal of our project is to offer a local best-practice data infrastructure to give the region greater control of its data resources – we want to encourage data-sovereignty for the longer term.

Now, we will never force this infrastructure on those that don't need it – but it will be a robust regional option for those who are interested.

(SLIDE 8)

Sustainability is absolutely central to the success of this project. And so we have 3 years to enact a pragmatic sustainability plan that can work in the Caribbean.

This sustainability is anchored on our University commitment to the region and its commitment to open-data and open-science. We want this ecosystem to increase the impact and usefulness of our own research, and as data scientists we want to be more useful to the Caribbean community of data producers.

So I guess what I'm saying is that this ecosystem will be as useful to my University as I hope it will be for you, which is an important step on the road to sustainability. We are committed to making this a long-term solution.

(SLIDE 9)

Here are our current partners, and we are actively looking to expand the CaribData partnerships.

(SLIDE 10)

And so thanks very much from us. We're looking forward to working with you all, and we'll be reaching out to better understand your current setups, your challenges, your needs.