Slide 1	**Introduction**
	    Name. I'm leading the CaribData project.
	
	What we're developing and why
	    Spend a few minutes expanding on WHAT we're planning for CaribData, and WHY we think this project is needed and is good for the region.
	
	Tagline

Slide 2	We've been working with Caribbean data for over 2 decades and here are some of the reasons we've started our CaribData **initiative**:
	
	The data industry is moving fast.
	    Some of us can remember when handling of data used to be a simple affair. Collect it, clean it, use it, store it.
	
	Big data
	    *The move towards handling 'big data' – **so 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.*
	
	Regulations
	    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.
	
	GOAL:
	    *So one goal of our project is to help Caribbean data professionals keep up with **best practice data handling.***

Slide 3	Well documented / easily available
	    For data to be useful in the long term it must be accurate, and the data must be well documented and easily available. I'm sure that many of us at some point have received datasets that are less than perfect.
	
	Data graveyards
	    Data that don't meet these basic criteria are stored at institutions across the world – the idea of 'data graveyards'. And these graveyards use resources that would be better spent elsewhere. Important in our resource limited region.
	
	GOAL:
	    So another goal of our project is to reduce data wastage across the Caribbean. And this is all about making data accessible.

Slide 4	**Collaborations:** 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. Loss of data sovereignty 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. GOAL: 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. **Optional** 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.
Slides 5 and 6	Limited availability/accessibility Last and perhaps most important for our project, data availability and accessibility in the Caribbean remains limited. And our goal is to improve that situation. Example There are lots of examples of this problem – much of it anecdotal. [SLIDE] But here is an analysis from our recent CaribData special edition on <u>Caribbean Open Data</u> And the analysis uses World Bank Statistical Performance Indicators from 2022.
Slide 7	**Data Availability** So we're looking at data availability. First for the SDGs. Graph of the 7 world regions, and at the bottom we see that between 2018 and 2022, SDG data availability for the SIDS was lower than for any other world region, roughly equivalent to availability in sub-Saharan Africa (SSA), despite very different development profiles. Now I reckon that the SDGs probably represent a best-case scenario for data availability, given the level of support globally for the SDGs. So let's now look at data availability across 900 gender-stratified metrics.

Slide 8	**Data Availability (Gender-stratified data)** And again the SIDS are lower than for any world region, and this time the shortfall is more substantial. One-third of gender-stratified indicators were available for SIDS in 2022; other world regions having roughly two-thirds and higher availability. So some strong evidence for the lack of data availability in the SIDS.
Slide 9	**Overview of initiative** So on to the CaribData initiative. And broadly. We're building a set of resources to enable and encourage Caribbean data sharing and data re-use. • NEXT. We'll be setting up an online infrastructure to enable easier data collection, data sharing and re-use. • NEXT. We'll be developing a training program and longer-term mentoring to increase capacity in data handling and data communication. And we're also developing a data communications web platform. This is all about actively communicating the stories around the data we all produce. And at the same time, advocating for data re-use. • NEXT. We're documenting the legislative environment that encourages or hinders data sharing, and more formally auditing data availability in several fields. • NEXT. And we're now actively building collaborations to develop a sustainable team of data storytellers. We have our RPG funding from the IDB to the end of 2025, and our official project beneficiaries are regional NSOs.