# **Project Description**

**Brief Overview**

Planet Indonesia prioritizes ecosystem conservation through community-led governance to support environments while also empowering local and indigenous peoples. Therefore, it works with many governance institutions, non-governmental organizations, and village partners over multiple landscapes.

Today, there are many organizations wanting to work with Planet Indonesia, but there is limited funding to support all of them and their environmental goals. Therefore, Planet Indonesia is seeking a data-focused approach to determine target areas of concern where funding is most beneficial. There is already much data on terrestrial ecosystems from Global Forest Watch and other publically-available data sets, as well as approaches to determine at-risk terrestrial ecosystems based on various attributes. However, there is less relevant data on marine and coastal environments due to the relative ease of collecting data on land compared to the ocean. Therefore, the goal of this project is to collect marine data so that Planet Indonesia can evaluate best allocation of funding using both marine and terrestrial environmental data.

These areas would ideally have lots of biodiversity, support carbon sequestration, provide coastal protection, or offer other important environmental advantages. Data would be divided by Indonesian province, and this environmental data would be used to develop a method to decide environmental risk categories and other labels for each Indonesian province. Data would first be used to create a mathematical formula that attaches weights to different environmental factors to determine need for funding. If there is enough time, a machine learning model would be created. This model would take environmental data as input and output a number indicating need for funding; however, it will only be created if enough training data is found. Data would be handled mostly using R and Python libraries such as Pandas and TensorFlow.

**Key Deliverables**

* Multiple organized datasets that can be used to create a machine learning model
  + Includes ocean data, satellite imagery, risk factors, level of control that community has over lands, and other publicly-available datasets to show environmental attributes of Indonesian provinces
* Mathematical formula and methodology using data to determine risk category and environmental importance of each province
* Strategy report indicating which province(s) should be allocated funding with methodology and explanation included
  + Reasons can include biodiversity data, carbon sequestration rates, species risk factors, number of hectares of ecosystem, potential for coastal protection
* If time permits: machine learning model to determine provincial need for funding