

# Help & Documentation for the Marine Biodiversity Observation Network (MBON)

*2017-03-24*



# Contents

<b>1</b>	<b>Prerequisites</b>	<b>5</b>
<b>2</b>	<b>Introduction</b>	<b>7</b>
<b>3</b>	<b>Setup Software for Reproducible Research</b>	<b>9</b>
3.1	Git, Github . . . . .	9
3.2	R, Rstudio . . . . .	9
3.3	Diving Deeper . . . . .	9
<b>4</b>	<b>Methods</b>	<b>11</b>
<b>5</b>	<b>Applications</b>	<b>13</b>
5.1	Interactive Infographics . . . . .	13
<b>6</b>	<b>Final Words</b>	<b>19</b>



# Chapter 1

## Prerequisites

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports, e.g., a math equation  $a^2 + b^2 = c^2$ .

For now, you have to install the development versions of **bookdown** from Github:

```
devtools::install_github("rstudio/bookdown")
```

Remember each Rmd file contains one and only one chapter, and a chapter is defined by the first-level heading #.

To compile this example to PDF, you need to install XeLaTeX.



## Chapter 2

# Introduction

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter 4.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2016) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).



Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa



## Chapter 3

# Setup Software for Reproducible Research

### 3.1 Git, Github

- OHI Manual - Setup a Github Account and Git
  - 5.1 Set up a Github Account and git
  - 5.1.1 Install git software
  - 5.1.2 Set up your Git Identity
  - 5.1.3 Create a folder called github on your computer
  - 5.1.4 Update permissions
  - 5.2 Download R and RStudio
  - 5.3 Cloning a repository to your local computer

### 3.2 R, Rstudio

- OHI Manual - Download R, RStudio
  - **R** statistical programming language
  - **RStudio** integrated development environment (IDE)

### 3.3 Diving Deeper

- the “bible of R”: R for Data Science
- course materials
  - Software Carpentry workshop, particularly Day 2
  - Environmental Informatics course
- GitHub Guides
  - Hello World · GitHub Guides
  - Getting Started with GitHub Pages · GitHub Guides
  - Mastering Markdown · GitHub Guides
- getting started with R

- R tutorials | R-bloggers
- Online Learning – RStudio
- R Markdown

## Chapter 4

# Methods

We describe our methods in this chapter.



# Chapter 5

## Applications

Here we describe applications with conceptual process overviews and technical implementation details.

### 5.1 Interactive Infographics

#### 5.1.1 Product Overview

Jenn Brown conceived of a vision for making complex information readily available to managers and the public through an interactive infographic, that is navigable via:

1. Overview of Habitats
2. Species by Habitat
3. Time Series

This is a data-driven framework in support of National Marine Sanctuaries Condition Report and Management Plans.

Here are picture examples:

1. Overview of Habitats
2. Species by Habitat
3. Time Series

We've identified 3 phases of development to successfully generate this produce for any given setting:

1. **Design**

audience: managers, scientists

2. **Implementation**

audience: data and web technicians

3. **Feedback**

Gather feedback (like A/B testing) from users and iterate on Design and Implementation.

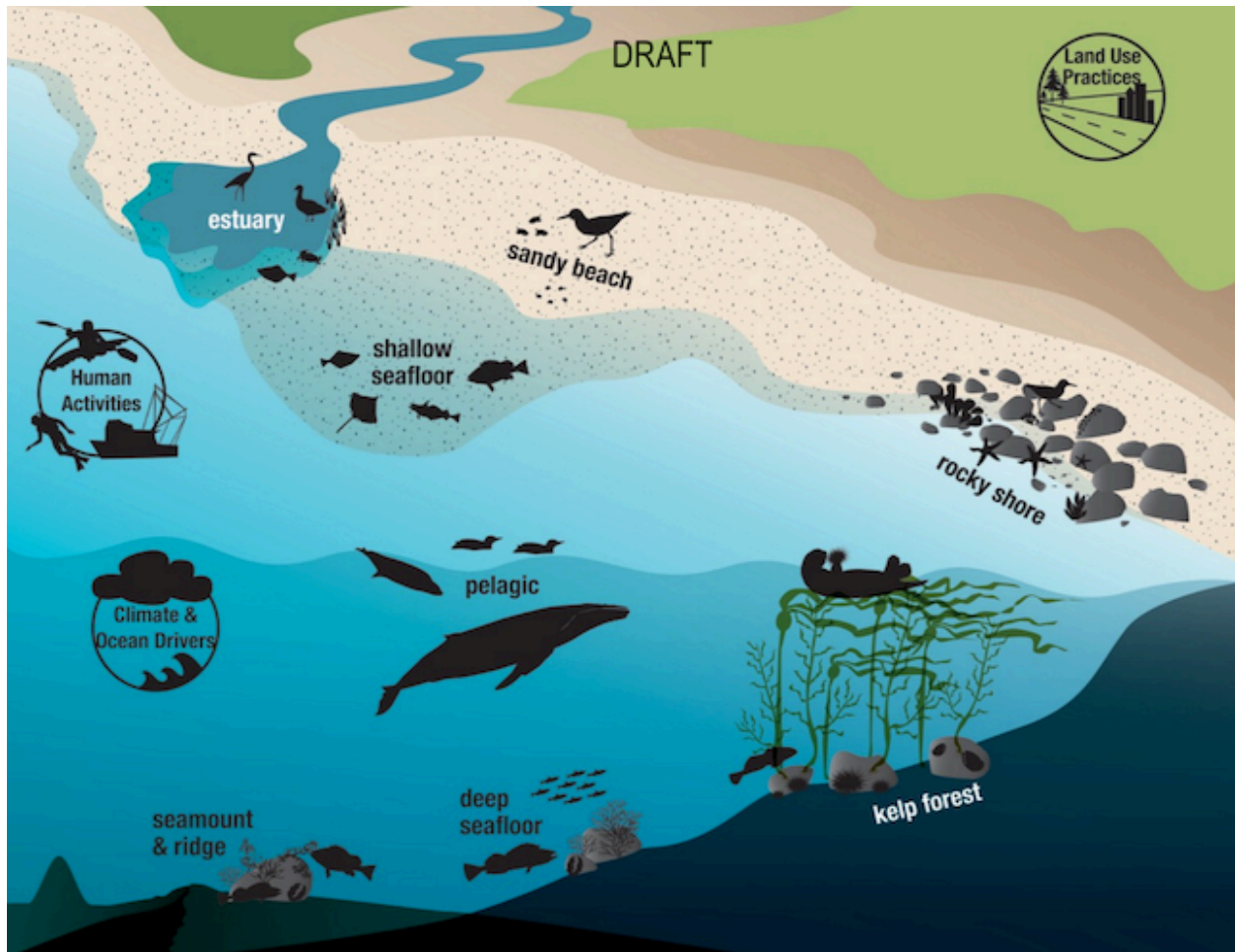


Figure 5.1:

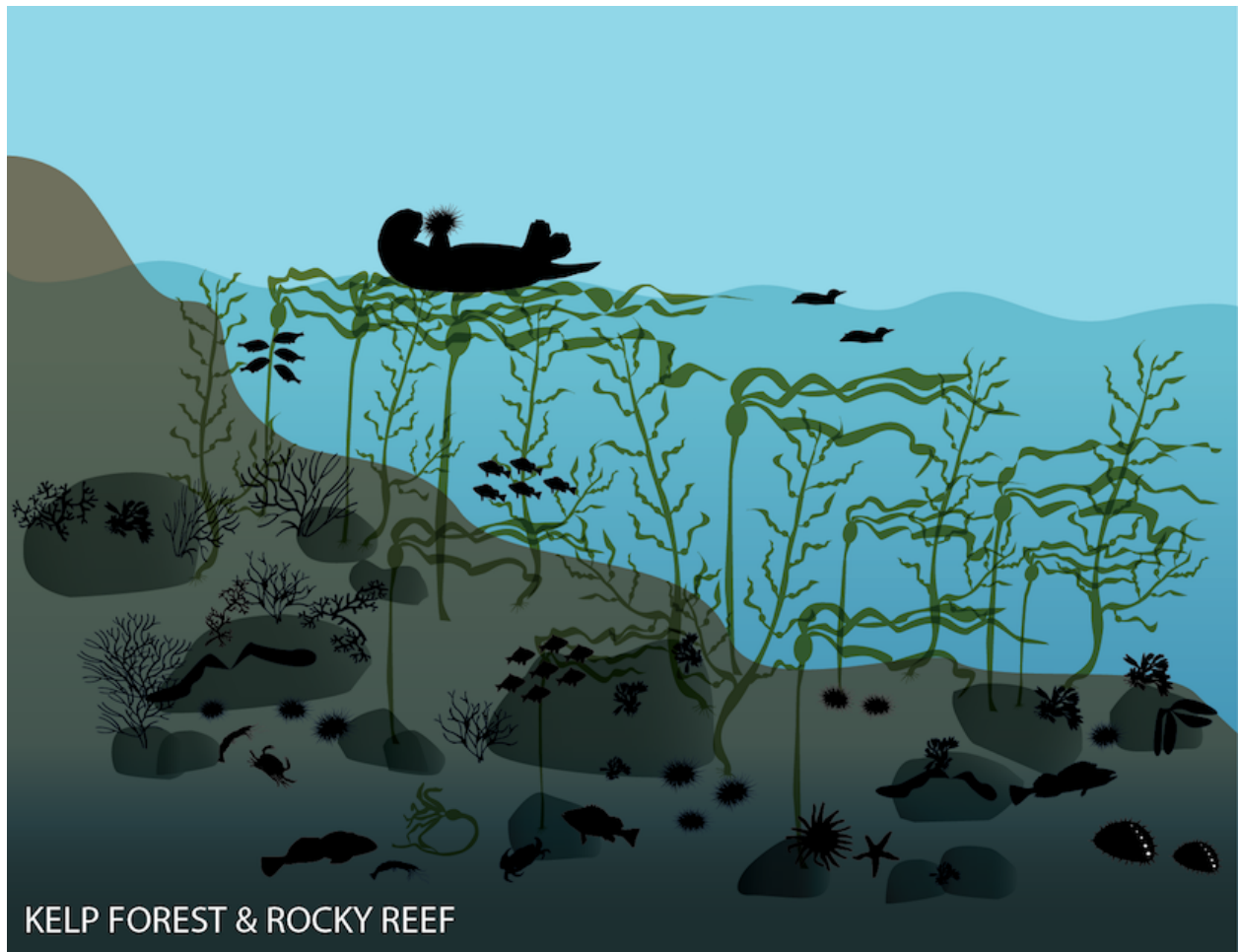


Figure 5.2:

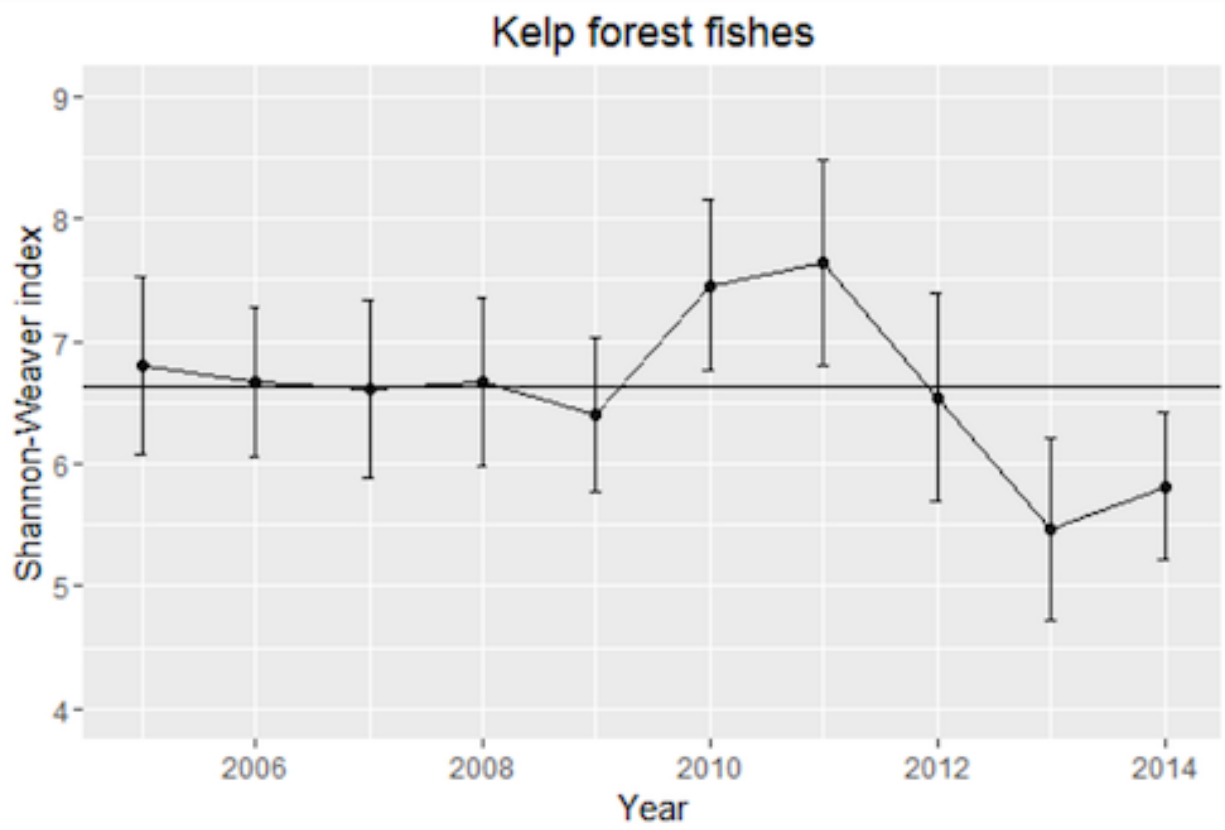


Figure 5.3:



### 5.1.2 Design Process

audience: managers, scientists

- What are the key habitats for your study area?
- Within these habitats, what are the key species and indicators?
- For these species and indicators, what datasets are available to inform them?
- What is the minimum criteria for inclusion of dataset? eg > 5 years of surveys covering > 50% of study area
- For a given dataset what are the key metrics and how do the values get binned into different categories?  
Example categories:

- unprecedented decreasing
- decreasing
- stable
- increasing
- unprecedented increasing
- data deficient

### 5.1.3 Implementation Details

audience: data and web technicians

This product is intended for hosting on a simple web server, without need for backend server-side programming. It will be generated from data products on a daily basis and involves the following elements:

- **Infographic.** These will be rendered using scalable vector graphics (SVG) icons color coded and clickable.
- **Time-series plots.** These interactive plots will be pre-generated from parameterized Rmarkdown documents.

### 5.1.4 Gather Feedback and Iterate

Gather feedback (like A/B testing) from users and iterate on Design and Implementation.



## Chapter 6

# Final Words

We have finished a nice book.



# Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2016). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.3.9.