

Tutorial 6: Refactoring R Code

Table of contents

1	1. Introduction	2
1.1	1.1 Load Libraries and Data	2
2	2. Methods	3
3	3. Model	4
4	4. Results	5
5	5. Conclusion	6
Sayana Imash		

Chapter 1

1. Introduction

In this tutorial, you will refactor the code into separate scripts corresponding to each section. The dataset we will use comes from the `palmerpenguins` package, which contains measurements of penguins from three species.

1.1 1.1 Load Libraries and Data

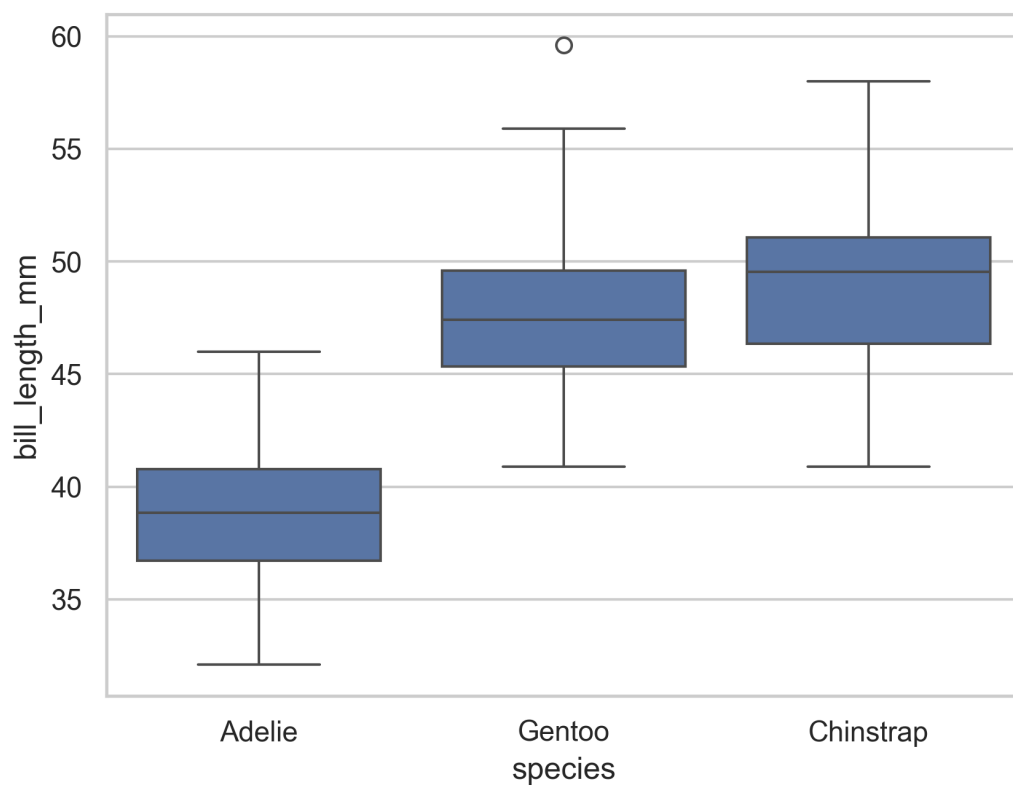
```
species      island bill_length_mm bill_depth_mm flipper_length_mm \
0 Adelie Torgersen      39.1          18.7          181.0
1 Adelie Torgersen      39.5          17.4          186.0
2 Adelie Torgersen      40.3          18.0          195.0
3 Adelie Torgersen      36.7          19.3          193.0
4 Adelie Torgersen      39.3          20.6          190.0

body_mass_g   sex  year
0      3750.0 male 2007
1      3800.0 female 2007
2      3250.0 female 2007
3      3450.0 female 2007
4      3650.0 male 2007
```

Chapter 2

2. Methods

In this section, we perform exploratory data analysis (EDA) and prepare the data for modeling.



Chapter 3

3. Model

We will fit a classification model using `tidymodels` to predict the species of a penguin based on its physical characteristics.

	precision	recall	f1-score	support
Adelie	0.97	1.00	0.99	37
Chinstrap	1.00	0.94	0.97	17
Gentoo	1.00	1.00	1.00	30
accuracy			0.99	84
macro avg	0.99	0.98	0.99	84
weighted avg	0.99	0.99	0.99	84

Chapter 4

4. Results

We evaluate the performance of the model using the test dataset.

Chapter 5

5. Conclusion

In this tutorial, we:

- Loaded and cleaned the `palmerpenguins` dataset.
- Performed exploratory data analysis.
- Built a k-Nearest Neighbors classification model using `tidymodels`.
- Evaluated the model's performance.