Reproducible Documentation with Quarto

Manual for Python Users

Dr Thiyanga S. Talagala

Table of contents

# 3. Set up

1. Install Python on your machine.
2. Install Quarto
3. Go to cmd and type
   * py --version
   * py -m pip install pandas plotly statsmodels
   * py -m pip install palmerpenguins
   * py -m pip install plotnine

* Note: The reason for typing py -m instead of just pip install is to avoid issues where you might have multiple versions of Python installed, and pip could be tied to the wrong version. By running py -m pip, you’re explicitly telling the Python launcher to use the right Python interpreter and its associated pip.

# 4. Loading packages

import pandas as pd  
import plotnine as p9   
from plotnine import ggplot, geom\_point, aes, stat\_smooth, facet\_wrap, labs  
from palmerpenguins import load\_penguins

# 5. Data

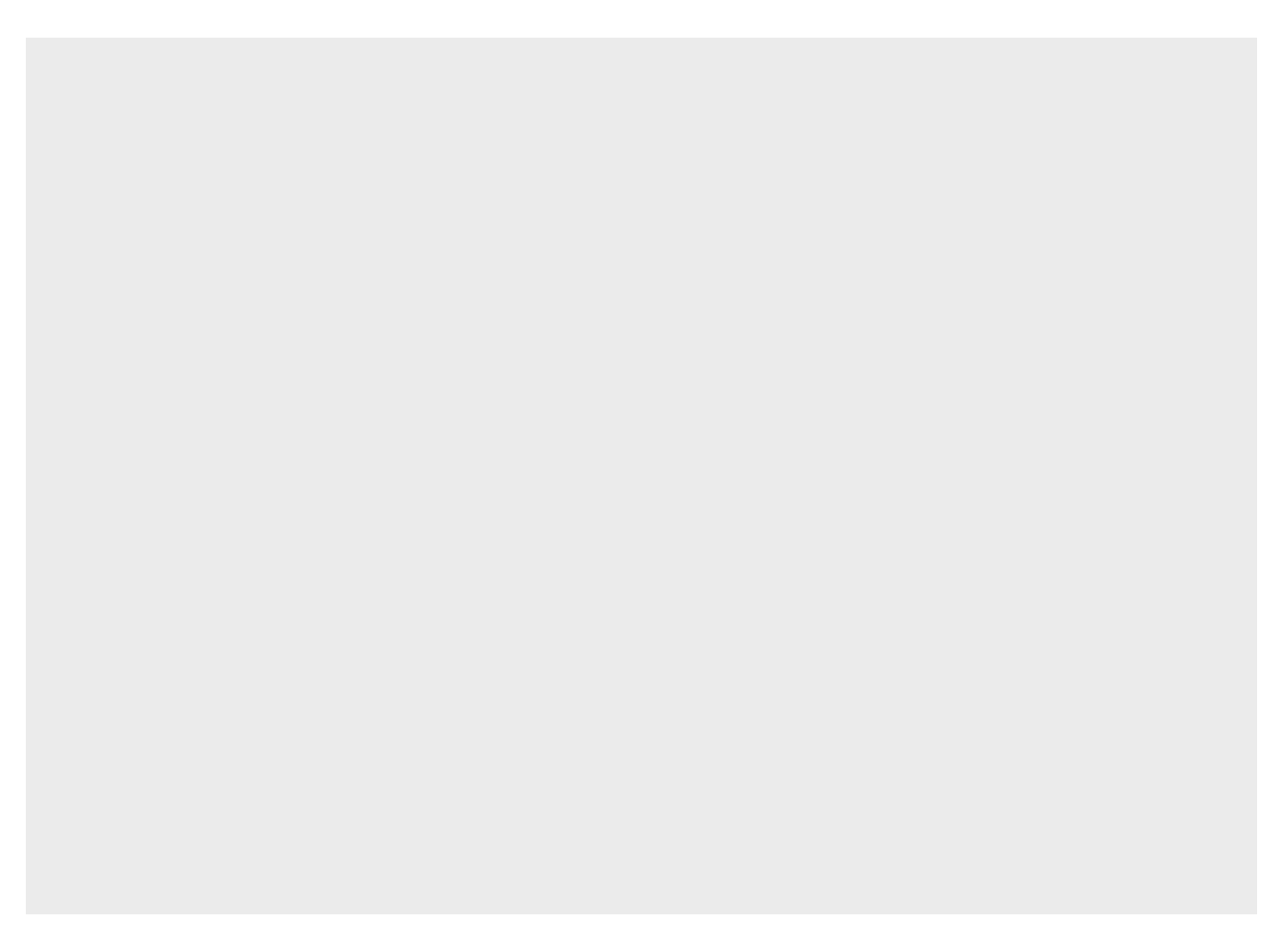
penguins = load\_penguins()  
penguins.head()

|  | species | island | bill\_length\_mm | bill\_depth\_mm | flipper\_length\_mm | body\_mass\_g | sex | year |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | Adelie | Torgersen | 39.1 | 18.7 | 181.0 | 3750.0 | male | 2007 |
| 1 | Adelie | Torgersen | 39.5 | 17.4 | 186.0 | 3800.0 | female | 2007 |
| 2 | Adelie | Torgersen | 40.3 | 18.0 | 195.0 | 3250.0 | female | 2007 |
| 3 | Adelie | Torgersen | NaN | NaN | NaN | NaN | NaN | 2007 |
| 4 | Adelie | Torgersen | 36.7 | 19.3 | 193.0 | 3450.0 | female | 2007 |

# 6. Data Visualisation with plotnine

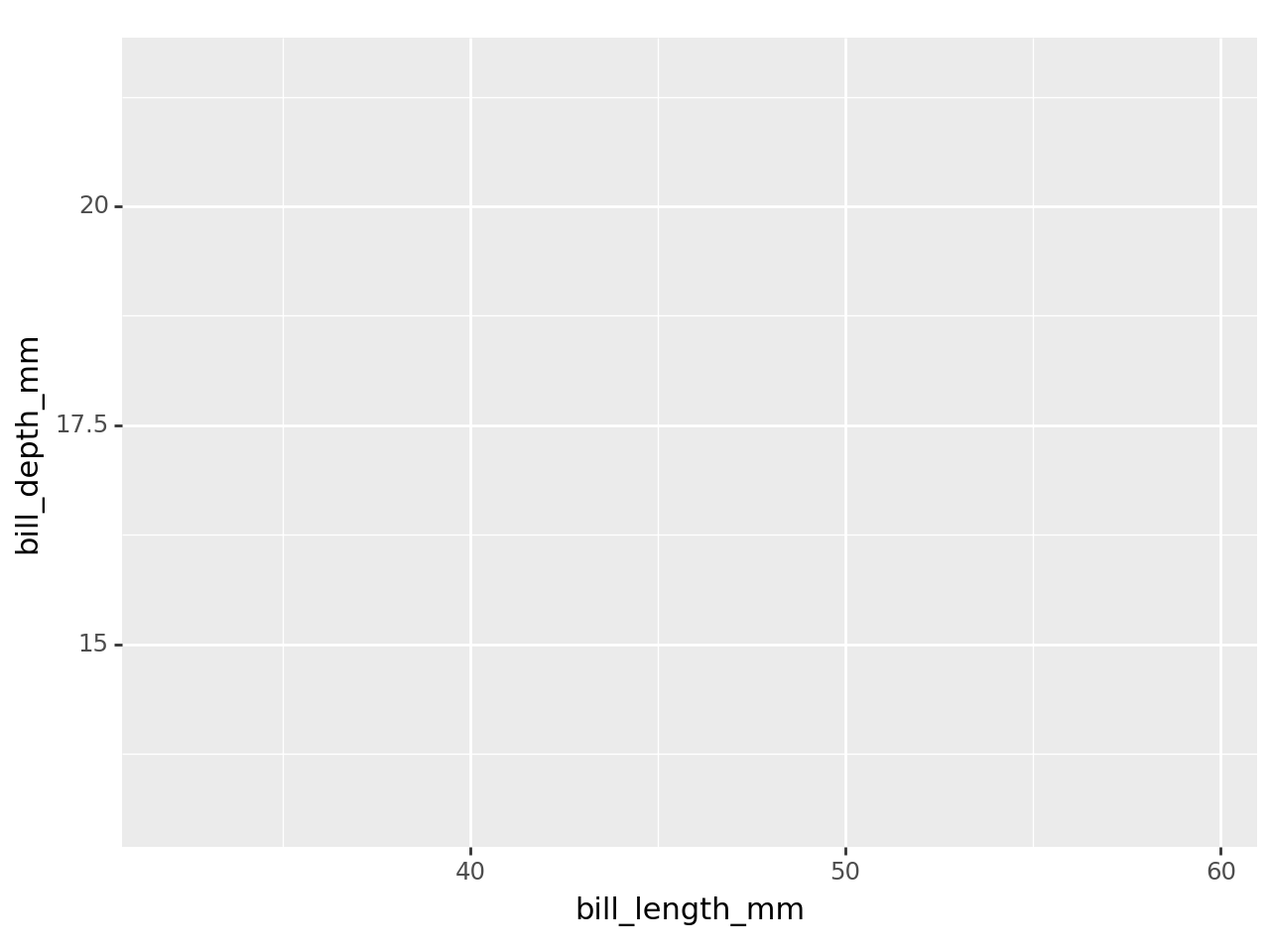
## 6.1 Step 1: Take data and obtain a canvas for plotting

ggplot(penguins)



## 6.2 Step 2: Defining Aesthetics

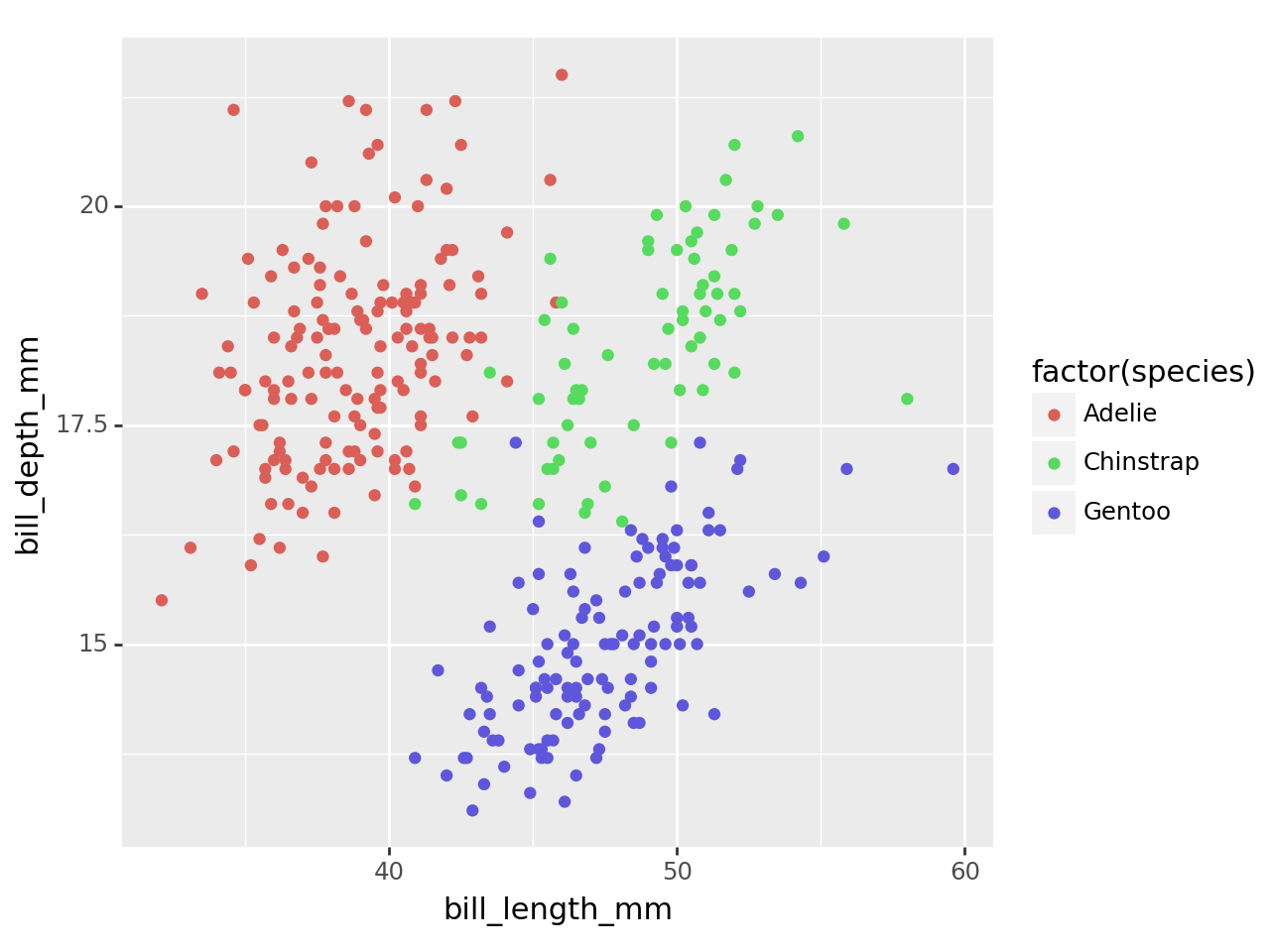
ggplot(penguins, aes("bill\_length\_mm", "bill\_depth\_mm", color="factor(species)"))



## 6.3 Step 3: Add Gemetries

ggplot(penguins, aes("bill\_length\_mm", "bill\_depth\_mm", color="factor(species)")) + geom\_point()

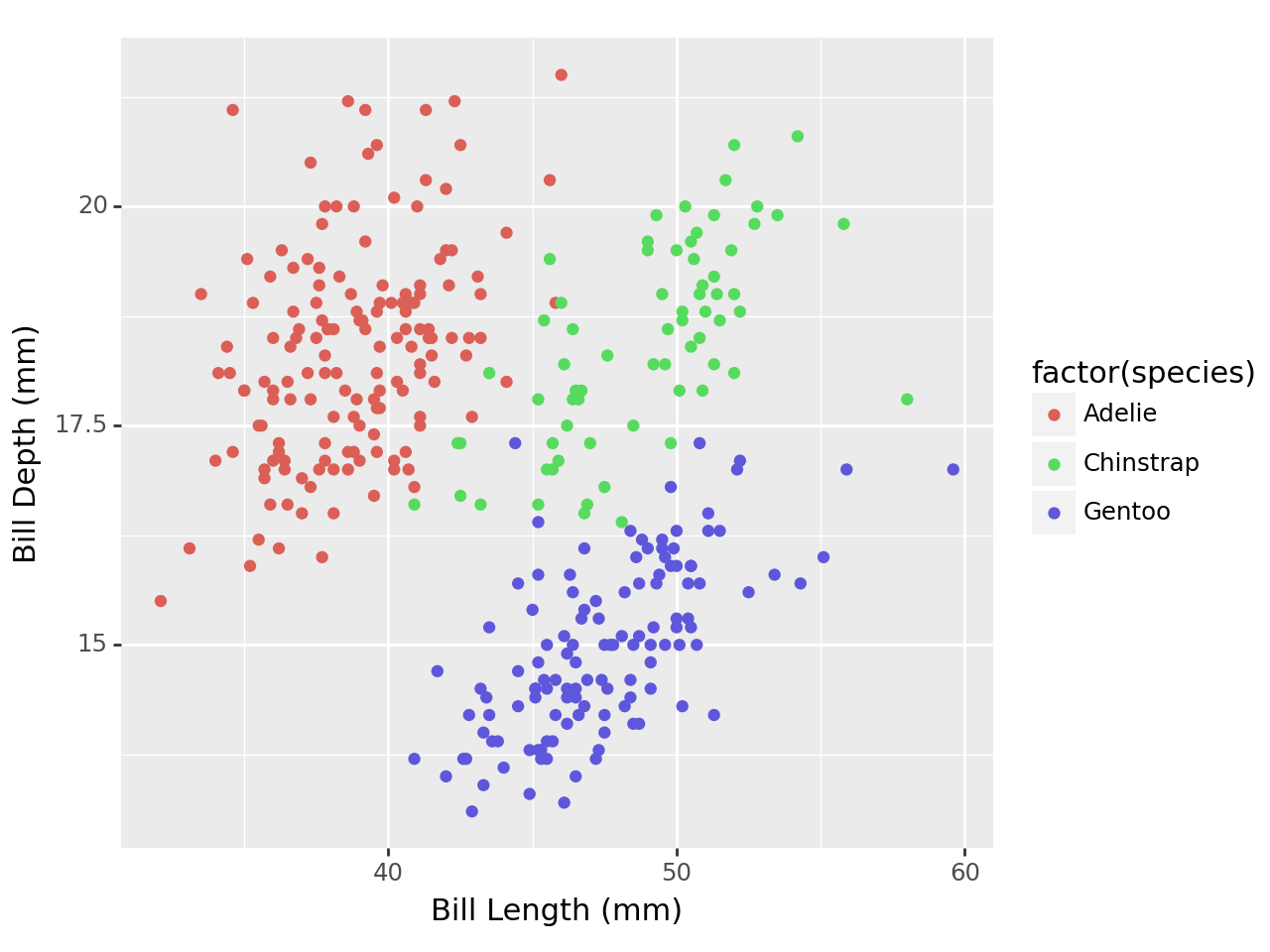
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_point : Removed 2 rows containing missing values.



## 6.4 Step 4: Create the plot with customized axis labels

ggplot(penguins, aes("bill\_length\_mm", "bill\_depth\_mm", color="factor(species)")) + geom\_point() + labs(x="Bill Length (mm)", y="Bill Depth (mm)")

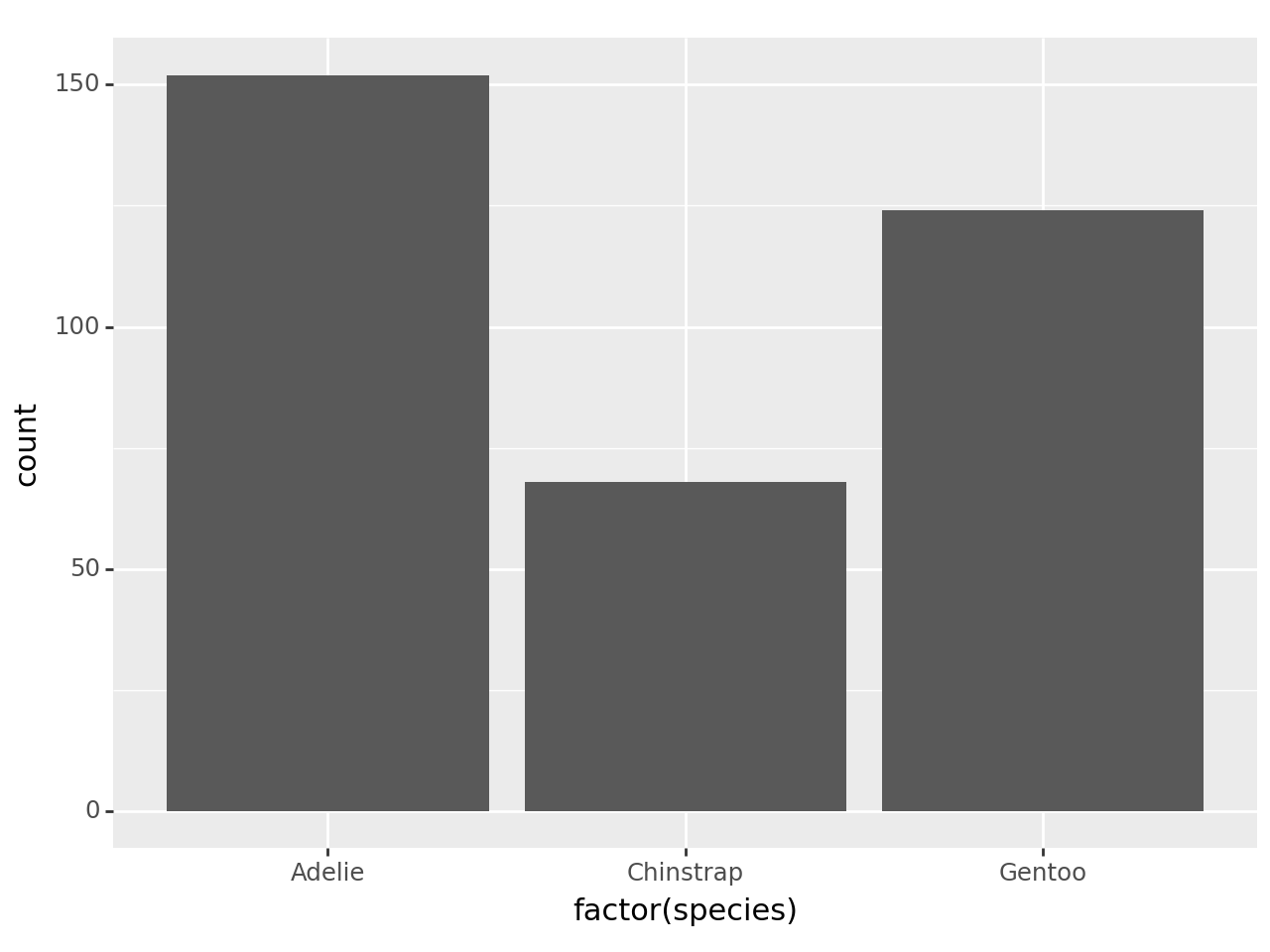
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_point : Removed 2 rows containing missing values.



# 7. Change themes

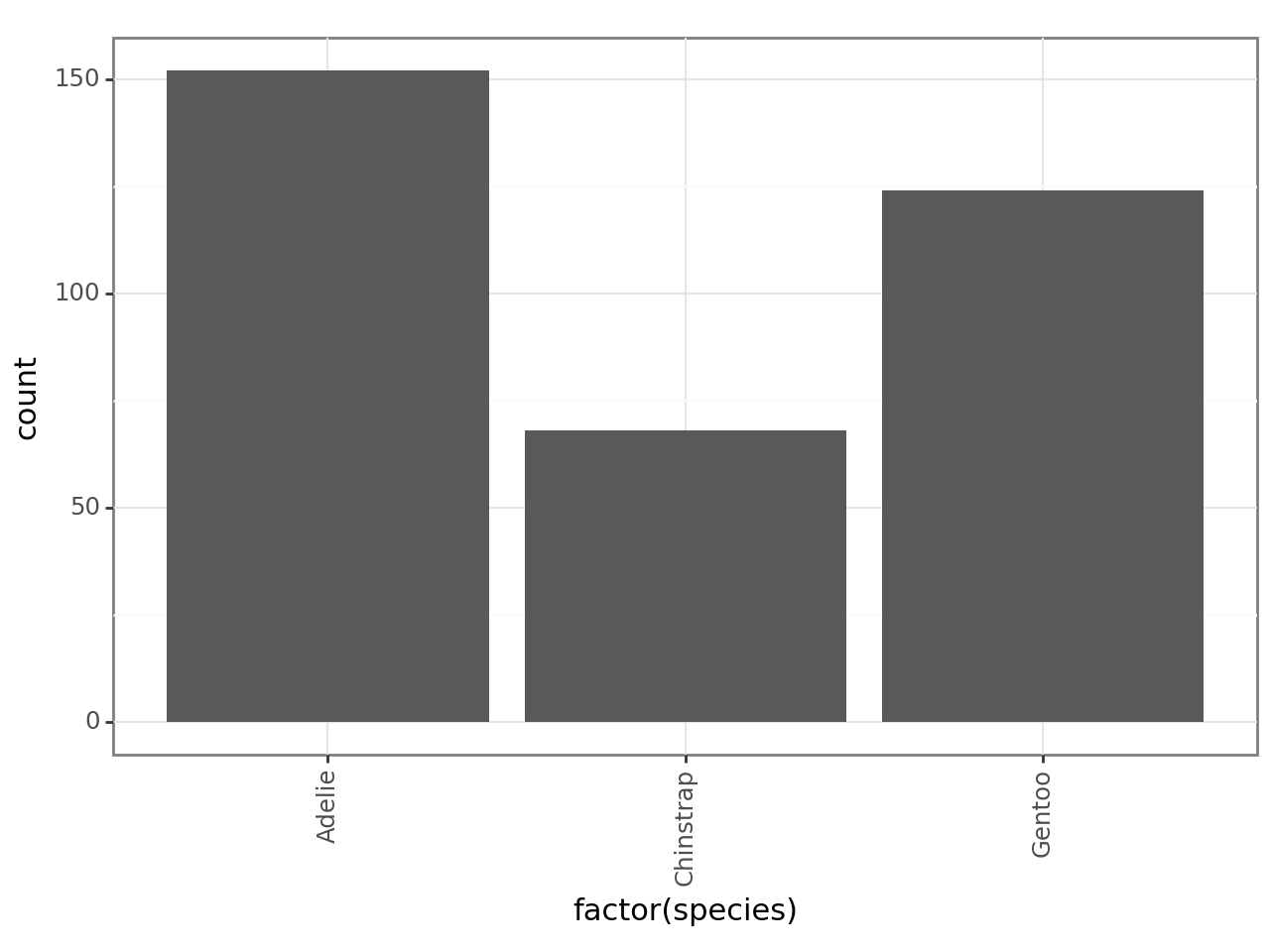
## 7.1 Without customized themes

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)')) + p9.geom\_bar()



## 7.2 With customized themes

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)')) + p9.geom\_bar() + p9.theme\_bw() + p9.theme(axis\_text\_x = p9.element\_text(angle=90))

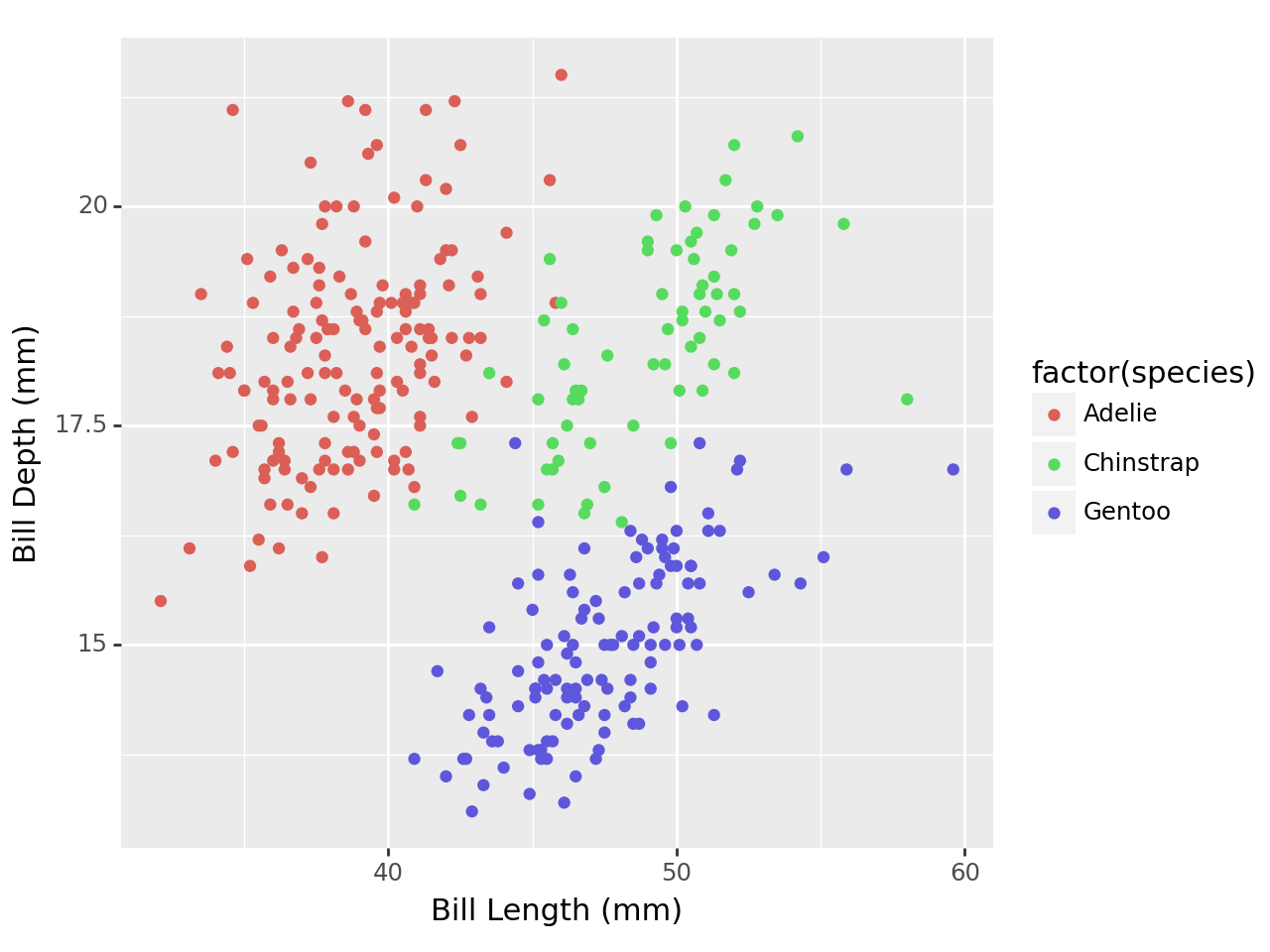


# 8. Faceting

## 8.1 Without faceting

ggplot(penguins, aes("bill\_length\_mm", "bill\_depth\_mm", color="factor(species)")) + geom\_point() + labs(x="Bill Length (mm)", y="Bill Depth (mm)")

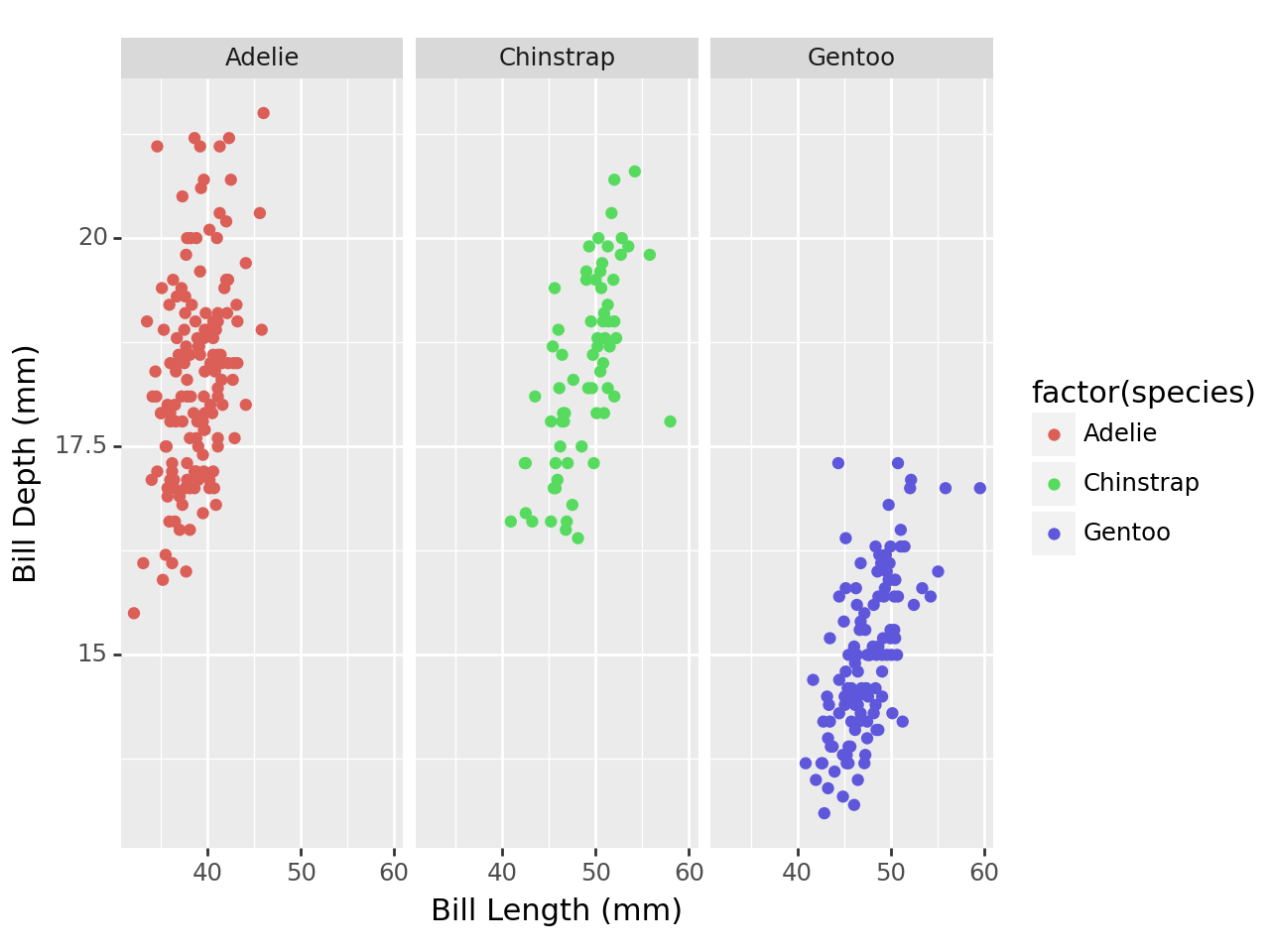
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_point : Removed 2 rows containing missing values.



## 8.2 With faceting

ggplot(penguins, aes("bill\_length\_mm", "bill\_depth\_mm", color="factor(species)")) + geom\_point() + labs(x="Bill Length (mm)", y="Bill Depth (mm)") + facet\_wrap("species")

C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_point : Removed 2 rows containing missing values.

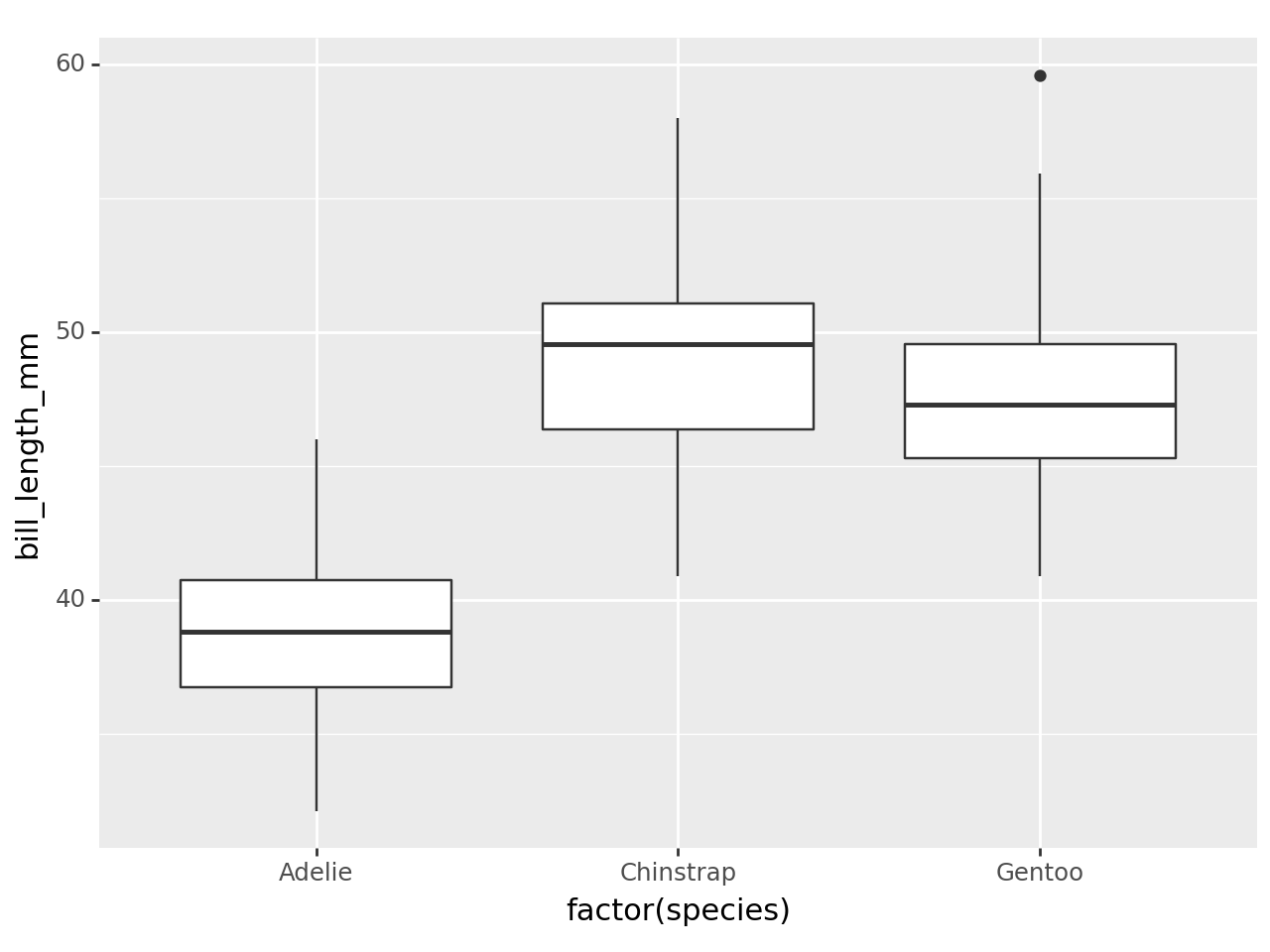


# 9. Other geoms

## 9.1 geom\_boxplot

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm')) + p9.geom\_boxplot()

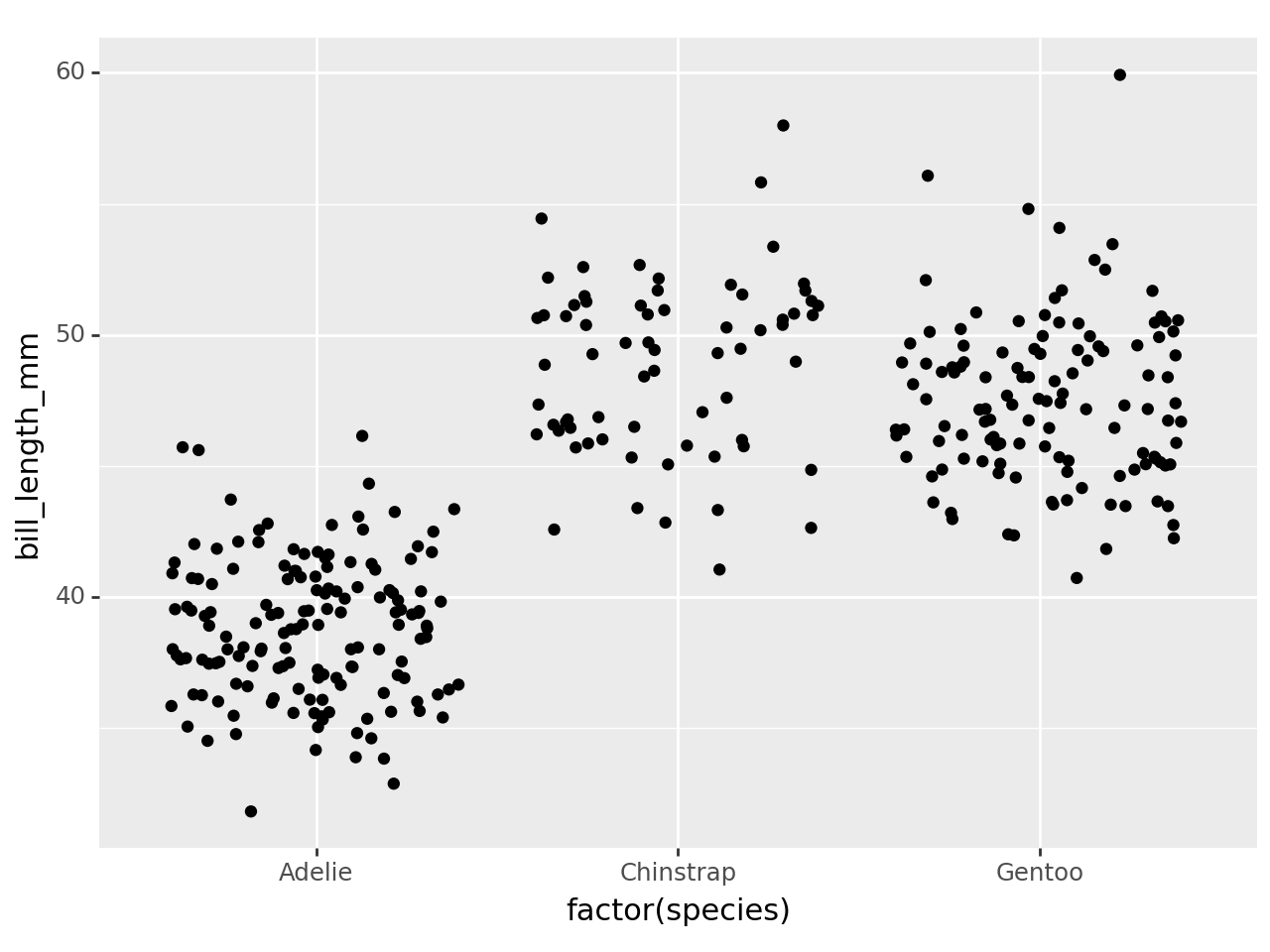
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.



## 9.2 geom\_jitter

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm')) + p9.geom\_jitter()

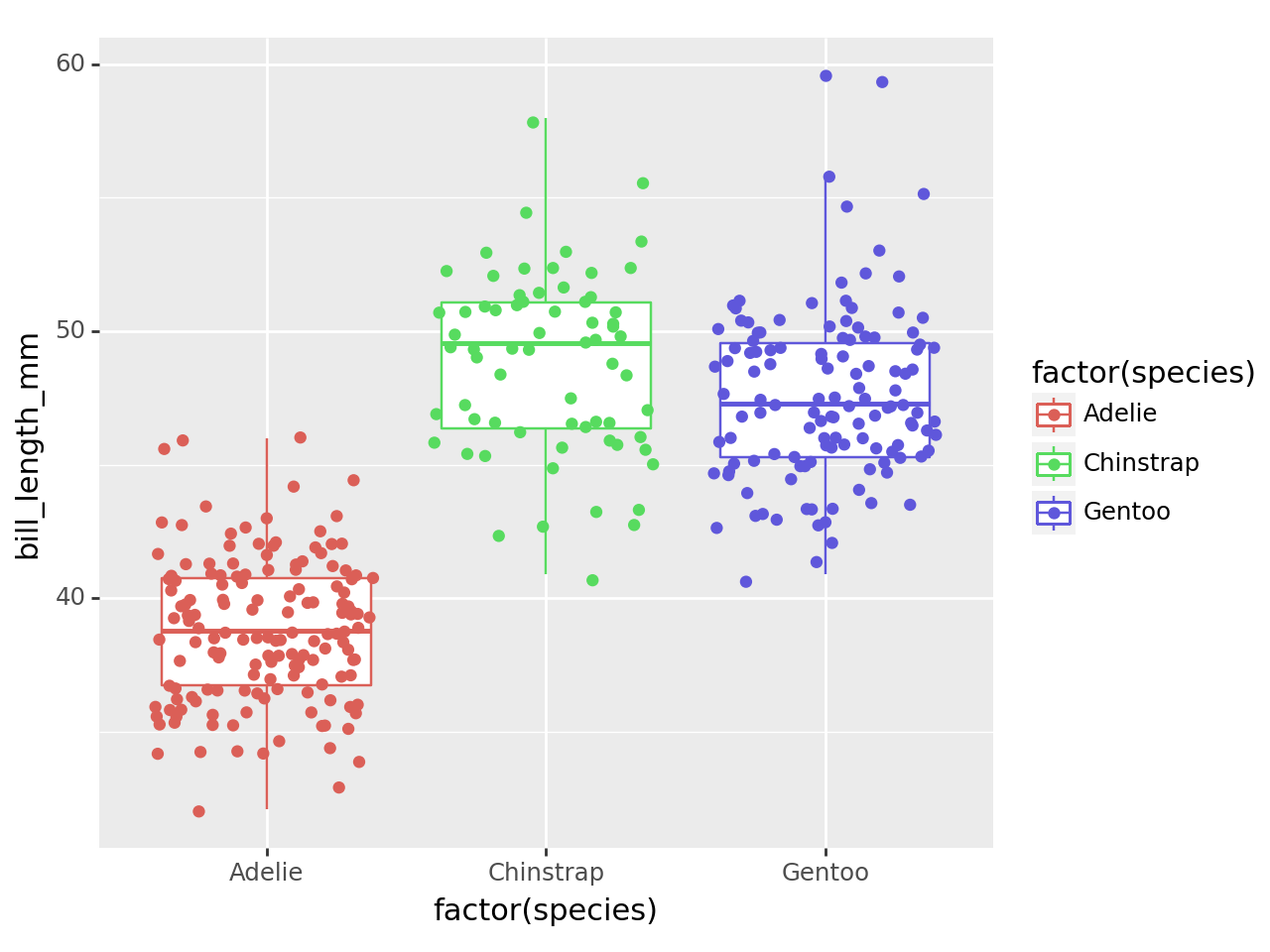
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



## 9.3 geom\_boxplot + geom\_jitter

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm', color='factor(species)')) + p9.geom\_boxplot()+ p9.geom\_jitter()

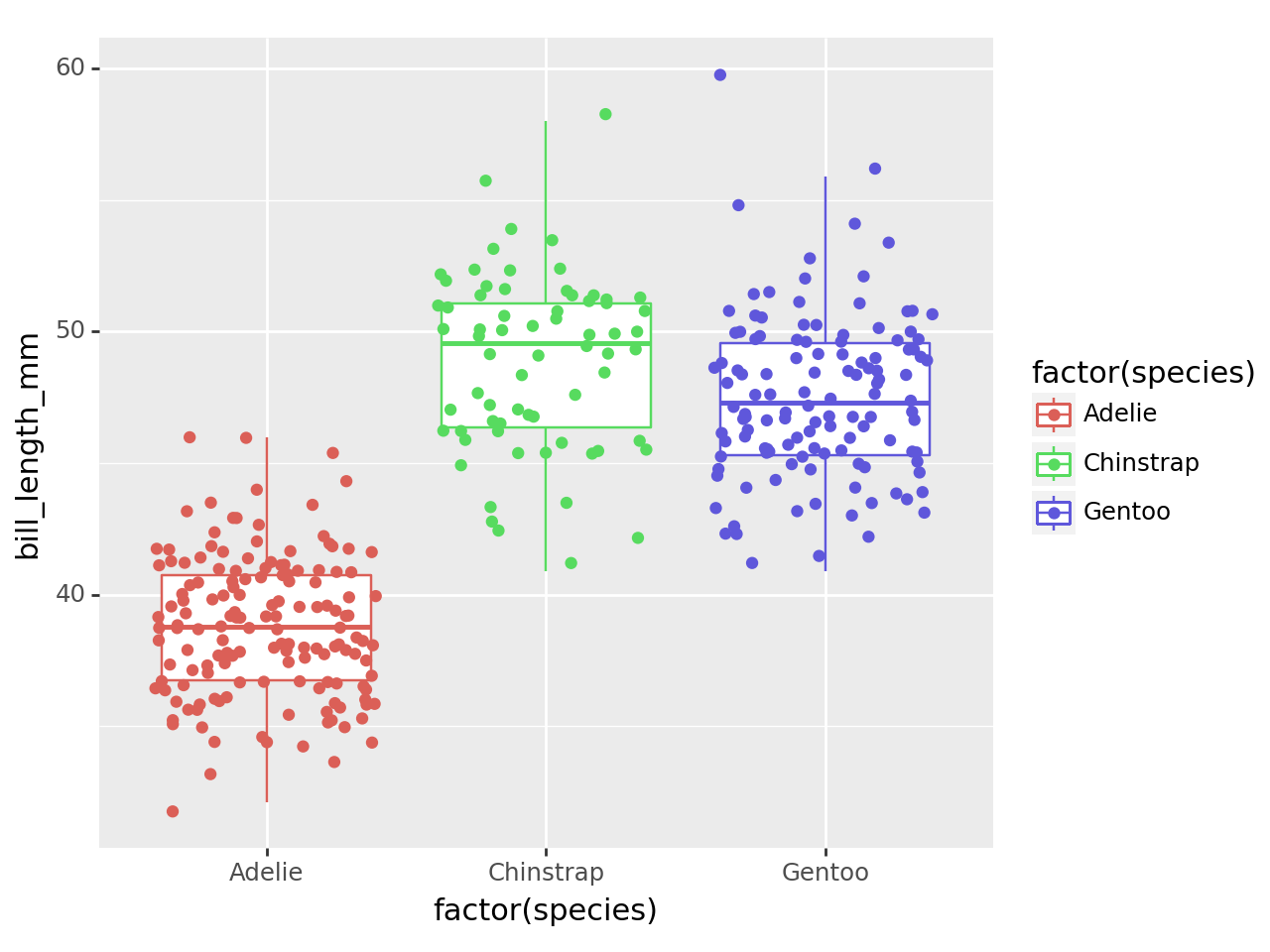
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.  
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



## 9.4 To remove outliers from the boxplot in your ggplot code

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm', color='factor(species)')) + p9.geom\_boxplot(outlier\_shape = "None")+ p9.geom\_jitter()

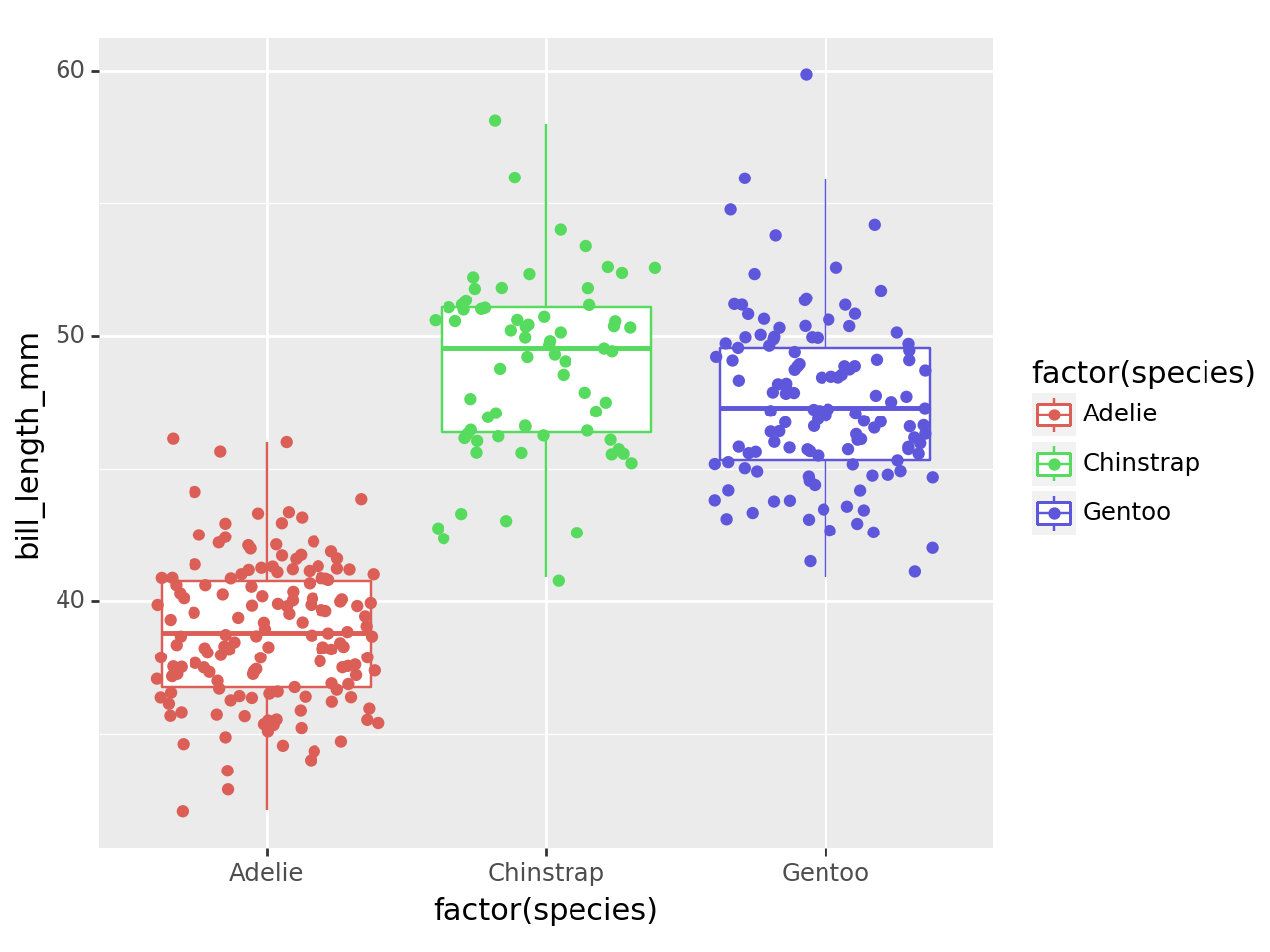
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.  
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



# 10. Working with chunk options

## 10.1 echo: false

C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.  
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



## 10.2 fig-cap

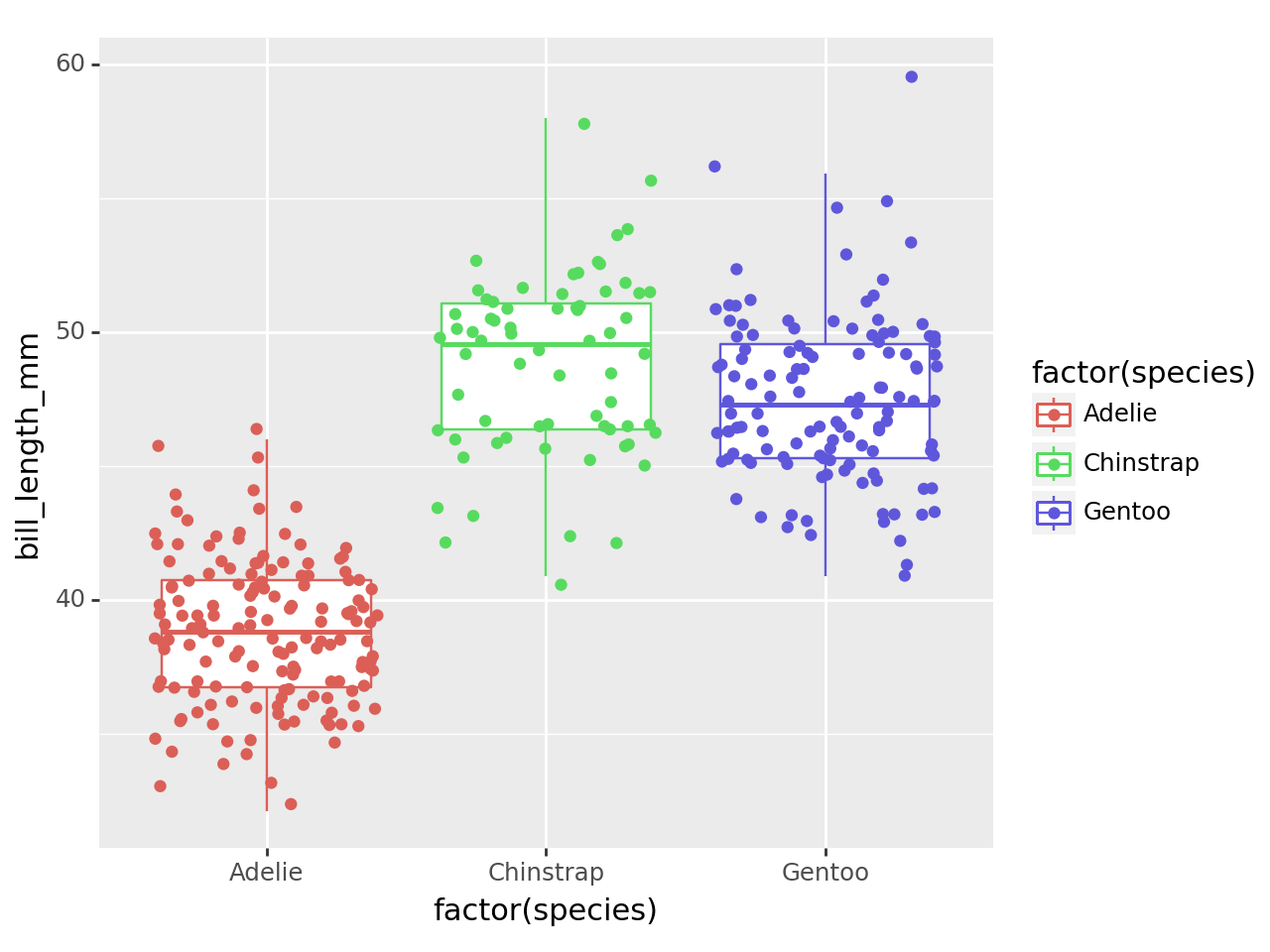
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.  
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



Distribution of Bill length by Species

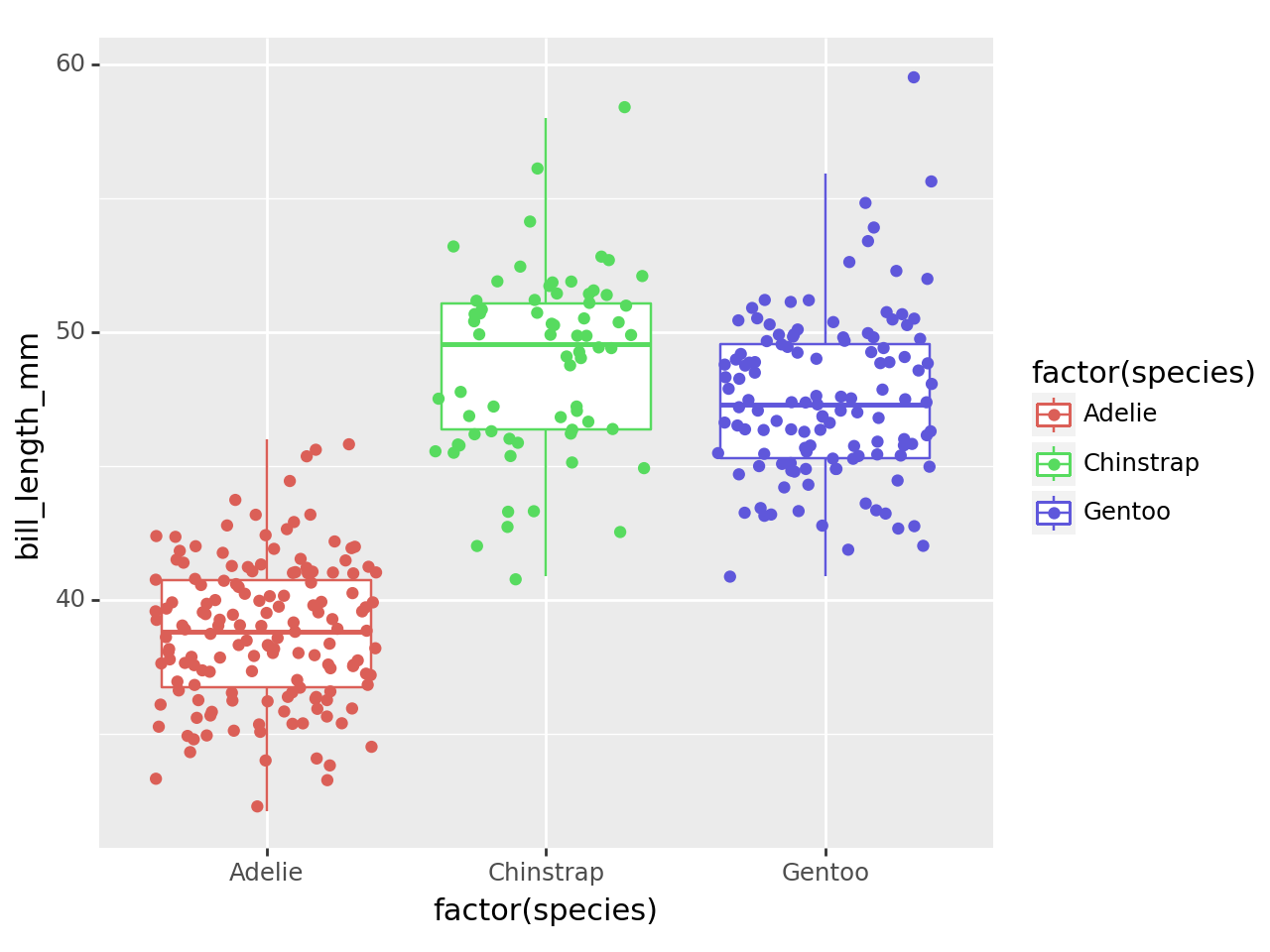
## 10.3 fig width

C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.  
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:364: PlotnineWarning: geom\_jitter : Removed 2 rows containing missing values.



Distribution of Bill length by Species

## 10.4 Suprress warnings and messages



Distribution of Bill length by Species

## 10.5 Fenced Echo

```{python}  
1 + 1  
```

2

1 + 1

2

## 10.6 output and code-overflow

```{python}  
#| output: false  
#| code-overflow: wrap  
1 + 1  
```

1 + 1

# 11. Highlighting

1 + 1

# 12. Themes

format:  
 html:  
 theme: united

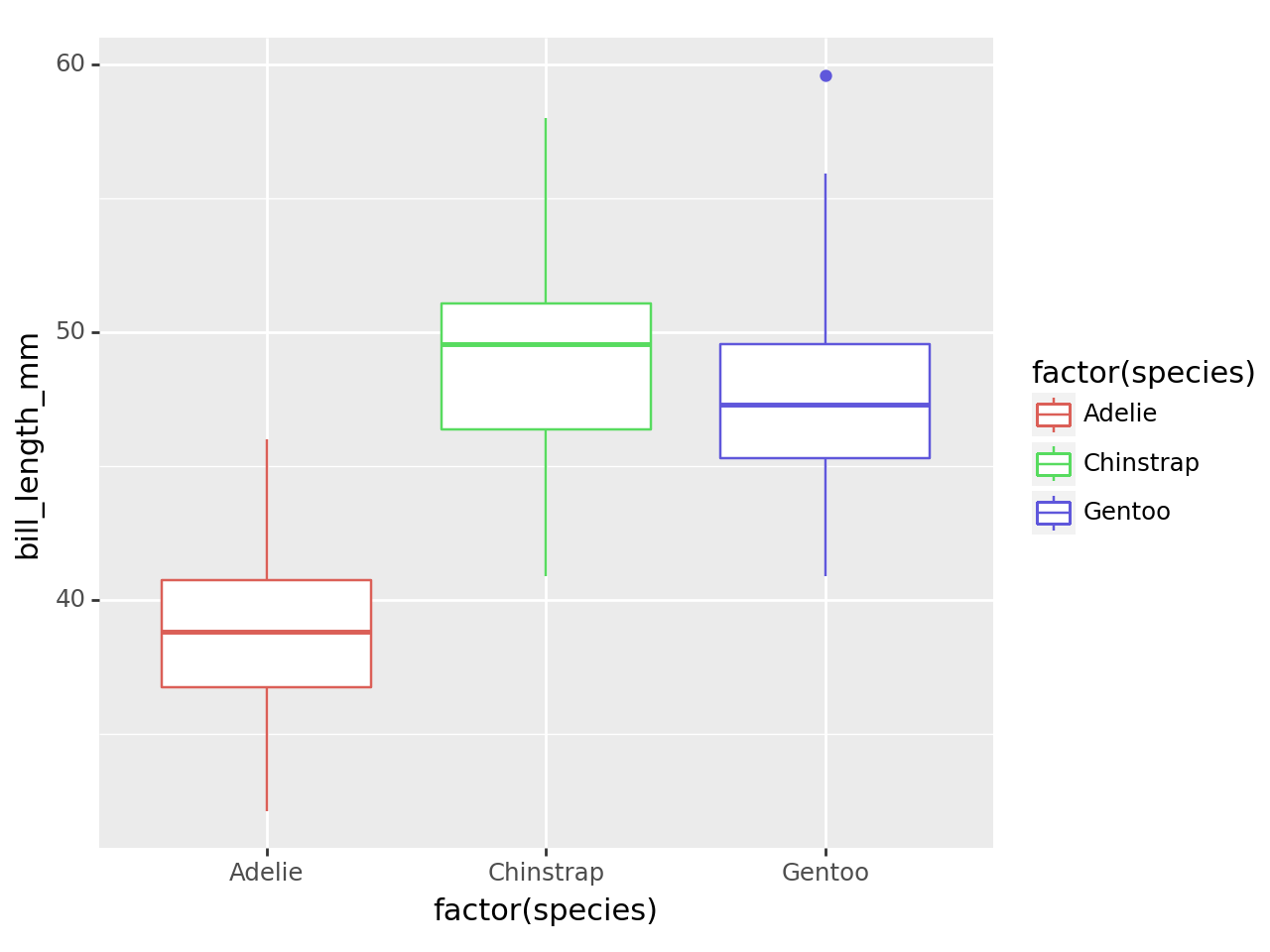
format:  
 html:   
 theme: cosmo  
 fontsize: 1.1em  
 linestretch: 1.7

format:  
 html:   
 theme: darkly

# 13. Folding Code

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm', color='factor(species)')) + p9.geom\_boxplot()

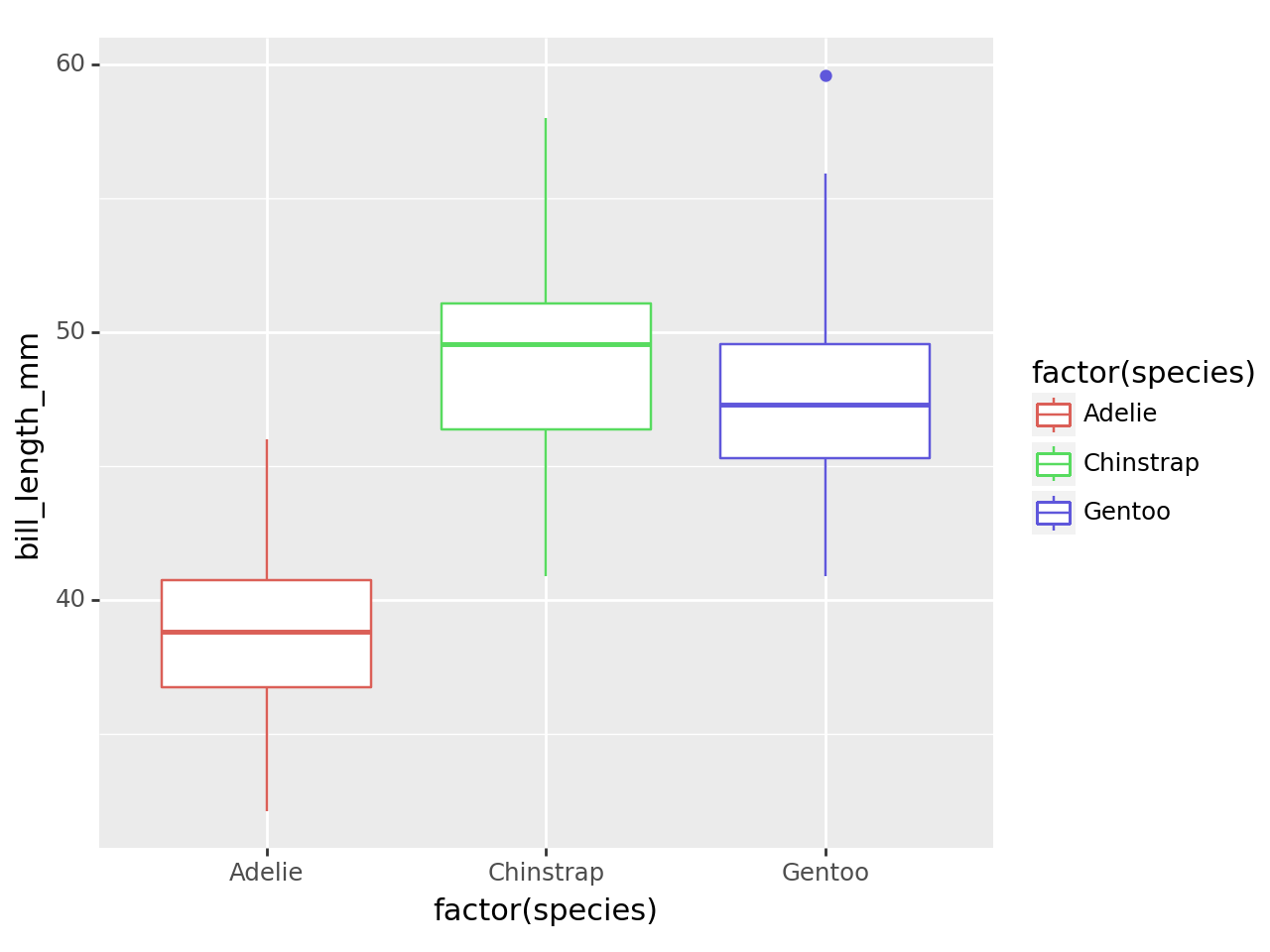
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.



# 14. Folding Code with code-summary

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm', color='factor(species)')) + p9.geom\_boxplot()

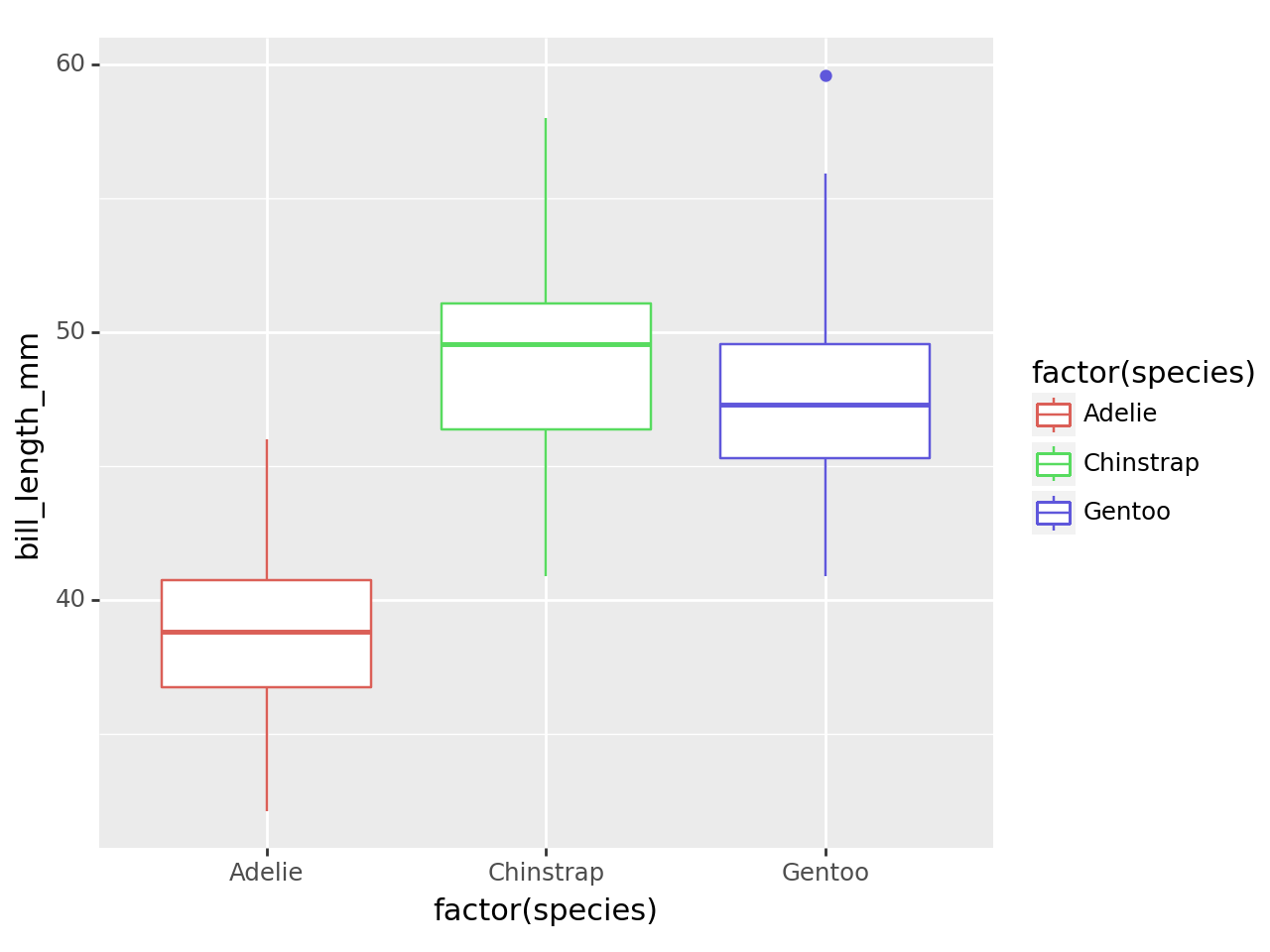
C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.



# 15. Code-overflow

p9.ggplot(data=penguins,  
 mapping=p9.aes(x='factor(species)', y='bill\_length\_mm', color='factor(species)')) + p9.geom\_boxplot()

C:\Users\DELL\AppData\Local\Programs\Python\Python312\Lib\site-packages\plotnine\layer.py:284: PlotnineWarning: stat\_boxplot : Removed 2 rows containing non-finite values.



# 16. Table of content

toc: true  
toc-depth: 2

# 17. Title banner

title-block-banner: true