

Test Book

Invalid Date

Table of contents

Preface	3
1 Data Load	4
2 Data Load	5
3 Penguins EDA	6
3.1 Penguin Size and Mass by Sex and Species	6
3.2 Penguin Size vs Mass by Species	6

Preface

This is a Test Quarto book.

1 Data Load

2 Data Load

Load data to DuckDB

```
import duckdb
from palmerpenguins import load_penguins

con = duckdb.connect('my-db.duckdb')
df = load_penguins()
con.execute('CREATE TABLE IF NOT EXISTS penguins AS SELECT * FROM df')
con.close()
```

3 Penguins EDA

3.1 Penguin Size and Mass by Sex and Species

```
import duckdb
import seaborn as sns
import matplotlib.pyplot as plt

con = duckdb.connect('my-db.duckdb')
df = con.execute("SELECT * FROM penguins").fetchdf().dropna()
con.close()

summary_df = (
    df.groupby(['species', 'sex'])
      .agg({col: 'mean' for col in df.columns if col.endswith('mm') or col.endswith('g')})
      .reset_index()
)
summary_df
```

	species	sex	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g
0	Adelie	female	37.257534	17.621918	187.794521	3368.835616
1	Adelie	male	40.390411	19.072603	192.410959	4043.493151
2	Chinstrap	female	46.573529	17.588235	191.735294	3527.205882
3	Chinstrap	male	51.094118	19.252941	199.911765	3938.970588
4	Gentoo	female	45.563793	14.237931	212.706897	4679.741379
5	Gentoo	male	49.473770	15.718033	221.540984	5484.836066

3.2 Penguin Size vs Mass by Species

```

sns.lmplot(
    data=df,
    x='bill_length_mm',
    y='body_mass_g',
    hue='species',
    aspect=1.5,
    height=6,
    legend=False
)
plt.legend(loc='upper left', bbox_to_anchor=(0.05, 0.95), title='Species')
plt.show()

```

