

Chương 7: Trực quan hóa Dữ liệu với Plotly

Contents

- Trực quan hóa Dữ liệu với Plotly
- 7.1 Giới thiệu về Plotly
- 7.2 Tạo Biểu Đồ Cơ Bản với Plotly
- 7.3 Các Biểu Đồ Tương Tác
- 7.4 Tùy chỉnh Biểu Đồ
- 7.5 Tạo Biểu Đồ Nâng Cao
- 7.6 Xuất và Chia Sẻ Biểu Đồ
- 7.7 Các Ví Dụ Thực Tế
- 7.8 Tài liệu tham khảo
- 7.9 Tổng kết chương 7

Trực quan hóa Dữ liệu với Plotly

Chương này sẽ hướng dẫn cách sử dụng thư viện Plotly để tạo các biểu đồ tương tác trong Python, từ những biểu đồ cơ bản đến nâng cao. Nội dung chương có thể được tổ chức thành các phần sau:

7.1 Giới thiệu về Plotly

Plotly là một thư viện mã nguồn mở giúp tạo ra các biểu đồ tương tác đẹp mắt và dễ sử dụng trong Python. Nó hỗ trợ nhiều loại biểu đồ khác nhau và cho phép xuất biểu đồ dưới dạng

HTML hoặc hình ảnh.

Lợi ích của việc sử dụng Plotly:

- **Tính tương tác cao:** Người dùng có thể tương tác với biểu đồ (zoom, di chuyển, hover, click).
- **Hỗ trợ nhiều loại biểu đồ:** Bao gồm biểu đồ đường, cột, phân tán, hộp, bản đồ, biểu đồ 3D, và nhiều loại khác.
- **Khả năng xuất biểu đồ:** Plotly hỗ trợ xuất biểu đồ dưới dạng HTML hoặc hình ảnh.

Cách cài đặt và thiết lập:

```
pip install plotly
```

Để xuất biểu đồ dưới dạng hình ảnh, bạn cần cài thêm kaleido:

```
pip install kaleido
```

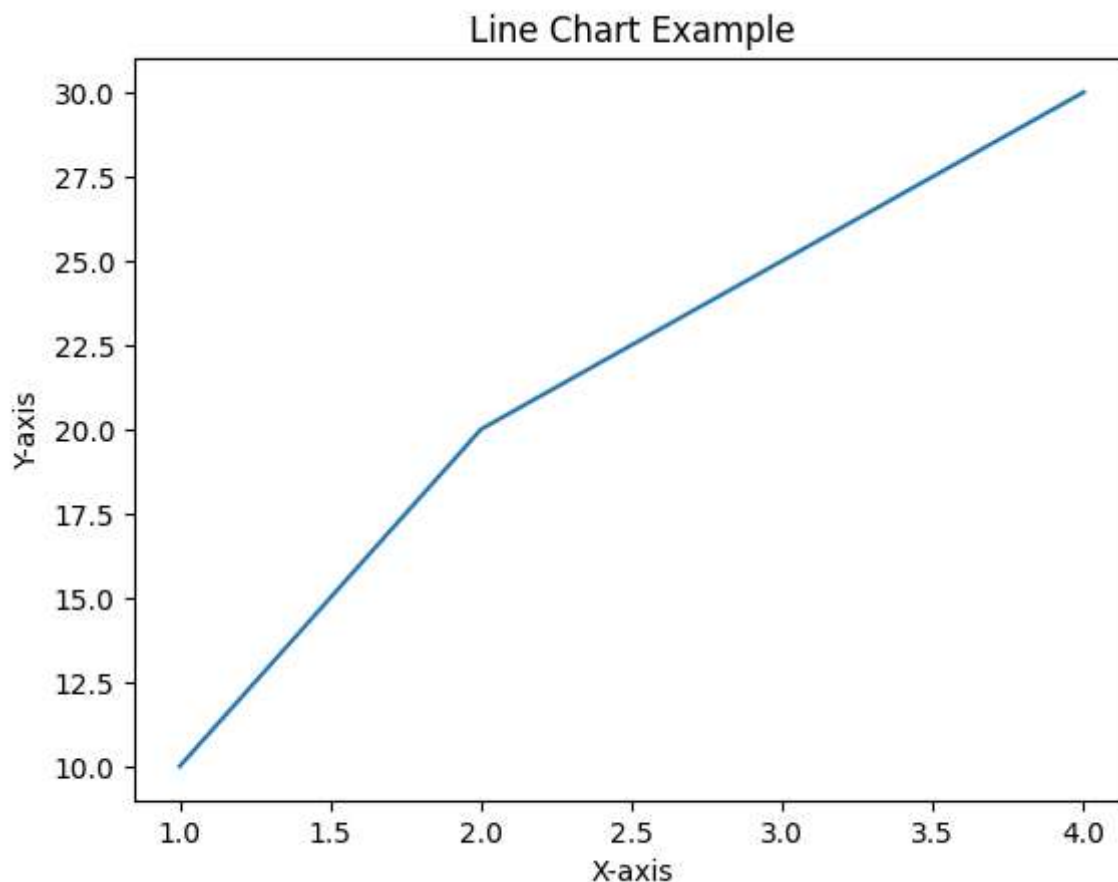
7.2 Tạo Biểu Đồ Cơ Bản với Plotly

7.2.1 Biểu đồ đường (Line Chart):

- Biểu đồ đường giúp bạn trực quan hóa sự thay đổi của một biến theo thời gian hoặc các giá trị liên tục.

```
import matplotlib.pyplot as plt

# Tạo biểu đồ đường
x = [1, 2, 3, 4]
y = [10, 20, 25, 30]
plt.plot(x, y)
plt.title("Line Chart Example")
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.show()
```

**Chú thích:**

- `go.Figure()`: Tạo một đối tượng Figure.
- `go.Scatter()`: Dùng để vẽ biểu đồ đường. `mode='lines+markers'` giúp vẽ cả đường và các điểm đánh dấu trên đó.
- `fig.update_layout()`: Cập nhật các thông số như tiêu đề và nhãn trục.

7.2.2 Biểu đồ cột (Bar Chart):

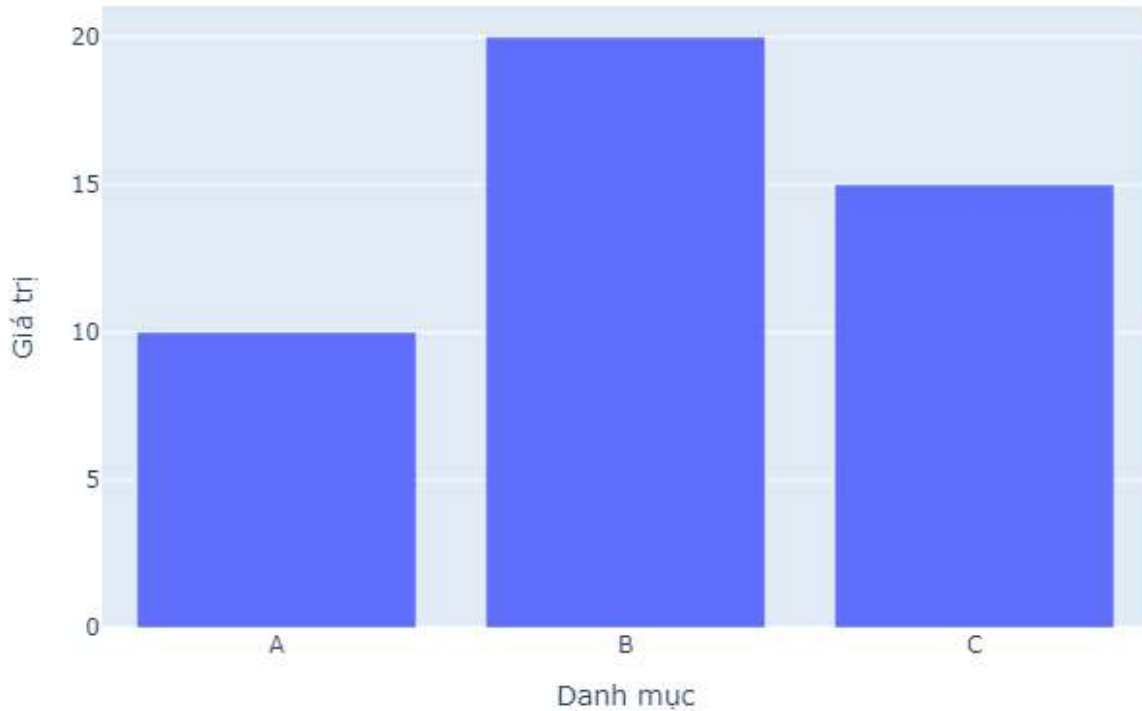
- Biểu đồ cột giúp bạn so sánh giá trị giữa các danh mục.

```
import plotly.graph_objects as go

fig = go.Figure(data=go.Bar(x=['A', 'B', 'C'], y=[10, 20, 15]))
fig.update_layout(title="Biểu đồ Cột", xaxis_title="Danh mục", yaxis_title="Giá trị")
fig.write_image("chart.png")

from IPython.display import Image
Image("chart.png")
```

Biểu đồ Cột

**Chú thích:**

- `go.Bar()`: Dùng để vẽ biểu đồ cột. Trục x là các danh mục và trục y là các giá trị tương ứng.

7.3 Các Biểu Đồ Tương Tác

Plotly nổi bật với khả năng tạo các biểu đồ tương tác, nơi người dùng có thể zoom in/out, di chuyển, và hover để xem thông tin chi tiết.

7.3.1 Biểu đồ phân tán (Scatter Plot):

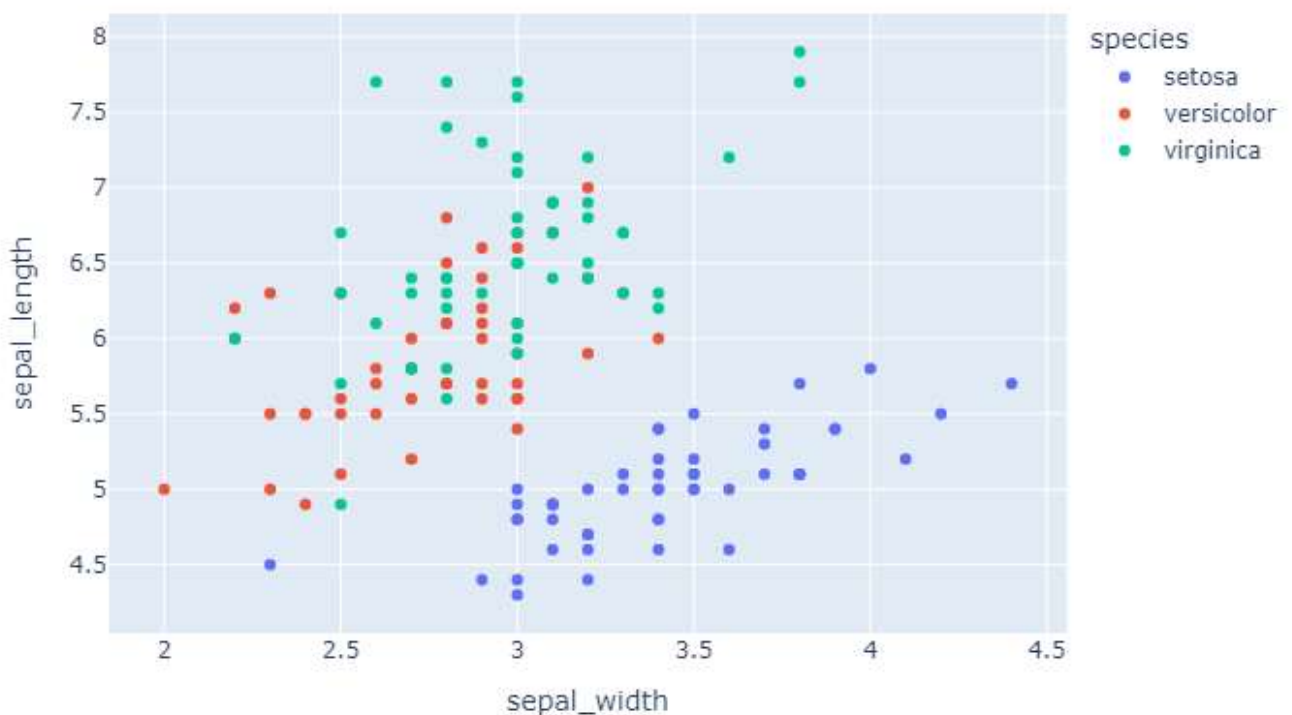
Biểu đồ phân tán giúp bạn tìm mối quan hệ giữa hai biến.

```
import plotly.express as px

df = px.data.iris()
fig = px.scatter(df, x="sepal_width", y="sepal_length", color="species",
                 title="Biểu đồ Phân tán")
fig.write_image("chart.png")

from IPython.display import Image
Image("chart.png")
```

Biểu đồ Phân tán

**Chú thích:**

- `px.scatter()`: Dùng để tạo biểu đồ phân tán, trong đó các điểm được tô màu theo từng nhóm (species).

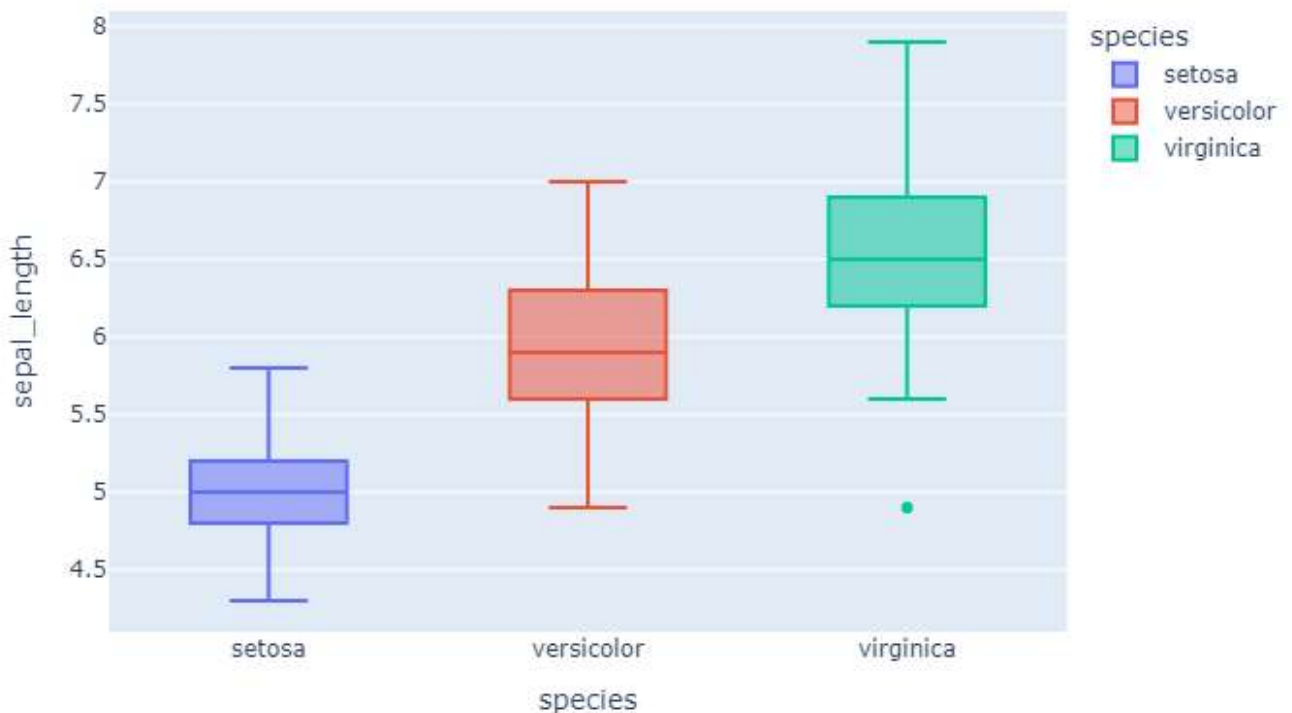
7.3.2 Biểu đồ hộp (Box Plot):

Biểu đồ hộp giúp bạn quan sát sự phân bố của dữ liệu và phát hiện các giá trị ngoại lai.

```
fig = px.box(df, x="species", y="sepal_length", color="species",
             title="Biểu đồ Hộp")
fig.write_image("chart.png")

from IPython.display import Image
Image("chart.png")
```

Biểu đồ Hộp

**Chú thích:**

- `px.box()`: Tạo biểu đồ hộp cho dữ liệu, cho thấy các phần tử cơ bản như trung vị, phần tư, và giá trị ngoại lai.

7.4 Tùy chỉnh Biểu Đồ

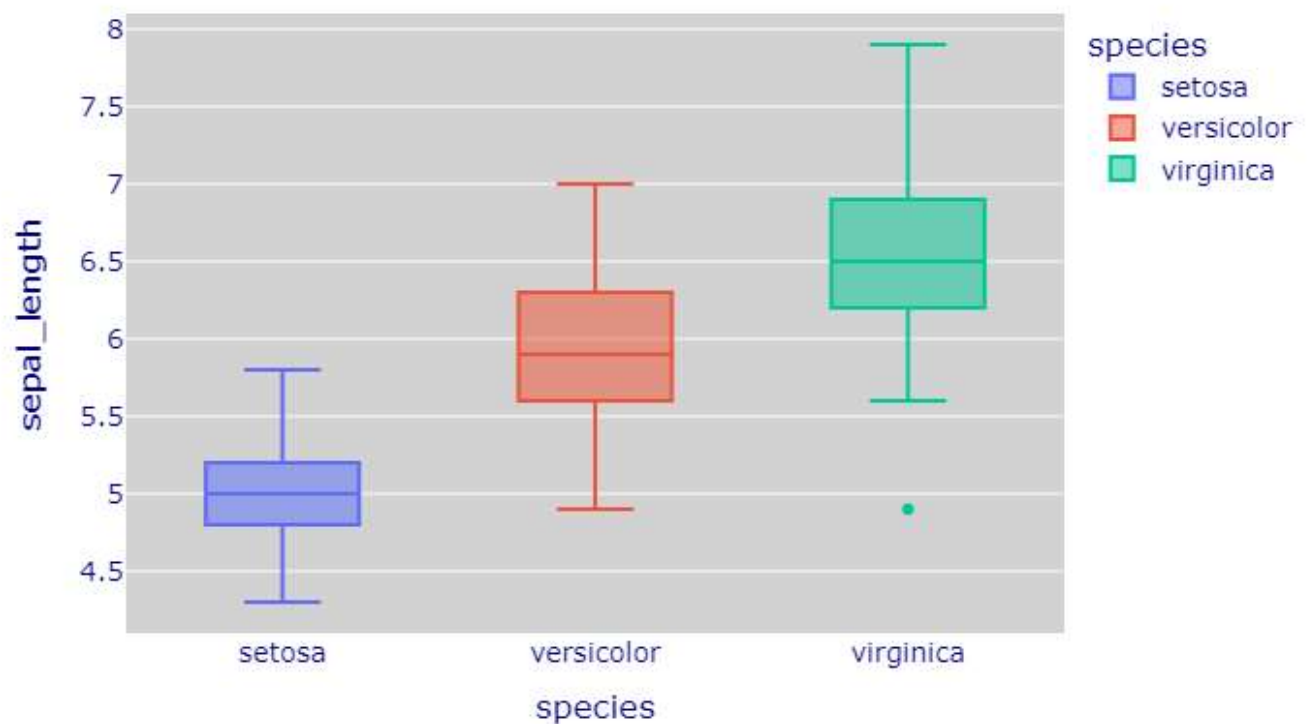
Thay đổi màu sắc, kích thước, và kiểu dáng

Plotly cho phép bạn dễ dàng tùy chỉnh các yếu tố của biểu đồ như màu sắc, kích thước, và kiểu dáng.

```
fig.update_layout(
    title="Tùy chỉnh Biểu đồ",
    font=dict(size=14, color="darkblue"),
    plot_bgcolor="lightgray"
)
fig.write_image("chart.png")

from IPython.display import Image
Image("chart.png")
```

Tùy chỉnh Biểu đồ



Chú thích:

- `update_layout()`: Tùy chỉnh giao diện của biểu đồ, bao gồm màu sắc nền, phông chữ, và tiêu đề.

7.5 Tạo Biểu Đồ Nâng Cao

Plotly cũng hỗ trợ các loại biểu đồ nâng cao, ví dụ như biểu đồ bản đồ và biểu đồ 3D.

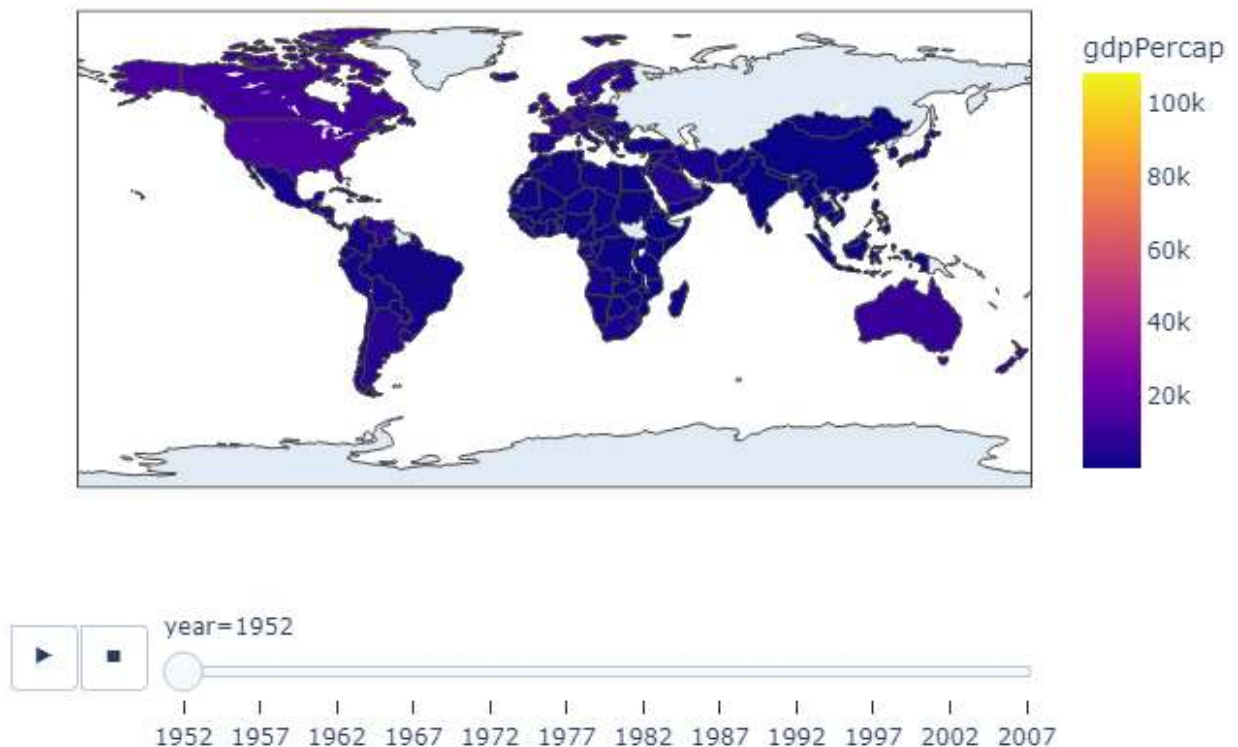
7.5.1 Biểu đồ bản đồ (Choropleth Map):

Biểu đồ bản đồ giúp bạn trực quan hóa dữ liệu theo vị trí địa lý.

```
fig = px.choropleth(px.data.gapminder(),
                    locations="iso_alpha",
                    color="gdpPercap",
                    hover_name="country",
                    animation_frame="year",
                    title="Biểu đồ Bản đồ")
fig.write_image("chart.png")

from IPython.display import Image
Image("chart.png")
```

Biểu đồ Bản đồ



Chú thích:

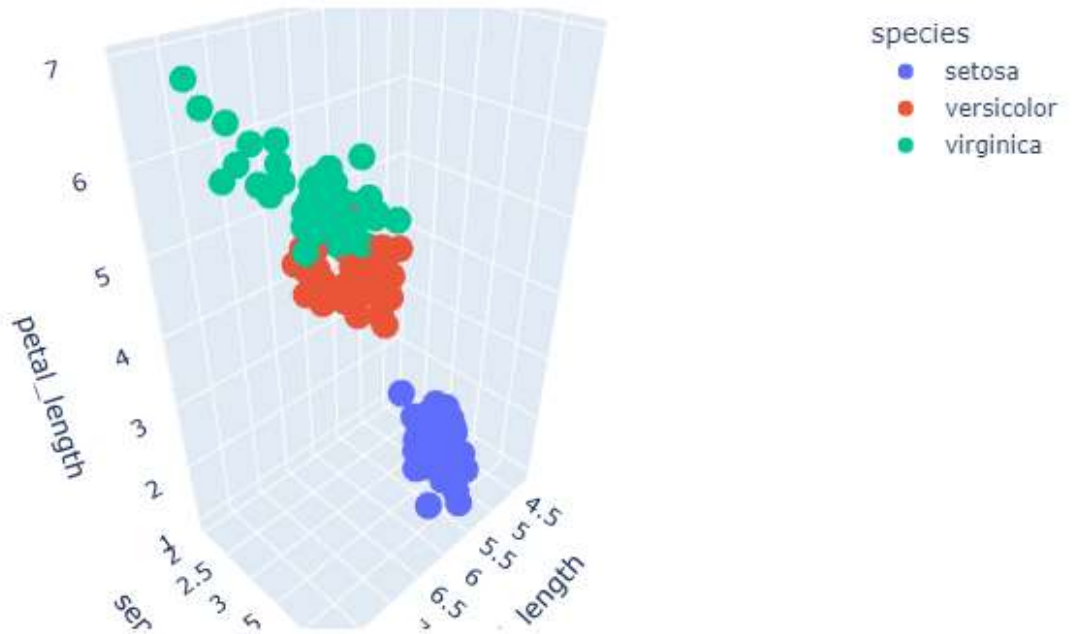
- `px.choropleth()`: Tạo biểu đồ bản đồ, trong đó các quốc gia được tô màu theo GDP.

7.5.2 Biểu đồ 3D (3D Scatter):

Biểu đồ 3D giúp bạn quan sát mối quan hệ giữa ba biến.


```
fig = px.scatter_3d(df, x="sepal_length", y="sepal_width", z="petal_length",  
                    color="species", title="Biểu đồ 3D")  
fig.write_image("chart.png")  
  
from IPython.display import Image  
Image("chart.png")
```

Biểu đồ 3D



Chú thích:

- `px.scatter_3d()`: Tạo biểu đồ phân tán 3D với ba trục không gian.

7.6 Xuất và Chia Sẻ Biểu Đồ

Lưu biểu đồ dưới dạng HTML:

- Lưu biểu đồ dưới dạng ảnh

```
fig.write_image("chart.png")
```

- Hiển thị trong Jupyter Book

```
from IPython.display import Image  
Image("chart.png")
```

7.7 Các Ví Dụ Thực Tế

- Phân tích dữ liệu doanh thu bán hàng.
- Theo dõi dữ liệu môi trường.
- Trực quan hóa dữ liệu thống kê trong lĩnh vực y tế.

7.8 Tài liệu tham khảo

- Liệt kê các tài liệu tham khảo:
- [Trang chủ Plotly](#)
- Tài liệu Python chính thức của Plotly.

7.9 Tổng kết chương 7

Nội dung: Trong chương 7 này chúng ta đã làm biểu đồ từ cơ bản đến nâng cao chỉ với những đoạn code đơn giản, trong chương 8 chúng ta sẽ đến các Kỹ thuật Trực quan hóa Nâng cao.