Visualizing Data with ggplot2: Takeaways

by Dataquest Labs, Inc. - All rights reserved © 2021

Syntax

Creating a plot using a DataFrame:

```
data %>%
ggplot()
```

Defining the variables you want to map to your graph:

```
data %>%

ggplot(aes(x = variable_1, y = variable_2))
```

Adding a line to your graph:

```
data %>%
  ggplot(aes(x = variable_1, y = variable_2))+
  geom_line()
```
```

Adding graph titles and axis labels:

```
data %>%

ggplot(aes(x = variable_1, y = variable_2)) +

geom_line() +

labs(

title = "Title of Graph",

x = "new x label",
```

```
y = "new y label"
)
...
```

Changing the background of a graph:

```
data %>%

ggplot(aes(x = variable_1, y = variable_2)) +

geom_line() +

labs(
 title = "Title of Graph",
 x = "new x label",
 y = "new y label"
) +

theme(
 panel.background = element_rect(fill = "background_color")
)

```
```

Concepts

- The ggplot2 package is a systematic, consistent, and efficient tool you can use to create high-quality visualizations.
- Plots are visual representations that use graphics like dots, lines, and bars to help you recognize patterns in data.
- A line chart is a type of plot that is especially useful for visualizing changes over time. A line chart displays information as a series of data points connected by a line. Line charts are useful for depicting data that is continuous.
- The gg in ggplot2 stands for Grammar of Graphics, which refers to a system for data visualization first described by Leland Wilkinson.
- The variable that changes depending on another variable is called the dependent variable. The dependent variable is usually placed on the y-axis.
- The variable that changes independent of another variable is called the independent variable. The independent variable is usually placed on the x-axis.
- We can add layers to the graph using the + character followed by another layer.

Resources

- Grammar of Graphics
- ggplot2 Documentation

Takeaways by Dataquest Labs, Inc. - All rights reserved $\ @ 2021$