

## Day 17: *Data Visualization in Python*

When Dot arrived in Tokyo, they breathed a sigh of relief. A hypermodern city, far from the frightening and stark nature of the mountain terrain. The ground was flat, and if they needed to get higher, they could take an escalator or elevator — no need to hike up and risk slipping to their death. Still, Dot worried about the odd woman on her own in an unforgiving, rocky landscape. From the Tokyo airport, they made their way into Shibuya square so they could people-watch and take in the city's bright lights. A giant TV screen against one of the skyscrapers caught their eye — there, looming over the milling crowds of the square, was an image of the woman! She was standing bundled in her snow gear, arms held up to the sky victoriously. Underneath was some text written in Japanese characters. Dot turned to the people next to them, speaking in English. “Excuse me,” they said cordially, “do you know what it says on the TV?” The group turned to the screen and then back to Dot. “It says this woman just set a record by climbing Mount Everest in less than a day,” the stranger said incredulously.

Dot felt regret at not tagging along with the woman — they, too, could have been a record-setter! But in Tokyo, they could reach great heights without expending so much energy or risking mortal injury. Dot made their way over to the Tokyo Skytree, one of the tallest buildings in the world. The elevator to the top seemed to take forever, and Dot’s stomach dropped as they got higher and higher and higher still. At the top, Dot stared out over Tokyo and took in the astonishing view. They wanted to meditate wistfully on existential questions about human life and the purpose of being, but they were distracted. Part of the problem was the crying child next to them, who wouldn’t be consoled by his father. “C’mon, Jimmy, cheer up,” the father said mournfully. He looked over at Dot apologetically. “I just can’t seem to get him to stop crying! He usually cheers up when I tell him all the facts I know about Toyota cars, but it’s not enough for him today. You don’t happen to know much about Toyota cars, do ya, stranger?” Dot thought for a moment. Can you help them quiet the child with some facts about Toyota vehicles?

## Tutorial

**Data visualization** refers to the graphical representation of information and data. By using visual elements such as charts and graphs, data visualization tools provide us with an accessible way to understand trends, outliers, and relationships in our data. It enables us to communicate data insights to others and understand the data better ourselves. Understanding how to use a data visualization tool is essential for anyone who uses data in their work.

One of the most popular visualization packages for Python is **Matplotlib**, a plotting library. Matplotlib is popular for how easy it makes building statistical plots, as well as for the excellent community support and the wide availability of tutorials online.

Read this [article](#) to increase your understanding of best practices in data visualization.

## Importing Matplotlib

Just like with pandas and other tools we have worked with so far, before we can use Matplotlib we have to import it. It's typical within data science to give Matplotlib the alias *plt* when importing it.

As a reminder, the code would look like this:

```
import matplotlib.pyplot as plt
```

### Basic Plot-Building in Matplotlibs

The process for building a plot in Matplotlib can be divided into three parts: the frame, the plot, and the stylistic features.

```
plt.figure() #The Frame: We start our plot with a figure  
plt.bar(x = data, height = data) #The Body: Declaring the specific bar plot  
plt.title("Example Bar Plot") #Stylistic Features: Adding the title  
plt.show() #To show our plot, we need to end our plot with a plt.show()
```

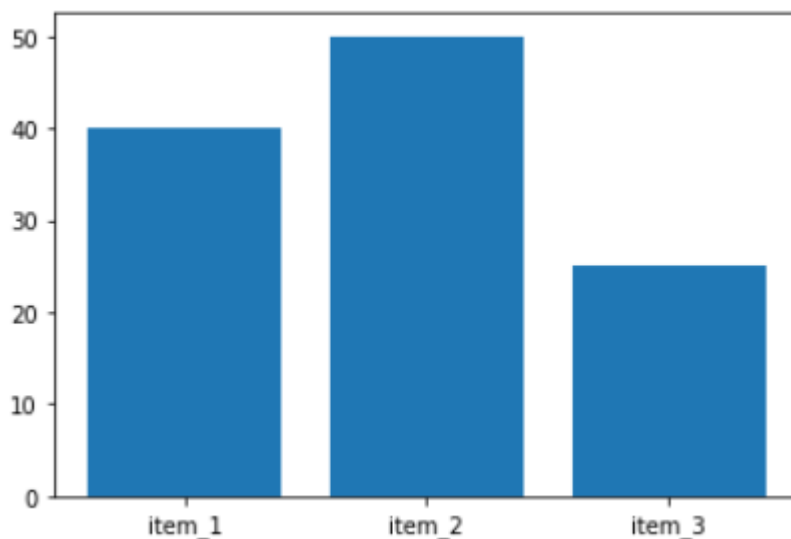
### Example

Below is example code that produces a simple bar chart. Try it out yourself!

```
data = {'item_1':40, 'item_2':50, 'item_3':25}  
items = list(data.keys())  
quantity = list(data.values())
```

```
plt.figure()  
plt.bar(x = items, height = quantity)  
plt.show()
```

output



### Pandas and Matplotlib

Pandas and Matplotlib can be integrated very easily. For example, we can feed specific pandas columns to Matplotlib as our data input.

To read more on the documentation for the bar chart in Matplotlib, click [here](#). For the matplotlib user guide, click [here](#).

## Challenge

Use the bar chart from matplotlib to answer the following question:

**What were the two most popular Toyota's car categories in 2020?**

```
In [2]: import pandas as pd
import matplotlib.pyplot as plt
```

```
In [3]: df = pd.read_csv('toyota_cars.csv')
```

```
In [4]: # SOLUTION

df = df[df.Year == 2020]
grouped_df = df.groupby('Category').size()
```

```
In [6]: plt.figure()
plt.bar(x = grouped_df.index, height = grouped_df.values)
plt.show()
```

