# Practise Creating Bar Charts in Python

This planning document is intended to support teachers who are delivering the NPA/PDA Data Science or for students who are learning independently. It also aligns with the Data Skills for Work framework.

## Contents

## Version Control

| Version number | Purpose/Change | By | Date |
|---|---|---|---|
| 1.0 | Published by Effini | John Bell | 31 May 2022 |
| | | | |
| | | | |
| | | | |
| | | | |

## Lesson Description

| | |
|---|---|
| **Lesson Overview** | Creation and modification of bar charts in Python using the seaborn package.<br><br>This lesson is intended to follow the **Creating Bar Charts in Python** and **Creating Other Graphs in Python** lessons.<br><br>This is a consolidation activity to give learners the chance to apply the data visualisation skills they have learned in these lessons.<br><br>Note that the accompanying Jupyter notebooks contain too many tasks for a single lesson; teachers are advised to select those tasks where learners would most benefit from consolidating their learning. |
| **Topic** | Visualisation and Storytelling |
| **Book Chapter(s)** | Interpreting data |

| | |
|---|---|
| **NPA level** | 5, 6 |
| **PDA level** | 7, 8 |
| **Data skills for work level** | Core, Analysis |

## Lesson Contents

This lesson consists of:

- A lesson plan (this document)
- A PowerPoint presentation, 'Practise Creating Bar Charts in Python'
- 2 Jupyter notebooks:
    - 'practise_creating_bar_charts.ipynb' (for learners)
    - 'practise_creating_bar_charts_with_answers.ipynb' (for teachers)

## Learning Intentions

We will be learning about **creating bar charts in Python,** specifically,

- creating **vertical and horizontal** bar charts
- displaying **chart titles** and **axis titles**

- changing the **colours of bars**
- changing the **tick mark labels/intervals** used in the axes, and
- changing the **order of bars**

## Success Criteria

I can *create* vertical and horizontal bar charts
I can *create* titles and axis titles
I can *modify* the colours of bars
I can *modify* the tick mark labels/intervals used in axes
I can *modify* the order of the bars

## Knowledge Prerequisites

Learners should know:

- Python programming to at least the level defined in SQA Computer Programming Level 5 (HY2C 45)
- How to use a Jupyter notebook to write, edit and run Python code
- How to create and modify bar charts in Python using the seaborn package, as covered in **Creating Bar Charts in Python**
- How to modify tick mark labels/intervals, as covered in **Creating Other Graphs in Python**

# Lesson Requirements

| | PDA | NPA | Data Skills for work |
|---|---|---|---|
| **Qualification** | Yes | Yes | Yes |
| **Outcome ID(s)** | CD7.3e, CD8.2h | DC5.3c, DC6.3d | c2.1, a2.1, a2.2, a3.1, a3.2 |
| **Outcome description(s)** | CD7.3e Creating visualisations using software<br><br>CD8.2h Creating visualisations using software | DC5.3c Create appropriate visualisations from data.<br><br>DC6.3d Create appropriate visualisations from data. | c2.1 Vocabulary used in data science and analytics<br><br>a2.1 Use of tools to analyse data<br><br>a2.2 Selecting suitable visualisations<br><br>a3.1 Visualisation of data to provide insight<br><br>a3.2 Visualisation of data to tell stories |
| **Level** | 7, 8 | 5, 6 | Core, Analysis |
| **Software language** | Python | Python | Python |
| **Required equipment /software for student** | Lesson: PowerPoint<br><br>Python notebook: Jupyter notebook environment | Lesson: PowerPoint<br><br>Python notebook: Jupyter notebook environment | Lesson: PowerPoint<br><br>Python notebook: Jupyter notebook environment |

# Jupyter Notebook

There is a Jupyter notebook for this lesson that provides examples and programming tasks for learners, drawn from the examples in the lesson PowerPoint.
The notebook uses Python 3.x and the following packages:

- numpy – for scientific computing
- pandas - for data manipulation
- s3fs - an API to AWS S3 (Simple Storage Service), used to import datasets
- seaborn – for data visualisation
- matplotlib – for data visualisation

The tasks are described in the table below.

| Notebook section | Task | Description |
|---|---|---|
| Creating Vertical and Horizontal Bar Charts | Task 1 - Loch Length | Create a bar chart showing the length of each loch in a small, simple dataset. |
| | Task 2 - Population of continents (2007) | Create a bar chart showing the population of each continent in 2007. The code required for aggregation is provided. |
| | Task 3 - Fraud clear up rates in 2019-2020 | Create a bar chart showing, for each Local Authority in Scotland, the percentage of reported frauds that were cleared up in 2019-2020. |
| Changing Tick Mark Labels and Intervals | Task 4 - How many tickets did each movie sell? | Create a horizontal bar chart using the **fashion** data frame, that shows the number of transactions for each product (e.g. jacket or gloves). |
| | Task 5 - Have more women than men been awarded the Nobel Peace Prize? | Create a bar chart that shows how many tickets each movie sold. Increase the size of the plot to ensure bar labels are legible. |
| | Extension Task 1 - Has the situation changed? | Recreate the bar chart from **Task 5** but this time filter the dataset to include more recent years. |
| Creating Bar Charts for Data that | Task 6 - Average crime clear up rates for all years | Create a bar chart that displays the mean clear up rate for each Local Authority in Scotland. |

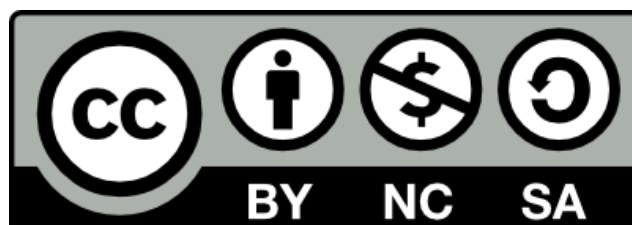| | | |
|---|---|---|
| Needs Aggregated | Task 7 - Which crime had the highest clear up rate? | Create a bar chart showing the average mean clear up rate for each crime across Scotland in the years 2017 to 2019. |
| | Task 8 - How far did Skyler go? | Create a bar chart that shows the total distance a single athlete in a Strava dataset has travelled for each activity they undertake. Order the bars alphabetically. Order the bars alphabetically. |
| Ordering Bars | Task 9 - On which continent were most of the Nobel peace prize winners between 2000 and 2016 born? | Create a bar chart to answer this question. Order the bars alphabetically. |
| | Task 10 - The Longest Run | Create a bar chart to discover which of the athletes in the Strava data frame ran for the longest time (in a single run), as measured by the amount of time they were moving. Order the bars by value. Colour the bars using a Strava branding colour. Requires consideration of the correct variable to use. |
| | Task 11 - In which country were most female Nobel Peace Prize winners born in? | Create a bar chart to answer this question. Order the bars by value. Modify the tick mark intervals to an appropriate value. |

## Datasets

The following datasets are used in this lesson.

| Dataset name | Description | Link |
|---|---|---|
| lochs | the volume, length and maximum depth of some Scottish lochs. | https://datasets.learn-data.science/lochs.csv |
| strava_activities_simple | running and cycling activity data for 2 Strava users. | https://datasets.learn-data.science/strava_activities_simple.csv |
| nobel_peace_prize_winners | all winners of the Nobel Peace Prize from 1901 to 2016. | https://datasets.learn-data.science/nobel_peace_prize_winners.csv |
| marvel_movies | budget and revenue figures for recent Marvel movies. | https://datasets.learn-data.science/marvel_movies.csv |
| highest_grossing_usa_movies_1995_2021 | for each year between 1995 and 2021, information about the movie which made the most money that year at the U.S. box office. | https://datasets.learn-data.science/highest_grossing_usa_movies_1995_2021.csv |
| police_scotland_clearup_rates | crimes cleared up by Police Scotland as a percentage of those recorded. Includes clear up rates for robbery, fraud and housebreaking between 2017 and 2019. | https://datasets.learn-data.science/police_scotland_clearup_rates.csv |
| gapminder | Life Expectancy, population and GDP per capita for each country, for every 5 years from 1952 to 2007. | https://ed-uni-data-lessons.s3.eu-west-2.amazonaws.com/data/gapminder.csv |

## How you can use this lesson

## Alternative format

**If you require this document in an alternative format, such as large print or a coloured background, please contact**
**hello@effini.com**
**or**
**4th Floor, The Bayes Centre**
**47 Potterrow**
**Edinburgh**
**EH8 9BT**