



TASK

Capstone Project IV

Visit our website

Introduction

WELCOME TO THE DATA VISUALISATION CAPSTONE PROJECT!

Great job! You have reached another capstone project in this bootcamp. For this task, we will be applying both our data analytics and visualisation skills to a real-life dataset.



Get in touch
Connect for support

Remember that with our courses, you're not alone! You can contact an expert code reviewer to get support on any aspect of your course.

The best way to get help is to login to Discord at <https://discord.com/invite/hyperdev> where our specialist team is ready to support you.

Our team is happy to offer you support that is tailored to your individual career or education needs. Do not hesitate to ask a question or for additional support!



As you have already learnt, as a Data Scientist you need to be able to display a very important set of skills. You need to be able to write code that can be used to clean, organise, analyse and visualise data. These coding and analytical skills need to be combined with the ability to communicate your findings to stakeholders effectively.

In this task, you will be given the opportunity to showcase all the skills which you have been developing throughout this bootcamp. You will be tasked with analysing a dataset of your choice. Furthermore, you will create a report that contains several data visualisations that will clearly communicate your findings. Finally, you will present your report for assessment.

You are free to choose any dataset of your choice to analyse. Here are some sources of datasets:

- [United Nations datasets](#)
- [South African Open Data](#)
- [World Bank Data](#)

Your report should include the following:

- A brief description of the dataset you have chosen
- A description of how your dataset was cleaned
- A description of how you managed missing data
- A detailed description of your findings accompanied by appropriate visualisations
- A wide selection of appropriate data visualisations. Besides Matplotlib, we encourage you to explore and use libraries with more interesting visualisations

A template for the report (**EDA_doc.odt**) accompanies this task.

Ensure that your report is professional! Remember, such a report can be added in your professional portfolio.

Capstone Project

- Select a dataset. **Note:** Once you've selected a dataset please make sure to check it with a reviewer before starting.
- You can choose from the following datasets contained in this folder:
 - Diabetes
 - Wine
 - Hourly Wage Data

- Penguins
- COVID-19 Clinical Trials
- Forbes Highest Paid Athletes 1990-2020
- Anything from the following sources:

- [United Nations datasets](#)
- [South African Open Data](#)
- [World Bank Data](#)

- Create the report as described in this task.
- When you submit your report for review, also book a call with a reviewer so that you can present your report and findings to them.

If you are having any difficulties, please feel free to contact our specialist team [on Discord](#) for support.

Completed the task(s)?

Ask an expert to review your work!

[Review work](#)



Rate us

Share your thoughts

HyperionDev strives to provide internationally-excellent course content that helps you achieve your learning outcomes.

Think that the content of this task, or this course as a whole, can be improved or think we've done a good job?

[Click here](#) to share your thoughts anonymously.

