



Welcome to this session
Skills Bootcamp:

Exploratory Data Analysis
Practical

The session will start shortly...

Questions? Drop them in the chat.
We'll have dedicated moderators
answering questions.



Skills Bootcamp Data Science Housekeeping

- The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly. **(Fundamental British Values: Mutual Respect and Tolerance)**
- No question is daft or silly - **ask them!**
- There are **Q&A sessions** midway and at the end of the session, should you wish to ask any follow-up questions. We will be answering questions as the session progresses as well.
- If you have any questions outside of this lecture, or that are not answered during this lecture, please do submit these for upcoming Academic Sessions. You can submit these questions here: **Questions**

Skills Bootcamp Data Science Housekeeping

- For all **non-academic questions**, please submit a query: www.hyperiondev.com/support
- Report a safeguarding incident: www.hyperiondev.com/safeguardreporting
- We would love your feedback on lectures: [Feedback on Lectures.](#)
- Find all the lecture **content** in your [Lecture Backpack](#) on GitHub.
- If you are hearing impaired, kindly use your computer's function through Google chrome to enable captions.

Safeguarding & Welfare

We are committed to all our students and staff feeling safe and happy; we want to make sure there is always someone you can turn to if you are worried about anything.

If you are feeling upset or unsafe, are worried about a friend, student or family member, or you feel like something isn't right, speak to our safeguarding team:



Ian Wyles
Designated Safeguarding
Lead



Simone Botes



Nurhaan Snyman



Rafiq Manan



Ronald Munodawafa



Tevin Pitts

Scan to report a
safeguarding concern



or email the Designated
Safeguarding Lead:
Ian Wyles
safeguarding@hyperiondev.com

Skills Bootcamp Progression Overview

✓ Criterion 1 - Initial Requirements

Specific achievements **within the first two weeks** of the program.

To meet this criterion, students need to, by no later than **01 December 2024 (C11)** or **22 December 2024 (C12)**:

- **Guided Learning Hours (GLH):** Attend a **minimum of 7-8 GLH per week** (lectures, workshops, or mentor calls) for a total minimum of **15 GLH**.
- **Task Completion:** Successfully complete the **first 4 of the assigned tasks**.

✓ Criterion 2 - Mid-Course Progress

Progress through the successful completion of tasks **within the first half** of the program.

To meet this criterion, students should, by no later than **12 January 2025 (C11)** or **02 February 2025 (C12)**:

- **Guided Learning Hours (GLH):** Complete at least **60 GLH**.
- **Task Completion :** Successfully complete the **first 13 of the assigned tasks**.

Skills Bootcamp Progression Overview

✓ Criterion 3 – End-Course Progress

Showcasing students' progress nearing the completion of the course.

To meet this criterion, students should:

- **Guided Learning Hours (GLH):** Complete the **total minimum required GLH**, by the **support end date**.
- **Task Completion : Complete all mandatory tasks**, including any necessary resubmissions, by the end of the bootcamp, **09 March 2025 (C11)** or **30 March 2025 (C12)**.

✓ Criterion 4 - Employability

Demonstrating progress to find employment.

To meet this criterion, students should:

- **Record an Interview Invite:** Students are required to record proof of invitation to an interview by **30 March 2025 (C11)** or **04 May 2025 (C12)**.
 - **South Holland Students** are required to proof and interview by **17 March 2025**.
- **Record a Final Job Outcome :** Within 12 weeks post-graduation, students are required to record a job outcome.

Learning Outcomes

- Load and explore a dataset using Python.
- Perform univariate, bivariate, and multivariate EDA.
- Generate and interpret visualisations for data insights.
- Identify key patterns and trends in the Iris dataset and transfer this knowledge to other datasets they explore.

Lecture Overview

- This lecture will focus on code examples demonstrating the theoretical concepts covered in the first lecture.

Initial Poll Assessment





How well did you understand the EDA concepts covered on Monday?

1. Very well – everything's clear and I'm ready for the practical.
2. I understood almost everything, but I'm hoping the practical will clarify things for me.
3. I'm struggling to understand EDA.
4. I did not attend Monday's lecture.



Let's Code!



CoGrammar

Final Poll Assessment



Which plot is best to understand the relationship between two numerical variables?

1. Scatter plot
2. Histogram
3. Box plot

Which plot is best to understand the relationship between two numerical variables?

1. **Scatter plot**
2. Histogram
3. Box plot

What does a correlation coefficient of 1 indicate?

1. No correlation
2. Perfect positive correlation
3. Perfect negative correlation

What does a correlation coefficient of 1 indicate?

1. No correlation
2. **Perfect positive correlation**
3. Perfect negative correlation



What is the purpose of PCA?

1. To remove missing values
2. To make data more variable for better representation
3. To cluster data
4. To reduce dimensionality while keeping key information



What is the purpose of PCA?

1. To remove missing values
2. To make data more variable for better representation
3. To cluster data
4. **To reduce dimensionality while keeping key information**

Questions and Answers



Thank you for attending



CoGrammar



Department
for Education