### **COHORT 4 - 2025**

### (Please copy this document and rename it using your official name)

### Mentee’s Name: Joshua Kimani

Mentor’s Name:

Pair’s Name:

DataCamp Profile: https://www.datacamp.com/portfolio/joshuakimani192

### [Calendar](https://drive.google.com/file/d/15QS2A5elP5wK_NRtASwbYxFfMVtrxdd7/view?usp=sharing)

Join the weekly meeting [here](https://meet.google.com/ist-izkn-qeo)

**Program Expectations.**

1. Attend all onboarding meetings.
2. Respond to all forms as requested by your mentor/ coordinator.
3. Attend Tuesday and Thursday meetings at 7.30 p.m.
4. Attend the Mondays spaces on X under #DataMondays.
5. Attend Expert and Feedback sessions as scheduled.
6. Attend our Physical Meet & Greet on Saturday July 19th, 2025.
7. Collaborate with your assigned peer in handling assessments.
8. Prepare and present your individual project on Week 12.
9. Attend your graduation ceremony on Saturday October 18th, 2025.
10. Update your workbook as you learn and upload it on Everything Data drive as instructed by your mentor.

### 

**CONTENTS**

**Month 1 - 41 hrs**

A. Week 1 Assessment ( 1,000 Experience Points )

B. Week 2

C. Week 3

D. Week 4

**Month 2 - 32 hrs**

A. Week 5

B. Week 6

C. Week 7

D. Week 8

**Month 3** **- 13 hrs**

A. Week 9

B. Week 10

C. Week 11

D. Week 12 PROJECT PRESENTATION

**Month 4**

A. Week 13 GitHub - 10 hrs

B. Week 14 Certification Assessment ( Data Scientist Associate )

C. Week 15

D. Week 16

**Additional Resources**

### 

### **[Assessment 1](https://assessment-v2.datacamp.com/exploratory-analysis-python?returnUrl=https://app.datacamp.com/learn/assessments) - Data Manipulation with Python**

### **[Assessment 2](https://assessment-v2.datacamp.com/statistical-experimentation-python?returnUrl=https://app.datacamp.com/learn/assessments) - Statistics Fundamentals with Python**

### **[Assessment 3](https://assessment-v2.datacamp.com/exploratory-analysis-theory-learn) - Exploratory Analysis Theory**

### **[Assessment 4](https://assessment-v2.datacamp.com/statistical-experimentation-theory-learn) - Statistical Experimentation Theory**

### 

### 

### **Month 1: Python, Pandas & Data Visualization**

**Week 1:** [**Introduction to Python**](https://app.datacamp.com/learn/courses/intro-to-python-for-data-science)

[**Intermediate Python**](https://app.datacamp.com/learn/courses/intermediate-python)

[**Guided Project**](https://app.datacamp.com/learn/projects/investigating_netflix/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:

* Understand basic Python syntax and programming concepts.
* Work with variables, data types, operators, and control structures (loops, conditionals).
* Write functions and understand scope and modules in Python.

**Activities:**

* **Practice Exercises:** Solve basic programming problems, such as list manipulation, string formatting, and simple functions.
* **Assignment:** Build a simple Python program to process and analyze a text file (e.g., word count).

**Week 2:** [**Introduction to Importing Data in Python**](https://app.datacamp.com/learn/courses/introduction-to-importing-data-in-python)

[**Cleaning Data with Python**](https://app.datacamp.com/learn/courses/cleaning-data-in-python)

[**Data Manipulation with Pandas**](https://app.datacamp.com/learn/courses/data-manipulation-with-pandas)

[**Guided project**](https://app.datacamp.com/learn/projects/exploring_nyc_public_school_test_result_scores/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:

* Create and manipulate DataFrames and Series in pandas.
* Import data from various formats (CSV, Excel) into pandas.
* Conduct basic data cleaning, such as removing duplicates and handling missing values.

**Activities:**

* **Practice Exercises:** Import sample datasets and perform basic manipulations.
* **Assignment:** Analyze a small dataset (e.g., customer records) to practice loading, cleaning, and basic data manipulation.

**Assessment:** Quiz on Python basics and introductory pandas.

**Week 3:** [**Guided project**](https://app.datacamp.com/learn/projects/exploring-airbnb-market-trends/guided/Python)

[**Introduction to Functions in Python**](https://app.datacamp.com/learn/courses/introduction-to-functions-in-python)

[**Python Toolbox**](https://app.datacamp.com/learn/courses/python-toolbox)

[**Joining Data with Pandas**](https://app.datacamp.com/learn/courses/joining-data-with-pandas) **Learning Outcomes:** By the end of this week, you should be able to:

* Apply filtering, sorting, and grouping methods on DataFrames.
* Perform advanced operations like merging, joining, and reshaping data.

**Activities:**

* **Practice Exercises:** Merge multiple datasets and apply group by for summary statistics.
* **Assignment:** Analyze a multi-table dataset, applying joins and aggregations to answer specific questions.

**Week 4:** [**Data Communication Concepts**](https://app.datacamp.com/learn/courses/data-communication-concepts)

[**Introduction to Data Visualization - Matplotlib**](https://app.datacamp.com/learn/courses/introduction-to-data-visualization-with-matplotlib)

[**Introduction to Data Visualization - Seaborn**](https://app.datacamp.com/learn/courses/introduction-to-data-visualization-with-seaborn)

[**Guided project**](https://app.datacamp.com/learn/projects/visualizing_the_history_of_nobel_prize_winners/guided/Python) **Learning Outcomes:** By the end of this week, you should be able to:

* Conduct data cleaning, including handling outliers and advanced transformations.
* Create basic visualizations (e.g., histograms, bar charts, scatter plots) using matplotlib and seaborn.
* Interpret visual data representations to enhance exploratory data analysis (EDA).

**Activities:**

* **Practice Exercises:** Visualize various relationships and distributions within sample datasets.

**GitHub Link:**

**How would you use Python in your daily life or business?**

# 

# **Month 2: Statistics for Data Science**

## **Week 5:** [**Exploratory Data Analysis in Python**](https://app.datacamp.com/learn/courses/exploratory-data-analysis-in-python)

## [**Guided Project**](https://app.datacamp.com/learn/projects/1876)

## [**Introduction to Statistics in Python**](https://app.datacamp.com/learn/courses/introduction-to-statistics-in-python)

**Learning Outcomes:** By the end of this week, you should be able to:  
 - Understand measures of central tendency (mean, median, mode).  
 - Calculate measures of dispersion (range, variance, standard deviation).  
 - Summarize data distributions using statistical tools.

**Activities:**  
 - Practice Exercises: Calculate descriptive statistics for various datasets.

## **Week 6:** [**Working with Dates and Times in Python**](https://app.datacamp.com/learn/courses/working-with-dates-and-times-in-python)

## [**Working with Categorical Data in Python**](https://app.datacamp.com/learn/courses/working-with-categorical-data-in-python)

## [**Guided Project**](https://app.datacamp.com/learn/projects/customer_analytics_preparing_data_for_modeling/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:

- Work with datetime objects and perform time-based calculations.

- Handle and preprocess categorical variables effectively.

## **Week 7:** [**Introduction to Regression with statmodels**](https://app.datacamp.com/learn/courses/introduction-to-regression-with-statsmodels-in-python)

## [**Sampling in Python**](https://app.datacamp.com/learn/courses/sampling-in-python)

## [**Guided Project**](https://app.datacamp.com/learn/projects/modeling_car_insurance_claim_outcomes/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:

- Conduct hypothesis testing (e.g., t-tests, chi-square tests).

- Calculate and interpret confidence intervals.

- Understand sampling methods and their impact on data analysis.

## **Week 8:** [**Hypothesis Testing in Python**](https://app.datacamp.com/learn/courses/hypothesis-testing-in-python)

## [**Experimental Design in Python**](https://app.datacamp.com/learn/courses/experimental-design-in-python)

[**Guided Project**](https://app.datacamp.com/learn/projects/hypothesis_testing_with_mens_and_womens_soccer_matches/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:  
- Calculate and interpret correlation coefficients.

- Understand and apply simple linear regression using statsmodels.

- Design simple experiments and understand key experimental design principles.

**How would you use Statistics in your daily life or business?**

**Month 3: Machine Learning**

## **Week 9:** [**Supervised Learning**](https://app.datacamp.com/learn/courses/supervised-learning-with-scikit-learn)

[**Guided project**](https://app.datacamp.com/learn/projects/1772)

**Learning Outcomes:** By the end of this week, you should be able to:  
- Understand regression models such as linear and polynomial regression.

- Build and evaluate regression models using scikit-learn.

- Interpret model results using metrics like RMSE and MAE.

## **Week 10:** [**Unsupervised Learning**](https://app.datacamp.com/learn/courses/unsupervised-learning-in-python)

[**Guided project**](https://app.datacamp.com/learn/projects/1809)

**Learning Outcomes:** By the end of this week, you should be able to:  
 - Understand clustering algorithms such as k-means and hierarchical clustering.

- Apply clustering techniques for exploratory data analysis.

- Interpret clustering results and understand their practical applications.

## **Week 11:** [**Machine Learning with tree-based models**](https://app.datacamp.com/learn/courses/machine-learning-with-tree-based-models-in-python)

[**Guided project**](https://app.datacamp.com/learn/projects/predicting-movie-rental-durations/guided/Python)

**Learning Outcomes:** By the end of this week, you should be able to:  
 - Understand the fundamentals of decision trees and ensemble methods like random forests.

- Implement and evaluate tree-based models using scikit-learn.

- Analyze model results and feature importance for better interpretability.

## **Week 12: PROJECT PRESENTATIONS**

### 

### 

### 

### 

**Month 4: Certification & GitHub**

**Week 13:** [**GitHub for Data Analysts**](https://app.datacamp.com/learn/skill-tracks/github-foundations)

By the end of this week, you should be able to;

* Understand the basics of **Git** version control.
* Create and manage a repository on **GitHub**.
* Understand the benefits of version control for data analysis projects

**GitHub Link:**

**Week 14/15: Certification (Data Analyst Associate)**

By the end of this week, you should be able to;

* Complete your certification assessment

**Week 16: CV and Portfolio Advancement**

By the end of this week, you should be able to;

* Reflect on the entire learning process and evaluate your growth.
* Apply the lessons from the capstone to future projects.
* Revamp your CV and add your new skills.
* Create/advance your portfolio.

**Updated CV Link:**

**What are your key take-aways from this program?**

### **Additional Resources:**

* **Recommended Readings**: [Joel\_Grus\_Data\_Science\_from\_Scratch\_First\_Princ.pdf](https://drive.google.com/file/d/1kU6OHdQcnP81Tv1s-tvz__Z7ji6JuTEt/view?usp=sharing) [Python\_Datascience.pdf](https://drive.google.com/file/d/14gjiE6twhDURQzgyLVP8jCEsz1_xjDDz/view?usp=sharing)[Ng\_MachineLearningYearning.pdf](https://drive.google.com/file/d/1I8UBJ9690rZis5t_m9ijlpI8Bt8nWUnV/view?usp=sharing)

[Alex the Analyst](https://www.analystbuilder.com/?via=ED)

* **Discussion Forums**: [Everything Data Community](https://chat.whatsapp.com/BtjJzn6ZgSM1HckRu1bT6n)