

DEPARTMENT: PURE AND APPLIED SCIENCES**PROGRAMME:** BACHELOR OF SCIENCE IN APPLIED STATISTICS**YEAR:** 4**SEMESTER:** II**UNIT CODE:** SPS 2471 **UNIT TITLE:** STATISTICAL COMPUTING III**LECTURE HOURS:** 45 Hrs.**CREDIT HOURS:** 45Hrs**PRE-REQUISITES:** STATISTICAL COMPUTING II**LECTURER:** KAROMO J.N**LECTURER CONTACTS:** EMAIL jkaromo@kyu.ac.ke **Mobile phone:** 0750 903 935**Purpose**

To equip students with the skills to enable them to perform basic statistical analysis using standard statistical software and interpret the results

To equip students with the skills to enable them analyse data using Python programming language.

Objectives

By the end of this course the student should be able to;

1. Perform different programming techniques using Python Programming Language.
2. Load data to Python and perform the analysis.
3. Understand the concept of Data Science.
4. To use the Google Colab, Jupyter Notebook or Anaconda Distributed Environment to perform data analysis.
5. Develop Data Science Regression and Classification models.

Course work plan

Week	Topic	Sub- Topic	Remark
1	Introduction to Python programming	<ul style="list-style-type: none"> • Python installation • Python IDE (eg PyCharm) • Creation of new projects • Version control 	
2	Python - Data types	<ul style="list-style-type: none"> • Integers, Float, Boolean Strings 	
	Python - Data Structures and their methods.	<ul style="list-style-type: none"> • Lists, Tuples, Arrays • Dictionaries • Data frames 	
3	Python - Operators	<ul style="list-style-type: none"> • Arithmetic operators 	

		<ul style="list-style-type: none"> • Relational operators • Logical operators 	
4	Python - Control Structures	<ul style="list-style-type: none"> • If else statement • Nested if else statement 	
5	CAT I	<ul style="list-style-type: none"> • A class sitting CAT I 	
6	Python - Loops	<ul style="list-style-type: none"> • For loop • While loop • Loops with control structures or sequence control. 	
7	Python - Loops	<ul style="list-style-type: none"> • Break statement in Python Loops • Continue statement in Python loops 	
8	Python - Functions	<ul style="list-style-type: none"> • Creating Python functions • Parameterised functions • Calling and applying the user built functions. 	
9	Introduction to Data Science	<ul style="list-style-type: none"> • General overview of data Science. • Key Libraries (NumPy, Scipy, Pandas, Matplotlib etc.) 	
10	Data Science Coding Platforms	<ul style="list-style-type: none"> • Google Colab • Jupiter Notebook • Anaconda Distributed environment. 	
11	CAT II	<ul style="list-style-type: none"> • Practical CAT II (Student to do a random Python question based on the course outline – 10 minutes for each student) 	
12	Data Science Process and Concepts.	<ul style="list-style-type: none"> • Data wrangling process and steps. 	
13	Machine Learning Algorithms - Regression	<ul style="list-style-type: none"> • Fitting, Training and testing ML models 	
14	Machine Learning Algorithms - Classification	<ul style="list-style-type: none"> • Fitting, Training and testing ML models 	
15/16	EXAMINATION	<ul style="list-style-type: none"> • Revision and Exam 	

Teaching Methods

1. Lecture: Oral presentation generally incorporating additional activities, e.g. writing on a white-board, exercises, class questions and discussions, or student presentations.
2. Practical: Computer Lab session as a means of further actively involving students to improve the skills.
3. Tutorial: To give the students more attention.

Instructional Material/Equipment

Include course notes, white-board, white-board marker, duster, computer and projector.

Assessment

1. Continuous Assessment Tests within the semester comprising 30% of the total marks (Tests 15%, Practical 10%, Assignment 5%)
2. Written end of semester Examination comprising 70% of the total marks

Course Text Books

1. Sheppard, K. (2014). Introduction to python for econometrics, statistics and data analysis.
2. McKinney, W. (2011). pandas: a foundational Python library for data analysis and statistics. Python for high performance and scientific computing, 14(9), 1-9.
3. Raschka, S. (2015). Python machine learning. Packt publishing ltd.