

**Computer Science Department****CS 504 – Programming Languages for Data Analysis****Course outline****Winter 2021**

---

**Professor:**

Dr. Mohammed Ayoub Alaoui Mhamdi

Email: Mohammed.Ayoub.Alaoui.Mhamdi@ubishops.ca

Office: Bishop's University, POL 314

Office hours: Online via Microsoft Teams

---

**Schedule:**

Lecture:

Online and Asynchronous

---

**Description**

Objective: To use different programming languages for data analysis.

Content In this course, students will be introduced to the most popular languages and software environments used in statistical computing and visualization. The course will involve significant programming projects in Julia, Python, R, Octave, Scala, and some opensource software for data analysis<sup>1</sup>.

Credits 3

Organization Three hours of lecture per week  
Six hours of personal work per week.

**1 Topics**

<sup>1</sup> <https://www.predictiveanalyticstoday.com/top-data-analysis-software/>

The topics of this course are the following:

1. Introduction
2. Basic probability
3. Probability distributions
4. Processing and Summarizing Data
5. Statistical Inference Concepts
6. Confidence Intervals
7. Hypothesis Testing
8. Linear Regression and Extensions
9. Machine Learning Basics
10. Simulation of Dynamic Models

## **2 Organization**

### **2.1 Teaching method**

A week includes three hours of lectures. Most online presentations will be done using slides available on Moodle. Throughout the semester, the student will implement the concepts seen in class through five assignments, a final project, and a final exam.

### **2.2 Evaluation**

Assignments: 40%

Final project: 20%

Final exam: 40%

Unless otherwise stated, the assignments can be done in teams of three students. Special instructions will be given for each assignment. The submission of assignment must be performed through Moodle. Subjects will be available on Moodle. The delivery of work is done before midnight of the due day. Failure to submit an assignment before the deadline will result in a penalty of 10% per day late, including Saturday and Sunday.

No assignment can be submitted by email. Plagiarized or undelivered assignment will be automatically attributed the grade zero.

### 3 Course Material

The course's slides will be available on Moodle.

No manual is required.

### 4 Bibliography

- [1] McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. In O'Reilly Media, Inc.
- [2] Lortie, C. (2017). R for Data Science. Journal of Statistical Software.  
<https://doi.org/10.18637/jss.v077.b01>
- [3] Guller, M., & Guller, M. (2015). Programming in Scala. In Big Data Analytics with Spark.  
[https://doi.org/10.1007/978-1-4842-0964-6\\_2](https://doi.org/10.1007/978-1-4842-0964-6_2)
- [4] Linge, S., & Langtangen, H. P. (2016). Programming for Computations - MATLAB/Octave. In Springer.
- [5] Helmreich, J. E. (2020). Data Science with Julia. Journal of Statistical Software.  
<https://doi.org/10.18637/jss.v094.b01>