



IB/CS 385 INTRODUCTION TO BUSINESS PROGRAMMING

IES Abroad Milan

DESCRIPTION: This course enables students to acquire basic skills in one of the most fundamental skills of the present period: computer programming. The course introduces students to data analysis programs with the aim of growing their skills in using those programs to predict future outcomes. Students will first learn how to write programming codes from scratch using R, one of the most intuitive and popular programming languages. The course will then familiarize students with the most important predictive techniques – from linear regressions to machine learning – and ask them to implement what they’ve learned theoretically in applications related to management and finance.

CREDITS: 3

CONTACT HOURS: 45

LANGUAGE OF INSTRUCTION: English

PREREQUISITES: An introductory knowledge of statistics is required. Students must bring their own laptop in class

METHOD OF PRESENTATION: Theoretical lectures illustrated by examples followed by programming sessions. **Students must bring their own laptops (at least 2 GB of RAM) to every class.** In the first session some concepts in calculus and statistics will be introduced to students. Since the conceptual material is dense and requires both theoretical and practical modules in order to be adequately processed and understood by students, the course pairs theoretical sessions with lab components during which students will get hands-on practice with the week's material by working on assigned lab activities.

REQUIRED WORK AND FORM OF ASSESSMENT:

- Course Participation - 10%
- Midterm Exam - 25%
- Final Exam - 30%
- Homework - 15%
- Group Project - 20%

Course Participation

Students are expected to actively participate in class and to demonstrate respect and responsibility in the classroom. A rubric for participation will be available on Moodle and distributed on the first day of class.

Midterm Exam

Students will be asked to solve exercises on the techniques covered in the first half of the class.

Final Exam

Students will be asked to solve exercises on the techniques covered in the second half of the class.

Homework

Students will be asked to perform weekly exercises that require them to practice and apply the techniques discussed in class.

- Two of the weekly assignments (in Week 4 and Week 9) will receive a letter grade and together will count for 10% of the grade. These graded assignments will require students both to complete problem sets and to give written explanation of their answers and methods. Students will receive feedback from the instructor on their work.
- During all other weeks, homework exercises will be graded for completion and will be worth 5% of the grade. The solutions to the problems will be discussed in class.

Group Project

Students will work in groups to apply techniques used in the course to a real-world case study involving a publicly available dataset. They will be asked to write a final report presenting their results of between 1500 and 2000 words (excluding diagrams, graphs, and

Global brilliance begins here.®

Institute for the International Education of Students
Africa | Asia Pacific | Europe | Latin America

tables) due the last week of the course. The project will require students to move from a business-related objective to the creation of a research question and ultimately a full statistical report. Students will be expected to:

- identify important variables;
- perform statistical analyses to address their question;
- produce tabular and graphical summaries to support their findings; and
- write a summary report detailing their methodological approach, findings, and potential limitations of their results.

LEARNING OUTCOMES:

By the end of the course, students will be able to:

- Apply the logic of computer programming to real-world problems;
- Write code that collects, manipulates, and visualizes data in R;
- Evaluate a data set and identify analytical techniques that can extract relevant information from it;
- Implement the most important predictive techniques in R and apply them to financial or business concerns.

ATTENDANCE POLICY:

Students are expected to be responsible for their own class attendance and to give academic responsibilities priority over other activities. Regular class attendance is mandatory. IES Abroad Milano allows a maximum of THREE (3) absences in each Italian language course, TWO (2) absences in each Area Studies course, and ONE (1) absence in each Seminar course. Every absence beyond this allowance will automatically result in a penalty of 2 percentage points off the final grade. SEVEN (7) absences will result in a failing grade in Italian language and Area Studies courses. FOUR (4) absences will result in a failing grade in Seminar courses. Failure to attend a scheduled exam, test, quiz, or presentation will automatically result in an F grade on that assessment.

CONTENT:

Week	Content	Assignments
Week 1 <i>Session 1</i>	Statistical learning and model accuracy. • Crash course in calculus and statistics • What is statistical learning?	
<i>Session 2</i>	Lab session: R • Introduction to R (bring your own laptop)	• Ch 2, pp. 15-52
Week 2 <i>Session 1</i>	Linear regression • Introduction to simple and multiple linear regressions	• Ch 3, pp. 59-120
<i>Session 2</i>	Lab session: Linear regression • Exercises using linear regression	
Week 3 <i>Session 1</i>	Classification, logistic regressions, and Linear Discretion Analysis • An overview of classification and its distinction from linear regression • The logistic model • Linear discretion analysis	• Ch 4, pp. 127-168

<i>Session 2</i>	Lab session: Classification Data set requiring use and comparison of classification methods	
Week 4 <i>Session 1</i>	Resampling Methods - Cross-validation - The Bootstrap	- Ch 5, pp. 175-197 1st graded homework assignment due.
<i>Session 2</i>	Lab session: Resampling methods Exercises on cross-validation and the bootstrap	
Week 5 <i>Session 1</i>	Linear model selection and regularization - Subset selection - Shrinkage methods - Dimension reduction methods - High dimensions	Ch 6, pp. 203-259
<i>Session 2</i>	Lab session: Model selection Choosing among models using validation	
Week 6 <i>Session 1</i>	MIDTERM EXAM	
Week 7 <i>Session 1</i>	Moving beyond linearity - Splines - Generalized additive models	Ch 7, pp. 265-297
<i>Session 2</i>	Lab session: Non-linear modeling Exercises in applying non-linear modelling	
Week 8 <i>Session 1</i>	Tree-based methods - Decision trees - Random forest - Boosting	Ch 8, pp. 303-332
<i>Session 2</i>	Lab session: Decision trees Exercises in fitting trees, random forests, and boosting	
Week 9 <i>Session 1</i>	Support Vector Machines - Support vector classifiers - The support vector machine	- Ch 9, pp. 337-368 2nd graded homework assignment due.

Session 2	Lab session: Support vector machines Application of support vector machines	
Week 10 Session 1	Unsupervised learning - Principal component - Clustering	Ch 10, pp. 373-413
Session 2	Lab session: Unsupervised learning Exercises applying unsupervised learning	
Week 11 Session 1	Introduction to Simulations - Text as data - Network science	Instructor slides
Session 2	R and Python Similarities and differences	
Week 12	Group project presentation	Group project due
	FINAL EXAM	

REQUIRED READINGS:

- James, Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani. *An Introduction to Statistical Learning with Application in R*. Springer, 2013.

RECOMMENDED READINGS:

- Larose, Chantal D. and Daniel T. Larose. *Data Science Using Python and R*. (Hoboken, NJ: Wiley, 2019).