

PGE – MSc in Management

Python Data Analysis




Course Code

MK347

Academic Year

2017-2018

1. COURSE SPECIFICATIONS*

TEACHING LANGUAGE	English	
ECTS CREDITS	5	
OFFICE HOURS	Office, Phone, Email	
INSTRUCTIONAL DESIGNER	CHERNY Lynn	
ADMINISTRATIVE ASSISTANT	ALESSI Marine	
LEVEL	beginner	
PRE-REQUISITES	MK34 – Python Bootcamp	
RULES	Attendance is required. Copying work from your peers or previous year solutions will result in a 0 for the course for both the copier and source work. Do not put your friends' grades at risk. (We have computational methods to detect this.)	



2. COURSE DESCRIPTION*

This course prepares the student for introductory data science roles that use programming for data analysis. An understanding of basic data analysis with code and data access tools like SQL are critical for a variety of market research and business strategy positions, including data science jobs.

3. PROGRAMME LEARNING GOALS AND OBJECTIVES*



3.1. Programme learning goals *(this section does not need to be completed)*

- 1 Our graduates are aware/sensitive to a disrupted world
- 2 Our graduates master business basics and disruptive business competencies
- 3 Our graduates develop a creative mindset and are able to make ideas happen
- 4 Our graduates reflect on themselves and project themselves in the future



3.2. Links of the course to the programme objectives* *(Please check the learning objectives of the program to which the course contributes the most (at least 1 to maximum 4))*

Learning objectives (LO) of the programme	Course Matching
PGE - MSc in Management	
Graduates of the programme are able to identify key relevant technological, scientific and societal disruptions and apply their insights to business situations, projects or issues.	<input type="checkbox"/> Yes
Graduates of the programme are able to adopt a global perspective, mindset and competences when dealing with various business situations, projects or issues	<input type="checkbox"/> Yes
Graduates of the programme know and effectively use basic concepts, knowledge and theories in the main business subjects and are able to integrate them in a systemic approach	<input checked="" type="checkbox"/> Yes
<i>Pour les tracks IDEA et ETD uniquement :</i> - ETD : Our students acquire advanced knowledge and skills to operate in different European countries - IDEA : Our students acquire advanced knowledge and skills to develop innovative ideas and projects	<input checked="" type="checkbox"/> Yes
Graduates of the programme are able to identify alternative/emergent business concepts, knowledge and theories and are able to apply them to various business situations, projects or issues.	X Yes
Graduates of the programme are able to question proven approaches and solutions to business problems and apply creativity and design techniques to enlarge the scope of their analysis, enabling them to bring new creative solutions.	<input checked="" type="checkbox"/> Yes
Graduates of the programme are able to demonstrate a sense of responsibility and upholds their values and principles when dealing with colleagues, instructors, employers or business partners	<input checked="" type="checkbox"/> Yes

Graduates of the programme are able to systematically appraise their learning/working experiences and take active steps to know themselves better in order to develop self-improvement strategies

☑ Yes



4. LEARNING HOURS ANALYSIS*

Type of learning Hours	Distribution (hours)
Face-to-Face (lecture) (24 hours max)	24
Online working Hours	0
Individual working hours (with mentoring or not)	24
Team working hours (with mentoring or not)	0
Evaluation	2
TOTAL	=0

5. LEARNING OBJECTIVES / OUTCOMES

5.1. Academic outcomes (identify a maximum of 5 learning objectives)



5.1.1. Concepts and theories to which participants are exposed during the course*

1	Computer programming principles
2	Data exploration and analysis principles
3	
4	
5	



5.1.2. Knowledge acquired during the course *

After this course, participants know or master	Python programming language concepts for data analysis using pandas, the most sophisticated library
After this course, participants know or master	Exploratory data analysis and basic data statistics
After this course, participants know or master	File read/write in various data formats
After this course, participants know or master	SQL database access from code
After this course, participants know or master	API access from code

5.1.3. Competencies acquired during the course *

After this course, participants are able to	Acquire data from various data sources (files, APIs, databases)
---	---

After this course, participants are able to	Clean and analyse data in code and summarize/describe it
After this course, participants are able to	Make charts using data in code
After this course, participants are able to	Write out summary statistics in formats such as Excel

5.2. Professional outcomes



5.2.1 Productions/Deliverables participants will be able to present after taking this course*

1	Jupyter notebooks demonstrating Python and pandas basic skills
2	Data analysis on files using code including charts and graphs
3	



5.2.2 The people that participants will have met or contacted during the course: (professors, guest speakers, tutors, companies, external experts, participants...)

Guest vacataire and data scientist Hugo Herrou, student at Ecole de Mines
RP : Lynn CHERNY, Professor at emlyon



6. ASSESSMENT OF LEARNING OBJECTIVES* *(please re-enter the n° of the learning objectives checked in part 3.2 and explain in a few words the assessment method)*

Students will do weekly homework assignments that teach import and data analysis skills with code. Final project will be a data analysis assignment that uses multiple skills taught in the course.



7. SEQUENCES DESCRIPTION, LEARNING ACTIVITIES AND ASSIGNMENTS

Content Plan	Useful References (more provided each week)
S1. Python concepts again – read/write files, tuples.	<p>https://www.packtpub.com/books/content/basics-jupyter-notebook-and-python</p> <p>https://www.tutorialspoint.com/python/python_tuples.htm .</p> <p>[Note that it was written in Python 2 -- the difference is in the print statement. In Python 3, we use print(thing), in Python 2, print thing.]</p> <p>Then read the chapter in Trinket, to remind yourself of the dictionaries .items() and .keys() as well, which return tuples: https://books.trinket.io/pfe/10-tuples.html</p> <p>Quiz on tuples and other basic concepts.</p>
S2. Intro to Pandas basic functionality : Selecting, Counting, Columns.	<p>Backup readings : http://www.dataschool.io/easier-data-analysis-with-pandas/</p> <p>http://pandas.pydata.org/pandas-docs/stable/10min.html</p>
S3. Pandas GroupBy, Aggregate, String data	
S4. Pivot and Unpivot. Timeseries.	
S5. SQLite, Join/Merge tables	
S6. SQL access remotely, Public APIs for data, Twitter API. JSON Parsing.	
S7. Charts and Graphs in Matplotlib/Seaborn	
S8. Review, Project data.	

Grading:

DELIVERABLE	DESCRIPTION	% FINAL GRADE	GRADING
Individual assignments corresponding to 8 sessions	Each analysis project will be begun in class.	55%	10/20 is minimal, 20/20 excellence
Quizzes based on the lecture and readings	You must pay attention to the lecture and readings. The quizzes will assess whether you did. They will not appear every week.	20%	10/20 minimal, 20/20 excellence
Final Project	Final PPT report plus notebook code and data	25%	50/100 minimal, 100/100 excellence

Final projects will be due 2 weeks after the end of the class lectures.

Extra credit will be given to students who help answer questions on the Discussion forum in Brightspace.



8. PEDAGOGICAL RESOURCES AND ACTIVITIES (videos, books, articles, links, etc)*

PEDAGOGICAL RESOURCE/ACTIVITY	RESOURCE LINK
Using the Jupyter Notebook :	https://www.packtpub.com/books/content/basics-jupyter-notebook-and-python Also Safari Online videos : https://www.safaribooksonline.com/learning-paths/learning-path-jupyter/9781491995648/9781771375030-video241997
Review your python bootcamp : Python for Everybody : Trinket online book on Python 3, and other online tutorials	https://books.trinket.io/pfe/index.html Another good online tutorial: http://introtopython.org/hello_world.html
Intro to Python – Supplemental, More Advanced	A WhirlWind Tour of Python by Jake Van Der Plas

Software Carpentry Lesson on Command Line Use and Bash Programming. Very useful.	https://swcarpentry.github.io/shell-novice/01-intro/
Pandas introductions. There are many, here are 2...	http://www.dataschool.io/easier-data-analysis-with-pandas/ http://pandas.pydata.org/pandas-docs/stable/10min.html
Pandas documentation. Get in the habit of looking at it.	http://pandas.pydata.org/pandas-docs/stable/#



9. OTHER USEFUL INFORMATION

Class attendance is required or your grade will be reduced.

Cheating policy: Copying work from other students (or sharing your work for them to copy) is not allowed. It will result in a 0 for all concerned parties. The proper way to work together is to explain it, ask questions, and make it your own. If you cannot customize code to make it distinct, you don't understand it.

---END OF DOCUMENT---

Date* : 16/08/2017

Document status* : **Draft**