



Εθνικό Μετσόβιο Πολυτεχνείο  
Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών  
Τομέας Τεχνολογίας Πληροφορικής & Υπολογιστών  
Εργαστήριο Τεχνολογίας Λογισμικού

## Προγραμματιστικά Εργαλεία και Τεχνολογίες για Επιστήμη Δεδομένων Ενημερωτικό Φυλλάδιο

### Γενικές πληροφορίες

Μάθημα	: Προγραμματιστικά Εργαλεία και Τεχνολογίες για Επιστήμη Δεδομένων (716)
Εξάμηνο	: χειμερινό, μεταπτυχιακό
Κατανομή ωρών	: 2 ώρες θεωρία, 1 ώρα εργαστήριο
Ακαδημαϊκό έτος	: 2020–2021
Διδάσκοντες	: Νίκος Παπασπύρου (nickie@softlab.ntua.gr) Κτίριο Ηλεκτρολόγων, 1.1.21, τηλ. 210-772-3393 Δημήτρης Φουσκάκης (fouskakis@math.ntua.gr) Κτίριο E, γρ. 205, τηλ. 210-772-1702
Διαλέξεις	: Πέμπτη 15:15–18:00, MS Teams, κωδικός sy61n1k
Ώρες γραφείου	: Παρασκευή 10:00–12:00
Ιστοσελίδα	: <a href="https://courses.softlab.ntua.gr/progds/">https://courses.softlab.ntua.gr/progds/</a> <a href="http://www.math.ntua.gr/~fouskakis/Programming_R/progr.html">http://www.math.ntua.gr/~fouskakis/Programming_R/progr.html</a>

### Σύντομη περιγραφή

Επεξεργασία και ανάλυση δεδομένων με τις γλώσσες Python και R. Εξερεύνηση δεδομένων και στατιστική μοντελοποίηση.

Εισαγωγή στην γλώσσα Python. Τύποι δεδομένων: Λίστες, λεξικά, κλπ. Ροή ελέγχου. Συναρτήσεις και πακέτα λογισμικού. Αντικείμενα και αντικειμενοστρέφεια. Βιβλιοθήκες επεξεργασίας και διαχείρισης δεδομένων (numpy, pandas, κλπ). Βιβλιοθήκες απεικόνισης δεδομένων (matplotlib, κλπ.). Βασικές βιβλιοθήκες μηχανικής μάθησης (scikit-learn). Εφαρμογές.

Εισαγωγή στην στατιστική γλώσσα προγραμματισμού R. Αριθμητικοί τελεστές. Βασικοί τύποι και δομές δεδομένων: διανύσματα, πίνακες, λίστες, dataframes. Δημιουργία, εισαγωγή, προ-επεξεργασία, διαχείριση και βασικοί μετασχηματισμοί δεδομένων. Ροή ελέγχου. Βιβλιοθήκες απεικόνισης δεδομένων (ggplot, κλπ.). Βασικές βιβλιοθήκες μηχανικής μάθησης (e1071, gpart, κλπ.). Εφαρμογές.

Οι παραδόσεις του πρώτου μέρους του μαθήματος (Python) αρχίζουν εντατικά περίπου από τις 15/9 και ολοκληρώνονται σε τρεις εβδομάδες.

### Εξετάσεις, εργασία και τελική βαθμολογία

Το μάθημα *δε θα έχει τελική εξέταση*. Ο τελικός βαθμός θα προκύψει ως το άθροισμα των βαθμών των δύο μερών του μαθήματος (Python και R), καθένα από τα οποία μετρά συνολικά πέντε (5) μονάδες. Για καθένα από τα δύο μέρη, η μία (1) μονάδα προκύπτει από την *εργαστηριακή εξέταση* και οι υπόλοιπες τέσσερις (4) από μία ατομική εργασία. Οι ημερομηνίες των εργαστηριακών εξετάσεων και τα θέματα των εργασιών ανακοινώνονται στην ιστοσελίδα του μαθήματος. Οι εργασίες θα έχουν ημερομηνία παράδοσης που θα τηρείται αυστηρά.