

INTRODUCTION TO DATA ANALYSIS WITH PYTHON

Σχετικά με τη γλώσσα προγραμματισμού Python

Η Python είναι η δημοφιλέστερη γλώσσα προγραμματισμού παγκοσμιώς^{1,2,3,4} λόγω της απλότητας και της ευελιξίας της. Τόσο η ίδια η γλώσσα, όσο και η πλειοψηφία των πολυπληθέστατων βιβλιοθηκών της είναι «ανοιχτού κώδικα» και διατίθενται εντελώς δωρεάν.

Εκτός από τη δημιουργία εφαρμογών γενικού σκοπού, η Python ξεχωρίζει για τις δυνατότητές της στα πεδία Data Science, Machine Learning και Artificial Intelligence, που γνωρίζουν ραγδαία ανάπτυξη τα τελευταία χρόνια.

Ο πολύ απλός τρόπος σύνταξης της γλώσσας την καθιστά ιδιαίτερα εύκολη στην εκμάθηση και χρήση, με αποτέλεσμα οι αρχάριοι χρήστες της να μπορούν να γίνουν πολύ γρήγορα παραγωγικοί. Σε ένα τυπικό επαγγελματικό περιβάλλον, η δημιουργία αναφορών, η ενημέρωση/επεξεργασία αρχείων, η αποστολή email, η συλλογή δεδομένων από αρχεία ή το internet είναι μόνο κάποιες από τις συνήθεις καθημερινές διεργασίες που μπορούν να αυτοματοποιηθούν και να αναβαθμιστούν με τη γλώσσα Python.

Η χρήση της Python καθώς και της βιβλιοθήκης για ανάλυση δεδομένων "Pandas", έχουν αρχίσει πλέον να υιοθετούνται από πολλούς κλάδους που επεξεργάζονται και αναλύουν δεδομένα κάθε είδους και μεγέθους, ενώ έχουν ήδη αρχίσει να καθίστανται industry-standards σε κλάδους που παραδοσιακά χρησιμοποιούσαν μεγάλα και σύνθετα spreadsheets.

Ένα χαρακτηριστικό παράδειγμα είναι ο κλάδος των Χρηματοοικονομικών/Χρηματιστηριακών, όπου τα spreadsheets κυριαρχούσαν για δεκαετίες. Στελέχη του κλάδου υποστηρίζουν ότι η Python έχει ήδη αντικαταστήσει το Excel⁵, με τους χρήστες της να περιλαμβάνουν μεταξύ άλλων traders⁶, asset managers, αναλυτές και quantitative researchers, καθώς οι επενδυτικές στρατηγικές που βασίζονται σε αλγόριθμους έχουν πάρει εκρηκτικές διαστάσεις. Μεγάλοι επενδυτικοί οίκοι διοργανώνουν πλέον σεμινάρια Python για το προσωπικό τους^{7,8,9}, ενώ παράλληλα, έχει διαμορφωθεί μια ξεκάθαρη τάση στον κλάδο για προσλήψεις στελεχών με γνώσεις Python ή κάποιας άλλης συναφούς γλώσσας^{10,11}.

```
self.fingerprints = self.  
.logdupes = True  
.debug = debug  
.logger = logging.getLogger  
path:  
self.file = open(os.path  
self.file.seek(0)  
self.fingerprints.update  
  
method  
from settings(cls, setting  
debug = settings.getbool(  
return cls(job_dir(settin  
  
request_seen(self, request  
fp = self.request_finger  
if fp in self.fingerprin  
return True  
self.fingerprints.add(fp  
if self.file:  
self.file.write(fp  
  
request_fingerprint(se  
return request_fingerp
```



24, 26, 31 Μαρτίου
& 2 Απριλίου 2020



18:00 - 22:00



17 Μαρτίου 2020



Ελληνικά



300 €
(συμπεριλαμβανομένου Φ.Π.Α.)



Χρηματιστήριο Αθηνών
Λεωφ. Αθηνών 110
T.K. 104 42



ΝΙΚΟΣ
ΧΡΙΣΤΟΔΟΥΛΟΠΟΥΛΟΣ

Διευκρινήσεις

Το Επίπεδο Ύλης του Σεμιναρίου:
«Εξειδικευμένο - Εισαγωγικό»

Η παρακολούθηση του σεμιναρίου
δεν έχει ως προϋπόθεση γνώσεις
προγραμματισμού.

Τα πλεονεκτήματα της Python έναντι των spreadsheets

Ταχύτητα

Είναι εντυπωσιακά γρηγορότερη, ιδιαίτερα όταν απαιτούνται σύνθετοι υπολογισμοί σε μεγάλο όγκο δεδομένων.

Πανίσχυρες υπολογιστικές δυνατότητες

Μια γραμμή κώδικα Python μπορεί να εκτελέσει κάτι που σε ένα spreadsheet θα χρειαζόταν πολύς κόπος και χρόνος για να υλοποιηθεί, ακόμα και με χρήση VBA.

Περιορισμός και ανιχνευσιμότητα λαθών

Η χρήση του απλού και ευανάγνωστου κώδικα -αντί παρεμβατικών ενεργειών όπως copy/paste δεδομένων, συναρτήσεων κλπ-, ελαχιστοποιεί την πιθανότητα λάθους και ταυτόχρονα απλοποιεί αφάνταστα την ανιχνευσιμότητά του.

Μεγάλες δυνατότητες κλιμάκωσης

Στην Python υπάρχει πλήρης διαχωρισμός μεταξύ των πρωτογενών δεδομένων και του κώδικα επεξεργασίας & ανάλυσής τους, δίνοντας έτσι τη δυνατότητα πρακτικά απεριόριστης προσθήκης επιπλέον δεδομένων (σε όρους "στηλών" ή "γραμμών") χωρίς καμία τροποποίηση του κώδικα.

Επαναληψιμότητα & αυτοματισμός

Η Python αποτελεί ιδανική λύση σε περιπτώσεις όπου απαιτείται τακτική -π.χ. καθημερινή- ενημέρωση δεδομένων και διενέργεια υπολογισμών, αφού όλη η διαδικασία (από τη συλλογή των νέων δεδομένων, την ανάλυσή τους, μέχρι και την αποστολή των τελικών αναφορών ή αρχείων) μπορεί να αυτοματοποιηθεί πλήρως.

Υποστήριξη βάσεων και πηγών δεδομένων

Η Python μπορεί να επικοινωνεί απευθείας με κάθε είδους βάσεις δεδομένων, αρχεία ή ιστοσελίδες και να αντλεί νέα δεδομένα σε προκαθορισμένα χρονικά διαστήματα.

Συμβατότητα και ευελιξία

Η Python είναι συμβατή με όλα τα λειτουργικά συστήματα (Windows, macOS, Linux) και ο κώδικας (όχι κατ'ανάγκη και τα δεδομένα) μπορεί να στεγάζεται και να εκτελείται σε cloud. Λόγω της ευελιξίας που της δίνει η ιδιαίτερη δυνατότητά της να επικοινωνεί με διαφορετικά συστήματα αποκαλείται και "glue language".

Επεκτασιμότητα

Οι δυνατότητες της Python δεν σταματούν στη διενέργεια υπολογισμών και στην ανάλυση δεδομένων, καθώς πρόκειται για μια ολοκληρωμένη και πανίσχυρη γλώσσα προγραμματισμού και όχι μια απλή εφαρμογή όπως ένα spreadsheet. Επομένως, οι δυνατότητες προσθήκης επιπλέον λειτουργιών (πχ η άμεση & αυτόματη δημοσίευση της ανάλυσης δεδομένων στο internet μέσω ενός Python Webserver) είναι πρακτικά απεριόριστες καθώς υπάρχει πλήθος δωρεάν, open-source βιβλιοθηκών σε πεδία όπως Machine Learning, Big Data, Artificial Intelligence, Web Development, Natural Language Processing κ.α. που αναπτύσσονται συνεχώς.

Κόστος

Η Python διατίθεται δωρεάν, όπως και η πλειοψηφία των βιβλιοθηκών της.

Σε ποιους απευθύνεται

Η παρακολούθηση του σεμιναρίου **δεν έχει ως προϋπόθεση γνώσεις προγραμματισμού.**

Το σεμινάριο απευθύνεται σε επαγγελματίες όλων των κλάδων οι οποίοι επιθυμούν να αξιοποιήσουν δεδομένα και να δώσουν προστιθέμενη αξία στην εργασία τους.

Αφορά όσους ασχολούνται ή θέλουν να ασχοληθούν με:

- επεξεργασία και ανάλυση δεδομένων
- συγκέντρωση και διαχείριση δεδομένων
- ενημέρωση αρχείων που περιέχουν δεδομένα
- δημιουργία αναφορών με στατιστικά δεδομένα ή γραφήματα

και θέλουν να αναπτύξουν νέες δεξιότητες σε εργαλεία ανάλυσης δεδομένων τεχνολογίας αιχμής.

Το σεμινάριο είναι ιδιαίτερα χρήσιμο για οποιαδήποτε ειδικότητα η οποία βασίζει μεγάλο μέρος της εργασίας της σε σύνθετα και μεγάλα spreadsheets, όπως:

- Διαχειριστές κεφαλαίων
- Traders / επενδυτές
- Σύμβουλοι επενδύσεων
- Risk managers
- Financial analysts
- Business analysts
- Business analytics specialists
- MIS analysts
- Financial controllers
- Αναλογιστές
- Στατιστικοί
- Δημοσκόποι
- Μηχανικοί
- Αναλυτές / ερευνητές που χρησιμοποιούν δεδομένα κάθε είδους, μεγέθους και πολυπλοκότητας.

Σκοπός

Σκοπός του σεμιναρίου είναι να δώσει τα βασικά εφόδια για την κατανόηση της φιλοσοφίας και του τρόπου σύνταξης της γλώσσας Python, ώστε οι συμμετέχοντες να είναι σε θέση να την αξιοποιήσουν στην πράξη ανάλογα με το αντικείμενό τους.

Επιχειρεί να εξηγήσει βήμα-βήμα τις πανίσχυρες τεχνικές ανάλυσης και επεξεργασίας δεδομένων που παρέχει η εξειδικευμένη βιβλιοθήκη Pandas, οι οποίες δίνουν τη δυνατότητα διεξοδικής ανάλυσης δεδομένων υψηλής προστιθέμενης αξίας με ακρίβεια και ταχύτητα.



Στόχοι

Ολοκληρώνοντας το σεμινάριο, οι συμμετέχοντες θα έχουν αποκτήσει τις βάσεις για να γράφουν κώδικα Python και να χρησιμοποιούν την εξειδικευμένη βιβλιοθήκη Pandas ώστε να μπορούν:

- Να διενεργούν διεξοδική ανάλυση δεδομένων υψηλής προστιθέμενης αξίας με ακρίβεια και ταχύτητα,
- Να συλλέγουν δεδομένα από αρχεία, βάσεις δεδομένων ή ιστοσελίδες,
- Να επεξεργάζονται, να καθαρίζουν, να τροποποιούν και να ομαδοποιούν δεδομένα,
- Να δημιουργούν αξιοποιήσιμες αναφορές και χρήσιμα γραφήματα,
- Να αποστέλλουν emails με εξατομικευμένες αναφορές,
- Να αυτοματοποιήσουν όλες τις παραπάνω διεργασίες.

Επιπλέον θα είναι σε θέση να πειραματίζονται και να περιεργάζονται τα δεδομένα με τέτοιο τρόπο, ώστε να μπορούν να εξαγάγουν χρήσιμα συμπεράσματα (exploratory data analysis).

Μεθοδολογία

Η παρουσίαση και διδασκαλία γίνεται με τη χρήση Jupyter Notebook, μια διαδραστική ιστοσελίδα που είναι σε θέση να εκτελέσει κομμάτια κώδικα. Οι συμμετέχοντες μπορούν μετά το σεμινάριο να τρέχουν το αρχείο της παρουσίασης στον υπολογιστή τους, κάνοντας έτσι πρακτική εξάσκηση πειραματιζόμενοι με τον κώδικα.

Πρόγραμμα σεμιναρίου

Python Basics

Installation

Jupyter Notebook

Basic Concepts

Data types

Control flow

Functions

Error handling

Pandas (εξειδικευμένη βιβλιοθήκη Python για ανάλυση δεδομένων)

Pandas Basics

Loading and Storing data

Selecting and Filtering data

Cleaning and Transforming data

Merging data

Aggregating data

Plotting data

Special topics

Παραπομπές

1. The Top Programming Languages 2019
IEEE.org
2. The most popular programming languages on Stack Overflow
Stackoverflow.com
3. Python is becoming the world's most popular coding language
Economist.com
4. Python has brought computer programming to a vast new audience
Economist.com
5. Python already replaced Excel in banking
efinancialcareers.com
6. Why traders need to start learning Python, and what they will do with it
efinancialcareers.com
7. Barclays is pushing its equities traders to learn how to code in Python and it shows where stock trading is heading
Businessinsider.com
8. Citi's Investment Bank Plans to Hire 2,500 Coders This Year
Bloomberg.com
9. JPMorgan's requirement for new staff: coding lessons.
Hundreds of new investment bankers and asset managers undergo mandatory training
FT.com
10. Graduates with tech and finance skills in high demand
FT.com
11. Goldman Plans Hiring Spree in Trading (Only Coders Need Apply)
Bloomberg.com

Πληροφορίες

Τρόπος Πληρωμής

Τα δίδακτρα κατατίθενται στην Alpha Bank, αριθμός λογαριασμού 101-00-2002-046556 ή IBAN GR65 0140 1010 1010 0200 2046 556. Το καταθετήριο της Τράπεζας, με το ονοματεπώνυμο του καταθέτη, θα πρέπει να αποστέλλεται στην ATHEXacademy μέσω ηλεκτρονικού ταχυδρομείου. ATHEXacademy@athexgroup.gr

Επιδότηση ΟΑΕΔ

Το σεμινάριο μπορεί να καλυφθεί με το 0,24% της εργοδοτικής εισφοράς για τα προγράμματα κατάρτισης & εκπαίδευσης από τον ΟΑΕΔ. Ενημερωθείτε για τη νέα εγκύκλιο ΛΑΕΚ και δείτε πως μπορείτε να ωφεληθείτε μέσω της ηλεκτρονικής διεύθυνσης www.oaed.gr. Καταληκτική ημερομηνία ηλεκτρονικής υποβολής των δικαιολογητικών στον Ο.Α.Ε.Δ. 5 εργάσιμες ημέρες πριν την έναρξη του σεμιναρίου.

Παρεχόμενα σεμιναρίου

Με την διεξαγωγή

του σεμιναρίου παρέχονται:

- Βεβαίωση παρακολούθησης σεμιναρίου
- Η παρουσίαση του σεμιναρίου σε ηλεκτρονικό αρχείο

Για περισσότερες πληροφορίες μπορείτε να επικοινωνείτε στα στοιχεία επικοινωνίας

Εισηγητής

ΝΙΚΟΣ

ΧΡΙΣΤΟΔΟΥΛΟΠΟΥΛΟΣ

Ο Νίκος Χριστοδουλόπουλος είναι επικεφαλής της Ειδικής Διαπραγμάτευσης Παραγώγων της Πειραιώς ΑΕΠΕΥ. Έχει 20 χρόνια εμπειρίας στο χώρο των χρηματοοικονομικών, 15 εκ των οποίων σε market-making παραγώγων προϊόντων που διαπραγματεύονται στο Χρηματιστήριο Αθηνών. Διαθέτει εκτενή και πολυετή εμπειρία σε προγραμματισμό, ανάλυση δεδομένων και spreadsheet development με χρήση VBA. Έχει εργαστεί σε διάφορους τραπεζικούς ομίλους σε Ελλάδα και Μεγάλη Βρετανία, είναι κάτοχος MSc in International Securities, Investment & Banking από το ICMA Centre, University of Reading, UK και BA in Accounting & Finance από το University of Essex, UK.