



Ροή+ Οργανισμός Εξειδικευμένης Εκπαίδευσης Ι.Κ.Ε.

ΣΕΜΙΝΑΡΙΑ

Ρυθμοί για Στατιστική Ανάλυση, Περιγραφή, Συμπερασματολογία & Μοντελοποίηση

2026

1. ΕΙΣΑΓΩΓΗ

Το παρόν πρόγραμμα σπουδών προσφέρει μια ολοκληρωμένη εισαγωγή στη χρήση της Python ως κύριου εργαλείου για στατιστική ανάλυση δεδομένων. Σχεδιασμένο για αρχάριους στον προγραμματισμό, θέτει γερά θεμέλια στη γλώσσα Python και οδηγεί μεθοδικά τους συμμετέχοντες στην κατανόηση και εφαρμογή περιγραφικών στατιστικών μετρικών, στην εκτέλεση στατιστικών ελέγχων για εξαγωγή συμπερασμάτων και στη δημιουργία στατιστικών μοντέλων για πρόβλεψη και ερμηνεία σχέσεων. Η φιλοσοφία του προγράμματος είναι η σύνδεση της στατιστικής θεωρίας με την πρακτική εφαρμογή, χρησιμοποιώντας πραγματικά δεδομένα και τις πλέον σύγχρονες βιβλιοθήκες της Python, όπως Pandas, SciPy, Statsmodels και Seaborn.

Συνολικές Ώρες: 20

2. ΕΝΟΤΗΤΕΣ

1. Θεμέλια της Python & Εισαγωγή στο Οικοσύστημα Στατιστικής Ανάλυσης (2 ώρες)

Σκοπός: Η θεμελίωση των βασικών αρχών προγραμματισμού στην Python και η πρώτη επαφή με το περιβάλλον και τα εργαλεία που χρησιμοποιούνται για στατιστική ανάλυση.

Θεωρητικό Υπόβαθρο:

- Η Python στην Επιστήμη Δεδομένων: Επισκόπηση του ρόλου της Python ως του κυρίαρχου εργαλείου για στατιστική ανάλυση, διερευνητική ανάλυση και μοντελοποίηση, τόσο στην ακαδημαϊκή έρευνα όσο και στη βιομηχανία.
- Διαμόρφωση Εργαστηρίου Στατιστικής Ανάλυσης: Λεπτομερής εγκατάσταση της σουίτας Anaconda, η οποία περιλαμβάνει το Jupyter Notebook ως διαδραστικό υπολογιστικό περιβάλλον, και τις θεμελιώδεις βιβλιοθήκες: NumPy, Pandas, SciPy, Statsmodels, Matplotlib και Seaborn .
- Βασικές Έννοιες Προγραμματισμού: Εισαγωγή στις μεταβλητές, τους τύπους δεδομένων (int, float, str, bool), τους τελεστές και τη σημασία της τεκμηρίωσης του κώδικα. Ιδιαίτερη έμφαση δίνεται στην αναγνωσιμότητα και την καλή πρακτική, ως θεμέλιο για την ανάπτυξη αξιόπιστων στατιστικών αναλύσεων .

Πρακτική Εφαρμογή:

- Πρώτο Πρόγραμμα: Δημιουργία και εκτέλεση ενός πρώτου προγράμματος Python για επιβεβαίωση της σωστής λειτουργίας του περιβάλλοντος.
- Εισαγωγή στις Βιβλιοθήκες: Φόρτωση βιβλιοθηκών (import) και εκτέλεση απλών αριθμητικών πράξεων.

Άσκηση Ενοποίησης: Υπολογισμός Περιγραφικών Στατιστικών Μέτρων με Απλά Δεδομένα.

Οι συμμετέχοντες δημιουργούν μικρές λίστες αριθμών (π.χ., βαθμολογίες φοιτητών) και υπολογίζουν "χειροκίνητα", χρησιμοποιώντας βασικές πράξεις και συναρτήσεις, στατιστικά μέτρα όπως το άθροισμα, το πλήθος, τον μέσο όρο και την τυπική απόκλιση. Αυτή η άσκηση στοχεύει στην κατανόηση του τι αντιπροσωπεύουν τα στατιστικά μέτρα, πριν αυτοματοποιηθούν.

2. Δομές Δεδομένων & Εισαγωγή στην Ανάλυση με NumPy (2 ώρες)

Σκοπός:

Η εμβάθυνση στην οργάνωση δεδομένων και η εισαγωγή στη βιβλιοθήκη NumPy ως θεμέλιο για στατιστικούς υπολογισμούς.

Θεωρητικό Υπόβαθρο:

- Οργάνωση Δεδομένων στην Python: Αναλυτική παρουσίαση των δομών δεδομένων (λίστες, πλειάδες, λεξικά) και η χρησιμότητά τους στην προ-στατιστική επεξεργασία.
- NumPy και η Δύναμη των Πινάκων: Η έννοια του ηδιάστατου πίνακα (`ndarray`) και τα πλεονεκτήματά του (ταχύτητα, διανυσματικές πράξεις) για στατιστικούς υπολογισμούς σε μεγάλα σύνολα δεδομένων. Παρουσίαση βασικών στατιστικών συναρτήσεων της NumPy (`np.mean`, `np.median`, `np.std`, `np.var`, `np.percentile`).
- Περιγραφική Στατιστική σε Επίπεδο Πίνακα: Η έννοια της εφαρμογής στατιστικών συναρτήσεων σε ολόκληρους πίνακες ή ανά άξονα (π.χ., υπολογισμός μέσου όρου για κάθε στήλη).

Πρακτική Εφαρμογή:

- Χειρισμός NumPy Arrays: Δημιουργία, τεμαχισμός (`slicing`), ευρετηρίαση και εφαρμογή βασικών πράξεων.
- Υπολογισμός Στατιστικών: Χρήση των στατιστικών συναρτήσεων της NumPy σε πραγματικά ή συνθετικά δεδομένα.

Άσκηση Ενοποίησης: Διερεύνηση Κατανομής Δεδομένων.

Οι συμμετέχοντες δημιουργούν έναν πίνακα NumPy με τυχαίους αριθμούς (π.χ., από κανονική κατανομή). Στη συνέχεια, υπολογίζουν και ερμηνεύουν τα βασικά περιγραφικά στατιστικά μέτρα (μέση τιμή, διάμεσο, τυπική απόκλιση, τεταρτημόρια) για να περιγράψουν την κεντρική τάση και τη διασπορά του συνόλου δεδομένων.

3. Έλεγχος Ροής & Θεμελιώδεις Αρχές Στατιστικής Συμπερασματολογίας (2 ώρες)

Σκοπός:

Η σύνδεση των δομών ελέγχου προγράμματος με τη λογική της στατιστικής συμπερασματολογίας και η εισαγωγή σε βασικές έννοιες πληθυσμού και δείγματος.

Θεωρητικό Υπόβαθρο:

Δομές Ελέγχου και Δειγματοληψία:

- Βρόχοι For/While: Η χρήση τους για προσομοίωση δειγματοληψίας, επαναλαμβανόμενης δειγματοληψίας (resampling) και τεχνικών bootstrap .
- If/else: Η λογική των συνθηκών ως βάση για τη δημιουργία υποσυνόλων δεδομένων βάσει κριτηρίων (π.χ., διαχωρισμός ηλικιακών ομάδων).

Από το Δείγμα στον Πληθυσμό:

- Πληθυσμός (Population) και Δείγμα (Sample): Η θεμελιώδης διάκριση και η σημασία της για τη στατιστική συμπερασματολογία.
- Κεντρικό Οριακό Θεώρημα (Central Limit Theorem): Μια διαισθητική εισαγωγή στο θεώρημα και τον ρόλο του στην εξαγωγή συμπερασμάτων για τον πληθυσμό από ένα δείγμα .

Πρακτική Εφαρμογή:

- Προσομοίωση Δειγματοληψίας: Χρήση βρόχων για επαναλαμβανόμενη δειγματοληψία από έναν γνωστό "πληθυσμό" (π.χ., ένα μεγάλο array) και υπολογισμός της κατανομής του δειγματικού μέσου.
- Οπτικοποίηση Κατανομής: Δημιουργία ιστογραμμάτων της κατανομής του δειγματικού μέσου και παρατήρηση της τάσης προς κανονικότητα, εμπεδώνοντας οπτικά το Κεντρικό Οριακό Θεώρημα.

4. Χειρισμός Δεδομένων με Pandas για Στατιστική Ανάλυση (2 ώρες)

Σκοπός:

Η εκμάθηση της βιβλιοθήκης Pandas ως του κύριου εργαλείου για φόρτωση, καθαρισμό, μετασχηματισμό και προετοιμασία δεδομένων για στατιστική ανάλυση.

Θεωρητικό Υπόβαθρο:

- Η Δομή DataFrame: Παρουσίαση του DataFrame ως θεμελιώδους δομής για την αποθήκευση και διαχείριση δομημένων (tabular) δεδομένων.
- Εισαγωγή/Εξαγωγή Δεδομένων: Φόρτωση δεδομένων από διάφορες πηγές, με έμφαση σε αρχεία CSV και Excel.
- Διερευνητική Ανάλυση Δεδομένων (EDA): Μέθοδοι για γρήγορη επισκόπηση της δομής και του περιεχομένου: `head()`, `info()`, `describe()`, `shape`.
- Καθαρισμός Δεδομένων (Data Cleaning): Αντιμετώπιση ελλιπών τιμών (missing values), εντοπισμός και διαχείριση διπλότυπων.
- Επιλογή και Φιλτράρισμα: Χρήση `loc`, `iloc` και boolean indexing για τη δημιουργία υποσυνόλων βάσει συνθηκών.

Πρακτική Εφαρμογή:

- Φόρτωση και Εξερεύνηση: Φόρτωση ενός πραγματικού dataset (π.χ., για ανάλυση πωλήσεων, δημογραφικά στοιχεία) και χρήση των μεθόδων EDA για την κατανόηση της δομής και των χαρακτηριστικών του.
- Καθαρισμός Δεδομένων: Εφαρμογή τεχνικών για τον εντοπισμό και τη διαχείριση ελλιπών τιμών.

Άσκηση Ενοποίησης: Προετοιμασία Dataset για Στατιστική Ανάλυση.

Οι συμμετέχοντες λαμβάνουν ένα "ακατέργαστο" dataset και καλούνται να το καθαρίσουν, να επιλέξουν τις σχετικές μεταβλητές και να δημιουργήσουν ένα καθαρό DataFrame, έτοιμο για τροφοδότηση σε στατιστικές διαδικασίες.

5. Οπτικοποίηση Δεδομένων για Στατιστική Ανάλυση (2 ώρες)

Σκοπός:

Η απόκτηση της ικανότητας δημιουργίας στατιστικών γραφημάτων για την εξερεύνηση δεδομένων, τον έλεγχο υποθέσεων και την παρουσίαση αποτελεσμάτων.

Θεωρητικό Υπόβαθρο:

- Αρχές Στατιστικής Οπτικοποίησης: Ο ρόλος των γραφημάτων στη διερευνητική ανάλυση, τον εντοπισμό προτύπων, ακραίων τιμών και την επικοινωνία των αποτελεσμάτων.
- Βιβλιοθήκες Οπτικοποίησης: Εισαγωγή στις βιβλιοθήκες Matplotlib (για βασική προσαρμογή) και Seaborn (για στατιστικά γραφήματα υψηλού επιπέδου).

Τύποι Στατιστικών Γραφημάτων:

- Ιστογράμματα (Histograms): Για την κατανόηση της κατανομής μιας συνεχούς μεταβλητής.
- Κουτιά Ανακάλυψης (Box Plots): Για την οπτικοποίηση της διάμεσου, των τεταρτημορίων και των ακραίων τιμών.
- Γραφήματα Πυκνότητας (Density Plots): Για την ομαλή αναπαράσταση της κατανομής.
- Διαγράμματα Διασποράς (Scatter Plots): Για την εξέταση της σχέσης μεταξύ δύο συνεχών μεταβλητών.
- Θερμικοί Χάρτες (Heatmaps): Για την οπτικοποίηση πινάκων συσχέτισης.

Πρακτική Εφαρμογή:

- Δημιουργία Γραφημάτων: Οι συμμετέχοντες δημιουργούν όλους τους παραπάνω τύπους γραφημάτων χρησιμοποιώντας το καθαρισμένο dataset της προηγούμενης ενότητας.
- Ερμηνεία Γραφημάτων: Κάθε γράφημα συνοδεύεται από μια σύντομη ερμηνεία των στατιστικών ευρημάτων που αναδεικνύει (π.χ., "το ιστόγραμμα δείχνει θετική ασυμμετρία", "το box plot αποκαλύπτει την ύπαρξη ακραίων τιμών").

Άσκηση Ενοποίησης:

Δημιουργία μιας ολοκληρωμένης αναφοράς με πολλαπλά γραφήματα για την περιγραφή ενός συνόλου δεδομένων, συνοδευόμενη από στατιστικά σχόλια.

6. Εκτιμητική & Διαστήματα Εμπιστοσύνης (Confidence Intervals) (2 ώρες)

Σκοπός:

Η εισαγωγή στην έννοια της εκτίμησης παραμέτρων ενός πληθυσμού με χρήση δειγματικών δεδομένων και ο υπολογισμός διαστημάτων εμπιστοσύνης.

Θεωρητικό Υπόβαθρο:

- Σημειακή Εκτίμηση (Point Estimation): Η χρήση στατιστικών μέτρων (π.χ., δειγματικός μέσος, δειγματική διακύμανση) για την εκτίμηση αντίστοιχων πληθυσμιακών παραμέτρων.
- Κατανομές Δειγματοληψίας (Sampling Distributions): Επανάληψη της έννοιας της κατανομής μιας στατιστικής συνάρτησης (π.χ., δειγματικού μέσου).
- Διαστήματα Εμπιστοσύνης (Confidence Intervals): Η θεμελιώδης έννοια του διαστήματος εμπιστοσύνης για μια παράμετρο. Υπολογισμός διαστημάτων εμπιστοσύνης για τον μέσο ενός πληθυσμού (με γνωστή και άγνωστη διακύμανση).

Ερμηνεία:

Η σωστή στατιστική ερμηνεία ενός διαστήματος εμπιστοσύνης και η διάκρισή του από μια απλή πιθανολογική δήλωση.

Πρακτική Εφαρμογή:

- Υπολογισμός Διαστήματος Εμπιστοσύνης: Χρήση των στατιστικών συναρτήσεων της SciPy (π.χ., `scipy.stats.t.interval``) για τον υπολογισμό διαστημάτων εμπιστοσύνης για τον μέσο ενός δείγματος.
- Προσομοίωση για Επαλήθευση: Διεξαγωγή προσομοίωσης Monte Carlo για την επαλήθευση της ερμηνείας: σε ποιο ποσοστό των επαναλαμβανόμενων δειγμάτων το υπολογιζόμενο διάστημα εμπιστοσύνης περιέχει την πραγματική πληθυσμιακή παράμετρο.

Άσκηση Ενοποίησης:

Εκτίμηση του μέσου ύψους ενός πληθυσμού με βάση ένα δείγμα υψών, υπολογισμός και ερμηνεία του 95% διαστήματος εμπιστοσύνης.

7. Έλεγχος Στατιστικών Υποθέσεων (Hypothesis Testing) (2 ώρες)

Σκοπός:

Η κατανόηση της λογικής του ελέγχου στατιστικών υποθέσεων και η εφαρμογή βασικών παραμετρικών και μη παραμετρικών ελέγχων.

Θεωρητικό Υπόβαθρο:

- Λογική του Ελέγχου Υποθέσεων: Η διατύπωση μηδενικής (H_0) και εναλλακτικής (H_1) υπόθεσης. Η έννοια του στατιστικού ελέγχου (test statistic), της τιμής p (p -value) και του επιπέδου σημαντικότητας (α).
- Σφάλματα Τύπου I και II (Type I & II Errors): Η πιθανότητα απόρριψης μιας αληθούς μηδενικής υπόθεσης και η πιθανότητα αποδοχής μιας ψευδούς.

Παραμετρικοί Έλεγχοι:

- T-test ενός δείγματος (One sample t-test): Σύγκριση του μέσου ενός δείγματος με μια γνωστή τιμή.
- T-test δύο ανεξάρτητων δειγμάτων (Independent two-sample t-test): Σύγκριση των μέσων δύο ανεξάρτητων ομάδων.
- T-test για ζευγαρωτές παρατηρήσεις (Paired t-test): Σύγκριση των μέσων δύο συσχετισμένων ομάδων (π.χ., πριν-μετά).

Μη Παραμετρικοί Έλεγχοι:

Εισαγωγή σε ελέγχους όπως ο Mann-Whitney U (εναλλακτική του t-test για μη κανονικά δεδομένα).

Πρακτική Εφαρμογή:

Εφαρμογή t-tests: Χρήση της `scipy.stats` για την εκτέλεση των διαφόρων τύπων t-test σε κατάλληλα δεδομένα.

Ερμηνεία Αποτελεσμάτων:

Εξαγωγή και ερμηνεία της στατιστικής συνάρτησης (statistic) και της τιμής p , και λήψη απόφασης για την απόρριψη ή μη της μηδενικής υπόθεσης.

Άσκηση Ενοποίησης: Σύγκριση Μέσων Όρων.

Χρήση ενός dataset (π.χ., που περιέχει δεδομένα για δύο διαφορετικές ομάδες, όπως άνδρες/γυναίκες, ή δύο διαφορετικές χρονικές στιγμές) και εφαρμογή του κατάλληλου t-test για να διερευνηθεί αν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των μέσων όρων των δύο ομάδων.

8. Ανάλυση Διακύμανσης (ANOVA) & Πίνακες Συνάφειας (2 ώρες)

Σκοπός:

Η επέκταση της σύγκρισης μέσων όρων σε περισσότερες από δύο ομάδες και η ανάλυση της σχέσης μεταξύ κατηγορικών μεταβλητών.

Θεωρητικό Υπόβαθρο:

- ANOVA (Analysis of Variance): Η λογική της ανάλυσης διακύμανσης για τη σύγκριση μέσω των όρων τριών ή περισσότερων ομάδων. Διάκριση μεταξύ ενός παράγοντα (one way) και δύο παραγόντων (two way) ANOVA.
- F-statistic: Το στατιστικό μέτρο που χρησιμοποιείται στην ANOVA.
- Post hoc Tests: Η ανάγκη για περαιτέρω ελέγχους (π.χ., Tukey HSD) όταν η ANOVA υποδεικνύει στατιστικά σημαντική διαφορά, για τον εντοπισμό των συγκεκριμένων ομάδων που διαφέρουν.

Ανάλυση Κατηγορικών Δεδομένων:

- Πίνακες Συνάφειας (Contingency Tables / Cross-tabulation): Πίνακες διπλής εισόδου για την περίληψη της σχέσης μεταξύ δύο κατηγορικών μεταβλητών.
- Έλεγχος χ^2 (Chi-Squared Test): Έλεγχος ανεξαρτησίας μεταξύ δύο κατηγορικών μεταβλητών.

Πρακτική Εφαρμογή:

- Εφαρμογή ANOVA: Χρήση της `scipy.stats.f_oneway` για oneway ANOVA. Οπτικοποίηση των αποτελεσμάτων με box plots για κάθε ομάδα.
- Δημιουργία Πινάκων Συνάφειας: Χρήση της `pd.crosstab()` για τη δημιουργία πινάκων συνάφειας από το dataset.
- Εφαρμογή Ελέγχου χ^2 : Εφαρμογή της `scipy.stats.chi2_contingency` για τον έλεγχο ανεξαρτησίας.

Άσκηση Ενοποίησης:

Ανάλυση της σχέσης μεταξύ εκπαιδευτικού επιπέδου και επαγγελματικής κατηγορίας ή σύγκριση της αποτελεσματικότητας τριών διαφορετικών διαφημιστικών καμπανιών (ANOVA).

9. Συσχέτιση & Απλή Γραμμική Παλινδρόμηση (2 ώρες)

Σκοπός:

Η ποσοτικοποίηση και μοντελοποίηση της γραμμικής σχέσης μεταξύ δύο συνεχών μεταβλητών.

Θεωρητικό Υπόβαθρο:

- Συντελεστής Συσχέτισης (Correlation Coefficient):
 - Pearson's r: Μέτρηση της έντασης και της κατεύθυνσης της γραμμικής σχέσης μεταξύ δύο συνεχών μεταβλητών.
 - Spearman's ρ : Μη παραμετρική εναλλακτική που βασίζεται στις τάξεις (ranks) και ανιχνεύει μονοτονικές σχέσεις.
- Απλή Γραμμική Παλινδρόμηση (Simple Linear Regression): Η έννοια της εύρεσης της "βέλτιστης" ευθείας (ελαχίστων τετραγώνων) που περιγράφει τη σχέση μεταξύ μιας ανεξάρτητης (X) και μιας εξαρτημένης (Y) μεταβλητής.

- Συντελεστής Παλινδρόμησης (Slope): Ερμηνεία ως η μεταβολή της Y για μια μονάδα μεταβολής της X .
- Συντελεστής Προσδιορισμού (R^2): Το ποσοστό της διακύμανσης της Y που εξηγείται από το μοντέλο.

Πρακτική Εφαρμογή:

- Υπολογισμός Συσχέτισης: Χρήση της `scipy.stats.pearsonr` και `scipy.stats.spearmanr` για τον υπολογισμό των συντελεστών συσχέτισης και των p -values τους.
- Οπτικοποίηση: Δημιουργία scatter plots με την ευθεία παλινδρόμησης (χρησιμοποιώντας `seaborn.lmplot`).
- Εφαρμογή Γραμμικής Παλινδρόμησης: Χρήση της `scipy.stats.linregress` ή του `statsmodels` (προετοιμασία για την επόμενη ενότητα) για την εκτίμηση των παραμέτρων του μοντέλου, την αξιολόγηση της στατιστικής σημαντικότητας και τον υπολογισμό του R^2 .

Άσκηση Ενοποίησης:

Ανάλυση της σχέσης μεταξύ διαφημιστικής δαπάνης και πωλήσεων ή μεταξύ ετών εκπαίδευσης και εισοδήματος, συμπεριλαμβανομένου του ελέγχου συσχέτισης, της δημιουργίας μοντέλου παλινδρόμησης και της ερμηνείας των αποτελεσμάτων.

10. Πολλαπλή Παλινδρόμηση & Τελικό Project (2 ώρες)

Σκοπός:

Η επέκταση της γραμμικής παλινδρόμησης σε πολλαπλές ανεξάρτητες μεταβλητές, η εισαγωγή στη βιβλιοθήκη Statsmodels για λεπτομερή στατιστική έξοδο και η σύνθεση όλων των γνώσεων σε ένα ολοκληρωμένο στατιστικό έργο.

Θεωρητικό Υπόβαθρο:

- Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression): Η έννοια της μοντελοποίησης μιας εξαρτημένης μεταβλητής ως γραμμικού συνδυασμού δύο ή περισσότερων ανεξάρτητων μεταβλητών.
- Ερμηνεία Συντελεστών: Η ερμηνεία κάθε συντελεστή, κρατώντας τις υπόλοιπες μεταβλητές σταθερές (*ceteris paribus*).
- Στατιστική Σημαντικότητα και Διαστήματα Εμπιστοσύνης για Συντελεστές: Αξιολόγηση της σημαντικότητας κάθε ανεξάρτητης μεταβλητής.
- Προϋποθέσεις (Assumptions) Γραμμικής Παλινδρόμησης: Μια σύντομη επισκόπηση (γραμμικότητα, ανεξαρτησία σφαλμάτων, ομοσκεδαστικότητα, κανονικότητα σφαλμάτων) και η σημασία τους.
- Εισαγωγή στη Βιβλιοθήκη Statsmodels: Παρουσίαση της βιβλιοθήκης που προσφέρει λεπτομερή στατιστική έξοδο, παρόμοια με εξειδικευμένα στατιστικά πακέτα (π.χ., R, SPSS).

Τελικό Project: Ολοκληρωμένη Στατιστική Ανάλυση Πραγματικού Dataset.

Στόχος:

Οι συμμετέχοντες (ατομικά ή σε ομάδες) αναλαμβάνουν ένα πλήρες έργο στατιστικής ανάλυσης, ακολουθώντας όλα τα βήματα που διδάχθηκαν.

1. Επιλογή και Φόρτωση Dataset: Επιλογή ενός σύνθετου, πραγματικού dataset (π.χ., από δημόσιες πηγές, ανοιχτά δεδομένα).

2. Διερευνητική Ανάλυση & Καθαρισμός: Εφαρμογή τεχνικών Pandas για καθαρισμό και προετοιμασία.

3. Οπτικοποίηση: Δημιουργία γραφημάτων που αναδεικνύουν βασικές κατανομές και σχέσεις.

4. Στατιστική Συμπερασματολογία:

Διατύπωση συγκεκριμένων ερευνητικών ερωτημάτων (π.χ., "Υπάρχει διαφορά στο μέσο εισόδημα μεταξύ ανδρών και γυναικών;").

Επιλογή και εφαρμογή κατάλληλων στατιστικών ελέγχων (t-test, ANOVA, χ^2).

Υπολογισμός και ερμηνεία διαστημάτων εμπιστοσύνης.

5. Μοντελοποίηση:

Διερεύνηση της σχέσης μεταξύ μεταβλητών με συσχέτιση.

Κατασκευή ενός μοντέλου πολλαπλής γραμμικής παλινδρόμησης (χρήση Statsmodels) για την πρόβλεψη μιας βασικής μεταβλητής ή για την κατανόηση των παραγόντων που την επηρεάζουν.

Ερμηνεία των συντελεστών και αξιολόγηση της στατιστικής σημαντικότητάς τους.

6. Συμπεράσματα & Παρουσίαση: Σύνθεση των ευρημάτων σε μια συνεκτική αναφορά (π.χ., Jupyter Notebook) που περιλαμβάνει κείμενο, κώδικα, γραφήματα και στατιστική ερμηνεία, απαντώντας στα αρχικά ερευνητικά ερωτήματα.

Νομική Σημείωση: Η παρούσα προσφορά υπηρεσιών αποτελεί αποκλειστική πνευματική ιδιοκτησία και πνευματικό δικαίωμα της Ροή+ Οργανισμός Εξειδικευμένης Εκπαίδευσης Ι.Κ.Ε (η Εταιρεία). Κάθε μορφή αναπαραγωγής, αναδιανομής, αντιγραφής, τροποποίησης, δημοσίευσης ή εκμετάλλευσης του περιεχομένου, εν όλω ή εν μέρει, χωρίς προηγούμενη έγγραφη άδεια της Εταιρείας απαγορεύεται ρητά και επισύρει νομικές κυρώσεις σύμφωνα με την ελληνική και ευρωπαϊκή νομοθεσία περί πνευματικής ιδιοκτησίας.

Το παρόν έγγραφο προορίζεται αποκλειστικά για τον σκοπό επικοινωνίας έναρξης τμημάτων σεμιναρίων της Ροή+ Οργανισμός Εξειδικευμένης Εκπαίδευσης Ι.Κ.Ε. Η αποκάλυψη, διαφυγή ή δημοσιοποίηση του περιεχομένου πέραν αυτού του σκοπού χωρίς ρητή συναίνεση της Εταιρείας απαγορεύεται. Η εταιρεία διατηρεί όλα τα νομικά δικαιώματα σε περίπτωση παραβίασης των ανωτέρω δικαιωμάτων.