

	<b>ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ</b>
	<b>ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ &amp; ΤΗΛΕΜΑΤΙΚΗΣ</b> <b>ΠΜΣ "ΕΦΑΡΜΟΣΜΕΝΗ ΠΛΗΡΟΦΟΡΙΚΗ"</b>
	<b>ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ: ΕΡΓΑΣΤΗΡΙΟ Python</b>
	ΕΡΓΑΣΙΑ : «Εισαγωγή στη Στατιστική: Περιγραφική Στατιστική- Στατιστική Συμπερασματολογία»

Στα πλαίσια της έρευνας που αφορούσε το πλήθος γραπτών μηνυμάτων SMS που αποστέλλουν μηνιαίως χρήστες κινητής τηλεφωνίας ηλικίας 18-25 κατοίκων μιας περιοχής της Δυτικής Αττικής, επιλέχθηκε ένα τυχαίο δείγμα 100 κατοίκων από τον αντίστοιχο πληθυσμό. Για τα άτομα του τυχαίου αυτού δείγματος, καταγράφηκε το πλήθος των γραπτών μηνυμάτων SMS που απέστειλαν από το κινητό τους τηλέφωνο ένα συγκεκριμένο μήνα, η εταιρεία τηλεπικοινωνιών TELECOM που ήταν συνδρομητές, το φύλο τους MF, η χρήση VIBER για αποστολή μηνυμάτων, η χρήση WhatsApp για τηλεφωνική επικοινωνία, τα μηνιαία λεπτά ομιλίας TIME, το μηνιαίο πλήθος των μηνυμάτων SMS\_AFTER και ο μηνιαίος χρόνος ομιλίας TIME\_AFTER μετά από μία διαφημιστική εκστρατεία υπέρ της χρήσης μηνυμάτων.

#### **A: Εισαγωγή και Κωδικοποίηση Στοιχείων**

A.1) Να εισάγετε με κατάλληλο τρόπο τα παραπάνω δεδομένα στην Colab\_Python από το αρχείο τύπου excel "Lab Data Python New".

#### **B: Περιγραφικά Μέτρα-Γραφικοί Τρόποι Παρουσίασης**

B.1) Να υπολογίσετε τα κυριότερα μέτρα θέσης (μέση τιμή, διάμεσος, ελάχιστη και μέγιστη παρατήρηση, πρώτο και τρίτο τεταρτημόριο), μέτρα διασποράς (διασπορά, τυπική απόκλιση και εύρος τιμών) και μέτρα μορφής (συντελεστής ασυμμετρίας) της μεταβλητής «TIME». Να κατασκευαστεί το αντίστοιχο ιστόγραμμα συχνοτήτων καθώς και το θηκόγραμμα. Σχολιάστε.

B.2) Να υπολογίσετε τα κυριότερα μέτρα θέσης (μέση τιμή, διάμεσος, ελάχιστη και μέγιστη παρατήρηση, πρώτο και τρίτο τεταρτημόριο), μέτρα διασποράς (διασπορά, τυπική απόκλιση και εύρος τιμών) και μέτρα μορφής (συντελεστής ασυμμετρίας) της μεταβλητής «TIME» για κάθε φύλο ξεχωριστά. Να κατασκευαστούν τα αντίστοιχα θηκογράμματα. Σχολιάστε.

**Γ: Στατιστική Συμπερασματολογία-Έλεγχοι Υποθέσεων (σε επίπεδο σημαντικότητας  $\alpha=5\%$ )**

Γ.1) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι ο μέσος χρόνος ομιλίας TIME στον αντίστοιχο πληθυσμό διαφέρει στατιστικά σημαντικά από τις 100 μονάδες. Σχολιάστε.

Γ.2) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι ο μέσος χρόνος ομιλίας TIME από γυναίκες στον αντίστοιχο πληθυσμό διαφέρει στατιστικά σημαντικά από τον μέσο χρόνο ομιλίας TIME από άνδρες. Σχολιάστε.

Γ.3) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι μηνιαίο ποσοστό χρήσης WhatsApp για τηλεφωνική επικοινωνία από γυναίκες στον αντίστοιχο πληθυσμό διαφέρει στατιστικά σημαντικά από το αντίστοιχο μηνιαίο ποσοστό χρήσης WhatsApp για τηλεφωνική επικοινωνία από άνδρες. Σχολιάστε.

Γ.4) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι η διαφημιστική εκστρατεία των εταιρειών κινητής τηλεφωνίας υπέρ της χρήσης SMS, μείωσε το μέσο χρόνο ομιλίας TIME. Σχολιάστε.

Γ.5) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι ο μέσος χρόνος ομιλίας TIME, διαφέρει στατιστικά σημαντικά ανάμεσα στις τρεις εταιρείες τηλεπικοινωνίας. Σχολιάστε.

**Δ: Πίνακες Συνάφειας (Διπλής Εισόδου)-Έλεγχος Ανεξαρτησίας- $\chi^2$**

Δ.1) Να κατασκευαστεί ο πίνακας διπλής εισόδου της μεταβλητής MF και της μεταβλητής WhatsApp.

Δ.2) Να δοθούν και να ερμηνευτούν οι περιθώριες σχετικές συχνότητες για τις δύο μεταβλητές.

Δ.3) Να δοθούν και να ερμηνευτούν οι δεσμευμένες σχετικές συχνότητες ως προς το φύλο.

Δ.4) Να διατυπώσετε τον κατάλληλο έλεγχο υποθέσεων για τον ισχυρισμό ότι η χρήση WhatsApp για τηλεφωνική επικοινωνία εξαρτάται από το φύλο. Σχολιάστε.

Δ.5) Να μετασχηματιστεί η ποσοτική μεταβλητή TIME σε πεντάτιμη ιεραρχική (κατηγορική) μεταβλητή TIMECATEGORICAL και να ελεγχθεί καταλλήλως η εξάρτηση της ιεραρχικής μεταβλητής TIMECATEGORICAL από την ποιοτική MF. Σχολιάστε.