**Εργασία #6 (27/12/2021)**

Χρησιμοποιήστε την ιστοσελίδα για τα Αρχεία Δημόσιας Χρήσης (https://www.statistics.gr/el/public-use-files) της Ελληνικής Στατιστικής Αρχής (ΕΛ.ΣΤΑΤ) και ειδικότερα τον σύνδεσμο <https://elstatforms.statistics.gr/open_micro/login> ακολουθώντας τις οδηγίες που παρέχονται στην συγκεκριμένη ιστοσελίδα με σκοπό να ανακτήσετε το αρχείο zip της Έρευνας Εργατικού Δυναμικού για το Δ' τρίμηνο του έτους 2020 (LFS\_2020D\_users). Στη συνέχεια αποσυμπιέζοντας το zip αρχείο σε τοπικό φάκελο του Η/Υ σας θα έχετε στην διάθεση σας τα πρωτογενή δεδομένα της Έρευνας Εργατικού Δυναμικού για το Δ' τρίμηνο του έτους 2020 (LFS\_2020D\_users.csv). Ακολουθώντας τα προηγούμενα βήματα για το Α’ τρίμηνο του έτους 2020 και αφού ανακτήσετε το αρχείο LFS\_2020A\_users.zip, θα δείτε ότι μετά την αποσυμπίεση του zip αρχείου θα έχετε πρόσβαση στο αρχείο LFS codebook\_1987-2020-gr.xls το οποίο περιγράφει τις μεταβλητές που περιλαμβάνονται στα csv αρχεία με τα πρωτογενής δεδομένα.

Χρησιμοποιώντας το Stata και πληκτρολογώντας την εντολή import delimited "C: ....your directory.... LFS\\_2019B'\\_users.csv", delimiter(";") clear μπορείτε να εισάγετε τα δεδομένα που θα επεξεργαστείτε στην συνέχεια.

Χρησιμοποιείστε τις ακόλουθες μεταβλητές:

Φύλο: a07

Ηλικία: age

Χώρα γέννησης: a13\_reu

Οικογενειακή κατάσταση: a11\_r

Περιφέρεια διαμονής: ypa\_r

Έτος ολοκλήρωσης σπουδών: e82

Ανώτατο ολοκληρωμένο επίπεδο εκπαίδευσης: e80\_2

Αντικείμενο εκπαίδευσης: e81

Τρέχουσα κατάσταση απασχόλησης: katap

Κύρια ασχολία 1 χρόνο πριν: e91

Θέση στην απασχόληση: e17\_r3

Ιδιοκτησιακό καθεστώς επιχείρησης: e15

Επάγγελμα κύριας απασχόλησης (1-ψήφιο isco-08): e16\_1

Κλάδος οικονομικής δραστηριότητας (1-ψήφιο ΣΤΑΚΟΔ-08): e14\_1

Συνήθεις ημέρες εργασίας την εβδομάδα: e25\_days

Ώρες συνήθους απασχόλησης την εβδομάδα: e25\_orr

Πλήρης ή Μερική Απασχόληση: e23

Συνολικές καθαρές αποδοχές τον τελευταίο μήνα e95

Πόσο συχνά εργάστηκε στο σπίτι: e43

Για όλες τις μεταβλητές (πλην της e14\_1) χρησιμοποιήστε την εντολή destring ώστε να μετασχηματίσετε τις τιμές σε αριθμητική μορφή. Για την e14 κατασκευάστε τις κατάλληλες ψευδομετβλητές που θα αναφέρονται στις κατηγορίες που περιγράφουν οι τιμές τις μεταβλητής αυτής. Για την εκπόνηση της εργασίας σας να επιλέξετε άτομα στην ηλικιακή ομάδα 25-54 ετών.

Ερωτήσεις

1. Έχοντας στη διάθεση σας τις παραπάνω μεταβλητές, και αφού κατανοήσετε το περιεχόμενο των μεταβλητών αυτών
	1. παρουσιάστε (σε πίνακα) για το δείγμα ατόμων στην ηλικιακή ομάδα 25-54 τα ακόλουθα ποσοστά για το σύνολο και για άνδρες και γυναίκες, ξεχωριστά (Ποσοστό συμμετοχής στην αγορά εργασίας, Ποσοστό απασχόλησης στον πληθυσμό, Ποσοστό ανεργίας). Επίσης παρουσιάστε στον ίδιο πίνακα τα αντίστοιχα ποσοστά και για κάθε μια από 6 πενταετίες ομάδων ηλικιών ανά φύλο.
	2. Ποιες είναι οι διαφορές στα παραπάνω ποσοστά μεταξύ αυτών που έχουν γεννηθεί στην Ελλάδα (γηγενείς) και σε οποιαδήποτε άλλη χώρα (μετανάστες).
	3. Ποια είναι η κατανομή του πληθυσμού σε κάθε περιφέρεια διαμονής ανά φύλο, οικογενειακή κατάσταση (έγγαμοι/άλλη κατάσταση), πενταετή ηλικιακή ομάδα και χώρα γέννησης (γηγενείς/μετανάστες).
2. Χρησιμοποιώντας την μεταβλητή για το έτος ολοκλήρωσης σπουδών και γνωρίζοντας την ηλικία του ατόμου το έτος 2020 (yborn=2020-age) κατασκευάστε μια νέα μεταβλητή για τα έτη εκπαίδευσης χρησιμοποιώντας την εξίσωση: yrsedu = e82 - (yborn + 6).
	1. Ποιος είναι ο μέσος όρος των ετών εκπαίδευσης (στρογγυλοποίηση δίχως δεκαδικά) για καθεμιά κατηγορία του ανώτατου ολοκληρωμένου επιπέδου εκπαίδευσης;
	2. Παρουσιάστε το μέσο όρο, τυπική απόκλιση και αριθμό παρατηρήσεων για τα έτη εκπαίδευσης ανά φύλο, οικογενειακή κατάσταση (έγγαμοι/άλλη κατάσταση), πενταετή ηλικιακή ομάδα και χώρα γέννησης (γηγενείς/μετανάστες).
	3. Ποια είναι η κατανομή του αντικειμένου εκπαίδευσης με βάση το φύλο και τη χώρα γέννησης;
3. Εφόσον γνωρίζετε την κατάσταση ασχολίας 1 χρόνο πριν αλλά και την τρέχουσα κατάσταση απασχόλησης,
	1. παρουσιάστε σε πίνακα τα ακόλουθα ποσοστά για άνδρες και γυναίκες, ξεχωριστά:
		1. Ποσοστό ανεργίας όλων όσων 1 χρόνο πριν ήταν απασχολούμενοι.
		2. Ποσοστό ανεργίας όλων όσων 1 χρόνο πριν ήταν άνεργοι.
		3. Ποσοστό ανεργίας όλων όσων 1 χρόνο πριν βρίσκονταν εκτός εργατικού δυναμικού.
	2. Εφόσον γνωρίζετε την θέση στην απασχόληση παρουσιάστε σε πίνακα τα ποσοστά της μισθωτής εργασίας ανά πενταετή ηλικιακή ομάδα για άνδρες και γυναίκες, ξεχωριστά.
	3. Ποια είναι η κατανομή της απασχόλησης των μισθωτών (ανά φύλο) με βάση το ιδιοκτησιακό καθεστώς της επιχείρησης; Ποια είναι τα ποσοστά απασχόλησης των μεταναστών με βάση το ιδιοκτησιακό καθεστώς της επιχείρησης;
	4. Ποια είναι η κατανομή της απασχόλησης των μισθωτών (ανά φύλο) με βάση το επάγγελμα κύριας απασχόλησης; Ποια είναι τα ποσοστά απασχόλησης των μεταναστών με βάση το επάγγελμα κύριας απασχόλησης;
	5. Ποια είναι η κατανομή της απασχόλησης των μισθωτών (ανά φύλο) με βάση τον κλάδο οικονομικής δραστηριότητας; Ποια είναι τα ποσοστά απασχόλησης των μεταναστών με βάση τον κλάδο οικονομικής δραστηριότητας;
4. Ποια είναι η κατανομή των ημερών απασχόλησης την εβδομάδα για μισθωτούς και αυτο-απασχολούμενους με βάση το φύλο (για άνδρες και γυναίκες, ξεχωριστά);
	1. Πόσες ώρες δουλεύουν κατά μέσο όρο την εβδομάδα οι άνδρες και οι γυναίκες ξεχωριστά; Είναι στατιστικά σημαντική αυτή η διαφορά; Παρουσιάστε τις διαφορές στις εβδομαδιαίες ώρες απασχόλησης και τη στατιστική σημαντικότητα αυτών των διαφορών μεταξύ ανδρών και γυναικών ανά πενταετή ομάδα ηλικίας.
	2. Ποια είναι τα επαγγέλματα με το μεγαλύτερο αριθμό εβδομαδιαίων ωρών εργασίας; Αρκετοί υποστηρίζουν ότι οι αυτο-απασχολούμενοι εργάζονται περισσότερες ώρες την εβδομάδα σε σχέση με τους μισθωτούς εργαζόμενους. Με βάση τα στοιχεία που έχετε στη διάθεση σας συμφωνείτε με αυτή την άποψη;
	3. Πώς διαμορφώθηκε το ποσοστό των μισθωτών που εργάστηκαν από το σπίτι (Τουλάχιστον τις μισές εργάσιμες μέρες ή Λιγότερο συχνά) ανά κατηγορία επαγγέλματος;
5. Με βάση τα στοιχεία που έχετε στη διάθεση σας για τις συνολικές καθαρές αποδοχές των μισθωτών εργαζόμενων παρουσιάστε τις μισθολογικές διαφορές (σε ποσοστά) μεταξύ ανδρών-γυναικών, γηγενών-μεταναστών, έγγαμων-μη έγγαμων και ατόμων ηλικίας 25-34 και ατόμων ηλικίας 45-54. Δείξτε ότι αυτές οι διαφορές είναι σημαντικές σε στατιστικούς όρους.
	1. Ποια είναι τα καλύτερα αμειβόμενα επαγγέλματα και ποιοι κλάδοι προσφέρουν τις καλύτερες μηνιαίες αποδοχές;
	2. Έστω ότι έχετε ένα δείγμα μισθωτών εργαζομένων με πλήρη απασχόληση για τους οποίους γνωρίζετε το ύψος των μηνιαίων συνολικών καθαρών αποδοχών και δεδομένου ότι οι συνολικές ώρες εργασίας το μήνα μπορούν να υπολογιστούν προσεγγιστικά ως το γινόμενο (4.2$xεβδομαδιαίες ώρες εργασίας) κατασκευάστε μια νέα μεταβλητή για το ωρομίσθιο (λόγος μηνιαίων αποδοχών προς σύνολο ωρών εργασίας το μήνα) και παρουσιάστε γραφικά την κατανομή της νέας μεταβλητής (σε λογαριθμική μορφή) για άνδρες και γυναίκες.
	3. Σε συνέχεια του προηγούμενου υπο-ερωτήματος και χρησιμοποιώντας ως εξαρτημένη μεταβλητή το ωρομίσθιο να εκτιμηθούν βήμα-βήμα τα ακόλουθα υποδείγματα με την μέθοδο των ελαχίστων τετραγώνων όπου θα χρησιμοποιούνται (προσθετικά) ως ανεξάρτητες μεταβλητές οι ακόλουθες:
* φύλο (γυναίκα =1), ηλικία (5ετείς ομάδες), οικογενειακή κατάσταση (έγγαμος =1)
* εκπαίδευση (πτυχίο τριτοβάθμιας εκπαίδευσης και άνω =1)
* χώρα γέννησης (γηγενείς =1)
* επάγγελμα και κλάδος
* περιφέρεια διαμονής
1. Έχοντας κατασκευάσει μια δυαδική μεταβλητή για την συμμετοχή στην αγορά εργασίας εκτιμήστε ένα υπόδειγμα Probit (και παρουσιάστε αναλυτικά τα ευρήματα σας) για την συμμετοχή στην αγορά εργασίας των γυναικών χρησιμοποιώντας την ακόλουθη δέσμη μεταβλητών:
* ηλικία (5ετείς ομάδες), οικογενειακή κατάσταση (έγγαμος =1)
* εκπαίδευση (πτυχίο τριτοβάθμιας εκπαίδευσης και άνω =1)
* χώρα γέννησης (γηγενείς =1)
* περιφέρεια διαμονής

Στη αναφορά (μορφή άρθρου) που θα μου επιστρέψετε (21 Ιανουαρίου 2022) να συμπεριλαμβάνεται στο email (ngias@upatras.gr) και το .do αρχείο με τις εντολές που έχετε χρησιμοποιήσει.