

19. Περιγραφή και Γενικές Οδηγίες Αξιοποίησης των Στατιστικών Δεδομένων της Έρευνας SHARE ή 'Οδηγός Ναυτιλλομένων' για την χρήση των στοιχείων

Θωμάς Γεωργιάδης

Πρόλογος

Το παρόν κείμενο απευθύνεται πρωτίστως σε ερευνητές που σκοπεύουν να χρησιμοποιήσουν τα πρωτογενή δεδομένα της έρευνας SHARE. Οι πληροφορίες για όλες τις χώρες και όλα τα κύματα είναι διαθέσιμες ανά άτομο ή νοικοκυριό, υπάρχουν οι πληροφορίες που έδωσαν οι ερευνώμενοι, αλλά και δευτερογενείς πληροφορίες και αναγωγές που έγιναν εκ των υστέρων. Αποτέλεσμα είναι ότι, ενώ η πρόσβαση είναι κατ' αρχήν ελεύθερη, η πολυπλοκότητα των στοιχείων και ο όγκος τους να προβάλλουν ως εμπόδια στην αξιοποίησή του, ακόμη και από έμπειρους ερευνητές. Σκοπός του κειμένου είναι να προσφέρει έναν 'οδηγό ναυτιλλομένων' ο οποίος να καταδεικνύει ότι τα προβλήματα που μπορεί να φανταστεί κανείς ότι υπάρχουν έχουν λύσεις, ενώ ερωτήσεις σε κοινές απαντήσεις μπορούν να αναζητηθούν σε συγκεκριμένες τοποθεσίες. Συνεπώς, το κείμενο που ακολουθεί παρουσιάζει τη δυνατότητα πρόσβασης στα δεδομένα (πώς είναι διαθέσιμα), την περιγραφή των τύπων των δεδομένων αυτών (τι δεδομένα είναι διαθέσιμα), επισημαίνοντας ορισμένες διευκρινήσεις και οδηγίες (απαντήσεις σε βασικά ερωτήματα πριν ξεκινήσει η ανάλυση των δεδομένων).

1. Πρόσβαση σε δεδομένα και ερευνητικό υλικό

Η επίσημη ιστοσελίδα της SHARE (Survey on Health, Ageing and Retirement) είναι η <http://www.share-project.org/> η οποία παρέχει πληροφόρηση και υλικό (ερωτηματολόγια, πρόσβαση σε δεδομένα, δημοσιεύσεις, δελτία τύπου κ.α.) της έρευνας SHARE για την Υγεία, τη Γήρανση και τη Συνταξιοδότηση στην Ευρώπη. Πρόσβαση στο ίδιο υλικό με έμφαση στις απαιτήσεις του Έλληνα χρήστη υπάρχει και από την Ελληνική ιστοσελίδα της SHARE (www.share-project.gr ή www.ageing.gr), στην οποία συγκεντρώνεται και γενικότερο υλικό που έχει σχέση με την γήρανση στη Ελλάδα.

1.1 Πρόσβαση στα στατιστικά δεδομένα

Τα στατιστικά δεδομένα από το πρώτο κύμα (2004), καθώς και από το δεύτερο (2006/2007) της έρευνας SHARE είναι διαθέσιμα στην ιστοσελίδα της έρευνας (χωρίς να απαιτείται χρέωση ή οικονομική επιβάρυνση). Μόνη προϋπόθεση αποτελεί η αποστολή από τον κάθε ερευνητή ενός fax με την απαιτούμενη δήλωση δικαιώματος χρήσης, η οποία απλώς πιστοποιεί ότι η πρόσβαση γίνεται για ερευνητικούς και όχι για εμπορικούς σκοπούς. Η συγκεκριμένη δήλωση είναι διαθέσιμη στην ιστοσελίδα της έρευνας. Στη συνέχεια ο οργανισμός ο οποίος αποθηκεύει τα στοιχεία (CenterData) αποστέλλει ένα ζεύγος κωδικών μοναδικών για κάθε ερευνητή. Έχοντας λάβει τους συγκεκριμένους κωδικούς, η πρόσβαση στα πρωτογενή δεδομένα γίνεται ως εξής: από την ιστοσελίδα <http://www.share-project.org/> επιλέγουμε την περιοχή "Access the Data" και στη συνέχεια "download the data". Έπειτα, ο ερευνητής εισάγει τους κωδικούς πρόσβασης στην ιστοσελίδα (<http://cdata8.uvt.nl/sharedatadissemination/users/login>).

Σημειώνεται ότι η Ελληνική ομάδα SHARE στο πλαίσιο ενθάρρυνσης της χρήσης των στοιχείων, εκπονεί μια σειρά βάσεων απλούστερων στην δομή και προσανατολισμένων σε χρήση από φοιτητές στο πλαίσιο εξοικείωσης με την χρήση μικρο-στοιχείων στην κοινωνική

έρευνα (EasySHARE-EL). Οι βάσεις αυτές θα είναι προσβάσιμες από τον Σεπτέμβριο 2009 μέσω της Ελληνικής ιστοσελίδας (www.share-project.gr).

1.2 Πρόσβαση σε λοιπό ερευνητικό υλικό

Η πρόσβαση σε άλλες πληροφορίες και ερευνητικό υλικό της έρευνας SHARE, όπως ερωτηματολόγια, κάρτες επίδειξης, δημοσιεύσεις, δελτία τύπου, περιγραφή και κριτήρια επιλογής δείγματος γίνεται μέσω της ιστοσελίδας <http://www.share-project.org/> χωρίς να απαιτείται καμιά διατύπωση. Τα ερωτηματολόγια και οι κάρτες επίδειξης τόσο για το πρώτο όσο και το δεύτερο κύμα είναι διαθέσιμα στα αγγλικά, καθώς και μεταφρασμένα στη γλώσσα κάθε χώρας που συμμετέχει στην έρευνα SHARE. Το ελληνικό ερωτηματολόγιο βρίσκεται και σε σχολιασμένη μορφή στην ελληνική ιστοσελίδα. Ένας ερευνητής θα ανατρέξει στα ερωτηματολόγια για δύο κυρίως λόγους:

- Πρώτον, θα βρει την ακριβή διατύπωση της ερώτησης σε όλες τις γλώσσες.
- Δεύτερον, θα διαπιστώσει τη «λογική ακολουθία» της συνέντευξης. Οι διαδικασίες “if” καθώς και τα “loops” που υπάρχουν σημαίνουν ότι δεν απαντήθηκαν όλες οι ερωτήσεις από όλους (π.χ. οι συνταξιούχοι δεν απαντούν για την ικανοποίησή τους από την εργασία). Το ερωτηματολόγιο πληροφορεί αναφορικά με το σε ποιους απευθύνεται η κάθε ερώτηση, πράγμα που επεξηγεί την ύπαρξη **κενών τιμών (missing values)**¹.

Μία ακόμη σημαντική πληροφορία που παρουσιάζεται στην ιστοσελίδα της έρευνας SHARE αφορά τις **αποκλίσεις μεταξύ των χωρών (country deviations)** σε συγκεκριμένα σημεία του ερωτηματολογίου. Οι αποκλίσεις αυτές, οι οποίες είναι περιορισμένες σε αριθμό, μπορεί να οφείλονται σε θεσμικές διαφορές μεταξύ των χωρών ή σε ιδιαιτερότητες των χωρών. Για παράδειγμα, στην γενική ερώτηση HC053 “Please look at card 18. What is your health insurance category in the National Health Insurance System?” οι κατηγορίες των απαντήσεων στην Ελλάδα αντιστοιχούν στα Ταμεία Ασφάλισης της Ελλάδας. Ως εκ τούτου, η συγκεκριμένη ερώτηση πρέπει να διερευνηθεί στην περίπτωση της Ελλάδας χωριστά, καθώς στις άλλες χώρες οι ίδιες κατηγορίες απαντήσεων αντιστοιχούν σε διαφορετικές πληροφορίες. Η ενημέρωση για αυτές τις αποκλίσεις μπορεί να γίνει μέσω των παρακάτω βημάτων: από την ιστοσελίδα <http://www.share-project.org/> επιλέγουμε την περιοχή “Item Correspondence” και στη συνέχεια “Country specific deviations within wave 1”.

2. Περιγραφή των διαθέσιμων στατιστικών δεδομένων

Τα στατιστικά δεδομένα της έρευνας SHARE είναι διαθέσιμα με δύο εναλλακτικές μορφές που επιλέγει ο ερευνητής: ως αρχεία δεδομένων του προγράμματος STATA (αρχεία με κατάληξη .dta), καθώς και ως αρχεία δεδομένων του προγράμματος SPSS (αρχεία με κατάληξη .sav). Έτσι, οι χρήστες των στατιστικών προγραμμάτων STATA ή SPSS μπορούν να κατεβάσουν τα δεδομένα αναλόγως με το πρόγραμμα που χρησιμοποιούν. Αν και οι δύο αυτοί τρόποι καλύπτουν την πλειοψηφία των ερευνητών, χρήστες άλλων οικονομετρικών / στατιστικών προγραμμάτων (π.χ. SAS) μπορούν να κατεβάσουν τα δεδομένα σε μορφή αρχείων STATA ή SPSS και έπειτα να τα «εισάγουν» στο πρόγραμμα που χρησιμοποιούν.

¹ Μετά την πρώτη απόδοση των αγγλικών όρων στα ελληνικά (παρουσιάζεται σε αντιπαράβολή με έντονα γράμματα), ο συντάκτης του άρθρου επιλέγει τη χρήση των αγγλικών όρων σε σχέση με τους ελληνικούς για θέματα που αφορούν βασικές έννοιες της έρευνας SHARE. Αυτό γίνεται για να διευκολυνθεί η χρήση των δεδομένων της έρευνας SHARE και για να αποφευχθεί η σύγχυση από την ενδεχομένως λανθασμένη απόδοση των βασικών όρων.

Ως προς το περιεχόμενο, τα στατιστικά δεδομένα SHARE μπορούν να ταξινομηθούν σε τρεις γενικές κατηγορίες:

- i) Τα δεδομένα του βασικού ερωτηματολογίου της έρευνας SHARE τα οποία ομαδοποιούνται σε 21 επιμέρους **θεματικές ενότητες (modules)**, ανάλογα με το τμήμα του κυρίως ερωτηματολογίου CAPI.
- ii) Τα δεδομένα του ερωτηματολογίου “drop off” και των ερωτηματολογίων “vignettes” (ειδικά ερωτηματολόγια που απαντήθηκαν ξεχωριστά από τον ερευνώμενο σε χαρτί).
- iii) Τα δεδομένα με τις **υπολογισθείσες μεταβλητές (generated variables)** από τις ομάδες εργασίας, αξιοποιώντας στοιχεία του ερωτηματολογίου (π.χ. συνολικό εισόδημα, κλίμακα κατάθλιψης EURO-D, κλπ).

2.1 Περιγραφή των στατιστικών δεδομένων του βασικού ερωτηματολογίου

Το ερωτηματολόγιο SHARE έχει προγραμματιστεί σε μια εφαρμογή προσωπικής συνέντευξης με υπολογιστή (CAPI) και αποτελείται από 21 modules (Πίνακας 1). Τα στατιστικά δεδομένα είναι διαθέσιμα για κάθε module χωριστά και κάθε module έχει παρατηρήσεις για όλες τις χώρες. Η πληροφορία για τις χώρες υπάρχει ως μεταβλητή σε όλα τα modules.

Με εξαίρεση το εισαγωγικό φύλλο σε επίπεδο νοικοκυριού (cv_h αρχείο), όλα τα modules του βασικού ερωτηματολογίου παρουσιάζουν δεδομένα σε επίπεδο ατόμων και όχι νοικοκυριών. Αυτό σημαίνει ότι κάθε οριζόντια γραμμή (παρατήρηση) αντιστοιχεί σε συνέντευξη με ένα άτομο. Στην περίπτωση που η ερώτηση αφορά πληροφορία σε επίπεδο νοικοκυριού (π.χ. καταναλωτική δαπάνη του νοικοκυριού τους τελευταίους δώδεκα μήνες για φαγητό στο σπίτι –βλέπε module κατανάλωσης “co”) καλείται να απαντήσει ένα μόνο άτομο σε κάθε νοικοκυριό. Ακόμη ωστόσο και σε αυτή την περίπτωση τα δεδομένα παρουσιάζονται σε επίπεδο ατόμων, έχοντας την απάντηση για το ύψος της καταναλωτικής δαπάνης του νοικοκυριού στο άτομο που απάντησε, ενώ στα υπόλοιπα μέλη του ίδιου νοικοκυριού υπάρχει missing value στη συγκεκριμένη μεταβλητή. Εφόσον η ανάλυση γίνεται σε επίπεδο νοικοκυριού οι missing values στα υπόλοιπα μέλη του νοικοκυριού δε συνιστούν πρόβλημα.

Πίνακας 1

Δεδομένα των βασικού ερωτηματολογίου της έρευνας SHARE, ανά ενότητα (module)

Ενότητες ερωτηματολογίου έρευνας SHARE			Ποιος απαντά σε κάθε ενότητα;	
#	Όνομα	Θεματική ενότητα (module)	Όλα τα μέλη του νοικοκυριού	Ένα μέλος σε κάθε νοικοκυριό
1	CV_H	Εισαγωγικό φύλλο (επίπεδο νοικοκυριού)		
2	CV_R	Εισαγωγικό φύλλο (επίπεδο ατόμων)		
3	DN	Δημογραφικά	X	
4	PH	Φυσική Υγεία	X	
5	BR	Ρίσκο Συμπεριφοράς	X	
6	CF	Γνωστική Λειτουργία	X	
7	MH	Ψυχική Υγεία	X	
8	HC	Ιατρική Φροντίδα	X	
9	EP	Εργασία και Συντάξεις	X	

10	GS	Δύναμη Χειρολαβής	X	
11	WS	Ταχύτητα Βαδίσματος	X	
12	CH	Παιδιά		X
13	SP	Κοινωνική Υποστήριξη	X	X
14	FT	Χρηματικές Μεταβιβάσεις		X
15	HO	Στέγαση		X
16	HH	Εισόδημα Νοικοκυριού		X
17	CO	Κατανάλωση		X
18	AS	Περιουσιακά Στοιχεία		X
19	AC	Δραστηριότητες	X	
20	EX	Προσδοκίες	X	
21	IV	Παρατηρήσεις Ερευνητή		

CV Εισαγωγικό Φύλλο: Η συνέντευξη αρχίζει με ένα "Εισαγωγικό Φύλλο" που αποτελεί την εισαγωγή στην έρευνα. Το Εισαγωγικό Φύλλο συλλέγει βασικές δημογραφικές πληροφορίες για κάθε έναν που ζει αυτήν την περίοδο στο νοικοκυριό (όνομα, φύλο, έτος και μήνας γέννησης, έτος που έγινε η συνέντευξη κλπ). Επισημαίνεται ότι το Εισαγωγικό Φύλλο σε επίπεδο ατόμων (δηλαδή το cv_r) έχει περισσότερες παρατηρήσεις σε σχέση με τα υπόλοιπα modules. Η διαφορά αυτή έγκειται στο γεγονός ότι στο Εισαγωγικό Φύλλο υπάρχουν όλα τα άτομα που διαμένουν σε ένα επιλέξιμο για συνέντευξη νοικοκυριό, ακόμη και αν αυτά είναι κάτω των 50. Ο ερευνητής πρέπει να μην λησμονεί τη «συγχώνευση» του Εισαγωγικού Φύλλου με άλλες θεματικές ενότητες. Τα δεδομένα του Εισαγωγικού Φύλλου μπορούν να αξιοποιηθούν προκειμένου να δημιουργηθούν χρήσιμες δημογραφικές μεταβλητές (αριθμός μελών στο νοικοκυριό, αριθμός ατόμων ενός ορισμένου εύρους ηλικίας που ζουν στο νοικοκυριό, δείκτης εξάρτησης), αφού αυτές αναφέρονται στο σύνολο των μελών του νοικοκυριού (επιλέξιμοι και μη).

DN Δημογραφικά Στοιχεία: Σε αυτή την ενότητα συλλέγονται δημογραφικές κυρίως πληροφορίες που αφορούν την οικογενειακή κατάσταση, τη χώρα γέννησης, την εκπαίδευση και την απασχόληση κάθε ατόμου που συμμετέχει στην έρευνα. Υπάρχουν, επίσης, λεπτομέρειες για τους γονείς, όπως επαγγελματική κατάσταση, κατάσταση υγείας και συχνότητα επαφής.

PH Φυσική υγεία: Καλύπτει τομείς της υγείας των ατόμων (μακροχρόνιες παθήσεις, προβλήματα στην όραση και ακοή, διαγνώσεις και συμπτώματα ασθενειών, δυσκολίες με δραστηριότητες της καθημερινής ζωής κλπ).

BR Συμπεριφορά κινδύνου: Παρέχει πληροφορίες για τις συμπεριφορές υγείας, όπως το κάπνισμα, η κατανάλωση ποτών και οι σωματικές δραστηριότητες.

CF Γνωστική λειτουργία: Περιέχει μετρήσεις βάσει τεστ τεσσάρων πτυχών της γνωστικής λειτουργίας των ατόμων της έρευνας: i) μνήμη, ii) συγκέντρωση, iii) επίπεδο αριθμητικών γνώσεων, και iv) λεκτική ευχέρεια.

MH Ψυχική υγεία: Συγκεντρώνει πληροφορίες που σχετίζονται με τη ψυχική υγεία, βάσει των οποίων υπολογίζεται η κλίμακα κινδύνου κλινικής κατάθλιψης (Euro-d).

HC Ιατρική φροντίδα: Αυτή η ενότητα σχετίζεται με τις πρόσφατες επισκέψεις σε γιατρούς και νοσοκομεία. Περιέχει, επίσης, ερωτήσεις για δαπάνες υπηρεσιών υγείας, όπως φάρμακα, πρωτοβάθμια περίθαλψη, νοσοκομειακή περίθαλψη κλπ. Για την περίπτωση της Ελλάδας μόνο, περιέχει πληροφορίες και για το ταμείο ασφάλισης των ατόμων (ΙΚΑ, ΟΓΑ κλπ).

EP Απασχόληση και Συντάξεις: Παρουσιάζονται πληροφορίες για την κατάσταση απασχόλησης καθώς και τις εργασιακές δραστηριότητες των ατόμων. Υπάρχουν, επίσης,

πληροφορίες για το εισόδημα από εργασία (για όσους εργάζονται) ή από συντάξεις (για όσους /ες έχουν συνταξιοδοτηθεί).

Οι ενότητες **GS Δύναμη χειρολαβής και WS Ταχύτητα βαδίσματος** παρέχουν άμεσες μετρήσεις της φυσικής κατάστασης που γίνονται κατά την διάρκεια της συνέντευξης. Η δοκιμασία GS μετρά την δύναμη χειρολαβής των ατόμων με τη βοήθεια ενός δυναμόμετρου, ενώ η δοκιμασία **WS** μετρά τον χρόνο που χρειάστηκε ένα άτομο για να διανύσει μια ορισμένη απόσταση.

CH Παιδιά: Σε αυτό το module συλλέγονται πληροφορίες για τα παιδιά των ατόμων του δείγματος της έρευνας SHARE. Οι πληροφορίες που παρουσιάζονται αφορούν την ηλικία του παιδιού, το επίπεδο εκπαίδευσης, τη συχνότητα επαφής, την εγγύτητα με τους γονείς κλπ.

SP Κοινωνική υποστήριξη: Σε αυτή η ενότητα συγκεντρώνονται οι πληροφορίες για διαστάσεις που αφορούν άτυπη βοήθεια και υποστήριξη που παρέχουν ή λαμβάνουν τα άτομα του δείγματος (όπως προσωπική φροντίδα, πρακτική βοήθεια, βοήθεια σε γραφειοκρατικές υποχρεώσεις, φροντίδα εγγονιών). Πρόσθετες πληροφορίες που μπορούν να αξιοποιηθούν περιλαμβάνουν i) τη συχνότητα βοήθειας, ii) τη σχέση με το άτομο που λαμβάνει ή παρέχει τη βοήθεια.

FT Χρηματικές μεταβιβάσεις: Σε αυτή την ενότητα περιλαμβάνονται οι ερωτήσεις για τακτικές οικονομικές μεταβιβάσεις και πληρωμές που μπορεί κάποιος να έχει δώσει σε ή να έχει λάβει από άτομα εκτός του νοικοκυριού (καθώς και τη σχέση με αυτά τα άτομα). Συγκεντρώνονται επίσης πληροφορίες για τις κληρονομιές.

HO Στέγαση: Συλλέγονται πληροφορίες για τις συνθήκες στέγασης, συμπεριλαμβανομένου του μεγέθους και της ποιότητας της κατοικίας. Οι ιδιοκτήτες ερωτώνται για την αξία της ιδιοκτησίας τους και, ανάλογα με τη διάρκεια μίσθωσης των ατόμων, υποβάλλονται ερωτήσεις για τις υποθήκες και τις πληρωμές ενοικίων. Πληροφορίες για άλλα περιουσιακά στοιχεία (δευτερεύουσα-εξοχική κατοικία, αγροτική γη, οικόπεδο κλπ) είναι επίσης διαθέσιμες.

HH Εισόδημα νοικοκυριού: Παρουσιάζονται πληροφορίες των εισοδημάτων του νοικοκυριού από τις διάφορες πηγές. Σημειώνεται ότι σε περιπτώσεις άρνησης ερωτάται η τάξη μεγέθους της κάθε πηγής εισοδήματος (‘Είναι μεγαλύτερο ή μικρότερο από χ;’)

CO Κατανάλωση: Περιέχει πληροφορίες για ορισμένες δαπάνες του νοικοκυριού, π.χ. φαγητό στο σπίτι, για τηλεφωνικές υπηρεσίες κλπ. Υπάρχουν επίσης πληροφορίες που σχετίζονται με υποκειμενικές αξιολογήσεις της οικονομικής επισφάλειας, όπως η δυνατότητα του νοικοκυριού να τα βγάλει πέρα οικονομικά, η σύγκριση της παρούσας οικονομικής κατάστασης του νοικοκυριού με αυτήν πριν από ένα χρόνο κλπ.

AS Περιουσιακά Στοιχεία: Παρουσιάζονται πληροφορίες για τα χρηματικά και μη χρηματικά περιουσιακά στοιχεία.

AC Δραστηριότητες: Υπάρχουν ερωτήσεις για το εύρος των δραστηριοτήτων που γεμίζουν τον ελεύθερο χρόνο κατά τον τελευταίο μήνα, καθώς επίσης και για τα κίνητρα που παρακινούν τα άτομα να συμμετέχουν σε αυτές τις δραστηριότητες.

EX Προσδοκίες: Παρουσιάζονται οι προσδοκίες των ανθρώπων η συνολική τους στάση για το μέλλον καθώς η στάση τους απέναντι στον κίνδυνο.

IV Παρατηρήσεις ερευνητών: Αφότου έχει ολοκληρώσει ο ερευνώμενος τη συνέντευξη υπάρχει μια σειρά ερωτήσεων για να απαντηθούν από τον συνεντευκτή, όπως το είδος του σπιτιού, ο αριθμός ορόφων κλπ.

2.2 Περιγραφή των στατιστικών δεδομένων “drop off” και “vignettes”

Το ερωτηματολόγιο drop-off συμπληρώθηκε σε χαρτί και δε βασίζεται στην εφαρμογή της Προσωπικής Συνέντευξης με τη Βοήθεια Υπολογιστή (CAPI) όπως το βασικό ερωτηματολόγιο. Περιέχει 18916 παρατηρήσεις για τις έντεκα χώρες του δείγματος SHARE και επιπλέον 1720 παρατηρήσεις για το Ισραήλ. Στην Ελλάδα, το ερωτηματολόγιο drop-off έχει 1986 παρατηρήσεις. Τα δεδομένα του ερωτηματολογίου drop off αναφέρονται σε επίπεδο ατόμων και μπορούν να συγχωνευτούν και να αξιοποιηθούν με τα δεδομένα του βασικού ερωτηματολογίου (για τις παρατηρήσεις εκείνες που είναι κοινές και στα δύο αρχεία δεδομένων). Τονίζεται ότι, σε αντίθεση με το βασικό ερωτηματολόγιο, το ερωτηματολόγιο drop-off δεν μπορεί να συμπληρωθεί από πληρεξούσιο.

Το drop-off περιέχει ερωτήσεις για τα παρακάτω: κοινωνική & ψυχολογική κατάσταση ικανοποίηση από τη ζωή, αντιλήψεις, στάσεις και προσδοκίες, αξιολογήσεις για τις οικογενειακές σχέσεις, ιατρική και προληπτική φροντίδα υγείας συμπεριλαμβανομένων εμβολιασμών, εξέταση οφθαλμών, ακτινογραφιών για καρκίνους κλπ., στέγαση, συνθήκες στέγασης όπως κεντρική θέρμανση, air condition, ασανσέρ ή κήπος κλπ., γενικές ερωτήσεις, καθώς και πληροφορίες για τη θρησκεία, την πολιτική τοποθέτηση.

Το ερωτηματολόγιο vignette χορηγήθηκε πειραματικά σε οχτώ χώρες του δείγματος (Βέλγιο, Γαλλία, Γερμανία, Ελλάδα, Ιταλία, Ολλανδία, Ισπανία και Σουηδία) περιλαμβάνοντας συνολικά 4544 παρατηρήσεις, αντί του ερωτηματολογίου drop-off. Τα ερωτηματολόγια vignettes προσπαθούν να αντιμετωπίσουν το πρόβλημα της συγκρισιμότητας υποκειμενικών απαντήσεων και αποτελούν μια από τις σημαντικότερες καινοτομίες που φέρνει η έρευνα SHARE στο πρόβλημα της συγκρισιμότητας των δεδομένων που βασίζονται σε υποκειμενικές απαντήσεις. (βλέπε το κεφάλαιο των O'Donnell κ.ά. σε αυτόν τον τόμο).

2.3 Περιγραφή των δεδομένων με generated variables

Τα στατιστικά δεδομένα με generated variables αφορούν μεταβλητές που υπολογίστηκαν από ειδικές ομάδες εργασίας, αξιοποιώντας τα πρωτογενή δεδομένα της έρευνας SHARE. Παραδείγματα τέτοιων περιλαμβάνουν την κωδικοποίηση των πρωτογενών απαντήσεων σε συγκεκριμένες κατηγορίες απαντήσεων (π.χ. σύμφωνα με το διεθνές πρότυπο ISCED-97 για την εκπαίδευση ή σύμφωνα με το ISCO και NACE για τα επαγγέλματα). Επισημαίνεται ότι τα συγκεκριμένα δεδομένα παρουσιάζονται όπως ακριβώς τα δεδομένα του βασικού ερωτηματολογίου (παρατηρήσεις σε επίπεδο ατόμων) επομένως μπορούν να χρησιμοποιηθούν συνδυαστικά. Οι generated variables της κάθε ομάδας εργασίας παρουσιάζονται σε διαφορετικά αρχεία ως εξής: εκπαίδευση, υγεία, στέγαση και τύπος διαμονής, επάγγελμα, υποστήριξη και σύνθεση νοικοκυριού, **συντελεστές στάθμισης (weights)**, εισόδημα, πηγές εισοδήματος και περιουσιακά στοιχεία.

3. Διευκρινήσεις για την αξιοποίηση των στατιστικών δεδομένων

3.1 Πριν την χρήση

Στο ξεκίνημα: Κάθε αρχείο με στατιστικά δεδομένα της έρευνας SHARE μπορεί να παρομοιαστεί σαν ένας πίνακας με κελιά. Οι γραμμές του πίνακα αντιστοιχούν σε παρατηρήσεις και οι κάθετες στήλες περιέχουν τις μεταβλητές.

Τι υπάρχει στα κελιά: Σε κάθε κελί υπάρχει η πληροφορία που περιγράφει η μεταβλητή (κάθετη στήλη) για κάθε άτομο του δείγματος. Για κάθε μεταβλητή υπάρχει ένα εύρος κατηγοριών απαντήσεων. Για παράδειγμα στην μεταβλητή gender (φύλο) οι κατηγορίες

απαντήσεων, αυτό δηλαδή που θα φαίνεται στις παρατηρήσεις (οριζόντια γραμμή) για αυτή την μεταβλητή θα είναι male ή female.

Τι σημαίνουν τα κενά κελιά (missing values): Είναι σημαντικό να αποσαφηνιστεί ότι εάν μια παρατήρηση έχει missing value αυτό ΔΕΝ σημαίνει ότι τα άτομα αυτά δεν απάντησαν τη συγκεκριμένη ερώτηση. Οι περιπτώσεις των ατόμων που αρνούνται να απαντήσουν ή δεν γνωρίζουν τι να απαντήσουν υποδεικνύονται ως “don’t know” ή “refusal”, αλλά όχι ως κενές παρατηρήσεις. Οι missing values στις έρευνες μικρο-δεδομένων σημαίνουν ότι η συγκεκριμένη ερώτηση ΔΕΝ έγινε στα συγκεκριμένα άτομα. Αυτό μπορεί να οφείλεται σε διαδικασίες “if” που ισχύουν κατά τη διάρκεια της συνέντευξης. Το παραπάνω μπορεί να γίνει πιο κατανοητό μέσω ενός υποθετικού παραδείγματος: στην ερώτηση “έχετε παιδιά;” οι κατηγορίες απαντήσεων είναι ναι /όχι. Σε επόμενη ερώτηση για την ηλικία του παιδιού, ο αριθμός των κενών παρατηρήσεων θα πρέπει να είναι ίσος με τον αριθμό των παρατηρήσεων που επέλεξαν “όχι” στην προηγούμενη ερώτηση. Σε ένα δεύτερο παράδειγμα, κενές τιμές μπορεί να εμφανίζονται σε ερωτήσεις που γίνονται σε επίπεδο νοικοκυριού: σε αυτές τις περιπτώσεις η απάντηση δίνεται από ένα μέλος του νοικοκυριού, ενώ στα υπόλοιπα μέλη του νοικοκυριού υπάρχει missing value.

Ποιες χώρες υπάρχουν στα δεδομένα του κάθε κύματος: Οι χώρες του πρώτου κύματος συμμετέχουν επίσης και στο δεύτερο κύμα της έρευνας SHARE. Αυτό σημαίνει ότι υπάρχουν διαχρονικά δεδομένα για την Αυστρία (AT), το Βέλγιο (BE), την Ελβετία (CH), τη Γερμανία (DE), τη Δανία (DK), την Ισπανία (ES), τη Γαλλία (FR), την Ελλάδα (GR), την Ιταλία (IT), την Ολλανδία (NL), και τη Σουηδία (SE). Επιπλέον, δύο νέες χώρες έχουν προστεθεί στο δεύτερο κύμα: η Δημοκρατία της Τσεχίας (CZ) και η Πολωνία (PL) για τις οποίες υπάρχουν δεδομένα μόνο για το δεύτερο κύμα.

Η έννοια του longitudinal και του refresher sample: Υπάρχουν δύο διαφορετικοί τύποι ερευνώμενων στο δεύτερο κύμα. Ένας τύπος περιλαμβάνει τα άτομα από το «παλαιό» δείγμα, δηλαδή όσους είχαν συμμετάσχει στο πρώτο κύμα, οι οποίοι επιστρέφουν στο δεύτερο κύμα (**longitudinal sample**). Ο δεύτερος τύπος αποτελείται όσους είχαν την πρώτη συνέντευξή τους στο δεύτερο κύμα (**refresher sample**). Το συνολικό δείγμα του δεύτερου κύματος έχει 33281 παρατηρήσεις. Από αυτές οι 18741 αφορούν longitudinal sample για τις χώρες που έχουν πάρει μέρος και στα δύο κύματα. Για την Ελλάδα το μέγεθος του δείγματος στο δεύτερο κύμα φθάνει τις 3243 παρατηρήσεις από τις οποίες οι 2280 αφορούν longitudinal sample, αντιστοιχούν δηλαδή σε άτομα που συμμετείχαν και στα δύο κύματα.

Οι εκδόσεις των στατιστικών δεδομένων: Τα στατιστικά δεδομένα ενός κύματος βελτιώνονται ανά τακτά διαστήματα και γίνονται διαθέσιμα μέσω νέων εκδόσεων (releases). Οι βελτιώσεις που γίνονται αφορούν ποιοτικούς ελέγχους (π.χ. outliers), κωδικοποίηση βασικών μεταβλητών, υπολογισμό τιμών για απαντήσεις «άρνησης» ή «δεν γνωρίζω» κλπ. Αυτή τη στιγμή, τα δεδομένα του πρώτου κύματος είναι στην version Release 2.0.1 η οποία είναι διαθέσιμη από τον Ιούλιο του 2007, ενώ τα δεδομένα από το δεύτερο κύμα είναι στην version Release 1.0.1 η οποία είναι διαθέσιμη από τον Δεκέμβριο του 2008.

3.2 Μετά την πρώτη επαφή

Οι μεταβλητές-κλειδιά της έρευνας: Στα modules του βασικού ερωτηματολογίου, των drop-off, vignettes, καθώς και στα αρχεία με τις generated variables υπάρχουν τρεις μεταβλητές-κλειδιά που ταυτοποιούν τις πληροφορίες ως αντιστοιχούσες σε συγκεκριμένο άτομο και επομένως είναι απαραίτητες για τον συνδυασμό πληροφοριών. Οι μεταβλητές αυτές στα δεδομένα του πρώτου κύματος είναι οι εξής:

- **sampid2:** Πρόκειται για έναν 13ψήφιο αριθμό, ο οποίος είναι μοναδικός για κάθε νοικοκυριό. Έτσι, άτομα που ανήκουν στο ίδιο νοικοκυριό έχουν το ίδιο αριθμό στην μεταβλητή sampid2.
- **cvid:** Πρόκειται για έναν μονοψήφιο αριθμό, ο οποίος ξεκινά με την τιμή 1, μοναδικός για κάθε άτομο του νοικοκυριού (για κάθε νοικοκυριό). Ο αριθμός αυτός προκύπτει μέσω του εισαγωγικού φύλλου, προσδίδεται δηλαδή τόσο στα άτομα ενός νοικοκυριού που επιλέγονται για τη συνέντευξη, όσο και στα άτομα που δεν είναι επιλέξιμα.
- **respid:** Πρόκειται για έναν μονοψήφιο αριθμό, ο οποίος ξεκινά με την τιμή 1, μοναδικός για κάθε άτομο του νοικοκυριού (για κάθε νοικοκυριό) και αφορά μόνο τα επιλέξιμα άτομα.

Οι νέες μεταβλητές-κλειδιά της έρευνας στο δεύτερο κύμα: Το δεύτερο κύμα δεδομένων της έρευνας SHARE έχει ορισμένες νέες μεταβλητές κλειδιά χρήσιμες ώστε να επιτελούν τη σύνδεση και αξιοποίηση των δεδομένων μεταξύ των κυμάτων σύμφωνα με το διαχρονικό χαρακτήρα της έρευνας. Οι νέες αυτές μεταβλητές που εμφανίζονται προς το παρόν μόνο στο δεύτερο κύμα είναι οι εξής:

- **mergeid:** Πρόκειται για μία μεταβλητή της οποίας η τιμή είναι μοναδική για κάθε άτομο και δεν αλλάζει μεταξύ των κυμάτων. Για κάθε άτομο, η μεταβλητή mergeid έχει την εξής μορφή «XX-ηηηηηηηη-rr». Για παράδειγμα, στη τιμή "GR-248123-02" τα δύο πρώτα γράμματα (GR) δηλώνουν ότι το άτομο είναι στο δείγμα της Ελλάδας, τα επόμενα ψηφία (248123) αντιστοιχούν στο κωδικό του νοικοκυριού που ανήκει το άτομο, ενώ τα δύο τελευταία ψηφία (02) δηλώνουν ότι είναι το δεύτερο επιλέξιμο άτομο για συνέντευξη μέσα στο νοικοκυριό (αντιστοιχεί στο respid). Η μεταβλητή mergeid υπάρχει σε όλα τα modules στο δεύτερο κύμα και χρησιμεύει στο να γίνονται συγχωνεύσεις των δεδομένων των διαφόρων modules.
- **hhid και hhid2:** Πρόκειται για δύο μεταβλητές που έχουν κωδικούς, μοναδικούς για κάθε νοικοκυριό. Η μεταβλητή hhid έχει τον κωδικό του νοικοκυριού στο οποίο ανήκει ένα άτομο τη στιγμή που συμπεριλήφθηκε στο δείγμα πάνελ. Οι τιμές της μεταβλητής hhid δεν αλλάζουν μεταξύ των κυμάτων. Η μεταβλητή hhid2 (όπου 2 σημαίνει δεύτερο κύμα) αναφέρει τον κωδικό του νοικοκυριού στο οποίο διαμένει το άτομο στο δεύτερο κύμα. Οι μεταβλητές hhid και hhid2 έχουν την εξής μορφή «XX-ηηηηηηηη-S». Για παράδειγμα, η τιμή "GR-248123-A" δηλώνει ότι το νοικοκυριό είναι στο δείγμα της Ελλάδας, το (248123) δηλώνει το κωδικό του νοικοκυριού, ενώ το τελευταίο γράμμα (A) δηλώνει ότι το νοικοκυριό δε διασπάστηκε σε σχέση με το προηγούμενο κύμα (διασπασμένα νοικοκυριά δηλώνουν τα γράμματα B και C).
- **waveid:** Πρόκειται για έναν διψήφιο κωδικό ο οποίος δηλώνει εάν ένα άτομο ανήκει στο διαχρονικό δείγμα (longitudinal) ή μόνο στο δεύτερο κύμα (refresher). Οι τιμές «42» και «51» δηλώνουν ότι πρόκειται για longitudinal sample, ενώ οι τιμές «61» «62» και «64» δηλώνουν refresher sample. Φυσικά, για τις δύο χώρες που μπήκαν στην έρευνα SHARE πρώτη φορά στο δεύτερο κύμα, δηλαδή την Τσεχία και την Πολωνία, οι τιμές της μεταβλητής waveid είναι μόνο «62».

Πώς χρησιμοποιούμε στοιχεία μόνο για την Ελλάδα; Τόσο στο πρώτο, όσο και στο δεύτερο κύμα, όλες οι θεματικές ενότητες των δεδομένων της έρευνας SHARE περιέχουν μια μεταβλητή με το όνομα **"country"**. Η μεταβλητή αυτή λαμβάνει τιμές διψήφιων αριθμών, κάθε ένας από τους οποίους αντιστοιχεί σε μία χώρα. Για παράδειγμα, η μεταβλητή "country" έχει για όλες τις παρατηρήσεις του Ελληνικού δείγματος την τιμή 19. Έτσι, εάν ένας ερευνητής θέλει να δουλέψει μόνο με τις παρατηρήσεις του Ελληνικού δείγματος, μπορεί να το κάνει επιλέγοντας να κρατήσει μόνο αυτές τις παρατηρήσεις (για παράδειγμα

στο STATA μπορεί να πληκτρολογήσει την εντολή «*keep if country = 19*». Στη πιο γενική περίπτωση που ένας ερευνητής θέλει να κάνει εκτιμήσεις μόνο για την Ελλάδα, δεν είναι απαραίτητο να διαγράψει τις παρατηρήσεις των υπολοίπων χωρών. Για παράδειγμα, εάν θέλουμε να εκτιμήσουμε την μέση τιμή μιας μεταβλητής μόνο για το Ελληνικό δείγμα SHARE μπορούμε να γράψουμε στο πρόγραμμα STATA την εντολή:

sum Όνομα_μεταβλητής if country= 19

Σε αυτή την περίπτωση οι εκτιμήσεις θα βασίζονται μόνο στις παρατηρήσεις για την Ελλάδα, ακόμη και εάν η βάση δεδομένων που χρησιμοποιούμε έχει παρατηρήσεις για όλες τις χώρες.

Συνδυάζοντας και συγχωνεύοντας πληροφορίες από διαφορετικά modules: Όπως αναφέρθηκε προηγουμένως, τα στατιστικά δεδομένα της έρευνας έχουν χωριστεί σε modules. Στις περισσότερες περιπτώσεις, ωστόσο, η εμπειρική ανάλυση απαιτεί αξιοποίηση πληροφοριών και δεδομένων που υπάρχουν σε διαφορετικά modules. Για παράδειγμα, επιχειρώντας την ανάλυση της κατάστασης υγείας των ατόμων ανά φύλο ή ηλικία απαιτείται να συνδυαστούν οι πληροφορίες της φυσικής υγείας “*ph module*” με πληροφορίες για δημογραφικά χαρακτηριστικά “*dn module*”. Στην προκειμένη περίπτωση συγχώνευση σημαίνει ότι στα δεδομένα της φυσικής υγείας “*ph module*”, προστίθενται στήλες (και όχι γραμμές) με μεταβλητές με δημογραφικά χαρακτηριστικά. Με αυτό τον τρόπο για κάθε άτομο (παρατήρηση) έχουμε πλέον πληροφορίες για φυσική υγεία και για δημογραφικά χαρακτηριστικά.

Ποια δεδομένα μπορούν να συγχωνευτούν; Η απάντηση είναι όλα με όλα. Οι θεματικές ενότητες του βασικού ερωτηματολογίου (modules) μπορούν να συγχωνευτούν μεταξύ τους, αλλά και με δεδομένα από drop-off, vignettes, καθώς και με τις generated variables. Πρέπει να έχει υπόψη ο ερευνητής ότι συγχωνεύοντας δύο αρχεία δεδομένων που έχουν διαφορετικό αριθμό παρατηρήσεων, μετά από τη συγχώνευση ο συνολικός αριθμός των παρατηρήσεων θα ισούται με τις παρατηρήσεις που υπήρχαν διαθέσιμες μόνο στο ένα αρχείο, συν αυτών που υπήρχαν μόνο στο άλλο συν των παρατηρήσεων που ήταν κοινές και στα δύο αρχεία. Στην περίπτωση που ο ερευνητής θέλει να κρατήσει μόνο τις παρατηρήσεις που ήταν κοινές και στα δύο αρχεία πριν τη συγχώνευση αυτό είναι σχετικά εύκολο σε κάθε στατιστικό πρόγραμμα.

Πώς μπορούν να συγχωνευτούν τα δεδομένα δύο διαφορετικών αρχείων του πρώτου κύματος; Η συγχώνευση δύο αρχείων στατιστικών δεδομένων σε επίπεδο ατόμων μπορεί να γίνει χρησιμοποιώντας ως συνδετικές μεταβλητές τις μεταβλητές “*samrid2*” και “*cvid*”, σύμφωνα με τη διαδικασία εντολών που υπάρχει σε κάθε στατιστικό πρόγραμμα. Για τη συγχώνευση σε επίπεδο νοικοκυριού (*cv_h module*) η συνδετική μεταβλητή είναι η μεταβλητή “*samrid2*”.

Πώς μπορούν να συγχωνευτούν τα δεδομένα δύο διαφορετικών αρχείων του δεύτερου κύματος; Η συγχώνευση δύο αρχείων στατιστικών δεδομένων σε επίπεδο ατόμων του δεύτερου κύματος μπορεί να γίνει χρησιμοποιώντας ως συνδετική μεταβλητή τη μεταβλητή **mergeid**, ακολουθώντας τη διαδικασία εντολών που υπάρχει σε κάθε στατιστικό πρόγραμμα.

Πώς μπορούν να συγχωνευτούν τα δεδομένα του πρώτου και του δεύτερου κύματος; Στο δεύτερο κύμα υπάρχει μια βάση δεδομένων με το όνομα “*x_over*”. Η συγκεκριμένη βάση έχει τις μεταβλητές *samrid2* και *respid2* και περιέχει παρατηρήσεις για όλα τα άτομα του δείγματος του πρώτου κύματος. Έτσι, ο ερευνητής θα πρέπει σε πρώτο στάδιο να συγχωνεύσει τα δεδομένα του πρώτου κύματος στη βάση “*x_over*”, χρησιμοποιώντας ως συνδετικές μεταβλητές τις *samrid2* και *respid2*. Έπειτα, χρησιμοποιώντας την μεταβλητή **mergeid**, η οποία είναι διαθέσιμη στη βάση “*x_over*” μπορεί να συγχωνεύσει τις

πληροφορίες του πρώτου κύματος που υπάρχουν στη νέα “x_over” βάση που δημιουργήθηκε με οποιαδήποτε θεματική ενότητα του δεύτερου κύματος.

Τι είναι οι συντελεστές στάθμισης (weights); Σε εφαρμογές της στατιστικής ανάλυσης σε έρευνες μικρο-δεδομένων (όπως για παράδειγμα ο υπολογισμός της μέσης τιμής) υπάρχει η υπόθεση ότι όλα τα άτομα του δείγματος έχουν την ίδια πιθανότητα να συμπεριλαμβάνονται στο δείγμα. Η συγκεκριμένη υπόθεση είναι για μια σειρά από λόγους πολύ περιοριστική στις έρευνες μικρο-δεδομένων. Οι συντελεστές στάθμισης του δείγματος χρησιμοποιούνται προκειμένου να σταθμιστούν οι διαφορές μεταξύ των ατόμων (ή πληθυσμιακών ομάδων) στην πιθανότητα επιλογής στο δείγμα. Στην έρευνα SHARE υπάρχουν διαθέσιμοι οι εξής τύπου συντελεστές στάθμισης:

- design weights: μπορούν να χρησιμοποιηθούν για να κάνει ένας ερευνητής τις δικές του εκτιμήσεις των συντελεστών στάθμισης (δηλαδή να κατασκευάσει calibrated weights).
- calibrated household weights: οι συντελεστές στάθμισης που χρησιμοποιούνται για σταθμισμένες εκτιμήσεις σε επίπεδο νοικοκυριού (πχ μέση καταναλωτική δαπάνη για φαγητό στο σπίτι).
- calibrated individual weights: οι συντελεστές στάθμισης που χρησιμοποιούνται για σταθμισμένες εκτιμήσεις σε επίπεδο ατόμου (πχ μέση ηλικία ατόμων του δείγματος).

Επιπλέον, στις χώρες που έχουν συμπληρωθεί ερωτηματολόγια vignette (Σουηδία, Βέλγιο, Ισπανία, Γαλλία, Γερμανία, Ελλάδα, Ιταλία και Ολλανδία) κάθε τύπου συντελεστών στάθμισης έχουν υπολογιστεί για:

- σταθμίσεις για το βασικό δείγμα
- σταθμίσεις για το δείγμα vignette
- σταθμίσεις για το βασικό δείγμα και το δείγμα vignette. Αυτό είναι απαραίτητο αφού ερωτήσεις vignettes έγιναν μόνο σε ένα τμήμα του δείγματος.

Ποιους συντελεστές στάθμισης (weights) χρειάζεται να χρησιμοποιήσω; Για εκτιμήσεις σε επίπεδο ατόμων χρησιμοποιούνται οι συντελεστές στάθμισης σε επίπεδο ατόμων (calibrated individual weights), ενώ για εκτιμήσεις σε επίπεδο νοικοκυριού χρησιμοποιούνται οι συντελεστές στάθμισης σε επίπεδο νοικοκυριών (calibrated household weights).

Γιατί ορισμένοι συντελεστές στάθμισης έχουν missing value; Η πιο τυπική περίπτωση να έχουμε missing value σε συντελεστή στάθμισης είναι στην περίπτωση που η συγκεκριμένη παρατήρηση αφορά άτομο ηλικίας κάτω των 50 ετών (συνήθως σύζυγος ενός ηλικιακά επιλέξιμου ατόμου του δείγματος). Άρα, ενώ υπάρχει η πληροφορία για να υπολογιστούν π.χ. το μέγεθος του νοικοκυριού, δεν υπεισέρχεται σε στατιστικές αναλύσεις όπου αυτομάτως του προσδίδεται μηδενική στάθμιση (δηλαδή εξαιρείται από την ανάλυση).

Πώς χρησιμοποιούνται οι συντελεστές; Τα περισσότερα στατιστικά / οικονομετρικά πακέτα επιτρέπουν τη χρήση των συντελεστών στάθμισης στις περισσότερες ρουτίνες – διαδικασίες εκτιμήσεων. Για παράδειγμα, στο στατιστικό πακέτο STATA ορισμένα παραδείγματα χρήσης των συντελεστών στάθμισης έχουν ως εξής:

Εκτίμηση ενός σταθμισμένου μέσου όρου μιας μεταβλητής σε επίπεδο νοικοκυριού

Απάντηση: `sum Όνομα_μεταβλητής [aw=wtgh]`

όπου wtgh είναι το όνομα της μεταβλητής των calibrated household weights

Εκτίμηση ενός σταθμισμένου μέσου όρου μιας μεταβλητής σε επίπεδο ατόμου

Απάντηση: `sum Όνομα_μεταβλητής [aw=wtgci]`

όπου *wgtaci* είναι το όνομα της μεταβλητής των calibrated individual weights

Εκτίμηση ενός OLS υποδείγματος με συντελεστές στάθμισης

Απάντηση: *regress Y X1 X2 [pw= wgtaci]*

όπου,

Y είναι το όνομα της εξαρτημένης μεταβλητής

X1 είναι το όνομα της πρώτης ερμηνευτικής μεταβλητής

X2 είναι το όνομα της δεύτερης ερμηνευτικής μεταβλητής, και

wgtaci είναι το όνομα της μεταβλητής των calibrated individual weights

3.3 Επιμέρους μη-τεχνικά ζητήματα

Τι είναι πληρεξούσιος και τι γίνεται όταν ο ερευνώμενος δεν είναι σε θέση να απαντήσει;

Η έρευνα SHARE επιτρέπει τη χρήση πληρεξούσιου κάτω από κατάλληλες συνθήκες. Η λέξη πληρεξούσιος δηλώνει ότι κάποιος αντικαθιστά τον επιλεγμένο ερευνώμενο στην παροχή απαντήσεων στις ερωτήσεις του ερωτηματολογίου. Η πιο κοινή αιτία για τη διεξαγωγή μιας συνέντευξης πληρεξούσιου είναι στην περίπτωση που ο ερευνώμενος του δείγματος αντιμετωπίζει φυσικές ή άλλες δυσκολίες να ολοκληρώσει τη συνέντευξη.

Τι είναι exit interview (τι γίνεται αν κάποιος πεθάνει); Η συνέντευξη σε περίπτωση θανάτου (end-of-life-interview) θεωρείται μια ιδιαίτερη διαδικασία συνέντευξης με πληρεξούσιο. Ενώ σε περιπτώσεις που ο ερευνώμενος είναι ζωντανός, αλλά αδυνατεί να συμμετέχει σε όλη ή σε μέρος της συνέντευξης ζητάμε τη βοήθεια του πληρεξούσιου (για όσα μέρη της συνέντευξης επιτρέπεται) για να ολοκληρωθεί η διαδικασία της συνέντευξης, στην περίπτωση που ο ερευνώμενος έχει πεθάνει, ο πληρεξούσιος καλείται να συμπληρώσει μια πολύ σύντομη, ξεχωριστή συνέντευξη. Το ερωτηματολόγιο της συνέντευξης σε περίπτωση θανάτου είναι ένα πολύ σύντομο τυπωμένο ερωτηματολόγιο, το οποίο περιέχει περίπου 30 ερωτήσεις. Το κύριο θέμα είναι υγεία και πλούτος κατά την διάρκεια του τελευταίου έτους ζωής.

Πού βρίσκω στοιχεία για την έρευνα HRS; Η πρόσβαση στα στατιστικά δεδομένα της έρευνας HRS που διεξάγεται στις Η.Π.Α. είναι ελεύθερη, χωρίς κόστος με μόνη προϋπόθεση την απαιτούμενη δήλωση δικαιώματος χρήσης. Στην ιστοσελίδα <http://hrsonline.isr.umich.edu/> επιλέγετε την περιοχή **Data Products** και έπειτα την περιοχή **Access to Public Data**. Μέσω της επιλογής **New Users** έχετε τη δυνατότητα να συμπληρώσετε ηλεκτρονικά τη δήλωση δικαιώματος χρήστη. Στη συνέχεια θα λάβετε μέσω e-mail ένα ζεύγος κωδικών μοναδικών για κάθε ερευνητή (userid και password), τους οποίους μπορείτε να τους συμπληρώσετε ώστε να αποκτήσετε πρόσβαση στα δεδομένα.