



Hướng Dẫn Phân Tích: Tạo Định Danh Khuyết Tật Sử Dụng Cú Pháp Stata cho Bộ Câu Hỏi Ngắn Gọn về Thực Hiện Chức Năng của Nhóm Washington (WG-SS)

Giới Thiệu

Khuyết tật được hiểu rõ nhất là một phổ liên tục. Về việc khó khăn khi thực hiện chức năng, ‘khó khăn’ có thể được diễn đạt thông qua những mô tả từ hoàn toàn không có khó khăn gì, hơi khó khăn và khó khăn nhiều đến hoàn toàn không thể thực hiện được hành động đó. Mỗi mô tả thể hiện điểm giới hạn hoặc ngưỡng giới hạn trong việc xác định định danh khuyết tật cuối cùng; ví dụ, để xác định những người bị và không bị khuyết tật. Những cấp độ thực hiện chức năng này cũng được thể hiện trong các các phạm trù trả lời cho Bộ Câu Hỏi Ngắn Gọn về Thực Hiện Chức Năng của Nhóm Washington (WG-SS).

Mức độ phổ biến của khuyết tật không phải là một thống kê đơn lẻ, nhưng có thể được tính toán ở các ngưỡng giới hạn khác nhau tùy thuộc vào mục đích của cả việc thu thập và báo cáo dữ liệu. Ví dụ, nếu mục đích là để mang lại sự tiếp cận công bằng những không gian công cộng – thì mức độ bao gồm của một nhận dạng khuyết tật có thể là *hơi khó khăn*, vì những người dù chỉ gặp khó khăn khi thực hiện chức năng ở mức độ rất ít cũng có khả năng sẽ được hưởng lợi từ những điều chỉnh được thực hiện để loại bỏ các rào cản và giúp dễ dàng tiếp cận. Việc lắp đặt thang cuốn thay cho cầu thang bộ chẳng hạn, là một yếu tố thiết kế phổ biến có lợi cho những người gặp nhiều mức độ khó khăn khác nhau trong việc đi lại. Ngoài ra, nếu mục đích là để trợ cấp hoặc phụ cấp – mức độ bao gồm của một nhận dạng khuyết tật có thể là *hoàn toàn không thể thực hiện được* do chỉ có những người có các hạn chế chức năng nghiêm trọng hơn mới đáp ứng được các tiêu chí đủ điều kiện nghiêm ngặt hơn.

Cú pháp Stata được mô tả trong tài liệu này hỗ trợ cho việc tính toán bốn nhận dạng khuyết tật ở bốn ngưỡng giới hạn. Nhóm người *khuyết tật* sử dụng bốn ngưỡng khác nhau này để lập nên bốn kiểu nhận dạng khuyết tật sau đây:

- **DISABILITY1:** mức độ bao gồm là có ít nhất một lãnh vực/ câu hỏi được mã hóa là HỜI KHÓ KHĂN hoặc KHÓ KHĂN NHIỀU hoặc HOÀN TOÀN KHÔNG THỂ THỰC HIỆN ĐƯỢC.
- **DISABILITY2:** mức độ bao gồm là có ít nhất 2 lãnh vực/ câu hỏi được mã hóa là HỜI KHÓ KHĂN hoặc bất kỳ 1 lãnh vực/ câu hỏi nào được mã hóa là KHÓ KHĂN NHIỀU hoặc HOÀN TOÀN KHÔNG THỂ THỰC HIỆN ĐƯỢC.

Các Tài Liệu Hướng Dẫn Thực Hiện Của Nhóm Washington

để cập đến các bộ công cụ do Nhóm Washington về Thống Kê Khuyết Tật (WG) soạn thảo để thu thập dữ liệu về khuyết tật có thể so sánh quốc tế từ các cuộc điều tra dân số và khảo sát. Những tài liệu này cung cấp những thực hành tốt nhất trong việc triển khai Bộ Câu Hỏi Ngắn Gọn, Bộ Câu Hỏi Mở Rộng, Bộ Câu Hỏi Ngắn Gọn – Nâng Cao, Các Mô-đun về Thực Hiện Chức Năng của Trẻ Em của WG/UNICEF dành cho trẻ em 2-4 tuổi và 5-17 tuổi, Mô-đun về Khuyết Tật Trong Khảo Sát Lực Lượng Lao Động (LFS-DM) của WG/ILO, cũng như các công cụ khác của WG. Các chủ đề bao gồm dịch thuật, đặc điểm của câu hỏi, hướng dẫn phân tích, mã lập trình để phân tích, việc sử dụng các công cụ nhằm mục đích phân tích, và các chủ đề khác nữa.

Để tìm Các Tài Liệu Hướng Dẫn Thực Hiện khác của Nhóm Washington và các thông tin khác, hãy truy cập trang web của Nhóm Washington:

<http://www.washingtongroup-disability.com/>.

- **DISABILITY3**: mức độ bao gồm là có bất kỳ 1 lãnh vực/ câu hỏi nào được mã hóa là KHÓ KHĂN NHIỀU hoặc HOÀN TOÀN KHÔNG THỂ THỰC HIỆN ĐƯỢC.

LƯU Ý: DISABILITY3 LÀ NGƯỠNG GIỚI HẠN DO WG KHUYẾN NGHỊ.

- **DISABILITY4** : mức độ bao gồm là có bất kỳ một lãnh vực nào được mã hóa là HOÀN TOÀN KHÔNG THỂ THỰC HIỆN ĐƯỢC (4).

LƯU Ý: Cú pháp Stata được dựa trên *nhãn biến* và *nhãn giá trị* được thể hiện trong bảng dưới đây. Đảm bảo rằng bạn sử dụng *các nhãn biến* và *nhãn giá trị* giống nhau HOẶC sửa lại cú pháp Stata để phản ánh *các nhãn* được sử dụng trong cơ sở dữ liệu của bạn.

WG-SS được sử dụng làm một phần của Khảo Sát Phỏng Vấn Sức Khỏe Quốc Gia Hoa Kỳ (NHIS). Dữ liệu được sử dụng để chuẩn bị các hướng dẫn này được lấy từ NHIS năm 2013.

Lưu ý dành cho người dùng NHIS: tên các biến trong tệp dữ liệu và tài liệu của NHIS có thể khác với các tên được sử dụng trong tài liệu này; ví dụ, biến về lãnh vực tự chăm sóc được gọi là SC-SS trong tài liệu này lại được đặt tên là UB_SS trong tài liệu và tệp dữ liệu của NHIS.

Các Câu Hỏi/ Lãnh Vực của Bộ Câu Hỏi Ngắn Gọn của Nhóm Washington	Nhãn Biến
1. Bạn có gặp khó khăn khi nhìn ngay cả khi đang đeo kính không?	VIS_SS
2. Bạn có gặp khó khăn khi nghe ngay cả khi đang dùng dụng cụ trợ thính không?	HEAR_SS
3. Bạn có gặp khó khăn khi đi bộ hoặc leo bậc cấp không?	MOB_SS
4. Bạn có gặp khó khăn khi nhớ hoặc tập trung không?	COG_SS
5. Bạn có gặp khó khăn khi (tự chăm sóc chẳng hạn như) tắm rửa toàn thân hoặc mặc quần áo không?	SC_SS
6. Khi dùng ngôn ngữ thông thường của bạn, bạn có gặp khó khăn khi giao tiếp (chẳng hạn như hiểu hoặc được người khác hiểu không)?	COM_SS

Các nhãn giá trị được sử dụng cho mỗi câu hỏi WG-SS là:

1. Không khó khăn
2. Có, hơi khó khăn
3. Có, khó khăn nhiều
4. Hoàn toàn không thể thực hiện được
7. Từ chối trả lời
8. Không biết chắc chắn
9. Không biết

Stata WG Short Set Syntax Annotated with Output Tables

Actual Stata syntax is indented and are in **Bold text**.

NOTE: For data analysis, use your standard weighting and estimation techniques.

Để biết thêm thông tin về Nhóm Washington về Thống Kê Khuyết Tật (WG), hãy truy cập:

<http://www.washingtongroup-disability.com/>.

The syntax below produces frequency distributions on each the six domains. Codes 7 (REFUSED), 8 (NOT ASCERTAINED) and 9 (DON'T KNOW) are RECODED as **MISSING**.

Step 1: Generate frequency distributions on each of the six domain variables.

VIS_SS is the WG-SS Vision question.

```
gen Vision=VIS_SS if inlist(VIS_SS, 1, 2, 3, 4)
replace Vision=. if inlist(VIS_SS, 7, 8, 9)
tabulate Vision
```

Vision: Degree of difficulty seeing

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13690	79.0	81.6	81.6
	Some difficulty	2708	15.6	16.2	97.8
	A lot of difficulty	333	1.9	2.0	99.8
	Cannot do at all	36	.2	.2	100.0
	Total	16767	96.8	100.0	
Missing		559	3.2		
Total		17326	100.0		

HEAR_SS is the WG-SS Hearing question.

```
gen Hearing=HEAR_SS if inlist(HEAR_SS, 1, 2, 3, 4)
replace Hearing=. if inlist(HEAR_SS, 7, 8, 9)
tabulate Hearing
```

Hearing: Degree of difficulty hearing

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13680	79.0	81.6	81.6
	Some difficulty	2753	15.9	16.4	98.0
	A lot of difficulty	310	1.8	1.8	99.9
	Cannot do at all	23	.1	.1	100.0
	Total	16766	96.8	100.0	
Missing		560	3.2		
Total		17326	100.0		

MOB_SS is the WG-SS Mobility question.

gen Mobility=MOB_SS **if inlist**(MOB_SS, 1, 2, 3, 4)
replace Mobility=. **if inlist**(MOB_SS, 7, 8, 9)
tabulate Mobility

Mobility: Degree of difficulty walking or climbing steps

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13424	77.5	80.1	80.1
	Some difficulty	2165	12.5	12.9	93.0
	A lot of difficulty	792	4.6	4.7	97.7
	Cannot do at all	380	2.2	2.3	100.0
	Total	16761	96.7	100.0	
Missing		565	3.3		
Total		17326	100.0		

COM_SS is the WG-SS Communication question.

gen Communication=COM_SS **if inlist**(COM_SS, 1, 2, 3, 4)
replace Communication=. **if inlist**(COM_SS, 7, 8, 9)
tabulate Communication

Communication: Degree of difficulty communicating using usual language

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	15874	91.6	94.7	94.7
	Some difficulty	745	4.3	4.4	99.2
	A lot of difficulty	94	.5	.6	99.7
	Cannot do at all	43	.2	.3	100.0
	Total	16756	96.7	100.0	
Missing		570	3.3		
Total		17326	100.0		

SC_SS is the WG-SS Self-care question.

gen Self_Care=SC_SS **if inlist**(SC_SS, 1, 2, 3, 4)
replace Self_Care=. **if inlist**(SC_SS, 7, 8, 9)
tabulate Self_Care

Self_Care Degree of difficulty with self-care

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	16029	92.5	95.7	95.7
	Some difficulty	544	3.1	3.2	98.9
	A lot of difficulty	114	.7	.7	99.6
	Cannot do at all	68	.4	.4	100.0
	Total	16755	96.7	100.0	
Missing		571	3.3		
Total		17326	100.0		

COG_SS is the WG-SS Cognition question.

```
gen Cognition=COG_SS if inlist(COG_SS, 1, 2, 3, 4)
replace Cognition=. if inlist(COG_SS, 7, 8, 9)
tabulate Cognition
```

Cognition: Degree of difficulty remembering or concentrating

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13719	79.2	81.9	81.9
	Some difficulty	2632	15.2	15.7	97.6
	A lot of difficulty	382	2.2	2.3	99.9
	Cannot do at all	20	.1	.1	100.0
	Total	16753	96.7	100.0	
Missing		573	3.3		
Total		17326	100.0		

Step 2: Calculate a variable, SUM_234

SUM_234 summates the number of domains coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) for each person. This new variable is used in the determination of disability identifiers: **DISABILITY1** and **DISABILITY2**.

The syntax below **counts** the number of domains/questions a person has that are coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4).

Possible range 0: no difficulties in any domain, to 6: all six domains coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4).

MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

```

gen SumPoints=0
foreach v of var Vision Hearing Mobility Cognition Self_Care Communication{
    replace SumPoints=SumPoints + inlist(`v',2,3,4)
}
replace SumPoints=. if missing(Vision) & missing(Hearing) & ///
missing(Mobility) & missing(Cognition) & missing(Self_Care) & missing(Communication)

gen SUM_234=. if SumPoints==.
replace SUM_234=1 if SumPoints==1
replace SUM_234=2 if SumPoints==2
replace SUM_234=3 if SumPoints==3
replace SUM_234=4 if SumPoints==4
replace SUM_234=5 if SumPoints==5
replace SUM_234=6 if SumPoints==6
replace SUM_234=0 if SumPoints==0
tabulate SUM_234

```

		SUM_234			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	9266	53.5	55.2	55.2
	1.00	3839	22.2	22.9	78.1
	2.00	1892	10.9	11.3	89.4
	3.00	989	5.7	5.9	95.3
	4.00	481	2.8	2.9	98.2
	5.00	232	1.3	1.4	99.5
	6.00	78	.5	.5	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		

Step 3: Calculate a variable, SUM_34

SUM_34 summates the number of domains coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) for each person. This new variable is used in the determination of disability identifier: **DISABILITY2**.

The syntax below counts the number of domains/questions a person has that are coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4)

Possible range 0: no difficulties coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) in any domain, to 6: all six domains coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4). MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

```

gen SumPoints2=0
foreach v of var Vision Hearing Mobility Cognition Self_Care Communication{
  replace SumPoints2=SumPoints2 + inlist(`v',3,4)
}
replace SumPoints2=. if missing(Vision) & missing(Hearing) & ///
missing(Mobility) & missing(Cognition) & missing(Self_Care) & missing(Communication)

gen SUM_34=. if Sumpoints2==.
replace SUM_34=1 if SumPoints2==1
replace SUM_34=2 if SumPoints2==2
replace SUM_34=3 if SumPoints2==3
replace SUM_34=4 if SumPoints2==4
replace SUM_34=5 if SumPoints2==5
replace SUM_34=6 if SumPoints2==6
replace SUM_34=0 if SumPoints2==0
tabulate SUM_34

```

		SUM_34			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	14905	86.0	88.8	88.8
	1.00	1367	7.9	8.1	97.0
	2.00	345	2.0	2.1	99.0
	3.00	117	.7	.7	99.7
	4.00	31	.2	.2	99.9
	5.00	9	.1	.1	100.0
	6.00	3	.0	.0	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		

Step 4: Calculate Disability Identifier: DISABILITY1

The syntax below calculates the first disability identifier: **DISABILITY1** where the level of inclusion is at least one domain/question is coded SOME DIFFICULTY or A LOT OF DIFFICULTY or CANNOT DO AT ALL.

MISSING (.) are those who have coded 7, 8 or 9 on all six domains.

```

gen Disability1=2
replace Disability1=1 if SUM_234 >=1
replace Disability1=. if missing(Vision) & missing(Hearing) & missing(Mobility) & ///
missing(Cognition) & missing(Self_Care) & missing(Communication)

```

NOTE: SUM_234 >= 1 means that at least one of the six domains is coded at least SOME DIFFICULTY (2).

tabulate Disability1

		Frequency	Percent	Valid Percent	Cumulative Percent	Weighted Percent*
Valid	without disability	9266	53.5	55.2	55.2	58.1
	with disability	7511	43.4	44.8	100.0	41.9
	Total	16777	96.8	100.0		100.0
Missing		549	3.2			
Total		17326	100.0			

*Weighted estimate provided – but is not part of the Stata syntax.

*Step 5: Calculate Disability Identifier: **DISABILITY2***

The syntax below calculates the second disability identifier: **DISABILITY2** where the level of inclusion is: at least 2 domains/questions are coded SOME DIFFICULTY or any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL. MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

gen Disability2=2

replace Disability2=1 **if** (SUM_234 >=2 | SUM_34==1)

replace Disability2=. **if** missing(Vision) & missing(Hearing) & missing(Mobility) & /// missing(Cognition) & missing(Self_Care) & missing(Communication)

NOTE: The above syntax identifies those with at least two of the six domains coded as at least SOME DIFFICULTY (2): SUM_234 >= 2, OR those who have one domain that is coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4): SUM_34 = 1.

tabulate Disability2

		Frequency	Percent	Valid Percent	Cumulative Percent	Weighted Percent*
Valid	without disability	12707	73.3	75.7	75.7	78.3
	with disability	4070	23.5	24.3	100.0	21.7
	Total	16777	96.8	100.0		100.0
Missing		549	3.2			
Total		17326	100.0			

*Weighted estimate provided – but is not part of the Stata syntax.

Step 6: Calculate Disability Identifier: **DISABILITY3**

The syntax below calculates the third disability identifier: **DISABILITY3** where the level of inclusion is: any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL. MISSING (9) are those who have coded 7, 8 or 9 on all six domains. **THIS IS THE CUT-OFF RECOMMENDED BY THE WG.**

```
gen Disability3=2
replace Disability3=1 if (inlist(Vision,3,4) | inlist(Hearing,3,4) | inlist(Mobility,3,4) | ///
inlist(Communication,3,4) | inlist(Self_Care,3,4) | inlist(Cognition,3,4))
replace Disability3=. if missing(Vision) & missing(Hearing) & missing(Mobility) & ///
missing(Cognition) & missing(Self_Care) & missing(Communication)
```

```
tabulate Disability3
```

		DISABILITY3				
		Frequency	Percent	Valid Percent	Cumulative Percent	Weighted Percent*
Valid	without disability	14905	86.0	88.8	88.8	90.5
	with disability	1872	10.8	11.2	100.0	9.5
	Total	16777	96.8	100.0		100.0
Missing		549	3.2			
Total		17326	100.0			

*Weighted estimate provided – but is not part of the Stata syntax.

Step 7: Calculate Disability Identifier: **DISABILITY4**

The syntax below calculates the fourth disability identifier: **DISABILITY4** where the level of inclusion is any one domain is coded CANNOT DO AT ALL (4). MISSING (.) are those who have coded 7, 8 or 9 on all six domains.

```
gen Disability4=2
replace Disability4=1 if ((Vision==4) | (Hearing==4) | (Mobility==4) | (Communication==4) | ///
(Self_Care==4) | (Cognition==4))
replace Disability4=. if missing(Vision) & missing(Hearing) & missing(Mobility) & ///
missing(Cognition) & missing(Self_Care) & missing(Communication)
```

```
tabulate Disability4
```

DISABILITY4

		Frequency	Percent	Valid Percent	Cumulative Percent	Weighted Percent*
Valid	without disability	16312	94.1	97.2	97.2	97.8
	with disability	465	2.7	2.8	100.0	2.2
	Total	16777	96.8	100.0		100.0
Missing		549	3.2			
Total		17326	100.0			

*Weighted estimate provided – but is not part of the Stata syntax.