



## Directrices analíticas: La creación de identificadores de la discapacidad mediante la sintaxis Stata de la lista breve de preguntas sobre funcionamiento del Grupo de Washington

### Introducción

La mejor manera de entender la discapacidad es plantearla como un continuo. Desde el punto de vista de la dificultad de funcionamiento, la “dificultad” se puede reflejar a través de una variedad de descriptores, desde “no tiene ninguna dificultad en absoluto” hasta “le resulta imposible”, pasando por “tiene cierta dificultad” y “tiene mucha dificultad” para llevar a cabo la acción. Cada uno de estos descriptores representa un punto de corte o umbral al determinar el identificador de la discapacidad final, por ejemplo, para definir a las personas con discapacidad y sin esta. Estos niveles de funcionamiento se plasman también en las categorías de respuesta a la lista breve de preguntas sobre funcionamiento del Grupo de Washington.

La prevalencia de la discapacidad no consiste en una sola estadística, sino que se puede calcular utilizando varios umbrales según el propósito de la recopilación de datos y la presentación de informes al respecto. Por ejemplo, si el propósito es ofrecer acceso equitativo a los espacios públicos, el nivel de referencia de la inclusión para un identificador de la discapacidad podría ser *cierta dificultad*, ya que incluso las personas que presenten un grado bajo de dificultad de funcionamiento se beneficiarán de las adaptaciones dirigidas a eliminar barreras y facilitar el acceso. La instalación de escaleras mecánicas en lugar de escalones, por ejemplo, es un elemento de diseño universal habitual que beneficia a personas con una amplia gama de dificultades de movilidad. De manera alternativa, si el propósito es facilitar subsidios o prestaciones, el nivel de inclusión para un identificador de la discapacidad podría ser *le resulta imposible*, ya que solo las personas con limitaciones funcionales más graves cumplirían los criterios de elegibilidad más estrictos.

La sintaxis Stata descrita en este documento se aplica al cálculo de los cuatro identificadores de la discapacidad en cuatro umbrales. El uso de estos cuatro umbrales diferentes en la población de personas *con discapacidad* da lugar a los cuatro identificadores de la discapacidad siguientes:

Los documentos sobre implementación del Grupo de Washington comprenden las herramientas desarrolladas por el Grupo de Washington sobre Estadísticas de la Discapacidad con el objeto de recopilar datos sobre la discapacidad internacionalmente comparables a través de censos y encuestas. En los documentos se abordan las mejores prácticas en materia de ejecución de la lista breve, la lista ampliada, la lista breve mejorada, los módulos del Grupo de Washington y el Fondo de las Naciones Unidas para la Infancia (UNICEF) sobre el funcionamiento en niños y niñas de 2 a 4 años y de 5 a 17 años, y el módulo sobre discapacidad para encuestas de población activa del Grupo de Washington y la Organización Internacional del Trabajo (OIT), así como otras herramientas del Grupo de Washington. Entre muchos otros temas, se tratan la traducción, las especificaciones de las preguntas, las directrices analíticas, los códigos de programación para los análisis, y el uso de las herramientas con fines de desglose.

Para consultar otros documentos sobre implementación del Grupo de Washington y obtener información adicional, visite su sitio web:

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

- **DISCAPACIDAD1:** el nivel de inclusión es al menos una esfera/pregunta codificada como TIENE CIERTA DIFICULTAD o TIENE MUCHA DIFICULTAD o LE RESULTA IMPOSIBLE.
- **DISCAPACIDAD2:** el nivel de inclusión es al menos dos esferas/preguntas codificadas como TIENE CIERTA DIFICULTAD, o una esfera/pregunta codificada como TIENE MUCHA DIFICULTAD o LE RESULTA IMPOSIBLE.
- **DISCAPACIDAD3:** el nivel de inclusión es al menos una esfera/pregunta codificada como TIENE MUCHA DIFICULTAD o LE RESULTA IMPOSIBLE.

**NOTA: DISCAPACIDAD3 ES EL PUNTO DE CORTE RECOMENDADO POR EL GRUPO DE WASHINGTON.**

- **DISCAPACIDAD4:** el nivel de inclusión es cualquier esfera codificada como LE RESULTA IMPOSIBLE (4).

**NOTA:** la sintaxis Stata se basa en las *etiquetas de variable* y *etiquetas de valor* indicadas en las tablas que aparecen más adelante. Asegúrese de utilizar las mismas *etiquetas de variable* y *valor* O BIEN de revisar la sintaxis Stata con el fin de que refleje las *etiquetas* de su base de datos.

La lista breve de preguntas sobre funcionamiento del Grupo de Washington se aplica como parte de la Encuesta Nacional de Entrevistas de Salud (NHIS) de los Estados Unidos. Los datos utilizados para elaborar estas directrices proceden de la encuesta de 2013.

*Nota dirigida a los usuarios de la Encuesta Nacional de Entrevistas de Salud: los nombres de variable del archivo de datos de dicha encuesta y la documentación pueden diferir de los indicados en este documento, p. ej., la esfera de cuidado personal indicada como SC\_SS, es UB\_SS en el archivo de datos y la documentación de la Encuesta Nacional de Entrevistas de Salud.*

<b>Preguntas o esferas de la lista breve del Grupo de Washington</b>	<b>Etiqueta de variable</b>
1. ¿Tiene dificultad para ver, incluso cuando usa lentes?	VIS_SS
2. ¿Tiene dificultad para oír, incluso cuando usa un audífono?	HEAR_SS
3. ¿Tiene dificultad para caminar o subir escalones?	MOB_SS
4. ¿Tiene dificultad para recordar o concentrarse?	COG_SS
5. ¿Tiene dificultad para lavarse o vestirse (gestionar su autosuficiencia para el cuidado personal)?	SC_SS
6. Cuando emplea su lenguaje habitual, ¿tiene dificultad para comunicarse —por ejemplo, para entender a los demás o para que lo entiendan—?	COM_SS

Las etiquetas de valor utilizadas para cada pregunta de la lista breve son:

1. No tiene ninguna dificultad
2. Sí, tiene cierta dificultad
3. Sí, tiene mucha dificultad
4. Le resulta imposible
7. No contesta
8. No comprobado
9. No sabe

## **Stata WG Short Set Syntax Annotated with Output Tables**

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

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

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
	<b>A lot of difficulty</b>	<b>333</b>	<b>1.9</b>	<b>2.0</b>	<b>99.8</b>
	<b>Cannot do at all</b>	<b>36</b>	<b>.2</b>	<b>.2</b>	<b>100.0</b>
	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
	<b>A lot of difficulty</b>	<b>310</b>	<b>1.8</b>	<b>1.8</b>	<b>99.9</b>
	<b>Cannot do at all</b>	<b>23</b>	<b>.1</b>	<b>.1</b>	<b>100.0</b>
	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
	<b>A lot of difficulty</b>	<b>792</b>	<b>4.6</b>	<b>4.7</b>	<b>97.7</b>
	<b>Cannot do at all</b>	<b>380</b>	<b>2.2</b>	<b>2.3</b>	<b>100.0</b>
	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
	<b>A lot of difficulty</b>	<b>94</b>	<b>.5</b>	<b>.6</b>	<b>99.7</b>
	<b>Cannot do at all</b>	<b>43</b>	<b>.2</b>	<b>.3</b>	<b>100.0</b>
	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
	<b>A lot of difficulty</b>	<b>114</b>	<b>.7</b>	<b>.7</b>	<b>99.6</b>
	<b>Cannot do at all</b>	<b>68</b>	<b>.4</b>	<b>.4</b>	<b>100.0</b>
	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
	<b>A lot of difficulty</b>	<b>382</b>	<b>2.2</b>	<b>2.3</b>	<b>99.9</b>
	<b>Cannot do at all</b>	<b>20</b>	<b>.1</b>	<b>.1</b>	<b>100.0</b>
	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

		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
```

		DISABILITY2		Valid	Cumulative	Weighted
		Frequency	Percent	Percent	Percent	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.