



## Directives analytiques : Créer des éléments d'identification du handicap à l'aide de la syntaxe STATA pour la brève série de questions du Groupe de Washington sur le fonctionnement

### Introduction

Le concept de handicap est mieux compris en tant que spectre. Concernant les difficultés fonctionnelles, le terme « difficulté » peut se caractériser par tout un éventail de descriptifs allant de « pas du tout de difficultés » à « n'y parvient pas du tout », en passant par « un peu » et « beaucoup [de difficultés] ». Chacun de ces descriptifs représente un seuil permettant de déterminer un élément final d'identification du handicap ; par exemple en vue de définir qui présente un handicap et qui n'en présente pas. On retrouve également ces niveaux de fonctionnement dans les catégories de réponses de la brève série de questions du Groupe de Washington sur le fonctionnement.

La prévalence du handicap n'est pas une statistique figée ; elle peut être calculée à partir de divers seuils en fonction de la finalité de la collecte de données ainsi que de la communication des informations. Par exemple, si la finalité consiste à fournir un accès équitable aux espaces publics, alors le niveau d'inclusion pour un élément d'identification du handicap pourrait être « *un peu de difficultés* », car les adaptations réalisées pour éliminer les obstacles et faciliter l'accès bénéficieraient probablement aussi aux personnes éprouvant des difficultés fonctionnelles mineures. Par exemple, l'installation d'escaliers mécaniques au lieu d'escaliers classiques est un élément de conception universel courant qui profite à toutes les personnes éprouvant des difficultés en matière de mobilité, quel qu'en soit le niveau. À l'inverse, si la finalité consiste à fournir des aides ou des allocations, le niveau d'inclusion pourrait être « *n'y parvient pas du tout* », car seules les personnes confrontées à des limites fonctionnelles plus graves répondraient à des critères d'octroi plus stricts.

La syntaxe STATA (pour *Statistics and data* – statistiques et données) décrite dans le présent document fournit des calculs pour quatre éléments d'identification du handicap à quatre seuils. Avec ces quatre seuils différents, on aboutit aux quatre éléments d'identification des personnes *présentant un handicap* suivants :

- **HANDICAP1** – le niveau d'inclusion est fixé à : au moins un domaine/une question est codé(e) UN PEU ou BEAUCOUP ou N'Y PARVIENT PAS DU TOUT.
- **HANDICAP2** – le niveau d'inclusion est fixé à : au moins deux domaines/questions sont codé(e)s UN PEU ou un des domaines/une des questions est codé(e) BEAUCOUP ou N'Y PARVIENT PAS DU TOUT.

Les **documents de mise en œuvre** du **Groupe de Washington** couvrent les outils élaborés par le Groupe de Washington pour recueillir des données sur le handicap comparables à l'échelle internationale dans le cadre des recensements et des enquêtes. Les documents portent sur les bonnes pratiques de mise en œuvre concernant la brève série de questions, le questionnaire détaillé, la version enrichie de la brève série de questions, les modules sur le fonctionnement de l'enfant du Groupe de Washington et du Fonds des Nations Unies pour l'enfance (UNICEF) pour les enfants de 2 à 4 ans et de 5 à 17 ans, le module sur le handicap de l'enquête sur la population active du Groupe de Washington et de l'Organisation internationale du Travail (OIT), ainsi que d'autres outils du Groupe de Washington. Ces documents couvrent notamment les sujets suivants : traduction, précision des questions, directives analytiques, code de programmation pour les analyses, utilisation des outils à des fins de ventilation, etc.

Vous trouverez d'autres documents de mise en œuvre du Groupe de Washington et des informations supplémentaires sur le site Internet du Groupe de Washington :

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

- **HANDICAP3** – le niveau d’inclusion est fixé à : un des domaines/une des questions est codé(e) BEAUCOUP ou N’Y PARVIENT PAS DU TOUT.

**REMARQUE : HANDICAP3 EST LE SEUIL RECOMMANDÉ PAR LE GROUPE DE WASHINGTON.**

- **HANDICAP4** – le niveau d’inclusion est fixé à : un des domaines est codé N’Y PARVIENT PAS DU TOUT (4).

**REMARQUE :** La syntaxe STATA s’appuie sur les *étiquettes de variable* et les *étiquettes de valeur* figurant dans les tableaux ci-dessous. Veillez à utiliser les mêmes *étiquettes de variable et de valeur* OU révisiez la syntaxe STATA en fonction des *étiquettes* employées dans votre base de données.

La brève série de questions du Groupe de Washington est administrée dans le cadre de l’enquête nationale de santé des États-Unis (NHIS). Les données utilisées pour préparer les présentes directives sont tirées de la NHIS de 2013.

*Remarque à l’intention des utilisateurs de la NHIS : les noms des variables dans le fichier de données et la documentation de la NHIS peuvent être différents de ceux employés dans le présent document ; par exemple, la variable relative au domaine des soins de soi portant la référence SC\_SS dans le présent document est désignée par UB\_SS dans le fichier de données et la documentation de la NHIS.*

Questions/domaines de la brève série de questions du Groupe de Washington	Étiquette de variable
1. Éprouvez-vous des difficultés à voir, même avec des lunettes ?	VIS_SS
2. Éprouvez-vous des difficultés à entendre, même avec une prothèse auditive ?	HEAR_SS
3. Éprouvez-vous des difficultés à marcher ou à monter des escaliers ?	MOB_SS
4. Éprouvez-vous des difficultés à vous rappeler certaines choses ou à vous concentrer ?	COG_SS
5. Éprouvez-vous des difficultés à prendre soin de vous, à vous laver ou à vous habiller, par exemple ?	SC_SS
6. Éprouvez-vous des difficultés à communiquer dans votre langue habituelle, à comprendre les autres ou à vous faire comprendre, par exemple ?	COM_SS

Les étiquettes de valeur utilisées pour chacune des questions de la brève série sont les suivantes :

1. Non, pas du tout.
2. Oui, un peu.
3. Oui, beaucoup.
4. N’y parvient pas du tout.
7. Refuse de répondre.
8. Incertain(e).
9. Ne sait pas.

## 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.

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

		<b>DISABILITY1</b>		Valid	Cumulative	Weighted
		Frequency	Percent	Percent	Percent	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

		<b>DISABILITY2</b>		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		Valid	Cumulative	Weighted
		Frequency	Percent	Percent	Percent	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.