MODULE 9
FINDING THE RIGHT GENDER DATA AND CONDUCTING BASIC ANALYSIS

EXERCISES

Curriculum on Gender Statistics Training
This product was developed under the guidance of the Subgroup on Gender Statistics Training, within the Asia-Pacific Network of Statistical Training Institutes.
**Exercise 1**

What are the characteristics of good gender data? (Select all that apply)

1. It should be accurate and reliable
2. It should come from a microdata source
3. It should only interview the household head
4. It should be accessible

**Solutions to Exercise 1**

1 & 4

**Exercise 2**

Which data source should you refer to if you want to conduct an analysis using internationally comparable estimates? (Select all that apply)

1. Census data
2. DHS STATcompiler
3. IPUMS International
4. SDG Global Database

**Solutions to Exercise 2**

2, 3 & 4

**Exercise 3**

Which of the following are national sources of gender data? (Select all that apply)

1. Ministries of education
2. Ministries of labour
3. Electoral management bodies
4. Official SDG database
5. Gallup surveys
6. Police, military
Solutions to Exercise 3

1,2,3 & 6

Exercise 4

Using data from the SDG Global Database, obtain the estimates for SDG Indicator 5.2.1 for Bangladesh, 2017.

(Indicator 5.2.1: Proportion of ever-partnered women and girls subjected to physical and/or sexual violence by a current or former intimate partner in the previous 12 months, by age (%))

Solution to Exercise 4

Step 1: Go to SDG Global Database https://unstats.un.org/sdgs/indicators/database/

Step 2: Select Indicator 5.2.1

Step 3: Select ‘Bangladesh’ from the list of countries

Step 4: Select ‘Download table’

- The following table will appear

<table>
<thead>
<tr>
<th>Goal</th>
<th>Target</th>
<th>Indicator</th>
<th>Series</th>
<th>GeoArea</th>
<th>GeoArea?</th>
<th>Time_Period</th>
<th>Time_Det</th>
<th>UpperBp</th>
<th>LowerBp</th>
<th>BaseUnit</th>
<th>Source</th>
<th>FootNote</th>
<th>Nature</th>
<th>Units</th>
<th>[Age]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>27.1</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>35-39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>21.6</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>40-44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>26.7</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>45-49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>28.8</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>15-49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>35.4</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>20-24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>38.4</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>15-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>42.2</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>25-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.2.2</td>
<td>5.2.2.1</td>
<td>VAW_P</td>
<td>Bangladesh</td>
<td>2015</td>
<td>30.8</td>
<td>2015</td>
<td>Violence against WmC</td>
<td>PERCENT</td>
<td>30-34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 5: Select the data and click on ‘Pivot Table’ from the ‘Insert’ option in MS Excel

- Select the full range of data table.

- Select the following ‘Row Labels’ options from ‘Pivot Table’ fields

  - Indicator
  - Value
  - Units
  - Age
  - Sex
Step 6: Select ‘PivotChart’ to visualize the data
Step 7: Format the Pivot Chart:

- Select ‘Add Chart Element’
  
  - Add ‘Chart’ title
  - Add data labels
  - Data Legends
  - Vertical axis (express in %)

Proportion of ever-partnered women and girls subjected to physical and/or sexual violence by a current or former intimate partner in the previous 12 months (across different age groups)
Exercise 5

Using pre-processed data from DHS STATcompiler, obtain the estimates for median years of education completed in Bangladesh by women and men in the years 2011.

Solutions to Exercise 5

Step 1: Go to DHS STATcompiler [https://www.statcompiler.com/en/](https://www.statcompiler.com/en/)

Step 2: Select Bangladesh from the list of countries

Step 3: Select ‘Median years of schooling completed’ for both men and women

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Median years of education completed [Women]</th>
<th>Median years of education completed [Men]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>2014 DHS</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2011 DHS</td>
<td>4.3</td>
<td>3.7</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2007 DHS</td>
<td>3.2</td>
<td>2.7</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2004 DHS</td>
<td>1.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1999-00 DHS</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1996-97 DHS</td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1993-94 DHS</td>
<td></td>
<td>1.3</td>
</tr>
</tbody>
</table>

Step 4: Export the table in ‘Table Format’

Step 5: Clean the data by deleting the row with missing data

- Delete row for 2014 DHS, 1999–00, 1996–97 and 1993–94 (as there is no data for men or women for these years).

- Delete column D, since we only want to compare women and men of the same age group (15-49).

- Rename column C as ‘Education Women’ and column E as ‘Education_Men’.

Step 6: Visualize the data and format the chart

- Choose a line graph because we want to show change over time

- Arrange the data in ascending order to see the change since 2004
Exercise 6

Obtain the estimates for proportion of women and girls who were married as children in Bangladesh using DHS raw survey data for the year 2014\(^1\).

Solutions to Exercise 6

- Open IR File (Bangladesh DHS 2014)

To calculate child marriage:

- Generate new variable ‘childmarriage’
  
  ```
  gen childmarriage=0
  ```

- Replace childmarriage with a value of 1 if the age at first marriage or cohabitation is less than 18 years
  
  ```
  replace childmarriage=1 if v511<18
  ```

- Replace all the missing values with a dot (.)
  
  ```
  replace childmarriage=. if v511==.
  ```

- Tabulate the result with appropriate weights
  
  ```
  tab childmarriage [iw=v005/1000000]
  ```

---

\(^1\) Note: Proportion of women ages 18-49 who were married below the age of 18 (defined here as child marriage) is calculated as a ratio of women ages 18-49 who married below the age of 18 to all women ages 18-49. In countries where surveys did not interview non-married women (e.g., DHS Bangladesh), the denominator has to be adjusted using all women’s factor (awfactt). While this is straightforward for one-way tabulations, the syntax’s complexity increases for two-way and three-way tabulations. Hence, for simplicity, awfactt has not been applied in this case.
Exercise 7

Using DHS STATcompiler:

a) Calculate the estimates on women unemployed in the past 12 months in Bangladesh
b) Compare women’s estimates with men’s estimates
c) Represent this information graphically

Solution to Exercise 7

Step 1: Go to STATcompiler [https://www.statcompiler.com/en/]

Step 2: Select Bangladesh as the country and select the indicator ‘employment status’. Make this selection for both women and men. Now, there are three options under this:

- Women who worked in the last 12 months and are currently
- Women who worked in the last 12 months, but not currently
- Women who did no work in the last 12 months

The command ‘gen weights’, generates women’s sample weights so that the results are representative of the population.

The command ‘tab v025’ tabulates women’s population distribution on type of location of residence.

In Bangladesh, 71.74% of women and girls live in rural households.

<table>
<thead>
<tr>
<th>type of place of residence</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>urban</td>
<td>5,047,3554</td>
<td>28.26</td>
<td>28.26</td>
</tr>
<tr>
<td>rural</td>
<td>12,815,6443</td>
<td>71.74</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>17,862,9997</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

In Bangladesh, 71.74% of women and girls live in rural households.
For the purpose of this exercise, select the third option: Women who did not work in the last 12 months and similarly for men: Men who did not work in the last 12 months

You will obtain the following result table:

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Total 15-49</th>
<th>Total</th>
<th>Total 15-49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>2014 DHS</td>
<td>64.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2011 DHS</td>
<td>85.0</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2007 DHS</td>
<td>65.0</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2004 DHS</td>
<td>77.5</td>
<td>10.7</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Reading the additional information about the data (referred to as the metadata) is crucial to avoid misinterpreting results. You will obtain the following indicator details through this:

This information indicates that the estimates are presented as percentages.

To represent this information graphically, first, export this result table in Excel by clicking the top right option.
You get the following Excel file:

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Total 15-49</th>
<th>Total</th>
<th>Total 15-49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>2014 DHS</td>
<td>64.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2011 DHS</td>
<td>85</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2007 DHS</td>
<td>65</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>2004 DHS</td>
<td>77.5</td>
<td>10.7</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Women who did no work in the 12 months preceding the survey
Men who did no work in the 12 months preceding the survey

To keep the same sample for comparison, only compare women and men in the 15–49 age group and for the years 2004, 2007 and 2011. Use a line graph.