

Data Standards for the Counter-Trafficking Community

Estimating the Prevalence of Human Trafficking in the United States

April 8-9, 2019

National Academies of Sciences, Engineering, and Medicine
Washington D.C.



DATA STANDARDS

Data standards are documented agreements on representation, format, definition, structuring, tagging, transmission, manipulation, use, and management of data.

DATA STANDARDS - EXAMPLES

The International Standard Industry Classification (ISIC)

- A. Agriculture, forestry and fishing
- B. Mining and quarrying
- C. Manufacturing
- D. Electricity, gas, steam and air conditioning supply
- E. Water supply; sewerage, waste management and remediation activities
- F. Construction
- G. Wholesale and retail trade; repair of motor vehicles and motorcycles
- H. Transportation and storage
- I. Accommodation and food service activities
- J. Information and communication
- K. Financial and insurance activities
- L. Real estate activities
- M. Professional, scientific and technical activities
- N. Administrative and support service activities
- O. Public administration and defence; compulsory social security
- P. Education
- Q. Human health and social work activities
- R. Arts, entertainment and recreation
- S. Other service activities
- T. Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
- U. Activities of extraterritorial organizations and bodies

DATA STANDARDS - EXAMPLES

The International Standard Industry Classification (ISIC)

Section	C	Manufacturing
Division	13	Manufacture of Textiles
Group	139	Manufacture of Other Textiles
Class	1393	Manufacture of Carpets and Rugs

DATA STANDARDS

OBSTACLES TO DATA SHARING AND USE

- Data privacy concerns
- Databases are diverse, dispersed, disconnected
- Data are not standardized or comparable
- Lack of technical resources and solutions

DATA STANDARDS

- Combination/interoperability
- Automation
- Replicability
- Transparency
- Efficiency
- Safer anonymization and sharing
- Developed by experts
- Creating a culture of greater data openness
- Can be a little painful but not expensive

Merging Datasets



Merging Datasets

POLARIS DATA

Based on “case records”
US legal definitions
Reported through crisis hotline
Big focus on reporting to law enforcement

Victim Records
73 Shared Variables
Aligned Values and Definitions

IOM DATA

Based on “victim records”
International definitions
Reported through client interviews
Big focus on info needed for migration services

LIBERTY SHARED DATA

Initial Data Contributions



55,000
Victim
Records



24,000
Victim
Records



13,000
Victim
Records



?

IOM UN
MIGRATION

Total Dataset: 91,000 records

- 169 Victim Nationalities
- 172 Countries of Exploitation

DATA STANDARDS - EXAMPLES

Alias	Data type	Length/Size	Format	Lookup	Notes
yearOfRegistration	Numeric		YYYY		
Datasource	Alphanumeric	20		Datasource	
gender	Alpha			Gender	
age	Alphanumeric		N--N	Age range 1	
ageBroad	Alphanumeric		N--N	Age range 2	
majorityStatus	Alpha			Age category	
majorityStatusAtExploit	Alpha			Age category	
ageAtExploit	Alphanumeric		N--N	Age range 1	
ageAtExploitBroad	Alphanumeric		N--N	Age range 2	
majorityStatusAtExploit	Alpha			Age category	
ageAtEntryIntoTrafficking	Alphanumeric		N--N	Age range 1	
ageAtEntryIntoTraffickingBroad	Alphanumeric		N--N	Age range 2	
MajorityEntry	Alpha			Age category	
ageDestination	Alphanumeric		N--N	Age range 1	
ageDestinationBroad	Alphanumeric		N--N	Age range 2	
citizenship	Alpha			COUNTRY	ISO code (two capital letters)
CountryLastResidence	Alpha			COUNTRY	ISO code (two capital letters)
PlaceLastResidence	Alphanumeric	1000			The largest administrative level
CountryBirth	Alpha			COUNTRY	ISO code (two capital letters)
PlaceBirth	Alphanumeric	1000			The largest administrative level
TraffickingDuration	Numeric				
MaritalStatus	Alphanumeric			Marital status	
NumberChildren	Numeric				
Education	Alphanumeric			Education level	
languagePrimary	Alpha	1000		Language	ISO code (three lower case letters)
languageSecondary	Alpha	1000		Language	ISO code (three lower case letters)
meansOfControlDebtBondage	Bool				Three possible values: "0", "1", "-99"
meansOfControlTakesEarnings	Bool				Three possible values: "0", "1", "-99"
meansOfControlRestrictsFinancialAccess	Bool				Three possible values: "0", "1", "-99"
meansOfControlThreats	Bool				Three possible values: "0", "1", "-99"
meansOfControlPsychologicalAbuse	Bool				Three possible values: "0", "1", "-99"

DATA STANDARDS - EXAMPLES

AN 'ISIC' FOR SEX WORK?

Type of sex work:

- Prostitution
- Pornography
- Remote interactive sexual services
- Private sexual services

Place of sex work:

- Bar/Club/Cantina
- Escort Service/Delivery Service
- Hostess/Strip Club
- Hotel/Motel-Based Commercial Sex
- Legal Brothel
- Commercial Front Brothel
- Residential Brothel
- Private residence
- Recording studio
- Street-Based Commercial Sex
- Truck Stop
- Other

Global Value	Existing IOM value	Existing Polaris values
Debt bondage	Debt bondage	Economic - Debt/Quota Related
Takes/Withholds Earnings	Withholding of wages	Takes/Withholds Earnings
Restricts access to finances		Limits Access to Finances
		Threat to blacklist
		Intimidation - Destroys Property
		Intimidation - Displays/Threatens Weapon
		Intimidation - Harms Others/Animals
		Intimidation - Other
		Threat to Expose or Shame Subject
		Threat to Harm Subject, Family or Other
		Threat - Other
Threats to individual or others	Threats to individual Threats to family	Emotional Abuse - Familial Related Emotional Abuse - Intimacy Related Emotional Abuse - Verbal/Manipulation Emotional Abuse - Other
Psychological abuse	Psychological abuse	
Physical abuse	Physical abuse	Physical abuse
	Denial of freedom to refuse client	
	Denial of freedom to refuse certain acts	
	Denial of freedom to use a condom	Reproductive Coercion
Sexual abuse	sexual abuse	Sexual Abuse
False promises/deception	False promises/deception	Isolation - Creates Distrust of Others Fraud/Misrepresentation of Job
Use of psychoactive substances	Giving of drugs Giving of alcohol	Induces/Exploits Substance Abuse Issues
		Isolation - Controls in Public
		Isolation - Denies Access to Support
		Isolation - Keeps Confined
		Isolation - Moves Subjects Frequently
Denied freedom of movement/kept in isolation	Denied freedom of movement	
Limits/restricts medical services	Denied medical treatment	Isolation - Limits Access to Medical
Excessive working hours	Excessive working hours	
Restricts access to/manipulates children		Restricts access to/manipulates children
Threat of action by law enforcement	Threat of action by law enforcement/deportation	Threat to Report to Immigration Threat to Report to Police
Withholds/denies basic necessities	Denied food/drink	Withholds/denies basic necessities
	Withholding of identity documents	
	Withholding of travel documents	
Withholds/Destroys Important Documents		Withholds/Destroys Important Documents
Other	Other	Other

Developing new standards

- Perpetrators
- Route data
- Country-level reporting

STANDARDS FOR ANONYMIZATION AND SHARING

- De-identification processes
 - No PII shared
 - Reduction in detail (Ex – Age range provided instead of exact age; locations)
 - K-Anonymization, k=11
 - Anonymization for visualizations, based on population data
 - Frequency of updates

Example of k-anonymity

Original DB

Birth day	gender	Zipcode
21/1/79	M	53715
10/1/79	F	55410
1/10/44	F	90210
21/2/83	M	02274
19/4/82	M	02237



2-anonymized DB

	Birth day	gender	Zipcode
group 1	*/1/79	human	5****
	*/1/79	human	5****
suppress	1/10/44	F	90210
group 2	*/*/8*	M	022**
	//8*	M	022**

DATA STANDARDS - EXAMPLES

The International Standard Industry Classification (ISIC)

Section	C	Manufacturing
Division	13	Manufacture of Textiles
Group	139	Manufacture of Other Textiles
Class	1393	Manufacture of Carpets and Rugs

STANDARDS FOR ANONYMIZATION AND SHARING

- De-identification processes
 - No PII shared
 - Reduction in detail (Ex – Age range provided instead of exact age; locations)
 - K-Anonymization, k=11
 - Anonymization for visualizations, based on population data
 - Frequency of updates

NIGERIA

COUNTRY OF CITIZENSHIP

0.65% VICTIMS % 501 VICTIMS

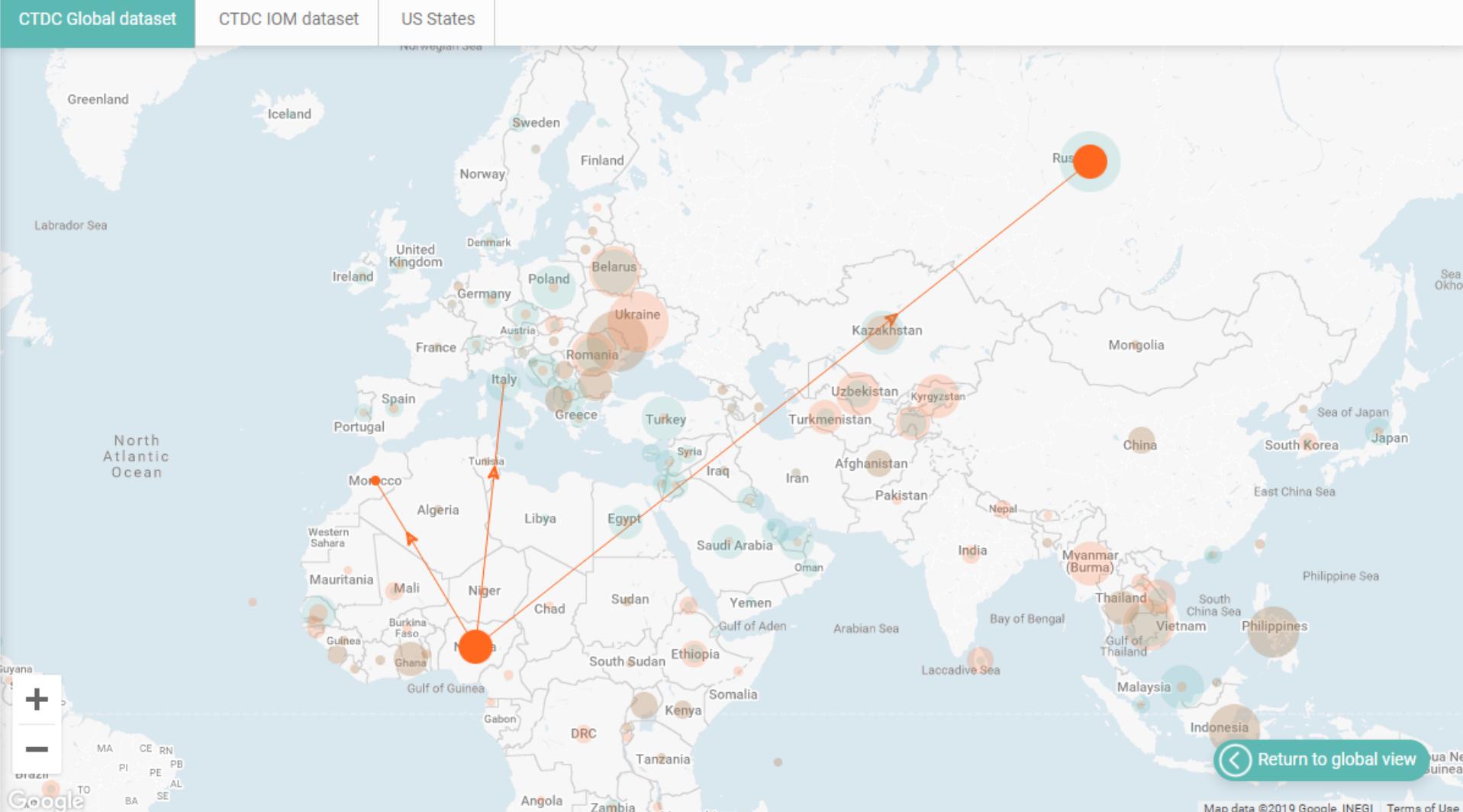
GENDER OF VICTIMS
Female 91% Male 9% Transgender -%

TYPE OF EXPLOITATION
20.71% Labour 1.52% Other 77.78% Sexual

COUNTRY OF EXPLOITATION
39.32% Russian Federation 11.58% Italy
9.78% Morocco

AGE BREAKDOWN

0 - 8	6.15%
9 - 17	5.27%
18 - 20	11.87%
21 - 23	21.54%
24 - 26	20.22%
27 - 29	16.92%



INDONESIA

COUNTRY OF ORIGIN



UNITED ARAB EMIRATES

COUNTRY OF EXPLOITATION

Indonesia Cambodia Myanmar Indonesia

United Arab Emirates Malaysia

0.54%

VICTIMS % i

416

VICTIMS

GENDER OF VICTIMS

99% Female

1% Male

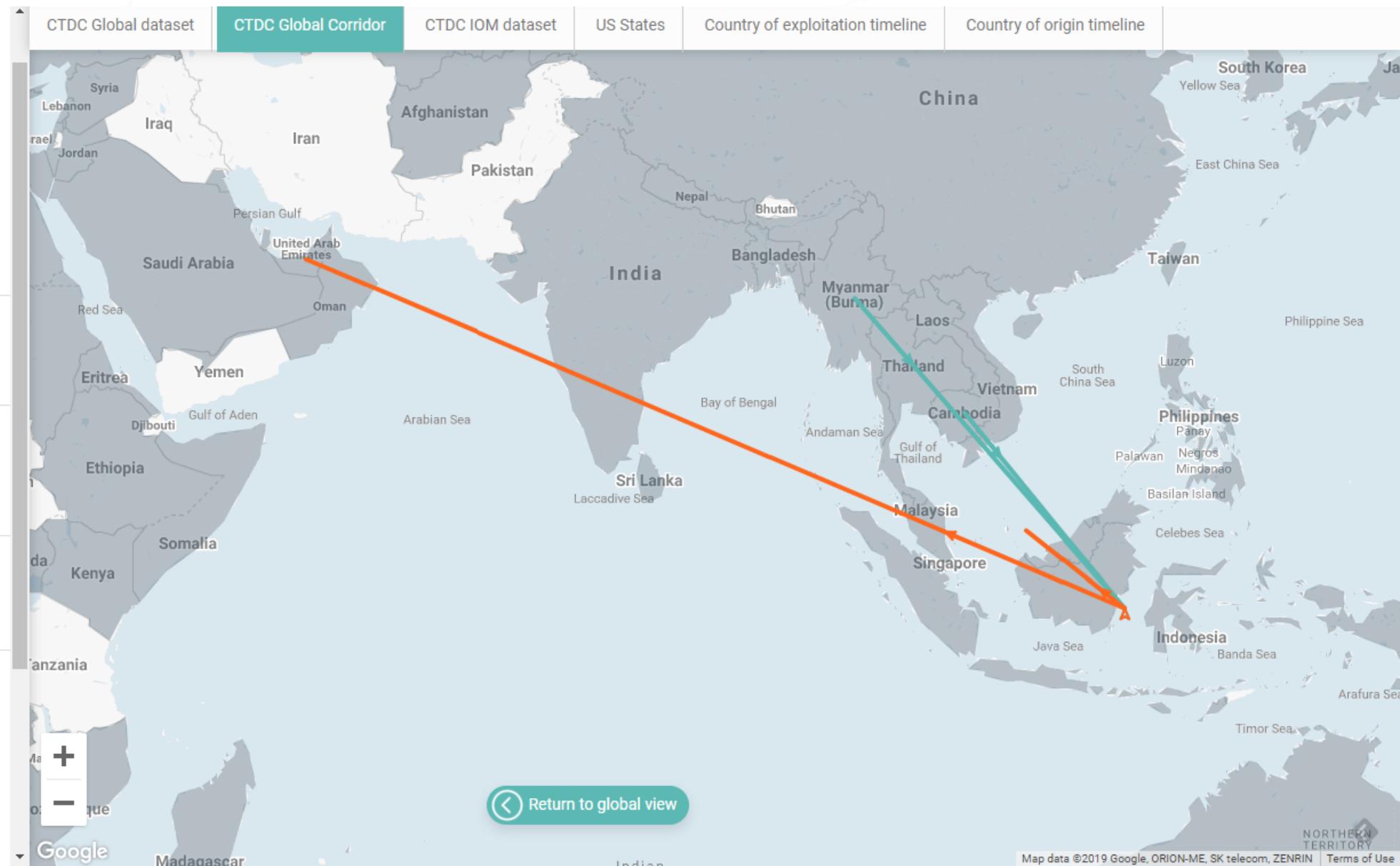
-% Transgender

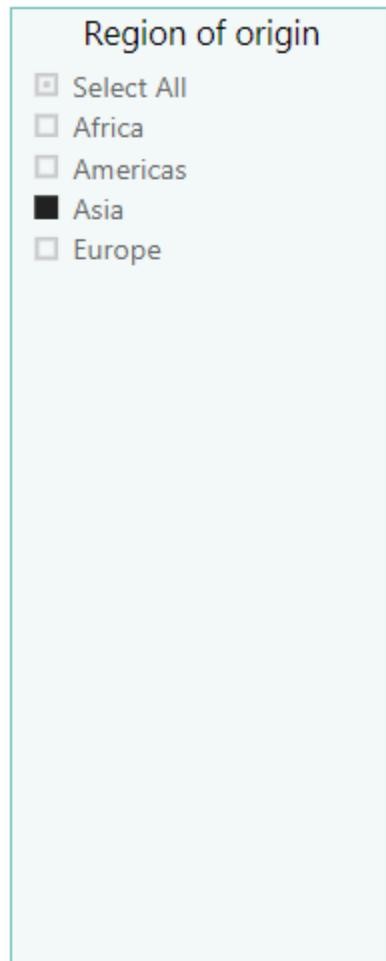
TYPE OF EXPLOITATION

100% Labour

AGE BREAKDOWN

0 - 8		1.20%
9 - 17		0.24%
18 - 20		1.68%
21 - 23		7.21%
24 - 26		6.01%
27 - 29		15.38%

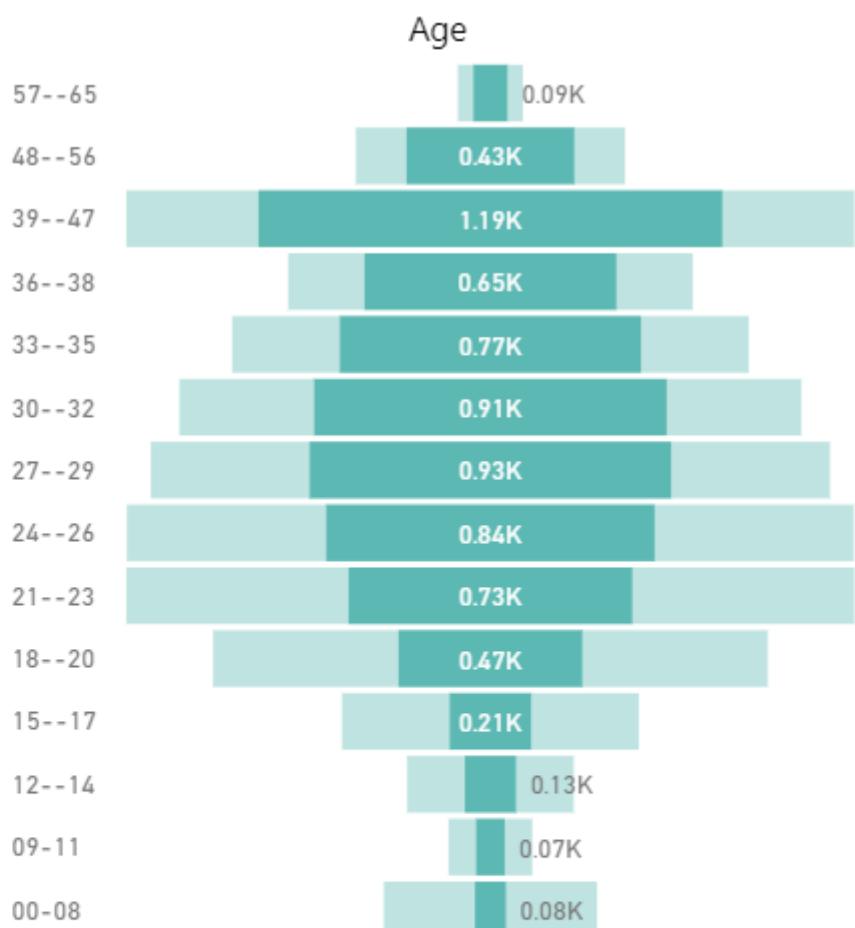




Victims count

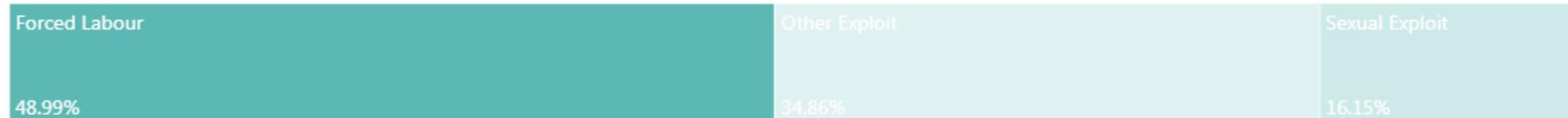
72.29K

Gender



Type of exploitation

Forced Labour ● Other Exploit ● Sexual Exploit



Differential privacy

Data governance

Some Applications for Case Data and Relationship with Prevalence Work

- Aggregation/reporting/supporting national government reporting mechanisms
- Imputation
- Modelling vulnerability/risk
- Complementing survey and administrative data
- Multiple Systems Estimation (MSE)

CONTACT

Harry Cook

Data Management & Research Specialist
Migrant Assistance Division

IOM Geneva

hcook@iom.int

+41 795822059