Data-driven decision making and storytelling

Unit 5

Nepal Data Literacy Program, 2019

Organized by

Supported by

THE WORLD BANK
IBRD • IDA

UKaid
from the British people

Objectives of the Unit

- Familiarize with the process of conceptualizing a hypothesis and mapping a story
- Understand how to summarize and simplify data insights to emphasize your finding
- Familiarize with the process of using analysis/story to make impact
- Work on a team-based data-driven analysis/story and share with the larger group

Image source: Cookerly.com

Module 1

Story Mapping

Objectives of data-driven storytelling

- Proactive process that digs deeply into an issue

- Producing new information

- Putting known information together

- Revealing secrets or uncovering issues

- Bearing witness, and investigating ideas

- Providing nuanced context, and explaining 'why'

- Not always about bad news

Telling a data-driven story is about analysing data and synthesizing insights, not merely reporting facts

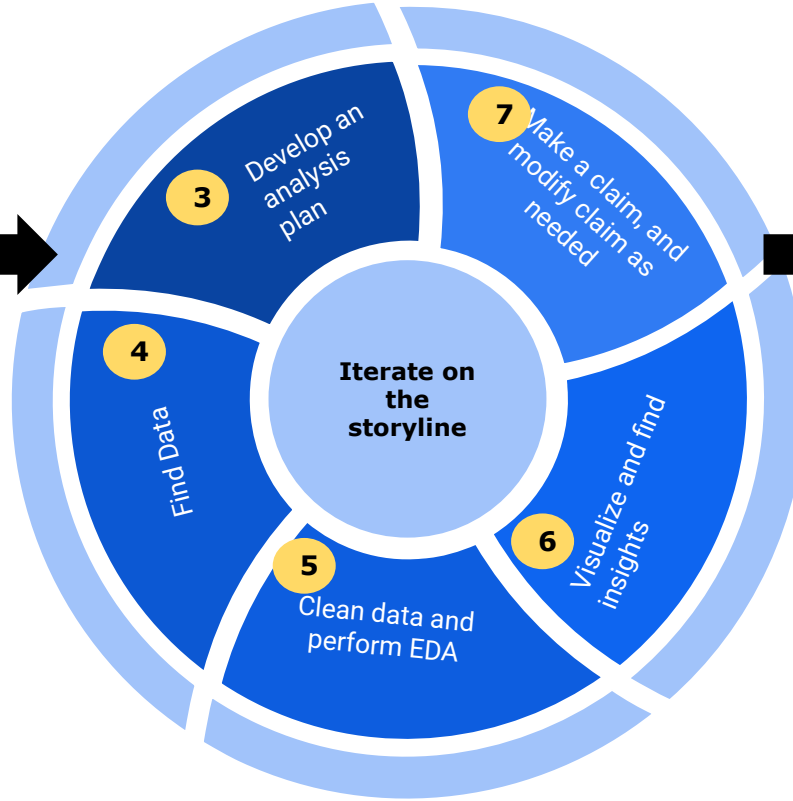| Reporting	Data driven story
Descriptive	Prescriptive
What?	Why?
Backward looking	Forward looking
Raise questions	Answer questions
Data → Information	Data + Information → Insights
Reports, dashboards, alerts	Findings, recommendations, predictions
Context	Context + solutions + storytelling

Story mapping process visualized

1
Identify an issue you want to analyze or story you want to narrate

2
Come up with a hypothesis or question to test

3 Develop an analysis plan

7 Make a claim, and modify claim as needed

4 Find Data

5 Clean data and perform EDA

6 Visualize and find insights

Iterate on the storyline

8
Develop your storyline based on your claims

9
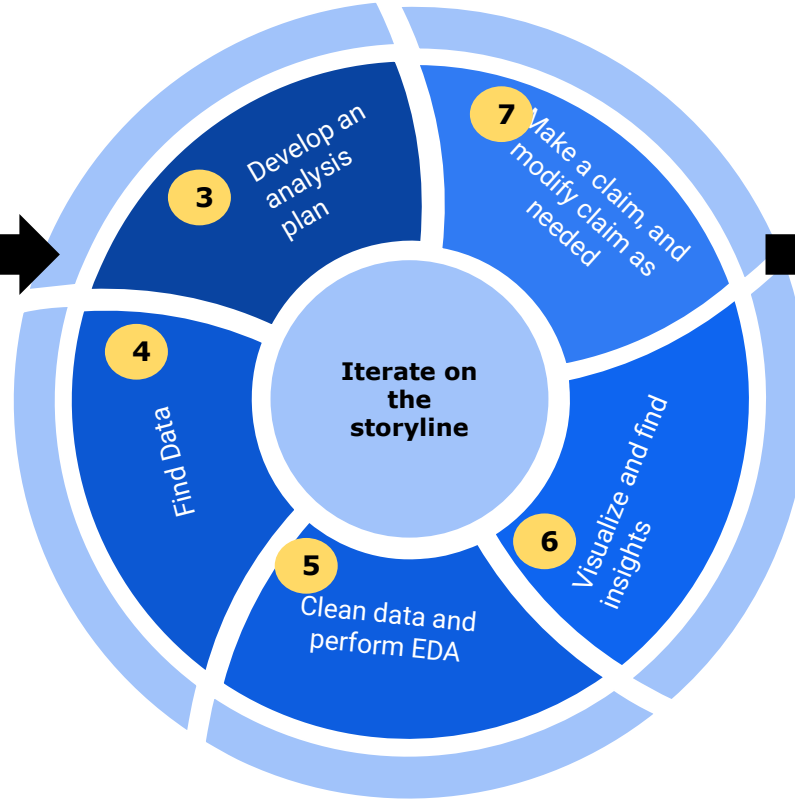Sharpen your storyline, provide context and share with your audience

Story mapping process visualized

1 Identify an issue you want to analyze or story you want to narrate

2 Come up with a hypothesis or question to test

3 Develop an analysis plan

4 Find Data

5 Clean data and perform EDA

6 Visualize and find insights

7 Make a claim, and modify claim as needed

Iterate on the storyline

8 Develop your storyline based on your claims

9 Sharpen your storyline, provide context and share with your audience

Can you tell what aspects of this story mapping process was covered where in the workshop?

Story mapping is a detailed and iterative process where you:

1. Identify an issue you want to analyze or a story you want to narrate

2. Come up with a research question/hypothesis to test

3. Develop an analysis plan for your hypothesis

4. Find related data: must be timely and relevant *(Unit 2)*

5. Clean and scrub data and perform EDA *(Unit 3)*

6. Visualize, find insights and related limitations from data *(Unit 3)*

7. Make a claim, and modify claims as needed *(Unit 4)*

8. Develop and sharpen your storyline, provide context and share with your audience

Identifying an issue to investigate

- Read/reflect intensively to gather information and inspiration
- Come up with a list of questions
- Understand the question/problem(s)
  - Redefine the problem definition
  - Ask WHY to broaden your horizon
  - Revise the problem definition and identify underlying problem
- Come up with multiple potential answers/ideas

Illustrative example

Let's say we are an NGO working to empower women in Nepal and we are trying to answer the following question: **How can we increase the literacy rate of women?**

1. **Original Problem:** "Increase the literacy rate of women"

2. **Why?** To decrease gender related violence in the rural society

    a. *Insights: Could we conduct workshops targeting women informing them about the policies (short term win)*

3. **Why? -** To increase the number of productive workers in the society

4. **Why? -** Educating a girl child has a high ROI and positive externality on all aspects

Developing an analysis plan to investigate your hypotheses

1. Write all your ideas/hypotheses and pick the top 2-3 hypotheses that resonates with you the most

2. Read analyses that have already been done on the topic to understand the issues better

3. Write down all the datasets you would need to investigate those hypotheses. Since availability of data is key, based on what you find, you might have to streamline/cut some hypotheses out

4. Write down all the questions you would need to ask the data to prove or disprove your hypothesis.

5. Interview the experts, people affected, decision-makers and stakeholders who can explain and put a human face to the issue

Selecting right data visualization design to tell your story
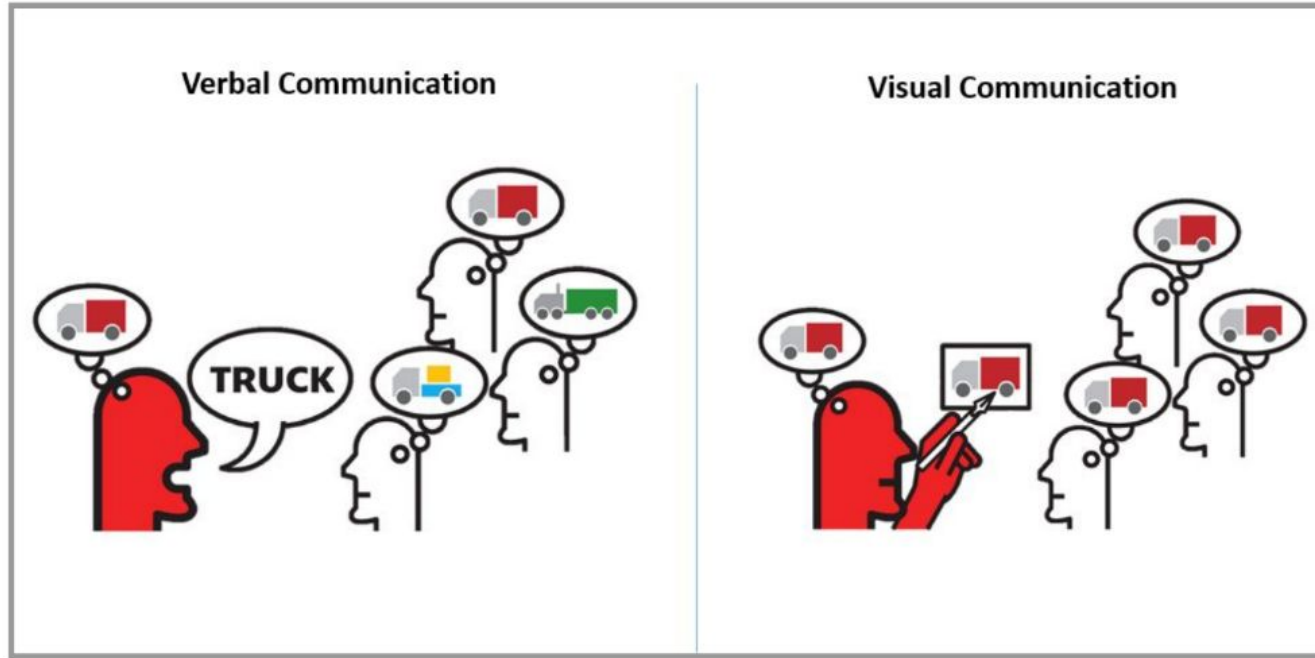
Unit 2, 3 and 4 had lessons on how to match data and graph types. Please keep those lessons in mind!

2. Coming up with a Hypothesis

- Hypotheses give you something to verify
- Hypotheses increase chances of discovering the unknown
- Hypotheses make it easier to manage your project
- Hypotheses are a tool you can use again and again
- Hypotheses ensures you deliver a story, not just a mass of data

Difference between Hypothesis and Claim:
- A hypothesis is a starting point, a statement that can be experimentally tested
- Claim is a well-founded statement that has been tested

Question/ hypothesis examples

- In Kathmandu, babies born between 2010 and 2019 are more likely to have respiratory illness

- Nepalis in Province 2 or 6 are more likely to be multidimensionally poor than Nepalis in Province 3

Module 2

Summarizing and Simplifying Data Insights

Simplifying Percentages: Examples

Percentage	Fraction	Population rate
47.1% of adult males in Nepal use the internet	Around half of Nepalese men use/have used the internet	One in two Nepalese men use/have used the internet
72.6% of women in Nepal use mobile telephone	Around three quarters of women in Nepal use mobile telephone	Three out of four women in Nepal use mobile telephone

Source: NDHS 2016

Rounding Off Numbers: Examples

Decimals	Rounded Off
Nepal's population was 29.3 million in 2017	Nearly 30 million people live in Nepal
Nepal's GDP was $24.88 billion in 2017	Nepal's GDP was nearly ~25 billion in 2017
Nepal's fertility rate was 2.08 in 2017	Women in Nepal had 2 children on average in 2017

Source: World Bank Database

Simplifying Percentages: Exercise

- In Nepal, 4.8% of people from Province 6 have migrated abroad
- In Nepal, 29.2% women who are abroad were reported to be between age 20-24 years

Source: NDHS 2016

Comparing Numbers: Example

1. Let's simplify this statement:
"49.7% of births are registered in rural areas and 84.5% of births are registered in urban areas"

2. To do this convert the percentages in the statement to fractions:
- 49.7% = 50/100 =5/10 = **10/20**
- 84.5% = 85/100 = **17/20**

3. So now we can say:
*"Compared to **10 out of 20** births registered in rural areas, **17 out of 20** births are registered in urban areas"*

Comparing Numbers: Exercise

Can you try playing with these statistics and seeing if you can come up with a better way of comparing the two numbers?

1.

   6.5% kids under 5 who had diarrhea had access to an improved toilet facility
   9.1% kids under 5 who suffered from Diarrhoea didn't have access to a proper toilet facility

2.

   69.1% of households in province 2 lived within less than 30 minutes walking distance from a health facility in 2016
   23.6% of households in province 6 lived within less than 30 minutes walking distance from a health facility in 2016

3.   In Nepal, 9.6% women and 78.3% men migrate abroad for work

Source: NDHS 2016

Module 3

Increasing Impact of Your Story

How ideas spread

- "Diffusion is essentially a social process through which people talking to people spread an innovation," -- Everett Rogers, an eminent American communication theorist and sociologist

- "Mass media can introduce a new idea to people. But, Rogers showed, people follow the lead of other people they know and trust when they decide whether to take it up. Every change requires effort, and the decision to make that effort is a social process." [1]

1.   Source: https://www.newyorker.com/magazine/2013/07/29/slow-ideas

Goals of open data stories and analysis

- Widespread public debate

- Overcoming indifference or inertia by stakeholders

- Engaging public, politicians and donors

- Demand for more data from government by citizens for better governance

Examples of reaching your audience

World Happiness Report 2019

MARCH 20, 2019

The World Happiness Report is a landmark survey of the state of global happiness that ranks 156 countries by how happy their citizens perceive themselves to be. This year's World Happiness Report focuses on happiness and the community: how happiness has evolved over the past dozen years, with a focus on the technologies, social norms, conflicts and government policies that have driven those changes.

Read the Report

WHR 2019

Source: https://worldhappiness.report/ed/2019/

Study reveals city potable water rife with bacteria

- Post Report, Kathmandu

Source: https://kathmandupost.ekantipur.com/news/2018-05-16/study-reveals-city-potable-water-rife-with-bacteria.html

Red Meat and Processed Meat

IARC Monographs on the Evaluation of Carcinogenic Risks to Humans
IARC Monographs, Volume 114
IARC

ISBN-13 9789283201526   ISBN-10 9283201523

Order Number 17200114   Format Paper Back

Price CHF 55.00 / US$ 66.00   Developing countries: CHF 38.50

English 2018 513 pages

Summary
Table of contents

Summary

This volume of the IARC Monographs provides evaluations of the consumption of

Source: https://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=72&codcch=114
And

Yes, bacon really is killing us

▲ Photograph: Guardian Design Team

Decades' worth of research proves that chemicals used to make bacon do cause cancer. So how did the meat industry convince us it was safe? By Bee Wilson

Thu 1 Mar 2018
06.00 GMT

There was a little cafe I used to go to that did the best bacon sandwiches. They came in a soft and pillowy white bap. The bacon, thick-cut from a local butcher, was midway between crispy and chewy. Ketchup and HP sauce were served in miniature jars with the sandwich, so you could dab on the exact amount you liked. That was all there was to it: just bread and bacon and sauce. Eating one of these sandwiches, as I did every few weeks, with a cup of strong coffee, felt like an uncomplicated pleasure.

Source: https://www.theguardian.com/news/2018/mar/01/bacon-cancer-processed-meats-nitrates-nitrites-sausages

Examples of reaching your audience

RAPE STATISTICS, BY AGE AND PROVINCE

Five ways to promote your story

- Find a storyline for your analysis that will resonate with your audience

- Publish a story or op-ed or blog in a media outlet or your website

- Use social media and contacts to spread the word

- Find an influencer in your network who can serve as spokesperson for your analysis or story

- Organize an event when applicable and possible to discuss your findings

Reaching Your Audience -- Exercise

- In your teams, discuss ideas, instances, and experiences you've had in the past with disseminating your stories to target audience. Focus on things that worked for you (10 mins)
- Choose one person from your team to share ideas with the larger group (5 mins)

Module 4

Data Privacy

Data Privacy
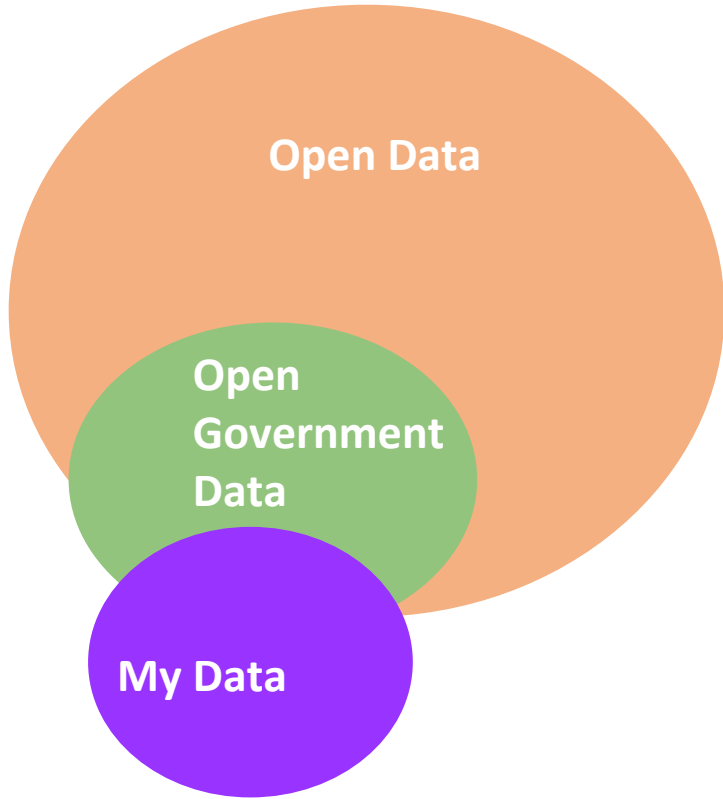
- Open data democratizes information as opposed to confining the power of data to its producers and those in a position to pay for access but it can also pose risks to privacy

Examples of privacy breach:
- **Adhaar Data Leak**
- **Target's Marketing Analytics**

Team exercise: Data Privacy

- What are the arguments in favor of and against releasing data in the following situations?
  - Neighborhood crime data (Teams 1,5,9,13,17)
  - List of individuals infected by Bird Flu with street addresses (Teams 2,6,10,14,18)
  - List of families receiving government financial assistance (Teams 3,7,11,15)
  - Names of locations of hospitals and prisons with the highest fatality rates (Teams 4,8,12,16)

Discuss your topic in teams for 10 mins and share key insights with everyone in the room

Ways to ensure data privacy

PRIVACY-AWARE OPEN DATA

PURPOSE ASSESSMENT

PRIVACY IMPACT ASSESSMENT

PSEUDONYMOUS   ANONYMIZED   AGGREGATED

ACCESS RESTRICTIONS

RE-USE RESTRICTIONS
↳ LICENCE

RELEASE DATA
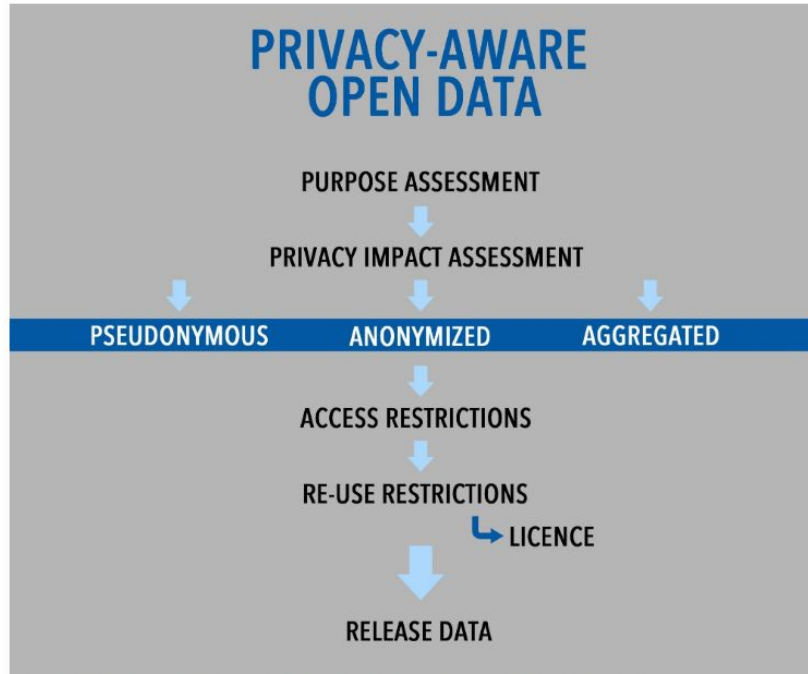
An overview of the steps to consider to achieve a privacy-aware open data initiative.

Technical Tools:
- Pseudonymous Data
- Anonymized Data
- Aggregated Data

Policy Approaches:
- Access Restrictions
- Re-Use Restrictions

For Data Users:
- Avoid using data with personally identifying information
- Report to relevant authorities in case you unknowingly came across information with personally identifiable details

Module 5

Data Expedition

Data Driven Decision Making - Capstone Project

- An opportunity to practice the skills you have been learning during the workshop and experience a more hands on approach to solving a data problem
- **<u>Goal:</u>**
  - Work with your teams to **come up with a data-driven story** (an infographic, poster,or a 2-page document etc.)
  - You will get to **present your work** (5 minutes) in front of a panel of judges on the last day of the workshop
  - Selected work will be published on the World Bank Blog and Nepal Data literacy site
- **<u>Things needed:</u>**
  - 8-10 hours to work on the project (6 hours allocated during the workshop)
  - Knowledge from the workshop, Creativity and Dedication!

Suggested project timeline

Friday, June 28	Monday, July 1	Tuesday, July 2	Wednesday, July 3	Thursday, July 4
Topic finalization + analysis plan	**Find, clean and Scrape data**	**Visualize data, Perform EDA, and make claims**	**Draw meaningful conclusions**	**Final Presentation**
+ Pick out a topic of interest, specifically related to Nepal **+ Formulate a hypothesis** or a question related to the topic **+ Decide what kind of data is needed**	**+ Find the data** **+Scrape/Clean the data** as needed to make it ready for analysis	**+ Use data visualization techniques** to learn about patterns and structure in the data **+ Perform exploratory data analysis** to gain meaningful insights **+ Make claims from the analysis** and **assess the validity of claims** using various analytical techniques	**+ Draw meaningful conclusions** by providing evidence that supports/negates the hypothesis **+ Create a data driven story** *Deadline to submit projects by 5:00pm*	**+ Teams to present their works** to the judges and their cohort in 5 minutes

Guidelines

- **Start early** with your team to brainstorm on topics or issues you are interested in writing about or finding more about
- Ensure your **topics are inline with the SDGs/related to Nepal** and there is **enough data** out there to support your story
- Ensure you **apply concepts from** each of **the 5 units** you have been taught
- **Prepare your visualizations** using any of the tools you have been taught
- **Prepare a short presentation** of your visualization/story
- **The instructors are available for support**, so feel free to ask them questions when needed.
- Share your final presentation through a public URL. Here's a [useful document](#) that shows you how to make your doc public
- Give **room for creativity** but most importantly, **HAVE FUN!!**

Example of Output

FACTSHEET 2018
NEPAL

World Health Organization
Regional Office for South-East Asia

Heart disease and stroke are the commonest ways by which tobacco kills people

QUIT TOBACCO USE NOW - FOR A HEALTHIER HEART

Gross national income per capita (lower-income country)	Total population	Youth population (13–17 years)	Economically productive population (30–69 years)
US$ 730	29.3 million	3.4 million = 11%	10.6 million = 36%

Tobacco[1] kills **27 137** people each year **14.9%** of all deaths

The most common way tobacco kills is from cardiovascular diseases (CVDs)[2]

Others 6%
CMNND[a] 12%
CRD[b] 21%
CVD 53%
Cancers 8%

Distribution of tobacco deaths by cause

CVDs in younger people are more likely to be caused by tobacco use

37% 39% 32% 23%

% of CVD deaths caused by tobacco

30–44 45–59 60–69 70+

Age in years

CVDs are the number one cause of death, causing **51 028** each year (28.1% of all deaths), as well as of premature death

CVD deaths caused by tobacco use
14 432 deaths
28% of all CVD deaths each year

Top 5 causes of overall death
1 Ischemic heart disease
2 Chronic obstructive pulmonary disease
3 Cerebrovascular disease
4 Lower respiratory infection
5 Diarrheal disease

Top 5 causes of premature death (YLL–years of life lost)
1 Ischemic heart disease
2 Lower respiratory infection
3 Neonatal encephalopathy
4 Cerebrovascular disease
5 Chronic obstructive pulmonary disease

Tobacco control is essential for preventing and controlling deaths and disability caused by CVDs

● Noncommunicable diseases ● Communicable, maternal, neonatal, and nutritional diseases

6 m current tobacco users and a substantial number of people exposed to secondhand smoke are at increased risk of CVDs

Current tobacco use among adults (%) (15–64 years) STEPS–2013
Tobacco users[5]: 30.8 48.1 14.1
Tobacco smokers[4]: 18.5 27.0 10.3
Smokeless tobacco users[5]: 17.8 31.3 4.8

Current tobacco use among youth (%) (13–15 years)
Tobacco users GSHS–2015: 7.2 9.5 4.8
Cigarette smokers GSHS–2015: 5.0 6.8 3.0
Smokeless tobacco users GYTS–2011: 16.2 19.7 12.9

Exposed to secondhand[6] smoke at home (%)
Adult (15–64 years) STEPS–2013: 36.1 37.3 35.0
Youth (13–15 years)* GSHS–2015: 54.1 57.9 50.7
*people smoked in their presence

● Total ● Male ● Female

Claim: *Heart diseases and stroke are the commonest ways by which tobacco kills people*

Facts to support the claim

Note: This is a sample infographic but we don't expect you to come up with advanced visualizations and graphic design as in this infographic.

Source: https://apps.who.int/iris/bitstream/handle/10665/272687/wntd_2018_nepal_fs.pdf?sequence=1&isAllowed=y

Examples of Output

Claim: *Pneumonia is the deadliest infection...*

Facts to support the claim

Pneumonia
The Deadliest Childhood Disease

Pneumonia is the deadliest infection: responsible for more deaths of children under 5 than any other infectious disease

Between 2000 and 2015 pneumonia mortality rate declined by **51%**

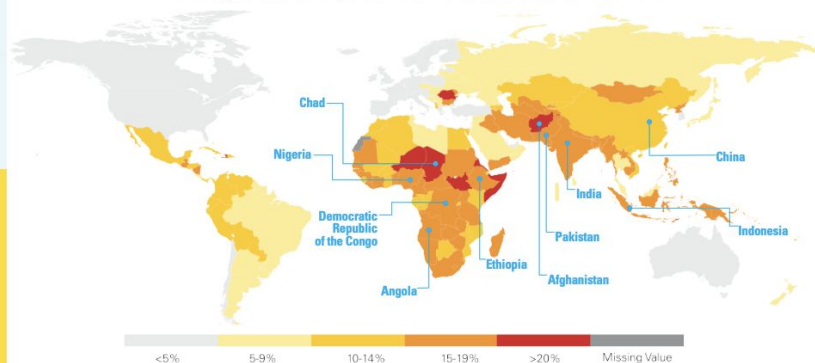**But** it is one of the lowest rates of decline as compared to other diseases

1 in 6 childhood deaths were due to pneumonia in 2015

922,000 per year
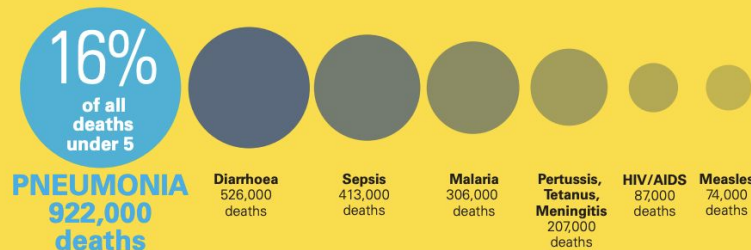2,500 per day
100 per hour

6 out of 10 childhood pneumonia deaths are concentrated in 10 countries

Percentage of deaths among children under-five attributable to pneumonia in 2015

Chad
Nigeria
Democratic Republic of the Congo
Angola
Ethiopia
China
India
Pakistan
Afghanistan
Indonesia

<5%   5-9%   10-14%   15-19%   >20%   Missing Value

Note: The boundaries and the names shown in and the designations used on these maps do not imply official endorsement or acceptance by the United Nations

Pneumonia is the number 1 killer among the leading infectious diseases

16% of all deaths under 5

PNEUMONIA 922,000 deaths

Diarrhoea 526,000 deaths

Sepsis 413,000 deaths

Malaria 306,000 deaths

Pertussis, Tetanus, Meningitis 207,000 deaths

HIV/AIDS 87,000 deaths

Measles 74,000 deaths

Source: https://data.unicef.org/resources/world-pneumonia-day-2015-infographic/

White papers are acceptable output too

METHODOLOGY

Stage I: Diagnostic:
- ☐ Literature Review
- ☐ Green Papers
- ☐ Research Studies
- ☐ District Consultations
- ☐ Education Conferences (National/Provi...)

Stage II: Prescriptive:
- ☐ Issue-based Consultations
- ☐ Organization-based Consultations
- ☐ Coordination with Development Partner...
- ☐ Thematic Papers

Stage III: Policy Development:
- ☐ The White Paper (December 2006)
- ☐ The White Paper (Revised- February 20...)
- ☐ The Policy Paper
- ☐ Strategy/strategies
- ☐ Implementation Plans

SAMPLE

HIGLIGHTS OF THE WHITE PAPER

The White Paper has been developed:

- In view of the socio-economic development policies of the state and international obligations through a evolutionary & participatory approach
- With the help of public and institutional incisive(ability to recognize) reflective views gained through a broad participatory process
- Based on commonalities of views, reflective of hopes of the people & views of experts for every sub-sector of education
- With an educational policy that is philosophical & thematic, not sectoral (about particular aspect of life)
- Proposing a number of policy options for debate and discussion; to be accepted, modified or discarded
- With feedback on the White Paper organized in the revised paper to ensure Stakeholders influence the final policy document
- To set a representative for a continued dialogue amongst all stakeholders in a reform process

SAMPLE

https://www.slideshare.net/EHTISHAMMANZOOR1/white-paper-on-education

Additional resources

Brainstorming topics

Sustainable Development Goals

Learn more about the SDGs

Sample topics:

Data & Gender Issues	https://data2x.org/
Data & Disaster Recovery/ Climate Change:	https://opendri.org/
Data & Oil, Gas, Mining:	https://resourcegovernance.org/ https://eiti.org/
Data & Health	https://www.who.int/healthinfo/civil_registration/en/
Data & Education	https://opendataimpactmap.org/education http://opendatanepal.com/dataset?oknp_category_label=Education
Data & Private Sector Growth:	http://www.opendata500.com/ https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information
Data & Agriculture / Food Security	https://www.godan.info/
Data & Transportation	http://opendatahandbook.org/solutions/en/Public-Transport-Data/ http://opendatanepal.com/dataset?oknp_category_label=Transportation
Data & Social Inclusion	https://www.worldbank.org/en/topic/social-inclusion

Sample topics

Data & Refugees / Migration:	https://migrationdataportal.org/themes/iom-data-overview https://data.humdata.org/organization/unhcr
Data & Foreign Aid to Nepal	https://amis.mof.gov.np/portal/ http://opendatahandbook.org/value-stories/en/effective-aid-in-nepal/
Data & Open Cities / Smart Cities	https://opencitiesproject.org/ https://www.opennorth.ca/
Data & Good Government	http://opengovimpact.org/ https://www.openduka.org/
Data & Transparent Government Contracts:	https://www.open-contracting.org/
Data & Elected Official Accountability	https://www.govtrack.us/congress/votes http://openparliament.pk/
Data & Elections	https://openelectiondata.net/en/ https://gottovote.cc/
Data & Crime / Public Safety	https://data.cityofnewyork.us/Public-Safety/Crime-Map-/5jvd-shfj http://crimemap.dc.gov/CrimeMapSearch.aspx
Data & Disability	http://www.d4dnepal.org/2018/12/01/improving-the-availability-of-disability-data/
Data & Indigenous Peoples:	https://www.stateofopendata.od4d.net/chapters/issues/indigenous-data.html

Helpful pointers for presenting

Start with the **ANSWER** first

Group your **SUPPORTING** statements

BACK-UP of the supporting statements

Create a compelling case, applying a powerful, overarching governing thought with strong, logically-structured supporting statements