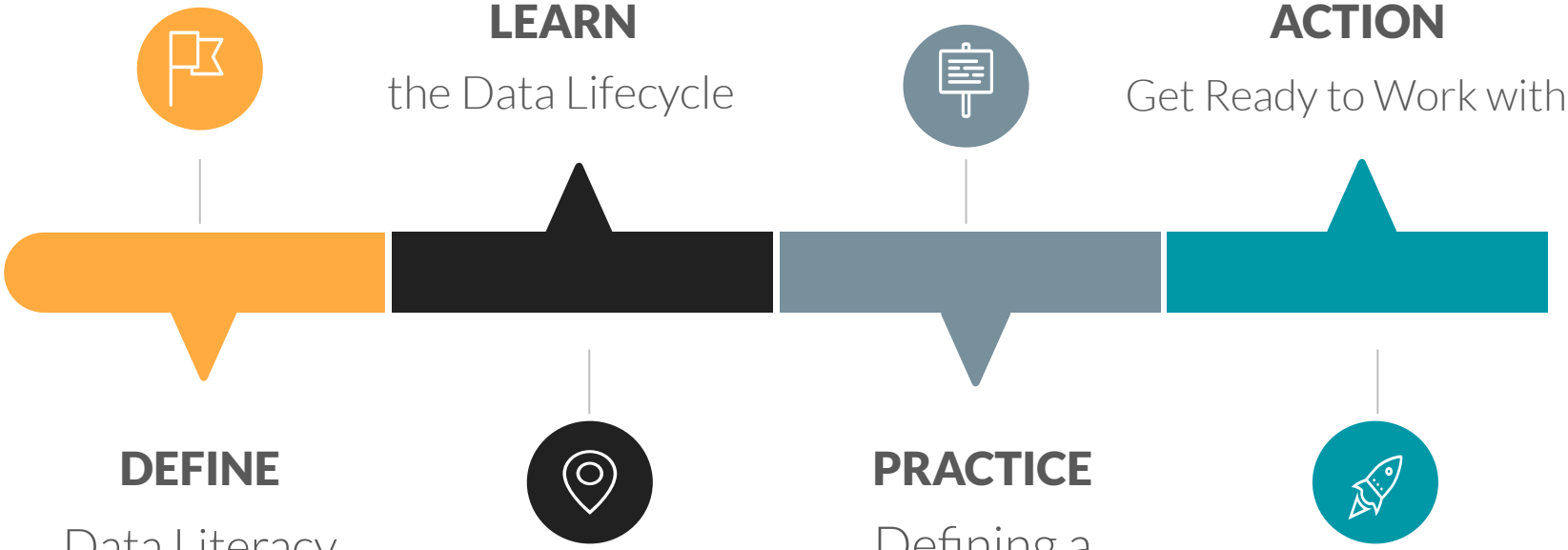
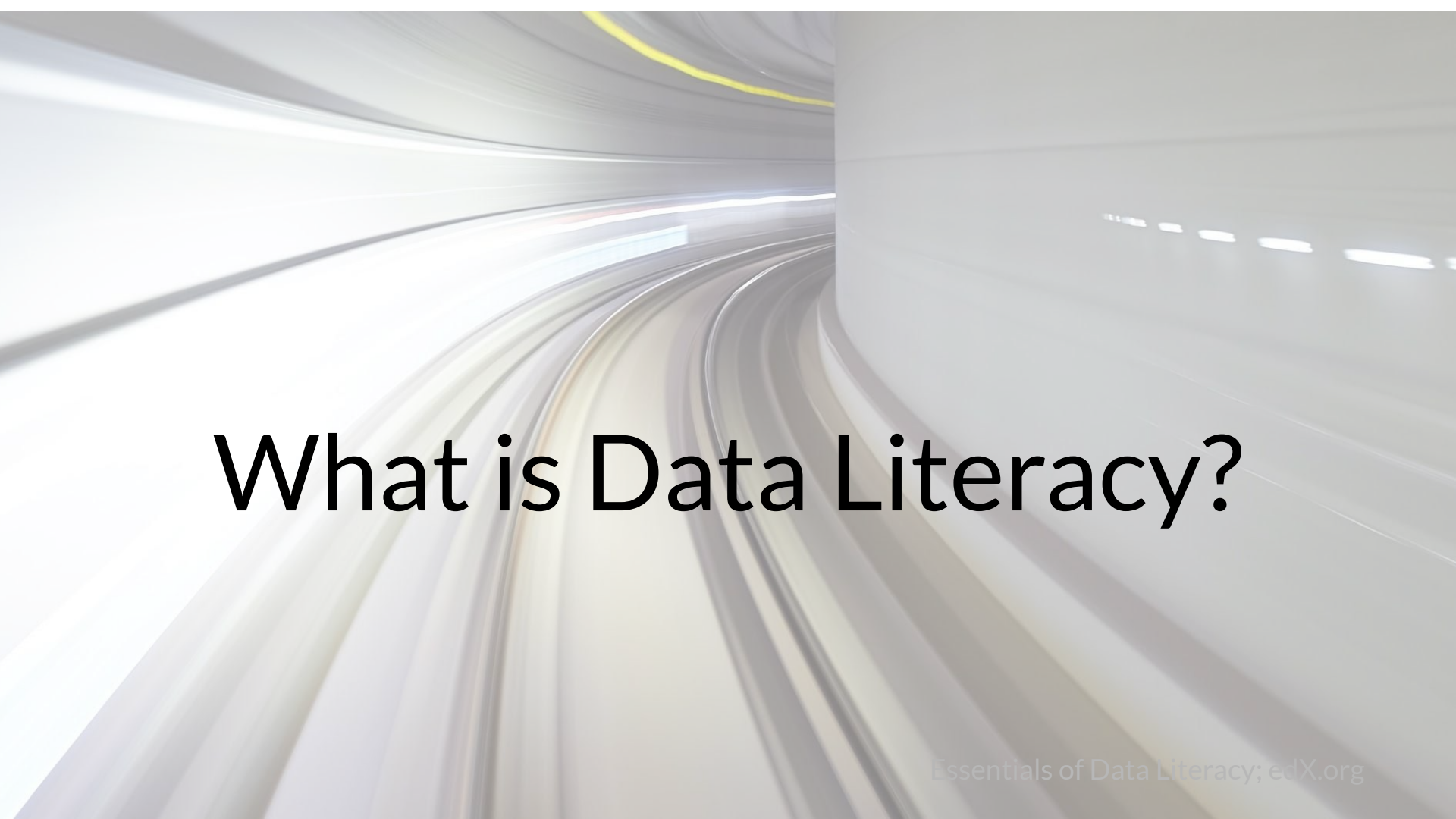
Welcome!

Let’s Get Started!


LEARN
the Data Lifecycle
ACTION
Get Ready to Work with R
DEFINE
Data Literacy
PRACTICE
Defining a
Research Question


What is Data Literacy?
Essentials of Data Literacy; edX.org

Defining Data Literacy

Understand

Work with

Communicate

Ways We Use Data

Planning & Programming

Experimentation & Intervention

Tracking

Reporting

Demonstrate Impact

Decision-making

Visualization

The Language of Data

Working with data is like cooking

Sorted


Arranged

Visually Presented

DATA

SORTED

ARRANGED

PRESENTED VISUALLY

Photography by Brandon Rossen Photography | www.brandonrossen.com

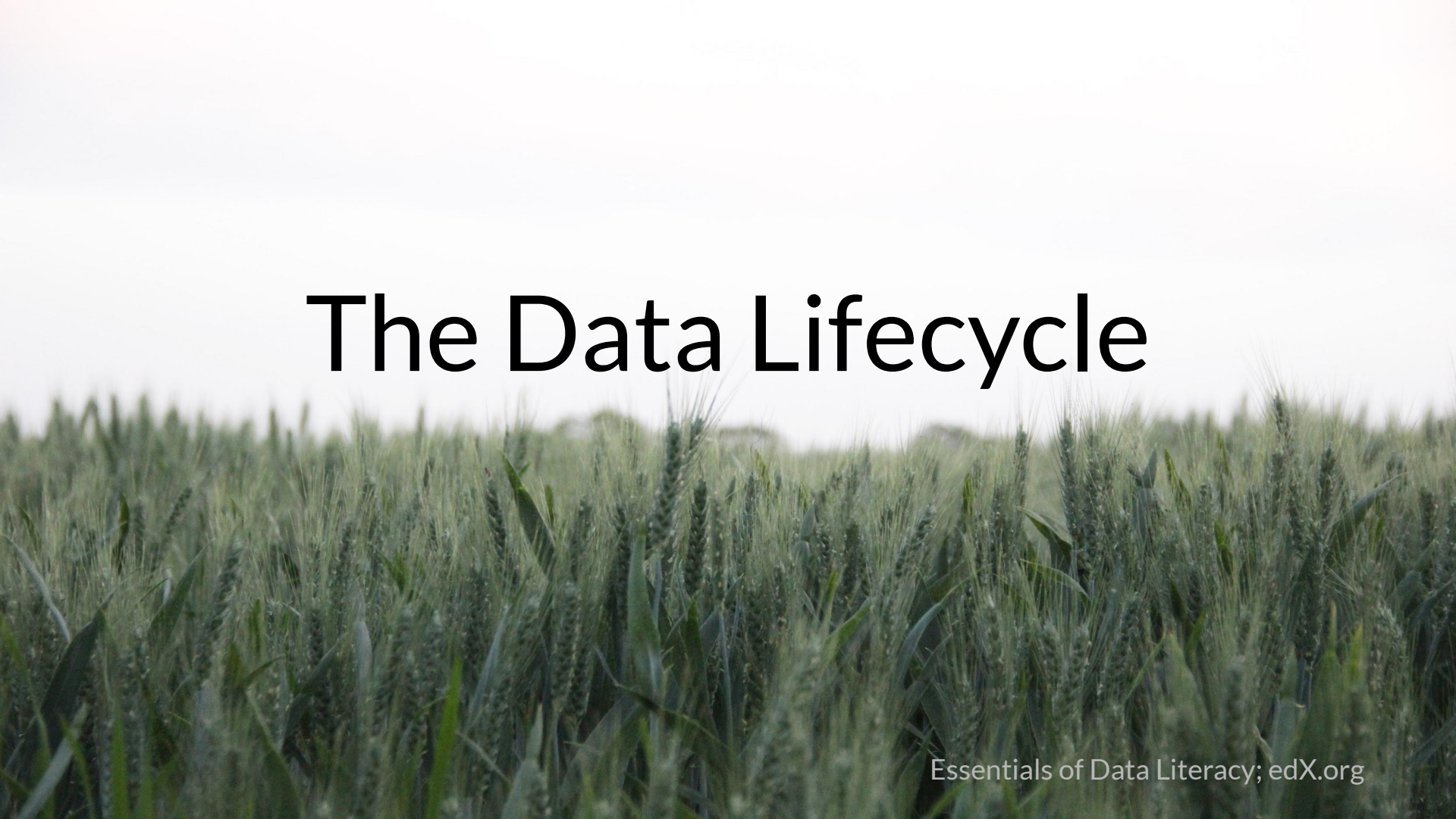
Take Home Messages

- **Know** the definition of data literacy
- **Recognize** that data is like a language, and it's also like cooking
- We will **practice** translating ideas into actionable data steps
- The **goal** is to know enough to be **conversant**, not necessarily fluent

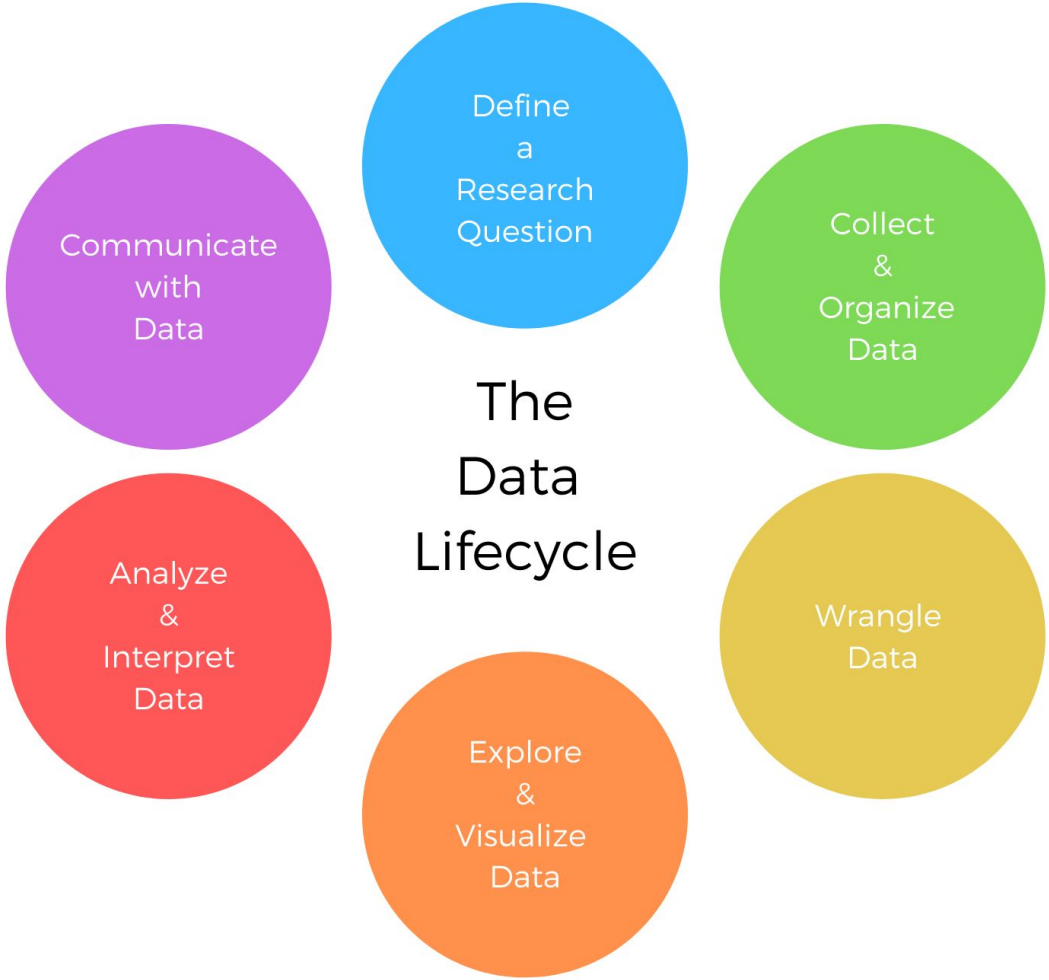
Take Home Messages

- You **don't have to become** a data scientist or data engineer
- The **benefits** of being data literate in your field/organization
  - You **help to bridge gaps** between technical and non-technical
  - You can **translate** a question or idea into a compelling story
  - You **help to make** others more data literate; **change the culture** of your organization to be more data-driven

The Data Lifecycle



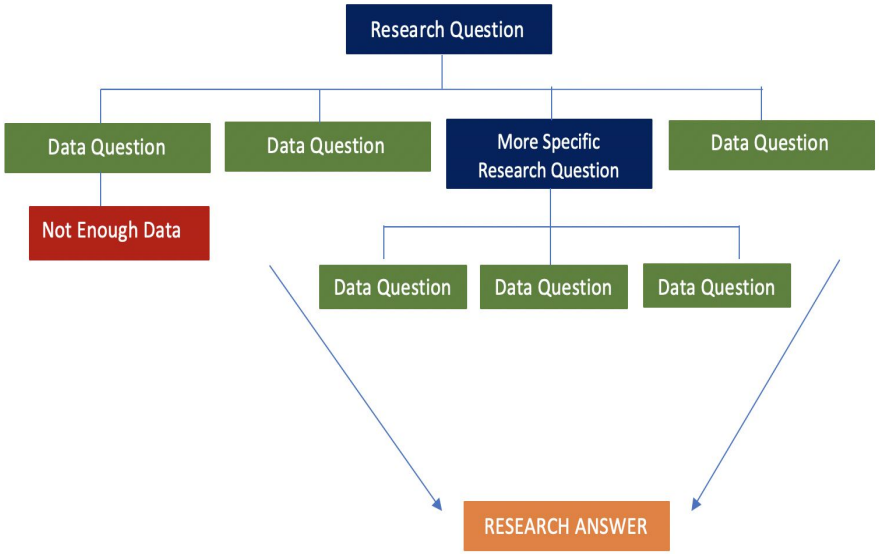
Essentials of Data Literacy; edX.org

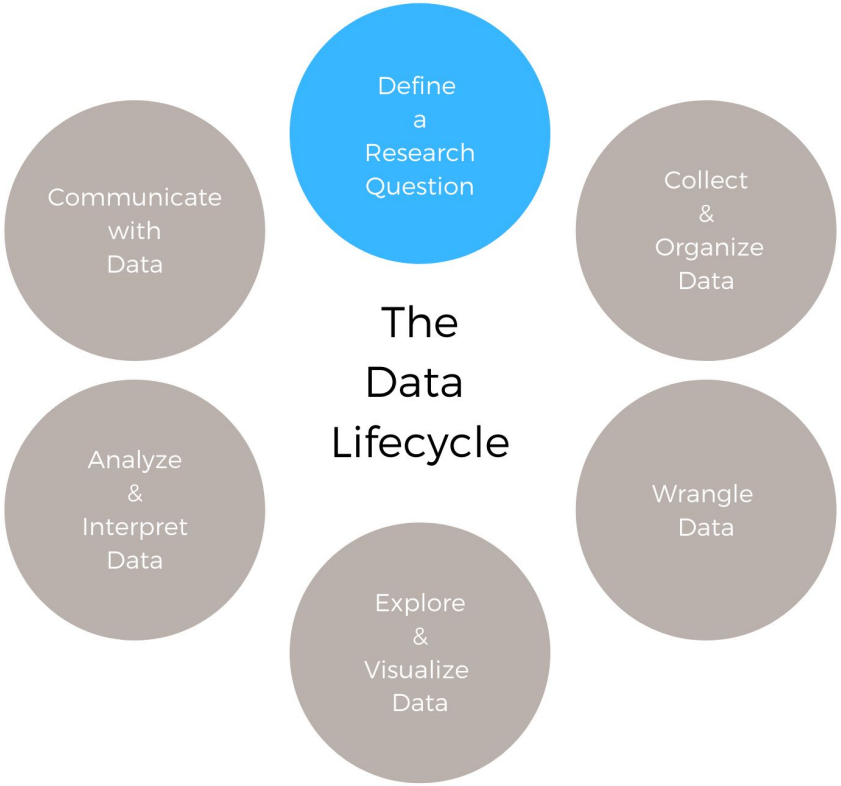

Define a Research Question
Collect & Organize Data
Wrangle Data
Explore & Visualize Data
Analyze & Interpret Data
Communicate with Data
The Data Lifecycle

The Data Lifecycle

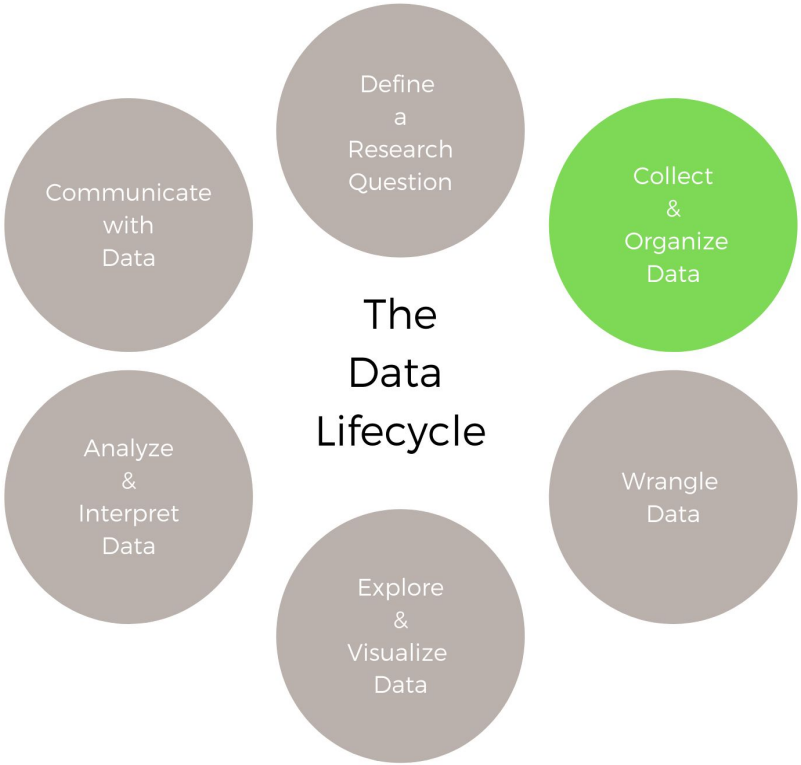
- Define a Research Question
- Collect & Organize Data
- Wrangle Data
- Explore & Visualize Data
- Analyze & Interpret Data
- Communicate with Data

The Data Lifecycle - Define a Research Question

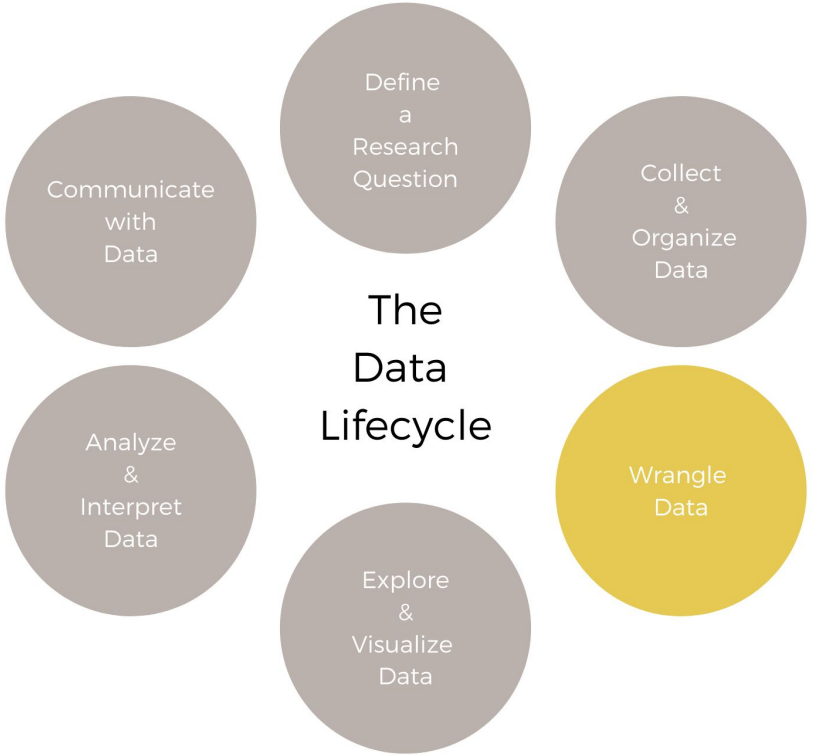




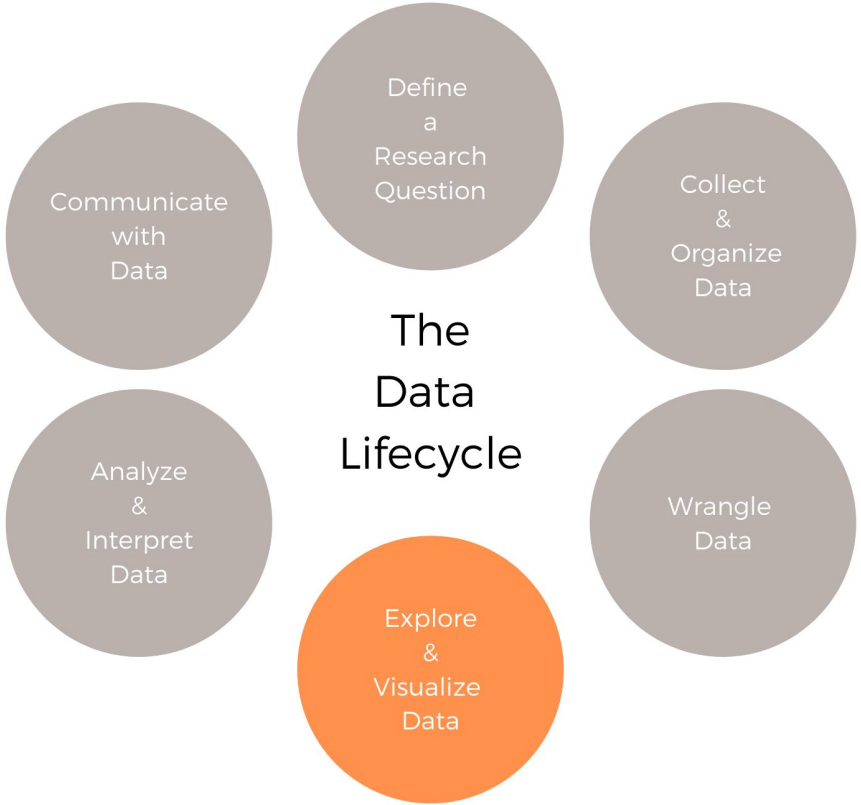
The Data Lifecycle - Collect & Organize Data



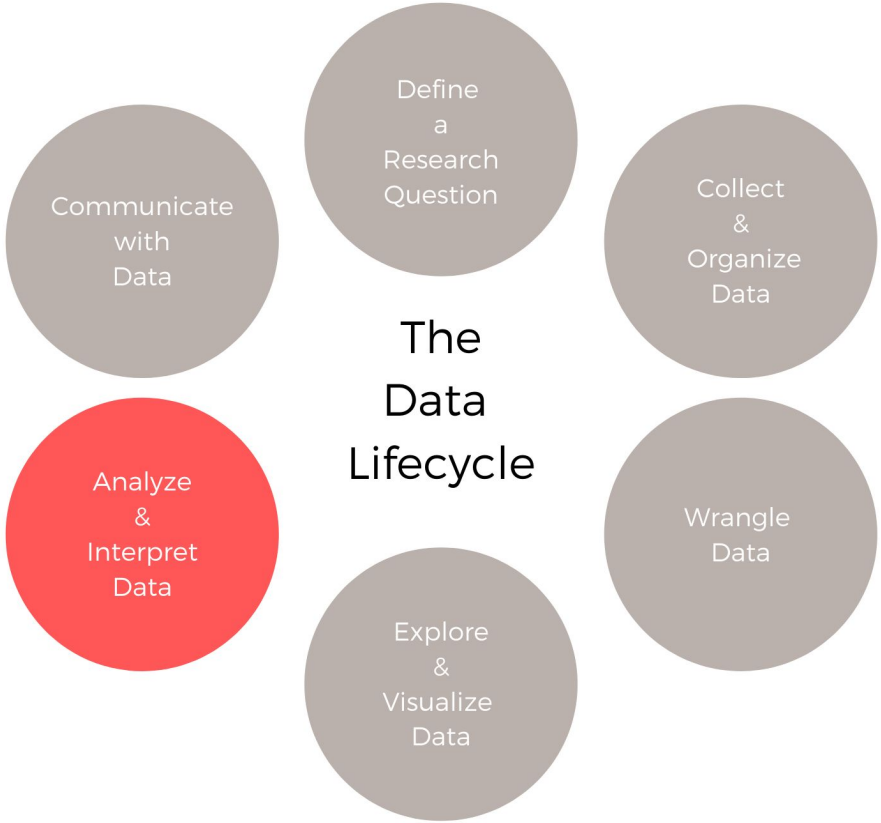
The Data Lifecycle - Wrangle Data



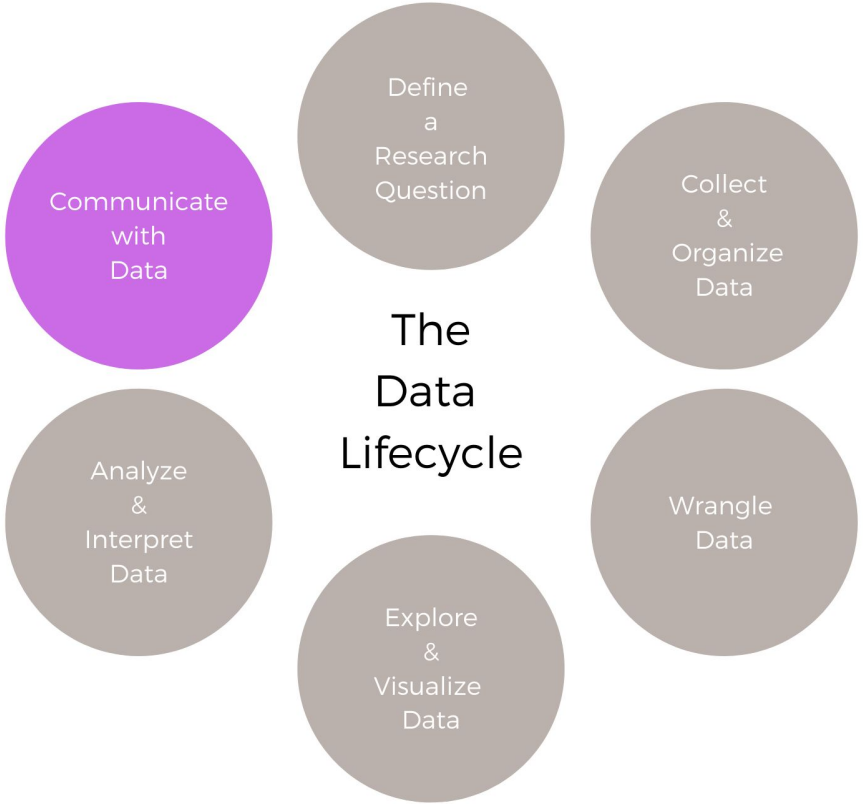
The Data Lifecycle - Explore & Visualize Data



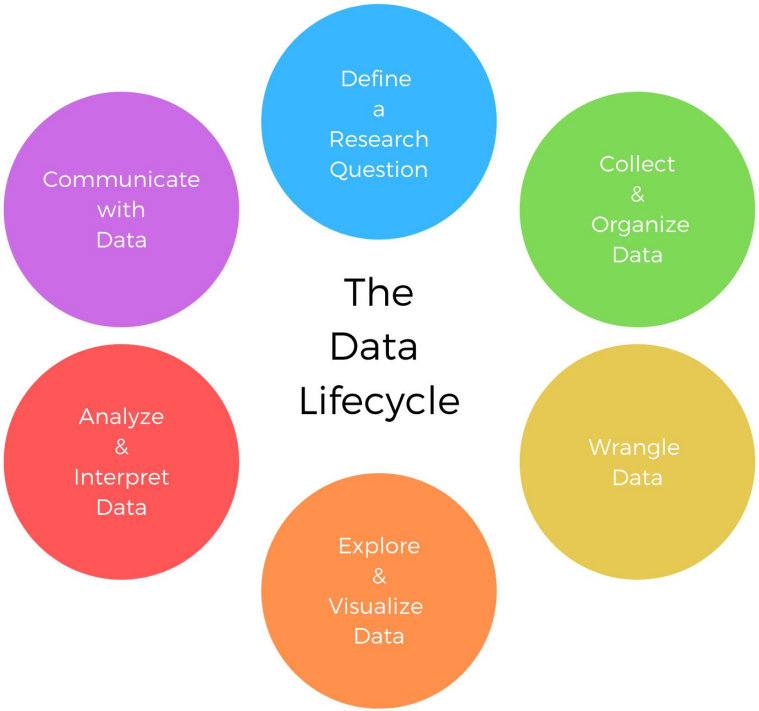
The Data Lifecycle - Analyze & Interpret Data



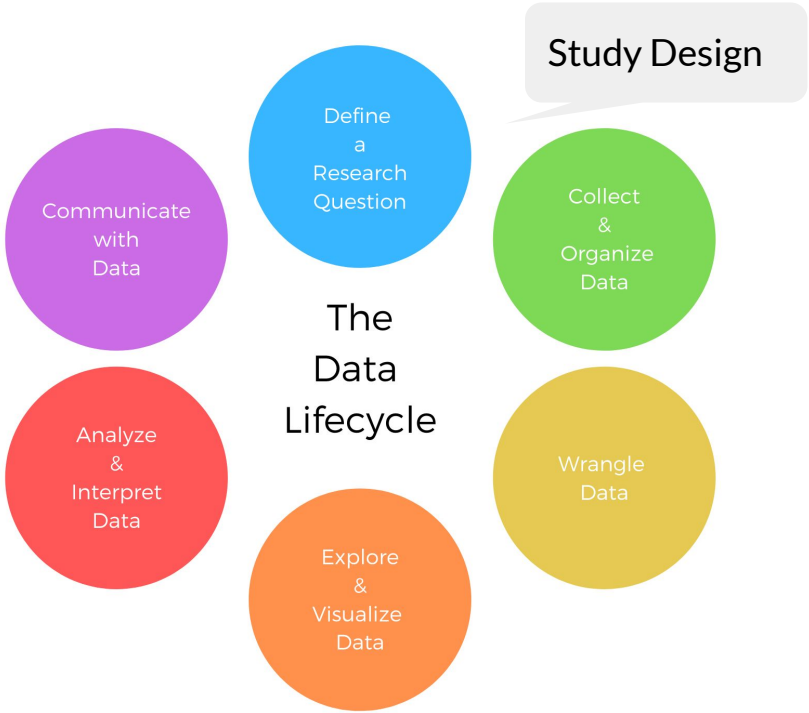
The Data Lifecycle - Communicate with Data



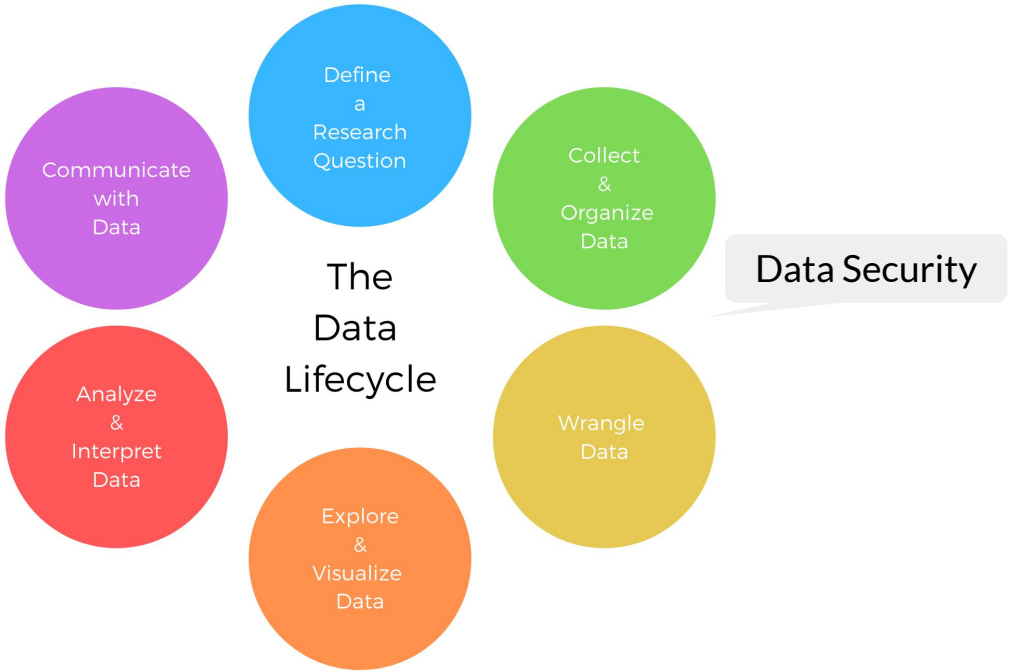
Hidden Steps in the Data Lifecycle



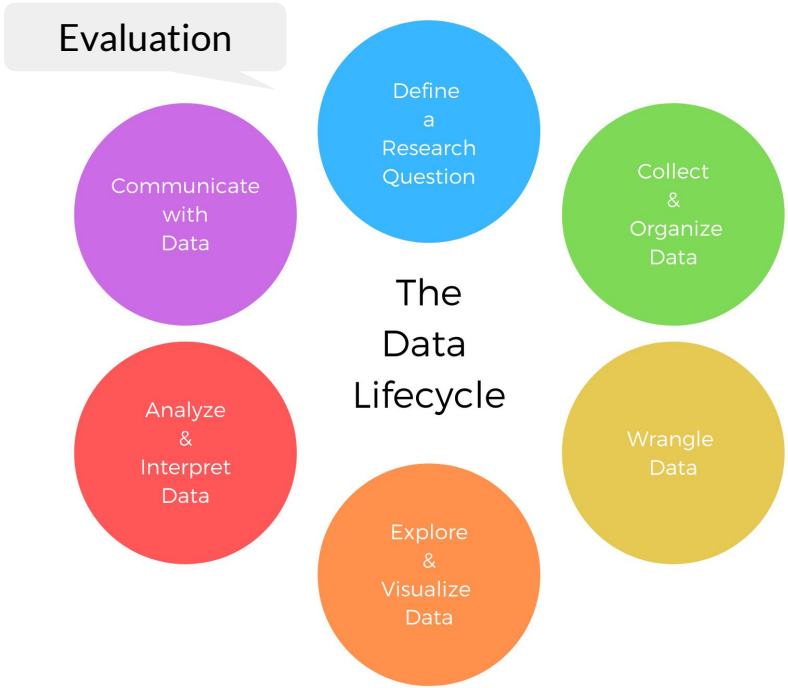
Hidden Steps in the Data Lifecycle



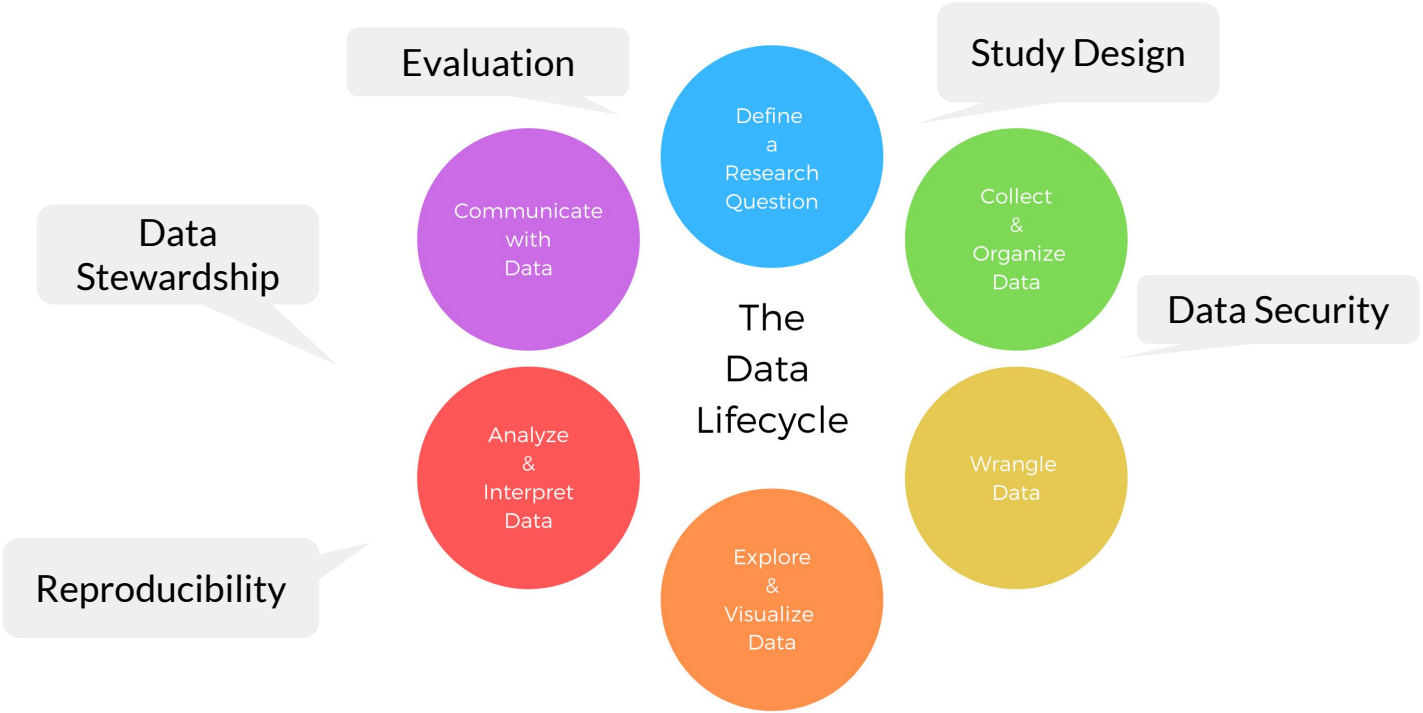
Hidden Steps in the Data Lifecycle



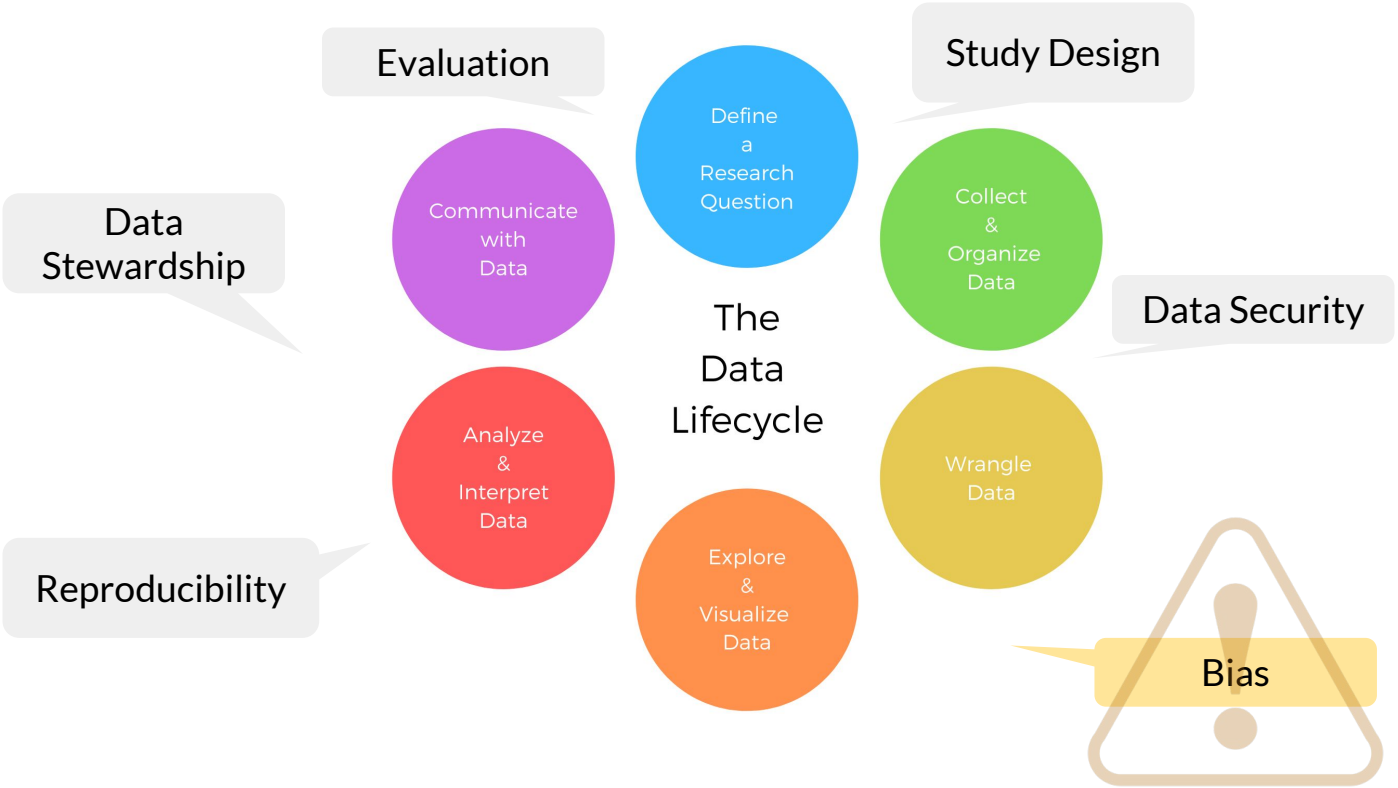
Hidden Steps in the Data Lifecycle

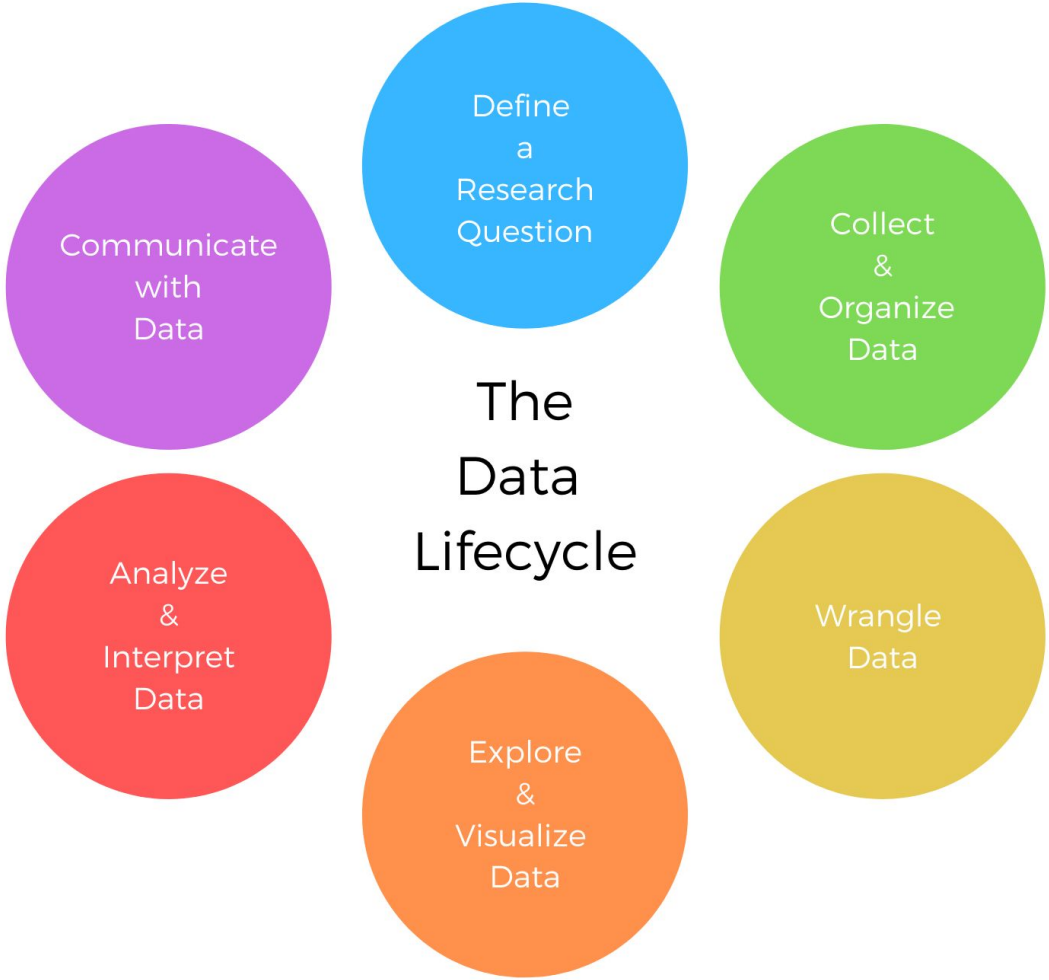


Hidden Steps in the Data Lifecycle

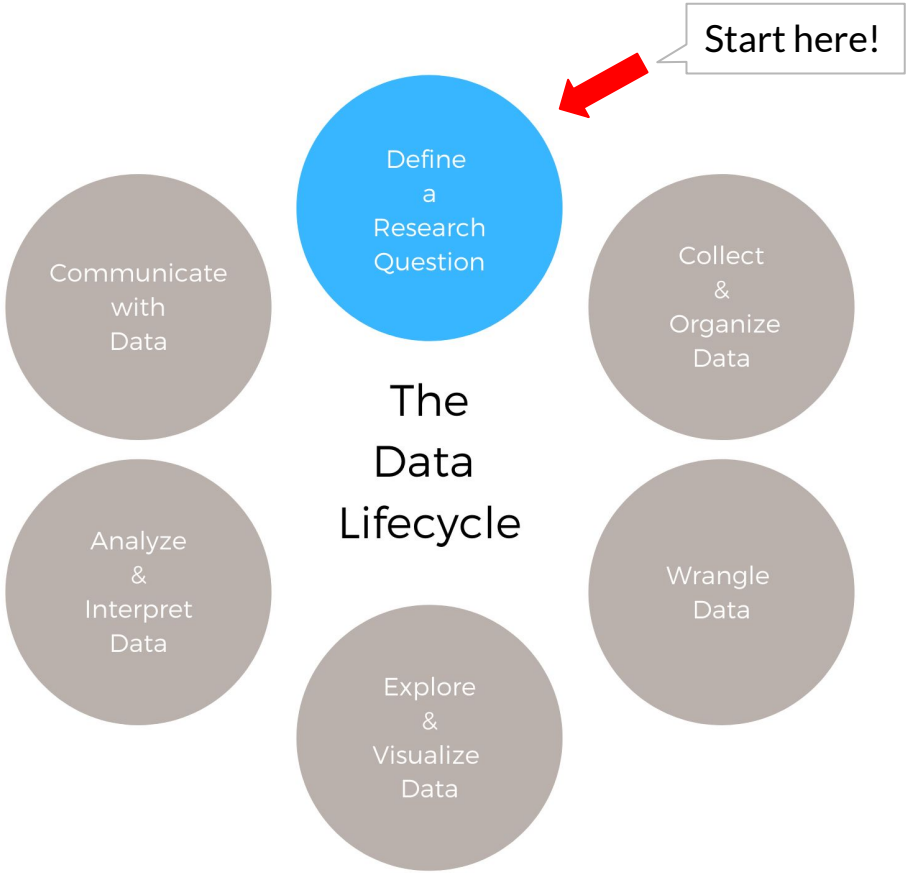


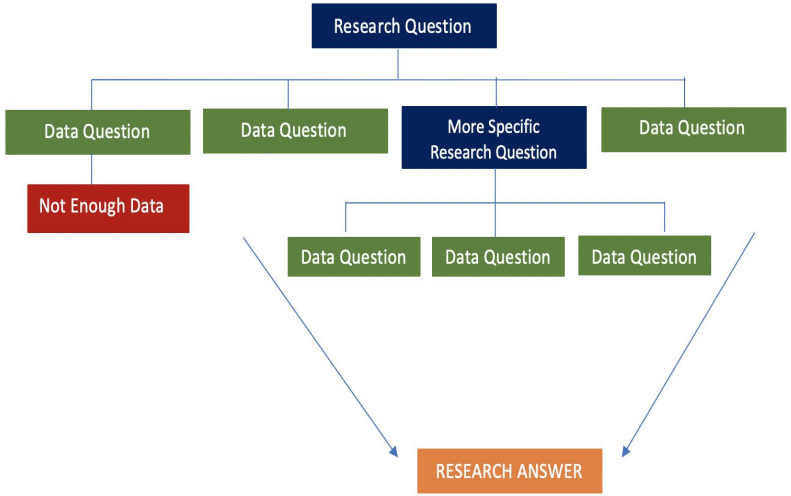
Hidden Steps in the Data Lifecycle




Define a Research Question
Collect & Organize Data
Wrangle Data
Explore & Visualize Data
Analyze & Interpret Data
Communicate with Data
The Data Lifecycle


What makes a good Research Question?


Start here!
Define a Research Question
Collect & Organize Data
Wrangle Data
Explore & Visualize Data
Analyze & Interpret Data
Communicate with Data
The Data Lifecycle


Research Question
Data Question
Data Question
More Specific Research Question
Data Question
Not Enough Data
Data Question
Data Question
Data Question
RESEARCH ANSWER

A good research question is...

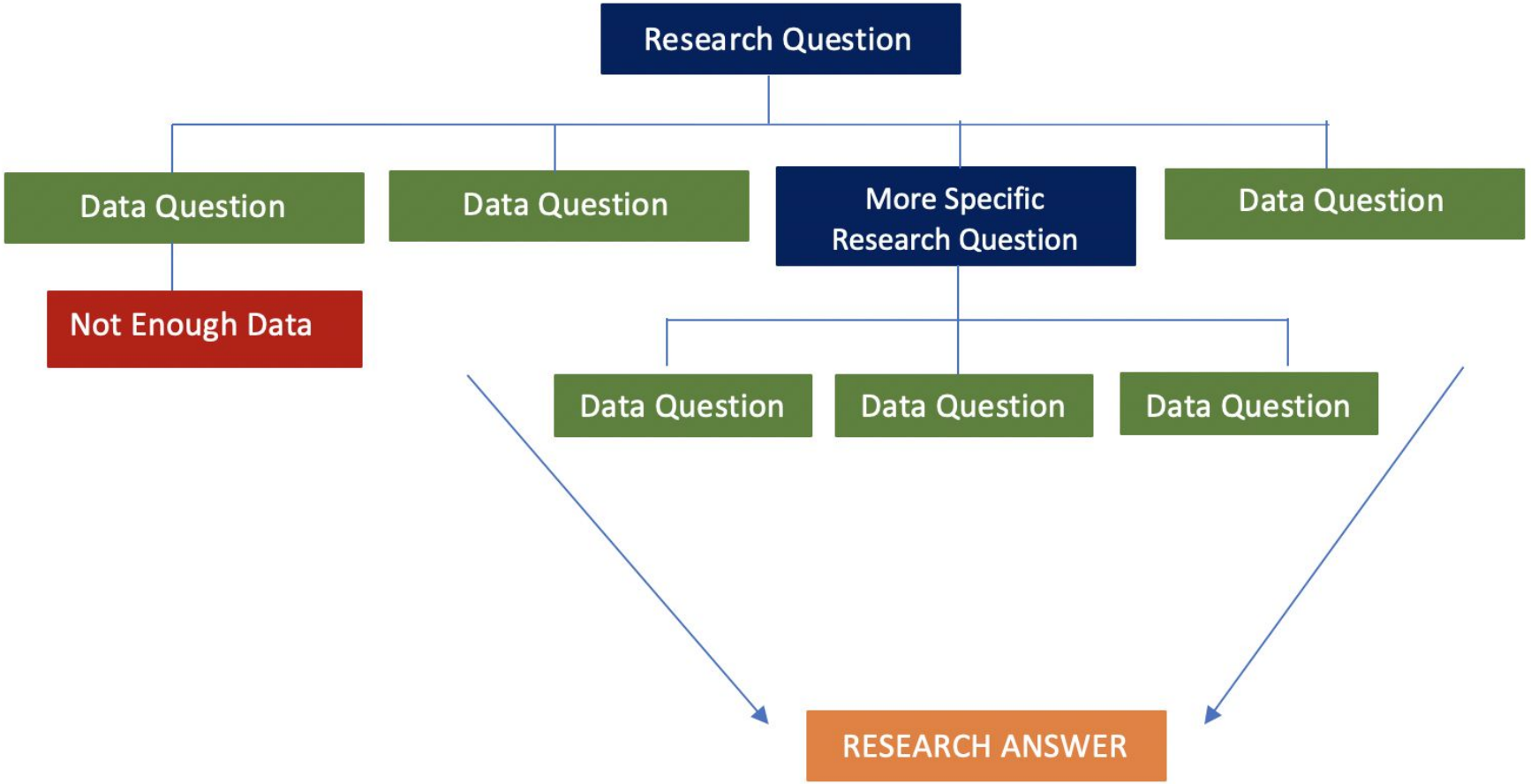
An actual question

Open-ended

Grounded in theory, prior research, or practice

Not too simple, not too complex

Figure adapted from: twitter.com/BecomingDataSci



Reference: https://dougenterprises.com/data-science/encouraging-data-analytics-questions/

Evaluating your research question

Does the question strike the **right balance** between simple and complex? ✔

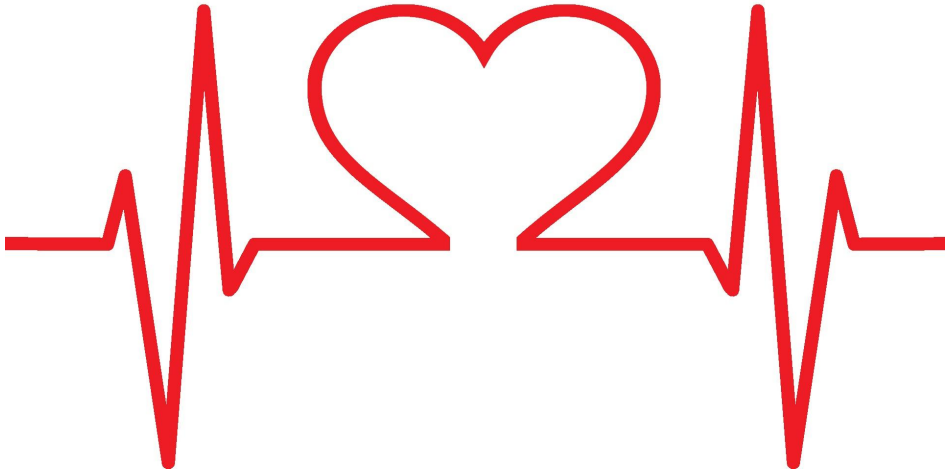
Will it **take more than** a simple data analysis or a single visualization to answer the question? ✔

Do we have **access** to the right amount and the right types of data to sufficiently answer the question? ✔

Will it be clear when we've **sufficiently answered** the question? ✔

Reference: Ratan, Anand, and Ratan (2019), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6322175/

Example: Life expectancy around the world



Dataset: Global Life Expectancy in 2015

	Country	BMI	Population	Life_expectancy_female	Life_expectancy	CHE GDP(%)	GDP
1	Afghanistan	22.9	34413603	64.7	63.2	10.1	1.990711e+
2	Albania	26.7	2880703	78.2	76.1	NA	1.138693e+
3	Algeria	25.5	39728025	77.2	76.2	7.0	1.659780e+
4	Andorra	27.0	78011	NA	NA	10.3	2.811489e+(
5	Angola	22.9	27884381	64.5	62.2	2.6	1.161940e+
6	Antigua and Barbuda	26.6	93566	77.4	75.0	4.7	1.336693e+(
7	Argentina	27.6	43131966	80.1	76.8	8.8	5.947490e+
8	Armenia	26.5	2925553	77.9	74.6	10.1	1.055334e+
9	Australia	27.4	23815995	84.5	82.6	9.3	1.351690e+
10	Austria	26.0	8642699	83.8	81.4	10.4	3.818180e+

Information on
194 countries

194 rows

9 columns

Dataset: Global Life Expectancy in 2015

	Country	BMI	Population	Life_expectancy_female	Life_expectancy	CHE GDP(%)	GDP
1	Afghanistan	22.9	34413603	64.7	63.2	10.1	1.990711e+
2	Albania	26.7	2880703	78.2	76.1	NA	1.138693e+
3	Algeria	25.5	39728025	77.2	76.2	7.0	1.659780e+
4	Andorra	27.0	78011	NA	NA	10.3	2.811489e+(
5	Angola	22.9	27884381	64.5	62.2	2.6	1.161940e+
6	Antigua and Barbuda	26.6	93566	77.4	75.0	4.7	1.336693e+(
7	Argentina	27.6	43131966	80.1	76.8	8.8	5.947490e+
8	Armenia	26.5	2925553	77.9	74.6	10.1	1.055334e+
9	Australia	27.4	23815995	84.5	82.6	9.3	1.351690e+
10	Austria	26.0	8642699	83.8	81.4	10.4	3.818180e+
11	Azerbaijan						

Defining a Research Question

- Which **factors** are associated with longer life expectancy?
- How do the GDP and population of a country **relate** to its life expectancy?

Defining a Research Question

- Which **factors** are associated with longer life expectancy? ✅

- How do the GDP and population of a country relate to its life expectancy?

Evaluating a Research Question

	Not too simple nor too complex?	***Will it take more than a simple analysis?***	***Do we have enough and the right kind of data?***	***What would be a sufficient answer?***
Which factors are associated with longer life expectancy?				

Evaluating a Research Question

	Not too simple nor too complex?	*Will it take more than a simple analysis?*	*Do we have enough and the right kind of data?*	*What would be a sufficient answer?*
Which factors are associated with longer life expectancy?	✔			

Evaluating a Research Question

	Not too simple nor too complex?	*Will it take more than a simple analysis?*	*Do we have enough and the right kind of data?*	*What would be a sufficient answer?*
Which factors are associated with longer life expectancy?	✔	✔		

Evaluating a Research Question

	Not too simple nor too complex?	*Will it take more than a simple analysis?*	*Do we have enough and the right kind of data?*	*What would be a sufficient answer?*
Which factors are associated with longer life expectancy?	✔	✔	✔	

Evaluating a Research Question

	Not too simple nor too complex?	*Will it take more than a simple analysis?*	*Do we have enough and the right kind of data?*	*What would be a sufficient answer?*
Which factors are associated with longer life expectancy?	✔	✔	✔	✔

Take Home Messages

- A good research question is **open ended**, and not easy to answer with a quick data analysis
- **Always revisit your research question**
- It's **okay to change/refine** your research question, **within reason**

Collect & Organize Data

The Data Lifecycle

- Define a Research Question
- Collect & Organize Data
- Wrangle Data
- Explore & Visualize Data
- Analyze & Interpret Data
- Communicate with Data

Looking Ahead

Image Credits:

Minimal Pencils on Yellow Background Photo by Joanna Kosinska on Unsplash
White Paint Photo by Joanna Kosinska on Unsplash
Autonomous Subway Photo by Mathew Schwartz on Unsplash
Hello by Tumisu on Pixaby
Cooking by Maria Elkind and licensed under CC BY-SA 2.0
Baking Ingredients by Marcel Kessler on Pixaby
Chocolate Cake Recipe with Ingredients by Marco Verch, and licensed under CC BY 2.0
Chocolate Cream Cake by Marco Verch , licensed under CC BY 2.0
Green Field Photo by Josh Silver on Unsplash
Laptop 1 Photo by Pixnio and under CC0
Laptop 2 Photo by Jesus Kiteque on Unsplash
Laptop 3 Photo by NordWood Themes on Unsplash
Laptop 4 Photo by Mika Baumeister on Unsplash
Laptop 5 Photo by Fernando Venturium on Unsplash
White Paper Photo by Brandi Redd on Unsplash
Data Lifecycle Figure adapted from: twitter.com/BecomingDataSci
Research Question Tree adapted from: https://dougenterprises.com/data-science/encouraging-data-analytics-questions/
White Ribbed Building Photo by Luca Bravo on Unsplash
Spices Photo by Calcum Lewis on Unsplash
Architectural Photo by Radoslav Bali on Unsplash