



ពិភាក្សាលេខាងក្រោម

SUNRISE INSTITUTE

MACHINE LEARNING

EMBARKING ON A JOURNEY
INTO DATA SCIENCE

YA MANON



You can have data without information but you
cannot have information without data.

-Daniel Keys Maran

Question

“In your experience so far, from the courses you’ve completed (Fundamental Data Science, Programming for Data Science, Statistics for Data Science, etc.), what are the key concepts or skills you’ve found most useful in your work? Could you give examples of how you applied them?”

- **Programming**
- **Data Analysis and Visualization**
- **Statistics Exploration**

Library We Used

General Purpose
Programming



Data Manipulation
& Analysis



Data Visualization



Machine Learning



Library We Will Be Using

General Purpose
Programming



Data Manipulation
& Analysis



Data Visualization



Machine Learning



Library We Will Be Using

General Purpose Programming



Data Manipulation & Analysis



Pandas helps us structure our data into dataset formats similar to that which you'd see in SQL or Excel. It also provides us with an arsenal of analytical functions that help us manipulate data and calculate the metrics we need to understand our data

Data Visualization



Matplotlib and **Seaborn** can create a wide array of visually appealing, static visualizations

Plotly can be used to create interactive visualizations and dynamic dashboards

Machine Learning



Scikit learn is among the most popular tools for building and testing machine learning models

Statsmodels provides a suite of tools for model building and statistical analysis

TensorFlow is the industry standard for developing deep learning models

INTRO TO DATA SCIENCE



In this section we'll **introduce the field of data science**, discuss how it compares to other data fields, and walk through each phase of the data science workflow

TOPICS WE'LL COVER:

What is Data Science

What is Artificial Intelligence AI?

Data Science vs. AI

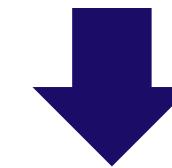
Data Science vs. Data Analytics

GOALS FOR THIS SECTION:

- Compare data science and machine learning with other common data analytics fields
- Introduce supervised and unsupervised learning, and examples of each technique
- Review the machine learning landscape and commonly used algorithms
- Discuss essential skills, and review each phase of the data science workflow

What is Data Science?

What's Data Science?



- 1 **The study of data to extract meaningful insights for business.**
- 2 **The process of creating physical models to simulate real-world events without using any data.**
- 3 **A branch of engineering focused solely on designing hardware components for data storage systems.**
- 4 **The study of astronomical phenomena using telescopes and computational tools.**

What is Data?

Data is a collection of information gathered by observations, measurements, research or analysis.

Data Types

Structures data

Dimension (rows, columns),
csv, excel , ..

Example: phone number, zip codes,
identity number, gender,..

Unstructures data

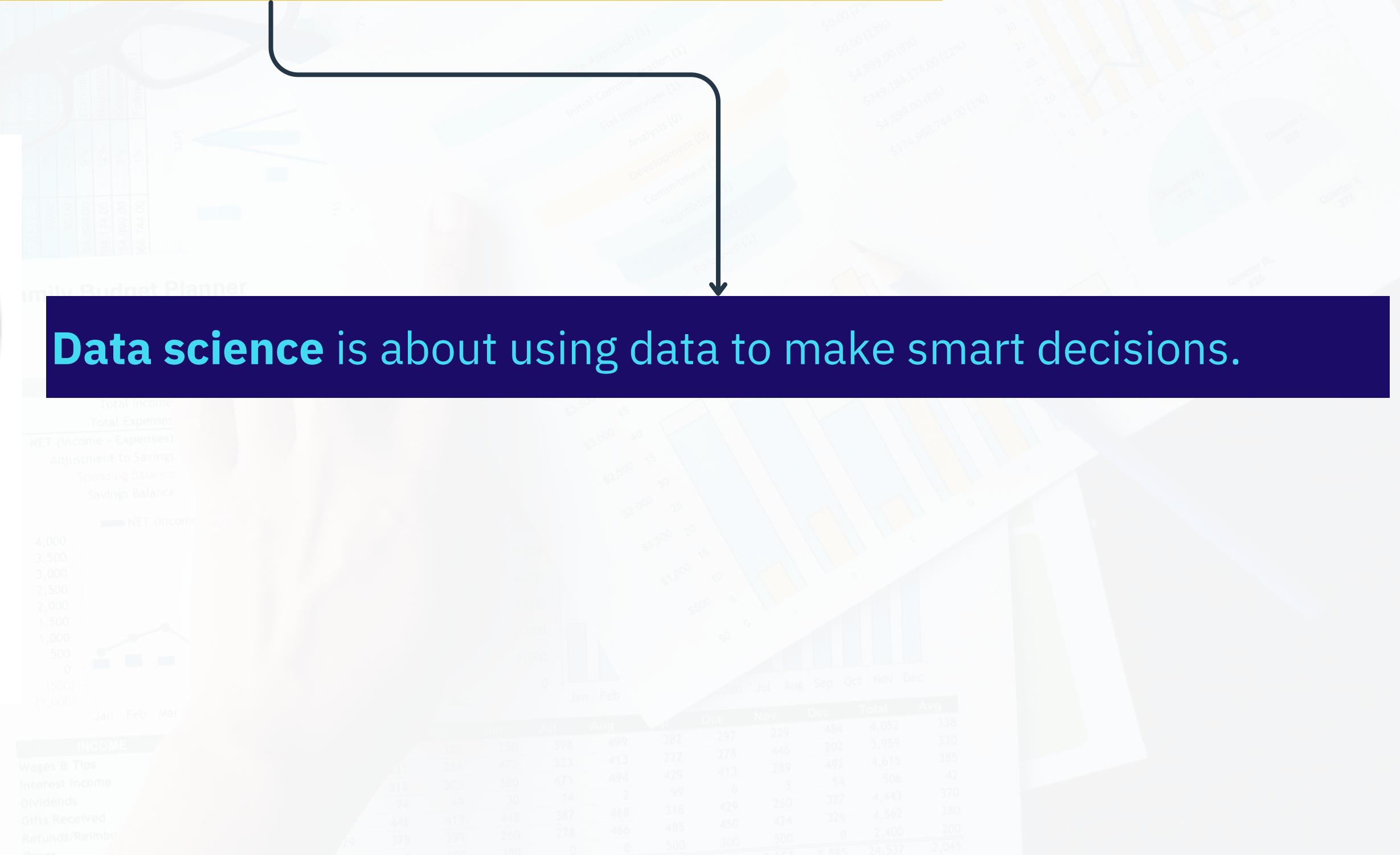
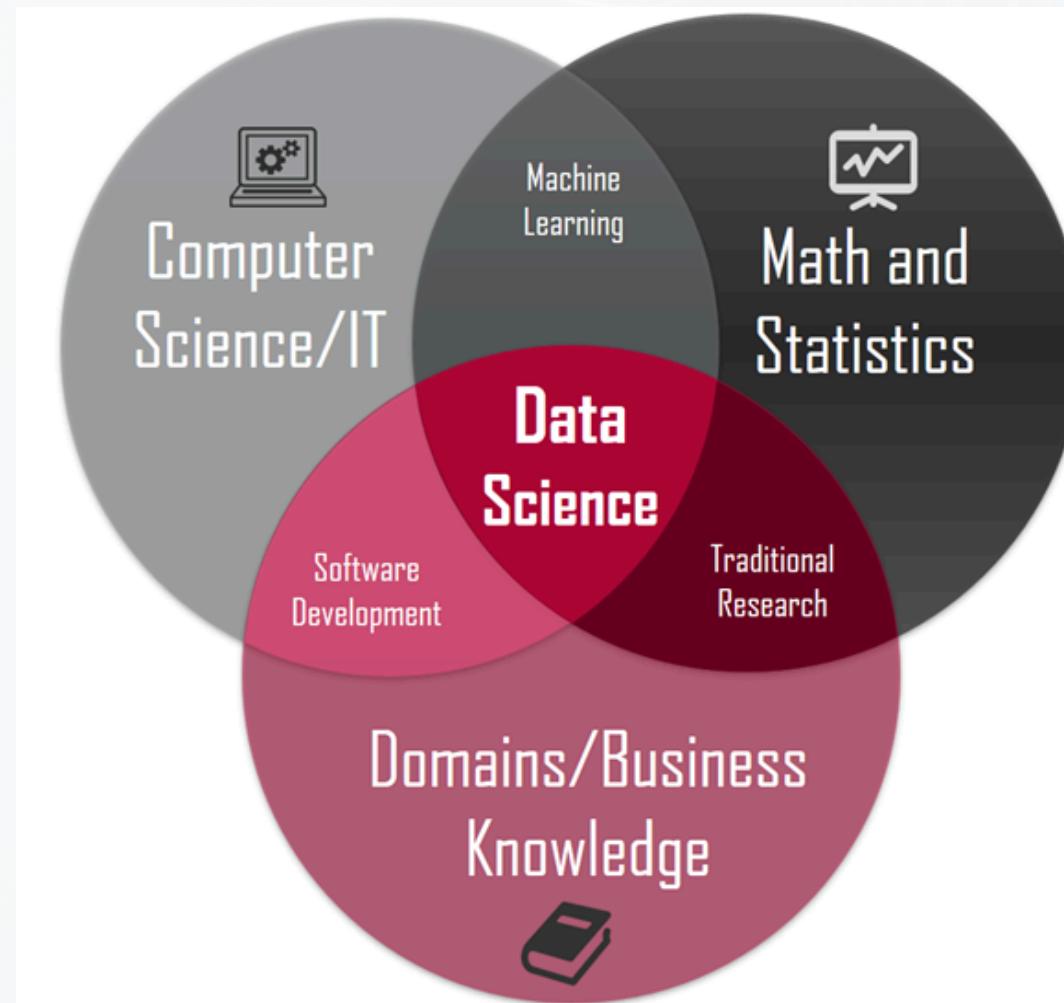
Everything else (no rows,
columns)...

Example: emails, social media post, video,
sensor, data...

Data Science

1

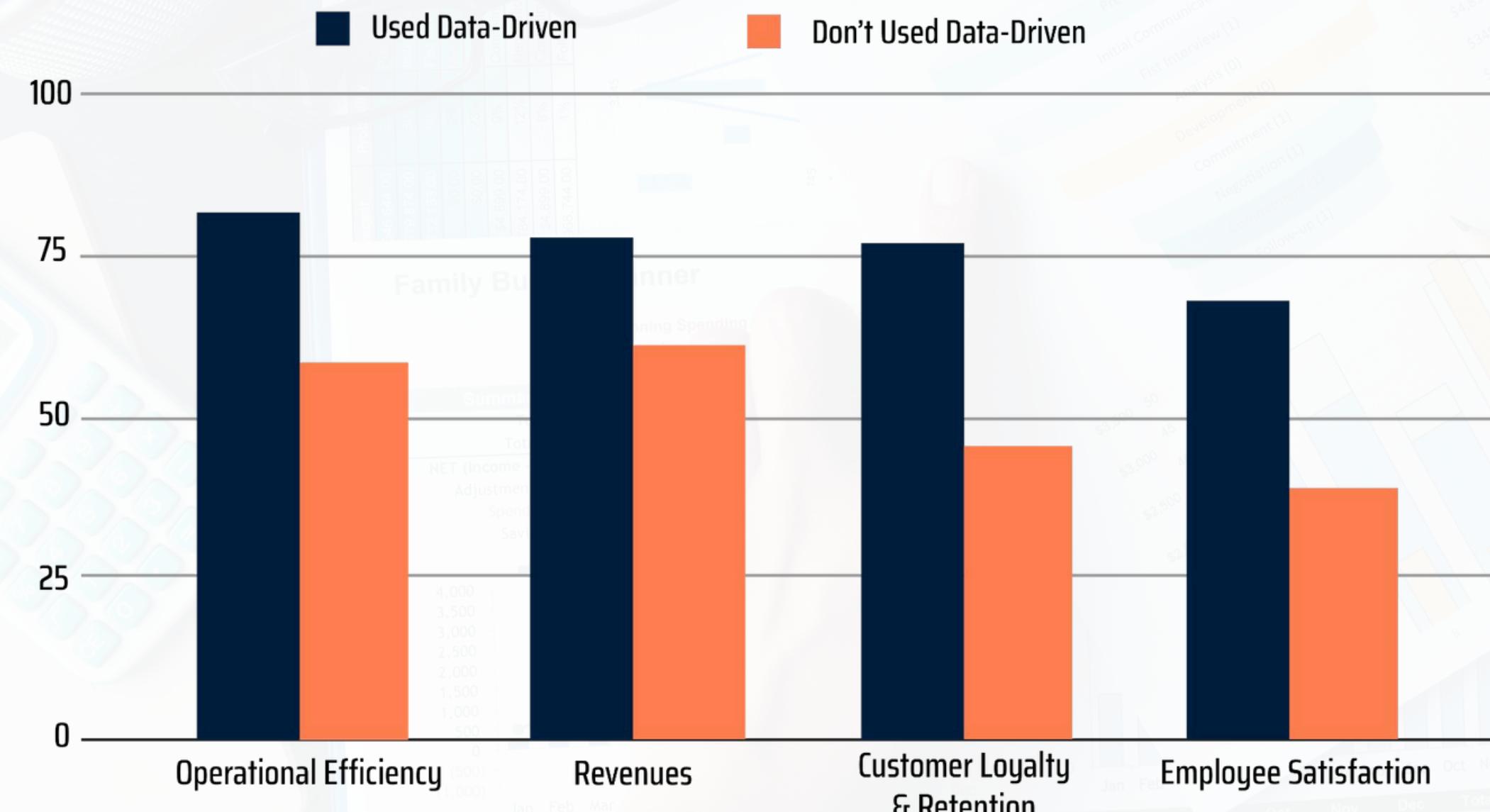
The study of data to extract meaningful insights for business.



Data Science

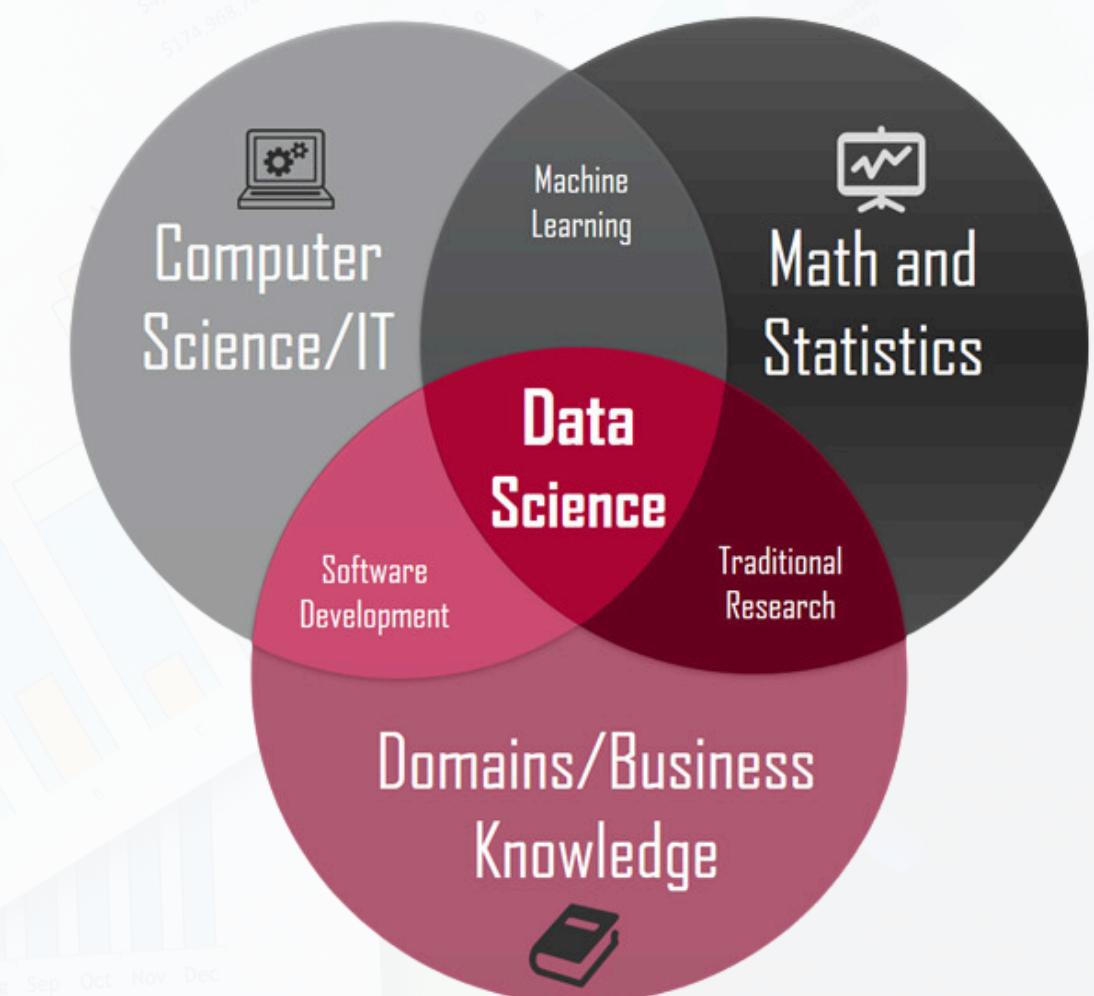
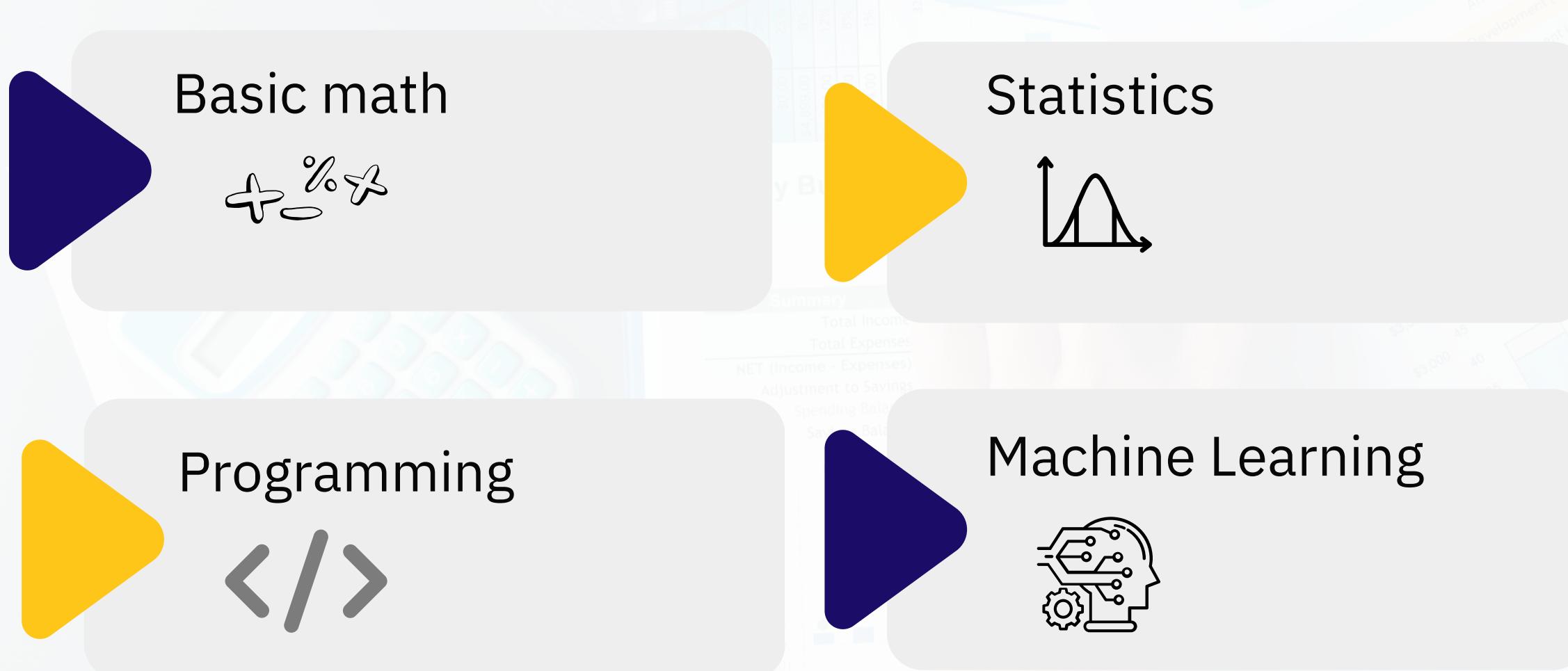
The global significance of data utilization

Business Success Metrics: 2021 Harvard Business Study (%)

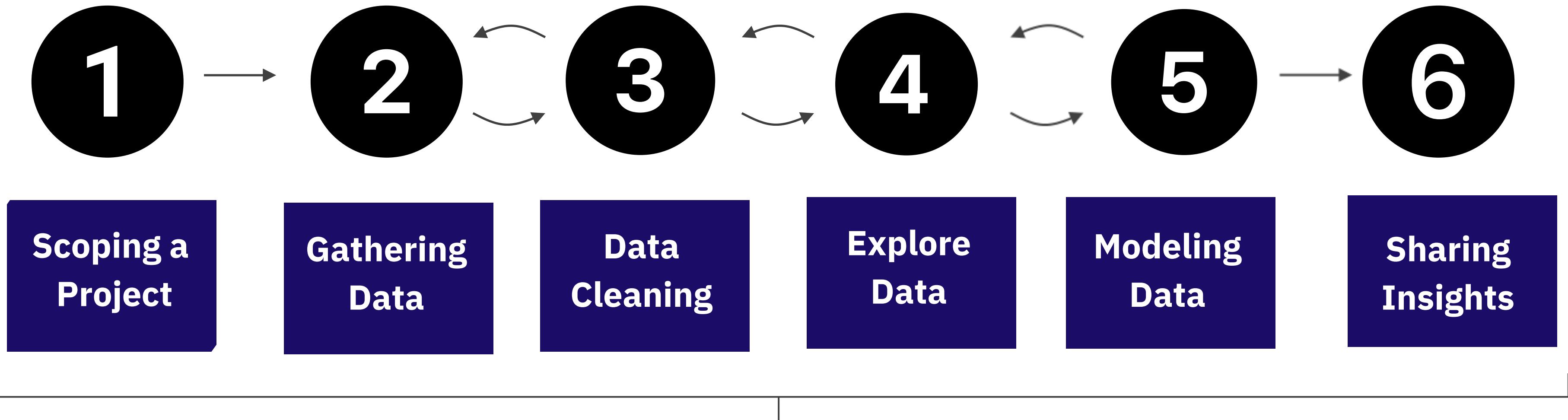


Data Science

Data science combines **math** and **statistics**, specialized **programming**, **advanced analytics**, **a** and **machine learning** with specific subject matter expertise to uncover actionable insights hidden in an organization's data.

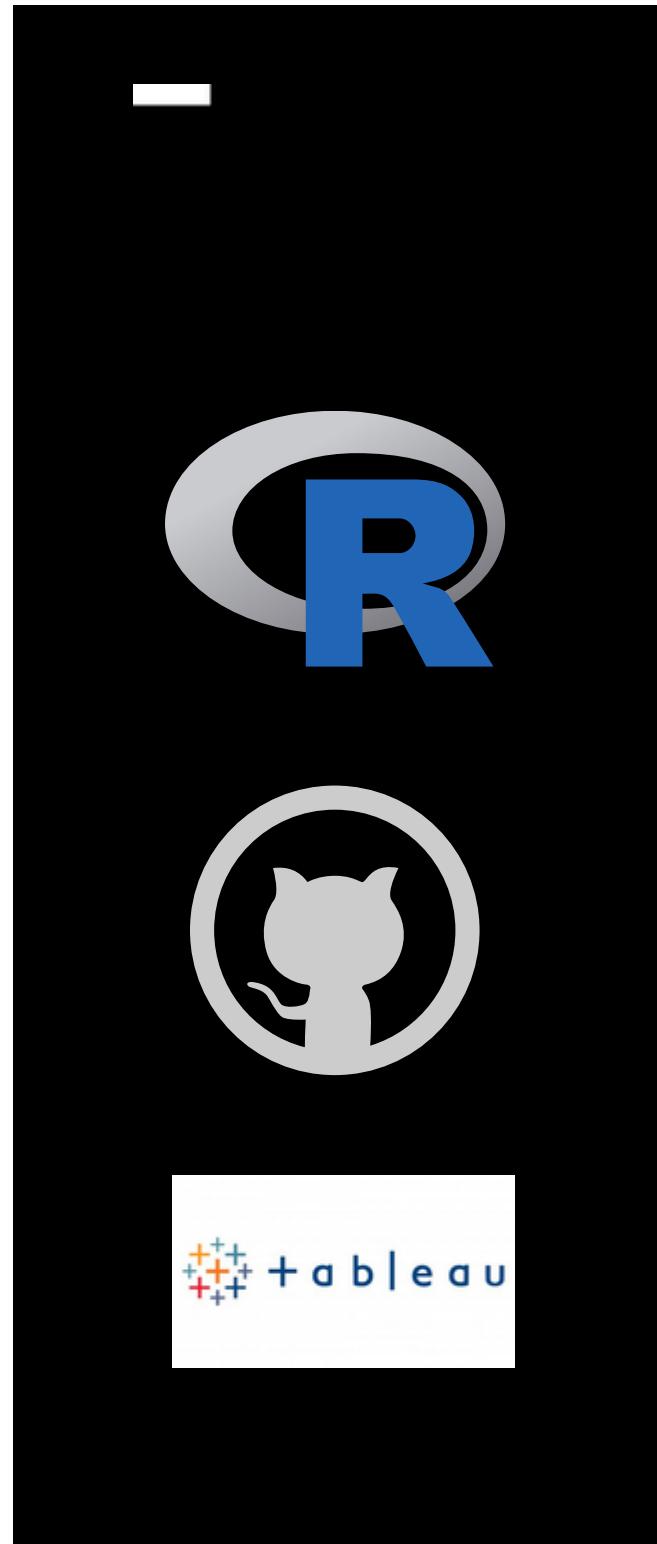


Data Science Workflow



This is not a linear process! You'll likely go back to further gather, clean and explore your data

Data Science tools



ផ្សេង python ដែចជា
បកពន្លាន



Why Does Programming Play Such a Big Role?

-  Data Manipulation and Cleaning
-  Statistical and Mathematical Modeling
-  Building Machine Learning Models
-  Data Visualization
-  Scalability and Automation
-  Collaboration and Reproducibility

Why Data Science

Data science is a vital part of every business

You can minor in data science

Job opportunities in all kinds of businesses

Work with diverse teams

Be a part of innovation and progress

Be part of a growing community

Data science jobs are in high demand

Data scientists earn good money

Job security in the field of data science

Use the power of data for social change

Application of Machine Learning

Social Media Features

Sentiment Analysis

Image Recognition

Product Recommendations

NLP Detection

Language Translation

Customer Churn