NASA SPACE APPS Cairo 2022

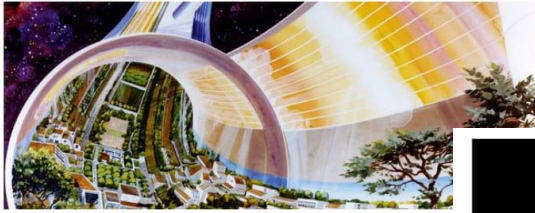## Data Science Workshop
By: Nada Makram

Outlines

- What is data science?
- Forms of data
- Data Science project lifecycle
- A closer look at NASA data
- Tips and tricks to get a data-driven MVP
- Technical Demo

Challenges

THE ART IN OUR WORLD

INTERMEDIATE ADVANCED

DIVERSITY & EQUITY SPACE EXPLORA

ARTS

MAKE

AR

ME

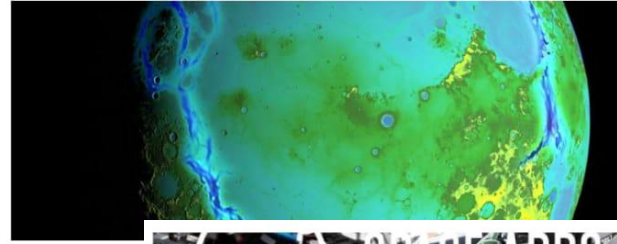## TAKE FLIGHT: MAKING THE MOST OF NASA'S AIRBORNE DATA

CLIMATE GAMES INTERMEDIATE

BEGINNER ADVANCED

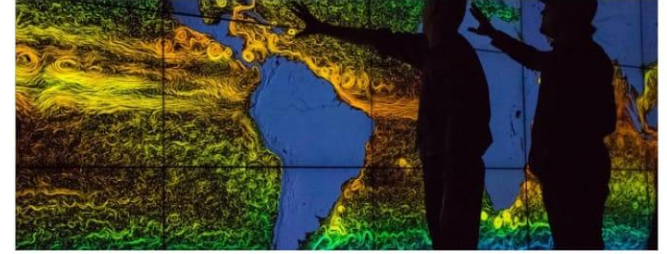## CALLING ALL RADIO ENTHUSIASTS!

ADVANCED EARTH
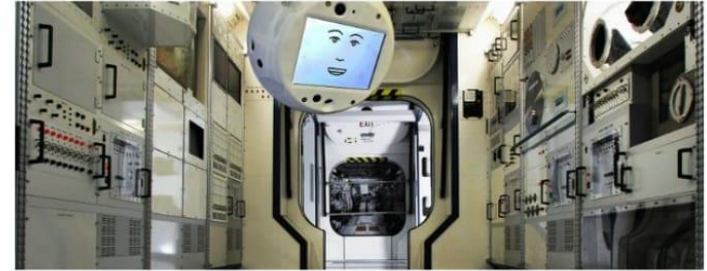
SPACE EXPLORATION SUN SOF

## CAN AI PRESERVE OUR SCIENCE LEGACY?

SOFTWARE ADVANCED

What's the data science?

Data science is the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data.

- Machine learning
- Deep learning
- Models & algorithms
- Data science toolkits

Statistics/ Math

- Applied statistics
- Bayesian inference
- Logistic & linear regression
- Differential & integral calculus
- Experiment design

Data Science

Computer Science

Domain Expertise
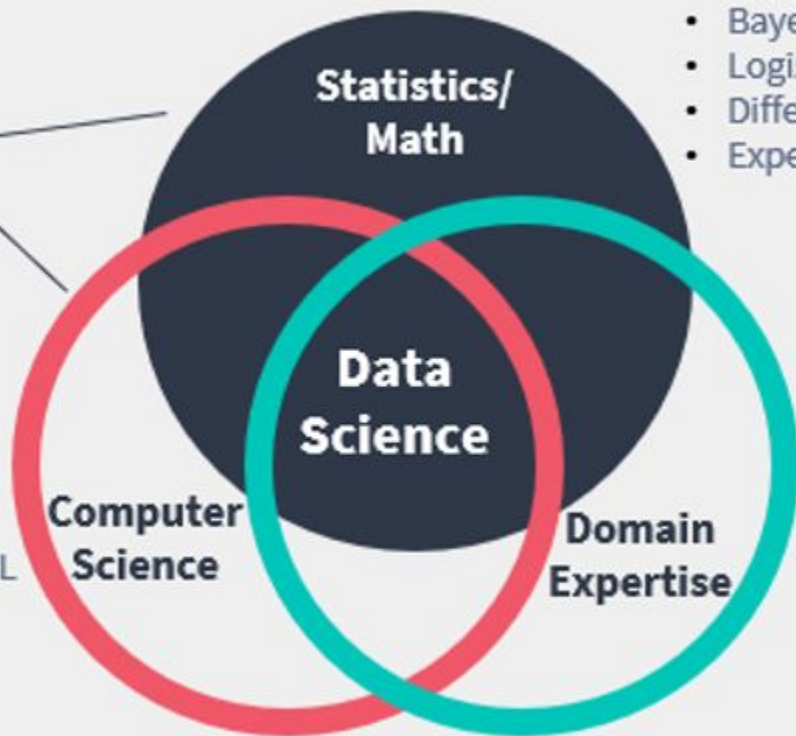
- Scripting languages
- Databases: SQL & NoSQL
- Data visualization
- Computational power
- Algorithm design

- Business acumen
- Business strategy
- Business Intelligence
- Communication skills

What are the applications of data science?

- **Healthcare:** identify and predict disease, and personalize healthcare recommendations.
- **Transportation:** optimize shipping routes in real-time.
- **Sports:** evaluate athletes' performance, and accurate player profiling
- **E-commerce:** digital ad placement and personalized marketing.
- **Social media:** people you may know and ads recommendations, tagging.
- **Fintech:** detect fraud and create a credit score.

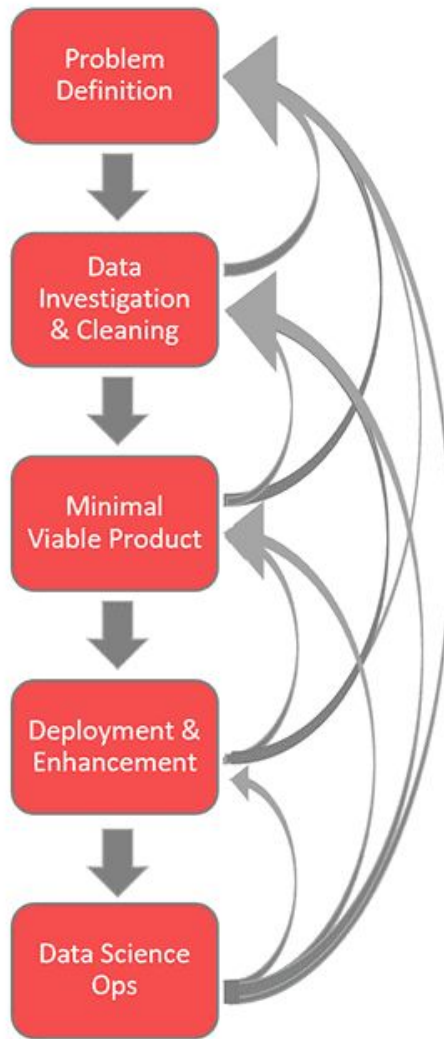# What are the applications of data science in **Space**?

- Health monitoring of spacecraft
- Autonomous navigation
- Intelligent object detection
- Environmental solutions (e.g. climate change prediction)
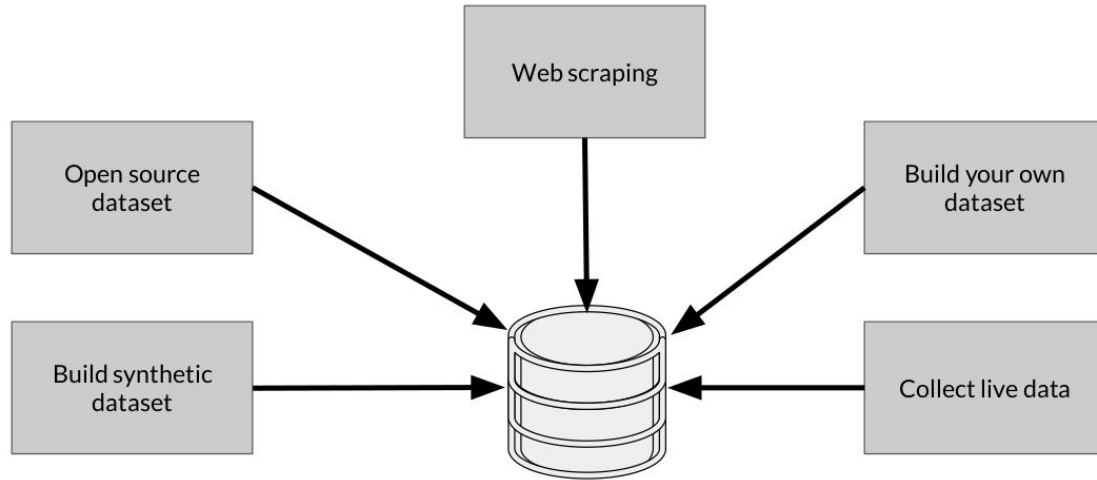- Forest fire detection

Forms of data

- Structured
- Semi-Structured
- Unstructured
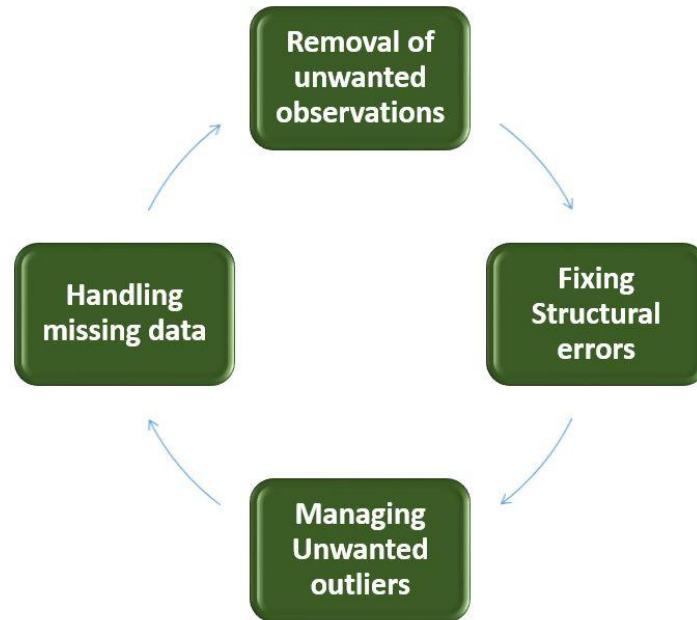
DS project lifecycle

Data Collection

NASA Data

some resources:

[NASA's Open Data Portal](#) (API)

[NASA Power](#)

Data Cleaning

Removal of unwanted observations

Fixing Structural errors

Managing Unwanted outliers

Handling missing data

$$f\left(\text{🗑} \right) = \text{🗑}$$

Feature Engineering

Feature Engineering

- Feature Selection
- Feature Transformation
- Feature Creation (Encoding, Binning)

Technical Demo

[Google Colab Notebook](https://bit.ly/3xuqhx0)

https://bit.ly/3xuqhx0

Resources

Vaex

Flask API, Fast API

Plotly

Dash

Kaggle (datasets, tutorials)

For learning: - Practical Statistics for Data Scientists 50+ Essential Concepts Using R and Python

- Andrew Ng Coursera and Hesham Assem Youtube

Thank you.