Understanding the Goals and Benefits of Exploratory Data Analysis (EDA)

Guillermo Fernández
DATA SCIENTIST

@guillermo_ai

Summary

Define what EDA is

Understand EDA goals and benefits

Give a quick real-life example overview

Know and understand the steps involved in the EDA workflow

What is EDA?

Exploratory Data Analysis

Is an approach, not a set of techniques, but an attitude or philosophy about how a data analysis should be carried out.

[...] is actively incisive rather than passively descriptive [...]

[...] a willingness to look for those things that we believe are not there, as well as those we believe to be there [...]

Sources:
Exploratory and Multivariate Data Analysis. Michael Jambu.
Exploratory data analysis as part of a larger whole. John Wilder Tukey.

EDA Goals and Benefits

Understand and improve your knowledge

Draw valid conclusions

Aid in decision making and planning

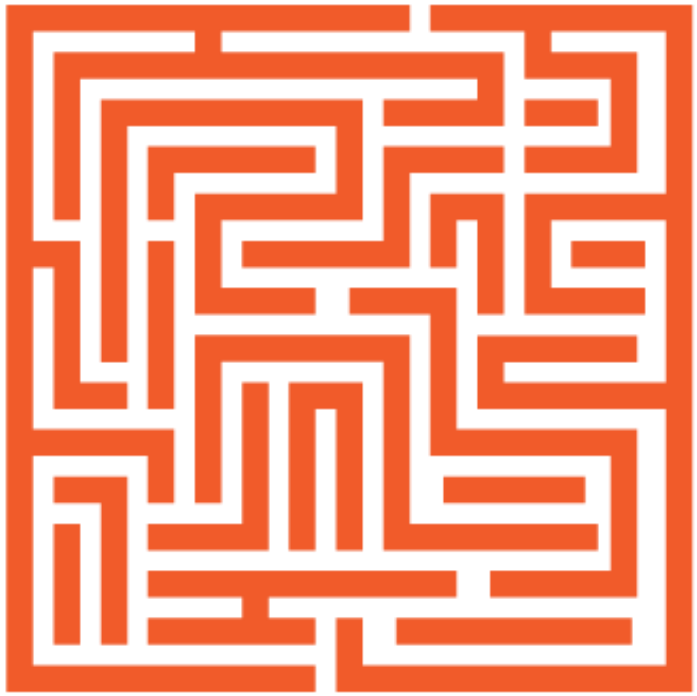
Help in causal analysis

To generate and confirm hypothesis.

To build intuition and gain insight about data.

Where to begin?

FORMULATE A HYPOTHESIS !!!



"An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem."

John Wilder Tukey

Enigma during WWII

Goodbye pattern at the end of transmissions

Check assumptions

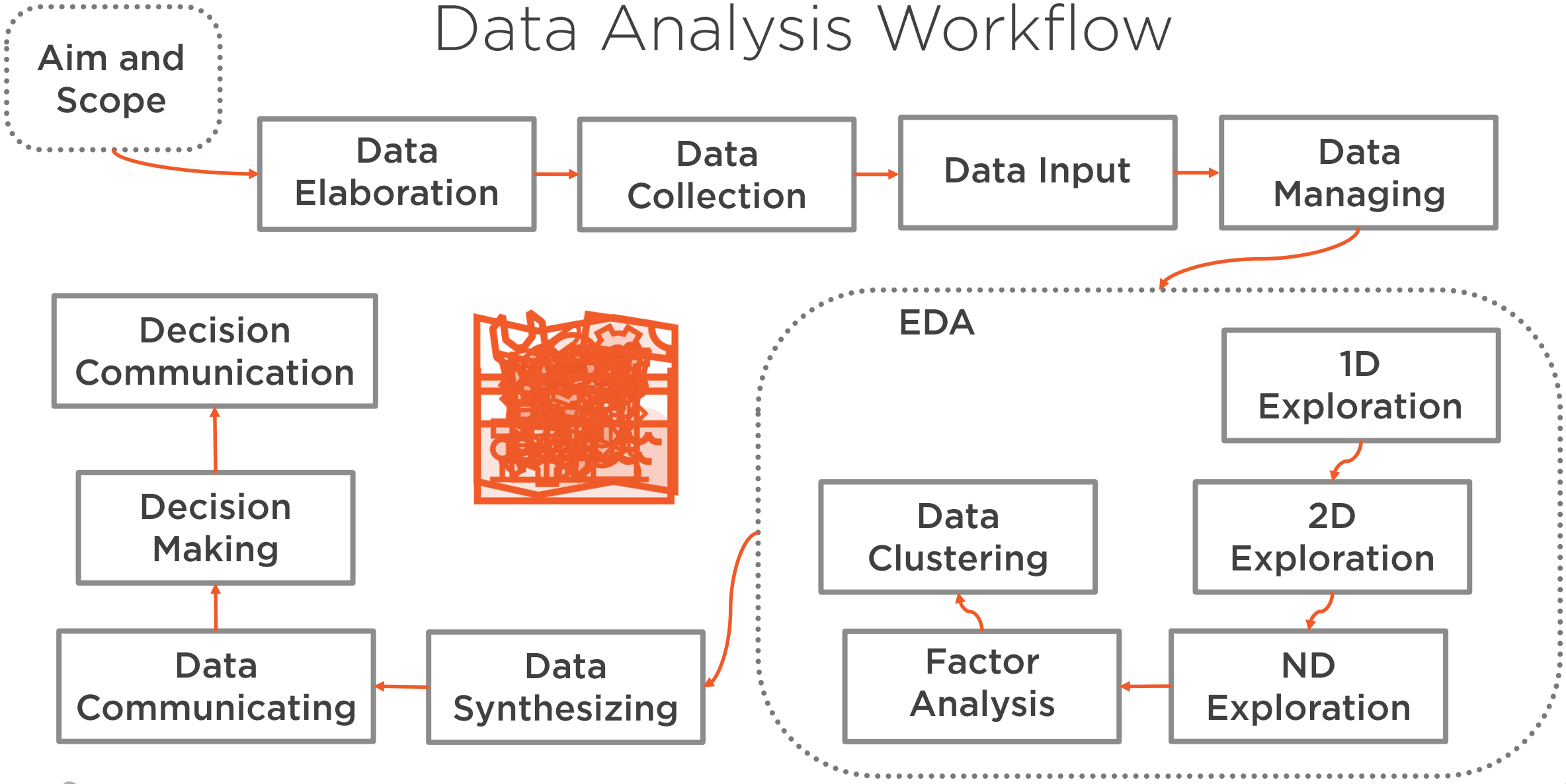
Discover structure in data

Understand and learn

The EDA Workflow

Where does EDA belong?

Data Analysis Workflow



Sources:
Exploratory and Multivariate Data Analysis. Michael Jambu.

Data Analysis Techniques

Exploratory DA

Problem

Data

Analysis

Model

Conclusions

Classical DA

Problem

Data

Model

Analysis

Conclusions

Bayesian DA

Problem

Data

Model

Prior Distribution

Analysis

Conclusions

Summary DA

Passive

Historic

Sources:
www.itl.nist.gov