# DATA ANALYSIS AND DATA SCIENCE- 101

Hugo Estrada S.

# About me:

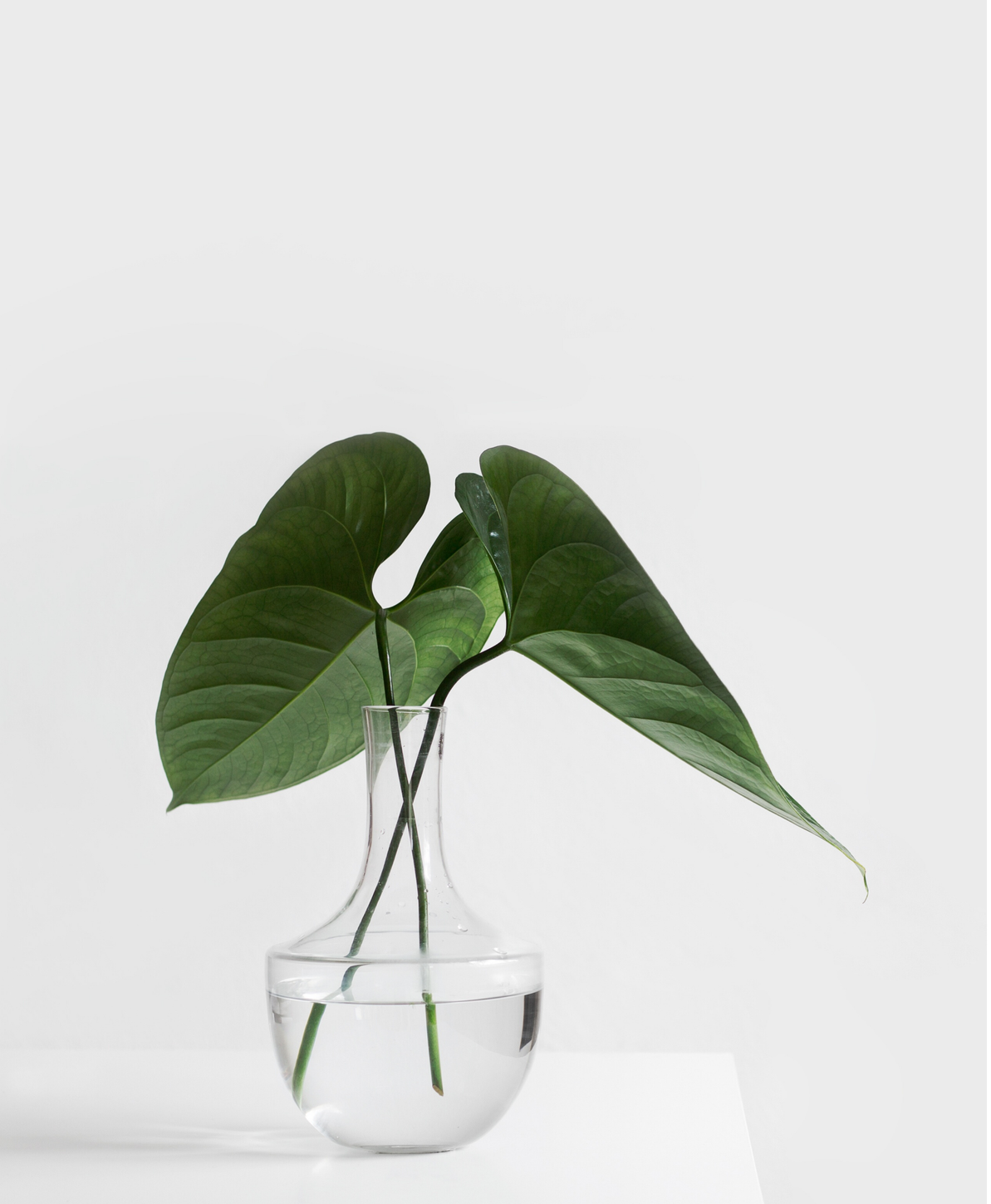
hugo_estradas@outlook.com

github.com/hugoestradas
linkedin.com/in/hugoestradas
dev.to/hugoestradas

IG | TW: @hugoestradas

INTERNATIONAL SCRUM INSTITUTE

Microsoft CERTIFIED
AZURE FUNDAMENTALS



# Summary

- Some theory.
- Datascience tools in Microsoft Azure.
- Azure DataBricks.
- How to setup your first scientific environment with Docker-Compose.
- Datascience basics with Python.
- Solve your first Datascience Requirement.

# What is Data Analysis <?>

Is the process of evaluating data using analytical or statistical tools to discover useful information.

# What is Data Science <?>

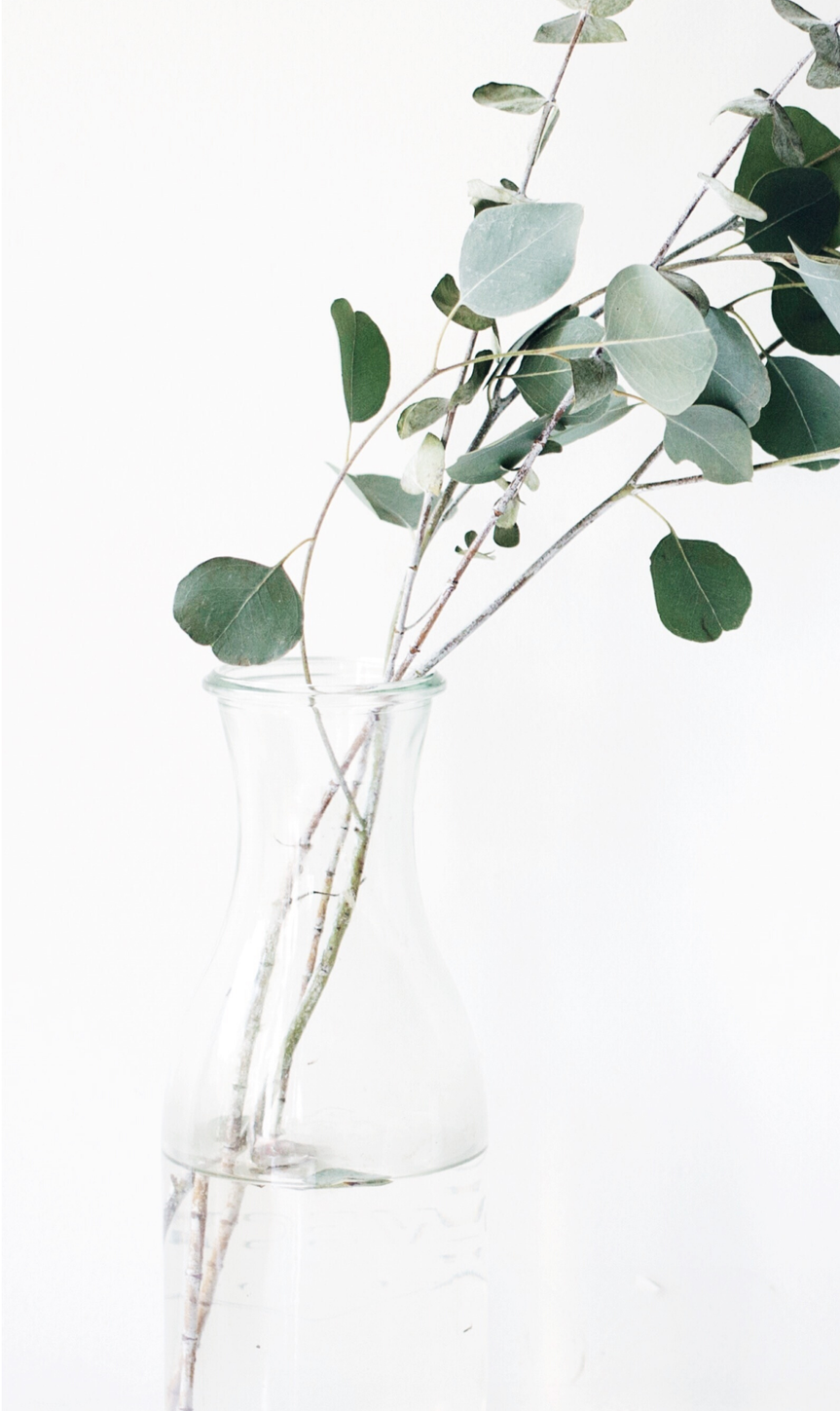
Is the study of data.
It involves developing methods of recording, storing, and analyzing data to effectively extract useful information.
The goal of data science is to gain insights and knowledge from any type of data.

# Data Science VS. Data Analytics

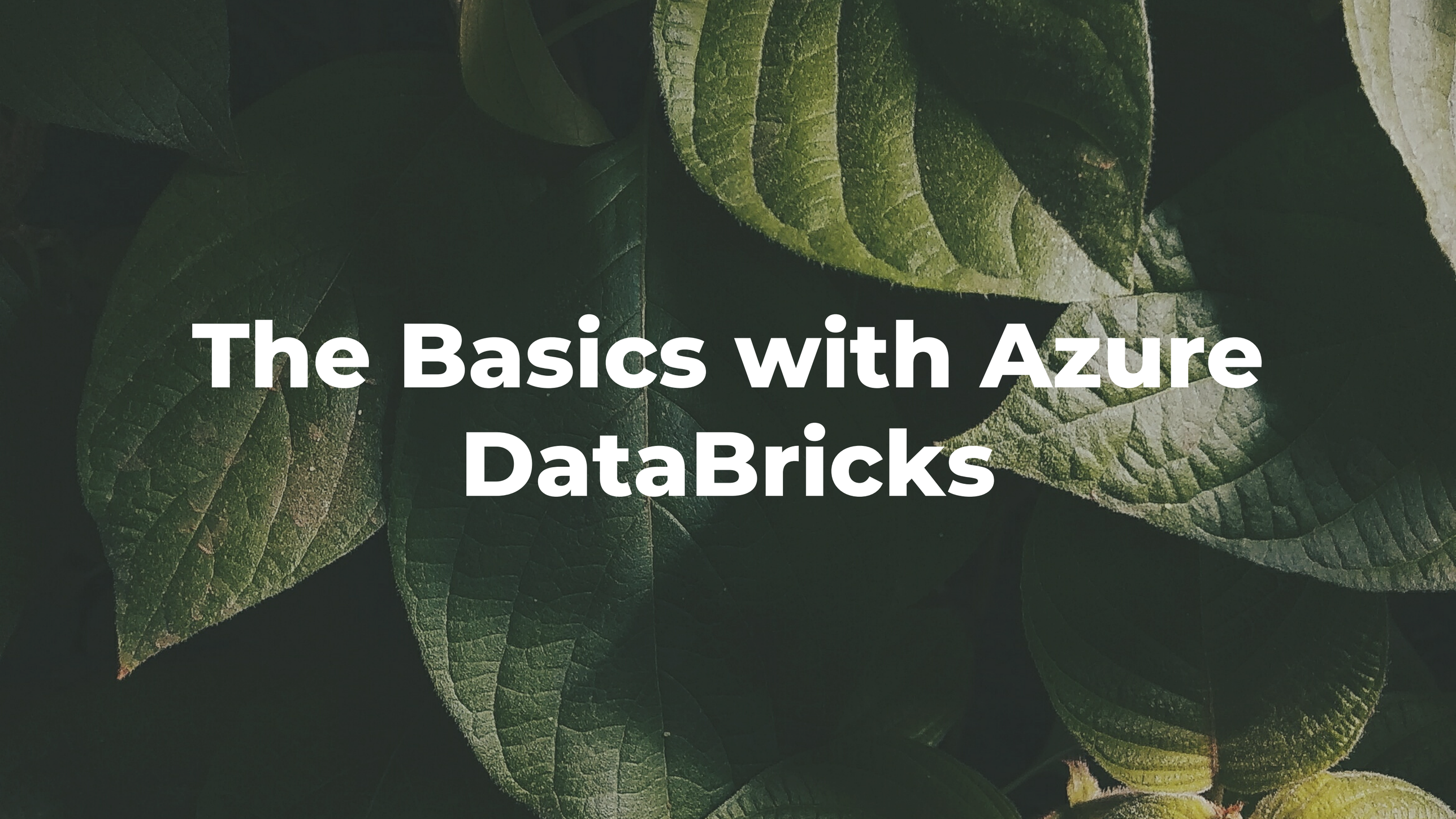
- Macro Scope.
- Discovering new things.
- Machine Learning and AI.
- Searching Engines.
- Corporate Analytics.
- Big Data usage.
- Data Modeling.
- Predictive Analysis.

- Micro Scope.
- Using existing things.
- Healthcare, Gaming, Travel.
- Immediate data needs.
- Corporate Analytics.
- Big Data usage.
- Statistics.
- Data visualization.

# What Microsoft Azure Offers for DS and DA <?>



- Data Science VM's (Linux and Wiondows)
- Microsoft's Machine Learning
- Azure Notebooks
- Azure HDInsights
- Data Lake Analytics
- Azure Power BI
- Azure Synapse Analytics
- Data Factory
- Azure DataBricks
- Text Analytics
- Web App Bot


The Basics with Azure
DataBricks

# Build your own Scientific Environment with Docker-Compose

- Computation: Jupyter Notebook.
- Structured data persistence: PostgreSQL.
- Non-structured data persistence: MinIO.
- Data visualization: Apache Superset.