

DATA SCIENCE ROAD MAP

2022



Arockia Liborius
Talk ML with AL

Why? Where? How? to begin your transition in Data Science

DATA SCIENCE JOURNEY

Four Phases of Data Science Road Map



EXPLORATION

- Why Data Science?
- Self realize if Data Science is your cup of tea
- Think Twice
- Ready For Life Long Learning?
- BI Tools



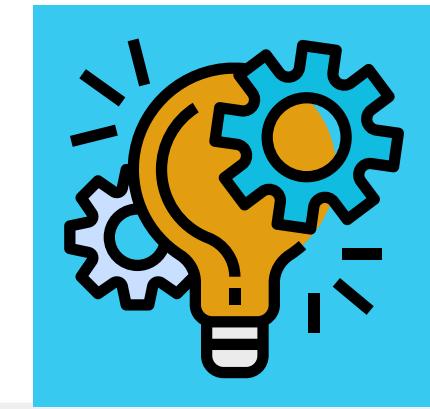
FOUNDATIONAL

- Programming
- Statistics
- Mathematics
- Data Pre-processing
- Data Wrangling
- EDA
- Data Storytelling



MACHINE LEARNING

- Classic ML
- Ensemble Modelling
- Hyper Parameter Tuning
- Deep Learning
- Big Data

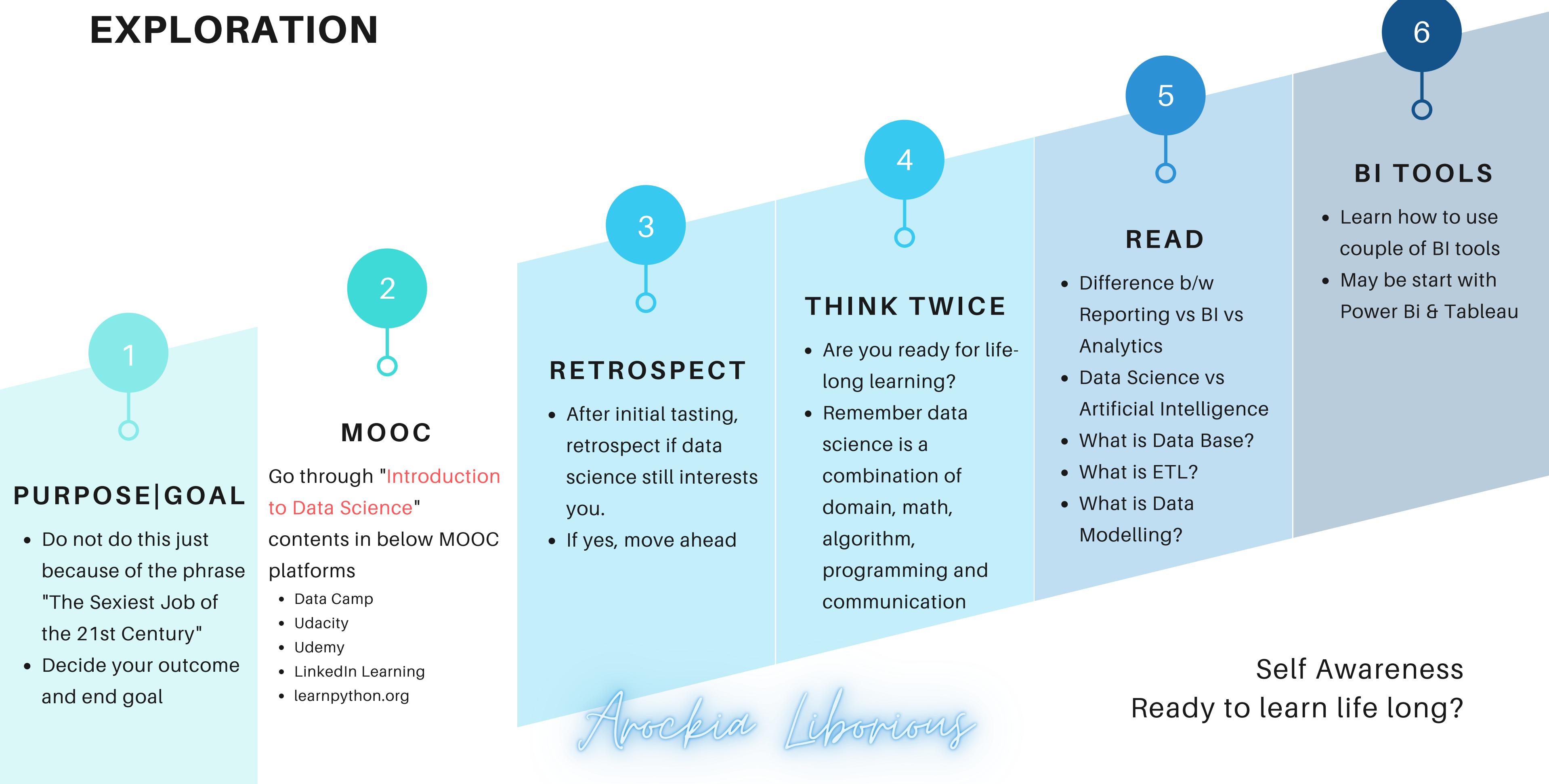


PROJECTS & BRANDING

- Hackathons
- Kaggle Projects
- Auto ML
- Resume/LinkedIn
- GitHub
- Interview Preparation

Arockia Liborious

EXPLORATION



FOUNDATIONAL PHASE

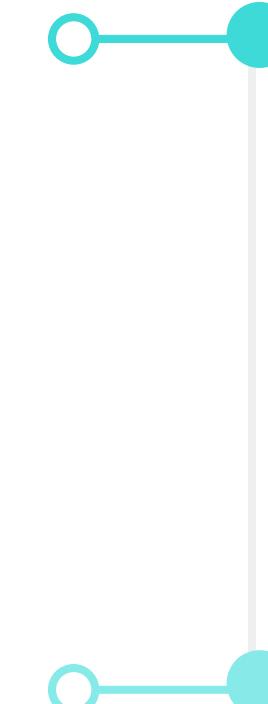
Read, Learn & Practice

PROGRAMMING

- Python (Numpy, Pandas, Matplotlib, and Seaborn)
- SQL DML
- R
- Excel

READ

- Blogs
- Articles
- You Tube Channels
- Top Data Science Use Cases in Business



DATA HANDLING

- Data Extraction
- Data Formatting
- Slice, Dice, Join, Index & Transform
- Missing Value Treatment

Arockia Liborios



MATH

- Descriptive Statistics
- Inferential Statistics - Sampling, Hypothesis Testing
- Probability - Bayes Theorem
- Distributions - Central Limit Theorem
- Correlation Co-efficient
- Linear Algebra



DATA STORYTELLING

- Ask right questions to your data
- Make sure it is aligned with business goals
- Communicate in natural language - easy to understandable

EDA

- Histograms, Box Plots, Scatter Plots, Heat Maps
- Univariate / Bivariate Analysis
- Outlier Detection
- Data Visualisation using Plotly
- Auto EDA Libraries - SweetWiz, Pandas Profiling

MACHINE LEARNING

Half Way / Turning Point

CLASSIC ML

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning (Optional)
- Model Evaluation Metrics
- Feature Selection
- Bias Variance Trade-off

ADV. ML

- Hyperparameter Tuning
- Adv. Regression
- Tree Models
- Bagging & Boosting
- Ensemble Models

TIME SERIES

- Time Series Analysis
- Simple Moving Average
- Exponential Smoothing
- Auto Regressive
- Vector Autoregression

DEEP LEARNING

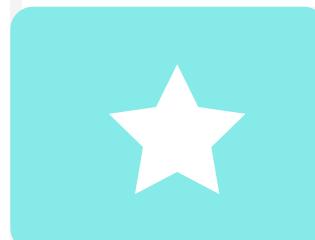
- Perceptron
- ANN
- TensorFlow
- CNN
- RNN
- LSTM
- Computer Vision
- GAN

NLP

- Lexical Processing
- Syntactic Processing
- Semantic Processing
- Chatbots

BIG DATA

- Command-Line Interface (CLI)
- ETL pipelines
- Hadoop & Spark
- AWS or Google Big Query or Microsoft Azure
- Model Deployment - (Heroku, Flask, Docker)



Arockia Liborios

DL - Deep Learning

NLP - Natural Language Processing

ANN - Artificial Neural Network

CNN - Convolutional Neural Network

RNN - Recurrent Neural Network

LSTM - Long Short Term Memory Networks

GAN - Generative adversarial networks

PROJECTS & BRANDING

Metamorphosis



Arockia Liborius

DATA IS MORE THAN OIL!

Arockia Liborios

FOLLOW ME

www.linkedin.com/in/arockialiborous

www.arockialiborous.com

