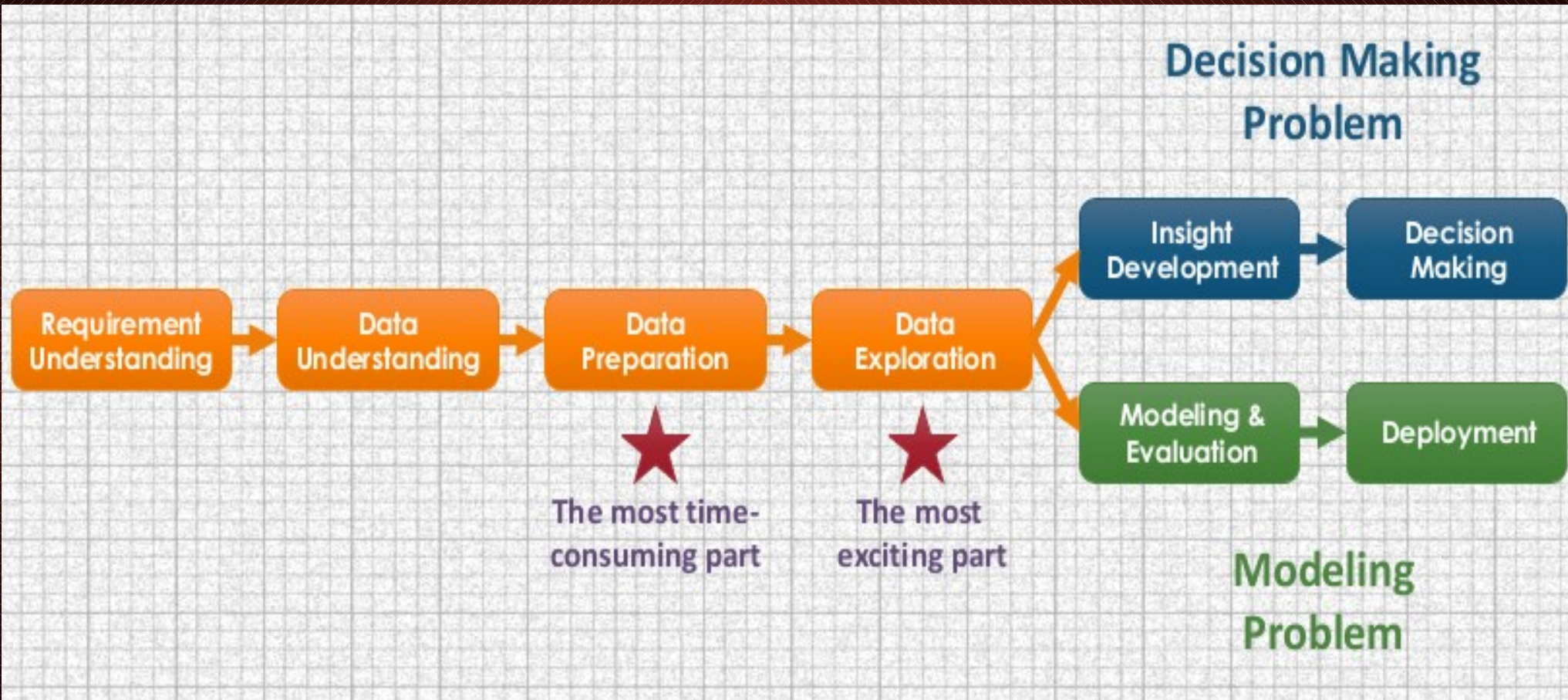# Gist of the workshop

Dr.Sm.Badhusha
Professor/CSE,
Alva's Institute of Engineering and Technology,
Mijar,Mangalore

Data analytics is the process and methodology of analyzing data to draw meaningful insight from the data

Machine Learning is the practice of applying algorithms to extract data and learn from it and forecast the future trends for a particular field

Typical Process of Data Analytics

Python as a Programming Language...

Python is a general-purpose
high-level programming language

Web development
Networking
Scientific computing
Data analytics

Libraries in Python for hot research areas

Library	Usage
numpy, scipy	Scientific & technical computing
pandas	Data manipulation & aggregation
mlpy, scikit-learn	Machine learning
theano, tensorflow, keras	Deep learning
statsmodels	Statistical analysis
nltk, gensim	Text processing
networkx	Network analysis & visualization
bokeh, matplotlib, seaborn, plotly	Visualization
beautifulsoup, scrapy	Web scraping

The flow of content delivery in the workshop ....

1. Python Basic Programming
2. Identifying the interesting dataset from the web
3. Load the dataset into a Pandas dataframe
4. Check the column names and see the first few rows
5. Data wrangling (cleaning, harmonizing, transposing, merging and filtering)
6. Do analysis with Explonatory Data Analysis (EDA)
7. To find the accuracy of the data analytical model by applying the Standard ML algorithms.

REFERENCES

Python for Data Analysis by Wes McKinney, Orielly publications

https://www.wikipedia.org/

https://uiowa.edu/datascience/resources

Let us dive
into
the demo and practical sessions 1