

Data Science Workshop

Safouene Kallel

2023-11-14

Session 0

- ① Introductions
- ② Data Science
- ③ Planning
- ④ Setup

Welcome

I am excited to have you join for the Data Science Training. The format of the workshop will be a mix of presentation slides, examples, and hands-on exercises. To ensure you are ready for the journey, please follow the installation instructions.

Introduction

- Name: Safouene Kallel
- Journey:
 - Worked in a Semiconductor Wafer Fab (ex-IBM, France, Malaysia)
 - Worked in an Automotive (Sensata, ex-TI, The Netherlands)
 - Worked in Semiconductor Design & RF-GaN Technology (Ampleon, ex-Philips/NXP, The Netherlands)
 - Currently working as a Data Scientist within the GaN Systems BL

Data Science

What is Data Science?

Data Science consists of:

- **Data Engineering:** Databases, Pipelines, Warehousing
- **Data Analytics:** Data Mining, Visualization, monitoring
- **Machine Learning:** Modelling, Feature selection, Experimentation, Deployment
- **Data Entrepreneurship:** Converting Data into Business, Solving Business Problems (8D) with data, Delivering results, Managing DS-Projects

Examples

Data sources:

Customer Returns, Equipment data, Characterization Data

Analytics:

Abnormal behavior detection, Process & Equipment Control

Business Value:

Reduce non Quality Costs, Early Warning, Efficient Maintenance, OEE
(Overall Equipment Efficiency) Increase

CRISP-DM

CRoss Industry Standard Process for Data Mining.

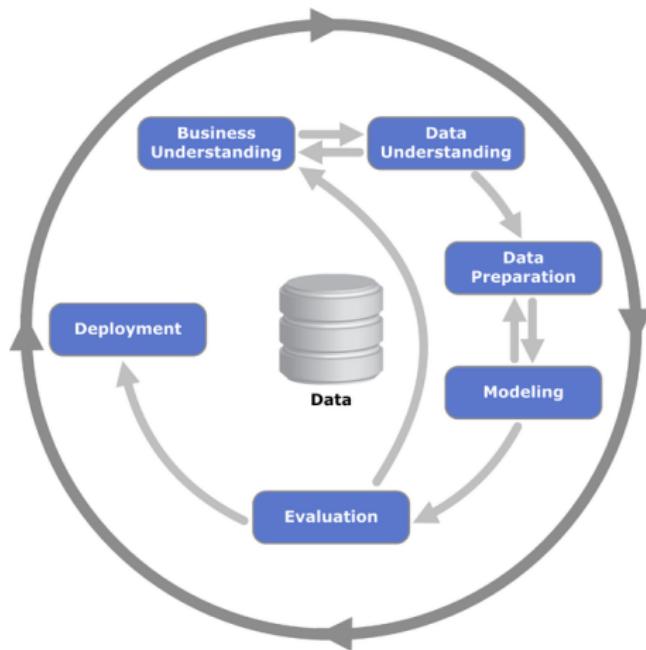


Figure 1: CRISP-DM

CRISP-DM

It is a standard, a model, or a set of guidelines to explore data in order to solve problems.

Close to DMAIC (Define, Measure, Analyse, Improve, Control) used in Six Sigma Projects.

This method is centered around data. That is where we are going to start.

Stage 0:

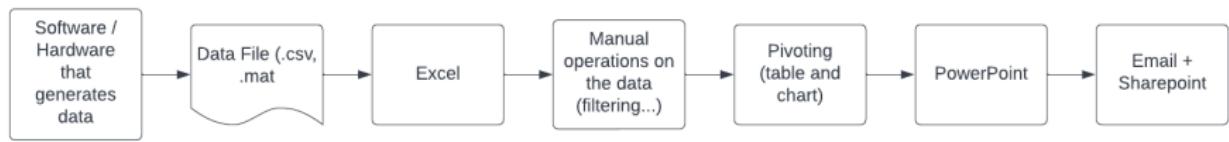


Figure 2: Stage 0

Stage 1:



Figure 3: Stage 1

Stage 2:

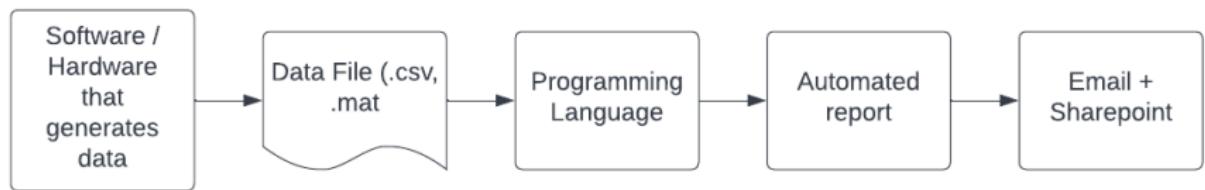


Figure 4: Stage 2

Stage 3:

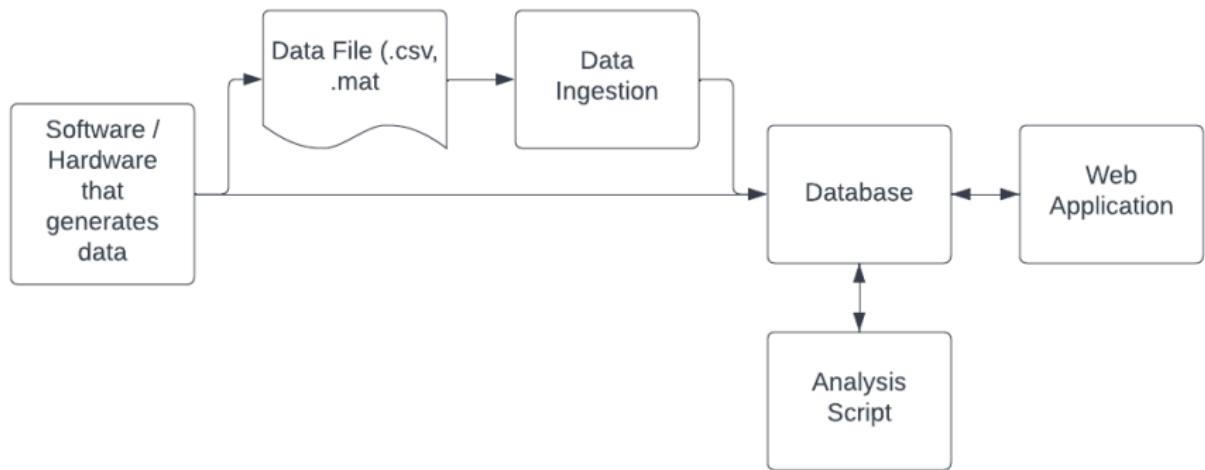


Figure 5: Stage 3

Planning

- Session 1: SQL
- Session 2: R
- Session 3: Python
- Session 4: Web Applications
- Session 5: Artificial Neural Networks

Confluence Page

Confluence Page

MySQL

- ① Please install MySQL Workbench (MySQL Client):
[MySQL Workbench](#)

R and RStudio

Local setup:

- R:
[download R](#)
- RStudio:
[download RStudio](#)
- Rtools:
[download Rtools](#)

Python

- Please install Anaconda:
[Anaconda](#)
- Open Anaconda Navigator:
click on CMD.exe

```
conda create --name Training
conda activate Training
pip install pandas
pip install tensorflow
pip install keras
```

Assignment

Create a working group with few colleagues and think about an idea on how to help Infineon through Data Science. Otherwise please check an open source data online.

Session 1

We get introduced to **MySQL**.