MODULE 4
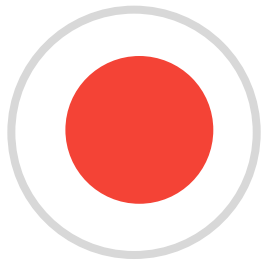
Deep Analytics, Visualization and Modelling

Let's Go

Structure of the module
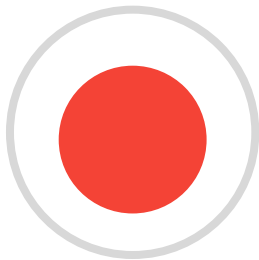
3 tasks
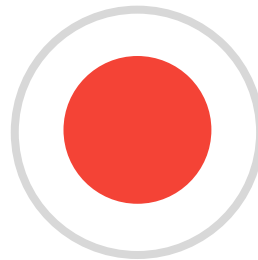
Task 1

28.01.2022

Task 2

23.02.2022

Task 3

23.03.2022
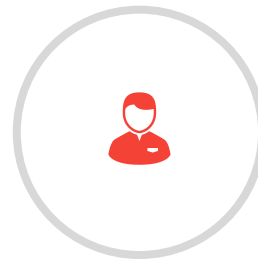
Who **is** who?

Which is your role in this module?

Company you work for

IOT Analytics

Goal

Build energy habits visualization for a Smart Home, create predictive model for energy consumption and develop model using wifi fingerprint to predict position

You

Congrats! You were promoted to junior data analyst

Business case

A large regional residential developer is designing a large 'Smart Home' apartment housing development and is **looking for evidence or positive reasons for adopting the use of electrical sub-metering devices used for power management in Smart Homes.**

Our job over the next few weeks will be to analyze this data to determine what kind of analytics and visualizations can be created that would empower Smart Home owners with greater understanding and control of their power usage.

Task 1 + Task 2

Task 1

1 Understand the Data → 2 Basic Data Exploration → 3 Present Findings

Develop subject matter expertise as fast as we can with basic data exploration

Task 2

4 Preprocessing → 5 Forecasting → 6 Create compelling charts → 7 Final deliverable
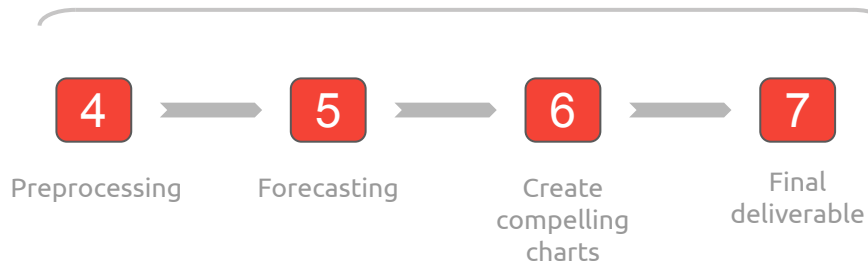
Perform analysis, forecasting and visualization

ubiqum code academy

Module 4 - Task **1**

Customer Segmentation

Tasks

- Understand the data:

Familiarize yourself with

concepts and electrical field

- Basic Data Exploration:

Check distribution, NA values

-Present findings:

Present finding and introduce

next steps

Deliveries
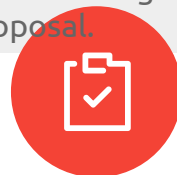
PPT Presentation

Skills

Understand the data
Introduction to electrical
Unit transformation

Basic Data Exploration
Lubridate
Tydiverse
DPLYR
SQL Basic

Present findings

Convert technical insight into
business proposal.

Deliverables - 1st task

1. Cover Slide
2. Agenda – What will be discussed in the presentation?
3. Background – What is the context of this project?
4. Objective/Goals – What is IOT Analytics being asked to do? What does the client want from this work?
5. Data Management - An explanation of how you will manage the data during the project.
6. Descriptions and location of related data.
7. Any known issues with the data and how you plan to address them.
8. **Descriptive statistics you gathered as an initial step in analysis.**
9. **High-Level Recommendations – Three or more recommendations about the existing data. What would you change or add to the data?**
10. **Example of potential dashboard/visualization you aim to create**

Deliverables - 1st task

Our job over the next few weeks will be to analyze this data to determine what kind of analytics and visualizations can be created that would empower Smart Home owners with greater understanding and control of their power usage
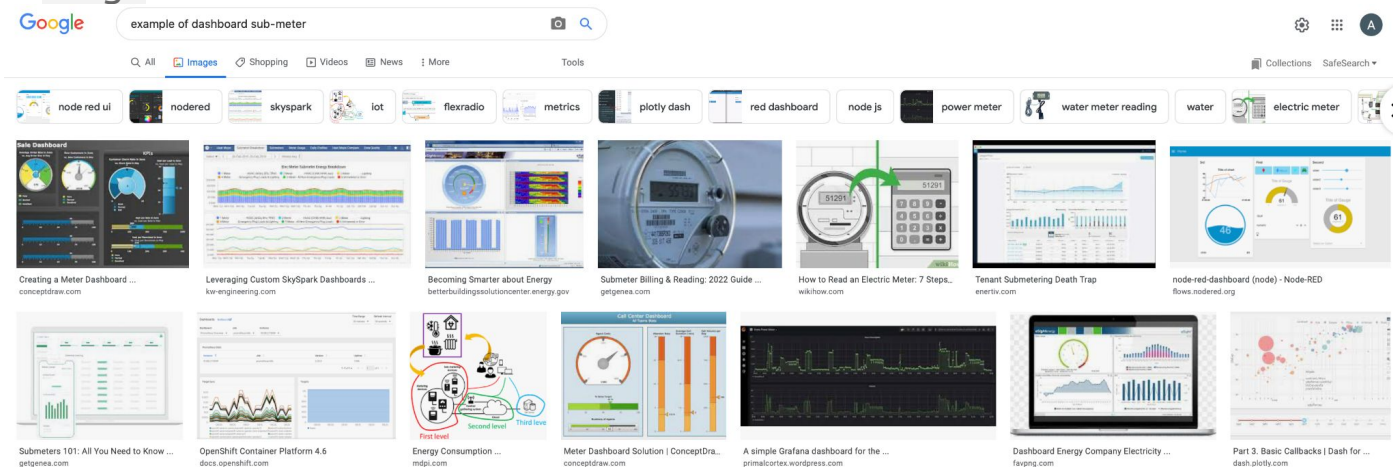
Example of Dashboard for submeters

The dataset

Attribute Information:

1.**date**: Date in format dd/mm/yyyy

2.**time**: time in format hh:mm:ss

3.**global_active_power**: household global minute-averaged active power (in kilowatt)

4.**global_reactive_power**: household global minute-averaged reactive power (in kilowatt)

5.**voltage**: minute-averaged voltage (in volt)

6.**global_intensity**: household global minute-averaged current intensity (in ampere)

7.**sub_metering_1**: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered).

8.**sub_metering_2**: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light.

9.**sub_metering_3**: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.

Module 4 - Task **2**

Tasks

- Modelling:

Use forecasting technique to forecast energy consumption

- Modelling:

Create compelling visualization

Deliveries

Report + script

Skills

Forecasting
Datetime, Regression, Exponential smoothing, others

Dashboard

Convert technical insight into business proposal.

Who **is** who? - Task 3

Which is your role in this module?

Company you work for

IOT Analytics
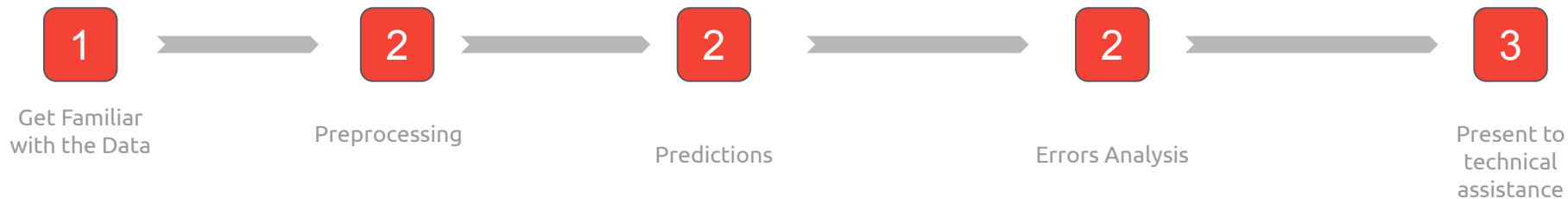
Goal

Build an indoor geo-localisation system

Module 3 **Plan**

Task 3

1	2	2	2	3
Get Familiar with the Data	Preprocessing	Predictions	Errors Analysis	Present to technical assistance

Module 3 - Sprint **1**

Customer Segmentation

Tasks

- **Preprocessing:**

- **Prediction:**

- **Error Analysis:**

Deliveries

Jupyter Notebook
or R script

Skills

Preprocessing
Logarithmic scale
Exponential transformation
Feature selection
Plotly

Prediction
Cascade model
Split modelline
Iterative modelling

Error Analysis
Error distribution
Error location
Confidence interval of errors

Dataset

Attribute 001 (WAP001): Intensity value for WAP001. Negative integer values from -104 to 0 and +100. Positive value 100 used if WAP001 was not detected.

....

Attribute 520 (WAP520): Intensity value for WAP520. Negative integer values from -104 to 0 and +100. Positive Vvalue 100 used if WAP520 was not detected.

Attribute 521 (Longitude): Longitude. Negative real values from -7695.9387549299299000 to -7299.786516730871000

Attribute 522 (Latitude): Latitude. Positive real values from 4864745.7450159714 to 4865017.3646842018.

Attribute 523 (Floor): Altitude in floors inside the building. Integer values from 0 to 4.

Attribute 524 (BuildingID): ID to identify the building. Measures were taken in three different buildings. Categorical integer values from 0 to 2.

Attribute 525 (SpaceID): Internal ID number to identify the Space (office, corridor, classroom) where the capture was taken. Categorical integer values.

Attribute 526 (RelativePosition): Relative position with respect to the Space (1 - Inside, 2 - Outside in Front of the door). Categorical integer values.

Attribute 527 (UserID): User identifier (see below). Categorical integer values.

Attribute 528 (PhoneID): Android device identifier (see below). Categorical integer values.

Attribute 529 (Timestamp): UNIX Time when the capture was taken. Integer value.

Objective

The objective is to be able to predict :

Building ID

Floor

Longitude and Latitude

Thank you

Deep Analytics, Visualization and Modelling