Assessment

What do I need to do?
Where do I start?

What tools do I need?
Template Notebooks

Read the assessment
Read data description
What can I figure out?

Review what was covered
Review Notes
Review Notebooks
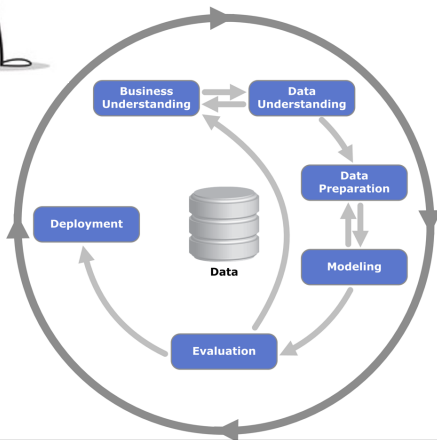

Remember the Process
(see Notebook layout)

1. Business Understanding
What's the problem?
What is required?
What is needed?
Week 1 + 2

2. Data Understanding
What data do you have?
Can you load the data?
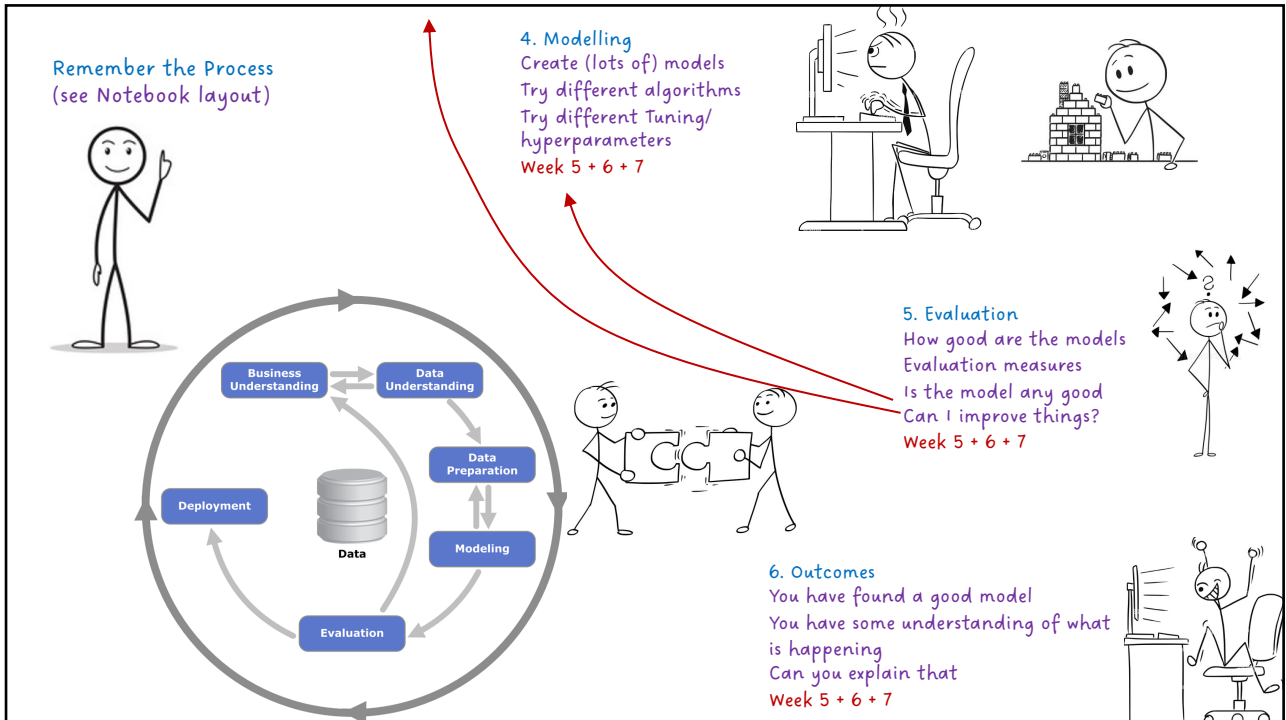Can you do some Charts?
Can you do some Stats?

What can you learn about the data
from doing this?
What's your understanding of the data?
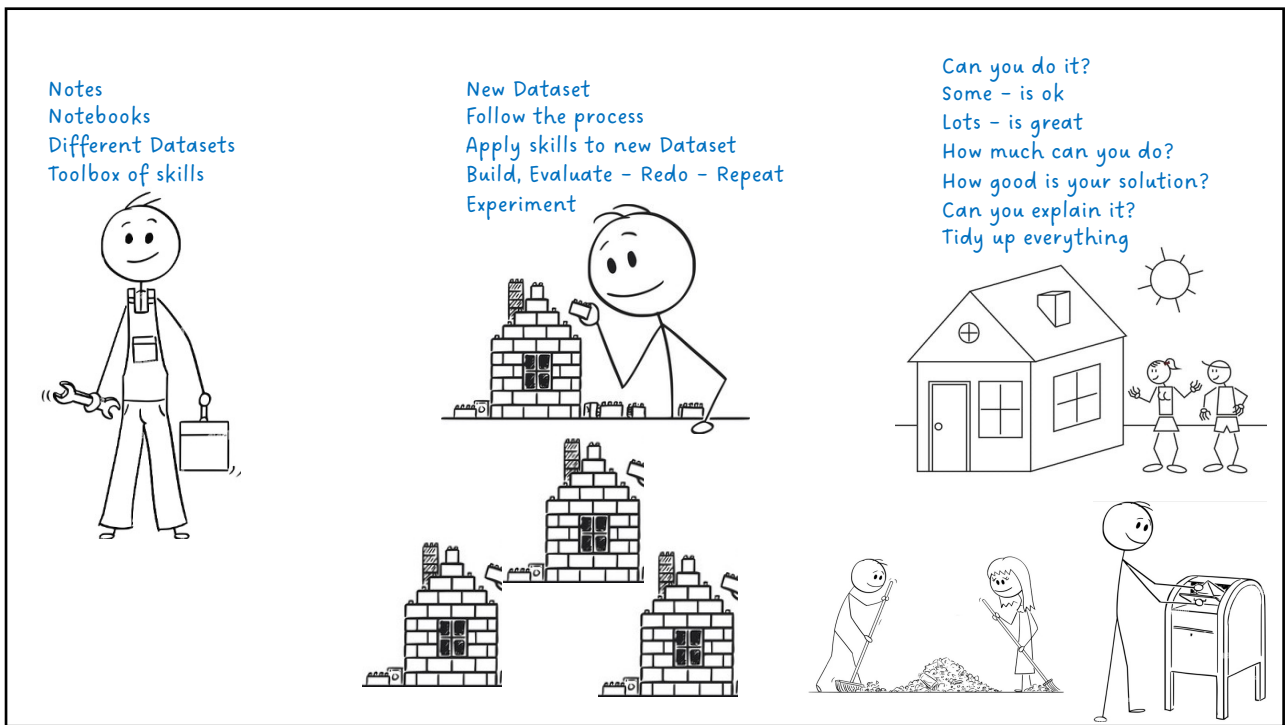Any low-hanging fruits?
Week 2 + 4

3. Data Preparation
Is the Data Clean?
Do I need to do any Data Cleaning?
Do I need to do any Data Transformations?
Are there any features I need to remove & why?
Anything else I need to do with the data? & why?
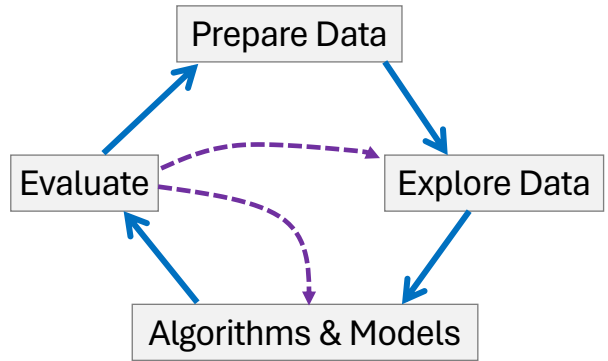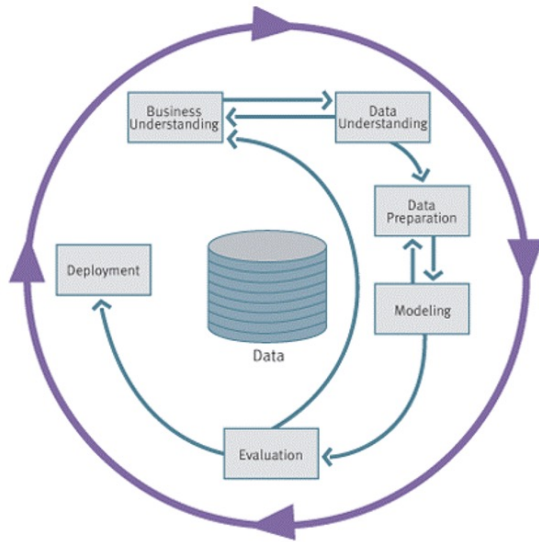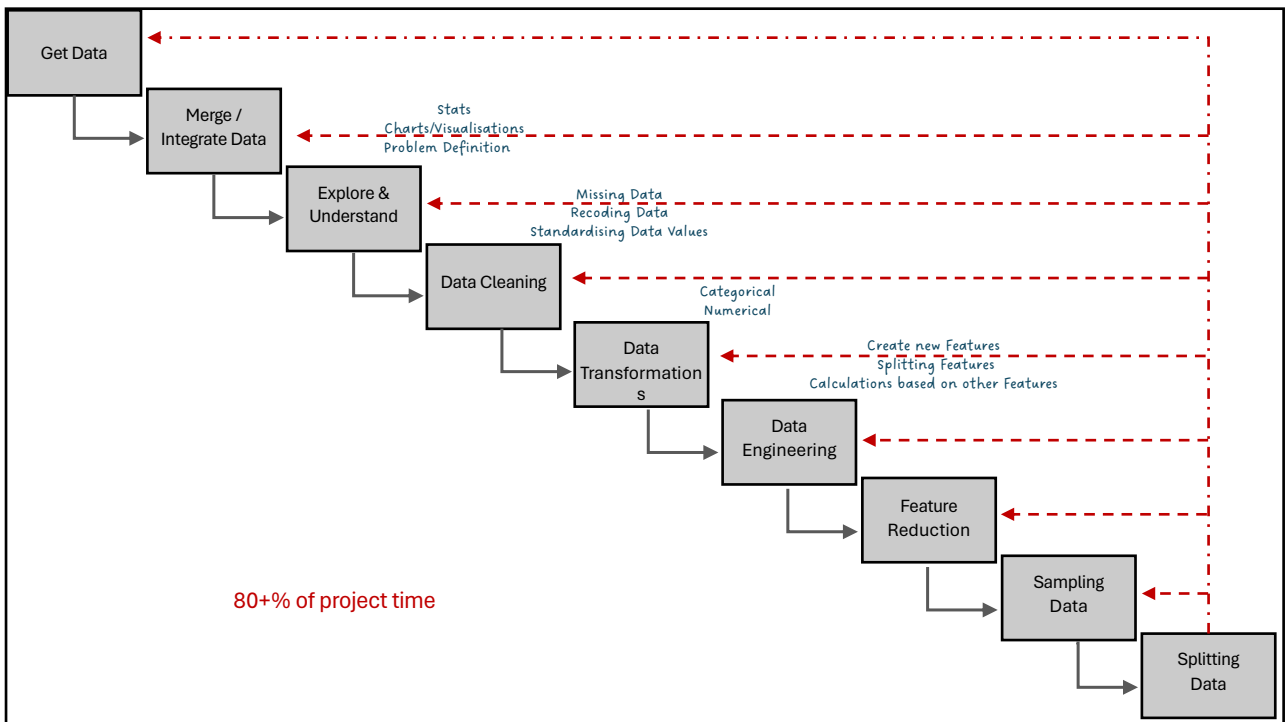Training & Test datasets.
Week 4 + 5

Business Understanding
Data Understanding
Data Preparation
Modeling
Evaluation
Deployment
Data

Remember the Process
(see Notebook layout)

4. Modelling
Create (lots of) models
Try different algorithms
Try different Tuning/
hyperparameters
Week 5 + 6 + 7

5. Evaluation
How good are the models
Evaluation measures
Is the model any good
Can I improve things?
Week 5 + 6 + 7

6. Outcomes
You have found a good model
You have some understanding of what
is happening
Can you explain that
Week 5 + 6 + 7

Business Understanding
Data Understanding
Data Preparation
Modeling
Evaluation
Deployment
Data

Notes
Notebooks
Different Datasets
Toolbox of skills

New Dataset
Follow the process
Apply skills to new Dataset
Build, Evaluate – Redo – Repeat
Experiment

Can you do it?
Some – is ok
Lots – is great
How much can you do?
How good is your solution?
Can you explain it?
Tidy up everything

The work is Iterative

80+% of project time

Business Problem

7

Document, Explain, How clear can you make it. Only include what is necessary

What is the outcome of your experimentation? What does it mean?

How can you improve things? Hyperparameter Tuning, better Data Prep

Iteratively create models & Evaluate. Compare. Document findings

What Data to include / Exclude. Document Why

Data Transformations, Cleaning, Recoding.

Document what you have learned

What Data Insights can you identify

Collect & Explore Data

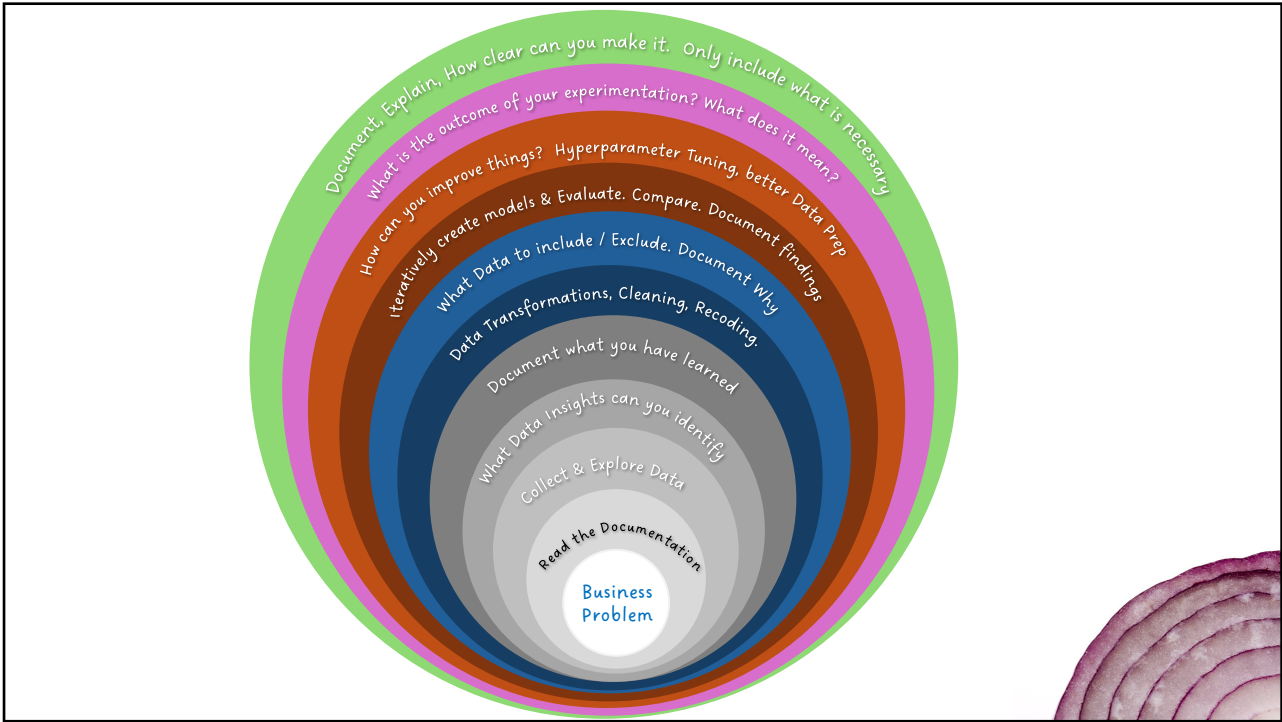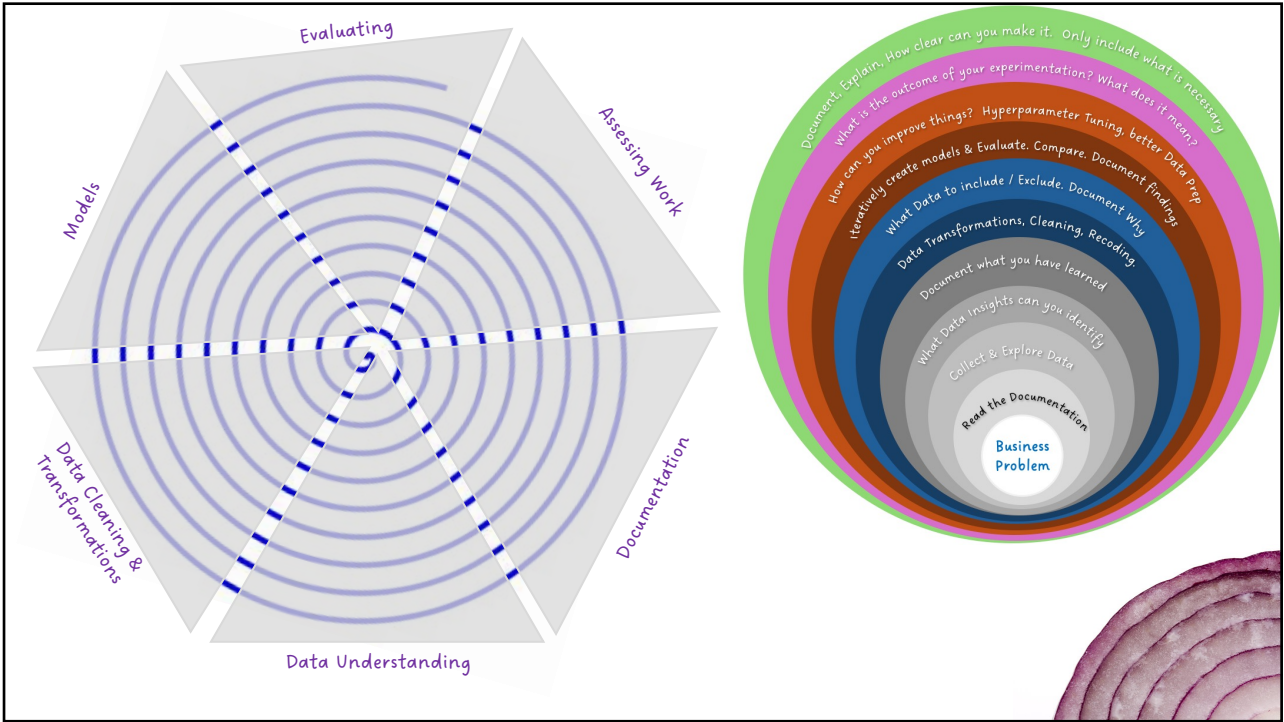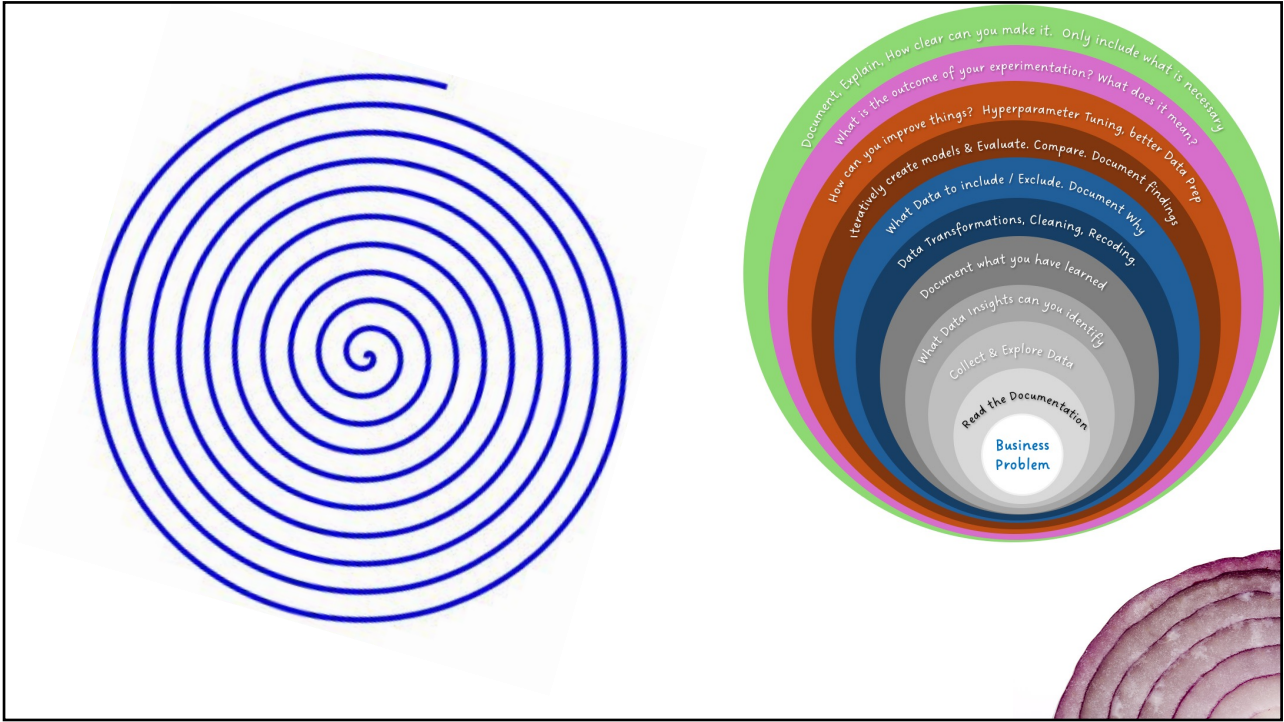Read the Documentation

Business Problem

8

9

10

Slide 11

Documentation is Key

- Explain each step
- Explain why you included/excluded features or other things
- Explain what is shown in a table/diagram/chart etc
- If you do something, can you explain <u>Why</u> and what you <u>Learned</u>
- Explain what you learned or discovered at each step
- How would you explain this to your Boss? If this was for work?

- You don't have to document everything – just the cleaned-up process
- Add as Comments or as Markdown
    - Comments – for Code
    - Markdown - for discussion

- Experimentation is Key
    - Iterate - Non-linear
    - Evolve with each Iteration (you might Discover something and have to go back to fit it in and re-do)
    - Discuss with Team/Group – You will learn from each other – This is part of the learning process
    - There isn't (necessarily) one correct solution. There could be many. "It Depends!"

Slide 12