

Introduction to managing data projects and products

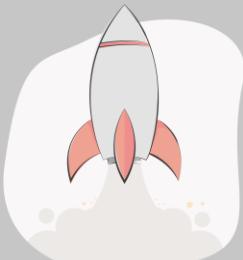


L5 Data Engineer Higher Apprenticeship
Module 1 / 12 (“Data Fundamentals”)
Topic 3 / 4

The challenges of traditional project management

The need for a more data-centric approach

- Data science and data engineering projects are dynamic and unpredictable
- Constantly evolving requirements due to exploration of unstructured data, identification of new trends, and real-time predictions
- Incompatibility of Waterfall approach with the dynamic nature of data projects



A data-centric approach means:

- Adoption of Agile, Lean, and Six Sigma principles
- Emphasis on fostering a culture that supports data-driven decision-making
- Focus on delivering tangible value

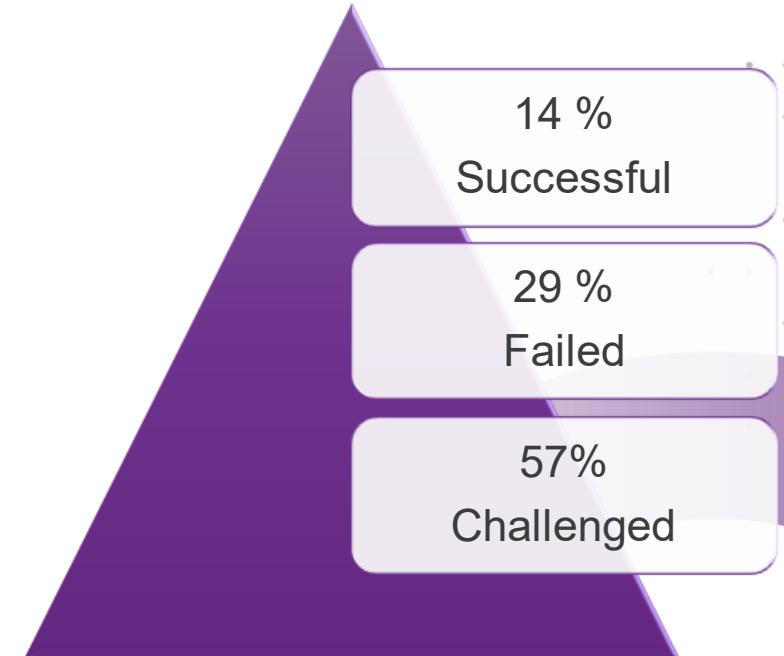


Image source: Standish Group, 2012

Embracing the principles of Agile, Lean and Six Sigma

Empowering data-driven decision-making

Agile projects are 28% more successful than traditional projects (PwC)

76% of business in the Netherlands and Belgium believe that Agile projects will outnumber Waterfall projects by 2020

Over a quarter (27.4%) of manufacturing organisations rely solely on Agile, whereas 56.6% rely on a 'combination of methodologies' (Liquid Planner)

Almost three-quarters (71%) of organisations report using Agile approaches sometimes, often, or always (Project Management Institute)



Organisations are embracing Agile working practices to be more adaptive and collaborative

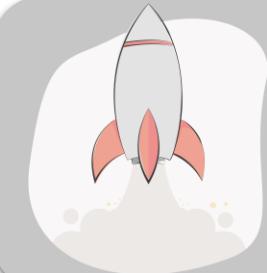
Module Case Study

A reminder

- Credit Bank Corporation aimed to gain insights into employee performance, satisfaction, and HR processes
- The HR team planned to build a comprehensive analytics dashboard to track key metrics
- Data engineers must integrate diverse data and follow best practices
- The team must determine data types, build a suitable data ecosystem, visualise results, and articulate business value



Case study: Credit Bank Corporation

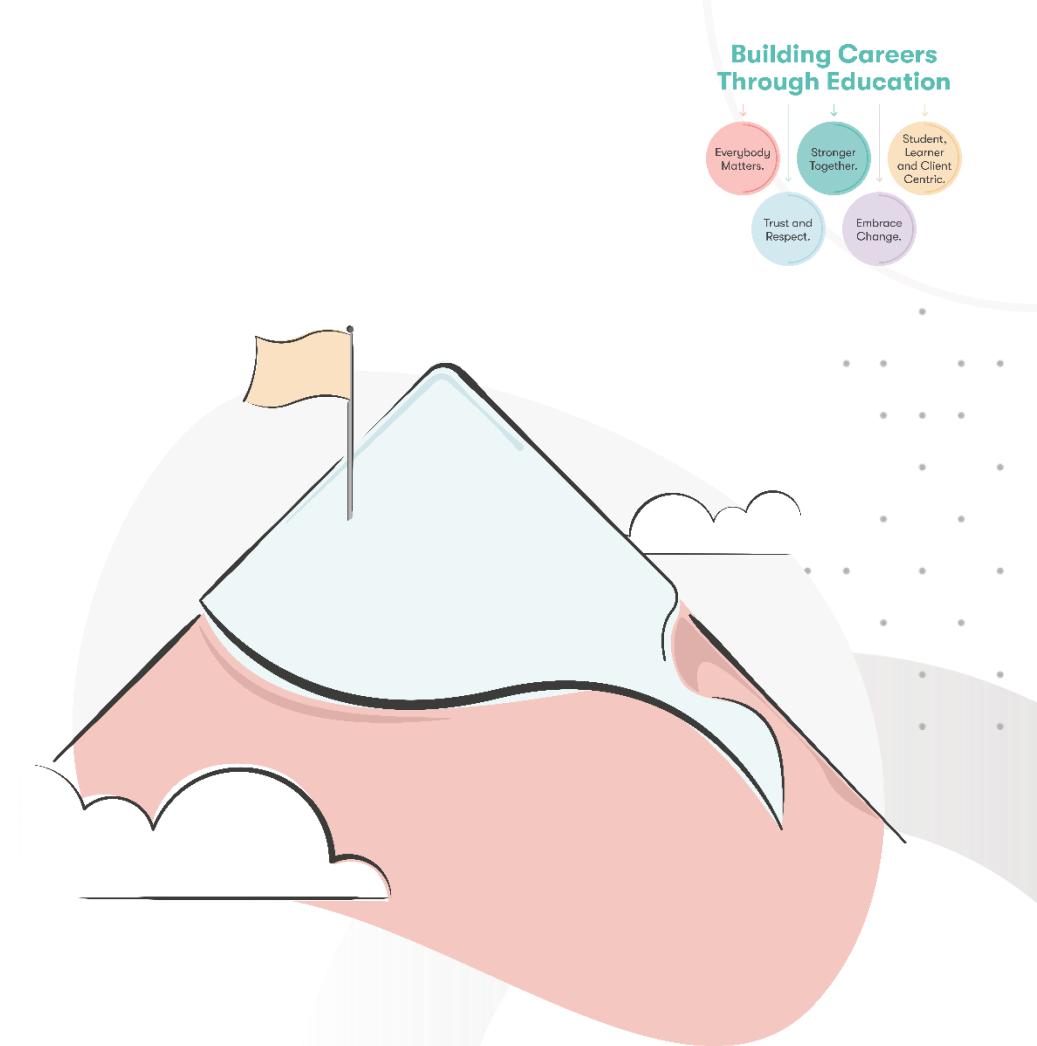


This 'Data Fundamentals' module is designed to provide you with the knowledge and skills to overcome these challenges!

Learning outcomes

This webinar is designed to support the following learning outcomes:

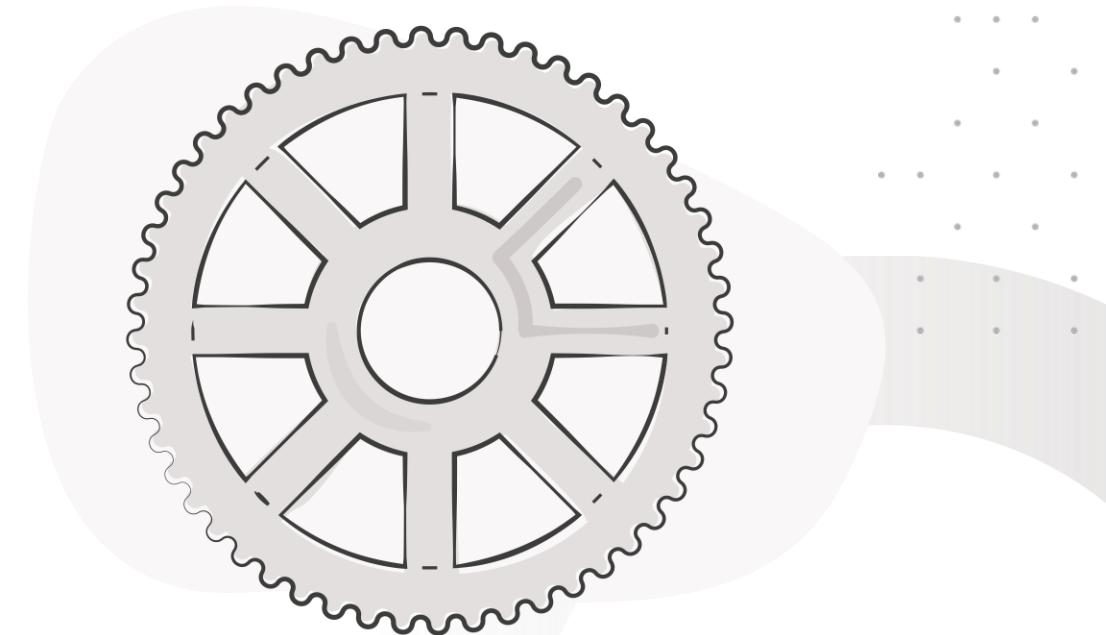
- List the limitations of the Waterfall model in the context of data science projects
- Explain how Agile principles and frameworks can foster adaptability and collaboration
- Identify which Lean and Six Sigma practices can drive continuous improvement
- Evaluate approaches to collaborating with diverse stakeholders in a way that aligns data-driven initiatives with business objectives
- Argue how a data-driven culture can be championed within your organisation by leveraging the synergies between Agile, Lean, and Six Sigma approaches



Webinar agenda

This webinar will cover the following:

- The limitations of Waterfall project management
- Embracing Agile principles
- Navigating roles and responsibilities in Agile Scrum
- Integrating Lean and Six Sigma Principles
- Practical lab



Webinar length: 3 hours

Limitations of Waterfall project management



Section introduction

What does the Waterfall model look like?

- Waterfall is a sequential, linear methodology used in software development projects
- It comprises distinct phases from requirements gathering to deployment
- Despite its structured nature, it has limitations:



- Difficulty accommodating change
- Lengthy development cycles
- Lack of stakeholder feedback
- Challenges with data complexity

Requirements

Design

Development

Testing

Deployment

Stages of Waterfall project management

Knowledge Check Poll

Credit Bank Corporation aimed to construct an analytics dashboard to enhance insights into HR metrics, aiming to improve workforce productivity, retention, and efficiency.

However, during the project, new data sources emerged, posing challenges for integration.

Which limitation of the Waterfall model does this scenario illustrate?

- A) Difficulty accommodating change
- B) Lengthy development cycles
- C) Lack of stakeholder feedback
- D) Challenges with data complexity

Feedback: D – A difficulty accommodating change.



**Submit your responses to
the chat!**

Challenges with Waterfall

A real-world example: BigBox Corp

Background: BigBox Corp. launched a data-driven transformation project in 2018

Objective: Aimed to create a comprehensive data warehouse and analytics platform for customer behavior insights

Approach: Initially used a traditional Waterfall approach, focusing on detailed requirements and design

Challenges:

- New data sources and evolving requirements from shifting retail landscape
- Waterfall's rigidity posed challenges in adapting to changing needs and data landscape
- Lengthy development cycles resulted in outdated requirements upon deployment

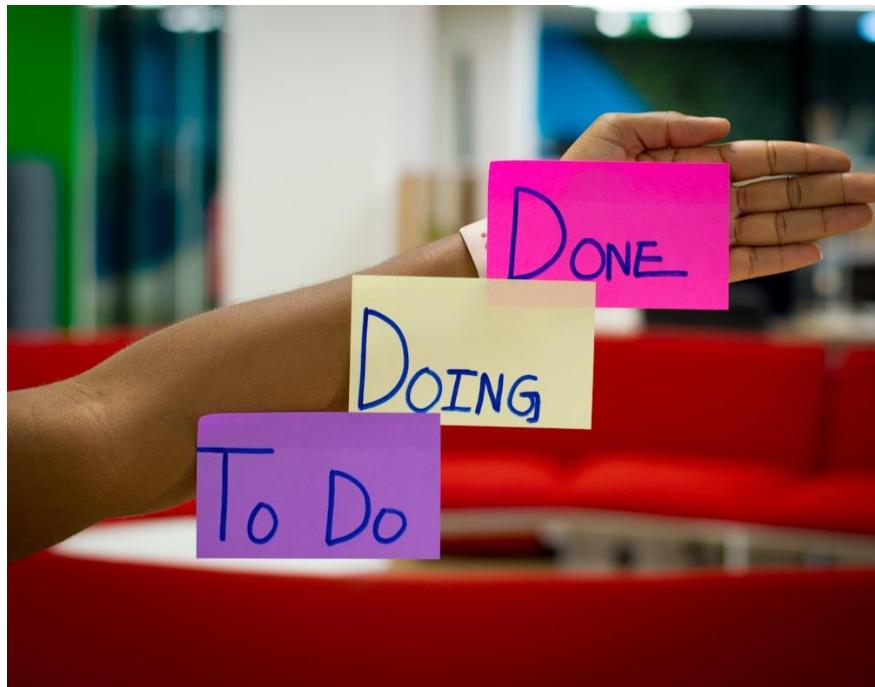


Case study: BigBox Corp

Waterfall challenges

A summary

Understanding the challenges associated with Waterfall is crucial for project managers and teams to mitigate risks and optimise project outcomes.



The challenges associated with Waterfall

Embracing Agile principles



Section Introduction

Agile principles

- Agile methodology is gaining traction in data science projects, moving away from rigid Waterfall approaches
- Agile aligns with the dynamic nature of data science projects and evolving requirements
- Frameworks like Scrum and Kanban revolutionise PM in data engineering, with emphasis on flexibility, responsiveness and value



The core values of Agile include:

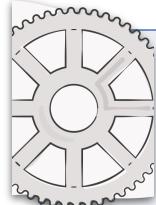
- Interactions over processes and tools
- Iterative development
- Close collaboration
- Responding to change



The benefits of Agile in data engineering

In the context of our case study

Let's explore the benefits of adopting Agile methodology in data engineering projects, using the case study of Credit Bank Corporation as a practical example.



Faster iteration



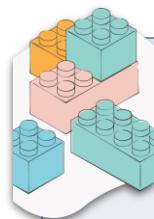
Effective feedback loops



Enhanced collaboration



Responsiveness to change



Flexible planning

The benefits of Agile



Case study:

Credit Bank Corporation

Exploring Agile frameworks

What is Scrum?

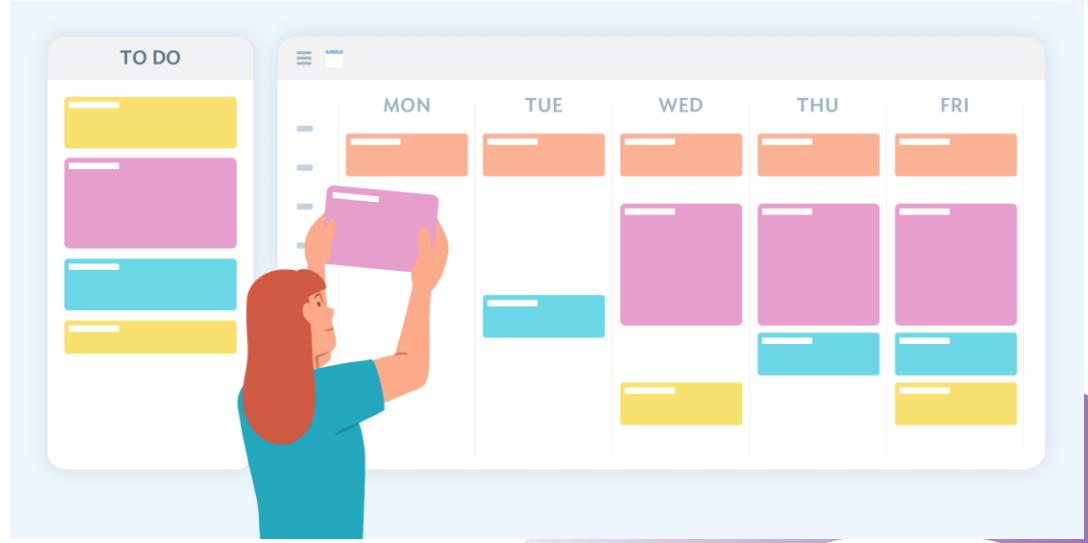
- A lightweight and practical Agile methodology (framework)
- Whereas Waterfall often uses tooling such as MS Project, Scrum uses tooling such as Atlassian Jira
- Used in companies such as:

Google

Spotify®

IBM

3M



Timebox iterations: Sort tasks by priority, estimate relative effort

Image source: getclockwise.com

Exploring Agile frameworks

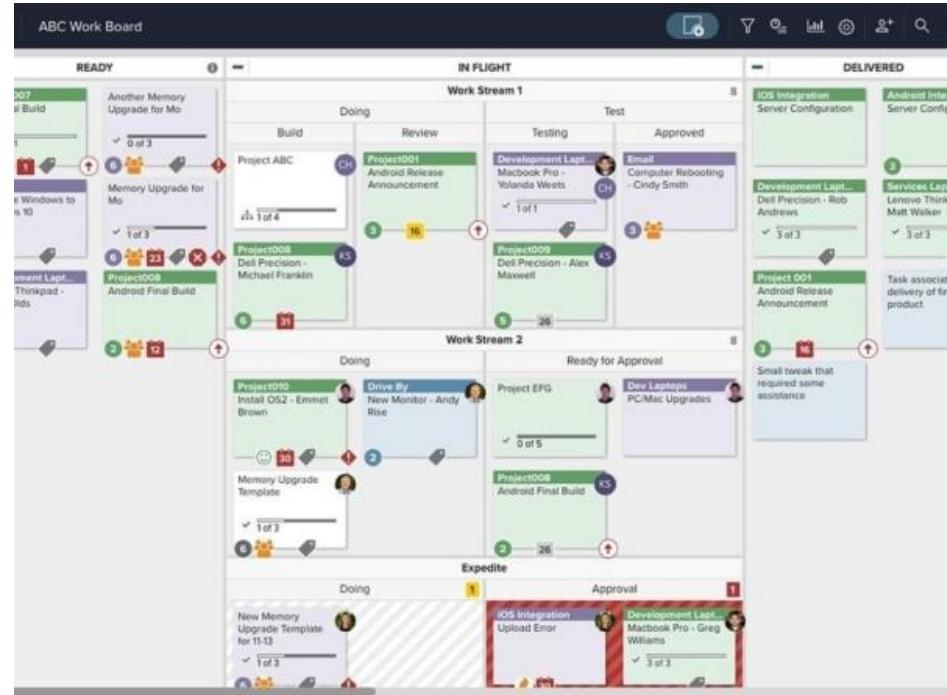
What is Kanban?

- In the context of data engineering, Kanban can be particularly useful for managing the complex workflows and diverse tasks involved in the project lifecycle
- Kanban focuses on visualising work, limiting work in progress (WIP), and enhancing project visibility

This screenshot shows a Kanban board interface for SwiftKanban Development 4.0. The board is organized into columns representing different stages of the workflow: Ready for development, Design, Coding + Junit, Func Automation, Pending Review, Code Review, Dev Complete, In-Progress, Done, and Validation. Each column contains a list of tasks with their respective status and identifiers.

Ready for development (17/10 - 26)	Design (1/8 -)	Coding + Junit (12/12 -)	Func Automation (1/12 -)	Pending Review (3/5 -)	Code Review (1/3 -)	Dev Complete (4)	In-Progress (0/5 -)	Done (0/10 -)	Validation (0/15 -)
TASK36, UST679, UST684, ENHC402, UST680, UST685, UST690, UST663, TASK533, UST688, UST691, ISS159, UST686, CDEF446, UST669	UST683	CDEF444, UST687, TASK602, DEF1043, UST661, UST682, TASK594, ENHC401, UST662, TASK504, ISS175, ENHC442	ISS165, UST681, UST664, UST670	TASK556, UST659, ISS163	ISS176, TASK601, ISS164, ISS166, ISS160, DEF1038, CDEF443				

Kanban example 1



Kanban example 2

Exploring Agile frameworks

What are sprints?

- Sprints are fixed-time iterations, between 2 and 4 weeks in length
- 3-week sprints are typical
- Each sprint starts with planning and ends with retrospective
- A sprint produces a demonstration of something that works

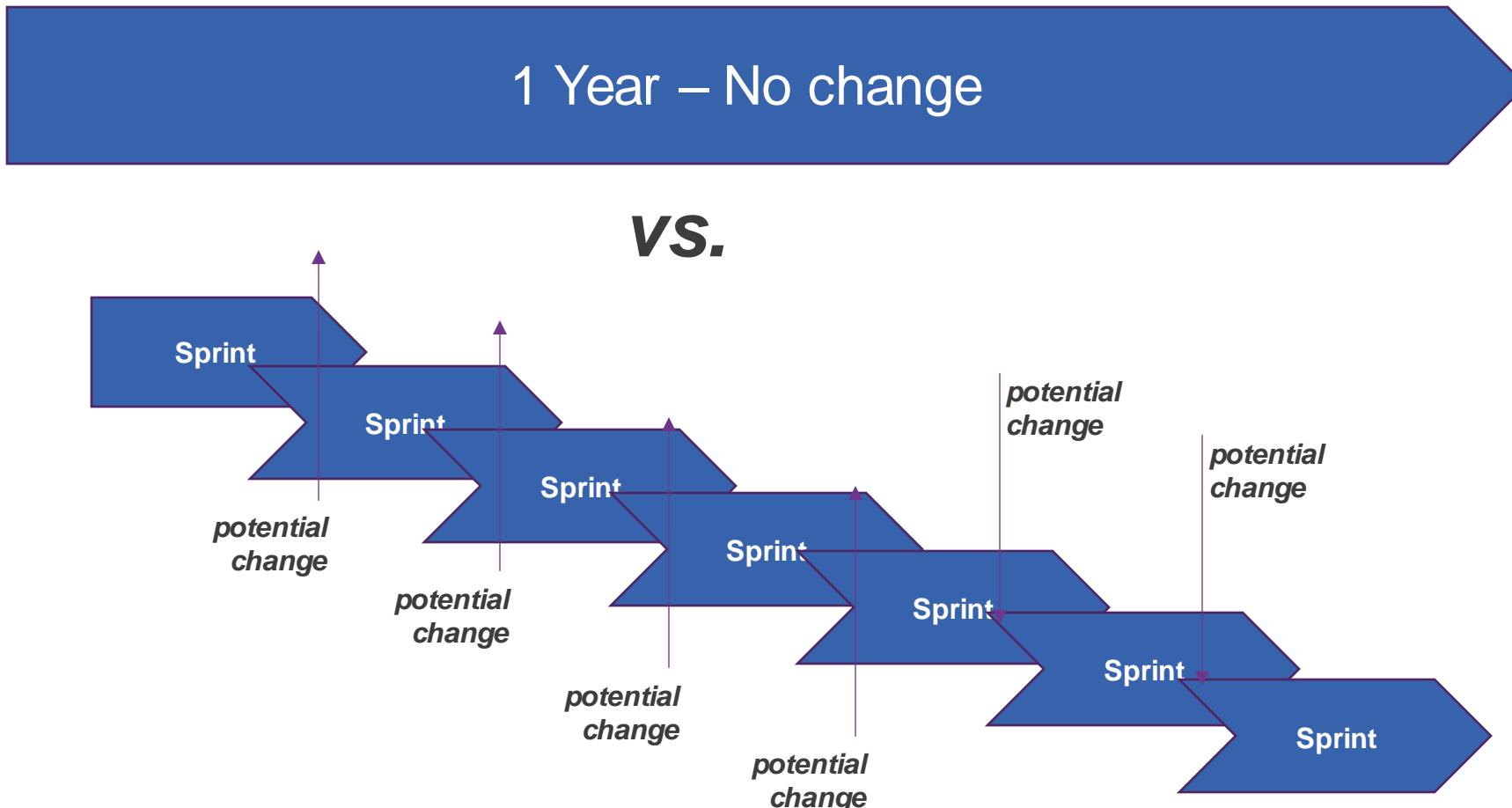


Sprint planning

Image source: Unsplash

Exploring Agile frameworks

Linear Waterfall vs Agile iteration



*The iterative nature of Agile compared
to Waterfall*

Exploring Agile frameworks

What are daily stand-ups?

These are short daily team meetings (usually in the morning).

They are designed to:

- Discuss what you achieved the previous day
- Discuss the plans for the rest of the day
- Discuss any blockers



Daily stand-ups

Image source: MeetingNotes.com

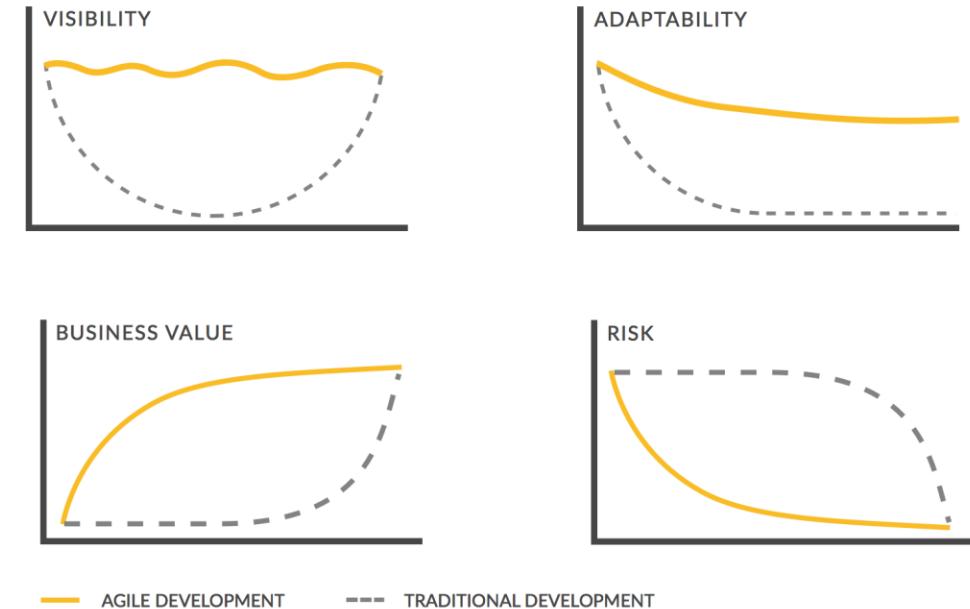
The value proposition of Agile

In summary

Here's a comparison of the value proposition of Agile versus traditional (Waterfall) development approaches:

Metric	Waterfall	Agile
Visibility	Limited visibility due to milestone-based progress tracking.	Provides ongoing insight into project progress
Adaptability	Less adaptable to changing requirements.	Offers flexibility to respond to changing requirements
Business value	Delays value delivery until project completion.	Delivers incremental value throughout the project
Risk management	Higher risk of issues due to inflexible planning and execution.	Identifies and addresses issues early in the development process

AGILE DEVELOPMENT
VALUE PROPOSITION



The value proposition of Agile

Image source: MeetingNotes.com

Take a break!



Navigating roles and responsibilities in Agile Scrum



Section Introduction

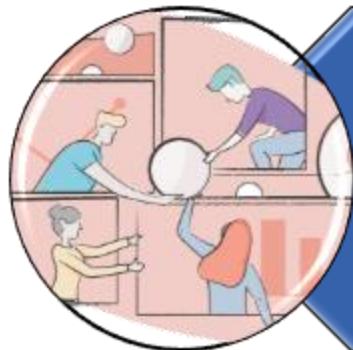
Scrum roles

In this section, we'll explore the individual responsibilities of Scrum roles how they collaborate within the Scrum.

Here's the basics:



Scrum roles have specific responsibilities and collaborate for project success



Scrum ceremonies, like daily standups and sprint planning, ensure effective communication and alignment

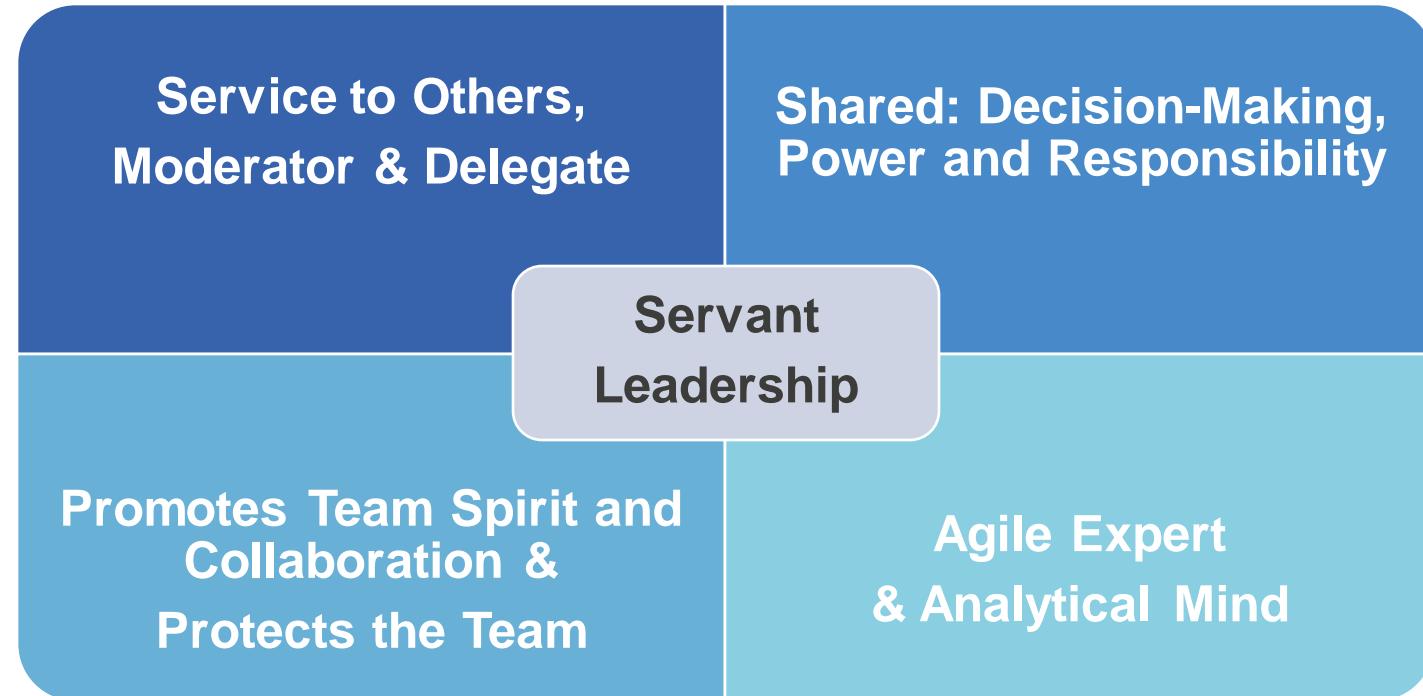


Image source: Teamly

The Scrum Master

Servant leadership

The Scrum Master is a pivotal role in Scrum methodology, embodying the principles of servant leadership. Here's a breakdown of the key aspects of servant leadership as it relates to the Scrum Master:



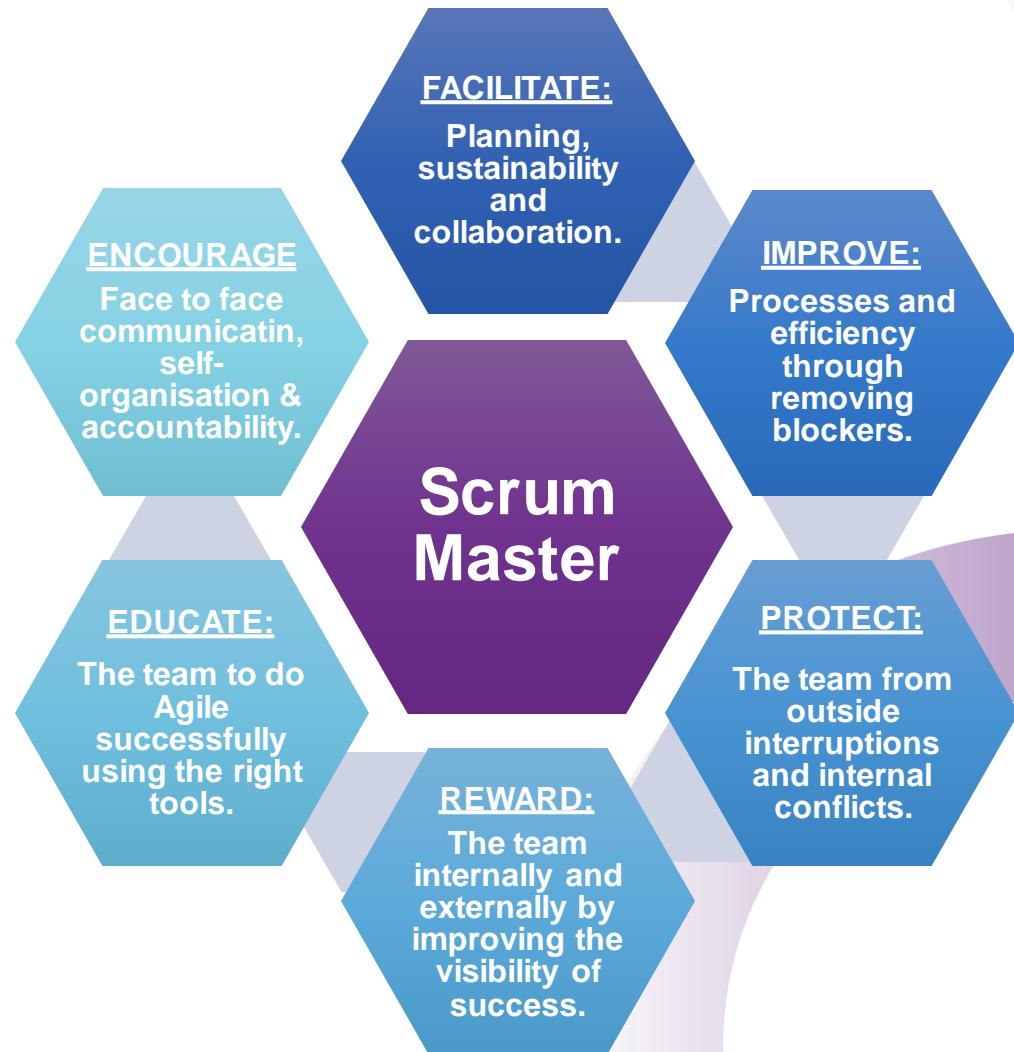
Key aspects of Scrum Master servant leadership

The Scrum Master

Understanding their role

The role of the Scrum Master encompasses several critical functions aimed at facilitating, improving, protecting, rewarding, and educating the Scrum Team.

Here's a breakdown of each aspect:



The critical functions of the Scrum Master

The Product Owner

Understanding their role

Agile Product Owners play a pivotal role in ensuring the success of Agile projects by taking on various responsibilities, including:



The critical functions of the Scrum Master

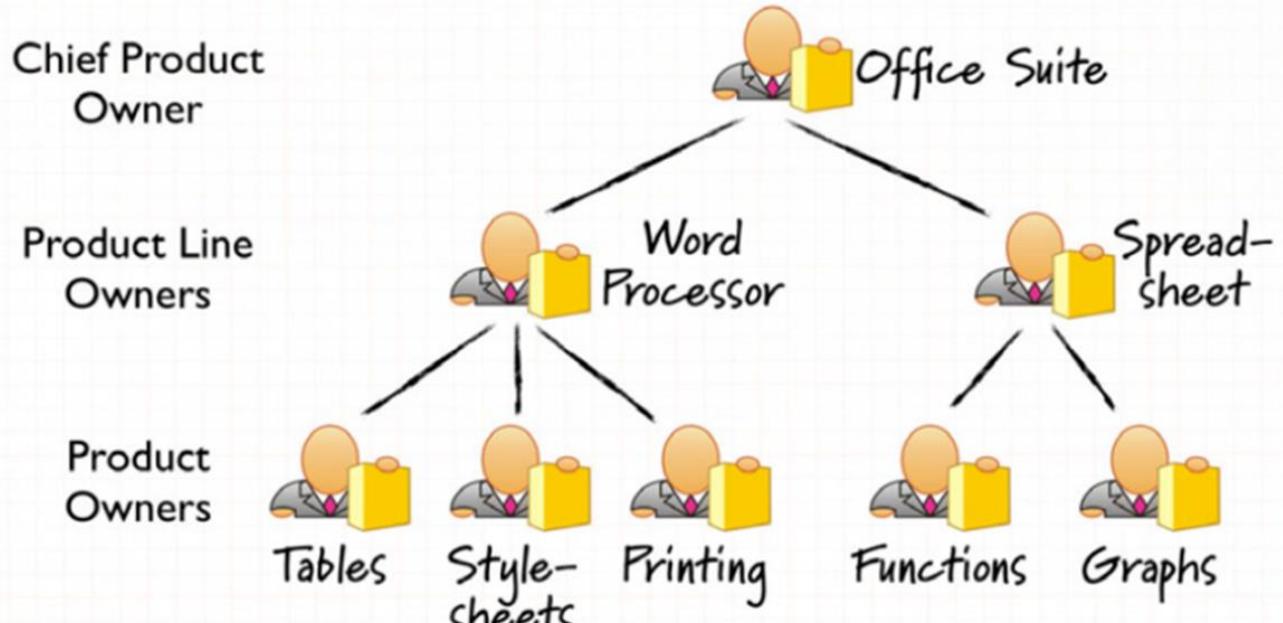
Multi-level Product Owners

Managing and ensuring alignment

In Agile development, the concept of multi-level Product Owners can be employed to effectively manage product development and ensure alignment with business goals.

The following roles are involved:

- Chief Product Owner
- Product Line Owners
- Product Owners

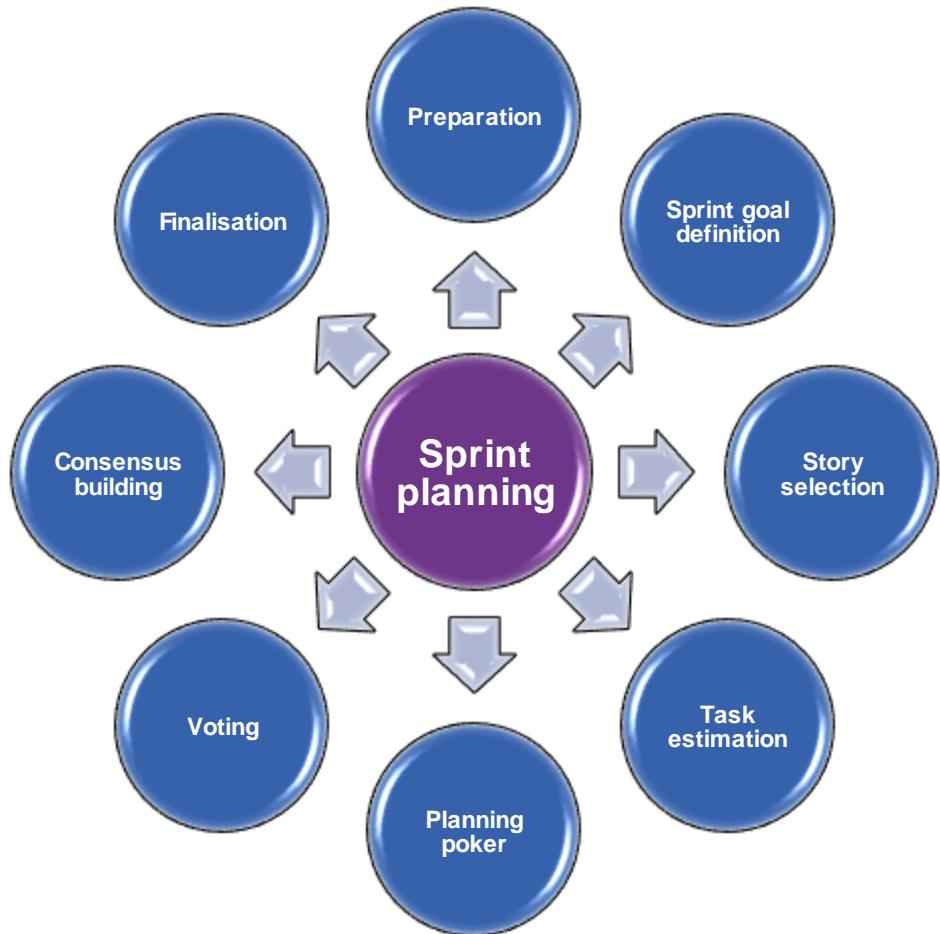


Multi-level Product Owners

Sprint planning

What is involved?

Agile sprint planning is a vital Scrum ceremony where the team collaborates to define the upcoming sprint's work.



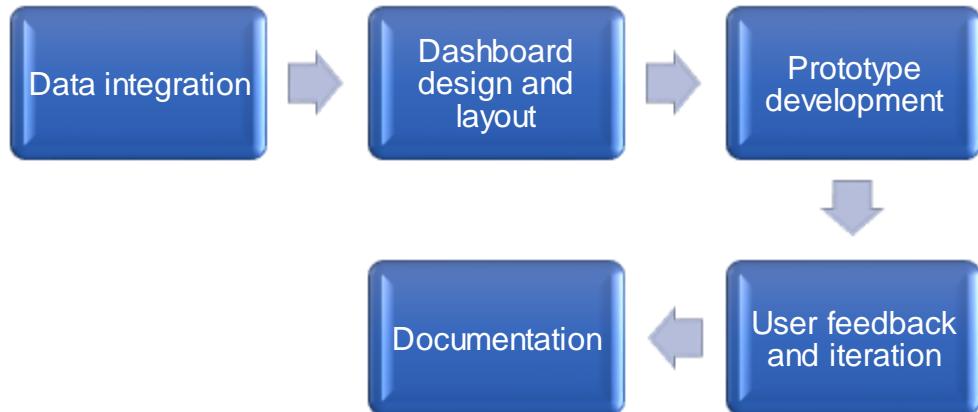
Sprint planning

A case study example

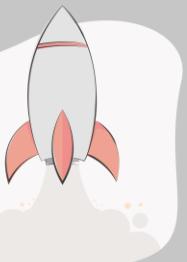
Sprint duration: 3 weeks

Sprint goal: Create a functional prototype of the analytics dashboard by integrating and processing HR data from various sources

Tasks:



Case study: Credit Bank Corporation



End of sprint deliverables:

- Functional prototype of the analytics dashboard
- Documentation on data integration process and dashboard usage
- User feedback report and iteration plan for future sprints

Burndown charts

Visualising the remaining work

A Burndown chart is a visual representation of the work remaining in a sprint or project over time.

Burndown charts:

- Track remaining work against time axis
- Help monitor progress and adherence to sprint goals
- Facilitate identification of deviations from planned trajectory
- Enable early detection of issues and bottlenecks
- Promote transparency, accountability, and informed decision-making

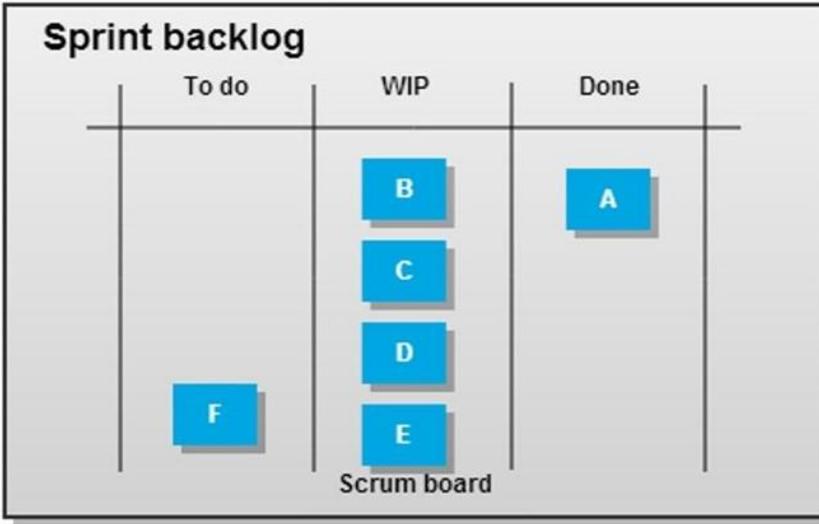
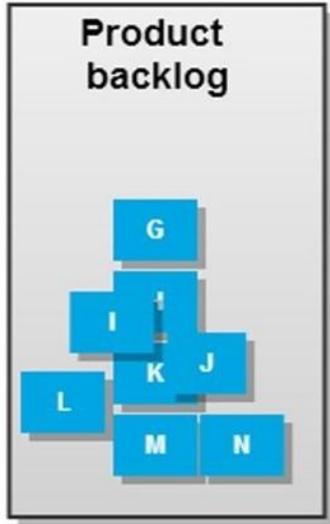


Multi-level Product Owners

Scrum boards and sprint backlog

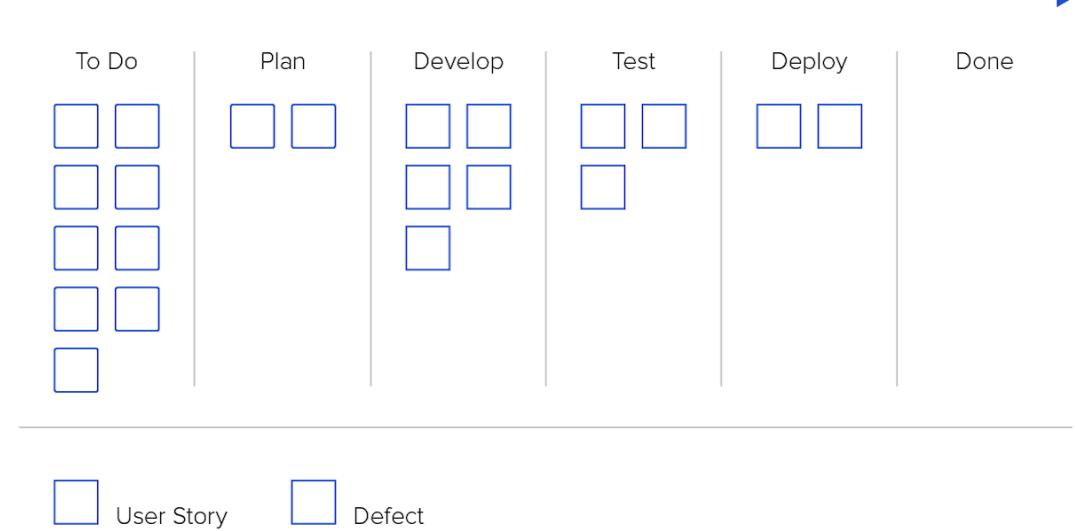
What are they and how can they help?

A Scrum board visually represents the sprint backlog with columns like "To Do," "In Progress," and "Done," enabling team members to track progress and address bottlenecks.



Typical Scrum board

Image source: ?



Scrum board variation

Backlog grooming

What does this involve?

Once the initial product backlog is created and prioritised, the ongoing backlog groom process takes over.

- Ongoing grooming ensures top backlog items are sprint-ready
- Meetings involve product owner, team, and Scrum Master
- Team reviews stories before the meeting
- Criteria, complexity, and strategy are discussed
- Details are refined until everyone agrees
- Stories are estimated using planning poker

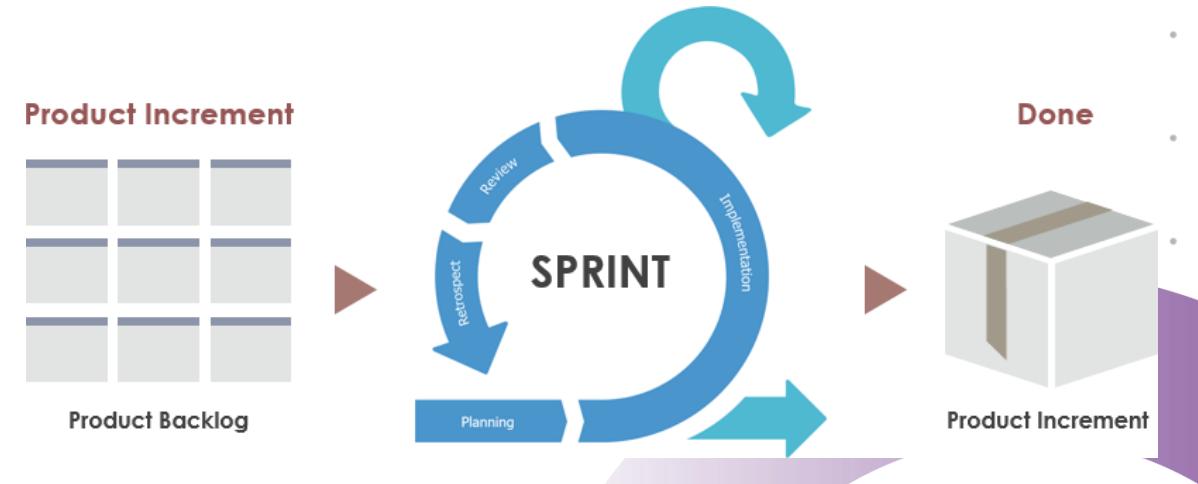


Image source: Visual paradigm

Take a break!



Knowledge Check Poll

Sprint reviews vs sprint retrospectives

- A) True
- B) False



Sprint review

What? Product



Sprint retrospective

How? Process



**Submit your responses to
the chat!**

Feedback: B - Sprint review and sprint retrospective are actually two distinct ceremonies in Agile methodology.



Sprint reviews

What do they look like?

A sprint review is a crucial ceremony in Agile methodology that occurs at the end of each sprint cycle

Its primary purpose is to demonstrate completed work to stakeholders

During sprint reviews the Scrum team showcases the product developed during the sprint

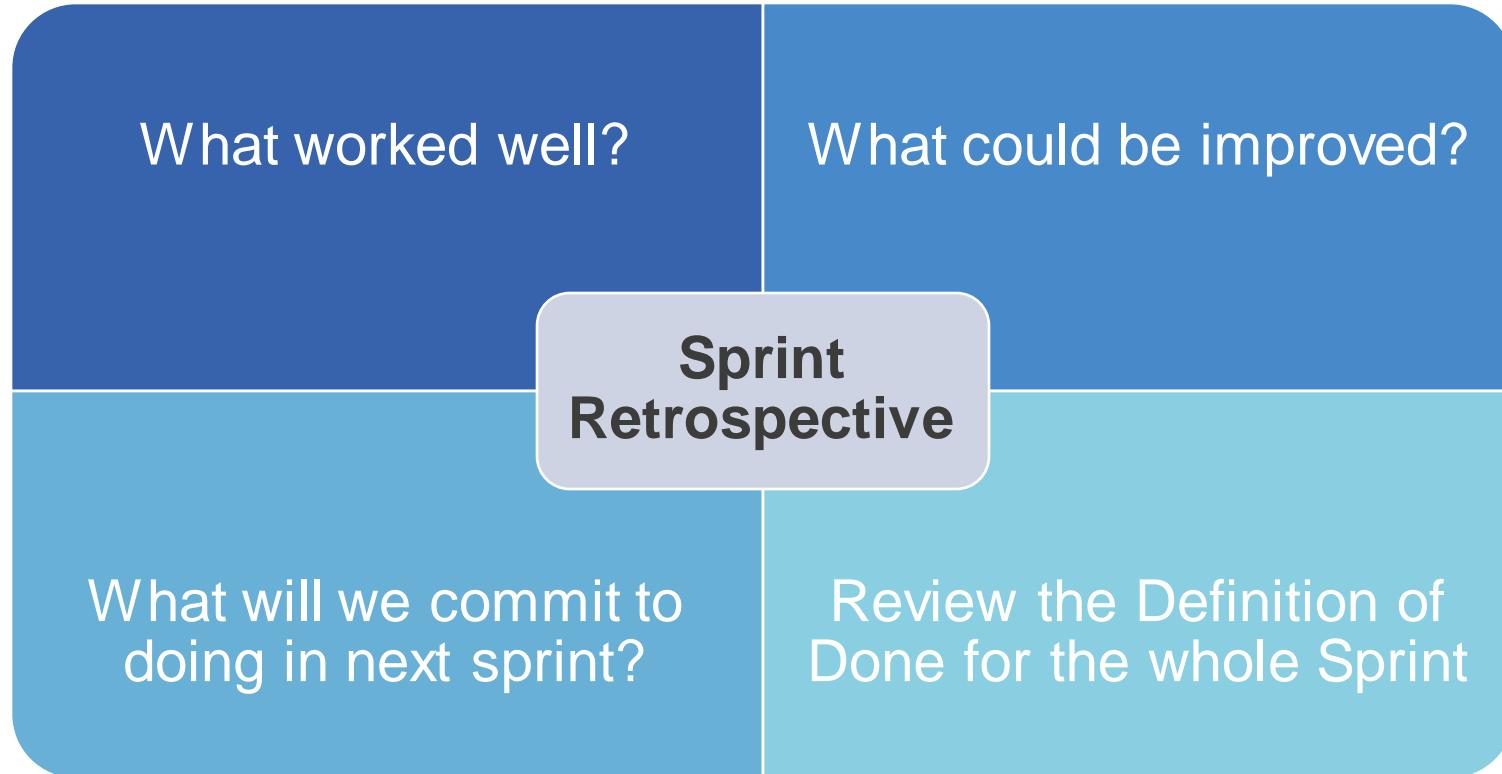


Sprint review

What? Product

Sprint retrospectives

What do they look like?



Sprint retrospective

How? Process

Writing User Stories

What does this involve?

Writing user stories is a key aspect of Agile planning, as they serve as the foundation for defining requirements and guiding development efforts.

Here are some key points about writing user stories in Agile planning:

- User-centric
- Template
- INVEST Criteria
- Acceptance criteria
- Iterative refinement

OFTEN, **USER STORIES** ARE FORMULATED LIKE THIS:

As a

..... (type of user/persona/role),

I want

..... (action),

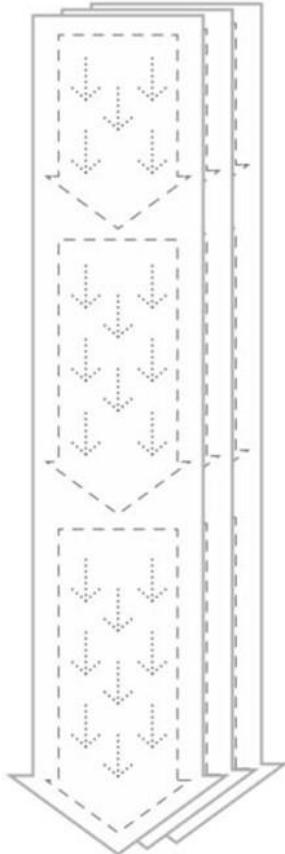
so that

..... (outcome).

An example User Story

Writing User Stories

What does this involve?



"As a [persona],
I want to [do something]
so that I can [realize a reward]"

Who is this user?
What makes them tick?
Who's an example of such a person?

Why do they want to do this?
What's the benefit/reward?
How will we know if it's working?

An example User Story

Image source: ?

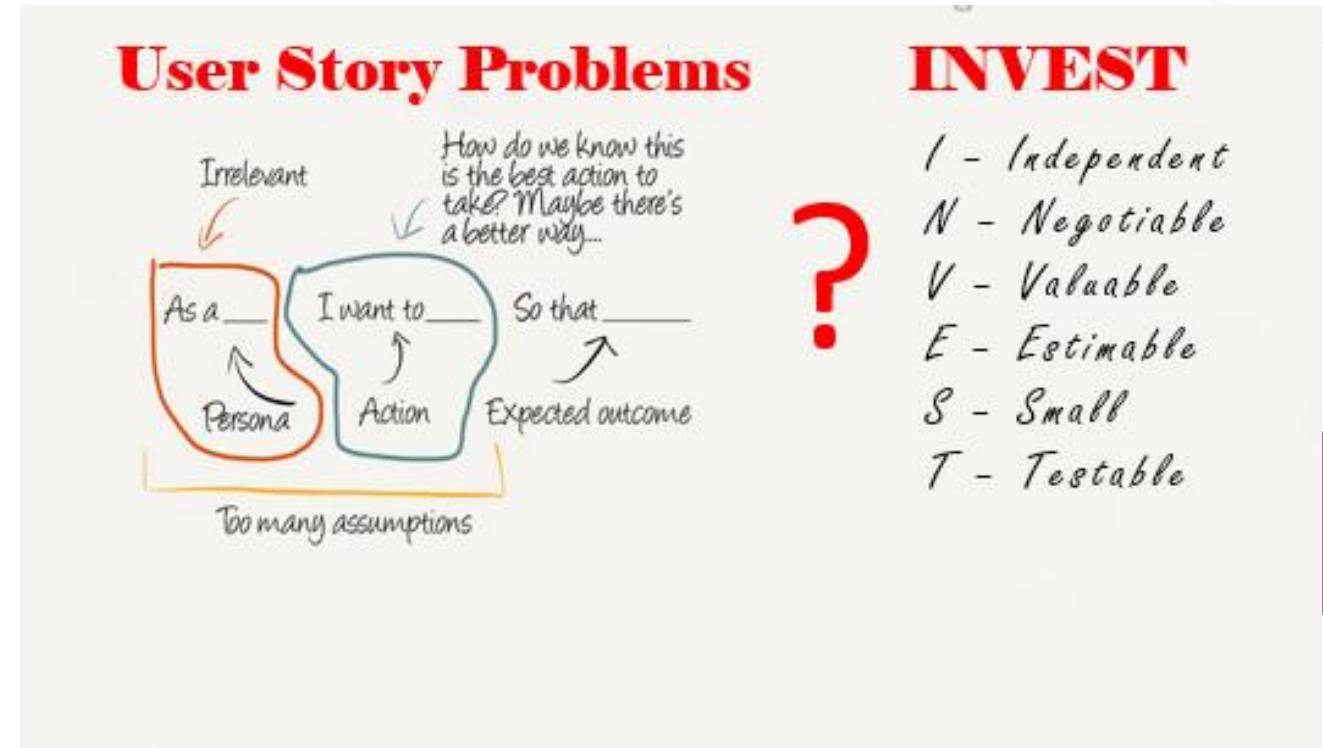
'Invest' User Stories

Defining well-written user stories

"Invest" is an acronym used to define well-written user stories in Agile development.

Each letter in "INVEST" represents a key attribute that user stories should possess:

- **I** - Independent
- **N** - Negotiable
- **V** - Valuable
- **E** - Estimable
- **S** - Small
- **T** - Testable



An example User Story

Image source: LinkedIn

Understanding Epic User Stories

Used to group related stories

Epic Criteria	Stories	Acceptance
As a marketing lead , I want to have content management system so that I can manage and provide quality content and experience to my readers.	As a Content Owner , I want to be able to create product content so that I can provide information and market to customers	Ensure the Content Owner is able to: log in to the content management system create a page of content edit / update an existing page of content save changes assign content page to Editor for review
	As an Editor , I want to review content before it is published so that I can assure it is optimized with correct grammar and tone.	Ensure the Editor is able to: log in to the content management system view existing content page edit / update page of content add markup comments- save changes save changes re-assign to Content Owner to make updates schedule content publish

An example Epic

Evaluating User Stories

Group discussion

What is wrong with these user stories?

Consider:

- Specificity
- Acceptance criteria
- User perspective
- Objectives



Group discussion



Exercise: What's wrong with these stories?

story one

As a customer service rep, I
need an interface so that I can
access customer data.

story two

As a user, I need to administer
accounts so that I can control
account access.

Example user story

User Story Checklist

When writing your checklists you may wish to consider:

-  **Keeping them short**
-  **Keeping them simple**
-  **Writing from the perspective of the user**
-  **Make the value/benefit of the story clear – what is the reason for the story?**
-  **Describe one piece of functionality. If you have to write and break it into two stories**
-  **Use stories as a team**
-  **Use acceptance criteria to show an MVP**



Definition of Done (DoD)

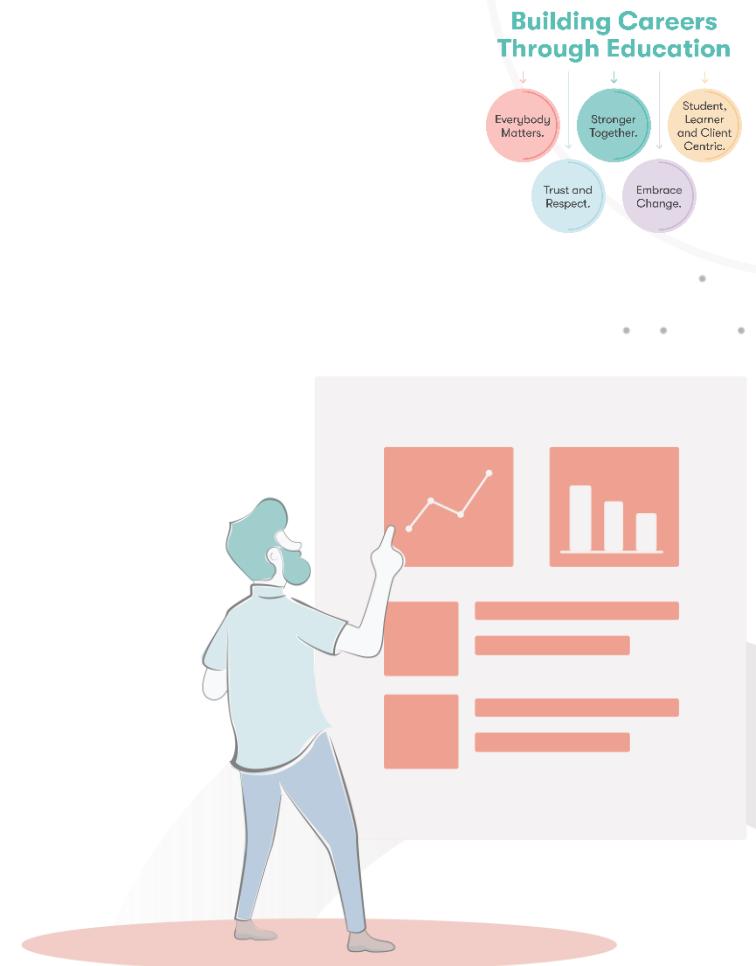
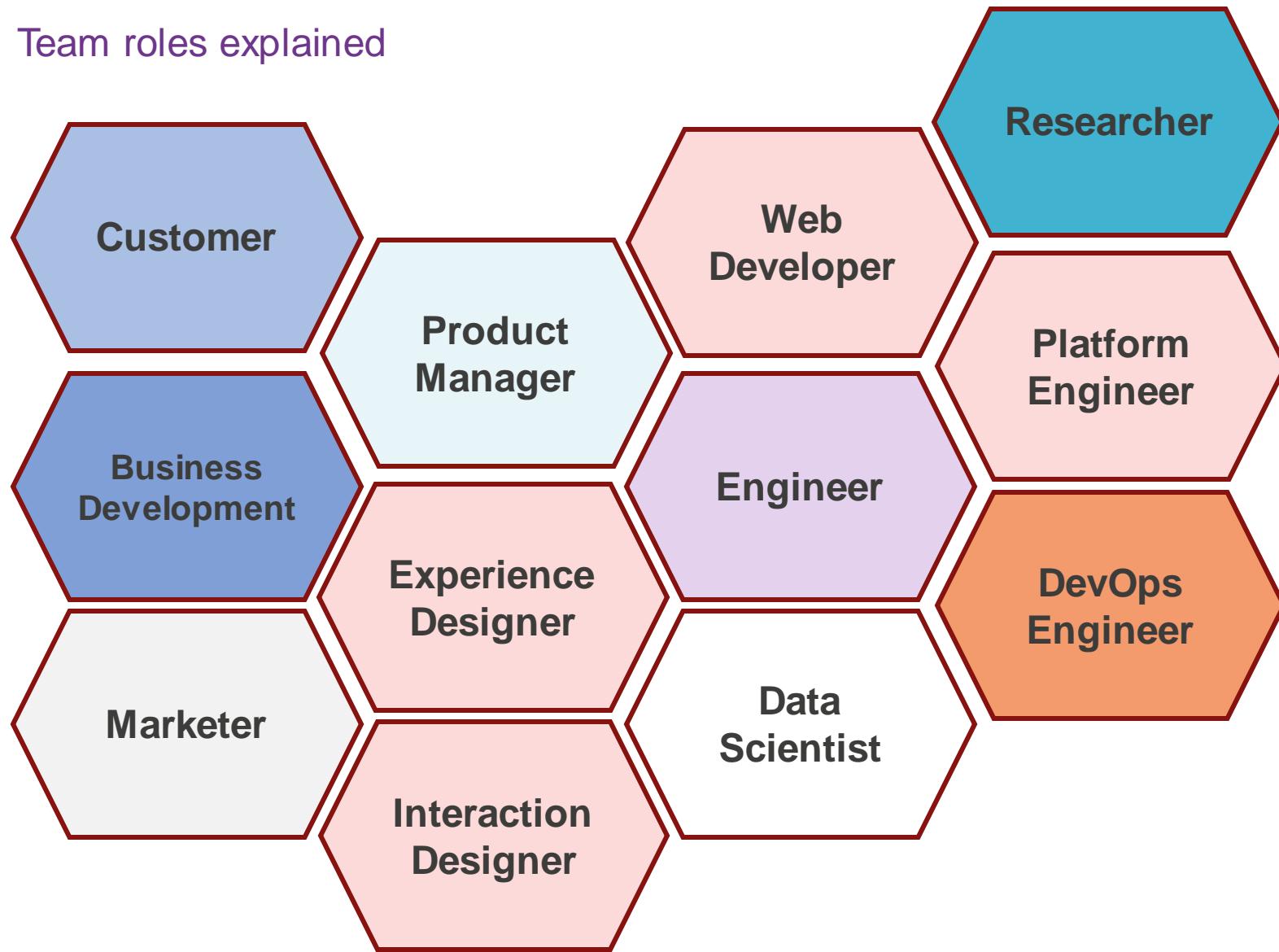
What criteria is typically used in Data Engineering?

Criteria	Criteria (cont)
Produced code for presumed functionalities	QA performed & issues resolved
Assumptions of User Story met	Feature is tested against acceptance criteria
Project builds without errors	Feature ok-ed by Product Owner
Unit tests written and passing	Refactoring completed
Project deployed on the test environment identical to production platform	Any configuration or build changes documented
Tests on devices/browsers listed in the project assumptions passed	Documentation updated
Feature ok-ed by UX designer	Peer Code Review performed



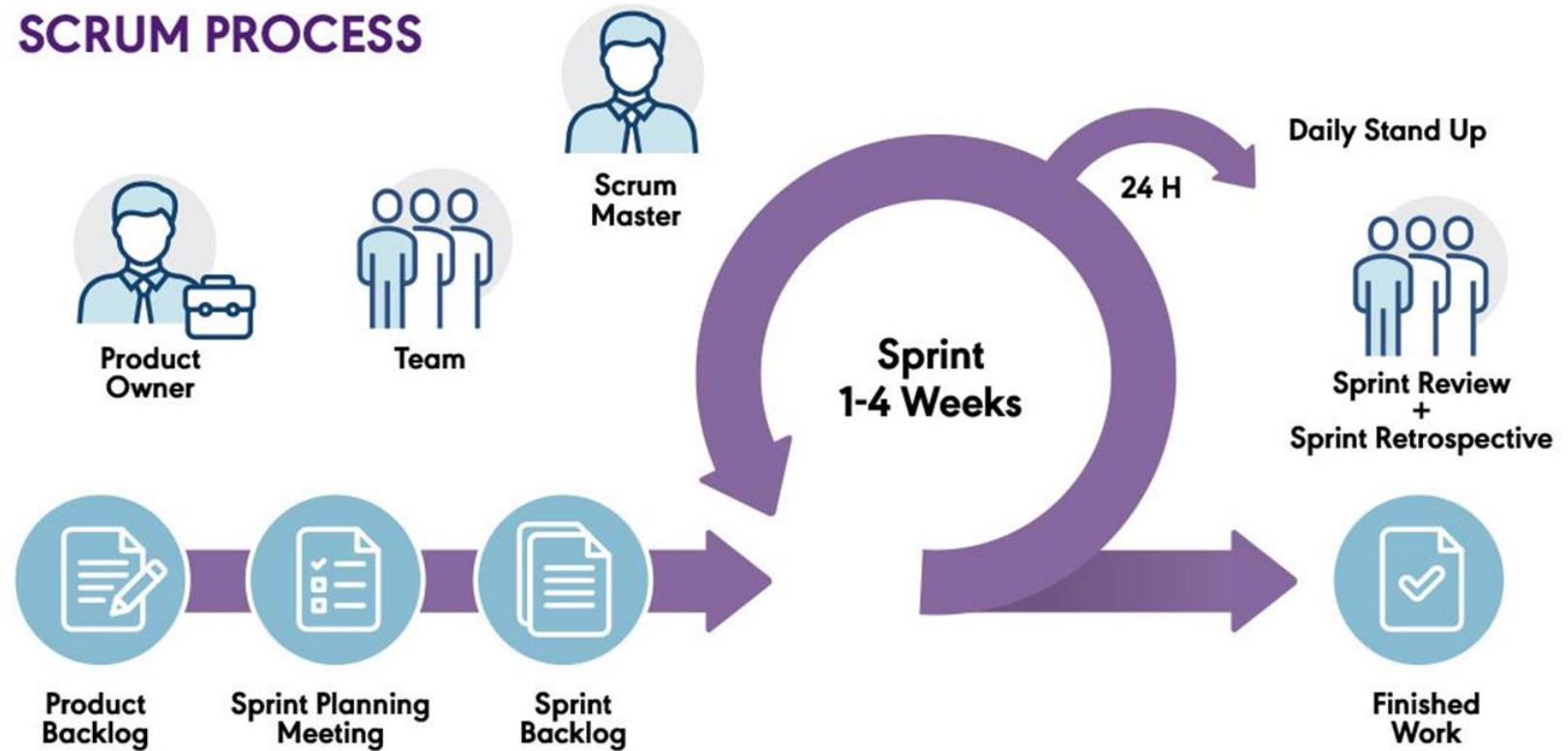
Team roles in data science

Team roles explained



Agile Scrum

A summary of roles, responsibilities and processes



The scrum process from start to finish

Image source: PM Partners

Agile backlog for Credit Bank Corporation

Collaborate activity 1

Your brief is as follows:

Remember Credit Bank Corporation wants to gain better insights into employee performance, satisfaction, and internal HR processes.

The goals are to improve workforce productivity, retention, and efficiency.

1. As a group, identify the key epics (large user stories) for the HR analytics dashboard project
2. For each epic, define a set of user stories, along with their acceptance criteria

Note: Example Epics and User Stories are provided in the collaborate brief resource for this activity on the Hub.



Case study: Credit Bank Corporation

Take a break!



Integrating Lean and Six Sigma Principles



Section Introduction

The Lean (Six-Sigma) framework

- Inspired by Ford, invented in the Toyota factory
- Focused on the elimination of all waste
- Toyota identified types of waste and wastefulness
 - Defects
 - Overproduction
 - Useless waiting
 - Underutilised talent
 - Excess processing



Toyota worked to eliminate the above wastage by applying:

- “Automation with a human touch” (*Jidoka*)
- “Making only what is needed and the amount needed” (JIT)



The Six-Sigma DMAIC process

Lean pillars

What are they?

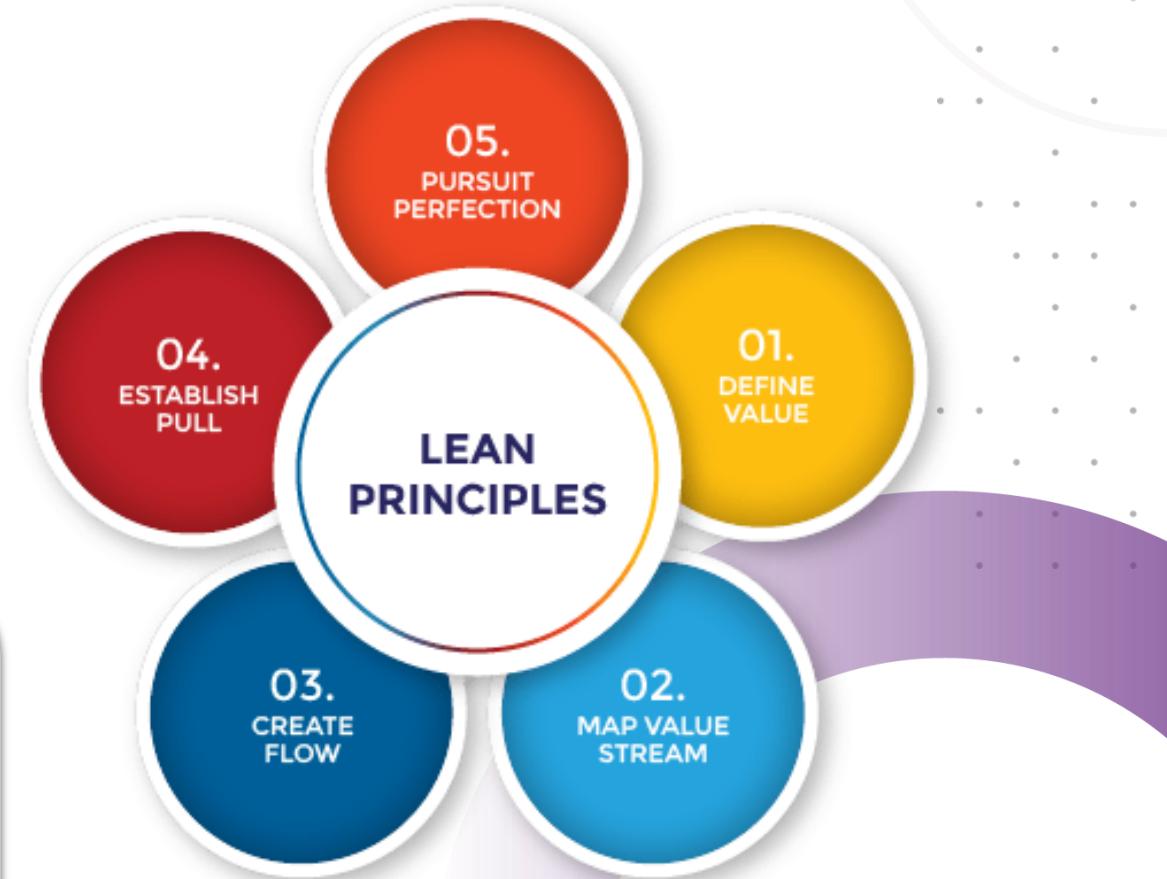
The Lean methodology is built upon a foundation of five core principles.

These are often referred to as "Lean pillars," which serve as guiding principles for organisations seeking to implement Lean practices.

What you need to know:



- Lean pillars guide organisations to identify and eliminate waste
- They focus on improving value delivery and customer satisfaction
- Embracing Lean principles enhances efficiency, quality, and competitiveness



The five pillars of Lean

Image source: The Lean Way

Lean software development

Optimising efficiency and minimising waste

- Lean software development aims to optimise efficiency and minimise waste in the software development process
- Dr. Robert Charette introduced the concept in 1993, focusing on principles adapted from lean manufacturing
- In 2003, Charette outlined the "12 Principles of Lean Software Development" to provide a structured framework for applying lean principles in software development

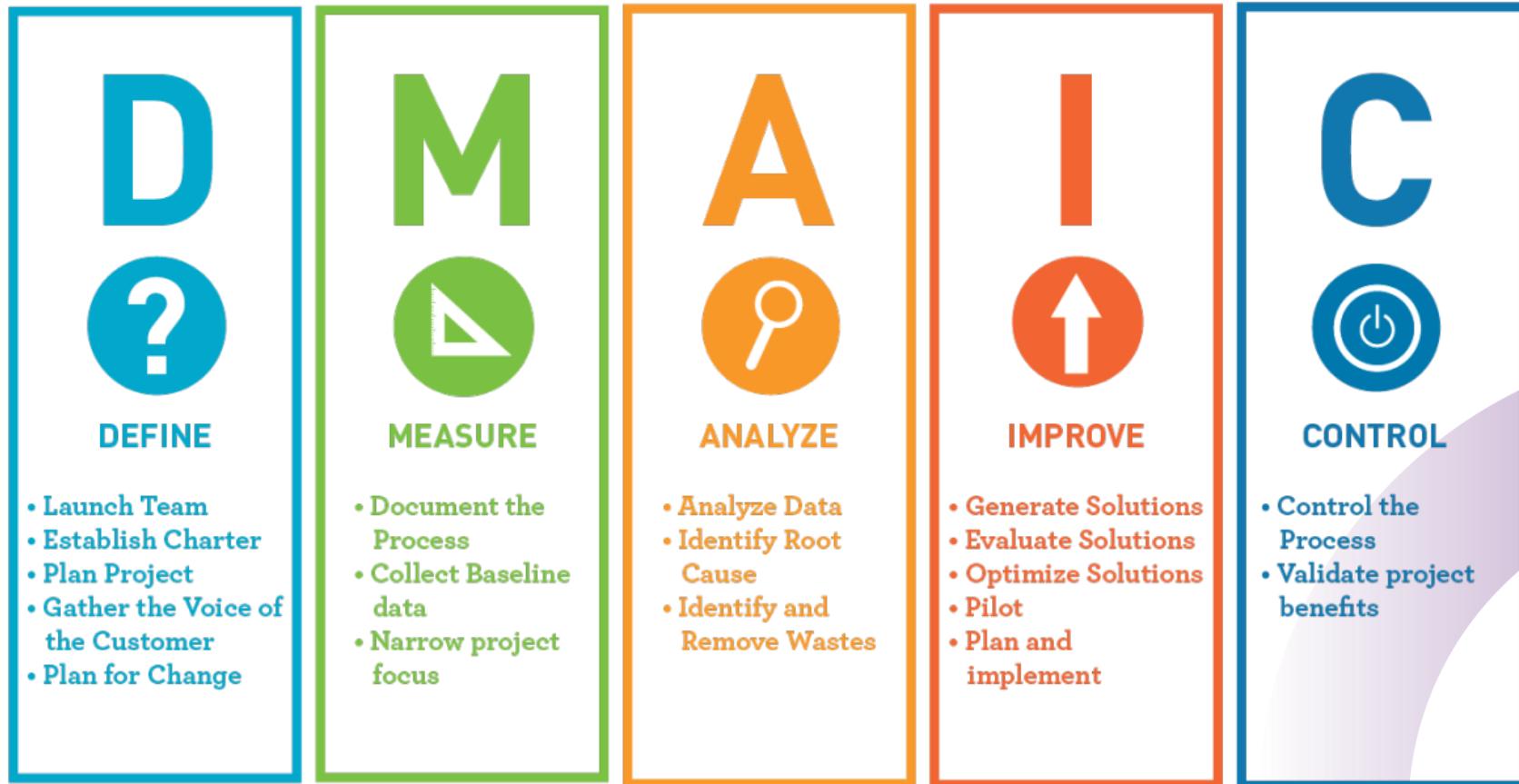
7 LEAN SOFTWARE DEVELOPMENT PRINCIPLES



The seven principles of Lean software development

Lean + Six Sigma = DMAIC

Combining the principles of Lean and Six Sigma methodologies



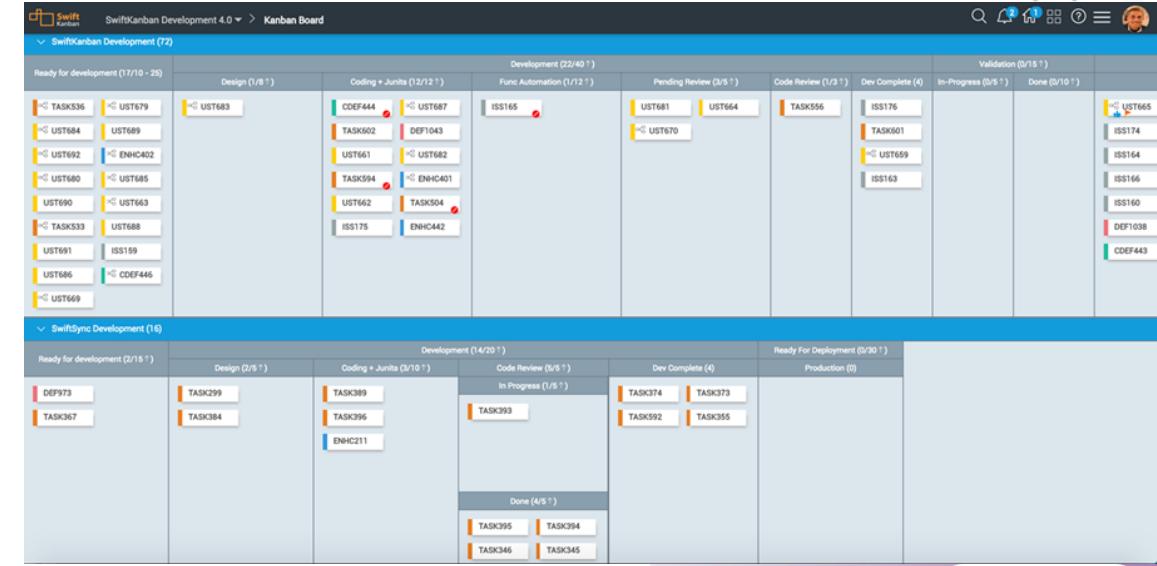
Define, Measure, Analyse, Improve, Control (DMAIC)

Image source: LinkedIn

Integrating Agile and DMAIC with Kanban

How does this work?

- Not a methodology, more of a technique
 - A linear manufacturing-line approach
- Recommended for maintenance work such as DevOps
 - It also projects that require event-driven responses
- Scrum can be practiced with Kanban (ScrumBan)
 - Makes the planning board more detailed
 - Unlike pure Scrum, we try to measure time (not points)
 - We fill in vacant slots in the board using the “pull” approach



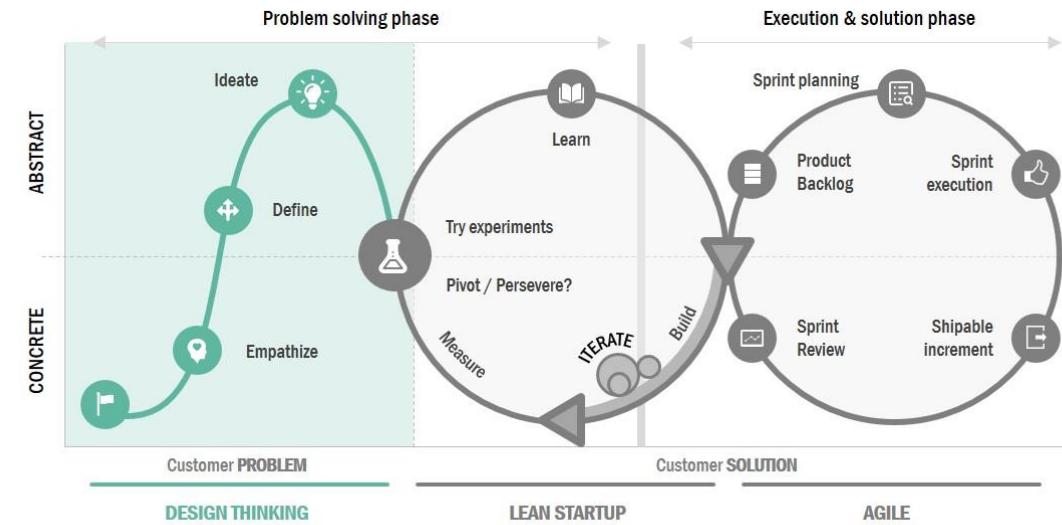
Kanban example

Agile vs Lean

How do they compare?

Metric	Lean	Agile
Obsessed with	Waste	Users
Manages	Process	Uncertainty
Delivers	Improvement	Working Product
Applies	Heuristics	Principles
Requirements	Hypotheses	User stories
Approach	Test and Learn	Sprint
Delivers	MVP	Working versions

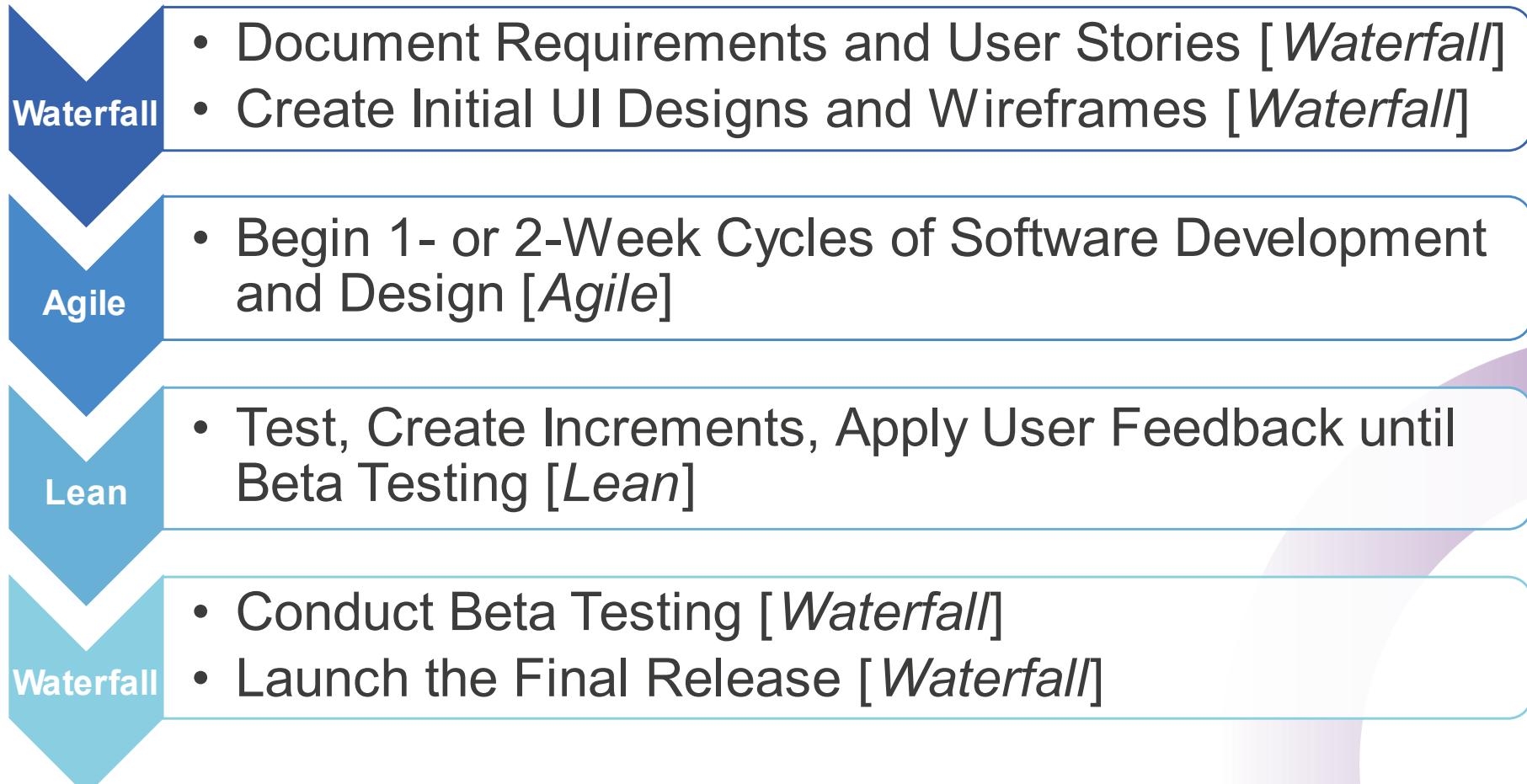
Design Thinking + Lean Startup + Agile Diagram



Combining Lean and Agile methodologies enhances product development

Combining Lean and Agile methodologies

A hybrid approach



A hybrid approach to combining Lean and Agile methodologies

Minimum Viable Product (MVP)

What you need to know:

- MVP is a strategy for delivering a product with the minimum set of features to satisfy early customers and gather feedback
- Used to quickly validate assumptions, test hypotheses, and gather feedback
- Reduce time to market, minimise costs, and mitigate risks



Minimum

Rudimentary,
Bare-bones
but functional



Viable

Something
you can
produce and
customers
will want



Product

MVP =
minimum
viable
product

The characteristics of MVP

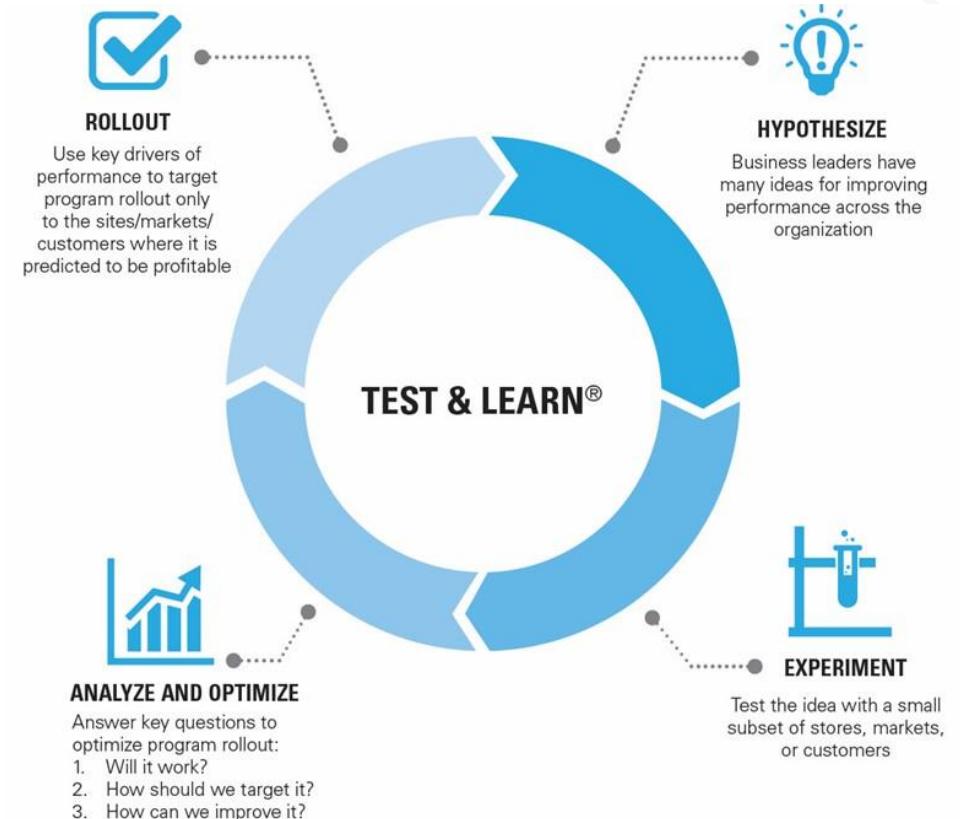
Understanding 'test and learn'

In the context of Agile and MVP

In the context of Agile working practices and MVP, the concept of "test and learn" emphasises a continuous cycle of testing hypotheses, gathering data, and learning from the results to inform decision-making and product development.



Test and learn practices in Agile data working enable teams to make informed decisions, reduce uncertainty, and deliver value to users efficiently.



The test and learn cycle

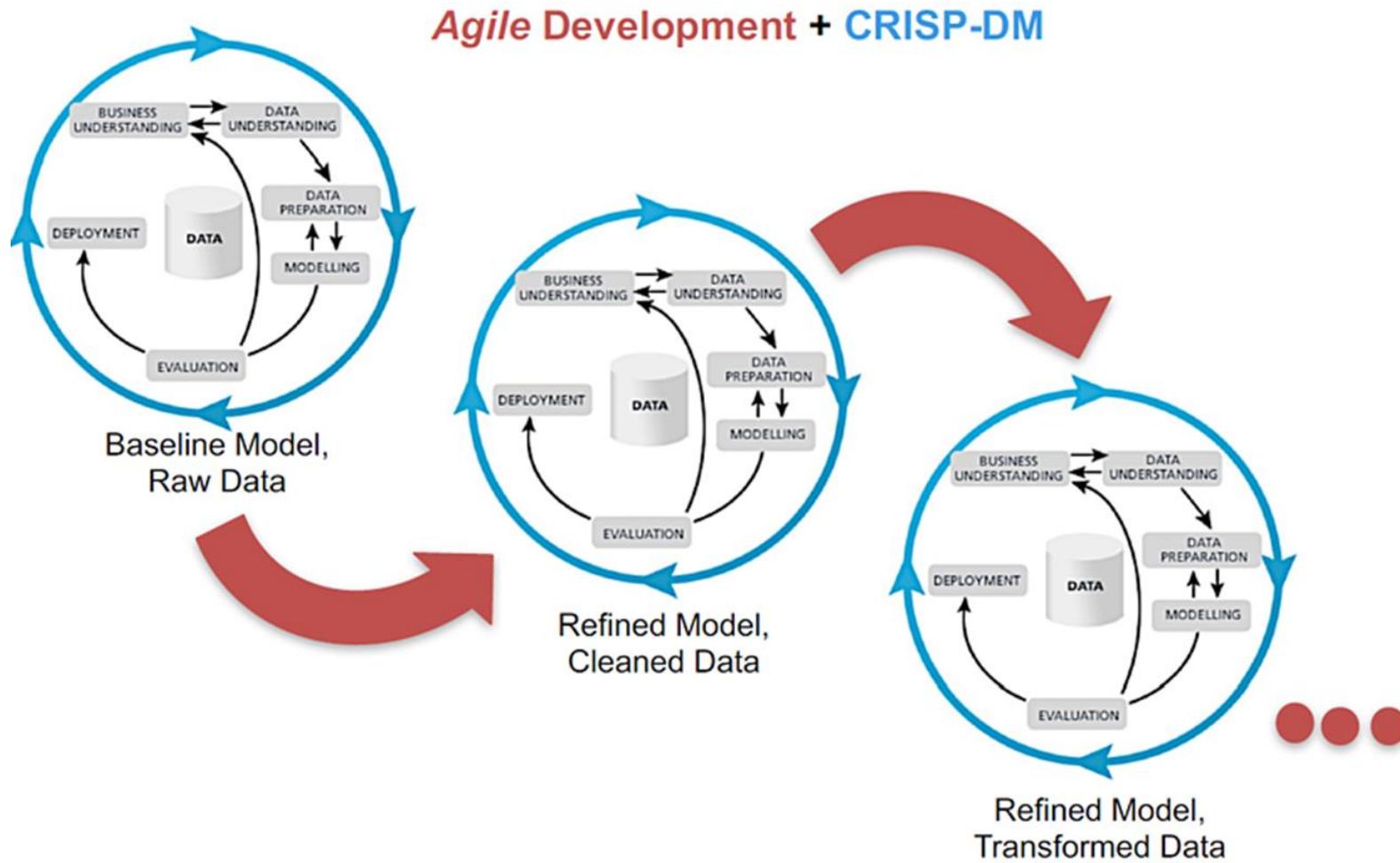
Image source: ?

Guidance for working with data scientists



Collaborative data workflow

Navigating integration with data scientists withing CRISP-DM



The Crisp-DM Framework

Image source: LinkedIn

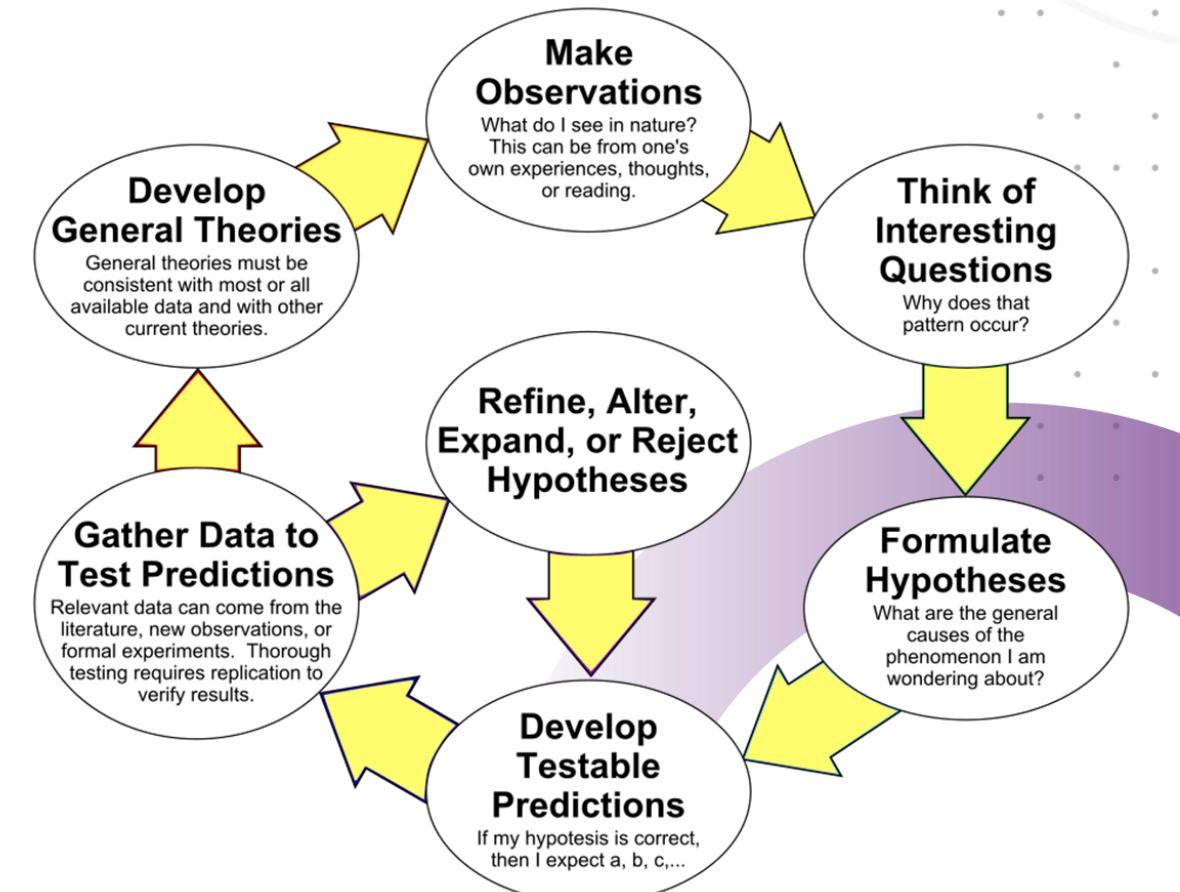
Integration of Scientific Method in data projects

Aligning sprint activities with the scientific method

When data engineers collaborate with data scientists within the context of Agile sprint schedules, they can expect the steps of the scientific method to manifest in various stages of the project lifecycle.



Scientific method integration in Agile projects enhances collaboration for informed decision-making.



Steps of the scientific method

Image source: ?

Take a break!



Practical lab



Sprint planning for Credit Bank Corporation

Collaborate activity 2

In this activity, you'll work as a team to plan a sprint for the development of the HR analytics dashboard at Credit Bank Corporation.

Instructions:

1. Log into Jira using the provided credentials
2. Assign the Scrum roles within your team
3. Review the Agile backlog from the previous activity and identify the user stories for this sprint
4. Create the user stories in Jira, ensuring they meet the 'INVEST' criteria
5. Estimate the effort required for each user story and assign story points
6. Start a new sprint in Jira and add the selected user stories
7. Assign the user stories to your team members
8. Review the sprint goal, timeline, and any potential risks or dependencies.
9. Finalise the sprint plan and be prepared to present it



Apply hands-on exercise

Your homework (on-the-job)

Championing a data-driven culture

1. Identify key stakeholders

- Key business leaders, IT, data analysts, end-users

2. Analyse effective stakeholder collaboration strategies

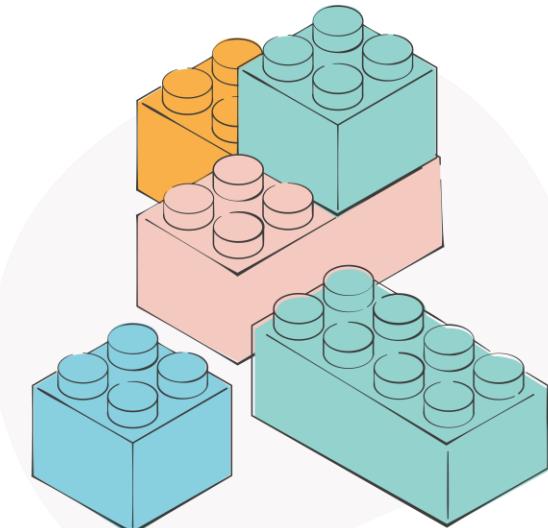
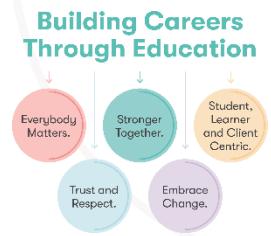
- Align data initiatives with business objectives
- Document in a stakeholder engagement plan

3. Examine synergies between Agile, Lean, and Six Sigma

- How can these methodologies be integrated?
- Develop a proposal for championing a data-driven culture

4. Secure buy-in from leadership and teams

- Outline specific Agile, Lean, Six Sigma practices
- Define success metrics and KPIs



Apply exercise

Apply hands-on exercise

Your homework (off-the-job)

Imagine you are the Data Protection Officer (DPO) at Cosmetics International Limited, London

- They have a new makeup line: Shine-N-B-Fine with variable colour palette
- Initial sales are below forecasts

Marketing meeting outcome:

- Sales team recommends geographical variations for marketing

Epic: Shine-N-B-Fine soft launch (100 points)



"I want: an analytics dashboard to classify Shine-N-B-Fine user numbers geographically, so that: I can optimise marketing plan for the product "

- The VP of Marketing at Shine-N-B-Fine

The VP Marketing has added this user story to the Jira Backlog

Apply hands-on exercise

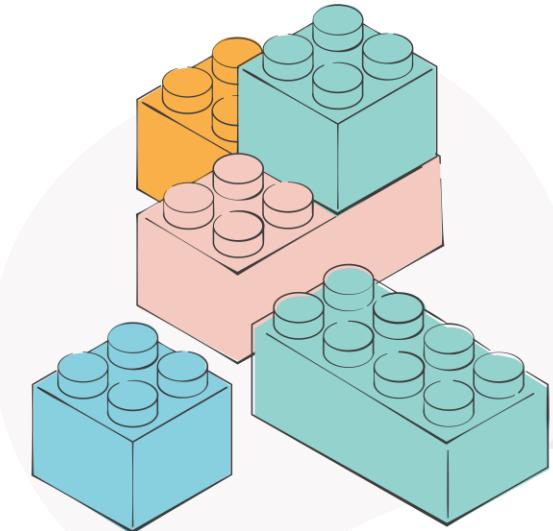
Your homework (off-the-job)

Based on this situation, the VP has added a user story to your Jira backlog.

A rumour ensued that this user story may cause data protection issues.

You are called in to a backlog refinement meeting to answer the following questions:

- Will any GDPR restrictions apply?
- What rights can the data subjects exercise? Enumerate and explain in detail.
- Is a DPIA required? Justify why or why not.
- What other recommendations do you have?
- Do you include the story in this sprint?
- Do you suggest any additional acceptance criteria?
- Review the User Story points



Apply exercise

Key Learning Summary

The key takeaways from this session are as follows:

- Waterfall project management struggles with rapidly changing requirements, lengthy development cycles, lack of stakeholder feedback, and challenges with data complexity
- Agile principles emphasise adaptability, iterative development, and close collaboration
- Scrum is a popular Agile framework that introduces specific roles (Scrum Master, Product Owner, Team Members) and ceremonies (Sprint Planning, Daily Standup, Sprint Review, Retrospectives)
- The Scrum Master embodies servant leadership and is responsible for facilitating, improving, protecting, rewarding, and educating the Scrum team
- The Lean methodology is built upon five core principles: Continuous Improvement, "Kaizen", "Genchi Genbutsu", Respect, and Teamwork
- The DMAIC (Define, Measure, Analyse, Improve, Control) framework combines Lean and Six Sigma principles
- The Minimum Viable Product (MVP) is a strategy for delivering the minimum set of features to validate assumptions and gather feedback





Thank you

**Do you have any questions,
comments, or feedback?**

