





Data Analysis in Industry

Session 1



Session Outline

The Data Lifecycle

The Data Analytics Lifecycle

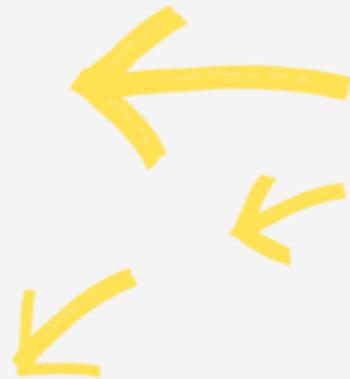
Identifying Stakeholders

Dealing with Stakeholders

Methods of Communication

Project Briefs

Recap

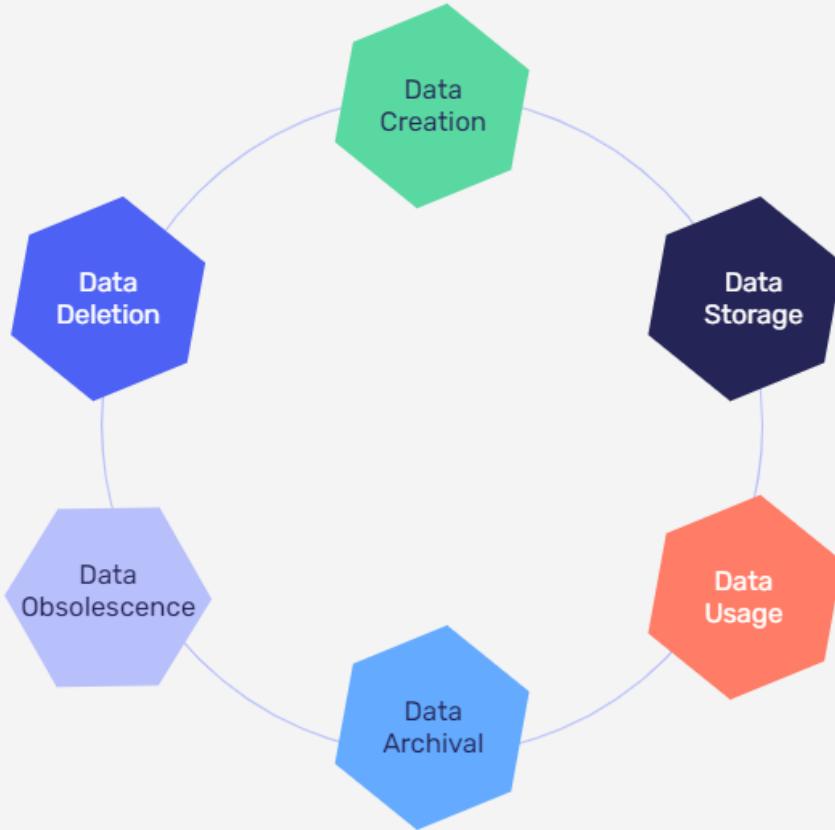


Learning Objectives



- Describe the **data analytics lifecycle** and set out details, activities and roles at each stage in a real-world example
- Design effective **project briefs** which **communicate expectations to stakeholders**
- Identify **business specific rules** related to datasets and data characteristics that will influence project design and analysis

The Data Lifecycle



Data LifeCycle

Is the data structured or unstructured?

Is it quantitative or qualitative?

Where has the data come from?

What size is the dataset?

Is it created continuously or in batches?

Data
Creation

Data LifeCycle

Where is the data stored?

How many people require access?

What are the access requirements?

What rules relate to the data storage?



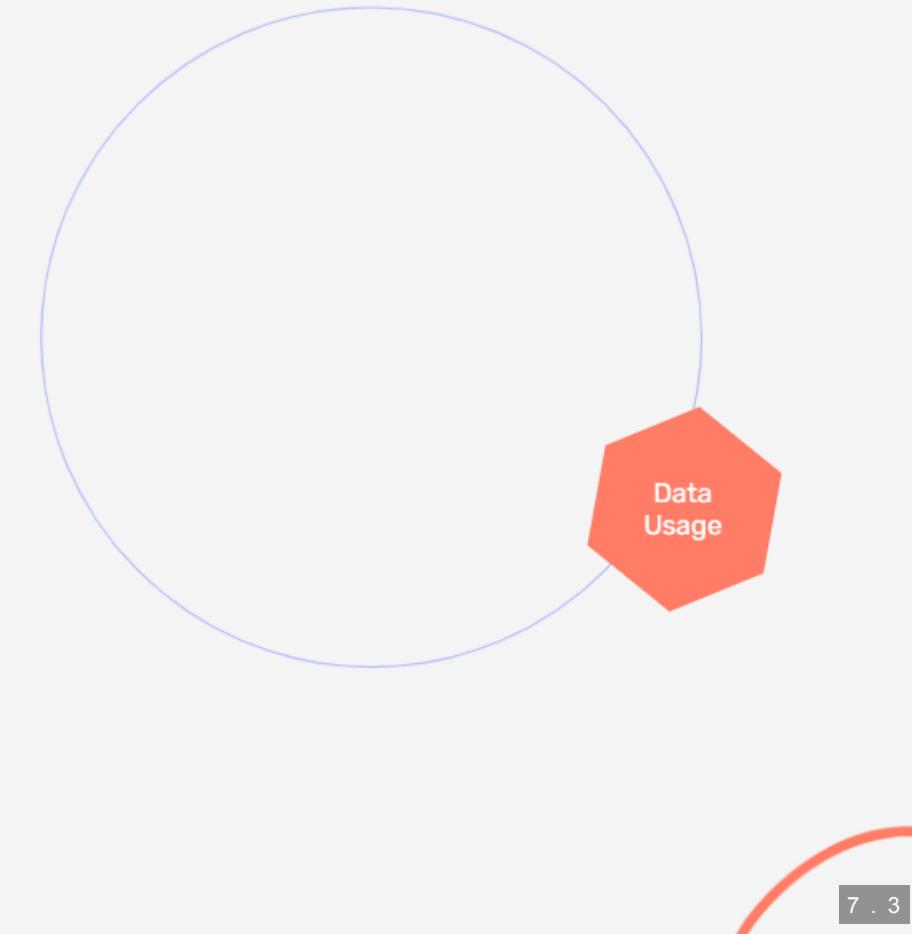
Data LifeCycle

How is the data being used?

How often is it being used?

Does anyone external to your organisation have access?

Are there any business rules?



Data LifeCycle

Is the data always needed?

What is the working life of the data?

Can it be made available given prior notice?



Data LifeCycle

Is the data still being used?

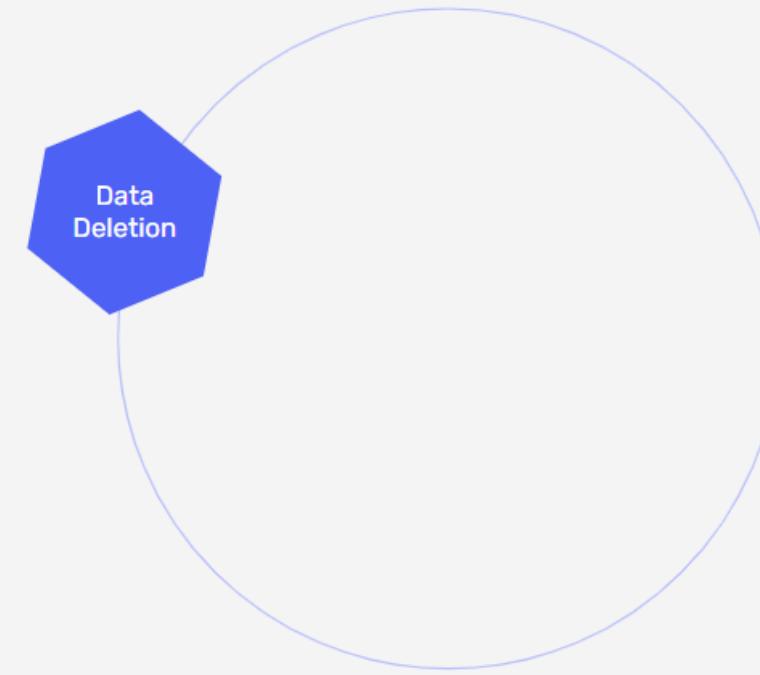
Is the data relevant?

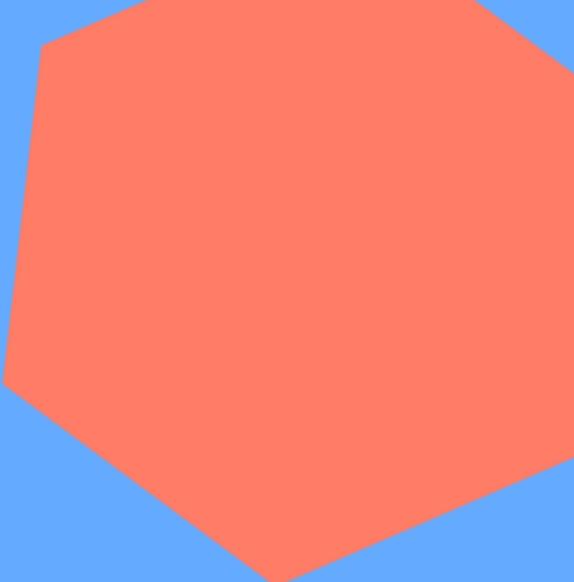
Is the data needed?

Data
Obsolescence

Data LifeCycle

How can the data be securely and permanently removed?

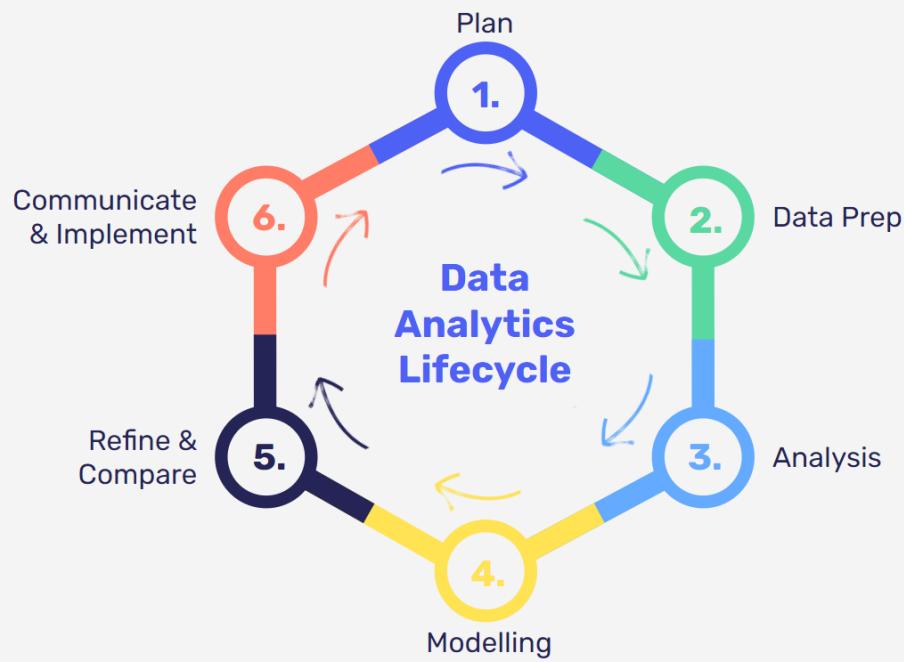




Activity

In pairs, walk your partner through the data life cycle for a piece of data you have used in your role.

The Data Analytics Lifecycle





What is the context of business problem?

Why does the problem matter?

Who are the stakeholders?

What assumptions have you made about the problem?

What is your proposed solution and issues you have considered?

What data will you require?



Does the data require cleaning?

How have you summarised the data?

Is there any missing data or outliers?

What are the data structures?

What are the data types?

What features are in your data?



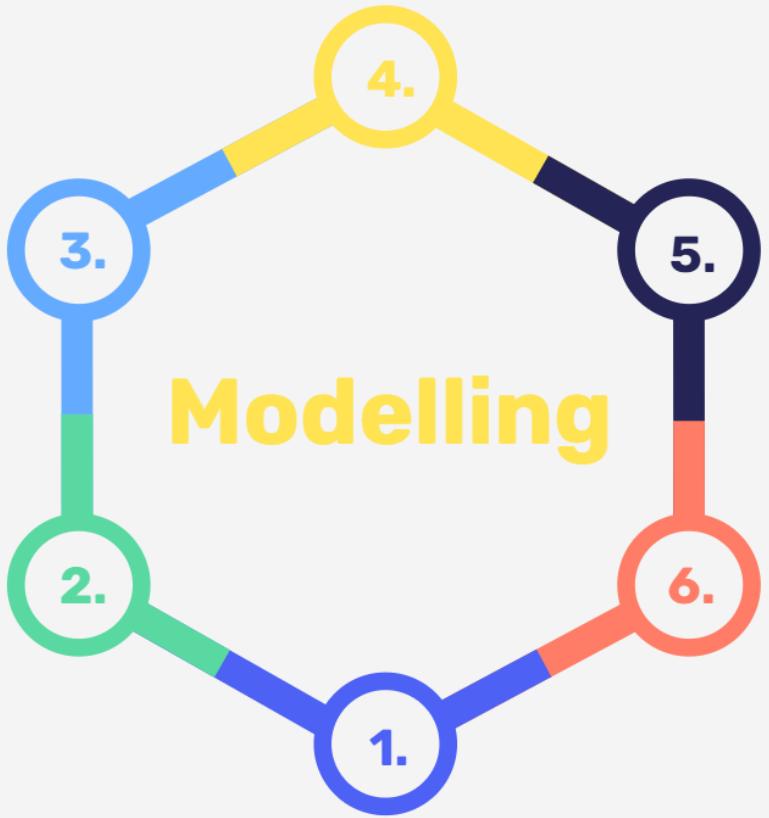
What aggregates have you calculated?

How have you visualised the data?

What trends have you observed?

Are there any correlations between features?

What insights have you gained from the data?



What is your solution to the business problem?

Is it a prototype?

Based off your analysis, how have you come to this solution?



Have you considered other solutions?

How do you know yours is the best one?

How have you optimised your solution?

How have you verified that the project requirements have been met?



How have you presented your solution to the stakeholders?

What are your recommendations?

What are the next steps or future opportunities gained from this project?



Activity

You work for a recruiting firm and have been asked to produce a model that predicts the salary of a data analyst from a job board website. The problem is that not all jobs advertised list a salary range, so your line manager wants you to find a way of estimating these salaries based off features such as location, job title and keywords in the description.

In groups identify where each of the steps on the following slide would fall into in the data analytics lifecycle.

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

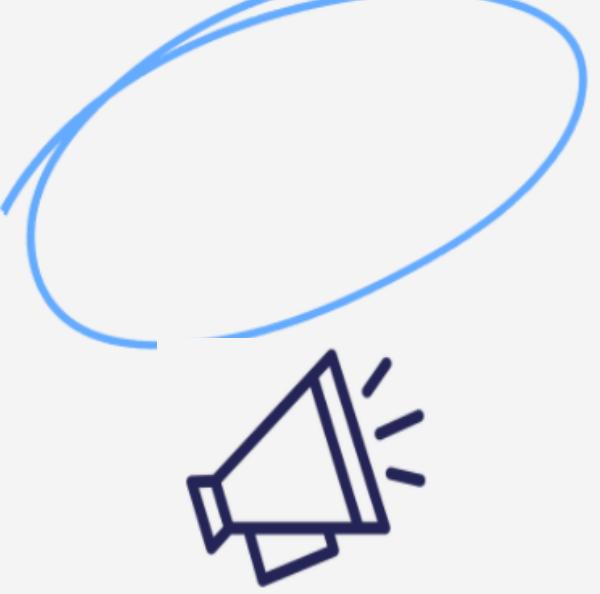
Ensure location data is consistent	Establish project deadline	Build predictive model
Fine tune model to optimise results	Move model to production	Understand stakeholder expectations
Extract keywords from job descriptions	Remove currency symbols from data	Identify adverts which list salaries
Plot visualisation of average salary per location	Communicate results to stakeholders	Aggregate mean salary per location
Extract location information from job adverts	Interpret results	Decide whether you are predicting continuous salary values or salary bands

Identifying Stakeholders

What is a stakeholder?

**Who could be a
stakeholder?**

Stakeholders are individuals or groups who are involved in your project from whom you require input, or who will be influenced or impacted by the results.



Sponsor

Examples: Directors, Senior Leadership, Heads of Department

Team

Examples: Project Managers, Team Leads, Line Managers, Sales Teams

Users

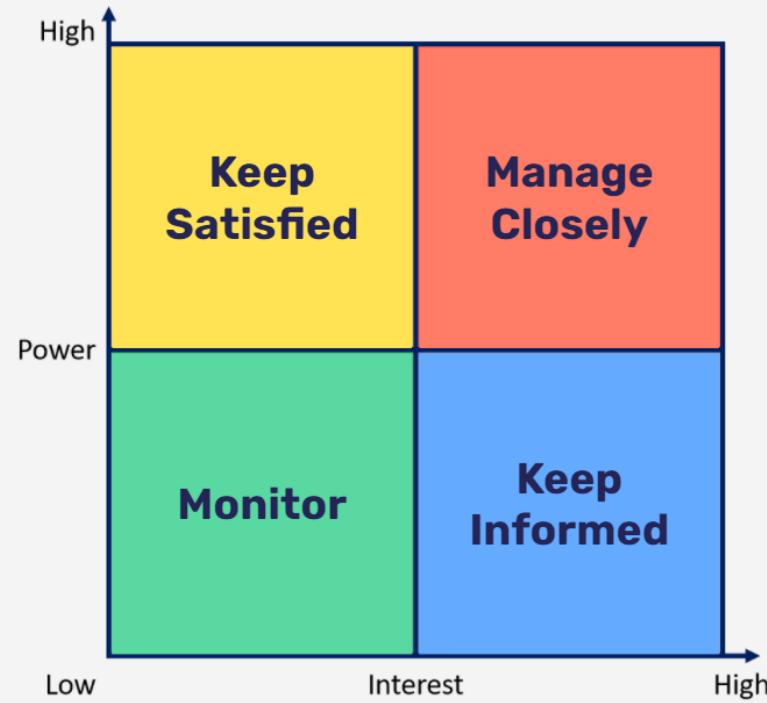
Examples: Customers, Clients, Business Owner

Suppliers

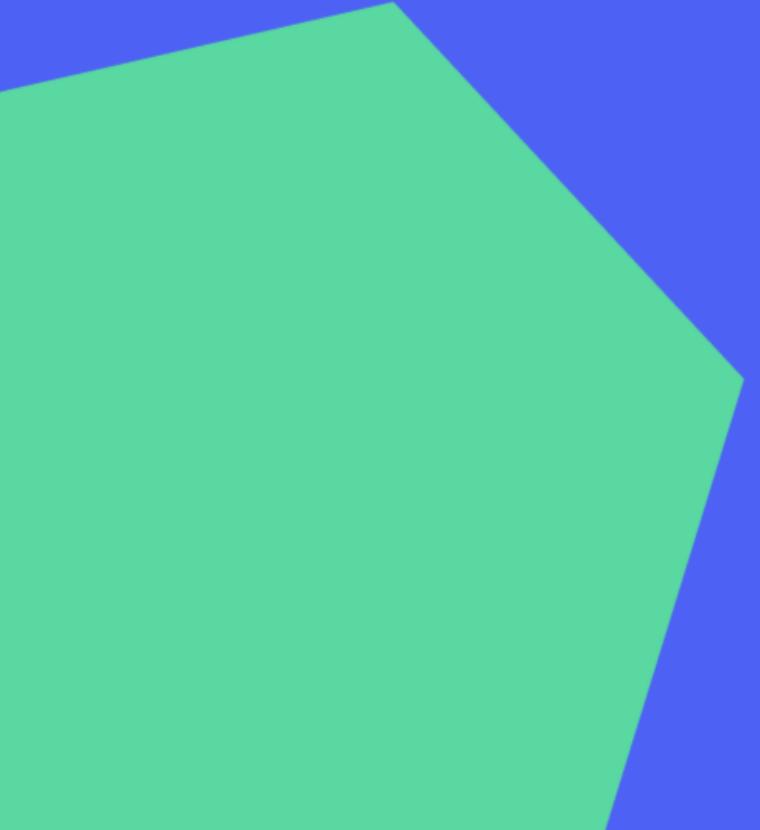
Examples: Contractors, Consultants, Subject Matter Experts

1. What motivates them?
2. How can you influence them?
3. What information would they like to receive?
4. How would they like to receive it?
5. What information do we require from them?
6. What dependencies are they responsible for?

Questions to consider in
your communications with
each stakeholder



Dealing with Stakeholders



Do you have enough information to draft a project brief to share with stakeholders?

Are you accurately capturing the necessary information through effective questioning?

A good question is one that lets
you obtain the **type, quality** and
quantity of the information you
need

**Open ended questions
encourage people to reflect and
reveal what is important to
them**

TYPE	DESIRED OUTCOME	EXAMPLE
Explorative	Expand on new points of view and uncovered areas	Have you thought of...?
Affective	Reveal stakeholder's feelings about something	How do you feel about...?
Reflective	Encourage elaboration	What do you think causes...?
Probing	Invite a deeper examination	Can you describe how...?
Analytical	Find the root of a problem	What are the causes of...?
Clarifying	Help align and avoid misunderstandings	You mean that...?

Other things to consider

What return do they get?

What do they need?

What do you need?

What's their opinion of you?

What is their strategy?



Activity

In groups you will be assigned a business case.

Brainstorm a list of questions you would ask to get more insight into the problem

Think who else might be a stakeholder in the situation

Scenario 1

Yusef owns an online shoe e-commerce website. He isn't happy with how sales are going, especially the beachwear section of his store
He has contracted you to create an analysis to help him with this issue.

I want you to look at the day-on-day sales to see where this is going wrong. I want moving averages, and a linear regression of the shopping baskets! I spent a lot of money buying 20,000 pairs of flip flops. They need to sell.

Scenario 2

Danielle runs a DJ store selling vinyl records and cassette tapes. Recently, she has found that she is running out of stock of certain albums and songs. Customers are asking for them but all her copies are gone! She's contracted you to help.

I've got a list of the ages of my customers. I want you to explore who the customers are, how old are they? What is their vibe like?

Scenario 3

You are the data analyst at an advertising firm and have just implemented a new strategy to boost the ROI of a client's advertising campaign, with the sign-off of both the account manager and your line-manager. However, later in the day, you receive an angry email from someone at the client company. It turns out they are temporarily in charge of the campaigns and don't like your changes, asking you to immediately revert them. It seems they are unaware of how you usually operate.

Communication Plan



To help keep track of stakeholders you should maintain a communication plan to help you:

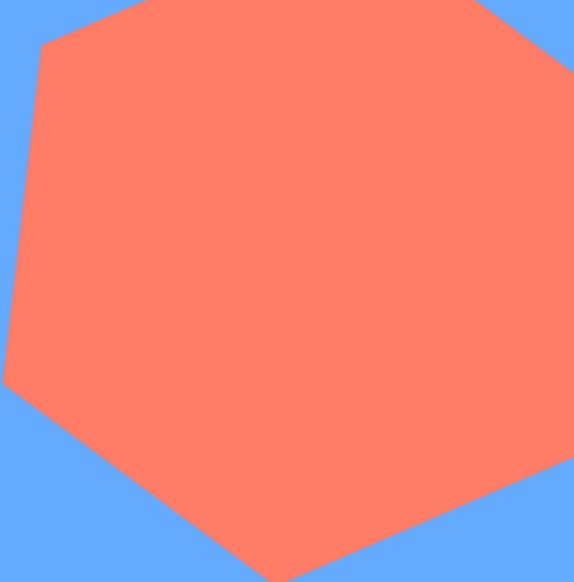
- Determine how often you will be in contact with stakeholders
- Decide on your messages (ignore, follow, inform, consult, involve, etc)
- Decide on the medium of contact (email, phone call, etc)
- Assign a responsible project member



#	Name	Role	Interest	Power	Status	Message	Medium	Frequency	Who
1	Jintana Roberts	Project sponsor	High	High		Inform, consult	Bilateral, Steering Cie	Bi-weekly	PM
2	Richard Shah	CFO	None	High		Involve	Bilateral Workgroup	Bi-weekly Ad hoc	PM
3	...								



Methods of Communication

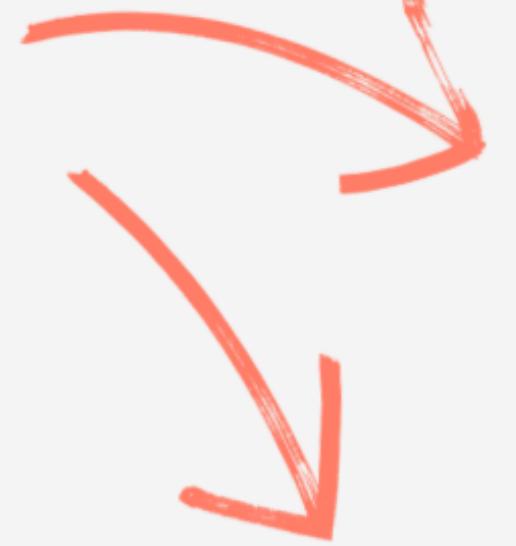


Activity

Communication problems lead to delays, misunderstandings, frustration, workplace conflicts and a mismatch in stakeholder expectations.

- What are some examples of effective communication methods?
- In which context and with what type of audience are those effective?

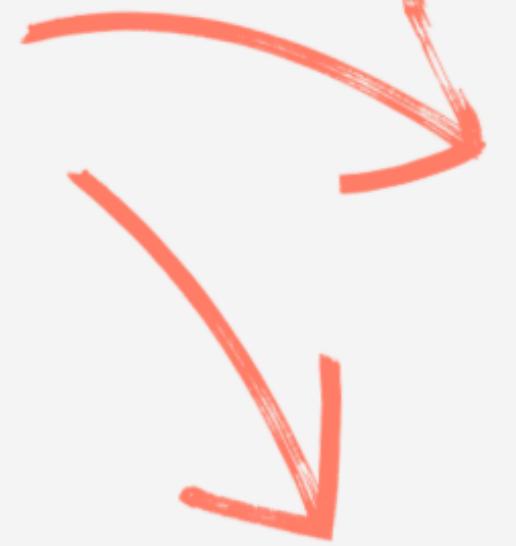
Two types of Communication



Two types of Communication

Verbal

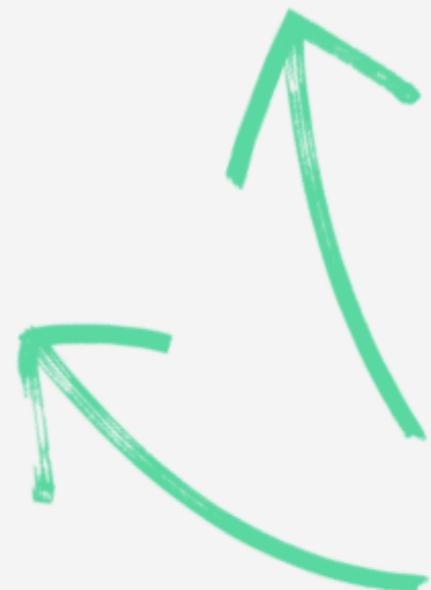
No Actions



Two types of Communication

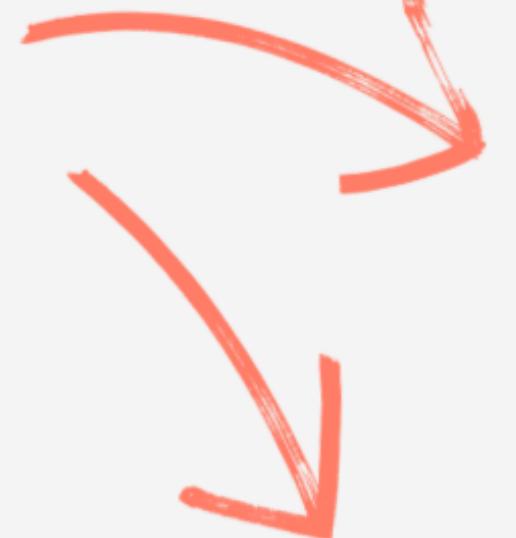
Verbal

No Actions

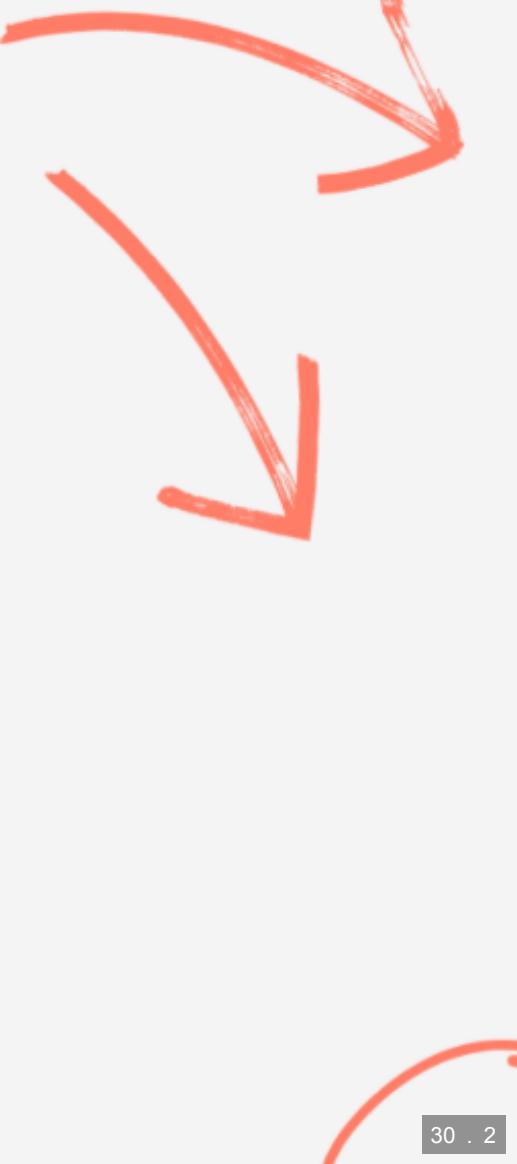
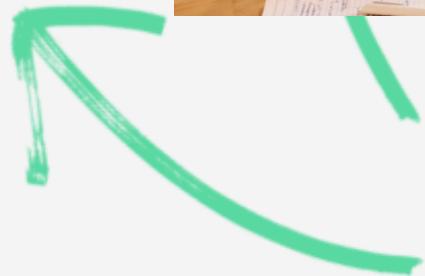


Non-Verbal

Actions



Email



Email



Face to Face



30 . 2





30 . 3

A good data analyst will
accurately **frame** a problem to
be solved and **share** it with key
stakeholders



Project Briefs

A project brief is the document
that sets out the requirements
for a piece of analysis



Who
What
When
Why



Project Sponsors

Those who set the analytical requirements and decides if your work has fulfilled their needs:

- What questions do you need to ask them?
- What do you need to produce for them?
- What format are they expecting?
- Do they understand the limitations of their data?



Metrics, Outputs and Format

What are the key pieces of information you need to agree on?

- What will you be delivering?
- What tools do you intend to use?
- What will the final product look like?
- How will you know that you have been successful?



Who

What

When

Why

Deadlines

What is the timeline for completion on your project?

- When will your project be completed ?
- What will be the check-in dates where you will give updates?
- How often do you intend to communicate with stakeholders?



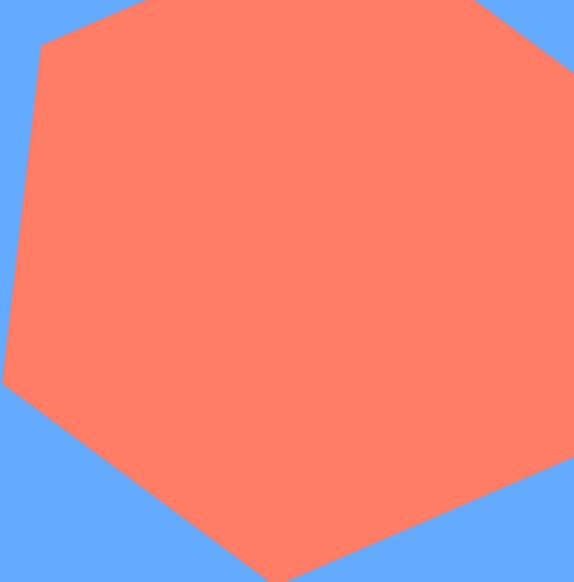
Who
What
When
Why

Context

Why does your project matter?

- What is the business problem?
- What will happen if the problem is left unchecked?
- How will your solution fix this problem?

**Does your industry have
additional requirements?**

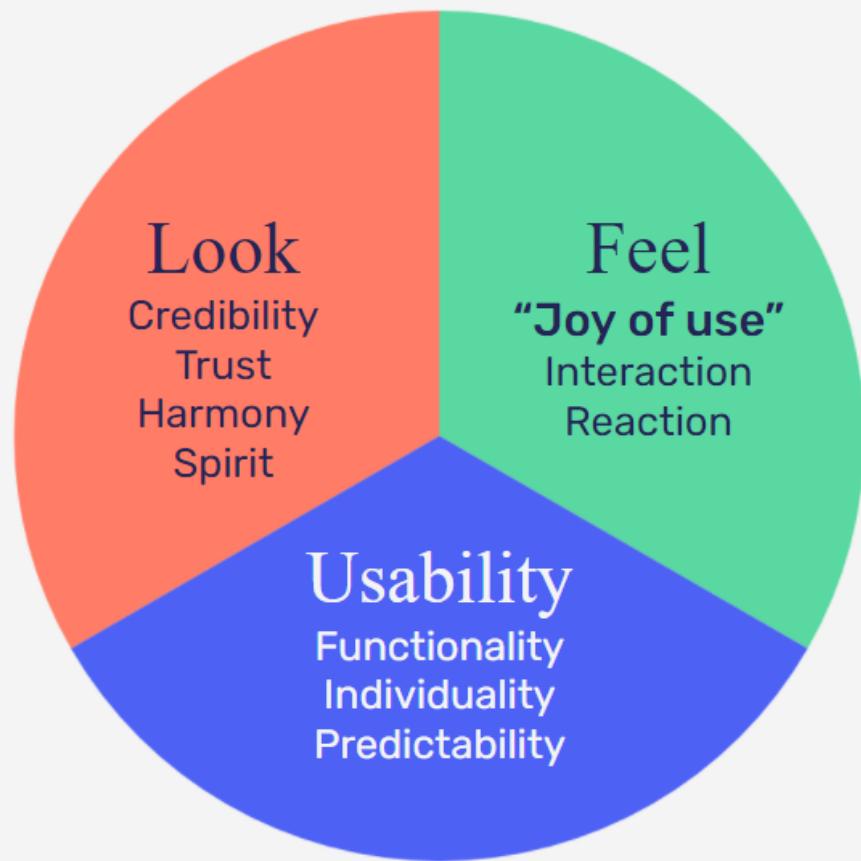


Activity

Review these project briefs

- What is good or bad and why?
- Is there any missing information?
- What would you always want to see in a brief?
- Who should supply the brief information?
- What other questions should you be asking of these briefs?

14



Domain Context



Domain Context

What are the industry requirements?



Domain Context

What are the industry requirements?

How does it fit into the wider business structure?



Domain Context

What are the industry requirements?

How does it fit into the wider business structure?

What will be the commercial impact?



Business Need for Analysis



Descriptive Analytics

Diagnostic Analytics

Predictive Analytics

Prescriptive Analytics

Looking at past data to tell what has happened

E.g. KPIs, revenue, sales leads



Descriptive Analytics

Diagnostic Analytics

Predictive Analytics

Prescriptive Analytics

Determining the cause of an observed event



Descriptive Analytics

Diagnostic Analytics

Predictive Analytics

Prescriptive Analytics

Predicting what will happen in the future



Descriptive Analytics

Diagnostic Analytics

Predictive Analytics

Prescriptive Analytics

Creating an action plan based off your analysis



Good Project Briefs:



Good Project Briefs:

Are concise and clear



Good Project Briefs:

Are concise and clear

Set time expectations clearly



Good Project Briefs:

Are concise and clear

Set time expectations clearly

Explicitly state the intended outputs



Good Project Briefs:

Are concise and clear

Set time expectations clearly

Explicitly state the intended outputs

State the scope and purpose of the analysis



Good Project Briefs:

Are concise and clear

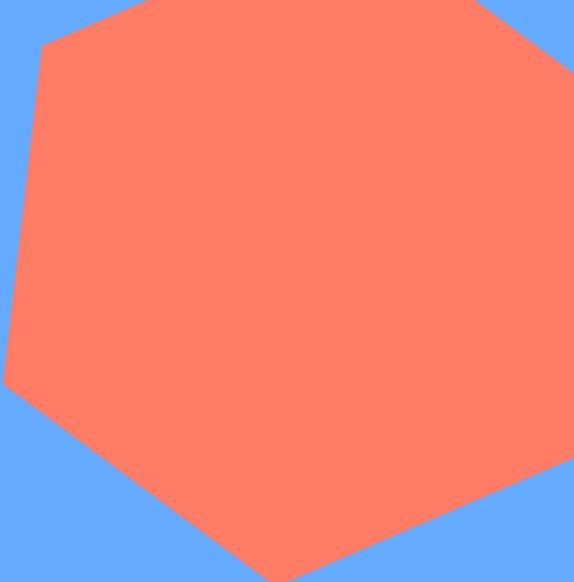
Set time expectations clearly

Explicitly state the intended outputs

State the scope and purpose of the analysis

Are agreed and signed off by all parties

BRIEF TITLE	<i>Come up with a project name</i>
Situation	<i>Share a brief description of what you hope to achieve during your project - what is the problem you are trying to solve? (max 400 words)</i>
Relevance	<i>How is the problem relevant to your role? Is the project relevant in the context of your day-to-day work? What impact could this project potentially have on your role/team/organisation? (max 400 words)</i>
Tasks	<i>What is your approach? Have you identified an existing source of data or are you going to create a new one? (max 400 words)</i>
Challenges	<i>What challenges have you identified going into this project? Is it access to data? Data protection? Anonymisation? IT? How do you envisage overcoming them, and what is your backup plan? (max 400 words)</i>



Activity

Draft a project brief for the following stakeholder:

A senior manager at your firm needs to access KPIs quickly, but the current process takes too long. You have been asked to make a dashboard in PowerBI.

Verification

The internal process of evaluating a data project to determine whether the products of a given development phase satisfy the conditions imposed at the start. This includes:

- Inspecting
- Design Analysis
- Specification Analysis

Validation

The external process of evaluating a data project during or at the end of the development process to determine whether it satisfies specific external requirements. This includes:

- Evaluating if the product meets the client expectations and requirements

Learning Objectives



- Describe the **data analytics lifecycle** and set out details, activities and roles at each stage in a real-world example
 - Design effective **project briefs** which **communicate expectations to stakeholders**
 - Identify **business specific rules** related to datasets and data characteristics that will influence project design and analysis
- 



ASSIGNMENT

PART 1- DATA ANALYTICS LIFE CYCLE

Use a work-related example to identify the stages of the Data Analytics Lifecycle. Describe what happened in each stage and highlight what was your role in the process. In the end, add a summary of the project/analysis including the main findings, what went well and what could have been improved.

Word Count	Max 1500 words
Deadline	3 weeks
Deliverables	Word Document or PowerPoint presentation



ASSIGNMENT

PART 2- PROJECT BRIEF

Use a work-related example to create a project brief. This could be related to a project you are about to start or something new. Your brief should contain a business problem, the wider context of the analysis and a plan of action to solve the problem.

Word Count	Max 1500 words
Deadline	4 weeks
Deliverables	Word Document

Complete Session Attendance Log and Update Your OTJ

