EXPLORE AI ACADEMY

A successful project is one where we understand the problem - Workshop 3

# Creating a Project Landscape

# Overview

01. Context

02. Define a problem landscape

03. Templates and examples

04. Conclusion and deliverable

EXPLORE AI ACADEMY

01. Context

02. Define a problem landscape

03. Templates and examples

04. Conclusion and deliverable

# The need

"The most reliable way to forecast the future is to try to understand the present." - John Naisbitt

- Determine data accessibility
- Understanding the domain and data
- Identify passed and potential solutions
- Determine areas of white space and uncertainty
- Understand value and impact on business
- Identify tools and expertise needed
- Assess risk and mitigations

EXPLORE AI ACADEMY

01. Context

02. Define a problem landscape

03. Templates and examples

04. Conclusion and deliverable

# What can we GATHER

> The following is aimed to guide you through the project landscape journey

| | |
|---|---|
| **Data** | • **Who owns the data** and who signs it off?<br>• **Where** is it?<br>• What **data format** is it structured or unstructured? |
| **Information** | • What does the **data dictionary** say?<br>• How is the data **gathered and updated**?<br>• Is it a problem that we can **solve using Data Science**? (what should we try and what was done in the past) |
| **Knowledge** | • Have we solved a **similar type of problem** before?<br>• Is there a s**ubject matter expert** on the team?<br>• Approach will be informed by our **tech stack and experience** |

EXPLORE AI
ACADEMY

01. Context

02. Define a problem landscape

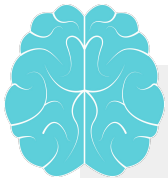
03. Templates and examples

04. Conclusion and deliverable

Exercise 1

# How to Landscape a Problem

Improving Underwriting at Infinity Life Assurance.
Use the problem landscape technique to assess based on the scenario below:



## Scenario:

You are new to the data team at a life assurance company, Infinity Life Assurance. They have asked you to identify the -14% drop in the revenue for the last business year. They have dropped in their underwriting percentages to 3%; the industry average is 15%. The company can potentially sell 75,000 policies but have only sold 30,000 policies at R450 million per annum.

You were tasked to create a data solution to tackle this problem and you are presenting a plan to the CEO on the problem.

Exercise 1 - Solution

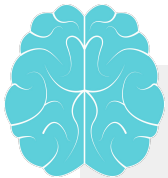
# Landscaping

| | Data | Information | Knowledge |
|---|---|---|---|
| **Underwriting Engine** | What is the current underwriting system? | What information does it currently use to price risk? | What other systems exist and how do they differ in their approach? |
| **Risk Price** | Can we get access to all of the previous prices produced by the current system? | What is the distribution of prices produced by the current system? | Price only matters relative to claims... what claims data is available? |
| **Margin** | What is the current underwriting margin? | What factors, other than the risk price, that might affect margin? | What are other companies margins and risk prices? |

Exercise 2

# How to Landscape a Problem

Reducing faults at Viber Fibre.
Use the problem landscape technique to assess based on the scenario below:



## Scenario:

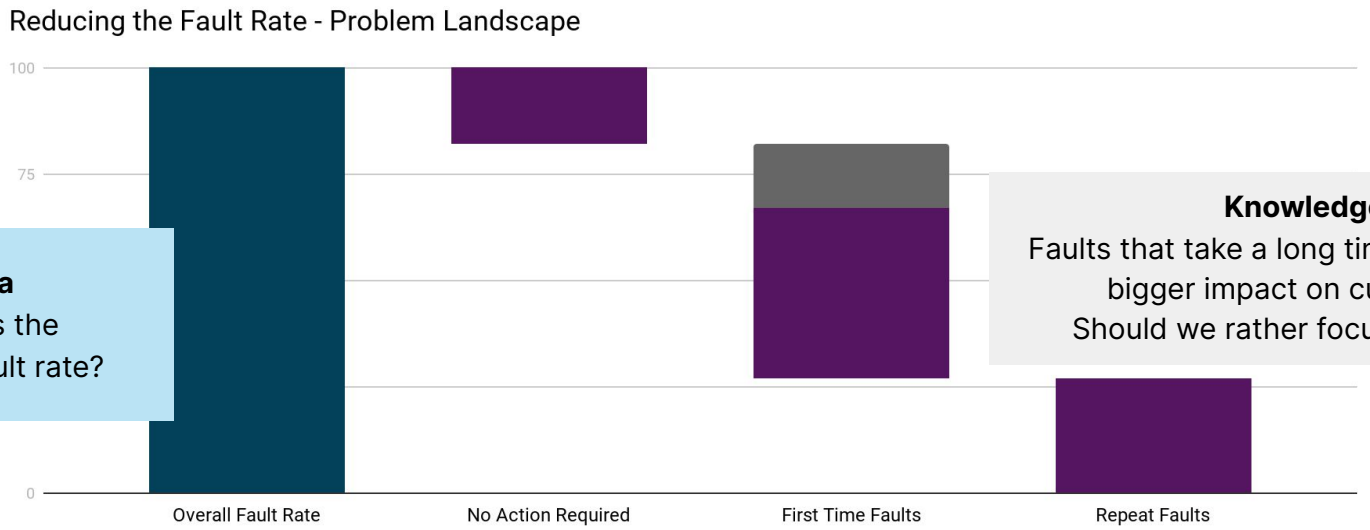
You are consulting to a fibre company, Viber Fibre, as a data scientist solutions expert. They sent you last year's operational report. In the report, it explains how the company has a 10% fault report by customers, whereas the industry standard is only 5%. There are over 1 million customers, but the company is losing an average of 100,000 each year, and only gaining an average of 10,000 new signups yearly. This churn rate is above the industry standard of 5%. The company charges R1400 for a callout fee, and the marginal revenue is R20,000.

You need to build a data-driven process to reduce the fault rate and customer churn. You will have a deep dive with the CTO around the problem.

Exercise 2 - Solution

# Landscaping

**Data** | **Information** | **Knowledge**




**Data**
What is the current fault rate?

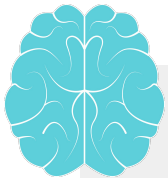
**Knowledge**
Faults that take a long time to fix have a bigger impact on customers.
Should we rather focus on those?

**Information**
What is the best way to split the fault rate?

Exercise 3

# How to Landscape a Problem

Find out what customers think of CloudBank SA.
Use the problem landscape technique to assess based on the scenario below:



## Scenario:

CloudBank SA has had a rainy start to the new marketing campaign. Their churn rate has increased from 1% to 6%. They have also spent R100 per acquisition, totalling 200,000 new users. Compared to last year, they had achieved an acquisition cost of R35 per new signup and had 500,000 people take on their services. Although, they had a reach 25 million people during the new campaign, and 20 million in the previous campaign, the number of interested clients dropped by 10%.

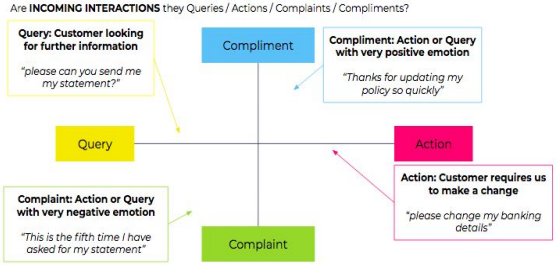
You have been called in as the lead data scientist to analyse the audience and marketing campaigns. You will need to explain the problem using data analysis to the marketing team.

Exercise 3 - Solution

# Landscaping

| Data | Information | Knowledge |
| --- | --- | --- |
| Logged on our website and online application process<br><br>Publically available<br><br>Sent to our mail servers | Do we have existing pipelines to format and process the data to get insights<br><br>Currently manual<br>(potential process automation) | Clustering<br>Natural Language Processing<br>Sentiment Analysis |

hellopeter.com

EXPLORE AI ACADEMY

01. Context

02. Define a problem landscape

03. Templates and examples

04. Conclusion and deliverable

# Conclusion and deliverable

> Design and upload a project landscape. This is a 1 page powerpoint or google slides.

**What you have learnt:**

- **The role of raw, unprocessed data in the Problem Landscape**
- **The importance of access to the right information**
- **The importance of identifying appropriate skills and expertise to solve the problem**