# Approaching the Loan Default Problem: Data Driven Decision Making

5/14/2021

# What does our process look like?

1. Identify the business questions.
2. Plan for the analysis.
3. Data collection.
4. Pull out insights.
5. Provide actionable recommendations.

# Identifying the Business Questions

1. How have loan amounts been previously decided?
2. Who are the people defaulting on loans? Are there trends that can help us understand this issue better? (i.e.: seasonality, demographic details)
3. Is there a model that can help us lower default rates by providing insight around who should be approved for loans and to what amount?

# Analysis Plan

- We want to look at the problem from a full 360 degree view.
- Let’s look at what’s been done in the past, what is currently happening, and what we want to see in the future.

Past: Use past data to determine how loan amounts have been decided previously.

# Analysis Plan con.

Present Issue: Once we understand how loan amounts have been decided previously, we can see what percentage of loans over the last 6 months have been defaulted on.

Some things to consider:

How is default defined?

How are minimum payments determined?

Let’s look for trends in the data that could help us understand the why.

# Analysis Plan con.

Future: Is there a model that can help us better decide how to determine who is more likely to default on loans?

Let’s relate this model back to findings from past data to create a full picture of the problem.

# Data Collection

- There is 6 months worth of data being offered to the team that is broken out per consumer per month.
- The data is currently being stored in a data warehouse and needs to be pulled down by users.
- The data requires some cleansing and is primarily made up of volumes. The team will have to do the following:
  - Remove duplicates
  - Convert anything that is text to a number (statistical processes cannot process text)
  - Create rates that will help us see rates (example: default rates)

# Data Collection con.

Recommendation: The team recommends at least a full year of consumer data so that seasonality can be accounted for in the analysis.

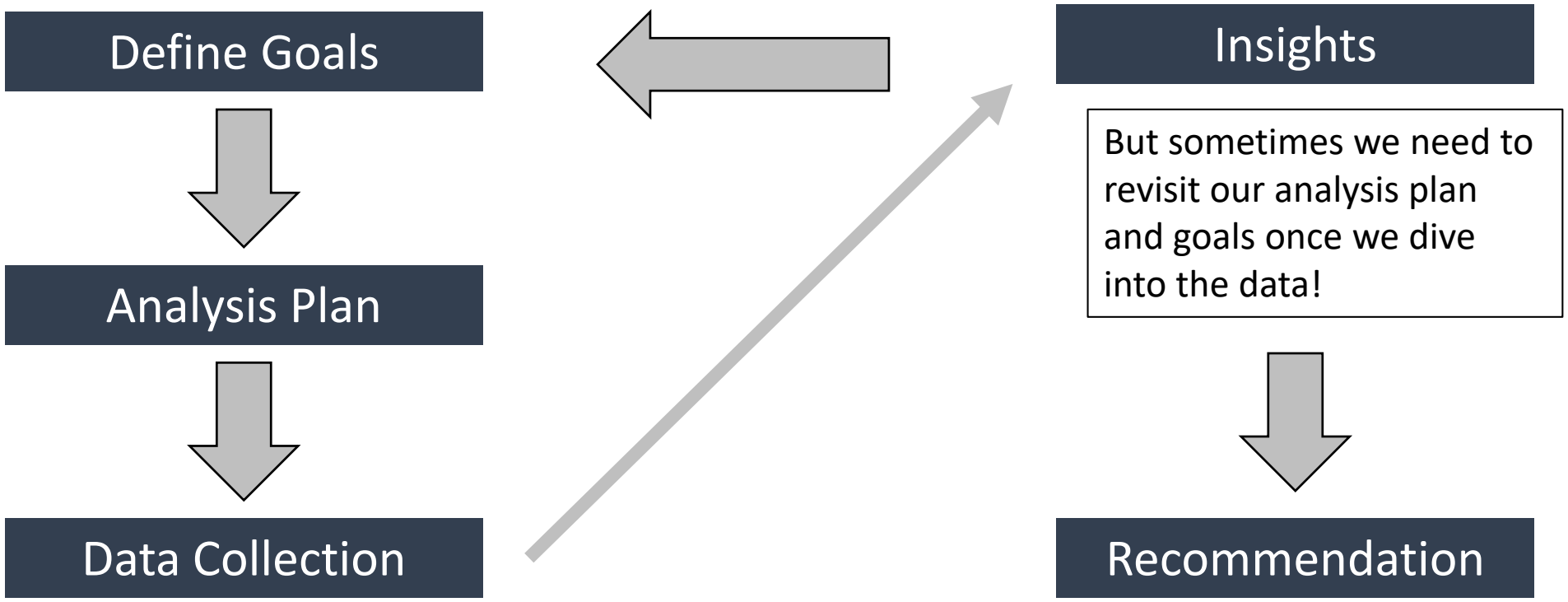
# Insights and Recommendations

This will be the final and most important steps.

As we work through the data we can decide what format is best for the patterns we are seeing and who should be in the meeting.

Implementation will also be discussed in this section.

# The Process Visualized



# Initial Insights

1. Females are offered larger loans for every level of education except for Graduate School.
2. Men and women who are married are offered the largest loans on average while those who were divorced received the smallest loans.
3. On average, the loan amounts increase as age increases, however there is a decrease in the age group of 51 – 60 when compared to the previous and subsequent decade.

# Thank you