Credit One Project

CARLOS E PENA
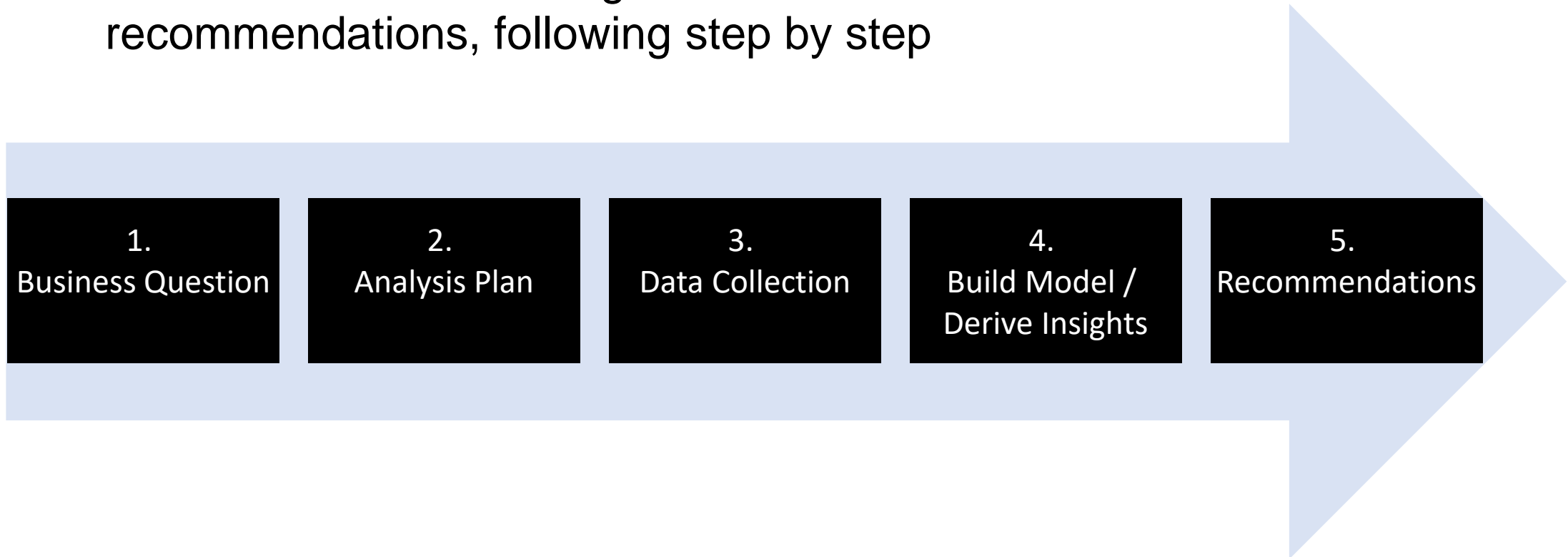
Data Scientist

Define the Goal

- ✓ Credit One is at risk of losing business because its <u>credit score service</u> is getting less effective:

  - ➢ Increase of customers who have defaulted on loans

- ✓ **Credit One needs to improve its understanding on how much credit is allowed to someone to use.**

Data Science Framework

✓ The framework to be followed for this project will be (BADIR):

➢ From clear understanding of the issue to the recommendations, following step by step

| 1.
Business Question | 2.
Analysis Plan | 3.
Data Collection | 4.
Build Model /
Derive Insights | 5.
Recommendations |

Collect Data

✓ Historical data has been collected:

  ➢ *Amount of the given credit*, between $10K and $1,000K

  ➢ *Gender*, 60% female – 40% male

  ➢ *Education*, High school, University, Graduated school and other

  ➢ *Marital Status*, married, single, divorce, other

  ➢ *Age,* goes from 21 to 79 years old

  ➢ *History of past payments*

  ➢ *Amount of previous payment*

  ➢ *Customer behavior*, default and not default

Manage Data

- ✓ Data is in well shape although some work need to be done:
  - ➤ Check on missing data
  - ➤ Handling inconsistent data
  - ➤ Delete duplications
  - ➤ Data reduction
  - ➤ Check relationship between data

Potential issues with Data

- ✓ There are some concerns with the given data:

  - ➢ Relationship between variables

  - ➢ Information need to be numeric, therefore some text data need to be changed

  - ➢ Too much data

  - ➢ What information will be meaningful for the analysis

Initial Insights

✓ Glean from quick look at the data:

➤ Divorce people is less likely to default

➤ People less than 30 years old, high school degree with high amount of credit

➤ Customer behavior, 22% of people expected to default

➤ From those 22%, 55% has a credit lower or equal to $100K and 37% lower or equal than $50K