# LOGISTICS REGRESSION

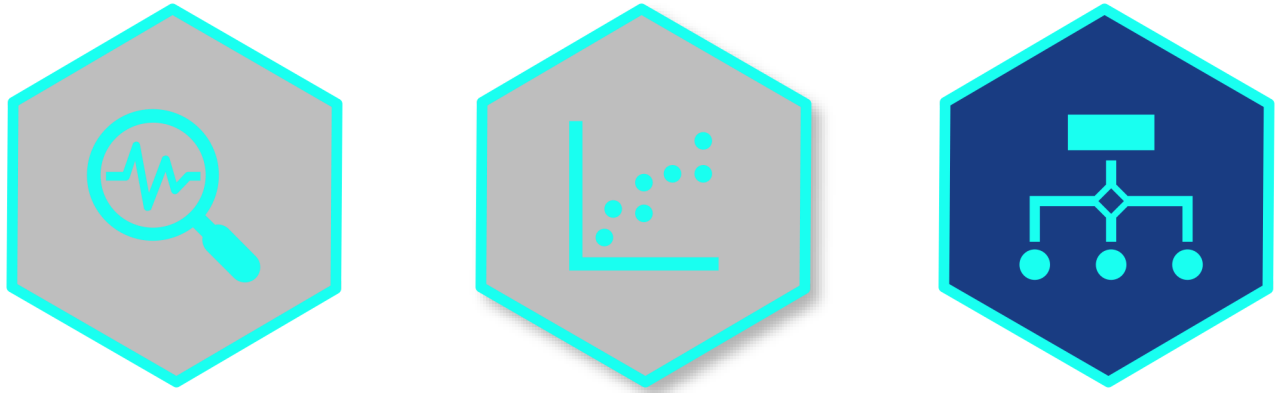
EMBARKING ON A JOURNEY INTO DATA SCIENCE

YA MANON

You can have data without information but you cannot have information without data.

-Daniel Keys Maran

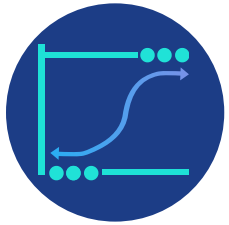
# ABOUT THIS SERIES



**PART 1** → **PART 2** → **PART 3**

Data Prep & EDA | Regression | Classification

# LOGISTICS REGRESSION



In this section we'll cover the a technique that predict a categorical dependent variable based on one or more independent variables

**TOPICS WE'LL COVER:**

- Logistics Regresson
- Goals of Regression
- Evaulation

**GOALS FOR THIS SECTION:**

# LOGISTICS REGRESSION

Logoistics Regression

Goals of Regression

Evaulation

**Logistic Regression** is a classification technique used to predict the probability of a binary (true/false)outcome

- In its simplest form, logistic regression forms an **S-shaped curve between 0 -1**, which represents the probability of a TRUE outcome for any given value of X
- The **likelihood function** measures how accurately a model predicts outcomes, and is used to optimize the "shape" of the curve
- Although it has the word *"regression"*in its name, logistic regression is not used for predicting numeric variables

## Example use cases:

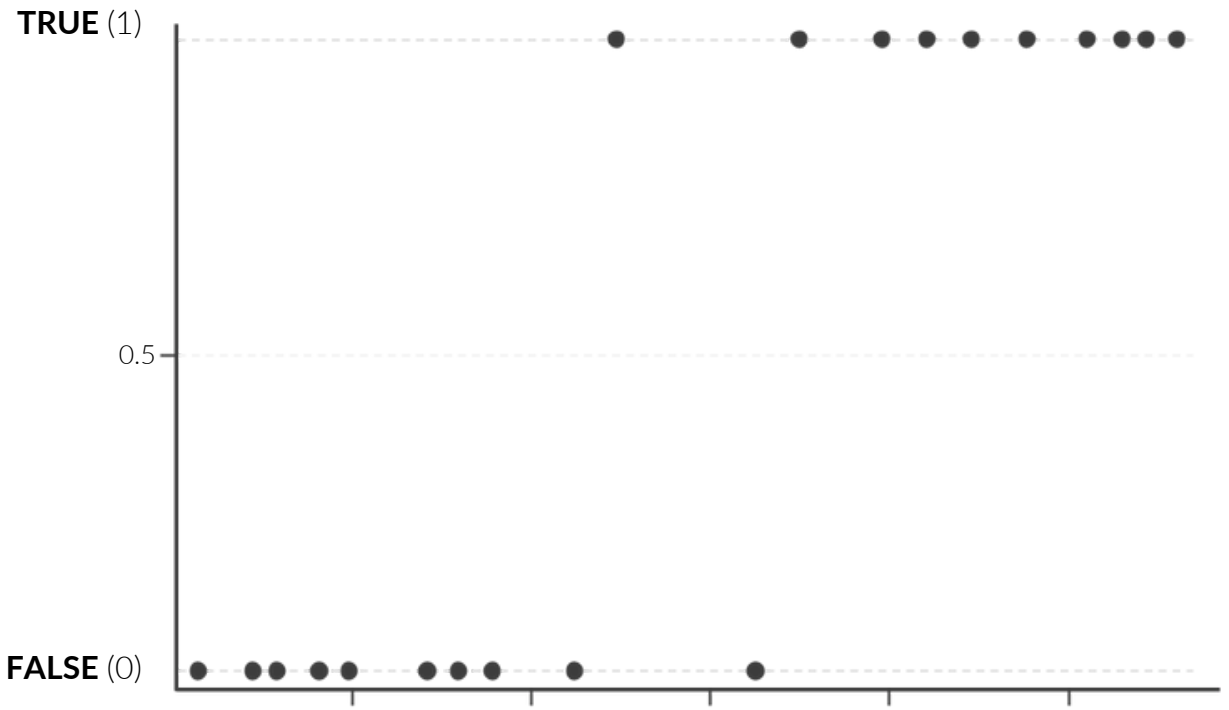
Flagging spam emails or fraudulent credit card transactions

# LOGISTICS REGRESSION

Logoistics Regression

Goals of Regression

Evaulation



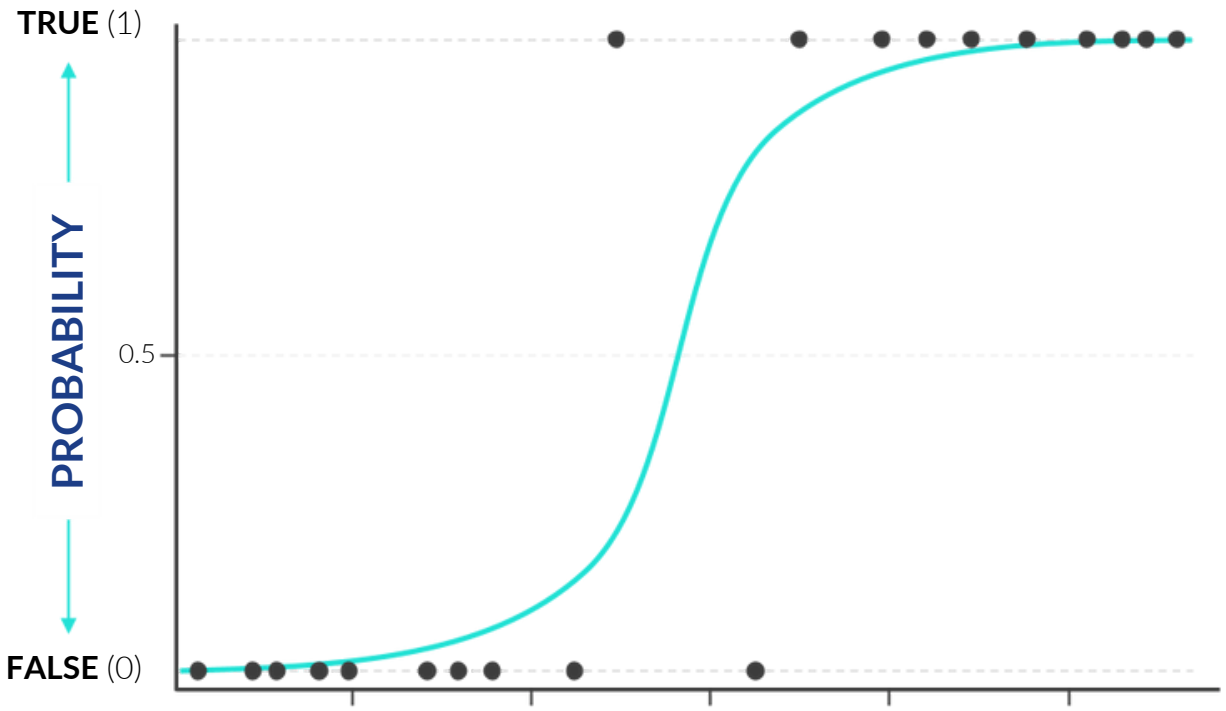

- Each dot represents an observed value, where **X is a numerical independent variable** and **Y is the binary outcome** *(true/false)*that we want to predict

# LOGISTICS REGRESSION

Logoistics Regression

Goals of Regression

Evaulation



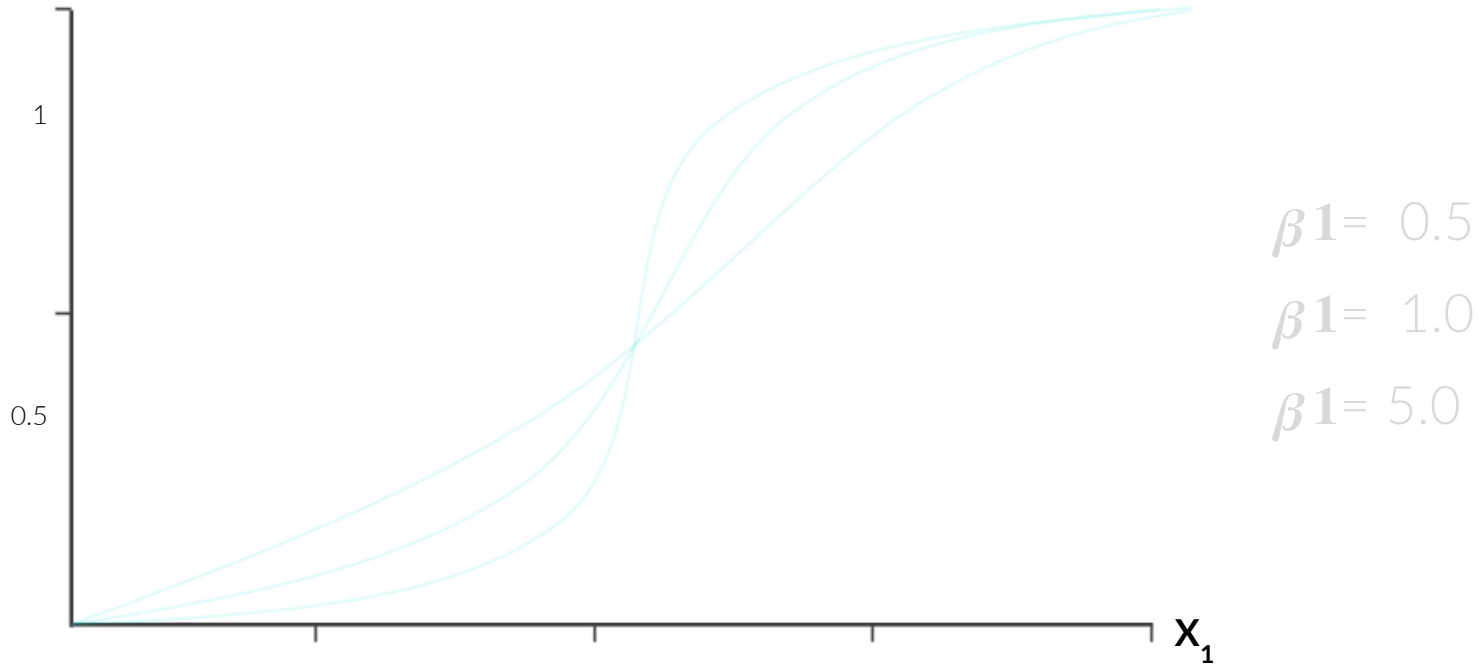

- Logistic regression plots the **best-fitting curve between 0 and 1**, which tells us the probability of Y being TRUE for any given value of X1

# LOGISTICS REGRESSION

- Logoistics Regression
- Goals of Regression
- Evaulation

Makes the output fall between **0** and **1**

$$\frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}}$$




- Logistic regression plots the **best-fitting curve between 0 and 1**, which tells us the probability of Y being TRUE for any given value of X1

# LOGISTICS REGRESSION

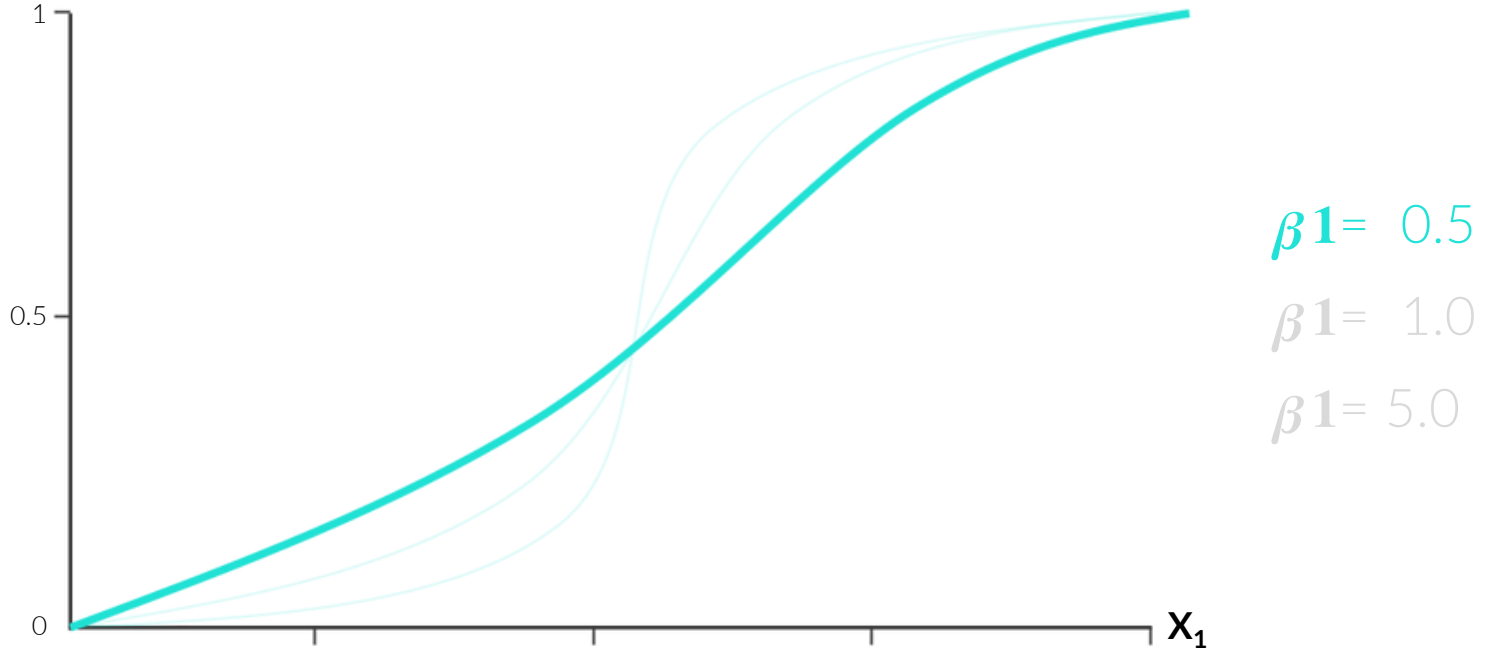
Logoistics Regression

Goals of Regression

Evaulation

Makes the output fall between **0** and **1**

$$\frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}}$$




- Logistic regression plots the **best-fitting curve between 0 and 1**, which tells us the probability of Y being TRUE for any given value of X1

# LOGISTICS REGRESSION

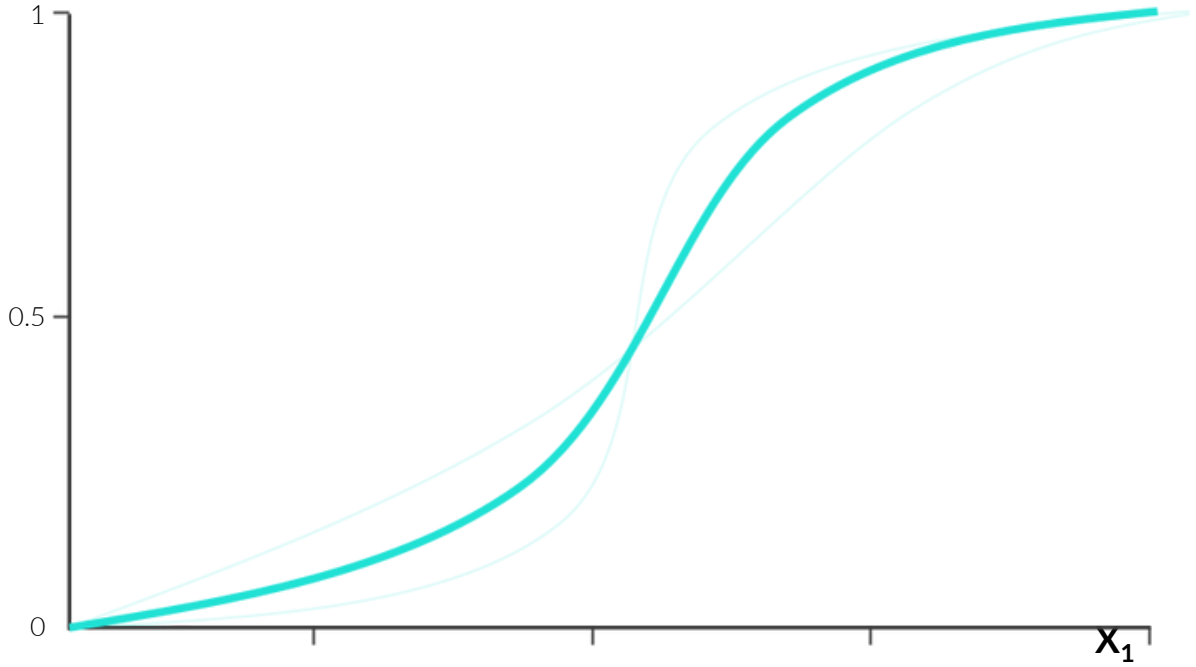
Logoistics Regression

Goals of Regression

Evaulation

Makes the output fall between **0** and **1**

$$\frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}}$$




$\beta 1 = 0.5$

$\beta 1 = 1.0$

$\beta 1 = 5.0$

- Logistic regression plots the **best-fitting curve between 0 and 1**, which tells us the probability of Y being TRUE for any given value of X1

# GOALS OF REGRESSION

Logoistics Regression

Goals of Regression

Evaulation

Regression models are used for two primary goals: **prediction** and **inference**

The goal shapes the modeling approach, including the regression algorithm used, the complexity of the model, and more

## PREDICTION

Used to **predict** the target as accurately as possible

*"What is the predict charges for a client given their age?"*

## INFERENCE

- Used to **understand the relationships** between the features and target
- *"How much do a age impact its charges?"*

# EVALUATION

- Logoistics Regression
- Goals of Regression
- Evaulation