

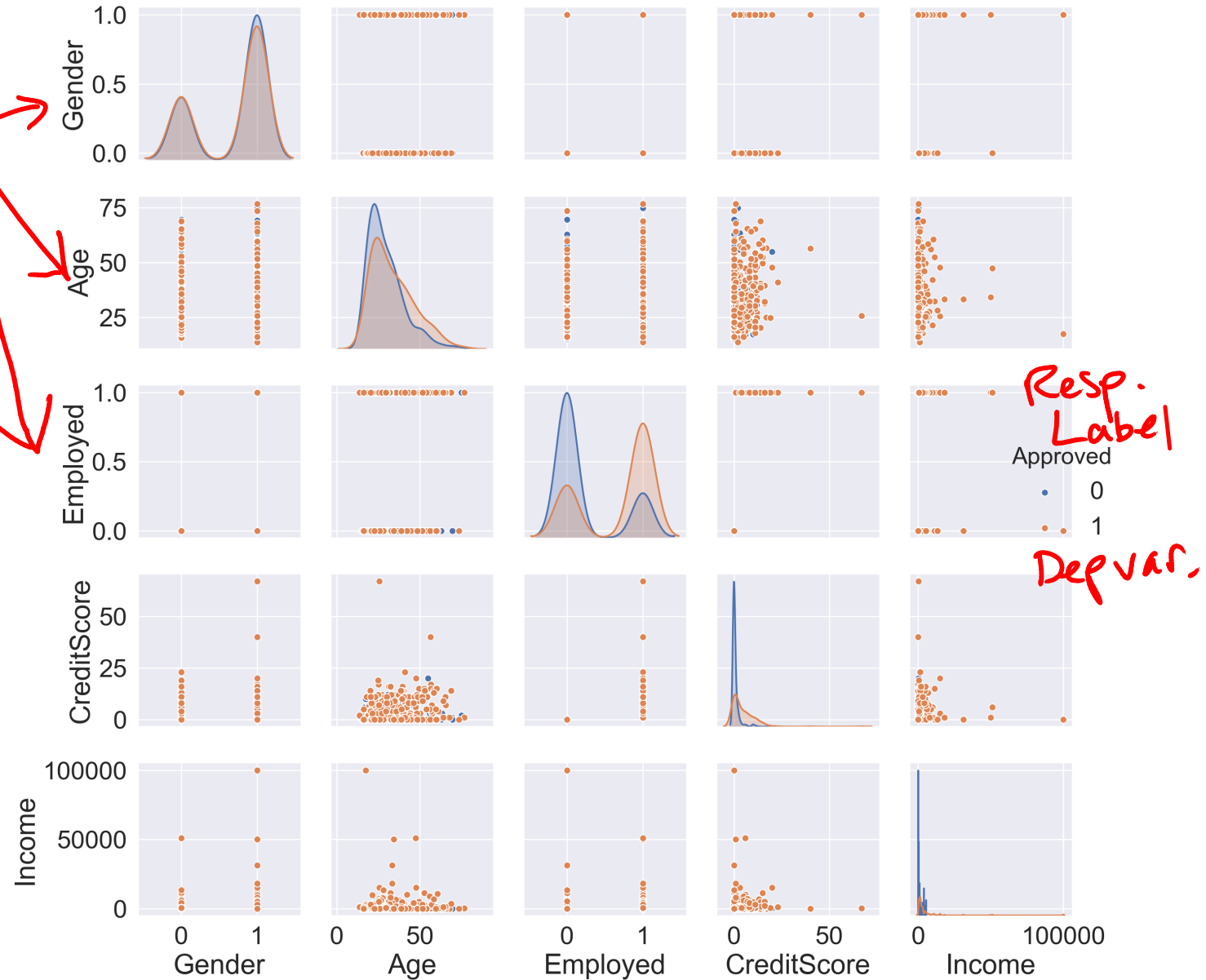
The background of the slide features a stylized globe on the left side, partially obscured by a dense pattern of binary code (0s and 1s) that recedes into the distance, creating a sense of depth and digital connectivity. The overall color palette is a mix of light blues, purples, and whites.

# Foundations of Data Science: Logistic regression - Principle of logistic regression

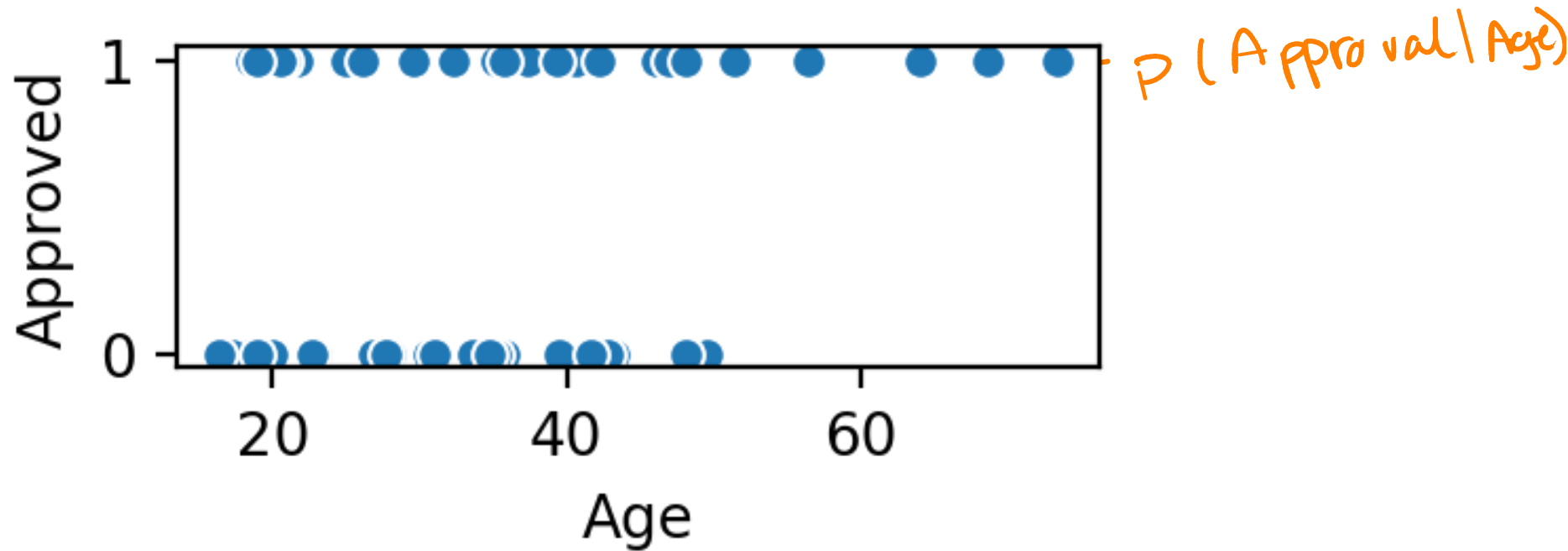
# Supervised classification

Indep. vars  
Features  
Predictor vars

Logistic regression  
predict  
probability



# Logistic regression task on one continuous variable



# Binary variables: odds and odds ratios

$$P(Y=y | X=x)$$

	Y Approved	Not approved	Approval odds
X Employed			
0	0.25	0.75	0.34
1	0.71	0.29	2.42

odds ratio  
 $OR(\text{Employed}) = \frac{2.42}{0.34} = 7.09$

} Effect Size  
609%

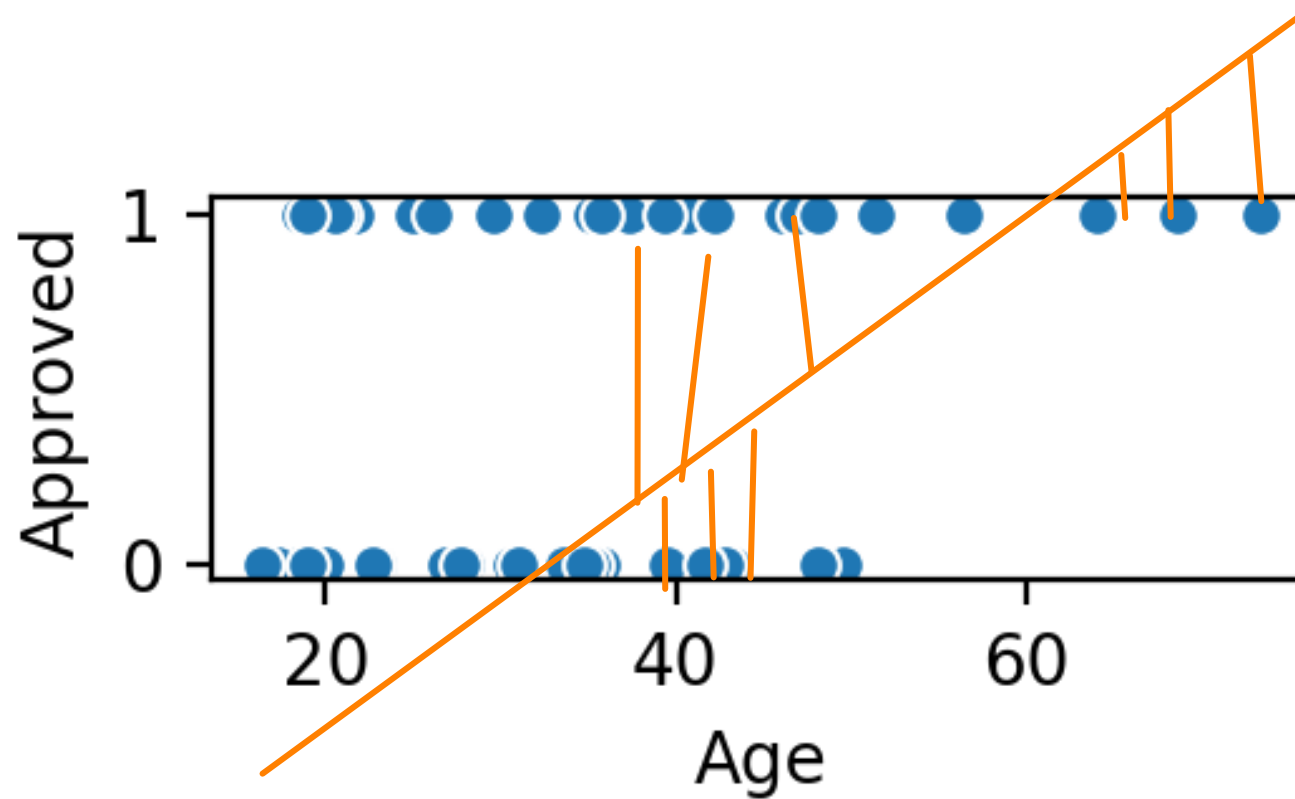
$y \in \{ \text{Approved, Not approved} \}$

$x \in \{ \text{Employed, Not Employed} \}$

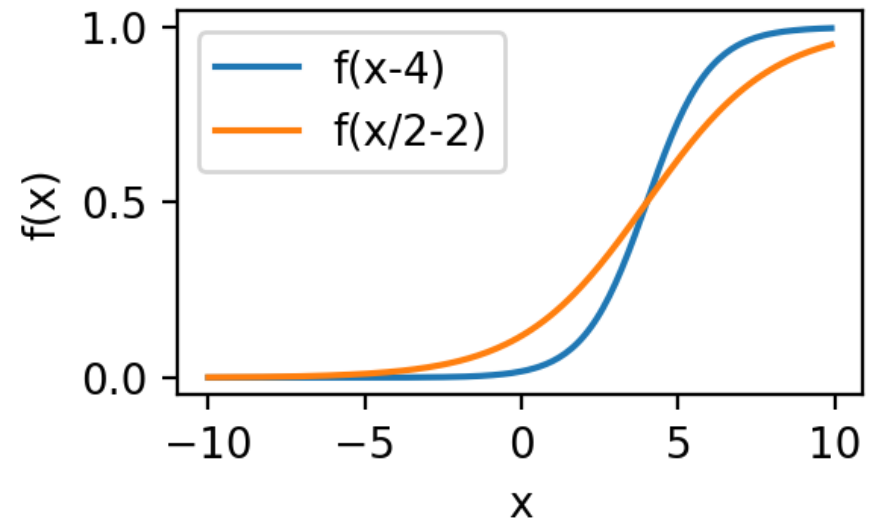
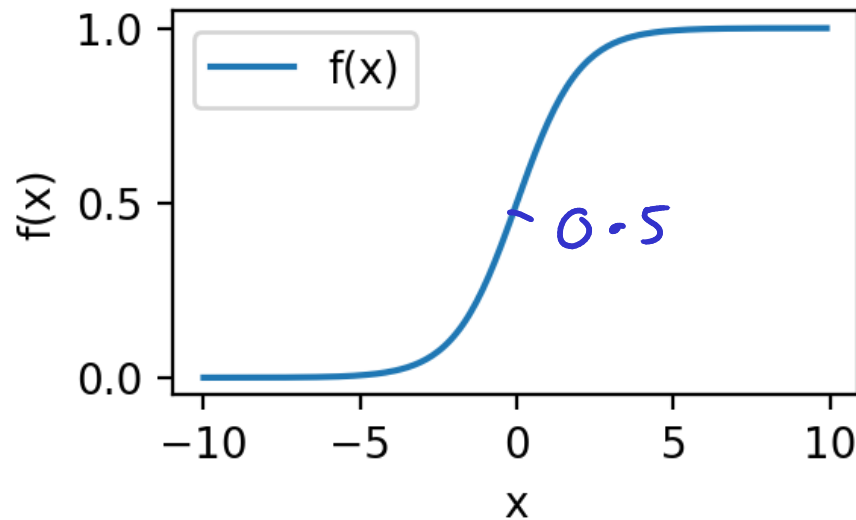
$$\text{Odds (Success)} = \frac{P(\text{Success})}{P(\text{Failure})} = \frac{P(\text{Success})}{1 - P(\text{Success})}$$

Odds ratio  $OR(x) = \frac{\text{Odds (Success} | x = \text{True})}{\text{Odds (Success} | x = \text{False})}$

# Principle of logistic regression with one variable



# Logistic function



$$S(x) = \sigma(x) = f(x) = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}$$

$$P(Y=1 | x) = f(\beta_0 + \beta_1 x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

$\hat{\beta}_0 \quad \hat{\beta}_1$ 

 $\uparrow \quad \uparrow$   
 coeffs

# Application to continuous variable in credit example

---

