To explain and derive logistic regression based on the video you shared, here's a summarized version of the key concepts, followed by a Python implementation.

Summary of Key Concepts:

1. **Why Logistic Regression is Needed**: Linear regression is not suitable for classification tasks as it can predict values outside the range of [0, 1], which doesn't make sense for probabilities. Logistic regression solves this by modeling the probability of an outcome using a logistic (sigmoid) function.

2. **Logistic Function**: The logistic regression model uses the sigmoid function:

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

   This gives a probability value between 0 and 1, which is ideal for classification tasks.

3. **Maximum Likelihood Estimation (MLE)**: The goal is to find the coefficients (β) that maximize the likelihood function. This ensures the model's predictions are as close to the true class labels as possible. The likelihood function is the product of conditional probabilities for all samples:

$$L(\beta) = \prod_{i=1}^{n} P(Y_i|X_i)$$

   The log-likelihood is then used to make optimization easier: