2.

With logistic regression we model the probability that the qualitative output Y belongs to one particular class based on the given input X, i. e for the binary 0/1 value of Y it can be written as: $P(Y = 1/0|\boldsymbol{X})$

Relationship between the obtained probabilities and the variables $\boldsymbol{X}$ is not linear (it is *sigmoidal*), and can be represented by the equation below:

$p(\boldsymbol{X}) = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$

It can be also rewritten as:

$p(\boldsymbol{X}) = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}}$

3.

Value of *odds* allows to asses very quickly whether probability is low or high. We can calulate the odds with the following formula:

$odds = \frac{p(X)}{1-p(X)} = e^{\beta_0+\beta_1 X}$

As we can see, we calculate the odds by dividing the probability of an event by the probability of a complementary event.

If the value of odds is represented by a fraction $\frac{n}{p}$ it means that on average, out of $p + n$ values of Y with the given input $\boldsymbol{X}$, n of them will belong to the class the prediction is being calcuted for.

4.

$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$

$log(\frac{p(X)}{1-p(X)}) = log(\frac{\frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}}{1-\frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}})$

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

$= log(\frac{\frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}}{\frac{1}{1+e^{\beta_0+\beta_1 X}}})$

$= log(e^{\beta_0+\beta_1 X}) = \beta_0 + \beta_1 X$

The logit function is a linear function of X in contrast to logistic function. It is much easier to estimate the rate of change of logit function than logistic function with unit change of X. It can by done by increasing the value of logit by β_1.

5.

$\frac{odds(Y_{\boldsymbol{x}^T\boldsymbol{\beta}+\beta_i\delta})}{odds(Y_{\boldsymbol{x}^T\boldsymbol{\beta}})} = \frac{e^{\boldsymbol{x}^T\boldsymbol{\beta}+\beta_i\delta}}{e^{\boldsymbol{x}^T\boldsymbol{\beta}}} = e^{\beta_i\delta}$

This equality shows, that if we increase the value of X_i by δ, the value of odds becomes $e^{\beta_i\delta}$ times greater/smaller, where β_i is a corresponding parameter. It is a very simple way to asses the change of probablility of Y, in comparison with calculating the value of $P(Y = 1|X)$ all over again.