Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

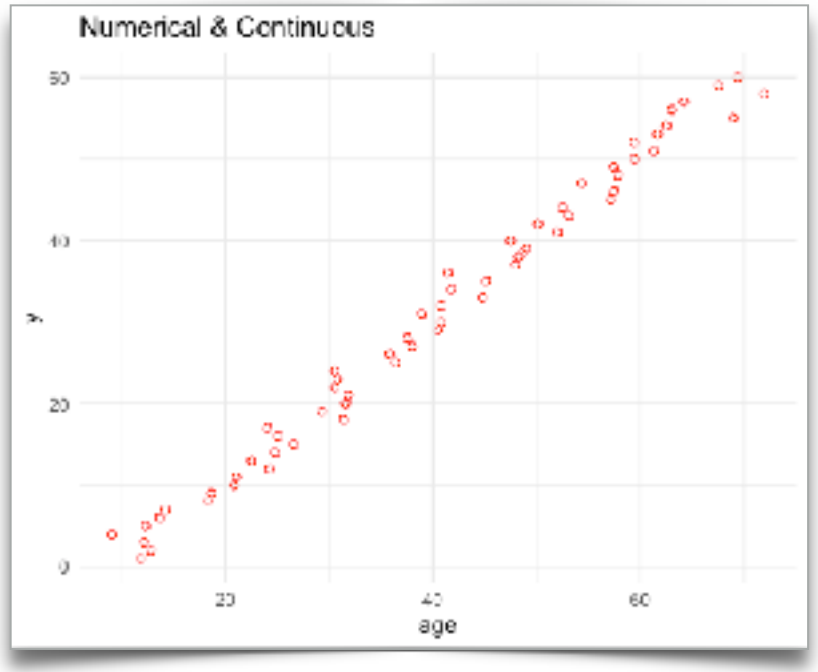
PREDICTING ORDINAL DATA

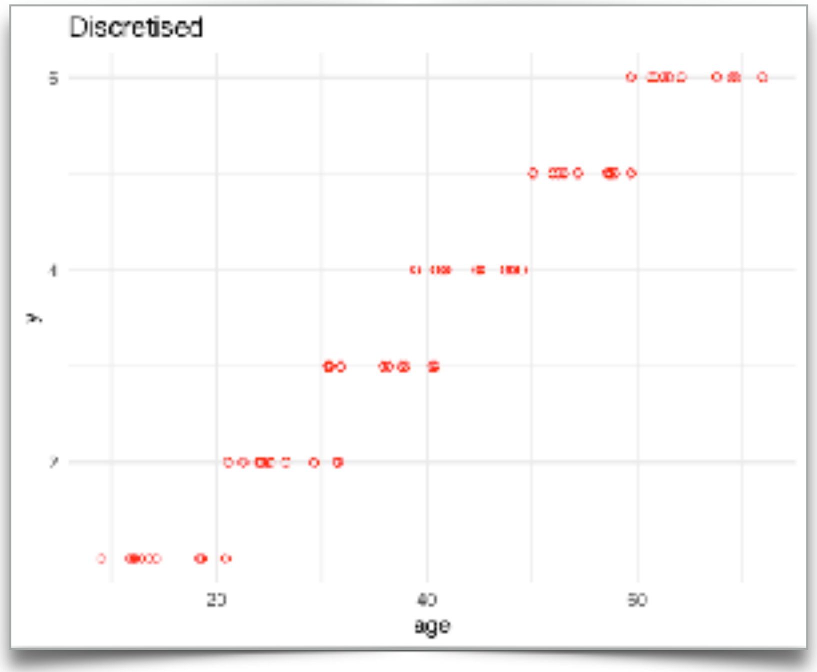
EMMA BEAUXIS-AUSSALET

e.m.a.l.beauxis@hva.nl

ORDINAL DATA

Ordinal data have continuous relationship between their ‘categories’.
For instance, numerical data are often discretised (e.g., ratings).

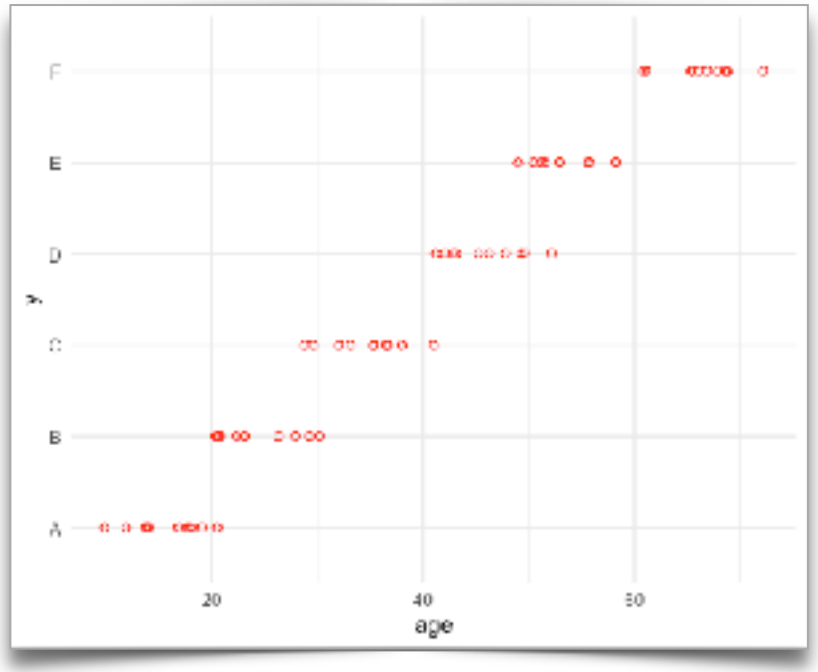


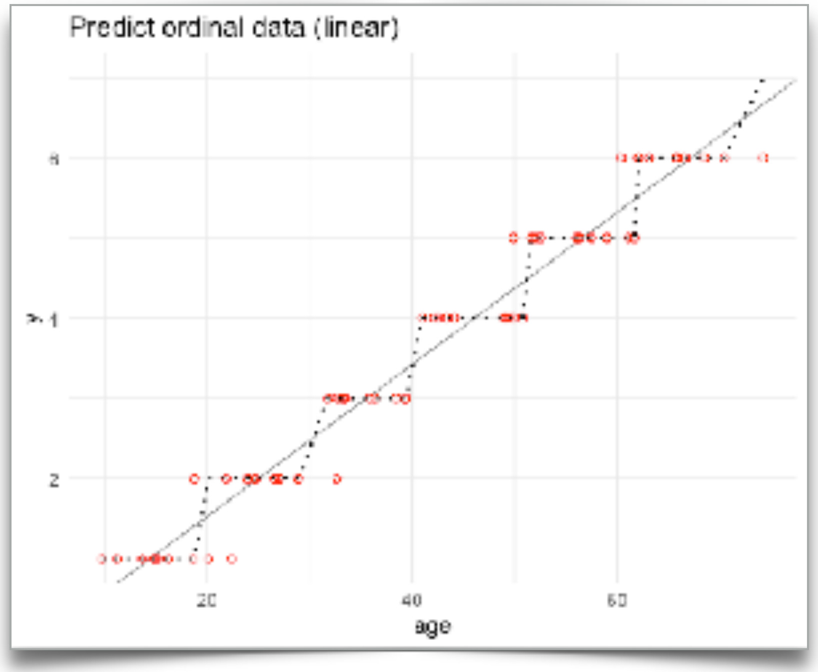


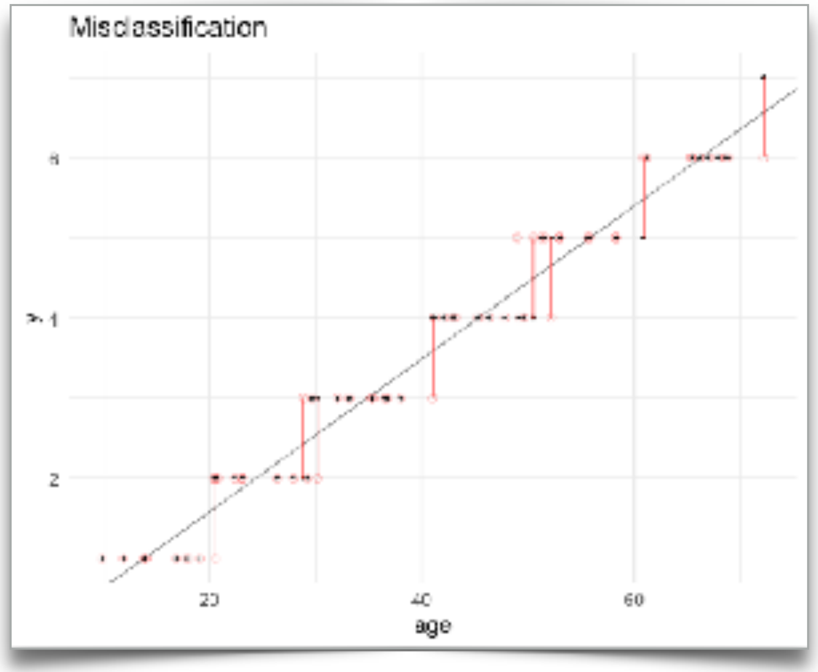
PREDICT ORDINAL DATA

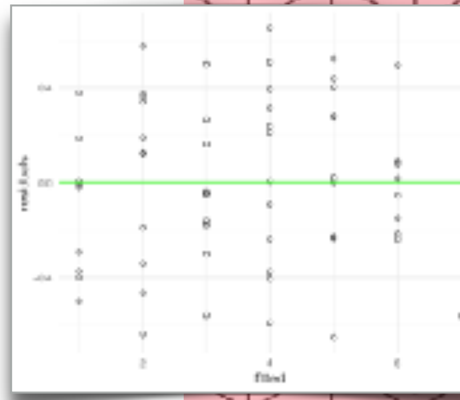
Linear regressions can be fitted on ordinal data.

```r
###### Linear regression
model <- lm(y ~ age)
```



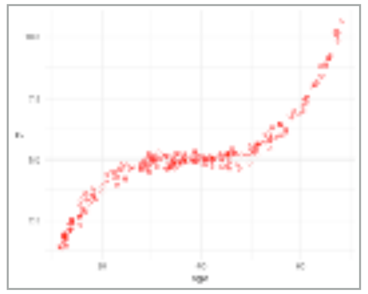




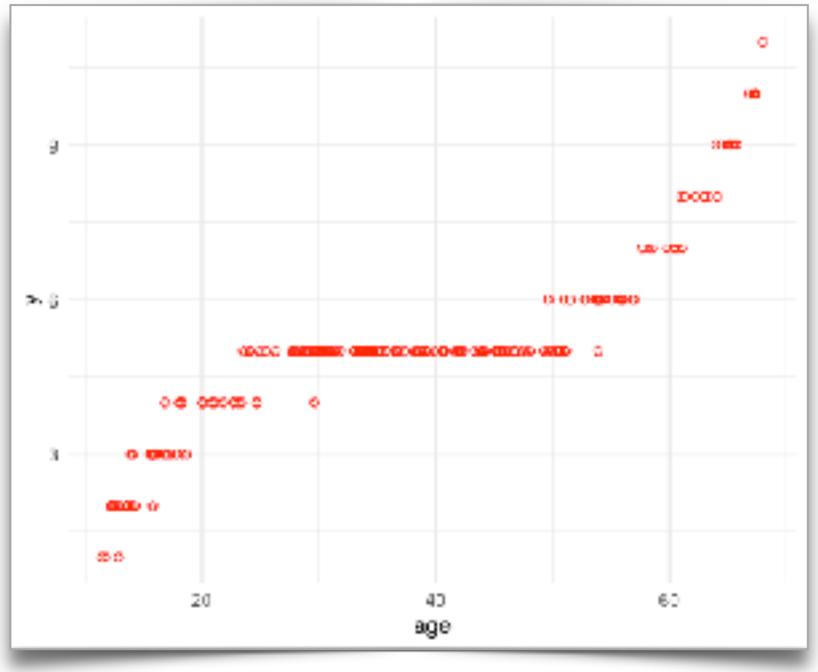


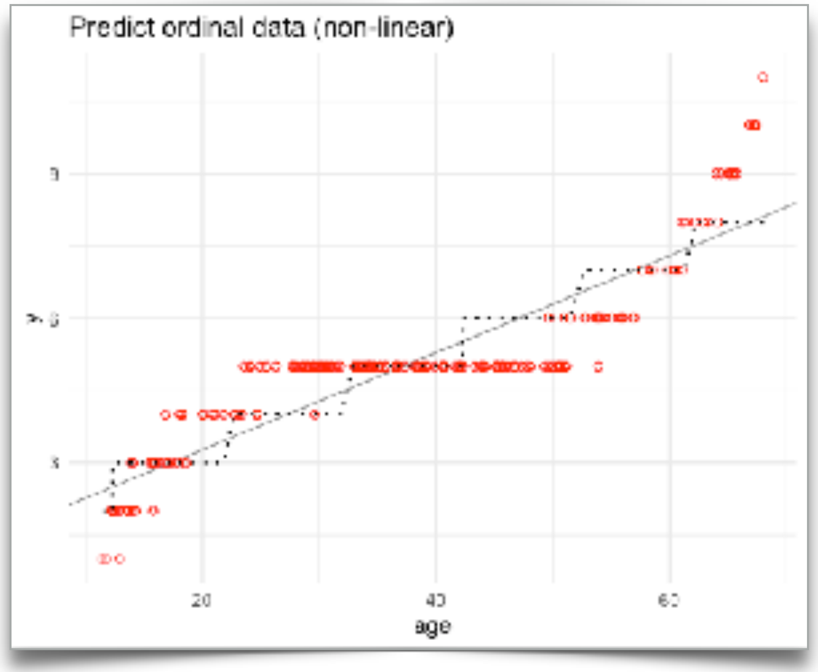
PREDICT ORDINAL DATA

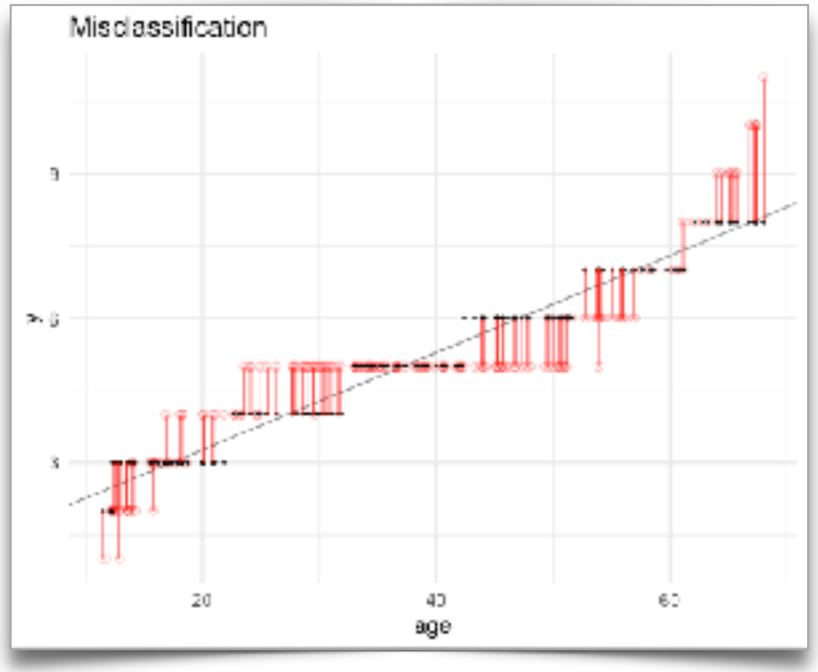
Linear regressions can be fitted on ordinal data.

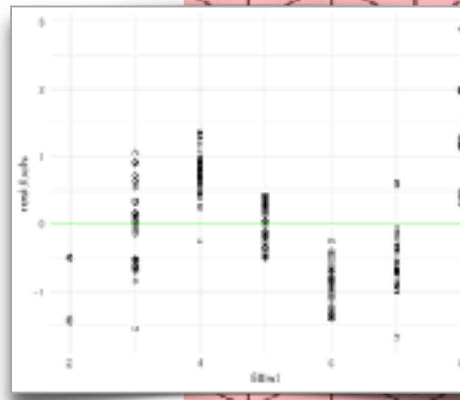


```r
###### Linear regression
model <- lm(y ~ age)
```



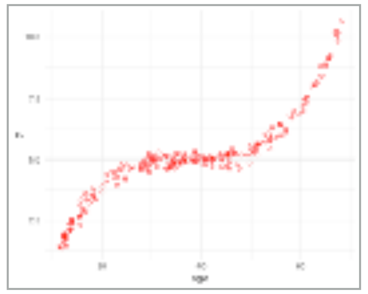




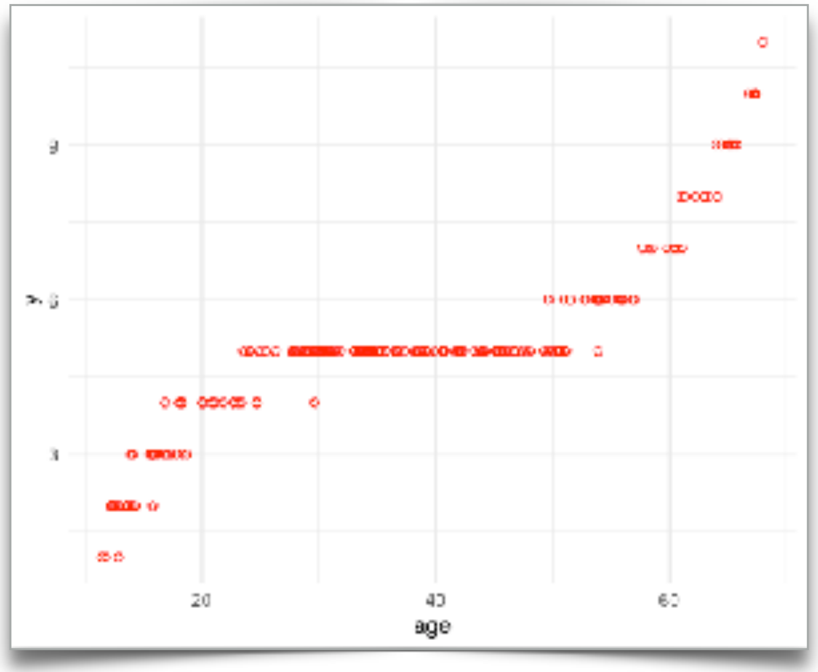


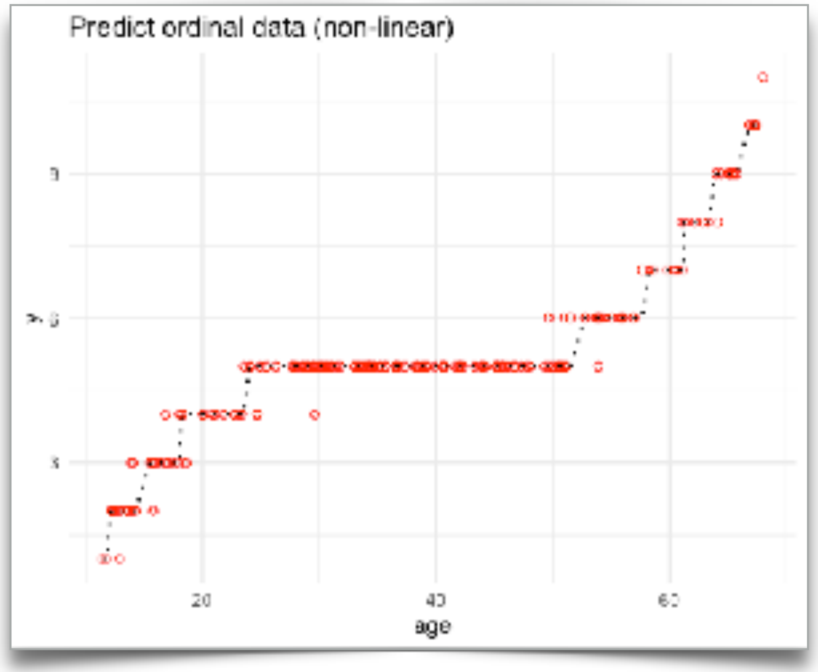
PREDICT ORDINAL DATA

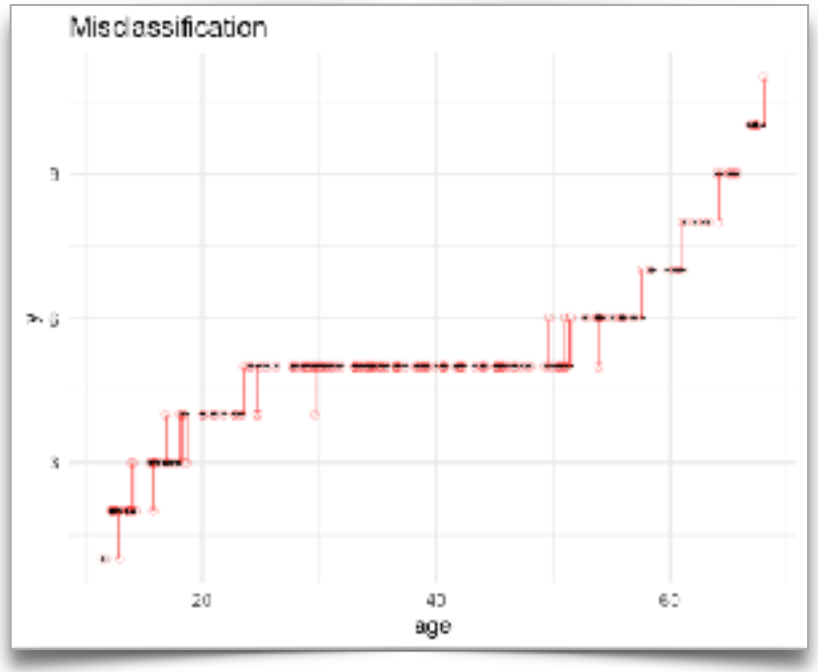
Linear regressions can be fitted on ordinal data.

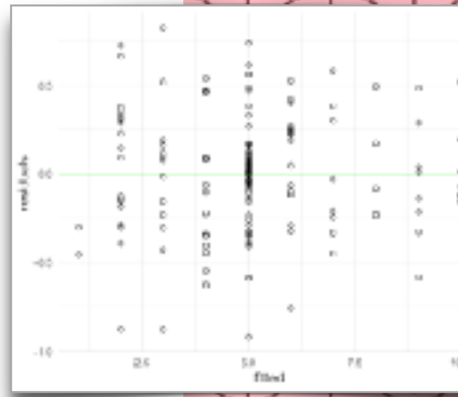


```r
###### Polynomial regression
model <- lm(y ~ poly(age,3))
```









Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

PREDICTING CATEGORICAL DATA

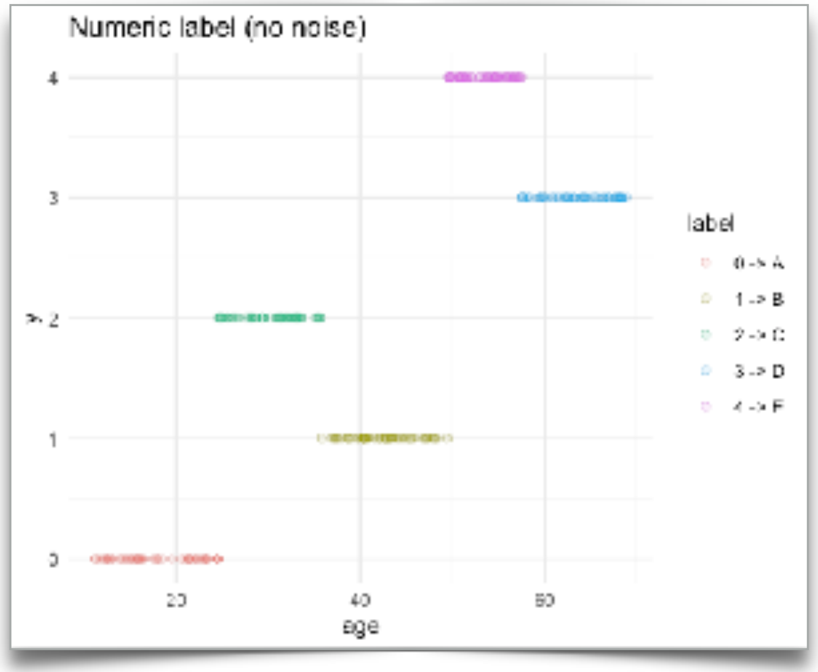
EMMA BEAUXIS-AUSSALET

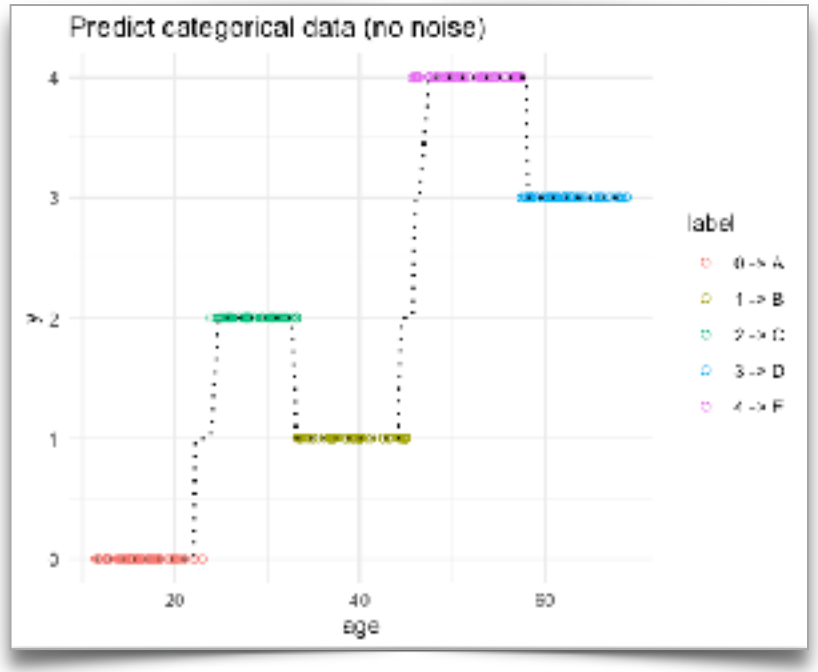
e.m.a.l.beauxis@hva.nl

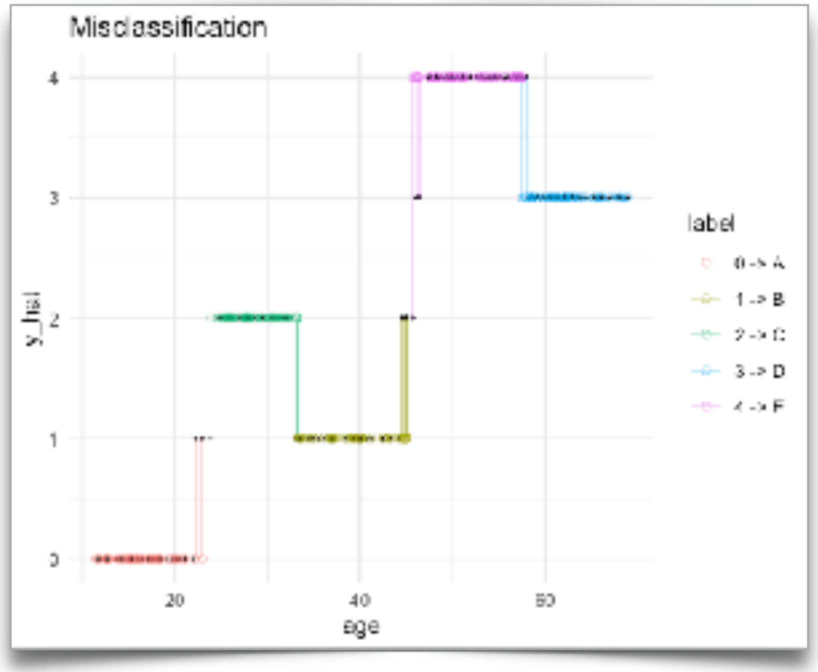
PREDICT CATEGORICAL DATA

Polynomial regressions may be fitted on categorical data, but with many parameters… but they are not interpretable.

```
###### Polynomial regression - Degree 20
model <- lm(y ~ poly(age,20), data=data)
```



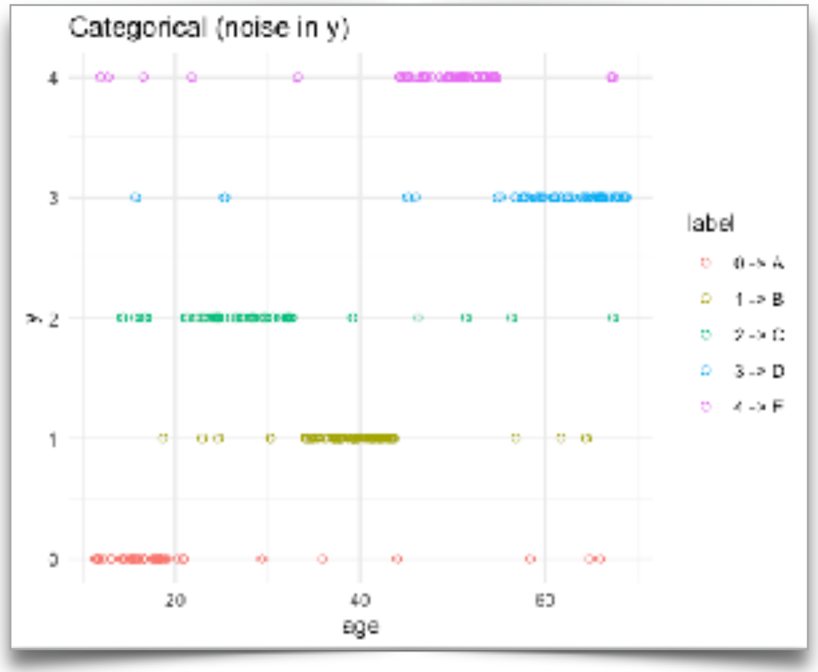


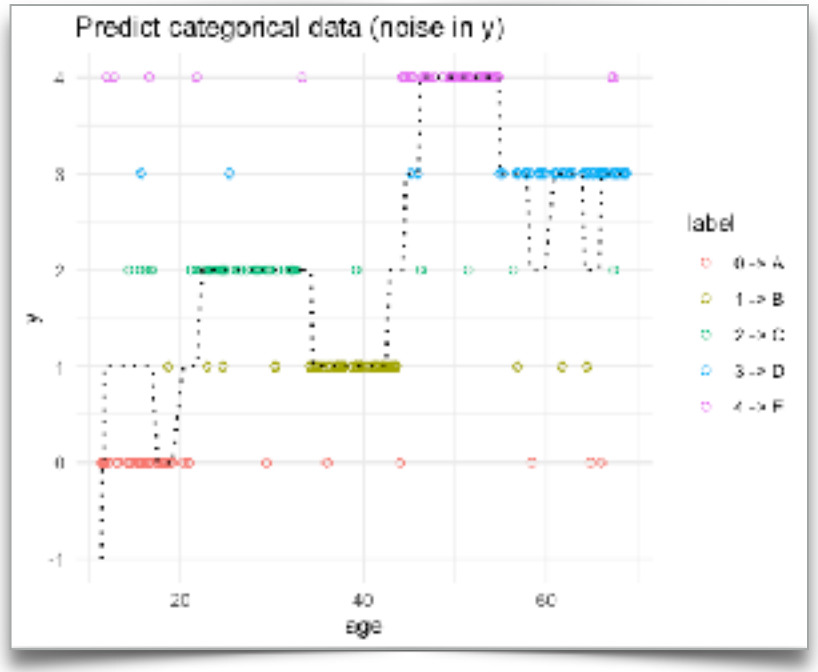


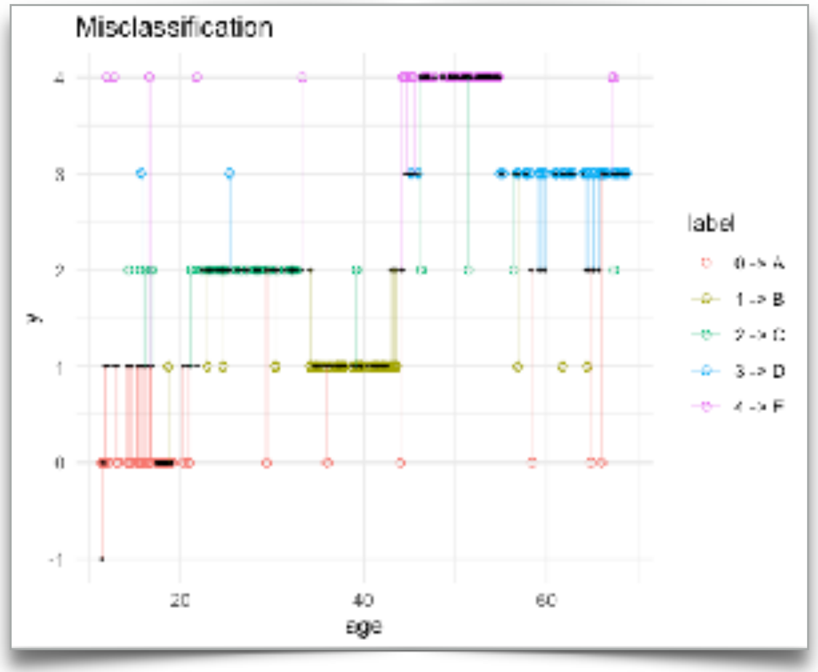
PREDICT CATEGORICAL DATA

Polynomial regressions may be fitted on categorical data, but with many parameters… but they are not interpretable… and **not robust to noise!**

```
###### Polynomial regression - Degree 20
model <- lm(y ~ poly(age,20), data=data)
```





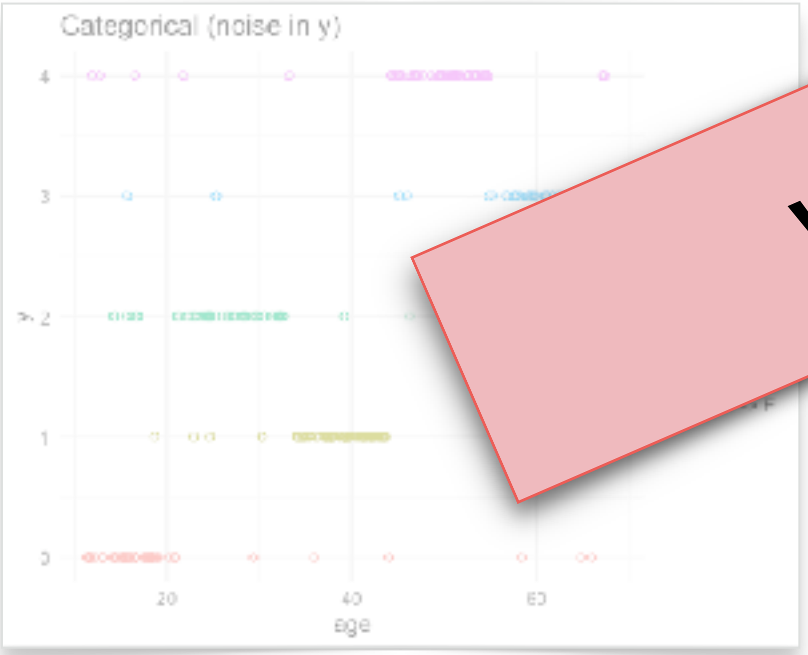


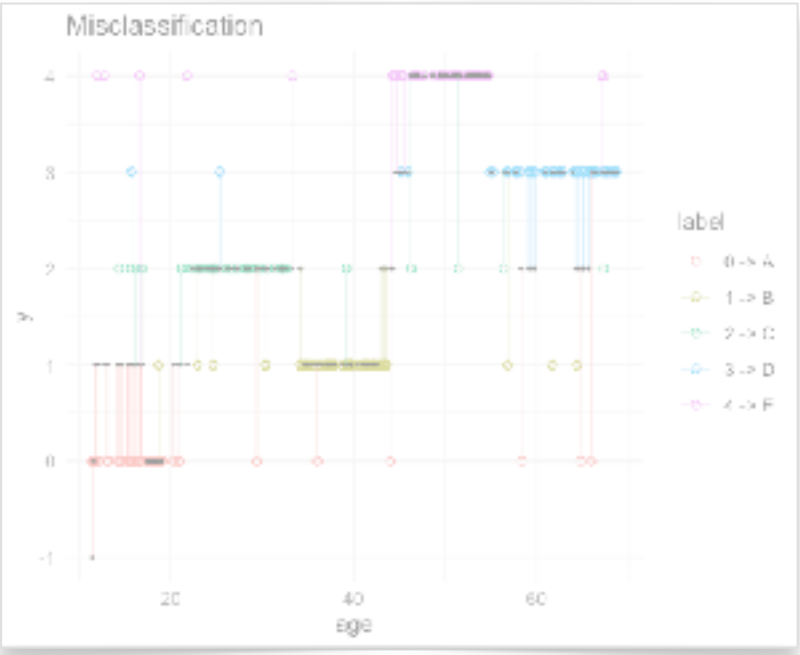
PREDICT CATEGORICAL DATA

Polynomial regressions may be fitted on categorical data ... many parameters…
but they are not interpretable… and **not robust to no**...

We need something different !

```
###### Polynomial re
model <- lm(y
```





Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

LOGISTIC REGRESSION

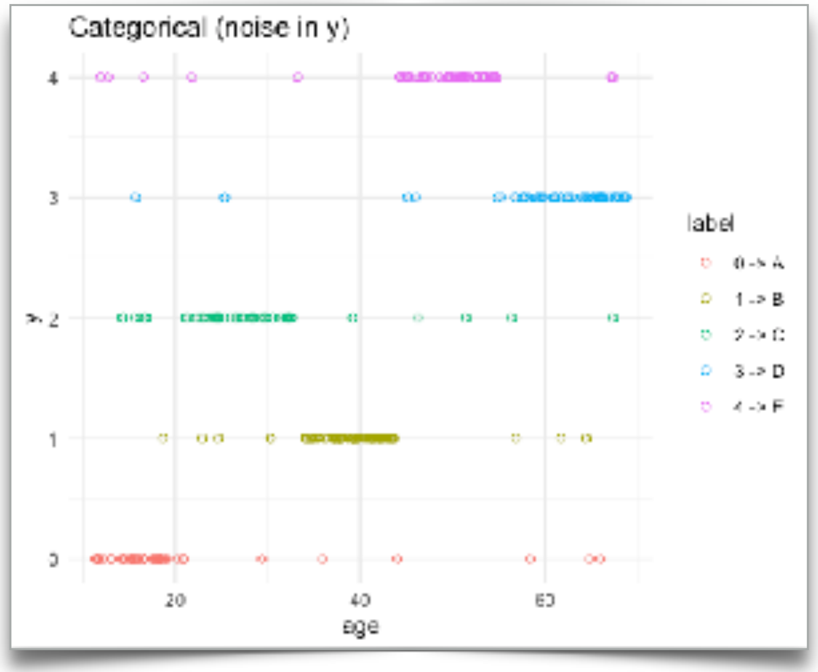
EMMA BEAUXIS-AUSSALET

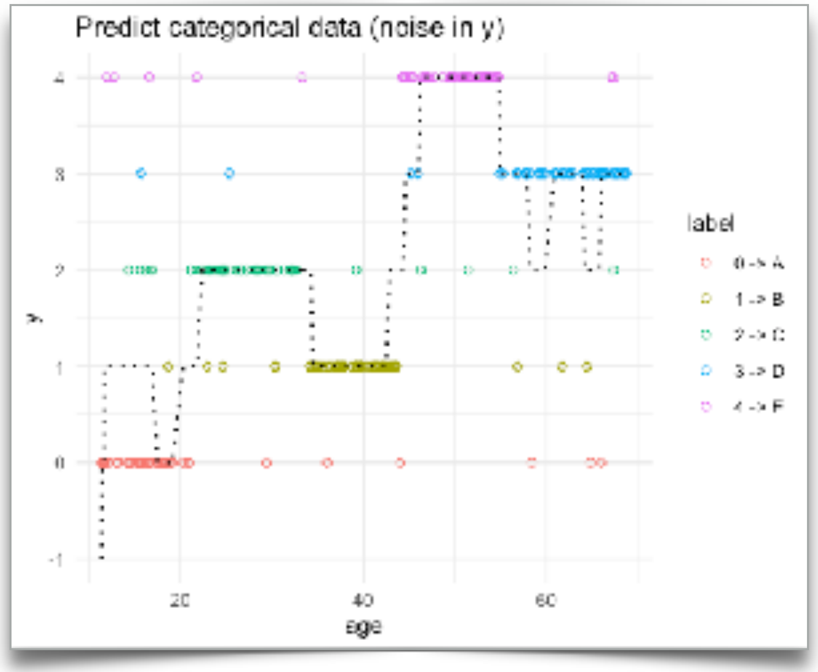
e.m.a.l.beauxis@hva.nl

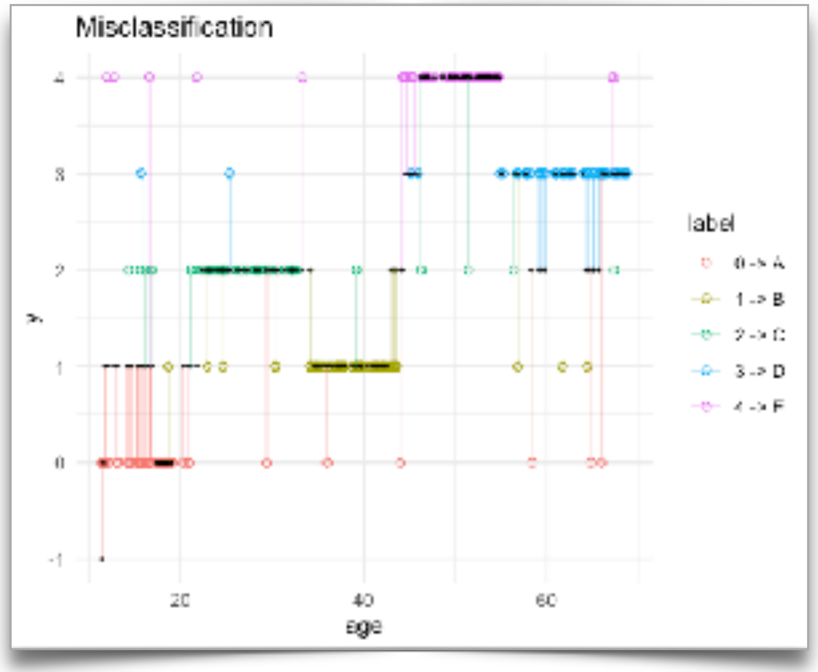
REGRESSION ISSUES

With multiclass data, linear or polynomial regressions have **large bias**.

Such bias depend on how numerical class labels are ordered (arbitrarily), and on random noise.



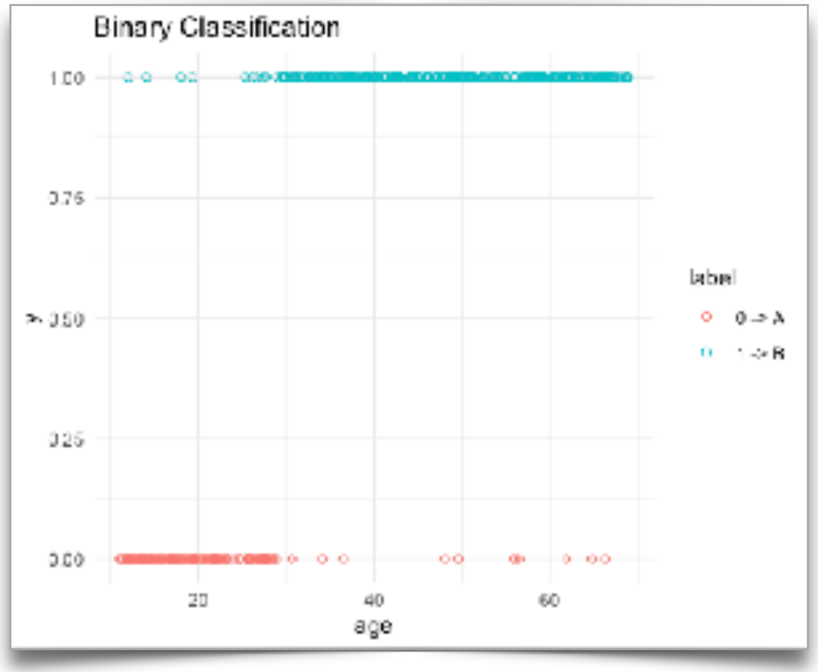


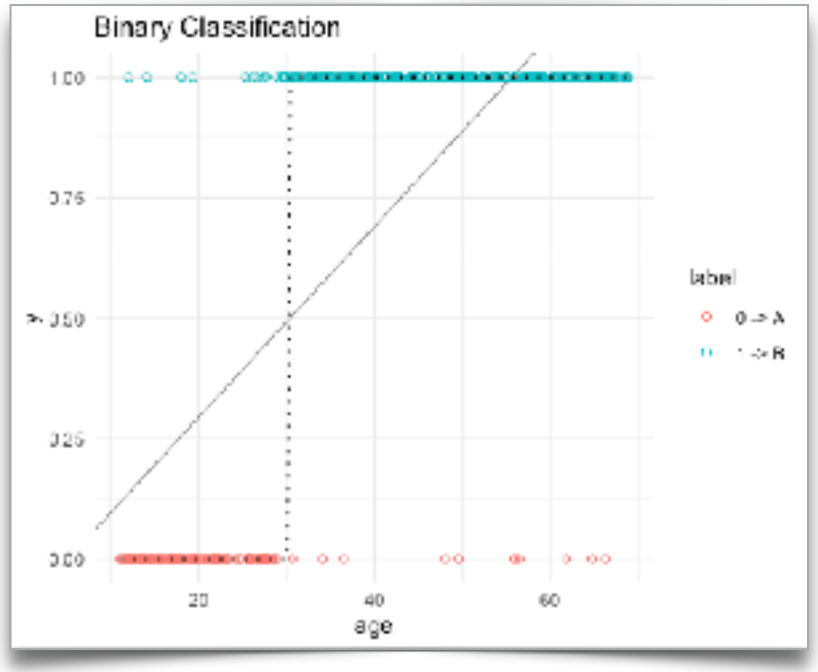


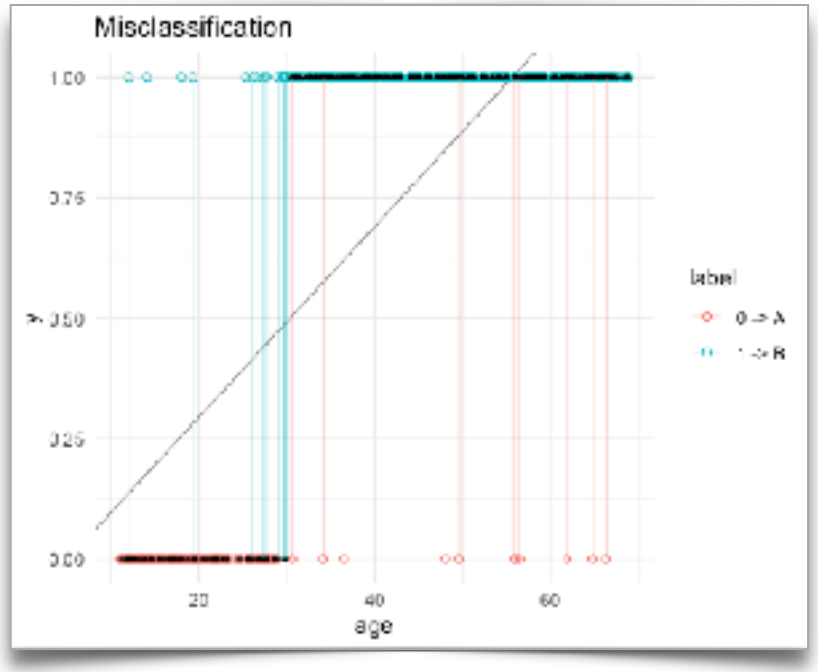
REGRESSION ISSUES

Predictions from **linear or polynomial regressions are unbounded**, thus hard to interpret.

For instance, such models can **predict a class that does not exist**.





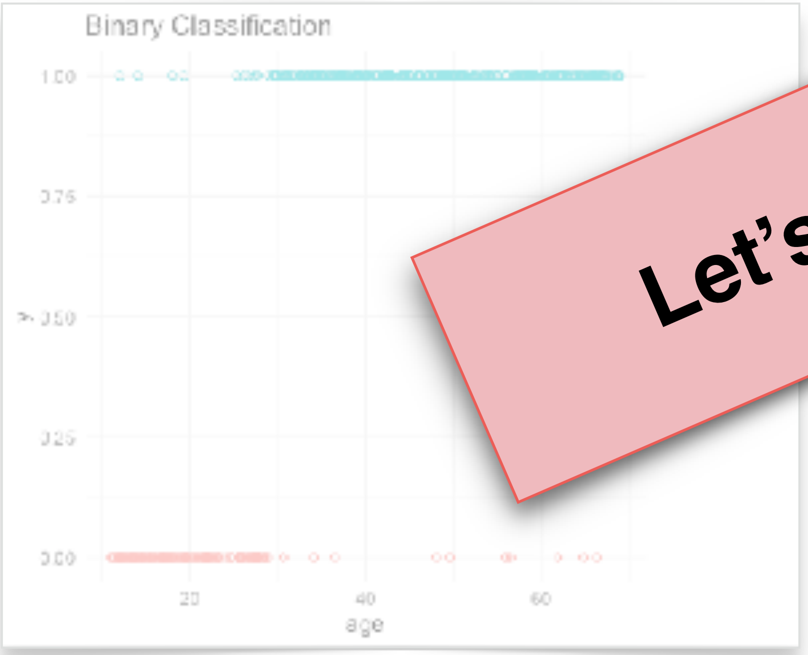


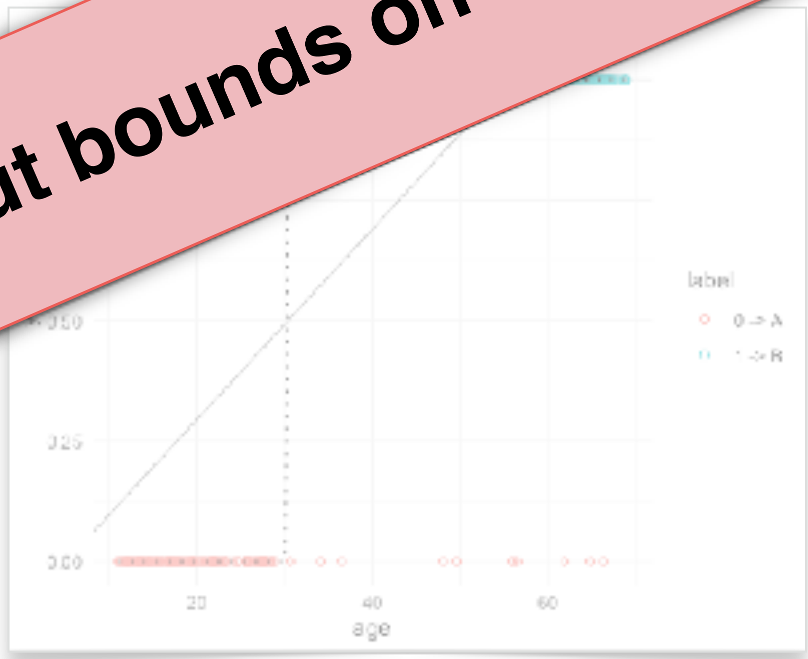
REGRESSION ISSUES

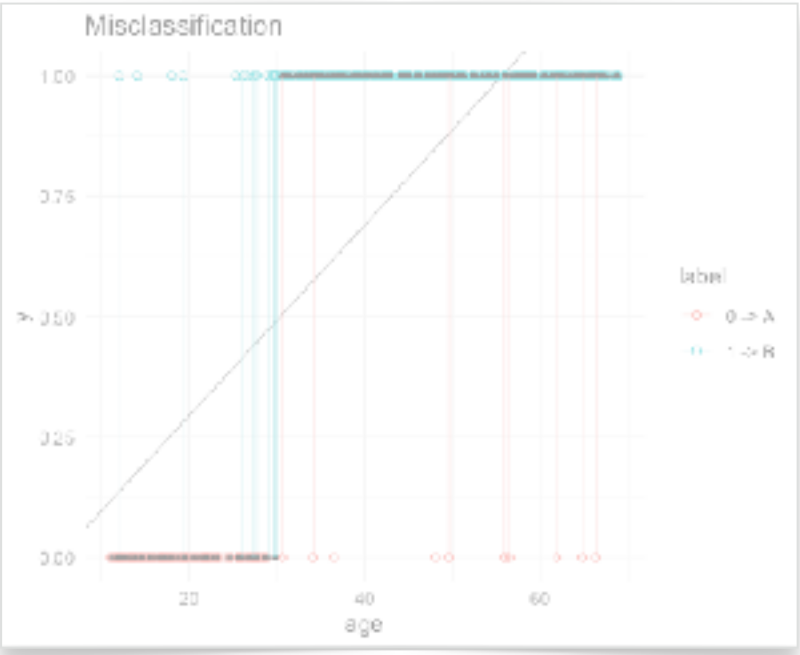
Predictions from **linear or polynomial regressions are unbou** thus hard to interpret.

For instance, such models can **predict a cla** **xist**.

Let’s put bounds on the predictions….

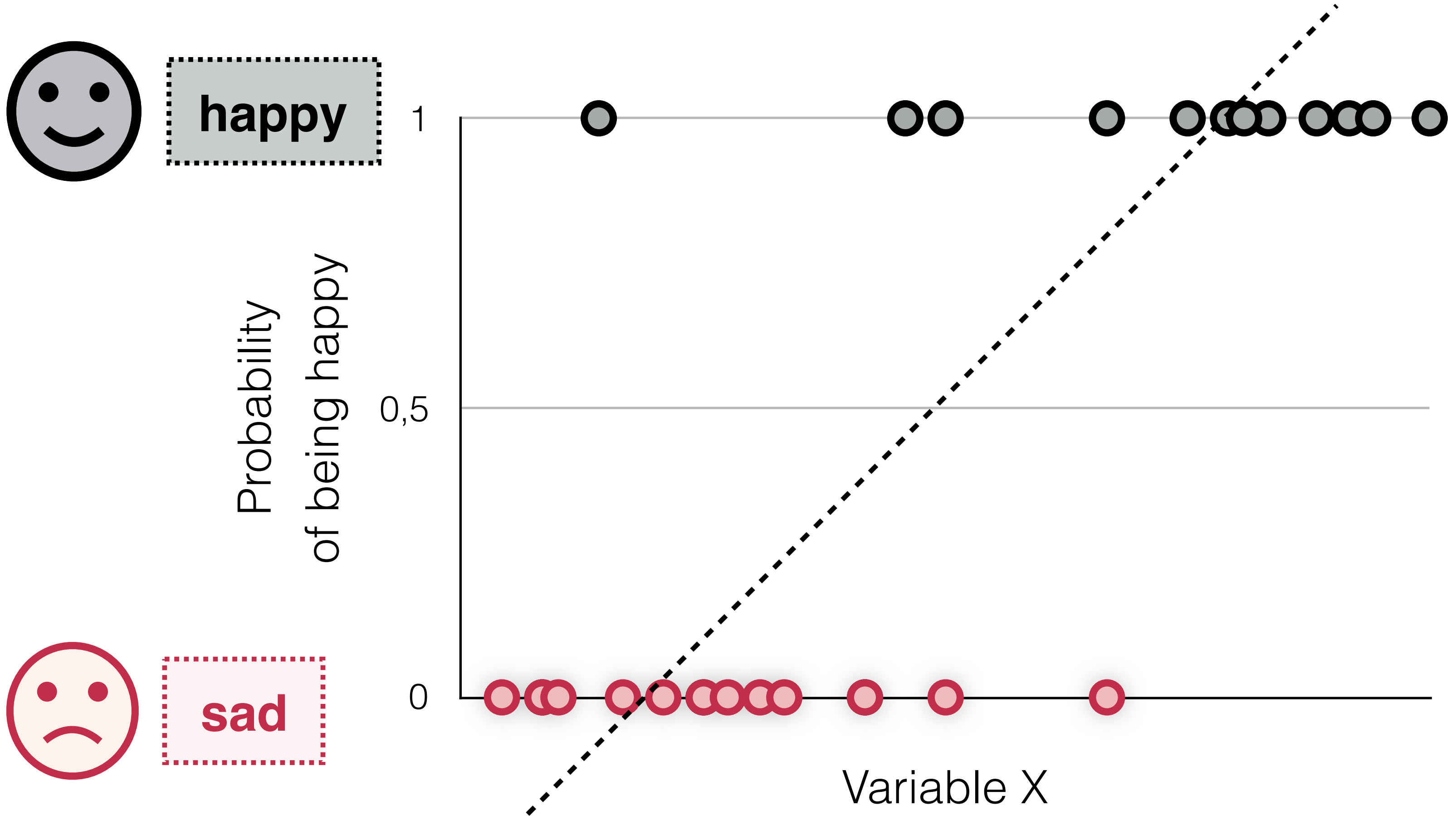






LOGISTIC REGRESSION

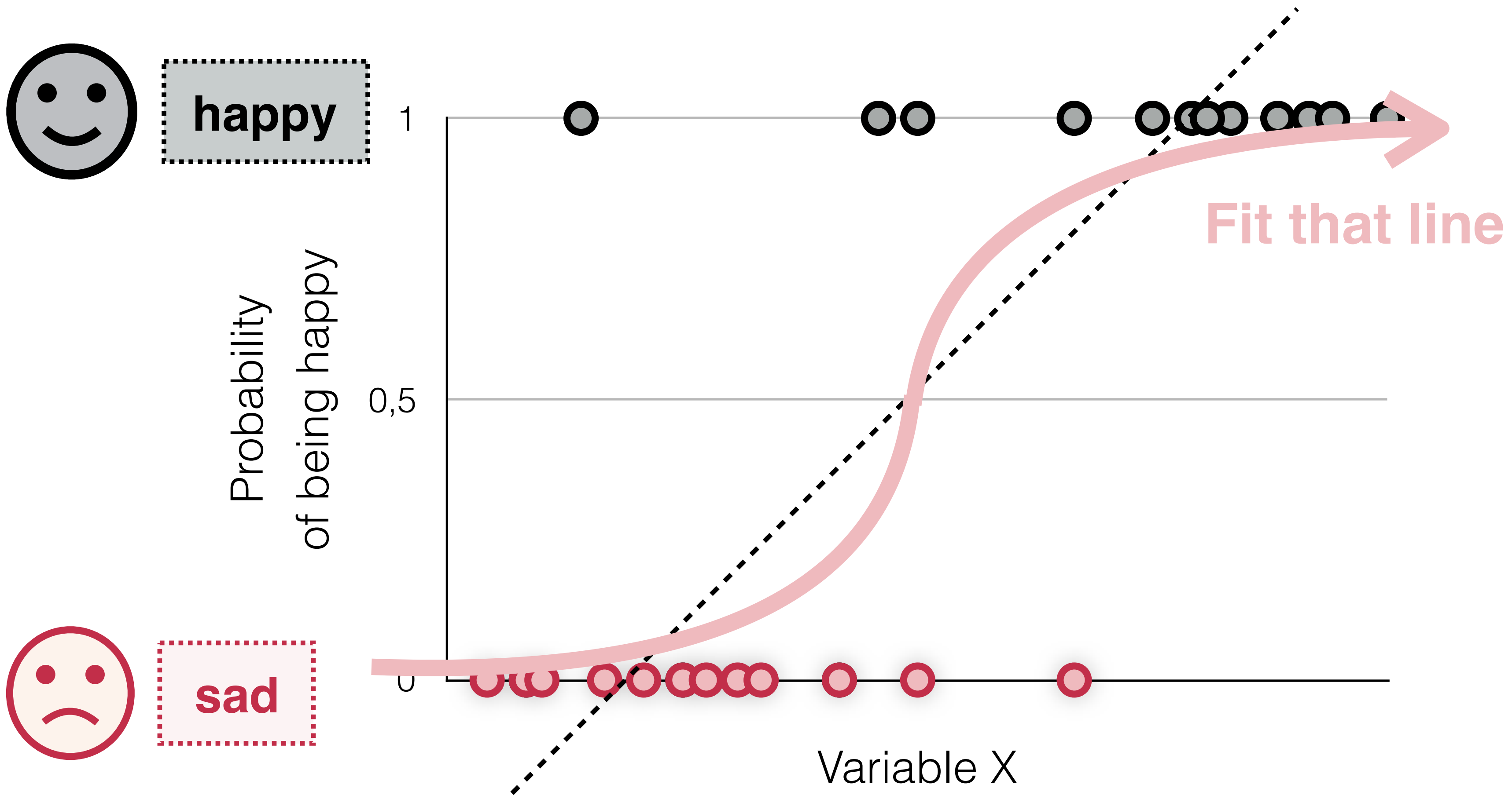
We can transform linear regressions into a **sigmoid function**, a s-shaped line, using the slope and intercept of the straight line.



$$\hat{y} = a\mathbf{x} + b$$

LOGISTIC REGRESSION

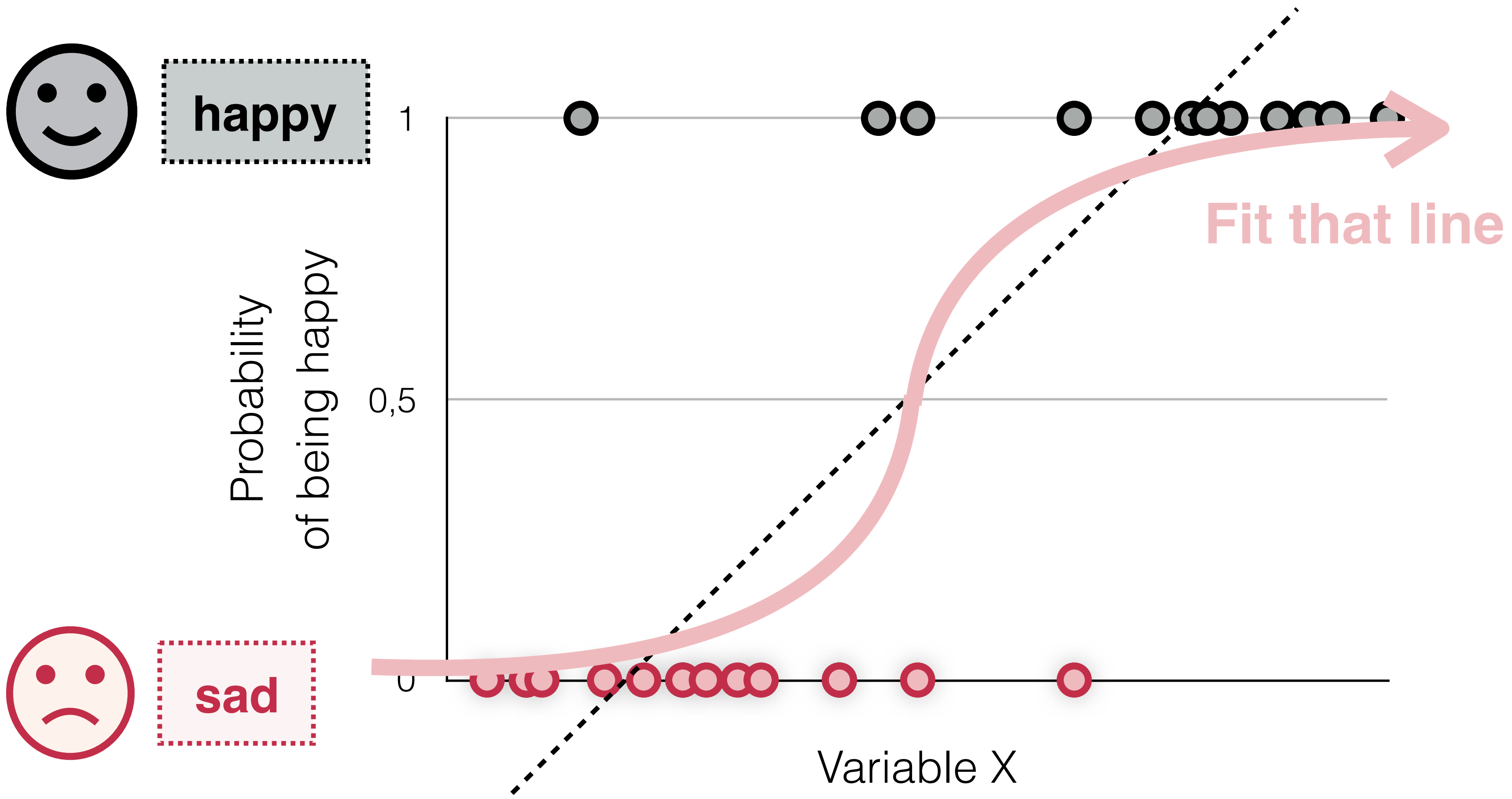
We can transform linear regressions into a **sigmoid function**, a s-shaped line, using the slope and intercept of the straight line.



$$\hat{y} = a\mathbf{x} + b$$

LOGISTIC REGRESSION

We can transform linear regressions into a **sigmoid function**, a s-shaped line, using the slope and intercept of the straight line.

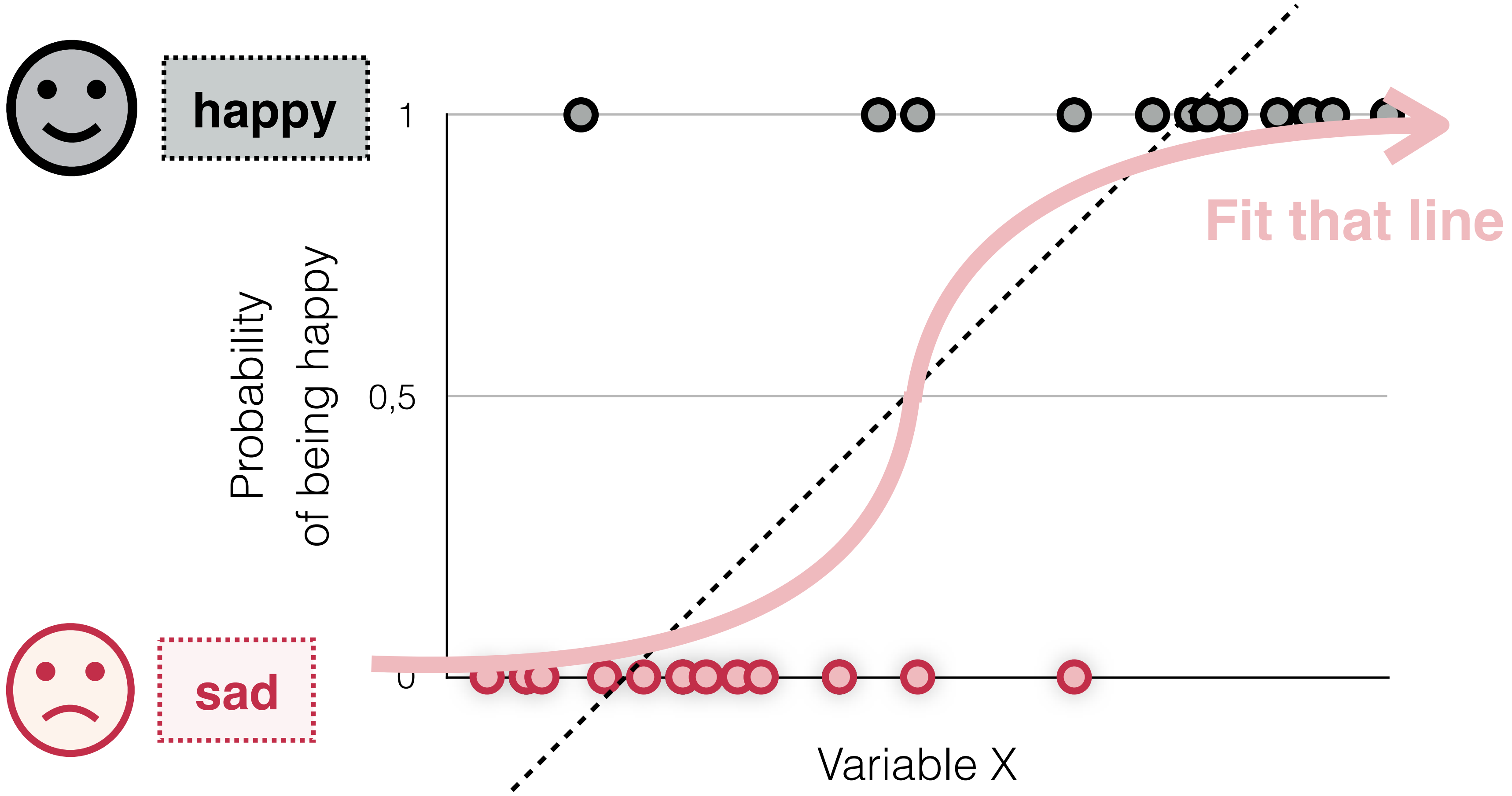


$$\hat{y} = a\mathbf{x} + b$$

$$\widehat{P}\,(happy) = \frac{1}{1 + e^{-(a\mathbf{x}+b)}}$$

LOGISTIC REGRESSION

We can transform linear regressions into a **sigmoid function**, a s-shaped line, using the slope and intercept of the straight line.

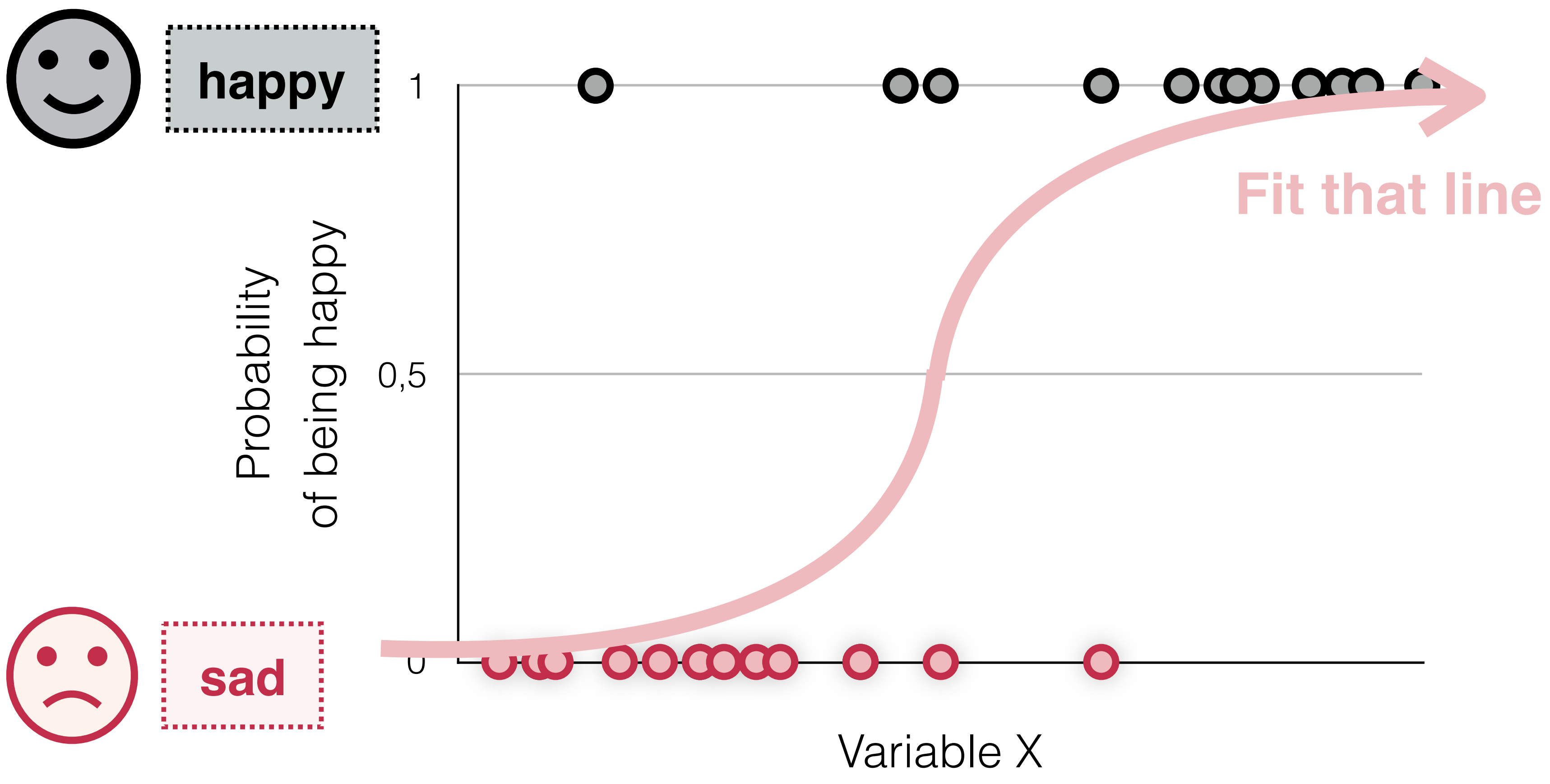


$$\hat{y} = a\mathbf{x} + b$$

$$\widehat{P}\,(happy) = \frac{1}{1 + e^{-\boxed{(a\mathbf{x}+b)}}}$$

LOGISTIC REGRESSION

The sigmoid function has **bonded results** that can be **interpreted as probabilities** of class membership (the probability that a data point belongs to a class).



$$\widehat{P}\,(happy) = \frac{1}{1 + e^{-(a\mathbf{x}+b)}}$$

LOGISTIC REGRESSION

$$\widehat{P}(neutral) = \frac{1}{1 + e^{-(\alpha_0 + \alpha_1 \mathbf{x})}}$$

$$\widehat{P}(happy) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \mathbf{x})}}$$

$$\widehat{P}(sad) = \frac{1}{1 + e^{-(\zeta_0 + \zeta_1 \mathbf{x})}}$$

Multiclass problems require **one regression model per class**.

The predicted class is the one having the maximum probability.

LOGISTIC REGRESSION

$$\widehat{P}(neutral) = \frac{1}{1 + e^{-(\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots)}}$$

$$\widehat{P}(happy) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots)}}$$

$$\widehat{P}(sad) = \frac{1}{1 + e^{-(\zeta_0 + \zeta_1 x_1 + \zeta_2 x_2 + \dots)}}$$

Multivariate logistic regressions add variables **within the embedded linear regression**.

Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

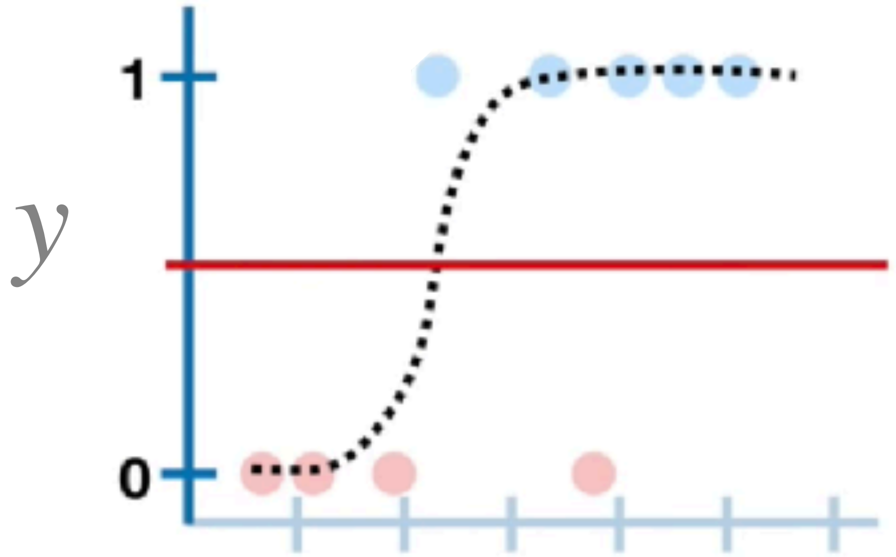
DEEP DIVE IN LOGISTIC REGRESSION

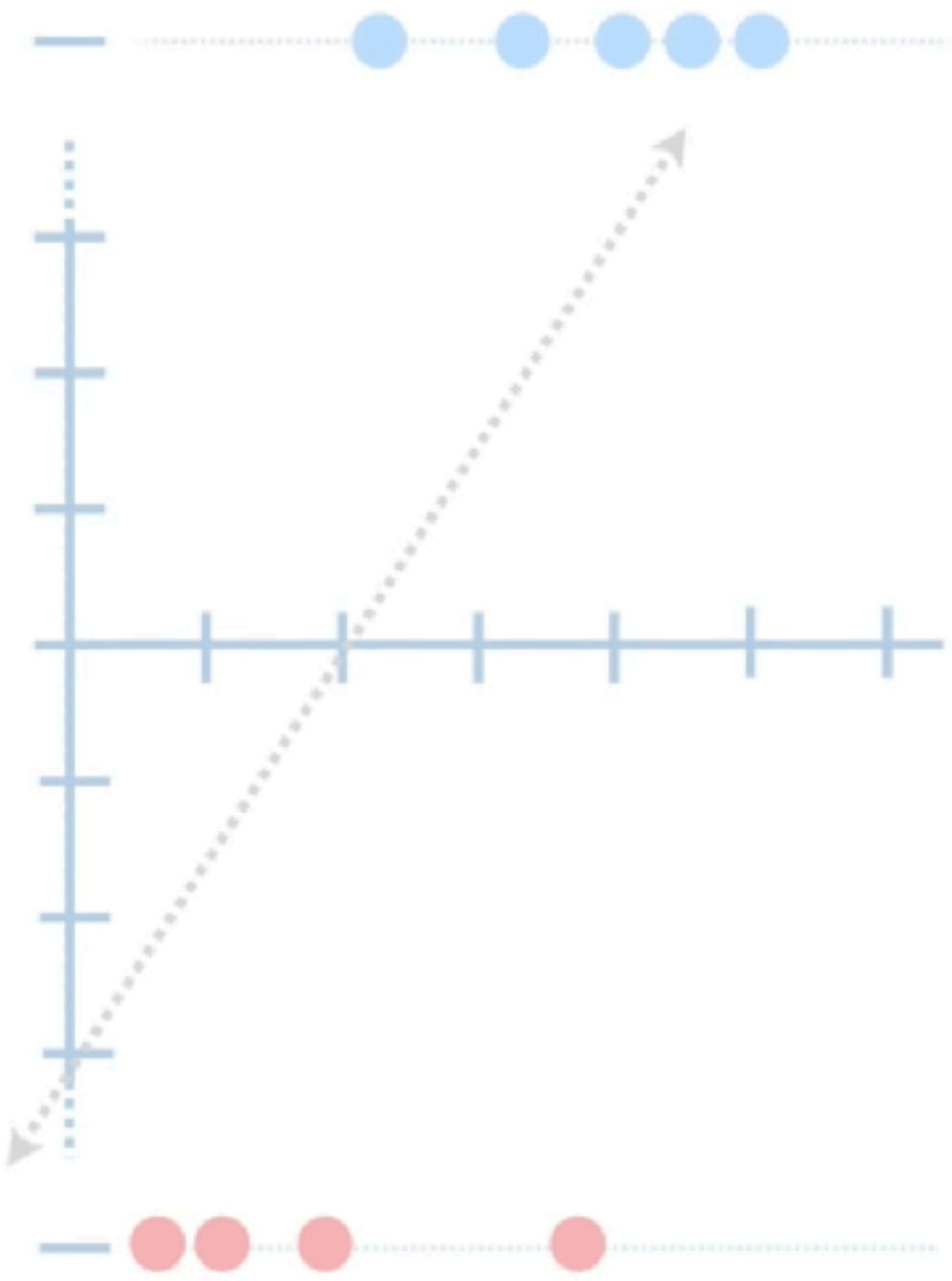
EMMA BEAUXIS-AUSSALET

e.m.a.l.beauxis@hva.nl

HOW DOES IT WORK ?

It is just fitting a **straight regression line**,
…but after transforming the data.

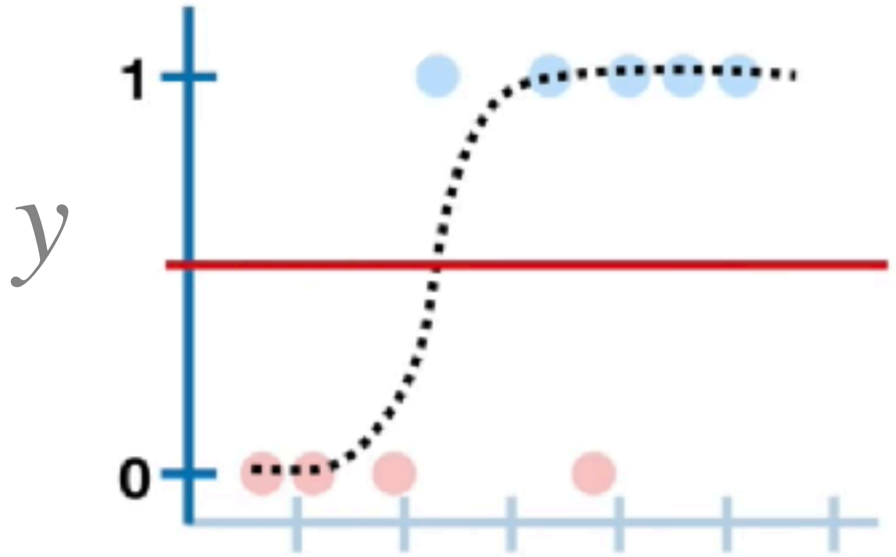


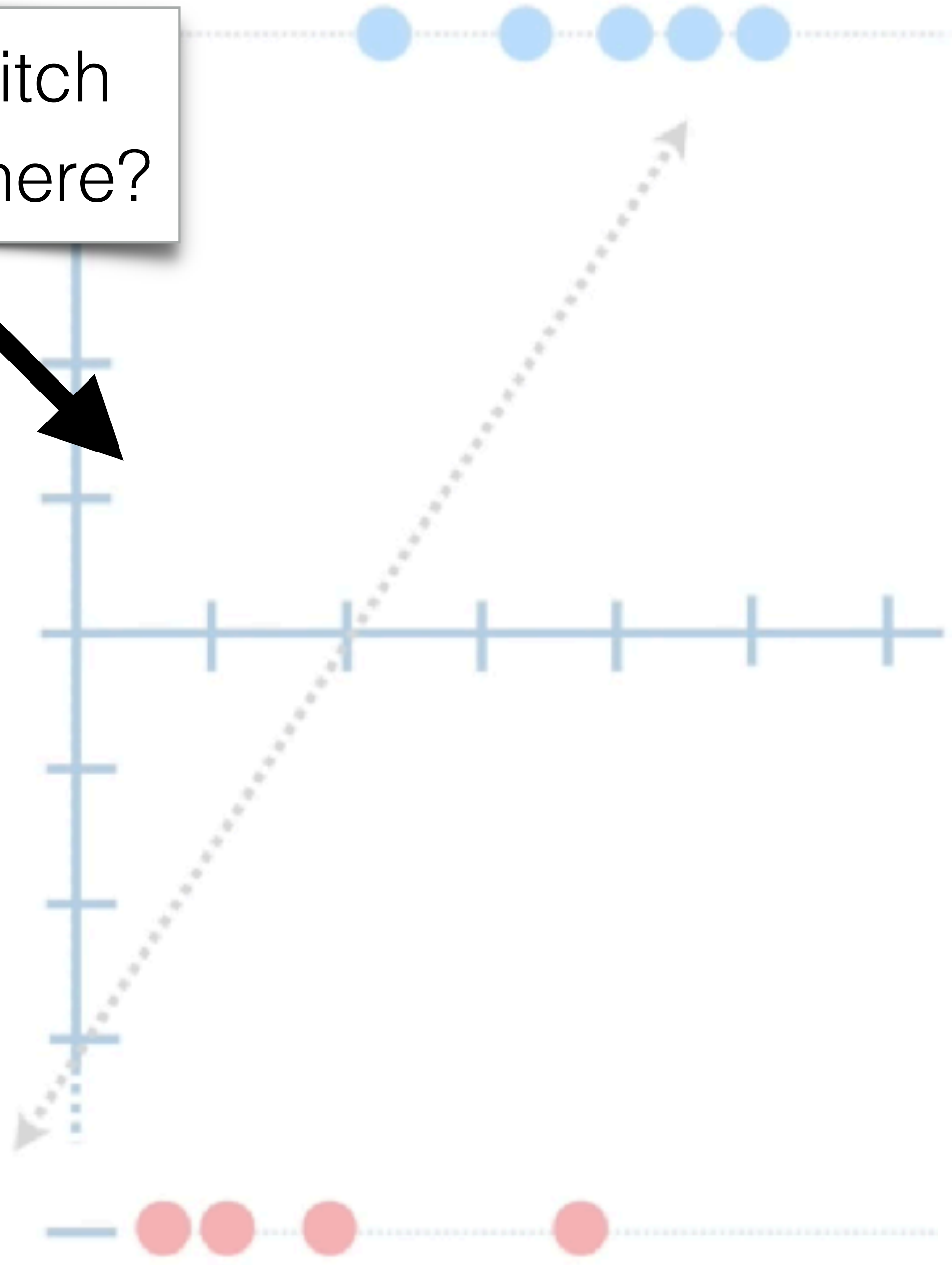


https://www.youtube.com/watch?v=vN5cNN2-HWE

HOW DOES IT WORK ?

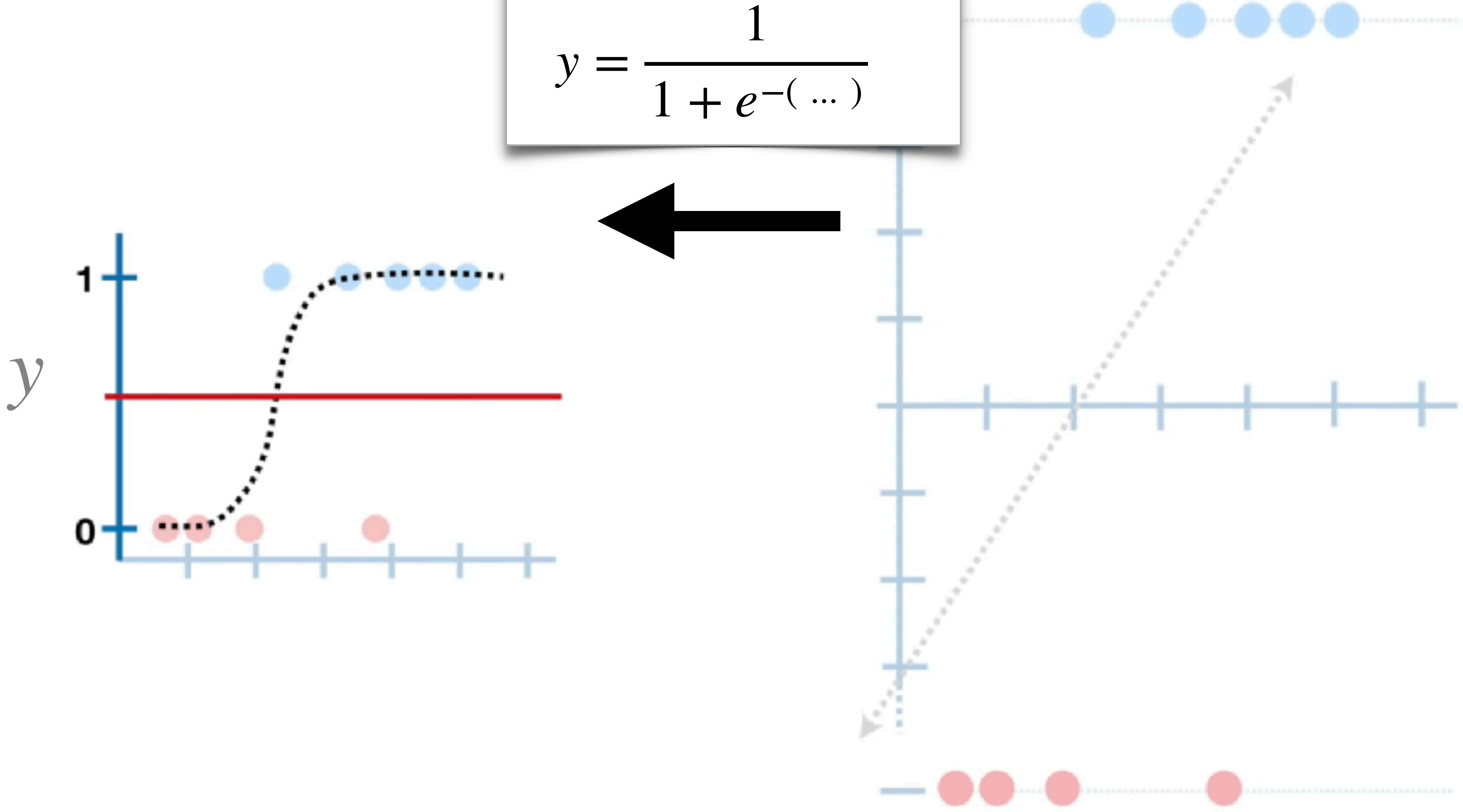
How do we switch from there to there?





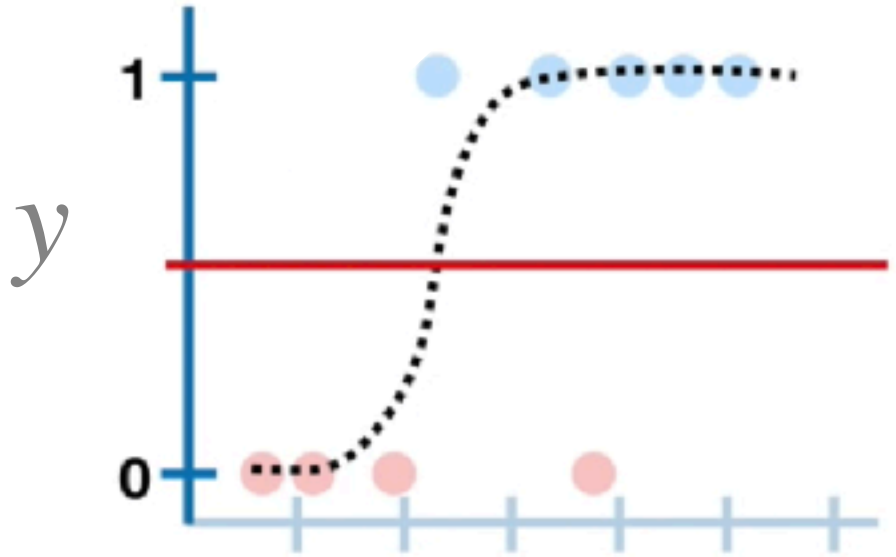
HOW DOES IT WORK ?

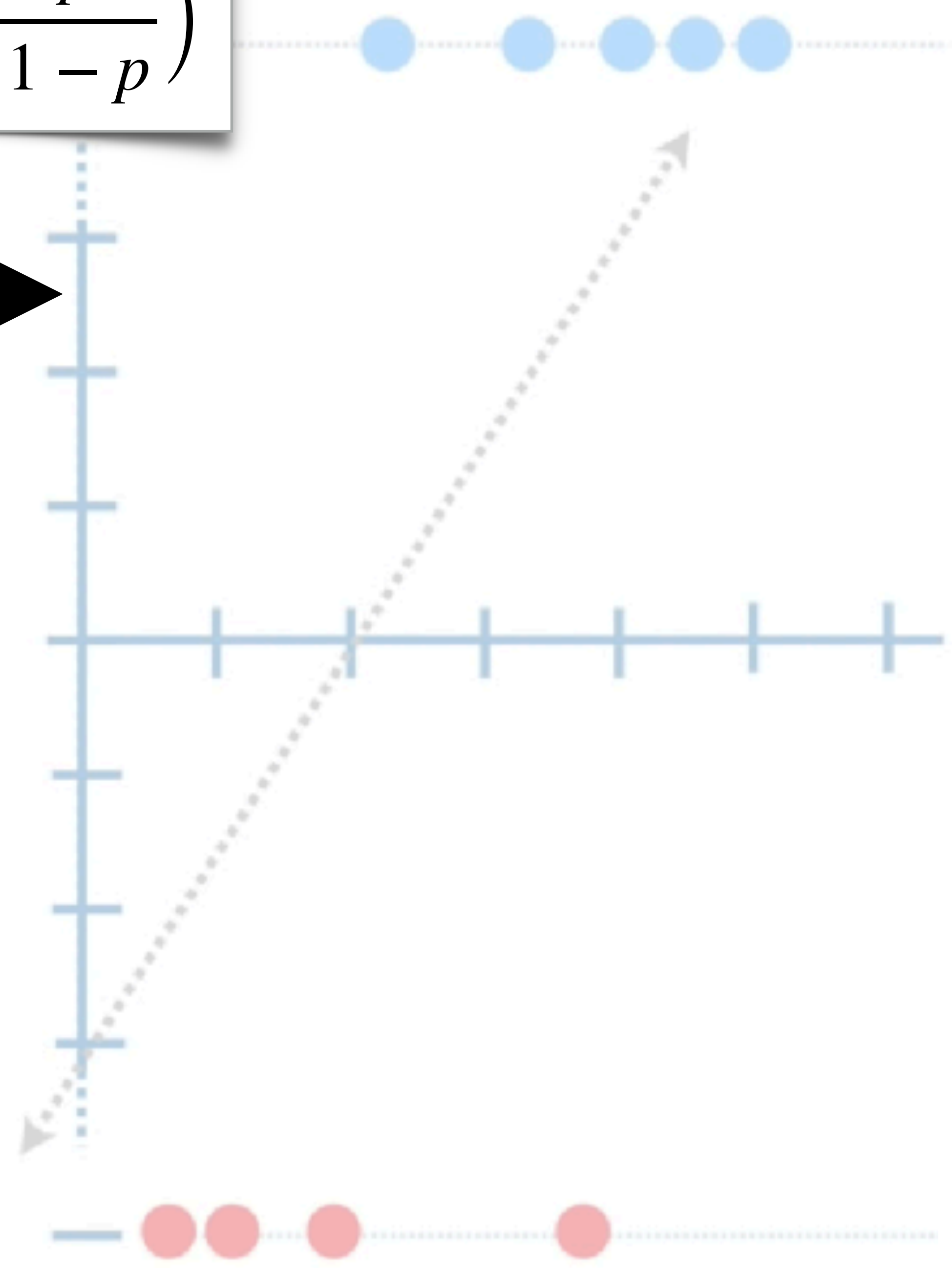
$$y = \frac{1}{1 + e^{-(\ \dots\ )}}$$



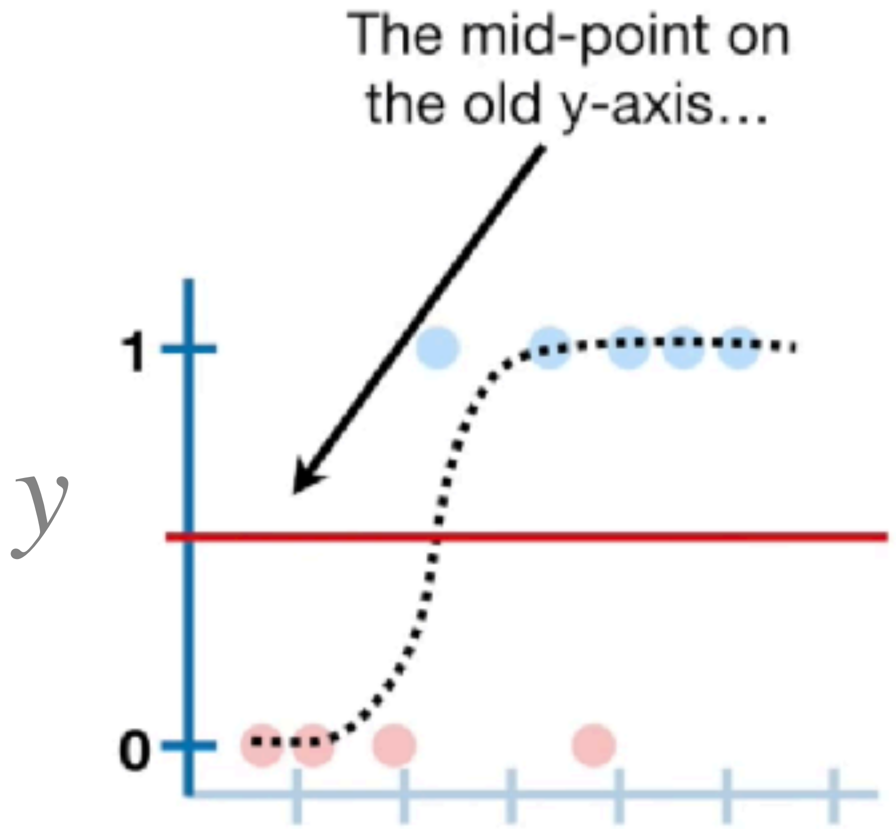
HOW DOES IT WORK ?

$$y_{trans} = log\left(\frac{p}{1-p}\right)$$

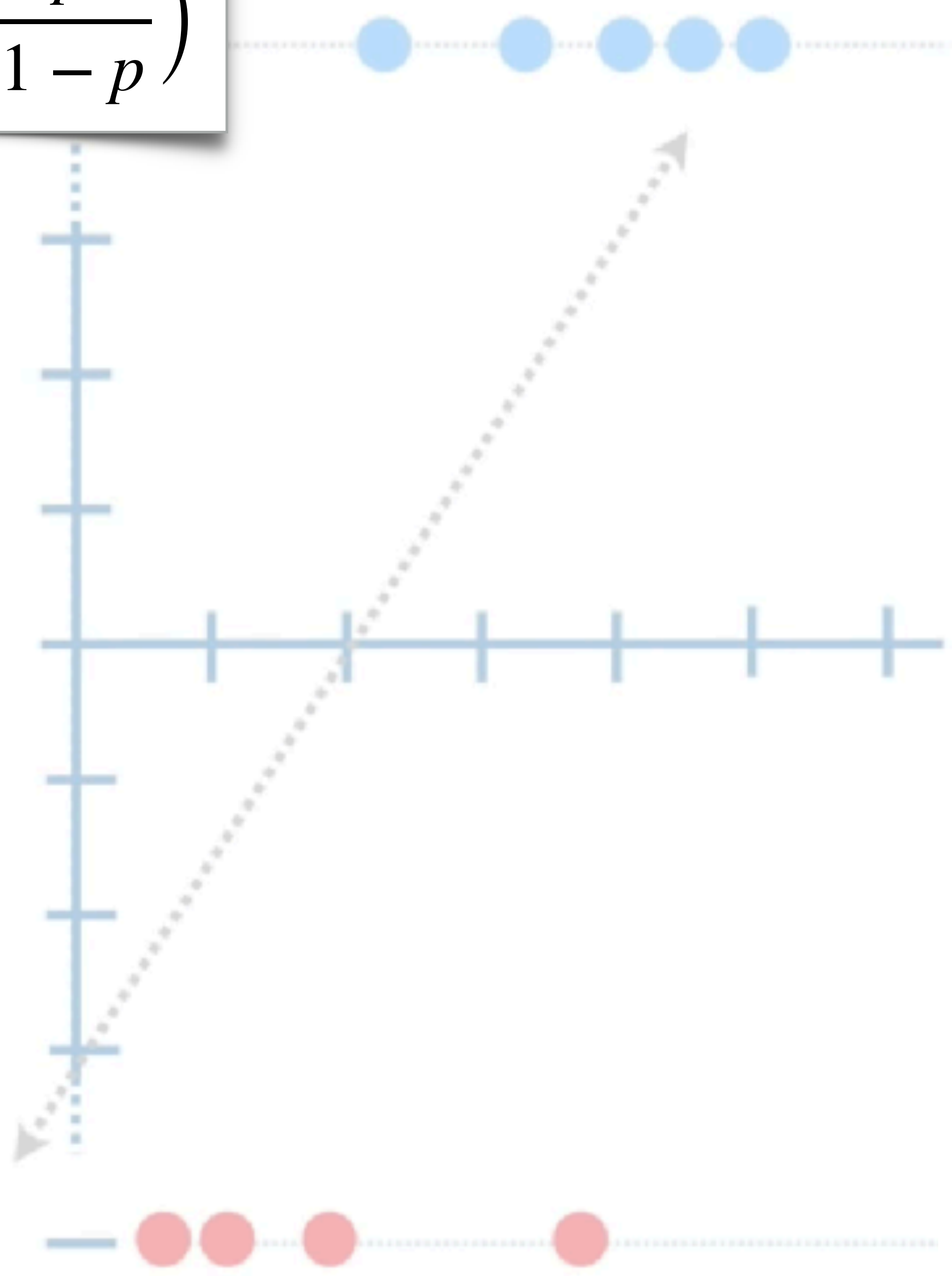




HOW DOES IT WORK ?

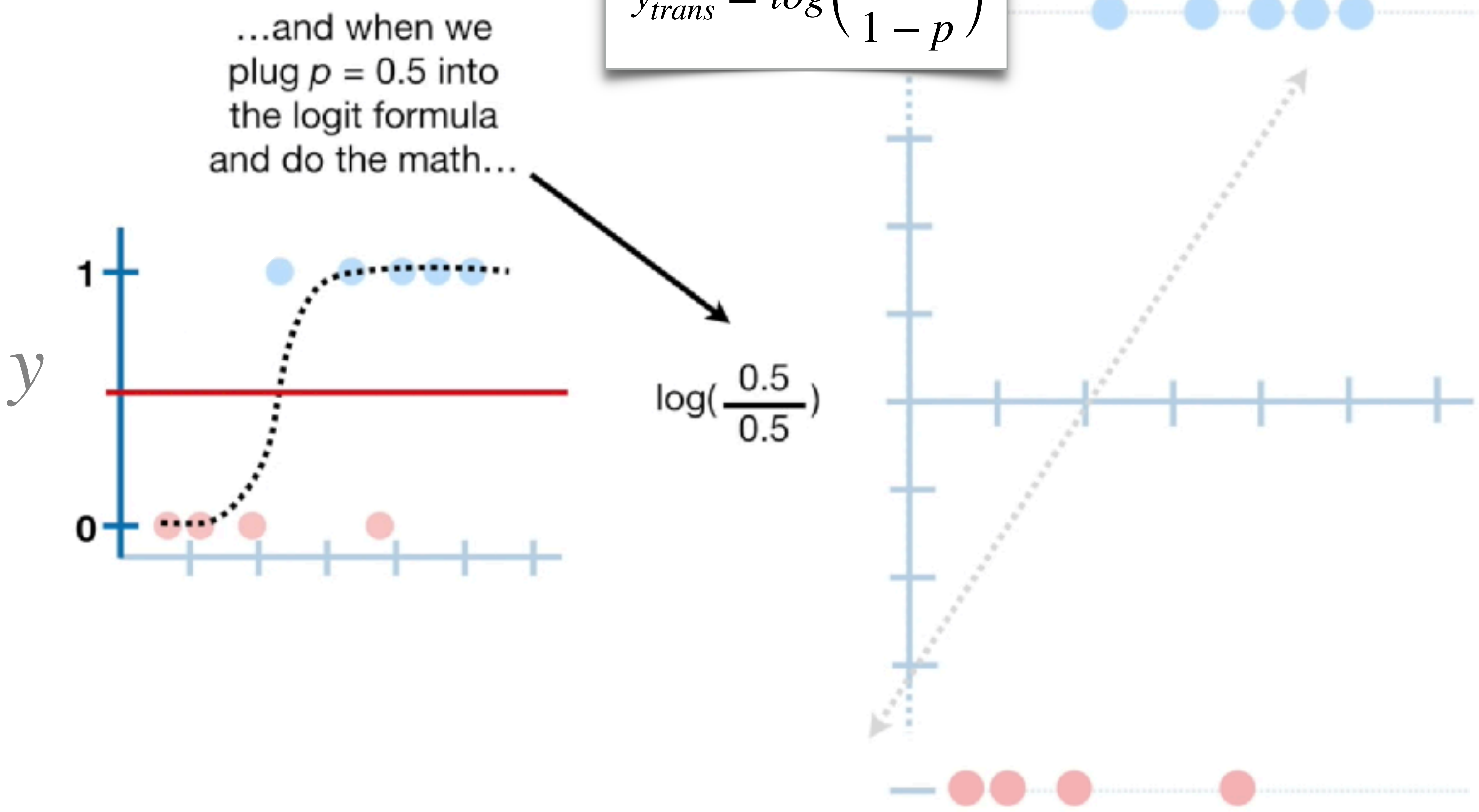


$$y_{trans} = log\left(\frac{p}{1-p}\right)$$



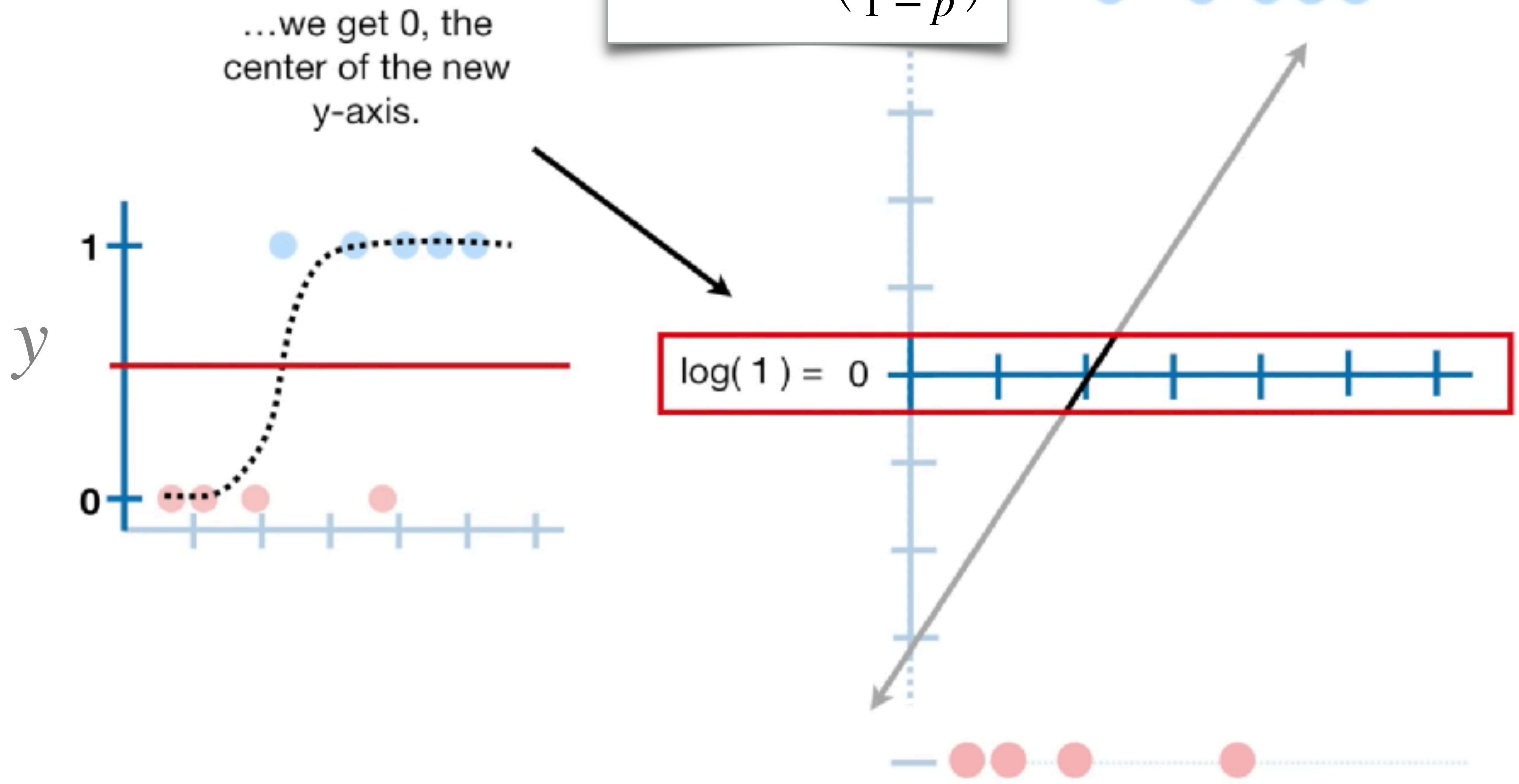
HOW DOES IT WORK ?

$$y_{trans} = log\left(\frac{p}{1-p}\right)$$

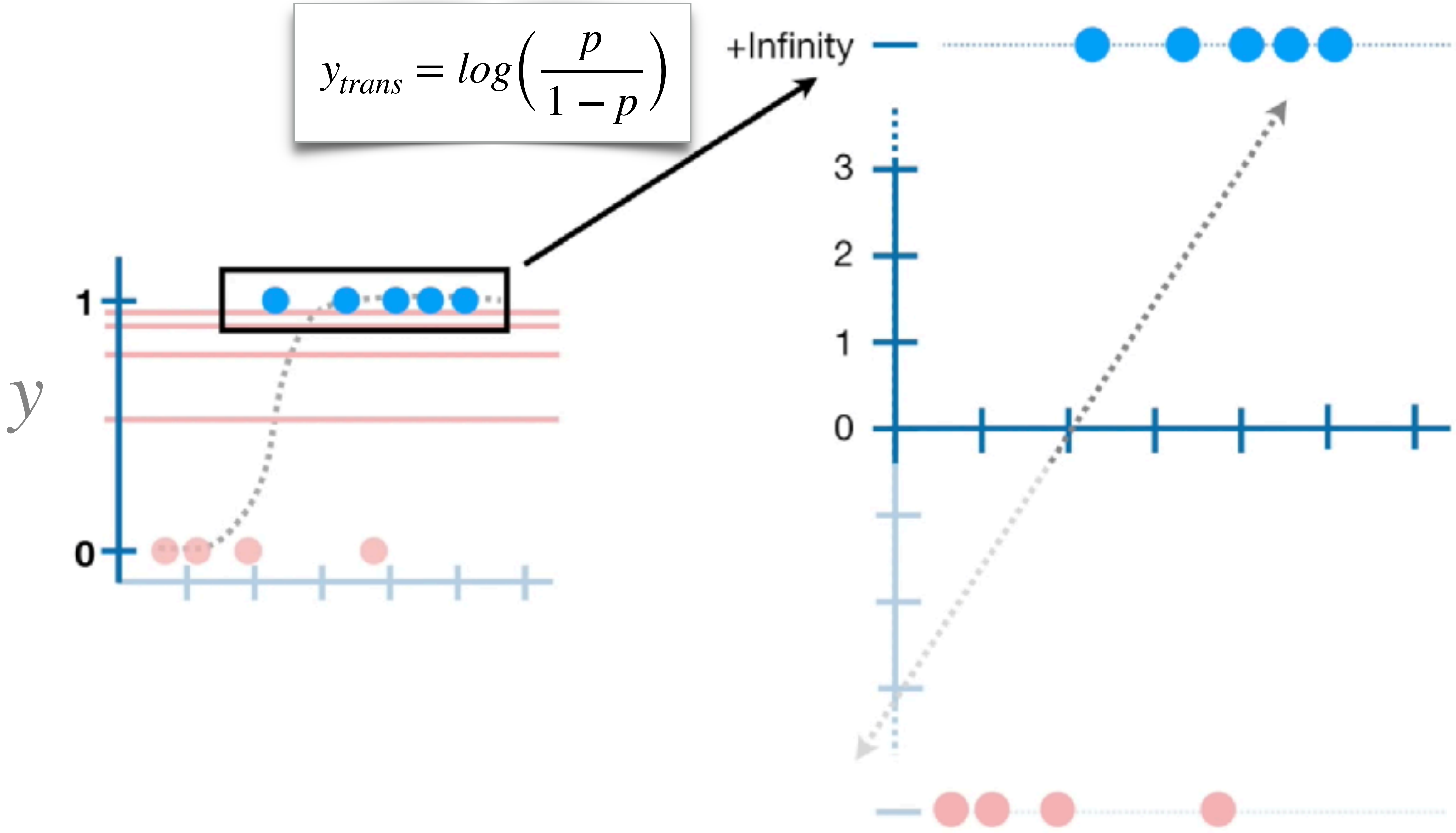


HOW DOES IT WORK ?

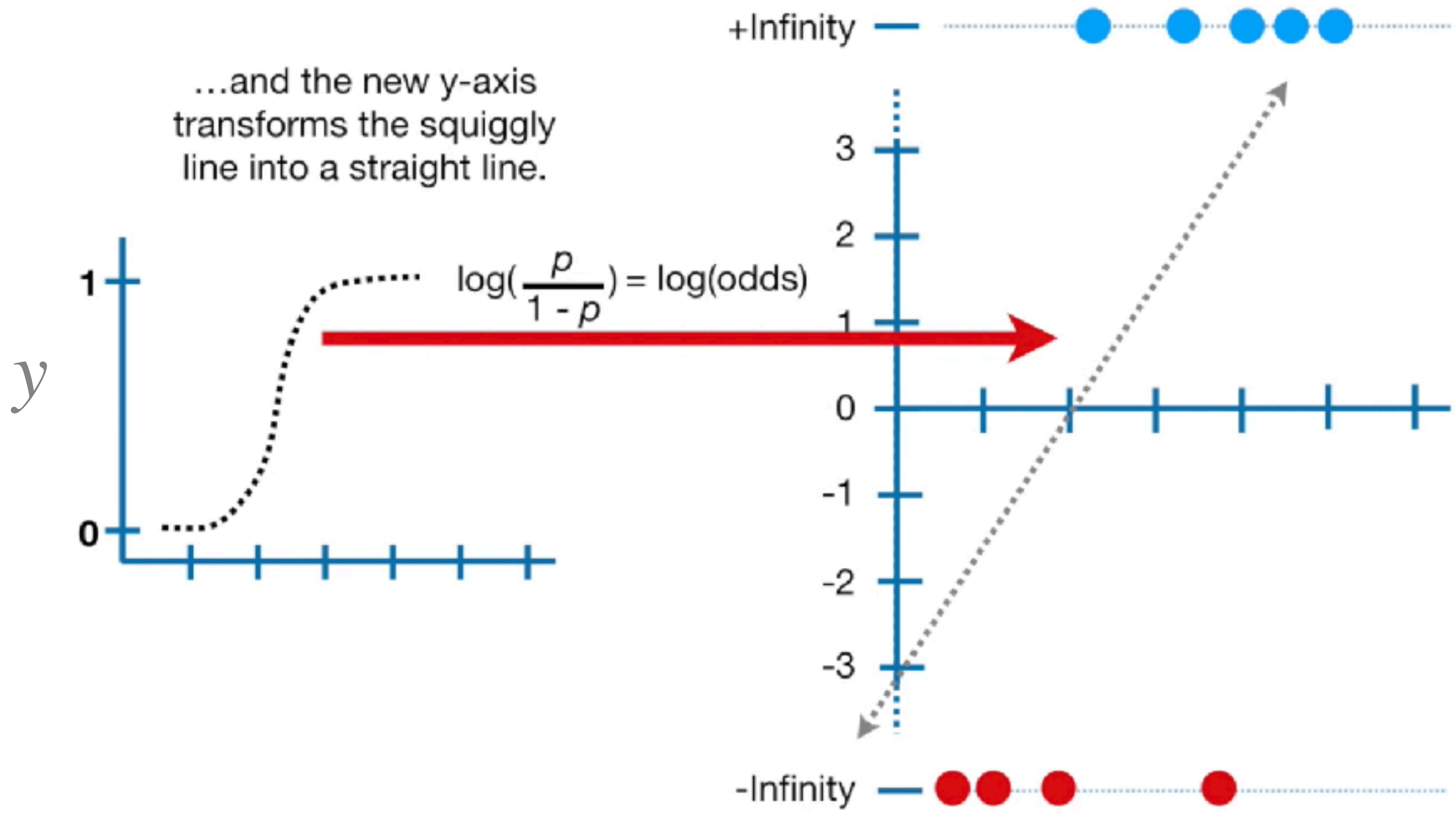
$$y_{trans} = log\left(\frac{p}{1-p}\right)$$



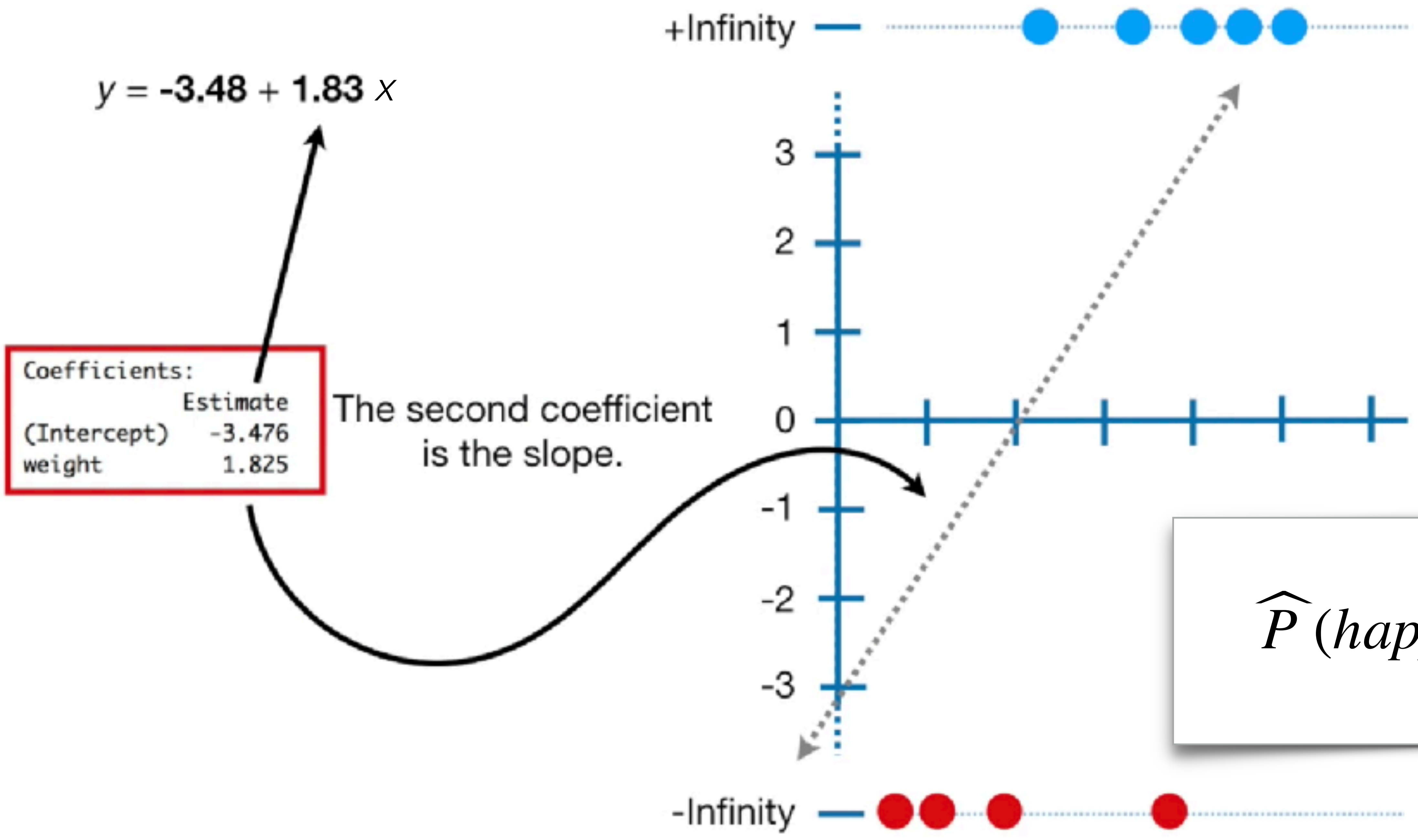
HOW DOES IT WORK ?



HOW DOES IT WORK ?

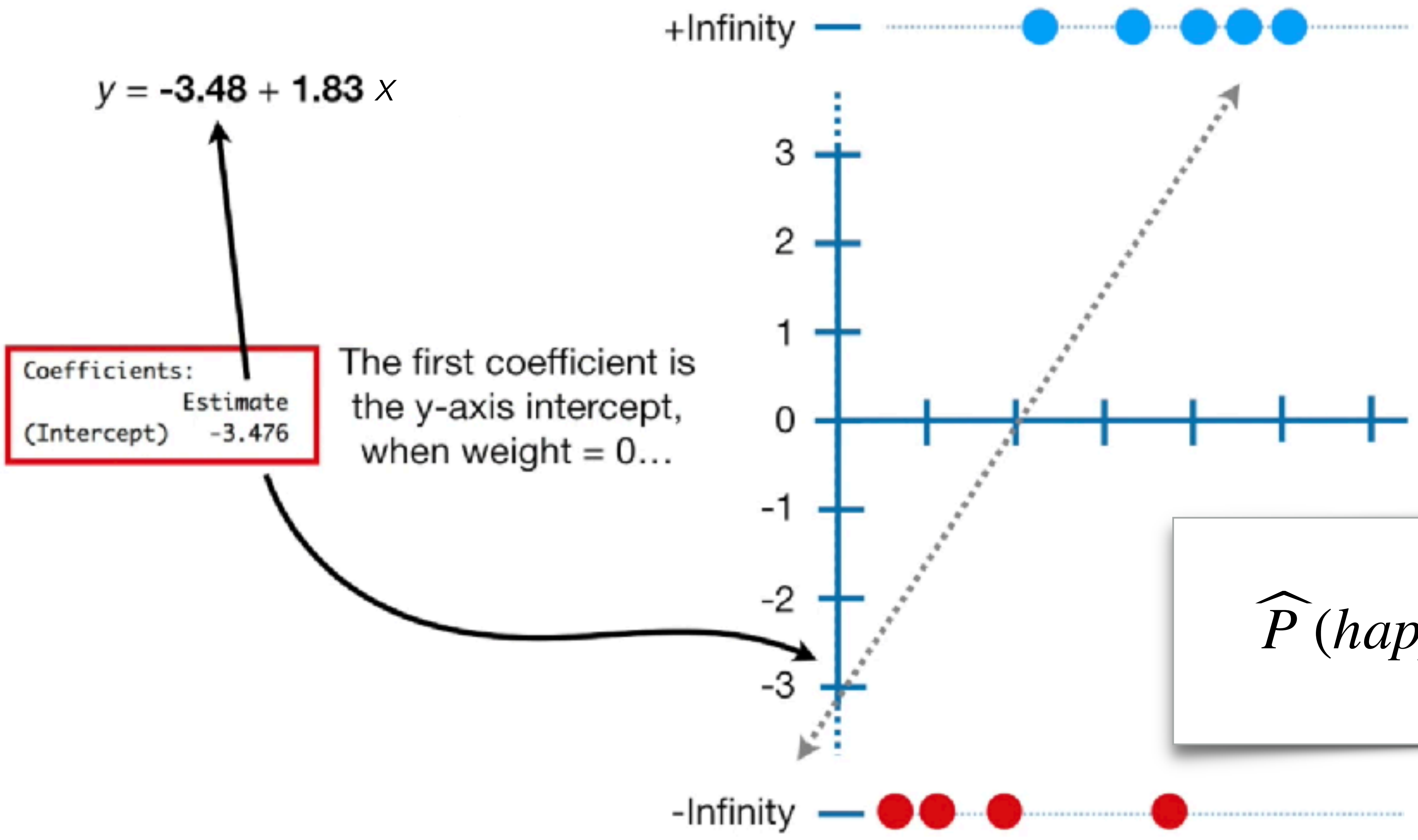


HOW DOES IT WORK ?



$$\widehat{P}\,(happy) = \frac{1}{1+e^{-(\boxed{a\mathbf{x}}+b\ )}}$$

HOW DOES IT WORK ?



$$\widehat{P}\,(happy) = \frac{1}{1 + e^{-(\,a\mathbf{x}+\boxed{b}\,)}}$$

Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

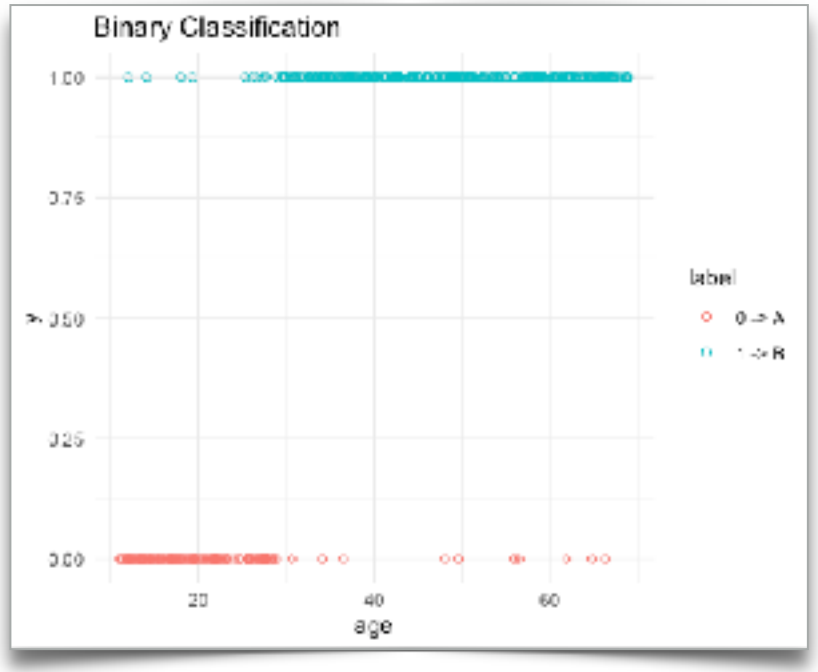
EXAMPLE

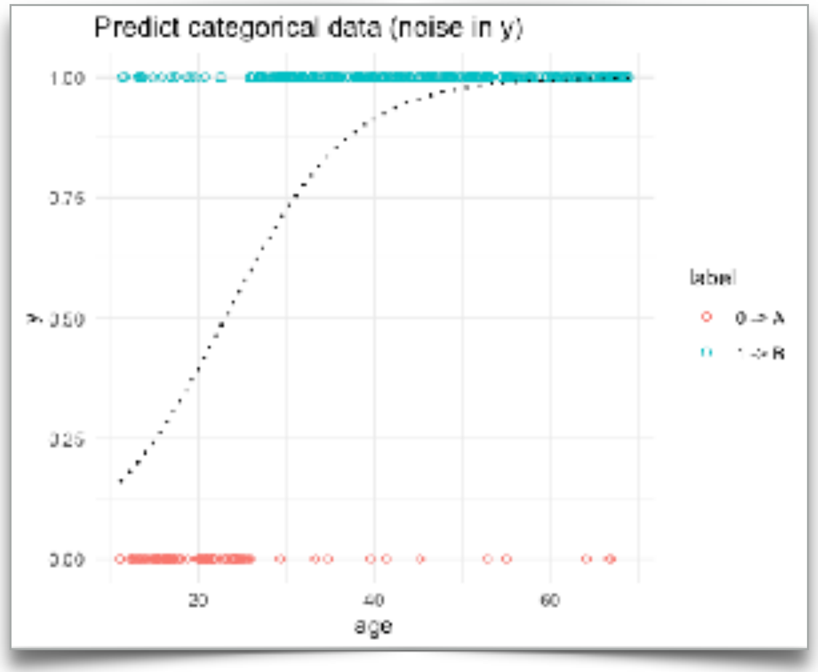
EMMA BEAUXIS-AUSSALET

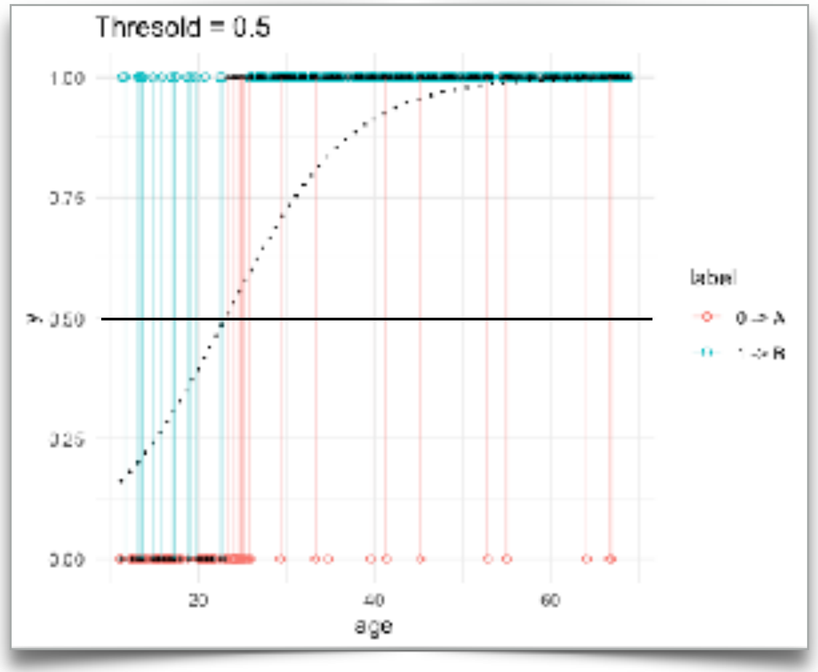
e.m.a.l.beauxis@hva.nl

EXAMPLE WITH BINARY DATA

```r
###### Logistic regression ######
model <- glm( formula = y ~ age,
              family = binomial,
              data = data)
```

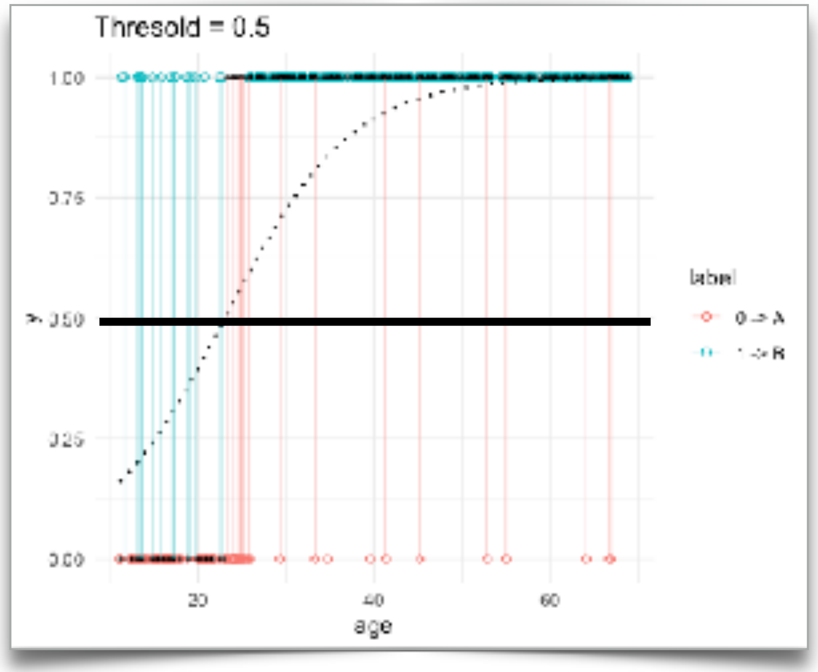


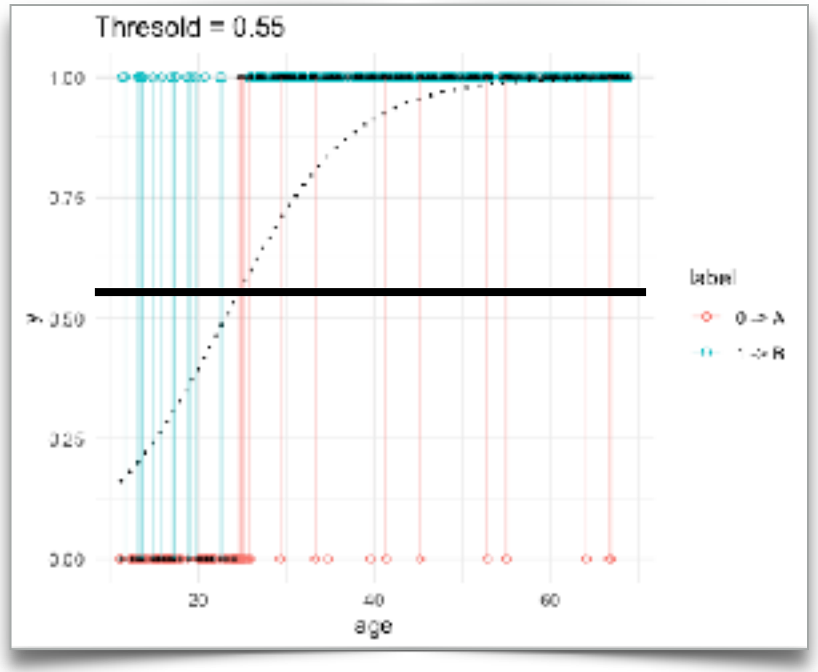


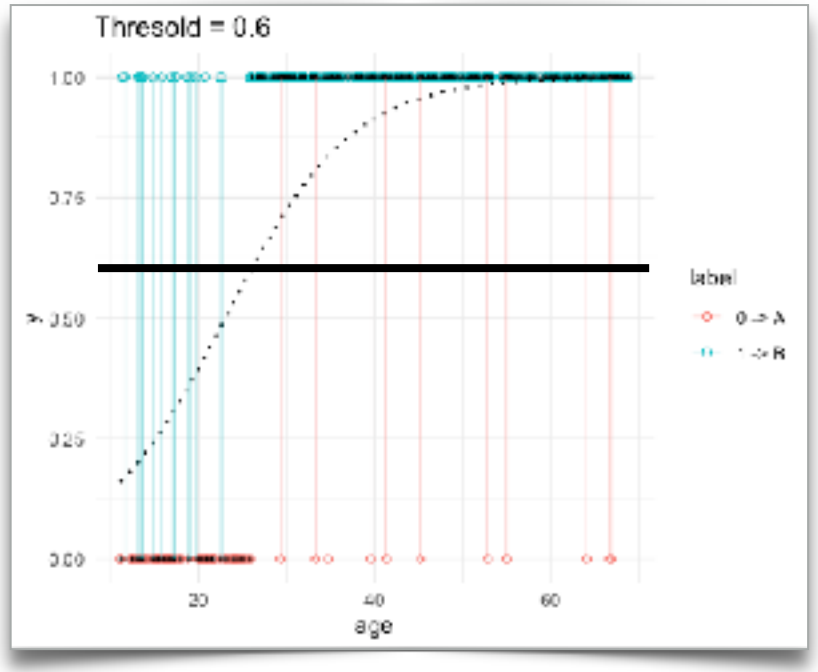


EXAMPLE WITH BINARY DATA

Different **thresholds can be used** to assign the final class to each data points.







Amsterdam University of Applied Sciences | DIGITAL SOCIETY SCHOOL

DATA-DRIVEN TRANSFORMATION

QUESTION / DISCUSSION