AIFS
AI INSTITUTE FOR FOOD SYSTEMS

AIBridge

Lecture 5

Let’s talk about the last lab!

Let's talk about the last lab!

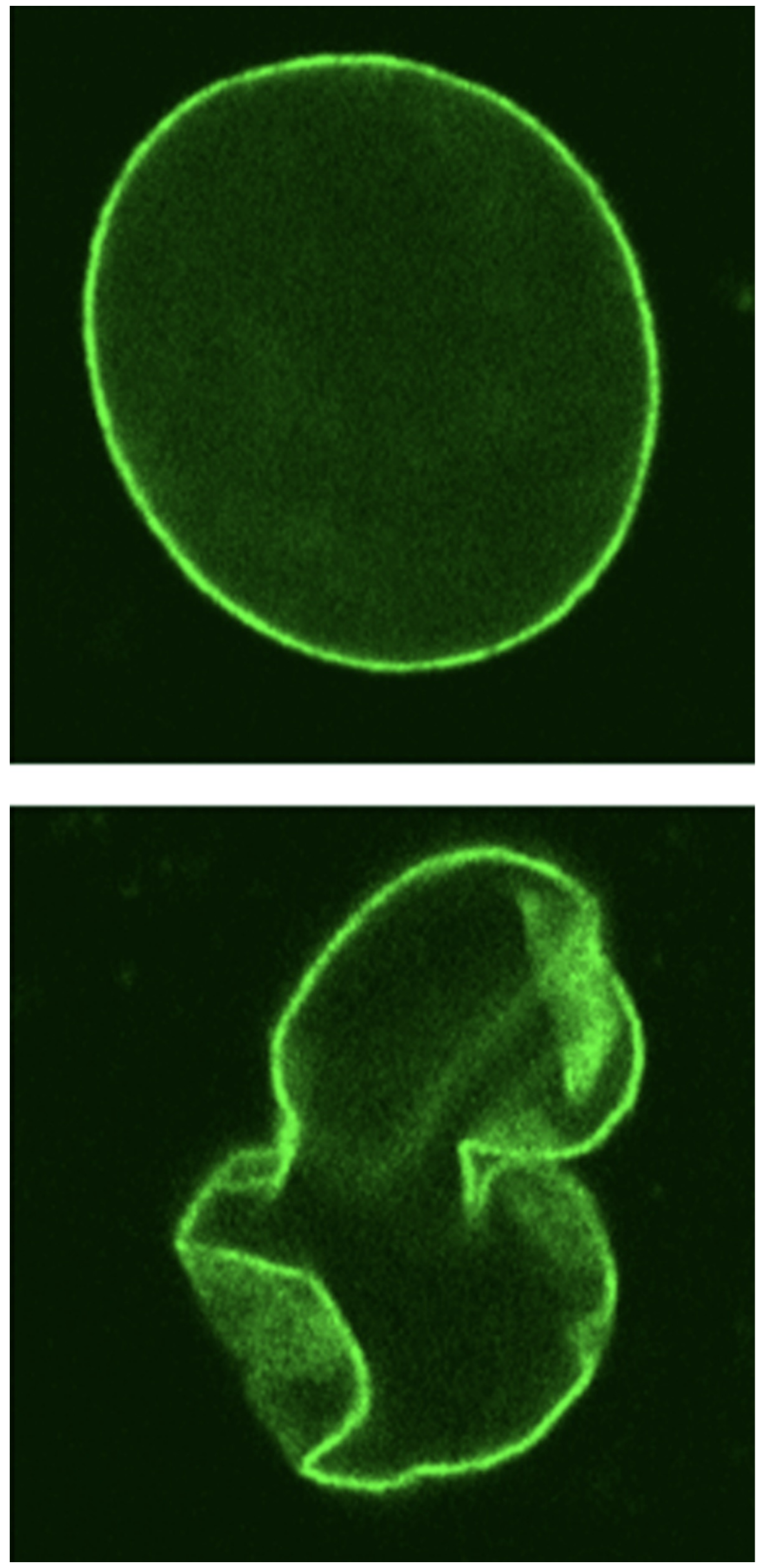
what does this even mean?

What circumstances made the model fit better?

worse?

Accuracy

"Why is it not enough?"



Progeria affects ~159 patients in the US

we have a dataset of all American pediatric patients

Q: If my model predicts with 99.99% accuracy, is it good enough?

protein shape → Model → **progeria** (yes or no)

Progeria affects ~159 patients in the US

we have a dataset of all American pediatric patients

a proposed model:

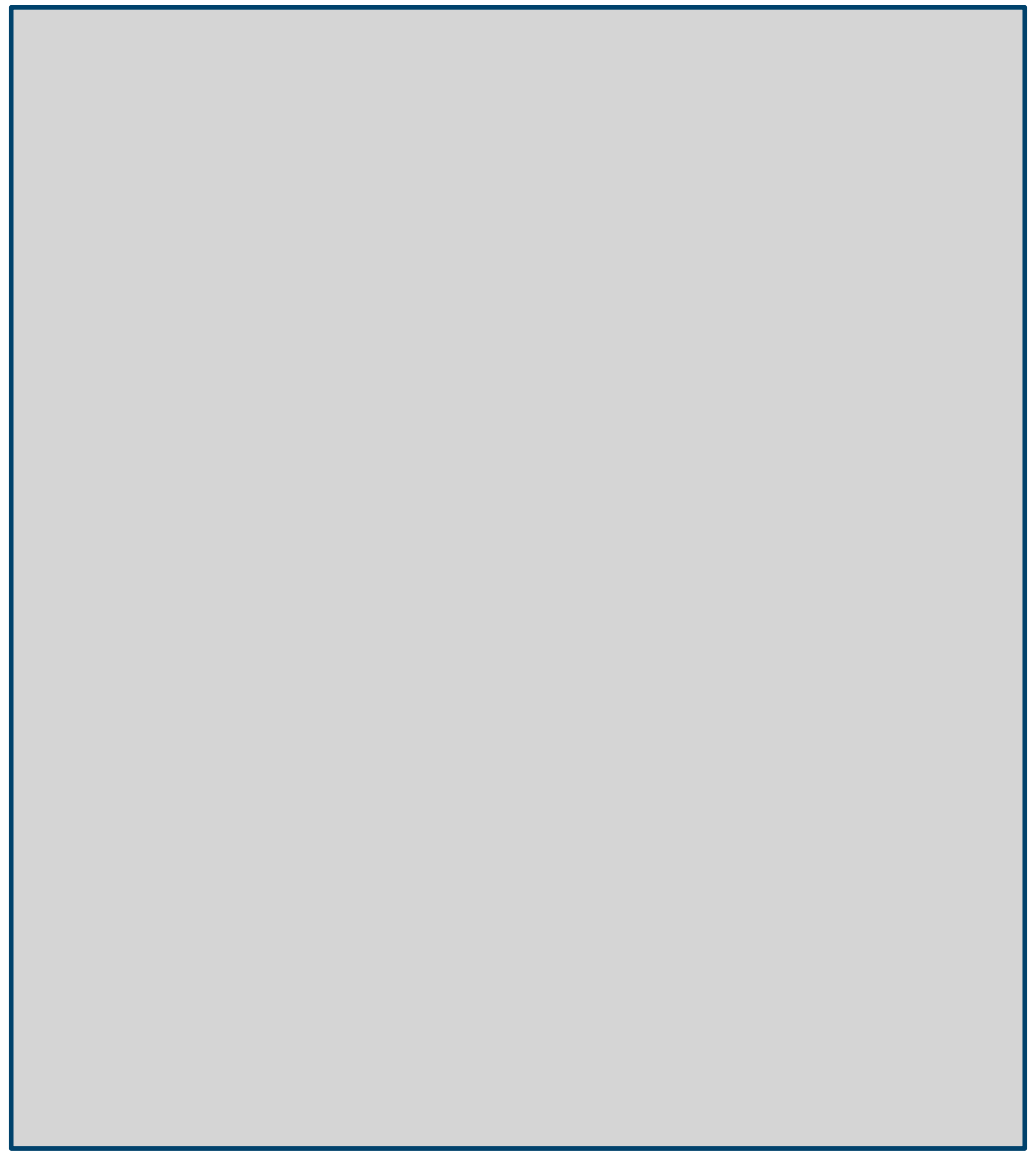
protein shape → Model → no progeria regardless

actually pretty accurate!

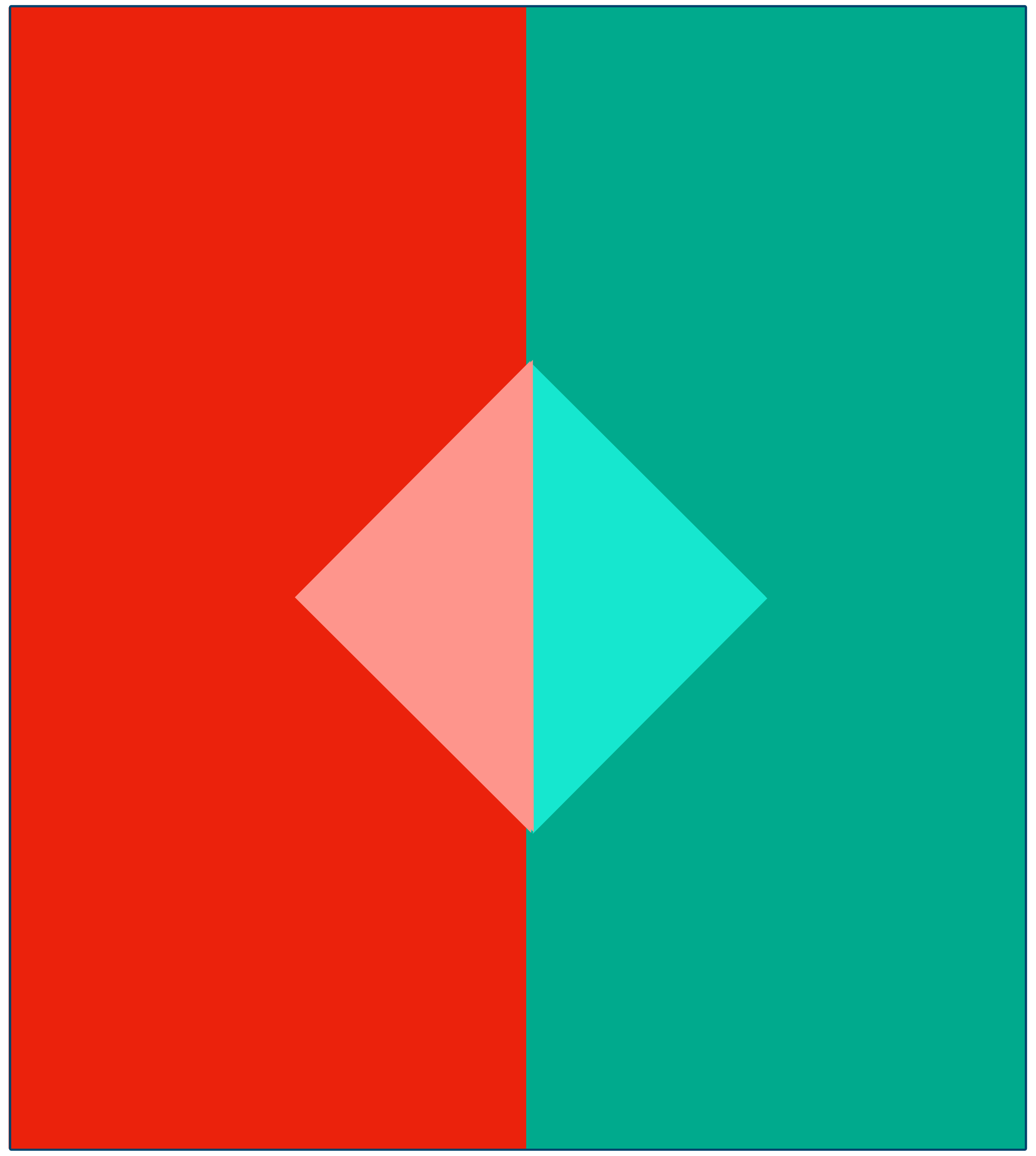
Progeria affects ~159 patients in the US

we have a dataset of all American pediatric patients

Accuracy , Precision, and Recall



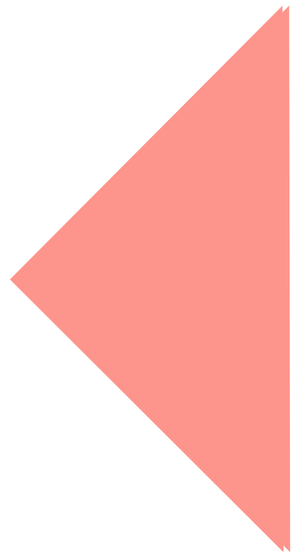
“Selection space”



“Selection space”

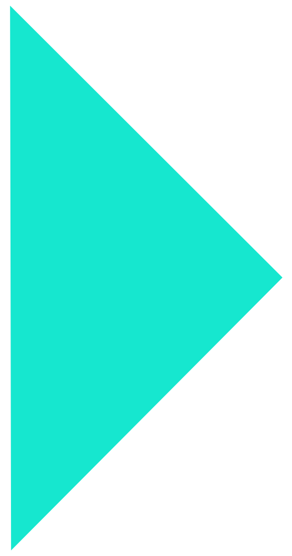
TRUE **P**OSITIVE

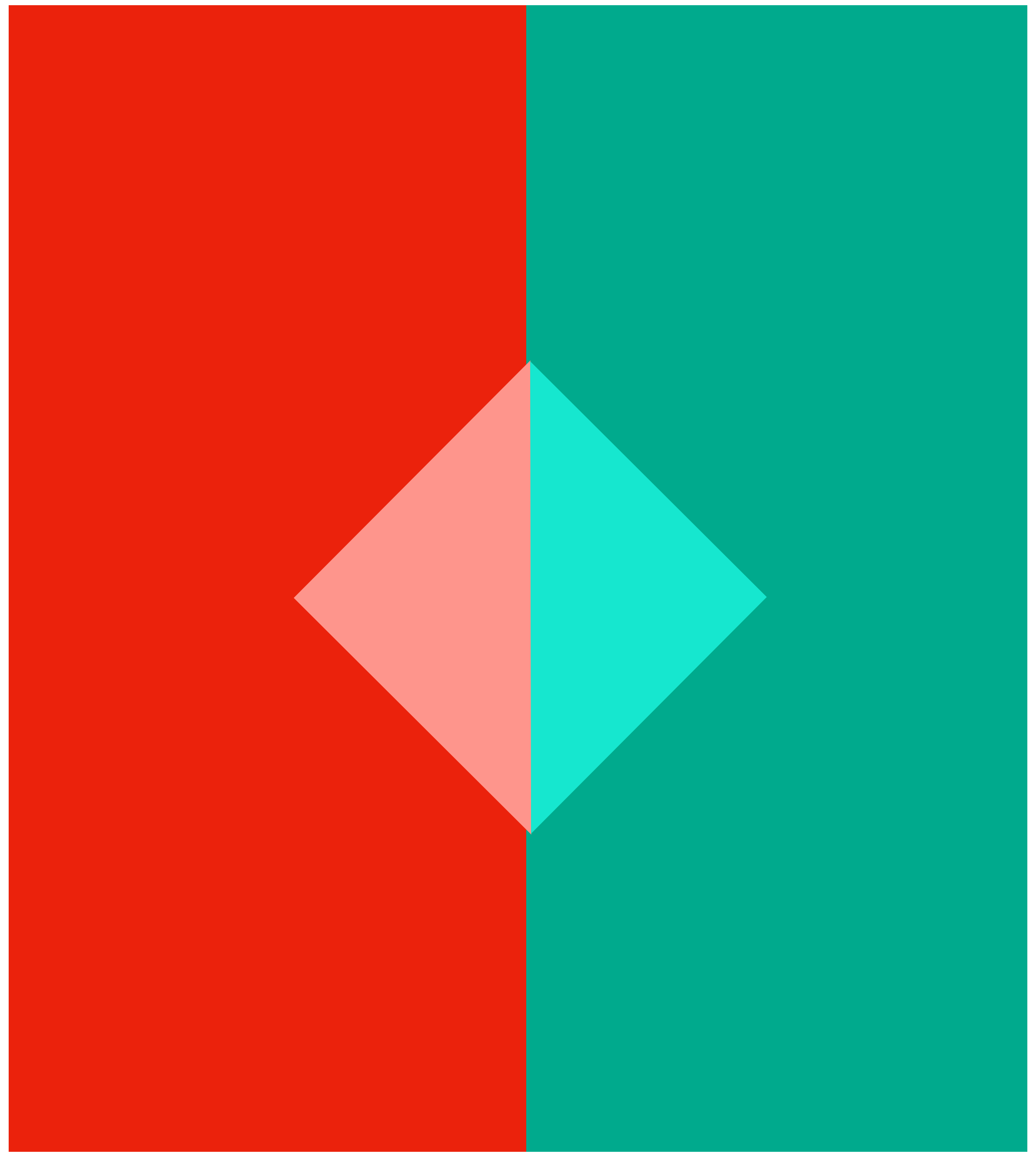
TP: Model selects **positive** and patient is **positive**



FALSE **P**OSITIVE

FP: Model selects **positive** and patient is **negative**





“Selection space”

FALSE **N**EGATIVE

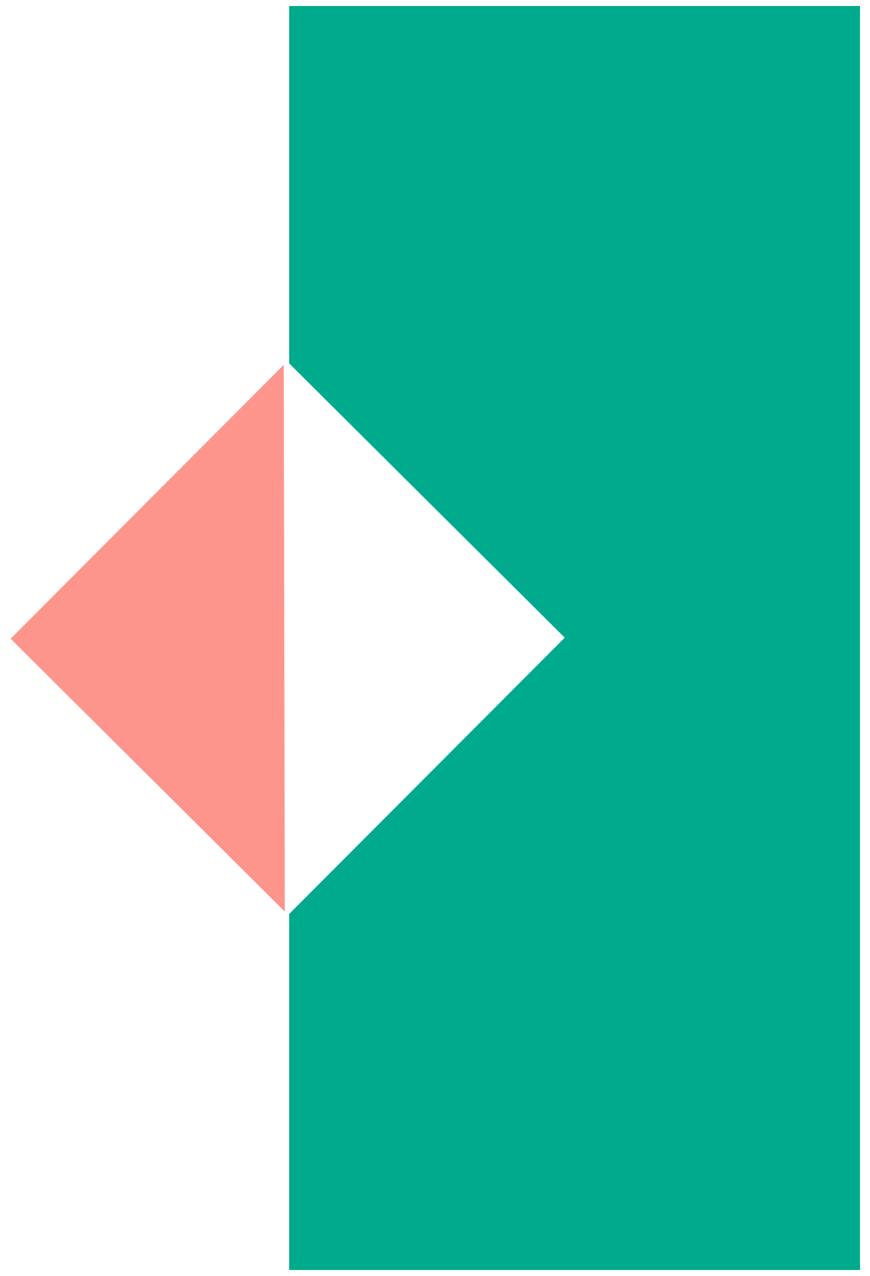
FN: Model selects **negative** and patient is **positive**

TRUE **N**EGATIVE

TN: Model selects **negative** and patient is **negative**

Accuracy

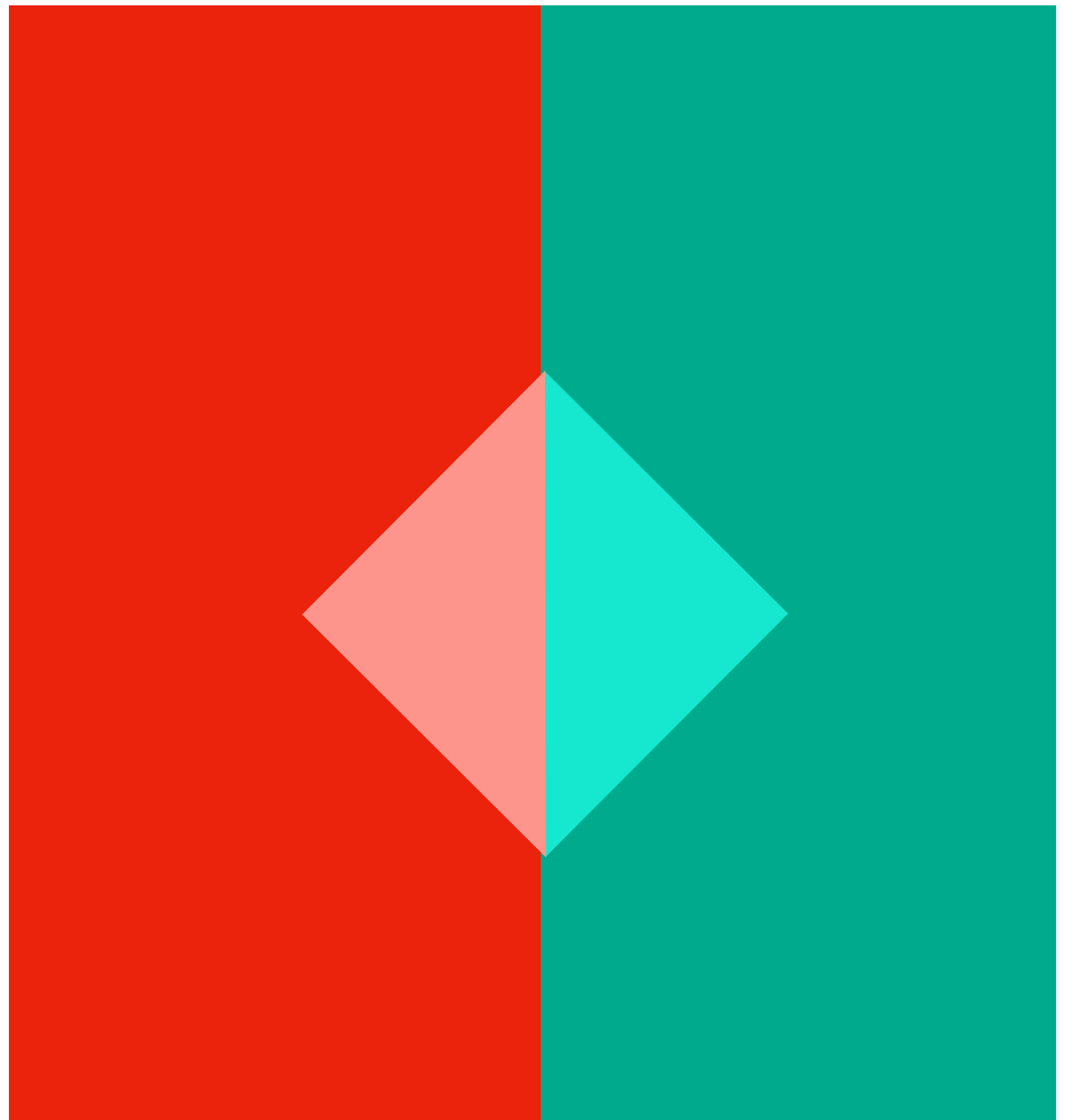
Overall ability of model



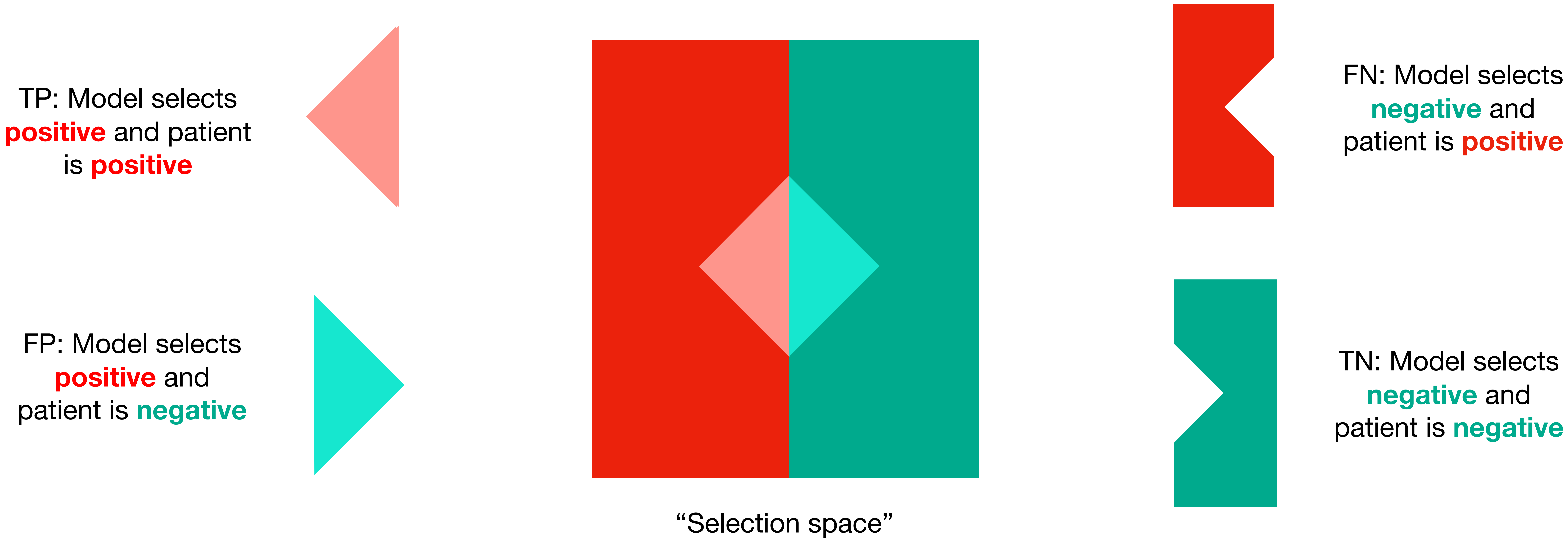
“Number of cases where **we chose** **positive** **when patient is** **positive**

and

Number of cases where **we chose** **negative** **when patient is** **negative**”



“Everything”

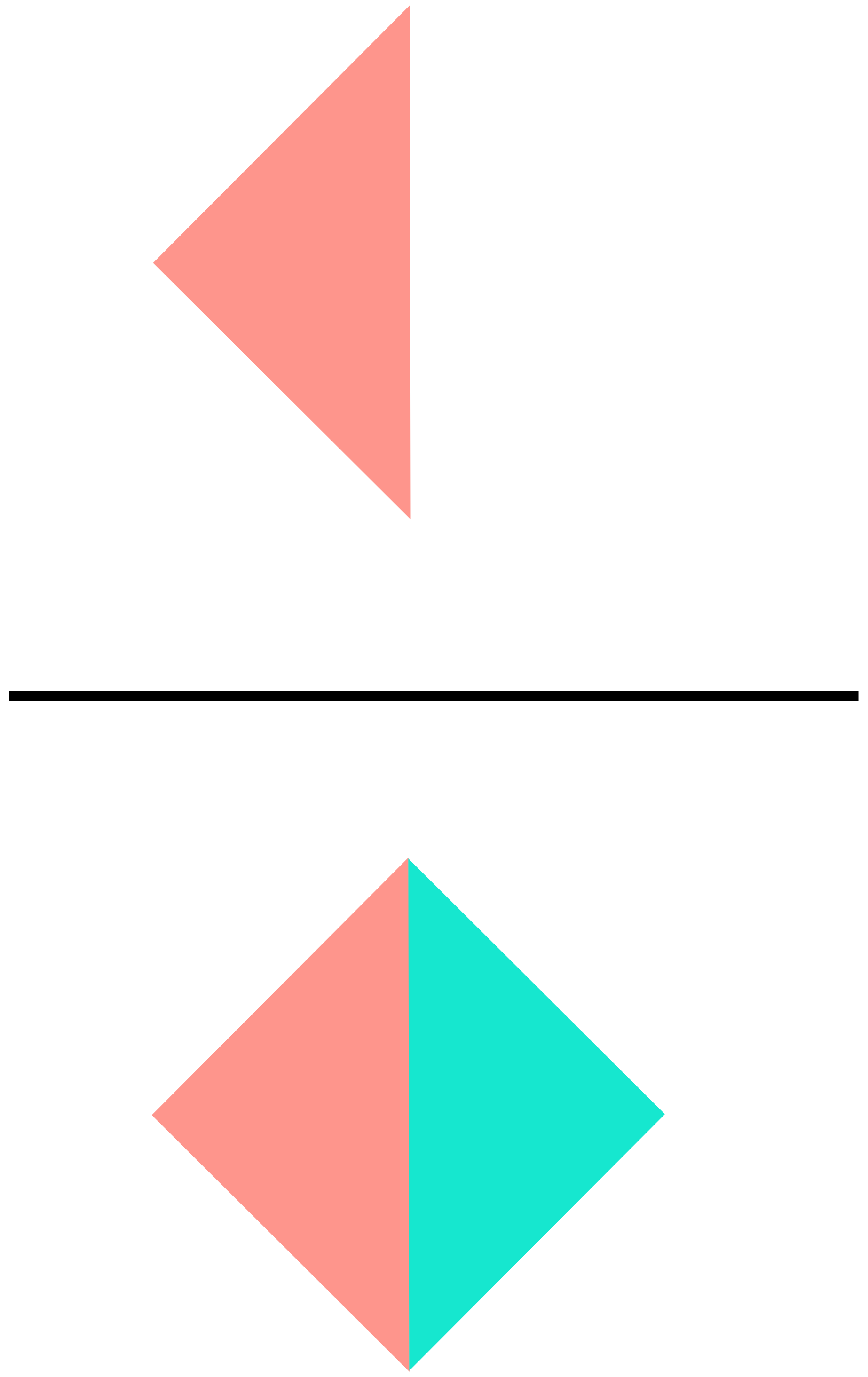


Accuracy

Overall ability of model

Precision

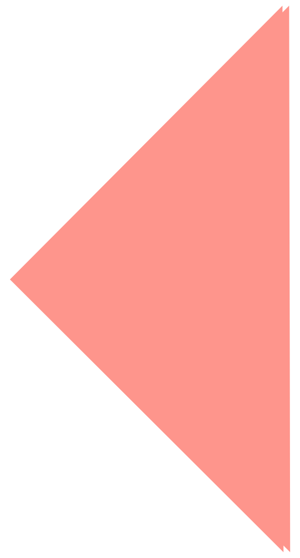
Accuracy of what we selected.
Or amount of selection that's actually correct.



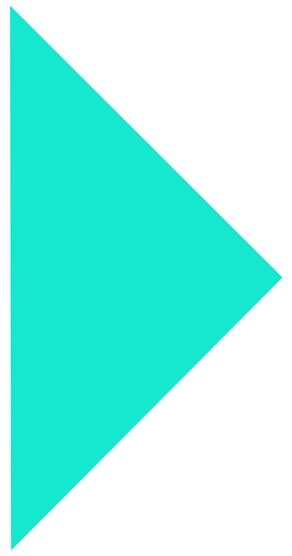
"Number of cases where **we chose positive when patient is positive**"

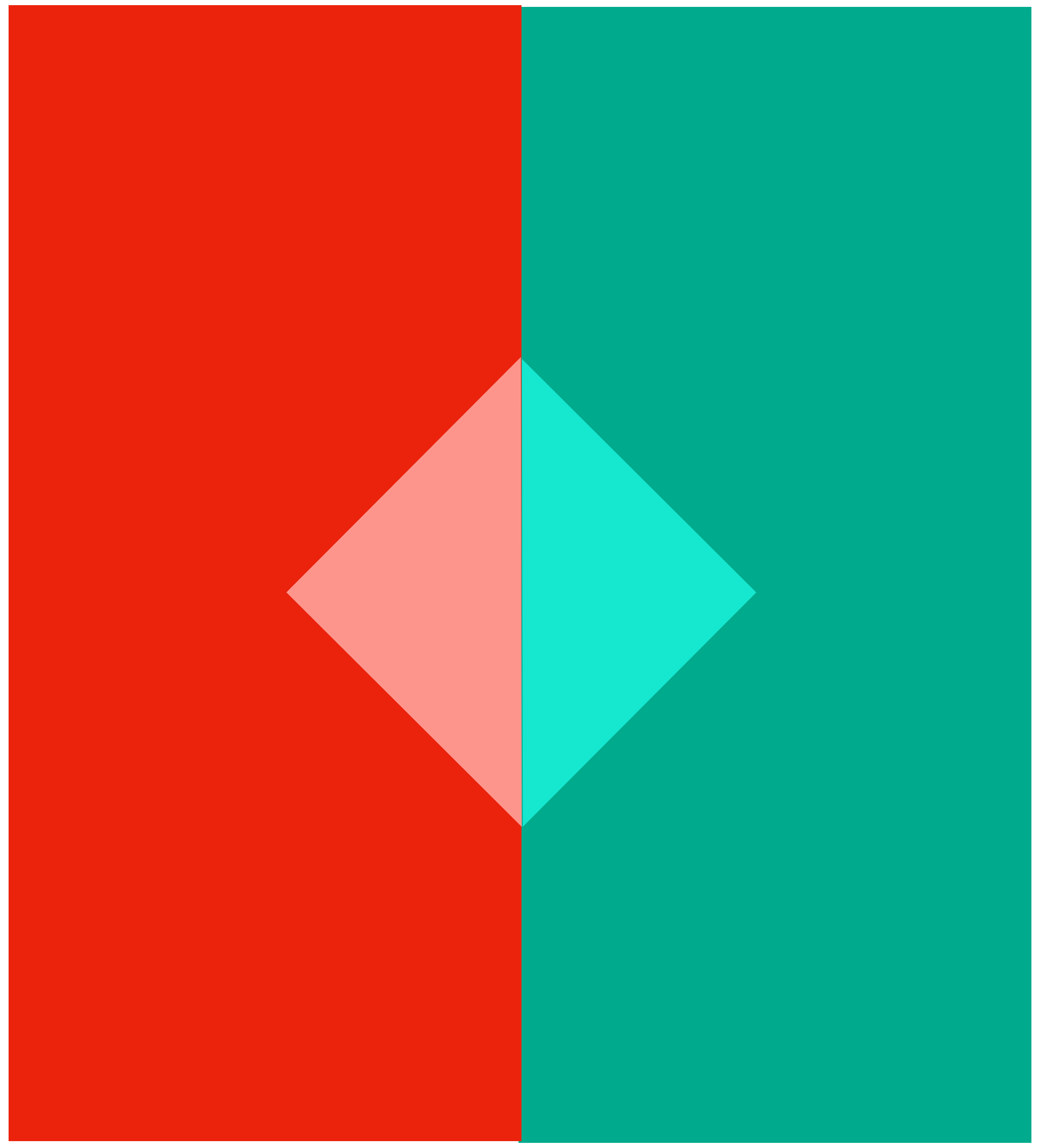
"All selected **positive** by the model"

TP: Model selects **positive** and patient is **positive**



FP: Model selects **positive** and patient is **negative**





“Selection space”

FN: Model selects **negative** and patient is **positive**

TN: Model selects **negative** and patient is **negative**

Accuracy

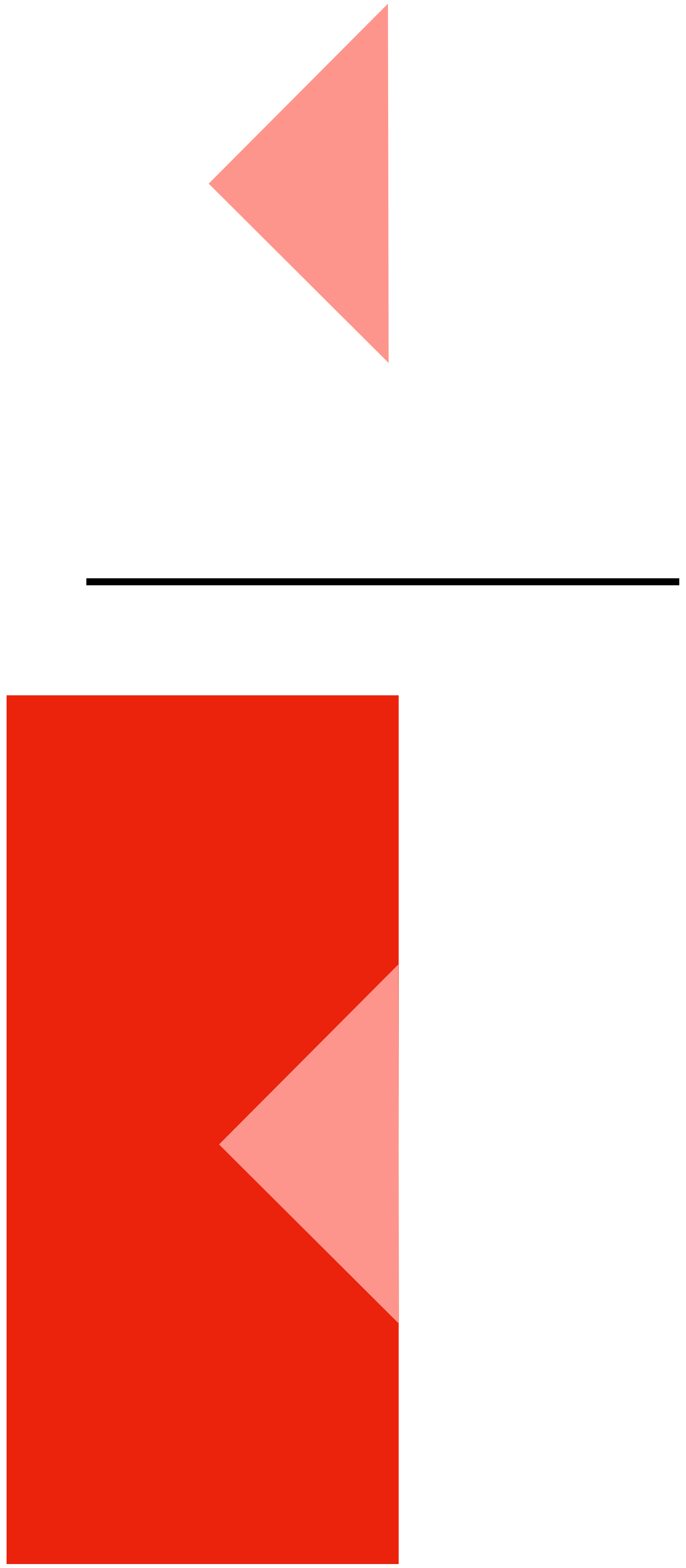
Overall ability of model

Precision

Amount of selection that’s actually correct.

Recall

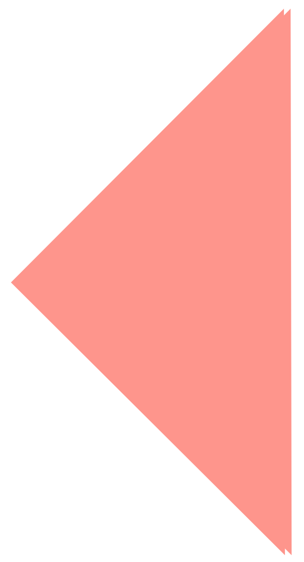
Accuracy of what we should select.
Or amount of what needs to be selected that is selected



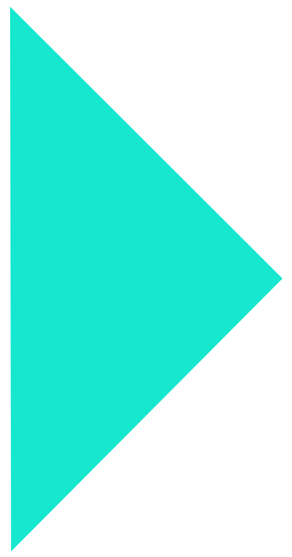
“Number of cases where **we chose positive when patient is positive**”

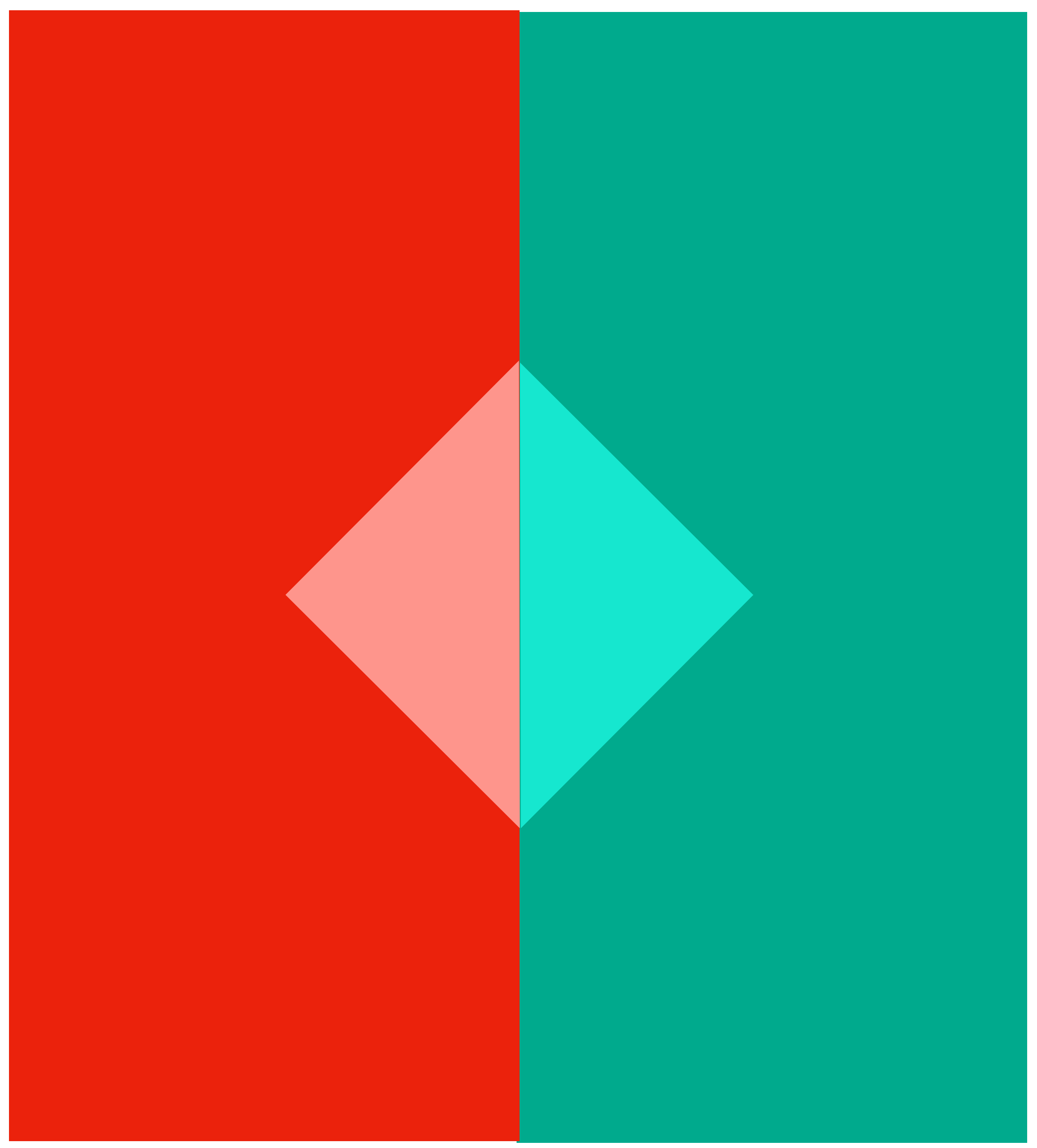
“All cases that the patients are **positive**”

TP: Model selects **positive** and patient is **positive**



FP: Model selects **positive** and patient is **negative**





"Selection space"

FN: Model selects **negative** and patient is **positive**

TN: Model selects **negative** and patient is **negative**

Accuracy

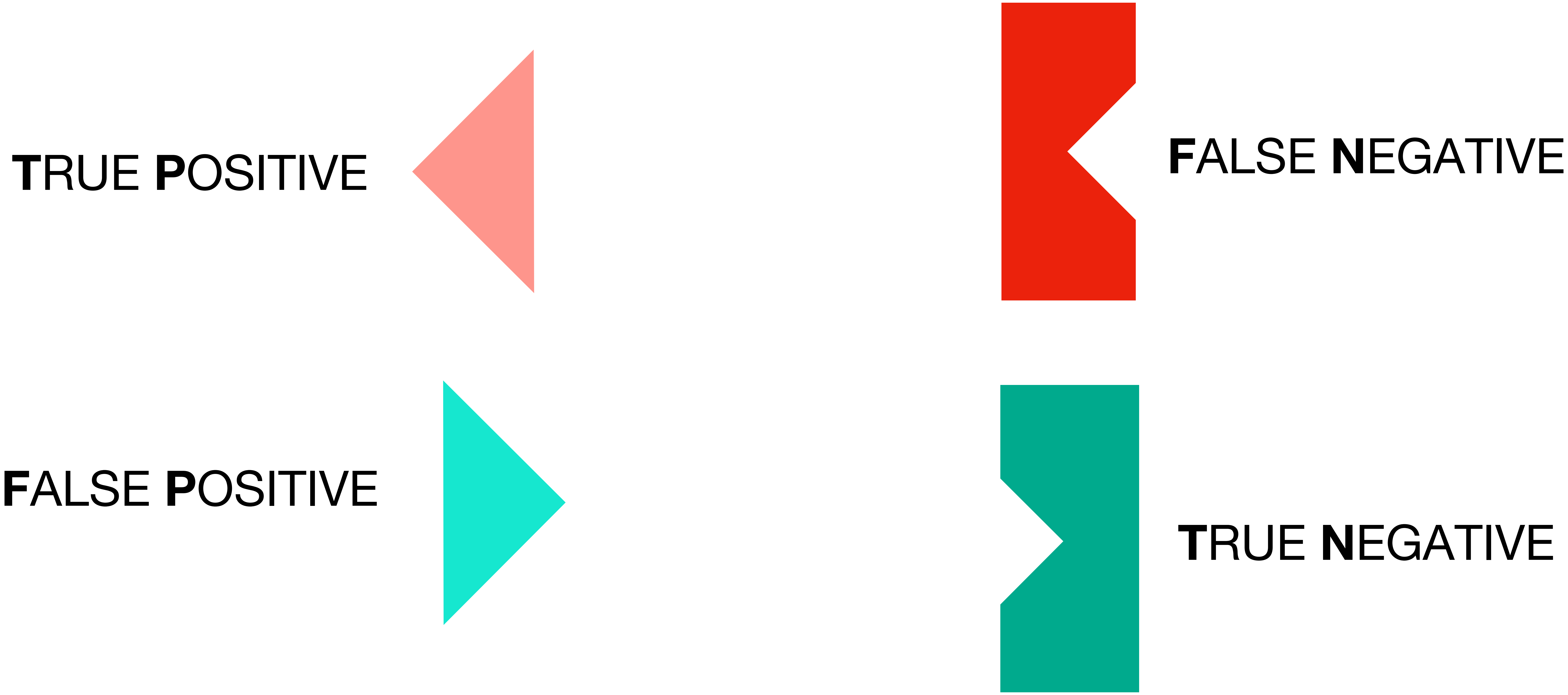
Overall ability of model

Precision

Amount of selection that's actually correct.

Recall

Amount of what needs to be selected that is selected



Accuracy

Overall ability of model

Precision

Amount of selection that’s actually correct.

Recall

Amount of what needs to be selected that is selected

Sources: [6][7][8][9][10][11][12][13][14] view · talk · edit

		Predicted condition			
	Total population $= P + N$	**Positive (PP)**	**Negative (PN)**	Informedness, bookmaker informedness (BM) $= TPR + TNR - 1$	Prevalence threshold (PT) $= \frac{\sqrt{TPR \times FPR} - FPR}{TPR - FPR}$
Actual condition	**Positive (P)**	**True positive (TP), hit**	**False negative (FN), type II error, miss, underestimation**	True positive rate (TPR), recall, sensitivity (SEN), probability of detection, hit rate, power $= \frac{TP}{P} = 1 - FNR$	False negative rate (FNR), miss rate $= \frac{FN}{P} = 1 - TPR$
	Negative (N)	**False positive (FP), type I error, false alarm, overestimation**	**True negative (TN), correct rejection**	False positive rate (FPR), probability of false alarm, fall-out $= \frac{FP}{N} = 1 - TNR$	True negative rate (TNR), specificity (SPC), selectivity $= \frac{TN}{N} = 1 - FPR$
	Prevalence $= \frac{P}{P + N}$	Positive predictive value (PPV), precision $= \frac{TP}{PP} = 1 - FDR$	False omission rate (FOR) $= \frac{FN}{PN} = 1 - NPV$	Positive likelihood ratio (LR+) $= \frac{TPR}{FPR}$	Negative likelihood ratio (LR–) $= \frac{FNR}{TNR}$
	Accuracy (ACC) $= \frac{TP + TN}{P + N}$	False discovery rate (FDR) $= \frac{FP}{PP} = 1 - PPV$	Negative predictive value (NPV) $= \frac{TN}{PN} = 1 - FOR$	Markedness (MK), deltaP (Δp) $= PPV + NPV - 1$	Diagnostic odds ratio (DOR) $= \frac{LR+}{LR-}$
	Balanced accuracy (BA) $= \frac{TPR + TNR}{2}$	F_1 score $= \frac{2 PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$	Fowlkes–Mallows index (FM) $= \sqrt{PPV \times TPR}$	Matthews correlation coefficient (MCC) $= \sqrt{TPR \times TNR \times PPV \times NPV} - \sqrt{FNR \times FPR \times FOR \times FDR}$	Threat score (TS), critical success index (CSI), Jaccard index $= \frac{TP}{TP + FN + FP}$

https://en.wikipedia.org/wiki/Precision_and_recall

Accuracy

Overall ability of model

$$\frac{TP + TN}{Total}$$

exactly zero

Precision

Amount of selection that's actually correct.

$$\frac{TP}{TP + FP}$$

Recall

Amount of what needs to be selected that is selected

$$\frac{TP}{TP + FN}$$

scaled properly!

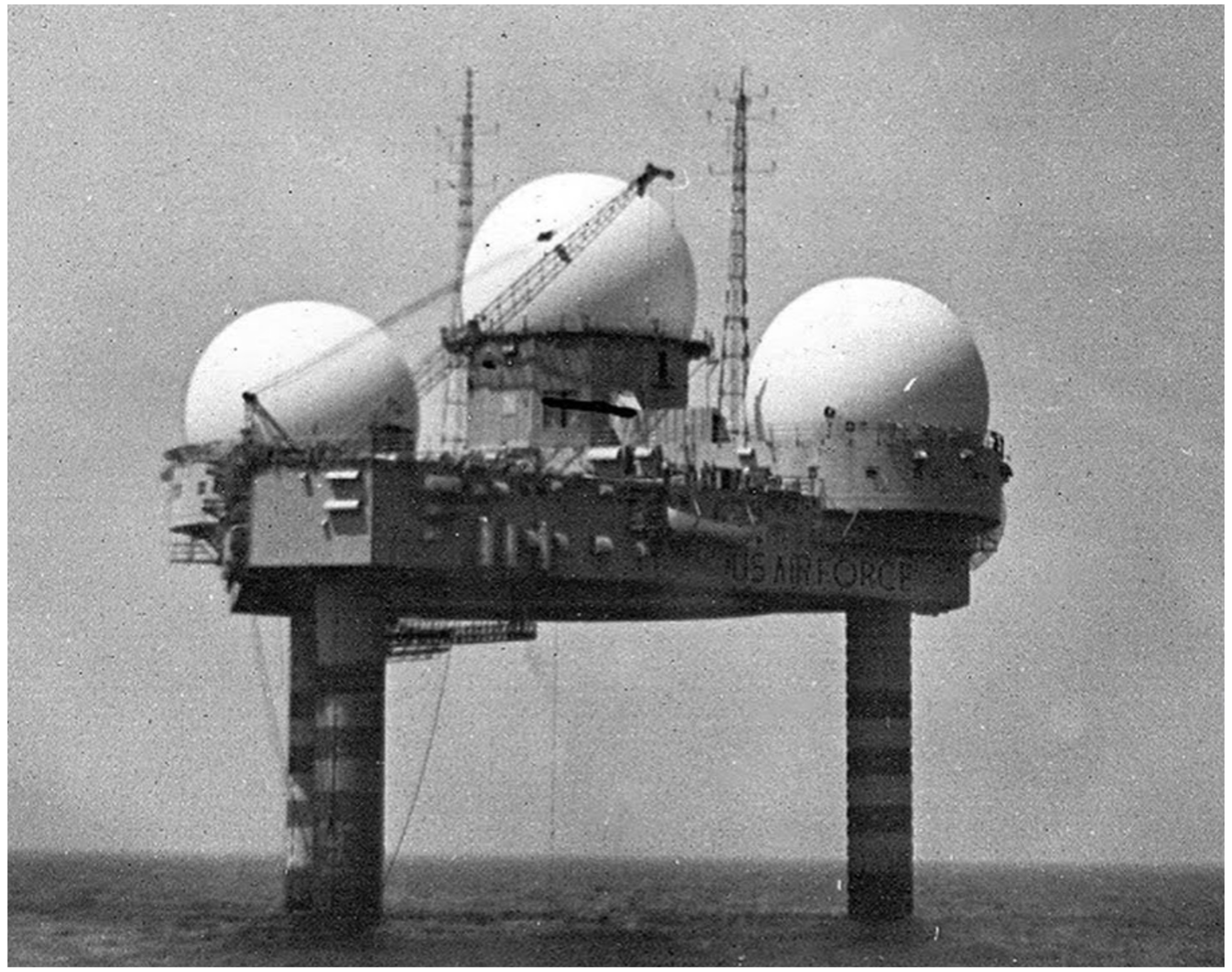
Model → **no progeria regardless**

Progeria affects ~159 patients in the US

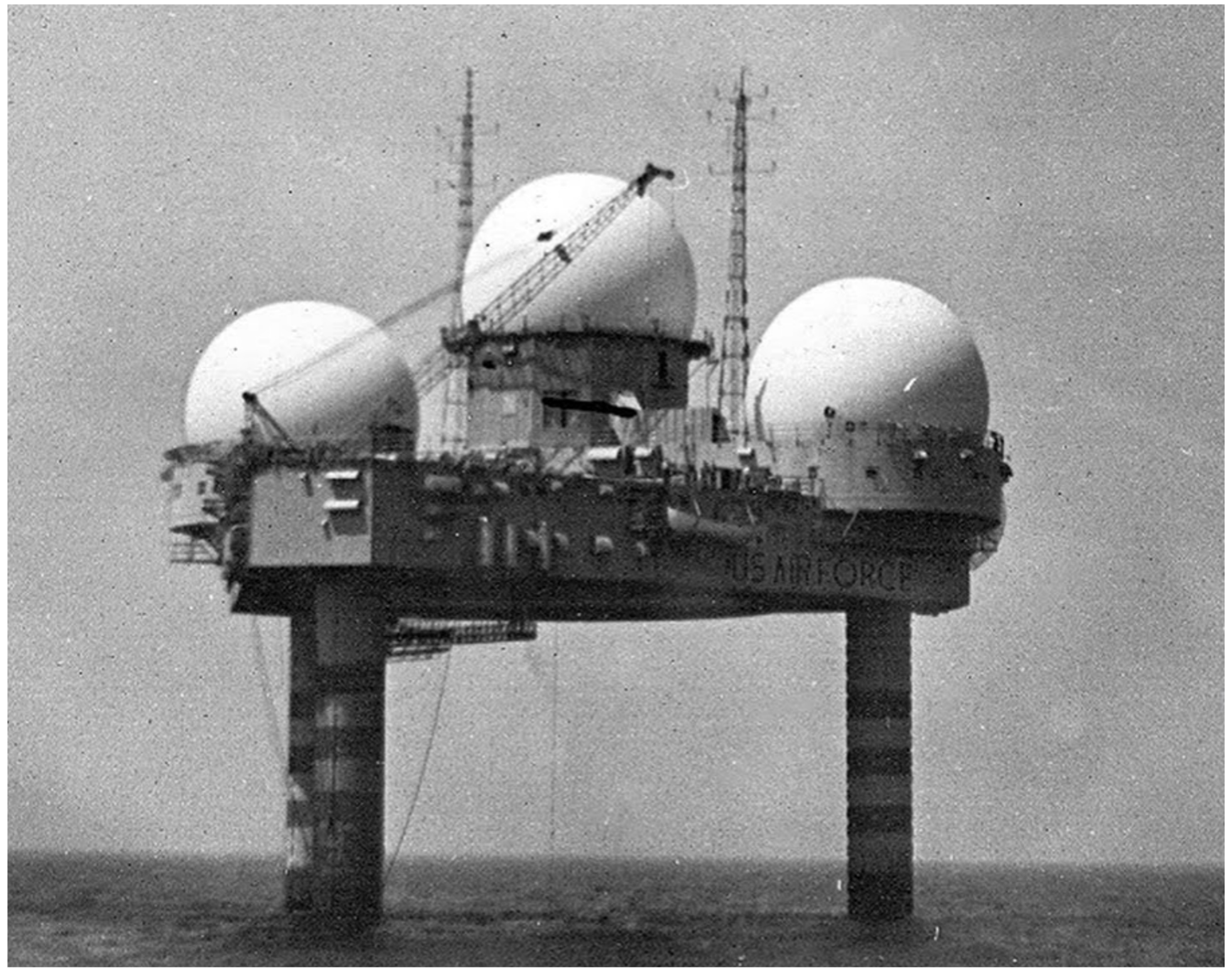
we have a dataset of all American pediatric patients

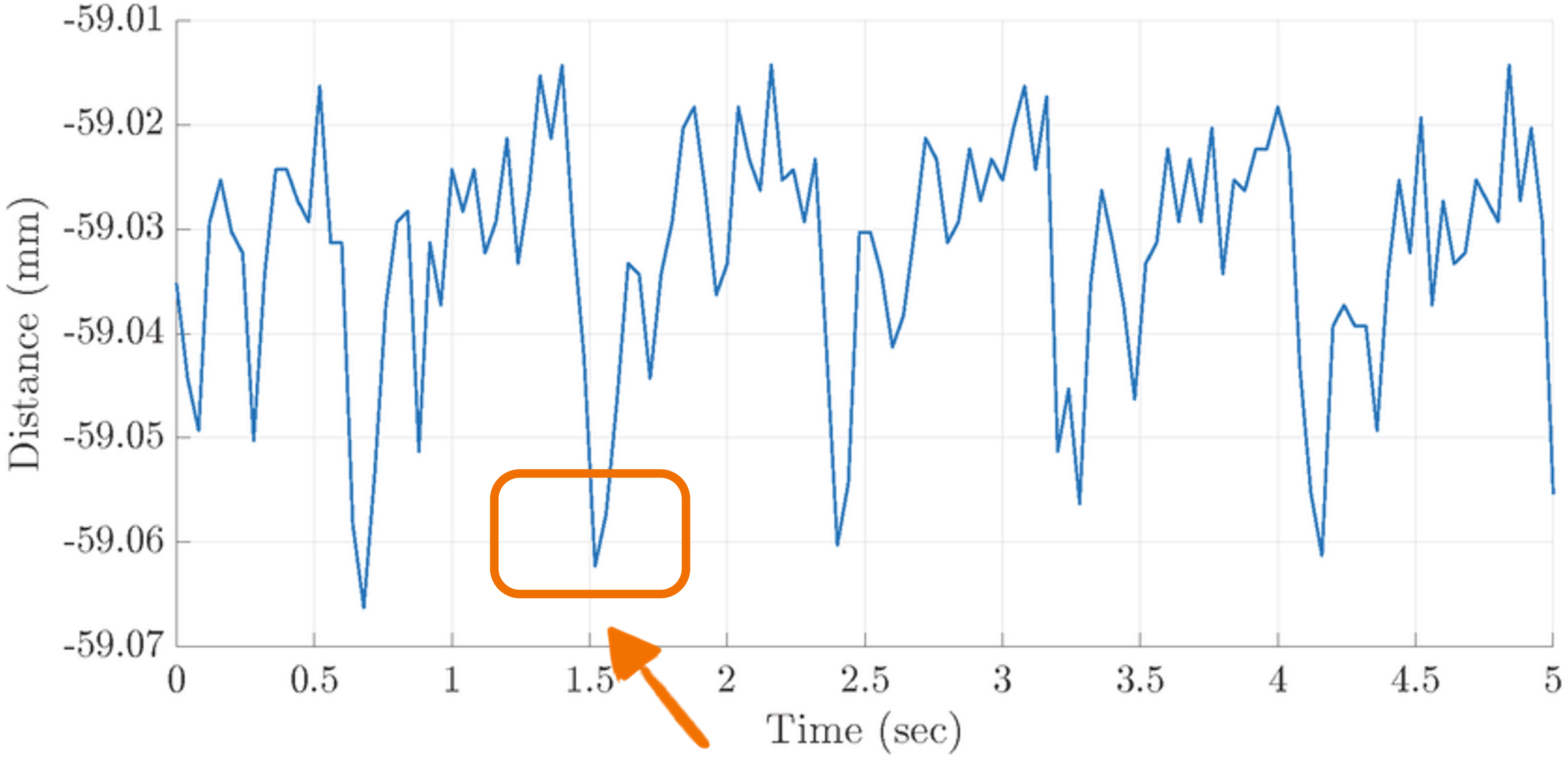
storytime!

storytime!



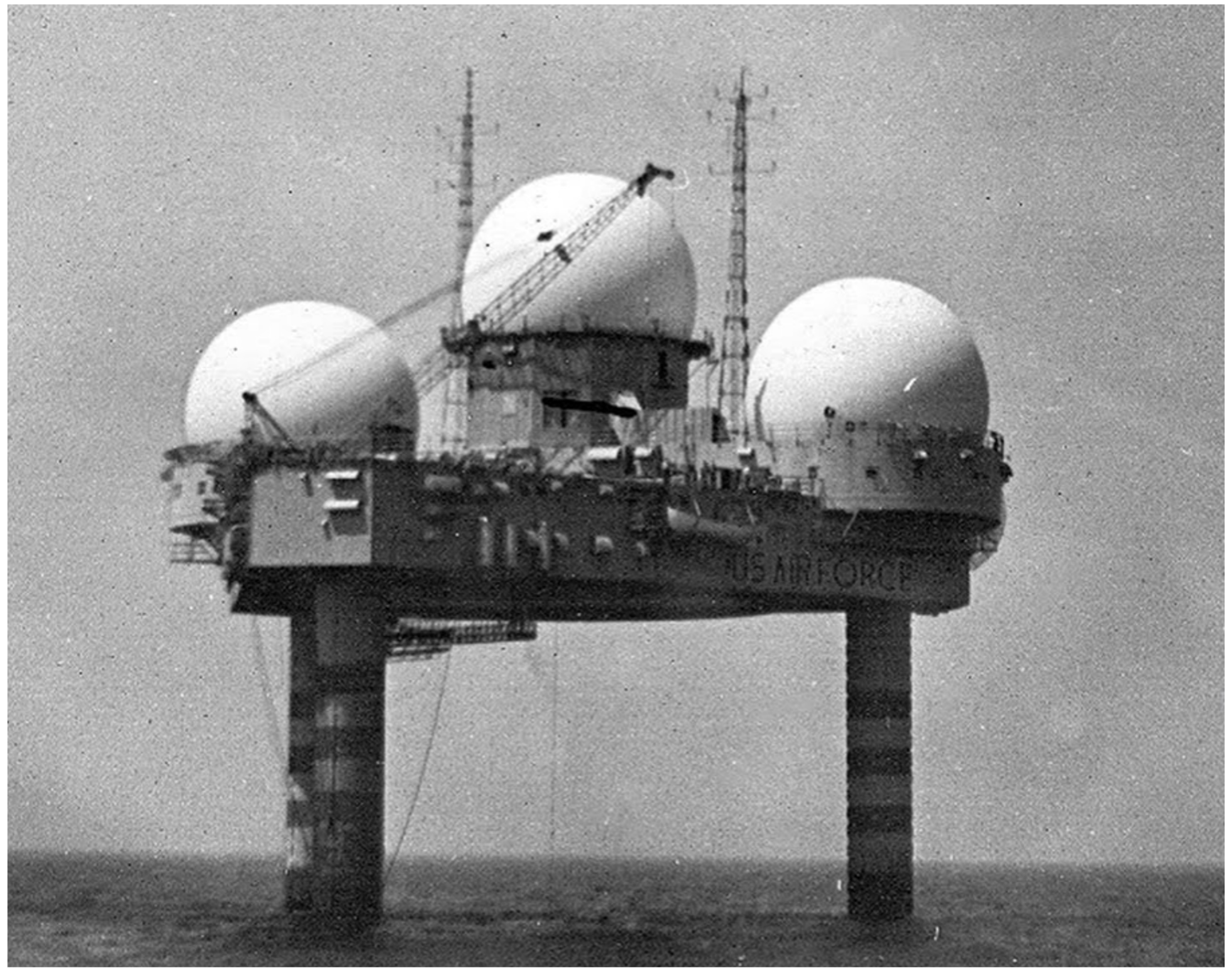
storytime!

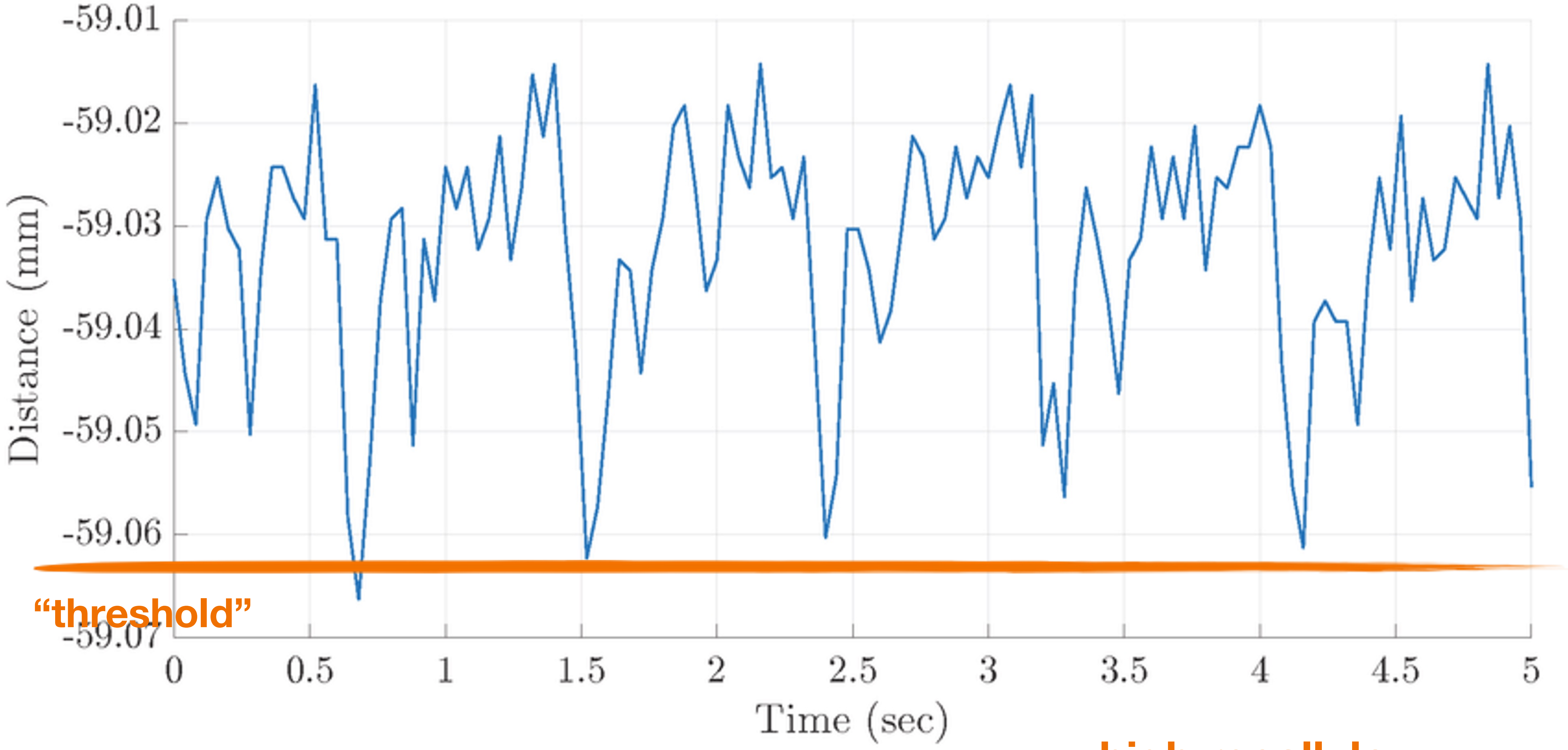




does this count as detection?

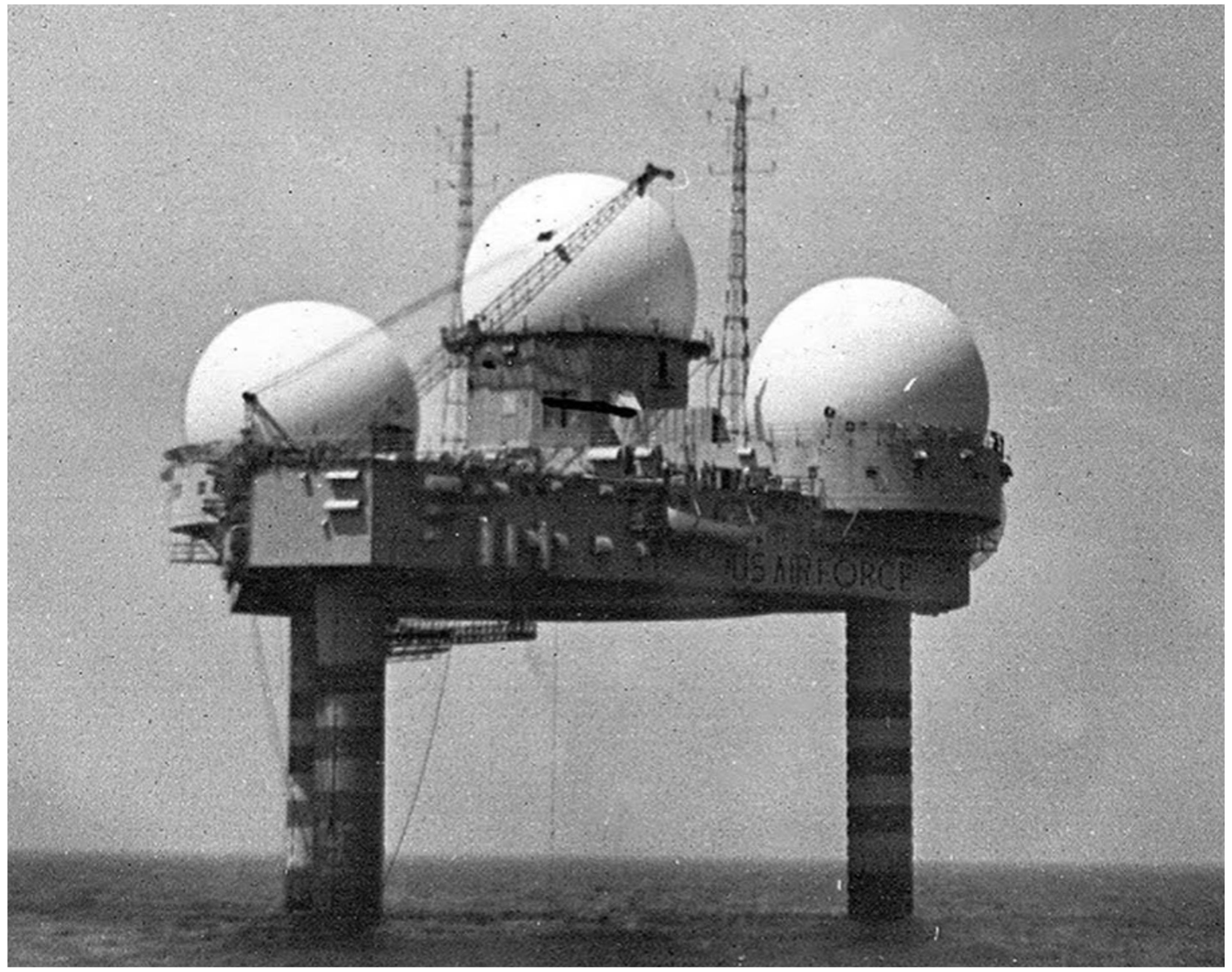
storytime!

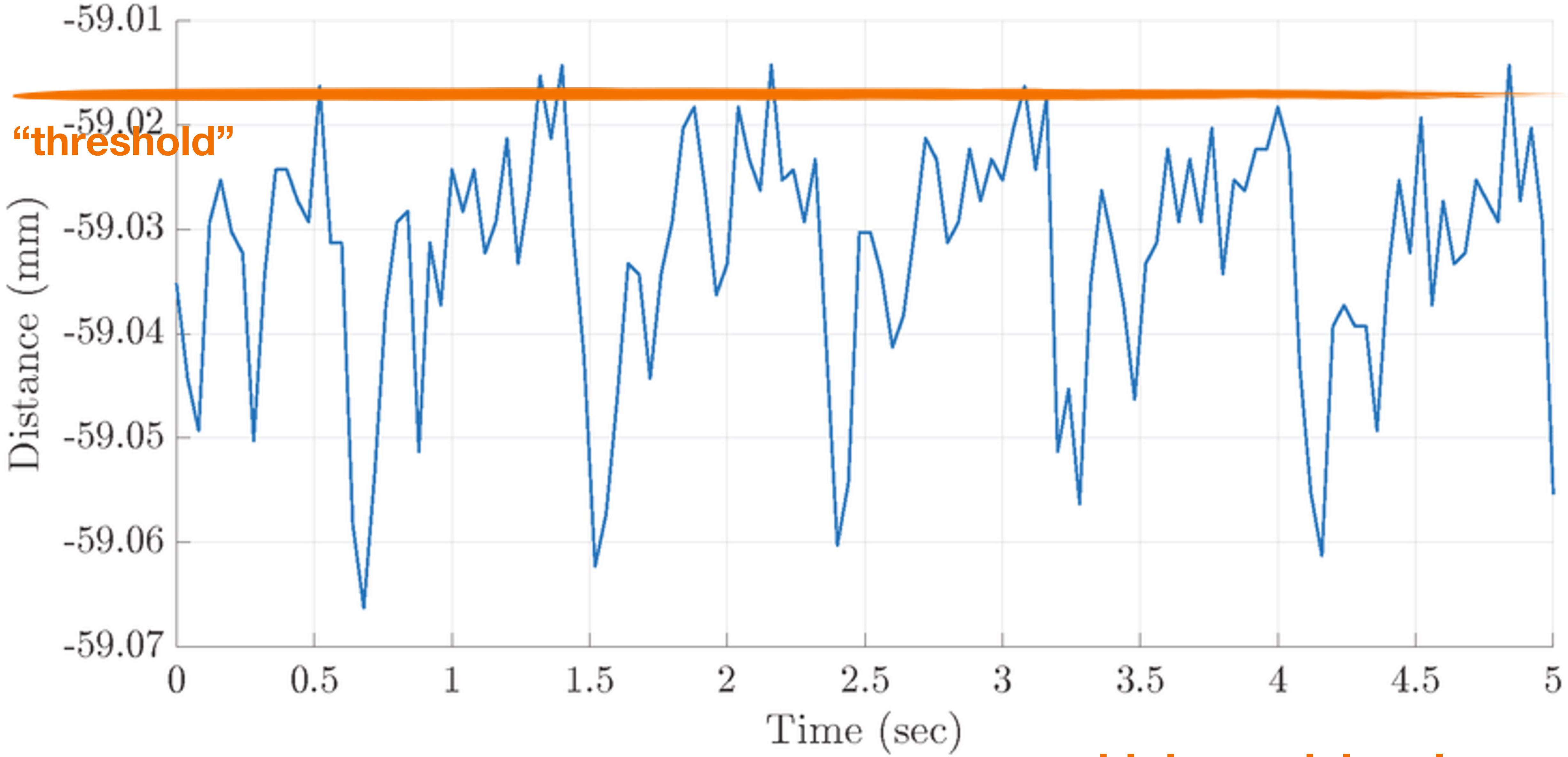




high recall, low precision

storytime!



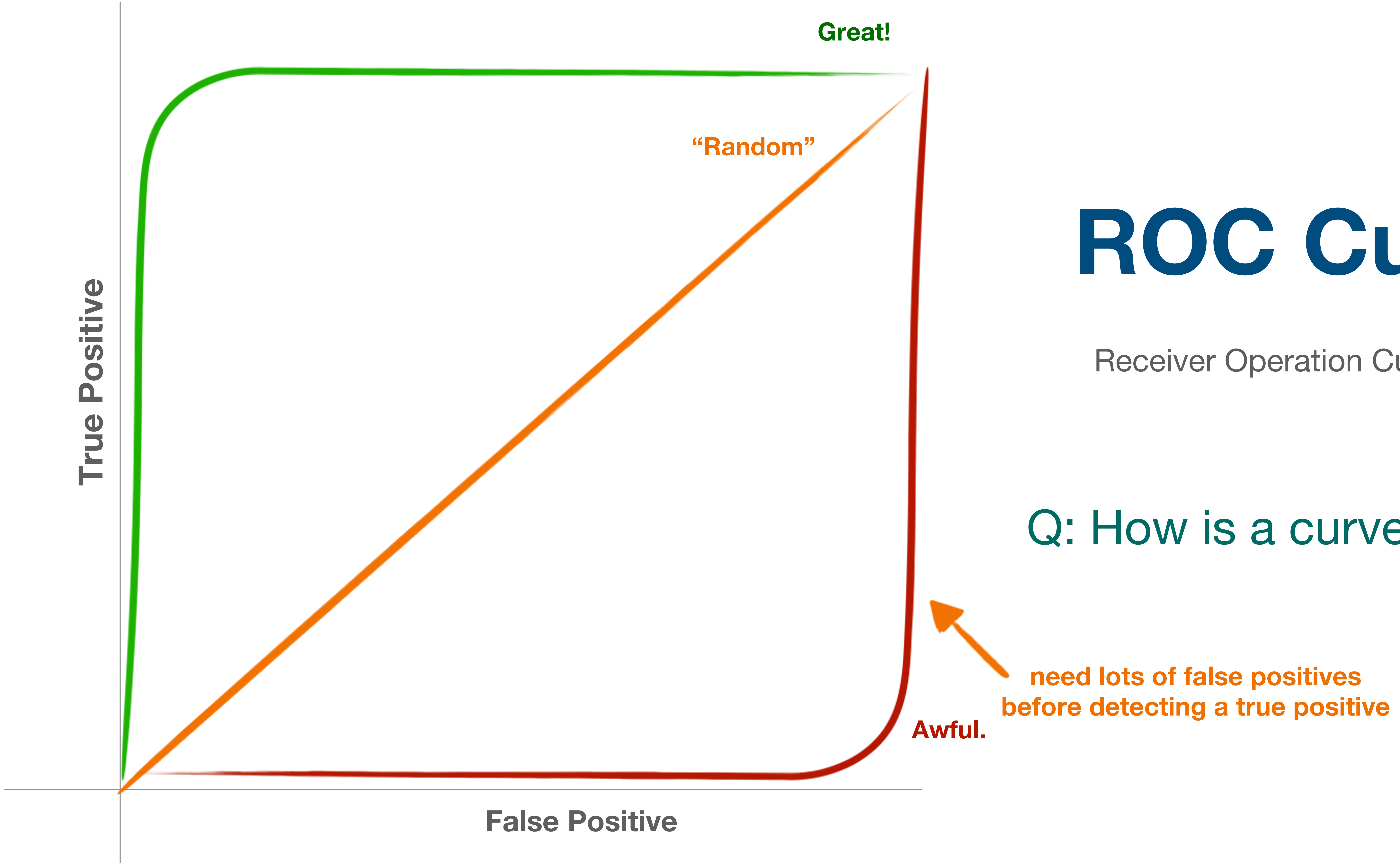


high precision, low recall

quantifying
“threshold”

quantifying
"threshold"

ROC Curve!

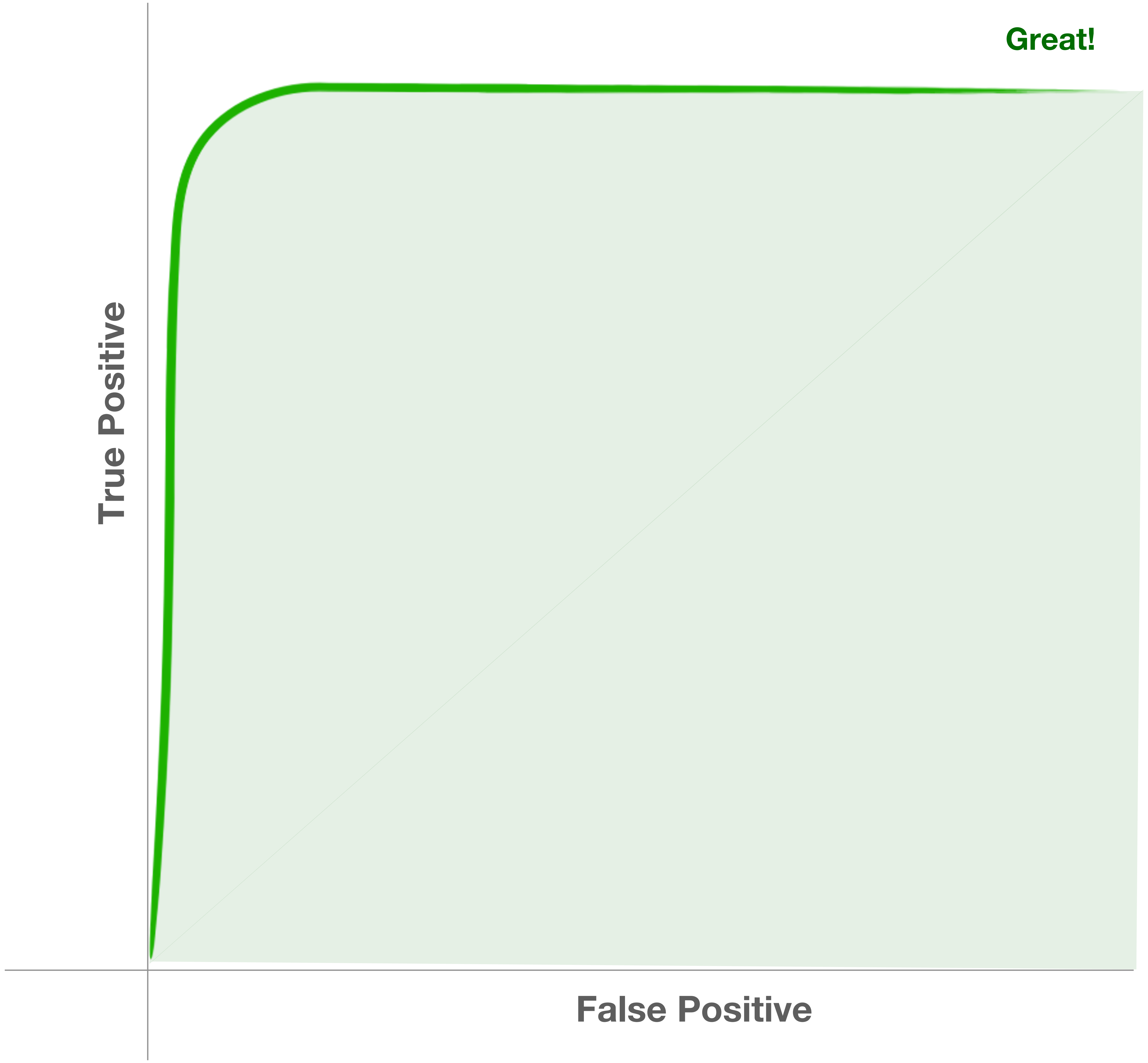


ROC Curve!

Receiver Operation Curve

Q: How is a curve generated?

■ **ROC Curve** quantify the amount of "error"/noise that is necessary for a classifier to make a good prediction



AUC

area under [the ROC] curve

Q: how do you compare these points

Self-test

Precision–recall AUC

Q: When do we really need it?
Q: what would it look like?

Especially for unbalanced datasets

what makes models fit better

more data

balanced data

normalized data

quality data

more data **balanced** data **normalized** data **quality** data

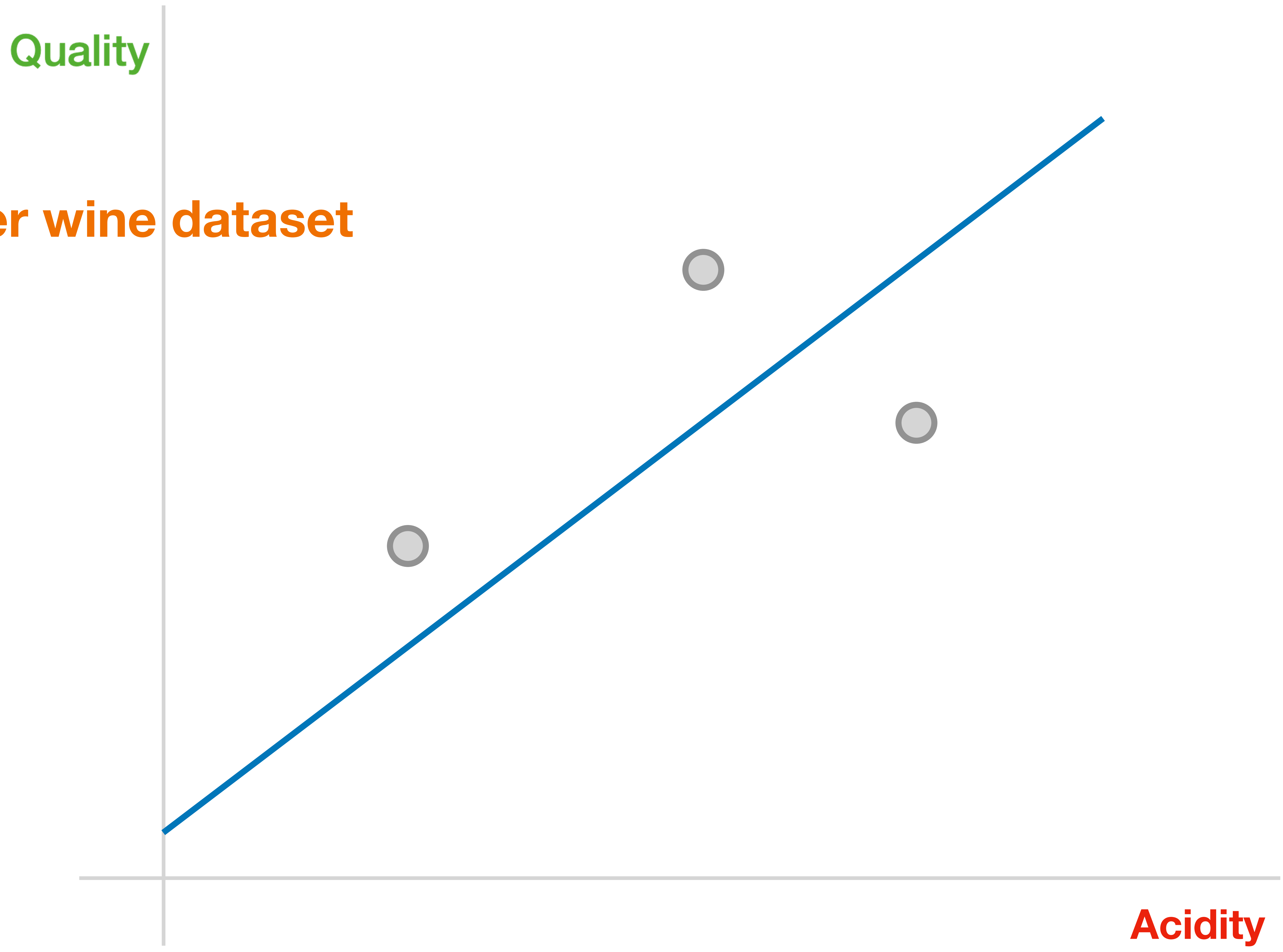
more data **balanced** data **normalized** data **quality** data

more data

let's say we have a simpler wine dataset

Quality on the y axis

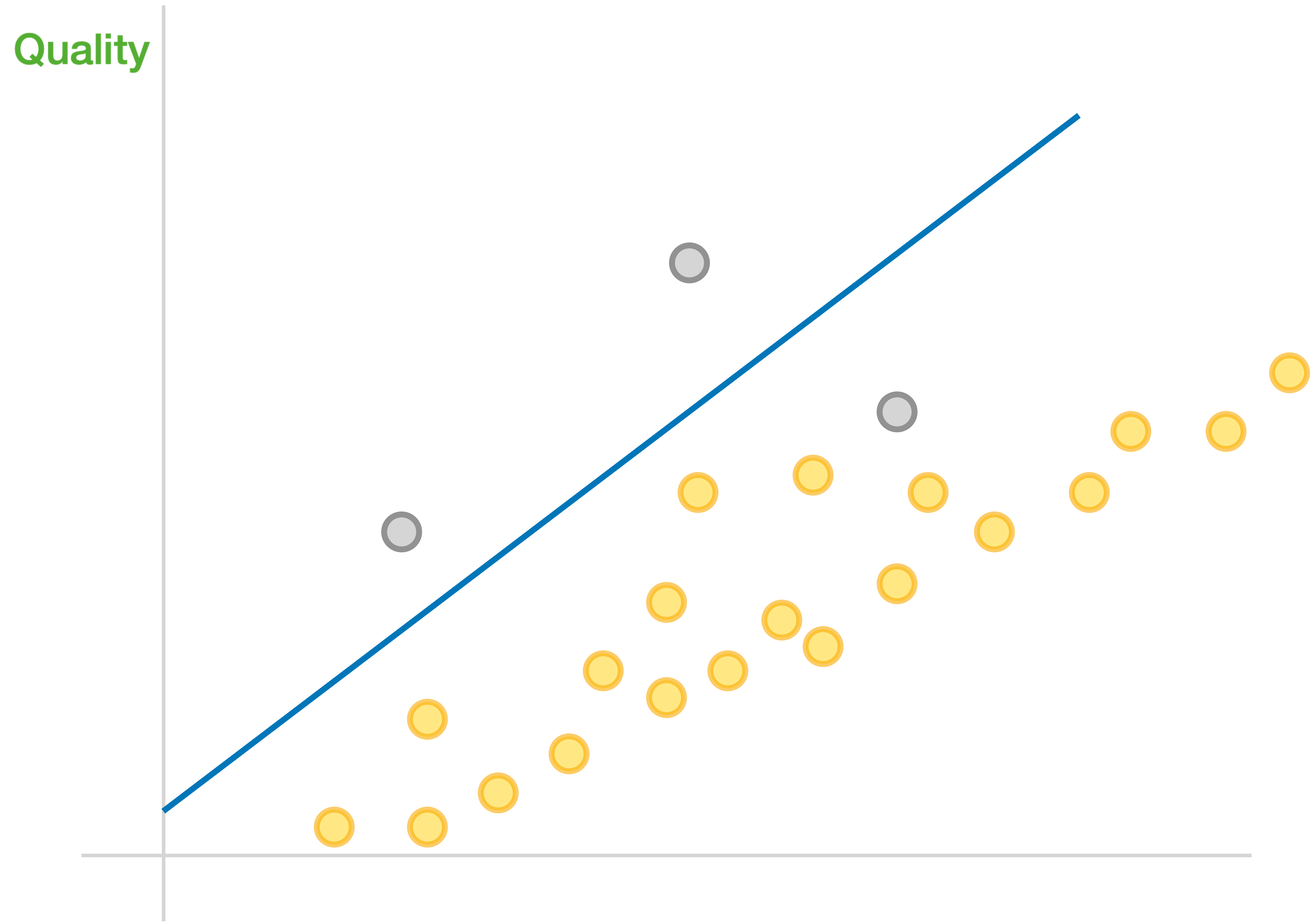
Acidity on the x axis



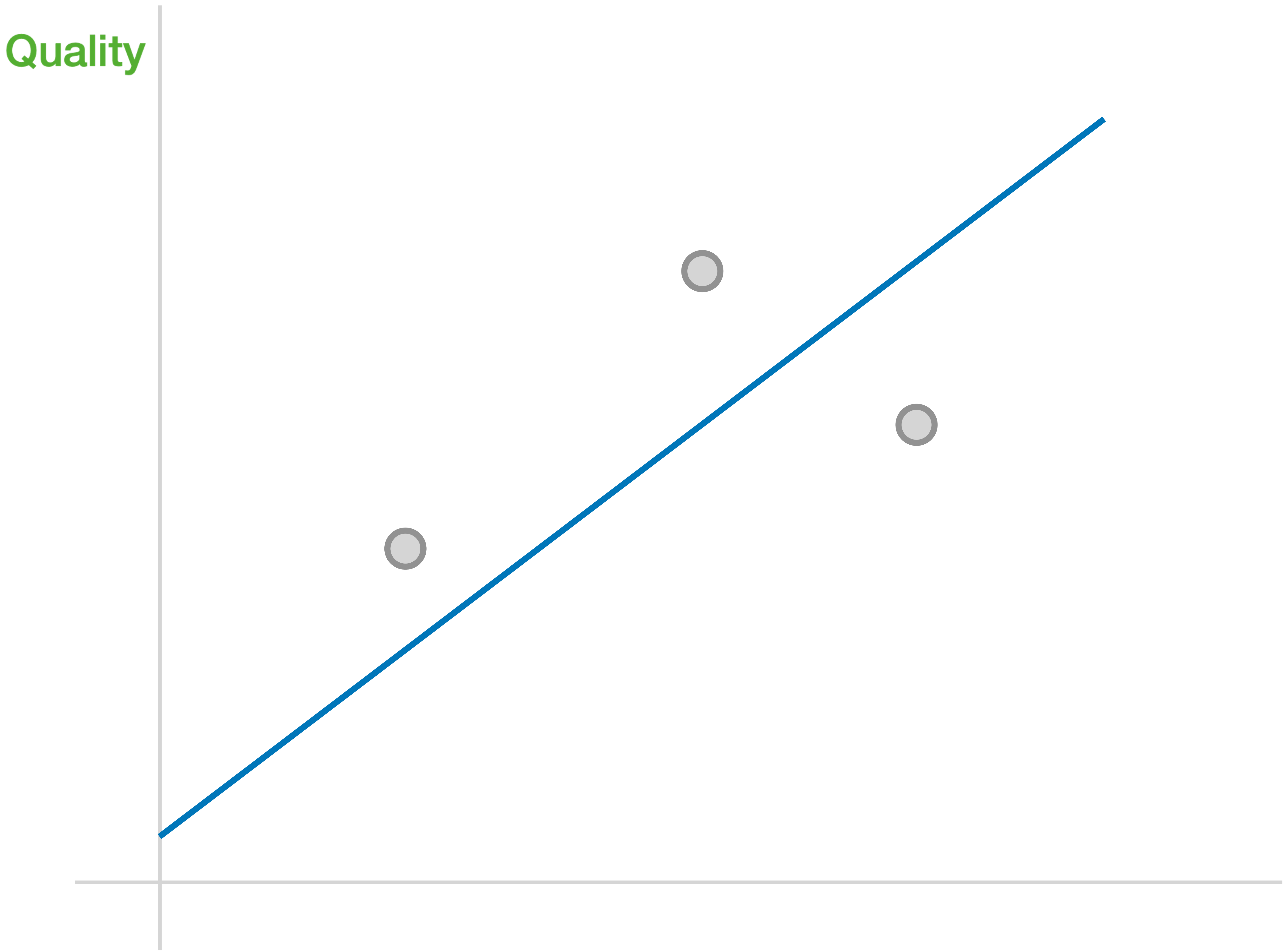
more data

Quality on the y axis

Acidity on the x axis



more data



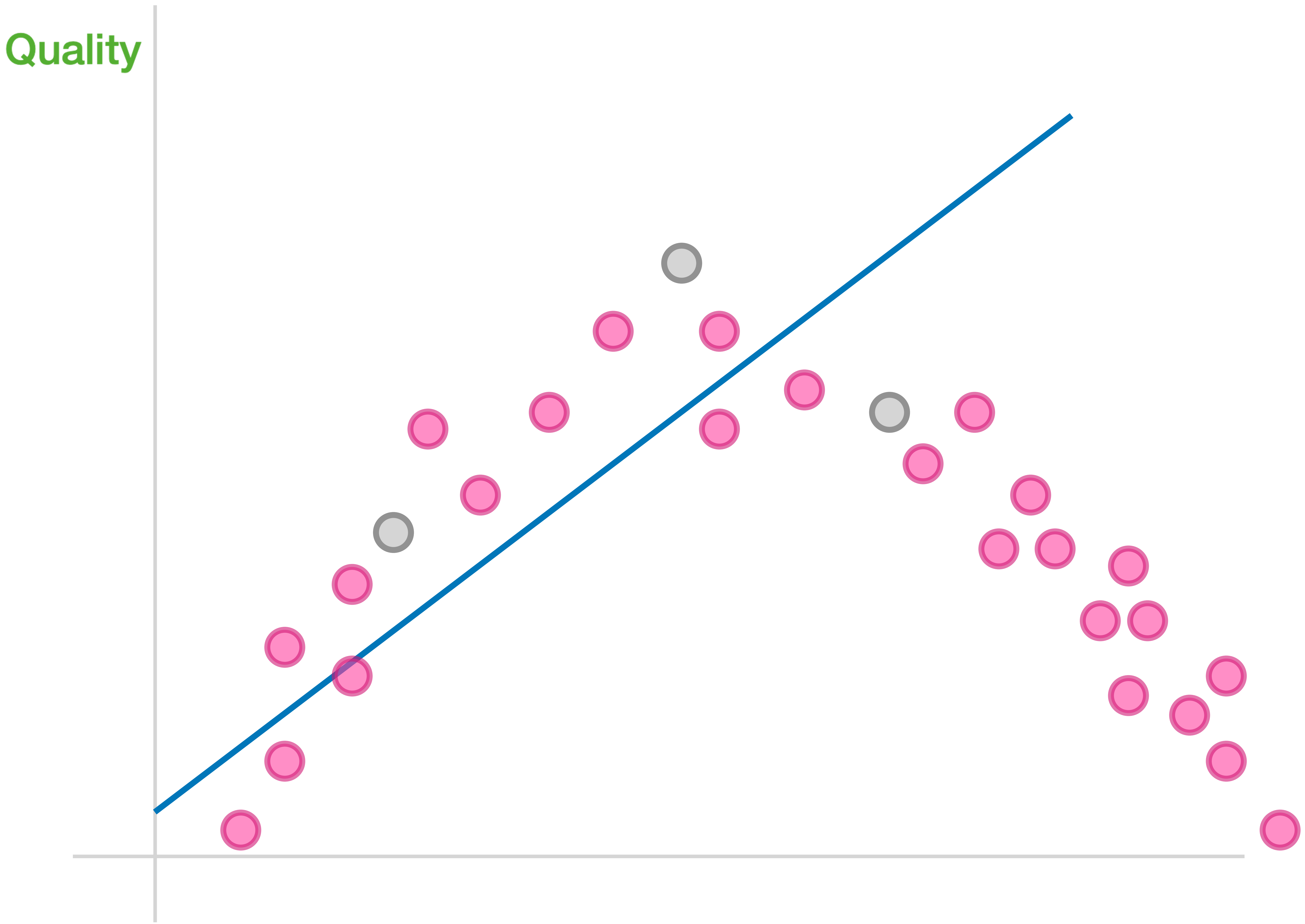
Quality on the y axis

Acidity on the x axis

more data

Quality **on the y axis**

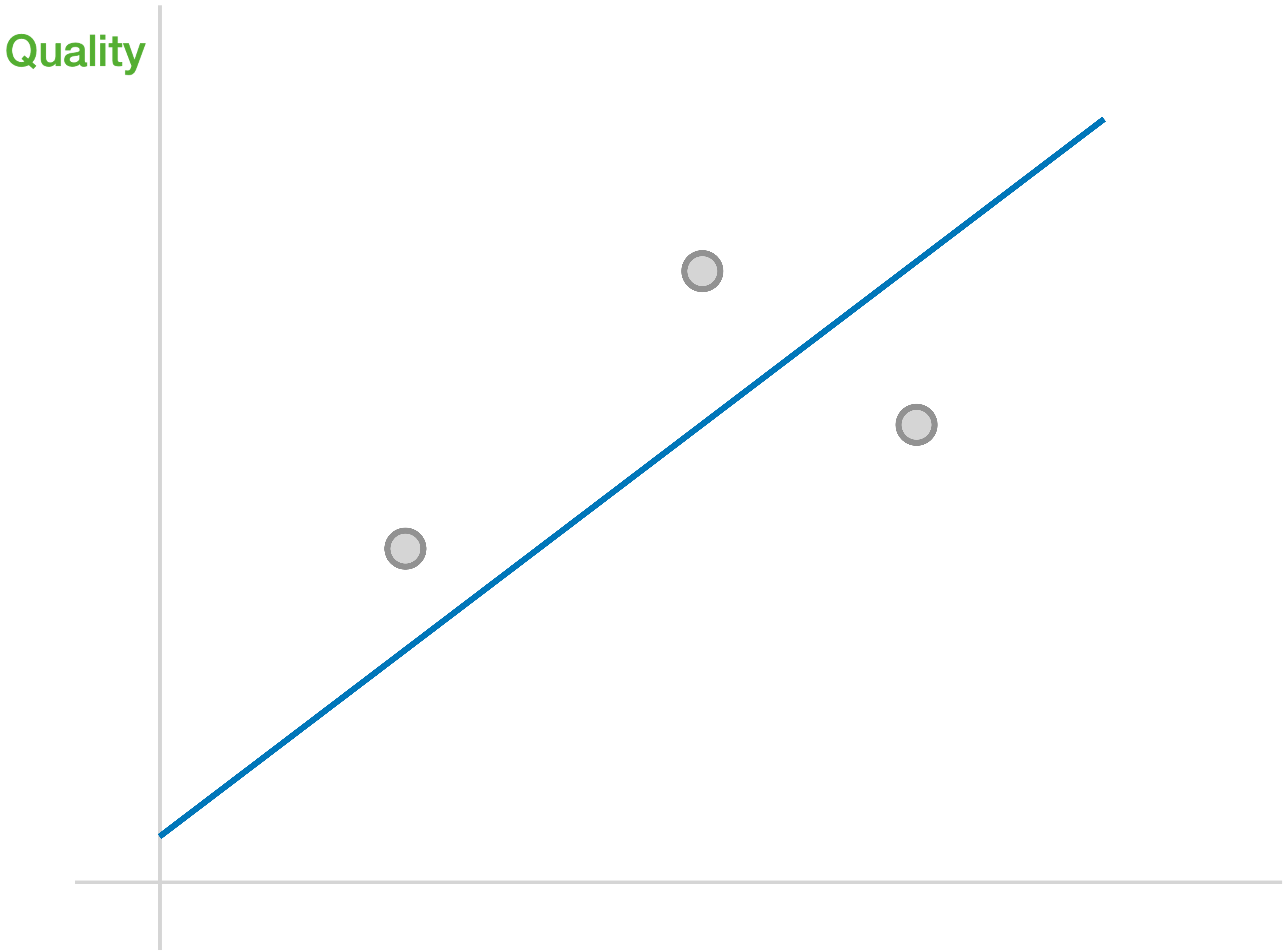
Acidity on the x axis



more data

Quality on the y axis

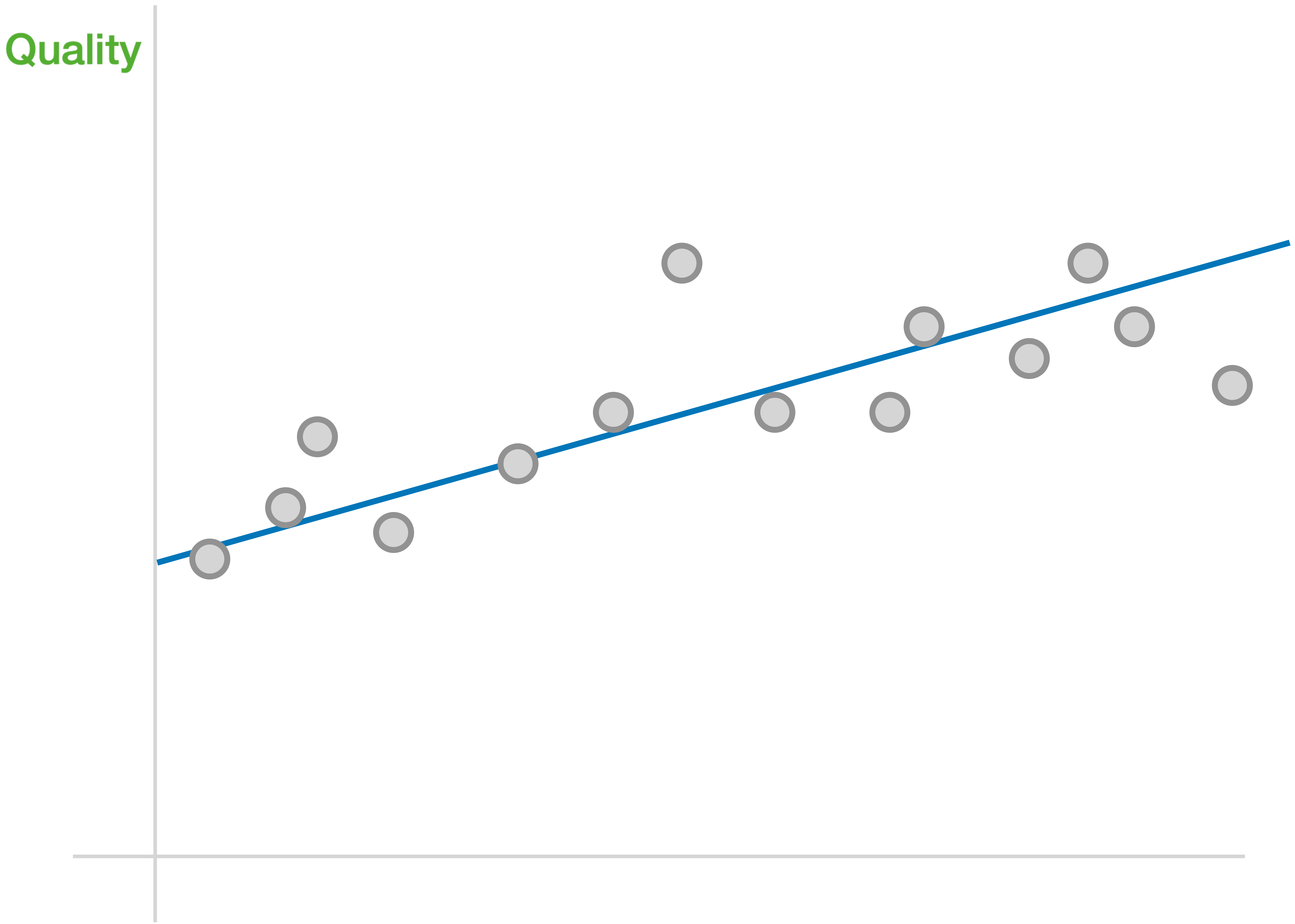
Acidity on the x axis



more data

Quality **on the y axis**

Acidity on the x axis



■ **use more data, get more accurate results**

more data **balanced** data **normalized** data **quality** data


Product


Product

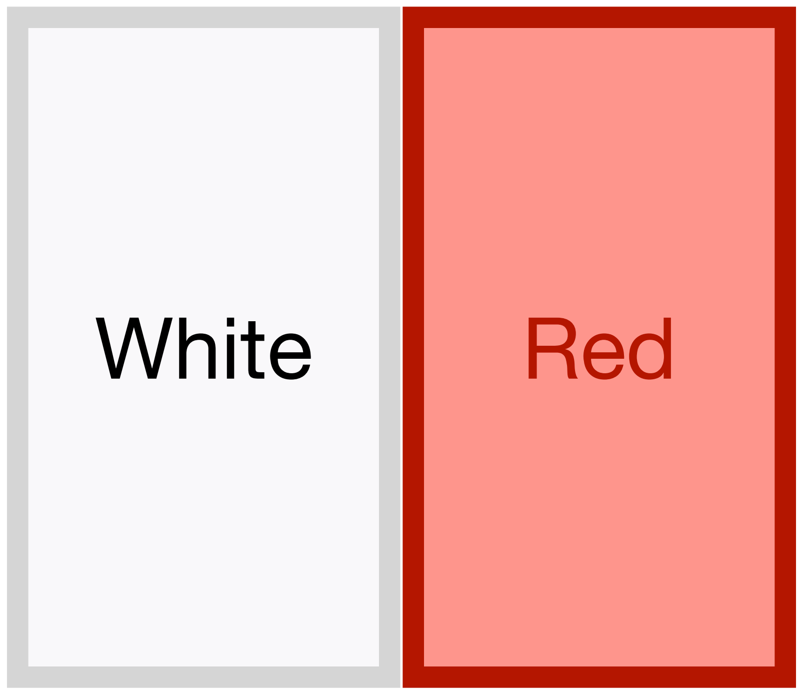

Amber Colored
Angora Super Sweet
Black Ethiopian
Burbank Slicing
Dona
Sophie's Choice
White Bush
Ace 55

more data **balanced** data **normalized** data **quality** data

more data **balanced** data **normalized** data **quality** data

Let's think about logistic functions!

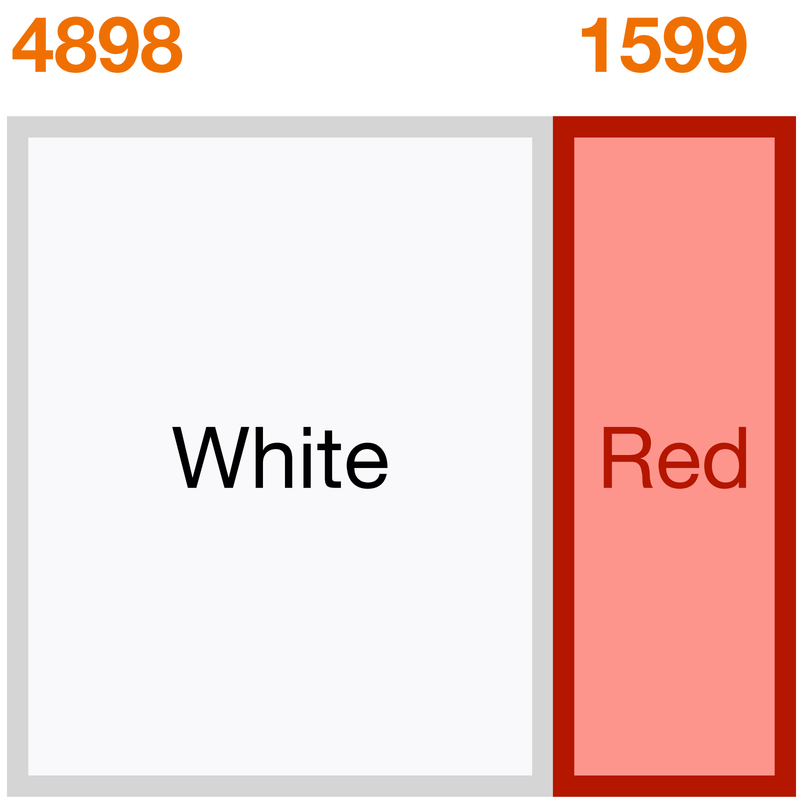
balanced data



in an ideal world

...but no

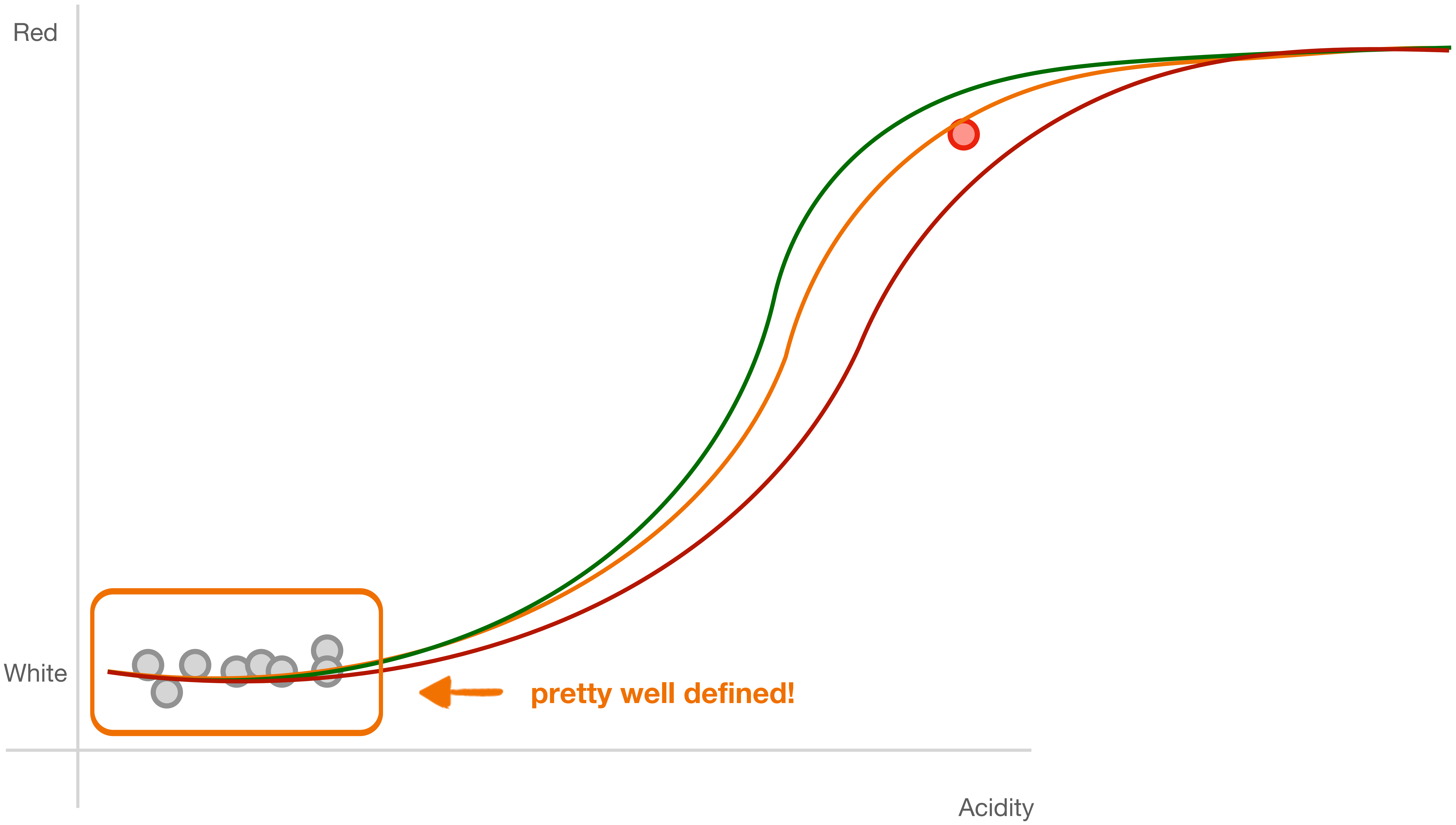
balanced data



What happens when we fit this dataset entirely?

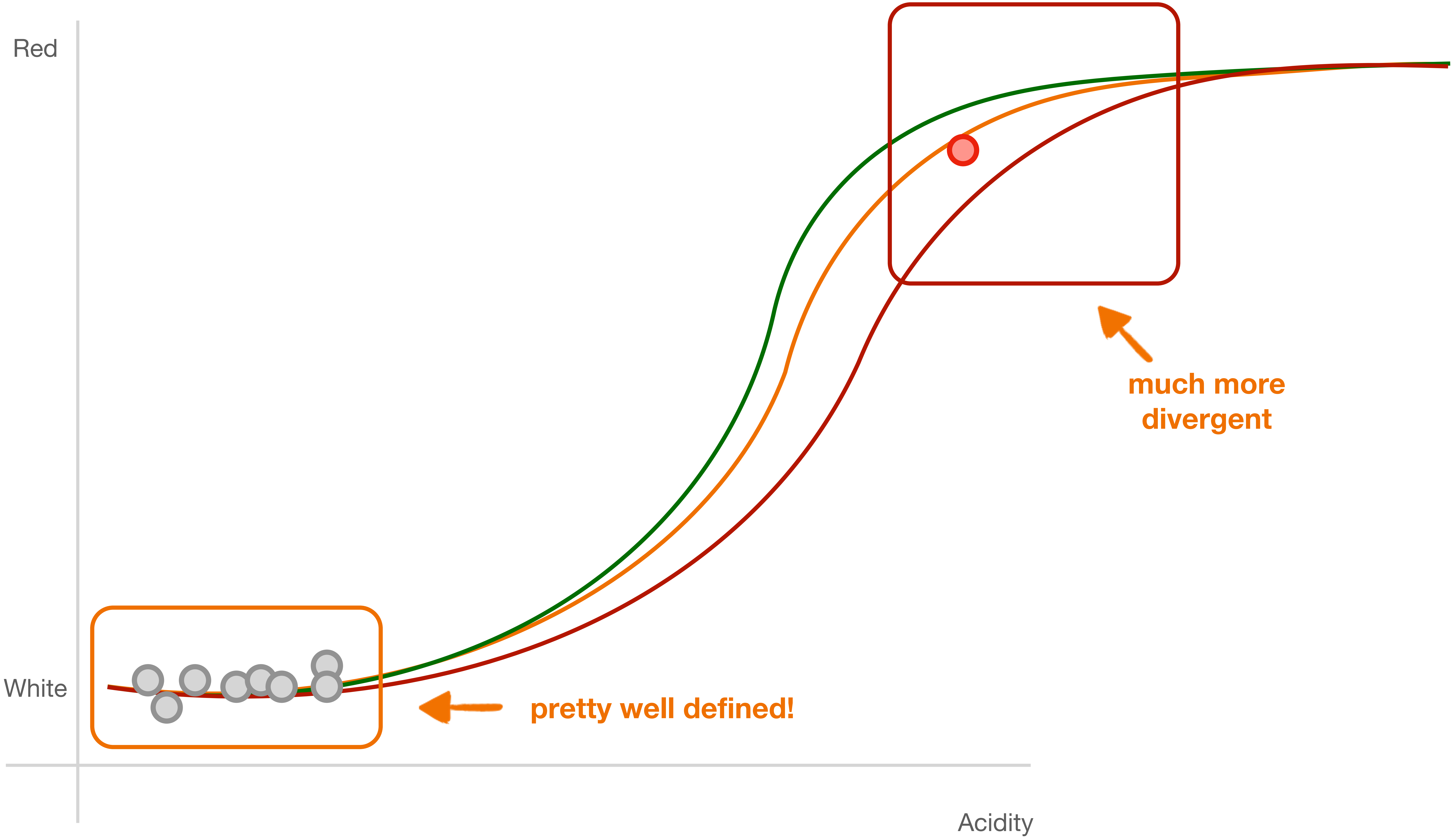
balanced data

Let's think about logistic functions!



balanced data

Let's think about logistic functions!

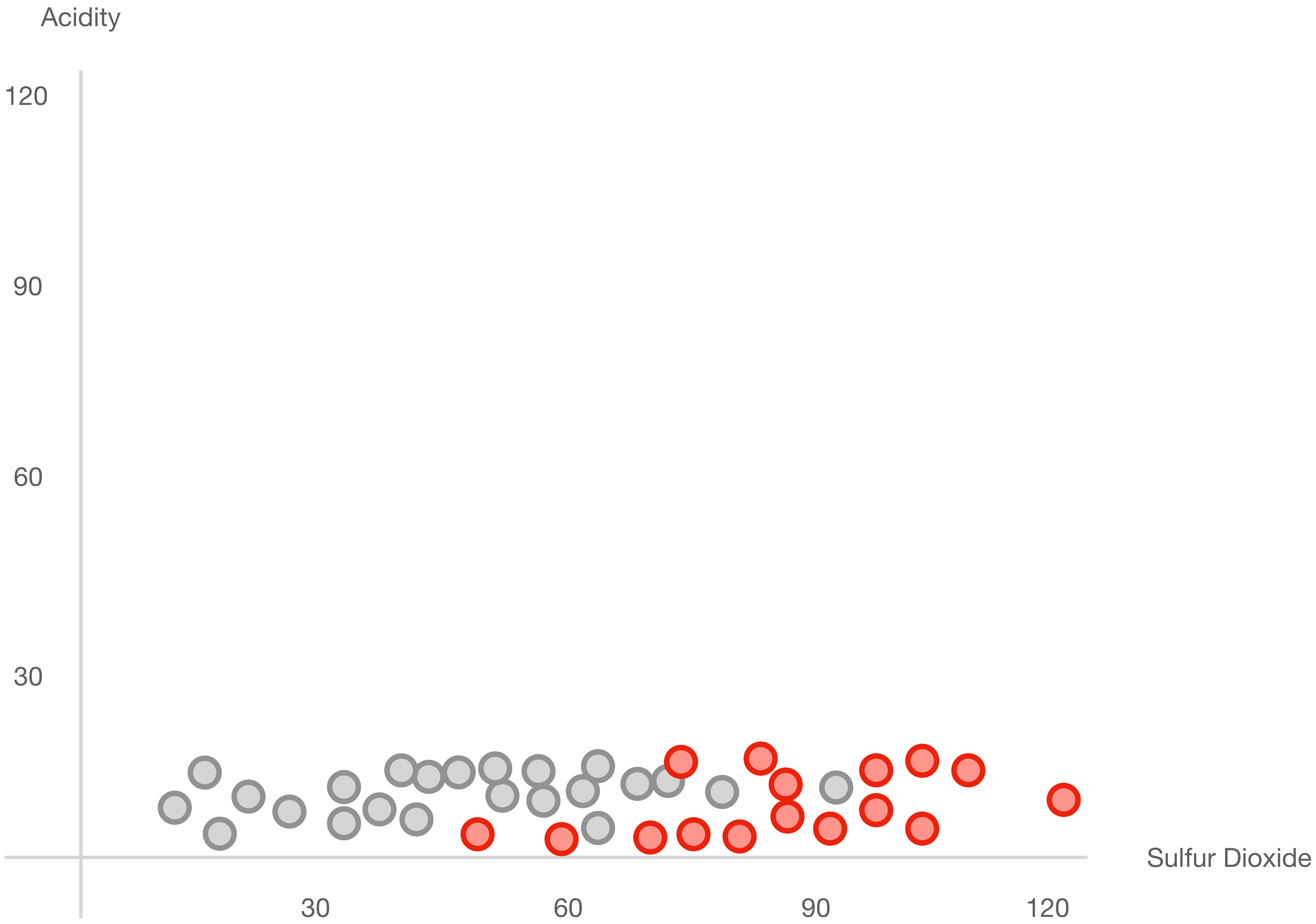


■ **balanced data, more accurate results**

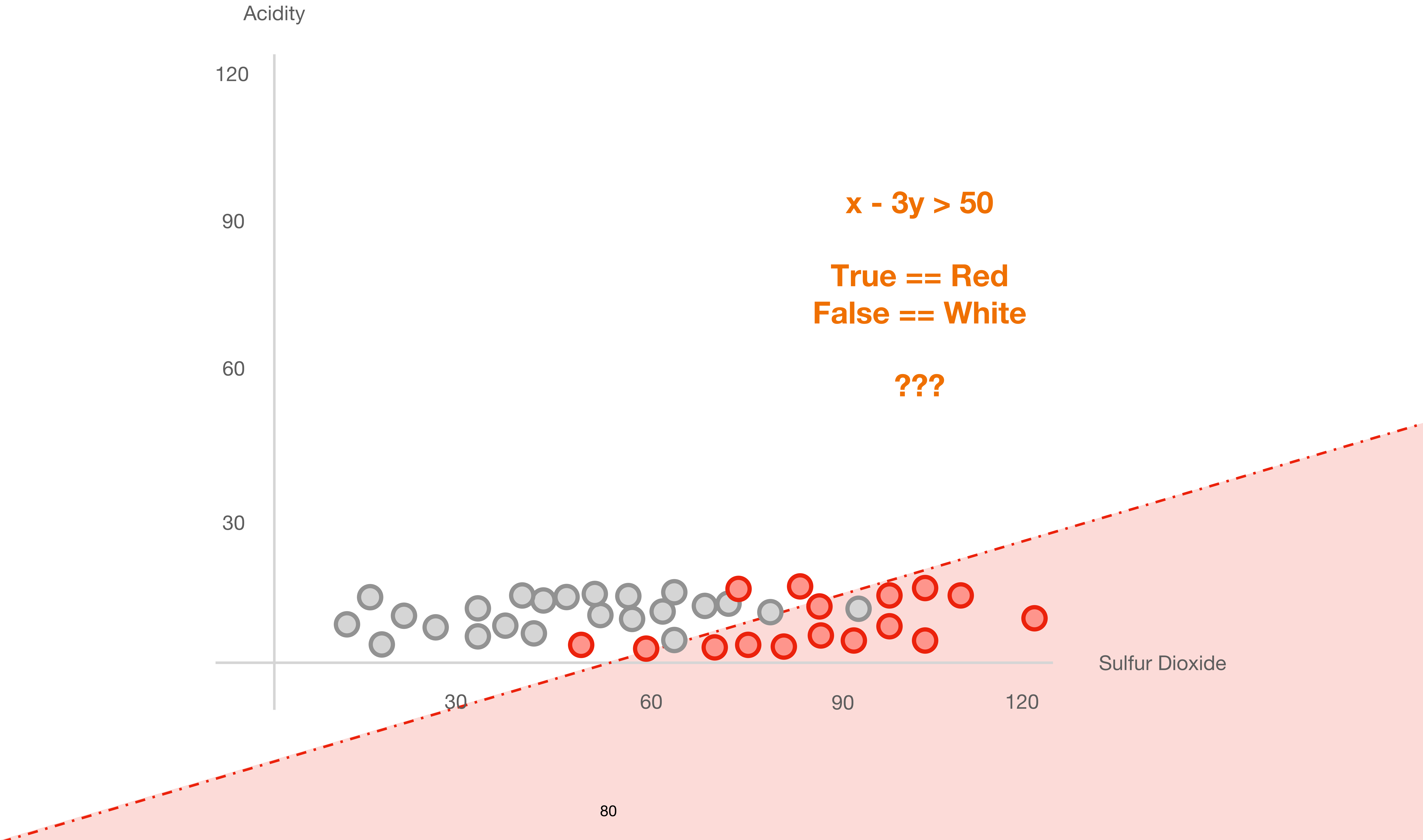
more data **balanced** data **normalized** data **quality** data

more data **balanced** data **normalized** data **quality** data

normalized data



normalized data



normalized data

■ **normalized data, better generalization, faster convergence**

more data **balanced** data **normalized** data **quality** data

more data **balanced** data **normalized** data **quality** data

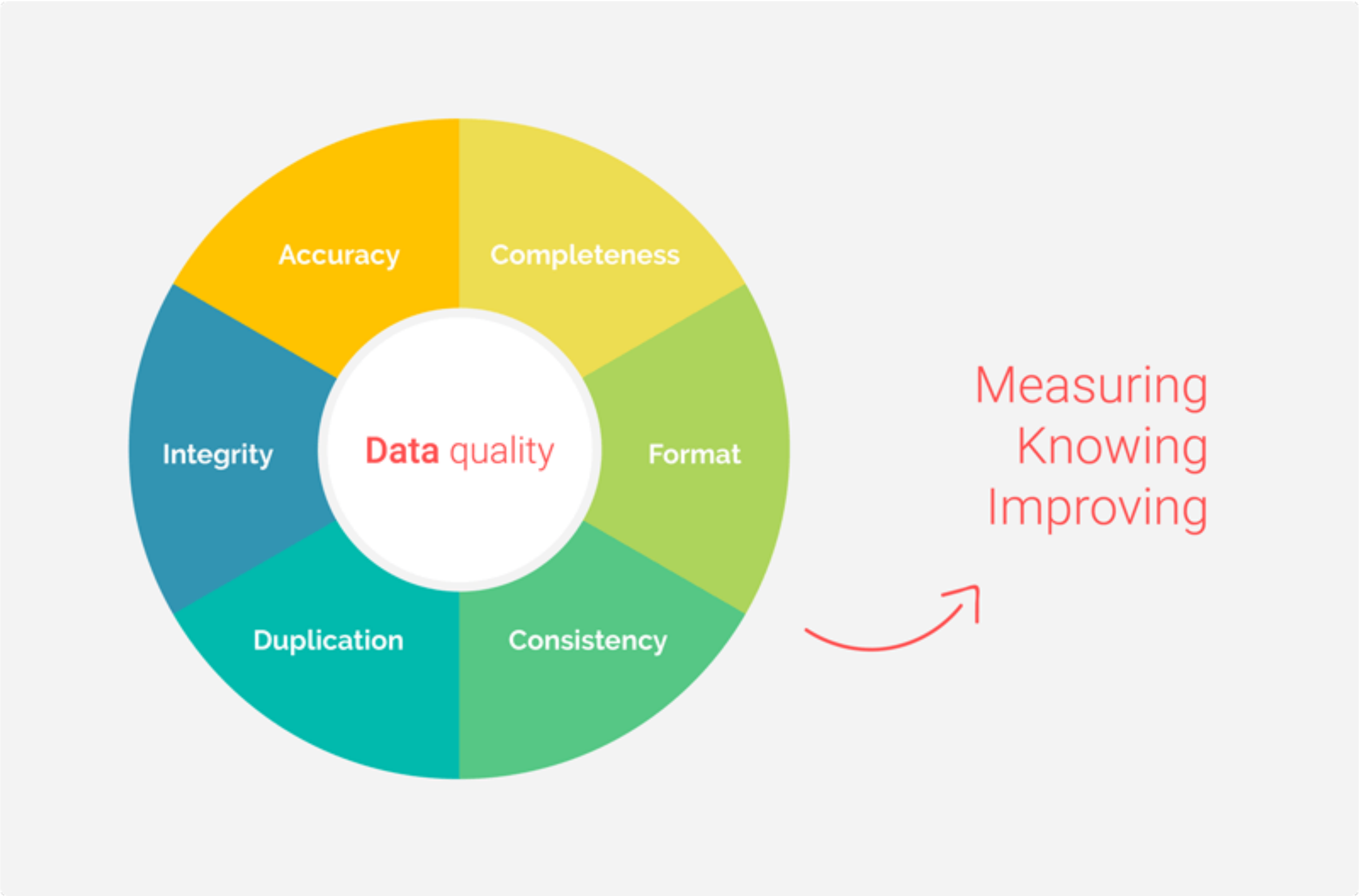


Image credit: Passionned Group

more data **balanced** data **normalized** data **quality** data

Missing Data

Missing Data

Missing **completely at random**

Missing **at random**

Missing **not at random**

Missing Data

remove

Use mean/most often

regression