1a) $n \to \infty$ flexible or inflexible?
number of predictors is small.
Flexible method should be better as the $n \to \infty$, fit the data better.

b). inflexible method. If we go with flexible methods, → overfitting.

c) is highly non-linear. → flexible. With more degrees of freedom, a flexible model would be a better fit.

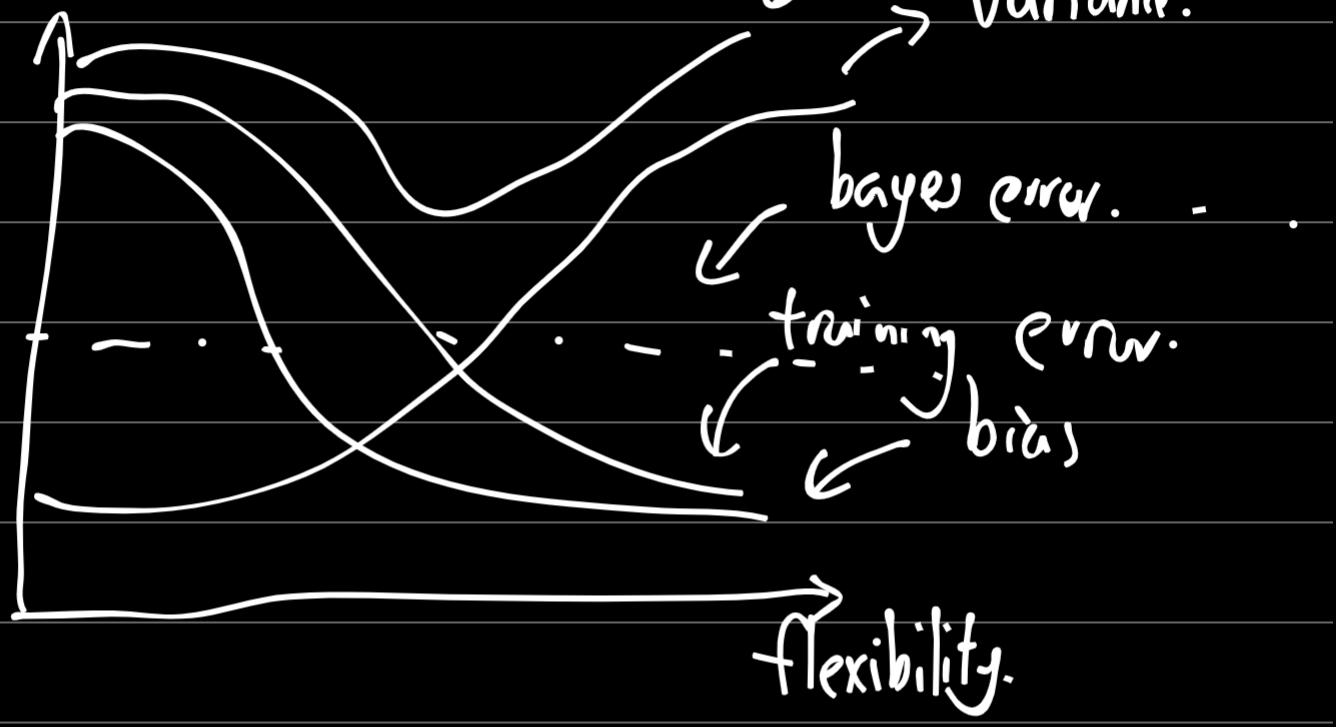
d) variance of error very high → bias is very low.
inflexible is better.

2a) $n = 500$. p = profit, number of employees, industry
Regression. Inference.
Quantitate output of CEO salary.

b) Classification of a qualitative success or failure.
Prediction,
$n = 20$
p = price, budget, competition and 10 other var.

c) p = % change in the US market, British, German.
n = 2012 daily data.
prediction. regression. Quantitative output of the % change.

3a)



5) flexible vs inflexible.

- complex data set with alot of strong features.
- A lot of data
- Interested in prediction.
- might cause overfit → low bias → high variance

- only a few features are relevant.
- little data.
- Interested in ~~reference~~ inference.

Parametric:

6) Advantages
- reduce the problem of estimating a function f down to a set of parameters.
- Might not capture the true f/s.

Non-parametric:
- does not depend on a set of parameters and require a large set of data.

7)a). Euclidean distance = straight line

Obs:	Euclidean
1	3.
2	2
3	$\sqrt{1^2+3^2} = \sqrt{10}$
4	$\sqrt{(1)^2+2^2} = \sqrt{5}$.
5.	$\sqrt{2}$.
6	$\sqrt{3}$.

b) when K=1, prediction: Red. Green.

c) K=3
1G 2R
∴ Red.

d) If the Bayes decision Boundary is highly non-linear → complicated. → irreducible error → high. → small.