

Week 3: Non-Parametric Models

Matthew Caldwell

COMP0088 Introduction to Machine Learning • UCL Computer Science • Autumn 2023

Admin

- Bartlett coursework students
- Pulse surveys

Week 3 Recap

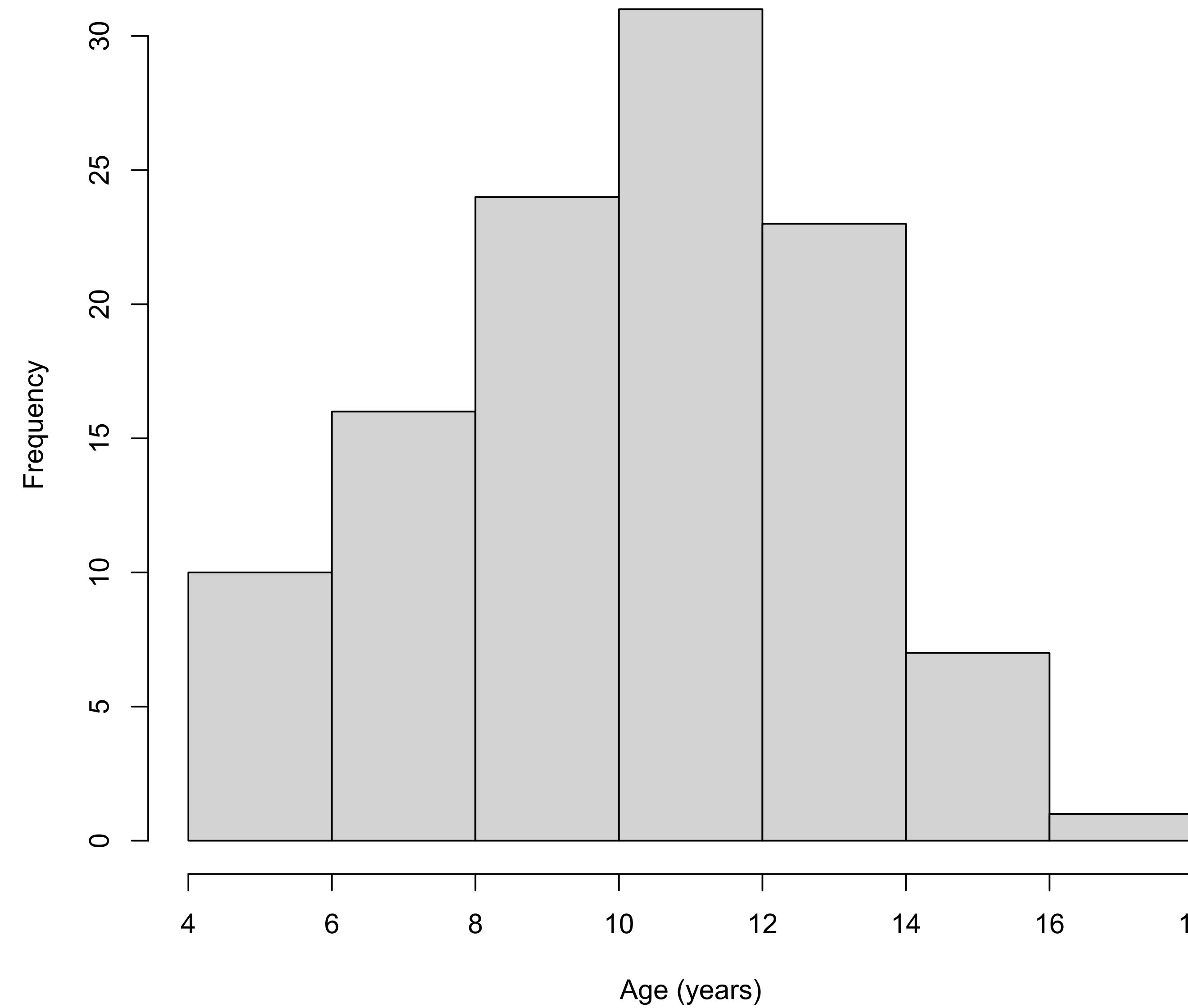
Crosseyed & Painless



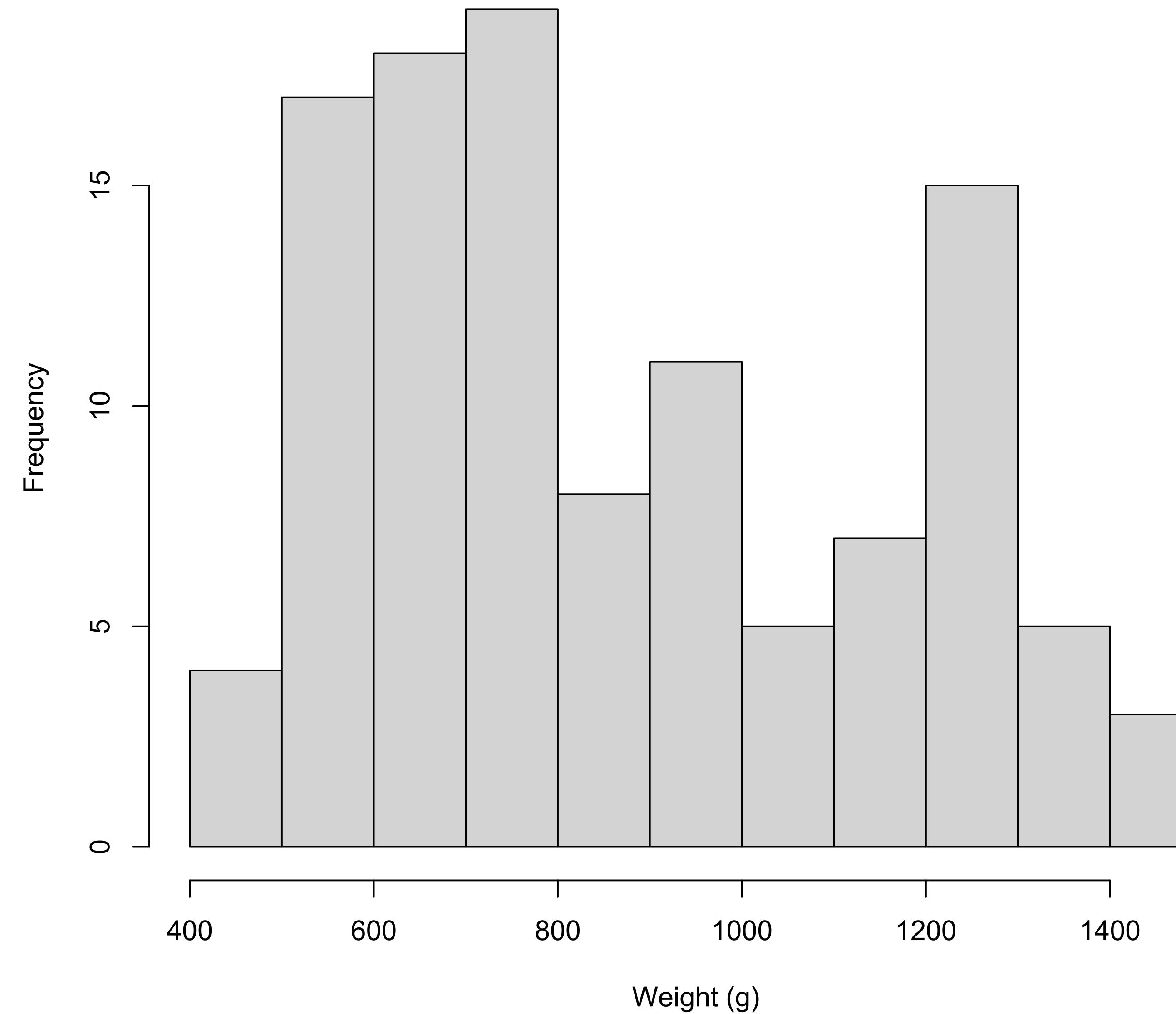
Meerkats

#	Age	Weight	Sex	Captive	...
1	11.8	815	1	0	
2	9.7	672	1	0	
3	8.9	446	1	0	
4	10.8	761	0	0	
5	8.3	1035	0	1	
6	11.7	930	1	1	
7	8.5	1027	0	1	
8	7.6	1234	0	1	
9	15.5	1461	0	1	
10	8.8	720	1	0	
11	12.8	1223	0	1	
12	7.7	711	1	0	
13	7.9	586			

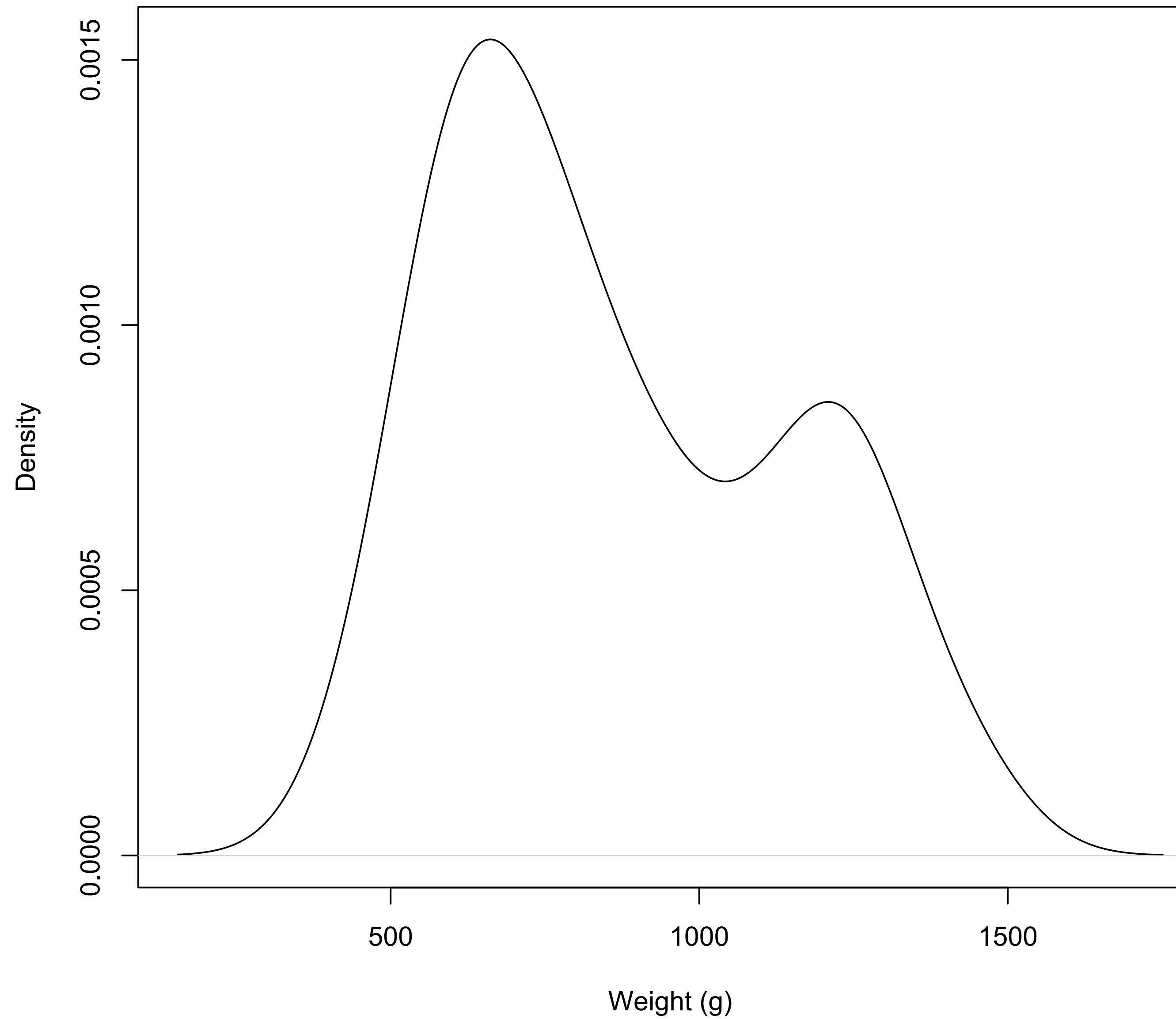
Meerkat Age Distribution



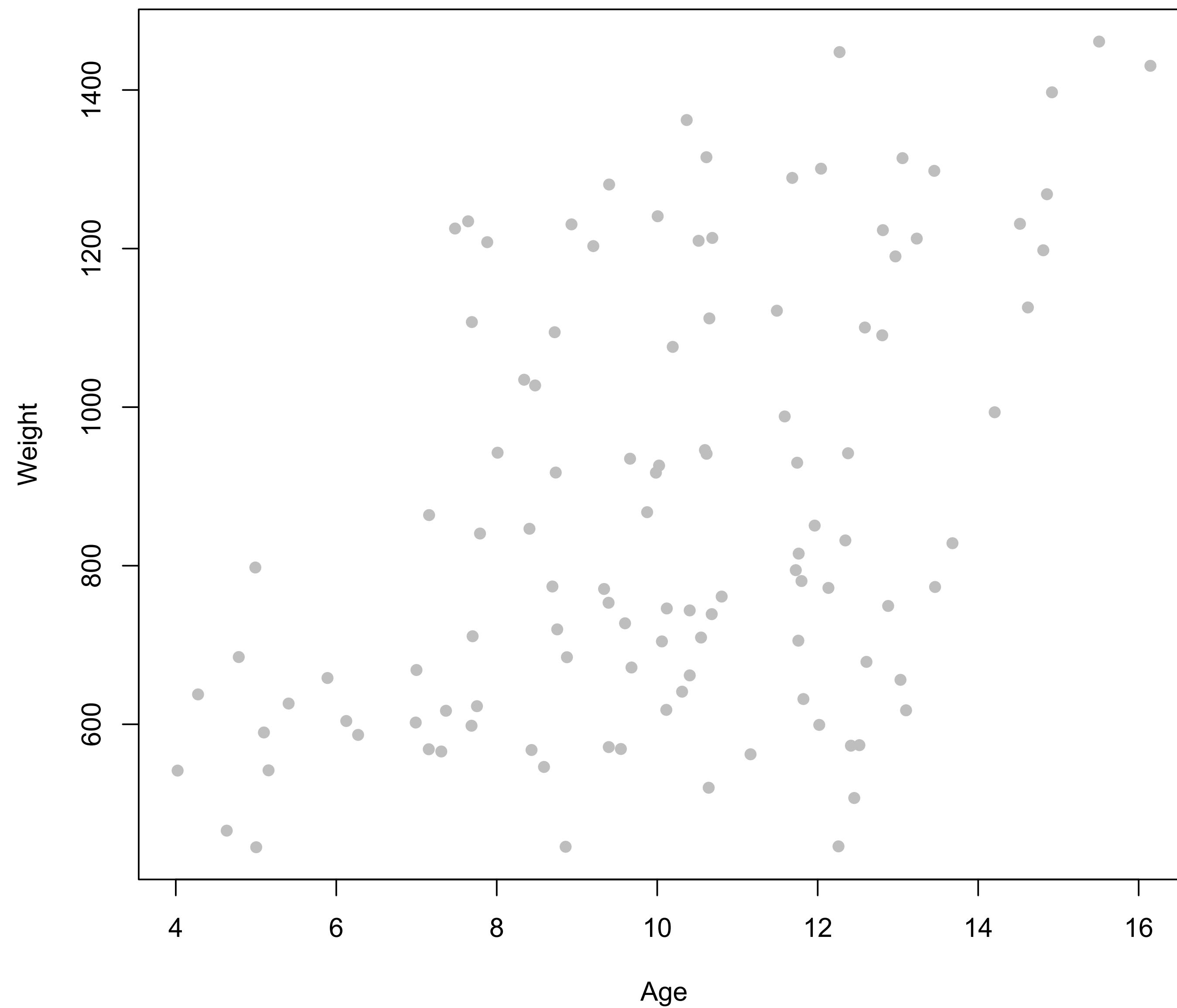
Meerkat Weight Distribution



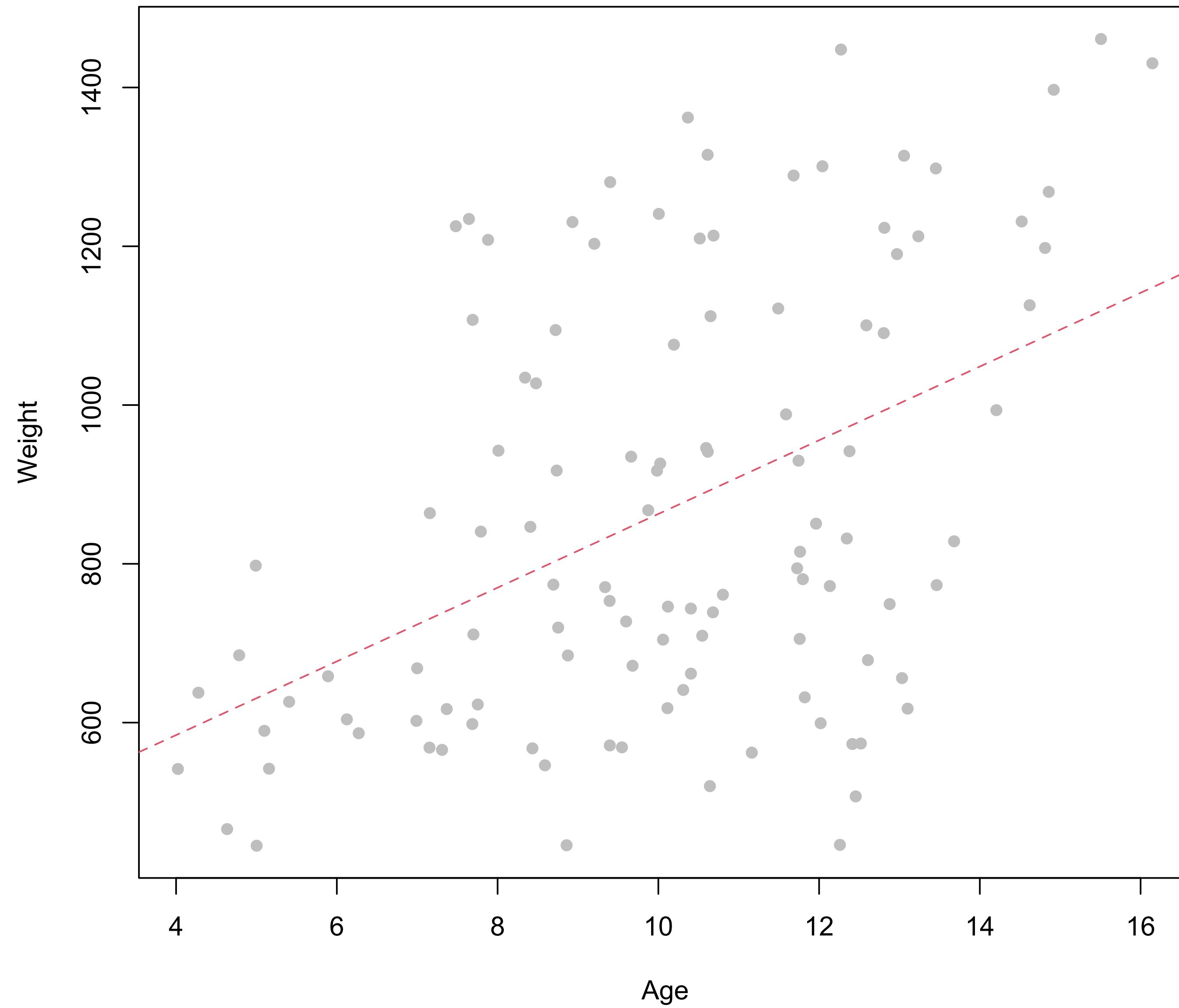
Meerkat Weight Distribution (KDE)



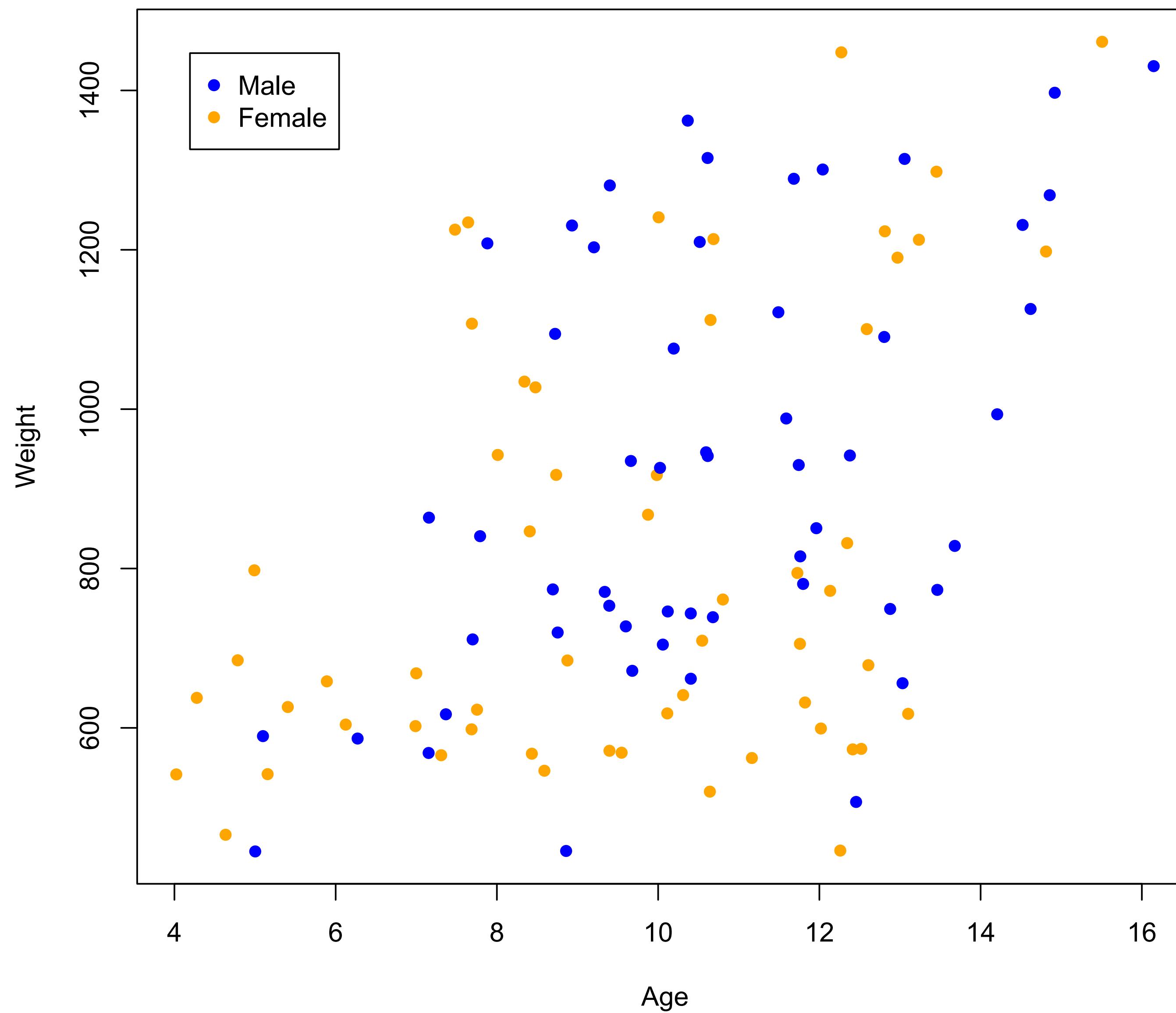
Meerkat Weight vs Age



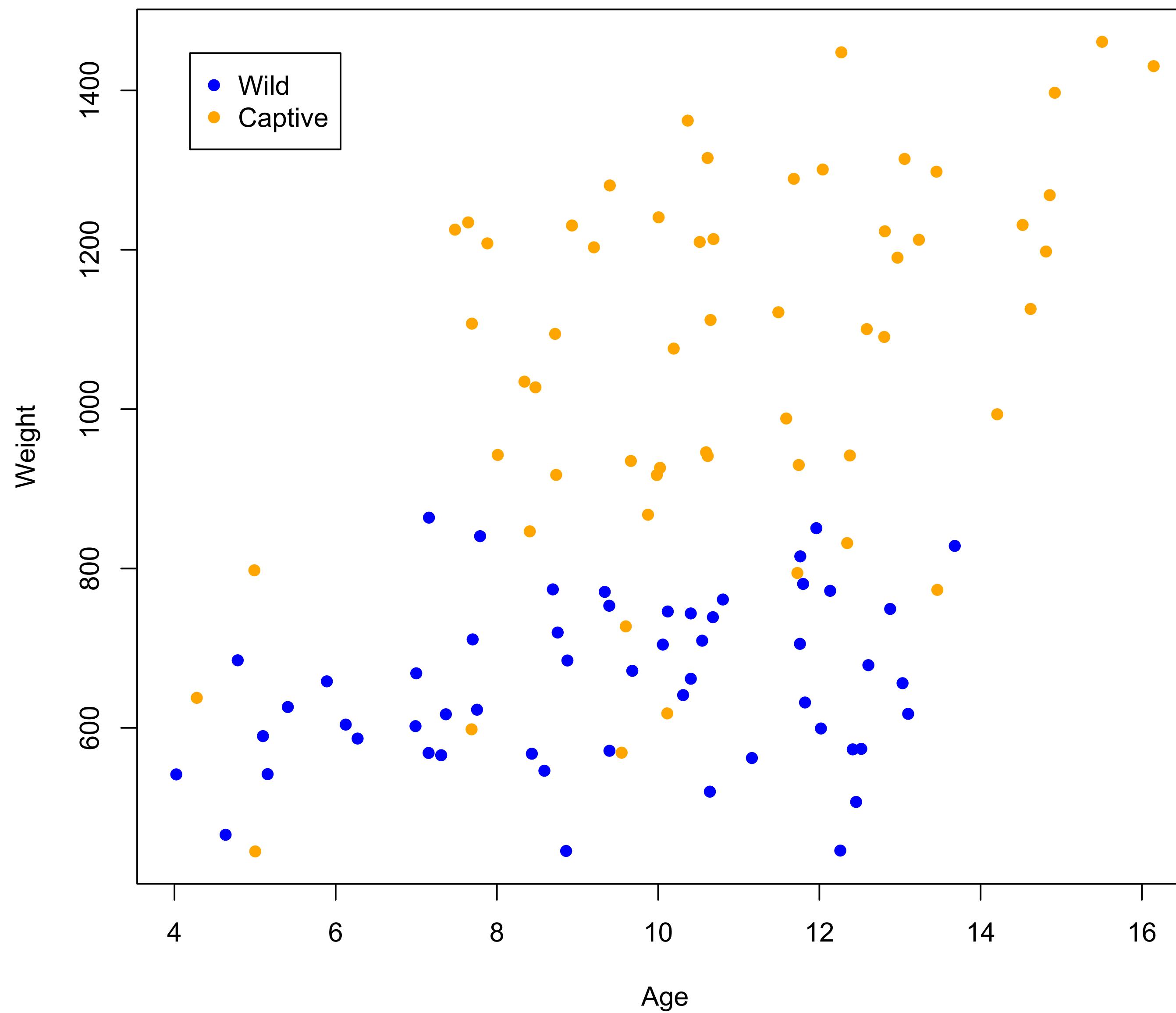
Meerkat Weight vs Age



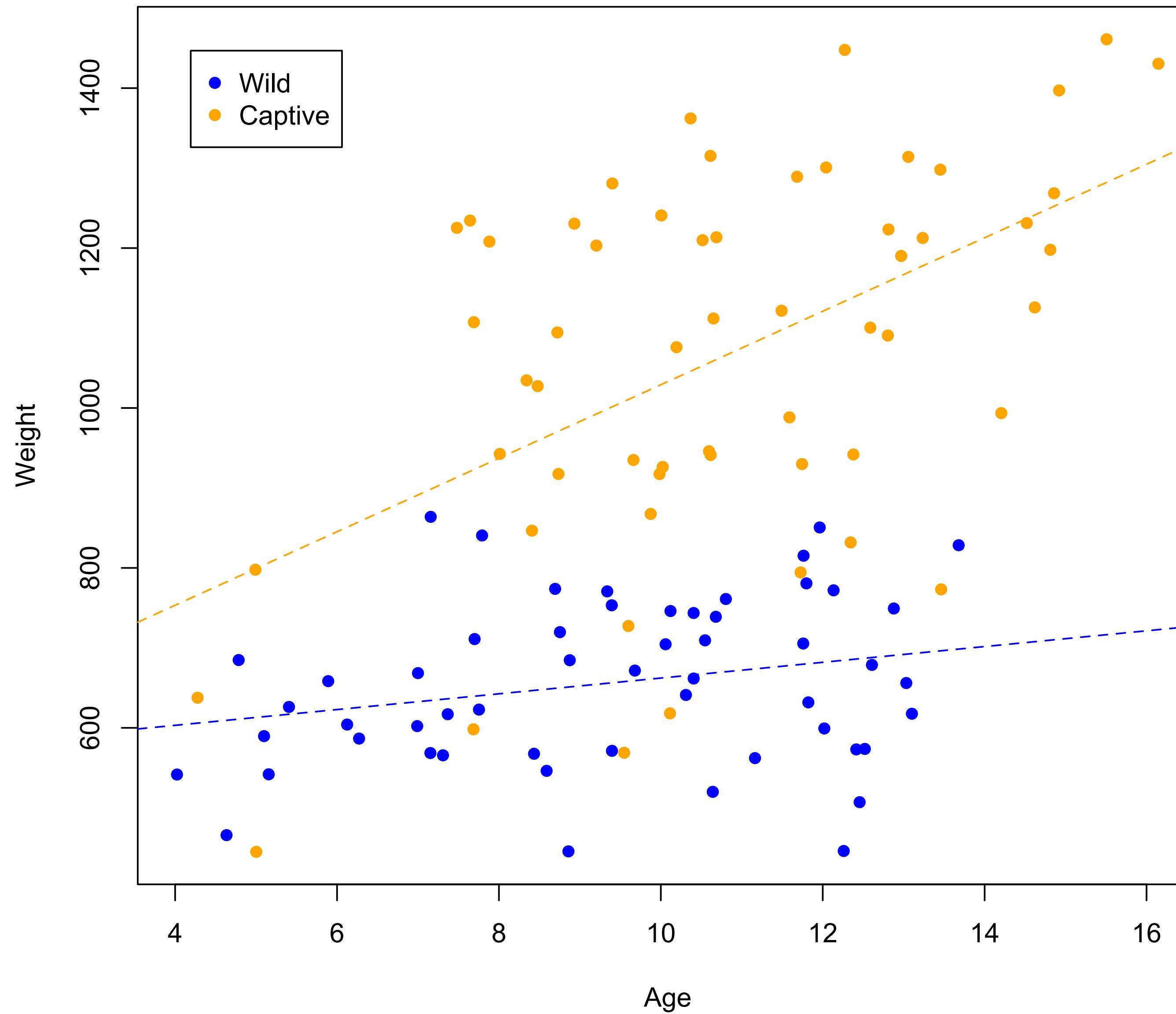
Meerkat Weight vs Age



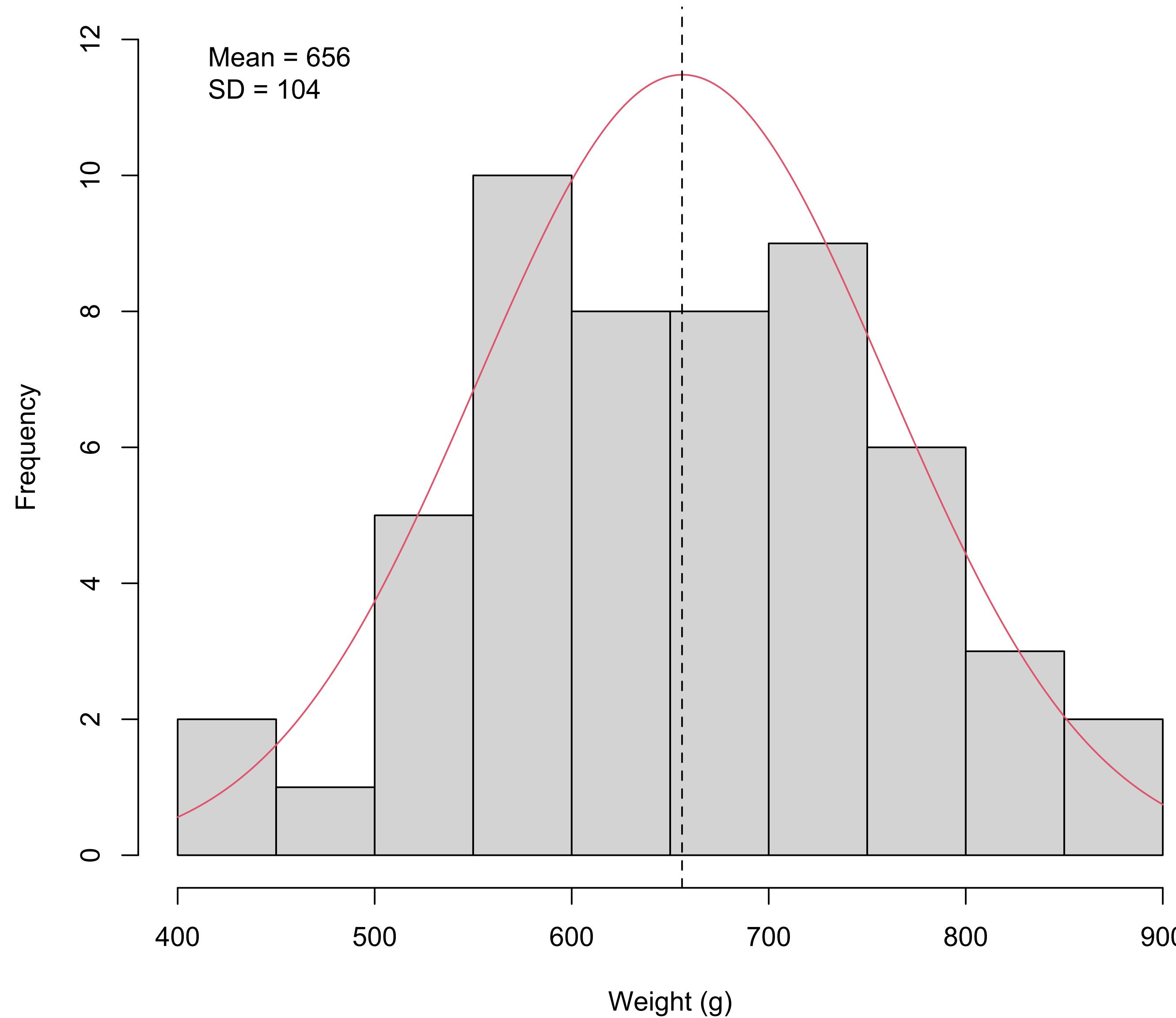
Meerkat Weight vs Age

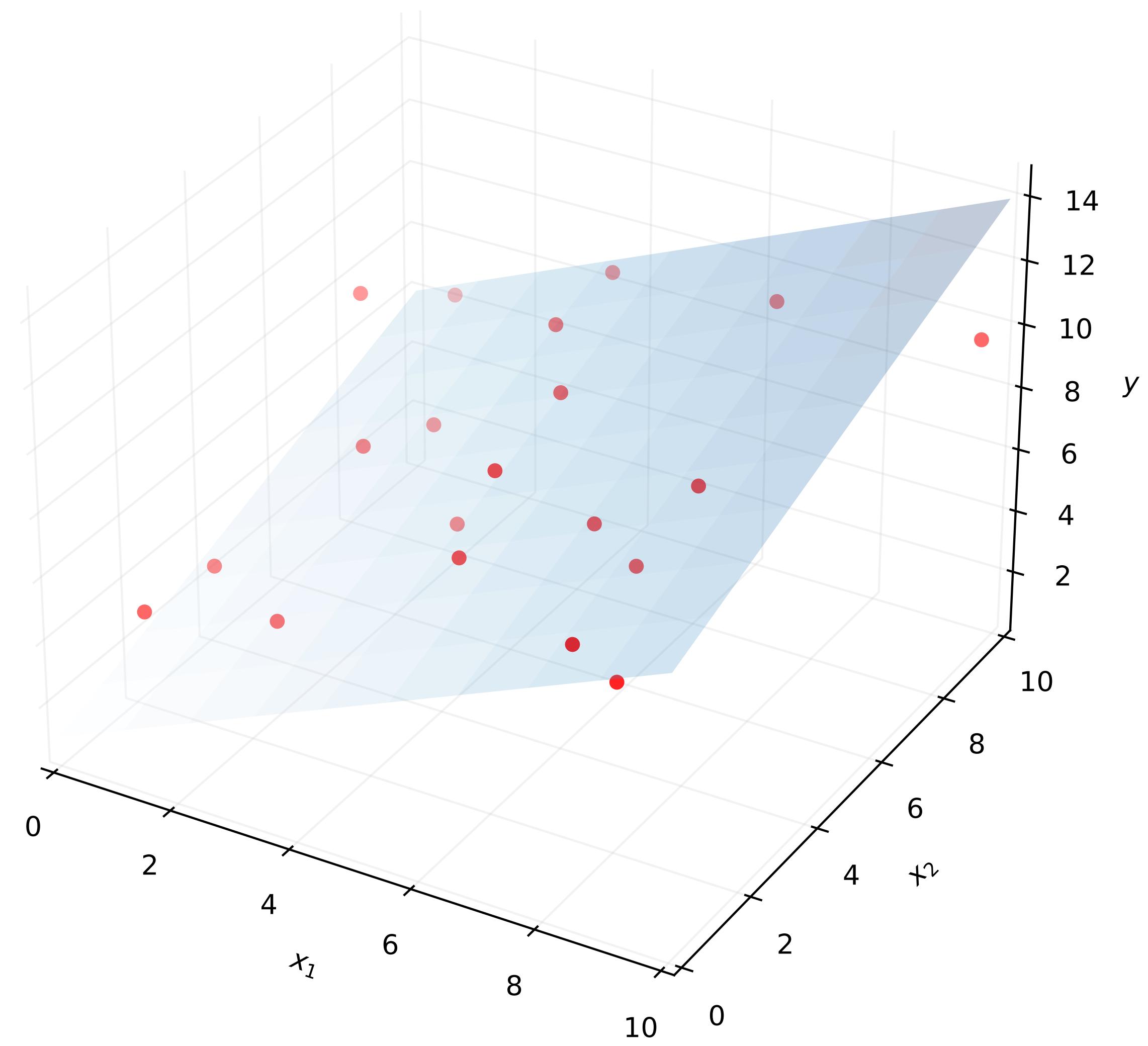
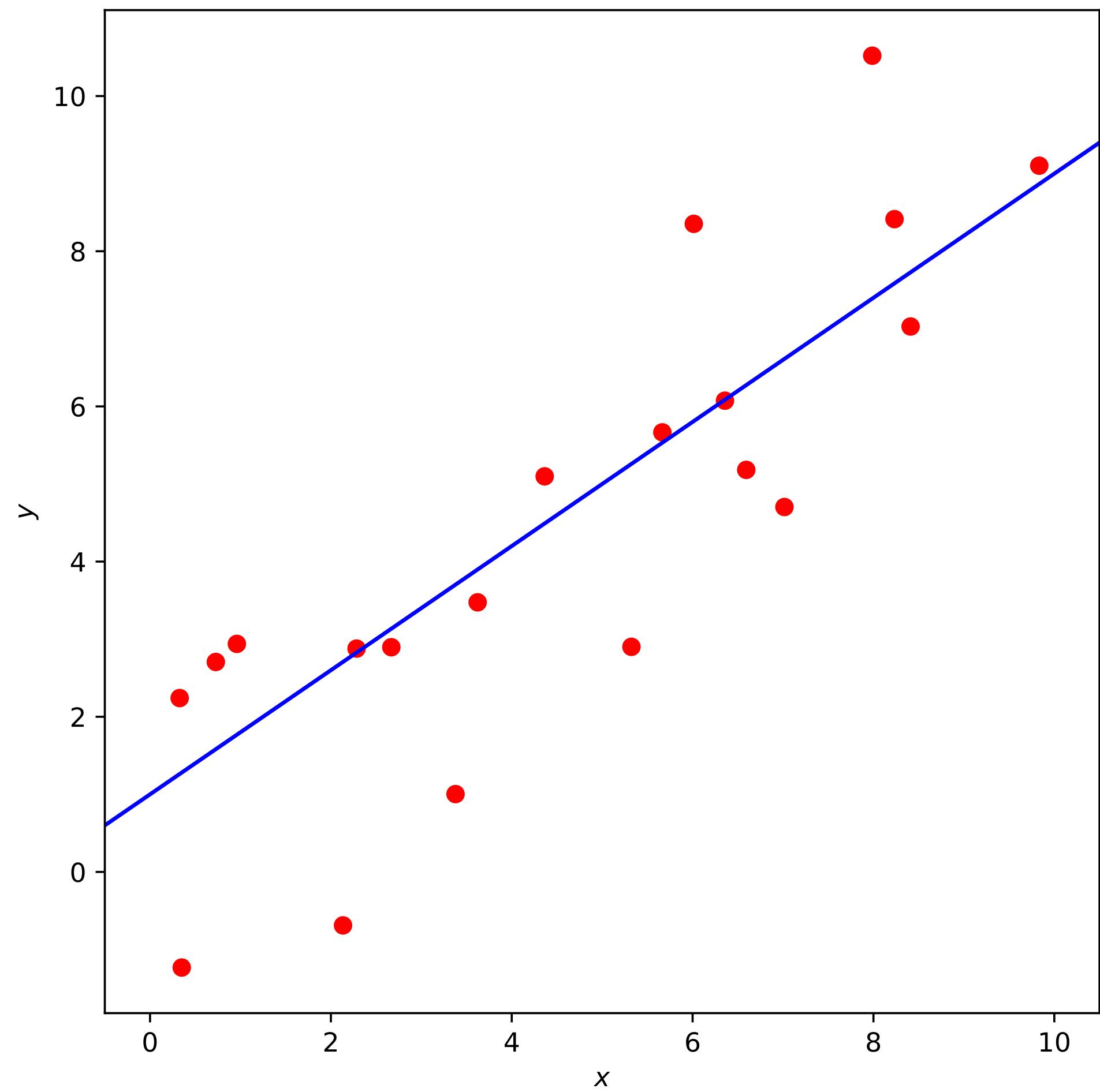


Meerkat Weight vs Age

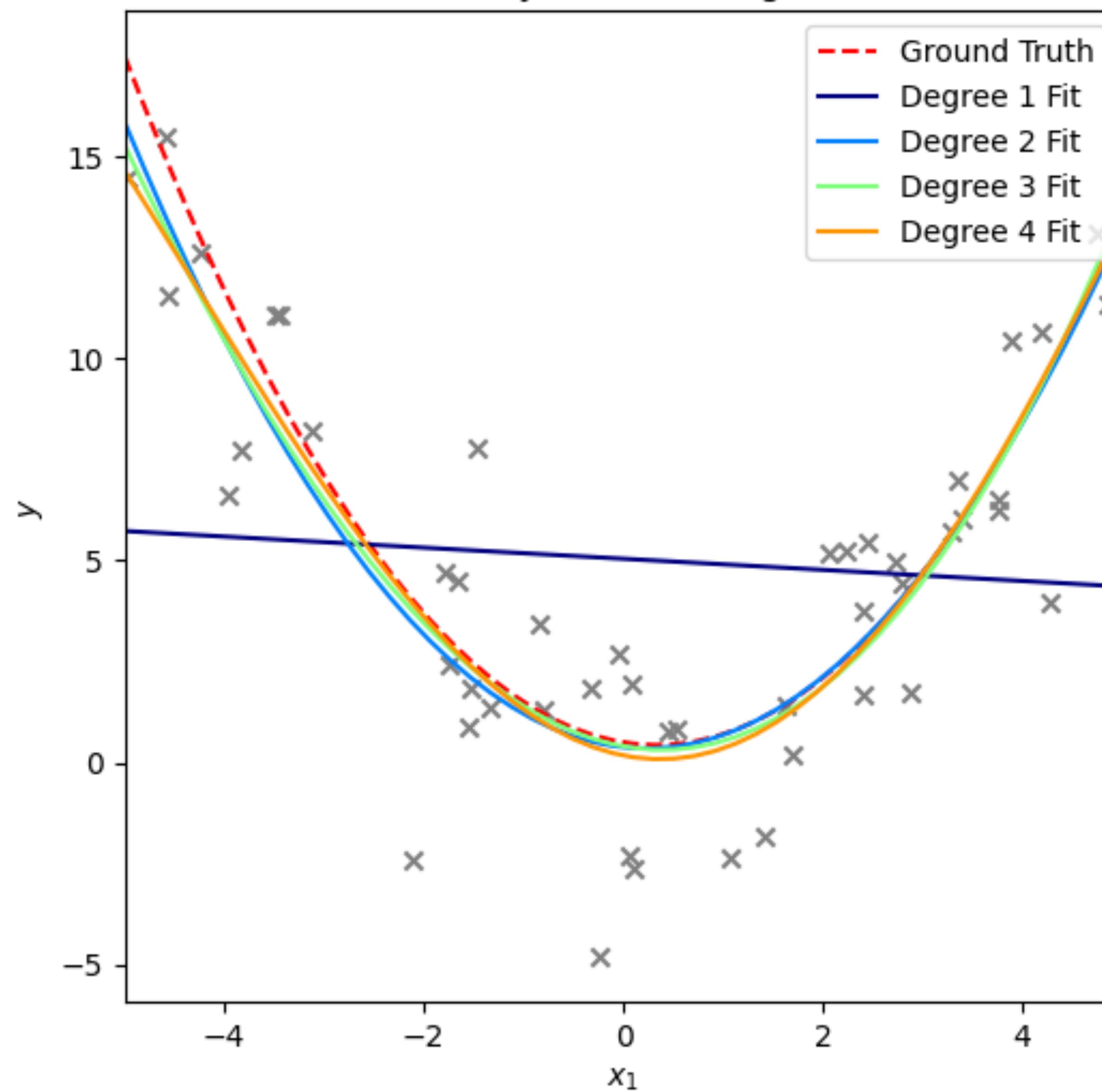


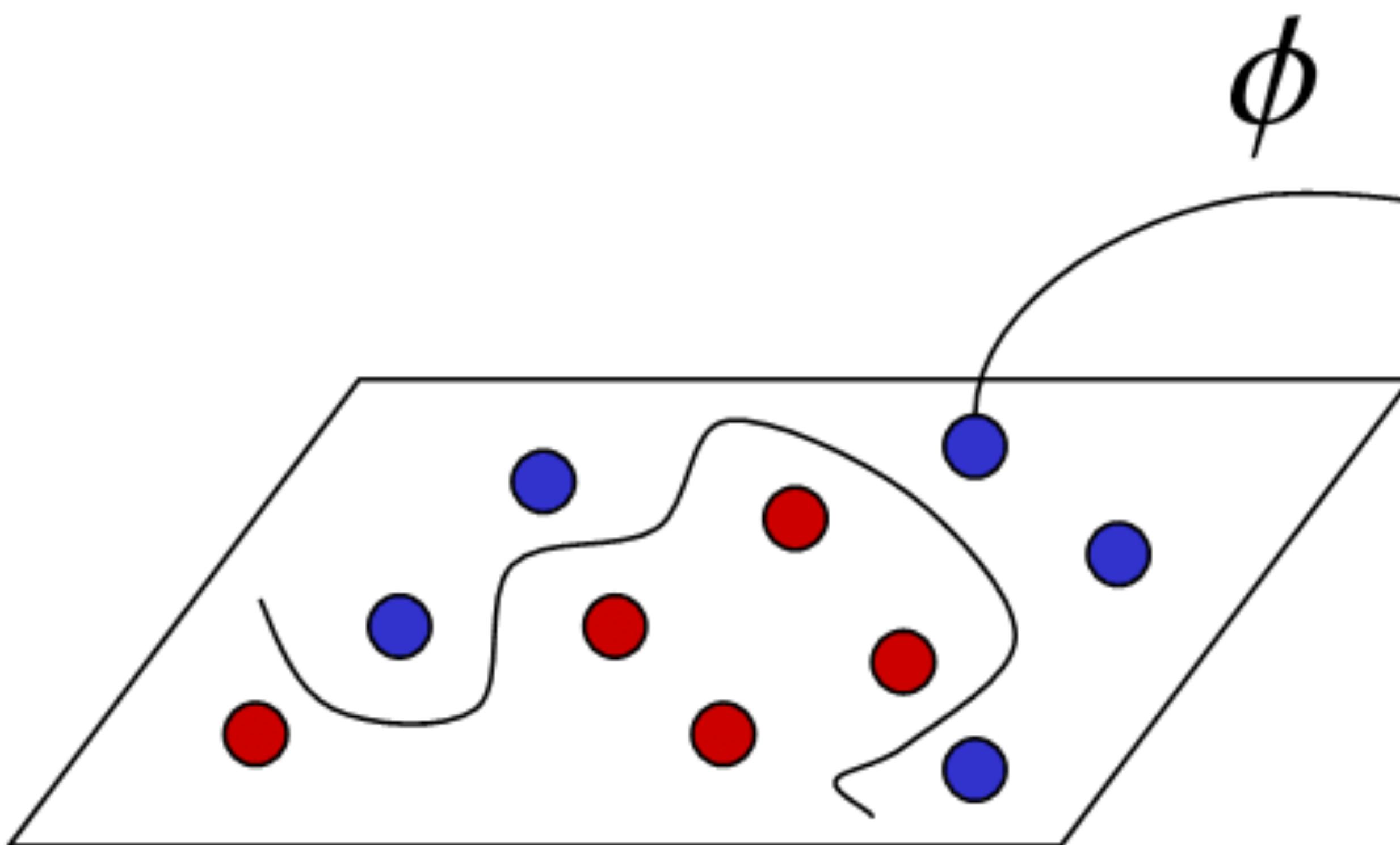
Wild Meerkat Weight Distribution



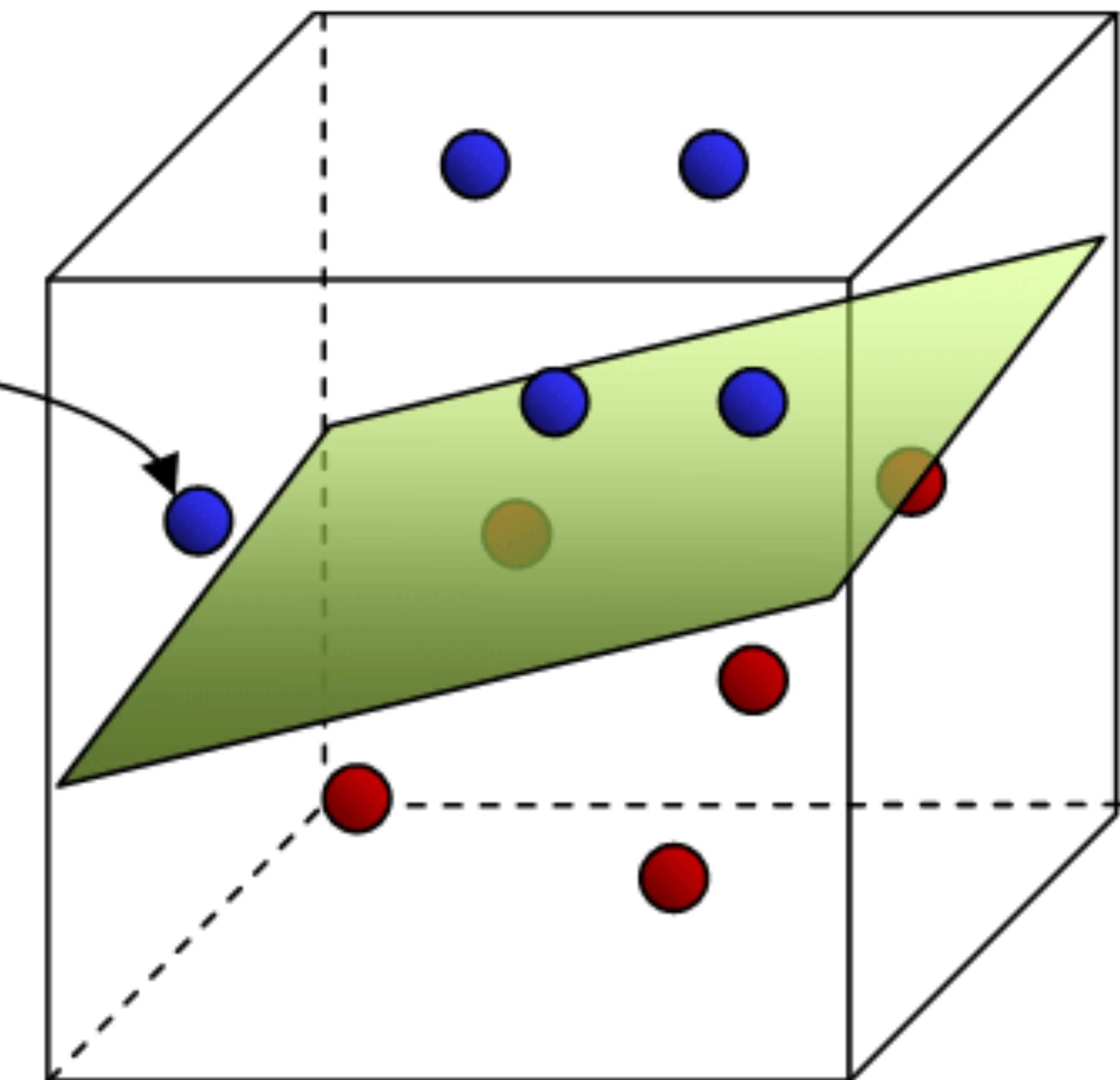


Polynomial Fitting

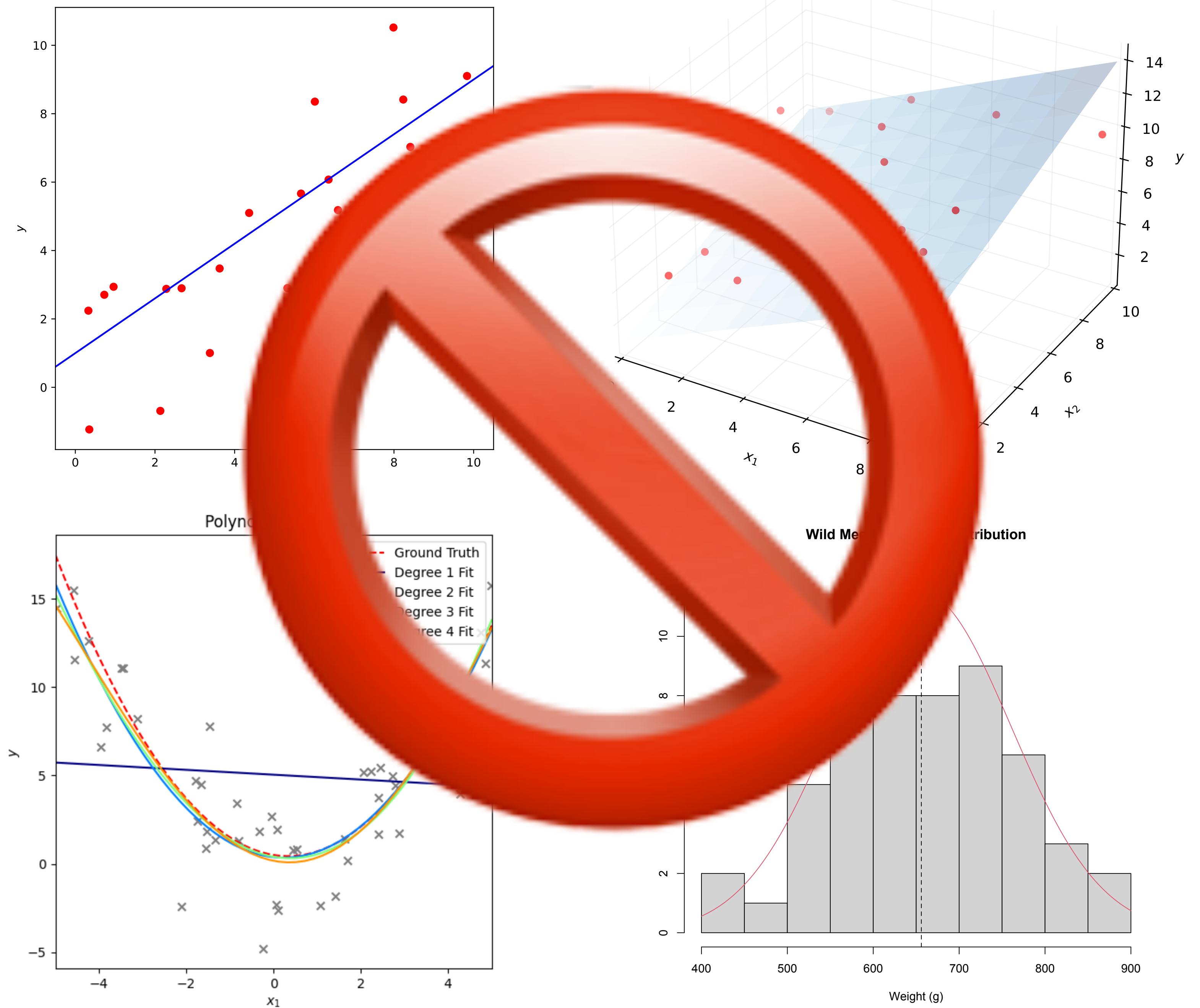




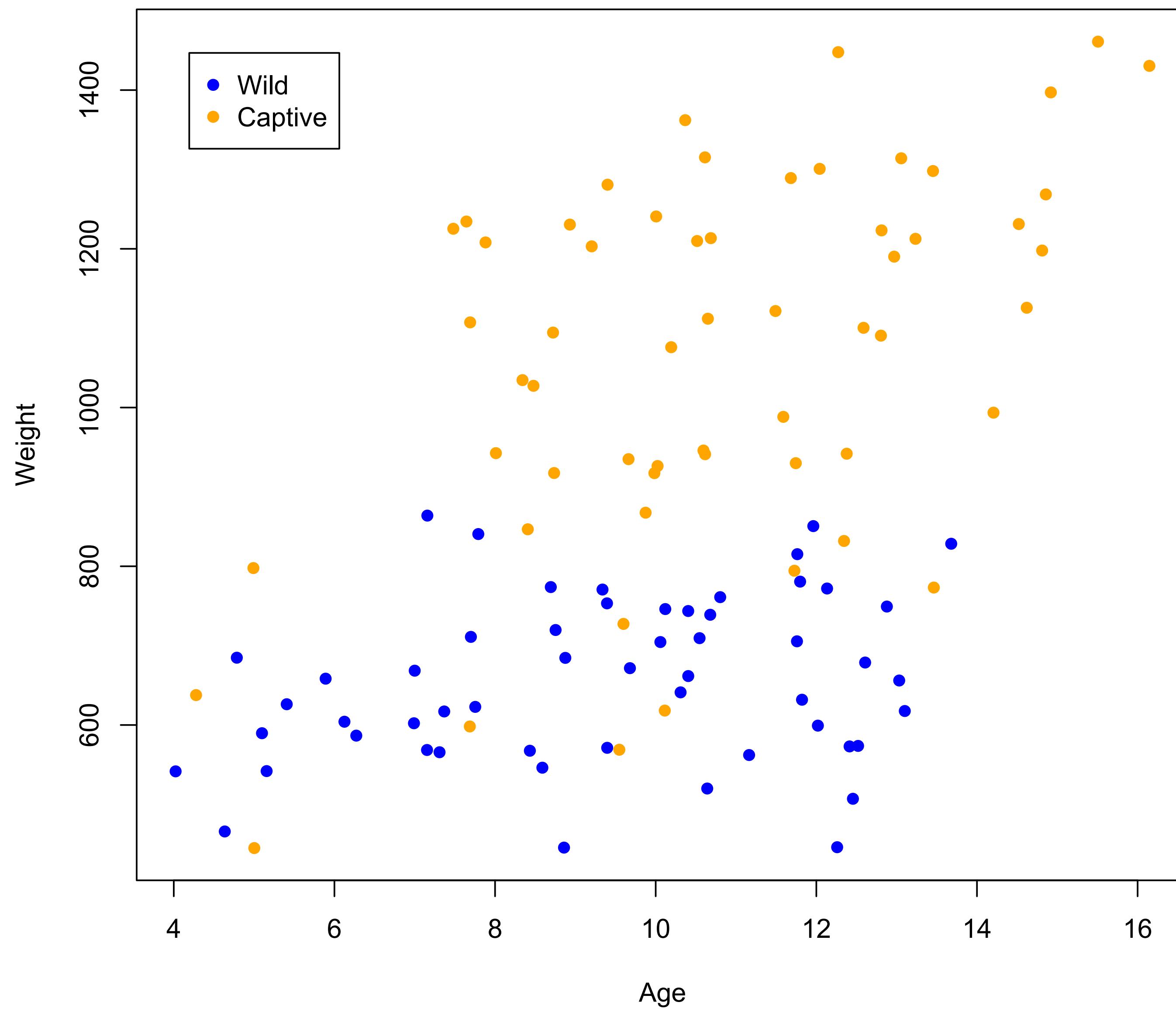
a) Input Space



b) Feature Space



Meerkat Weight vs Age

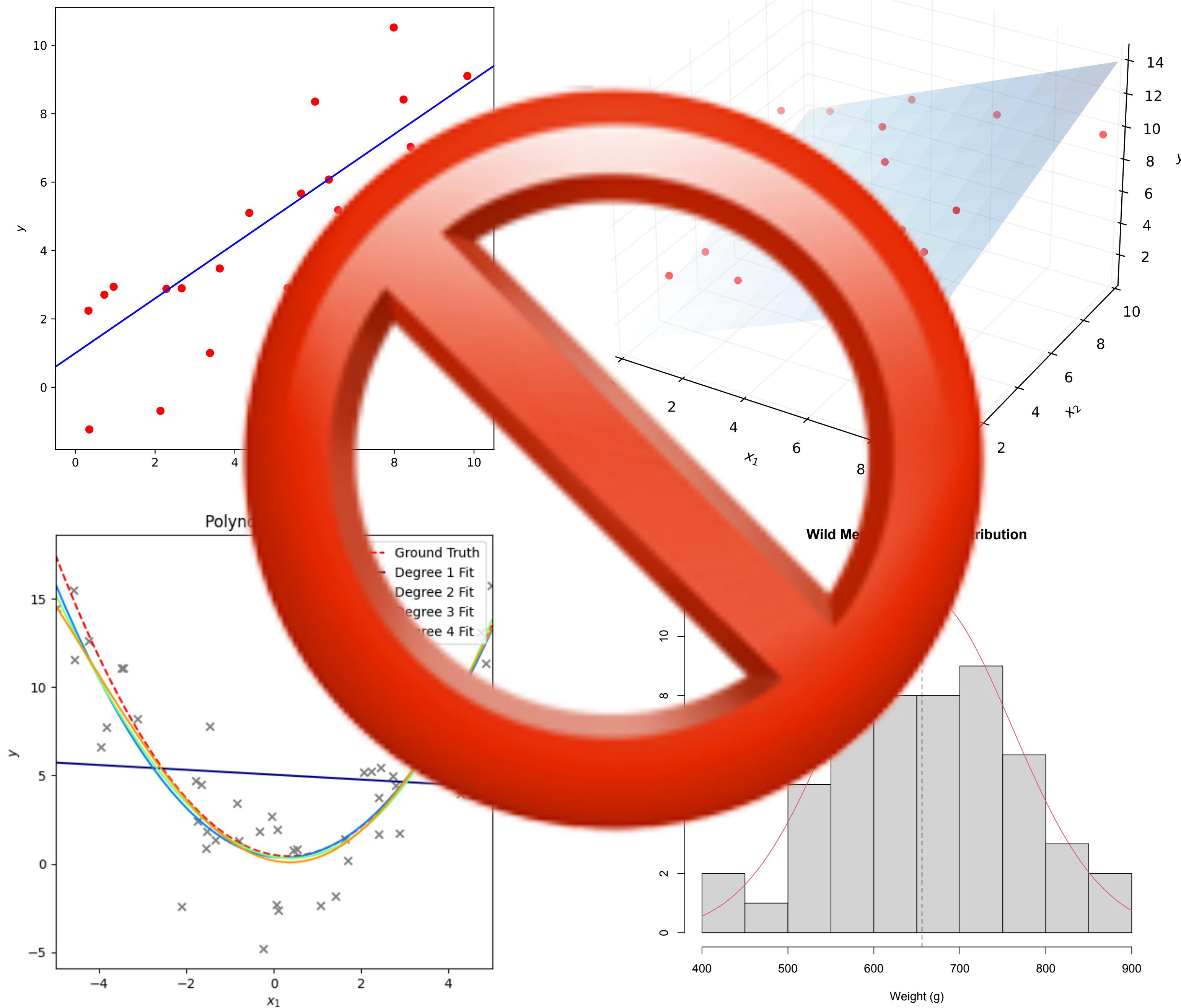




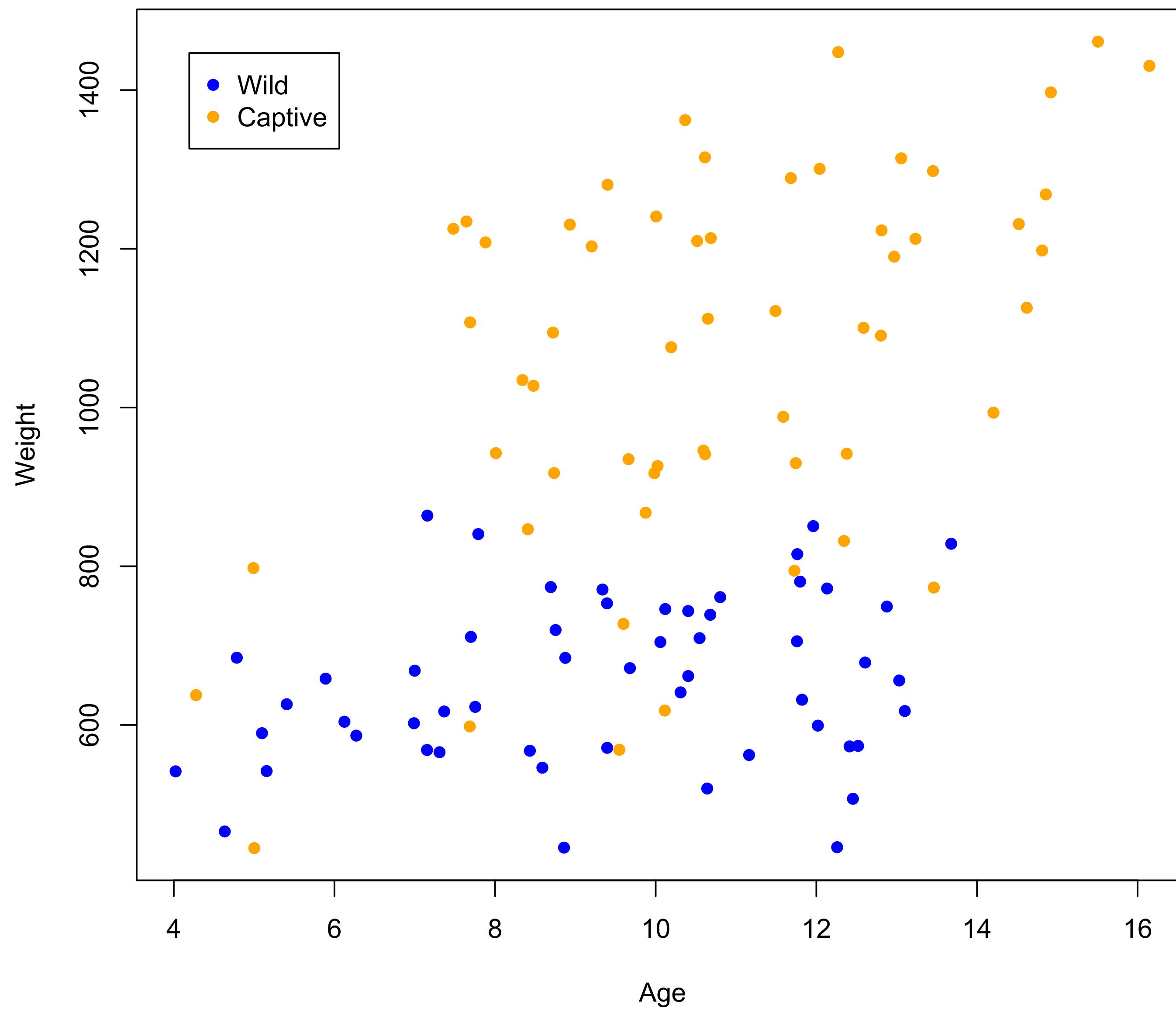
A close-up photograph of two people. On the left, an older man with a beard and mustache, wearing a dark jacket over a light-colored scarf, looks directly at the camera with a serious expression. On the right, a woman with long, dark hair, also in a dark jacket, looks towards the man. They appear to be in an intimate or serious conversation.

from a certain point of view

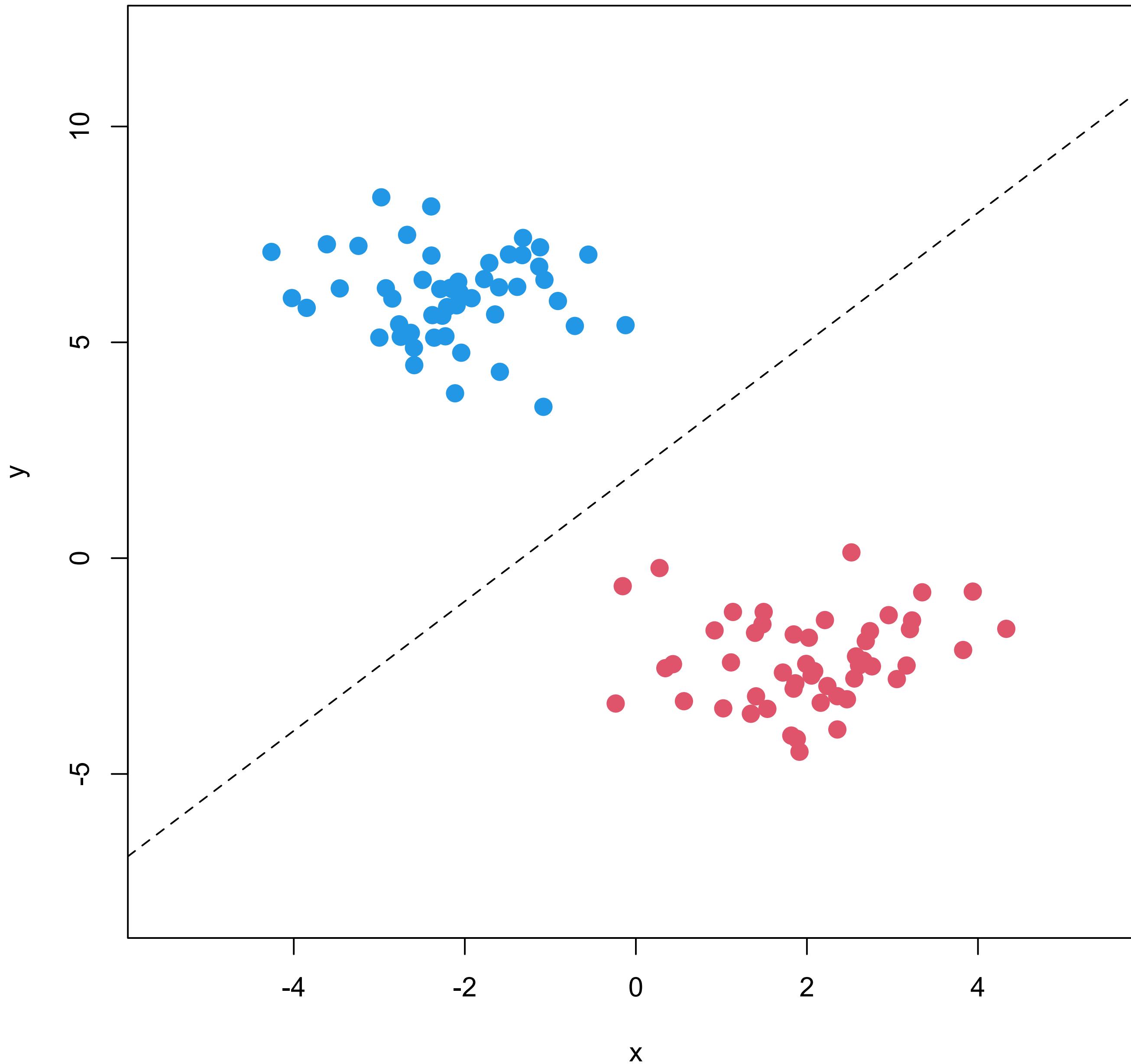
$$L(\mathbf{X}, \mathbf{Y}, \theta) = ??$$



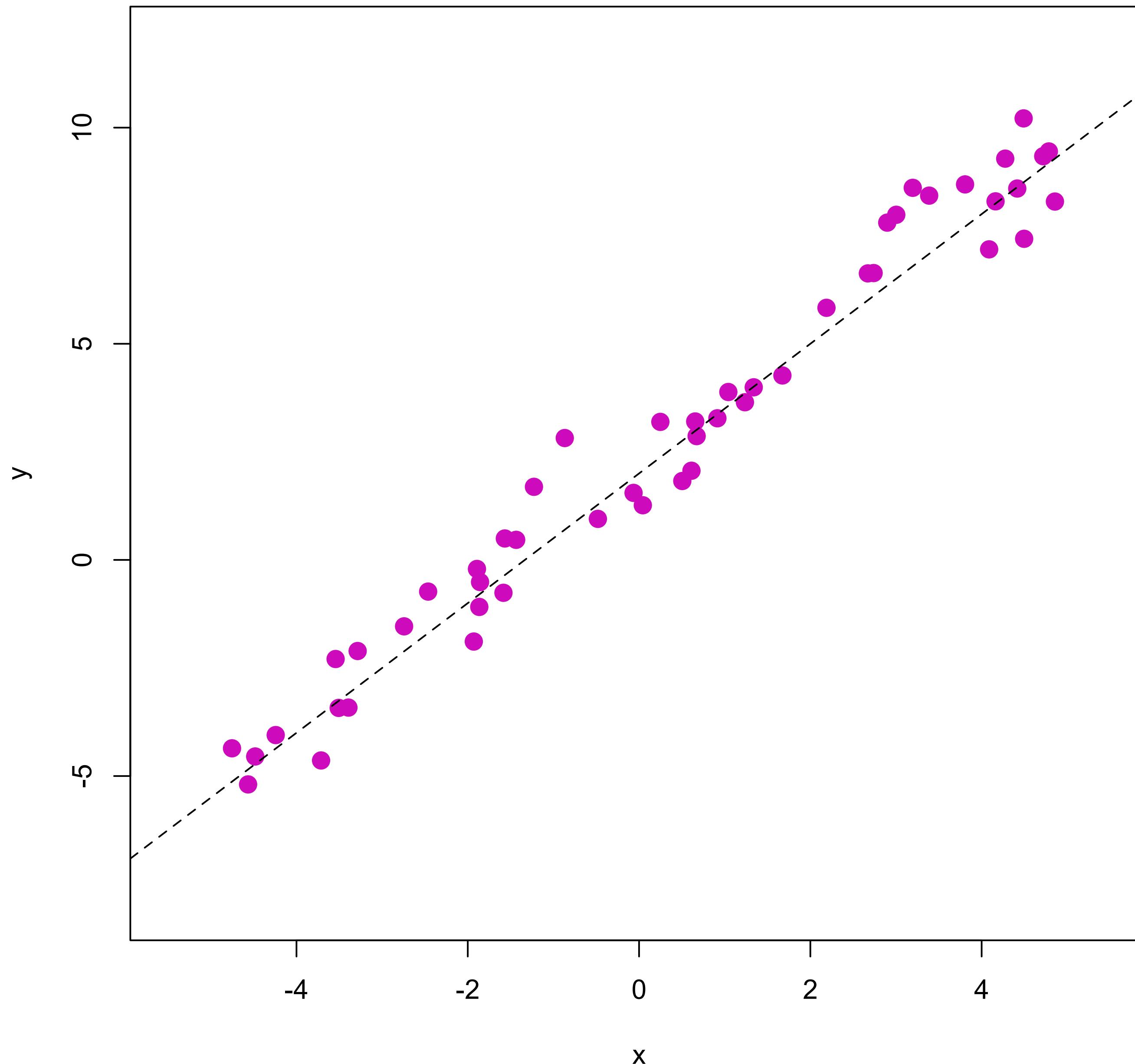
Meerkat Weight vs Age



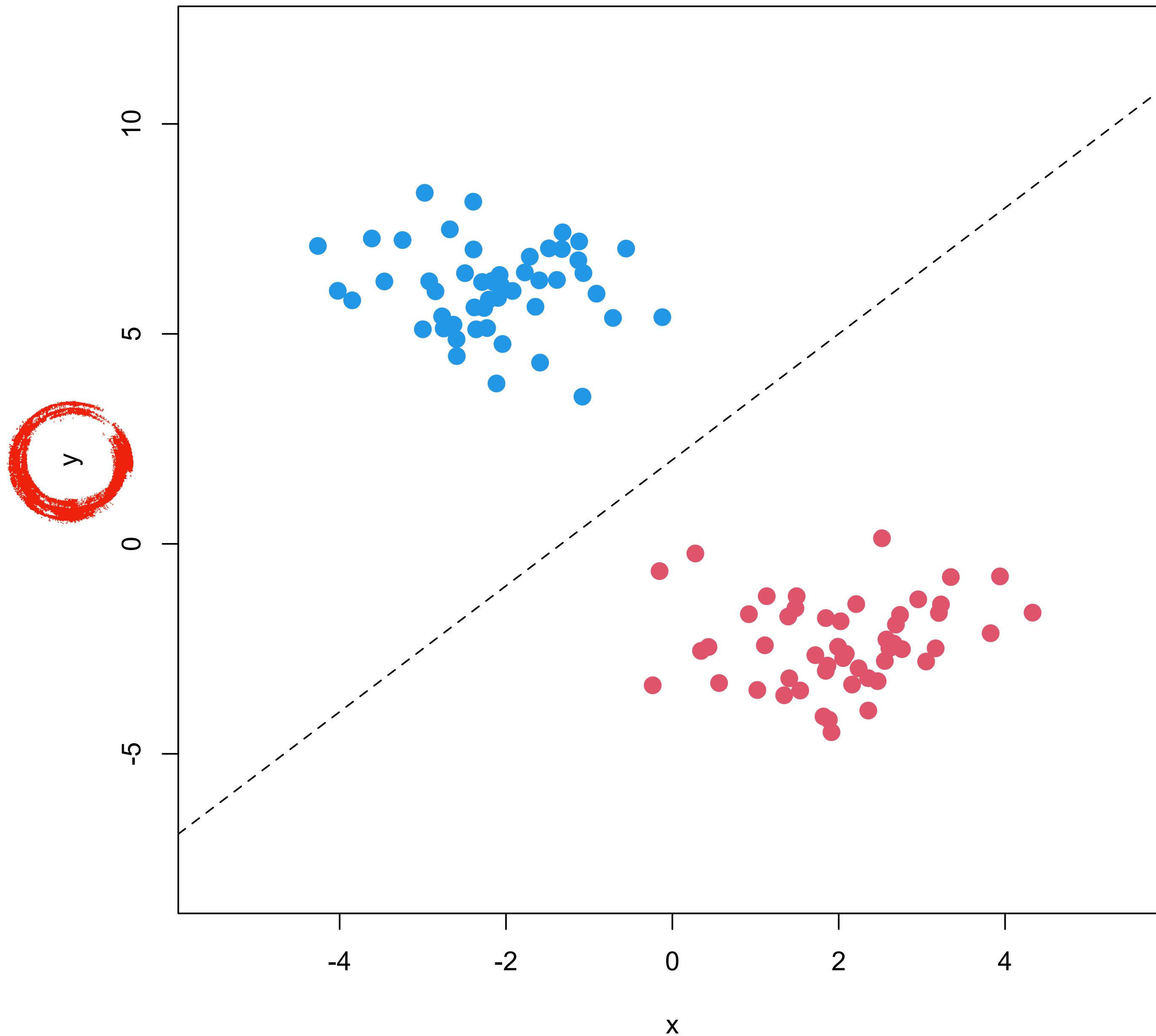
Classification

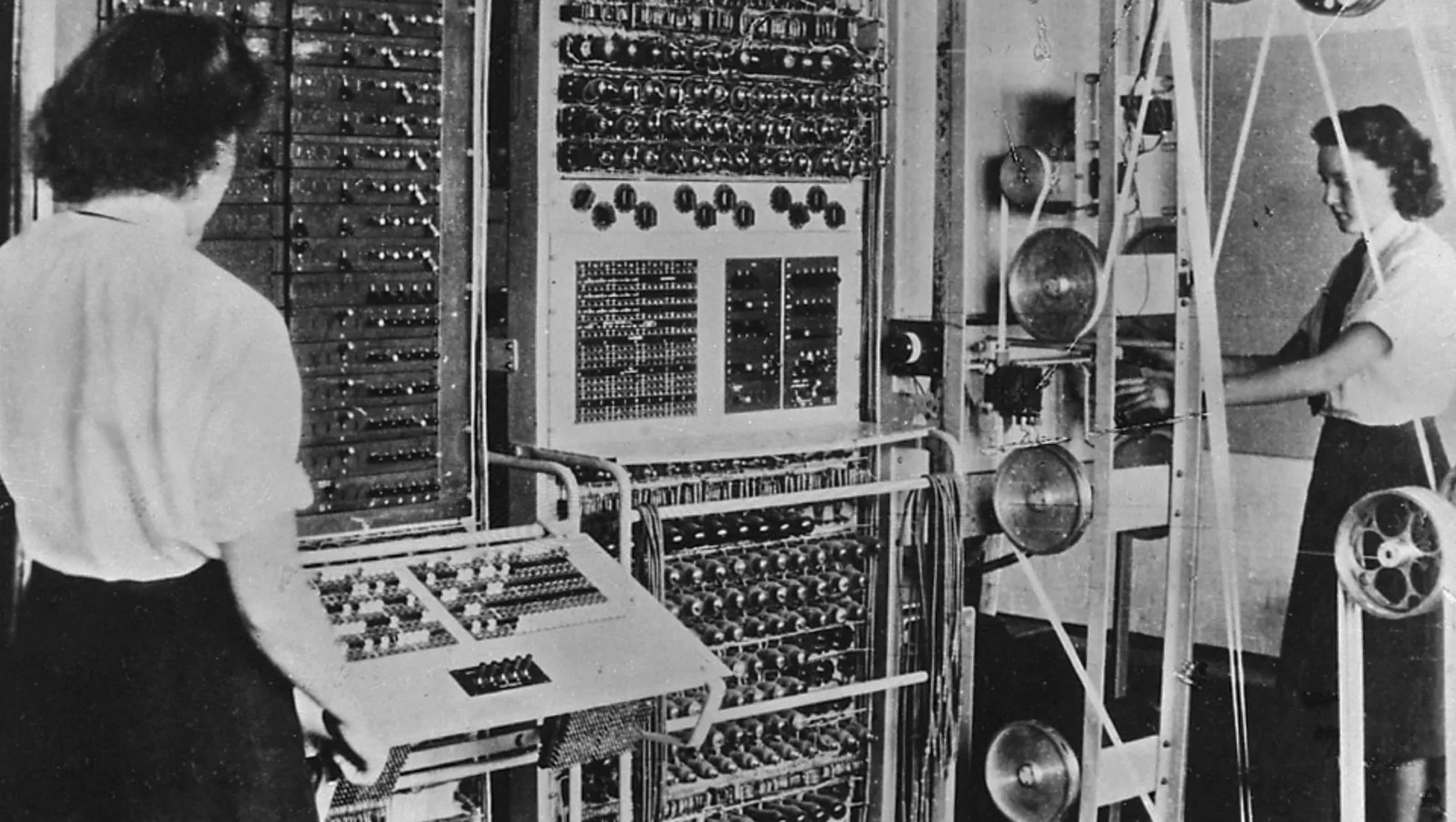


Regression

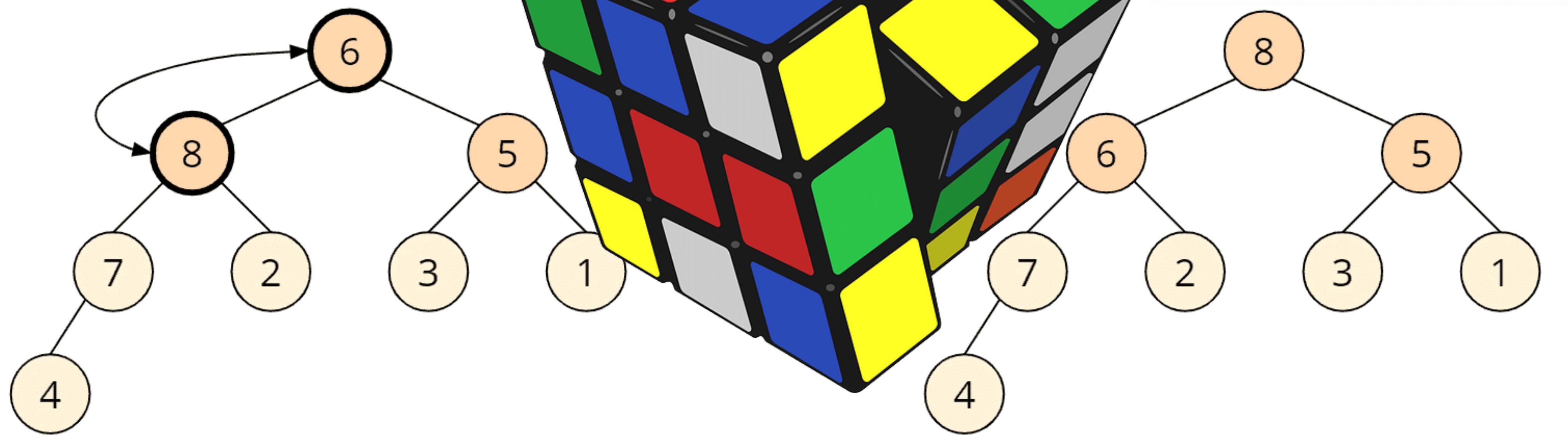
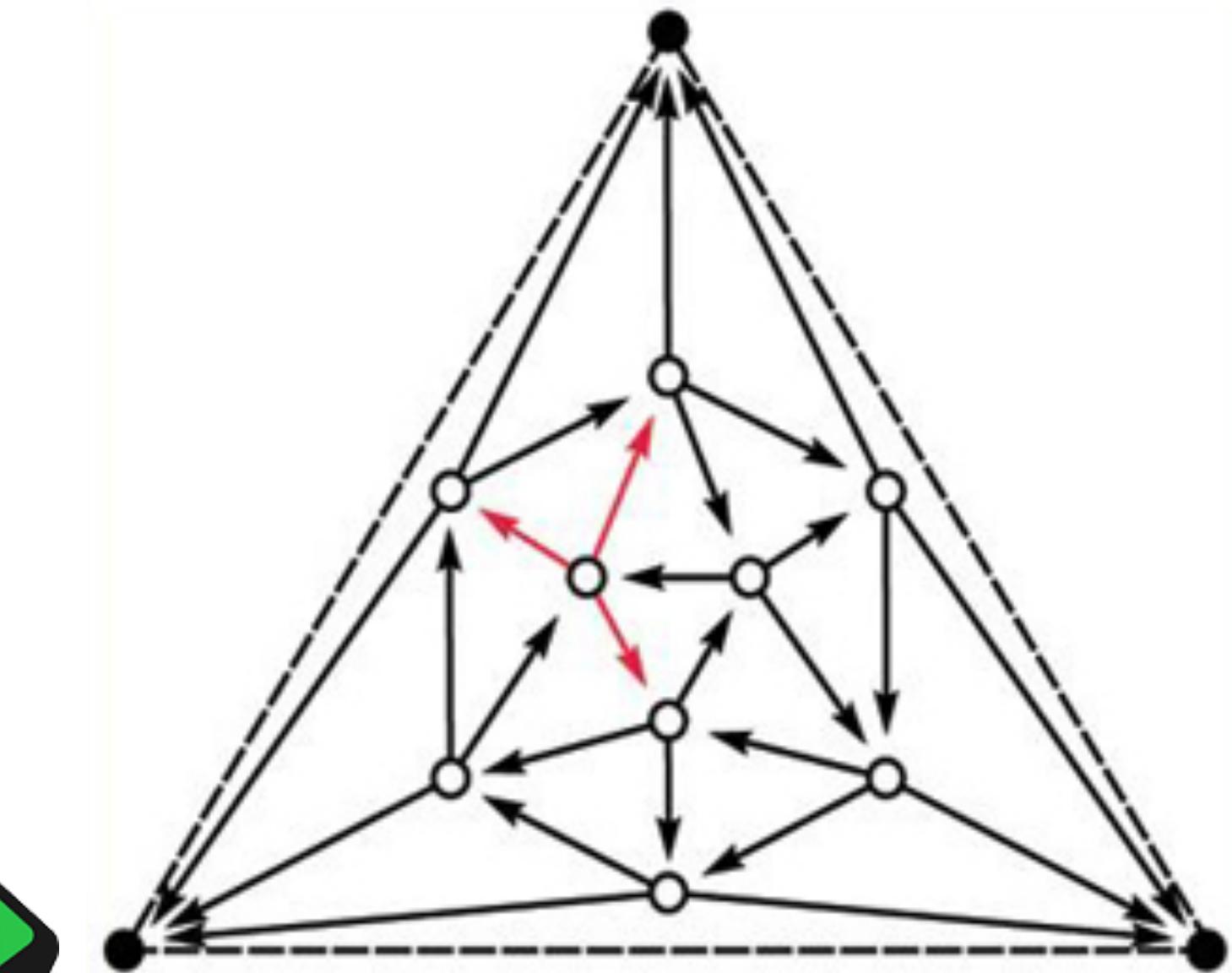
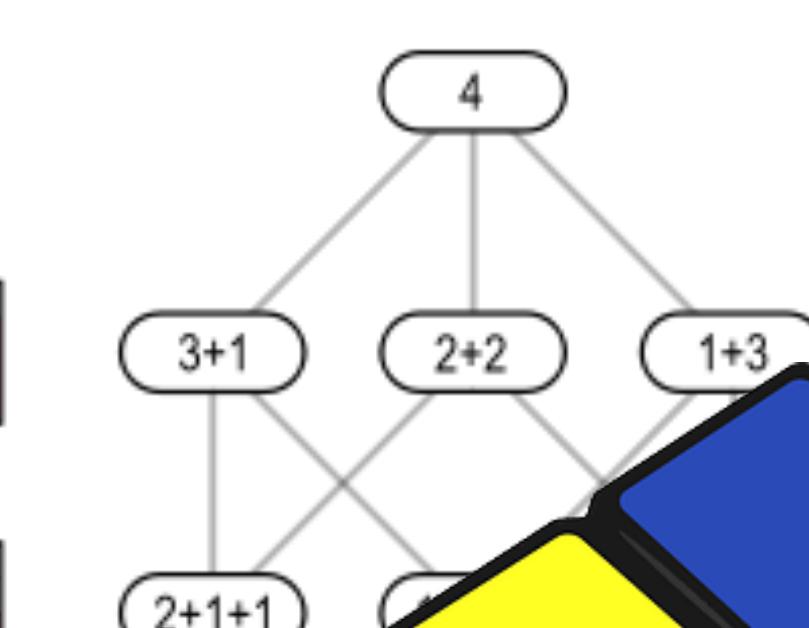
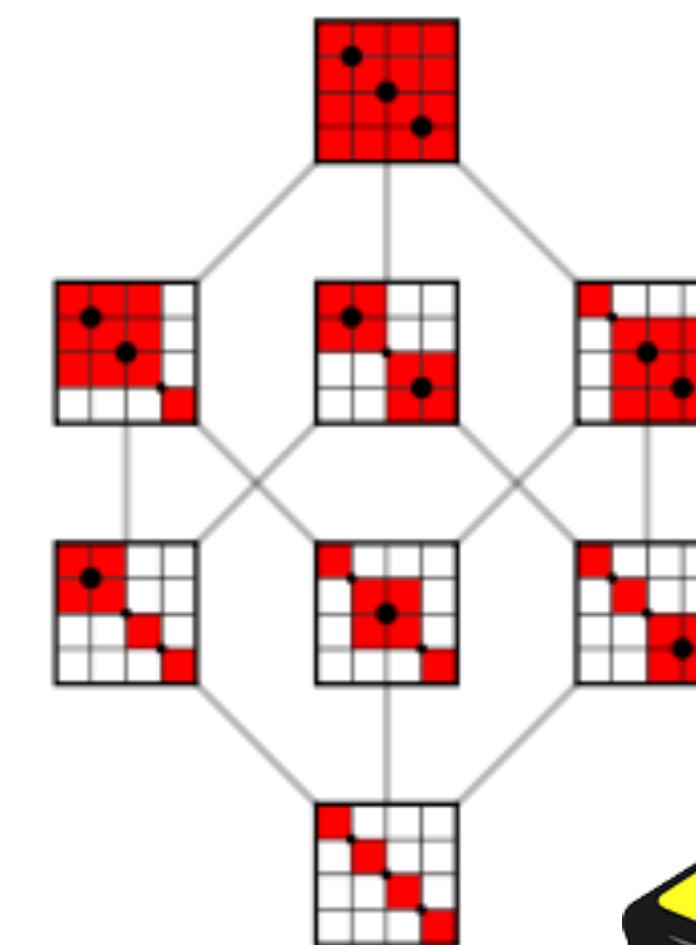
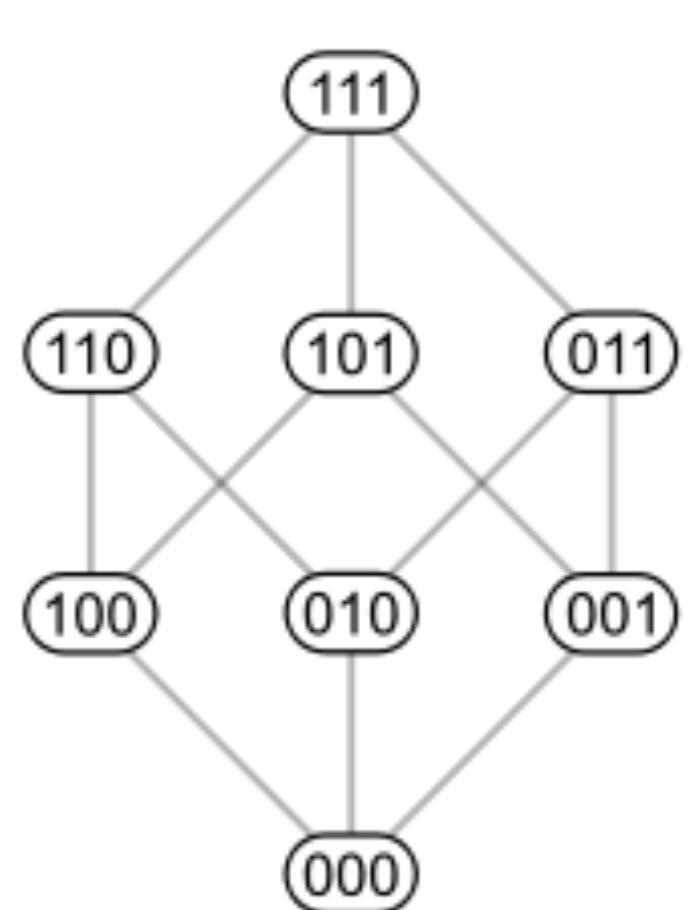


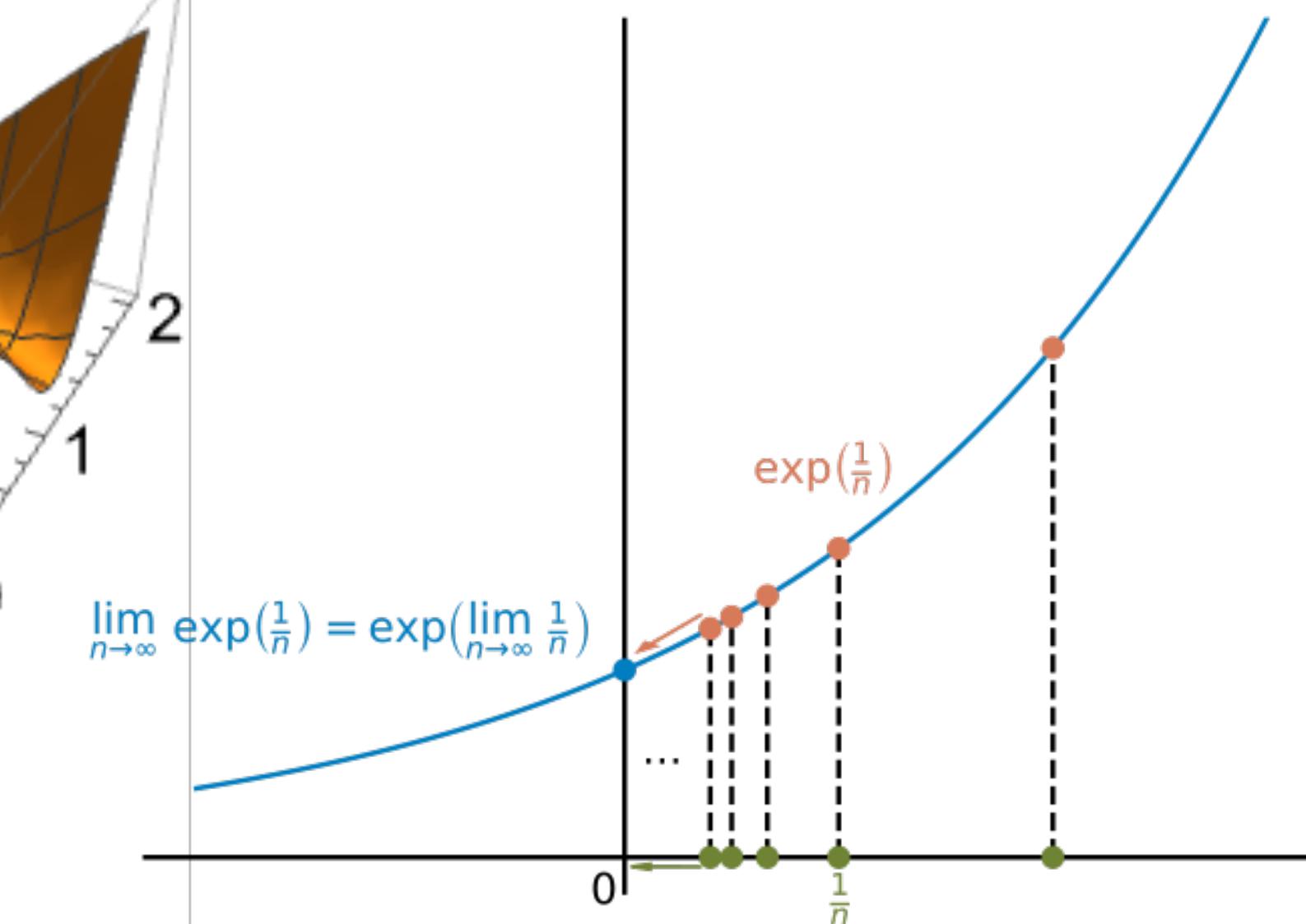
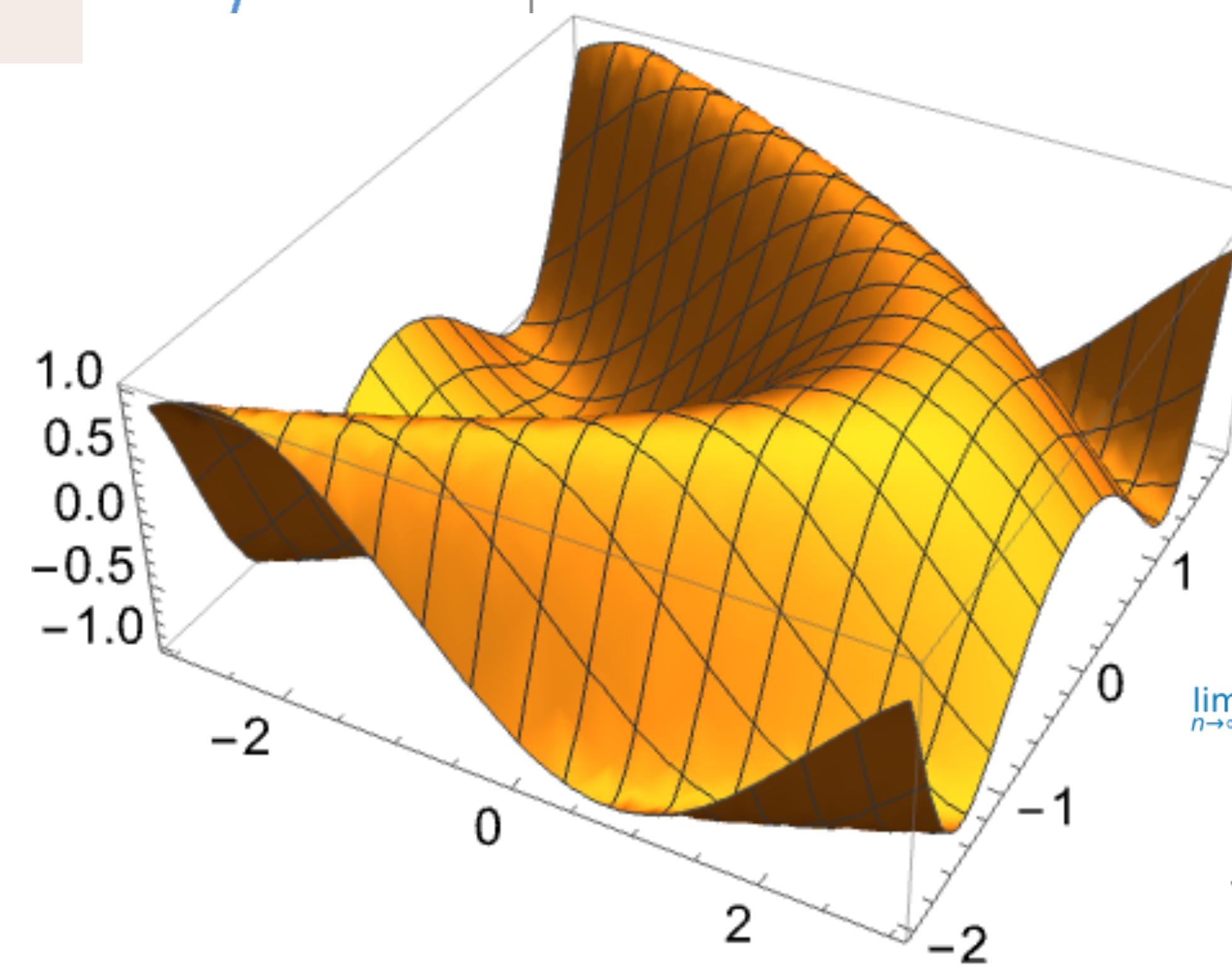
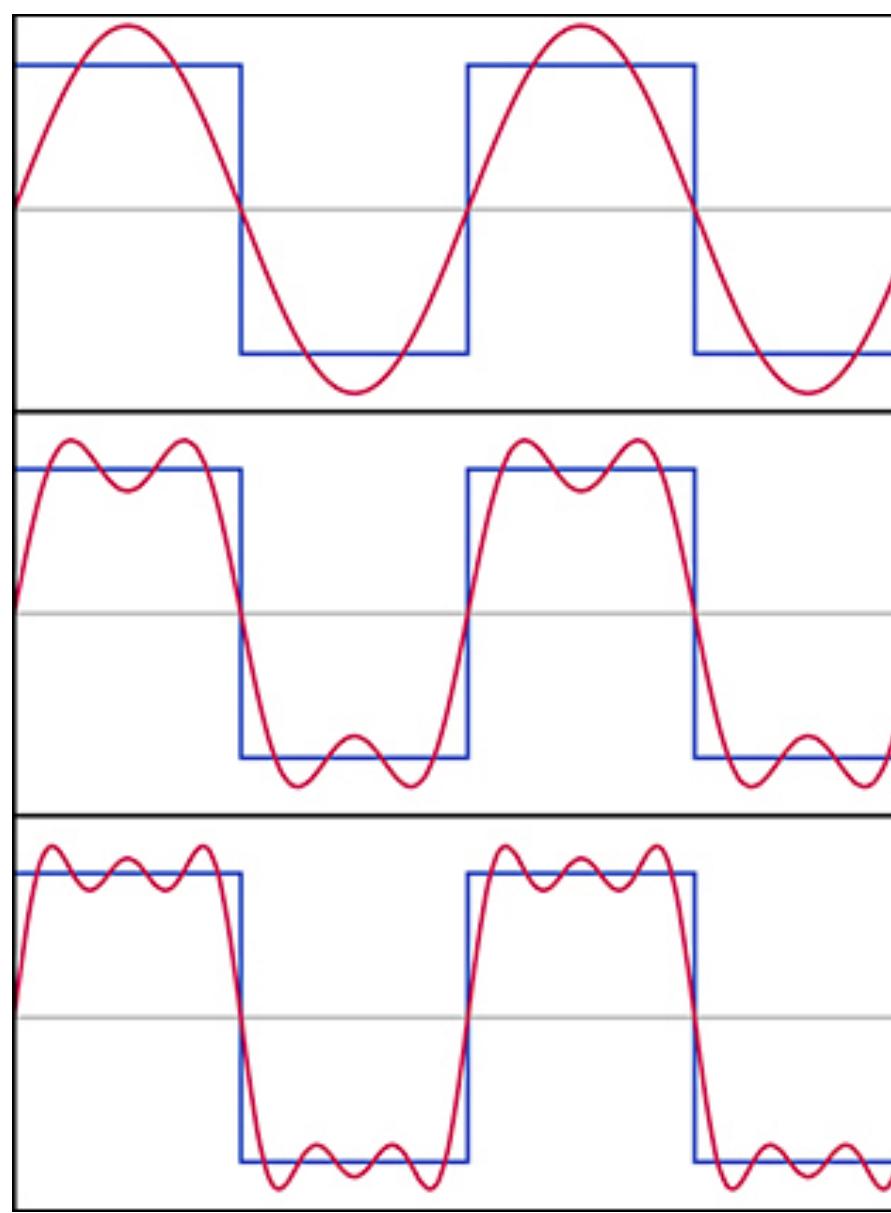
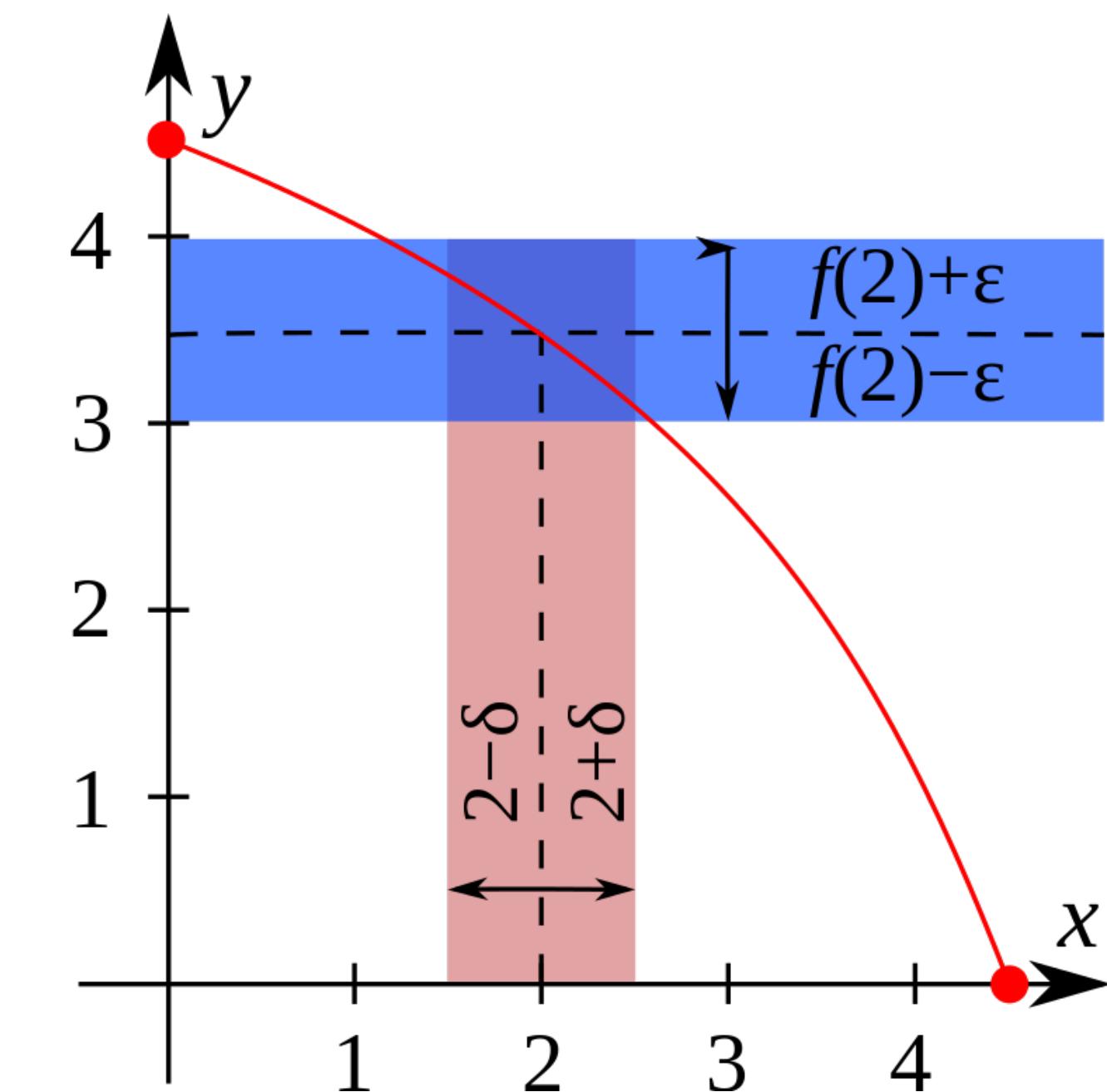
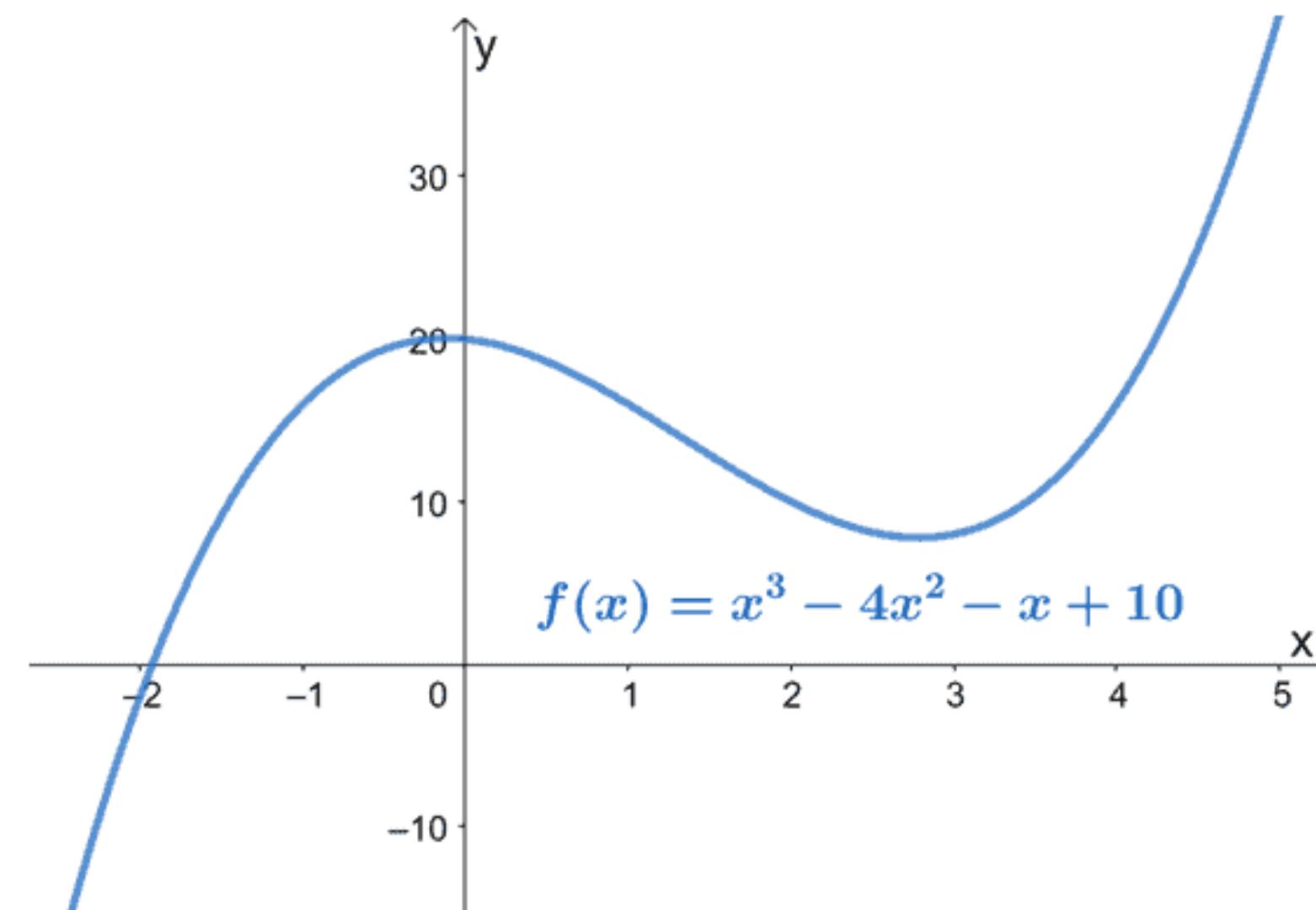
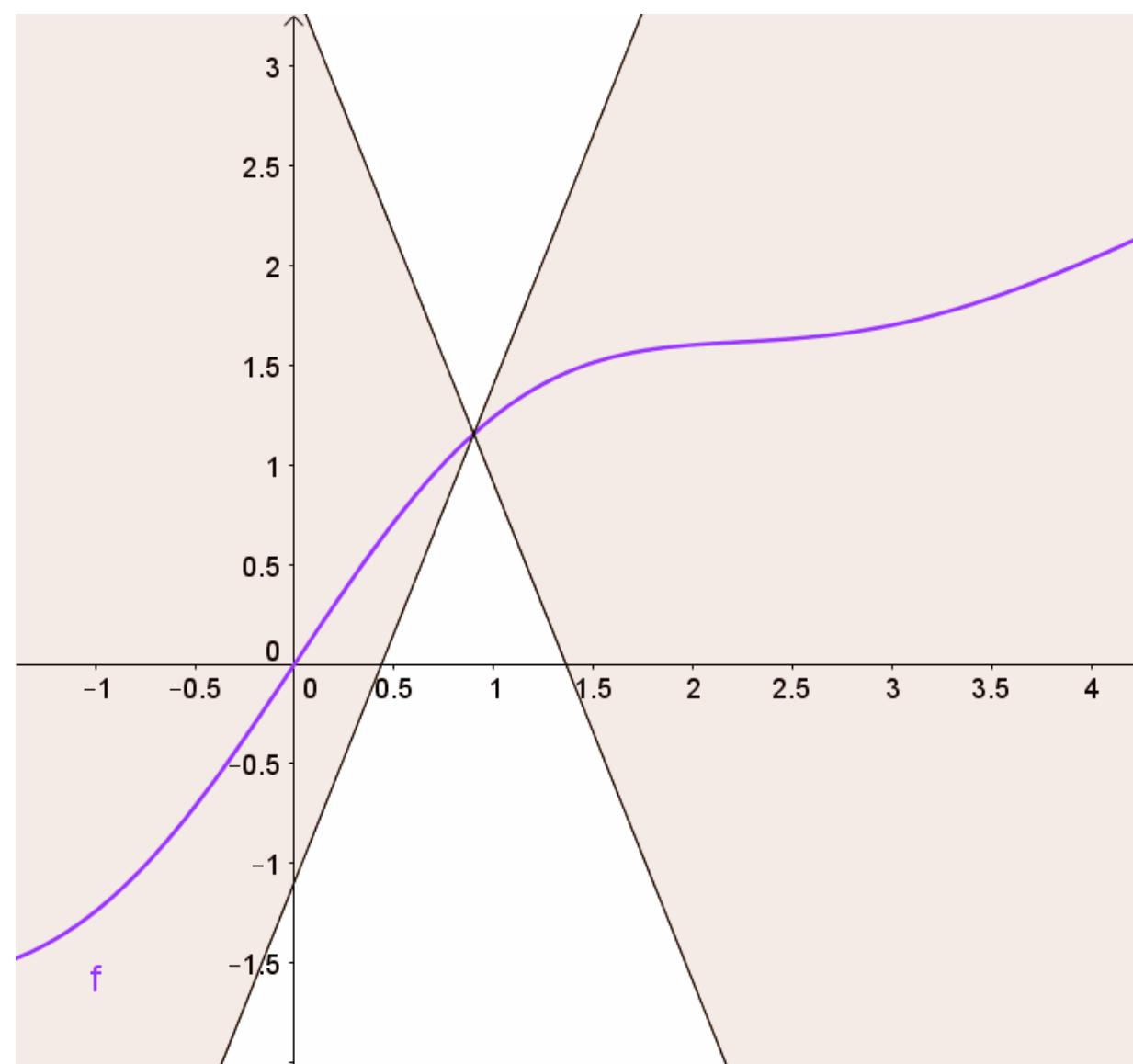
Classification



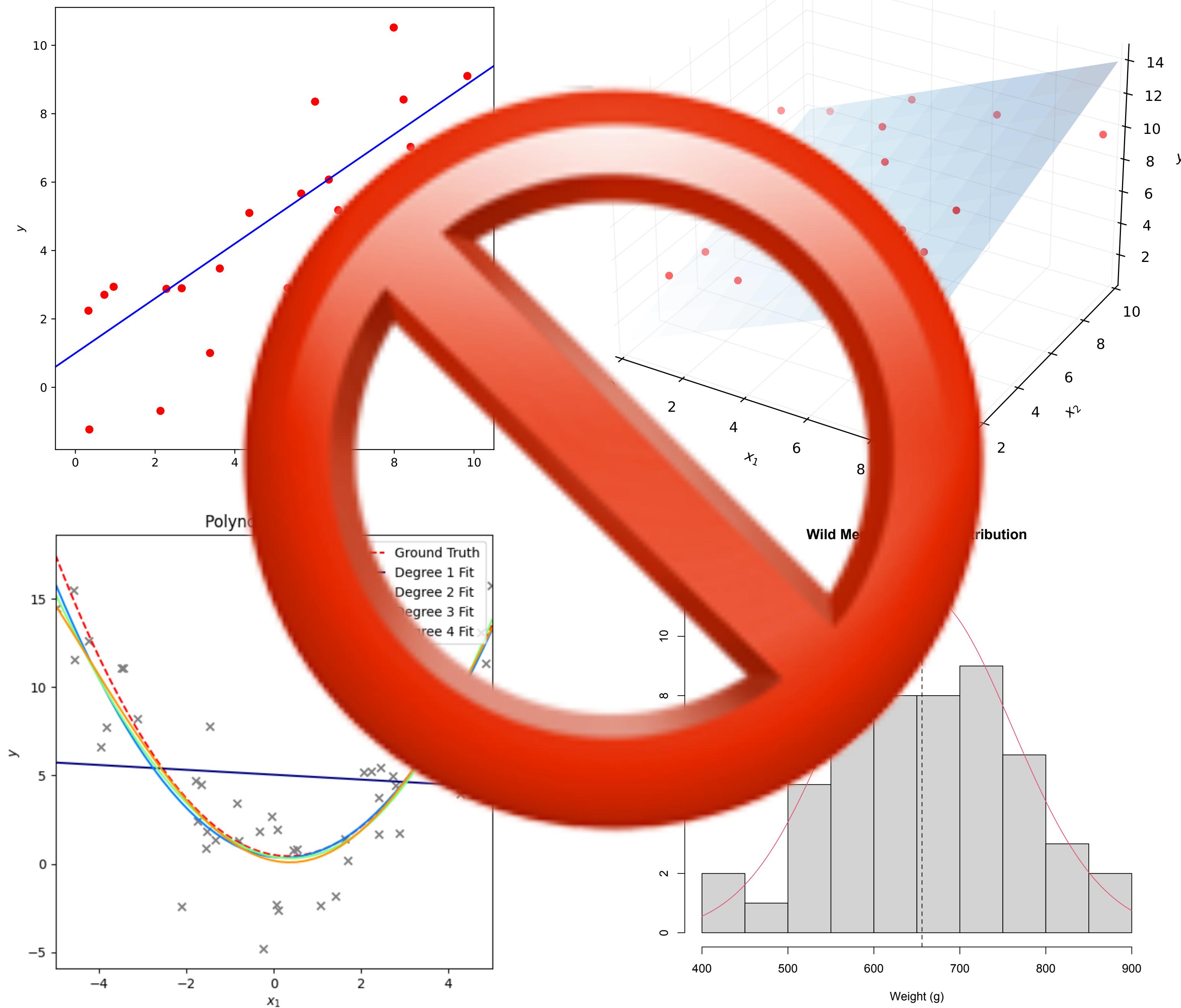


00000110011001100110000110
0100000000000000100000000000
0000110011001100110000000000
0011001100110011001000000000
000000110011001100110011001100
11001000000000000000000011001100
100111001000000000000000000000001



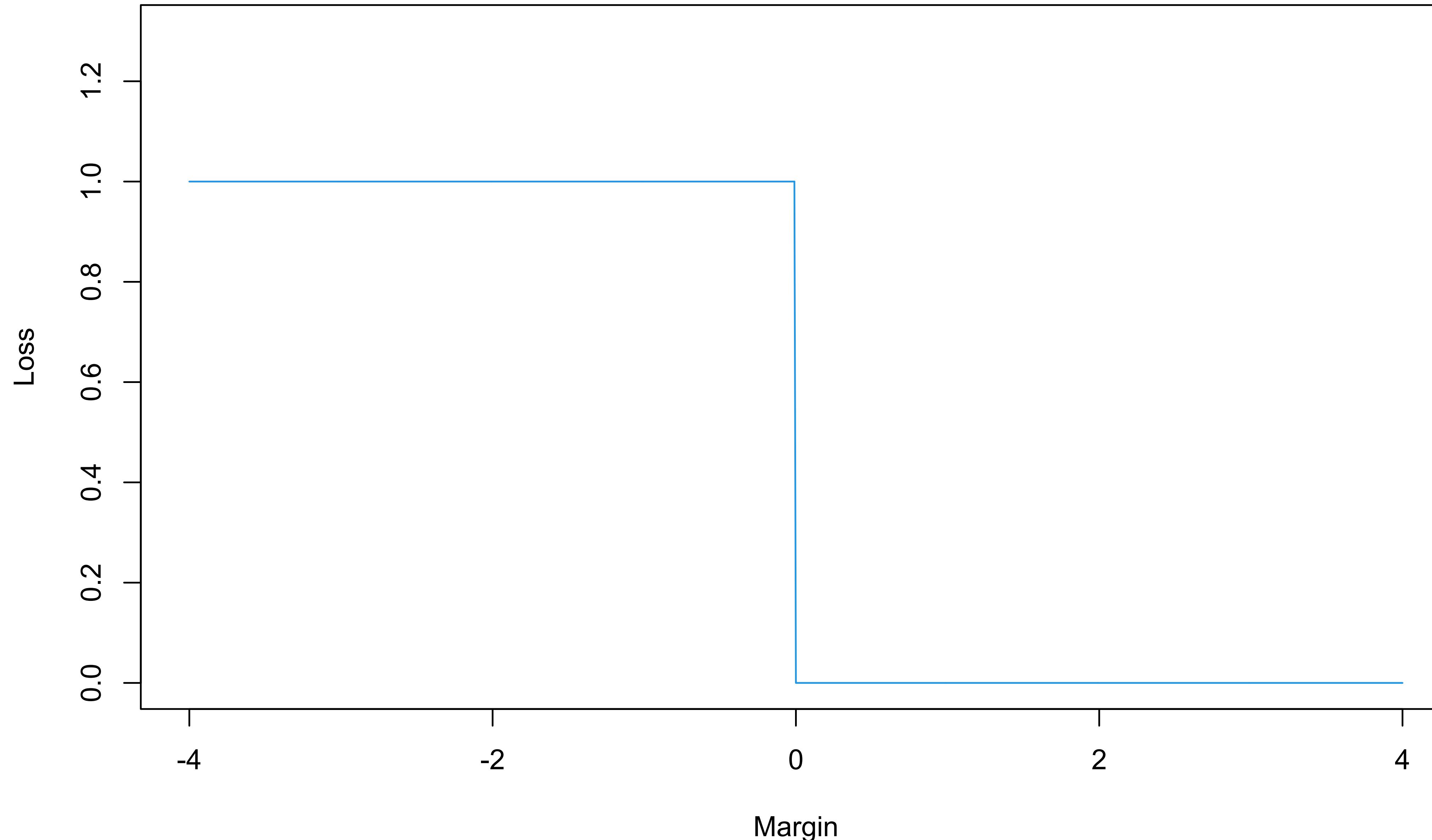


$$\frac{\partial x}{\partial y} = WTAF?$$

$$L(\mathbf{X}, \mathbf{Y}, \theta) = ??$$

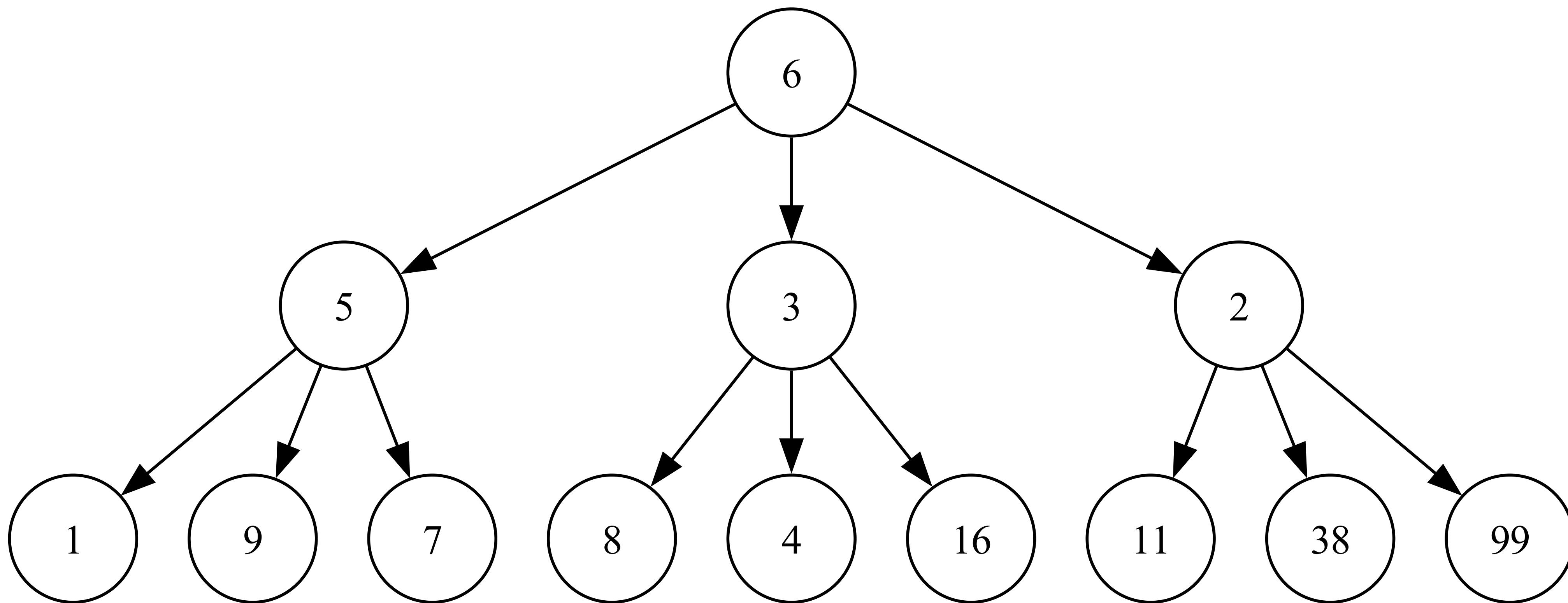
0-1 Loss

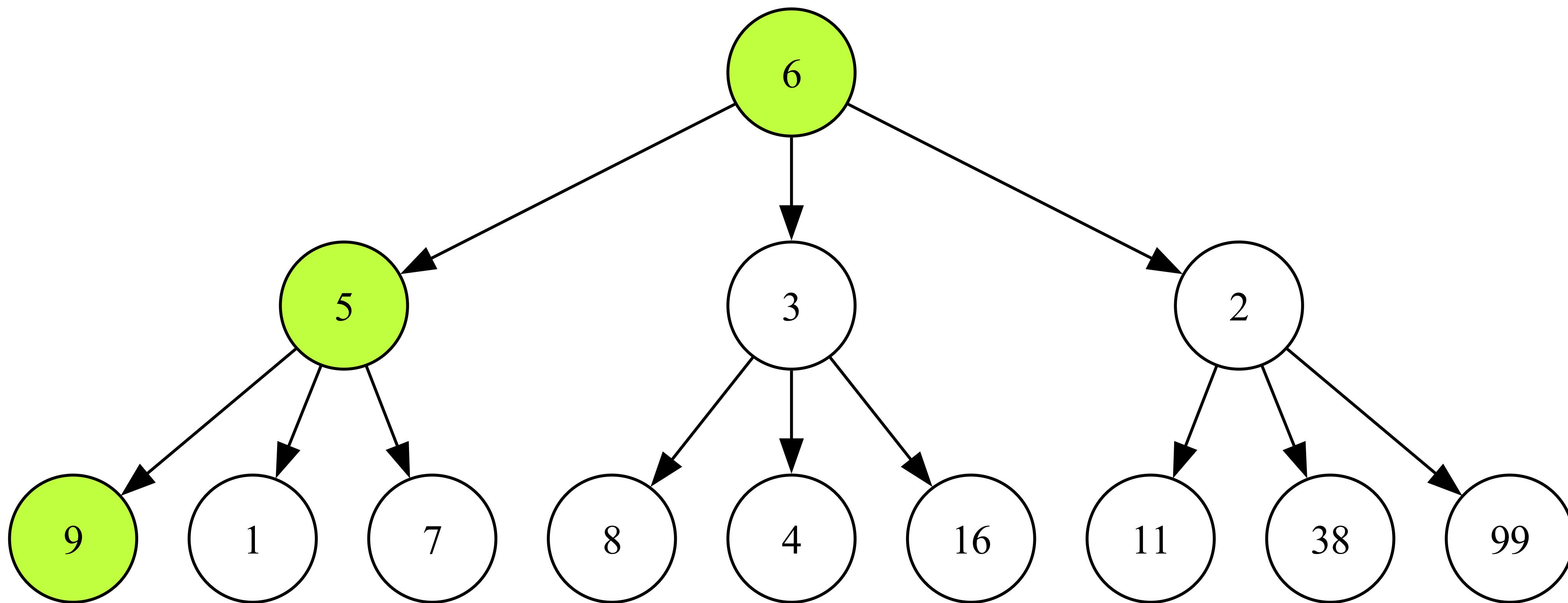


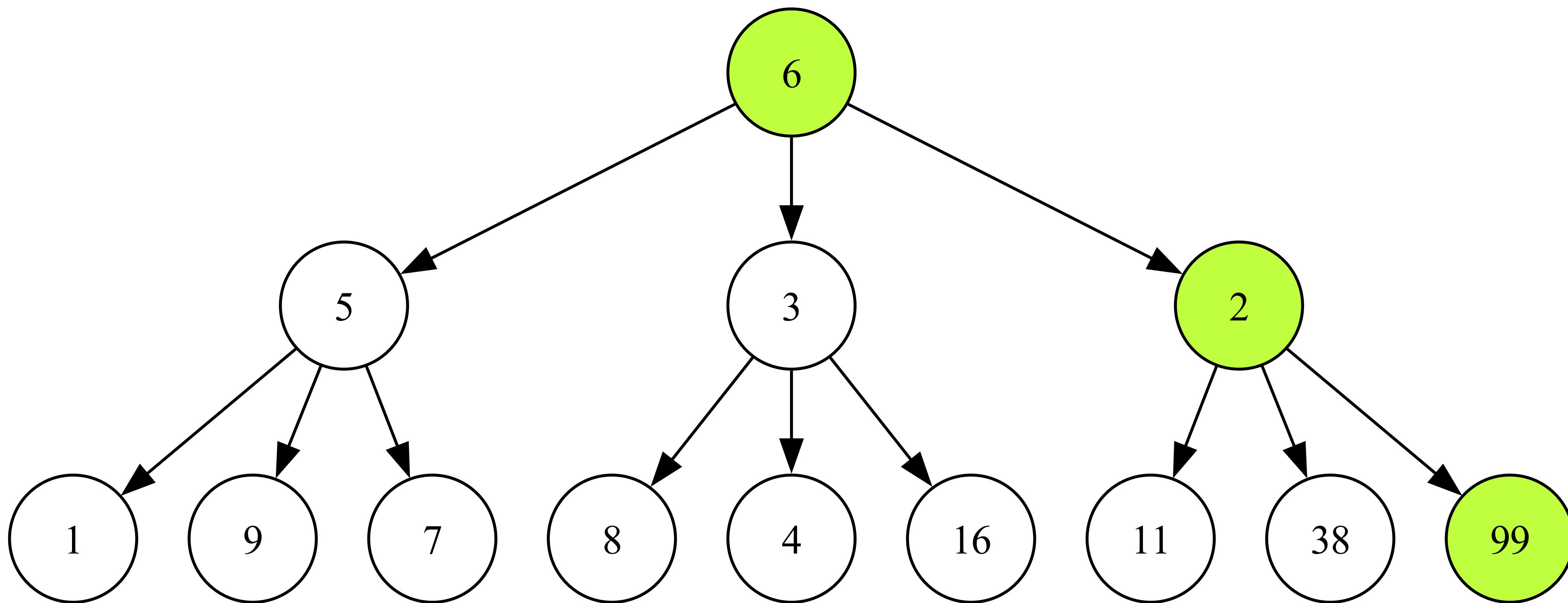
$$\nabla_{\theta} L(\mathbf{y}, \theta)$$

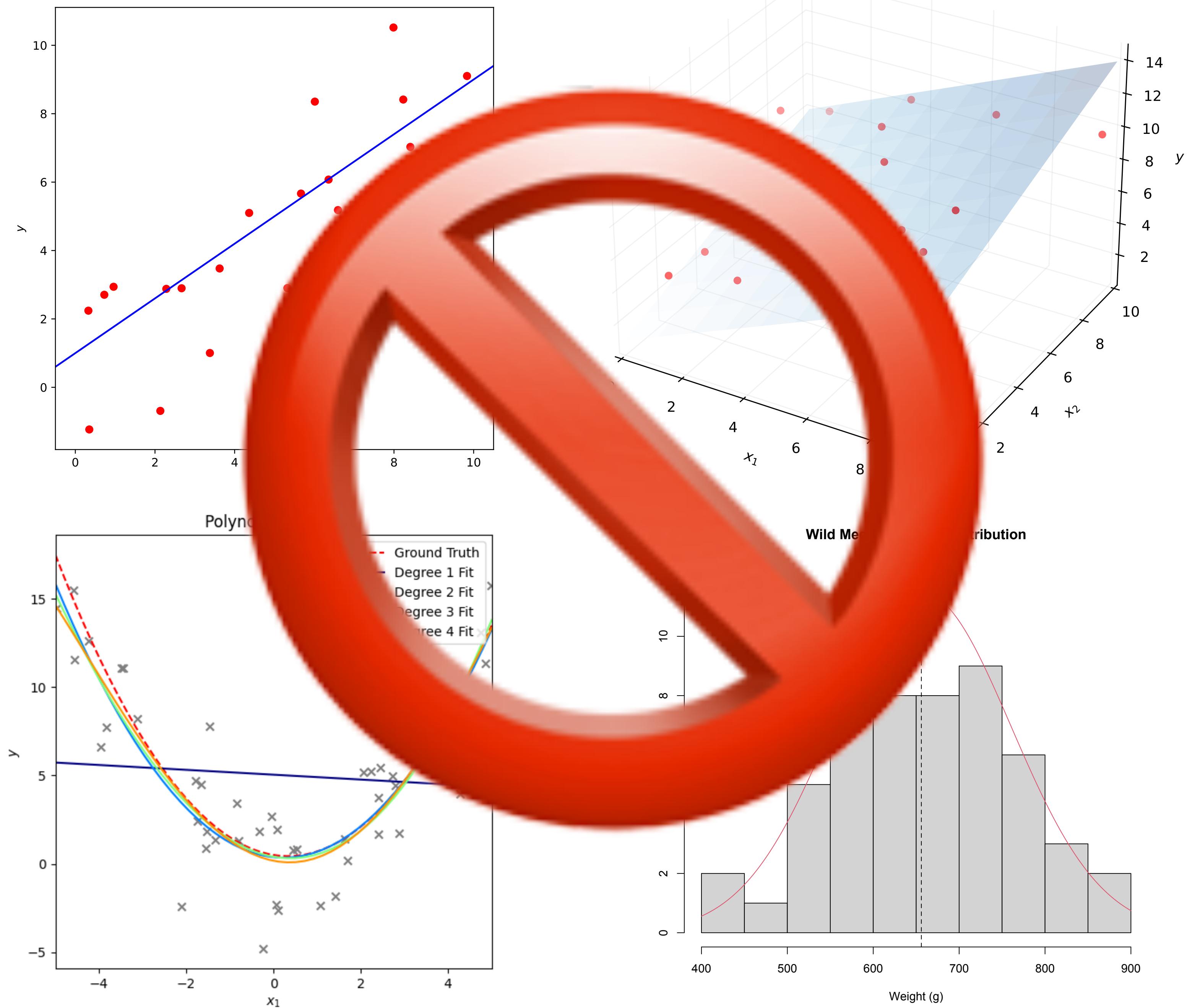



greed is good



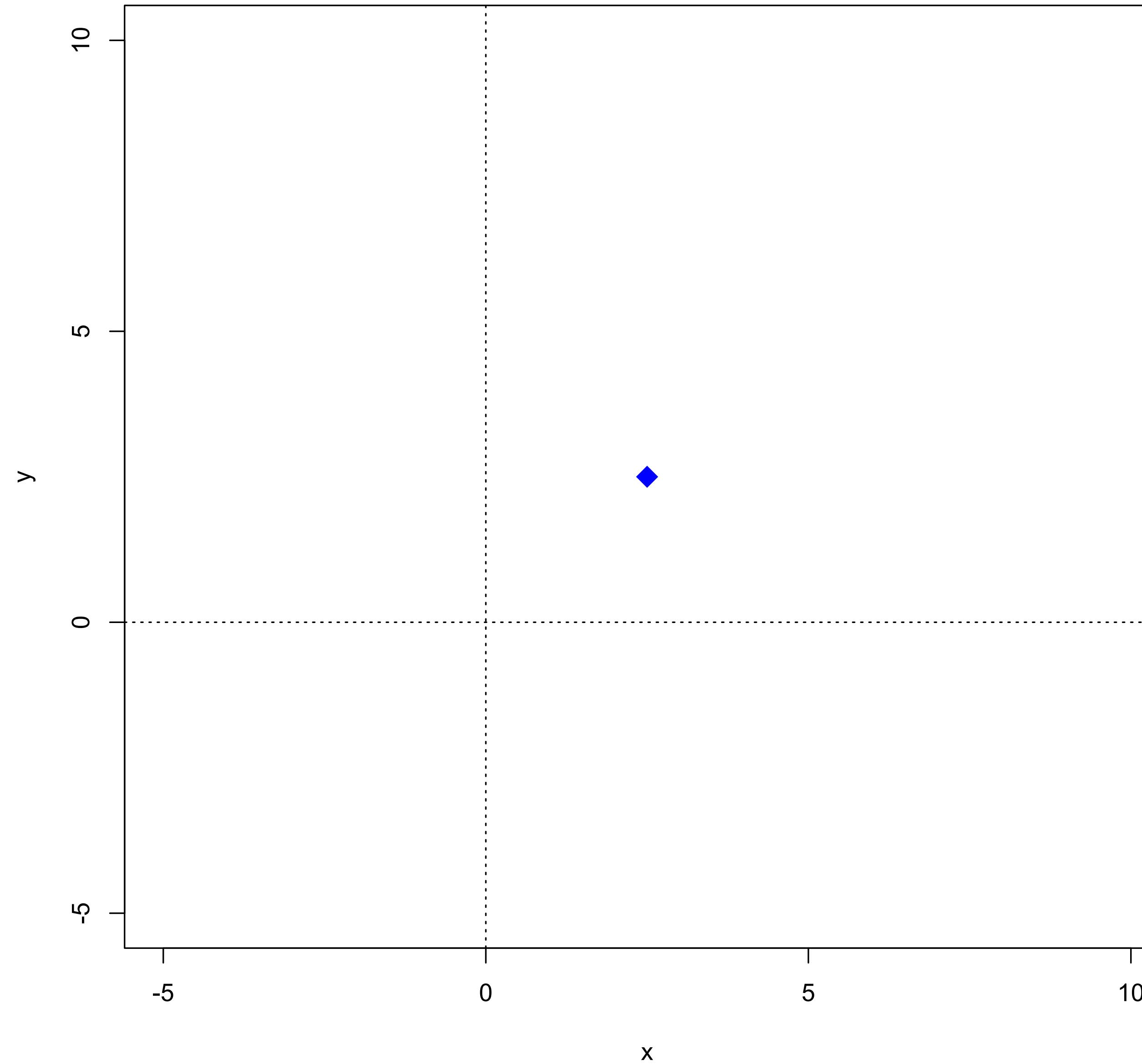




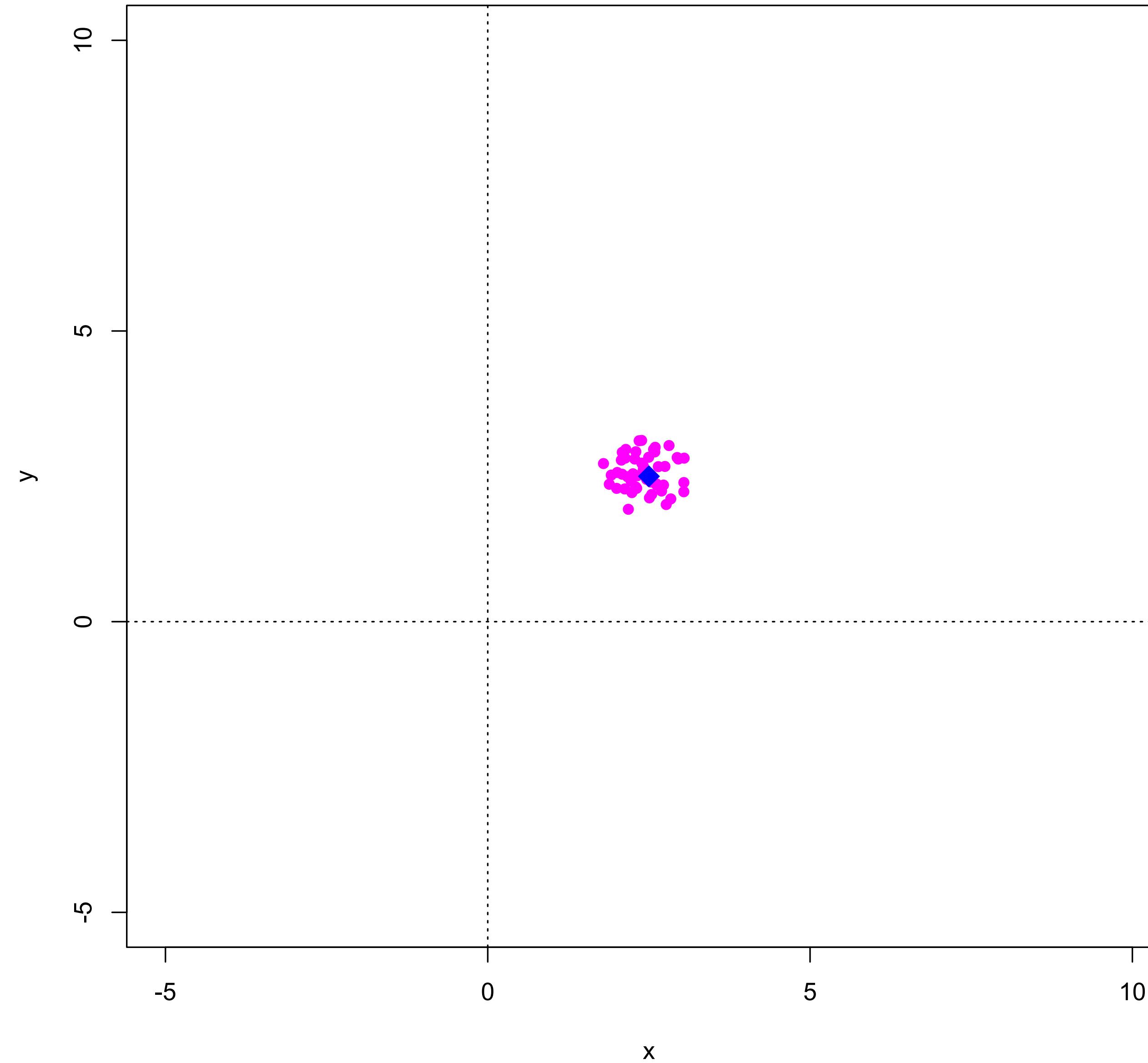


bias-variance tradeoff

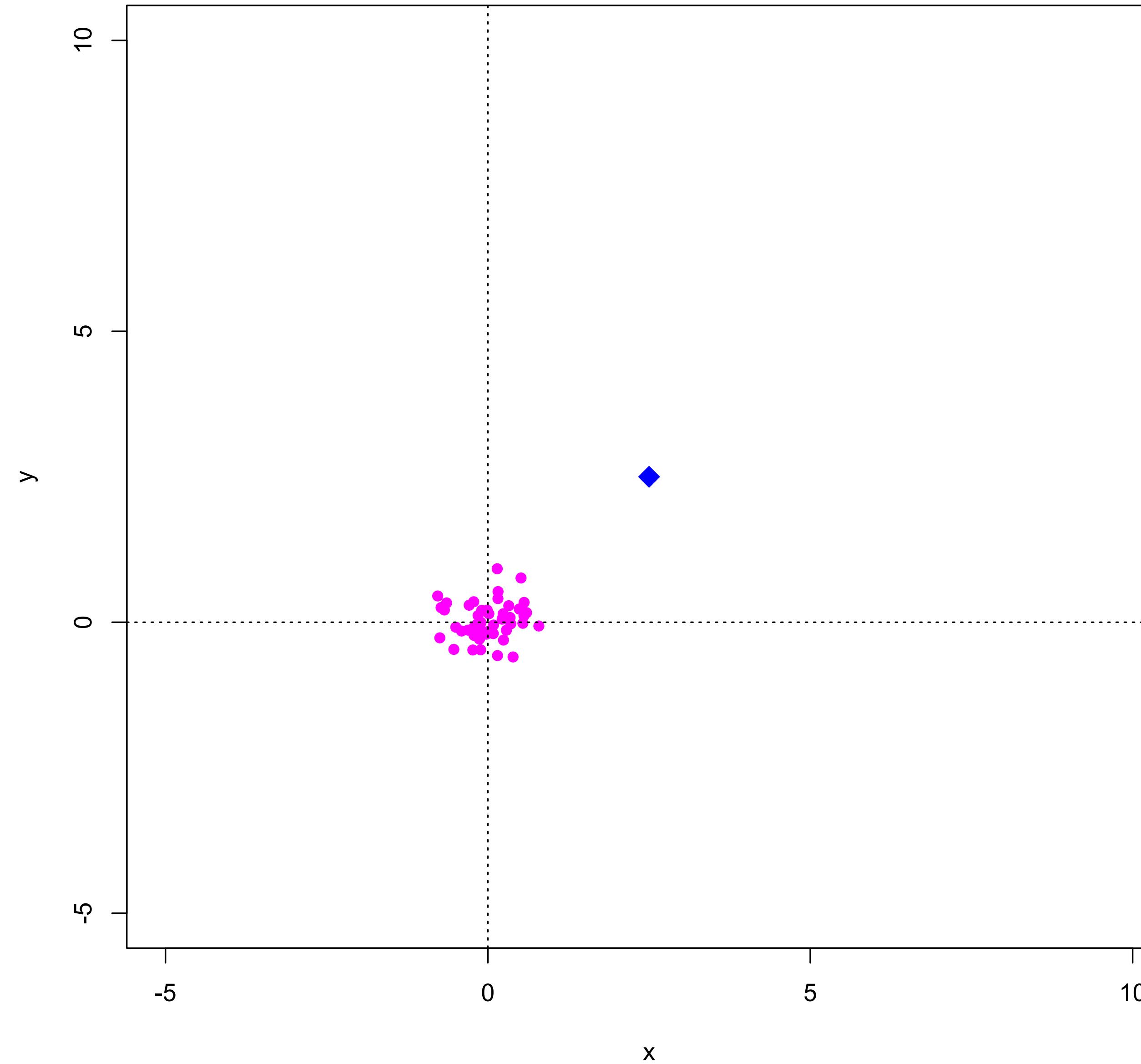
Ground Truth



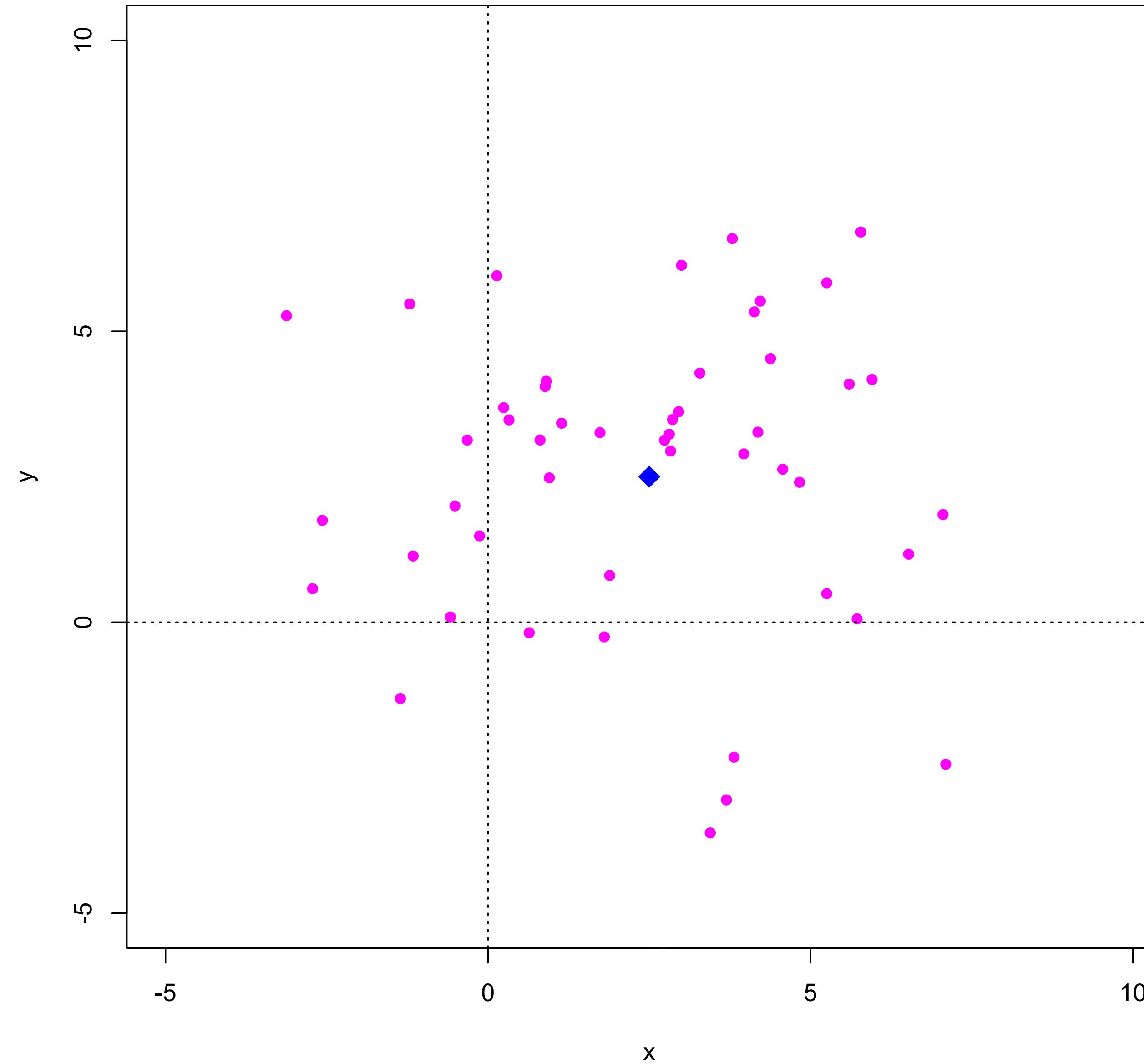
Low Variance, Low Bias



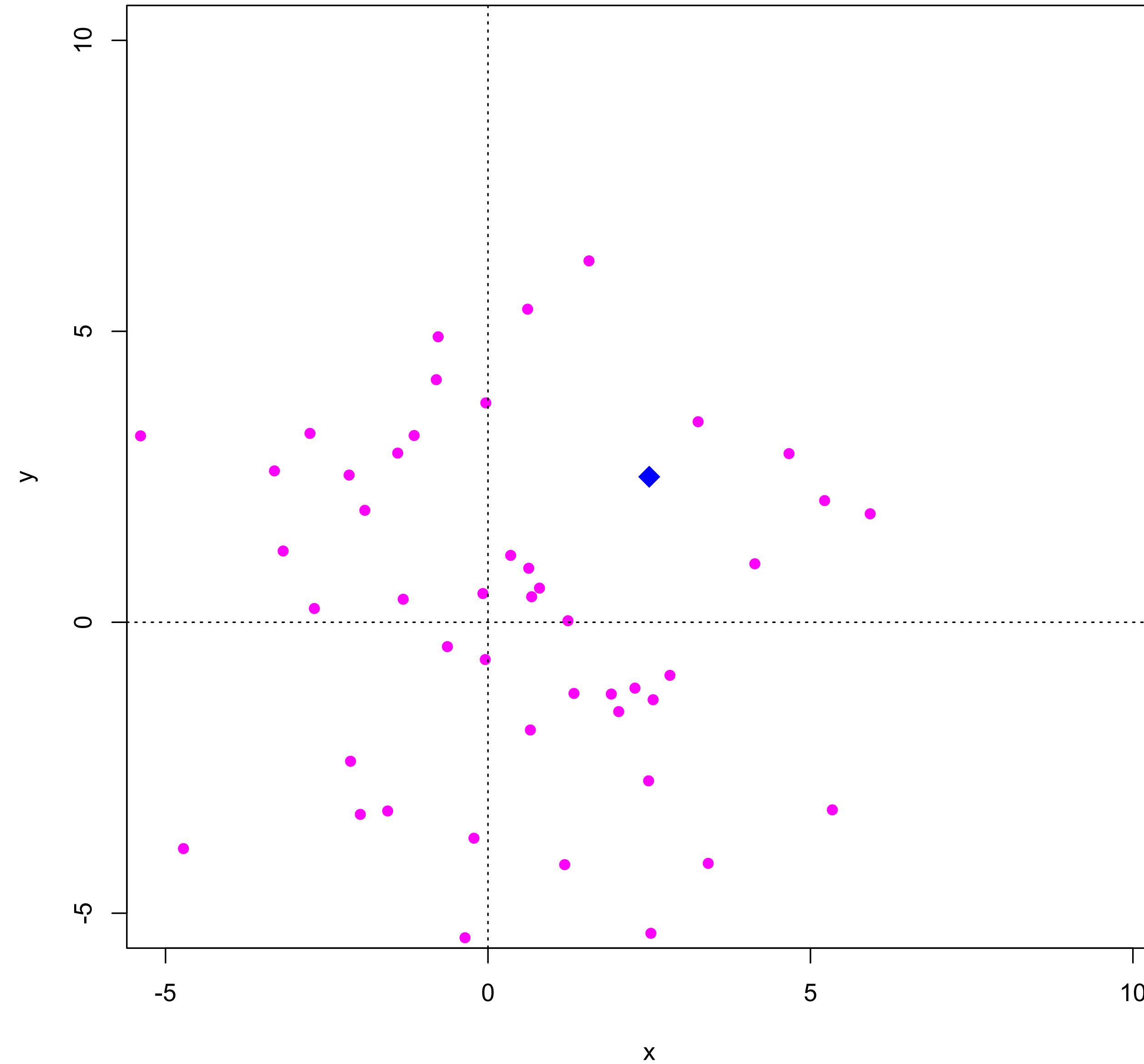
Low Variance, High Bias



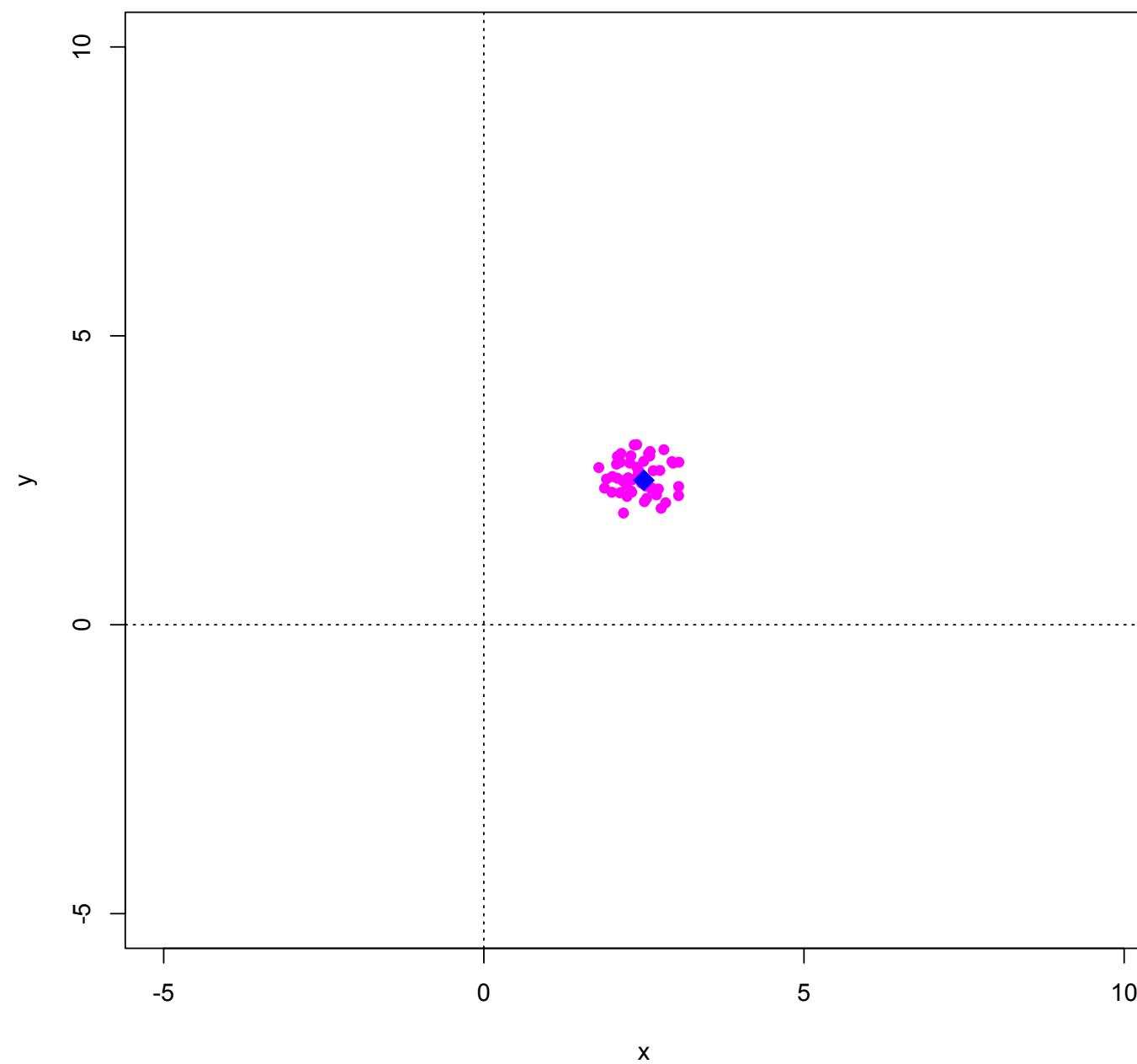
High Variance, Low Bias



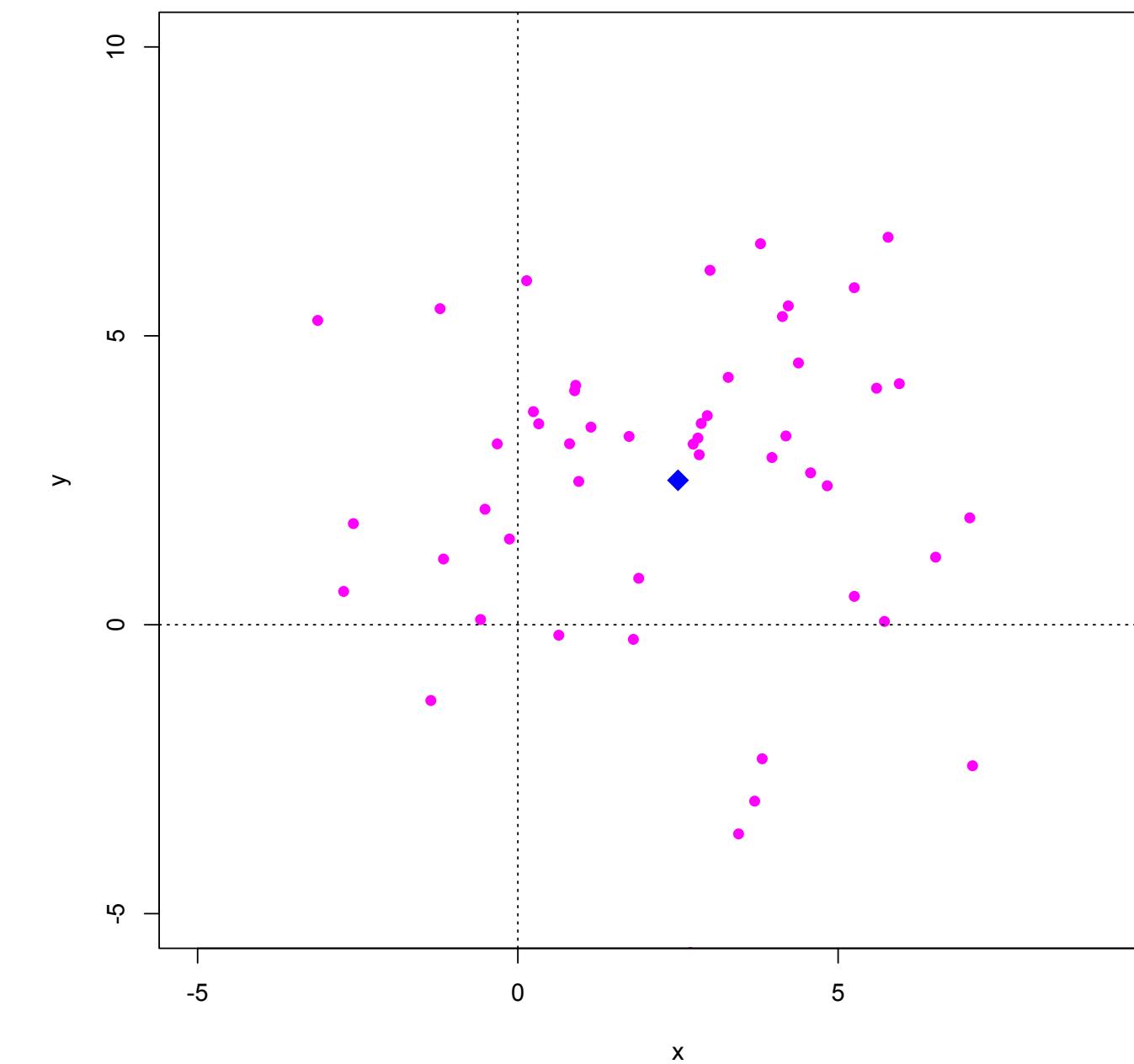
High Variance, High Bias



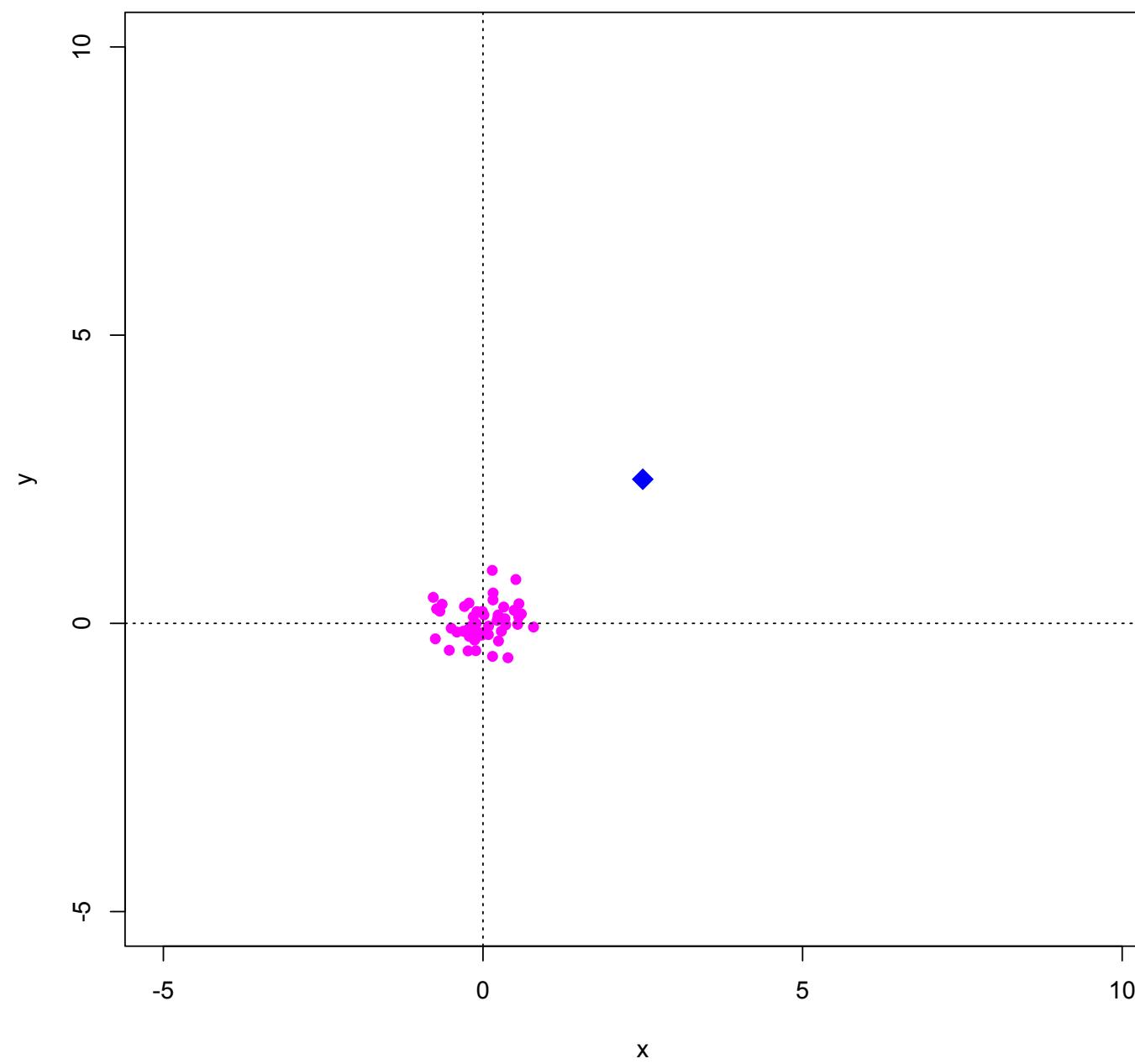
Low Variance, Low Bias



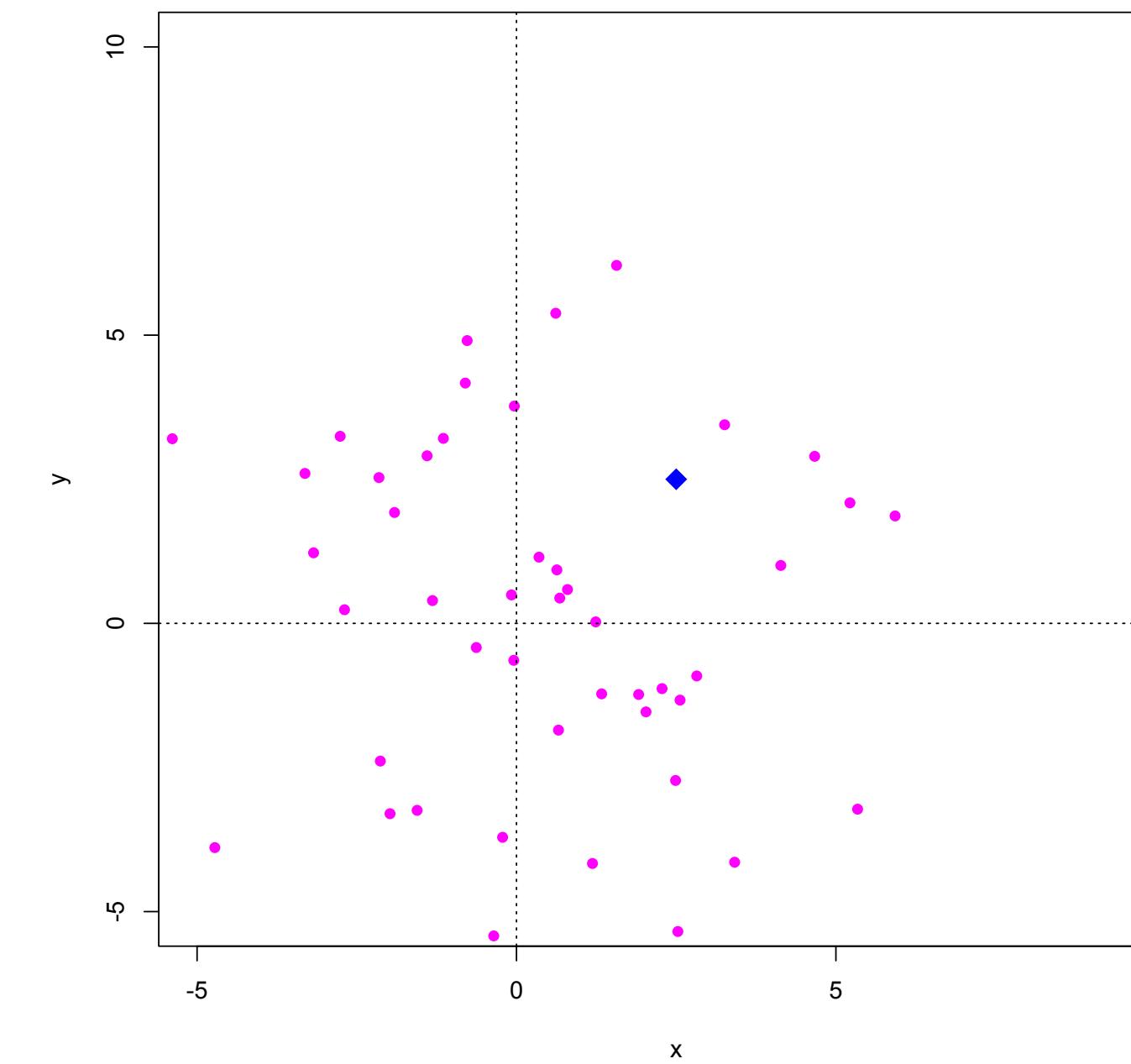
High Variance, Low Bias



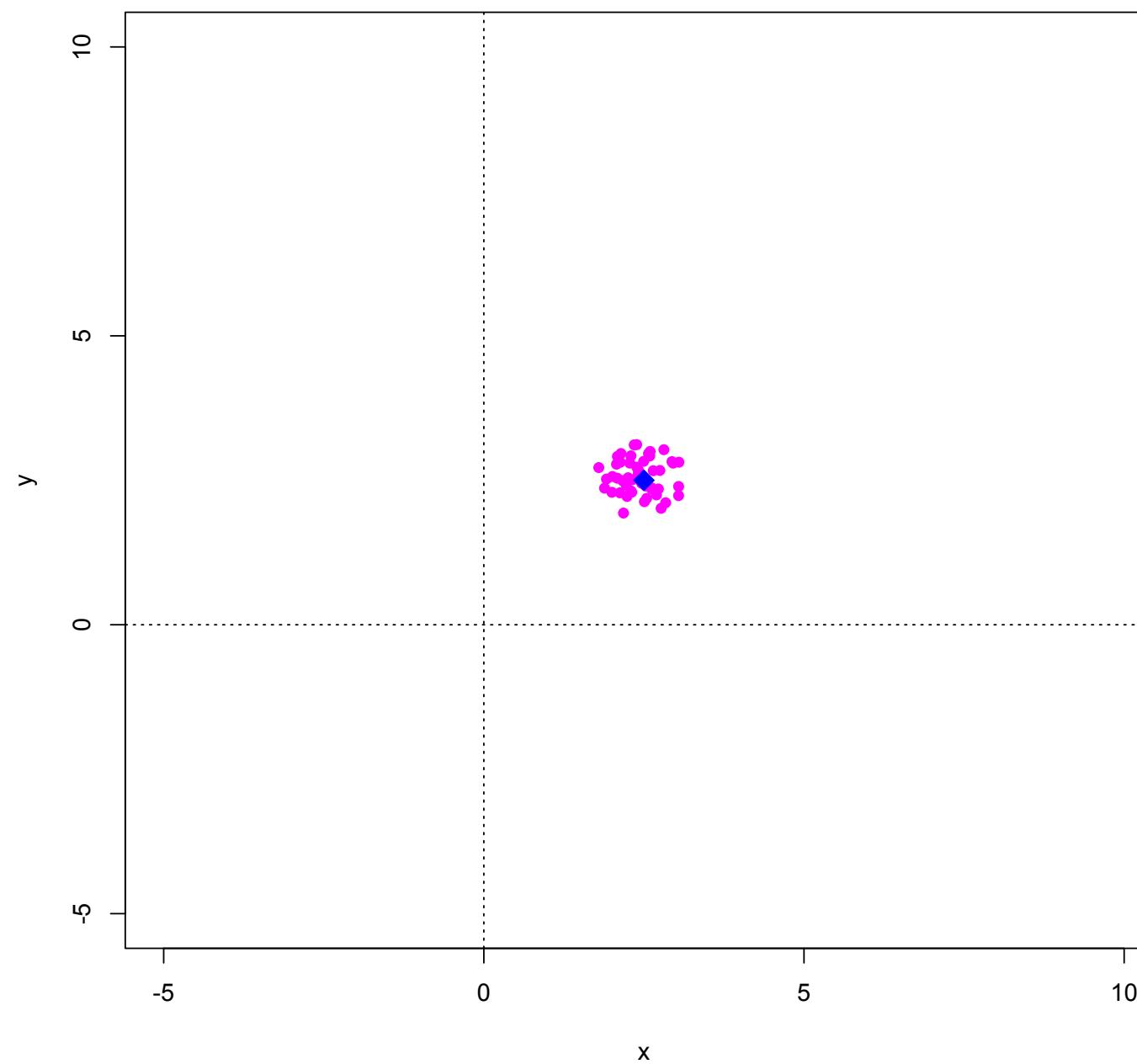
Low Variance, High Bias



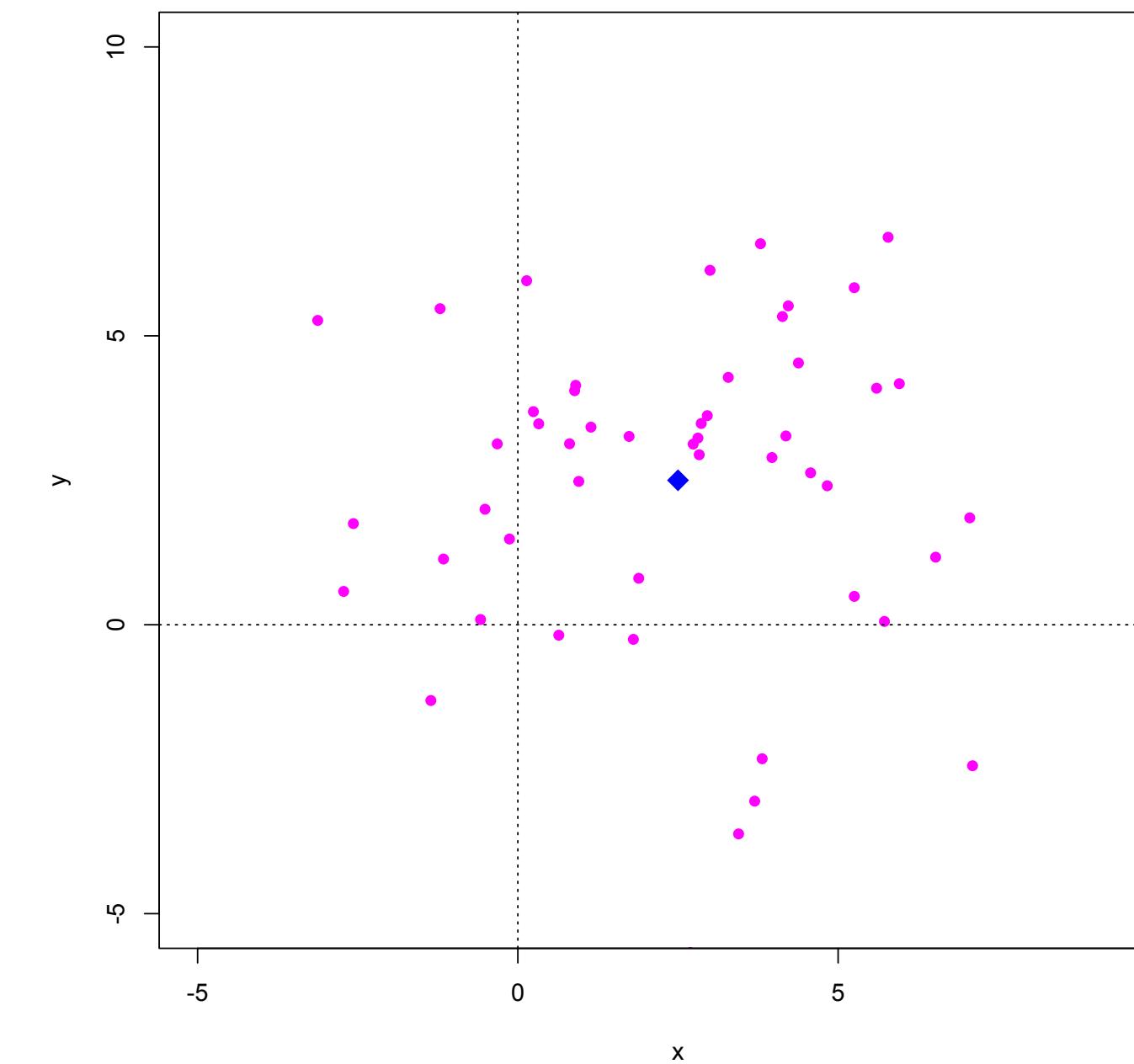
High Variance, High Bias



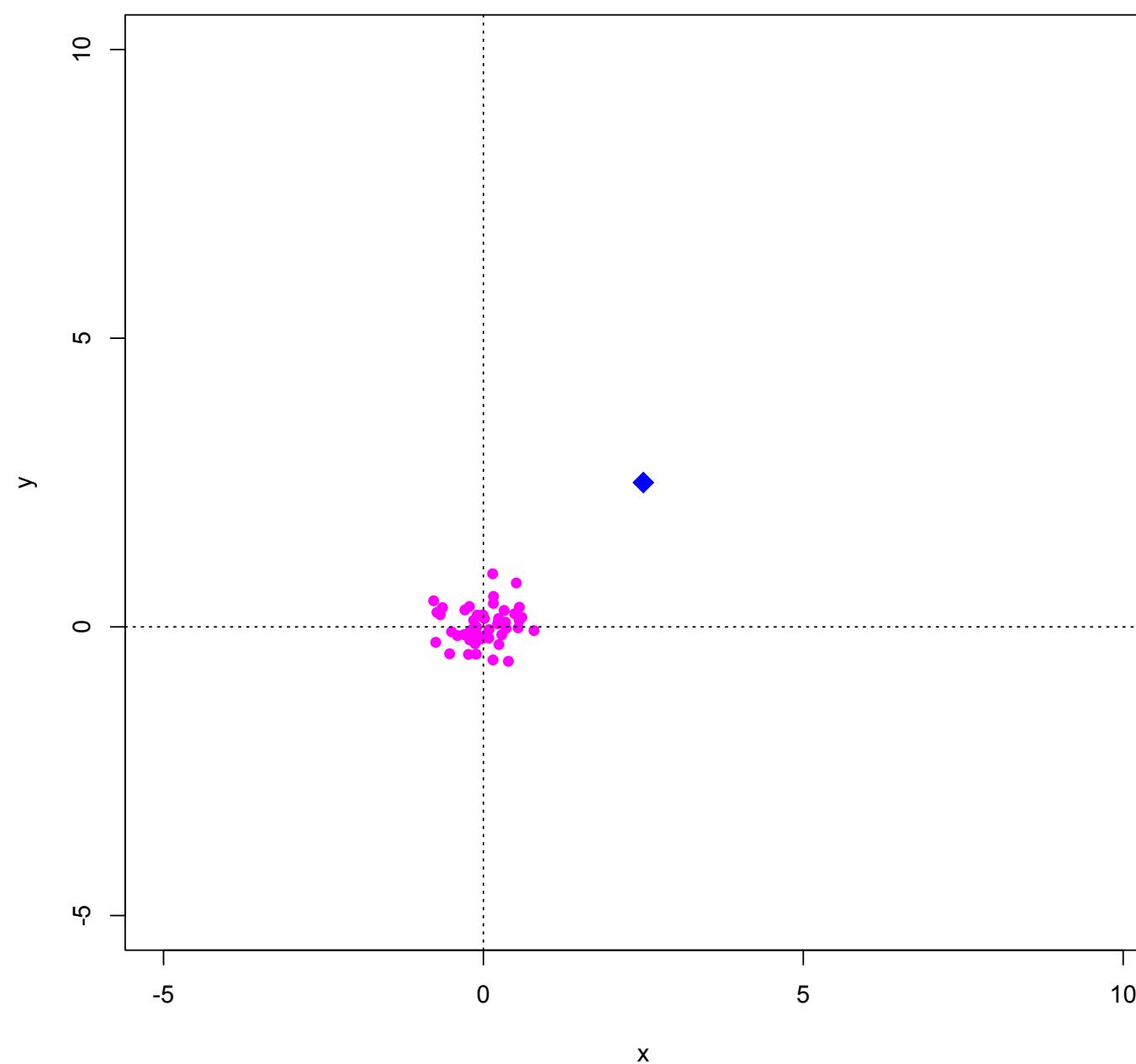
Low Variance, Low Bias



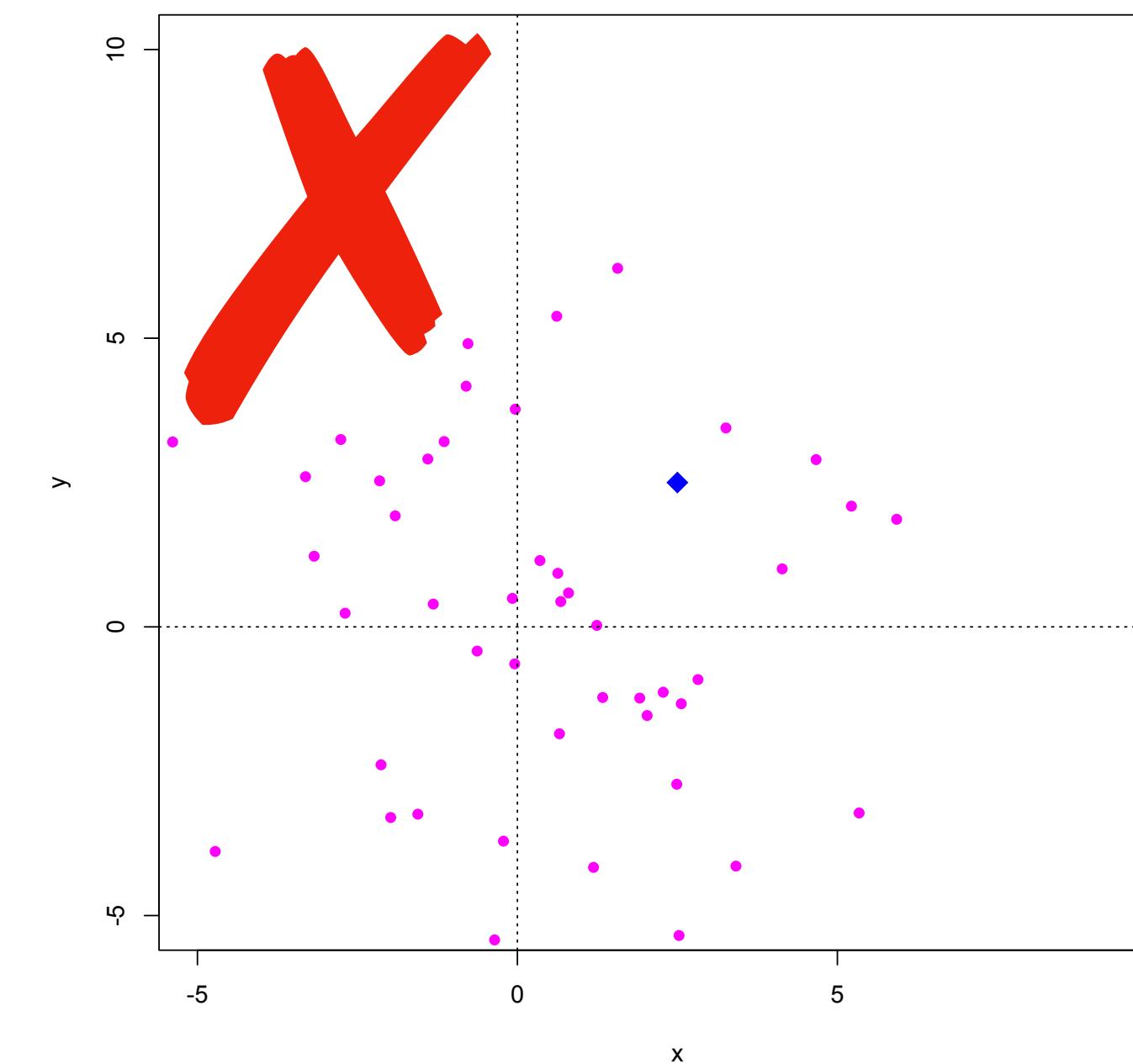
High Variance, Low Bias



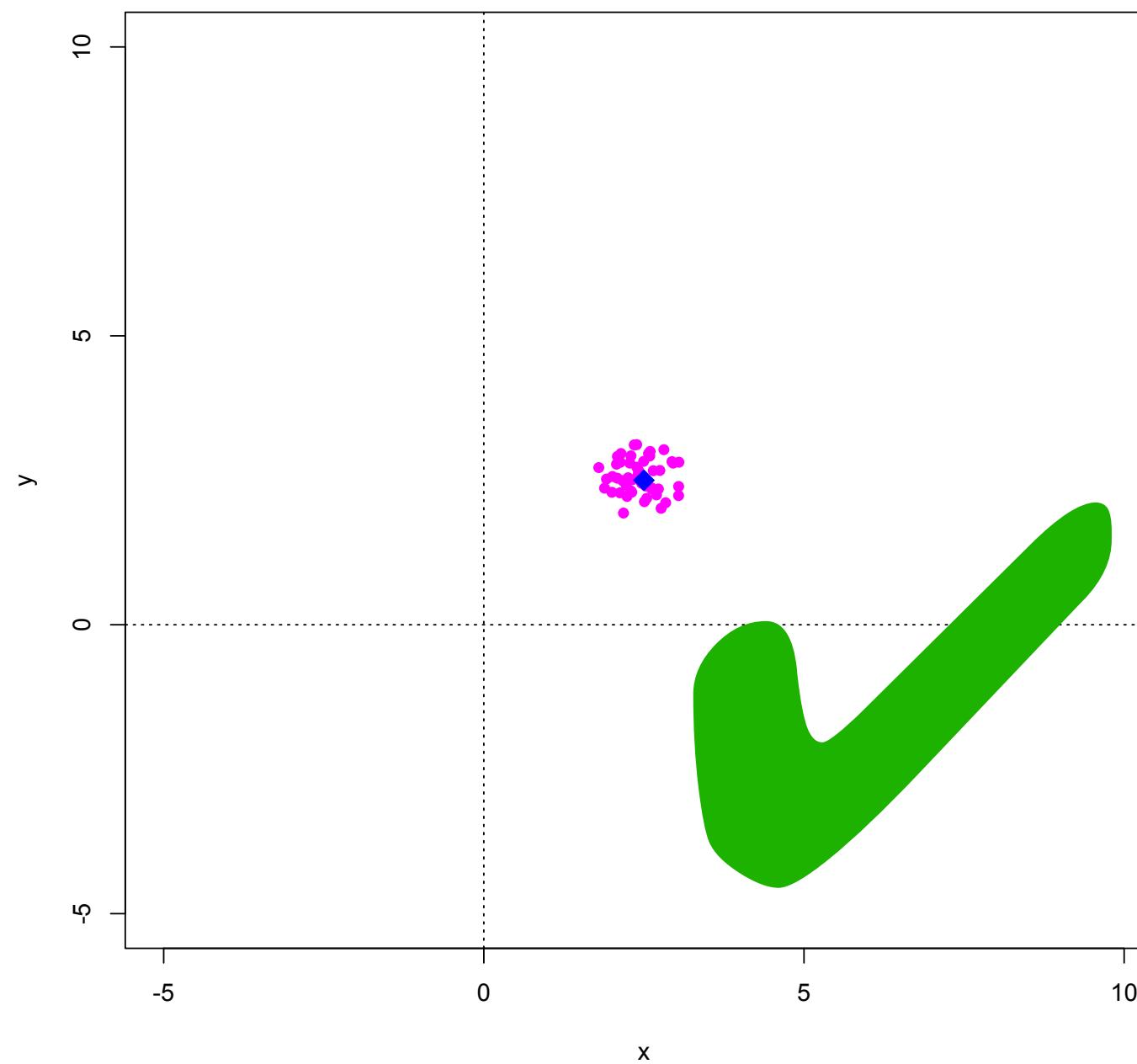
Low Variance, High Bias



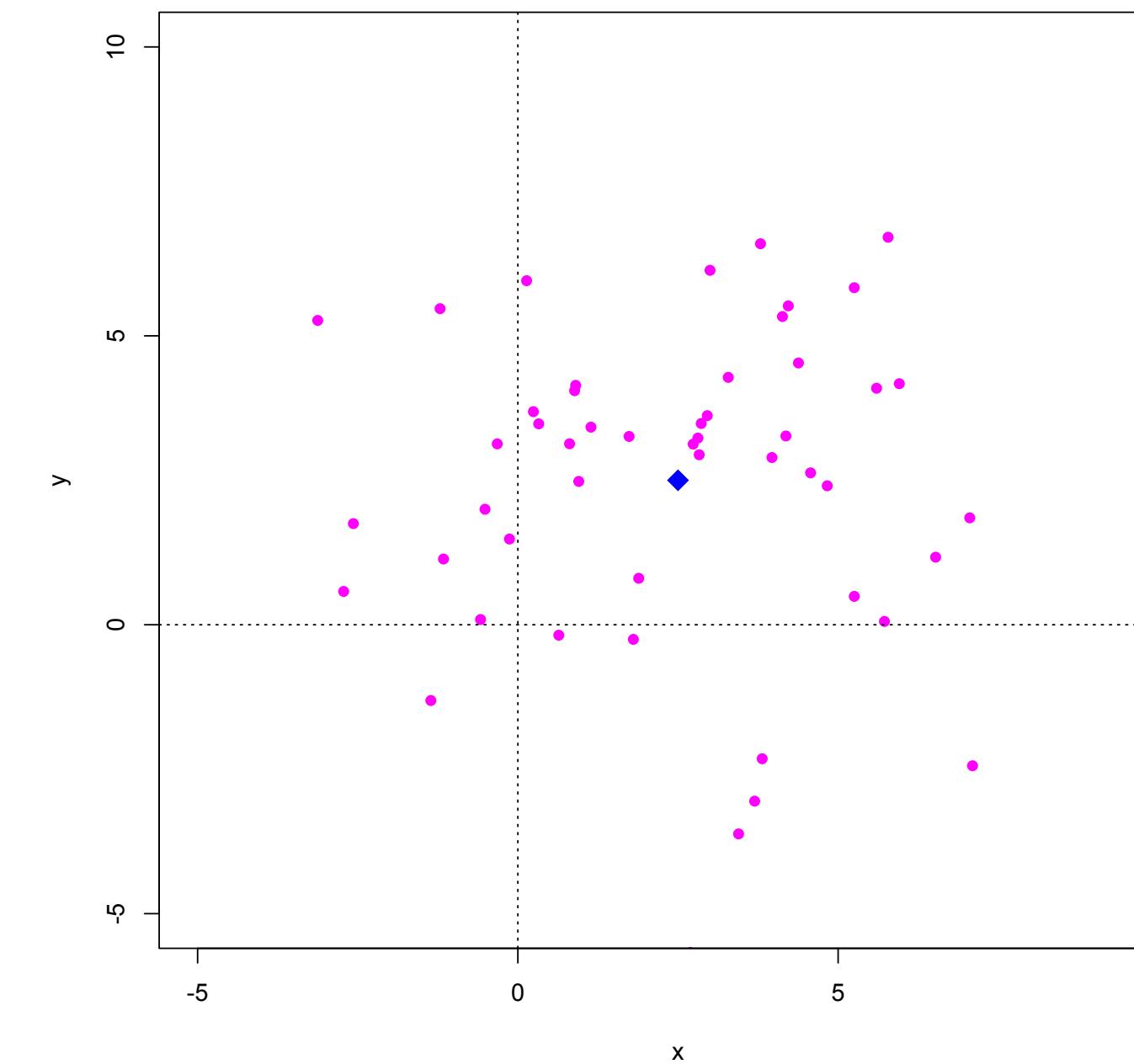
High Variance, High Bias



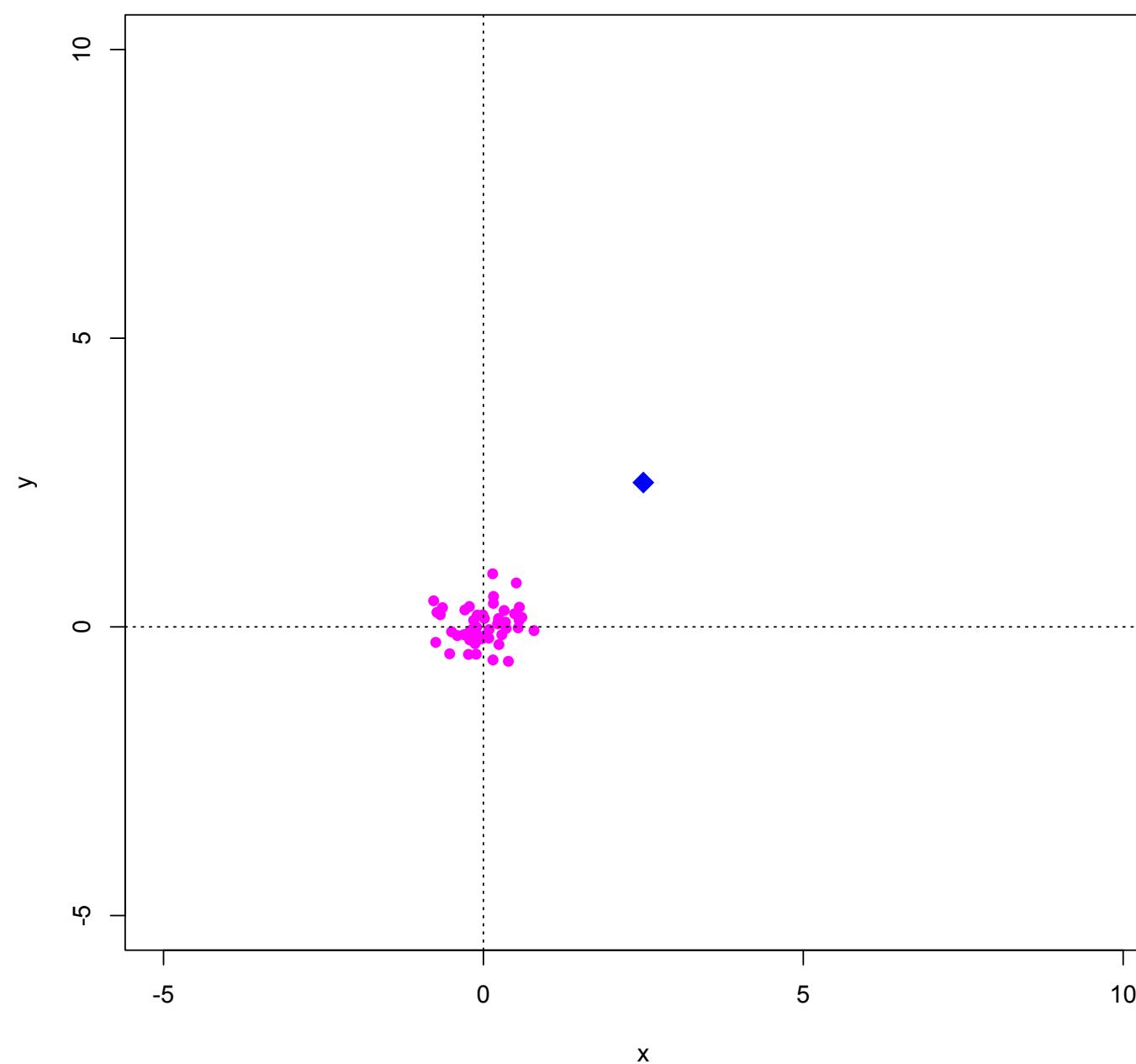
Low Variance, Low Bias



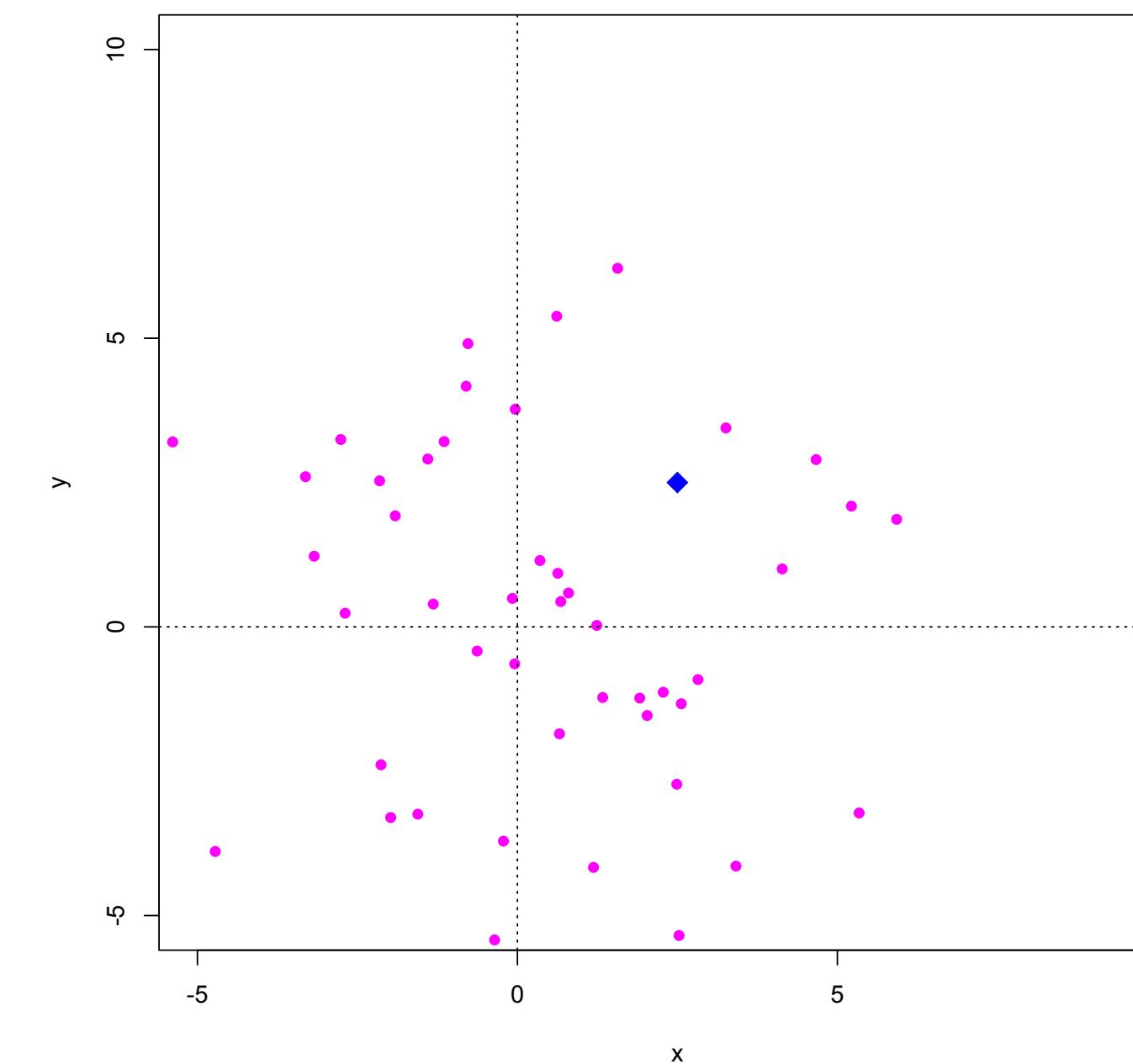
High Variance, Low Bias



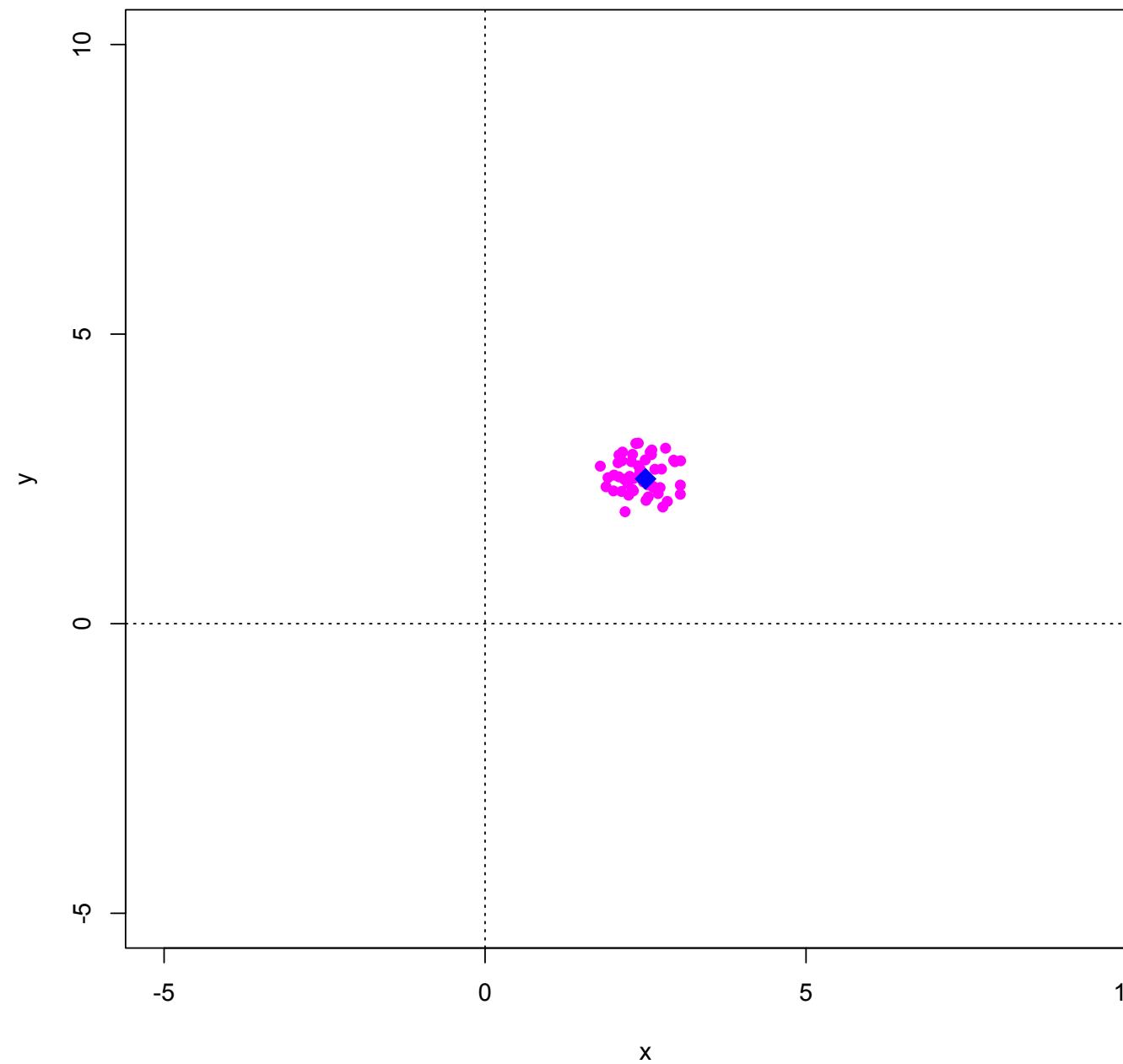
Low Variance, High Bias



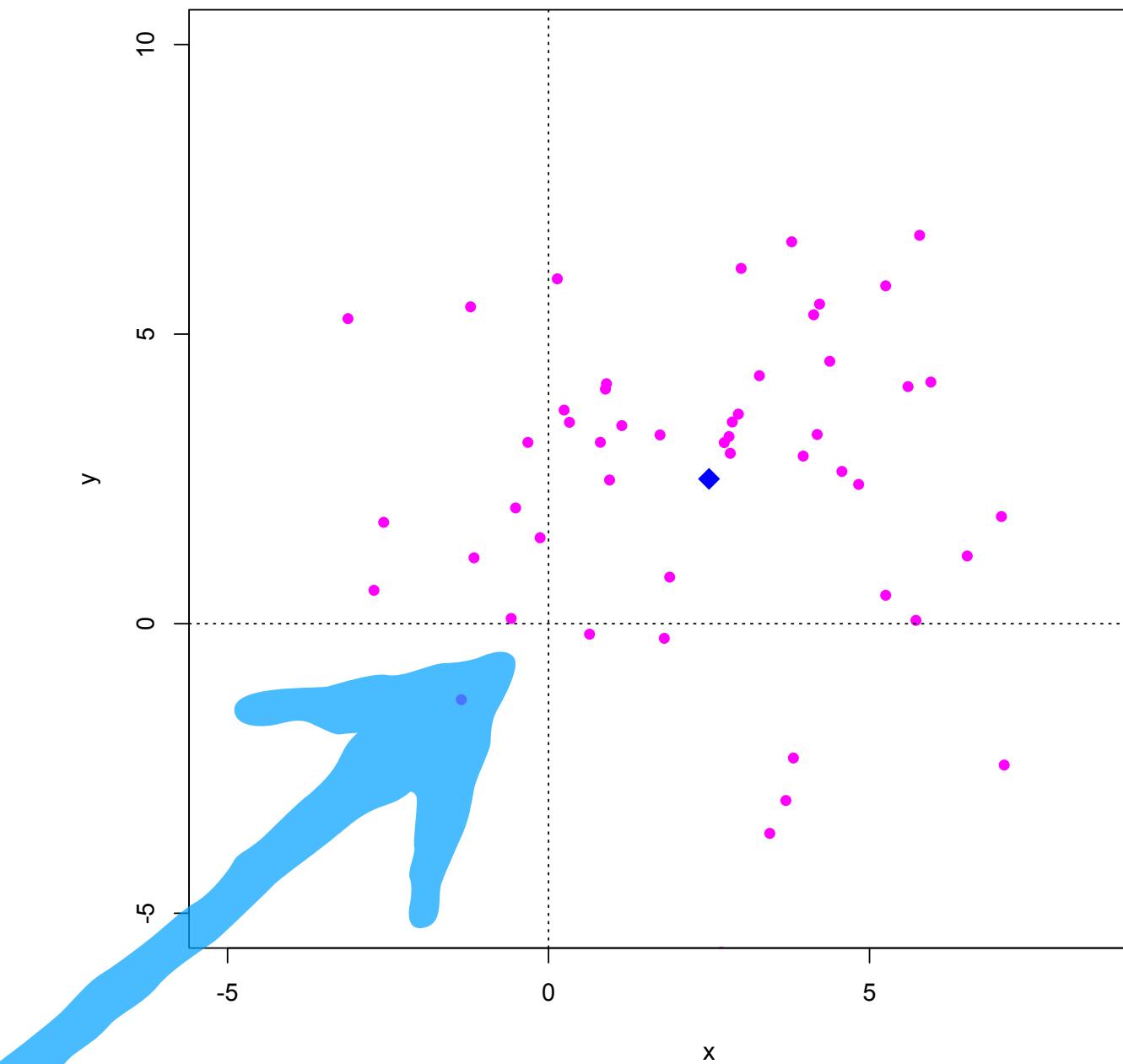
High Variance, High Bias



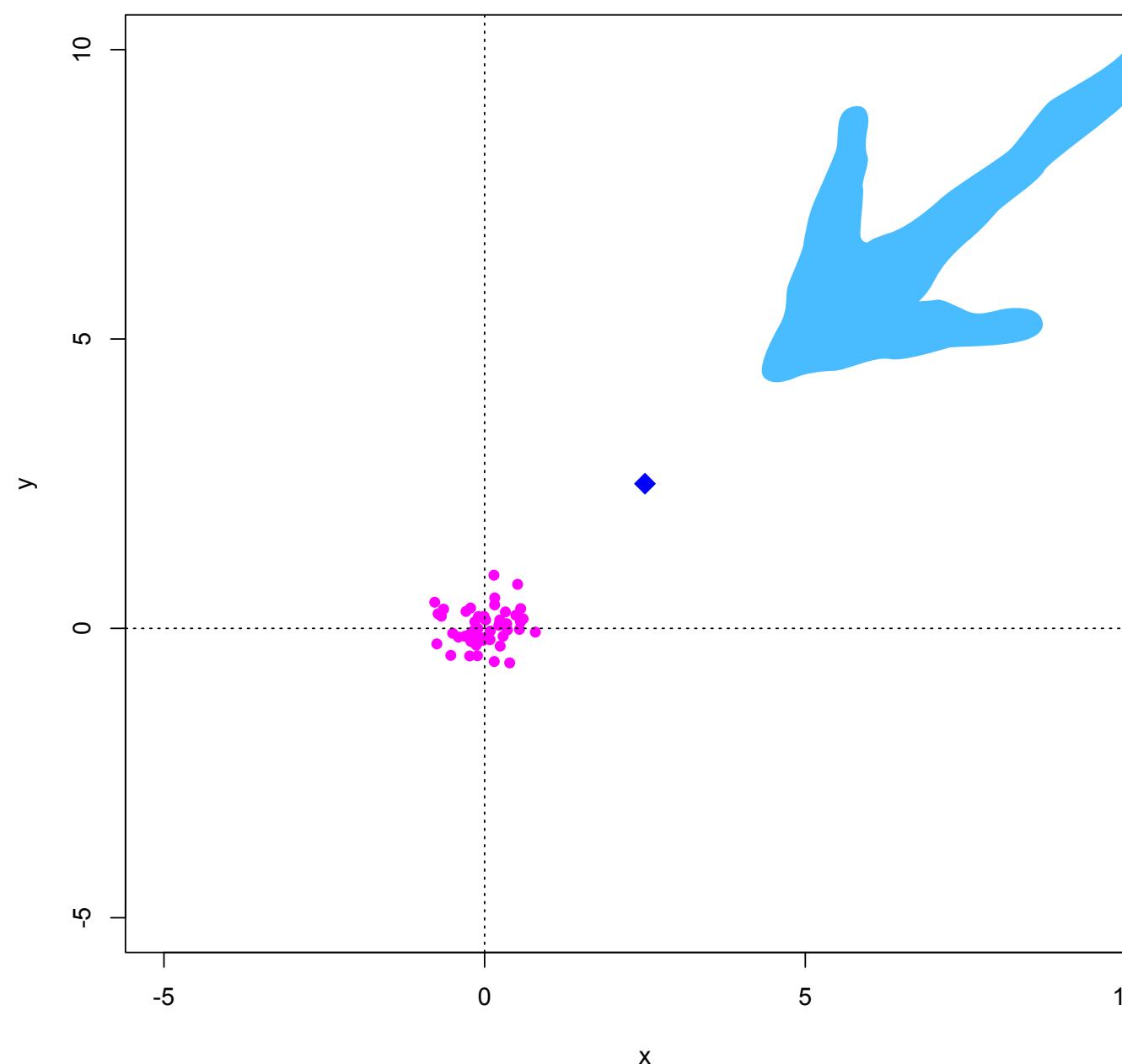
Low Variance, Low Bias



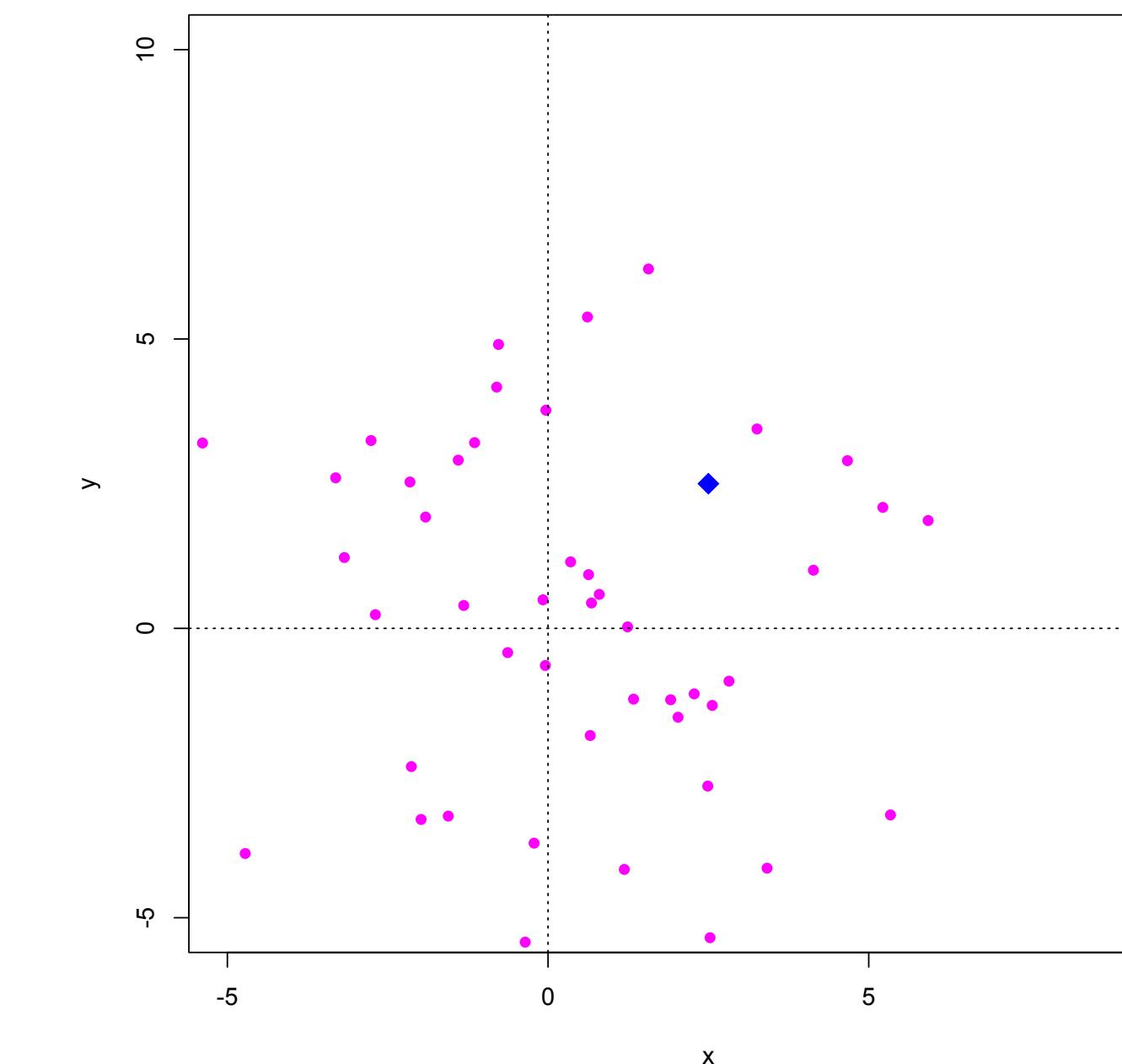
High Variance, Low Bias



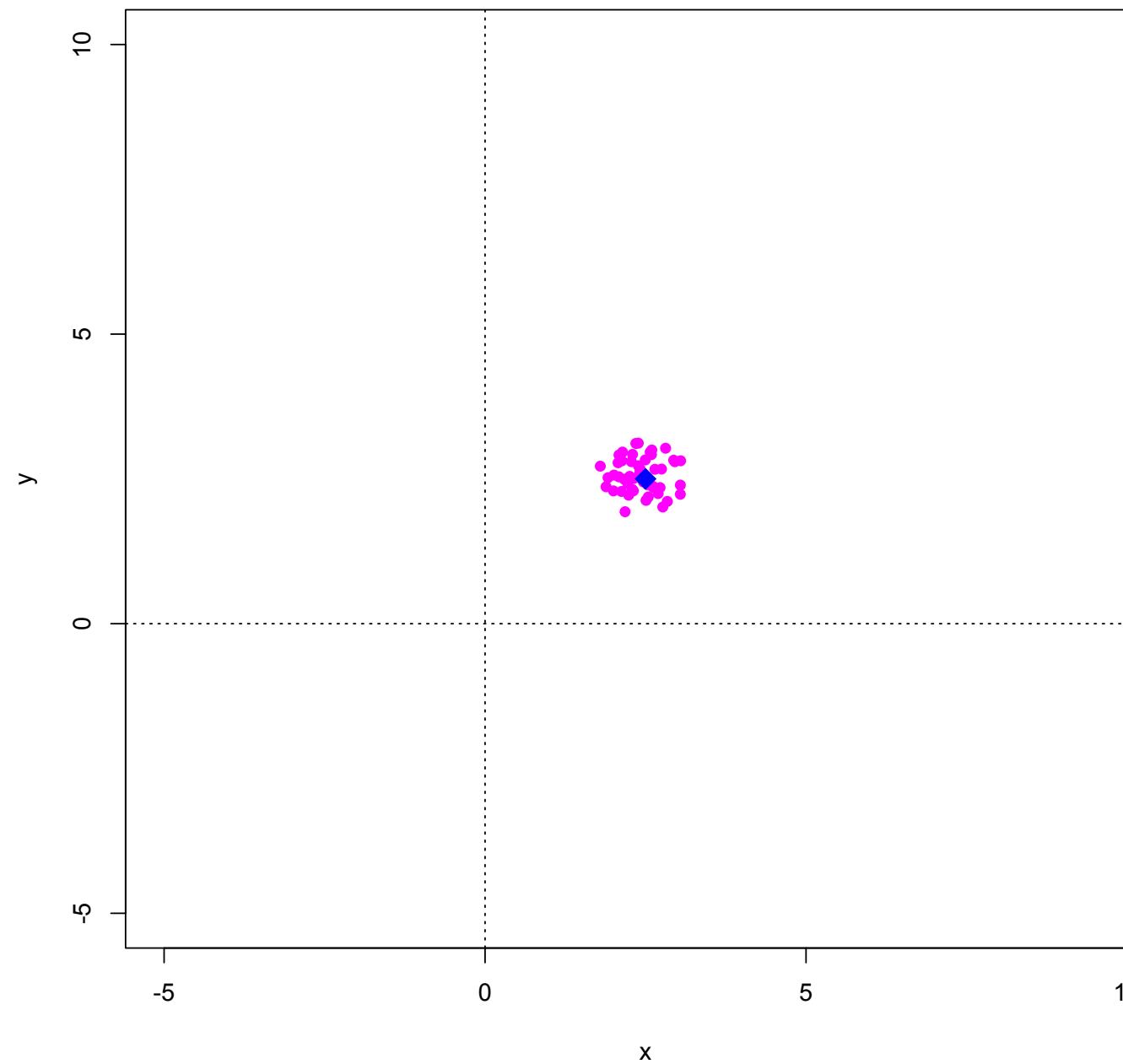
Low Variance, High Bias



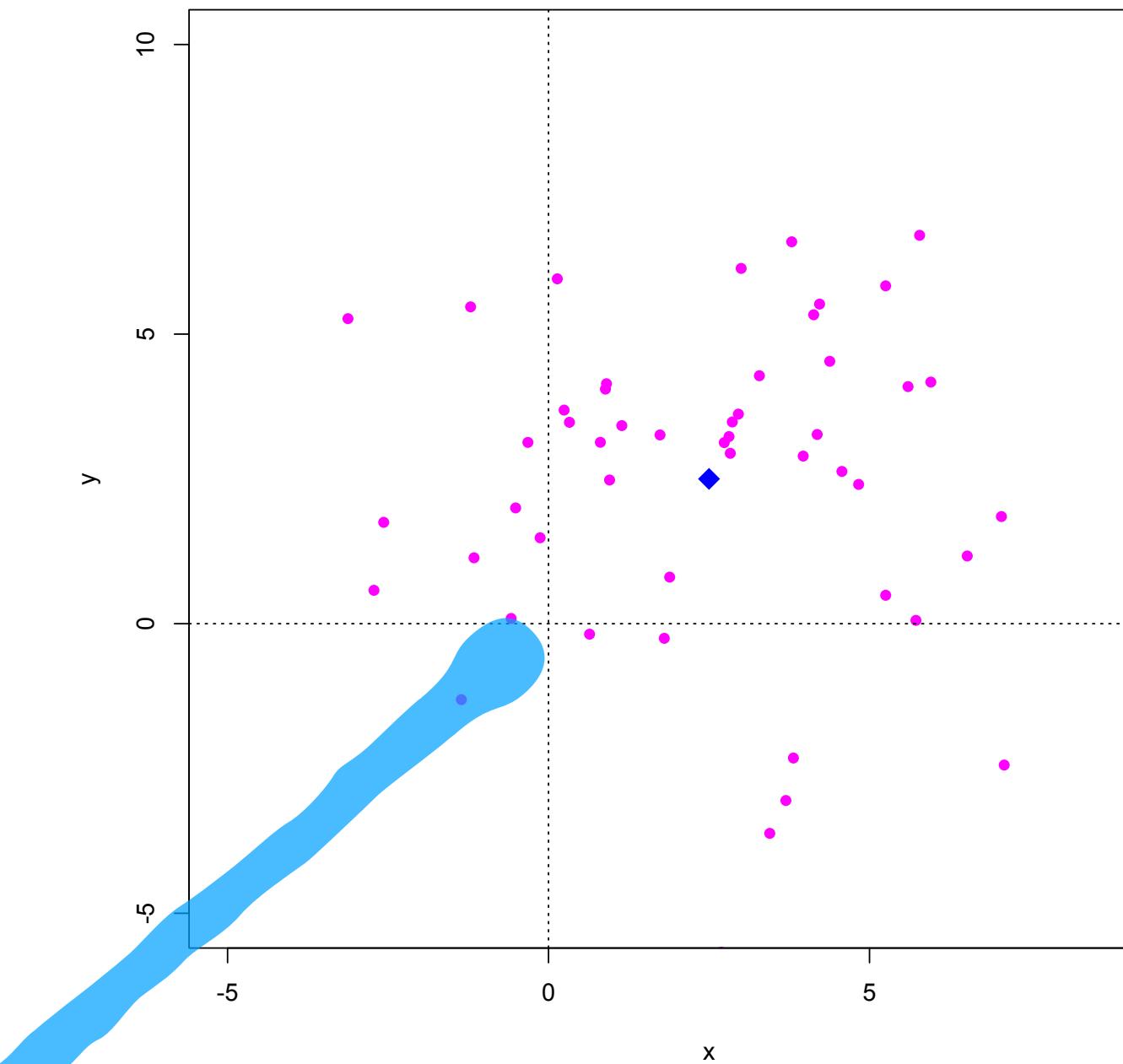
High Variance, High Bias



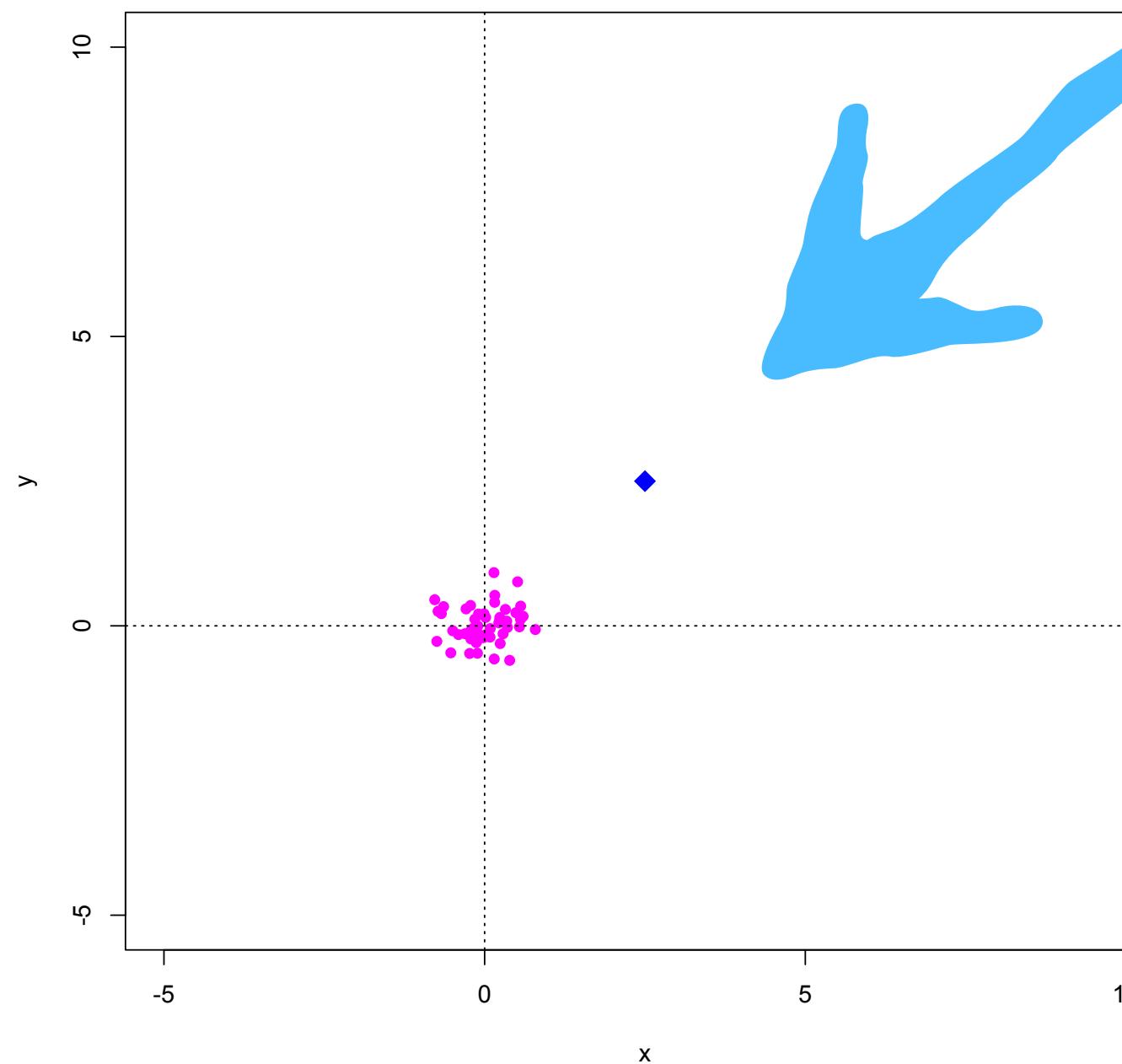
Low Variance, Low Bias



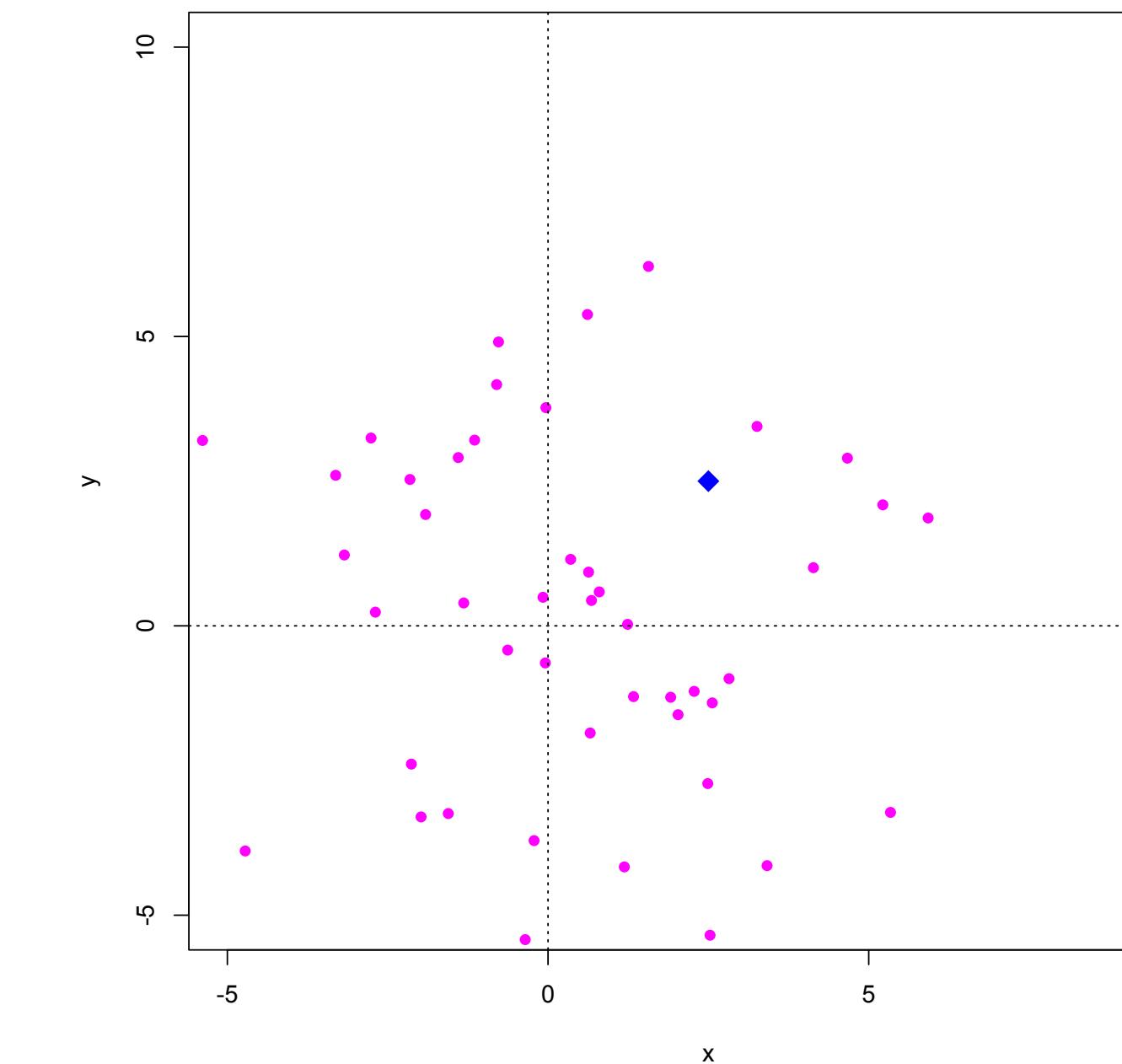
High Variance, Low Bias



Low Variance, High Bias



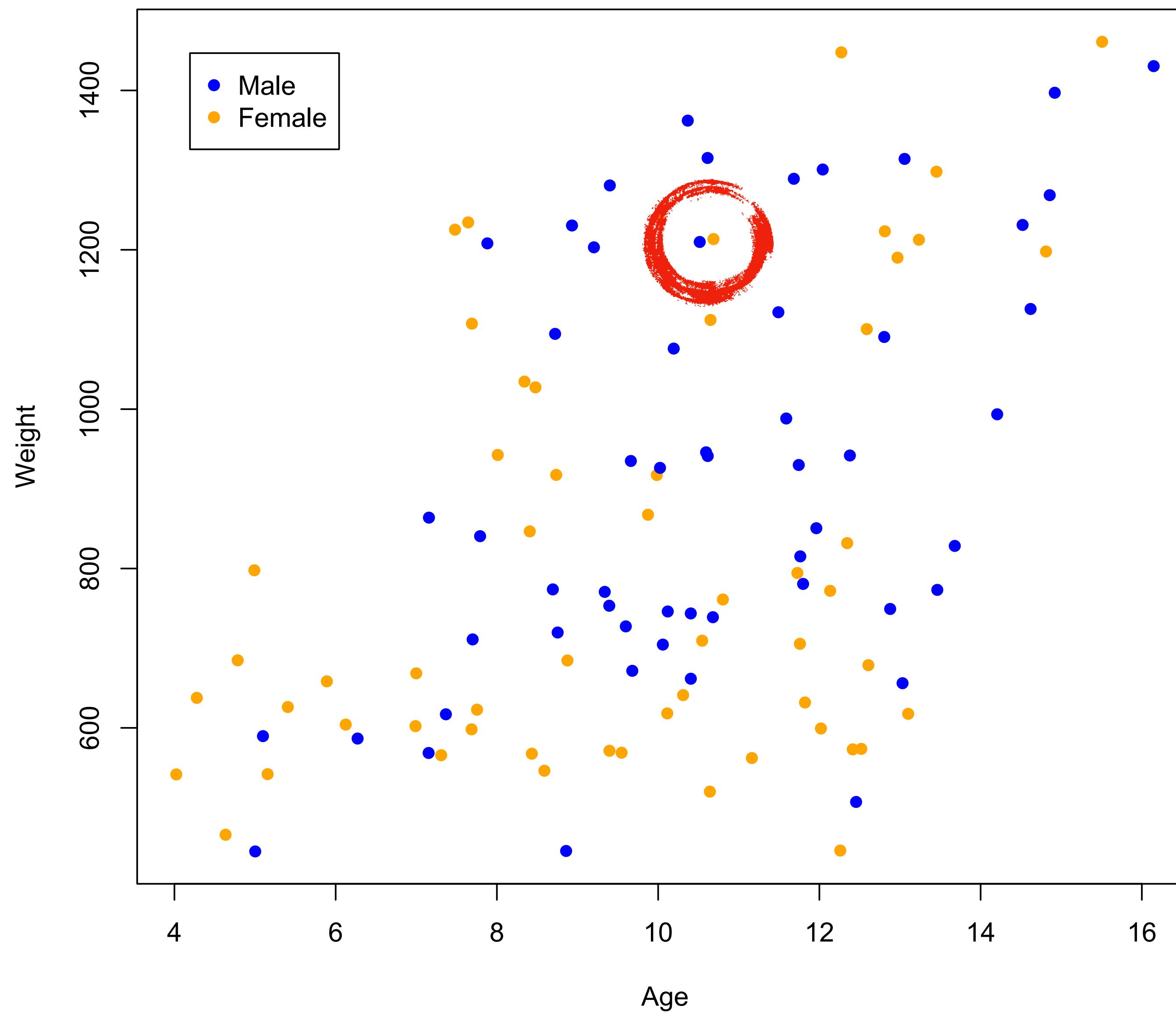
High Variance, High Bias



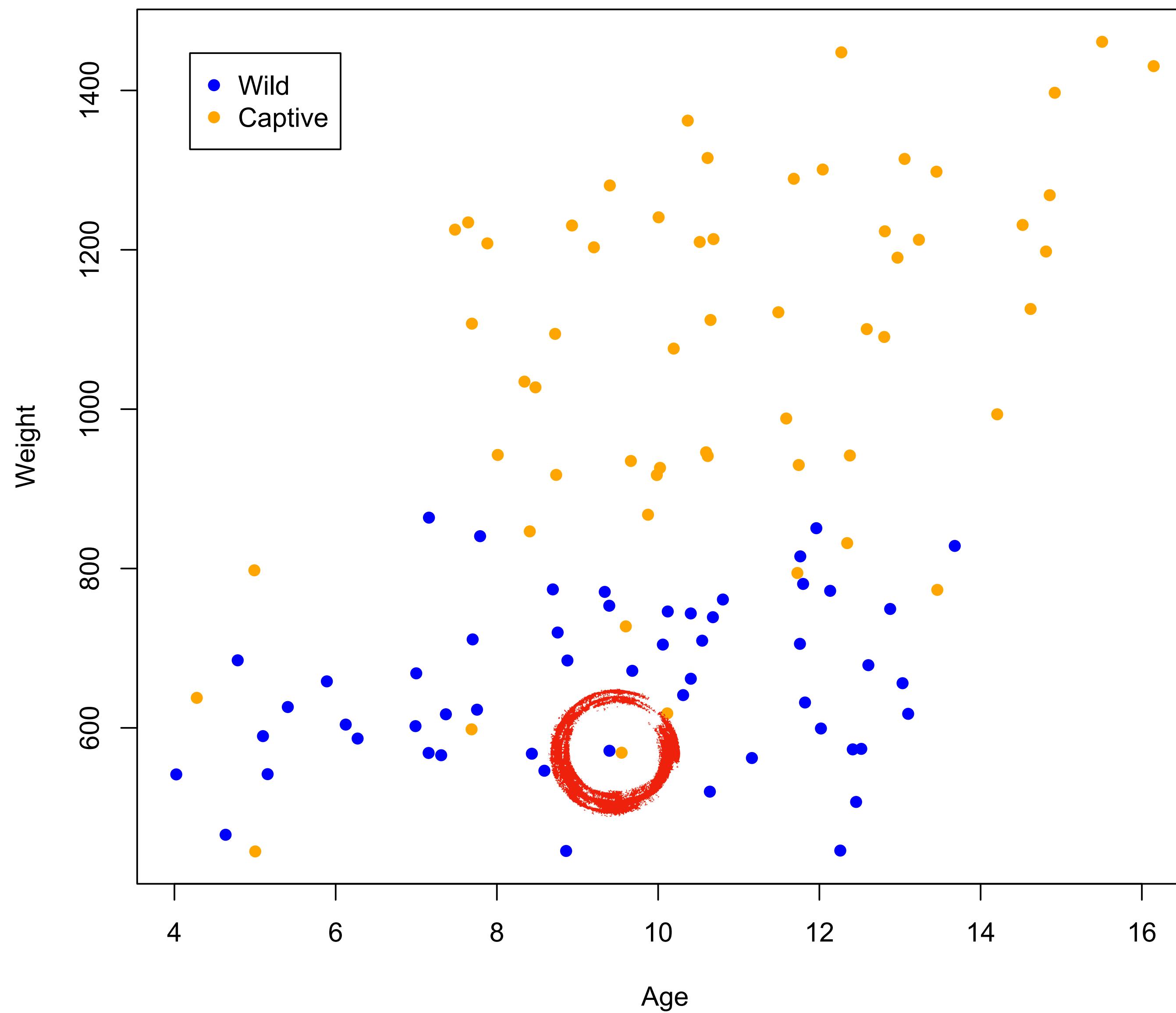


Neighbors

Meerkat Weight vs Age



Meerkat Weight vs Age



Meerkats

#	Age	Weight	Sex	Captive	...
1	11.8	815	1	0	
2	9.7	672	1	0	
3	8.9	446	1	0	
4	10.8	761	0	0	
5	8.3	1035	0	1	
6	11.7	930	1	1	
7	8.5	1027	0	1	
8	7.6	1234	0	1	
9	15.5	1461	0	1	
10	8.8	720	1	0	
11	12.8	1223	0	1	
12	7.7	711	1	0	
13	7.9	586			

Meerkats

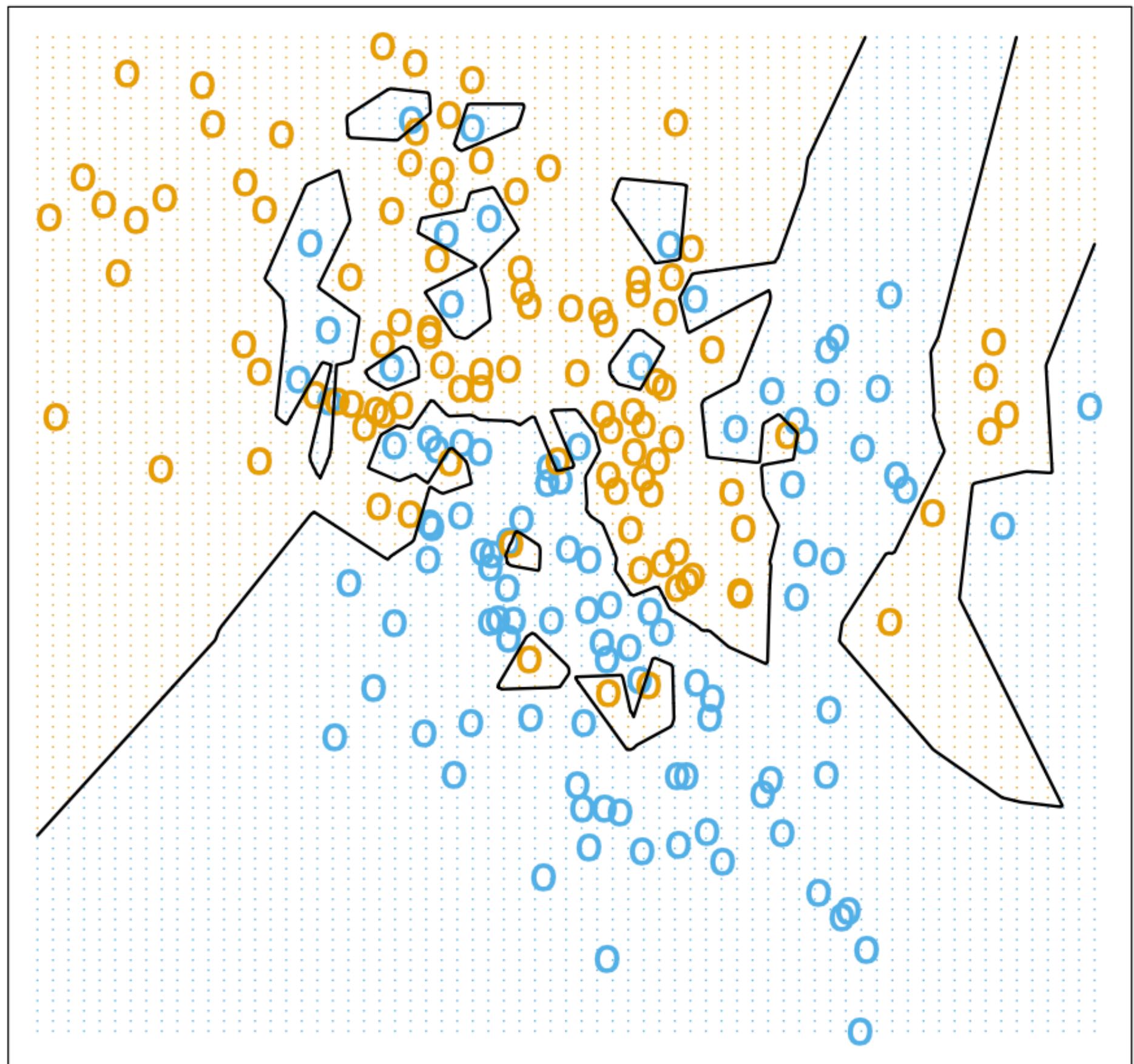
#	Age	Weight	Sex	Captive	...
1	11.8	815	1	0	
2	9.7	672	1	0	
3	8.9	446	1	0	
4	10.8	761	0	0	
5	8.3	1035	0	1	
6	11.7	930	1	1	
7	8.5	1027	0	1	
8	7.6	1234	0	1	
9	15.5	1461	0	1	
10	8.8	720	1	0	
11	12.8	1223	0	1	
12	7.7	711	1	0	
13	7.9	585			

Meerkats

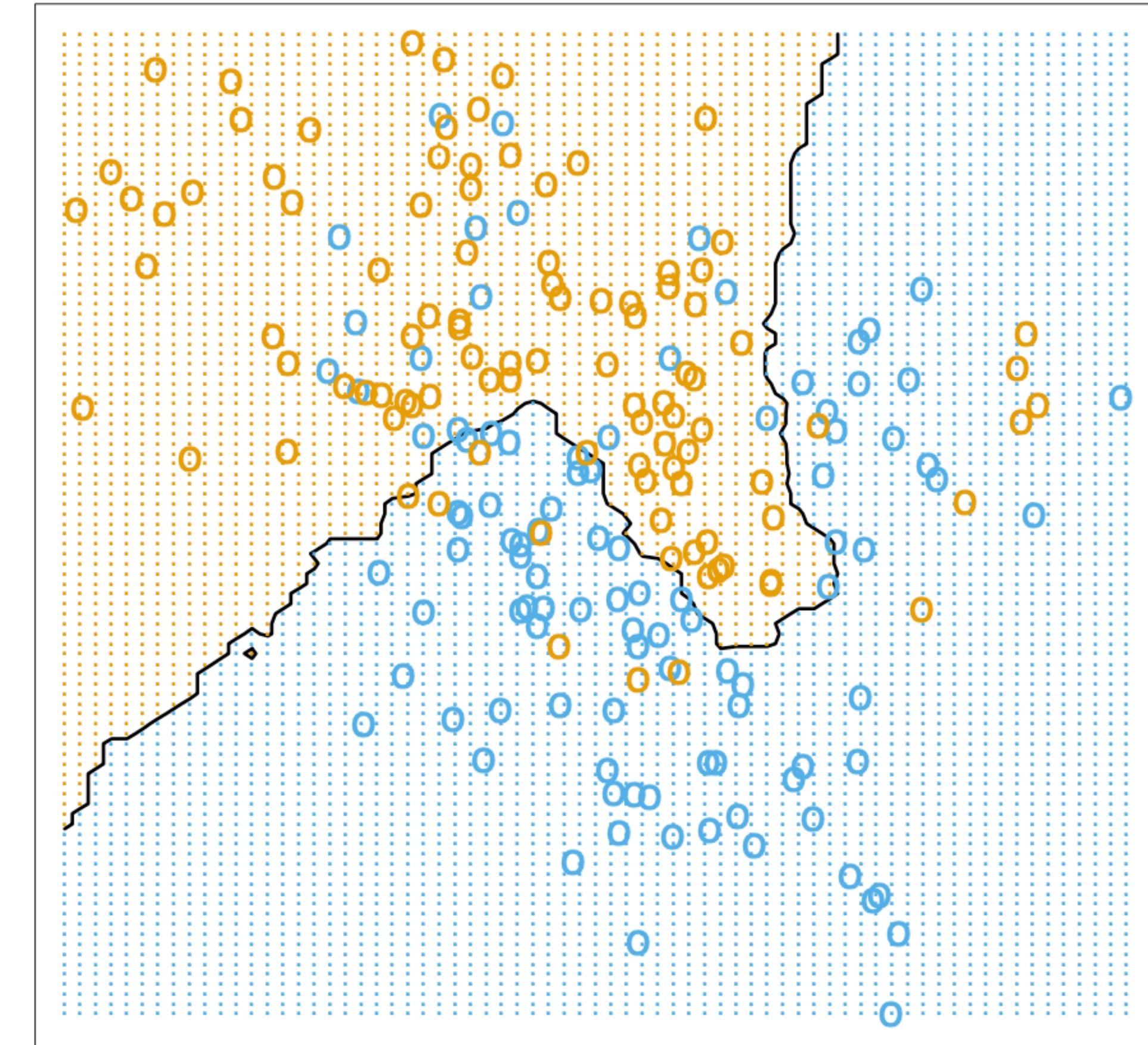
#	Age	Weight	Sex	Captive	...
1	11.8	815	1	0	
2	9.7	672	1	0	
3	8.9	446	1	0	
4	10.8	761	0	0	
5	8.3	1035	0	1	
6	11.7	930	1	1	
7	8.5	1027	0	1	
8	7.6	1234	0	1	
9	15.5	1461	0	1	
10	8.8	720	1	0	
11	12.8	1223	0	1	
12	7.7	711	1	0	
13	7.9	585			

Meerkats

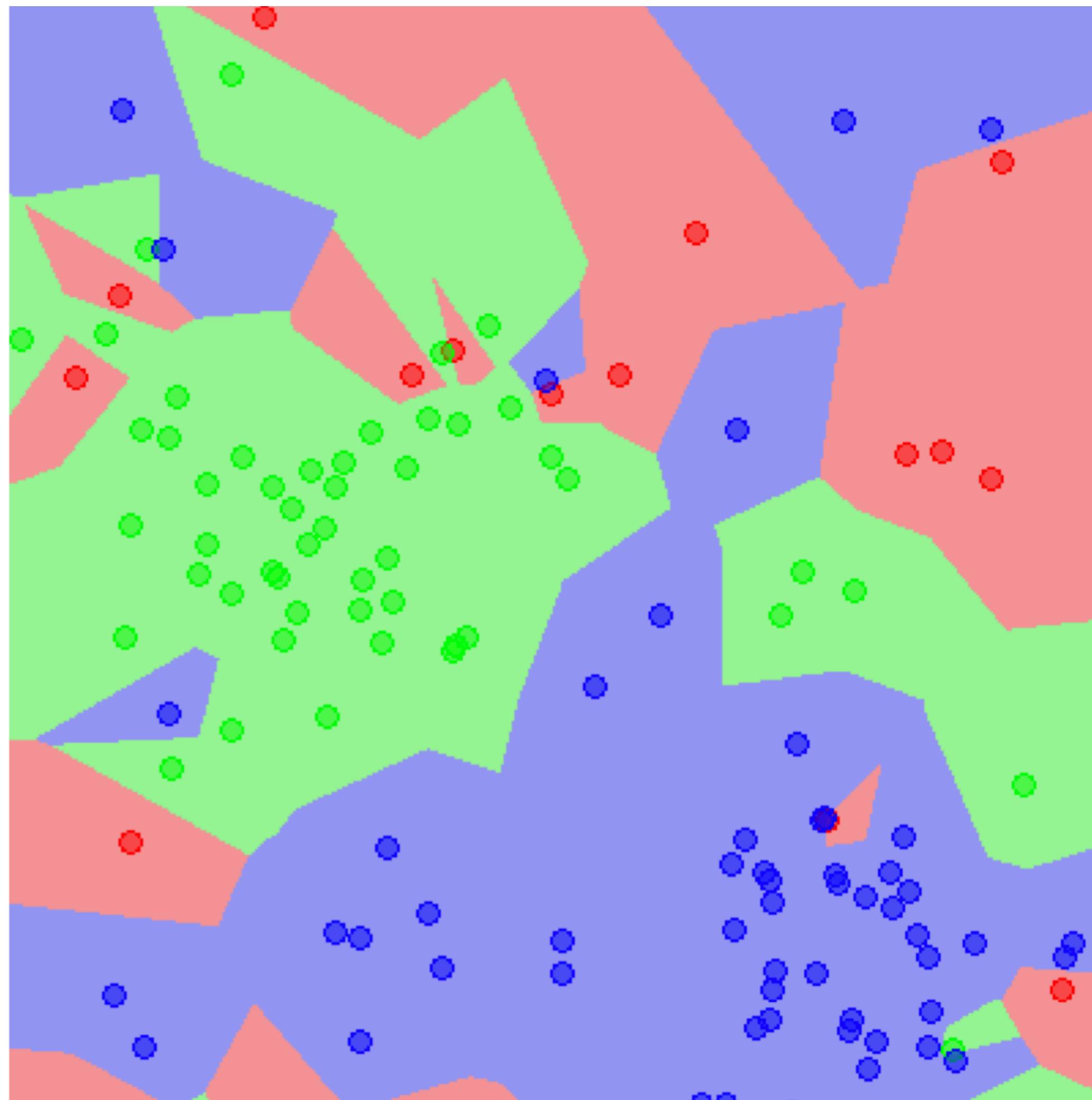
#	Age	Weight	Sex	Captive	...
1	11.8	815	1	0	
2	9.7	672	1	0	
3	8.9	446	1	0	
4	10.8	761	0	0	
5	8.3	1035	0	1	
6	11.7	930	1	1	
7	8.5	1027	0	1	
8	7.6	1234	0	1	
9	15.5	1461	0	1	
10	8.8	720	1	0	
11	12.8	1223	0	1	
12	7.7	711	1	0	
13	7.9	585			



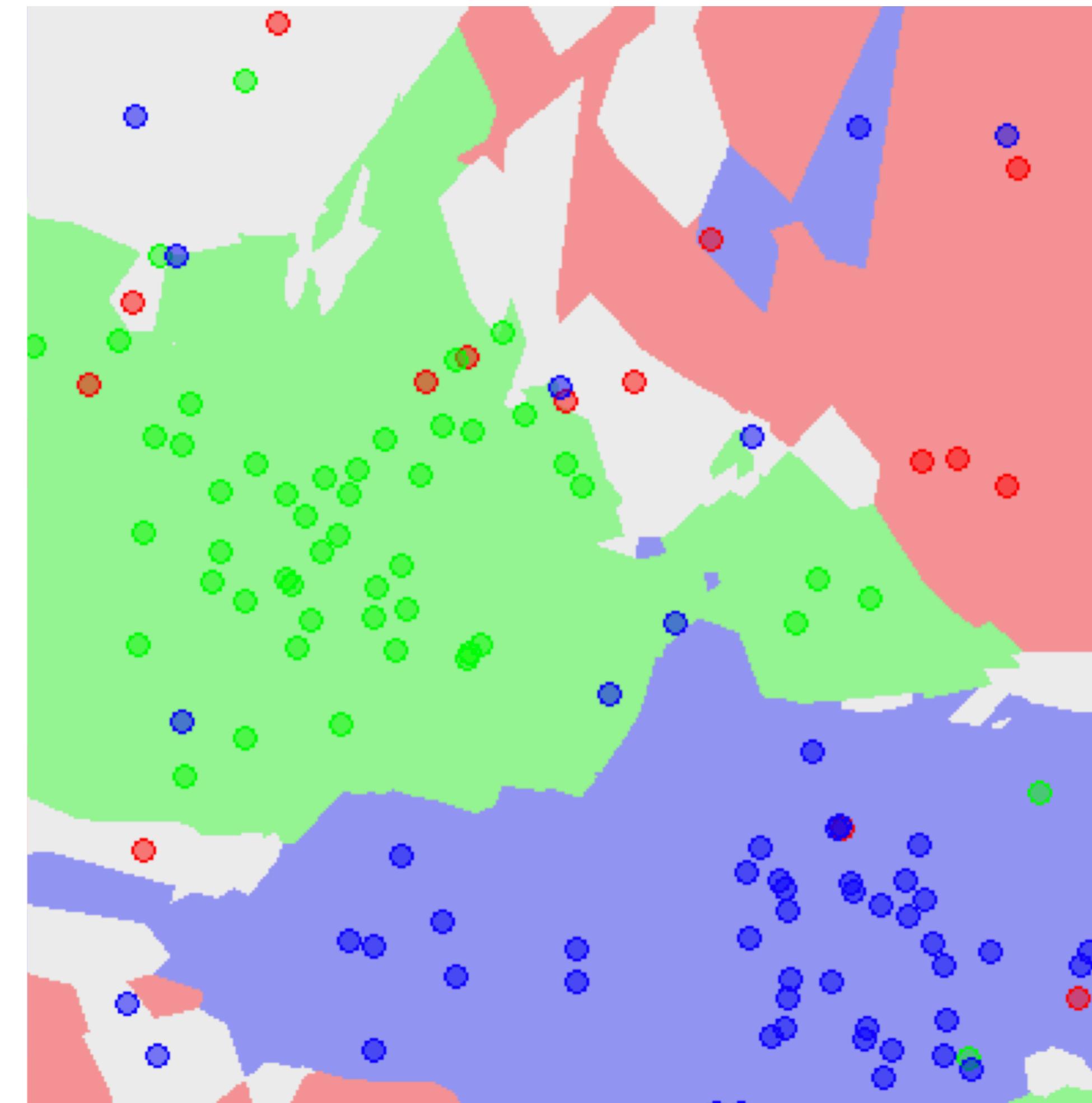
$k=1$



$k=15$



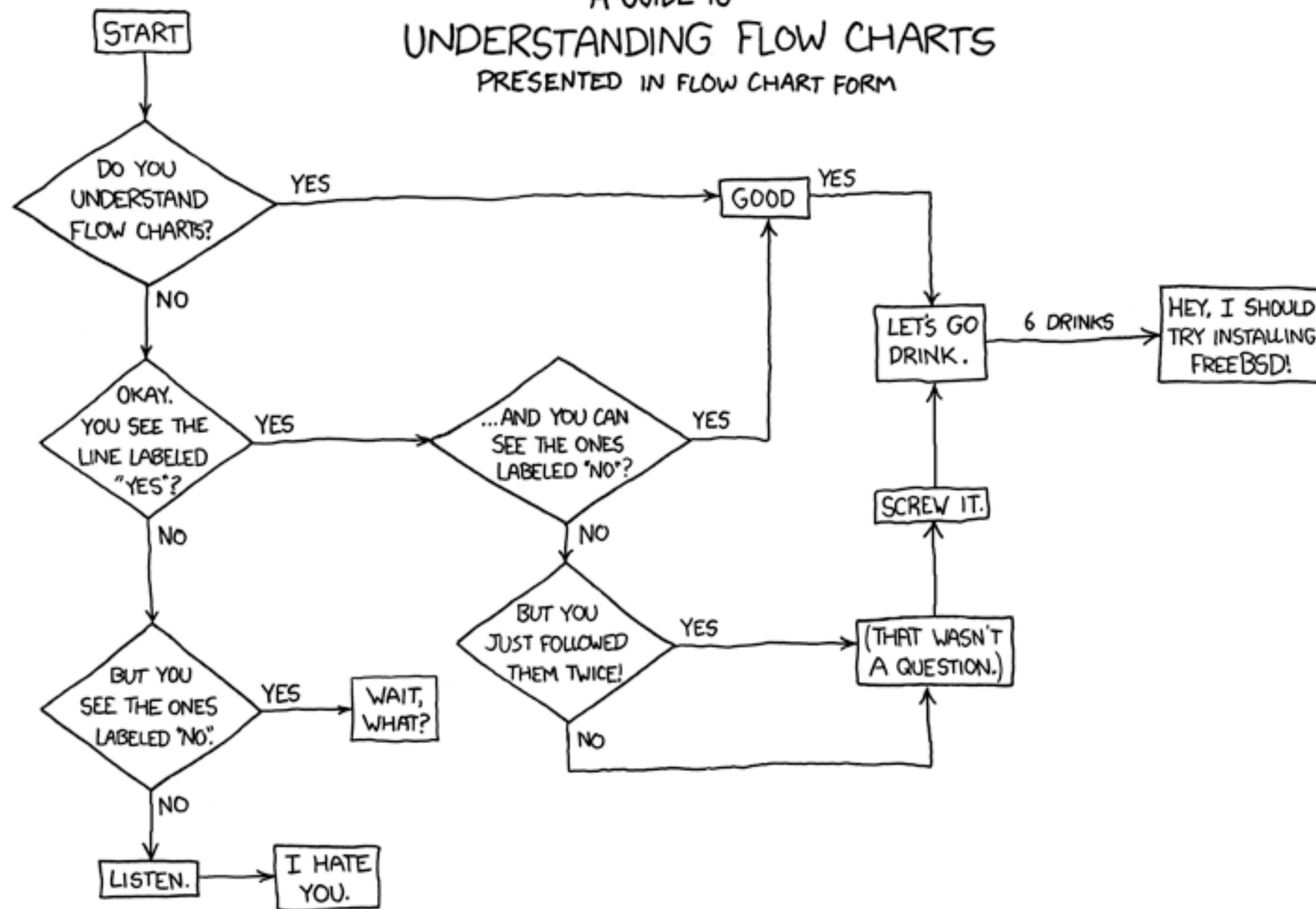
$k=1$

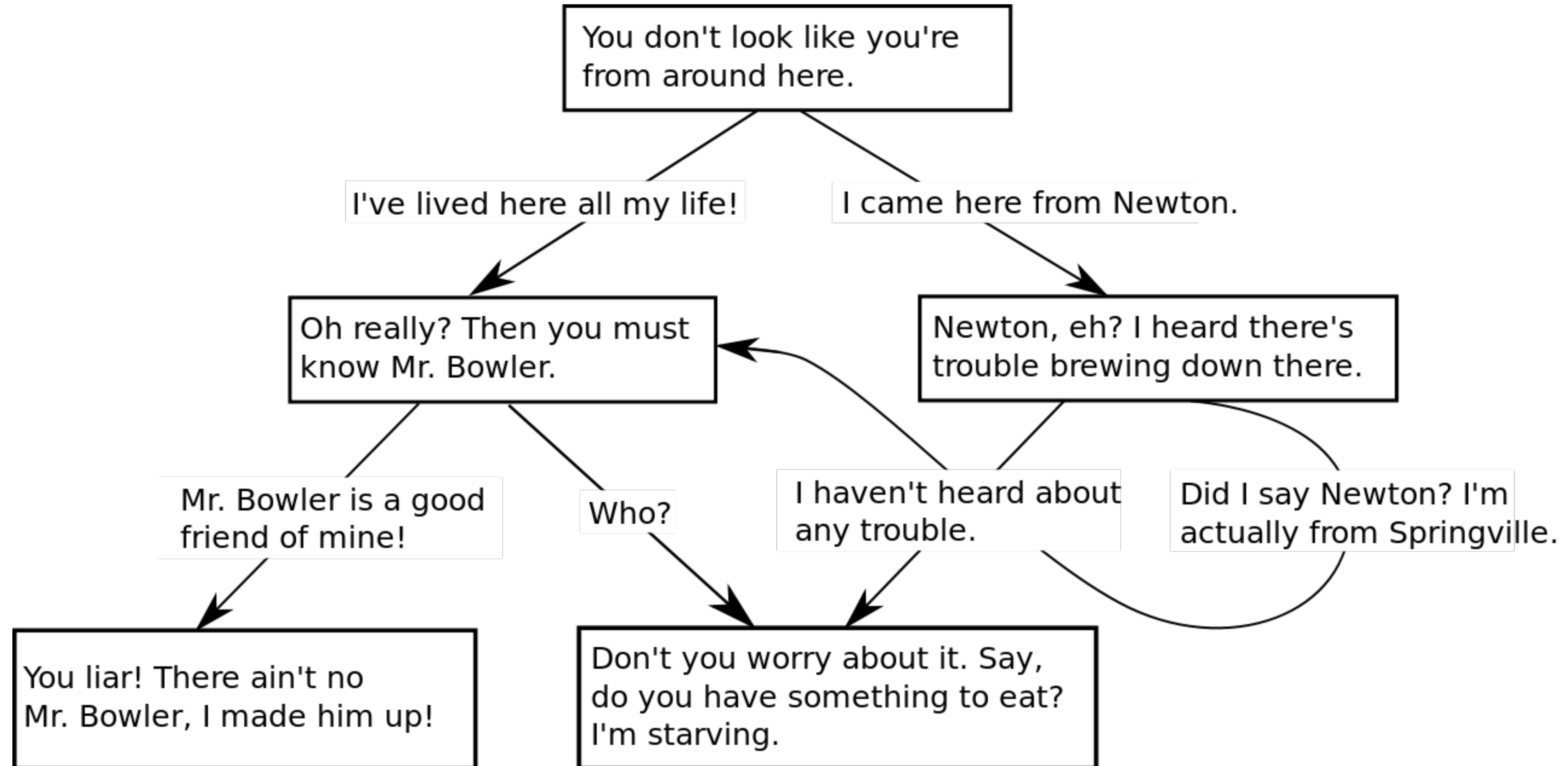


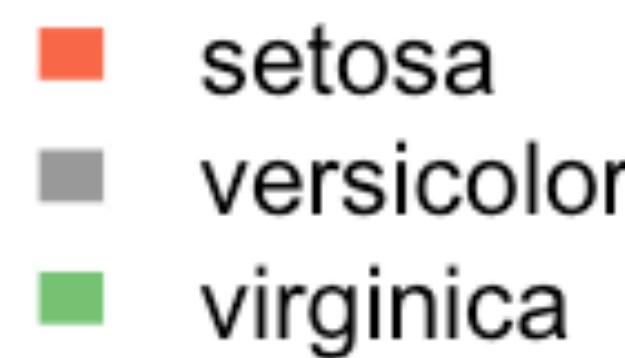
$k=3$



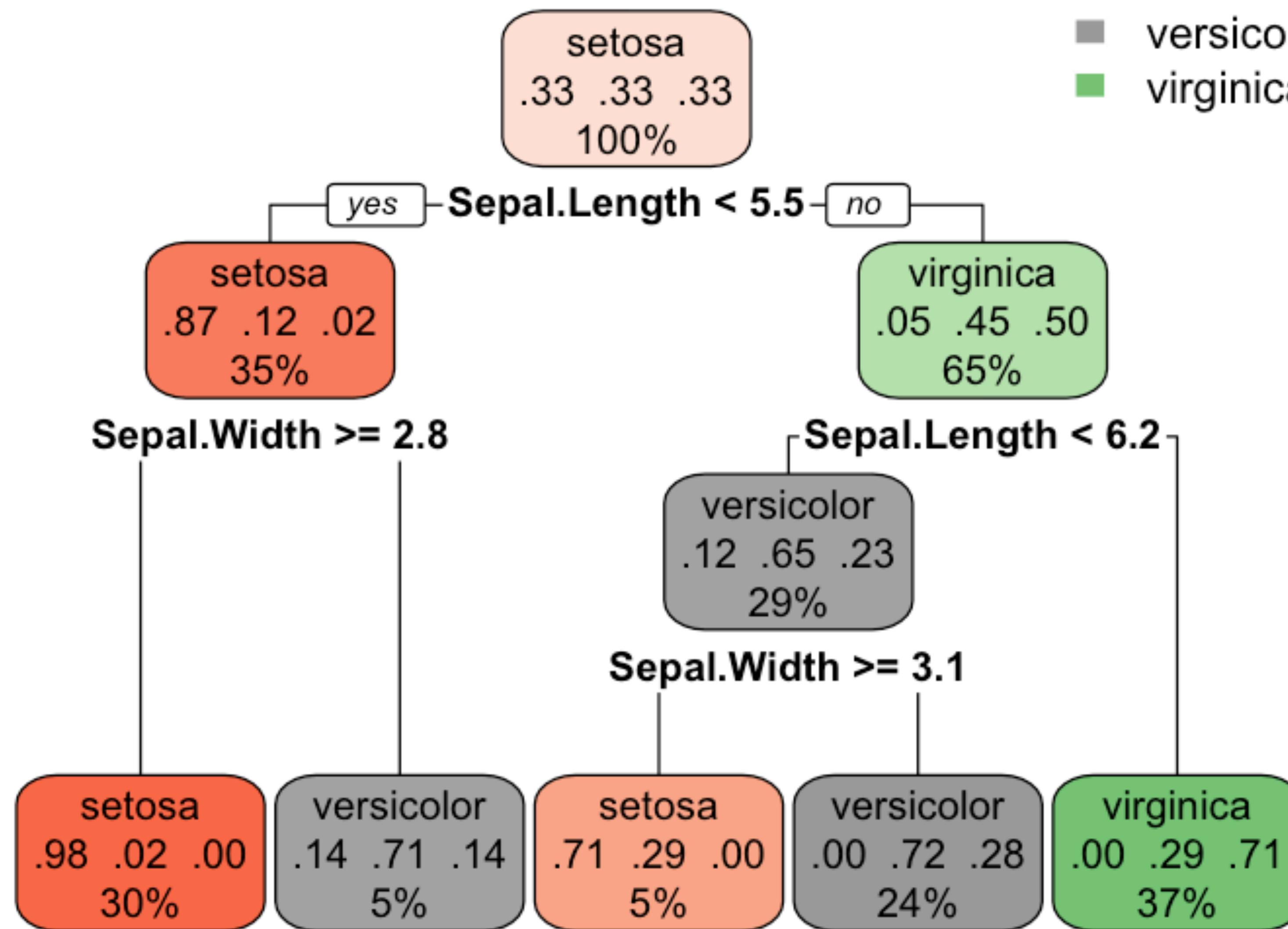
A GUIDE TO
UNDERSTANDING FLOW CHARTS
PRESENTED IN FLOW CHART FORM

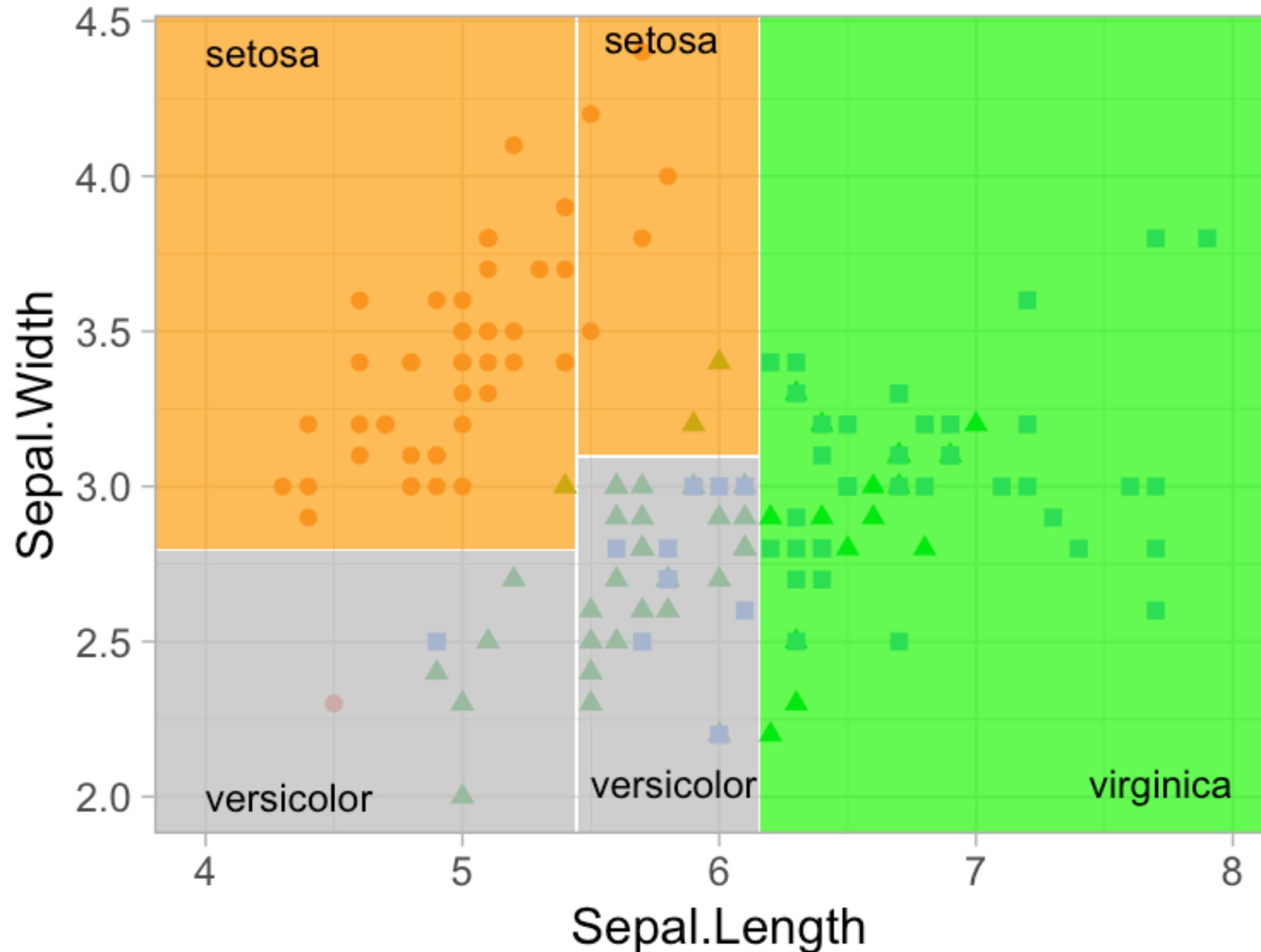






setosa
versicolor
virginica





$$\frac{1}{n_R} \sum_{i \in R} \mathbf{1}(y_i \neq \hat{y}_R)$$

Misclassification Error

$$-\sum_{k \in R} p_k \log p_k$$

(Cross) Entropy

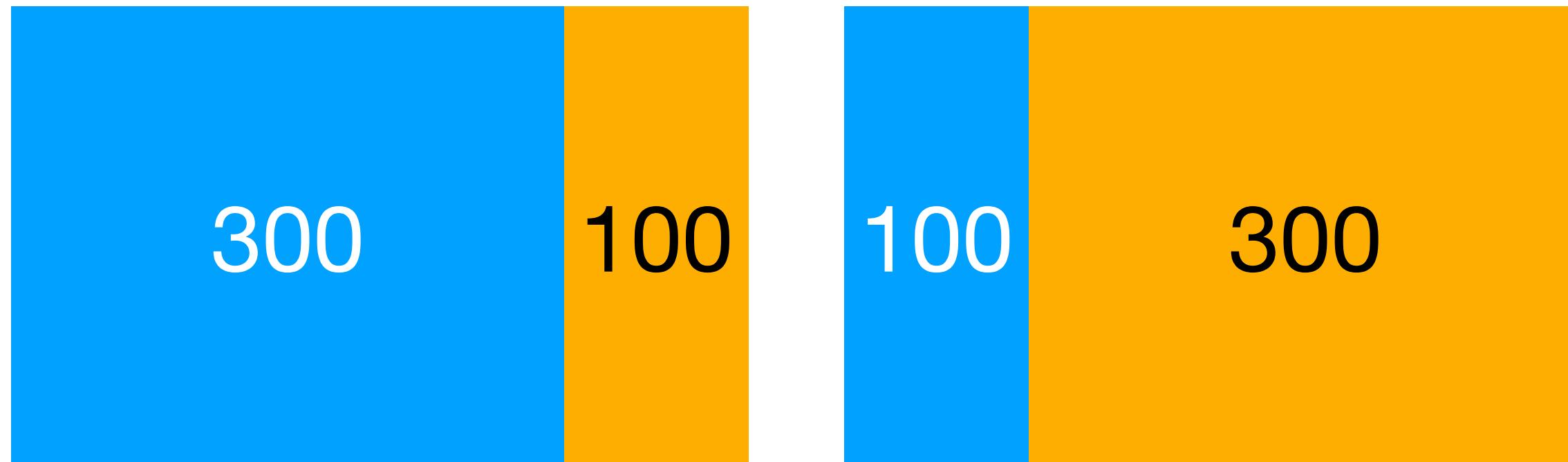
$$\sum_{k \in R} p_k(1 - p_k)$$

Gini Impurity

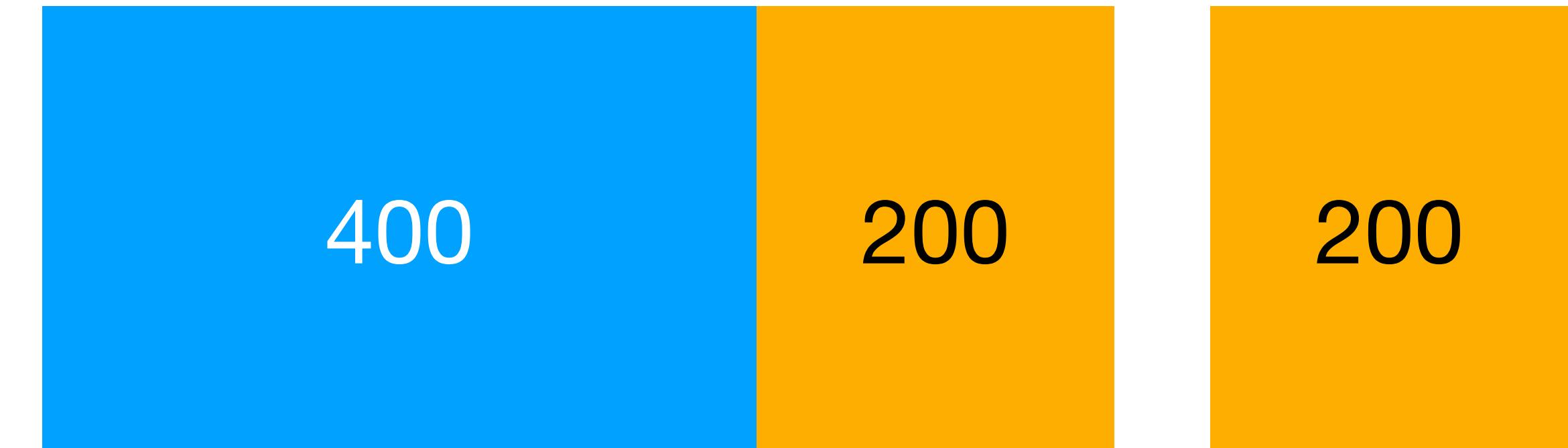
Parent



Split Option 1

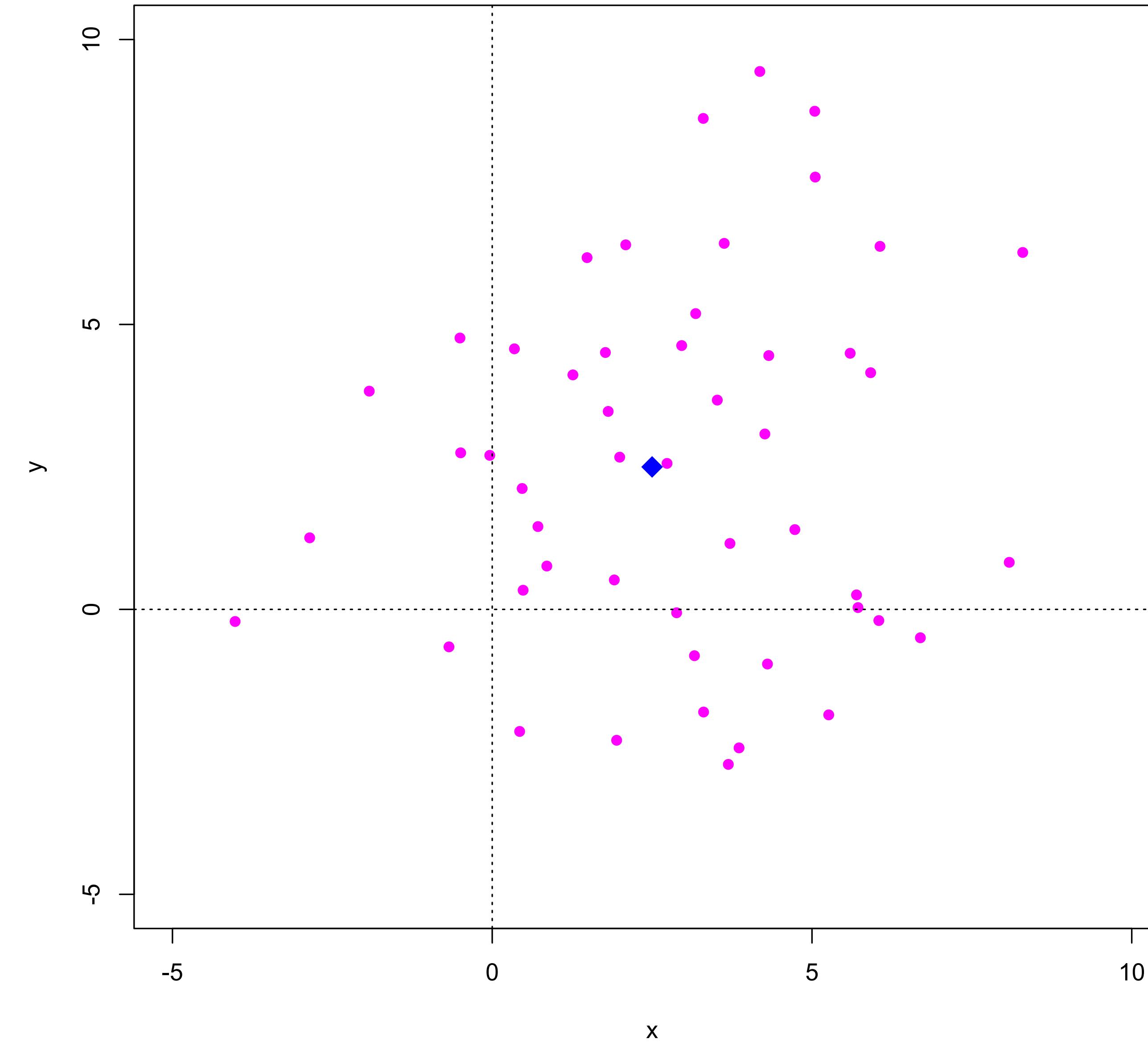


Split Option 2

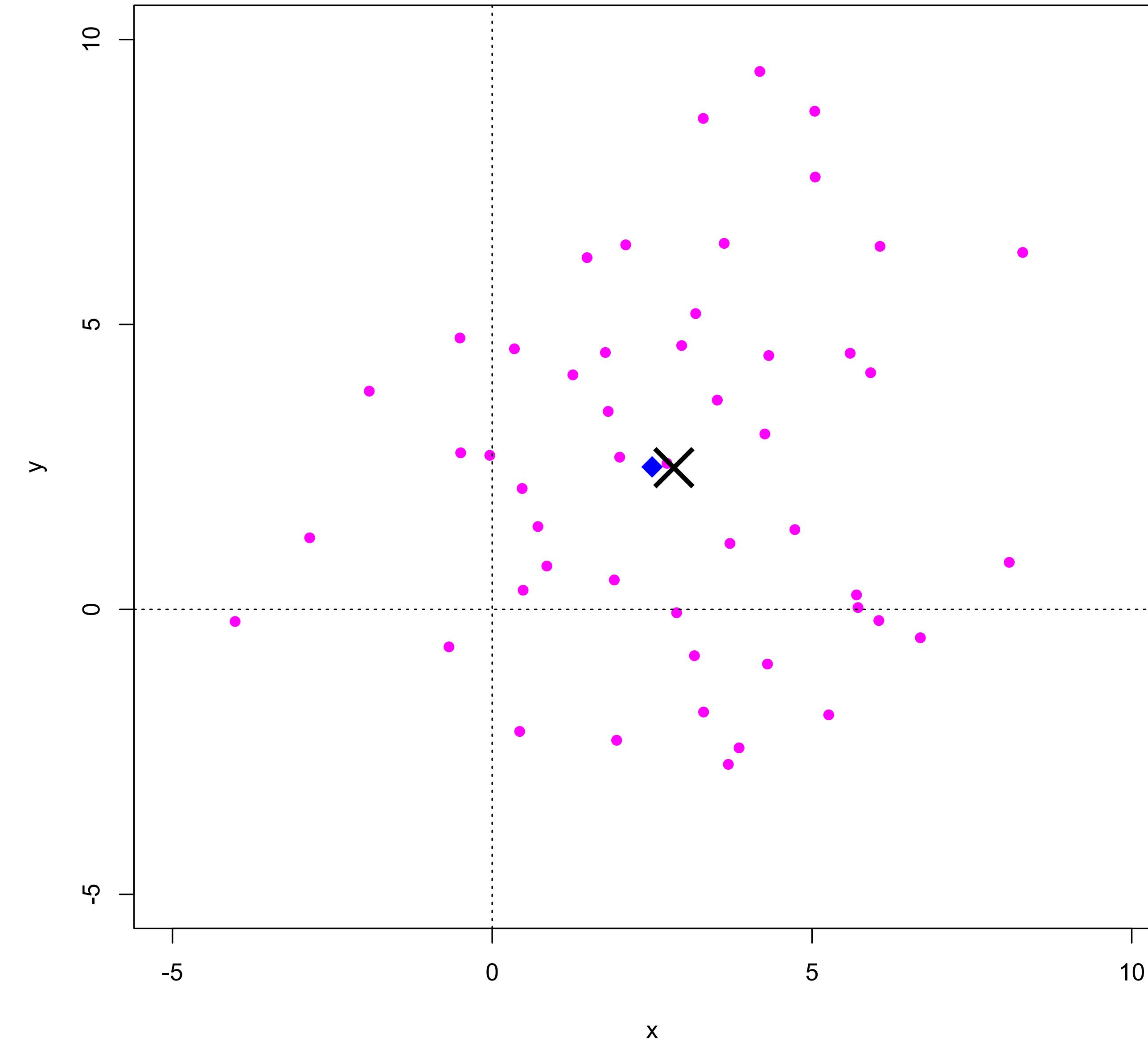




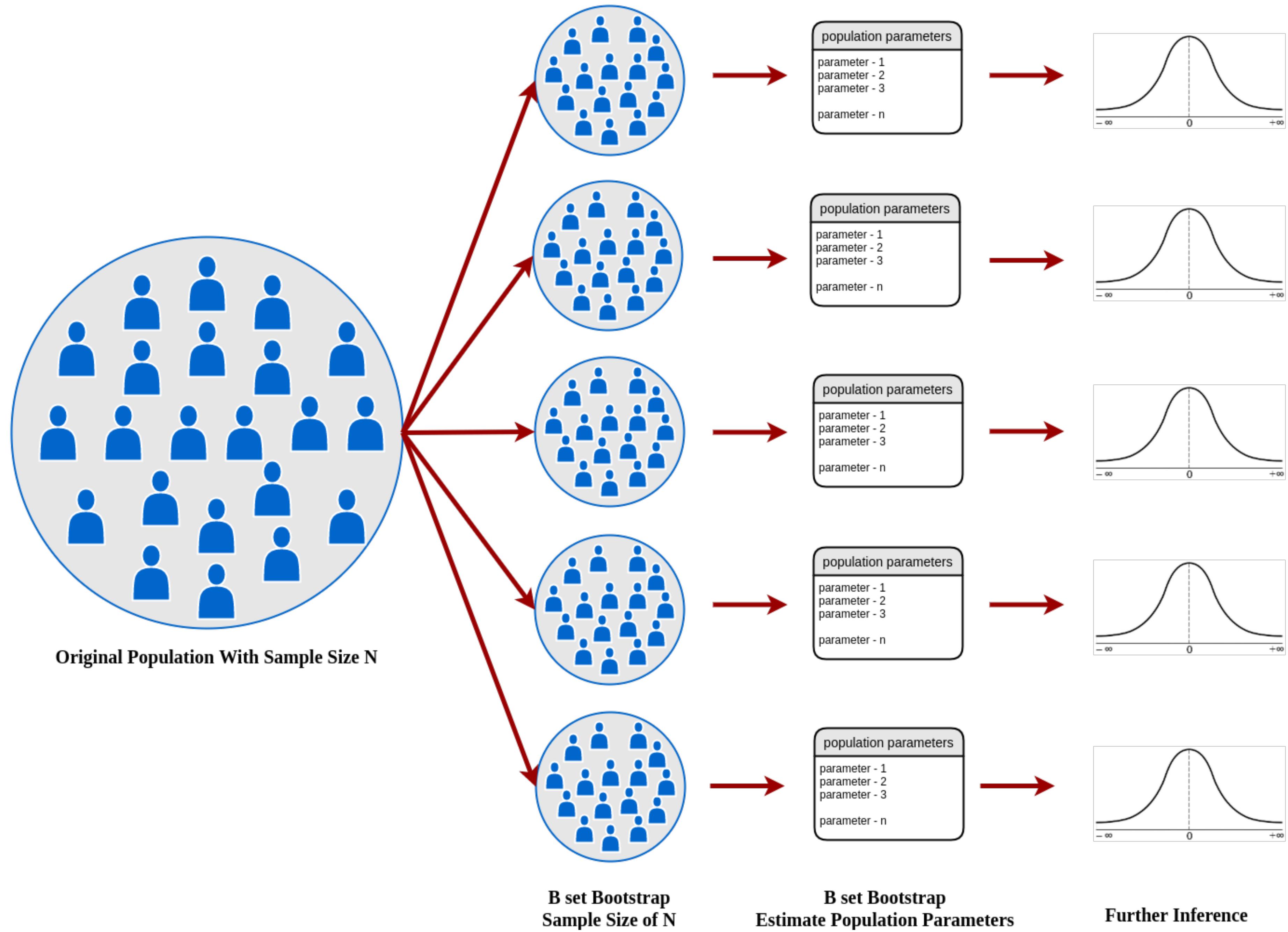
High Variance, Low Bias



High Variance, Low Bias

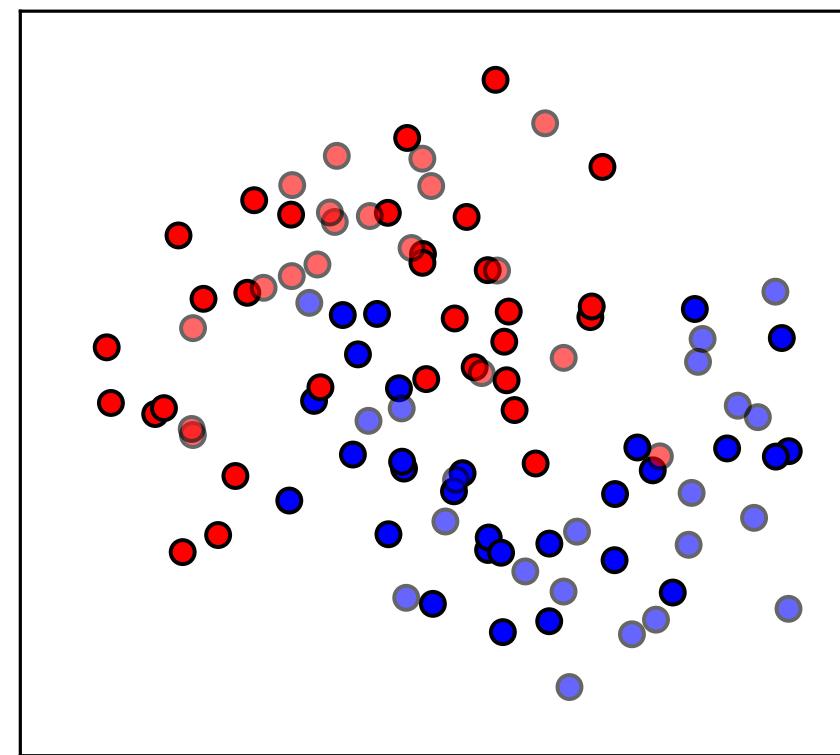




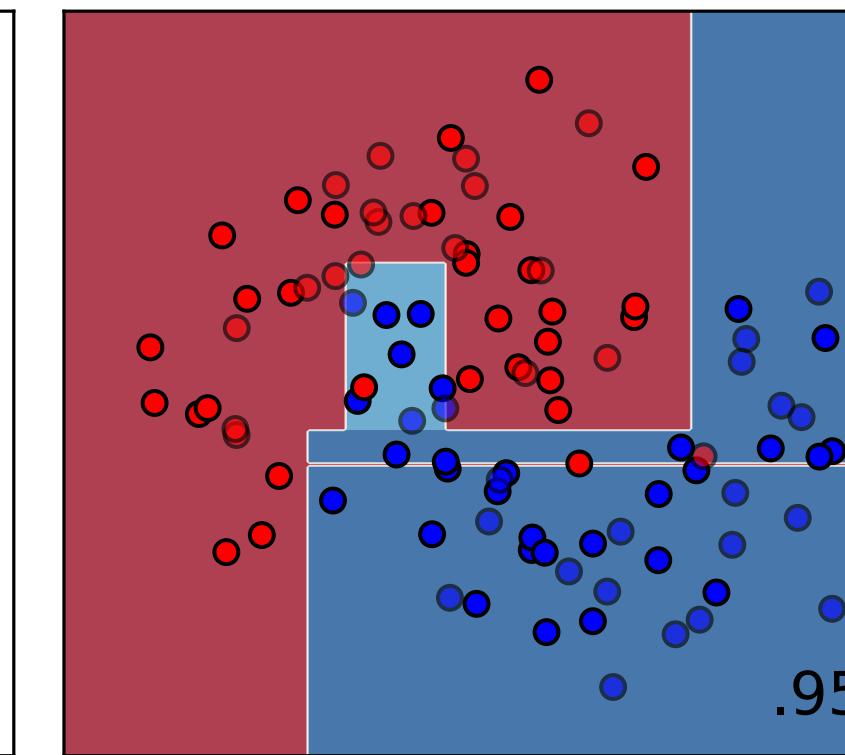




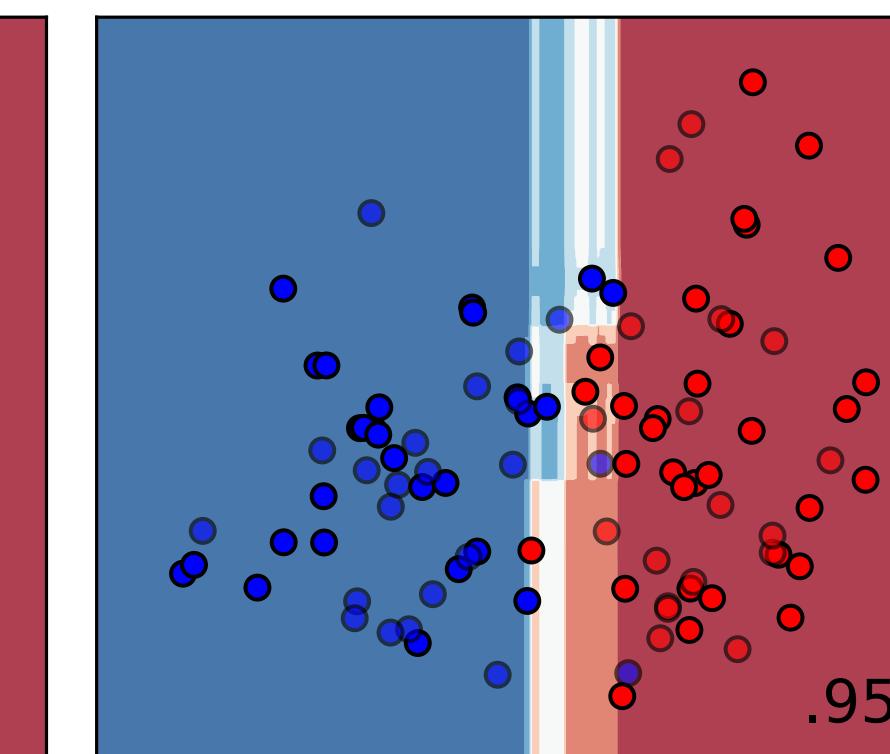
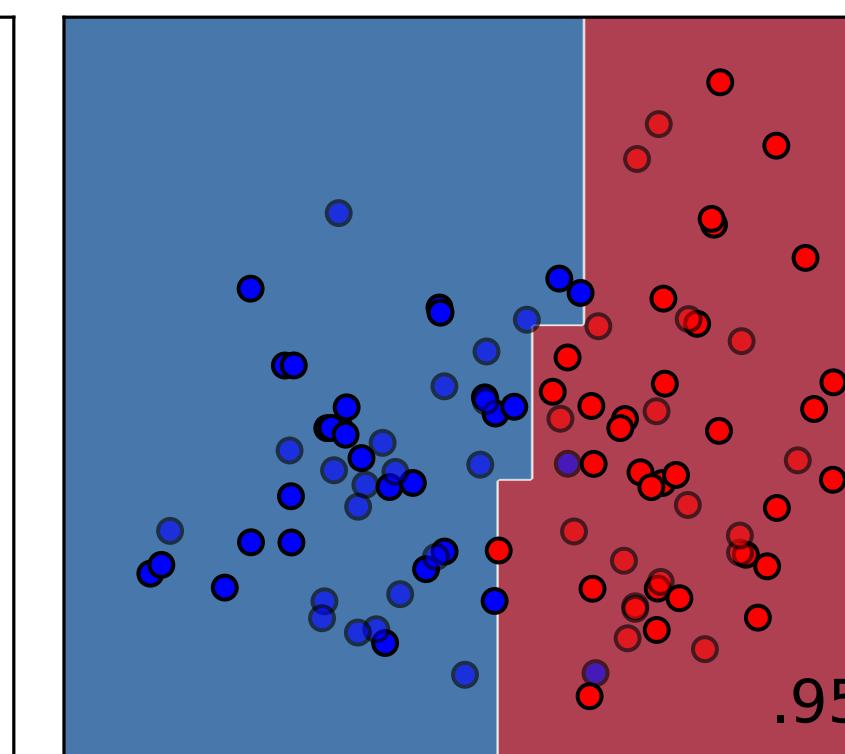
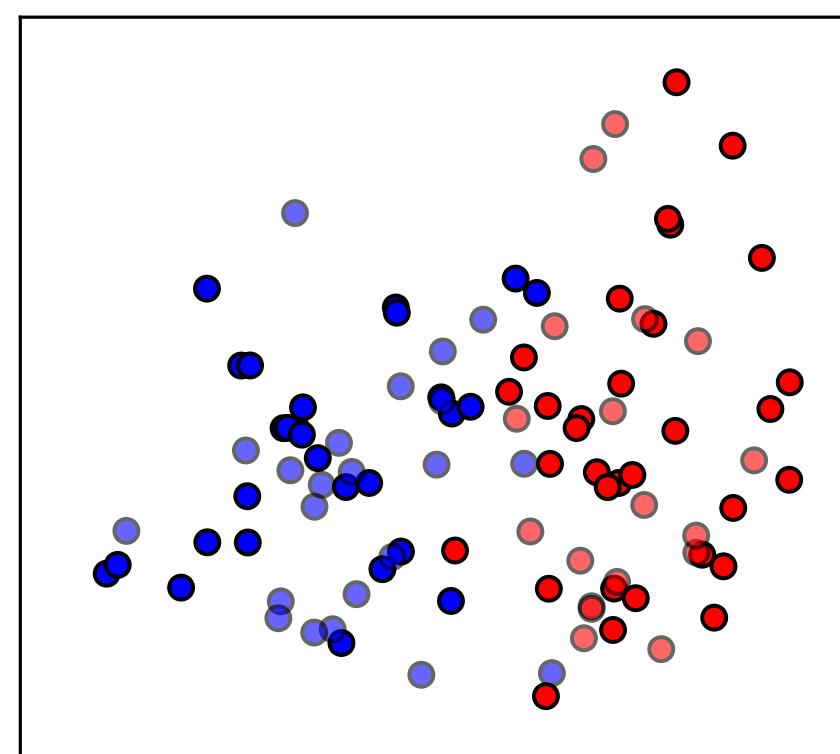
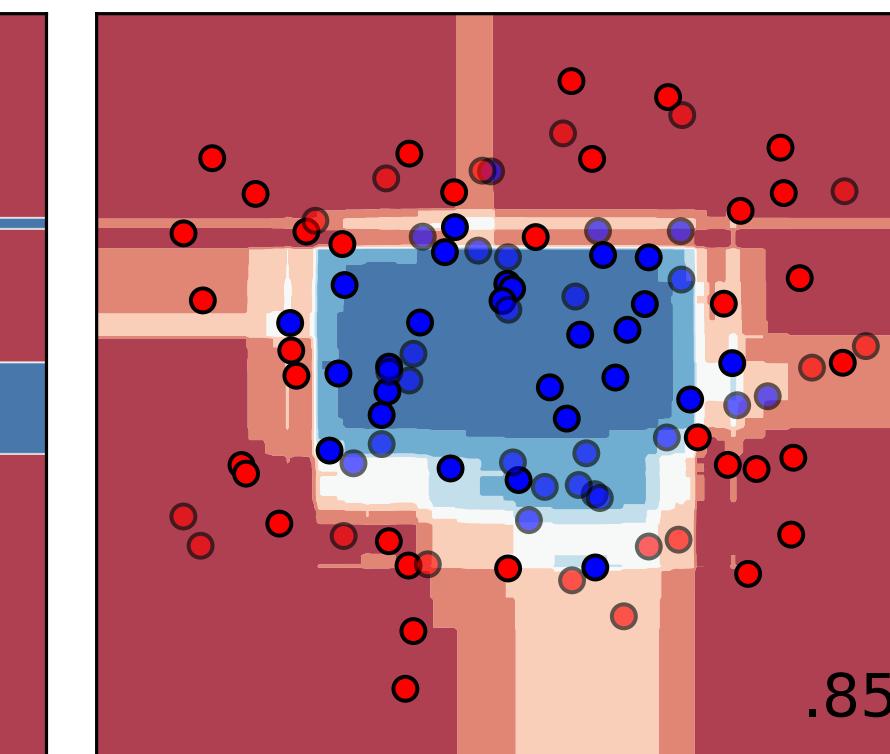
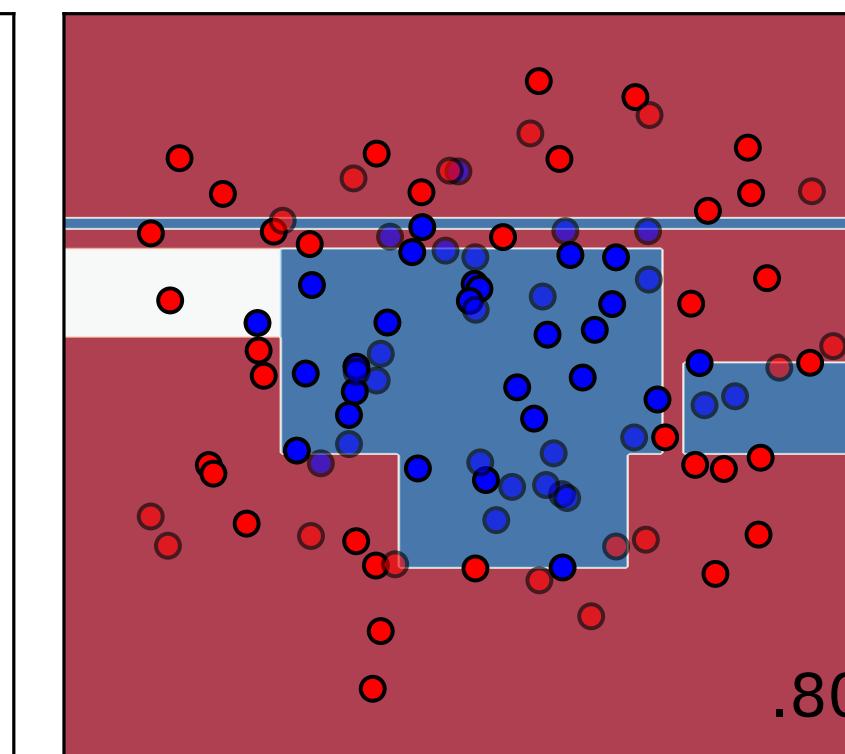
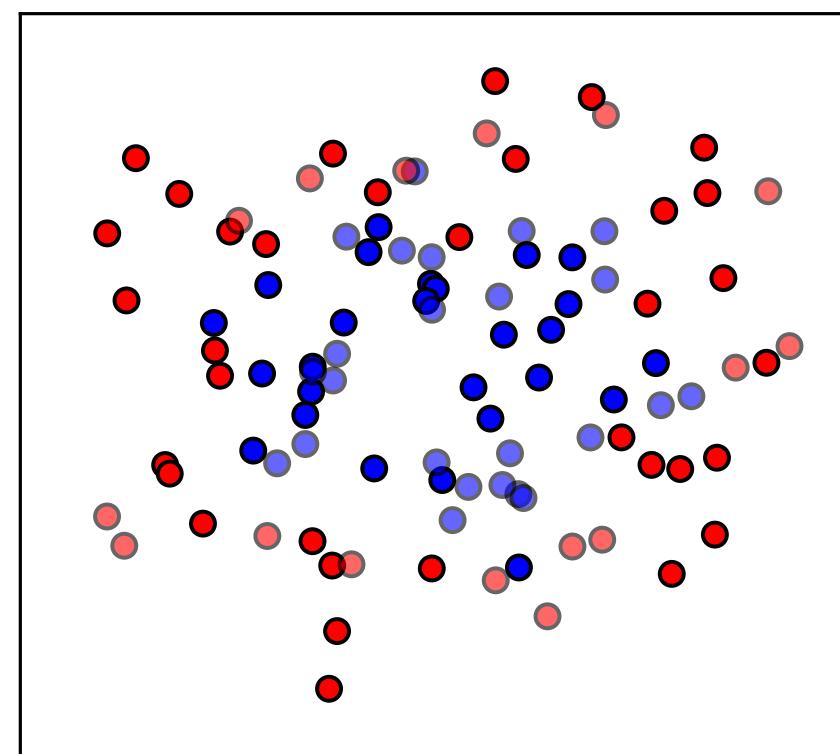
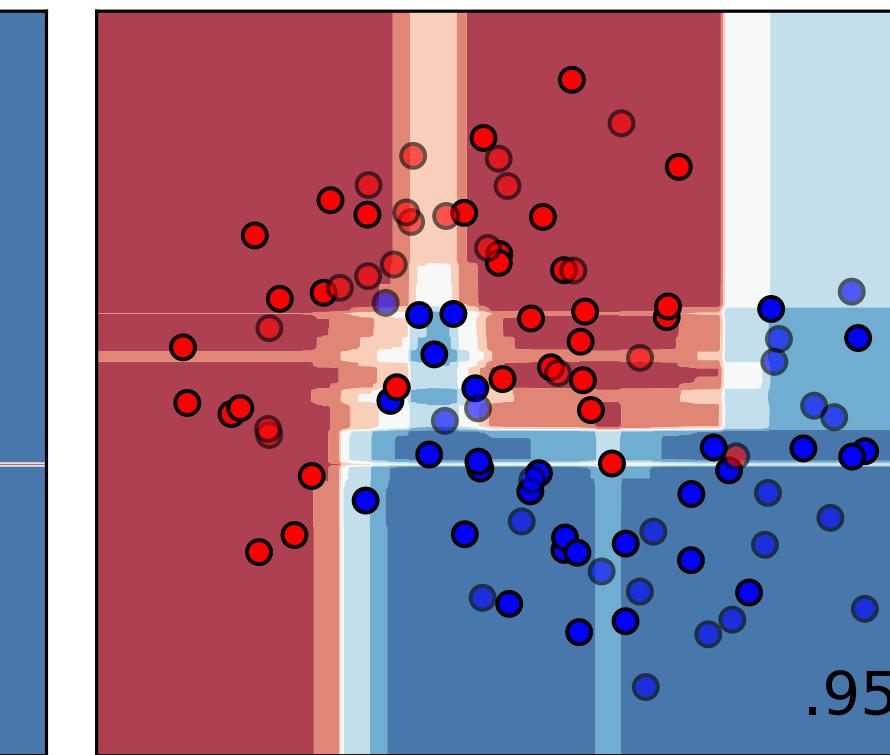
Input data

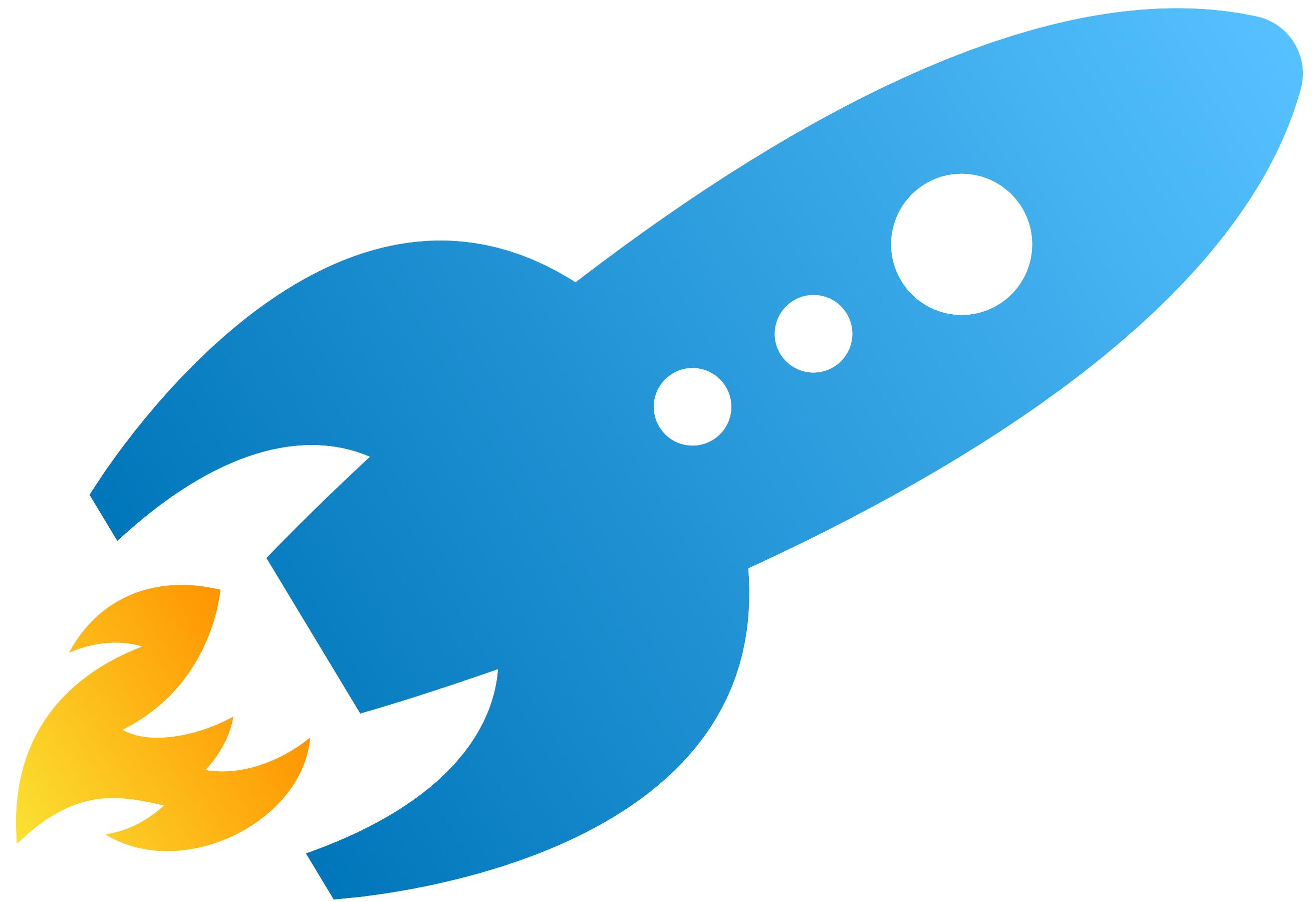


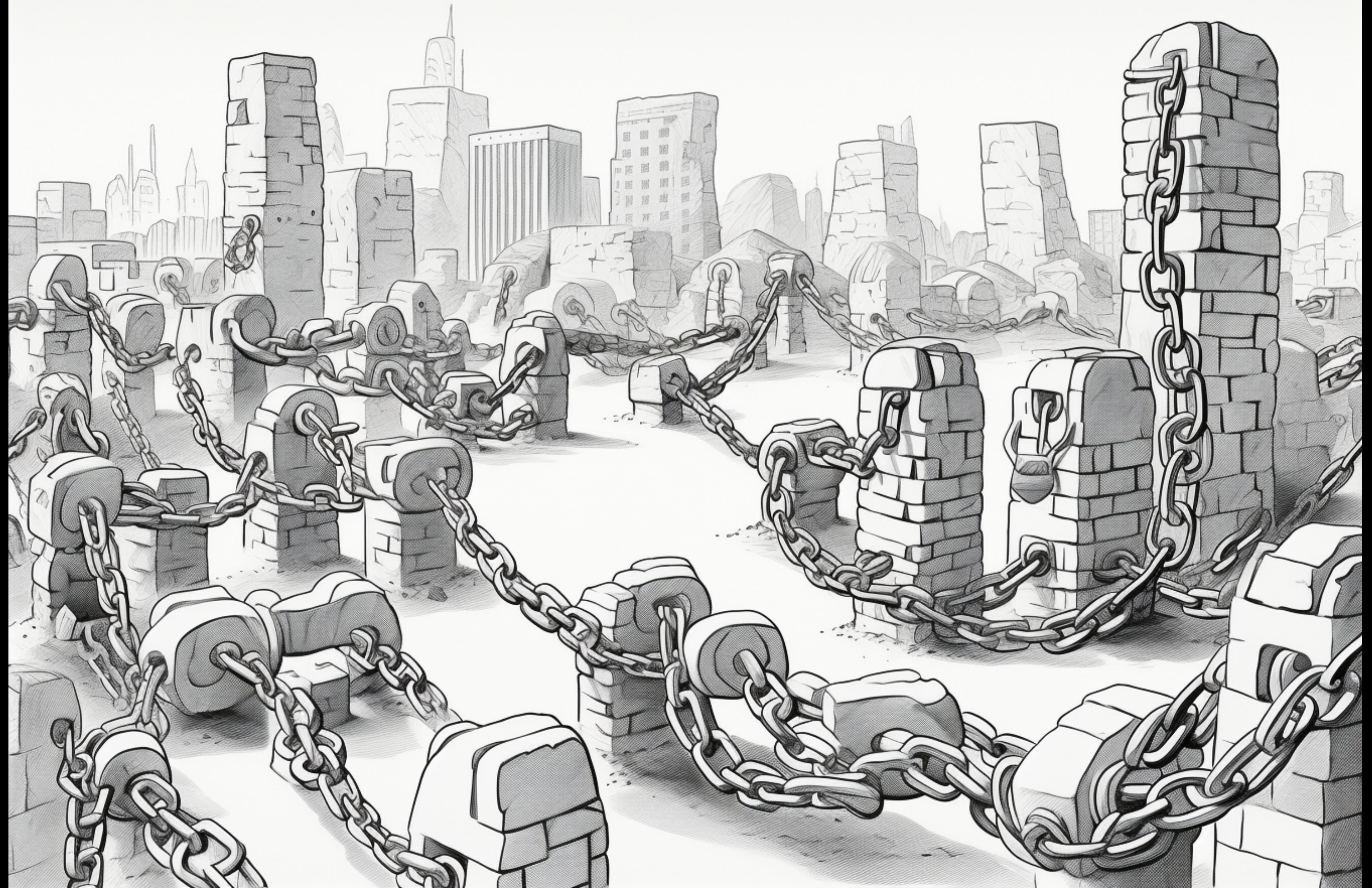
Decision Tree



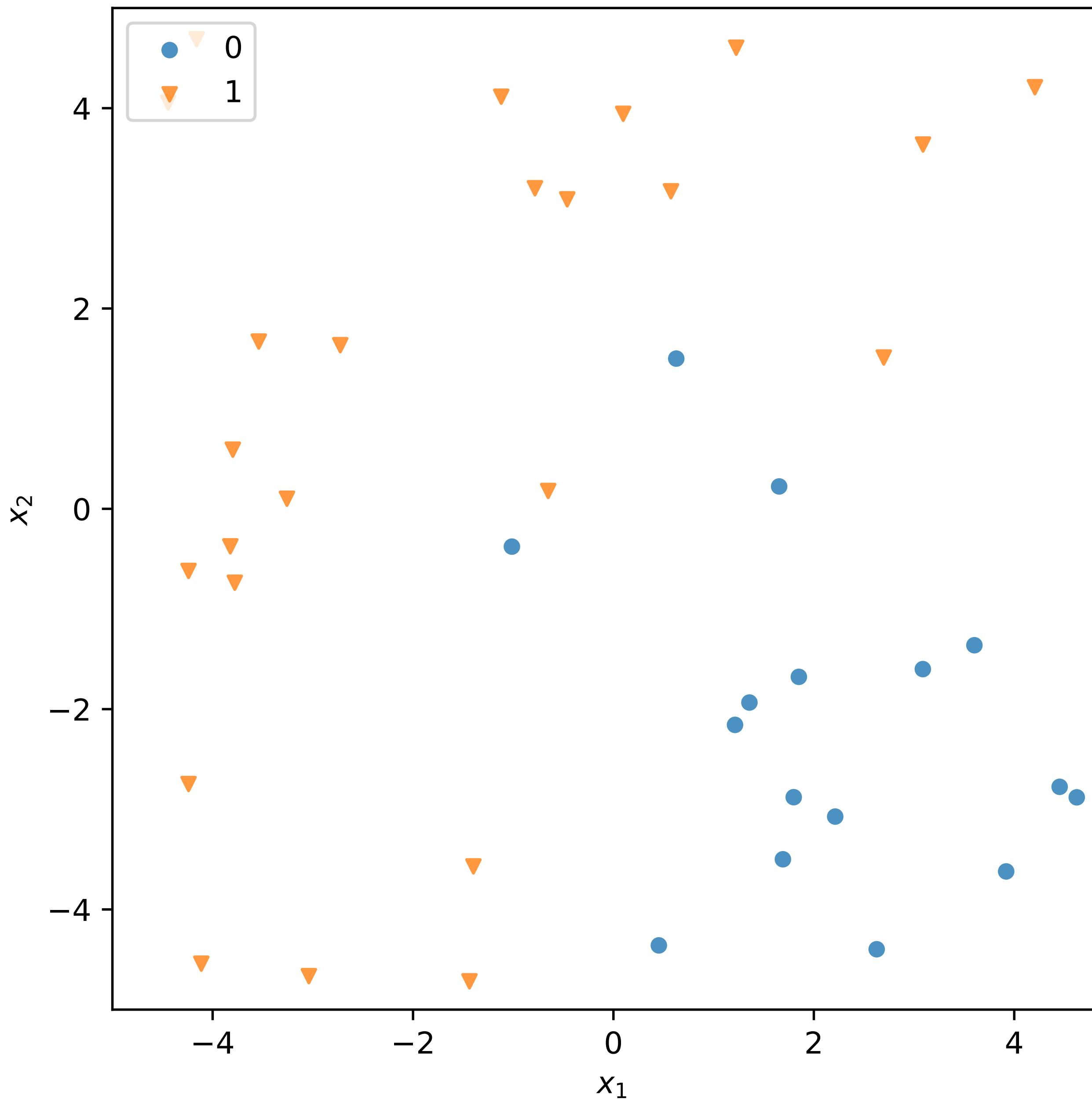
Random Forest



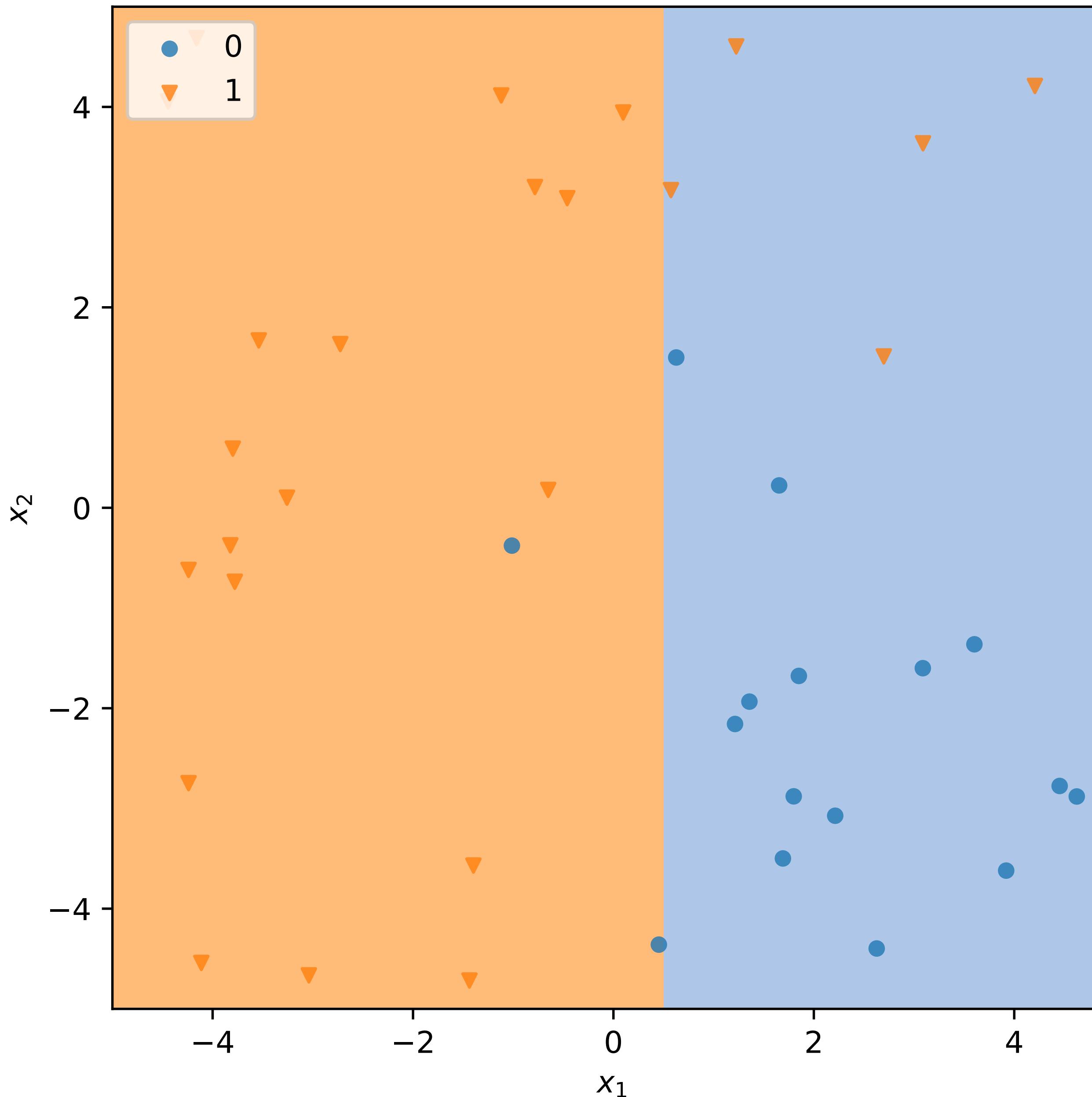




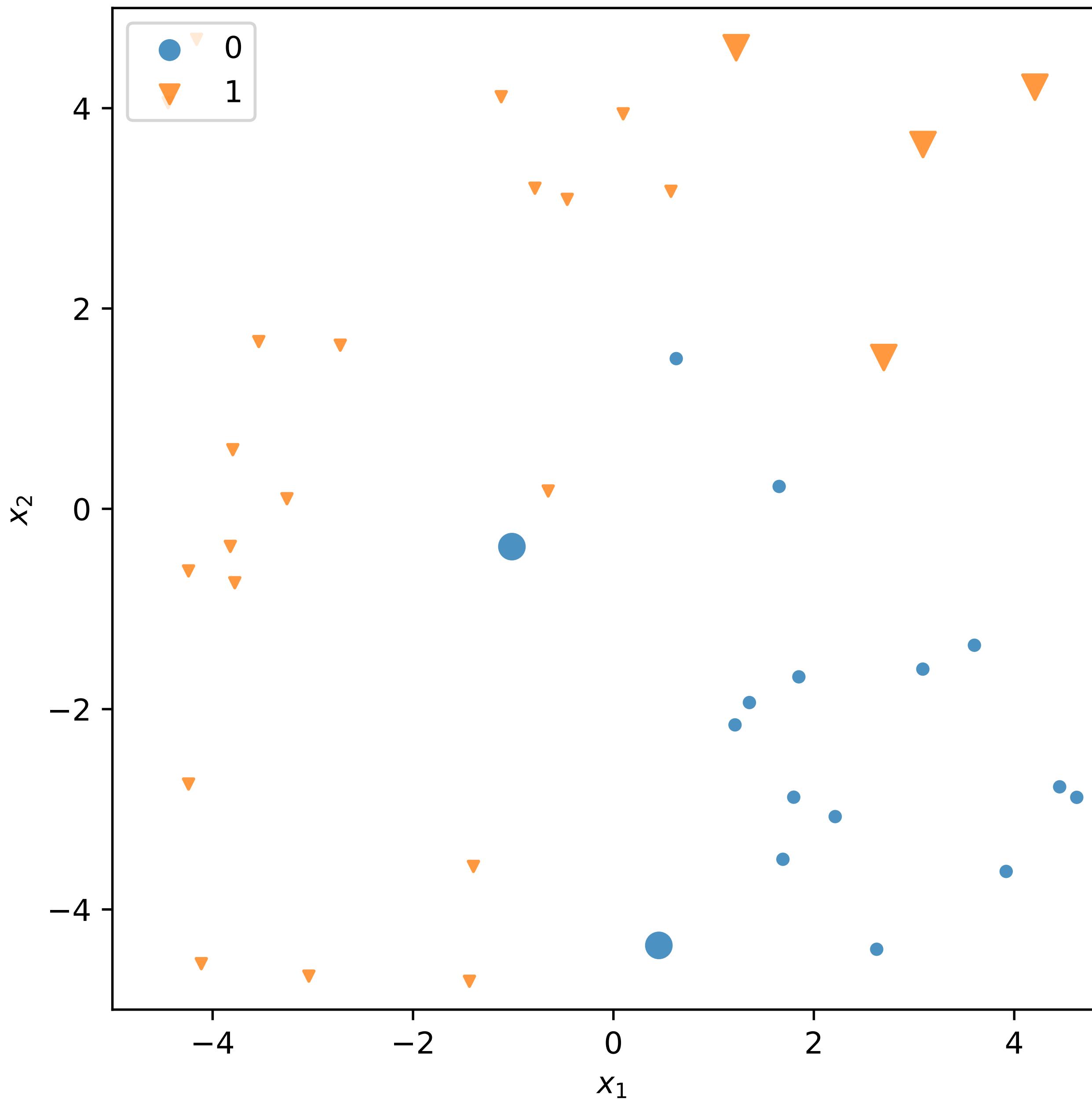
Weighted Samples 1



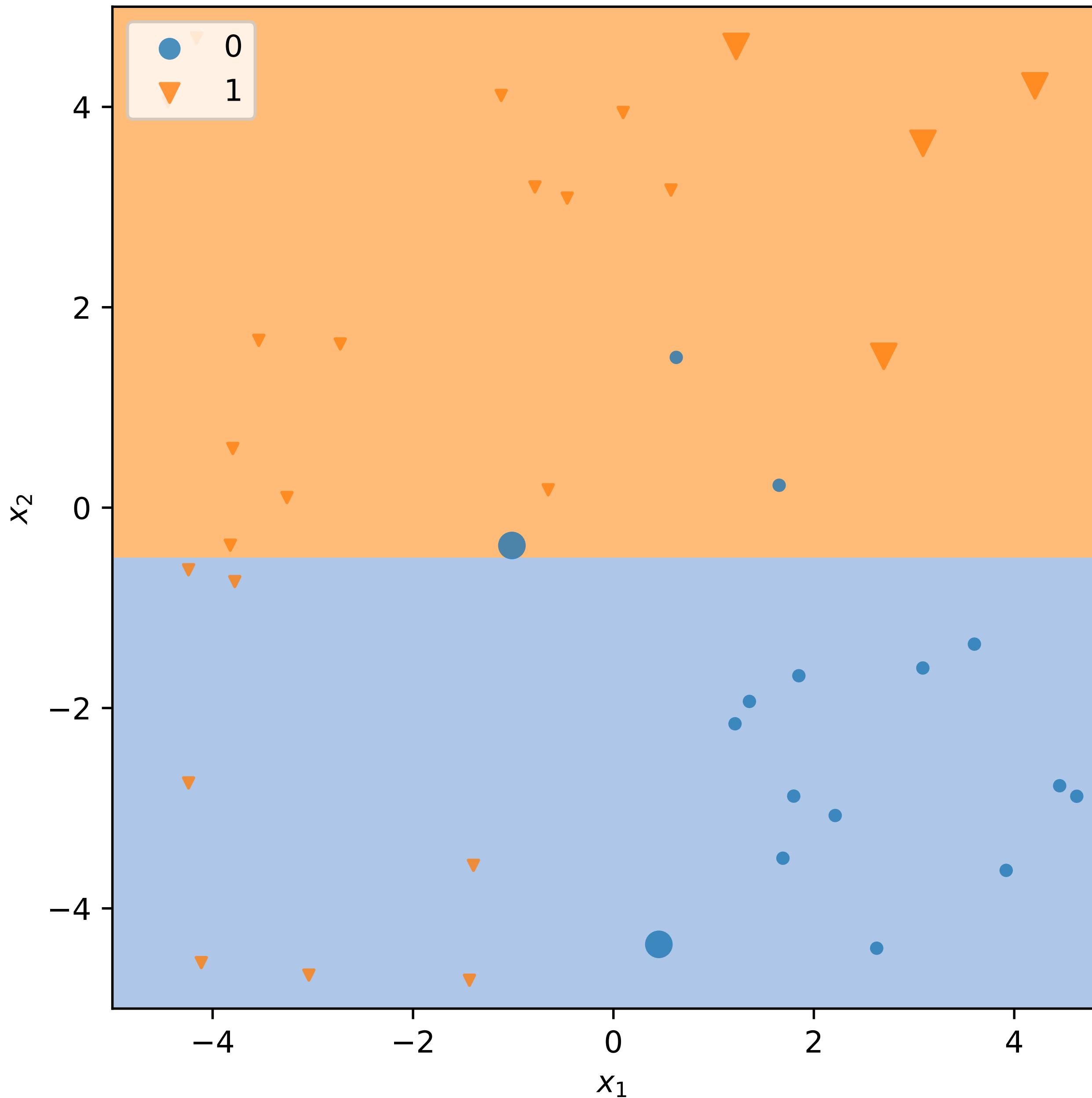
Decision Stump 1



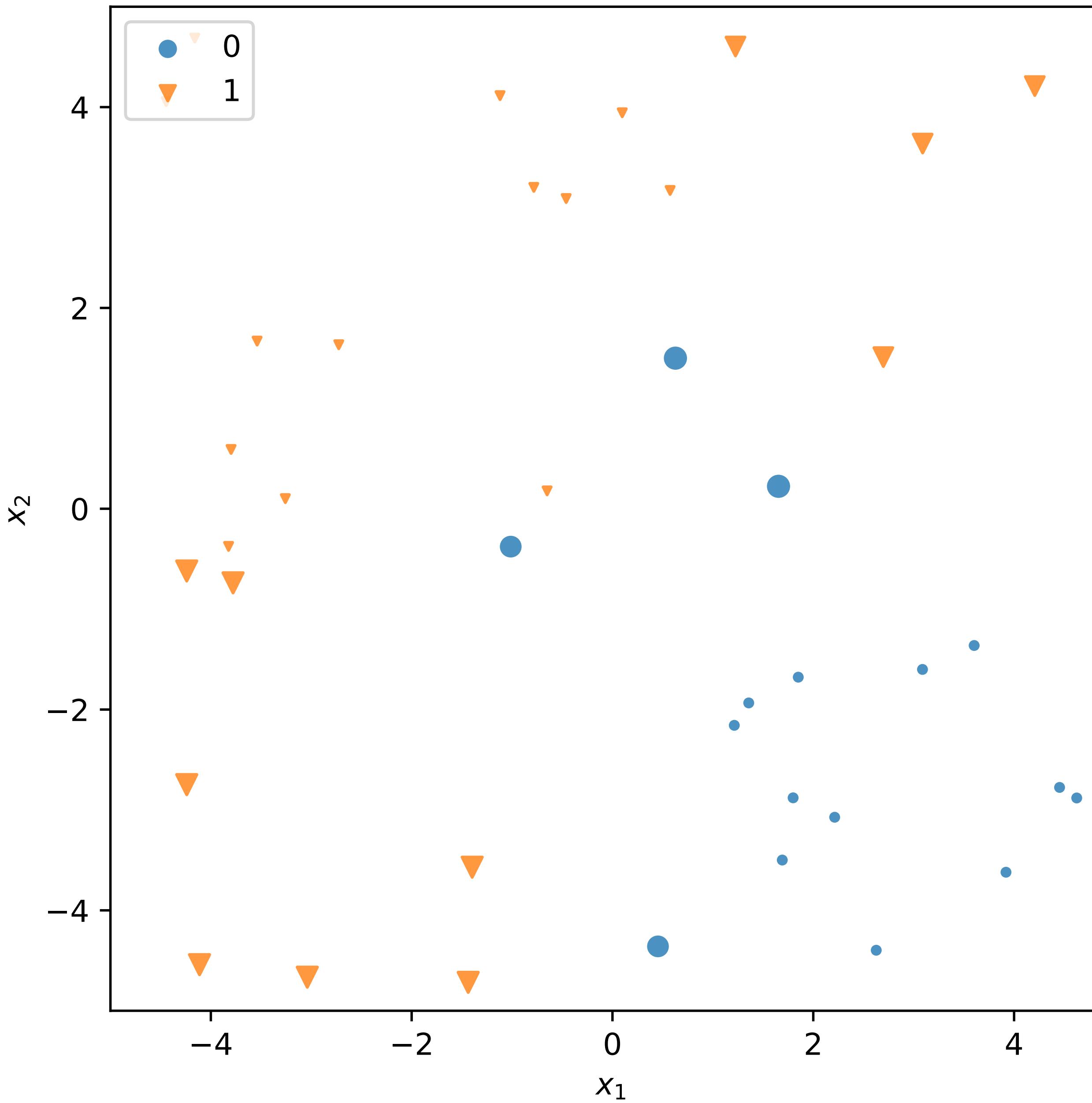
Weighted Samples 2



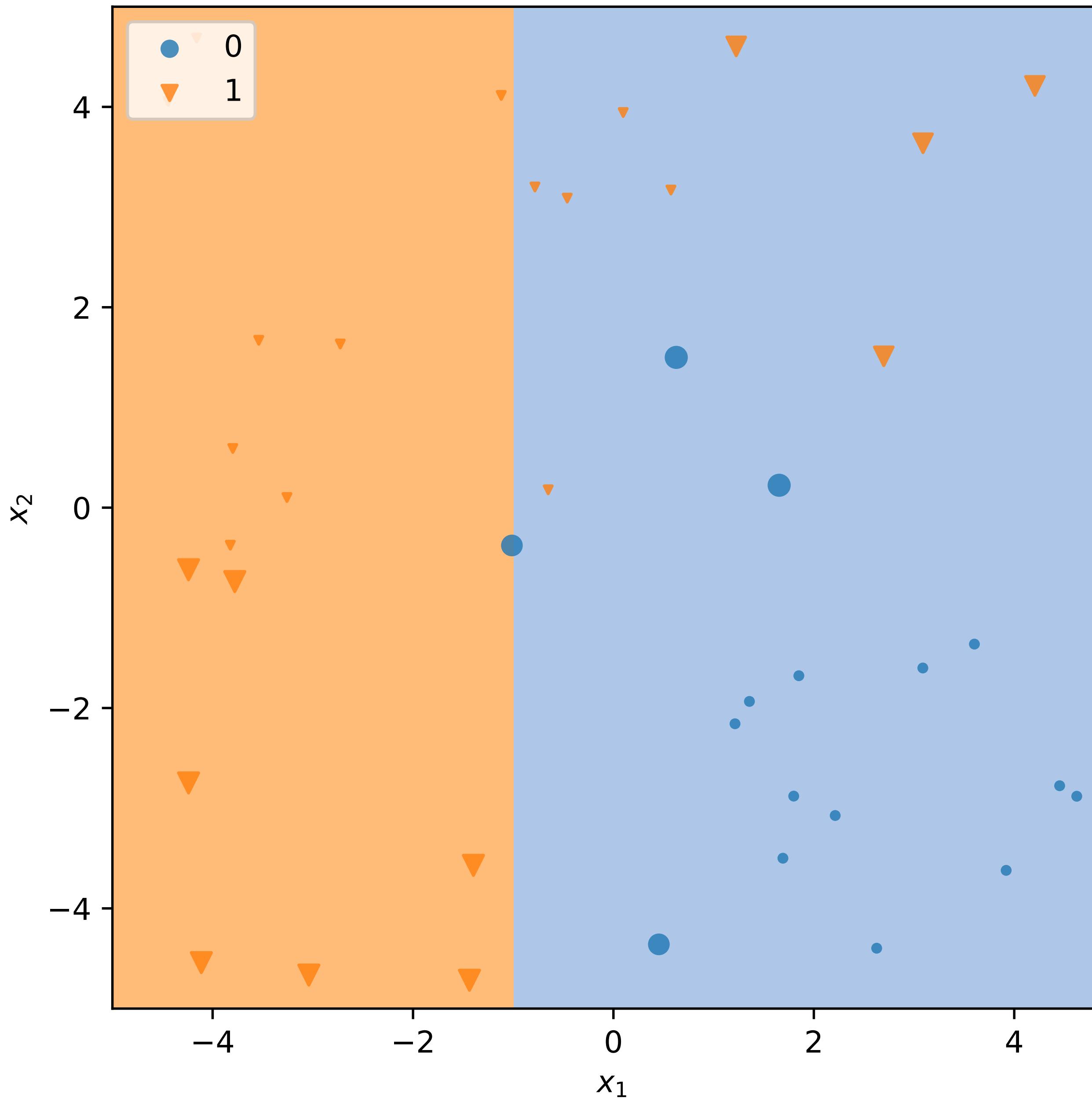
Decision Stump 2



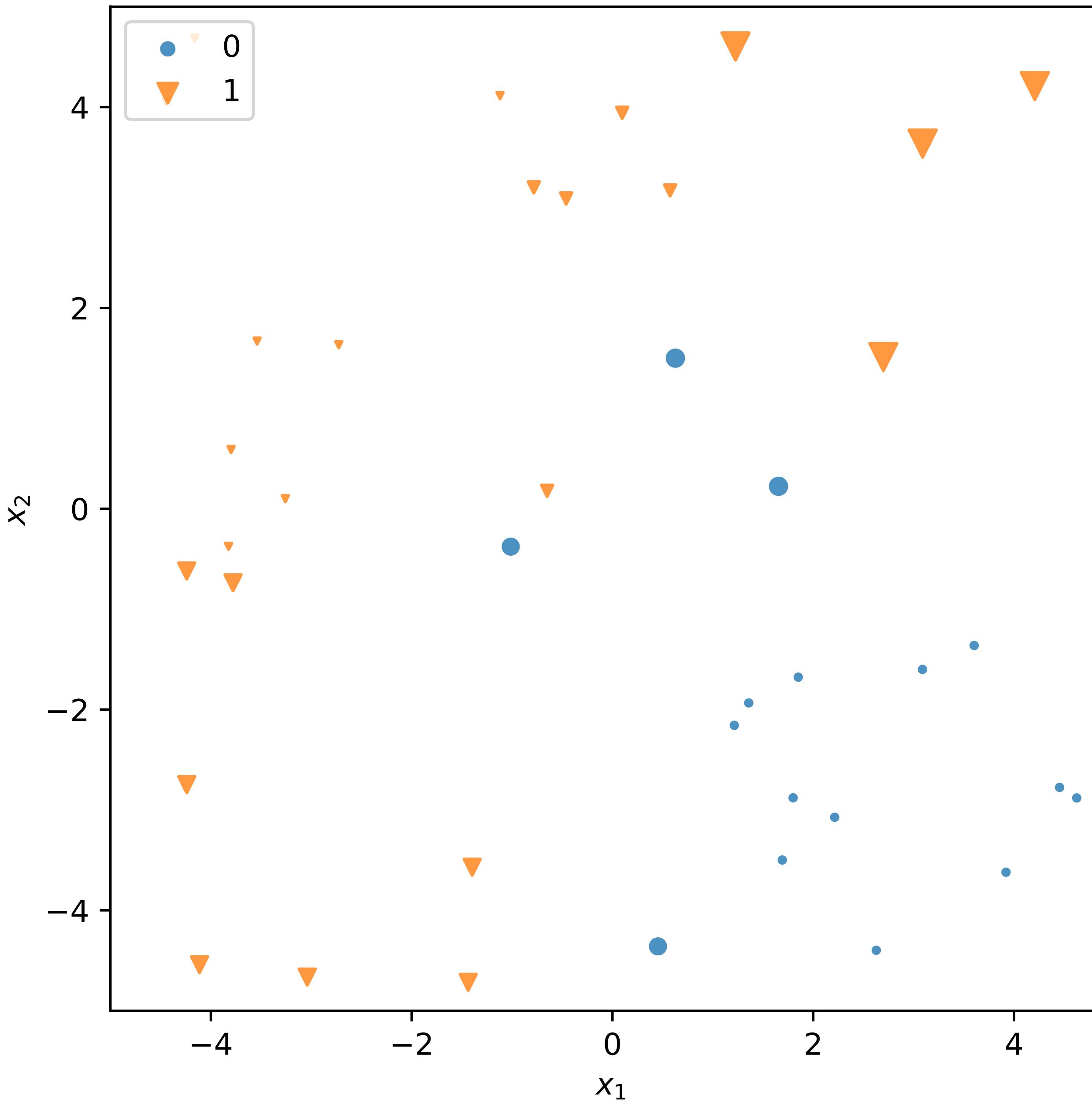
Weighted Samples 3



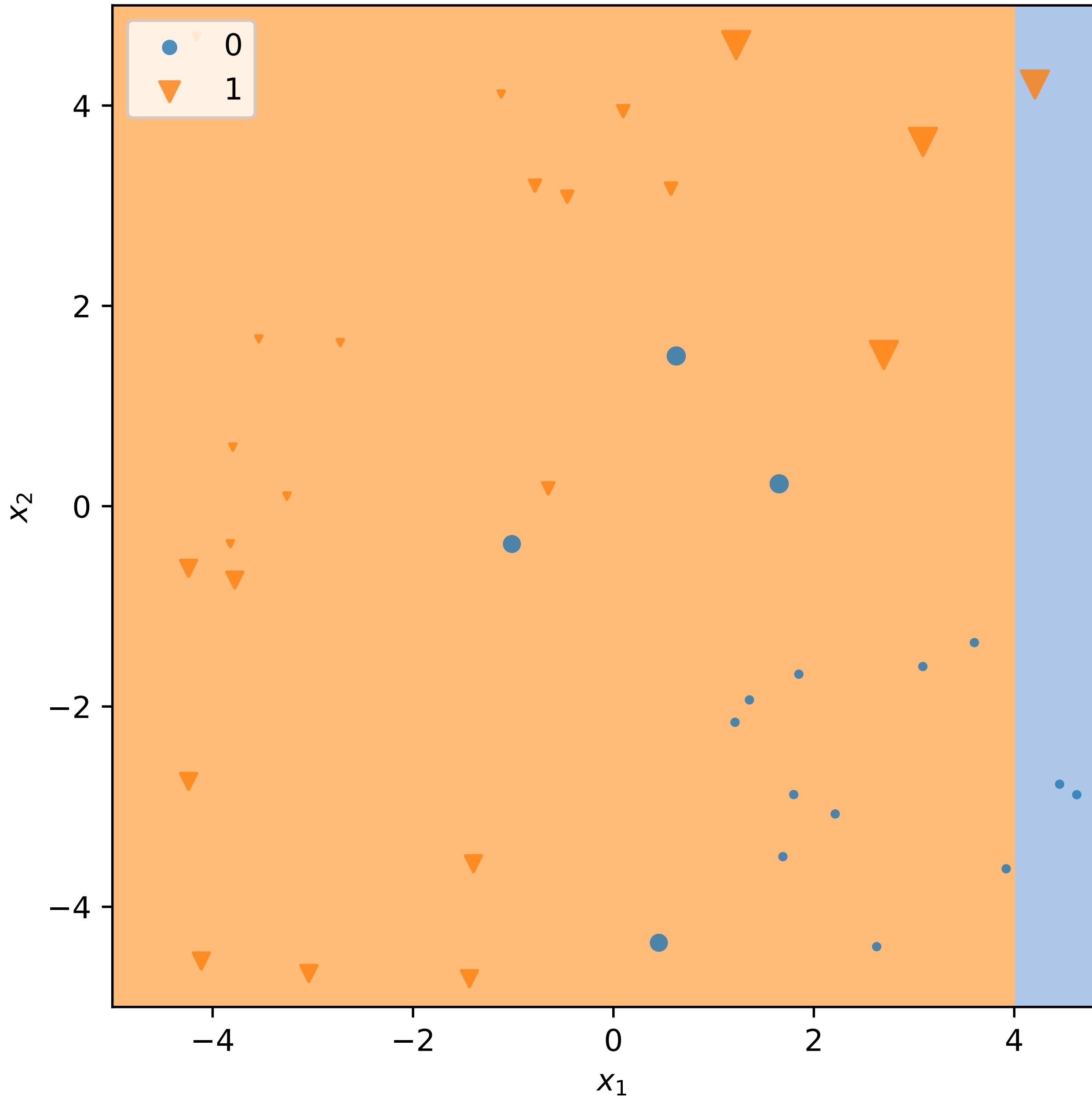
Decision Stump 3



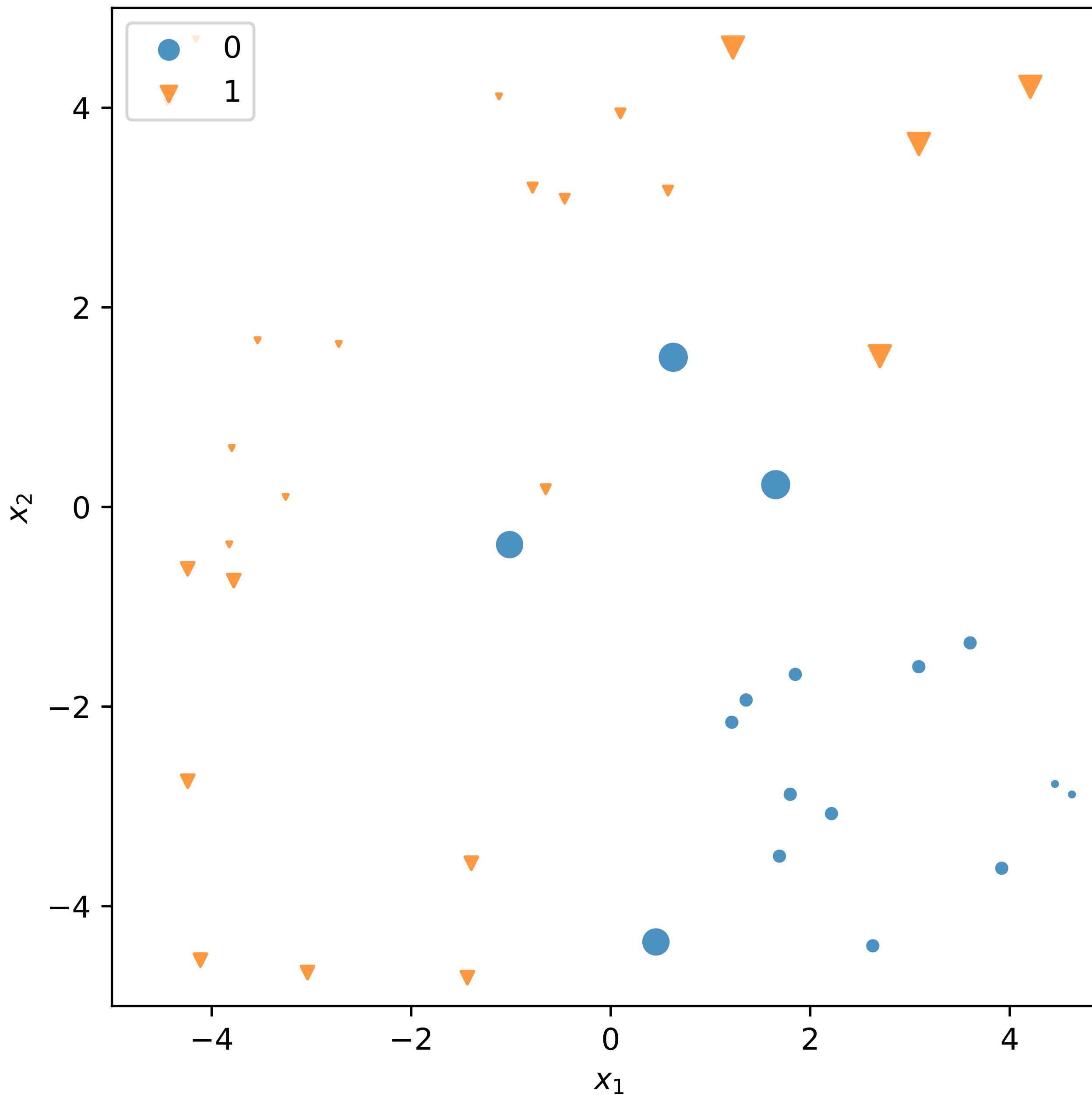
Weighted Samples 4



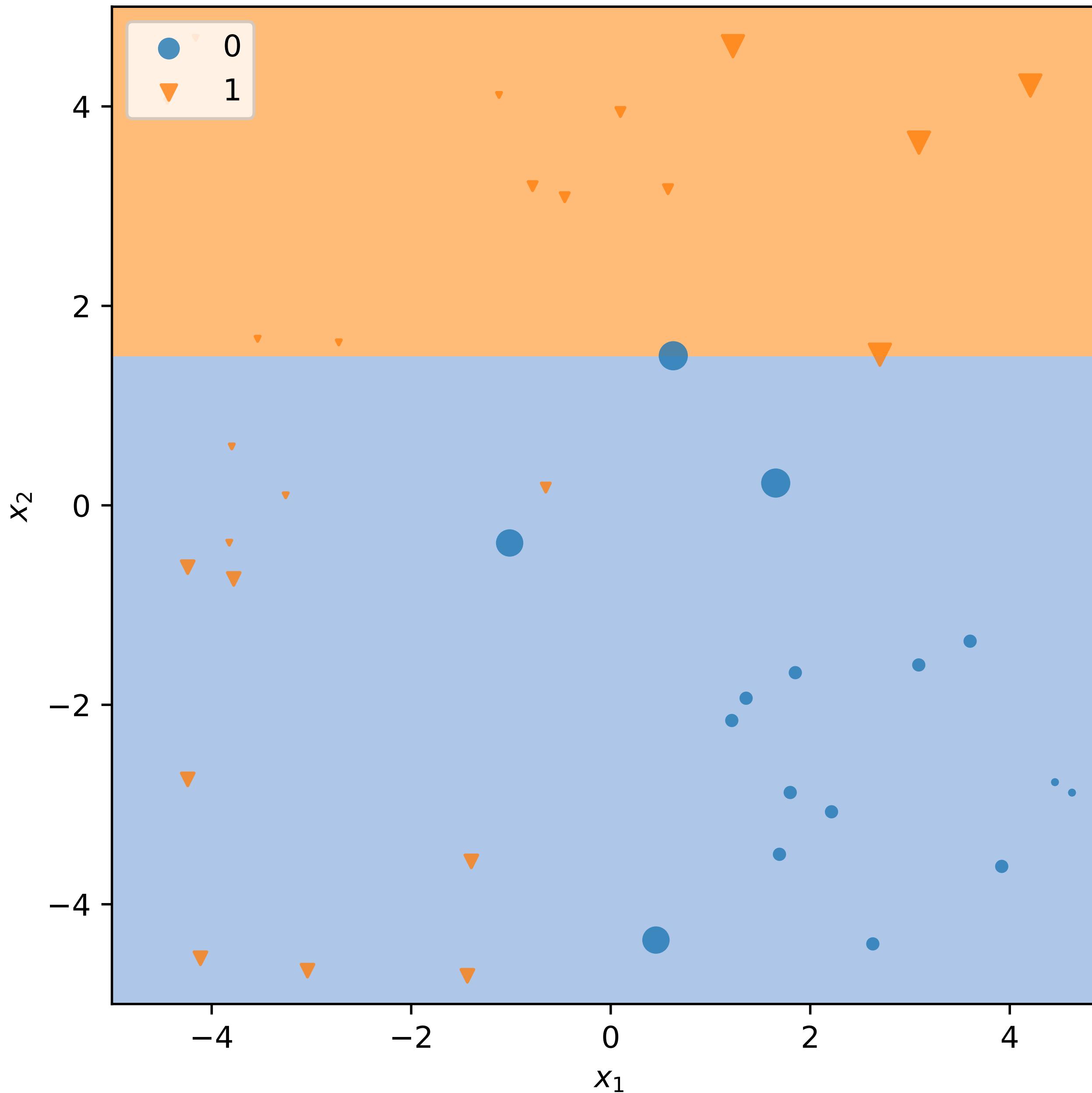
Decision Stump 4



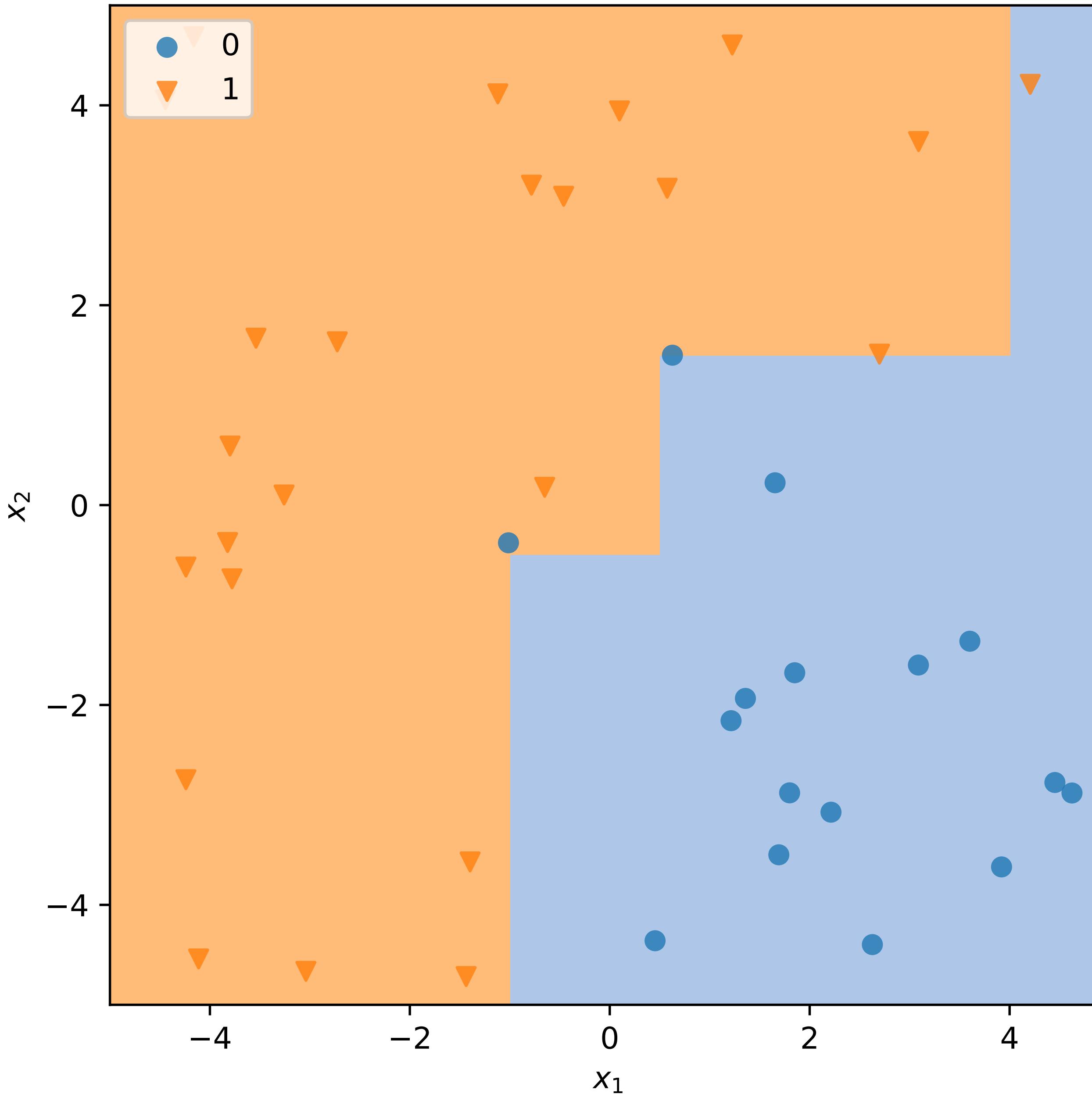
Weighted Samples 5



Decision Stump 5

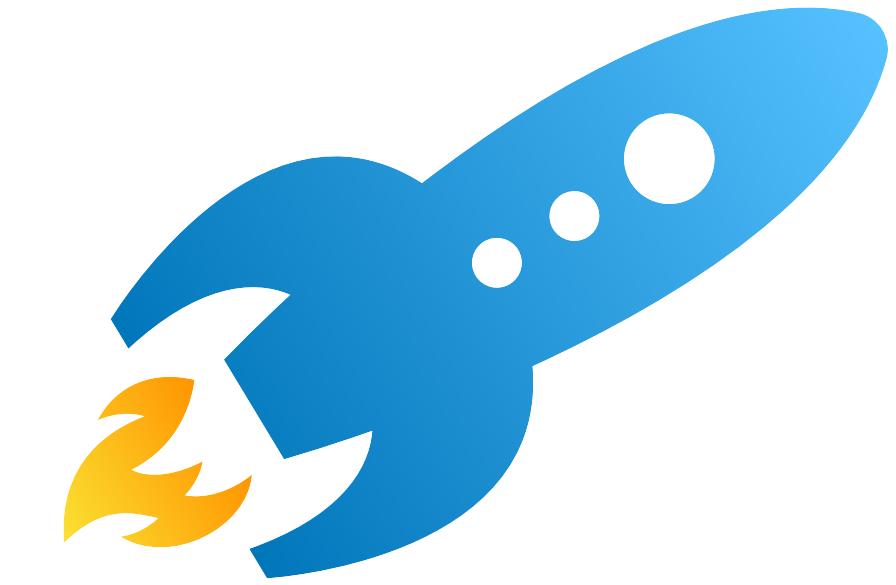


AdaBoost Combined





strong learners
parallel
reduce variance



weak learners
sequential
increase accuracy

Questions?

Next week: Linear Models Revisited

