CAS KAGGLE
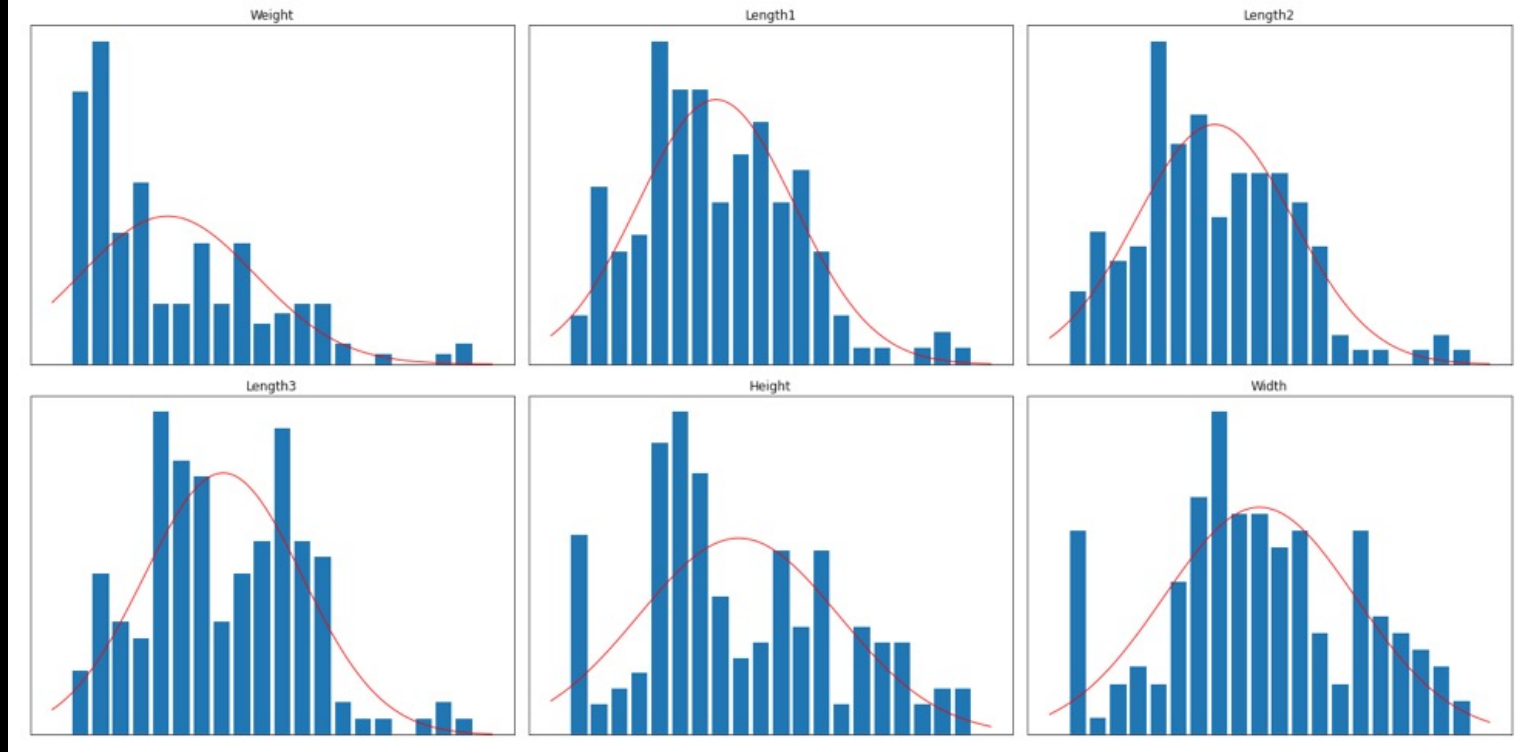
FISH MARKET

AUTOR: Lucía Carmona

NIU: 1606605

INDEX

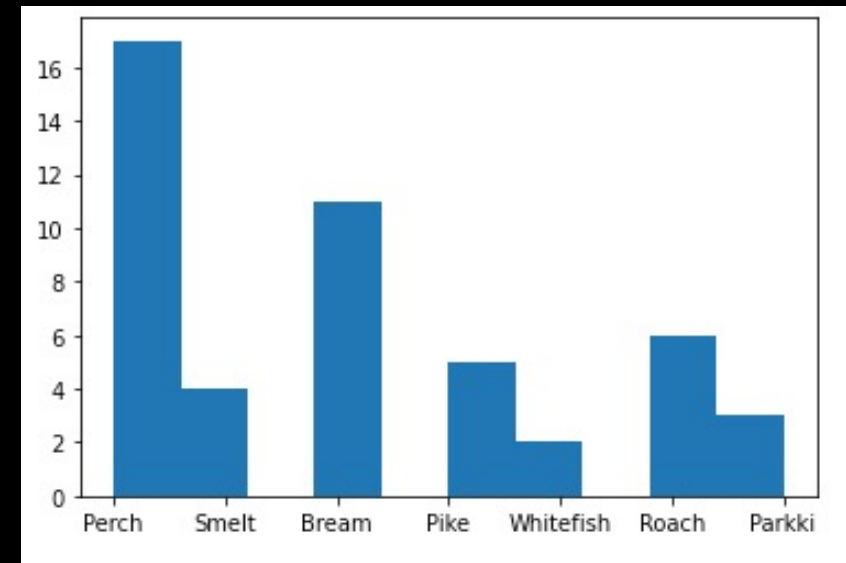# INTRODUCCIÓ I OBJECTIU
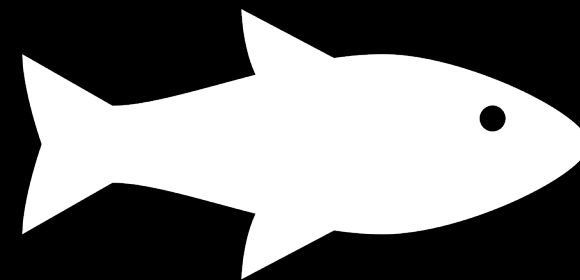
Millorar el Dataset
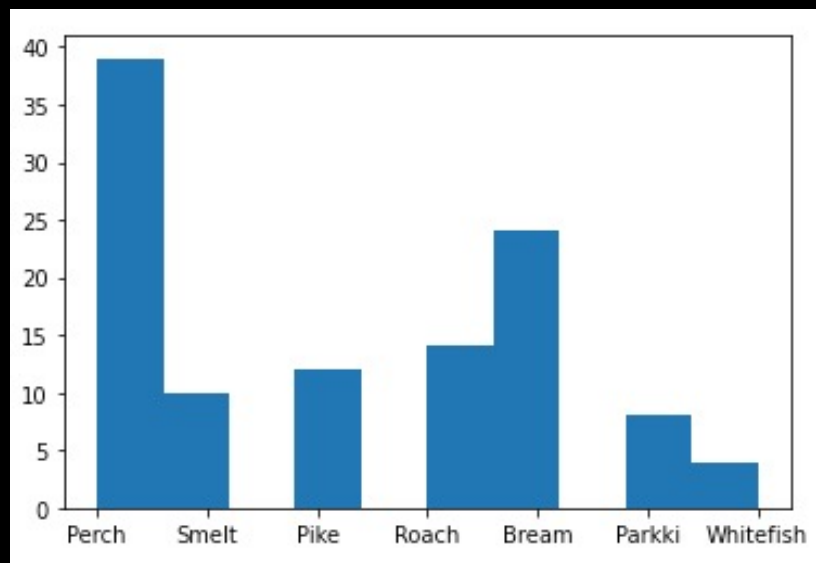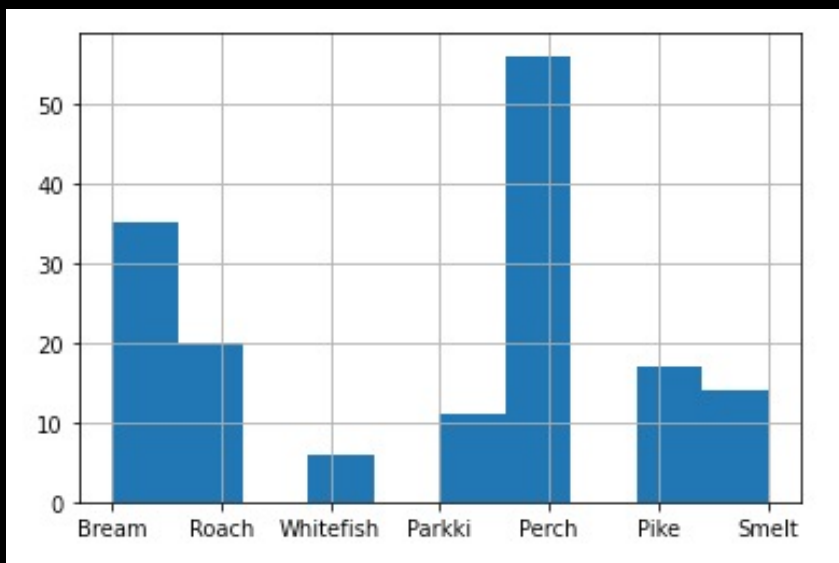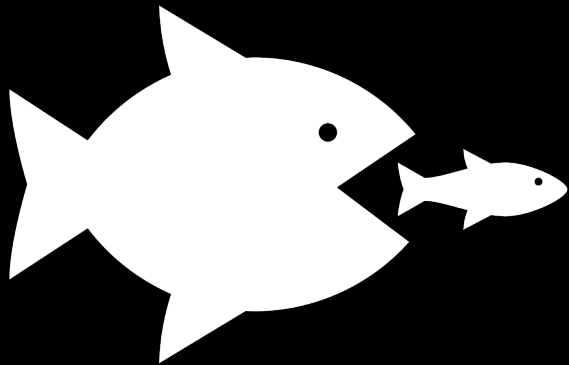
Predir una variable

TRACTAMENT DE LES DADES:
Com és el Dataset?

- Variables
    - Weight
    - Length1
    - Length2
    - Length3
    - Height
    - Width

- Tractament dels NaNs

ENTRENANT EL MODEL

- Repartiment de les dades

MODELS

1
- Regressió Logística
- One vs Rest

2
- Pesos

3
- Doble classificació (Other fishes)

4
- Eliminem Whitefish

MODELS

	precision	recall	f1-score	support
Bream	1.00	1.00	1.00	11
Parkki	1.00	1.00	1.00	3
Perch	0.80	0.94	0.86	17
Pike	1.00	0.60	0.75	5
Roach	0.67	0.67	0.67	6
Smelt	0.80	1.00	0.89	4
Whitefish	0.00	0.00	0.00	2
accuracy			0.85	48
macro avg	0.75	0.74	0.74	48
weighted avg	0.83	0.85	0.83	48

MODEL 1:   - Accuracy train: 96.3%
            - Accuracy test: 85.4%

	precision	recall	f1-score	support
Bream	1.00	0.91	0.95	11
Parkki	1.00	1.00	1.00	3
Perch	1.00	0.94	0.97	17
Pike	1.00	0.40	0.57	5
Roach	0.50	0.67	0.57	6
Smelt	0.80	1.00	0.89	4
Whitefish	0.50	1.00	0.67	2
accuracy			0.85	48
macro avg	0.83	0.85	0.80	48
weighted avg	0.90	0.85	0.86	48

MODEL 2:   - Accuracy train: 100%
            - Accuracy test: 85.41%

	precision	recall	f1-score	support
Bream	1.00	0.91	0.95	11
Parkki	0.75	1.00	0.86	3
Perch	1.00	0.94	0.97	17
Pike	1.00	0.80	0.89	5
Roach	0.75	0.50	0.60	6
Smelt	0.80	1.00	0.89	4
Whitefish	0.40	1.00	0.57	2
accuracy			0.88	48
macro avg	0.81	0.88	0.82	48
weighted avg	0.91	0.88	0.88	48

MODEL 3:   - Accuracy train: 98.18%
            - Accuracy test: 87.5%

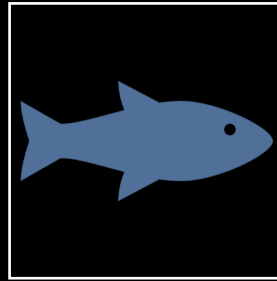|	precision	recall	f1-score	support
Bream	1.00	0.91	0.95	11
Parkki	1.00	1.00	1.00	3
Perch	0.89	0.94	0.91	17
Pike	1.00	0.40	0.57	5
Roach	0.50	0.67	0.57	6
Smelt	0.80	1.00	0.89	4
accuracy			0.85	46
macro avg	0.86	0.82	0.82	46
weighted avg	0.88	0.85	0.84	46

MODEL 4:   - Accuracy train: 100%
            - Accuracy test: 84.78%

CONCLUSIONS
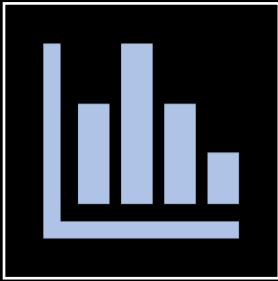
Millor model: Model 3

Dataset amb poques dades

Overfitting?