IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Features selection and Regression analysis to predict weight of Fish

[1]Aishwarya Rajesh Walkar, [2]Chintamani Madhukar Kasture.

[1]Mechanical Engineer, [2]MCA.
Ahmednagar, India

Abstract: The present data is of different features of fishes out of which Weight is a response variable and others, Length1, Length2, Length3, Height, and Width are predictors. The present study is carried out with two reasons:-

1. To point out the important predictors.
2. To obtain the appropriate regression model.

Keywords: Backward elimination, Best subset regression, Blue revolution, Forward selection.

INTRODUCTION

Blue Revolution, the Neel Kranti Mission has the vision to achieve economic prosperity of the country. The fishers and fish farmers contribute towards food and nutritional security through full potential utilization of water resources for fisheries development in a sustainable manner, keeping in view the bio-security and environmental concerns.

Foreseeing the huge scope for development of fisheries, the Prime Minister Shri Narendra Modi, in December, 2014, had called for “a revolution” in the Fisheries sector and named it as “Blue Revolution”.

Objectives:-

1. To increase the overall fish production in a responsible and sustainable manner for economic prosperity
2. To modernize the fisheries with special focus on new technologies
3. To ensure food and nutritional security
4. To generate employment and export earnings
5. To ensure inclusive development and empower fishers and aquaculture farmers

Common bream:-

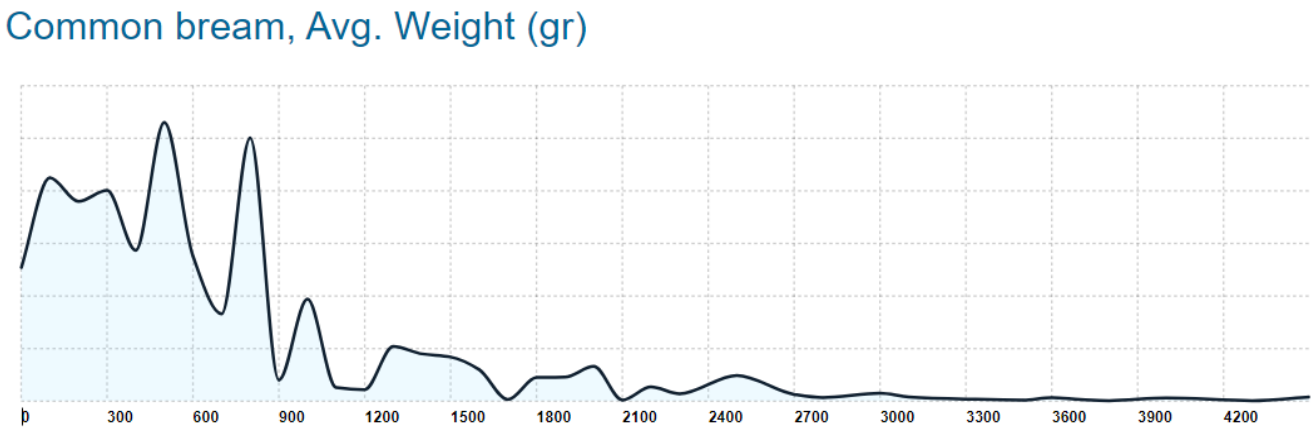
The common bream, freshwater bream, bream, bronze bream, carp bream is a European species of freshwater fish in the family Cyprinidae. It is now considered to be the only species in the genus Abramis.

Distribution and Habitat:-

The common bream's home range is Europe north of the Alps and Pyrenees, as well as the Balkans. It is found as far as east as the Caspian Sea, the Black Sea, and the Aral Sea. The common bream lives in ponds, lakes, canals, and slow-flowing rivers.

Size and Growth:-

The bream is usually 30 to 55 cm (12 to 22 in) long, though some specimens of 75 cm (30 in) have been recorded; it usually weighs 2 to 4 kg (4.4 to 8.8 lb). Its maximum length is 90 cm (35.5 in) with a maximum recorded weight around 9.1 kg (20 lb).



Fishing tips:-

During the summer, you can practice successful bream angling with a light tackle bait with worm or corn, or both, so-called "cocktail". When the bream are sitting high in the water under structure such as a bridge or pontoon, then using a lightly weighted jig head and drifting the lure slowly past the hot spots is a good option. If the fish are holding deep, fishing a heavy jig head is going to help get the lure down into the strike zone.

European perch:-

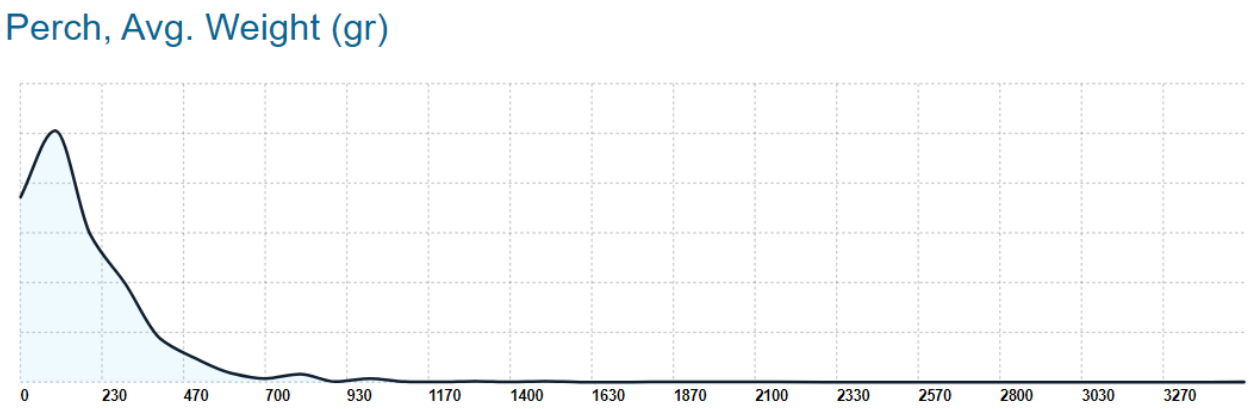
The European perch, also known as the common perch, red fin perch, big-scaled red fin, English perch, Eurasian perch, Eurasian river perch, Hatch or in Anglophone parts of Europe, simply the perch, is a predatory species of the freshwater perch native to Europe and northern Asia. The species is a popular quarry for anglers, and has been widely introduced beyond its native area, into Australia, New Zealand, and South Africa. They have caused substantial damage to native fish populations in Australia and have been proclaimed a noxious species in New South Wales.

Distribution and Habitat:-

The range of the European perch covers fresh water basins all over Europe, excluding the Iberian Peninsula. Their range is known to reach the Kolyma River in Siberia to the east. It is also common in some of the brackish waters of the Baltic Sea.

Size and Growth:-

When the perch is small it grows fast, and it is already over 10 cm long in the second year at best. After that its growth rate slows down, If you get a perch over 30 cm it is often more than ten years old. At most, the perch becomes just over 20 years old, and can reach a length of about 60 cm.



Fishing tips:-

The perch is a typical predator that often hunts in shoals. Larger perch prefer worms, other small fish and fish fry. Effective baits include lob worms, earthworms, maggots and casters.

Importance of Various features:-

Weight Weight of the fish (in grams)

Length1 Length from the nose to the beginning of the tail (in cm)

Length 2 Length from the nose to the notch of the tail (in cm)

Length3 Length from the nose to the end of the tail (in cm)

Height% Maximal height as % of Length3

Width% Maximal width as % of Length3

Length 1 and Length 2 turn out to be insignificant in any model in which length 3 is included, I.e Best subsets regression,

Forward selection and backward elimination.

Weight = length * width therefore width selected.

Aim and Objective:

Usually Weight or Volume of an object is based on length, width & height.
In this data set lengths at different points are given, so our aim to point out the most important measurement of length. Similarly among the Width and Height, point out the most sensible variable.

Method:-

2 Species studied		Variables taken
Bream (35*07)	Data type (Numeric)	Response Variable :- Weight (kg)
Perch (56*07)	Data type (Numeric)	Regressor Variable:- Length1 (cm) Length 2 (cm) Length 3 (cm) Height (cm) Width (cm)

Summary Statistics:-

Common bream:-

	Weight	***L1***	***L2***	***L3***	***Height***	***Width***
Mean	626.00	30.31	33.11	38.35	39.53	14.13
Standard Error	34.41	0.61	0.66	0.70	0.27	0.13
Median	620.00	30.40	33.00	38.50	39.80	14.10
Mode	500.00	26.80	35.00	36.20	40.00	13.80
Standard Deviation	203.54	3.59	3.91	4.16	1.62	0.77
Sample Variance	41430.00	12.91	15.30	17.29	2.63	0.59
Kurtosis	-0.64	-0.10	-0.25	-0.39	1.46	0.11
Skewness	0.19	0.11	0.16	0.05	0.38	-0.25
Range	758.00	14.80	15.60	16.50	8.30	3.50
Minimum	242.00	23.20	25.40	30.00	36.20	12.00
Maximum	1000.00	38.00	41.00	46.50	44.50	15.50
Sum	21910.00	1060.70	1158.80	1342.40	1383.40	494.60
Count	35.00	35.00	35.00	35.00	35.00	35.00

European perch:-

	Weight	*L1*	*L2*	*L3*	*Height*	*Width*
Mean	382.24	25.74	27.89	29.57	26.26	15.84
Standard Error	46.45	1.14	1.21	1.27	0.25	0.18
Median	207.50	23.25	25.30	26.75	26.25	15.55
Mode	130.00	20.00	22.00	23.50	24.00	15.00
Standard Deviation	347.62	8.56	9.02	9.53	1.91	1.36
Sample Variance	120838.08	73.30	81.39	90.81	3.64	1.85
Kurtosis	-0.87	-1.01	-0.95	-0.94	-0.21	2.13
Skewness	0.84	0.27	0.27	0.27	0.12	1.04
Range	1094.10	33.60	35.60	37.80	9.50	7.70
Minimum	5.90	7.50	8.40	8.80	21.30	13.20
Maximum	1100.00	41.10	44.00	46.60	30.80	20.90
Sum	21405.40	1441.20	1562.00	1656.00	1470.40	887.00
Count	56.00	56.00	56.00	56.00	56.00	56.00

COMMON BREAM

MatrixPlot 'Length1'-'Width';

Correlations: Weight, Length1, Length2, Length3, Height, Width

	Weight	Length1	Length2	Length3	Height
Length1	0.954				
	0.000				
Length2	0.961	0.998			
	0.000	0.000			
Length3	0.963	0.996	0.998		
	0.000	0.000	0.000		
Height	0.484	0.326	0.354	0.358	
	0.003	0.056	0.037	0.035	
Width	0.342	0.199	0.235	0.245	0.436
	0.044	0.253	0.174	0.157	0.009

Best Subsets Regression: Weight versus Length1, Length2, Length3, Height, Width

Response is Weight

					L	L	L		
					e	e	e	h	
					n	n	n	e	w
					g	g	g	i	i
					t	t	t	g	d
					h	h	h	h	t
Vars	R-sq	R-sq (adj)	C-p	S	1	2	3	t	h
1	92.7	92.5	14.1	55.677			X		
1	92.3	92.1	16.6	57.231		X			
2	95.0	94.7	2.2	47.034			X	X	
2	94.7	94.4	3.9	48.318		X		X	
3	95.3	94.8	2.2	46.245			X	X	X
3	95.1	94.7	3.2	47.025	X			X	X
4	95.3	94.7	4.0	46.848	X		X	X	X
4	95.3	94.7	4.1	46.963		X	X	X	X
5	95.3	94.5	6.0	47.645	X	X	X	X	X

From table it is observed that Length 3, height & width forms the best subset regression with maximum R-square adjusted 94.8 and least mallows CP = 2.2

Stepwise Regression: Weight versus Length1, Length2, Length3, Height, Width

Forward selection. Alpha-to-Enter: 0.25

Response is Weight on 5 predictors, with N = 35

Step	1	2	3
Constant	-1182	-1870	-1968
Length3	47.1	44.3	44
T - Value	20.53	21.34	21.43
P - Value	0.000	0.000	0.000
Height		20.1	17.0
T - Value		3.77	2.99
P - Value		0.001	0.005
Width			17
T - Value			2.99
P - Value			0.005
S	55.7	47	46.2
R-Sq	92.74	94.97	95.29
R-Sq (adj)	92.52	94.66	94.84
c-p	14.1	2.2	2.2

From table it is observed that Length 3, height & width forms the best Stepwise regression (Forward selection) with maximum R-square adjusted 94.84 and least mallows CP = 2.2

Stepwise Regression: Weight versus Length1, Length2, Length3, Height, Width

Backward elimination. Alpha-to-Remove: 0.1

Response is Weight on 5 predictors, with N = 35

Step	1	2	3	4
Constant	-1993	-1983	-1968	-1870
Length1	17	15		
T - Value	0.38	0.46		
P - Value	0.704	0.652		
Length2	-3			
T - Value	-0.07			
P - Value	0.945			
Length3	32.6	31.3	44.0	44.3
T - Value	0.94	1.11	21.43	21.34
P - Value	0.356	0.275	0.000	0.000
Height	17.6	17.5	17.0	20.1
T-Value	2.89	2.98	2.99	3.77
P-Value	0.007	0.006	0.005	0.001
Width	20	20	17	
T-Value	1.43	1.48	1.45	
P-Value	0.163	0.150	0.157	
S	47.6	46.8	46.2	47.0
R-Sq	95.33	95.33	95.29	94.97
R-Sq(adj)	94.52	94.70	94.84	94.66
C-p	6.0	4.0	2.2	2.2

From table it is observed that Length 3, height & width forms the best Stepwise regression (Backward elimination) with maximum R-square adjusted 94.84 and least mallows CP = 2.2.

Regression Analysis: Weight versus Length3, Height, Width

From three different Regression approaches, it is observed that Length 3, width & Height are three most important variables in analyzing the regression. The results of regression analysis with Weight as a response variable and three predictors namely Length 3, height and width are given below:-

The regression equation is:-

Weight = - 1968 + 44.0 Length3 + 17.0 Height + 16.7 Width

Predictor	Coef	SE Coef	T	P	
Constant	-1968.4	205.3	-9.59	0.000	
Length3	44.024	2.054	21.43	0.000	
Height	16.952	5.677	2.99	0.005	
Width	16.69	11.51	1.45	0.157	
S = 46.24	R-Sq = 95.3%	R-Sq(adj) = 94.8%			
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	3	1342324	447441	209.22	0.000
Residual Error	31	66296			
Total	34	1408620			
Source	DF	Seq SS			
Length3	1	1306321			
Height	1	31510			
Width	1	4492			

Conclusion:-

In Conclusion, it can be said that in the present example while applying different data analysis tools and getting certain conclusions it is necessary to analyze the data by different approaches and come up with common features observed in the analysis.
In the present study three different regression approaches gives same three predictors and hence gives strong support to the regression of weight on Length3, width and height**.**

EUROPEAN PERCH

MatrixPlot 'Length1'-'Width';

Correlations: Weight, Length1, Length2, Length3, Height, Width

	Weight	Length1	Length2	Length3	Height
Length1	0.958				
	0.000				
Length2	0.959	1.000			
	0.000	0.000			
Length3	0.960	0.999	1.000		
	0.000	0.000	0.000		
Height	0.595	0.545	0.545	0.545	
	0.000	0.000	0.000	0.000	
Width	0.563	0.485	0.484	0.485	0.628
	0.000	0.000	0.000	0.000	0.000

Best Subsets Regression: Weight versus Length1, Length2, Length3, Height, Width.

Response is Weight

					L	L	L		
					e	e	e	h	
					n	n	n	e	w
					g	g	g	i	i
					t	t	t	g	d
					h	h	h	h	t
Vars	R-sq	R-sq (adj)	C-p	S	1	2	3	t	h
1	92.1	91.9	9.1	98.822			X		
1	91.9	91.8	10.4	99.833		X			
2	93.3	93.0	1.6	91.652			X		X
2	93.2	92.9	2.6	92.582		X			X
3	93.4	93.0	2.6	91.642			X	X	X
3	93.4	93.0	3.0	92.030		X	X		X
4	93.5	93.0	4.0	91.999		X	X	X	X
4	93.5	92.9	4.4	92.328	X		X	X	X
5	93.5	92.9	6.0	92.905	X	X	X	X	X

From table it is observed that Length 3, height & width forms the best subset regression with maximum R-square adjusted 93.0 and least mallows CP = 2.6

Stepwise Regression: Weight versus Length1, Length2, Length3, Height, Width.

Forward selection. Alpha-to-Enter: 0.25

Response is Weight on 5 predictors, with N = 56

Step	1	2
Constant	-652.8	-1100.3
Length3	35.0	32.8
T - Value	25.03	22.08
P - Value	0.000	0.000
Width		32
T - Value		3.13
P - Value		0.003
S	98.8	91.7
R-Sq	92.07	93.30
R-Sq (adj)	91.92	93.05
c-p	9.1	1.6

From table it is observed that Length 3, height & width forms the best subset regression with maximum R-square adjusted 93.05 and least mallows CP = 1.6

Stepwise Regression: Weight versus Length1, Length2, Length3, Height, Width.

Backward elimination. Alpha-to-Remove: 0.1

Response is Weight on 5 predictors, with N = 56

Step	1	2	3	4
Constant	-1223	-1228	-1224	-1100
Length1	6			
T - Value	0.10			
P - Value	0.922			
Length2	-57	-51		
T - Value	-0.61	-0.77		
P - Value	0.547	0.443		
Length3	80.9	80.1	32.2	32.8
T - Value	1.28	1.29	20.31	22.08
P - Value	0.206	0.202	0.000	0.000
Height	9.1	9.1	8.9	
T-Value	1.01	1.02	1.01	
P-Value	0.318	0.314	0.319	
Width	26	26	26	32
T-Value	2.15	2.18	2.22	3.13
P-Value	0.036	0.034	0.031	0.003
S	92.9	92.0	91.6	91.7
R-Sq	93.51	93.51	93.43	93.30
R-Sq(adj)	92.86	93.00	93.05	93.05
C-p	6.0	4.0	2.6	1.6

From table it is observed that Length 3, height & width forms the best subset regression with maximum R-square adjusted 93.05 and least mallows CP = 1.6

Regression Analysis: Weight versus Length3, Width

From three different Regression approaches, it is observed that Length 3 and Width are two most important variables in analyzing the regression. The results of regression analysis with Weight as a response variable and two predictors namely Length 3 and width are given below:-

The regression equation is
Weight = - 1100 + 32.8 Length3 + 32.5 Width

Predictor	Coef	SE Coef	T	P	
Constant	-1100.3	148.7	-7.40	0.000	
Length3	32.751	1.483	22.08	0.000	
Width	32.45	10.38	3.13	0.003	
S = 91.65	R-sq = 93.3%	R-sq(adj) = 93.0%			
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	2	6200892	3100446	369.10	0.000
Residual Error	53	445203	8400		
Total	55	6646094			
Source	DF	Seq SS			
Length3	1	6118739			
Width	1	82152			

Conclusion:-

In Conclusion, it can be said that in the present example while applying different data analysis tools and getting certain conclusions it is necessary to analyze the data by different approaches and come up with common features observed in the analysis.

In the present study Best Subset Regression gives three predictors such as Length3, Height and Width and remained two approaches gives two predictors such as Length3 and Width hence gives strong support to the regression of weight on Length3 and Width.

Therefore in two different cases, we got two different insights, For Common bream there are 3 predictors Length3, Height and Width important, whereas for European perch Length3 and Width are important.