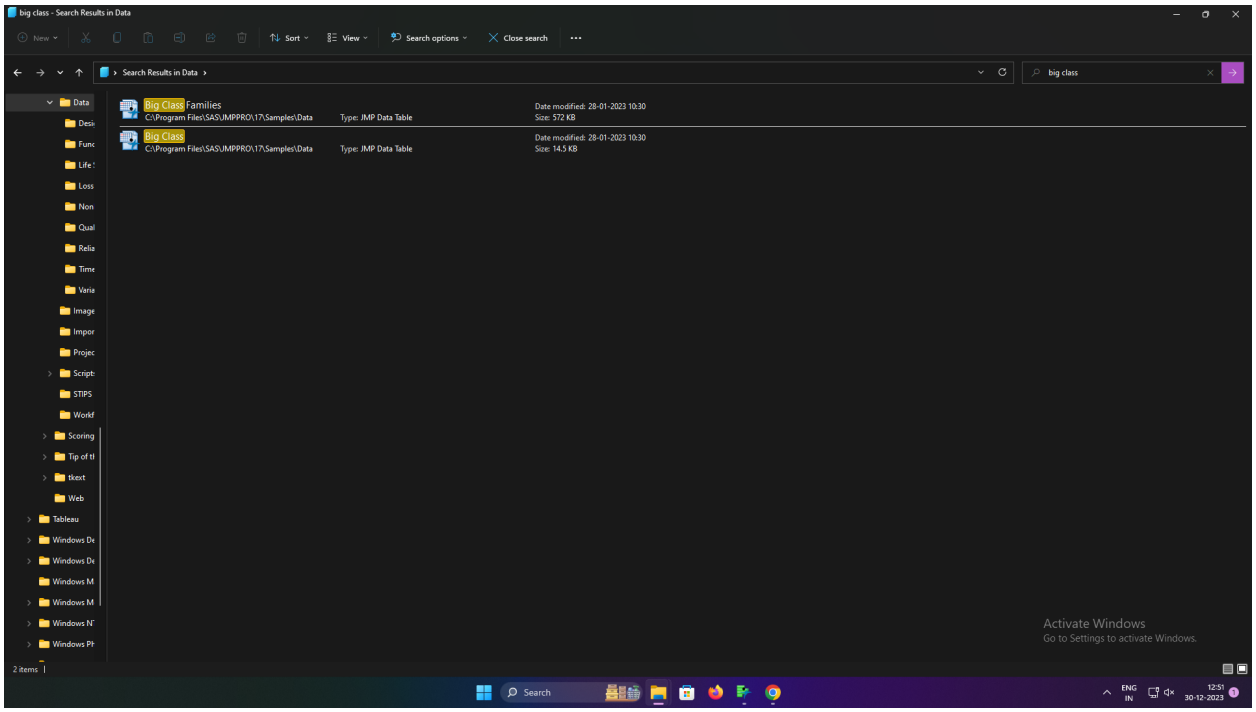
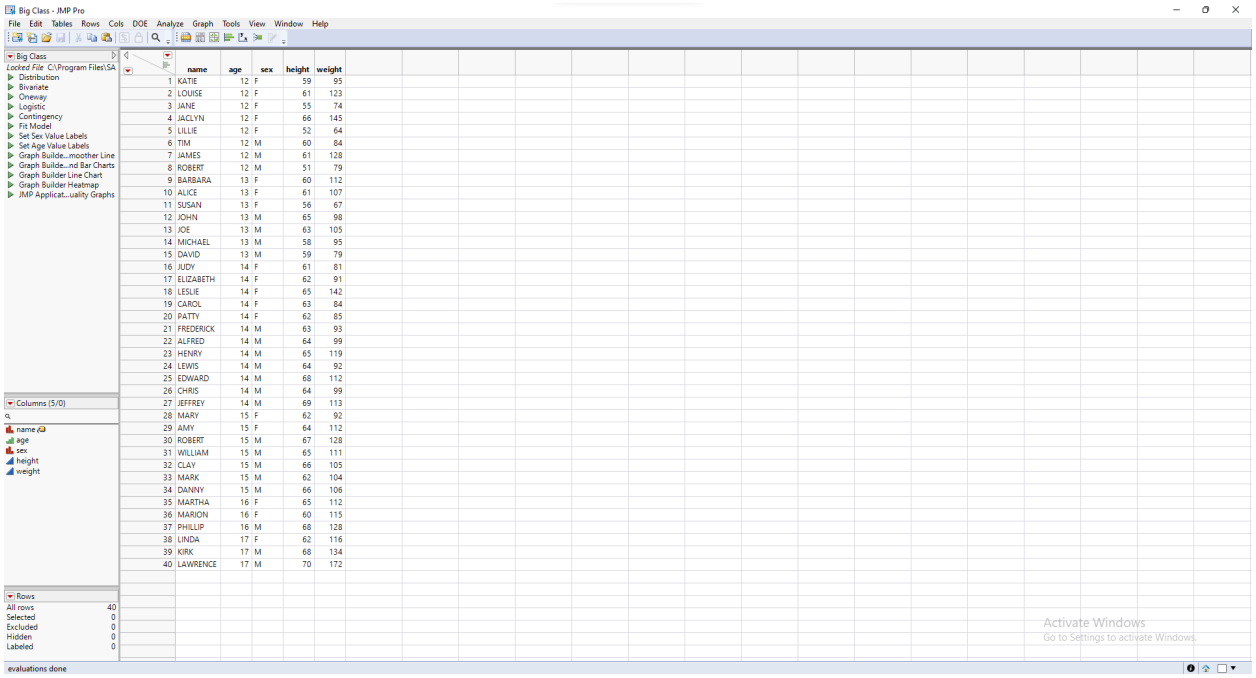
Program 4 : Analyze and predict using advanced regression & classification methods. Explore various ML algorithms to build, validate, compare, select, and deploy statistical methods

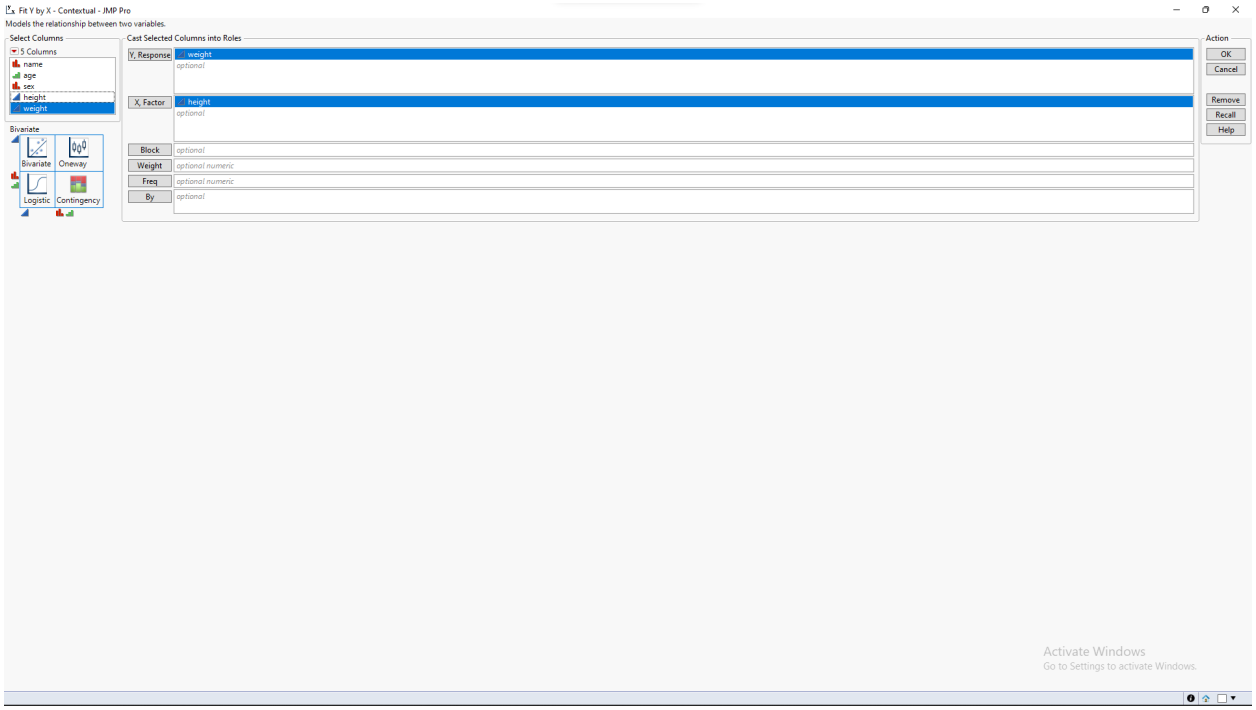
Step 1: Open dataset ‘ Big Class’ from sample data folder.



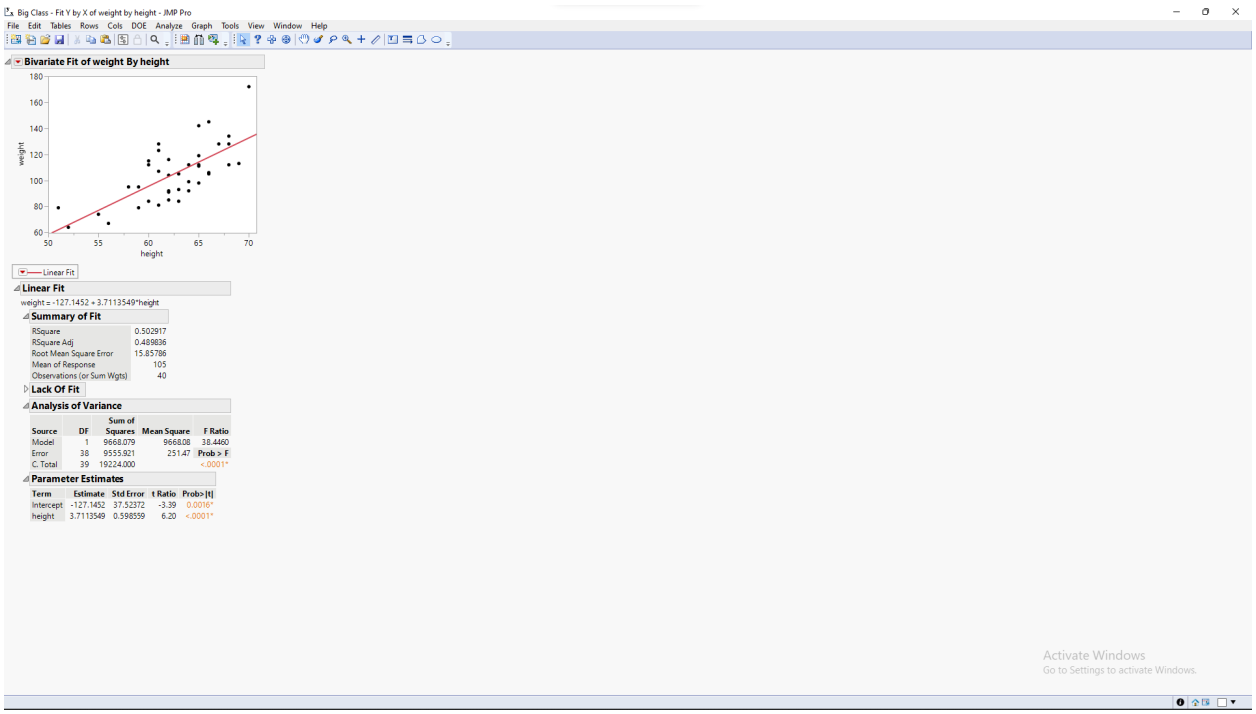


	name	age	sex	height	weight
1	KATIE	12	F	59	95
2	LOUISE	12	F	61	123
3	JANE	12	F	55	74
4	JACLYN	12	F	66	145
5	LILLIE	12	F	52	64
6	TIM	12	M	60	84
7	JAMES	12	M	61	128
8	ROBERT	12	M	51	79
9	BARBARA	13	F	60	112
10	ALICE	13	F	61	107
11	SUSAN	13	F	56	67
12	JOHN	13	M	65	98
13	JOE	13	M	63	105
14	MICHAEL	13	M	58	95
15	DAVID	13	M	59	79
16	JUDY	14	F	61	81
17	ELIZABETH	14	F	62	91
18	LESLIE	14	F	65	142
19	CAROL	14	F	63	84
20	PATTY	14	F	62	85
21	FREDERICK	14	M	63	93
22	ALFRED	14	M	64	99
23	HENRY	14	M	65	119
24	LEWIS	14	M	64	92
25	EDWARD	14	M	68	112
26	CHRIS	14	M	64	99
27	JEFFREY	14	M	69	113
28	MARY	15	F	62	92
29	AMY	15	F	64	112
30	ROBERT	15	M	67	128
31	WILLIAM	15	M	65	111
32	CLAY	15	M	66	105
33	MARK	15	M	62	104
34	DANNY	15	M	66	106
35	MARTHA	16	F	65	112
36	MARION	16	F	60	115
37	PHILLIP	16	M	68	128
38	LINDA	17	F	62	116
39	KIRK	17	M	68	134
40	LAWRENCE	17	M	70	172

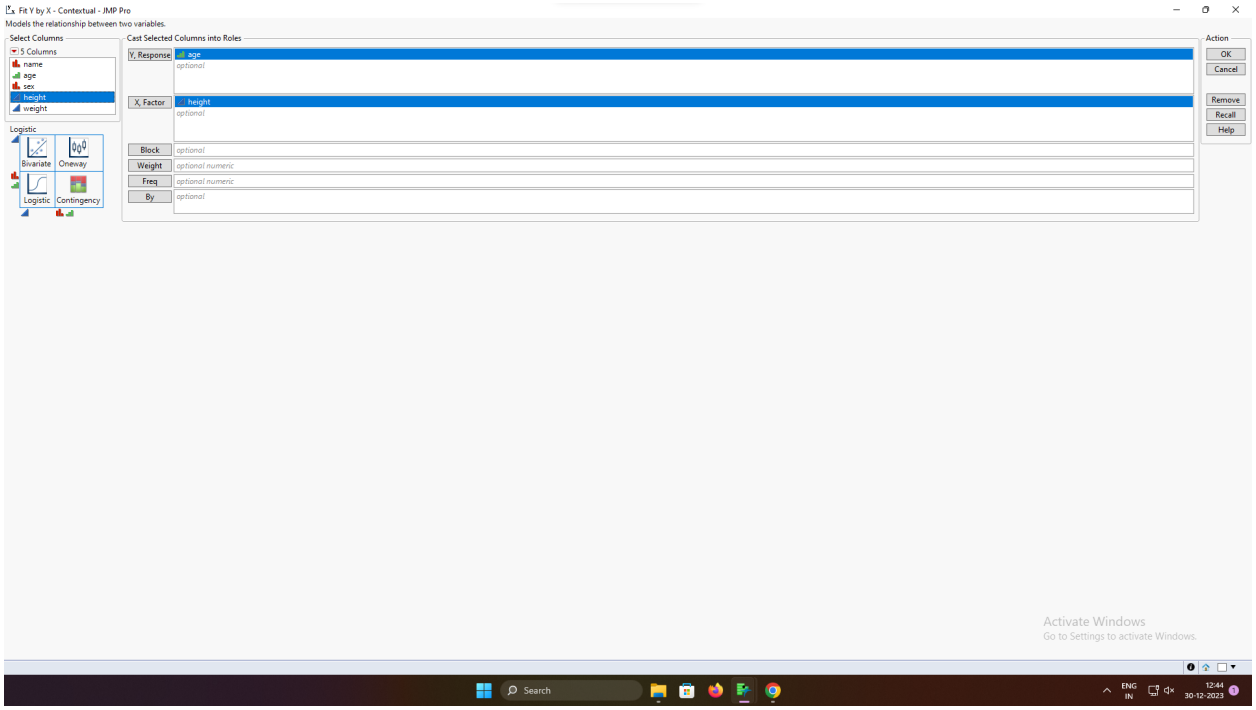
Step 2: In order to do simple linear regression, click on ‘analyze’ then click on ‘Fit Y by X’. Select height in X axis and weight in X axis.



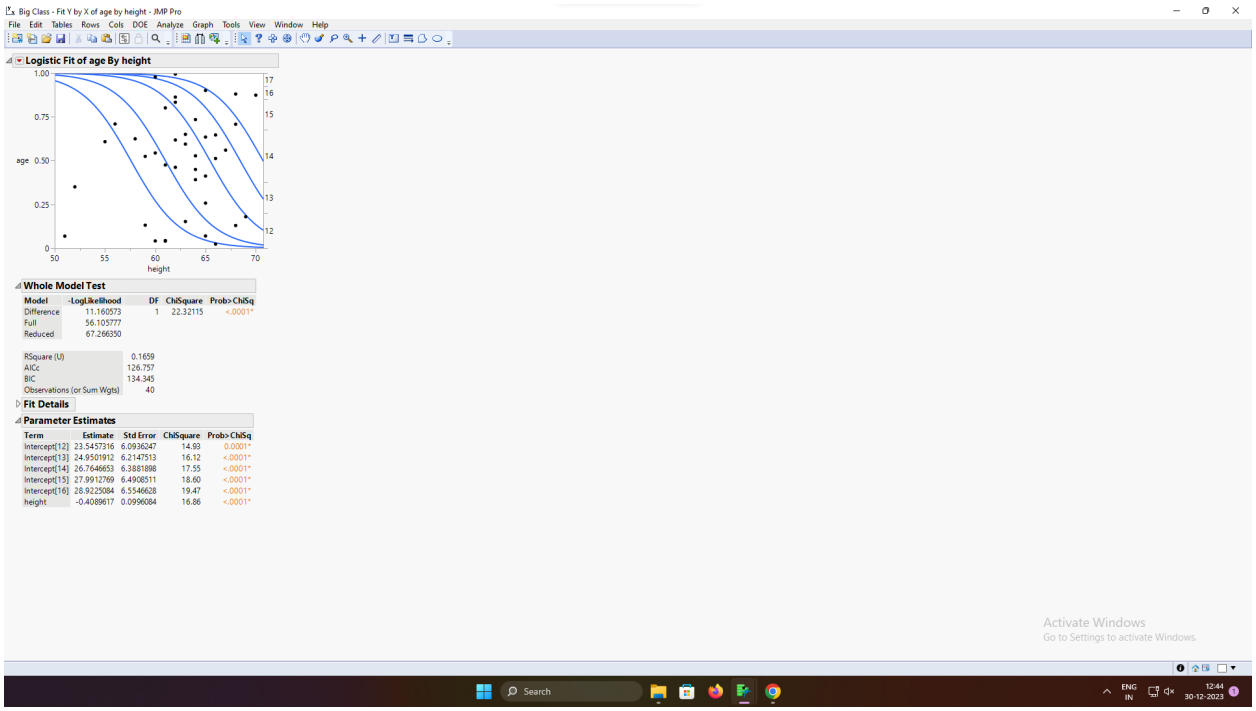
Step 3: In Order to plot the graph click on ‘OK’. Then click on fit line to add a straight line.



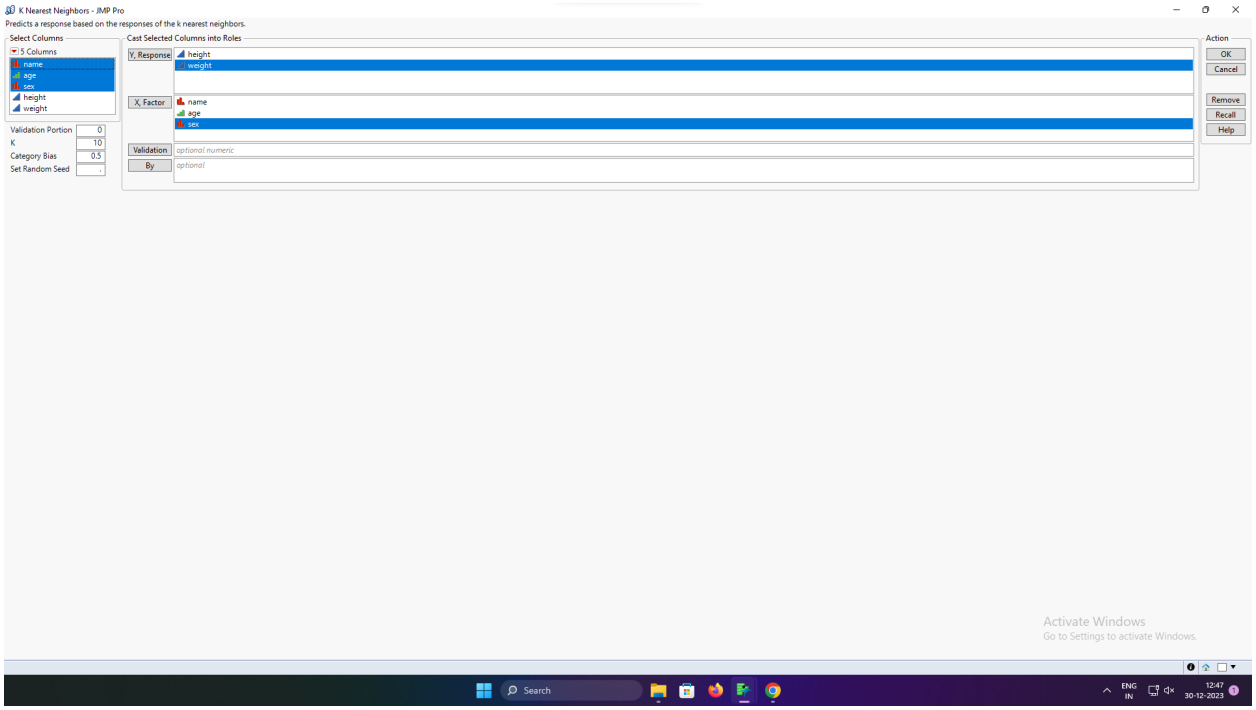
Step 4: Similarly, to make a logistic regression. Select ‘height’ and ‘weight’ in the y axis and select the rest of the factors in the y axis

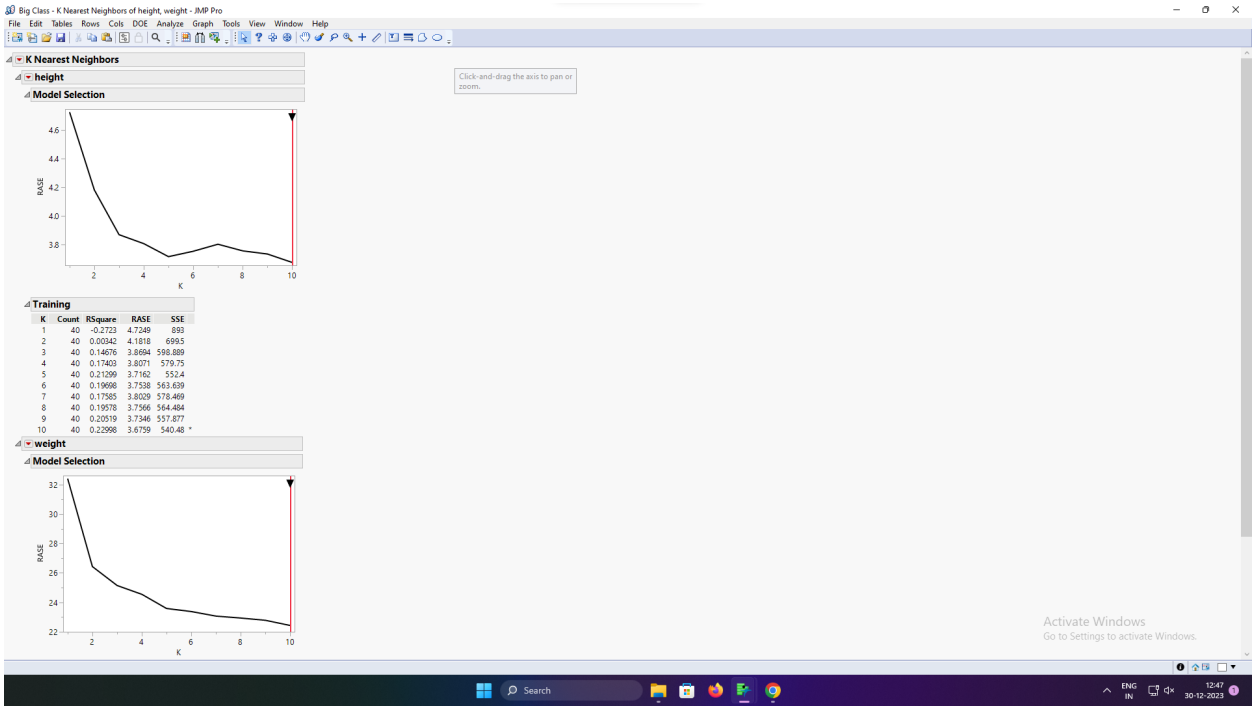


Step 5: This will provide a logistic graph



Step 6: For classification we use the k nearest neighbors algorithm to classify





Big Class - K Nearest Neighbors of height, weight - JMP Pro

File Edit Tables Rows Cols DOE Analyze Graph Tools View Window Help

K Nearest Neighbors

height

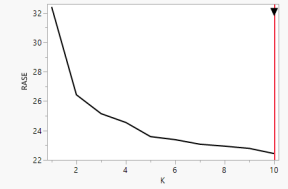
Model Selection

Training

K	Count	RSquare	RASE	SSE	
1	40	-0.2723	4.7249	893	
2	40	0.00342	4.1818	699.5	
3	40	0.14676	3.8694	598.889	
4	40	0.17403	3.8071	579.75	
5	40	0.21299	3.7162	552.4	
6	40	0.19698	3.7538	563.639	
7	40	0.17585	3.8029	578.469	
8	40	0.19578	3.7566	564.484	
9	40	0.20519	3.7346	557.877	
10	40	0.22998	3.6759	540.48	*

weight

Model Selection



Training

K	Count	RSquare	RASE	SSE	
1	40	-1.1849	32.405	42003	
2	40	-0.454	26.435	27951.5	
3	40	-0.316	25.149	252993	
4	40	-0.2542	24.551	241109	
5	40	-0.1578	23.589	22257.4	
6	40	-0.1376	23.382	21869.6	
7	40	-0.1075	23.071	212915	
8	40	-0.0951	22.941	21052	
9	40	-0.0804	22.787	207703	
10	40	-0.0475	22.437	201363	*