Report

Introduction

A study was conducted in 36 children diagnosed at a clinic with moderate-to-severe dehydration to determine the degree of recovery that takes place 90 minutes following treatment. Patients were treated with an electrolyte solution in popsicle-form at 3 different doses (measured in mEq/L). A rehydration scale score was assigned subjectively based on examination and parent report. The scores were scaled to potentially range from 0 to 100 (as a percent of recovery). In addition to dose, the age and weight of each child were recorded and reported. This data was recorded in a list format in a text file which was imported and read in SAS for analysis.

```
filename in1 '/home/u62275281/dehydration_s23.txt';

data one;
infile in1;
input ID rehydration_score dose age weight @@;
run;
```

First we run three multiple linear regressions predicting rehydration score with different predictor sets in each model.

A) Model A predicted rehydration score from dose, and weight.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2667.66870	889.22290	9.03	0.0002
Error	32	3151.22019	98.47563		
Corrected Total	35	5818.88889			

Root MSE	9.92349	**R-Square**	0.4584
Dependent Mean	71.55556	**Adj R-Sq**	0.4077
Coeff Var	13.86823		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	85.47636	5.96528	14.33	<.0001	.	0
dose	1	6.16969	1.79081	3.45	0.0016	0.97245	1.02833

Parameter Estimates							
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Tolerance**	**Variance Inflation**
age	1	0.27695	2.28474	0.12	0.9043	0.12250	8.16330
weight	1	-0.54278	0.32362	-1.68	0.1032	0.12244	8.16745

Collinearity Diagnostics (intercept adjusted)					
			Proportion of Variation		
Number	**Eigenvalue**	**Condition Index**	**dose**	**age**	**weight**
1	1.99054	1.00000	0.02518	0.02918	0.02918
2	0.94617	1.45044	0.97480	0.00337	0.00330
3	0.06329	5.60804	0.00002134	0.96745	0.96752

First we test the global model and then move onto individual effects.
The following global hypothesis is tested:
H_o: There is no linear association between *rehydration* score and dose and age and weight ($\beta_{dose} = \beta_{weight} = \beta_{age} = 0$).
H_1: There is a linear association between *rehydration* score and dose and age and weight ($\beta_{dose} \neq \beta_{weight} \neq \beta_{age} \neq 0$).

Level of significance: α=0.05

Estimates of interest:
F = 9.03
df = (3, 32)
$R^2 = 0.4077$
P=0.0002
Conclusion of global hypothesis test: Reject H_o and conclude that there is significant evidence of an linear association between *rehydration* score and dose and age and weight. The variables dose and age and weigh together explain 45.84 % of the variability in *rehydration* score.
Since the global test was significant, we now look at the individual level tests.

<u>*Dose:*</u>
H_o: There is no linear association between *rehydration* score and dose, adjusting for weight and age(β_{dose} =0|age,weight).
H_1: There is a linear association between *rehydration* score and dose, adjusting for weight and age(β_{dose} ≠0|age,weight)
Level of significance: α=0.05

Results:
$\hat{\beta}_{dose=}$ 6.16969
t = 3.45, df = 32
p=0.0016
SE= 1.79081

For every one unit increase in dose was associated with a increase in mean *rehydration score* of 6.16969 units (p=0.0016), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Reject H_o and conclude that there is significant evidence of an association between *rehydration* score and dose, adjusting for age and weight.

Age:
H_o: There is no linear association between *rehydration* score and age, adjusting for weight and dose(β_{age} =0|dose,weight).
H_1: There is a linear association between *rehydration* score and agee, adjusting for weight and dose(β_{age} ≠0|dose,weight)
Level of significance: α=0.05

Results:
$\hat{\beta}_{age=}$ 0.27695
t = 0.12, df = 32
p=0.9043
SE= 2.28474
For every one unit increase in age was associated with a increase in mean *rehydration score* of 0.27695 units (p=0.9043), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Fail to reject H_o and conclude that there is no significant evidence of an association between *rehydration* score and age, adjusting for dose and weight.

Weight:
H_o: There is no linear association between *rehydration* score and weight, adjusting for age and dose(β_{weight} =0|dose,age).
H_1: There is a linear association between *rehydration* score and weight, adjusting for age *and dose*(β_{weight} ≠0|dose,age)
Level of significance: α=0.05

Results:
$\hat{\beta}_{weight=}$ -0.54278
t = -1.68 df = 32
p=0.1032
SE= 0.32362
For every one unit increase in weight was associated with a decrease in mean *rehydration score* of 0.54278 units (p=0.9043), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Fail to reject H_o and conclude that there is no significant evidence of an association between *rehydration* score and weight, adjusting for dose and age.
The variation inflation factor (VIF) values for age and weight do not exceed 10 and there is no clear indication of a collinearity problem related to these variables.
The condition index for principal component 3 is approximately 5.6 and over 90% of the variance of age and weight is explained by it.

B) Model B predicted rehydration score from dose, and age

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	2	2390.66183	1195.33092	11.51	0.0002
Error	33	3428.22706	103.88567		
Corrected Total	35	5818.88889			

Root MSE	10.19243	**R-Square**	0.4108
Dependent Mean	71.55556	**Adj R-Sq**	0.3751
Coeff Var	14.24408		

Parameter Estimates							
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Tolerance**	**Variance Inflation**
Intercept	1	84.07228	6.06631	13.86	<.0001	.	0
dose	1	6.06709	1.83827	3.30	0.0023	0.97359	1.02713
age	1	-3.30580	0.83240	-3.97	0.0004	0.97359	1.02713

Collinearity Diagnostics (intercept adjusted)				
			Proportion of Variation	
Number	**Eigenvalue**	**Condition Index**	**dose**	**age**
1	1.16252	1.00000	0.41874	0.41874
2	0.83748	1.17819	0.58126	0.58126

First we test the global model and then move onto individual effects.
The following global hypothesis is tested:
H_o: There is no linear association between *rehydration* score and dose and age ($\beta_{dose} = \beta_{age} = 0$).
H_1: There is a linear association between *rehydration* score and dose and age and ($\beta_{dose} \neq\neq \beta_{age} \neq 0$).

Level of significance: α=0.05

Estimates of interest:
F = 11.51
df = (2, 33)

$R^2 = 0.3751$
P=0.0002
Conclusion of global hypothesis test: Reject H_o and conclude that there is significant evidence of an linear association between *rehydration* score and dose and age. The variables dose and age together explain 37.51 % of the variability in *rehydration* score.
Since the global test was significant, we now look at the individual level tests.

<u>*Dose:*</u>
H_o: There is no linear association between *rehydration* score and dose, adjusting for age(β_{dose} =0|age).
H_1: There is a linear association between *rehydration* score and dose, adjusting for age(β_{dose} ≠0|age)
Level of significance: α=0.05

Results:
$\hat{\beta}_{dose=}$ 6.06709
t = 3.30, df = 33
p=0.0023
SE= 1.83827
For every one unit increase in dose was associated with a increase in mean *rehydration score* of 6.06709 units (p=0.0023), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for dose: Reject H_o and conclude that there is significant evidence of an association between *rehydration* score and dose, adjusting for age.

<u>*Age:*</u>
H_o: There is no linear association between *rehydration* score and age, adjusting for dose(β_{age} =0|dose).
H_1: There is a linear association between *rehydration* score and age, adjusting for dose(β_{age} ≠0|dose)
Level of significance: α=0.05

Results:
$\hat{\beta}_{age=}$ -3.30580
t = -3.97, df = 33
p=0.0004
SE= 0.83240
For every one unit increase in age was associated with a decrease in mean *rehydration score* of 3.30580 units (p=0.0004), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for age: Reject H_o and conclude that there is significant evidence of an association between *rehydration* score and age, adjusting for dose.
The variation inflation factor (VIF) values for dose and age do not exceed 10 and there is no clear indication of a collinearity problem related to these variables. The condition index for principal component 3 is approximately 1.17 and over 58% of the variance of age is explained by it. The Rsquare value has decreased.
When we remove the weight variable we notice that the VIF's are much more in line with the model and it is now less concerning than the earlier model.

C) Model C predicted rehydration score from dose, and weight.

D) Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	2	2666.22169	1333.11084	13.95	<.0001
Error	33	3152.66720	95.53537		
Corrected Total	35	5818.88889			

Root MSE	9.77422	**R-Square**	0.4582
Dependent Mean	71.55556	**Adj R-Sq**	0.4254
Coeff Var	13.65962		

Parameter Estimates							
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Tolerance**	**Variance Inflation**
Intercept	1	85.59416	5.79705	14.77	<.0001	.	0
dose	1	6.17526	1.76329	3.50	0.0013	0.97309	1.02765
weight	1	-0.50610	0.11307	-4.48	<.0001	0.97309	1.02765

Collinearity Diagnostics (intercept adjusted)				
			Proportion of Variation	
Number	**Eigenvalue**	**Condition Index**	**dose**	**weight**
1	1.16404	1.00000	0.41798	0.41798
2	0.83596	1.18002	0.58202	0.58202

First we test the global model and then move onto individual effects.
The following global hypothesis is tested:
H_o: There is no linear association between *rehydration* score and dose and weight ($\beta_{dose} = \beta_{weight} = 0$).
H_1: There is a linear association between *rehydration* score and dose and age and ($\beta_{dose} \neq\neq \beta_{weight} \neq 0$).

Level of significance: α=0.05

Estimates of interest:
F = 13.95
df = (2, 33)
R^2 = 0.4254

$P<0.0001$
Conclusion of global hypothesis test: Reject H_o and conclude that there is significant evidence of an linear association between *rehydration* score and dose and weight. The variables dose and weight together explain 42.54 % of the variability in *rehydration* score.
Since the global test was significant, we now look at the individual level tests.

Dose:
H_o: There is no linear association between *rehydration* score and dose, adjusting for weight(β_{dose} =0|weight).
H_1: There is a linear association between *rehydration* score and dose, adjusting for weight(β_{dose} ≠0|weight)
Level of significance: $\alpha=0.05$

Results:
$\hat{\beta}_{dose=}$ 6.17526
t = 3.50, df = 33
p=0.0013
SE= 1.76329
For every one unit increase in dose was associated with a increase in mean *rehydration score* of 6.17526 units (p=0.0013), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for dose: Reject H_o and conclude that there is significant evidence of an association between *rehydration* score and dose, adjusting for weight.

Weight:
H_o: There is no linear association between *rehydration* score and weight, adjusting for dose(β_{weight} =0|dose).
H_1: There is a linear association between *rehydration* score and weight, adjusting for dose(β_{weight} ≠0|dose)
Level of significance: $\alpha=0.05$

Results:
$\hat{\beta}_{weight=}$ -0.50610
t = -4,48, df = 33
p<.0001
SE= 0.11307
For every one unit increase in weight was associated with a decrease in mean *rehydration score* of 0.50610 units (p<0.0001), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for age: Reject H_o and conclude that there is significant evidence of an association between *rehydration* score and weight, adjusting for dose.
The variation inflation factor (VIF) values for dose and weight do not exceed 10 and there is no clear indication of a collinearity problem related to these variables.
The condition index for principal component 3 is approximately 1.18 and over 58% of the variance of age and weight is explained by it.
When we remove the age variable we notice that the VIF's are much more in line with the model and it is now less concerning than the earlier model.

I will choose the **model with only dose and weight as predictors i.e. model C** because it has a VIF of 1.02765 and the unadjusted R2 value of 45.82% and adjusted R2 of 42.54%. When removing age from out model the VIF and model fit increases and the principal component 3 has a Condition index of 5.60804 in the model with all variables but the principal component 2 in model with dose and weight has a condition index of 1.18 explaining 58% variability.

For part 2 of the analysis we construct a grouping (categorical) variable with those who received of a dose of 0 through < 1 in group 1, those who received a dose of 1 through less than 2 in group 2, and those who received a dose greater than or equal to 2 in group and three new continuous variables called *dose1*, *dose2*, and *dose3* to perform a piecewise linear analysis using it.

```
data two;
set one;
if (0 <= dose < 1) then dose1=dose;
else if dose >= 1 then dose1=1;
if (0 <= dose < 1) then dose2=1;
else if (1 <= dose < 2) then dose2=dose;
 else if dose >= 2 then dose2=2;

if (0 <= dose < 2) then dose3=2;
else if dose >= 2 then dose3=dose ;

if (0 <= dose < 1) then dosegroup='d1';
 else if (1 <= dose < 2) then dosegroup='d2';
 else if dose >= 2 then dosegroup='d3';
 run;
```

Perform analyses examining the relationship of rehydration score and dose below. Statistically compare the mean rehydration scores between the dose groups and summarize your results.

Global model
H0: The mean *rehydration* score is same across the all doses.
H1: Atleast one of the level of dose group differs from others in mean *rehydration* score.

Dependent Variable: rehydration_score

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	971.701389	485.850694	3.31	0.0491
Error	33	4847.187500	146.884470		
Corrected Total	35	5818.888889			

R-Square	Coeff Var	Root MSE	rehydration_score Mean
0.166991	16.93732	12.11959	71.55556

Source	DF	Type I SS	Mean Square	F Value	Pr > F
dosegroup	2	971.7013889	485.8506944	3.31	0.0491

Source	DF	Type III SS	Mean Square	F Value	Pr > F
dosegroup	2	971.7013889	485.8506944	3.31	0.0491

Global model
F-statistic= 3.31
df= 2,33
p-value = 0.0491

R^2 = 0.1669

When we conducted the global test we found that there is statistical significance at level of α=0.05, as p=0.0491, this shows there is difference in mean *rehydration* score among the all dose groups.

dosegroup	rehydration_score LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
d1	62.2500000	4.2849223	<.0001	1
d2	72.2500000	3.4986244	<.0001	2
d3	75.6875000	3.0298976	<.0001	3

The mean for group 1 is 62.2500000 and standard error is 4.2849223.

The mean for group 2 is 72.2500000 and standard error is 3.4986244 .

The mean for group 3 is 75.6875000 and standard error is 3.0298976.

Least Squares Means for Effect dosegroup t for H0: LSMean(i)=LSMean(j) / Pr > |t| Dependent Variable: rehydration_score			
i/j	**1**	**2**	**3**
1		-1.80773 0.1827	-2.56053 0.0393
2	1.807726		-0.74272

Least Squares Means for Effect dosegroup t for H0: LSMean(i)=LSMean(j) / Pr > |t| Dependent Variable: rehydration_score			
i/j	1	2	3
	0.1827		0.7401
3	2.56053 0.0393	0.742722 0.7401	

Now we proceed to look at the pairwise comparisons between dose group 1 and 2, dose group 1 and 3 and dose group 2 and 3.

1) comparison between dose group 1 and 2

Null hypothesis: There is no statistically significant difference in the mean *rehydration* score between dose 1 and dose 2 adjusting for dose 3

Alternative hypothesis: There is a statistically significant difference in the mean *rehydration* score between dose 1 and dose 2 adjusting for dose 3

For the first pairwise comparison, we compare group 1 and 2 and observe the *rehydration* score, we fail to reject the null as the comparison is not significant with p-value for the pairwise comparison group 0.1827, and the t statistic of -1.80773.

2) comparison between dose group 1 and 3

Null hypothesis: There is no statistically significant difference in the mean *rehydration* score between dose 1 and dose 3 adjusting for dose 2

Alternative hypothesis: There is a statistically significant difference in the mean *rehydration* score between dose 1 and dose 3 adjusting for dose 2

For the first pairwise comparison, we compare group 1 and 3 and observe the *rehydration* score, we reject the null as the comparison is significant with p-value for the pairwise comparison group 0.0393, and the t statistic of -2.56053.

3) comparison between dose group 2 and 3

Null hypothesis: There is no statistically significant difference in the mean *rehydration* score between dose 2 and dose 3 adjusting for dose 1

Alternative hypothesis: There is a statistically significant difference in the mean *rehydration* score between dose 2 and dose 3 adjusting for dose 1

For the first pairwise comparison, we compare group 2 and 3 and observe the *rehydration* score, we reject the null as the comparison is significant with p-value for the pairwise comparison group 0.7401, and the t statistic of -0.74272.

Now we run a simple linear regression with the rehydration score and as the outcome and the dose (continuous) as the predictor.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	752.15170	752.15170	5.05	0.0313
Error	34	5066.73719	149.02168		
Corrected Total	35	5818.88889			

Root MSE	12.20744	**R-Square**	0.1293
Dependent Mean	71.55556	**Adj R-Sq**	0.1037
Coeff Var	17.06009		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	63.89576	3.97040	16.09	<.0001
dose	1	4.88058	2.17242	2.25	0.0313

H_0: There is no linear association between rehydration score and *dose*. $B_{dose} = 0$

H_1: There is a linear association between rehydration score and *dose*.. $B_{dose} \neq 0$

Level of significance: 0.05

Estimates of interest:

$$\hat{\beta}_{\text{dose}} = 4.88058(SE = 2.17242),\ p = 0.0313$$
$$t = 5.05,\ df = 34$$
$$R^2 = 0.1293$$

A one unit increase in dose was associated with a difference in mean *rehydration score* 4.88 (p=0.0313)

Conclusion of hypothesis test: We Reject H_0 and conclude that there is significant evidence at α=0.05 of a linear association between rehydration score and *dose*.

The R^2 of this model is 12.93% hence the *dose* is not a good predictor of *rehydration score.*

Now we run piecewise linear model using *dose1*, *dose2*, and *dose3* to predict rehydration score.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	951.62922	317.20974	2.09	0.1217
Error	32	4867.25967	152.10186		
Corrected Total	35	5818.88889			

Root MSE	12.33296	**R-Square**	0.1635
Dependent Mean	71.55556	**Adj R-Sq**	0.0851
Coeff Var	17.23550		

The following global hypothesis is tested:
H_o: There is no linear association between *rehydration* score and the three levels of dosages ($\beta_{dose1} = \beta_{dose2} = \beta_{dose3} = 0$).
H_1: There is a linear association between *rehydration* score and at least on of the level of dosages ($\beta_{dose1} \neq \beta_{dose2} \neq \beta_{dose3} \neq 0$).

Level of significance: α=0.05

Estimates of interest:
F = 2.09,
df = (3, 32)
$R^2 = 0.1635$
P=0.1217

Conclusion of global hypothesis test: Fail to reject H_o and conclude that there is no significant evidence of a linear association *rehydration* score and the three levels of dosages. The dose groups explain 16.35% of the variability in *rehydration* score.

The piecewise model generate a R2 of 16.35% which is higher than the r2 of simple linear model which was 12.93%
Now we look at individual effects of the dose groups.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Standardized Estimate**
Intercept	1	58.95375	13.84723	4.26	0.0002	0
dose1	1	8.66356	7.60065	1.14	0.2628	0.22715
dose2	1	7.44284	6.84121	1.09	0.2847	0.26671

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Standardized Estimate**
dose3	1	-2.71192	7.82333	-0.35	0.7311	-0.07275

Dose 1:
H_o: There is no linear association between *rehydration* score and dose 1 after adjusting for dose 2 and dose 3(β_{dose1} =0|dose2,dose3).
H_1: There is a linear association between *rehydration* score and dose 1 after adjusting for dose 2 and dose 3(β_{dose1} ≠0| dose2,dose3)
Level of significance: α=0.05

Results:
$\hat{\beta}_{dose1}$= 8.66356
t = 1.14 df = 32
p=0.2628
SE= 7.60065
For every one unit change in dose 1 the mean *rehydration score* will increase by 8.66356 (p=0.2628), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Fail to Reject H_o. and conclude that there is no significant evidence of an association between *rehydration* score and dose 1, adjusting for dose2 and dose 3.

Dose 2:
H_o: There is no linear association between *rehydration* score and dose 2 after adjusting for dose 1 and dose 3(β_{dose2} =0|dose1,dose3).
H_1: There is a linear association between *rehydration* score and dose 2 after adjusting for dose 1 and dose 3(β_{dose2} ≠0| dose1,dose3)
Level of significance: α=0.05

Results:
$\hat{\beta}_{dose2}$= 7.44284

t = 1.09 df = 32
p=0.2847
SE= 6.84121
For every unit change in dose 2 the mean *rehydration score* will increase by 7.44284 (p=0.2847), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Fail to Reject H_o. and conclude that there is no significant evidence of an association between *rehydration* score and dose 2, adjusting for dose 1 and dose 3.

Dose 3:
H_o: There is no linear association between *rehydration* score and dose 3 after adjusting for dose 1 and dose 2(β_{dose3} =0|dose1,dose2).
H_1: There is a linear association between *rehydration* score and dose 3 after adjusting for dose 1 and dose 2(β_{dose3} ≠0| dose1,dose2)
Level of significance: α=0.05

Results:
$\hat{\beta}_{dose3}$= -2.71192

t = -0.35 df = 32
p=0.7311
SE= 7.82333
For every one unit change in dose 3 the mean *rehydration score* will decrease by 2.71192 (p=0.7311), assuming all other variables in the regression model remain constant.

Conclusion of hypothesis test for doors: Fail to Reject H_o. and conclude that there is no significant evidence of an association between *rehydration* score and dose 3, adjusting for dose1 and dose 2.

Now we compare the slopes of dose 1 and dose 2
Null hypothesis– slope of dose1 is equal to slope of dose2
Alternative hypothesis– slope of dose1 not equal to slope of dose2

Test 1 Results for Dependent Variable rehydration_score				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	1	1.42538	0.01	0.9235
Denominator	32	152.10186		

Results:
F value = 0.01
df for numerator and denominator = 1,32
p=0.29235

We fail to reject the null and conclude that slope of dose 1 is equal to slope of dose 2

We will also compare the slopes of dose 2 and dose 3

Null hypothesis– slope of dose2 is equal to slope of dose 3
Alternative hypothesis– slope of dose2 not equal to slope of dose3

Test 2 Results for Dependent Variable rehydration_score				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	1	91.76415	0.60	0.4430
Denominator	32	152.10186		

Results:
F value = 0.60
df for numerator and denominator = 1,32
p=0.4430

We fail to reject the null and conclude that slope of dose 2 is equal to slope of dose 3.

From this analysis, we can see that there is no significant increase in rehydration score in dose 1 and dose 2, and there is a no significant increase in rehydration score in dose 2 and dose3.

I prefer the model 1 since the r2 is the highest with 16.69% and the p-value is significant. In the piecewise model it is confirmed that the slopes of the dose groups are equal so it is not the best model to select.

Conclusion:
In conclusion we see that the collinearity model with dose and weight is the best in this analysis and the piecewise linear model is not the best to choose.