PROJECT 2

Harsha Kankanamge

A: Randomized Complete Block (RCB) Design

In a certain study, three diets were assigned for a period of 3 days to each of 6 subjects in a randomized complete block design. The subjects, playing the role of blocks, were assigned the following 3 diets in a random order.

Diet 1: mixed fat and carbohydrates,
Diet 2: high fat,
Diet 3: high carbohydrates.

At the end of the 3-day period, each subject was put on a treadmill and the time to exhaustion, in seconds, was measured. The data is in **641PROJ2.sav**

a)Identify the response variable, treatments and blocks. Write the model for the RCB experiment. What are the values of 'b' and 'k'?

Response variable: time to exhaustion, in seconds

Treatments: Diet 1: mixed fat and carbohydrates, Diet 2: high fat, Diet 3: high carbohydrates.

Blocks: 6 subjects

K=3 b=6

Model:

$Y_{ij} = \mu + \alpha_i + ß_j + \mathcal{E}_{ij}$

i = 1….k

j= 1……b **Such that**

i=1…3

b=1….6

$$\sum_{i=1}^{k=3} \boldsymbol{\alpha i} = 0$$

$$\sum_{j=1}^{b=6} ß\mathbf{j} = 0$$

b)Test if 'time to exhaustion' in seconds is normally distributed in all three diet groups.

H0: 'time to exhaustion' in seconds is normally distributed among all three diet groups

H1: 'time to exhaustion' in seconds is not normally distributed among all three diet groups

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	diet groups	Statistic	df	Sig.	Statistic	df	Sig.
yij-time to exhaustion	mixed fat and carbohydrates	.245	6	.200*	.922	6	.517
	high fat	.210	6	.200*	.924	6	.536
	high carbohydrates	.203	6	.200*	.902	6	.386

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Time to exhaustion: among

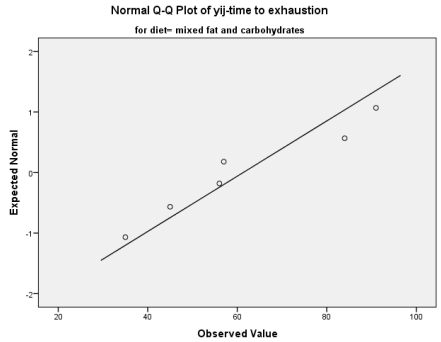
Mixed fat and carbohydrates KS statistic 0.245 p-value = 0.200>0.05

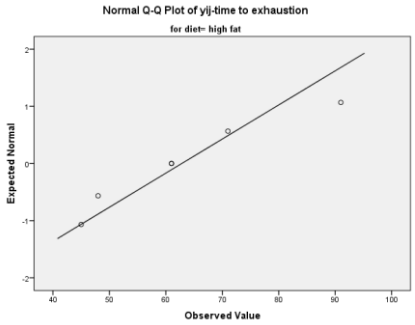
High fat KS statistic 0.210 p-value = 0.200>0.05

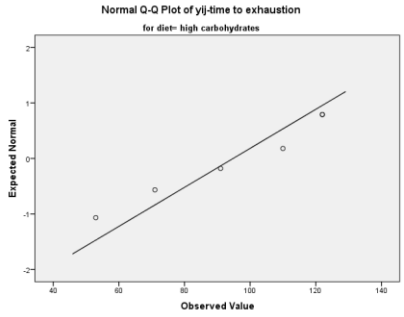
High carbohydrates KS statistics 0.203 p-value = 0.200>0.05

Do not reject H0

'time to exhaustion' in seconds is normally distributed among all three diet groups







All the data points are fitting onto the line. Which mean data are randomly distributed among all three diet groups.

c) Test if 'time to exhaustion' in seconds has homogeneous variance in all three diet groups. (Use Cochran's test)

H0: $\sigma^2_1 = \sigma^2_2 = \sigma^2_3$

H1: At least one of the σ^2_i is different among the three diet groups.

Descriptives

	diet groups			Statistic	Std. Error
yij-time to exhaustion	mixed fat and carbohydrates	Mean		61.33	8.947
		95% Confidence Interval for Mean	Lower Bound	38.33	
			Upper Bound	84.33	
		5% Trimmed Mean		61.15	
		Median		56.50	
		Variance		480.267	
		Std. Deviation		21.915	
		Minimum		35	
		Maximum		91	
		Range		56	
		Interquartile Range		43	
		Skewness		.425	.845
		Kurtosis		-1.419	1.741
	high fat	Mean		62.83	6.843
		95% Confidence Interval for Mean	Lower Bound	45.24	
			Upper Bound	80.42	
		5% Trimmed Mean		62.26	
		Median		61.00	
		Variance		280.967	
		Std. Deviation		16.762	
		Minimum		45	
		Maximum		91	
		Range		46	
		Interquartile Range		29	
		Skewness		.888	.845
		Kurtosis		.697	1.741
	high carbohydrates	Mean		94.83	11.603
		95% Confidence Interval for Mean	Lower Bound	65.01	
			Upper Bound	124.66	
		5% Trimmed Mean		95.65	
		Median		100.50	
		Variance		807.767	
		Std. Deviation		28.421	
		Minimum		53	
		Maximum		122	
		Range		69	
		Interquartile Range		56	
		Skewness		-.565	.845
		Kurtosis		-1.393	1.741

$S^2_1 = 480.267$

$S^2_2 = 280.967$

$S^2_3 = 807.767$

Total of $S^2_i = 1569.001$

$$G = \frac{\text{largest } S_i^2}{\sum_{i=1}^{k} S_i^2},$$

G=807.767/1569.001 = 0.5148 g 0.05 k=3 = 0.7071

G (0.5148) <g (0.7071) Do not reject H0

All variances of 'time to exhaustion' are same among the three diet groups.

d)Using an RCB design, test for differences in 'time to exhaustion' among the three diets.

H0: µ1. =µ2. =µ3. =µ..

H1: At least one mean value of 'time to exhaustion' among the three diets groups is different.

Tests of Between-Subjects Effects

Dependent Variable: yij-time to exhaustion

Source	Type II Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10330.333[a]	7	1475.762	8.146	.002
Intercept	95922.000	1	95922.000	529.468	.000
diet	4297.000	2	2148.500	11.859	.002
subject	6033.333	5	1206.667	6.661	.006
Error	1811.667	10	181.167		
Total	108064.000	18			
Corrected Total	12142.000	17			

a. R Squared = .851 (Adjusted R Squared = .746)

Descriptive Statistics

Dependent Variable: yij-time to exhaustion

diet groups	blocks	Mean	Std. Deviation	N
mixed fat and carbohydrates	1	84.00	.	1
	2	35.00	.	1
	3	91.00	.	1
	4	57.00	.	1
	5	56.00	.	1
	6	45.00	.	1
	Total	61.33	21.915	6
high fat	1	91.00	.	1
	2	48.00	.	1
	3	71.00	.	1
	4	45.00	.	1
	5	61.00	.	1
	6	61.00	.	1
	Total	62.83	16.762	6
high carbohydrates	1	122.00	.	1
	2	53.00	.	1
	3	110.00	.	1
	4	71.00	.	1
	5	91.00	.	1
	6	122.00	.	1
	Total	94.83	28.421	6
Total	1	99.00	20.224	3
	2	45.33	9.292	3
	3	90.67	19.502	3
	4	57.67	13.013	3
	5	69.33	18.930	3
	6	76.00	40.632	3
	Total	73.00	26.725	18

diet groups

Dependent Variable: yij-time to exhaustion

diet groups	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
mixed fat and carbohydrates	61.333	5.495	49.090	73.577
high fat	62.833	5.495	50.590	75.077
high carbohydrates	94.833	5.495	82.590	107.077

Diet

F statistic = 11.859 p-value = 0.002

α (0.05)> p-value (0.002)

Reject H0

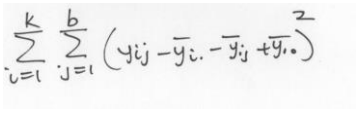
At least one mean value of 'time to exhaustion' among the three diets groups is different.

SSA= $6\sum_{i=1}^{3}(\bar{y}i. - \bar{y}..)^2$

6*((61.33-73)^2+(62.83-73)^2+(94.83-73)^2) =4297

SSB= $3\sum_{j=1}^{6}(\bar{y}.j - \bar{y}..)^2$

3*((99-73)^2+(45.33-73)^2+(90.67-73)^2+(57.67-73)^2+(69.33-73)^2+(76-73)^2) =6034

SSE=  =1812(from calculator)

Source of variation	Sum of square	D.F	mean square	F
Diet	4297	2	s_1^2 =2148.5	2148.5/181.2=11.85
Subject	6034	5	s_2^2 =1206.8	
Error	1812	10	s_3^2 =181.2	
Total	12143	17		

Table value $F_{(0.05,\ 2,15)}$ = 3.68

F calculated (11.85) >F table (3.68)

Reject H0

At least one mean value of 'time to exhaustion' among the three diets groups is different.

e)Conduct Tukey and Bonferroni test for multiple comparisons.

Multiple Comparisons

Dependent Variable: yij-time to exhaustion

	(I) diet groups	(J) diet groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	mixed fat and carbohydrates	high fat	-1.50	7.771	.980	-22.80	19.80
		high carbohydrates	-33.50*	7.771	.004	-54.80	-12.20
	high fat	mixed fat and carbohydrates	1.50	7.771	.980	-19.80	22.80
		high carbohydrates	-32.00*	7.771	.005	-53.30	-10.70
	high carbohydrates	mixed fat and carbohydrates	33.50*	7.771	.004	12.20	54.80
		high fat	32.00*	7.771	.005	10.70	53.30
Bonferroni	mixed fat and carbohydrates	high fat	-1.50	7.771	1.000	-23.80	20.80
		high carbohydrates	-33.50*	7.771	.005	-55.80	-11.20
	high fat	mixed fat and carbohydrates	1.50	7.771	1.000	-20.80	23.80
		high carbohydrates	-32.00*	7.771	.006	-54.30	-9.70
	high carbohydrates	mixed fat and carbohydrates	33.50*	7.771	.005	11.20	55.80
		high fat	32.00*	7.771	.006	9.70	54.30

Based on observed means.
The error term is Mean Square(Error) = 181.167.

*. The mean difference is significant at the .05 level.

H0: $\mu_i = \mu_j$ i, j=1,2,3
H1: $\mu_i \neq \mu_j$ i≠j

α=0.05

Tukey HSD: Means difference of “time to exhaustion “between “high carbohydrates”, “mixed fat and carbohydrates” (p-value= 0.004 <0.05 and difference= 33.5, Reject H0) is significant.

Means difference of “time to exhaustion “between” high carbohydrates”, “high fat” (p-value= 0.005<0.05and difference= 32.00, Reject H0) is significant.

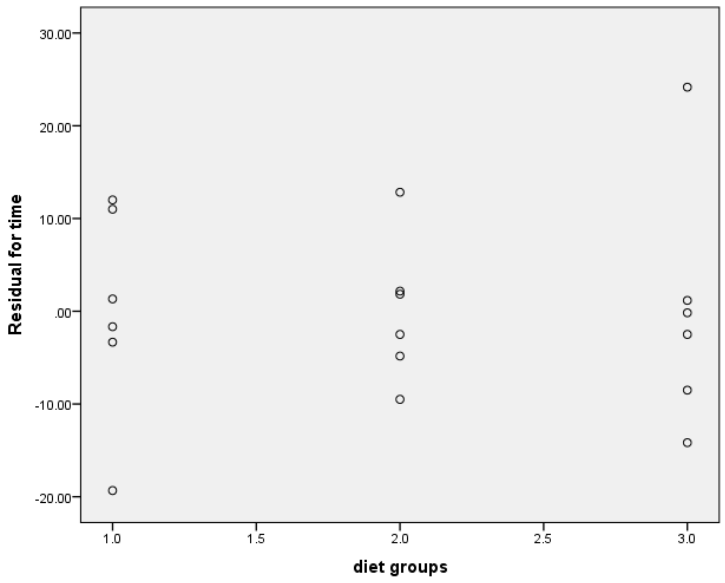
Means differences of “time to exhaustion “between” high fat”, “mixed fat and carbohydrates” (p-value= 0.980>0.05and difference= 1.5, Do not reject H0) are not significant.

Bonferroni: Means difference of “time to exhaustion “between “high carbohydrates”, “mixed fat and carbohydrates” (p-value= 0.005 <0.05 and difference= 33.50, Reject H0) is significant.

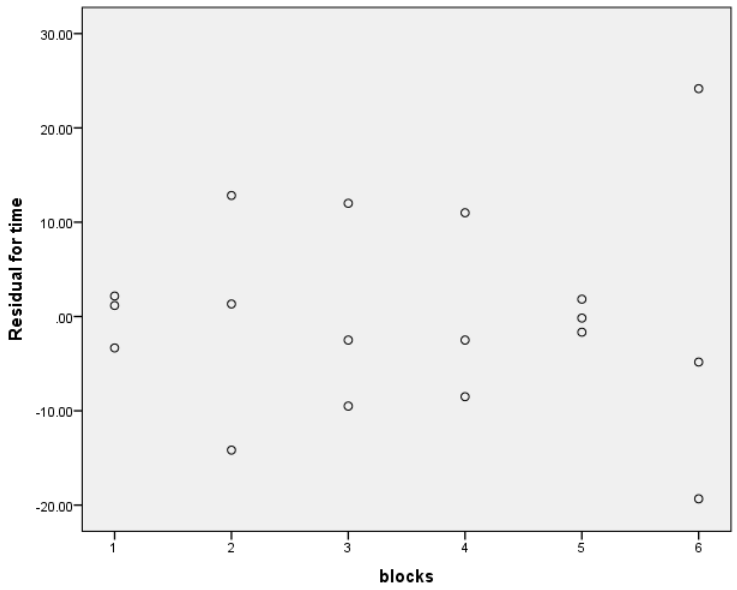
Means differences of “time to exhaustion “between” high carbohydrates”, “high fat” (p-value= 0.006<0.05and difference= 32.00, Reject H0) are significant.

Means differences of “time to exhaustion “between” high fat”, “mixed fat and carbohydrates” (p-value= 1.0>0.05and difference= 1.5, Do not reject H0) are not significant.

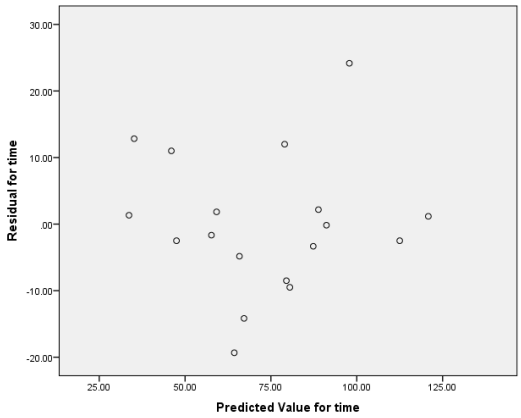
f)Comment on the homogeneity of variance of the residuals using the three residual plots.



The residual plot (diet group vs residual for time) show that most of the points are within the band -20 to +20. As well as residuals don't shows a big difference among three groups. Which means plot of residuals verses diet groups indicate error variance of residuals may be the same for all three diet groups. Therefore, variances of residuals are homogeneous among diet groups.



The residual plot (block vs residual for time) show that most of the points are within the band -20 to +20. As well as residuals show a big difference among six blocks. Residual under block six has higher variances and blocks 1 and 5 show lower variances. Which means plot of residuals verses blocks indicate error variance of residuals may not be the same for all six subjects. Therefore, variances of residuals are not homogeneous among the six blocks.



The residual plot (predicted value for time vs residual for time) show that most of the points are within the band -20 to +20. As well as residuals don't shows a pattern. Which means it is random. Therefore, variances of residuals are homogeneous.

Normality test for residuals

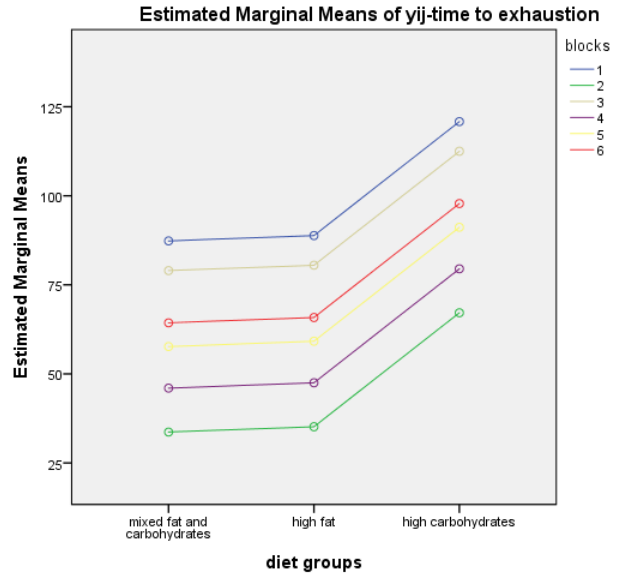
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for time	.195	18	.070	.961	18	.625

a. Lilliefors Significance Correction

KS test statistic is 0.195 not significant
p-value = 0.075>0.05
borderline significant
We can assume residuals are normally distributed.

g) Comment on the plot of mean time to exhaustion by diet separately for each block. In other words, interpret interaction effects of blocks and treatments using the interaction plot.



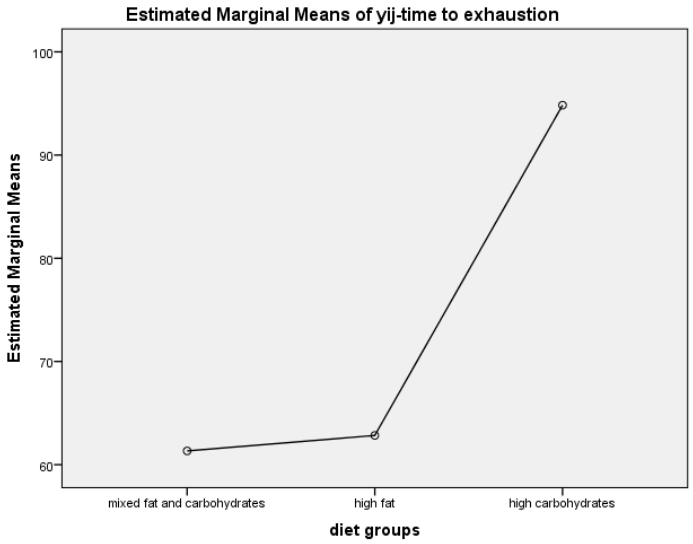
There are not interactions between six block on mean exhaustion time among three diet groups.
Block 1 has highest ‘time to exhaustion’ on mean compared with other blocks in all diet groups.
Block 2 has lowest ‘time to exhaustion’ on mean compared with other blocks in all diet groups.
Block 6 and 5 have mid ‘time to exhaustion’ on mean.

Diet group ‘high carbohydrates’ has higher ‘time to exhaustion’ on mean compared with other groups among all the blocks.
Diet group ‘mixed fat and carbohydrates’ has lower ‘time to exhaustion’ on mean compared with other groups among all the blocks.

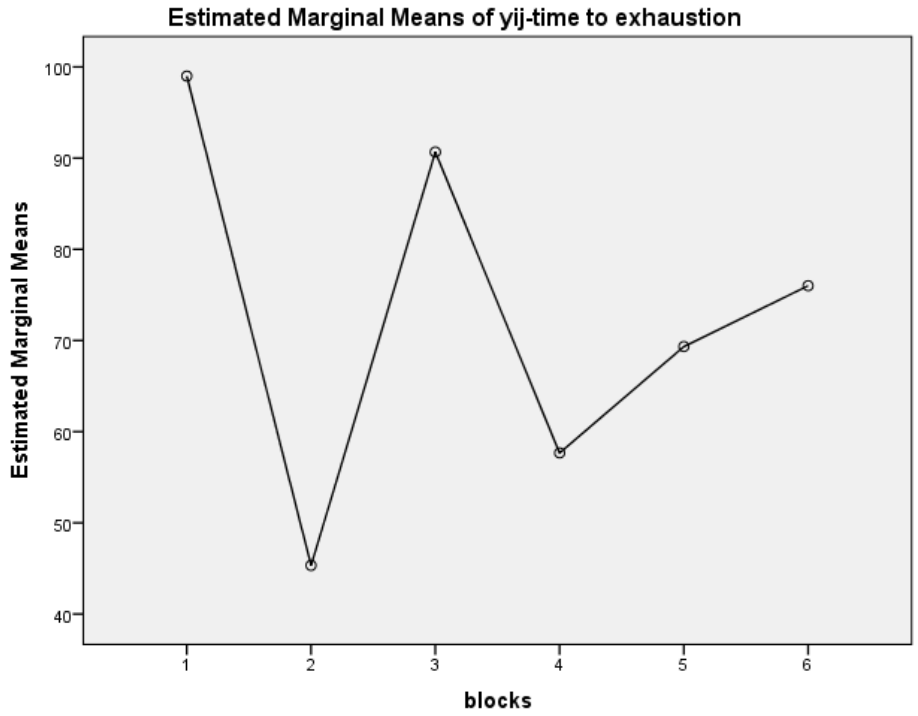
If the diet group is ‘high carbohydrates’ and block is one so that has highest mean ‘time to exhaustion’ compared with other groups and blocks.

If the diet group is ‘mixed fat and carbohydrates’ and block is two, so that has lowest ‘exhaustion time’ on mean compared with other groups and blocks.

Group ‘mixed fat and carbohydrates’ and ‘high fat’ have very close mean value for ‘time to exhaustion’ among all the blocks.



In this graph it is clearly show that there are differences in mean values of ‘time to exhaustion’ in three groups. We can see that ‘mixed fat and carbohydrates’ and ‘high fat’ groups have very closer mean values of ‘time to exhaustion’. Group ‘high carbohydrates’ has highest mean value for ‘time to exhaustion’.



In this graph it is clearly show that there are differences in mean values of variable 'time to exhaustion' among six block.

h) Is the RCB design better than the one-factor completely randomized design? Explain.

Descriptives

yij-time to exhaustion

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
mixed fat and carbohydrates	6	61.33	21.915	8.947	38.33	84.33	35	91
high fat	6	62.83	16.762	6.843	45.24	80.42	45	91
high carbohydrates	6	94.83	28.421	11.603	65.01	124.66	53	122
Total	18	73.00	26.725	6.299	59.71	86.29	35	122

ANOVA

yij-time to exhaustion

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4297.000	2	2148.500	4.108	.038
Within Groups	7845.000	15	523.000		
Total	12142.000	17			

Multiple Comparisons

Dependent Variable: yij-time to exhaustion

	(I) diet groups	(J) diet groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	mixed fat and carbohydrates	high fat	-1.500	13.204	.993	-35.80	32.80
		high carbohydrates	-33.500	13.204	.056	-67.80	.80
	high fat	mixed fat and carbohydrates	1.500	13.204	.993	-32.80	35.80
		high carbohydrates	-32.000	13.204	.069	-66.30	2.30
	high carbohydrates	mixed fat and carbohydrates	33.500	13.204	.056	-.80	67.80
		high fat	32.000	13.204	.069	-2.30	66.30
Bonferroni	mixed fat and carbohydrates	high fat	-1.500	13.204	1.000	-37.07	34.07
		high carbohydrates	-33.500	13.204	.068	-69.07	2.07
	high fat	mixed fat and carbohydrates	1.500	13.204	1.000	-34.07	37.07
		high carbohydrates	-32.000	13.204	.085	-67.57	3.57
	high carbohydrates	mixed fat and carbohydrates	33.500	13.204	.068	-2.07	69.07
		high fat	32.000	13.204	.085	-3.57	67.57

	One way ANOVA	RCB
SST	12142	12142
SSA	4297	4297
SSE	7845	1811.667
F/sig	4.108/ 0.038	11.859/0.002
Tukey/sig	High carbo and mixed fat 33.5 p-value 0.056 (borderline significant) High carbo and high fat 32.0 p-value 0.069 High fat and mixed fat 1.5 p-value 0.993 (not significant)	High carbo and mixed fat 33.5 p-value 0.004(significant) High carbo and high fat 32.0 p-value 0.005(significant) High fat and mixed fat 1.5 p-value 0.980 (not significant)

RCB design has lower SSE and higher F and pretty lower p-value. Tukey test for RCB design is showing pretty strong significant differences on mean ‘time to exhaustion’ among some diet groups. So RCB design is better than one way ANOVA.

B: Friedman's Test

Using Friedman's Test with the above data, test for differences in 'time to exhaustion' among the three diets at $\alpha = 0.05$.

Show all work AND use computer output. Include relevant computer output.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
trt1	6	61.33	21.915	35	91
trt2	6	62.83	16.762	45	91
trt3	6	94.83	28.421	53	122

Ranks

	Mean Rank
trt1	1.33
trt2	1.67
trt3	3.00

Test Statistics[a]

N	6
Chi-Square	9.333
df	2
Asymp. Sig.	.009

a. Friedman Test

H0: Population medians of 'time to exhaustion' among all three ('high carbohydrates', 'mixed fat and carbohydrates' and high fat) treatments' groups are the same.

H1: Population medians of 'time to exhaustion' among all three ('high carbohydrates', 'mixed fat and carbohydrates' and high fat) treatments' groups are not same.

Chi-Square test statistics =9.333 and p-value 0.009<0.05

Reject H0

Population medians of 'time to exhaustion' among all three ('high carbohydrates', 'mixed fat and carbohydrates' and high fat) treatments' groups are not same.

84/1	91/2	122/3
35/1	48/2	53/3
91/2	71/1	110/3
57/2	45/1	71/3
56/1	61/2	91/3
45/1	61/2	122/3
8	10	18

$g = 12/(6*(3)(4))(8^2+10^2+18^2) -3*6*4$

$g = 9.333 \quad > \quad X_{0.05,2}=5.991$

Reject H0

Population medians of ‘time to exhaustion’ among all three (‘high carbohydrates’, ‘mixed fat and carbohydrates’ and high fat) treatments’ groups are not same.