

> #1(a): Completely randomized design is appropriate because all the units of the different diets are homogenous, no extraneous variables to affect the design. It would eventually lead to a one-way analysis of the variance, which is what we're expecting.

>

> #1(b): We randomly assign the 16 animals into 4 groups, and conduct 4 different diets on each group. Then, we record the coagulation time and calculate its mean among each group. We denote the groups as A, B, C, D, and compare their means. We later run the ANOVA test between the coagulation time as the response, and each group as a categorical predictor, to determine whether there's a significant difference between each group.

>

> #1(c): The variance is constant, normality of errors, no association with the order in which experimental units are conducted, no outer influence.

~

```
> error=c(rnorm(16))
> mu=c(1,2,0,4,1,2,0,4,1,2,0,4,1,2,0,4)
> yij=mu+error
> cateogry=c(A,B,C,D,A,B,C,D,A,B,C,D,A,B,C,D)
> dfrm=data.frame(yij,category)
> lmod=lm(yij~category-1, dfrm)
> summary(lmod)
```

Call:
lm(formula = yij ~ category - 1, data = dfrm)

Residuals:

	Min	1Q	Median	3Q	Max
	-1.19549	-0.69035	0.08279	0.47766	1.41421

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
categoryA	1.053116	0.428933	2.455	0.03030	*
categoryB	1.759572	0.428933	4.102	0.00147	**
categoryC	-0.005856	0.428933	-0.014	0.98933	
categoryD	4.468642	0.428933	10.418	2.3e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8579 on 12 degrees of freedom
Multiple R-squared: 0.9163, Adjusted R-squared: 0.8884
F-statistic: 32.85 on 4 and 12 DF, p-value: 2.232e-06

```
> anova(lmod)
```

Analysis of Variance Table

Response: yij

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
category	4	96.696	24.1739	32.848	2.232e-06 ***
Residuals	12	8.831	0.7359		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> qf(0.95,4,12)
```

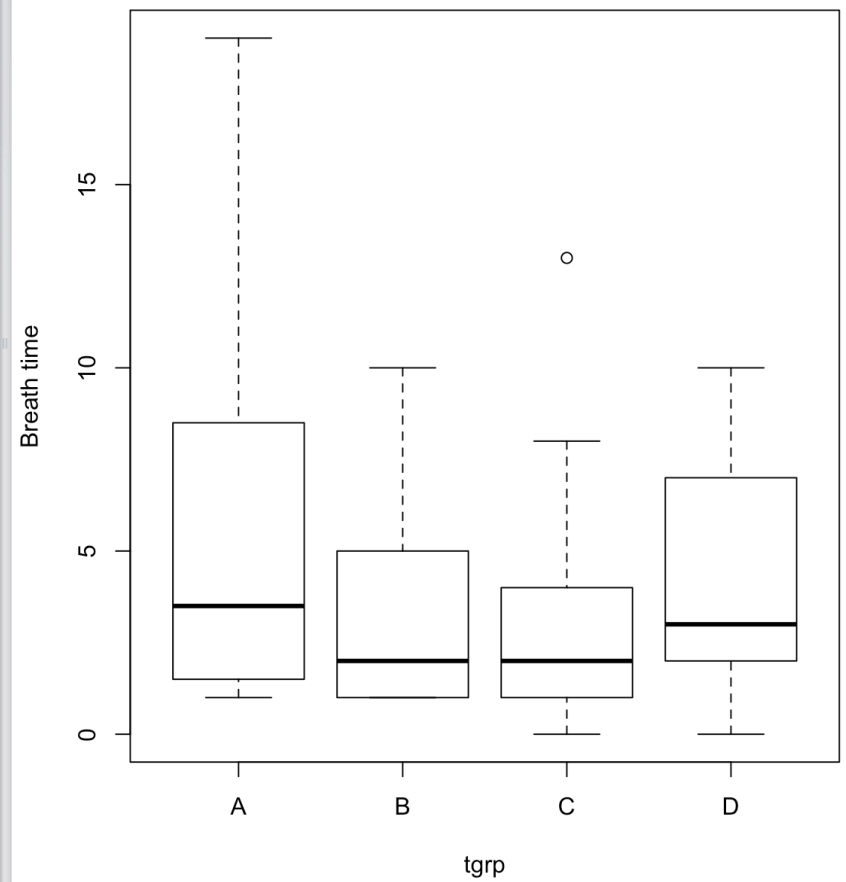
```
[1] 3.259167
```

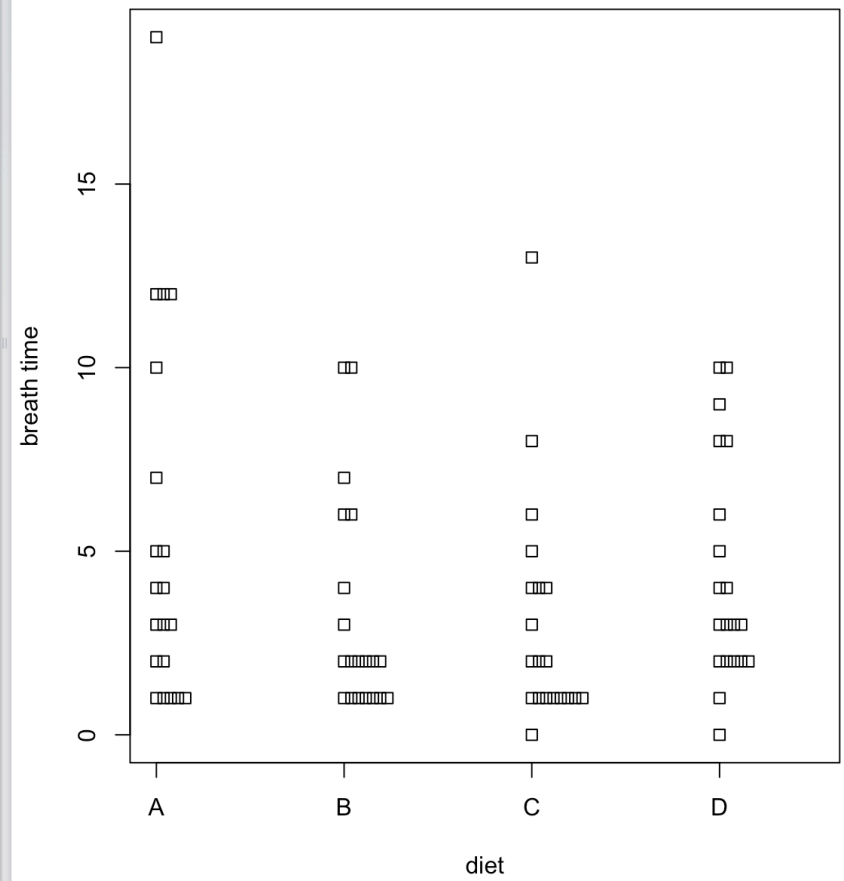
> #F-value>F0.05(4,12), so there's significant difference between the coagulation times between the different groups of diet.

```

mean the different groups of data.
> #2(a)
> par(mfrow=c(1,1))
> data(anaesthetic, package="faraway")
> plot(breath~tgrp,anaesthetic, ylab="Breath time")
> #We can see that group A has the largest range of breath time, followed by group D, group
B, then group C. On average Group A has the longest breath time, group C has the smallest.
There's an outlier on group C.

```



[illegible]

```
> #2(c)
> lmod=lm(breath~factor(tgrp)-1,anaesthetic)
> summary(lmod)
```

Call:

```
lm(formula = breath ~ factor(tgrp) - 1, data = anaesthetic)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.400	-2.250	-1.250	1.613	13.600

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
factor(tgrp)A	5.400	0.816	6.618	4.57e-09 ***
factor(tgrp)B	3.250	0.816	3.983	0.000155 ***
factor(tgrp)C	3.050	0.816	3.738	0.000357 ***
factor(tgrp)D	4.350	0.816	5.331	9.68e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.649 on 76 degrees of freedom

Multiple R-squared: 0.5732, Adjusted R-squared: 0.5507

F-statistic: 25.51 on 4 and 76 DF, p-value: 2.032e-13

```
> anova(lmod)
```

Analysis of Variance Table

Response: breath

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(tgrp)	4	1359	339.74	25.513	2.032e-13 ***
Residuals	76	1012	13.32		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> qf(0.95,4,76)
```

```
[1] 2.492049
```

```
> #F-value>F0.05(4,76),so we reject H0, there's significant difference between the four groups.
```

```
~
```

```
> #2(d)
```

```
> boxcox(lmod)
```

```
Error in boxcox.default(lmod) : response variable must be positive
```

```
> #You can't use boxcox because some response are 0, only positive response can be used in  
boxcox transformation
```

```
`
```

```
> #2(e)
> lmod1=lm(sqrt(breath)~tgrp-1, anaesthetic)
> summary(lmod1)
```

```
Call:
lm(formula = sqrt(breath) ~ tgrp - 1, data = anaesthetic)
```

```
Residuals:
```

```
      Min       1Q   Median       3Q      Max
-1.9333 -0.5547 -0.2013  0.5229  2.2561
```

```
Coefficients:
```

```
      Estimate Std. Error t value Pr(>|t|)
tgrpA    2.1028     0.1898  11.081 < 2e-16 ***
tgrpB    1.6543     0.1898   8.718 4.63e-13 ***
tgrpC    1.5547     0.1898   8.193 4.70e-12 ***
tgrpD    1.9333     0.1898  10.188 7.31e-16 ***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.8486 on 76 degrees of freedom
Multiple R-squared:  0.8295,    Adjusted R-squared:  0.8205
F-statistic: 92.43 on 4 and 76 DF,  p-value: < 2.2e-16
```

```
> anova(lmod1)
```

```
Analysis of Variance Table
```

```
Response: sqrt(breath)
```

```
      Df Sum Sq Mean Sq F value    Pr(>F)
tgrp    4 266.268   66.567   92.433 < 2.2e-16 ***
Residuals 76  54.732    0.720
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> qf(0.95,4,76)
```

```
[1] 2.492049
```

```
> #We can see that F-value>F0.05(4,76),so we reject H0, there's signifcant difference betwe
en the groups. It corresponds with the conclusion we got from above, so it's satisfactory.
```

```
> par(mfrow=c(2,2))
```

```
> plot(lmod1)
```

```
> #From the graphs, we can see indeed there's significance difference between the values f
rom the four groups.
```



```
> #3(b)
> lmod=lm(faft~fbef+rbef+raft+grp-1, hips)
> summary(lmod)
```

Call:

```
lm(formula = faft ~ fbef + rbef + raft + grp - 1, data = hips)
```

Residuals:

Min	1Q	Median	3Q	Max
-22.2573	-3.5275	-0.4578	3.9354	18.3737

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
fbef	0.44540	0.06997	6.366	1.52e-08	***
rbef	0.23370	0.13651	1.712	0.0911	.
raft	0.02838	0.13951	0.203	0.8394	
grpcontrol	58.19504	7.52018	7.739	4.36e-11	***
grptreat	65.41118	7.85054	8.332	3.35e-12	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.705 on 73 degrees of freedom

Multiple R-squared: 0.9971, Adjusted R-squared: 0.9969

F-statistic: 5086 on 5 and 73 DF, p-value: < 2.2e-16

> #We're predicting the effects on faft from all the variables without using the intercept. As we can see from the result, the only significant results are fbef, grpcontrol and grptreat. We can see the estimated value for grptreat is larger than grpcontrol, meaning the treatment does have an effect on the treatment group in terms of flexion angle.

```
> #3(c)
> critval=qt(0.025,df=df.residual(lmod)-1,lower=FALSE)
> which(abs(rstudent(lmod))>critval)
21 36 49 50 70
21 36 49 50 70
> #All the outliers are below 90 in this case, so we only need to consider the data which before treatment flexion measurements greater than 90, so the results wouldn't be skewed.
```

```
> #3(d)
> lmod1=lm(faft[fbef>90]~fbef[fbef>90]+rbef[fbef>90]+raft[fbef>90]+grp[fbef>90]-1,hips)
> summary(lmod1)
```

Call:

```
lm(formula = faft[fbef > 90] ~ fbef[fbef > 90] + rbef[fbef >
  90] + raft[fbef > 90] + grp[fbef > 90] - 1, data = hips)
```

Residuals:

Min	1Q	Median	3Q	Max
-18.4953	-3.2852	-0.2525	2.6106	15.3209

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
fbef[fbef > 90]	0.74945	0.07568	9.902	6e-15	***
rbef[fbef > 90]	0.10751	0.11465	0.938	0.35159	
raft[fbef > 90]	0.03981	0.11433	0.348	0.72872	
grp[fbef > 90]control	26.72318	7.98237	3.348	0.00131	**
grp[fbef > 90]treat	31.68713	8.42927	3.759	0.00035	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.489 on 70 degrees of freedom

Multiple R-squared: 0.9981, Adjusted R-squared: 0.998

F-statistic: 7314 on 5 and 70 DF, p-value: < 2.2e-16

> #We can see the estimated of treatment group is larger than the control group in this case, so the treatment does have an effect.

```

> #3(e)
> #From the summary, we can see the estimated size for the treatment effect for patients with before-treatment flexion angles greater than 90 degrees is 31.68713.
> confint(lmod1)

```

	2.5 %	97.5 %
fbef[fbef > 90]	0.5984974	0.9003930
rbef[fbef > 90]	-0.1211461	0.3361692
raft[fbef > 90]	-0.1882040	0.2678244
grp[fbef > 90]control	10.8028517	42.6435086
grp[fbef > 90]treat	14.8754754	48.4987913

```

> #The 95% confidence interval for the treatment effect size is 14.8754754, 48.4987913.

```