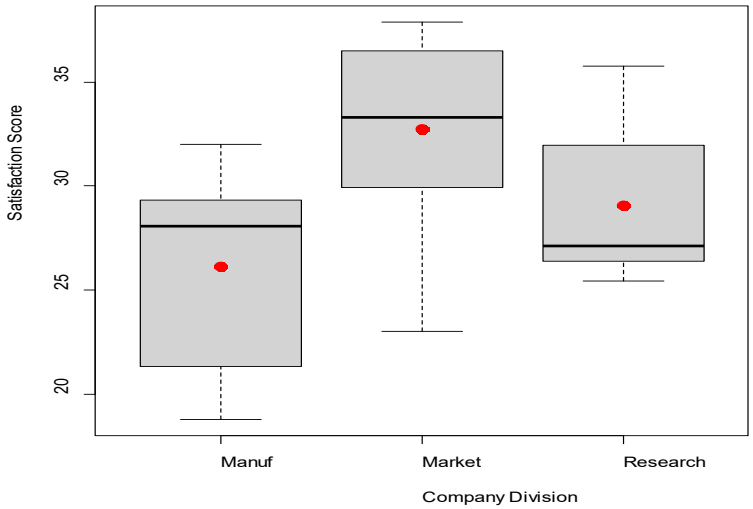
1.)

a.)

```
attach(empsat)

n.i<-apply(!apply(empsat, 2, is.na), 2, sum)

division<-as.factor(c(rep("Manuf",n.i[1]), rep("Market",n.i[2]), + rep("Research",n.i[3])))

boxplot(empsat, xlab="Company Division", ylab="Satisfaction Score")

points(1,mean(Manuf[!is.na(Manuf)]),col='red',cex=1.5,pch=19)

points(2,mean(Market),col='red',cex=1.5,pch=19)

points(3,mean(Research[!is.na(Research)]),col='red',cex=1.5,pch=19)
```

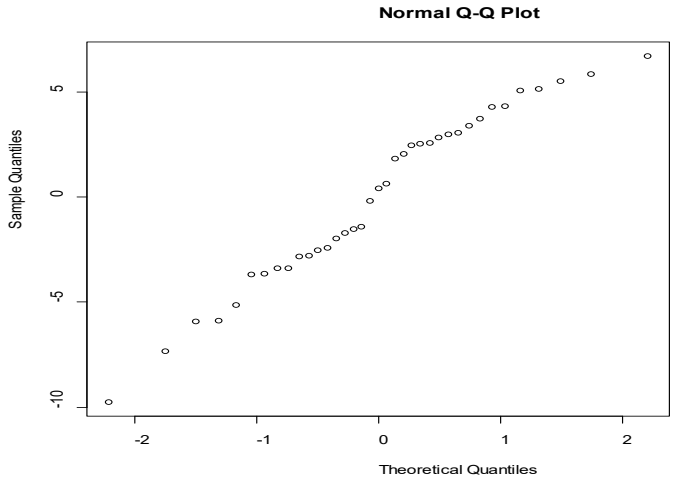


From the side-by-side boxplot above, it appears the distribution of data for the satisfaction scores in the manufacturing department is skewed left, in the marketing department is not skewed but is heavy-tailed, and is skewed right plus heavily-tailed in the research department. The variability in each dataset appears to be relatively equal. The sample medians for the manufacturing department data and the research department data appear to be nearly equal, while the sample median for the marketing department data is moderately larger. Means roughly follow the same pattern as the medians.

b.) For the ANOVA test, we are checking that we have independent random samples, the samples are normally distributed, and homogeneity of variances. In the question, we are told the samples were drawn randomly, and they are independent from each other. For the assumption of normality, we look at the normal qqlpot of the residuals given by:

```
means.sat<-
+c(rep(mean(Manuf[!is.na(Manuf)]),length(Manuf[!is.na(Manuf)])),rep(mean(Market),length(Market)),r
+ ep(mea+n(Research[!is.na(Research)]),length(Research[!is.na(Research)])))
```

```
all.sat<-c(Manuf[!is.na(Manuf)], Market[!is.na(Market)],
+ Research[!is.na(Research)])
res.sat<-all.sat-means.sat
qqnorm(res.sat)
```

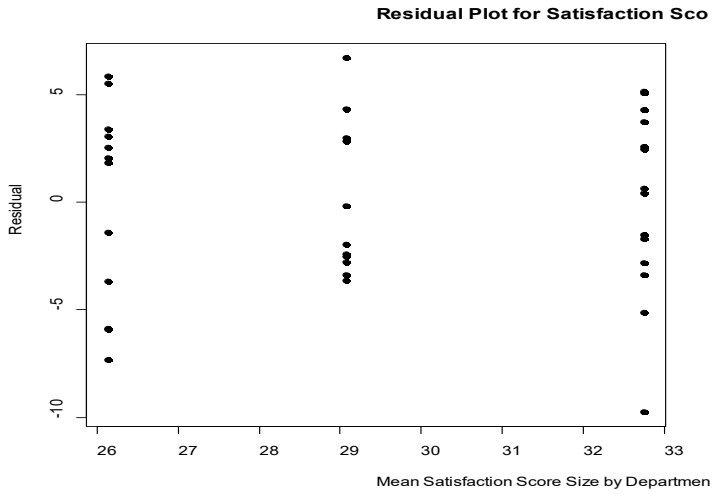


Here we see a fairly linear curve when plotting the residuals, so we can assume reasonable normality. As for the assumption of equal variances, we look to see the ratio of the largest and smallest standard deviations in the samples given by:

```
max.sd<-max(sd(Manuf[!is.na(Manuf)]), sd(Market), sd(Research[!is.na(Research)]))
min.sd<-min(sd(Manuf[!is.na(Manuf)]), sd(Market), sd(Research[!is.na(Research)]))
max.sd/min.sd
[1] 1.294883
```

This falls in our “reasonably equal” range of 0.5 and 2 since 0.5 < 1.294883 < 2. Additionally, let us take a look at the plot of residuals:

```
plot(means.sat, res.sat, xlab="Mean Satisfaction Score Size by Department",
+ ylab="Residual", main="Residual Plot for Satisfaction Scores", pch=19)
```



Here, the spread looks even enough to not have a doubt of homogeneity of variances, thus we can assume equal variances (Levene Test also corroborates this assumption with $p \approx 0.77$).

Even if the data is not exactly normal, or variances not exactly equal, ANOVA is very robust against moderate departures from these assumptions, so I will conclude that an ANOVA test is valid here.

c.)

Hypotheses:

$$H_0: \mu_1 = \mu_2 = \mu_3, \;\; H_a: Not\ all\ three\ means\ are\ equal$$

Since assumptions were said to be met in (b) ANOVA will be used

$F = 7.966, \;\; p = 0.00145$ (output posted below)

Since $p < \alpha\ (0.05)$ we reject H_0, and we can say that at the 0.05 significance level, we have sufficient evidence to conclude that not all three mean satisfaction scores for respective departments at this company are equal.

summary(aov(all.sat~division))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
division	2	286.3	143.16	7.966	0.00145	**
Residuals	34	611.1	17.97			

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

2.

a.) Assumptions for ANOVA were said to be met so they will not be checked here.

```
attach(pain)
all.pain<-c(LBL,DBL,LBR,DBR)
color<-as.factor(c(rep("LBL",5),rep("DBL",5),rep("LBR",5),rep("DBR",5)))
 pairwise.t.test(all.pain, color,  p.adjust.method="bonferroni")
```

Pairwise comparisons using t tests with pooled SD

data: all.pain and color

	DBL	DBR	LBL
DBR	0.1062	-	-
LBL	0.8696	0.0043	-
LBR	0.3307	1.0000	0.0144

P value adjustment method: Bonferroni

Let $p_{x,y}$ denote the p value when comparing means for x and y. Here, we have

$p_{DBR,DBL} = 0.1062 \;\; p_{LBL,DBL} = 0.8696 \;\; p_{LBL,DBR} = 0.0043 \;\; p_{LBR,DBL} = 0.3307$

$p_{LBR,DBR} = 1.0000 \;\; p_{LBR,LBL} = 0.0144$

The only adjusted p-values below $\alpha_E = 0.05$ are $p_{LBL,DBR}$ and $p_{LBR,LBL}$. Thus we conclude that there is a statistically significant difference in mean pain threshold between light blondes and dark brunettes, and between light brunettes and light blondes.

b.)

Here, I will utilize a Scheffé test of contrast hypotheses, working under the assumptions given in the question for the ANOVA test.

Hypotheses:

$H_0: \frac{\mu_{LBL} + \mu_{LBR}}{2} - \frac{\mu_{DBL} + \mu_{DBR}}{2} = 0 \;\; H_a: \frac{\mu_{LBL} + \mu_{LBR}}{2} - \frac{\mu_{DBL} + \mu_{DBR}}{2} \neq 0$

$l = -2.5 \;\; p = 0.9263$

```
output<-aov(all.pain~color)
a.vec<-c(1/2,-1/2,-1/2,1/2)
```

ScheffeTest(output, contrasts=matrix(a.vec,nrow=4))

Posthoc multiple comparisons of means: Scheffe Test

95% family-wise confidence level

$color

	diff	lwr.ci	upr.ci	pval
DBL,LBR-DBR,LBL	-2.5	-14.00604	9.006036	0.9263

Fail to reject H_0

Here, $p > \alpha = 0.05$, so we can say that at the 0.05 significance level, there is insufficient evidence to conclude that there is a difference in mean pain threshold scores between blondes and brunettes.

3.

a.)

Hypotheses:

$H_0: \mu_C = \mu_{C1A} = \mu_{C2A} = \mu_{C1B} = \mu_{C2B}$ $H_a: Not\ all\ five\ means\ are\ equal$

F = 81.67 p = 5.6e-14

attach(turkeys)

group<-as.factor(c(rep("Control",6), rep("C1A",6),rep("C2A",6),rep("C1B",6), rep("C2B",6)))

summary(aov(Gain~group))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
turkeys$Group	4	103.04	25.760	81.67	5.6e-14 ***
Residuals	25	7.88	0.315		

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

Reject H_0

Since $p < \alpha = 0.05$, we can say that at the 0.05 significance level, there is sufficient evidence to conclude that the mean weight gain of turkeys across all five different types of diet over the given time period are not equal.

b.)

```
pairwise.t.test(Gain, group, p.adjust.method="bonferroni")
```

Pairwise comparisons using t tests with pooled SD

data: Gain and group

	C1A	C1B	C2A	C2B
C1B	0.00098	-	-	-
C2A	0.00112	1.00000	-	-
C2B	7.5e-11	1.1e-06	9.4e-07	-
Control	0.00017	3.8e-09	4.2e-09	2.1e-14

All pairs of means are statistically significantly different, except μ_{C2A} and μ_{C1B}.

c.)

```
DunnettTest(Gain, group, control="Control")
```

Dunnett's test for comparing several treatments with a control :

95% family-wise confidence level

$Control

	diff	lwr.ci	upr.ci	pval	
C1A-Control	1.716667	0.8710423	2.562291	6.0e-05	***
C1B-Control	3.216667	2.3710423	4.062291	3.4e-10	***
C2A-Control	3.200000	2.3543756	4.045624	1.1e-09	***
C2B-Control	5.600000	4.7543756	6.445624	< 2e-16	***

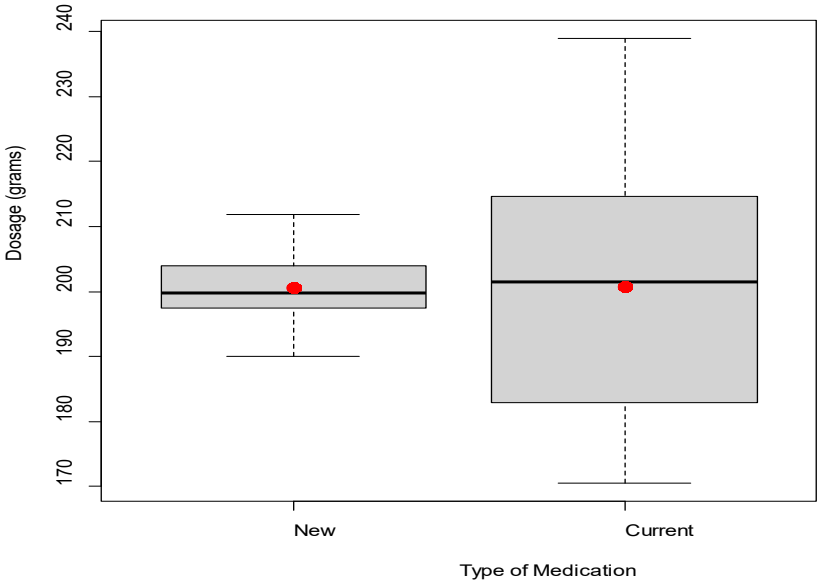
Based on the confidence intervals and p-values given by the test above, we can determine that all additive diets result in a significantly different mean weight gain compared to the control.

4.

a.)

To develop an idea of the data, we first look at a boxplot of the data for new and current medications

```
boxplot(meds, xlab="Type of Medication", ylab="Dosage (mg)")
 points(1,mean(New),col='red',cex=1.5,pch=19)
 points(2,mean(Current),col='red',cex=1.5,pch=19)
```



Here, we can see that both medication data seem to be symmetric, with the data for the new medication being much less variable, corresponding to a shorter range of actual dosages we see in the pills. The smallest and largest dose for the sample of new medication is 190mg and 211.9mg respectively. As for the current medication, its smallest and largest dose as taken from the sample data of 20 is 170.5mg and 239mg respectively. The boxplot shows the median of each dataset, with the mean labeled as red points. Their numerical values are given below.

mean(New)	median(New)
[1] 200.525	[1] 199.7
mean(Current)	median(Current)
[1] 200.755	[1] 201.45

As we can see, both sets of data have comparable measures of central tendency, or “average” value. To look at the numerical side of the spread, we can use corresponding values for standard deviation and variance as given below.

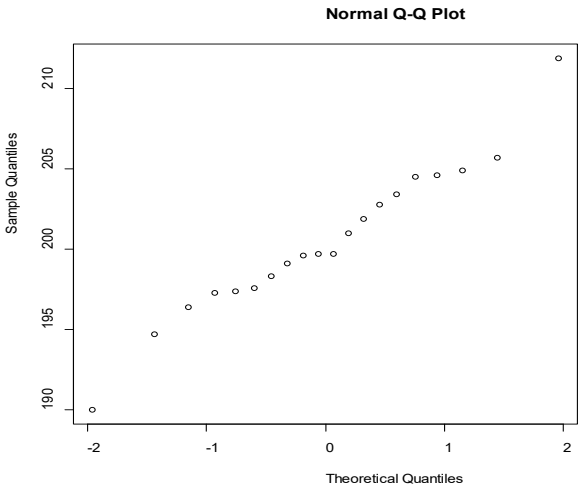
sd(New)	var(New)
[1] 4.717925	[1] 22.25882
sd(Current)	var(Current)
[1] 18.86374	[1] 355.8405

As a rule of thumb, we see if the ratio of the standard deviations falls in the range $0.5 < ratio < 2$ to gauge if the variances are nearly the same. In this case the ratio can be seen as $\frac{\sigma_{New}}{\sigma_{Current}} = \frac{4.718}{18.864} = 0.25$. This leads us to believe there is a significant difference in the variability of the dosage in the new medication vs. the current medication. This corresponds to there being a higher chance of the pill dosage being far away from 200mg for the current medication (not good!).

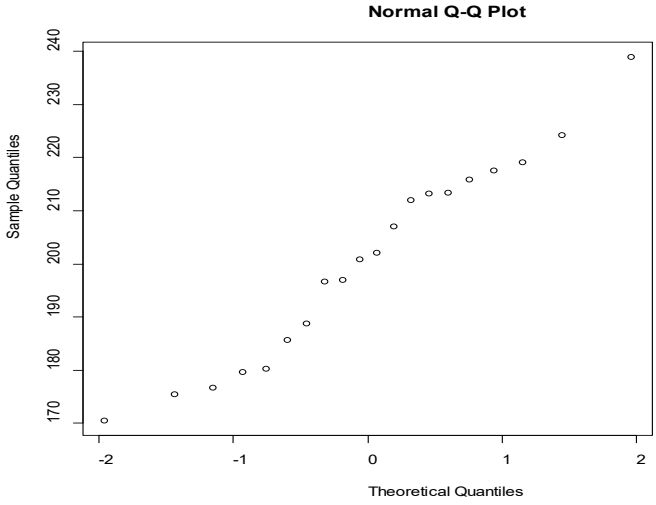
b.)

The goal of this study is to not only determine if the true "average" value of each pill dosage is 200mg, but also to assess the variability of the dosage. For assumption purposes, this study was conducted by collected independent random samples (as stated in the question). To assess the normality of each dataset we look at the normal qqplots given by

qqnorm(New)



qqnorm(Current)



Both appear to have some degree of nonlinearity, but no so much to doubt the assumption of nonnormality. Alongside this, both boxplots appear to be nearly symmetric.

Our first test here is to determine if the true "average" dosage of pills for new and current medications differs from 200mg. If we can find evidence to support a different average value than this, it is strong evidence in disfavor of the given type of medication, as we want to ensure people are generally getting what we say they are. Since both datasets appear to be normally distributed with moderate sample size (n=20), I will appeal to the robustness of the t test, and use a 95% confidence level ($\alpha = 0.05$). The hypotheses for both tests is the same and are given by

$$H_0: \mu = 200 \quad H_a: \mu \neq 200$$

t.test(New, mu=200) (one sample t test for new medication)

One Sample t-test

data: New

t = 0.49765, df = 19, p-value = 0.6244

alternative hypothesis: true mean is not equal to 200

95 percent confidence interval:

198.3169 202.7331

sample estimates:

mean of x

200.525

t.test(Current, mu=200) (one sample t test for current medication)

One Sample t-test

data: Current

t = 0.17899, df = 19, p-value = 0.8598

alternative hypothesis: true mean is not equal to 200

95 percent confidence interval:

191.9265 209.5835

sample estimates:

mean of x

200.755

In the results of our tests above, both tests gave a p-value greater than our level of significance with $p_{New} = 0.6244, p_{Current} = 0.8598$. This leads us to fail to reject the null in both cases. Thus we can say that at the 0.05 significance level, there was insufficient evidence in both cases to support the claim that

the mean pill dosage in mg differs from 200. This is good news for both medications, but the analysis is not complete. Let us look at the 95% confidence intervals for both.

95% confidence interval for new medication : $(198.3169, 202.77331)$

t.test(New, mu=200)$conf.int

[1] 198.3169 202.7331

95% confidence interval for current medication: $(191.9265, 209.5835)$

t.test(Current, mu=200)$conf.int

[1] 191.9265 209.5835

These intervals correspond to us being 95% confident that the true mean dosage for the new medication is between 198.3169mg and 202.77331mg, and that the true mean dosage for the current medication is between 191.9265mg and 209.5835mg. It is important to note that the interval for the new medication is significantly tighter, so we are more confident that the true mean dosage for the new medication lies closer to 200mg. Now, let us look at how the "average" values differ from each other using a two-sample t test with hypotheses:

$H_0: \mu_{New} - \mu_{Current} = 0 \quad H_a: \mu_{New} - \mu_{Current} \neq 0$

t.test(New,Current, var.equal=F)

Welch Two Sample t-test

data: New and Current

t = -0.052898, df = 21.368, p-value = 0.9583

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-9.262662 8.802662

sample estimates:

mean of x mean of y

200.525 200.755

The resulting conclusion is that we fail to reject the null hypothesis $p > \alpha$

So at the 0.05 significance level, there is insufficient evidence to conclude that the true mean dosage for the new medication and old medication differ.

Up to this point, we have concluded that there is insufficient evidence of either mean dosages being different from 200mg, and also insufficient evidence of the true mean dosages being different from each other. Now, let us take a look at the variability in each drug dosage, and compare it to one another. I will utilize the F test working under the assumption of normality with the hypotheses:

$$H_0: \frac{\sigma_{New}^2}{\sigma_{Current}^2} = 1 \quad H_a: Variances\ are\ not\ equal$$

var.test(New, Current)

F test to compare two variances

data: New and Current

F = 0.062553, num df = 19, denom df = 19, p-value = 1.227e-07

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.02475915 0.15803655

sample estimates:

ratio of variances

0.06255279

Reject H_0

Here, we have $p < \alpha = 0.05$, so we can say that at the 0.05 significance level, there is sufficient evidence to conclude that the variances for dosages (in mg) for the new medication and the old medication are not equal. Additionally, the 95% confidence interval provides more detail, and is given by

sqrt(var.test(New,Current)$conf.int)

[1] 0.1573504 0.3975381

Since the entire interval lies below 1, we can say with 95% confidence that $\sigma_{Current}$ is between 1.51548216 and 5.35524282 larger than σ_{New}.

This is a very significant result in that we can now be confident that the current medication used is more variable in dosage than the new medication, and by order up to 5.

c.)

Based on the statistical findings in part b.) I would recommend to the Board of Directors that they switch medication from current to new. This is because we can be much more confident that across the entire population of pills produced, the dosages much more accurately represent a dosage of 200mg than does the current medication being used. On top of this, since the current medication is much more variable,

very low and very high doses (relative to 200mg) can be seen much more often, and this can result in at least negative consumer feedback, and at worst severe health risks.