Homework 5

First we'll read in the data.

```
dat<-read.csv("homework5.csv")
```

Now let's check the structure of the data we have.

```
str(dat)
```

```
## 'data.frame':    64 obs. of  2 variables:
##  $ music: Factor w/ 2 levels "metal","none": 2 2 2 2 2 2 2 2 2 2 ...
##  $ wpm  : num  23.5 44.1 42.3 24.4 42.8 ...
```

The data contains the number of words per minute typed while listening to metal music or nothing.

Let's find the mean wpm for those who listened to metal and those who listened to nothing.

```
mean(dat$wpm[dat$music=="metal"])
```

```
## [1] 48.03665
```

```
mean(dat$wpm[dat$music=="none"])
```

```
## [1] 42.44641
```

The metal group typed approximately 48 wpm, while the group that listened to nothing typed approximately 42 wpm.

Let's run an independent t-test to see if the above differences are significant.

```
t.test(dat$wpm[dat$music=="metal"],
       dat$wpm[dat$music=="none"],
       var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  dat$wpm[dat$music == "metal"] and dat$wpm[dat$music == "none"]
## t = 2.3317, df = 62, p-value = 0.02298
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.7977514 10.3827312
## sample estimates:
## mean of x mean of y
##  48.03665  42.44641
```

The t-test reveals that the participants listening to metal typed more wpm than those not listening to anything, $t(62) = 2.3$, $p = .023$.

Although we should really only run one t-test, to demonstrate that I can run directional t-tests, I'm conducting one here with the null hypothesis, that the metal group types as fast or slower than the group with no music.

```
t.test(dat$wpm[dat$music=="metal"],
       dat$wpm[dat$music=="none"],
       var.equal = TRUE,
       alternative = "greater")
```

```
##
##  Two Sample t-test
##
## data:  dat$wpm[dat$music == "metal"] and dat$wpm[dat$music == "none"]
## t = 2.3317, df = 62, p-value = 0.01149
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.586923      Inf
## sample estimates:
## mean of x mean of y
##  48.03665  42.44641
```

As expected this also indicated that listening to metal increases typing wpm over no music, $t(62) = 2.3$, $p = .011$

In order to create a figure of the data, I need to load the ggplot2 and dplyr libraries.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.3
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##     filter, lag
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

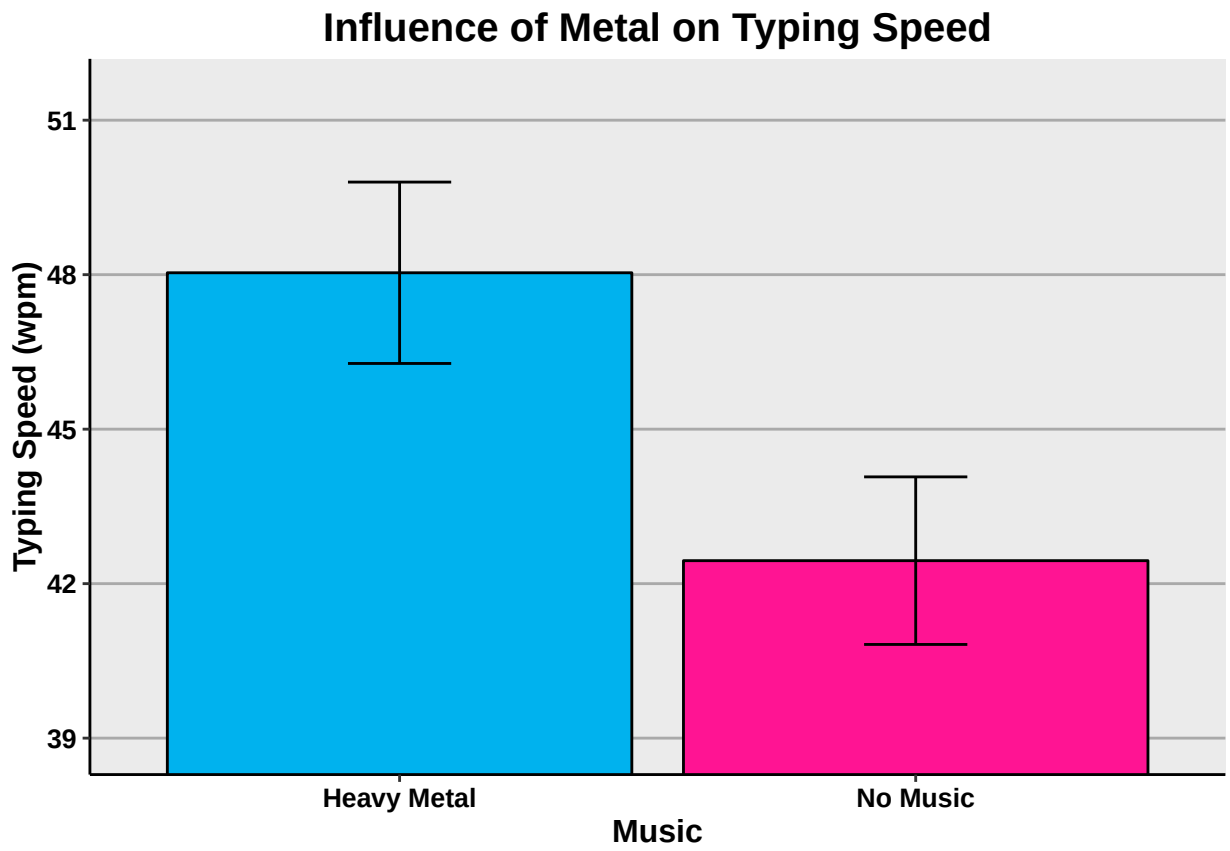
Next I need to use dplyr to summarize the data and find the means and standard errors of the means for the different groups.

```
temp<-dat%>%group_by(music)%>%
  summarize(means=mean(wpm), sems=sd(wpm)/sqrt(length(wpm)))
```

Armed with the summarized data I can now create a figure of the data.

```
f<-ggplot(temp,aes(x=factor(music),y=means))+
  geom_bar(stat="identity", color="black",
           fill=c("deepskyblue2", "deeppink"))+
  geom_errorbar(aes(ymax=means+sems,
```

```
                    ymin=means-sems),
                width=.2)+
  ggtitle("Influence of Metal on Typing Speed")+
  labs(x="Music", y="Typing Speed (wpm)")+
  scale_x_discrete(breaks=c("none","metal"),
                   labels=c("No Music","Heavy Metal"))+
  theme(plot.title=element_text(size=15,
                                face="bold",
                                vjust=.5))+
  theme(axis.title.x=element_text(size=12,
                                  face="bold",
                                  vjust=-.25))+
  theme(axis.title.y=element_text(size=12,
                                  face="bold",
                                  vjust=.25))+
  theme(axis.text.x=element_text(size=10,
                                face="bold",
                                color="black"))+
  theme(axis.text.y=element_text(size=10,
                                face="bold",
                                color="black"))+
  coord_cartesian(ylim=c(min(temp$means)-2*max(temp$sems),
                         max(temp$means)+2*max(temp$sems)))+
  theme(panel.border=element_blank(),
        axis.line=element_line())+
  theme(panel.grid.major.x=element_blank())+
  theme(panel.grid.major.y=element_line(color="darkgrey"))+
  theme(panel.grid.minor.y=element_blank())
f
```



Here, I'm running a two-tailed paired t-test to show that I know how to. I mention this because I really should only run a single test on the data. But hey, I'm the teacher and I can do that.

```
t.test(dat$wpm[dat$music=="metal"],
       dat$wpm[dat$music=="none"],
       var.equal = TRUE,
       paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  dat$wpm[dat$music == "metal"] and dat$wpm[dat$music == "none"]
## t = 12.153, df = 31, p-value = 2.51e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  4.652093 6.528390
## sample estimates:
## mean of the differences
##                5.590241
```

With the paired test, the difference between wpm and music is even more powerful, $t(31) = 12.2$, $p < .05$.