Table 1: Table 1: Fist lines of Carbon Fibre Dataset

X	additive	temperature	strength
1	a	Low	76
2	a	Low	90
3	a	Low	89
4	a	Low	105
5	a	Low	89
6	a	High	74

Table 2: Table 2: Anovatab for Main Effects Model

	df	Sum Sq	Mean Sq	F value	Pr(>F)
Model	3	8267.500	2755.8333	8.161278	0.0005418
Error	26	8779.467	337.6718	NA	NA
Total	29	17046.967	NA	NA	NA

Carbon Fibre Data

Introduction

This dataset have a response called strength which represents the strength of a carbon fibre material, and two categorical predictors: temperature (high and low), and additive(a, b, c or d).

Its sampling is equally divided between its predictors as it consists into 30 observations: 10 for each of the 3 additives, 15 for each temperature level. Below, on Table 1, you can see the first rows of the data:

The objective here is identify the relationship between those two predictors and strenght. Let's first start with a Main Effects Model to understand the relationship between those predictors and the response:

Model 1: Main Effects Model

Our Main Effects Model is the following:

$$y_i = \beta_0 + \beta_1\delta_{i,b} + \beta_2\delta_{i,c} + \beta_3\gamma_{i,low} + \varepsilon_i$$

Where $\delta_{i,b}$ will be 1 in case of using additive b, or zero otherwise, same for $\delta_{i,c}$ for additive c. If both δ are zero, then additive a is being used.
Also $\gamma_{i,low}$ will be 1 in case a low temperature was used, otherwise, it will be 0. We can start accessing the relevance of this model by checking if any of the predictors have an effect on strength.
$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

$H_\alpha : \beta_j \neq 0$ for some j ϵ{1,2,3}

Based on the anovatab above, with p-value 0.0005418, we reject H_0 and conclude that at least one predictor affects strength.
We can also check the coeficcients of the model and see if they are significant:

Based on the P values we can see that all predictors are significant on the current model and they all reduce the strength. An important detail is that this model assumes that temperature and additive don't iteract with each other. To consider that, we need to move to an Iteration model.

Table 3: Table 3: Coeficients for Main Effects Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	103.46667	6.709911	15.419976	0.0000000
additiveb	-17.60000	8.217929	-2.141659	0.0417614
additivec	-22.20000	8.217929	-2.701411	0.0119911
temperatureLow	-27.13333	6.709911	-4.043769	0.0004170

Table 4: Table 4: Coeficients for Iteraction Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	90.0	6.699751	13.4333346	0.0000000
additiveb	8.4	9.474879	0.8865549	0.3841164
additivec	-7.8	9.474879	-0.8232295	0.4184829
temperatureLow	-0.2	9.474879	-0.0211084	0.9833337
additiveb:temperatureLow	-52.0	13.399503	-3.8807411	0.0007118
additivec:temperatureLow	-28.8	13.399503	-2.1493335	0.0419023

Model 2, Iteraction Model

Our new model with iteration terms is the following:

$$y_i = \beta_0 + \beta_1\delta_{i,b} + \beta_2\delta_{i,c} + \beta_3\gamma_{i,low} + \beta_4\delta_{i,b}\gamma_{i,low} + \beta_5\delta_{i,c}\gamma_{i,low} + \varepsilon_i$$

Where this part $\beta_4\delta_{i,b}\gamma_{i,low} + \beta_5\delta_{i,c}\gamma_{i,low}$ is the iteration part of the model.

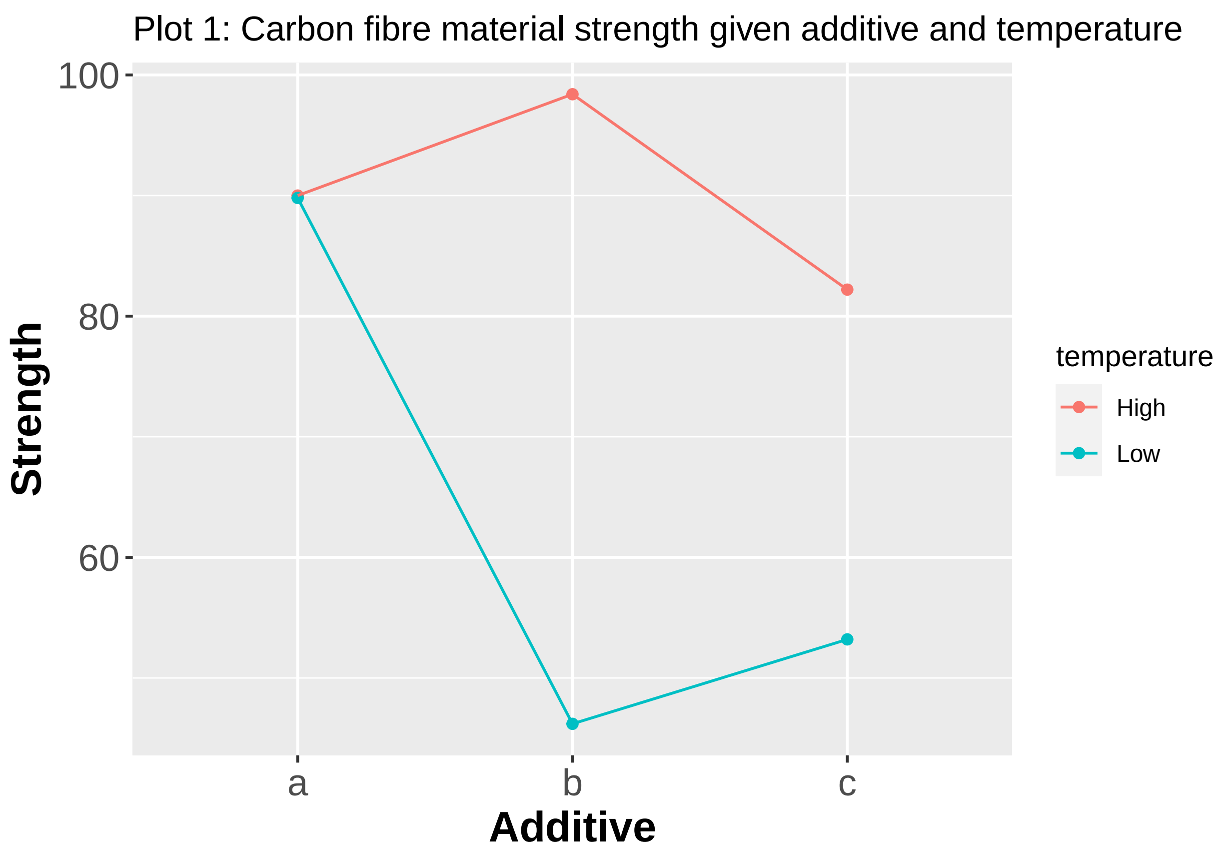
```
## Single term deletions
##
## Model:
## strength ~ additive + temperature + additive:temperature
##                      Df Sum of Sq    RSS    AIC F value   Pr(>F)
## <none>                            5386.4 167.71
## additive:temperature  2    3393.1 8779.5 178.37  7.5592 0.002844 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In this new model, we can see that the iteraction `additive:temperature` is significant (p-value 0.002844). Although we can see on the coefficients became somewhat problematic with $\beta_1, \beta_2, \beta_3$ with p-values > 0.05.

As the iteraction part works and its predictors are significantly higher in value, let's ignore for now the problematic $\beta_1, \beta_2, \beta_3$ and see a plot of this model.

Table 5: Table 5: Pairwise Contrasts for Iteraction Model

contrast	estimate	SE	df	t.ratio	p.value
a High - b High	-8.4	9.474879	24	-0.8865549	0.4965626
a High - c High	7.8	9.474879	24	0.8232295	0.4965626
a High - a Low	0.2	9.474879	24	0.0211084	0.9833337
a High - b Low	43.8	9.474879	24	4.6227503	0.0004288
a High - c Low	36.8	9.474879	24	3.8839546	0.0018616
b High - c High	16.2	9.474879	24	1.7097844	0.1670098
b High - a Low	8.6	9.474879	24	0.9076633	0.4965626
b High - b Low	52.2	9.474879	24	5.5093051	0.0001728
b High - c Low	45.2	9.474879	24	4.7705094	0.0004288
c High - a Low	-7.6	9.474879	24	-0.8021211	0.4965626
c High - b Low	36.0	9.474879	24	3.7995208	0.0018709
c High - c Low	29.0	9.474879	24	3.0607251	0.0100680
a Low - b Low	43.6	9.474879	24	4.6016418	0.0004288
a Low - c Low	36.6	9.474879	24	3.8628461	0.0018616
b Low - c Low	-7.0	9.474879	24	-0.7387957	0.5005664



Here we can see that additive a has nearly identical good performance in high or low temperatures, whereas additive b has the best performance of them all in a high temperature and the worst performance in a low temperature. We can also see that additive c performs badly on low temperature and better on high temperature but not with extreme difference as additive b.

On the pairwise contrast table we can see the statistic significance between different temperatures and additives. As seen in the chart, there is no difference between additive a with high or low temperatures. Also, most importantly to notice, there is no difference in strength between additive b with high temperature and additive a (either with high or low temperature), both pairs have a very high p-value of 0.496. In fact, there is not even evidence that b at high temperature is better than c at a high temperature (p-value 0.167)

\begin{table}

\caption{Table 6: Pairwise 95% CI for Iteraction Model}

additive	temperature	emmean	SE	df	lower.CL	upper.CL
a	High	90.0	6.699751	24	76.17239	103.82761
b	High	98.4	6.699751	24	84.57239	112.22761
c	High	82.2	6.699751	24	68.37239	96.02761
a	Low	89.8	6.699751	24	75.97239	103.62761
b	Low	46.2	6.699751	24	32.37239	60.02761
c	Low	53.2	6.699751	24	39.37239	67.02761

\end{table} We can also confirm looking at the 95% CI of each of each predictor combination how additive has almost identical confidence intervals for high and low temperature and how much a(high and low), b(high) and c(high) overlap on its confidence intervals.

Conclusion

We've concluded with our interaction model that all additives a, b, and c are equally good on high temperature. Additive a though, seems to perform also as good as the others in low temperature. There is big evidence also that additive b and c perform poorly in creating carbon fibre material strength in low temperatures

R Code:

```
# setwd as the folder where this script is in
current_path = rstudioapi::getActiveDocumentContext()$path
setwd(dirname(current_path))
```

```
# Setup that will be used across the rest of this notebook:
#Install and load kable extra to render tables in a nicer way
#install.packages("kableExtra")
library(kableExtra)
#install and load tidyverse to do dataset manipulations
#install.packages("tidyverse")
library(tidyverse)
source("anovatab.R")
# set a seed to gel always the same "random" results
set.seed(29011987)

# read carbonfibre cvs into a a dataframe
carbon_fibre <- read.csv("carbon_fibre.csv", header=T)
# summary of dataset just to see what is in the data:
summary(carbon_fibre)
```

```
head(carbon_fibre) %>%
  kbl(caption = "Table 1: Fist lines of Carbon Fibre Dataset") %>%
  kable_styling()
```

```
# Transform categorical variables into factor in the dataset
carbon_fibre$additive <- factor(carbon_fibre$additive)
carbon_fibre$temperature <- factor(carbon_fibre$temperature)
# Fit main effect model
fit1=lm(strength~additive+temperature, data = carbon_fibre)
anovaTab1 <- anovatab(fit1)
```

```
anovaTab1 %>%
  kbl(caption = "Table 2: Anovatab for Main Effects Model") %>%
  kable_styling()
```

```
summary(fit1)[["coefficients"]] %>%
  kbl(caption = "Table 3: Coeficients for Main Effects Model") %>%
  kable_styling()
```

```
# Add iteration to the model
fit2=update(fit1,.~.+additive:temperature)
drop1(fit2,test='F')
```

```
summary(fit2)[["coefficients"]] %>%
  kbl(caption = "Table 4: Coeficients for Iteraction Model") %>%
  kable_styling()
```

```r
library(emmeans);library(ggplot2)

g=emmip(fit2,temperature ~ additive)
g+theme(axis.text=element_text(size=14),
        axis.title=element_text(size=16,face="bold"))+
ylab("Strength")+
  xlab ('Additive')+
  labs(title="Plot 1: Carbon fibre material strength given additive and temperature")
```

```r
emmeansFit2 <- emmeans(fit2,pairwise~additive:temperature,adjust='fdr')
emmeansFit2$contrasts %>%
kbl(caption = "Table 5: Pairwise Contrasts for Iteraction Model") %>%
  kable_styling()
```

```r
emmeansFit2$emmeans %>%
kbl(caption = "Table 6: Pairwise 95% CI for Iteraction Model") %>%
  kable_styling()
```