Final_Report

1. Project Description

Moira Foster was tasked with designing an experiment that measured helmet foam durability over long periods of time. The experiment contained three different chemical compositions of foam with two different porosity levels, to create six types of foam. Each of these foams were compressed at different frequencies, strains, and stresses over a period of time. Dr. Foster has hired **enter group name here.** to deterimine which foam type has the most consistent static stiffness over time with impact. The sample contains 71 different foam samples across the 6 combinations.

1.1 Research Questions

Question 1: Which combination of porosity and chemical index is best to predict static stiffness over time after accounting for all other variables?

1.2 Variables

What is (are) possible explanatory and response variables? Relevant notes about how each is measured/recorded.

A table is recommended here.

To ensure that we are extracting the information needed, we created two new variables to be used in our analysis. First, we created a time variable that was calculated by taking the number of squishes performed on a foam sample in that current cycle and dividing it by the frequency of those squishes. Next, we created a relative static stiffness variable to act as our response variable. This was done by taking every measurement of static stiffness for a given foam sample and dividing it by the initial value of static stiffness for that foam sample. This helps with seeing how the static stiffness changes over time.

Below is a table of the variables we considered in our analysis:

Table 1: Variables used in Data Analysis

Name	Type	Notes
Sample Code	Categorical	Unique identifier for sample of foam
Strain	Numerical	Amount of strain tested on foam sample in kPa
Relative Porosity	Numerical	The measured porosity of the foam sample after the initial squish
Frequency	Numerical	The frequency at which the sample foam was squished
Amplitude	Numerical	The amplitude at which the sample foam was squished
Number of Squishes	Numerical	Number of squishes performed on foam at given time
Chemical Index	Categorical	Chemical index of given foam sample (Low, Medium, High)
Porosity	Categorical	Porosity of given foam sample (Low, High)
Relative Static Stiffness	Numerical	Static stiffness of given foam sample relative to the initial value of static stiffness
Time	Numerical	Time recorded at each measurement

2. Exploratory Data Analysis (EDA)

To explore what needs to be done for modeling, we perform some exploratory data analysis.

Figure 1 shows relative static stiffness over log(Time). We are using log(Time) because the relationship between time and static stiffness was not linear. See Figure 6 in the appendix for more information. From this figure, we can see that there is a negative relationship between log(Time) and relative static stiffness. It also appears that, depending on the foam type, there might be a different relationship. Some slopes are steeper than others, and some of the starting values of relative static stiffness are different than others which led us to believe that in the model we would need unique intercepts and slopes based on the foam type.

Next, we decided to look at the chemical index and porosity of the foams versus the relative static stiffness.

In Figure 2, the relationship between chemical index and relative static stiffness is different for the two levels of foam porosity. This indicates that there is likely a significant interaction between chemical index and porosity.

Next, we look at a correlation plot between all numerical variables.

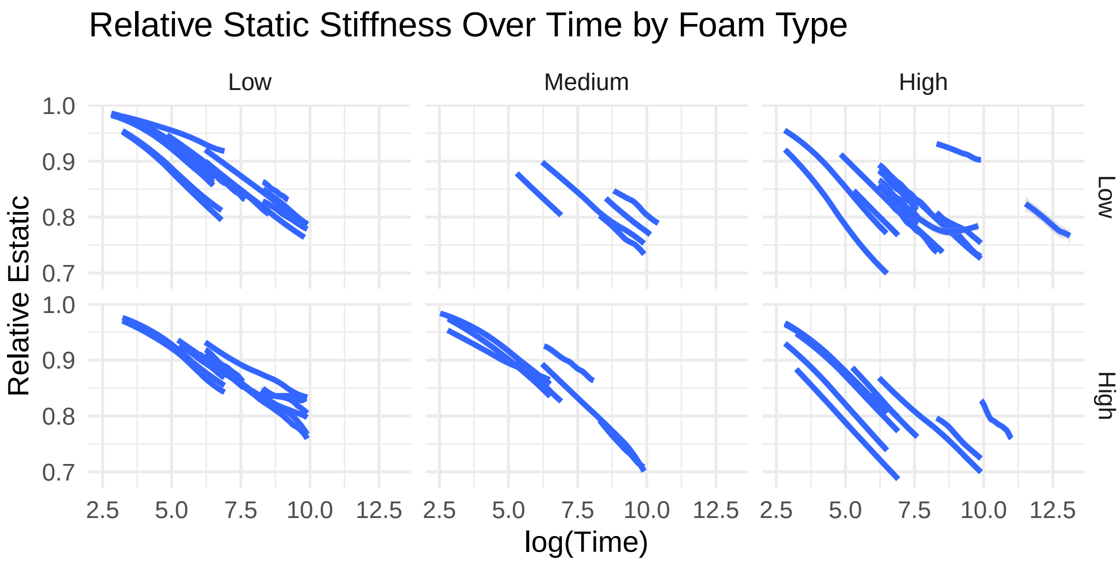


Figure 1: log(Time) vs. Relative Static Stiffness for each Foam Type

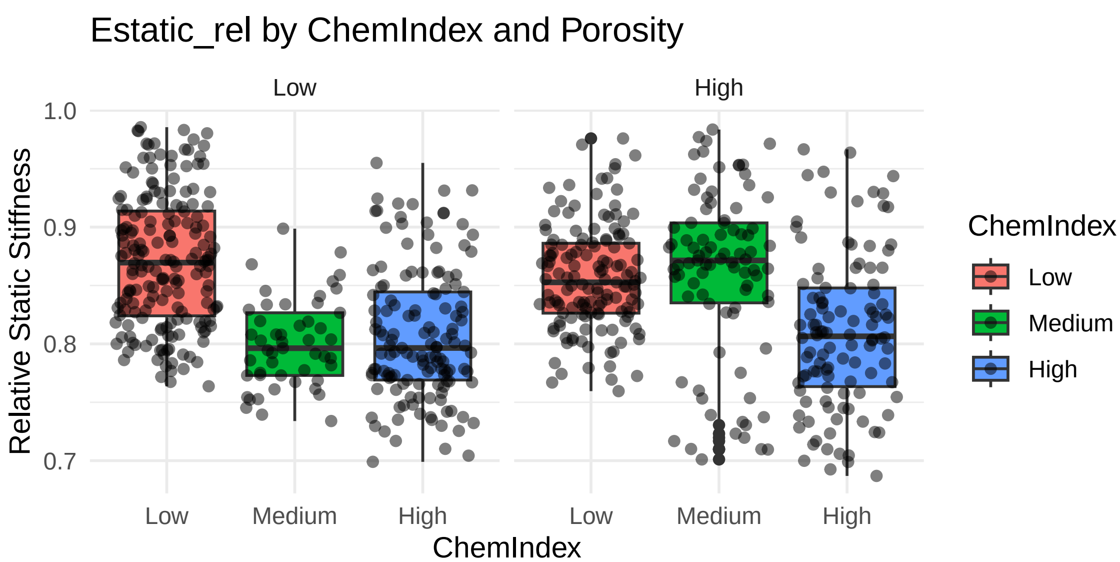


Figure 2: Chemical Index vs. Relative Static Stiffness for each level of Porosity

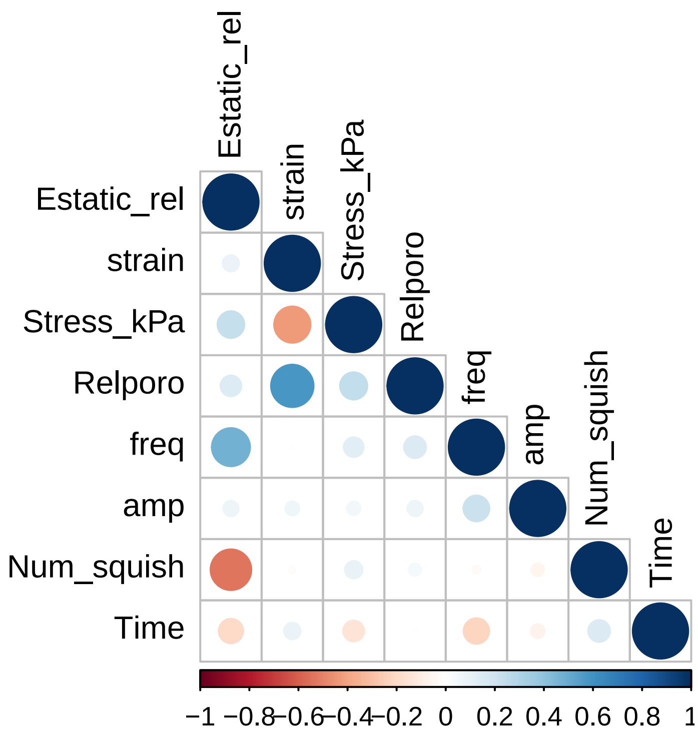


Figure 3: Correlation Matrix

Figure 3 gives us a good idea of what variables might need to be included in our final model. As we can see, number of squishes and frequency are highly correlated with relative static stiffness. This is obvious, as relative static stiffness is going to have a relationship with time as we saw in Figure 1. We also see that relative porosity and strain are correlated, as well as stress and strain.

Lastly, we look at how frequency is related to relative static stiffness over time.

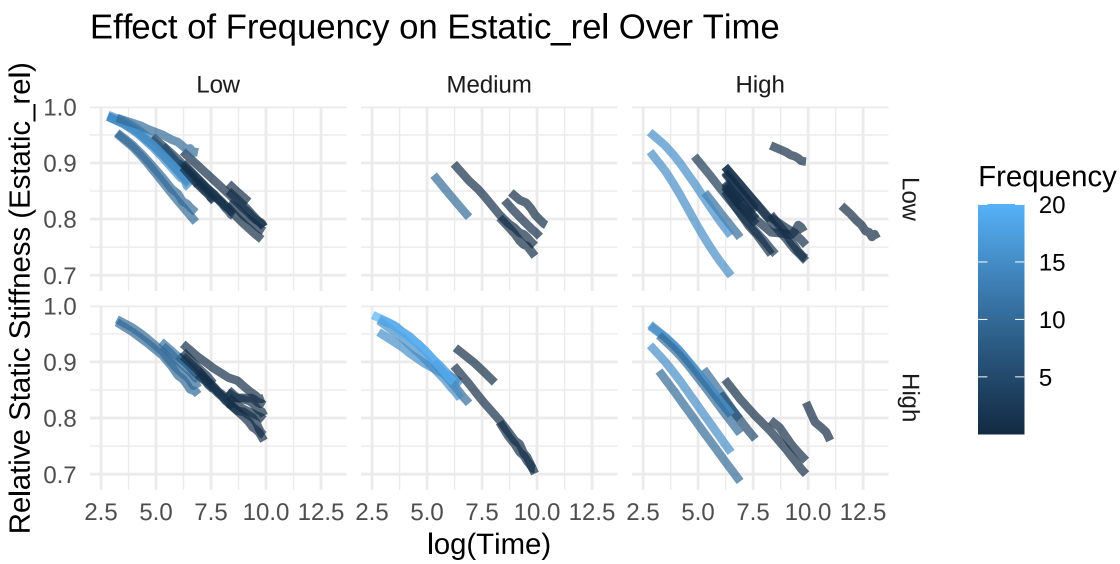


Figure 4: Frequency vs. Relative Static Stiffness over Time

In Figure 4, we can see that for lower frequencies, there is a higher relative static stiffness level recorded. This trend is the same across the different foam types. Frequency is a possible explanatory variable that we will consider in our final model.

3. Statistical Analysis

The linear mixed effects model is defined as:

$$
\begin{aligned}
Estatic\_rel_{ij} = {} & \underbrace{\beta_0}_{\text{fixed intercept}} + \underbrace{a_j}_{\text{random intercept for SampleCode}} \\
& + \beta_1 \left(\log(\text{Time}) \cdot \text{Porosity}\right)_{ij} \\
& + \beta_2 \left(\log(\text{Time}) \cdot \text{Porosity} \cdot \text{ChemIndex}\right)_{ij} \\
& + \beta_3 \cdot \text{freq}_{ij} + \beta_4 \cdot \text{numSquish}_{ij} + \varepsilon_{ij}
\end{aligned}
$$

where:

β_0 is the fixed intercept $a_j \sim \mathcal{N}(0, \sigma^2_{SampleCode})$ is the random intercept for group j (`SampleCode`) $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ is the residual error.

```
Linear mixed model fit by REML ['lmerMod']
Formula: Estatic_rel ~ log(Time):Porosity * ChemIndex + freq + Num_squish +
    (log(Time) | SampleCode)
   Data: df_clean

REML criterion at convergence: -5086.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-6.0458 -0.3427 -0.0347  0.3885  6.4327

Random effects:
 Groups     Name        Variance  Std.Dev. Corr
 SampleCode (Intercept) 2.735e-03 0.052299
            log(Time)   7.268e-05 0.008525 -0.75
 Residual               1.217e-05 0.003489
Number of obs: 700, groups:  SampleCode, 71

Fixed effects:
                                        Estimate Std. Error t value
(Intercept)                            1.116e+00  1.113e-02 100.208
ChemIndexMedium                        2.858e-02  1.799e-02   1.589
ChemIndexHigh                          2.821e-02  1.446e-02   1.950
freq                                  -3.864e-03  7.428e-04  -5.202
Num_squish                            -7.914e-07  1.258e-07  -6.293
log(Time):PorosityLow                 -3.231e-02  1.875e-03 -17.228
log(Time):PorosityHigh                -3.186e-02  1.990e-03 -16.009
log(Time):PorosityLow:ChemIndexMedium  -6.164e-03  3.652e-03  -1.688
log(Time):PorosityHigh:ChemIndexMedium -4.527e-03  3.360e-03  -1.347
log(Time):PorosityLow:ChemIndexHigh   -1.055e-02  2.651e-03  -3.979
log(Time):PorosityHigh:ChemIndexHigh  -1.298e-02  2.968e-03  -4.373

Correlation of Fixed Effects:
            (Intr) ChmInM ChmInH freq   Nm_sqs lg(T):PL lg(T):PH l(T):PL:CIM
ChemIndxMdm -0.452
ChemIndxHgh -0.613  0.367
freq        -0.413 -0.063  0.025
Num_squish   0.297  0.011 -0.035 -0.153
```

```
lg(Tm):PrsL -0.611  0.347  0.438 -0.010 -0.231
lg(Tm):PrsH -0.626  0.319  0.416  0.115 -0.227  0.431
l(T):PL:CIM  0.234 -0.603 -0.218  0.139 -0.005 -0.490   -0.184
l(T):PH:CIM  0.357 -0.644 -0.243 -0.112  0.006 -0.223   -0.569    0.373
l(T):PL:CIH  0.363 -0.252 -0.665  0.108  0.030 -0.680   -0.268    0.359
l(T):PH:CIH  0.416 -0.210 -0.595 -0.146  0.030 -0.257   -0.650    0.111
            l(T):PH:CIM l(T):PL:CIH
ChemIndxMdm
ChemIndxHgh
freq
Num_squish
lg(Tm):PrsL
lg(Tm):PrsH
l(T):PL:CIM
l(T):PH:CIM
l(T):PL:CIH  0.147
l(T):PH:CIH  0.390       0.379
fit warnings:
Some predictor variables are on very different scales: consider rescaling
```

In our model we allowed for a random intercept so that each sample could have their own unique baseline and we allowed for the change that each sample would experience over time to vary uniquely. Looking at our fixed effects we can see that chemical index is not significant. Frequency is significant and has a negative relationship with relative static stiffness. We also saw that the more squishes a foam indured the lower the stiffnes became. Looking at the relationship between log(Time) and porosity for both levels of porosity the stiffness will decrease over time. Now the terms we were most interested in was the three way interaction terms between log(Time), Porosity, and Chemical Index. In R it assigned the baseline to be the interaction between log(Time), high porosity and low chemical index. Thus the estimates seen for the interaction terms are in relation to this baseline. We can see that all of these estimates are negative thus there static stiffness decreases more over time in comparison to the baseline foam type. With the data being approximately normal from the central limit theorem, we determined anything with a t-value over $|2|$ to be significant. The t-values within the model all were significant and were kept within the final model.

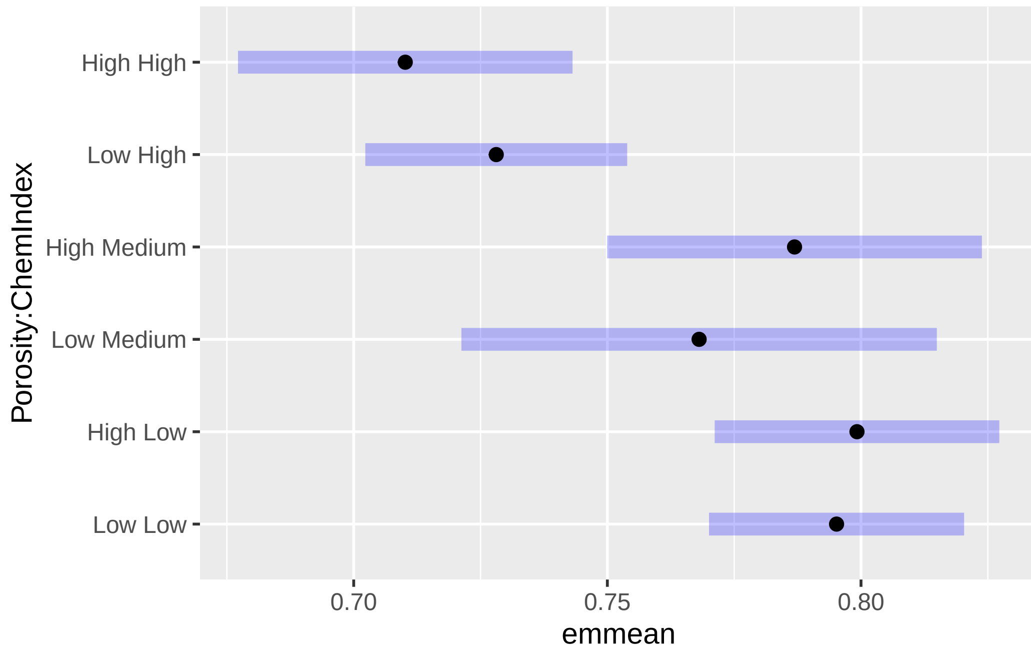


Figure 5: Emmeans Comparison by Porosity*ChemIndex

@em-plot shows a 95% confidence interval the Relative Static Stiffness based on the unique combination of porosity and chemical index, after accounting for all other variables. This allows us to see if there are significant differences between the different foam types. The foam type with Hight Porosity and Low Chemical Index (79) has the highest stiffness while also having a small confidence interval. That means that it will maintain a high static stiffness and also not fluctuate too much.

Table 2: Estimated Marginal Means for Static Stiffness.

ChemIndex	Porosity	emmean	SE	DF	Lower Conf-Lvl	Upper Conf-Lvl
Low	Low	0.7951780	0.0125892	64.83279	0.7700344	0.8203216
Medium	Low	0.7680857	0.0234669	65.59029	0.7212270	0.8149443
High	Low	0.7280985	0.0129362	68.37499	0.7022873	0.7539098
Low	High	0.7992060	0.0140531	66.37062	0.7711510	0.8272610
Medium	High	0.7868936	0.0184988	65.59176	0.7499553	0.8238320
High	High	0.7101639	0.0165089	65.14698	0.6771947	0.7431331

4. Recommendations

Question 1: Which combination of porosity and chemical index is best at keeping static stiffness stable over time, after accounting for all other variables?

After performing our analysis we believe that the foam type that has a porosity of 81% and a chemical index of 79 will maintain the static stiffness the best over time.

5. Resources

List resources that your client might find useful

6. Additional Considerations

Study Design

- *Limitations to the recommendations*
- *Concerns you may have about the study; suggestions for similar studies in future*
- *Technical comments*

Technical Appendix

Here we can see a clear nonlinear relationship between relative static stiffness and time. A logarithmic transformation is needed. See Figure 7.

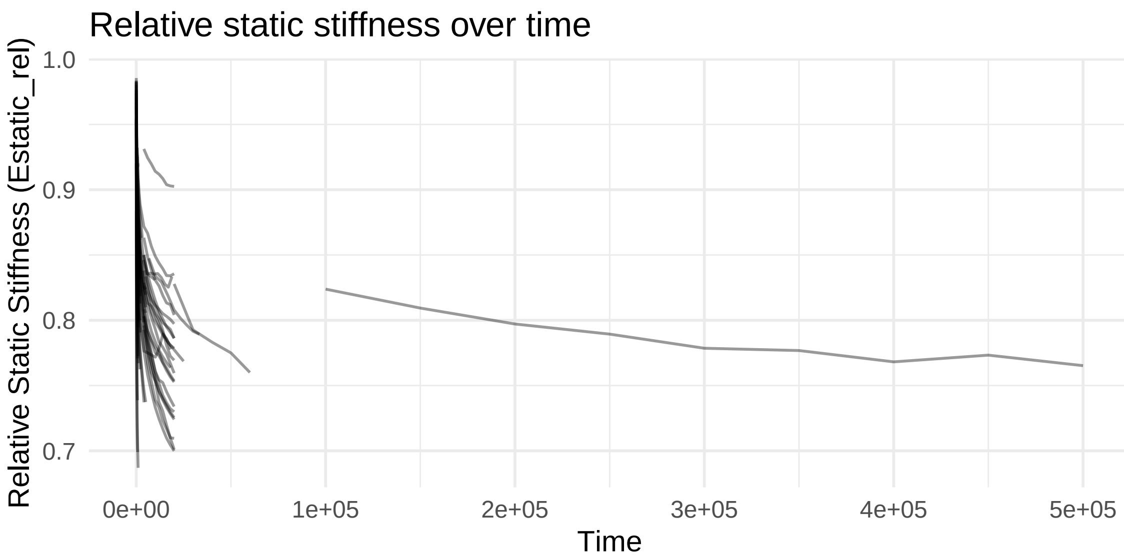


Figure 6: Time vs. Relative Static Stiffness

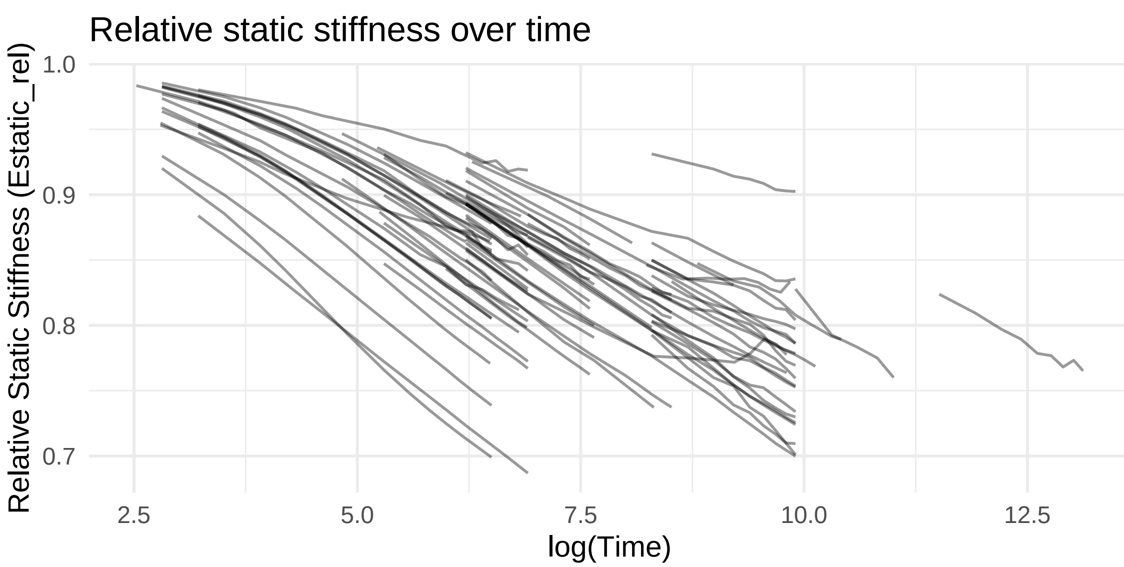
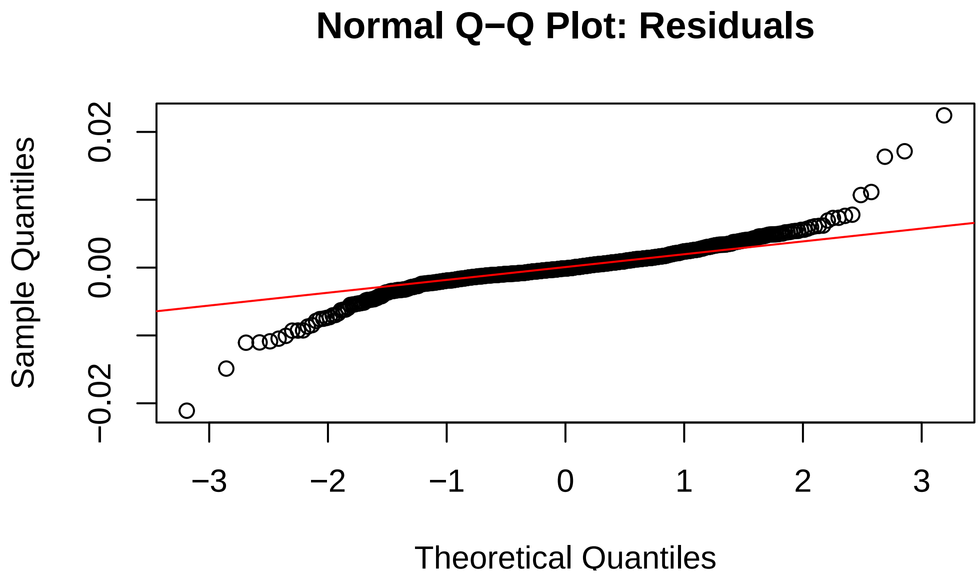
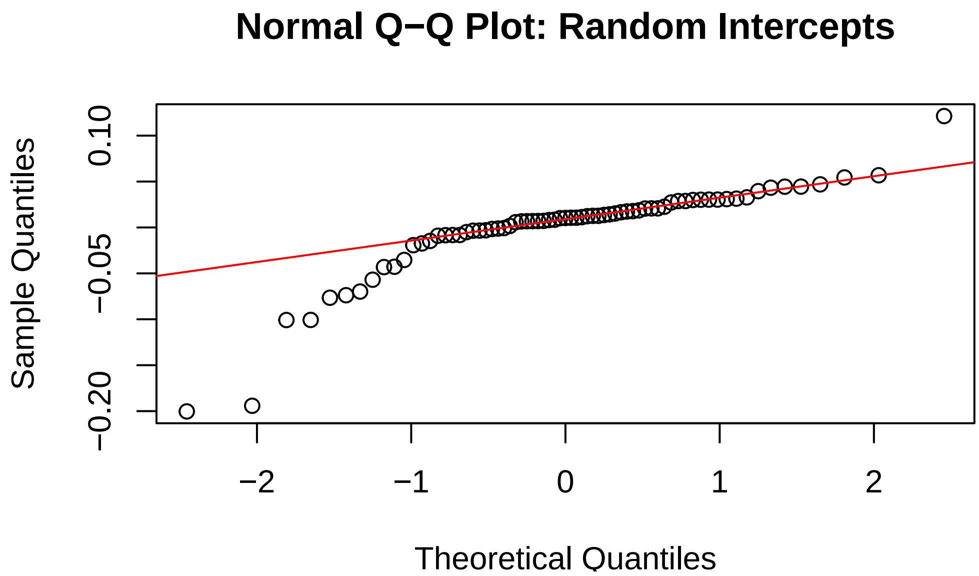
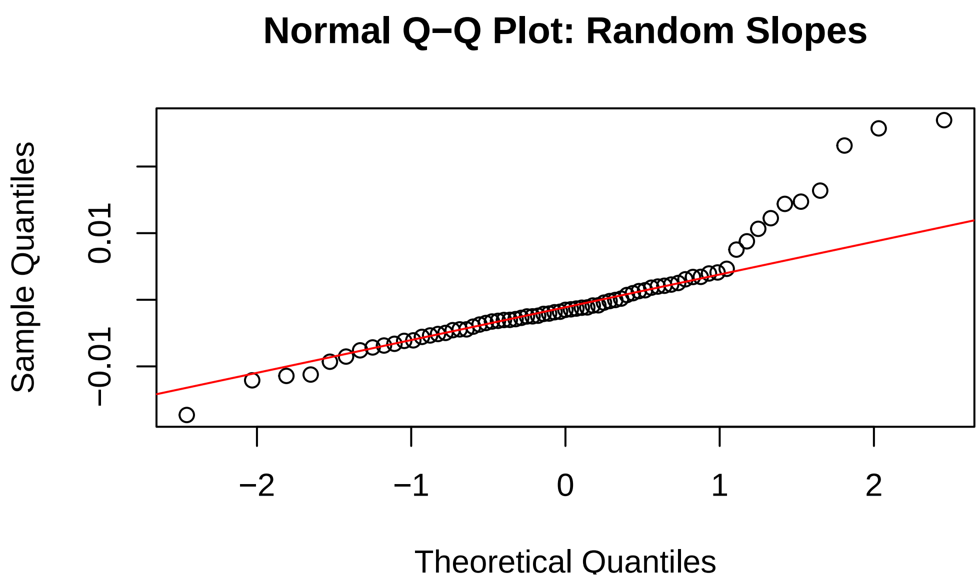


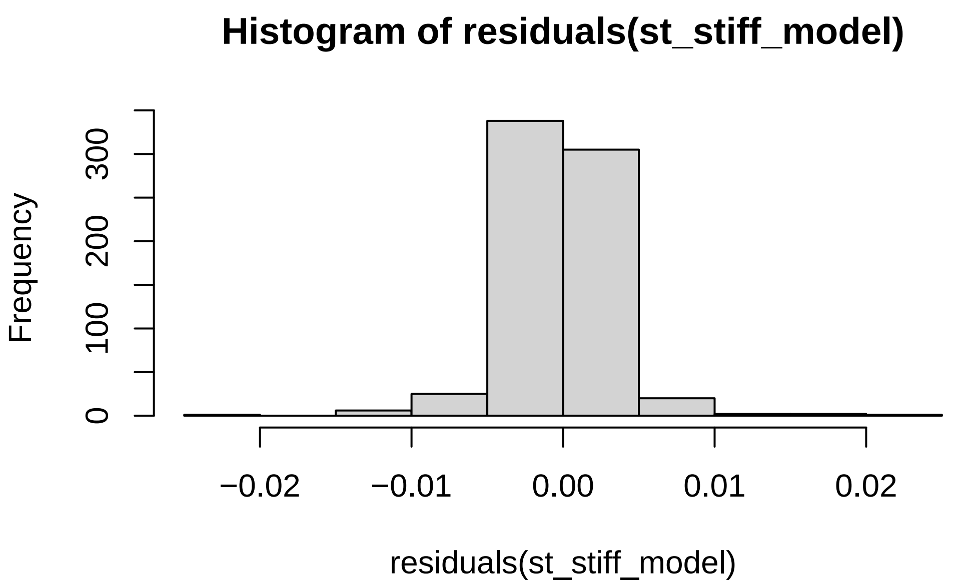
Figure 7: log(Time) vs. Relative Static Stiffness


Normal Q−Q Plot: Residuals
Sample Quantiles
−0.02
0.00
0.02
−3
−2
−1
0
1
2
3
Theoretical Quantiles

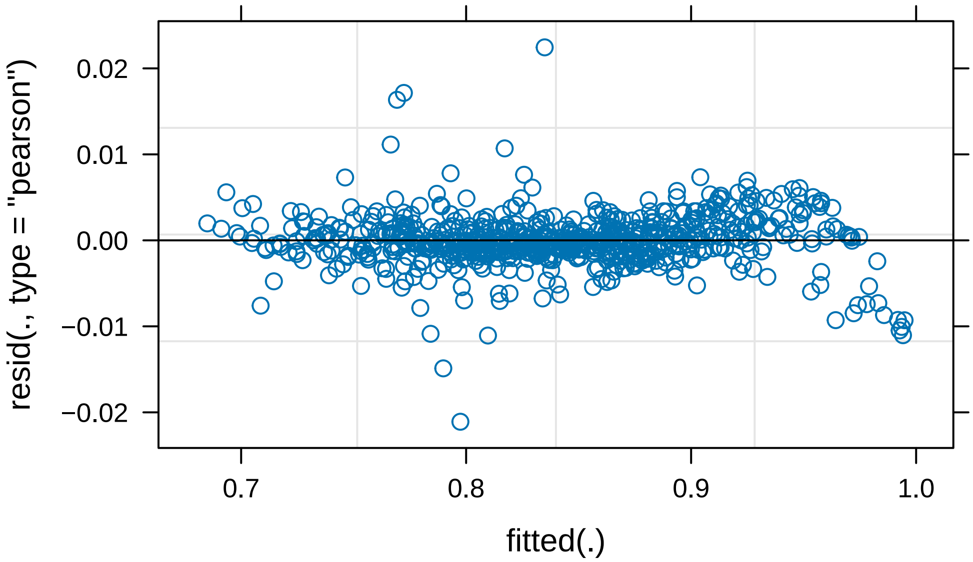

Normal Q−Q Plot: Random Intercepts
Sample Quantiles
−0.20
−0.05
0.10
−2
−1
0
1
2
Theoretical Quantiles



Normal Q−Q Plot: Random Slopes
Sample Quantiles
−0.01
0.01
−2
−1
0
1
2
Theoretical Quantiles



Histogram of residuals(st_stiff_model)
Frequency
0
100
200
300
−0.02
−0.01
0.00
0.01
0.02
residuals(st_stiff_model)




```
                               numDF denDF   F-value p-value
(Intercept)                        1   622 25384.644  <.0001
ChemIndex                          2    67     9.143   3e-04
freq                               1    67    24.626  <.0001
Num_squish                         1   622 10313.149  <.0001
log(Time):Porosity                 2   622  1143.459  <.0001
log(Time):Porosity:ChemIndex       4   622    63.594  <.0001
```

R Script

```
# clean up & set default chunk options
rm(list = ls())
knitr::opts_chunk$set(echo = FALSE)

# packages
library(tidyverse) # for example
library(mosaic)    # for example
library(ggformula) # for example
library(lme4)
library(ggplot2)
library(emmeans)
library(nlme)
library(readxl)
library(tinytex)
```

```r
library(xtable)
library(knitr)
library(ggthemes)
library(GGally)
library(kableExtra)
library(emmeans)
library(Stat2Data)
library(corrplot)


# read in data

# use this space to do any data processing you need

# load data
df <- read_xlsx("Foster_Mines_Capstone_Data_v2.xlsx")

#Pivot all cycle-related columns into long format
df_long <- df %>%
  pivot_longer(
    cols = matches("(^Cyc_[A-Z]$)|(_Cyc_[A-Z]$)"),
    names_to = "FullName",       # give it a neutral name
    values_to = "Value"
  )

#Extract the cycle (e.g., Cyc_A) and variable type (e.g., E1_MPA)
df_tidy <- df_long %>%
  mutate(
    Cycle = str_extract(FullName, "Cyc_[A-Z]$"),
    VariableType = str_remove(FullName, "_Cyc_[A-Z]$"),
    VariableType = ifelse(VariableType == "Cyc", "Cycle_time", VariableType)
  ) %>%
  select(-FullName) %>%
  pivot_wider(
    names_from = VariableType,
    values_from = Value
  ) %>%
  arrange(SampleCode, Cycle)

head(df_tidy)

#Pivot the actual cycle time columns only
```

```
df_lt <- df_tidy %>%
  pivot_longer(
    cols = matches("^Cyc_[A-Z]$"),  # Cyc_A to Cyc_N
    names_to = "Cycle_column",      # Temporary name to avoid collision
    values_to = "Num_squish"
  )

#Remove NA times
df_clean <- df_lt %>%
  filter(!is.na(Num_squish)) %>%
  select(-Cycle) %>%                #Drop the existing "Cycle" column and rename
  rename(Cycle = Cycle_column)

df_clean <- df_clean |>
  mutate(
    ChemIndex = as.numeric(str_split_fixed(SampleCode, "_", 4)[,2]),
    Porosity = as.numeric(str_split_fixed(SampleCode, "_", 4)[,3]))

df_clean <- df_clean %>%
  group_by(SampleCode) %>%
  mutate(
    Estatic_initial = first(Estatic_MPA),  # First Estatic value per SampleCode
    Estatic_rel = Estatic_MPA / Estatic_initial  # Relative to initial
  ) %>%
  ungroup()

df_clean <- df_clean |>
  select(-E1_MPA, -tandelta, -NVP)

df_clean <- df_clean %>%
   mutate(Porosity = case_when(
     Porosity == 71 ~ "Low",
     Porosity == 81 ~ "High"
   ))

 df_clean <- df_clean %>%
   mutate(ChemIndex = case_when(
     ChemIndex == 121 ~ "High",
     ChemIndex == 100 ~ "Medium",
     ChemIndex == 79 ~ "Low"
   ))
```

```r
df_clean$Porosity <- factor(df_clean$Porosity, levels = c("Low", "High"))

df_clean$ChemIndex <- factor(df_clean$ChemIndex, levels = c("Low", "Medium", "High"))

df_clean <- df_clean %>%
  mutate(Cycle = case_when(
    Cycle == "Cyc_A" ~ "A",
    Cycle == "Cyc_B" ~ "B",
    Cycle == "Cyc_C" ~ "C",
    Cycle == "Cyc_D" ~ "D",
    Cycle == "Cyc_E" ~ "E",
    Cycle == "Cyc_F" ~ "F",
    Cycle == "Cyc_G" ~ "G",
    Cycle == "Cyc_H" ~ "H",
    Cycle == "Cyc_I" ~ "I",
    Cycle == "Cyc_J" ~ "J",
    Cycle == "Cyc_K" ~ "K",
    Cycle == "Cyc_L" ~ "L",
    Cycle == "Cyc_M" ~ "M",
    Cycle == "Cyc_N" ~ "N"
  ))

df_clean$Cycle <- factor(df_clean$Cycle,
                         levels = c("A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "

df_clean <- df_clean %>%
  filter(Num_squish <= 10000)
df_clean <- df_clean %>%
  mutate(Time = Num_squish/freq)

df_clean <- df_clean |>
  filter(Estatic_rel < 1)

variable.desc <- data.frame(Name = c("Sample Code", "Strain", "Relative Porosity", "Frequency
variable.desc$Type <- c("Categorical", "Numerical", "Numerical", "Numerical", "Numerical", "N
variable.desc$Notes <- c("Unique identifier for sample of foam", "Amount of strain tested on
knitr::kable(variable.desc, format = "latex", booktabs = TRUE,
             col.names = c("Name", "Type", "Notes"),
             longtable = FALSE) %>%
  kable_styling( full_width = FALSE) %>%
  column_spec(3, width = "8cm") %>%
  add_header_above(c(" " = 3))
```

```r
ggplot(df_clean, aes(x = log(Time), y = Estatic_rel, group = SampleCode)) +
  geom_smooth(alpha = 0.3) +
 # stat_summary(aes(group = 1), fun = mean, geom = "line", color = "black", size = 1) +
  facet_grid(Porosity ~ ChemIndex) +
  labs(title = "Relative Static Stiffness Over Time by Foam Type",
       x = "log(Time)",
       y = "Relative Estatic") +
  theme_minimal()
# grey lines are individual measurements

ggplot(df_clean, aes(x = ChemIndex, y = Estatic_rel, fill = ChemIndex)) +
  geom_boxplot() +
  geom_jitter(alpha = 0.5) +
  facet_wrap(~ Porosity) +
  labs(title = "Estatic_rel by ChemIndex and Porosity",
       y = "Relative Static Stiffness") +
  theme_minimal()

df_numeric <- df_clean %>%
  select(Estatic_rel, strain, Stress_kPa, Relporo, freq, amp, Num_squish, Time)  # Adjust bas

# Compute the correlation matrix
corr_matrix <- cor(df_numeric, use = "complete.obs")

# Plot the correlation matrix
corrplot(corr_matrix, method = "circle", type = "lower", tl.col = "black")
ggplot(df_clean, aes(x = log(Time), y = Estatic_rel, color = freq)) +
  geom_line(aes(group = SampleCode), alpha = 0.7, linewidth = 1.5) +
  facet_grid(Porosity ~ ChemIndex) +
  labs(
    title = "Effect of Frequency on Estatic_rel Over Time",
    x = "log(Time)",
    y = "Relative Static Stiffness (Estatic_rel)",
    color = "Frequency"
  ) +
  theme_minimal()
emmean <- emmeans(st_stiff_model,~ Porosity *  ChemIndex)
plot(emmean)

ggplot(df_clean, aes(x = Time, y = Estatic_rel)) +
  geom_line(aes(group = SampleCode), alpha = 0.4) +
  labs(
```

```
    title = "Relative static stiffness over time",
    x = "Time",
    y = "Relative Static Stiffness (Estatic_rel)",
  ) +
  theme_minimal()

ggplot(df_clean, aes(x = log(Time), y = Estatic_rel)) +
  geom_line(aes(group = SampleCode), alpha = 0.4) +
  labs(
    title = "Relative static stiffness over time",
    x = "log(Time)",
    y = "Relative Static Stiffness (Estatic_rel)",
  ) +
  theme_minimal()
# model diagnostic plots
qqnorm(residuals(st_stiff_model),
       main = "Normal Q-Q Plot: Residuals")
qqline(residuals(st_stiff_model), col = "red")

qqnorm(ranef(st_stiff_model)$SampleCode[[1]],
       main = "Normal Q-Q Plot: Random Intercepts")
qqline(ranef(st_stiff_model)$SampleCode[[1]], col = "red")

qqnorm(ranef(st_stiff_model)$SampleCode[[2]],
       main = "Normal Q-Q Plot: Random Slopes")
qqline(ranef(st_stiff_model)$SampleCode[[2]], col = "red")

hist(residuals(st_stiff_model))

plot(st_stiff_model)
```