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

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.

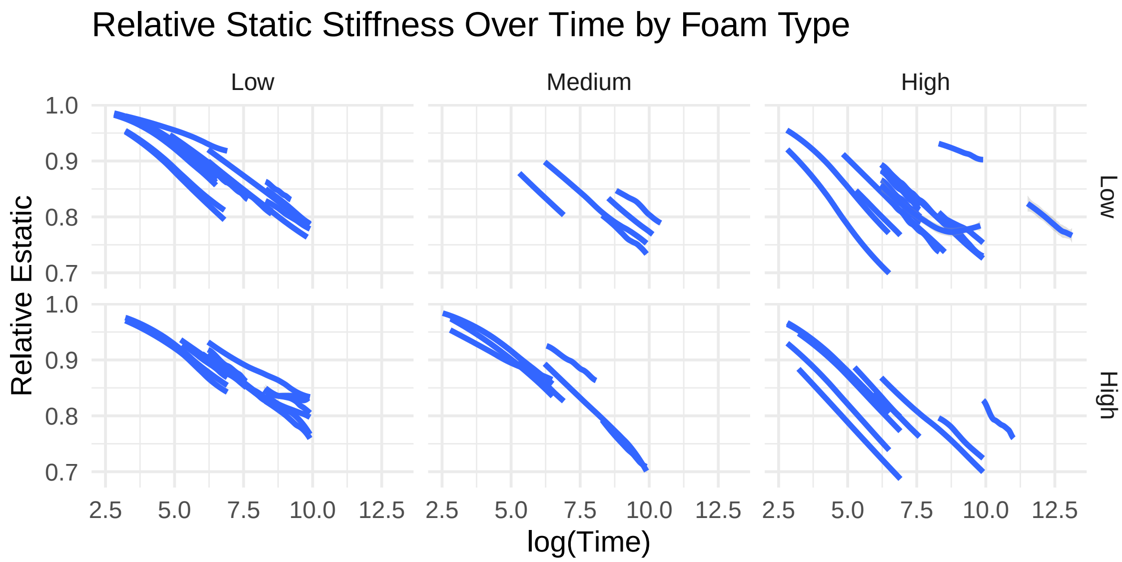


Figure 1: log(Time) vs. Relative Static Stiffness for each Foam Type

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.

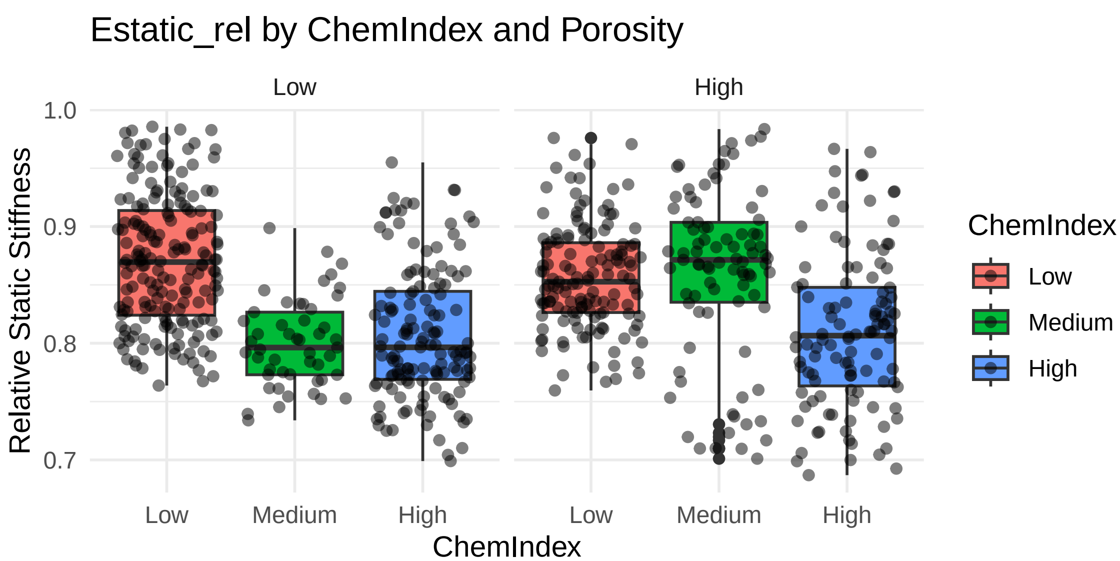


Figure 2: Chemical Index vs. Relative Static Stiffness for each level of Porosity

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.

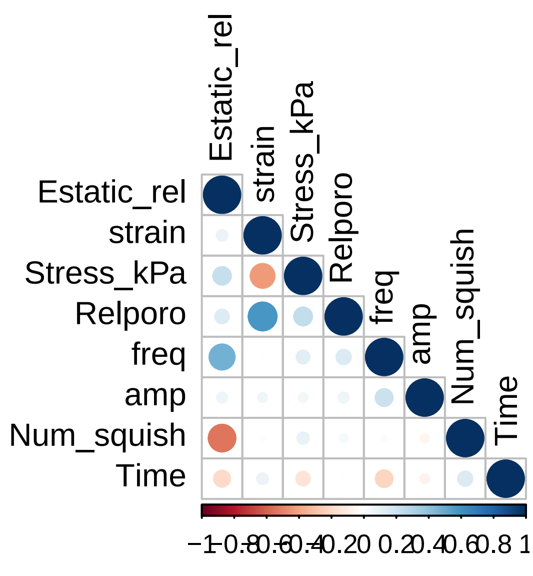


Figure 3: Correlation Matrix

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.

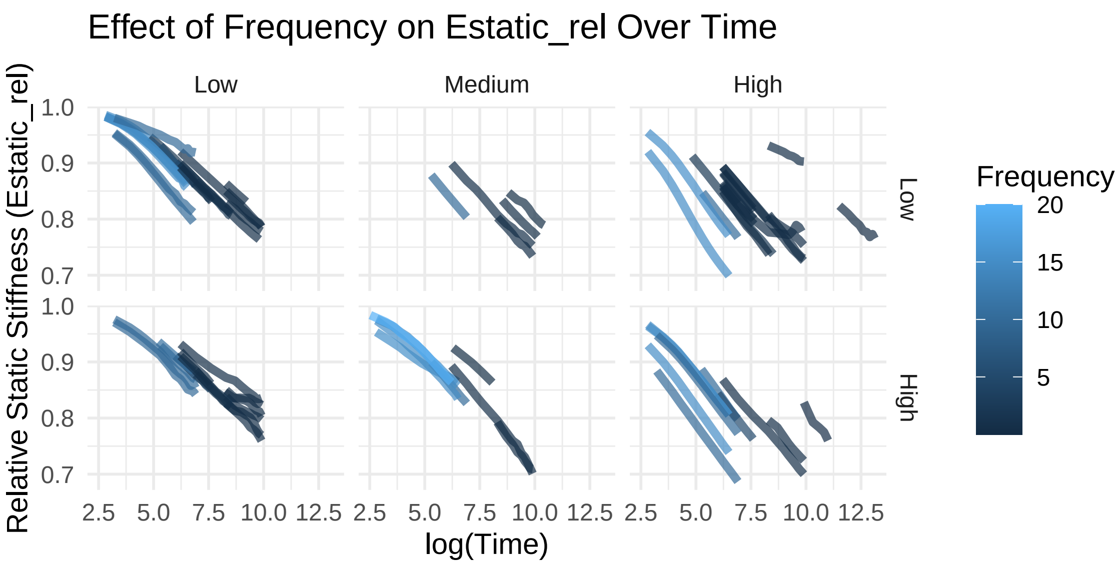


Figure 4: Frequency vs. Relative Static Stiffness over Time

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

We will be using a mixed model for our analysis. A mixed model is used when there is variation

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.

	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

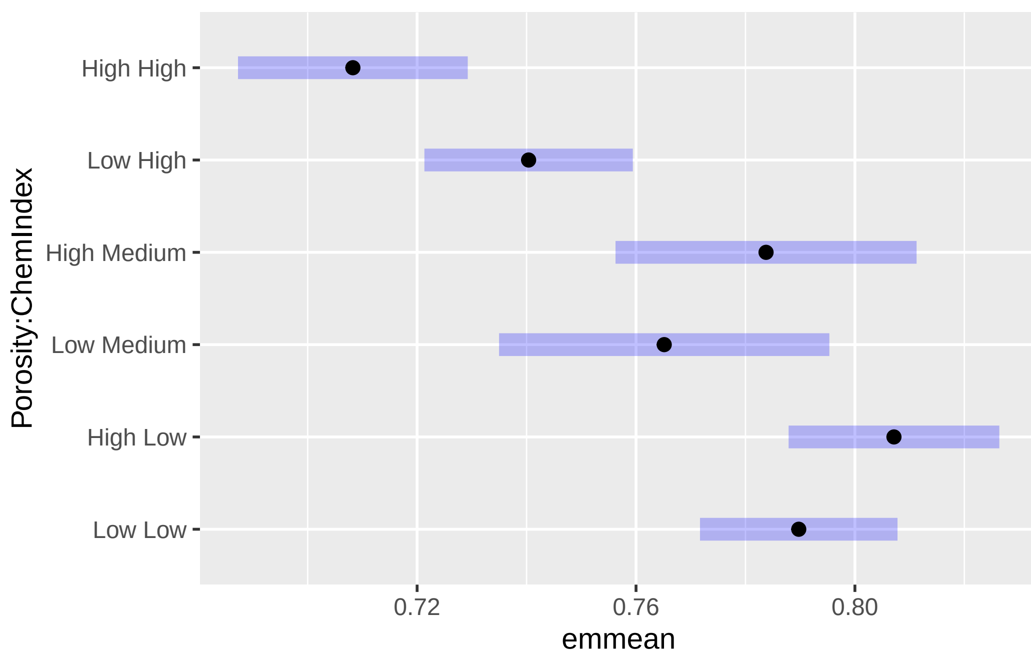


Figure 5: Emmeans Comparison by Porosity*ChemIndex

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?

5. Resources

6. Additional Considerations

Technical Appendix

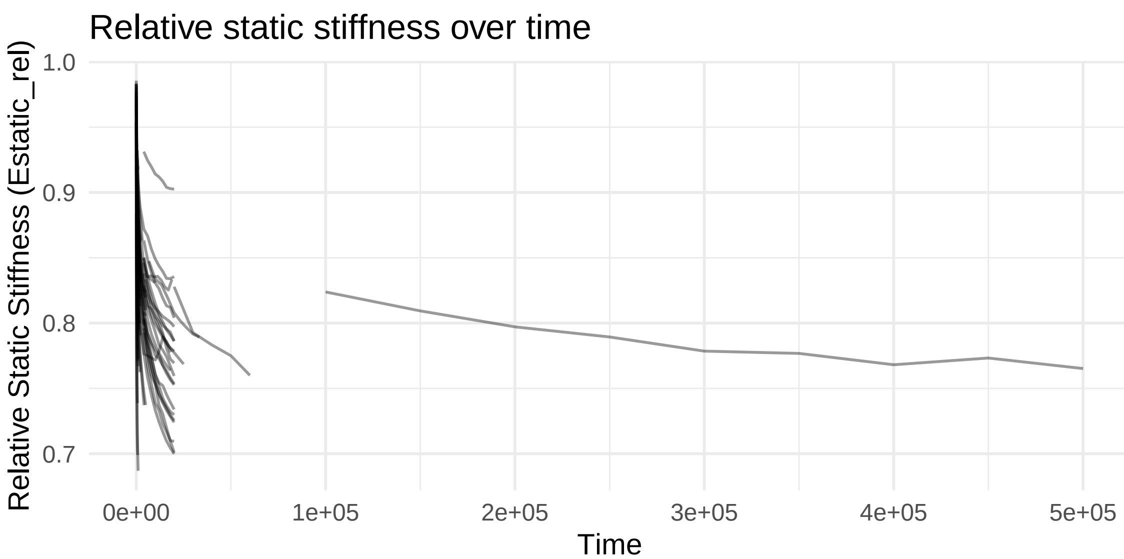


Figure 6: Time vs. Relative Static Stiffness

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

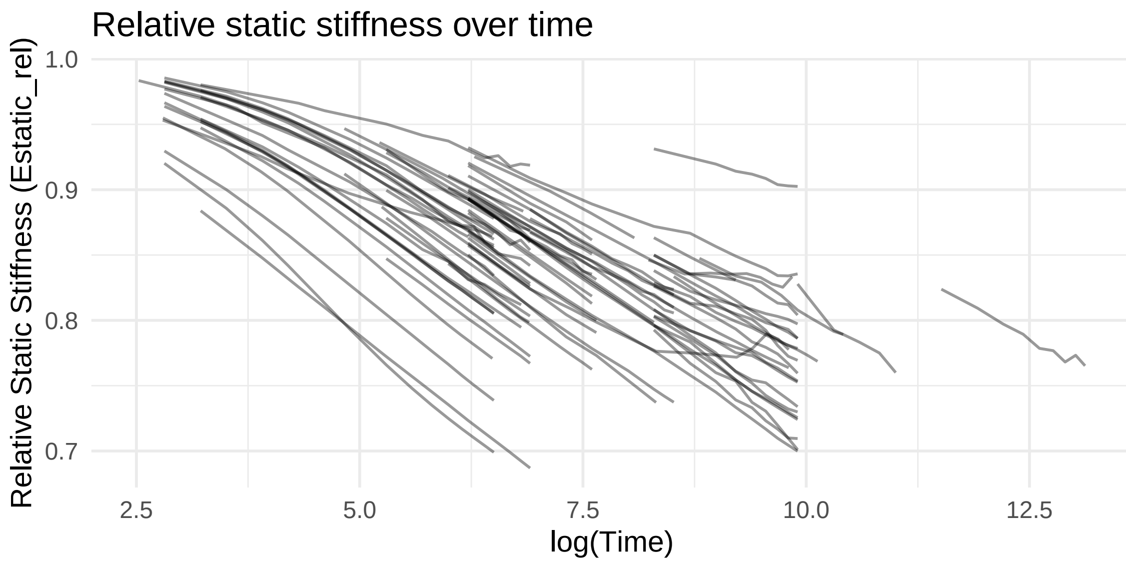
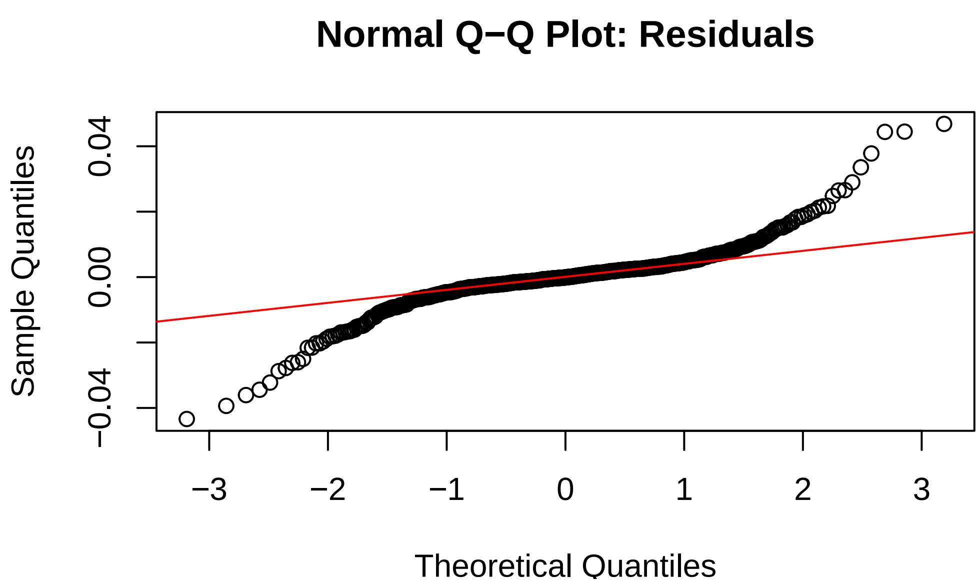
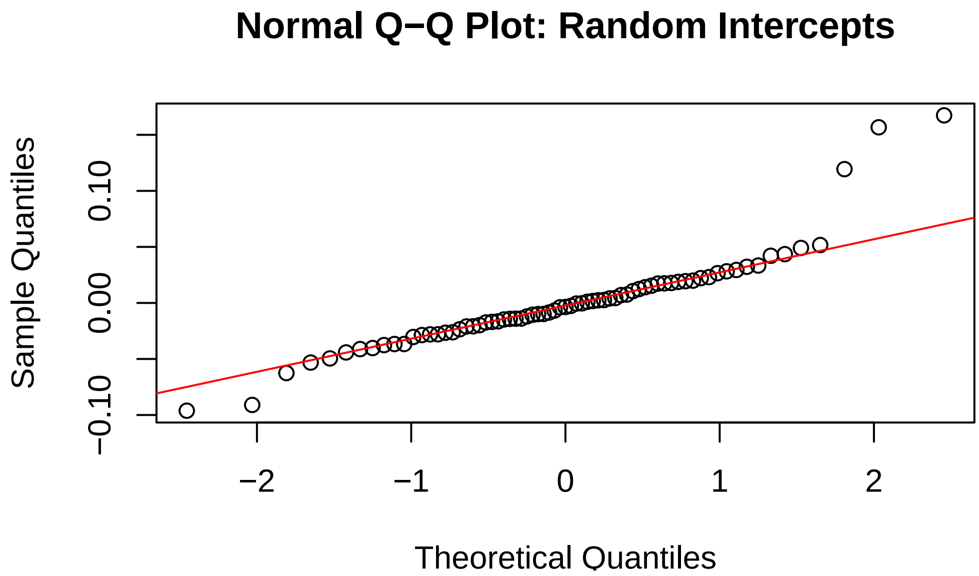
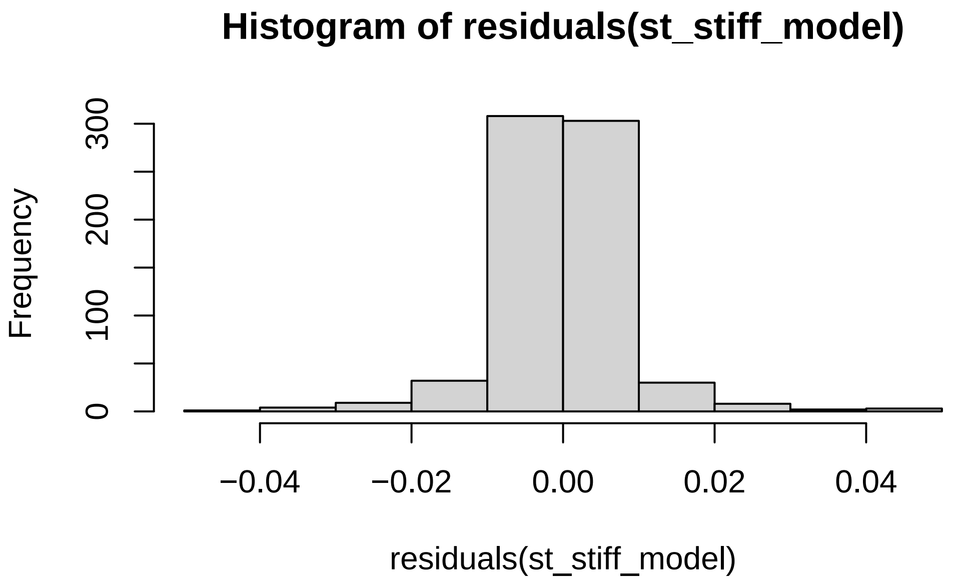
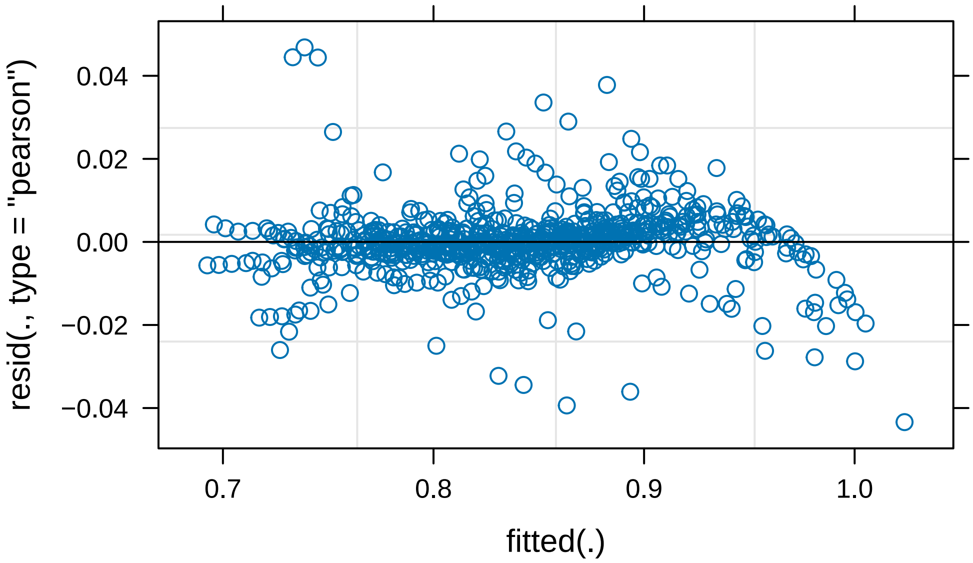


Figure 7: log(Time) vs. Relative Static Stiffness




Normal Q−Q Plot: Random Intercepts
Sample Quantiles
−0.10
0.00
0.10
−2
−1
0
1
2
Theoretical Quantiles


Histogram of residuals(st_stiff_model)
Frequency
0
100
200
300
−0.04
−0.02
0.00
0.02
0.04
residuals(st_stiff_model)



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

```r
# 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"
   ))

 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))
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(
```

```r
    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")

hist(residuals(st_stiff_model))

plot(st_stiff_model)
```