STA305 Final Project - Technical Report

Group 2

14/08/2021

1 Objectives

Through this experiment, we aim to determine:

- Do toilet paper brands have different abilities to withstand stresses?
- Does the wetness of toilet paper affect its durability?
- Does the strength of toilet paper depend on the brands differently for different levels of wetness (and vice-versa)?

2 Variables

We have two predictor variables:

1) Toilet paper brand (A, B, C, D, E) for (Cashmere, Charmin, Cottonelle, Kirkland, Royale) respectively. It is a factor (categorical predictor) with 5 levels.
2) The amount of water applied to the toilet paper (J, K, L) for (Low, Medium, High), respectively. It is a factor (categorical variable) that describes the amount of water that was applied to the toilet paper, with 3 levels: Low = 0.5mL, Medium = 1.0mL, High = 1.5mL.

Our response variable is the number of coins: a numerical, interval variable that describes the number of coins resting on the wet toilet paper before it breaks.

Therefore, we treat Cashmere and Low as the baselines for each model, and there are a total of $5 \times 3 = 15$ different treatments. Each of the treatments is possible and they all have the same number of combinations, making this a complete and balanced design where the factors are crossed.

3 Sample Size Calculation

To determine the sample size necessary for the experiment, the `ss.2way` function from the `pwr2` R package will be used. The function has the following parameters: the number of groups in each factor, significance level chosen for the hypothesis test, desired power, and the effect sizes for the factors. Table 1 summarizes the values assigned to the function and the output sample size. The table also precedes the justification for the input.

Table 1: Sample Size Calculation

Variable	Value
Brand Groups	5.00
Moisture Groups	3.00
Significance Level	0.05
Power	0.80
Effect size	0.25
Group Size	5.00
Total Sample	75.00

- There are 5 groups in the brand factor (Cashmere, Charmin, Cottonelle, Kirkland, Royale) and there are 3 groups in the moisture factor (Low, Medium, high).
- The significance level chosen for the hypothesis test is 0.05.
- Since the purpose of this study was to understand the properties of various toilet paper products and inform future consumer choice, the ability to replicate the findings of this study is a key objective. As a result, a high power is necessary as it corresponds to a low probability of Type II error. Thus, the power was set to 0.8, or 80%.
- In order to “fine tune” the experiment such that even relatively small effect gets detected, the effect size for each factor was set to 0.25.

4 Checking ANOVA Assumptions

In ANOVA, we need to check the five assumptions on the error terms before we can proceed with the analysis.

4.1 Independence

The independence assumption ensures that a trial does not have any effect on another trial’s response value. We take a look at the residuals versus fitted plot for the interaction and additive models, seen in Figure 1.

After having observed null plots and dispersed residuals (with the exception of notable points), there does not seem to be any violations of the independence assumption.

4.2 Identical Distribution (Within/Between groups) and Homogeneity of Variance

We will assume that all trials in the same group will have the same distribution of ϵ_{ij}’s. Due to the nature of this experiment, there is no reason to believe that the errors are correlated with each other, nor is there a time-dependent predictor. Also, our sample size is sufficiently large enough to affirm our assumption.

Furthermore, to check the homogeneity of variances, we will use a scale-location plot for both interaction and additive models to observe the standardized residuals and the fitted values, seen in Figure 2.

Again, there seems to be no pattern, along with the noteworthy points we saw earlier. Hence, we can assume homoscedasticity and proceed.

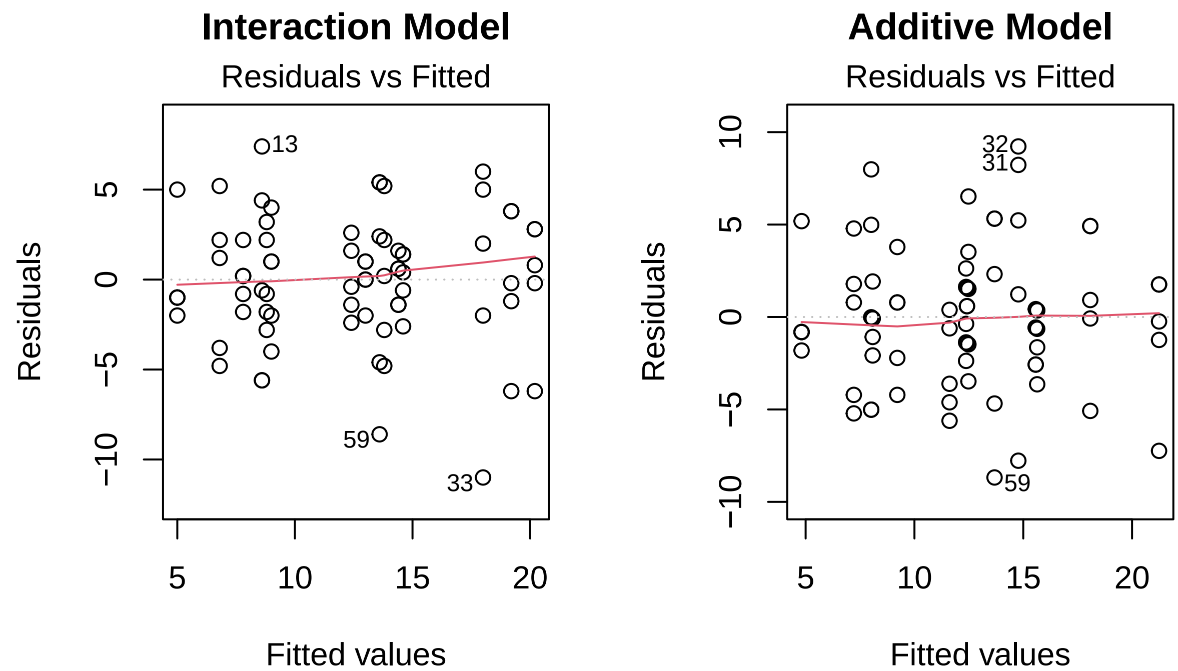

Figure 1: Residuals vs. Fitted Plots

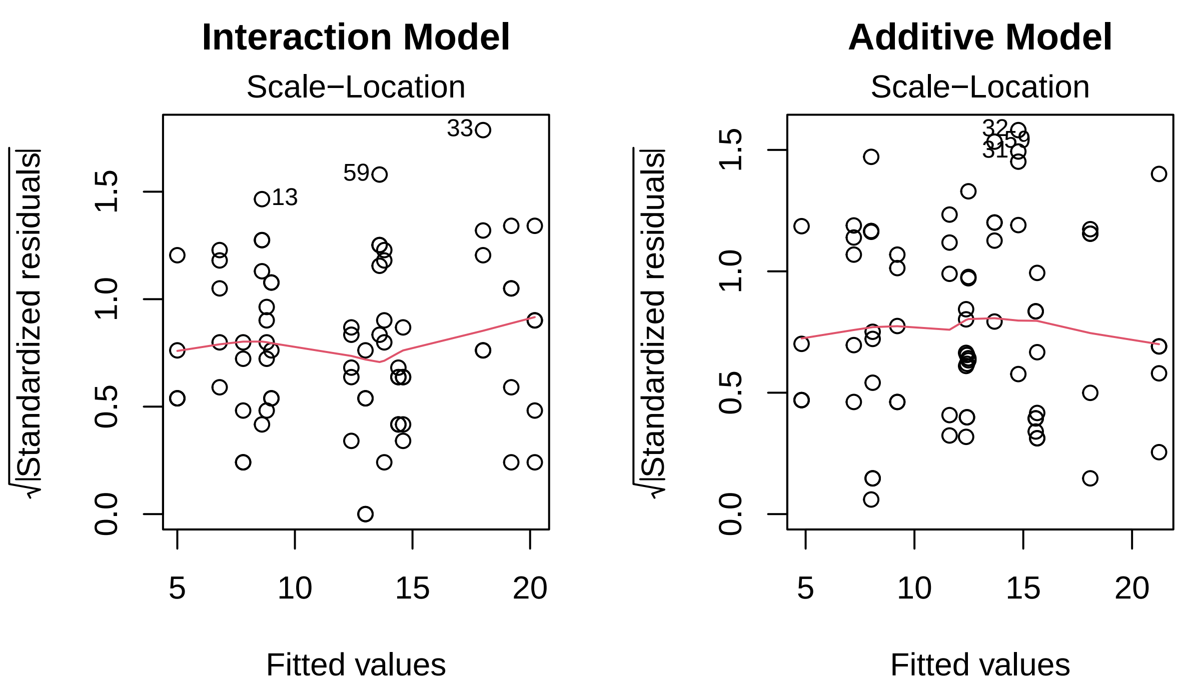

Figure 2: Scale-Location Plots

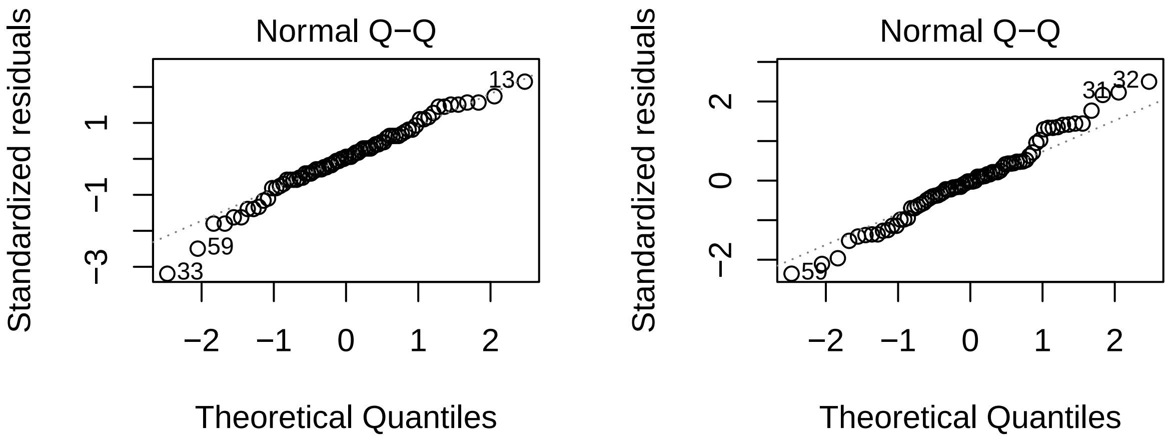

Figure 3: QQ Plots

4.3 Normality

Finally, we assess the normality of the errors by using a Q-Q plot and the Shapiro-Wilk test, seen in Figure 3 and Table 2.

Table 2: Interaction and Additive Model Shapiro-Wilk Tests

Interaction statistic	Interaction p value	Additive statistic	Additive p value
0.9802356	0.290978	0.9842141	0.4761434

Visually, almost all the points fall approximately on the normality line. In addition to this, we get a p-value of 0.291 and 0.4761 from the ANOVA residuals of the interaction and additive models respectively using the Shapiro-Wilk test. This is further evidence implying that normality is satisfied, since the distribution of the data are not statistically significant from a Normal distribution.

5 Interaction Plot

As seen in Figure 4, not all of the lines are parallel; notably, the lines between Cashmere and Charmin have wildly different slopes. However, none of them ever cross each other. More specifically, the line segment for the low moisture level in combination with the Cashmere brand on the interaction plot has a slope that is negative compared to the lines corresponding to the medium and high moisture levels. This indicates that there may be medium interaction between the low moisture level and the Cashmere brand of toilet paper. Furthermore, the three levels for the moisture factor always follow the order of: the toilet paper with low moisture having the greatest number of coins, followed by medium moisture, and finally high moisture, as expected. With these in mind, we use a two-way ANOVA F-test to determine whether the interaction effect is significant.

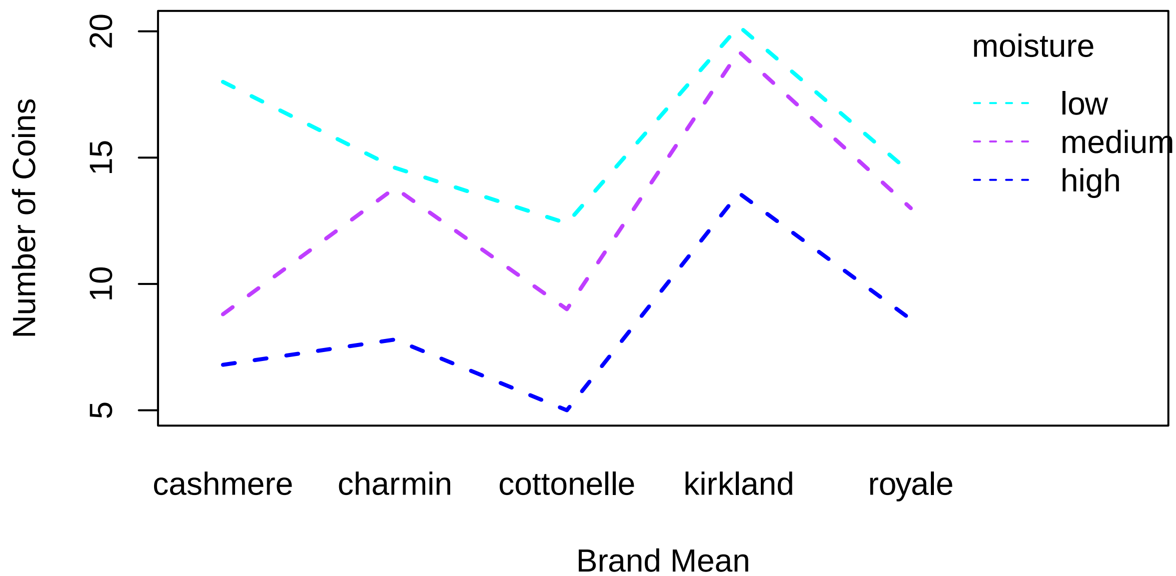

Figure 4: Interaction Plot for Brand and Moisture

6 Interaction + Main Effects Model

We define the following interaction model using dummy coding (for brevity, "interaction terms" is written "ITs" below):

$$
\begin{aligned}
E(Y_i) = \beta_0 &+ \underbrace{\beta_1 I_{B,i} + \beta_2 I_{C,i} + \beta_3 I_{D,i} + \beta_4 I_{E,i}}_{\text{Terms for brands}} \\
&+ \underbrace{\beta_5 I_{K,i} + \beta_6 I_{L,i}}_{\text{Terms for water applied}} \\
&+ \underbrace{\beta_7 (I_{B,i} \times I_{K,i}) + \beta_8 (I_{B,i} \times I_{L,i})}_{\text{Charmin-(Med/High) ITs}} + \underbrace{\beta_9 (I_{C,i} \times I_{K,i}) + \beta_{10} (I_{C,i} \times I_{L,i})}_{\text{Cottonelle ITs}} \\
&+ \underbrace{\beta_{11} (I_{D,i} \times I_{K,i}) + \beta_{12} (I_{D,i} \times I_{L,i})}_{\text{Kirkland-(Med/High) ITs}} + \underbrace{\beta_{13} (I_{E,i} \times I_{K,i}) + \beta_{14} (I_{E,i} \times I_{L,i})}_{\text{Royale-(Med/High) ITs}} + \epsilon_i
\end{aligned}
$$

where:

- The indicators $I_{a,i}$ where $a \in \underbrace{\{A, \dots, E\}}_{\text{brands}} \cup \underbrace{\{J, K, L\}}_{\text{moisture}}$ represent whether a treatment includes that factor level. For example, a treatment where Cottonelle is used and high moisture is applied would have $I_{C,i} = I_{K,i} = 1$ and all other indicators equal to 0.
- β_0 represents the mean number of coins supported for the baseline group: Cashmere toilet paper with low moisture (0.5mL applied)
- $\beta_1, \dots, \beta_4$ each represent the mean increase in the number of coins for each brand (Charmin, Cottonelle, Kirkland, Royale, respectively) compared to the baseline group, with low moisture

- β_5, β_6 each represent the mean increase in the number of coins for (medium moisture, high moisture respectively) compared to the baseline group, for Cashmere toilet paper.
- $\beta_7, \ldots, \beta_{14}$ are interaction terms that each represent the mean increase in the number of coins compared to the baseline group, allowing for both factors to vary. For example, if $I_{C,i} = I_{K,i} = 1$, then the term with β_9 would be non-zero, corresponding to the mean increase from the baseline group for the treatment where Cottonelle is used and high moisture is applied.
- $\epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ is an error term

To determine whether there are significant interactions between the brand and the amount of moisture on the number of coins that toilet paper can support, we perform the following hypothesis test:

$$H_0 : \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = 0$$
$$H_A : \text{at least one } \beta_i \neq 0 \text{ for } i \in \{7, \ldots, 14\}$$

6.1 Fitted Model

A summary for the fitted interaction + main effects model is in Table 3.

Table 3: Interaction Model Summary

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18.0	1.723175	10.4458318	0.0000000
brandcharmin	-3.4	2.436938	-1.3951935	0.1681005
brandcottonelle	-5.6	2.436938	-2.2979658	0.0250654
brandkirkland	2.2	2.436938	0.9027723	0.3702569
brandroyale	-3.6	2.436938	-1.4772637	0.1448350
moisturemedium	-9.2	2.436938	-3.7752295	0.0003689
moisturehigh	-11.2	2.436938	-4.5959315	0.0000227
brandcharmin:moisturemedium	8.4	3.446351	2.4373608	0.0177763
brandcottonelle:moisturemedium	5.8	3.446351	1.6829396	0.0975840
brandkirkland:moisturemedium	8.2	3.446351	2.3793284	0.0205416
brandroyale:moisturemedium	7.8	3.446351	2.2632636	0.0272501
brandcharmin:moisturehigh	4.4	3.446351	1.2767128	0.2066229
brandcottonelle:moisturehigh	3.8	3.446351	1.1026156	0.2745972
brandkirkland:moisturehigh	4.6	3.446351	1.3347452	0.1870027
brandroyale:moisturehigh	5.4	3.446351	1.5668748	0.1224033

Table 4: Interaction Model ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
brand	4	635.9200	158.98000	10.708127	0.0000013
moisture	2	720.8267	360.41333	24.275707	0.0000000
brand:moisture	8	129.4400	16.18000	1.089807	0.3828059
Residuals	60	890.8000	14.84667	NA	NA

6.2 Hypothesis Testing - model with interaction & main effect

We want to determine if there is a significant interaction effect between the toilet paper brands and the moisture levels on the number of coins that the toilet paper can support. That is, we want to perform the hypothesis test:

$$H_0 : \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = 0$$
$$H_A : \text{at least one } \beta_i \neq 0 \text{ for } i \in \{7, \dots, 14\}$$

From the ANOVA table of the interaction model (Table 4), we compute the test statistic $F = \frac{MS\,brand{:}moisture}{MS\;Residuals} = \frac{16.18}{14.8466667} \approx 1.09$.

We want to compare this to the distribution of $F_{8,60}$ since we are testing 8 coefficients and the degrees of freedom of error for the full model (Table 4) is 60.

Assuming the null hypothesis, we get a p-value of $P(F_{8,60} > 1.0898069) = 0.3828059$.

At 5% significance level, we fail to reject H_0: there is insufficient evidence to conclude that the effect of brand on coins varies by moisture level.

In other words, there is no significant interaction effects between the toilet paper brands and the moisture levels on the number of coins that the toilet paper can support. Thus, we can fit the simpler additive model, which only includes the main effects, as our final model.

7 Additive Model

We define the following additive model using dummy coding:

$$E(Y_i) = \beta_0 + \underbrace{\beta_1 I_{B,i} + \beta_2 I_{C,i} + \beta_3 I_{D,i} + \beta_4 I_{E,i}}_{\text{Terms for brands}}$$
$$+ \underbrace{\beta_5 I_{K,i} + \beta_6 I_{L,i}}_{\text{Terms for water applied}} + \epsilon_i$$

where the terms are the same as in the interaction model, but without the interaction terms:

- The indicators $I_{a,i}$ where $a \in \{A, \dots, E\} \cup \{J, K, L\}$ represent whether a treatment includes that factor level. For example, a treatment where Cottonelle is used and high moisture is applied would have $I_{C,i} = I_{K,i} = 1$ and all other indicators equal to 0.
- β_0 represents the mean number of coins supported for the baseline group: Cashmere toilet paper with low moisture (0.5mL applied)
- $\beta_1, \dots, \beta_4$ each represent the mean increase in the number of coins for each brand (Charmin, Cottonelle, Kirkland, Royale, respectively) compared to the baseline group, with low moisture
- β_5, β_6 each represent the mean increase in the number of coins for (medium moisture, high moisture respectively) compared to the baseline group, for Cashmere toilet paper.
- $\epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ is an error term

To determine whether each factor has significant effects on the number of coins that the toilet paper can support while holding the other fixed, we perform the following hypothesis tests:

- Brand Effect

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$
$$H_A : \text{at least one } \beta_i \neq 0 \text{ for } i \in \{1, \dots, 4\}$$

- Moisture Effect

$$H_0 : \beta_5 = \beta_6 = 0$$
$$H_A : \text{at least one } \beta_i \neq 0 \text{ for } i \in \{5, 6\}$$

7.1 Fitted Model

A summary for the fitted additive model is in Table 5.

Table 5: Additive Model Summary

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	14.7733333	1.183355	12.4842769	0.0000000
brandcharmin	0.8666667	1.414380	0.6127538	0.5420831
brandcottonelle	-2.4000000	1.414380	-1.6968567	0.0942970
brandkirkland	6.4666667	1.414380	4.5720860	0.0000210
brandroyale	0.8000000	1.414380	0.5656189	0.5735146
moisturemedium	-3.1600000	1.095574	-2.8843328	0.0052491
moisturehigh	-7.5600000	1.095574	-6.9004924	0.0000000

Table 6: Additive Model ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
brand	4	635.9200	158.98000	10.59617	9.868e-07
moisture	2	720.8267	360.41333	24.02190	1.280e-08
Residuals	68	1020.2400	15.00353	NA	NA

7.2 Hypothesis Testing - additive model, brand effect

From the ANOVA table of the additive model (Table 6), we compute the test statistic
$F = \frac{MS\,brand}{MSE} = \frac{158.98}{15.0035294} \approx 10.596.$

Assuming the null hypothesis, we get a p-value of $P(F_{4,68} > 10.5961734) = 9.8681288 \times 10^{-7}$.

At a 5% significance level, we reject H_0: strong evidence to conclude that brand affects coins when controlling for moisture level. (i.e., # coins supported vary from one brand to another within the same moisture level)

7.3 Hypothesis Testing - additive model, moisture effect

From the ANOVA table of the additive model (Table 6), we compute the test statistic
$F = \frac{MS\,Moisture}{MSE} = \frac{360.4133333}{15.0035294} \approx 24.022.$

Assuming the null hypothesis, we get a p-value of $P(F_{2,68} > 24.0219033) = 1.283 \times 10^{-8}$.

At a 5% significance level, we reject H_0: strong evidence to conclude that moisture level affects coins when controlling for brand.

8 Conclusion

In summary, the main purpose of this report was to find out whether or not toilet paper brands and/or different levels of moisture affect the strength of toilet paper, as well as noting any interactions between the two factors. After observing our interaction plot, model summary, and ANOVA table, we found that there was inconclusive evidence to show a significant effect between the toilet paper brands and the moisture levels on the number of coins the toilet paper can support.

With this first conclusion in mind, we then looked at the plain additive model and focused on the main effects. We concluded that each factor (while holding the other factor constant) had a statistically significant effect on the number of coins supported, at the 5% benchmark significance level.

9 Appendix - R code

9.1 Sample size calculation

```
sample_size_calc = ss.2way(a=5, b=3, alpha=0.05, beta=1-0.8, f.A=0.25, f.B=0.25, B=4)
```

9.2 Interaction plot

```
interaction.plot(brand, moisture, coins,
                 col=c("cyan", "darkorchid1", "blue"),
                 lty=2,
                 lwd=2,
                 xlab="Brand Mean",
                 ylab="Number of Coins")
```

9.3 Model generation

```
# Function to convert fitted model to df with convenient col/rownames
model_to_df <- function(model) {
  df <- tidy(model) %>% as.data.frame

  # Use cleaned predictors names as dataframe row names
  varnames <- df$term %>%
    lapply(function(x) gsub("[^A-Za-z0-9]", "", x)) # keep only alphanumeric chars in row names
  rownames(df) <- varnames
  df$term <- NULL
  return(df)
}

# Function to add useful terms to a named list with its model.
gen_model_summaries <- function(vec) {
  # Pull out useful values from model summaries to data frame and to vec scope
  vec$dataframe <- model_to_df(vec$model)
  vec$modelsummary <- summary(vec$model)
  vec$anovatable <- anova(vec$model) %>% as.data.frame

  return(vec)
}
```

```r
# Loading in the data & attaching col vectors to global scope, rename "some" -> "medium"
toiletPaperData <- read.csv("toilet_paper_data.csv") %>%
  mutate(moisture = str_replace(moisture, "some", "medium"))
  replace("some", "medium")
attach(toiletPaperData)

# Force an order for the data variables
toiletPaperData$moisture <- factor(toiletPaperData$moisture, levels=c("low", "medium", "high"))

# Model with interaction
int <- list(model=lm(coins ~ brand * moisture))

# Pull out useful values from model summaries to data frame and to object scope
int <- gen_model_summaries(int)

# Additive Model: model with main effects only
additive <- list(model=lm(coins ~ brand + moisture))

# Pull out useful values from model summaries to data frame and to object scope
additive <- gen_model_summaries(additive)
```

9.4 Assessing Model Assumptions

```r
# Assessing independence of the error terms using residual vs fitted plots
par(mfrow=c(1, 2))
plot(int$model, which = 1, main="Interaction Model")
plot(additive$model, which = 1, main="Additive Model")

# Assessing identical distribution of the error terms using scale-location plots
par(mfrow=c(1, 2))
plot(int$model, which=3, main="Interaction Model")
plot(additive$model, which=3, main="Additive Model")

# Assessing normality of the error terms using QQ plots
par(mfrow=c(1, 2))
plot(int$model, which=2)
plot(additive$model, which=2)

# Assessing normality of the error terms using Shapiro-Wilk tests
int$resid <- residuals(object = int$model)
additive$resid <- residuals(object = additive$model)
data.frame(
  Interaction=shapiro.test(int$resid) %>% tidy %>% select(-method),
  Additive=shapiro.test(additive$resid) %>% tidy %>% select(-method)
) %>%
  setNames(colnames(.) %>% lapply(function(x) gsub("\\.", " ", x))) %>%
  kable(caption="Interaction and Additive Model Shapiro-Wilk Tests")
```