Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Project 4: Understanding typing dynamics changes over time

Introduction:

Typing patterns vary significantly between individuals, making typing dynamics a promising biometric for enhancing password-based security systems. In their paper *"Comparing Anomaly-Detection Algorithms for Keystroke Dynamics"*, Dr. Roy Maxion and colleagues demonstrate how keystroke dynamics can provide an additional layer of security by analyzing unique typing behaviors.

The goal of this project is to investigate whether individuals' typing dynamics remain **consistent over time**. Specifically, we analyze the typing patterns of study participants as they repeatedly enter the passcode **".tie5Roanl"**.

For this study, **51 subjects** were recruited at Carnegie Mellon University (CMU) by Dr. Maxion and his team. Each participant completed **8 data collection sessions**, with each session occurring one day apart. Within each session, participants typed the passcode **50 times**, resulting in a robust dataset for evaluating typing consistency over multiple days.

According to the authors, the subject group consisted of **30 males and 21 females**, including **8 left-handed** and **43 right-handed** individuals. The age distribution ranged from **18 to 70 years**, with the **median age group being 31–40 years** (*Maxion et al.*).

The passcode **".tie5Roanl"** was designed to be representative of a typical strong password. The authors explain:

"To make a password that is representative of typical, strong passwords, we employed a publicly available password generator and password-strength checker. We generated a 10-character password containing letters, numbers, and punctuation, and then modified it slightly, interchanging some punctuation and casing to better conform with the general perception of a strong password."

The password-strength checker rated this password as **strong** because it meets key criteria: it includes **more than 7 characters**, a **capital letter**, a **number**, and **punctuation**. While the highest rating is reserved for passwords longer than 13 characters, the authors note that a **10-character password** is typical, as observed in prior studies (*Maxion et al.*).

Linear Mixed Effects Model for repeated measures data

Linear mixed models are an extension of simple linear models to allow for both fixed and random effects. This type of model is suitable for analyzing data with repeated measurements (subjects typing the password multiple times) because it accounts for the correlation between observations within the same subject.

Fixed Effects: A fixed effect is a parameter that does not vary. This is same as in linear regression where we assume that the data are random but parameters are fixed effect (coefficients). In our case, the fixed effect would be to check if typing speed changes across session for everyone.

Random Effects: It is parameter that are random variable themselves. This is where the individual differences come in, it is like recognizing each person have their own typing style and learning curve.

Random Intercept Models and Random Intercept and Slope Models are two commonly used Linear mixed effects model.

Random Intercept Models

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Assumes that each subject has a different baseline typing speed (intercept), but they all change at the same rate across sessions. The model assumes that random effects and residuals are randomly distributed with mean zero and variance σ_u^2 . Let y_{ij} represents the observation made at time t_j on individual i, the random intercept model for y_{ij} is $y_{ij} = \beta_0 + \beta_1 t_j + u_i + \varepsilon_{ij}$

Random Slope and Intercept Models: It capture the fact that individuals improve at different rates, i.e., allows for heterogeneity in both slope and intercept. In effect we are assuming that the repeated measures on each individual in our study can be characterized by their own individual regression model. Let y_{ij} represents the observation made at time t_j on individual i which has random slope v_i and random intercept u_i, the random intercept and slope model is given by, $y_{ij} = \beta_0 + \beta_1 t_j + u_i + v_i t_j + \varepsilon_{ij}$

The two random effects are assumed to have a bivariate normal distribution with zero means for both variables and variances σ_u^2 and σ_v^2 with covariance σ_{uv}.

By accounting for random effects, the model provides a more accurate picture of how typing speed changes over time, while recognizing different individual patterns

Dataset Summary:

We began the project by **loading and summarizing the dataset** while performing checks for missing values. The dataset contains **20,400 observations** and **34 variables**, providing a robust foundation for analysis.

Dataset Variables were concluded as:

H.period: The amount of time that the "." is held down.
DD.period.t: The time between pressing down the "." key to the time to press down the "t" key.
UD.period.t: The time between the "." key coming up to the time to press down the "t" key.
H.t: The amount of time that the "t" is held down.
DD.t.i: The time between pressing down the "t" key to the time to press down the "i" key.
UD.t.i: The time between the "t" key coming up to the time to press down the "i" key.
H.i: The amount of time that the "i" is held down.
DD.i.e: The time between pressing down the "i" key to the time to press down the "e" key.
UD.i.e: The time between the "i" key coming up to the time to press down the "e" key.
H.e: The amount of time that the "e" is held down.
DD.e.five: The time between pressing down the "e" key to the time to press down the "5" key.
UD.e.five: The time between the "e" key coming up to the time to press down the "5" key.
H.five: The amount of time that the "5" is held down.
DD.five.Shift.r: The time between pressing down the "5" key to the time to press down the "shift+r" key combination.
UD.five.Shift.r: The time between the "5" key coming up to the time to press down the "shift+r" key combination.
H.Shift.r: The amount of time that the "shift+r" key combination is held down.
DD.Shift.r.o: The time between pressing down the "shift+r" key combination to the time to press down the "o" key.
UD.Shift.r.o: The time between the "shift+r" key combination coming up to the time to press down the "o" key.
H.o: The amount of time that the "o" is held down.
DD.o.a: The time between pressing down the "o" key to the time to press down the "a" key.
UD.o.a: The time between the "o" key coming up to the time to press down the "a" key.
H.a: The amount of time that the "a" is held down.
DD.a.n: The time between pressing down the "a" key to the time to press down the "n" key.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

UD.a.n: The time between the "a" key coming up to the time to press down the "n" key.
H.n: The amount of time that the "n" is held down.
DD.n.l: The time between pressing down the "n" key to the time to press down the "l" key.
UD.n.l: The time between the "n" key coming up to the time to press down the "l" key.
H.l: The amount of time that the "l" is held down.
DD.l.Return: The time between pressing down the "l" key to the time to press down the "return" key.
UD.l.Return: The time between the "l" key coming up to the time to press down the "return" key.
H.Return: The amount of time that the "return" is held down.

During the summary review, we identified **outliers** in several variables. Notably:

- **DD.i.e** exhibited an outlier value of **25.9873**, which is **25.8279 above the mean**.
- **UD.i.e** exhibited an outlier value of **25.9158**, which is **25.8380 above the mean**.
- **DD.period.t** exhibited an outlier value of **12.5061**, which is **12.242 above the mean**.
- **UD.period.t** exhibited an outlier value of **12.4517**, which is **12.2809 above the mean**.
- **DD.five.Shift.r** exhibited an outlier value of **8.3702**, which is **7.9313 above the mean**.
- **UD.five.Shift.r** exhibited an outlier value of **8.2908**, which is **7.9288 above the mean**.

These and other outliers were carefully reviewed and addressed during the **Exploratory Data Analysis (EDA)** process (See Table 1 below).

Additionally, we conducted a review of the **51 subjects** within the dataset. It was observed that some subjects may have been either **removed or renamed**, resulting in a discrepancy between the dataset and reported subject count (See Table 2 below).
EDA:

Our initial **Exploratory Data Analysis (EDA)** involved calculating summary statistics for **three key response metrics**:

1. **Hold Times (H.*)**
2. **Down-Down Times (DD.*)**
3. **Up-Down Times (UD.*)**

We visualized the distributions of these metrics by creating corresponding **histograms** (See Tables 3-5 below).

To further understand the relationships between timing variables, we created a **correlation graph** for Down-Down Times (DD.*) and Up-Down Times (UD.*) (See **Table 6 below**). From this analysis, we observed that **most correlations are positive**, indicating that as one timing variable increases, the others tend to increase as well.

Response Variables:

We created two response variables: **TotalTypingTime** and **ud_sum**.

1. **TotalTypingTime**:
   a. This variable represents the total time to type the passcode, calculated as the sum of all **H*** (hold times) and **UD*** (up-down times).
   b. During our exploratory analysis, we noticed that the **DD*** (down-down) variables are equivalent to the combined total of **H*** and **UD***. To avoid redundancy, we focused on **H*** and **UD*** variables to calculate typing time more efficiently.

c. We then created a streamlined dataset, named passcode.total.dat, containing only the relevant variables:
    i. **subject** (participant ID)
    ii. **sessionIndex** (session number)
    iii. **rep** (trial within the session)
    iv. TotalTypingTime

2. **ud_sum**:
    a. This variable represents the sum of all **UD*** (up-down) variables for each participant across sessions.
    b. The **UD*** times were isolated because they provide insights into the transitions between key presses, independent of hold times.

Modeling:

To analyze these response variables, we fit **Linear Mixed-Effects Models** (LMMs) to account for repeated measures across multiple sessions and subjects (See **Random Intersect Model: Table 7 and Random Slope Model: Table 8** below):

1. **For TotalTypingTime (Comparison see Table 9 below)**:
    a. We modeled it as a function of sessionIndex to see how typing time changes across sessions.
    b. Random effects for **subjects** were included to capture individual variability.
    c. Checked for residuals (see **Table 10 below**)
    d. Due to large outliers, we had used log transformation to normalize the data (see **Tables 11 and 12 below**)
2. **For ud_sum**:
    a. Similarly, we examined how the sum of **UD*** values change across sessions while accounting for subject-level differences (See **Tables 13 and 14 below)**.

Pairwise Comparison:
To complete the analysis, we converted sessionIndex into a categorical factor for the **Total Typing Time** variable. This ensured that the model treated sessions as distinct categories rather than numeric values. We then refitted the linear mixed-effects model and performed pairwise comparisons to evaluate differences in typing times across sessions. This allowed us to identify significant differences between sessions while accounting for multiple comparisons (See **Tables 15 and 16 below**). Additionally, we checked the data distribution for Total Typing Time to confirm consistency across sessions.

For the **ud_sum** variable, sessionIndex was left as a numeric variable, as its progression over time better reflects trends in the sum of **Up-Down (UD)** times (See **Tables 17 and 18 below**). This approach allowed us to analyze the relationship between sessions and ud_sum without categorizing sessions explicitly.

Conclusion:

This analysis explored typing dynamics by investigating how two response variables, TotalTypingTime and ud_sum, change over multiple sessions. The results show significant reductions in both total typing time and UD times as participants completed repeated sessions, highlighting consistent improvement in typing speed and efficiency. Linear Mixed-Effects Models revealed that improvements vary among participants, with those starting slower showing greater gains. These findings emphasize the potential of typing dynamics to provide reliable and measurable behavioral patterns over time, which could enhance biometric applications and authentication systems.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 1: Descriptive Summary Statistics

```
    H.period          DD.period.t        UD.period.t            H.t                DD.t.i       
 Min.   :0.00140   Min.   : 0.0187   Min.   :-0.2358   Min.   :0.00930   Min.   :0.0011  
 1st Qu.:0.07440   1st Qu.: 0.1469   1st Qu.: 0.0498   1st Qu.:0.06600   1st Qu.:0.1136  
 Median :0.08950   Median : 0.2059   Median : 0.1087   Median :0.08100   Median :0.1404  
 Mean   :0.09338   Mean   : 0.2641   Mean   : 0.1708   Mean   :0.08573   Mean   :0.1691  
 3rd Qu.:0.10790   3rd Qu.: 0.3064   3rd Qu.: 0.2124   3rd Qu.:0.09980   3rd Qu.:0.1839  
 Max.   :0.37610   Max.   :12.5061   Max.   :12.4517   Max.   :0.24110   Max.   :4.9197  
     UD.t.i              H.i               DD.i.e             UD.i.e               H.e         
 Min.   :-0.16210   Min.   :0.00320   Min.   : 0.0014   Min.   :-0.16000   Min.   :0.00210  
 1st Qu.: 0.02720   1st Qu.:0.06200   1st Qu.: 0.0893   1st Qu.: 0.00740   1st Qu.:0.06860  
 Median : 0.05780   Median :0.07710   Median : 0.1209   Median : 0.04120   Median :0.08340  
 Mean   : 0.08336   Mean   :0.08157   Mean   : 0.1594   Mean   : 0.07781   Mean   :0.08914  
 3rd Qu.: 0.09640   3rd Qu.:0.09690   3rd Qu.: 0.1731   3rd Qu.: 0.09340   3rd Qu.:0.10270  
 Max.   : 4.79990   Max.   :0.33120   Max.   :25.9873   Max.   :25.91580   Max.   :0.32540  
   DD.e.five         UD.e.five            H.five        DD.five.Shift.r  UD.five.Shift.r 
 Min.   :0.0013   Min.   :-0.1505   Min.   :0.0014   Min.   :0.1694   Min.   :0.0856  
 1st Qu.:0.2166   1st Qu.: 0.1332   1st Qu.:0.0610   1st Qu.:0.3079   1st Qu.:0.2297  
 Median :0.2890   Median : 0.2004   Median :0.0742   Median :0.3775   Median :0.3020  
 Mean   :0.3774   Mean   : 0.2883   Mean   :0.0769   Mean   :0.4389   Mean   :0.3620  
 3rd Qu.:0.4568   3rd Qu.: 0.3694   3rd Qu.:0.0906   3rd Qu.:0.4860   3rd Qu.:0.4089  
 Max.   :4.9618   Max.   : 4.8827   Max.   :0.1989   Max.   :8.3702   Max.   :8.2908  
   H.Shift.r         DD.Shift.r.o      UD.Shift.r.o           H.o               DD.o.a      
 Min.   :0.00140   Min.   :0.0494   Min.   :-0.0865   Min.   :0.00690   Min.   :0.0012  
 1st Qu.:0.07020   1st Qu.:0.1565   1st Qu.: 0.0547   1st Qu.:0.07150   1st Qu.:0.1064  
 Median :0.09350   Median :0.2014   Median : 0.1022   Median :0.08630   Median :0.1316  
 Mean   :0.09594   Mean   :0.2509   Mean   : 0.1550   Mean   :0.08835   Mean   :0.1569  
 3rd Qu.:0.11670   3rd Qu.:0.2834   3rd Qu.: 0.1910   3rd Qu.:0.10190   3rd Qu.:0.1676  
 Max.   :0.28170   Max.   :4.1523   Max.   : 4.0120   Max.   :0.68720   Max.   :2.8567  
     UD.o.a              H.a             DD.a.n            UD.a.n               H.n        
 Min.   :-0.22870   Min.   :0.0040   Min.   :0.0011   Min.   :-0.23550   Min.   :0.0037  
 1st Qu.: 0.01700   1st Qu.:0.0821   1st Qu.:0.0961   1st Qu.:-0.00900   1st Qu.:0.0673  
 Median : 0.04440   Median :0.1019   Median :0.1250   Median : 0.02270   Median :0.0853  
 Mean   : 0.06858   Mean   :0.1063   Mean   :0.1507   Mean   : 0.04441   Mean   :0.0899  
 3rd Qu.: 0.08030   3rd Qu.:0.1223   3rd Qu.:0.1746   3rd Qu.: 0.06890   3rd Qu.:0.1079  
 Max.   : 2.81520   Max.   :2.0353   Max.   :3.3278   Max.   : 2.52420   Max.   :0.3577  
     DD.n.l           UD.n.l              H.l            DD.l.Return      UD.l.Return     
 Min.   :0.0013   Min.   :-0.1758   Min.   :0.00370   Min.   :0.0083   Min.   :-0.1245  
 1st Qu.:0.1276   1st Qu.: 0.0235   1st Qu.:0.07740   1st Qu.:0.2100   1st Qu.: 0.1141  
 Median :0.1725   Median : 0.0955   Median :0.09370   Median :0.2630   Median : 0.1603  
 Mean   :0.2026   Mean   : 0.1127   Mean   :0.09559   Mean   :0.3218   Mean   : 0.2263  
 3rd Qu.:0.2288   3rd Qu.: 0.1457   3rd Qu.:0.11110   3rd Qu.:0.3502   3rd Qu.: 0.2551  
 Max.   :4.0252   Max.   : 3.9782   Max.   :0.34070   Max.   :5.8836   Max.   : 5.8364  
    H.Return      
 Min.   :0.00290  
 1st Qu.:0.06990  
 Median :0.08550  
 Mean   :0.08831  
 3rd Qu.:0.10370  
 Max.   :0.26510  
```

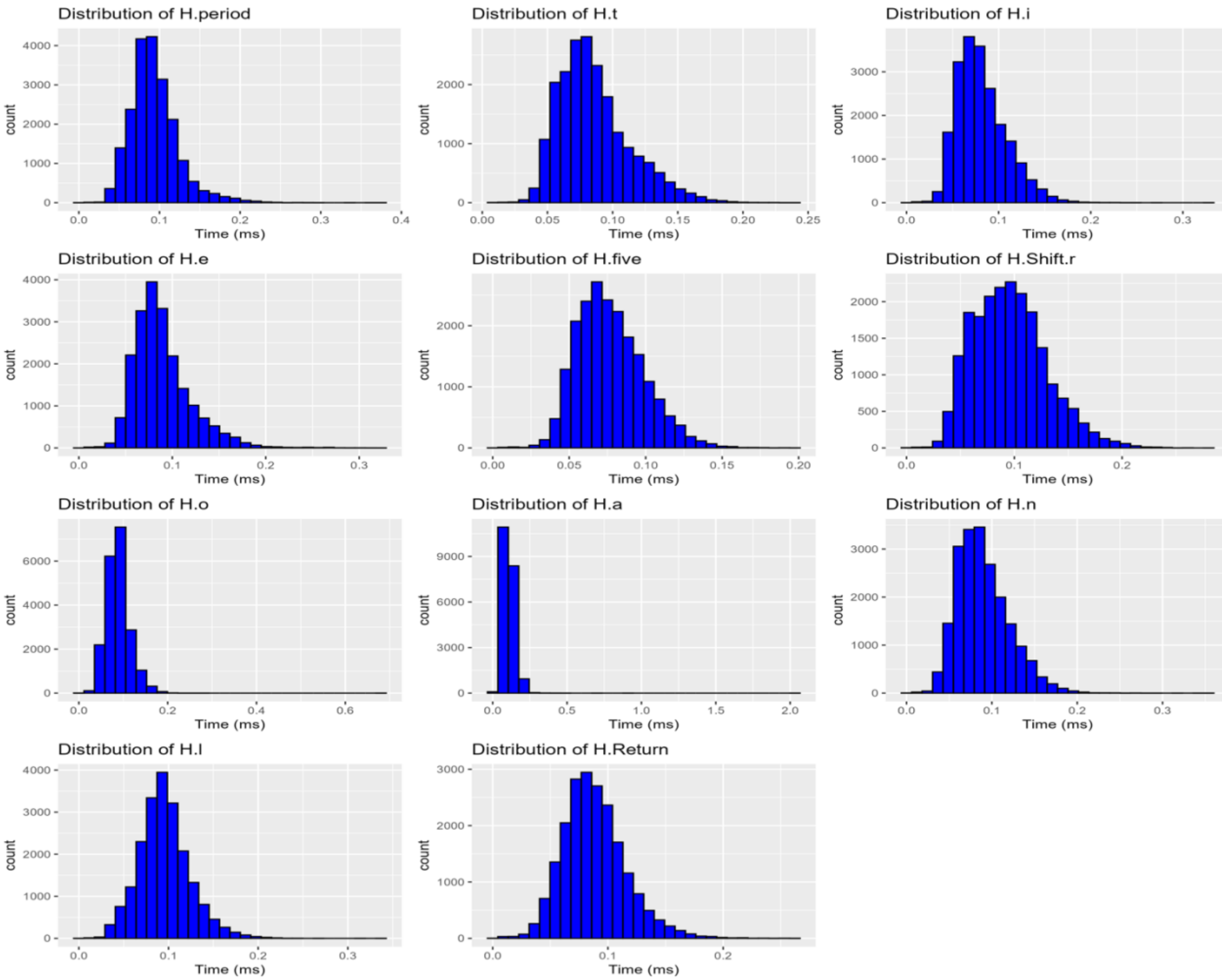
Table 2: List of subjects

```
Subjects:
 s002, s003, s004, s005, s007, s008, s010, s011, s012, s013, s015, s016, s017, s018, s019, s020, s021, s022,
s024, s025, s026, s027, s028, s029, s030, s031, s032, s033, s034, s035, s036, s037, s038, s039, s040, s041,
s042, s043, s044, s046, s047, s048, s049, s050, s051, s052, s053, s054, s055, s056, s057
```

There are 51 total subjects.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

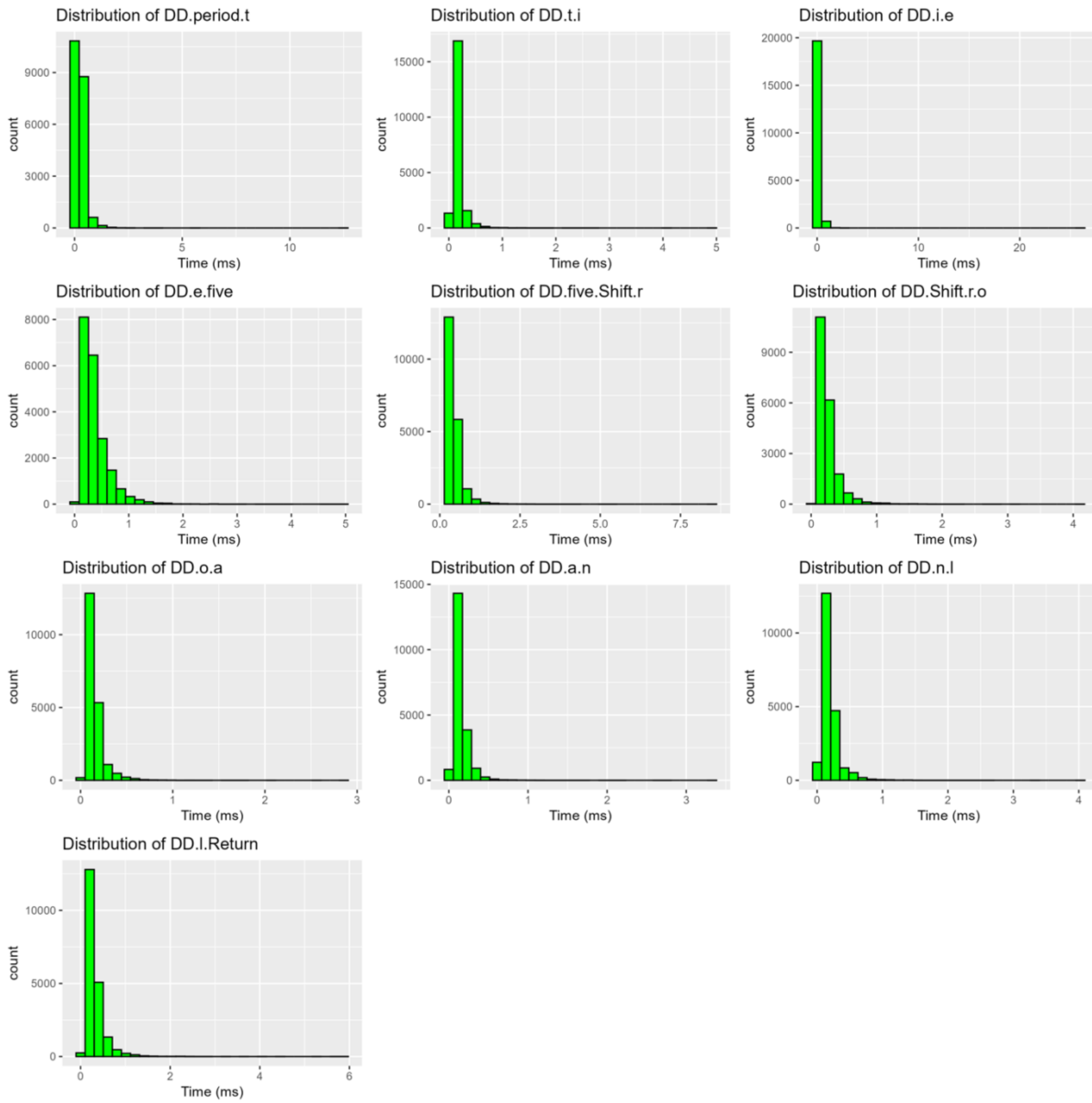
Table 3: Histogram for Distribution of Hold time



Comment on Histogram: The above histogram shows unimodal distribution of Hold time which are mostly skewed except for the H.Return and H.I which appears to be symmetric

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

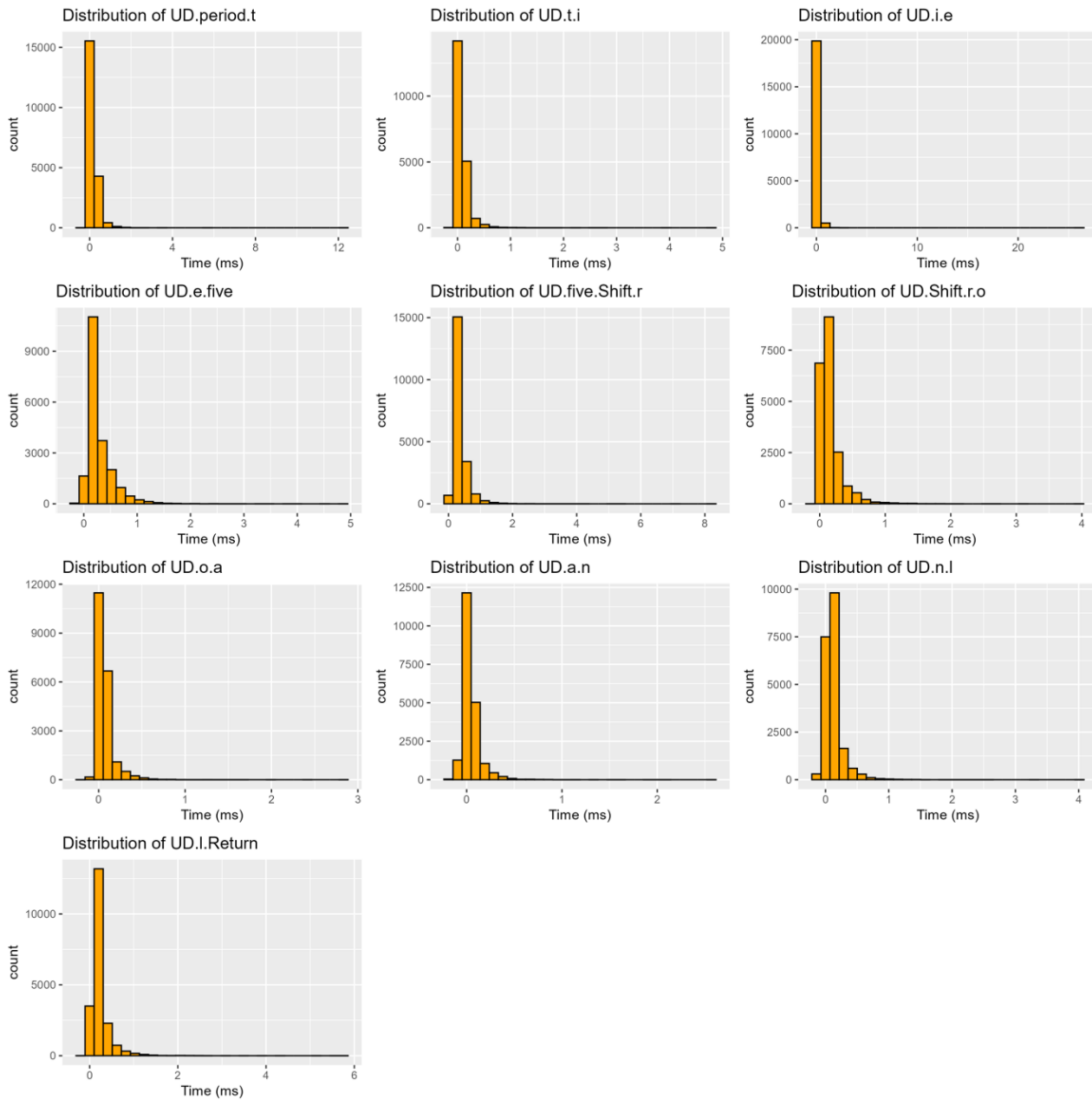
Table 4: Histogram showing Distribution of Down-Down Key time



Comment on the histogram: The above histogram shows unimodal distribution of Down-Down time represents a skewed distribution with potential outliers.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

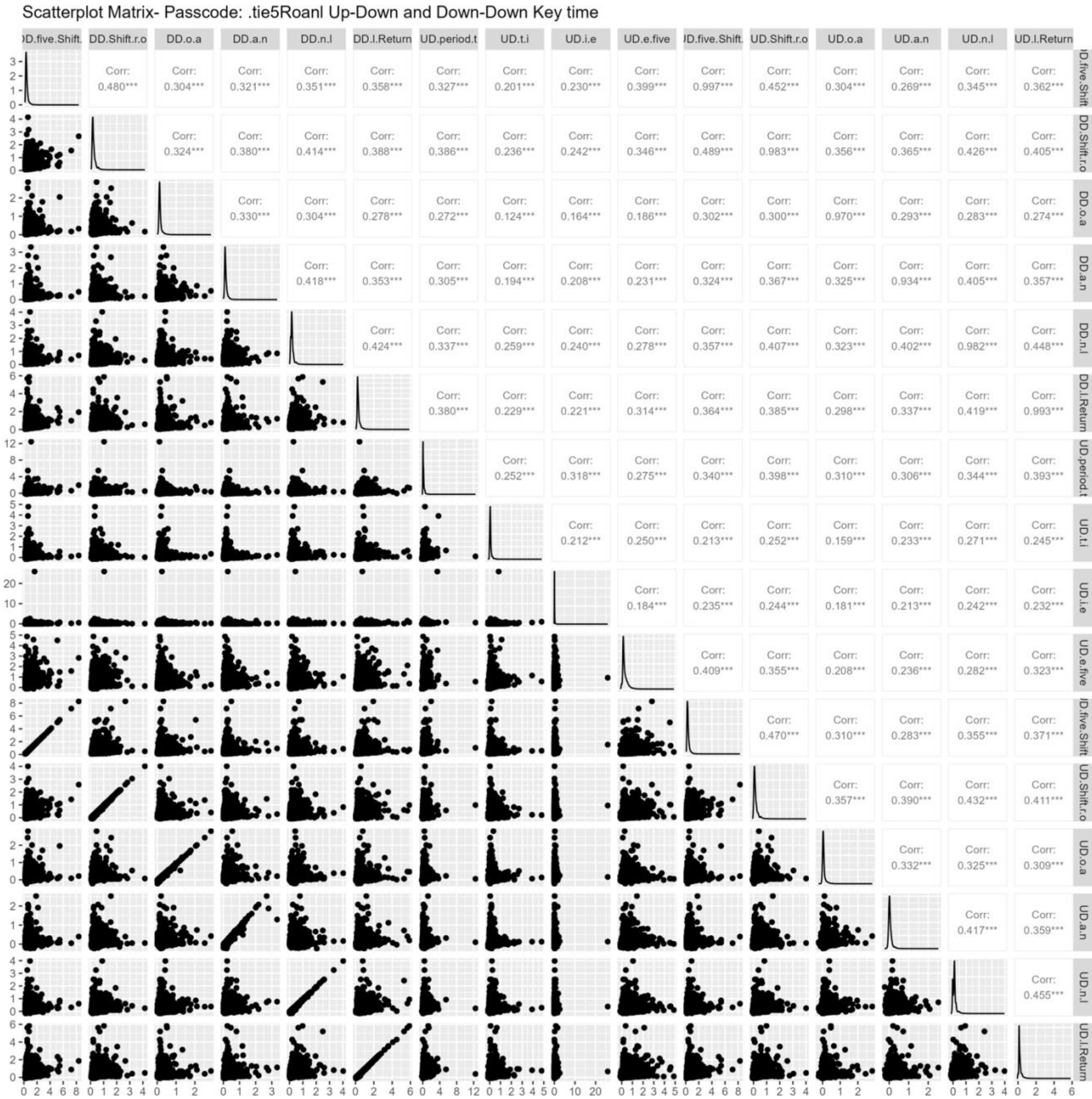
Table 5: Histogram showing Distribution of Up-Down Key time



Comment on Histogram: The above histogram for all ten Up-Down variables is highly skewed to the right.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 6: Scatterplot Matrix- Passcode: Up-Down and Down-Down Key time



- Most of the UD time is highly correlated with DD time. Such as UD.five.Shift--DD.five.Shift, UD.Shift.r.o--DD.Shift.r.o, UD.o.a--DD.o.a and other are highly correlated with each other.
- Additionally, looking at the timing variable we found that Down-Down (DD) Key is the sum of Up-Down (UD) and Hold (H) Key.
  - DD is the measure of the time between pressing down a certain key to pressing down another subsequent key.
  - UD is the measure of the time a certain key is coming up to the time another subsequent key is pressed down.
  - H is the amount of the time a certain key is held down

This leads us to believe that `Total Typing Time` can be calculated as the sum of all `UDs` and `Hs` and hence our response variable. Hence `Total Typing Time` is our response variable which is calculated for each rep and session.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Total Typing Time (Sum of Hs and UDs) as Response Variable

Table 7: Random Intercept Model for Total Typing Time

```
Error in install.packages : Updating loaded packages
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: TotalTypingTime ~ sessionIndex + (1 | subject)
   Data: passcode.total.dat

REML criterion at convergence: 43284.8

Scaled residuals: 
   Min     1Q Median     3Q    Max 
-3.252 -0.466 -0.130  0.273 43.972 

Random effects:
 Groups   Name        Variance Std.Dev.
 subject  (Intercept) 0.7515   0.8669  
 Residual             0.4806   0.6933  
Number of obs: 20400, groups:  subject, 51

Fixed effects:
               Estimate Std. Error         df t value Pr(>|t|)    
(Intercept)   3.144e+00  1.219e-01  5.062e+01   25.80   <2e-16 ***
sessionIndex -1.252e-01  2.118e-03  2.035e+04  -59.11   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr)
sessionIndx -0.078
```

Comment on Output for Random Intercept Model for Total Typing Time:
The model analyzes “TotalTypingTime” as the response variable which is the total time taken to type the given passcode. SessionIndex used as a predictor suggesting that the study looks at how typing time changes over multiple sessions. The result shows that on an average the total typing time is 3.144 seconds which significantly decreases by 0.12522 seconds per session.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 8: Random Slope and Intercept Model of Total Typing Time

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: TotalTypingTime ~ sessionIndex + (sessionIndex | subject)
   Data: passcode.total.dat

REML criterion at convergence: 40789.2

Scaled residuals:
   Min      1Q Median      3Q     Max
-5.247 -0.466 -0.148  0.256 43.727

Random effects:
 Groups   Name         Variance Std.Dev. Corr
 subject  (Intercept)  1.59117  1.2614
          sessionIndex 0.01142  0.1069   -0.88
 Residual              0.42168  0.6494
Number of obs: 20400, groups:  subject, 51

Fixed effects:
             Estimate Std. Error       df t value Pr(>|t|)
(Intercept)   3.14373    0.17692 50.00290  17.769  < 2e-16 ***
sessionIndex -0.12522    0.01509 50.00389  -8.296 5.82e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr)
sessionIndx -0.880
```

The output shows that the average starting total time is 3.143 second which significantly decreases by 0.12522 seconds per session.

The Random effect variance, sessionIndex 0.01142 represents that the effect of sessionIndex is different for different subject. The Random Intercept Variance subject (Intercept) 1.59117 represents that subjects have different starting points (baseline) for total typing time.

Table 9: Model Comparison of Random Intercept and Random Slope Models

```
Data: passcode.total.dat
Models:
lmer_model: TotalTypingTime ~ sessionIndex + (1 | subject)
lmer_model_slope: TotalTypingTime ~ sessionIndex + (sessionIndex | subject)
                 npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
lmer_model          4 43280 43312 -21636    43272
lmer_model_slope    6 40791 40839 -20390    40779 2492.5  2  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The results compare two models to see how Total Typing Time changes across sessions while accounting for differences between subjects. Both models show that typing time gets faster with each session. The simpler model assumes everyone improves at the same rate but starts at different typing times. The more flexible model allows each subject to have their own starting time and their own rate of improvement.

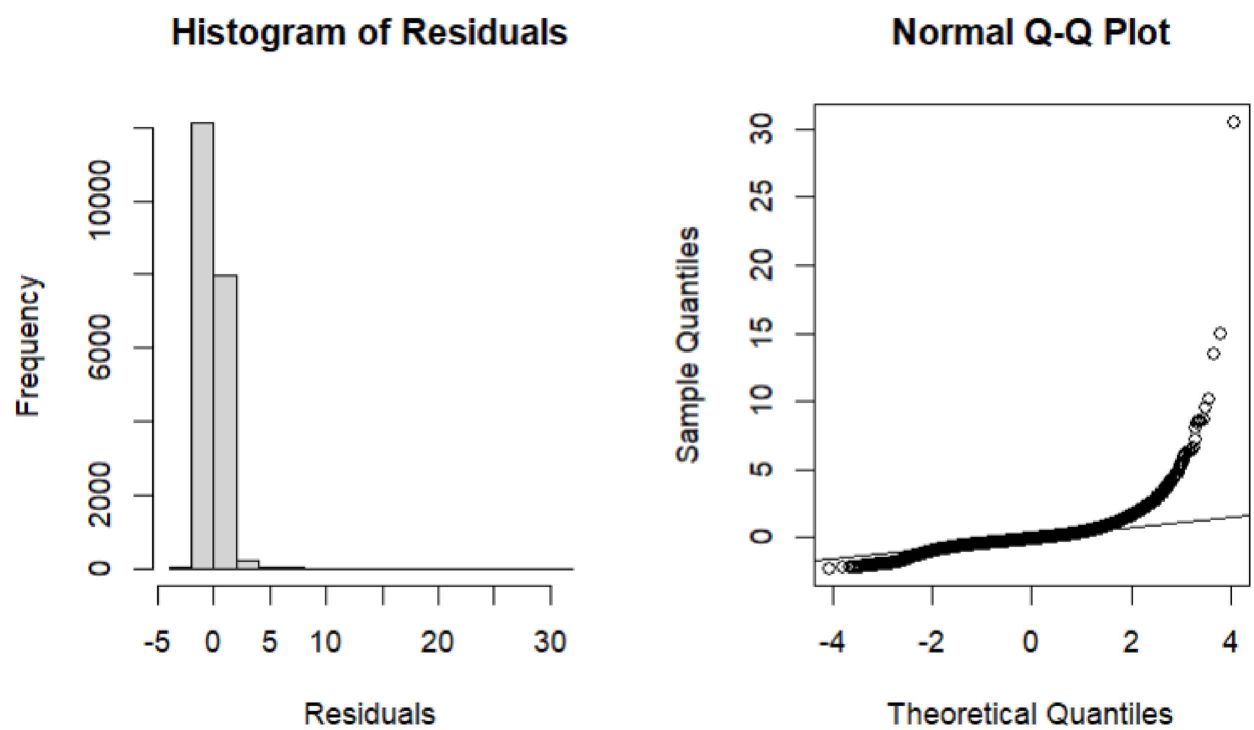
Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

The results show this flexible model fits the data much better (p-value < 2.2e-16), meaning some people improve faster than others. There's also a strong negative correlation (-0.88) between starting times and improvement rates, suggesting that people who started off slower tended to improve the most over time.

When we compare the model Random Slope and Intercept model has lower AIC and BIC and higher log likelihood, suggesting that lmer_model_slope fits better than the simpler lmer_model.

Post Hoc: Check Assumption of Normality

Table 10: Distribution of Residuals



Residuals don't seem to follow normal distribution as they deviate far from normal line.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Log Transformation of Total Typing Time as Response Variable

The response variable Total typing time was transformed with log transformation. This address skewness in the data and stabilizes the variance and make the relationship between Total Typing Time and session Index more linear.

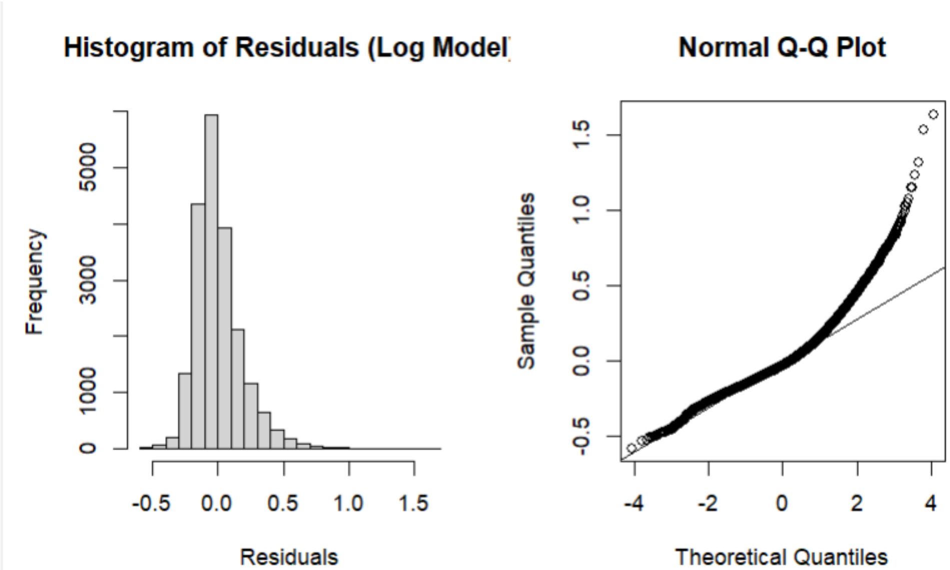
Table 11 Linear Slope and Intercept Model for Log transformed Total Typing Time

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log_TotalTypingTime ~ sessionIndex + (sessionIndex | subject)
   Data: passcode.total.dat
Control: lmerControl(optimizer = "nloptwrap", optCtrl = list(maxfun = 1e+05))

REML criterion at convergence: -12183.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2723 -0.6507 -0.1763  0.4627  9.2367

Random effects:
 Groups   Name         Variance  Std.Dev. Corr
 subject  (Intercept) 0.1013954 0.31843
          sessionIndex 0.0003607 0.01899  -0.47
 Residual             0.0313951 0.17719
Number of obs: 20400, groups:  subject, 51

Fixed effects:
              Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   1.073962   0.044672 50.009895   24.04   <2e-16 ***
sessionIndex -0.043441   0.002714 49.997597  -16.01   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr)
sessionIndx -0.469
optimizer (nloptwrap) convergence code: 0 (OK)
Model failed to converge with max|grad| = 0.00257757 (tol = 0.002, component 1)
```

The output shows that the average total typing time decreases over the session.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 12: Check Normality Assumption for lmer_model_log



The residual still appears to deviate from a normal distribution, as indicated by the tall, narrow peak in the histogram and deviation from the straight line in a normal Q-Q plot. This suggests the assumption of normality for the residuals is not met, indicating persistent issue with the model's fit despite log transformation.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 12 Estimated Marginal Means (EMM) for lmer_model_log

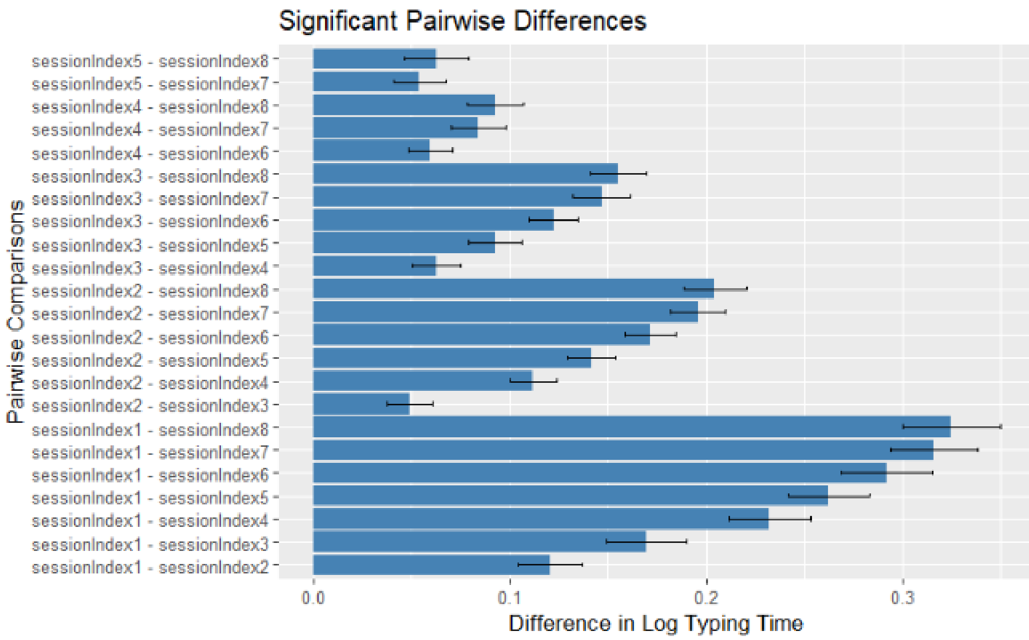
```
$emmeans
 sessionIndex emmean     SE  df asymp.LCL asymp.UCL
 1             1.093 0.0466 Inf     1.002     1.184
 2             0.973 0.0403 Inf     0.894     1.052
 3             0.924 0.0427 Inf     0.840     1.007
 4             0.861 0.0430 Inf     0.777     0.945
 5             0.831 0.0411 Inf     0.750     0.912
 6             0.801 0.0397 Inf     0.723     0.879
 7             0.777 0.0383 Inf     0.702     0.852
 8             0.768 0.0412 Inf     0.687     0.849

Degrees-of-freedom method: asymptotic
Confidence level used: 0.95

$contrasts
 contrast                      estimate      SE  df z.ratio p.value
 sessionIndex1 - sessionIndex2  0.12030 0.01620 Inf   7.438  <.0001
 sessionIndex1 - sessionIndex3  0.16935 0.02030 Inf   8.331  <.0001
 sessionIndex1 - sessionIndex4  0.23227 0.02100 Inf  11.080  <.0001
 sessionIndex1 - sessionIndex5  0.26211 0.02080 Inf  12.618  <.0001
 sessionIndex1 - sessionIndex6  0.29181 0.02290 Inf  12.749  <.0001
 sessionIndex1 - sessionIndex7  0.31603 0.02180 Inf  14.488  <.0001
 sessionIndex1 - sessionIndex8  0.32474 0.02470 Inf  13.139  <.0001
 sessionIndex2 - sessionIndex3  0.04904 0.01150 Inf   4.248  0.0006
 sessionIndex2 - sessionIndex4  0.11196 0.01200 Inf   9.327  <.0001
 sessionIndex2 - sessionIndex5  0.14180 0.01220 Inf  11.613  <.0001
 sessionIndex2 - sessionIndex6  0.17151 0.01280 Inf  13.382  <.0001
 sessionIndex2 - sessionIndex7  0.19573 0.01410 Inf  13.889  <.0001
 sessionIndex2 - sessionIndex8  0.20444 0.01580 Inf  12.970  <.0001
 sessionIndex3 - sessionIndex4  0.06292 0.01220 Inf   5.141  <.0001
 sessionIndex3 - sessionIndex5  0.09276 0.01370 Inf   6.764  <.0001
 sessionIndex3 - sessionIndex6  0.12246 0.01260 Inf   9.711  <.0001
 sessionIndex3 - sessionIndex7  0.14668 0.01450 Inf  10.135  <.0001
 sessionIndex3 - sessionIndex8  0.15539 0.01430 Inf  10.890  <.0001
 sessionIndex4 - sessionIndex5  0.02984 0.01060 Inf   2.802  0.1422
 sessionIndex4 - sessionIndex6  0.05954 0.01100 Inf   5.401  <.0001
 sessionIndex4 - sessionIndex7  0.08376 0.01390 Inf   6.012  <.0001
 sessionIndex4 - sessionIndex8  0.09247 0.01420 Inf   6.513  <.0001
 sessionIndex5 - sessionIndex6  0.02970 0.00987 Inf   3.008  0.0736
 sessionIndex5 - sessionIndex7  0.05392 0.01340 Inf   4.039  0.0015
 sessionIndex5 - sessionIndex8  0.06263 0.01610 Inf   3.890  0.0028
 sessionIndex6 - sessionIndex7  0.02422 0.01110 Inf   2.184  0.8106
 sessionIndex6 - sessionIndex8  0.03293 0.01420 Inf   2.322  0.5670
 sessionIndex7 - sessionIndex8  0.00871 0.01150 Inf   0.756  1.0000

Degrees-of-freedom method: asymptotic
P value adjustment: bonferroni method for 28 tests
```

Estimated marginal means is used to estimate the average value of a response variable at specific levels of a predictor variable, while accounting for other variables in the model. Here the EMM is the estimate of average of total typing time (log) at each session, while considering the individual differences captured by the random effects by the model lmer_model_log.

The output shows that the estimated mean for session 1 is 1.093 but it gradually decreases over the session. This shows that the typing time is not same as people progress through the session. The results show that the average log typing time decreases across sessions, starting at 1.093 in session 1 and dropping to 0.768 by session 8. Pairwise comparisons confirm significant differences between earlier sessions (e.g., session 1 - session 2 = 0.1203, $p < 0.0001$) but show smaller, insignificant differences between later sessions (e.g., session 7 - session 8 = 0.0087, $p = 1.0$). This indicates rapid initial improvement in typing speed that slows down over time.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

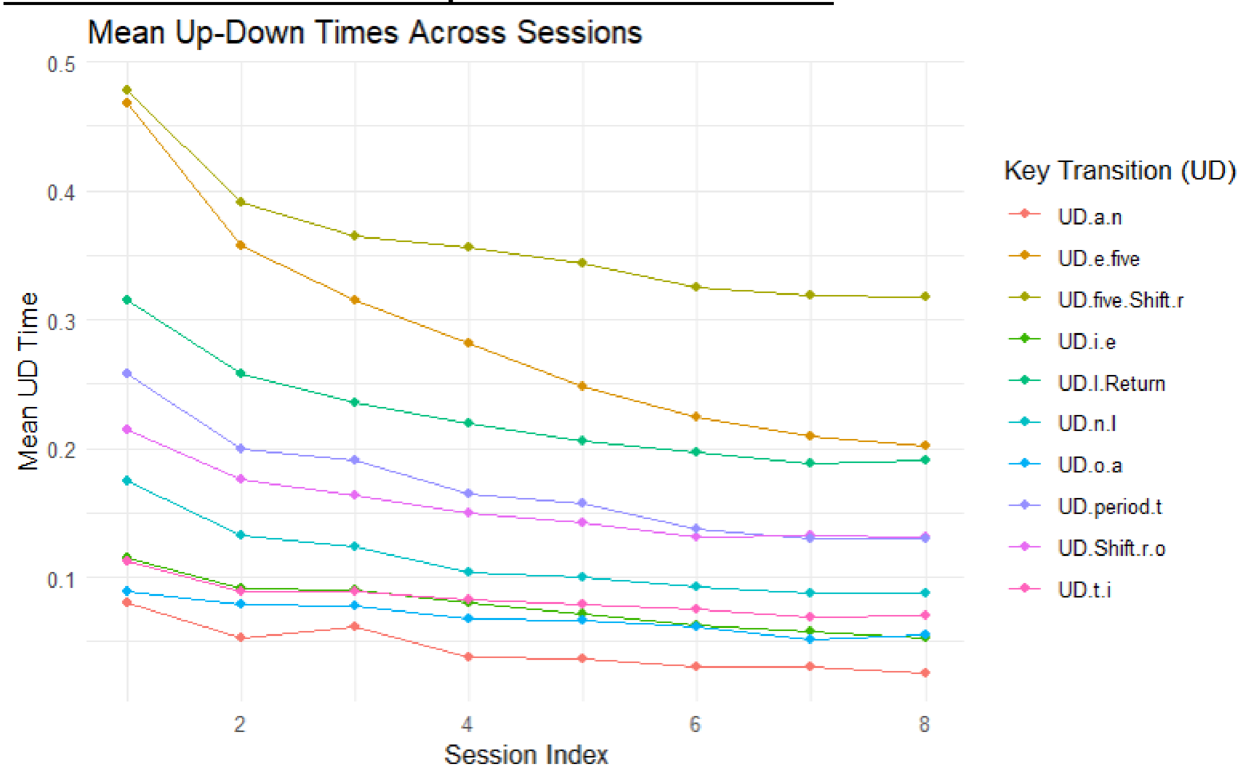
<u>Table 13 Significant Pairwise Difference for lmer_model_log</u>



The graph represents significant differences in typing behavior between each pair of session. The longest bar which is the comparison between session 1 and session 5 shows a substantial difference, suggesting significant improvement in typing speed from first to last session.

Mean Up Down Time as Response Variable

<u>Table 14 Distribution of Mean Up-Down Time across Session</u>



We tried to explore up-down time as the response variable as it represents the time that a typer would wait before pressing another key. The UD.five.Shift.r seems to be consuming the highest time out of all the up-down key time. This can be mainly due to shift key and number key which may be place significantly apart from the normal letter keys.

Table 15: Random Intercept Model for Sum_UD Values

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Sum_UD_Value ~ sessionIndex + (1 | subject)
   Data: ud_sum
Control: lmerControl(optimizer = "nloptwrap", optCtrl = list(maxfun = 1e+05))

REML criterion at convergence: 3671.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2165 -0.2891  0.0338  0.2721 12.1469

Random effects:
 Groups   Name        Variance Std.Dev.
 subject  (Intercept) 2037.9   45.14
 Residual              320.3   17.90
Number of obs: 408, groups:  subject, 51

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)    115.206      6.800  64.230  16.942  < 2e-16 ***
sessionIndex2  -23.783      3.544 350.000  -6.710 7.83e-11 ***
sessionIndex3  -29.653      3.544 350.000  -8.367 1.43e-15 ***
sessionIndex4  -38.054      3.544 350.000 -10.737  < 2e-16 ***
sessionIndex5  -42.668      3.544 350.000 -12.039  < 2e-16 ***
sessionIndex6  -48.262      3.544 350.000 -13.617  < 2e-16 ***
sessionIndex7  -51.512      3.544 350.000 -14.534  < 2e-16 ***
sessionIndex8  -52.045      3.544 350.000 -14.685  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) sssnI2 sssnI3 sssnI4 sssnI5 sssnI6 sssnI7
sessinIndx2 -0.261
sessinIndx3 -0.261  0.500
sessinIndx4 -0.261  0.500  0.500
sessinIndx5 -0.261  0.500  0.500  0.500
sessinIndx6 -0.261  0.500  0.500  0.500  0.500
sessinIndx7 -0.261  0.500  0.500  0.500  0.500  0.500
sessinIndx8 -0.261  0.500  0.500  0.500  0.500  0.500  0.500
```

The model provides statistical evidence for a significant decrease in Sum_UD_Value across multiple sessions. The results show that the average sum of UD times decreases across sessions, starting at 115.2 in session 1 and dropping to 63.2 by session 8. Pairwise comparisons confirm significant differences between earlier sessions (e.g., session 1 - session 2 = 23.783, $p < 0.0001$) but show smaller, insignificant differences between later sessions (e.g., session 7 - session 8 = 0.534, $p = 1.0$). This suggests that improvements in UD times are substantial in the initial sessions but taper off as participants become more consistent over time.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

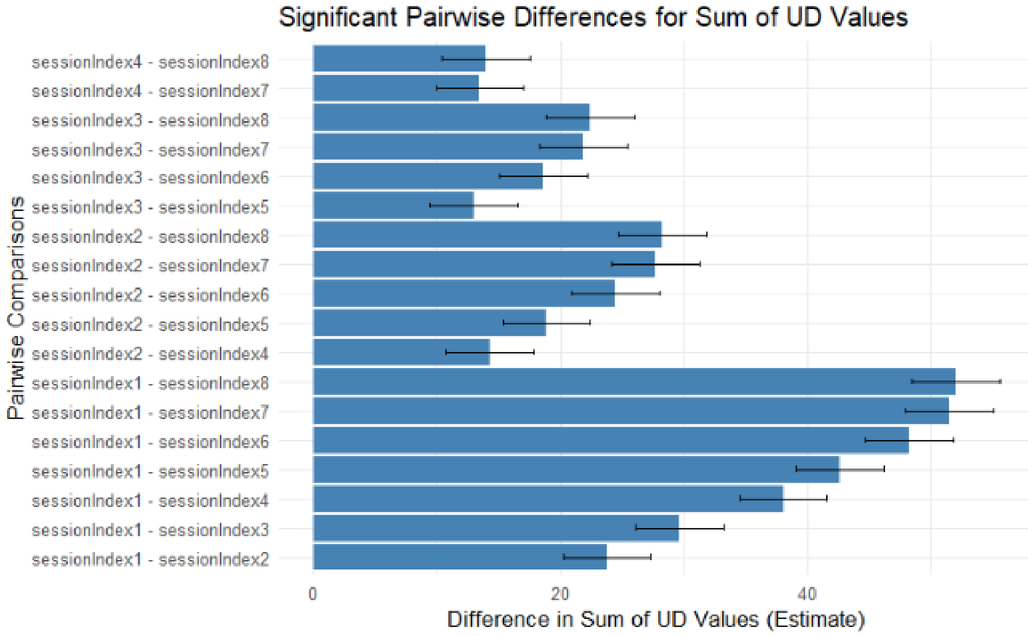
Table 16 Estimated Marginal Means (EMM) for Sum_UD Values

```
 sessionIndex emmean  SE   df lower.CL upper.CL
 1             115.2 6.8 64.2    101.6    128.8
 2              91.4 6.8 64.2     77.8    105.0
 3              85.6 6.8 64.2     72.0     99.1
 4              77.2 6.8 64.2     63.6     90.7
 5              72.5 6.8 64.2     59.0     86.1
 6              66.9 6.8 64.2     53.4     80.5
 7              63.7 6.8 64.2     50.1     77.3
 8              63.2 6.8 64.2     49.6     76.7

Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95
 contrast                    estimate   SE  df t.ratio p.value
 sessionIndex1 - sessionIndex2   23.783 3.54 350   6.710  <.0001
 sessionIndex1 - sessionIndex3   29.653 3.54 350   8.367  <.0001
 sessionIndex1 - sessionIndex4   38.054 3.54 350  10.737  <.0001
 sessionIndex1 - sessionIndex5   42.668 3.54 350  12.039  <.0001
 sessionIndex1 - sessionIndex6   48.262 3.54 350  13.617  <.0001
 sessionIndex1 - sessionIndex7   51.512 3.54 350  14.534  <.0001
 sessionIndex1 - sessionIndex8   52.045 3.54 350  14.685  <.0001
 sessionIndex2 - sessionIndex3    5.871 3.54 350   1.656  1.0000
 sessionIndex2 - sessionIndex4   14.272 3.54 350   4.027  0.0019
 sessionIndex2 - sessionIndex5   18.886 3.54 350   5.329  <.0001
 sessionIndex2 - sessionIndex6   24.480 3.54 350   6.907  <.0001
 sessionIndex2 - sessionIndex7   27.729 3.54 350   7.824  <.0001
 sessionIndex2 - sessionIndex8   28.263 3.54 350   7.974  <.0001
 sessionIndex3 - sessionIndex4    8.401 3.54 350   2.370  0.5126
 sessionIndex3 - sessionIndex5   13.015 3.54 350   3.672  0.0078
 sessionIndex3 - sessionIndex6   18.609 3.54 350   5.251  <.0001
 sessionIndex3 - sessionIndex7   21.858 3.54 350   6.167  <.0001
 sessionIndex3 - sessionIndex8   22.392 3.54 350   6.318  <.0001
 sessionIndex4 - sessionIndex5    4.614 3.54 350   1.302  1.0000
 sessionIndex4 - sessionIndex6   10.208 3.54 350   2.880  0.1181
 sessionIndex4 - sessionIndex7   13.457 3.54 350   3.797  0.0048
 sessionIndex4 - sessionIndex8   13.991 3.54 350   3.948  0.0027
 sessionIndex5 - sessionIndex6    5.594 3.54 350   1.578  1.0000
 sessionIndex5 - sessionIndex7    8.843 3.54 350   2.495  0.3653
 sessionIndex5 - sessionIndex8    9.377 3.54 350   2.646  0.2385
 sessionIndex6 - sessionIndex7    3.249 3.54 350   0.917  1.0000
 sessionIndex6 - sessionIndex8    3.783 3.54 350   1.067  1.0000
 sessionIndex7 - sessionIndex8    0.534 3.54 350   0.151  1.0000

Degrees-of-freedom method: kenward-roger
P value adjustment: bonferroni method for 28 tests
```

The results show that the average sum of UD times decreases across sessions, starting at 115.2 in session 1 and dropping to 63.2 by session 8. Pairwise comparisons confirm significant differences between earlier sessions (e.g., session 1 - session 2 = 23.783, $p < 0.0001$) but show smaller, insignificant differences between later sessions (e.g., session 7 - session 8 = 0.534, $p = 1.0$). This suggests that improvements in UD times are substantial in the initial sessions but taper off as participants become more consistent over time.

Group 5: Neha Karna, Prafulla Shrestha, Aidan Stewart, Josh Lefdal, Shivam Bhardwaj

Table 17 Significant Pairwise Differences for Sum_UD Values



The horizontal bars indicate the magnitude of the difference in the sum of UD values between the paired groups, with the direction of the difference also shown.

Conclusion

This analysis explored typing dynamics by investigating how two response variables, TotalTypingTime and ud_sum, change over multiple sessions. The results show significant reductions in both total typing time and UD times as participants completed repeated sessions, highlighting consistent improvement in typing speed and efficiency. Linear Mixed-Effects Models revealed that improvements vary among participants, with those starting slower showing greater gains.

References

Lecture Notes and Resources (STAT 541, STAT 600, STAT 601)

Killourhy, K. S., & Maxion, R. A. (2009). Comparing anomaly-detection algorithms for keystroke dynamics.

Hothorn, T., & Everitt, B. S. (2014). A Handbook of Statistical Analyses using R. In Chapman and Hall/CRC eBooks. https://doi.org/10.1201/b17081