Statistical Methods in Research

Analysis of How Stress Patterns Define Human Experience and Performance in Dexterous Tasks

May 3, 2018

Suchismitha Vedala

Lavanya Rao

Yashwanth Reddy Venati

Contents

Contributions

Suchismitha Vedala

1.Data Cleaning
2. Correlation.
3.Hypothesis Testing.
4.Box Plots
5.Report

Yashwanth Reddy Venati

1.Quality Control
2.Source Code Integration

Lavanya Rao

INTRODUCTION

The Microsurgery performance data represents the performance of 22 medical students in microsurgery activities. The 22 medical students or subjects in our analysis, participated in a longitudinal study regarding the relationship of sympathetic arousal and skill in learning micro-surgical tasks. The subjects had to pay five visits which we regard as sessions, lasting one hour each, in order to practice micro-surgical cutting and suturing in an inanimate simulator.A pre and post study questionnaire was also given to be completed by the subjects to know a little about their biography and anxiety.

During the main part of each session, the subjects underwent the following treatments:
1.Baseline: The subjects were relaxing for 5 min, listening to spa music. They were facially recorded by a thermal and visual camera.
2.Cutting: The subjects had to precision cutting in the inanimate simulator. They were facially recorded by a thermal and visual camera.
3.Suturing: The subjects had to perform suturing in the inanimate simulator. They were facially recorded by a thermal and visual camera.

Explicit accuracy scores per task is provided in the data. Hence, the cutting task has its own accuracy scores and so is the case with the suturing task. The perspiration values are recorded in all time frames for all subjects and sessions. The subjects were asked to fill out a NASA-TLX questionnaire after each task. he NASA-TLX instrument features five subscales measuring different aspects of the subjects' perceptions regarding task difficulty.We perform an analysis with this given data.

MOTIVATION

Given the Microsurgery Performance Data, We are interested in finding any new observations that can effect the performance. Statistical Tests and Analysis help in exploring the unexplored data and making assumptions and conclusions that can lead to better performance. The analysis can always lead to performing better experiments in the future.

DATA ORGANIZATION

The data was structured into main 3 main files.

1. MicroSurgeryPerformance file :

This file indicates the Age, Sex, Year, Time Taken for Cutting and Suturing in each session, Scores of Cutting and Suturing of both the Scorers in Each Session, Number of Sutures made in each session for each subject.

2. MasterFileMethodistSurgery:

This file indicates the presence and absence of the files of Baseline Perinasal Perspiration(PP), Cutting PP, Suturing PP, Cutting NASA,Suturing NASA, Mbg, and MPOST for each subject and each session.

3. TaiScore.txt:

This file indicates the tai scores of all subjects.

Along with this, there is a folder for each subject.Each Subject folder consists of :

1. Subject_tai.csv : which is a behavior analysis questionnaire.
2. Mbg and Mpost csv files which contain geographic and personal information of the subject before and after the tasks.
3. tp.csv, which gives pre and post Session values.
4.A folder. for each session.

Each Session Folder consists of five files:

1.Baseline.csv: Consists of the values of the Perspiration when the subjects were relaxed. This data is collected over every frame in every second for 5 minutes

2.Cutting.csv: Consists of the values of the Perspiration when the subjects were performing cutting task. This data is collected over every frame during the time taken to complete the task

3.Cutting_NASA.csv: Consists of Responses to the NASA TLX questionnaire.

4.Suturing.csv: Consists of the values of the Perspiration when the subjects were performing Suturing task. This data is collected over every frame during the time taken to complete the task

5.Suturing_NASA.csv: Consist of Responses to the NASA TLX questionnaire.

DATA CLEANING

To perform any analysis , we first need to clean the data to removed any redundant data and reshape the data to help us perform any statistical tests.

We perform our data cleaning on the MicroSurgeryPerformance.csv file.

1.For each row, where Sex value is 1 we rename to Male and 2 to Female.

2.We melt the data frame to obtain the Score values pertaining to each subject and each session.

3.We mention the Task to which the Scores belong.

4.With two Scorers being present, we extend the rows of the data to append Score values of the Scorer2.

5. For each subject, for each session and task, we take the Mean_Perspiration(mean persinal perspiration)value as Mean_(task)Perspiration-Mean(Baseline)Perspiration.

This is done because, to get the stress signal caused only by the task can be obtained by stress signal obtained during task - stress signal during relaxed.

6.We now normalize the Mean_Perspiration values, by taking the absolute minimum value(excluding NA) and adding to each value along with an error coefficient. We add this as Normalised_PP

7.The time taken for each task per session is totally converted into seconds .

The data now has the following format:

Age	Year	Sex	Session	Scorer	Task	Scores	Mean_Perspiration	Normalised_PP	Time	Subjects	Tai
22	1	Male	Session1	Scorer1	Cutting	18	-0.0000159794	0.007539787	595	S1	37
22	1	Male	Session2	Scorer1	Cutting	22	0.0011471670	0.008702933	359	S1	37
22	1	Male	Session3	Scorer1	Cutting	23	-0.0008589810	0.006696785	487	S1	37
22	1	Male	Session4	Scorer1	Cutting	22	NA	NA	153	S1	37
22	1	Male	Session5	Scorer1	Cutting	22	NA	NA	245	S1	37
23	1	Male	Session1	Scorer1	Cutting	17	-0.0059216820	0.001634084	405	S2	37

Figure 1: Data Cleaned for Analysis

INITIAL ANALYSIS

Correlation Matrix

With the modified data, we not try to understand the correlation between the values. Highly correlated values can be redundant and we can eliminate that in our classification. When taken correlation between all numeric columns of data , we get the following result, where
P-stands for the probable value
n stands for number of observations
r is the correlation value which can range from -1 to 1.
From the data we understand that Age and Year are positively correlated with a r value of 0.83 stating that, with increase in age , year increases.There is no association between sessions and Age, Year and Tai scores. All other attributes are neither highly positively or negatively correlated, i.e each correlation value is neither near to -1 or +1 which gives us a huge scope to perform analysis and interpret results

	Age	Year	Scores	Normalised_PP	Time	Tai	Sessions
Age	1.00	0.83	0.17	0.14	0.01	0.21	0.00
Year	0.83	1.00	0.20	0.11	0.01	0.11	0.00
Scores	0.17	0.20	1.00	0.17	-0.25	-0.02	0.55
Normalised_PP	0.14	0.11	0.17	1.00	0.04	0.11	0.26
Time	0.01	0.01	-0.25	0.04	1.00	0.05	-0.13
Tai	0.21	0.11	-0.02	0.11	0.05	1.00	0.00
Sessions	0.00	0.00	0.55	0.26	-0.13	0.00	1.00

n

	Age	Year	Scores	Normalised_PP	Time	Tai	Sessions
Age	300	300	300	272	300	300	300
Year	300	300	300	272	300	300	300
Scores	300	300	300	272	300	300	300
Normalised_PP	272	272	272	272	272	272	272
Time	300	300	300	272	300	300	300
Tai	300	300	300	272	300	300	300
Sessions	300	300	300	272	300	300	300

P

	Age	Year	Scores	Normalised_PP	Time	Tai	Sessions
Age		0.0000	0.0029	0.0188	0.8849	0.0003	1.0000
Year	0.0000		0.0004	0.0624	0.9214	0.0652	1.0000
Scores	0.0029	0.0004		0.0050	0.0000	0.7491	0.0000
Normalised_PP	0.0188	0.0624	0.0050		0.5141	0.0787	0.0000
Time	0.8849	0.9214	0.0000	0.5141		0.3586	0.0220
Tai	0.0003	0.0652	0.7491	0.0787	0.3586		1.0000
Sessions	1.0000	1.0000	0.0000	0.0000	0.0220	1.0000	

Figure 2: Correlation Matrix

Analysis of Mean Perspiration Based on Subject

We bar plot the mean perspiration (stress signal)values of all the subjects and of each task.We observe that subjects experienced more stress during the Suturing Task and Subject7 has displayed high level of stress.

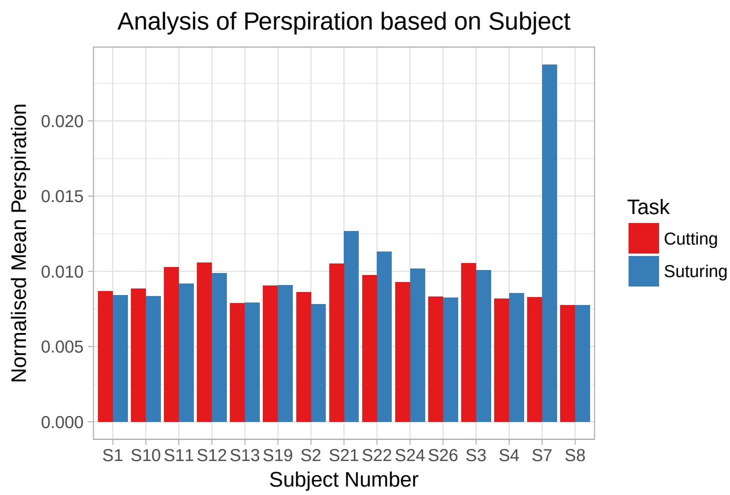

Figure 3: Accuracy of Mean Perspiration of all subjects

QUALITY CONTROL

1.Biographic Data

We draw a bar plot to see how gender defines data and histogram to see whether age has any effect on the data.

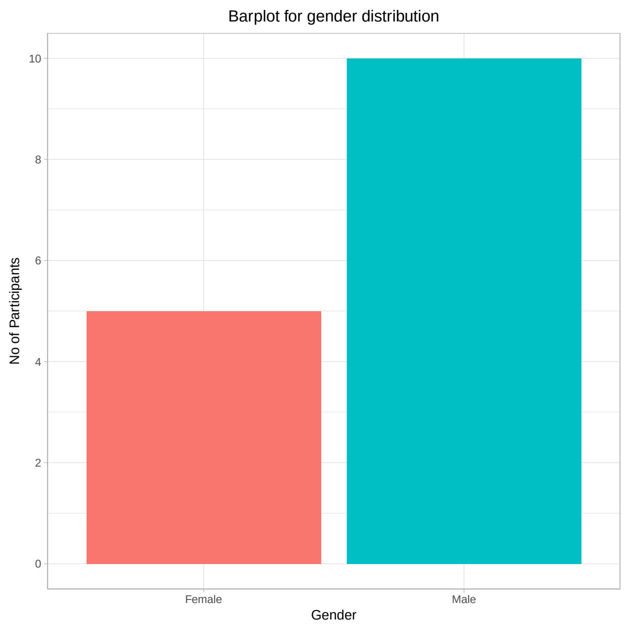

Figure 4: Barplot of Gender Distribution

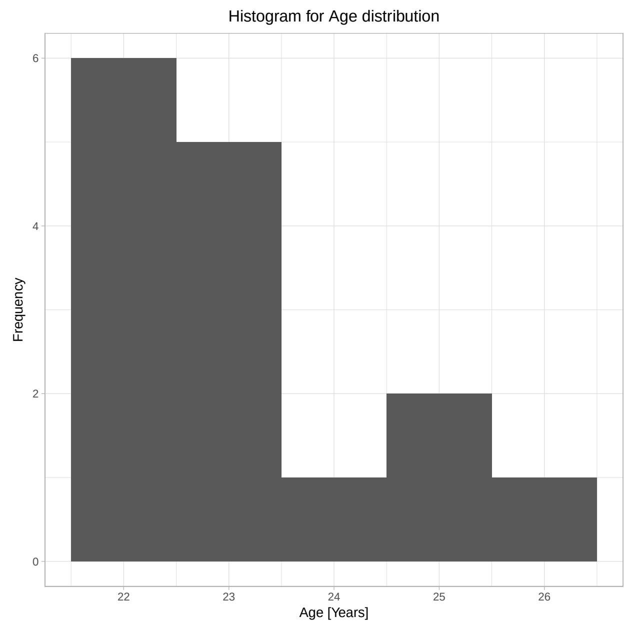

Figure 5: Histogram of age distribution

2.Trait Psychometric Data

We draw the histogram for Trait Anxiety Inventory(TAI) scores

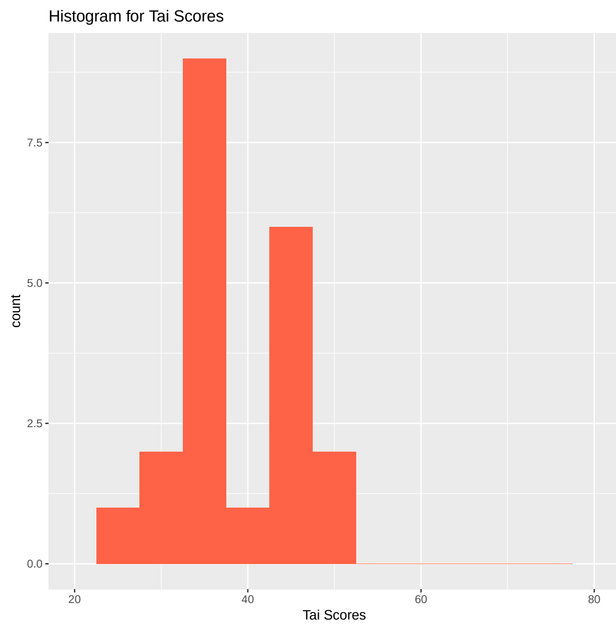

Figure 6: Histogram of Tai Scores

3.State Psychometric Data

For each subject draw the bar plots for all the NASA-TLX subscales per task. This will give two figures per subject per subscale, one for suturing and one for cutting, where the evolution of the scores from the initial to the final session will be evident.

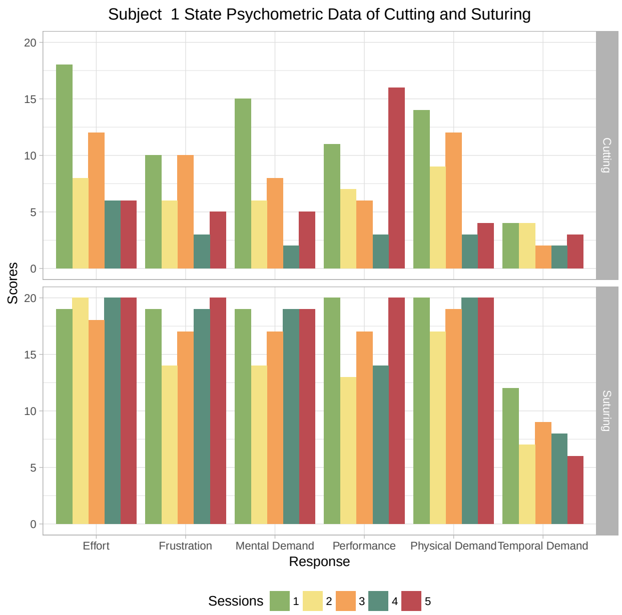

Figure 7: Subject 1

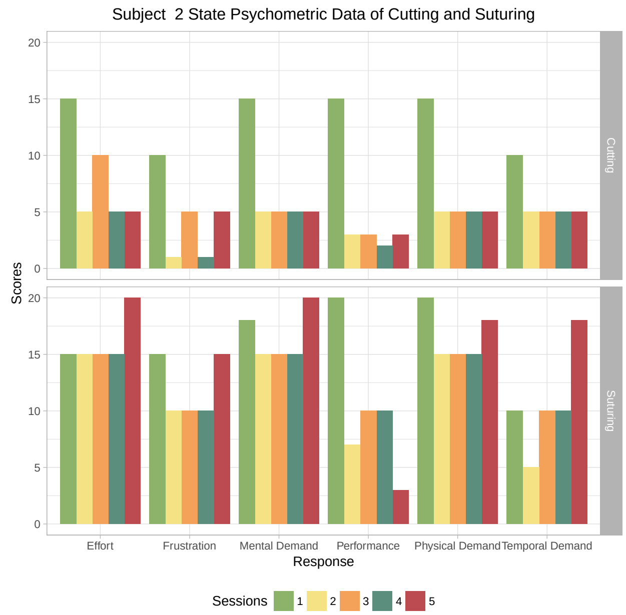

Figure 8: Subject 2

4.Perinasal Perspiration (Stress) Signal Data

For each session of each subject we draw the perspiration values using black for baseline, green for cutting, and red for suturing.

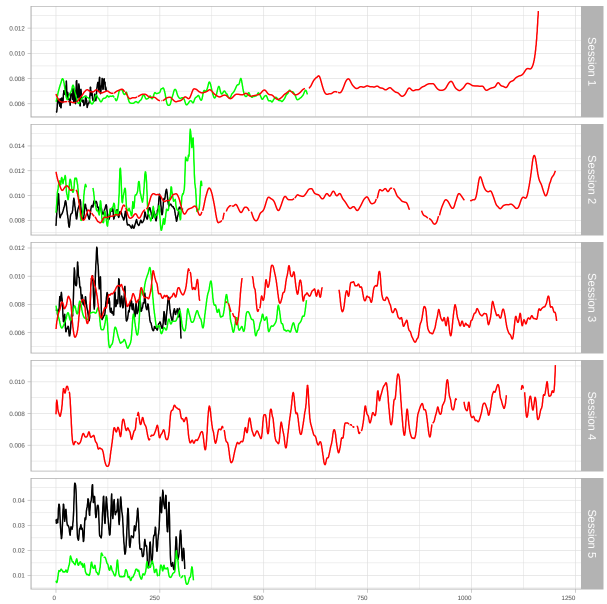

Figure 9: Subject 1

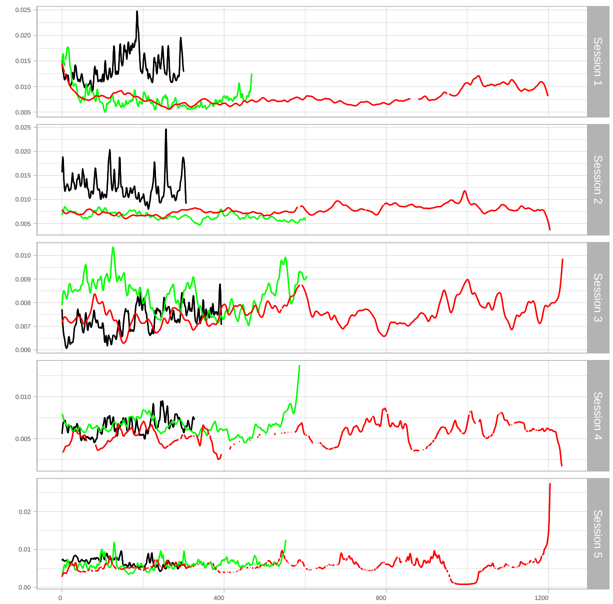

Figure 10: Subject 2

5.Performance Data

We draw the accuracy and time bar plots of each subject for each session and each task.

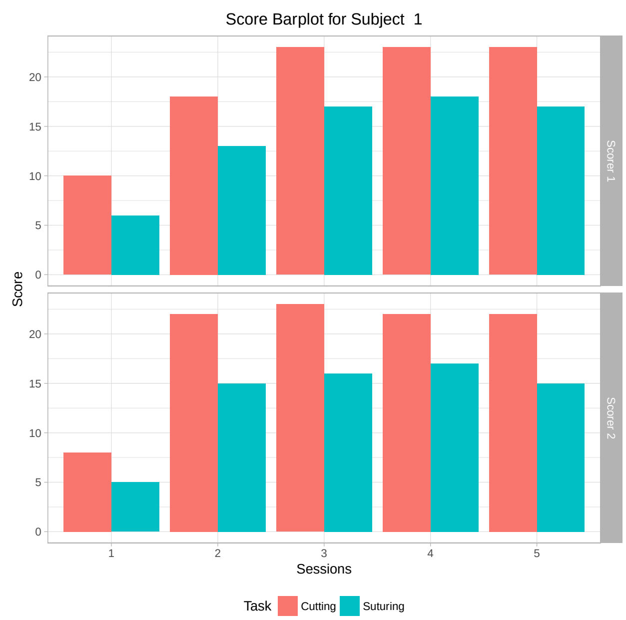

Figure 11: Subject 1 Score Barplot

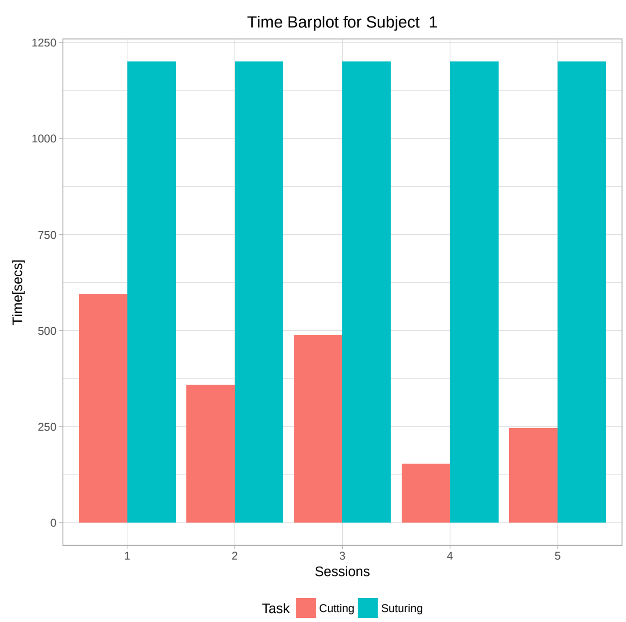

Figure 12: Subject 1 Time Barplot

Hypothesis Testing

We now perform various tests to determine significance of any factor on the other and make inferences about the performance of the subjects in tasks.

1. Analysis of effect of each attribute on Score

Hypothesis:
Null Hypothesis : H_0 = The score obtained does not depend on the demographics of the subject , session , age , year , sex and perspiration.
Alternate Hypothesis : H_1 = The score obtained depends on the demographics of the subject , session , age , year , sex and perspiration.

Approach:Linear Modelling:
Linear modeling gives the relationship between the dependent and independent variables. In our data set we are finding the hypothesis between each attribute such as Age, sex, year and mean perspiration with the scores of scorer.
When we performed the linear model and analyzed the results, we observe that the pvalue is way less than 0.05, thus rejecting the null hypothesis. We also observe that the Score is highly Significant on Sex and Sessions.

Inference:
From the plots and linear model, we infer that the score is highly dependent on Sex, Session number and nearly significant on the Task.

```
Call:
lm(formula = Scores ~ log(Normalised_PP) + Age + Sex + Task +
    Scorer + Session, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-9.0904 -2.0812  0.1778  2.4703  8.4800

Coefficients:
                   Estimate Std. Error t value            Pr(>|t|)
(Intercept)         0.43830    5.71336   0.077             0.93891
log(Normalised_PP) -1.13961    0.74582  -1.528             0.12772
Age                 0.49486    0.17195   2.878             0.00433 **
SexMale            -2.17630    0.47007  -4.630       0.00000576827 ***
TaskSuturing       -0.97517    0.42795  -2.279             0.02349 *
ScorerScorer2       0.02941    0.42648   0.069             0.94507
SessionSession2     4.26300    0.69164   6.164       0.00000000267 ***
SessionSession3     6.38075    0.69130   9.230 < 0.0000000000000002 ***
SessionSession4     7.68358    0.70093  10.962 < 0.0000000000000002 ***
SessionSession5     8.17553    0.71138  11.492 < 0.0000000000000002 ***

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
---
Residual standard error: 3.517 on 262 degrees of freedom
  (28 observations deleted due to missingness)
Multiple R-squared:  0.4606,    Adjusted R-squared:  0.442
F-statistic: 24.86 on 9 and 262 DF,  p-value: < 0.00000000000000022
```

Figure 13: Linear model of score vs all other attributes

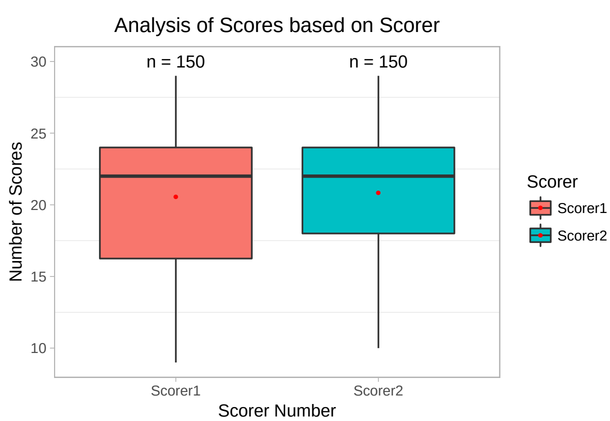

Figure 14: Analysis of Scores based on Score

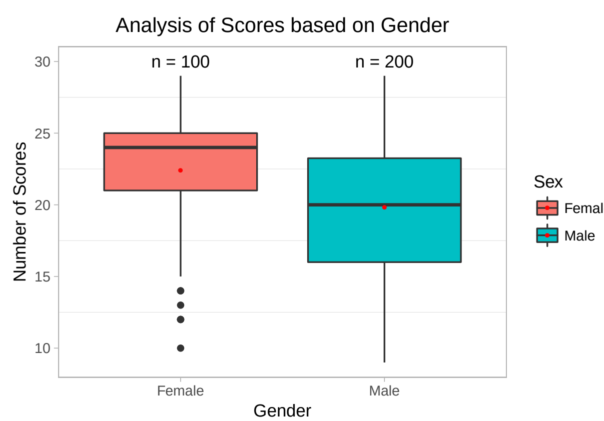

Figure 15: Analysis of Scores based on Gender

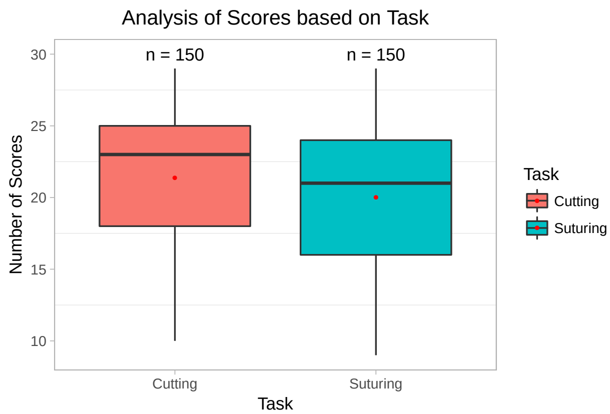

Figure 16: Analysis of Scores based on Task

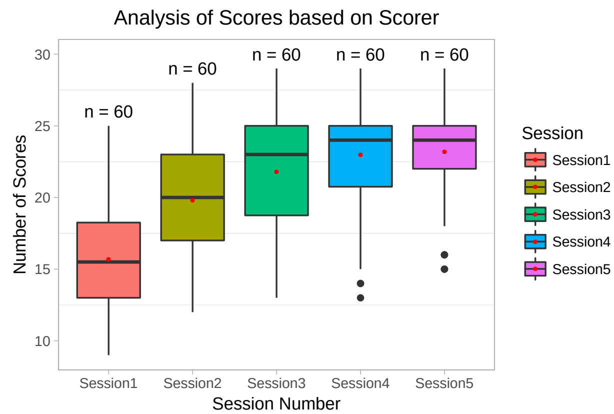

Figure 17: Analysis of Scores based on Session

2. Analysis of effect of each attribute on Time

Hypothesis:
Null Hypothesis : H_0= The time taken to do a task is not dependent of age, session, sex, Perspiration.
Alternate Hypothesis : H_1 = he time taken to do a task depends on age, session, sex, Perspiration.

Approach:Linear Modelling:
Linear modeling gives the relationship between the dependent and independent variables. In our data set we are finding the hypothesis between each attribute such as Age, sex,Scorer,Task,Session and mean perspiration with the time taken to do the task.

When we performed the linear model and analyzed the results, we observe that the pvalue is way less than 0.05, thus rejecting the null hypothesis. We also observe that the Time taken is highly Significant on Task and Sessions.

Inference:
We observe that the means of Time vs Scorer or Gender do not have highly significant difference. Although, Task does have significant difference, we do not take it as a notifiable inference, because the time limits are different for each tasks, with time to complete suturing task being. 20 minutes. Hence, From the summary and the plots we infer that the time taken highly depends on the sessions.

```
Call:
lm(formula = Time ~ log(Normalised_PP) + Age + Sex + Task + Session,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-340.68  -66.10    4.40   78.19  302.39

Coefficients:
                   Estimate Std. Error t value             Pr(>|t|)
(Intercept)          90.273    189.620   0.476                0.634
log(Normalised_PP)  -15.595     24.770  -0.630                0.530
Age                   7.915      5.711   1.386                0.167
SexMale              18.922     15.612   1.212                0.227
TaskSuturing        872.331     14.213  61.375 < 0.0000000000000002 ***
SessionSession2    -105.645     22.971  -4.599   0.0000065992793776 ***
SessionSession3    -113.626     22.959  -4.949   0.0000013335710345 ***
SessionSession4    -185.969     23.279  -7.989   0.0000000000000429 ***
SessionSession5    -170.059     23.626  -7.198   0.0000000000064042 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 116.8 on 263 degrees of freedom
  (28 observations deleted due to missingness)
Multiple R-squared:  0.9365,    Adjusted R-squared:  0.9345
F-statistic: 484.7 on 8 and 263 DF,  p-value: < 0.00000000000000022
```

Figure 18: Linear model of Time vs all other attributes

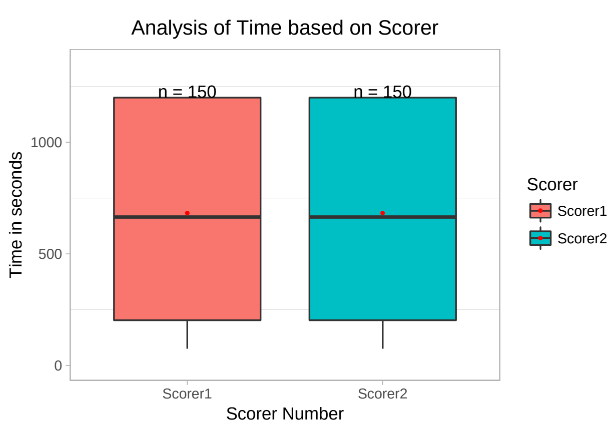

Figure 19: Analysis of Time based on Score

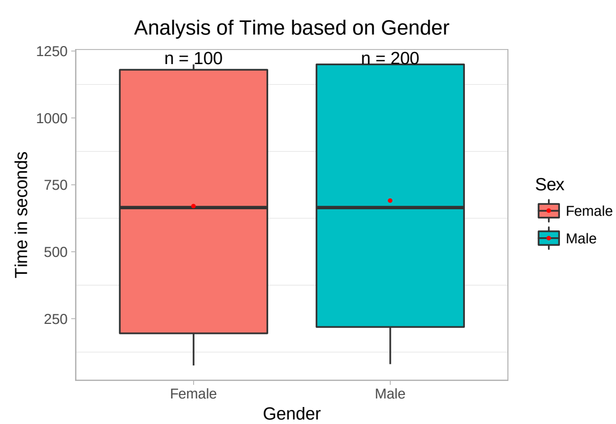

Figure 20: Analysis of Time based on Gender

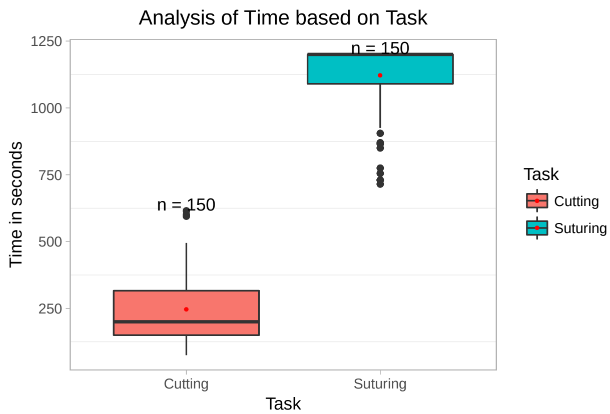

Figure 21: Analysis of Time based on Task

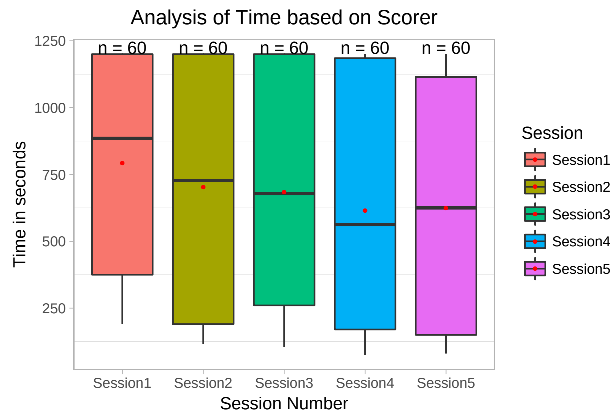

Figure 22: Analysis of Time based on Session

3. Performance Analysis of Suturing Task

We first perform a correlation between number of sutures made and subject, age, time taken scorer, scores and session. We observe that time and number of sutures made are moderately negatively correlated, which indicates that the subjects are performing better. Also, number of sutures made is highly correlated with the scores. The session is moderately positively correlated with Sutures, indicating a positive relationship.

Approach:Linear Model:

From the figures, we understand that the number of sutures increases with each session.
To summarize it over all the subjects, we draw the bar plot with values of all subjects summed under each session , which gives us an understanding that with each session, the performance increases across all subjects.
When creating the linear model, and performing the analysis, we observe that the number of sutures is highly dependent on sessions . Also, from the Performance analysis with respect to time plot, we infer that as the time decreases, the subjects are able to make more number of sutures and as number of session increases, the performance gets better.
Inference:
The number of sutures made increases with each session and takes less time, i.e the subjects are performing well.

```
        Subject   Age session Sutures Scores  Time
Subject    1.00  0.22    0.00    0.32   0.23 -0.08
Age        0.22  1.00    0.00    0.33   0.21 -0.06
session    0.00  0.00    1.00    0.62   0.59 -0.41
Sutures    0.32  0.33    0.62    1.00   0.82 -0.53
Scores     0.23  0.21    0.59    0.82   1.00 -0.56
Time      -0.08 -0.06   -0.41   -0.53  -0.56  1.00

n= 150


P
        Subject Age    session Sutures Scores Time
Subject         0.0068 1.0000  0.0000  0.0044 0.3289
Age     0.0068         1.0000  0.0000  0.0102 0.4472
session 1.0000  1.0000         0.0000  0.0000 0.0000
Sutures 0.0000  0.0000 0.0000          0.0000 0.0000
Scores  0.0044  0.0102 0.0000  0.0000         0.0000
Time    0.3289  0.4472 0.0000  0.0000  0.0000
```

Figure 23: Correlation Matrix of Sutures Data

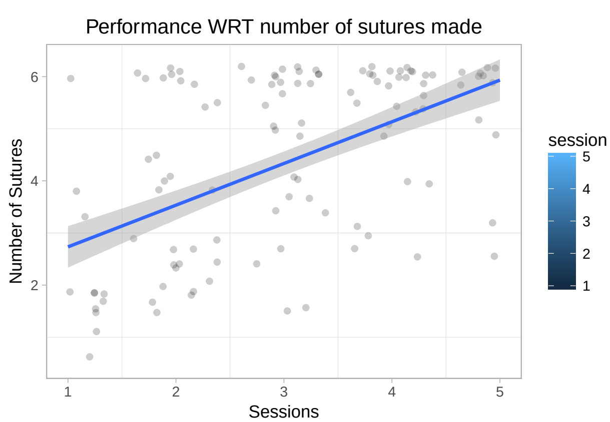

Figure 24: Performance WRT number of sutures made

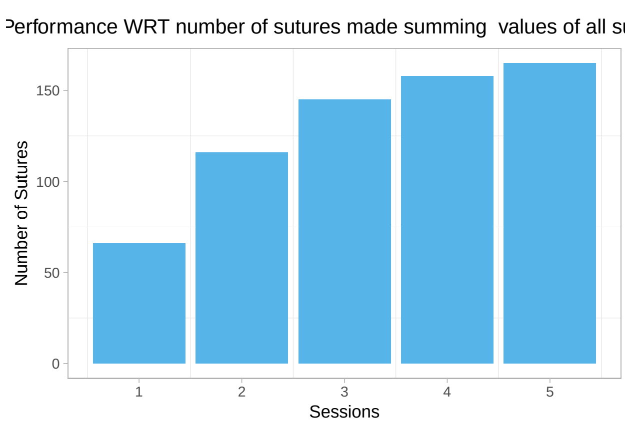

Figure 25: Performance WRT number of sutures summed across all subjects

```
Call:
lm(formula = Sutures ~ session + Scores + Sex + Age + Time, data = sdata)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.6218 -0.5798 -0.1343  0.5793  2.6515 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -5.6501540  1.6934588  -3.336  0.00108 ** 
session      0.3461851  0.0684093   5.060 1.26e-06 ***
Scores       0.2091886  0.0236815   8.833 3.19e-15 ***
Sex2         0.4381641  0.1735438   2.525  0.01266 *  
Age          0.2554761  0.0632943   4.036 8.79e-05 ***
Time        -0.0011590  0.0006741  -1.719  0.08771 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9234 on 144 degrees of freedom
Multiple R-squared:  0.7506,	Adjusted R-squared:  0.742 
F-statistic: 86.69 on 5 and 144 DF,  p-value: < 2.2e-16
```

Figure 26: Analysis of Linear Model

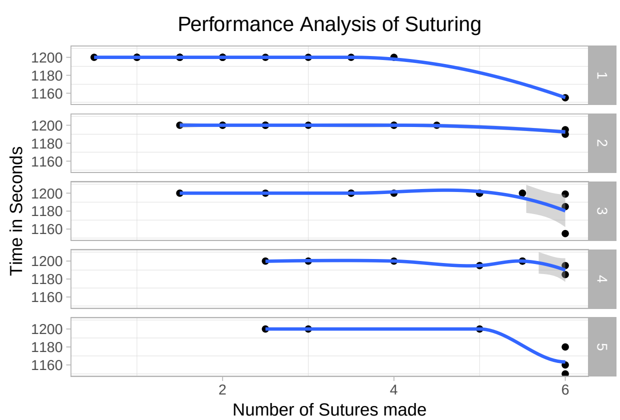

Figure 27: Performance Analysis of Suturing overTime

4. Analysis of Scorers on Task:

Hypothesis:

$NullHypothesis: H_0 =$ The mean of scores is same for both the Scorers

$AlternateHypothesis: H_1 =$ The mean of scores is different for both the Scorers

Assumptions:

We perform shapiro.test to determine whether the Scores of Cutting and Suturing are normalized or not.We observe that the p-value is less than 0.05 in both the cases, which tells us that the columns are not normalized . Hence we perform Wilcoxon Test.

Approach:Wilcox Test:

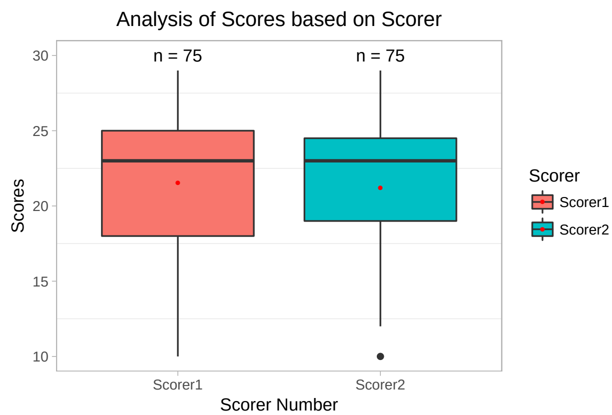

Figure 28: Cutting Scores

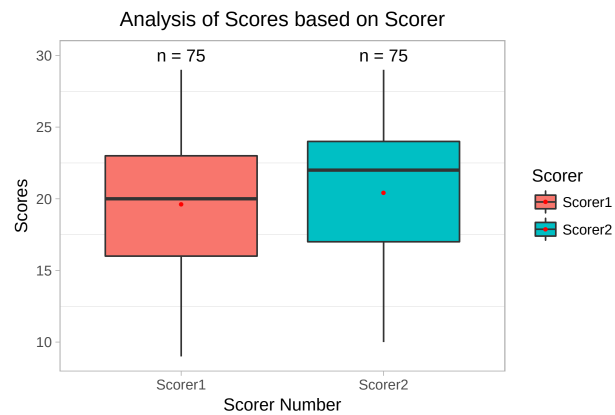

Figure 29: Suturing Scores

Cutting : When performed Wilcox test, p-value is 0.1676 which is greater than 0.05, which applies the means have not changed

Suturing: When performed Wilcox test, p-value is 0.00716, which less than 0.05, which states that the means of the scorers is different.

Inference:

Scorer has impact on Suturing Task but not on Cutting Task,

5. Analysis of Number of Sutures made with respect to sex and sessions:

Hypothesis:

Null Hypothesis : H_0= The number of sutures made is not significant on sex of the subject

Alternate Hypothesis : H_1 = The number of sutures made is significant on sex of the subject

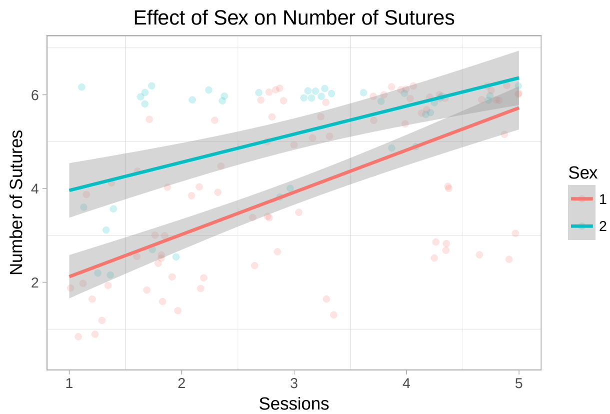

Figure 30: Interaction plot of Number of Sutures made wrt to Sex and Session

When performing the analysis for effect of Sutures on interaction of Sex and Sessions, we observe that, the p-value is 6.44e-11 which is way less than 0.05.
Also, from the interaction plot, we understand that the number of sutures made increases with the number of sessions and females make more number of sutures than males in average.

Inference:
Sex has an effect on Number of Sutures made and females made more number of Sutures in 20 minutes and the number os sutures increases with each session, i.e subject's performance gets better with each session.

6. Analysis of Performance With Respect to TAI

Assumptions:
We perform shapiro.test to determine whether the Scores Below and Above Tai are normalized or not.We observe that the p-value is less than 0.05 in both the cases, which tells us that the columns are not normalized . Hence we perform Wilcoxon Test.
Hypothesis:
Null Hypothesis : H_0 = The performance score does not depend on tai score
Alternate Hypothesis : H_1 = With increase in tai score, the performance score decreases

Approach:Wilxoxon Test:

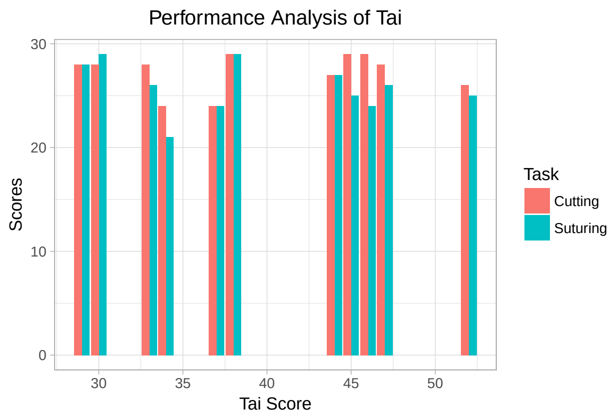

Figure 31: Analysis of Tai Vs Scores

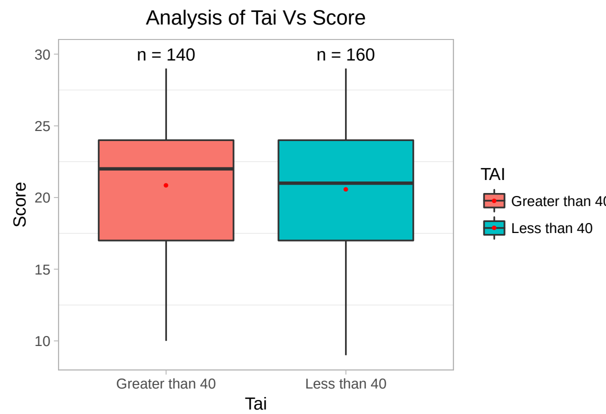

Figure 32: Analysis of Tai Vs Scores

Given, In the data, that Tai score above 40 means the subject is highly anxious. When performed T test, between all the values below or equal to 40 and above 40, we observe that the p-value is greater than 0.05, which indicates we accept the null hypothesis.
In addition, from the plots we can infer that the scores are pretty much the same, for both, above and below 40 Tai

Inference:
With this, we infer that the in general, performance does not depend on how anxious the subject is and the tai score.

Discussion

Conclusion

Appendix

List of Figures

1. State Psychometric Data of all Subjects

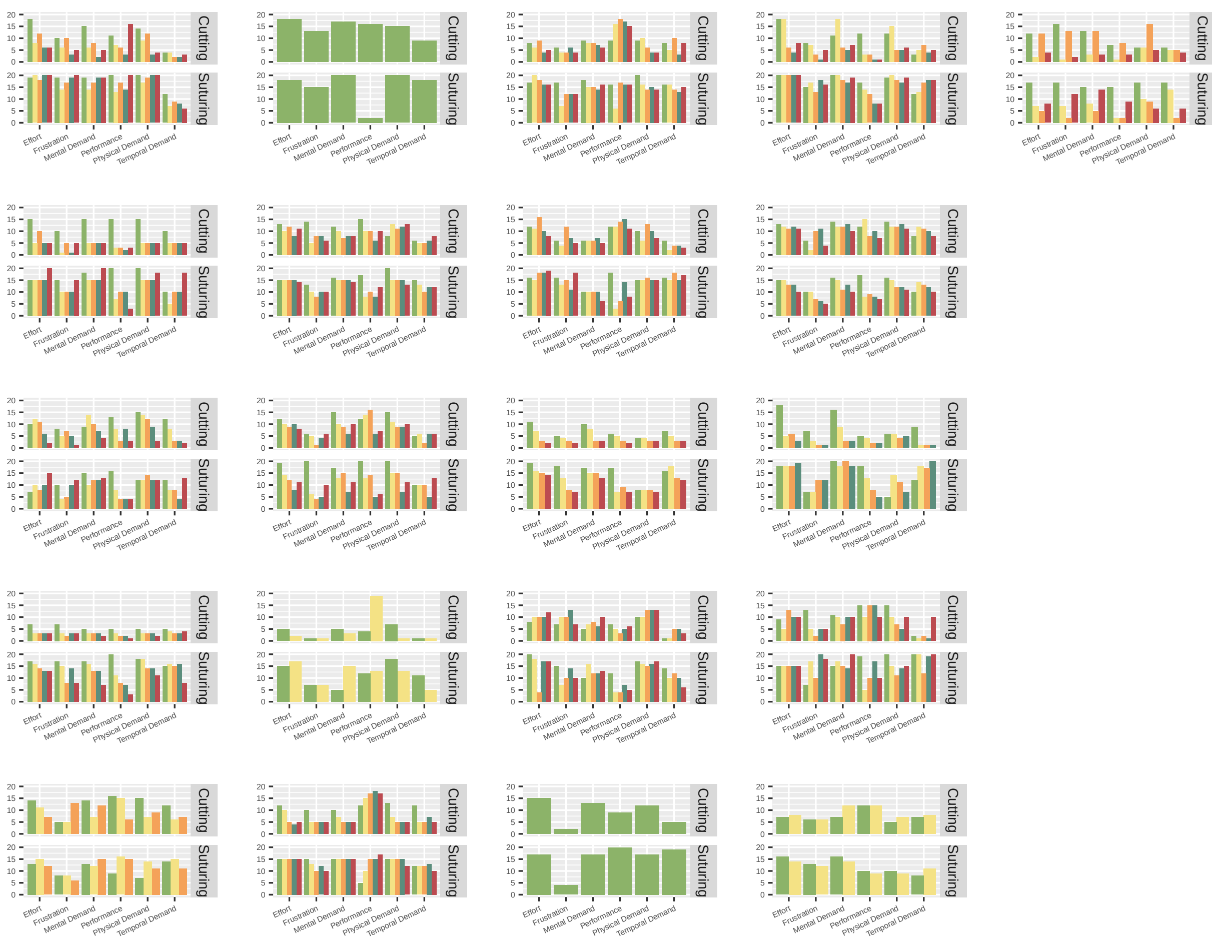

Figure 33: State Psychometric Data of all Subjects

2. Perinasal Perspiration (Stress) Signal Data of all subjects

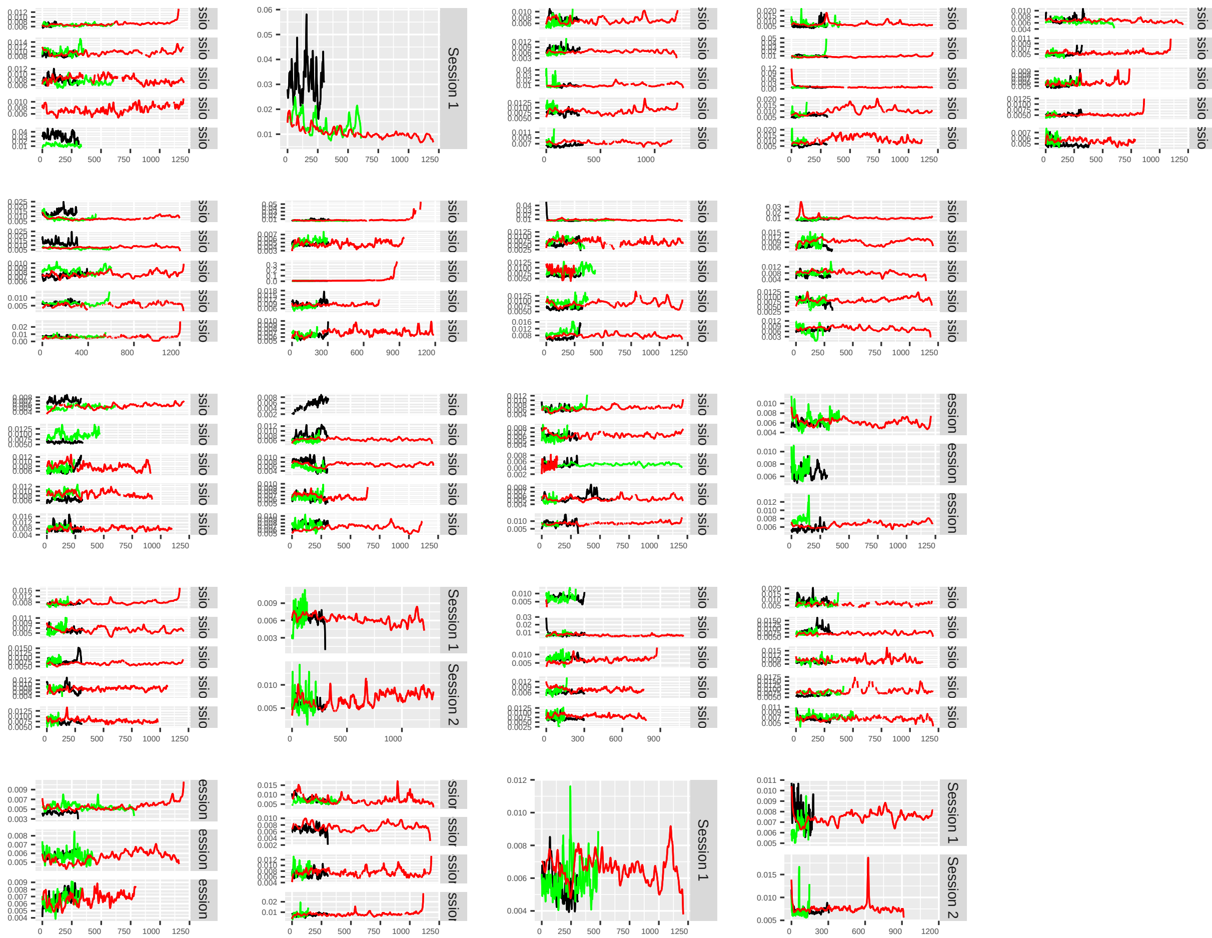

Figure 34: Perinasal Perspiration (Stress) Signal Data of all subjects

3. Suturing Time Performance of all Subjects

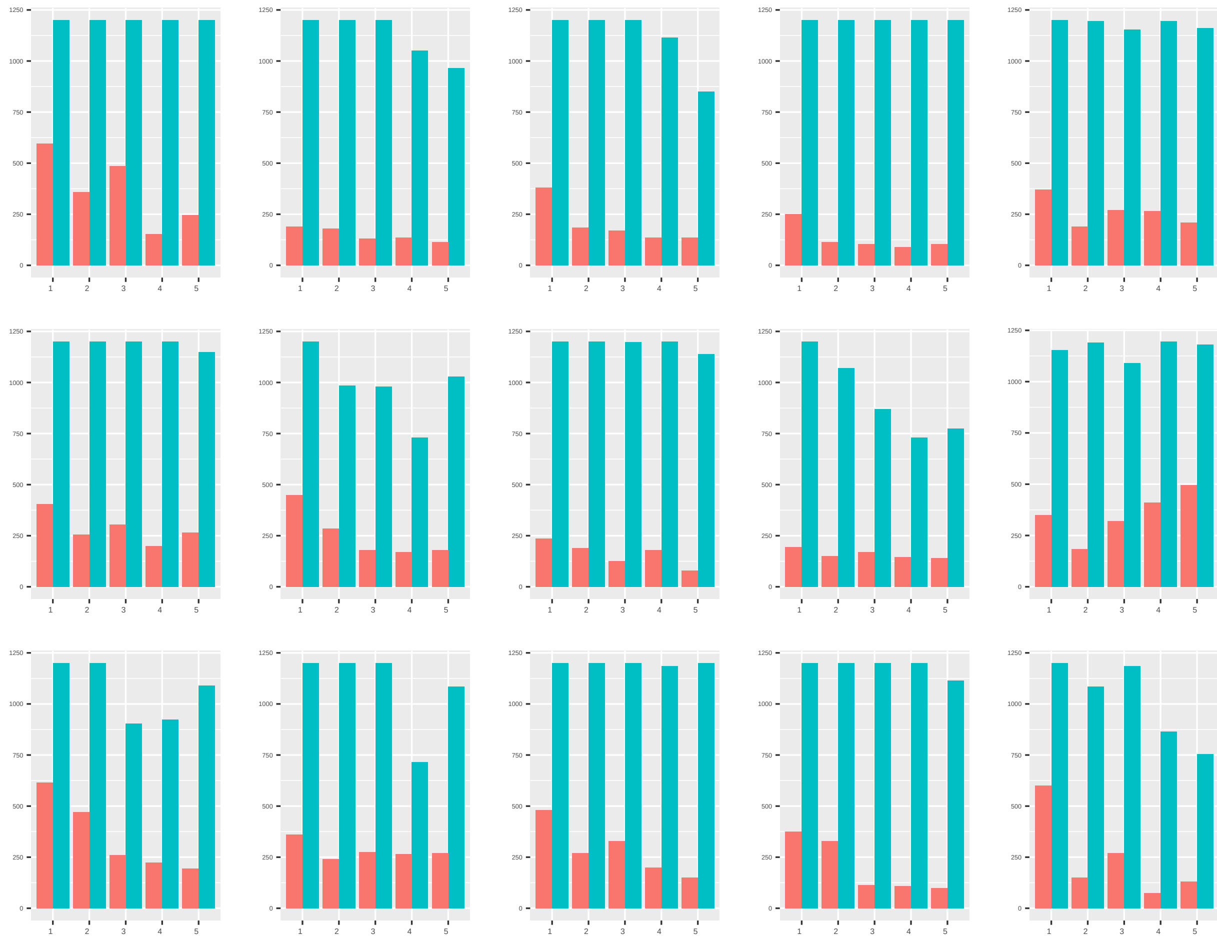

Figure 35: Suturing Time Performance of all Subjects

4.Accuracy of Scores of all subjects

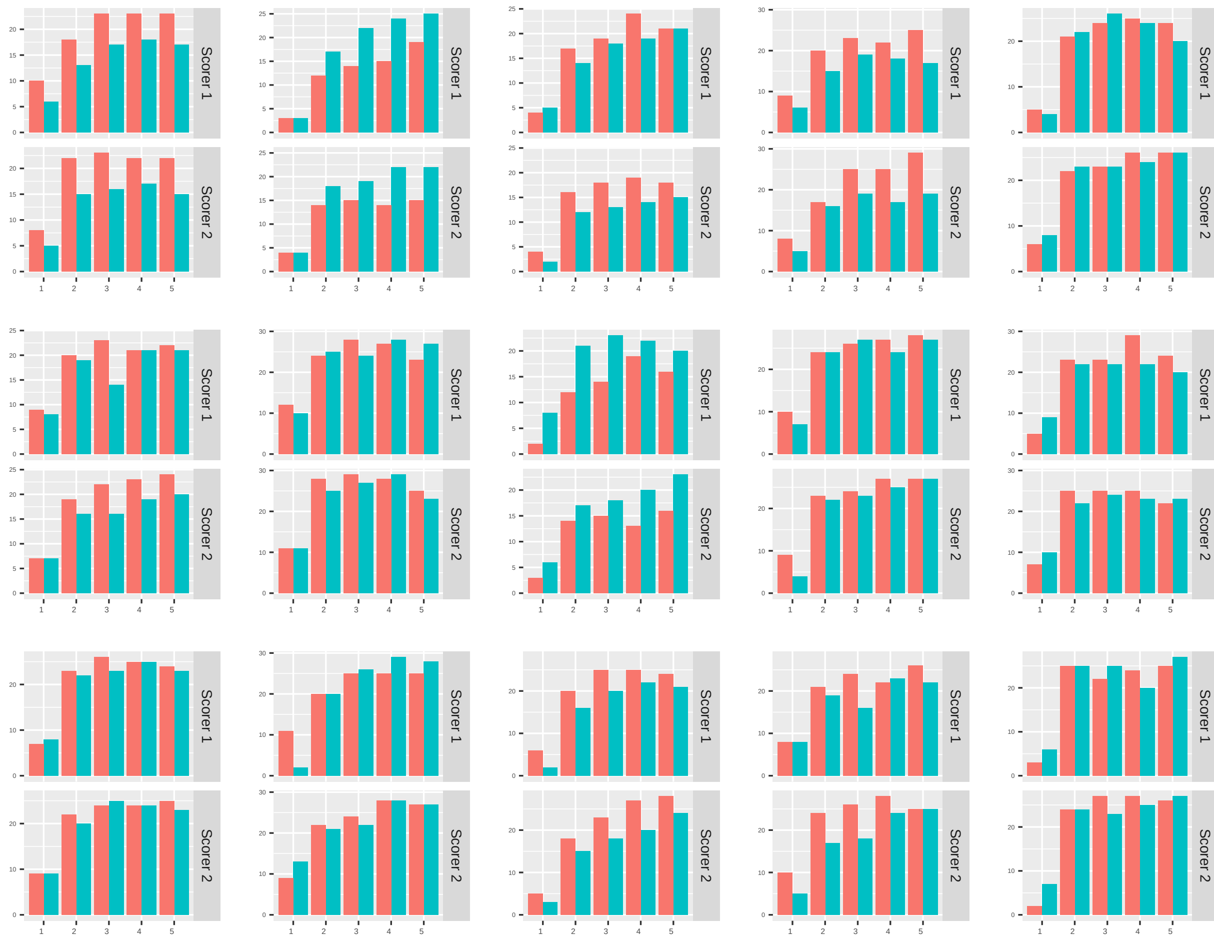

Figure 36: Accuracy of Scores of all subjects

5.Correlation Matrix Visualization of Plots of Data based on Task

Figure 37: Accuracy of Scores of all subjects

6.Correlation Matrix Visualization of Plots of Cutting Data based on scorer

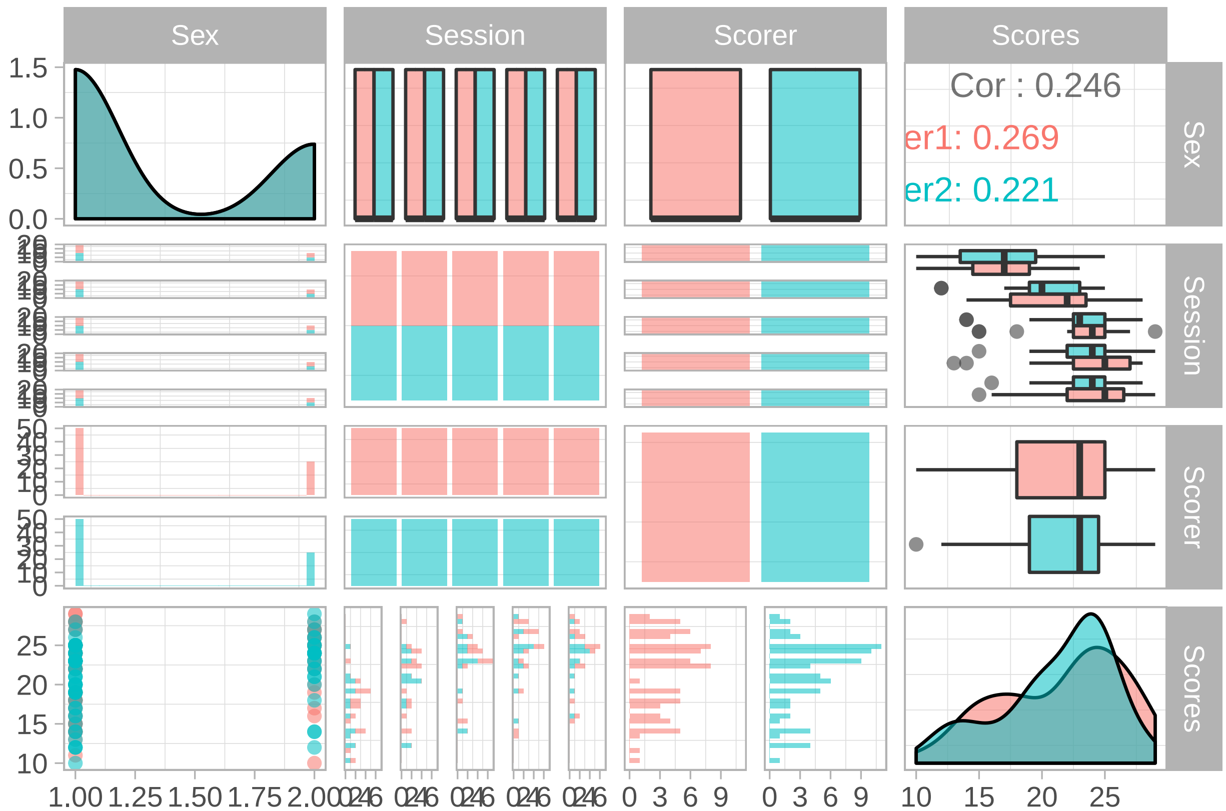

Figure 38: Accuracy of Scores of all subjects

7.Correlation Matrix Visualization of Plots of Suturing Data based on scorer

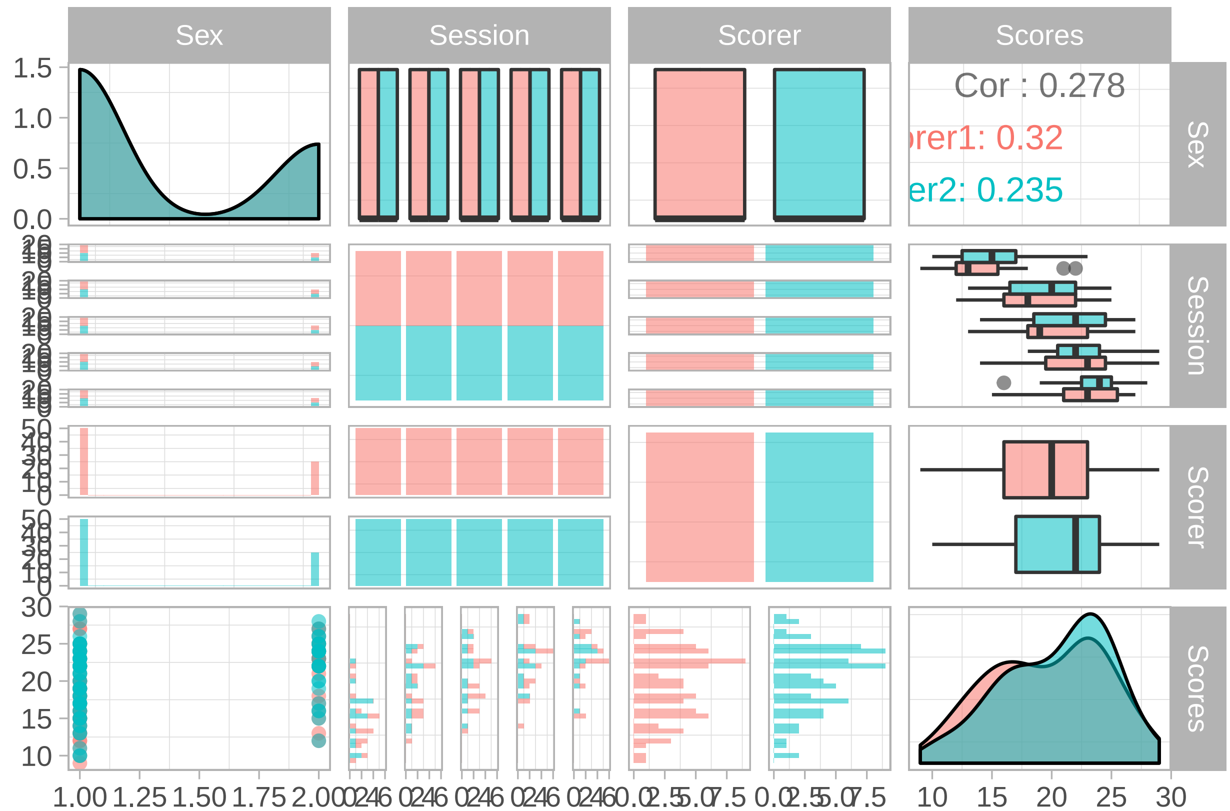

Figure 39: Accuracy of Scores of all subjects

REFERENCES