Title # Temperature Measurement Enhancement

Subtitle Assessing Regression of Variables

D. https://archive.ics.uci.edu/dataset/925/infrared+thermography+temperature+dataset
Please add from notes here in report [Facial and oral temperature data from a large set of human subject volunteers v1.0.0]

2.2.1. Intro (5 Points)

I am planning to research about contribution of 2 factors in 2 sheets, total of 2*2 = 4 cases, which are related with categorical and numerical data sets, but I will focus on numerical ones to research about regression, as the link also mentioned that: ”The class labels for the regression task are either the oral temperature measured in fast mode (aveOralF) or the temperature measured in monitor mode (aveOralM).”.

Data sets FLIR and ICI labeled with group 1 and 2 in the link consisted of multiple measurements, while I omitted local measurements and restricted my focus to aveOralF and aveOralM to make evaluations. This decision had a couple of consequences: First of all, I avoided a huge amount of time and coding to come up with a new indicator. Secondly, in the last columns in the main sheets on the link, as well as description on the page, aveOralF and aveOralM had other factors including atmosphere temperature, humidity, and distance to measure relations. Data Type checking and required data type conversion will be conducted in each section.

Oral temperature is a common measurement for elevated body temperature (EBT). Infrared thermographs (IRTs) have been applied for severe acute respiratory syndrome, Ebola virus disease, and coronavirus disease 2019. The aims of this clinical study were to (1) evaluate the performance of IRTs following consensus guidelines, (2) evaluate the effect of facial measurement location, (3) compare methods for IRT calibration, (4) identify best practices for assessing IRT clinical accuracy, and (5) compare IRT clinical accuracy to that of non-contact infrared thermometers.
Two evaluated infrared thermographs (IRTs) [4] were used for the clinical study – one from FLIR (IRT-1) and the other from ICI (IRT-2).
Data from 1020 subjects measured with IRT-1 and 1009 subjects measured with IRT-2 are available for analysis. About 11% of these subjects exhibited reference temperature above 37.5 °C.
FLIR (IRT-1) stands for Forward Looking Infrared (appendix) IRT-1 The IRT1 small-area infrared emitter is used as an assistive listening system in small areas, such as conference rooms, hospital rooms, and assisted living facilities.
ICI stands for Infrared Camera Inc. (1)
The Infrared Thermography Temperature Dataset contains temperatures read from various locations of inferred images about patients, with the addition of oral temperatures measured for each individual. The 33 features consist of gender, age, ethnicity, ambiant temperature,

humidity, distance, and other temperature readings from the thermal images. The dataset is intended to be used in a regression task to predict the oral temperature using the environment information as well as the thermal image readings. (2)
Website claims data has no missing values and has 1020 instances (rows) for FLIR and 1009 for ICI and 33 features (columns).
I'm planning to implement Option 1.1 and Option 2.2 models.

2.2.2. **Data Cleaning and Exploratory Data Analysis** (15 Points)

There are 36 data types represented in this page in a chart. [Infrared Thermography Temperature - UCI Machine Learning Repository]
Some don't require to be inserted in the data feed including ID. Also, we need to focus on a number of data sets because we are going to evaluate the behavior of some data in a mutual contribution.
I consider - gender
- age
- ethnicity
- ambiant temperature
- humidity
- distance
- temperature readings from the thermal images

Over
oral temperature measured in fast mode (aveOralF) or the temperature measured in monitor mode (aveOralM).
I ignore
T_offset1 (°C) , Cosmetics 1 0 , Date , and Time for my project.

A) SubjectID has a determined formatting of digits and dash, AveOralF and AveOralM are floating point numbers approximately ranging between 35 to 40 celsius (open interval). Gender, Age and Ethnicity are categorical data types, Age is categorical because it's age ranges rather than exact age. T_atm, Humidity and Distance are all floating point numbers as expected. Table 2 has a nice chart in the website
https://physionet.org/content/face-oral-temp-data/1.0.0/_Table2.pdf
B) Website mentions no N/A or NA values in the sheet, as data provided also pointed out that there's no missing value in the sheet, **whereas** I see some empty cells, i.e. FLIR has 1020 subjects and ICI has 1009 subjects evaluation showed that 1001 mutual SubjectID remains so less than 20 instances (rows) can be considered. I would prefer to omit those unbalanced (some SubjectIDs appeared only in one of the FLIR or ICI sheets) instances, because less than 20/appx. 1000 = 2% doesn't play a game changing role in the final results so it isn't worth our resources to do anything for them. I use subjectID to find and omit them accordingly.
Those aren't the only missing values, Round 1 and Round 2 also have some missing values but they aren't in my scope for modelling as I'm focusing on the other parameter's section.
Also, Both ICI and FLIR data sets, Among aveOralF, aveOralM, Gender, Age, Ethnicity, T_atm, Humidity, Distance, Cosmetics, Time, Date, I see missing values in cosmetics dimension (column) as I clarified formerly, they aren't in my main assessment scope. There are also 2

missing values in FLIR and ICI in **DR dimension** so I prefer to omit these ~2/1000 instances completely as it costs less in the sense of missing value handling. IDs 180313-05 and 180313-06 are those missing Distance. Moreover, it sounds like in numerating or entering data some SubjectIDs have been lost, i.e. 180313-04, 161118-3, 161121-2, and 161121-3. This issue doesn't affect the result because it's just labeling instances.
I did data cleaning and exploratory manually, but there are some code-based methods to do so which I couldn't find in the module context.
C) This implementation has done its mission to respond to this question properly. I utilized summary and plot syntaxes to make it work for this reason.
Here's a screenshot of what you may see on
[Assignment 2… (auto-N : 2) - JupyterLab]

```
  SubjectID           aveOralF         aveOralM          Gender         
 Length:1001        Min.   :35.75   Min.   :35.54   Length:1001       
 Class :character   1st Qu.:36.80   1st Qu.:36.74   Class :character  
 Mode  :character   Median :36.90   Median :36.94   Mode  :character  
                    Mean   :36.98   Mean   :37.03                     
                    3rd Qu.:37.10   3rd Qu.:37.14                     
                    Max.   :39.60   Max.   :40.34                     
     Age             Ethnicity             T_atm          Humidity    
 Length:1001        Length:1001        Min.   :20.2   Min.   : 9.90  
 Class :character   Class :character   1st Qu.:23.4   1st Qu.:17.60  
 Mode  :character   Mode  :character   Median :24.0   Median :26.10  
                                       Mean   :24.1   Mean   :28.59  
                                       3rd Qu.:24.7   3rd Qu.:35.90  
                                       Max.   :29.1   Max.   :61.20  
    Distance      
 Min.   : 0.5400  
 1st Qu.: 0.6000  
 Median : 0.6200  
 Mean   : 0.7316  
 3rd Qu.: 0.7000  
 Max.   :79.0000  
```

```
  SubjectID           aveOralF         aveOralM          Gender         
 Length:1001        Min.   :35.75   Min.   :35.54   Length:1001       
 Class :character   1st Qu.:36.80   1st Qu.:36.74   Class :character  
 Mode  :character   Median :36.90   Median :36.94   Mode  :character  
                    Mean   :36.98   Mean   :37.03                     
                    3rd Qu.:37.10   3rd Qu.:37.14                     
                    Max.   :39.60   Max.   :40.34                     
     Age             Ethnicity             T_atm         Humidity      
 Length:1001        Length:1001        Min.   :20.2   Min.   :   9.90  
 Class :character   Class :character   1st Qu.:23.4   1st Qu.:  17.60  
 Mode  :character   Mode  :character   Median :24.0   Median :  26.10  
                                       Mean   :24.1   Mean   :  31.58  
                                       3rd Qu.:24.7   3rd Qu.:  35.90  
                                       Max.   :29.1   Max.   :3021.10  
    Distance      
 Min.   : 0.5400  
 1st Qu.: 0.6000  
 Median : 0.6200  
 Mean   : 0.7316  
 3rd Qu.: 0.7000  
 Max.   :79.0000  
```

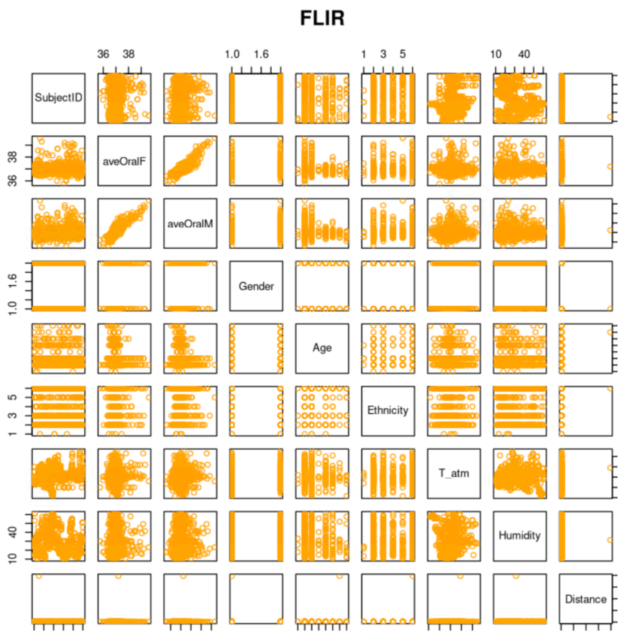

Fig. 1. Plotting FLIR with representing mutual correlations

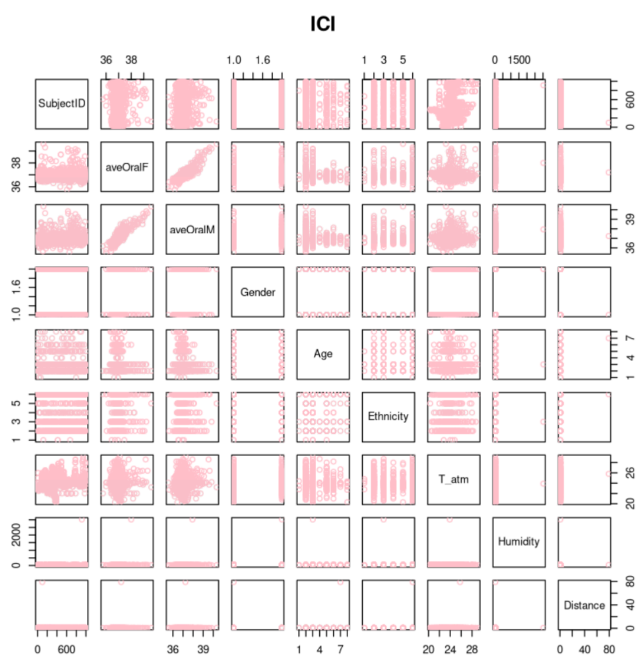

Fig. 2. Plotting FLIR with representing mutual correlations

D) A "response variable" is the variable you are trying to predict or explain, while a "predictor variable" is the variable used to make that prediction. (3) I took numerical variables including T_atm, Humidity and Distance as predictor variables and each time aveOralF and aveOralM may be assessed.

More Data Exploratory has been provided through charts in the following:

A data.frame: 10 × 9

	SubjectID	aveOralF	aveOralM	Gender	Age	Ethnicity	T_atm	Humidity	Distance
	<chr>	<dbl>	<dbl>	<chr>	<chr>	<chr>	<dbl>	<dbl>	<dbl>
491	171103-03	36.95	36.94	Male	21-25	White	27.2	48.5	0.60
649	180131-10	36.60	36.74	Female	21-25	White	24.5	11.5	0.60
330	170412-04	36.85	36.59	Female	18-20	Black or African-American	23.1	48.2	0.65
368	170421-03	36.90	37.14	Male	18-20	Hispanic/Latino	23.6	60.3	0.67
460	171027-04	37.00	36.99	Male	21-25	White	25.2	29.2	0.60
439	171020-10	36.90	37.09	Male	21-25	Asian	24.9	35.6	0.60
584	171208-06	37.25	37.04	Female	26-30	White	23.7	16.0	0.60
438	171020-09	37.15	37.14	Female	18-20	Asian	25.6	35.1	0.60
423	171018-04	36.85	36.69	Female	18-20	Black or African-American	26.5	28.3	0.60
511	171108-01	36.70	36.74	Male	18-20	Asian	25.4	28.2	0.60

Fig. 3. Random selection of FLIR chart to have a fast observation of details

A data.frame: 10 × 9

	SubjectID	aveOralF	aveOralM	Gender	Age	Ethnicity	T_atm	Humidity	Distance
	<chr>	<dbl>	<dbl>	<chr>	<chr>	<chr>	<dbl>	<dbl>	<dbl>
491	171103-03	36.95	36.94	Male	21-25	White	27.2	48.5	0.60
649	180131-10	36.60	36.74	Female	21-25	White	24.5	11.5	0.60
330	170412-04	36.85	36.59	Female	18-20	Black or African-American	23.1	48.2	0.65
368	170421-03	36.90	37.14	Male	18-20	Hispanic/Latino	23.6	60.3	0.67
460	171027-04	37.00	36.99	Male	21-25	White	25.2	29.2	0.60
439	171020-10	36.90	37.09	Male	21-25	Asian	24.9	35.6	0.60
584	171208-06	37.25	37.04	Female	26-30	White	23.7	16.0	0.60
438	171020-09	37.15	37.14	Female	18-20	Asian	25.6	35.1	0.60
423	171018-04	36.85	36.69	Female	18-20	Black or African-American	26.5	28.3	0.60
511	171108-01	36.70	36.74	Male	18-20	Asian	25.4	28.2	0.60

Fig. 4. Random selection of ICI chart to have a fast observation of details

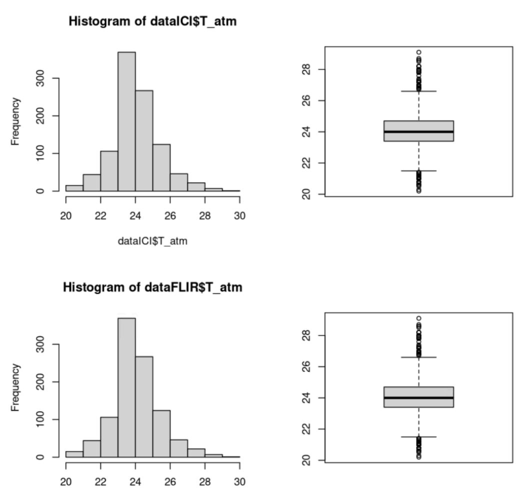

Fig. 5. Histogram and boxplot representation of atmosphere temperature of two datasets

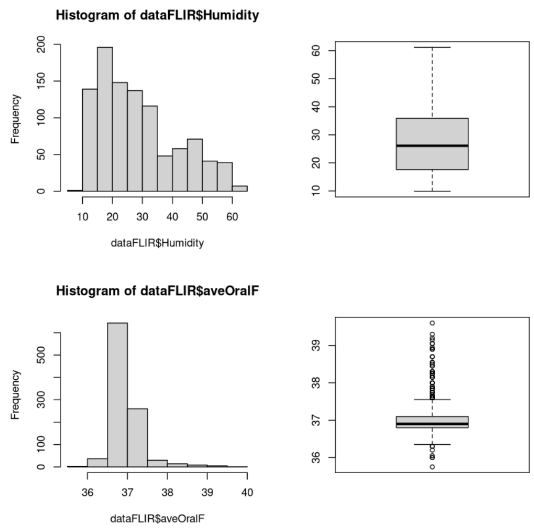

Fig. 6. Humidity and aveOralF of FLIR has been assessed by histograms and boxplots

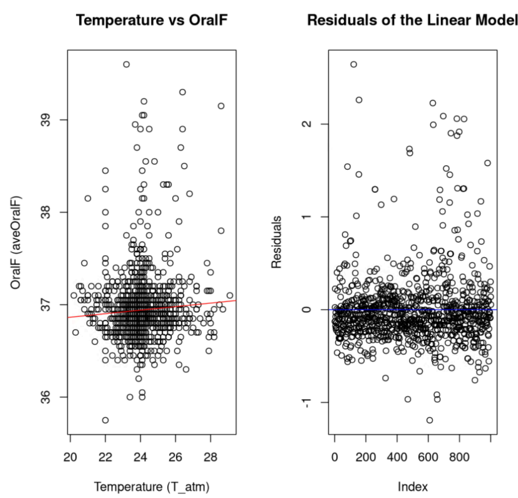

Fig. 7. A satisfactory visualization of trend and regression of Temperature and aveOralF as well as residuals in lm part.

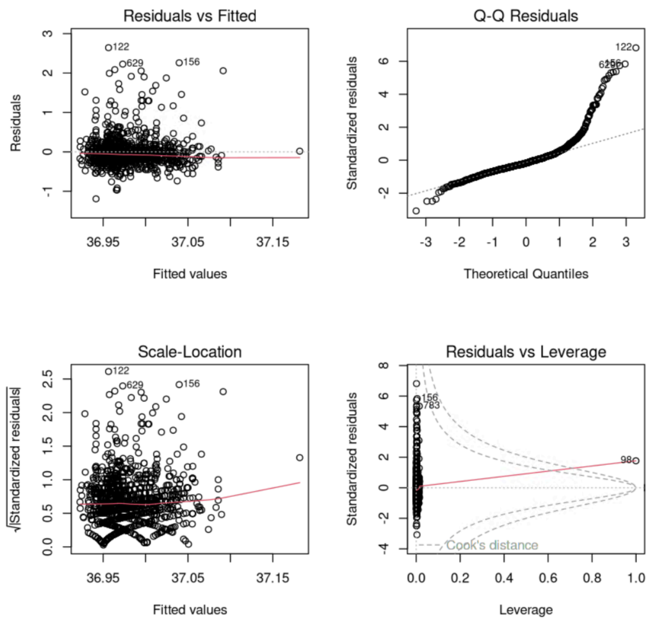

Fig. 8. Additional plotting to have a big picture of more favorable parts of the sheets.

2.2.3. **Modelling** 40

Option 1.1: Lasso, Ridge, and Elastic Net Regression

OLS (ordinary least square) Regression model is used to estimate the relationship between the dependent variable (here it is aveOralF) and the independent variables (here they named T_atm, Humidity, and Distance). It optimized the summation of the squared differences (as the original statistical formula suggests) between the predictor and response variables.

Recalling the favorable feature of Ridge, Ridge regression is used to address multicollinearity in linear regression models. It includes a penalty term to the loss function through lambda, which is proportional to the summation of the squared values of the coefficients. This penalty term helps to shrink the coefficients to zero (note: NOT getting zero like LASSO, it just converges them to the zero), reducing their variance and making the model more effective to evaluate.

Bearing in mind that LASSO is not multilear as Ridge is, Lasso Regression (Least Absolute Shrinkage and Selection Operator) is commonly implemented to perform both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model it produces. It minimizes the sum of the squared differences between the observed and predicted values, with an added penalty proportional to the sum of the absolute values of the coefficients.
LASSO isn't multilinear and recruits coefficients radically to omit less impressive factors totally, instead of tending them to the zero as Ridge does normally in its nature.

Fascinating nature of Elastic net is utilizing both LASSO and Ridge features in a balanced tradeoff, setting alpha parameter as 0.5 and initiating lambda as LASSO, this machine learning regression model brings insightful results to our canva.

In the upcoming Figures, a diverse range of implementation of data sets and results has been provided.

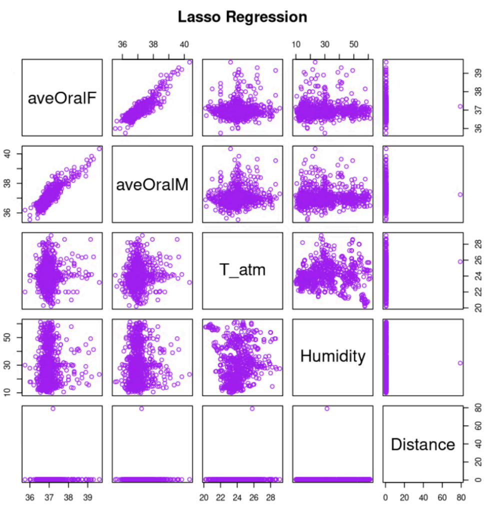

Fig. 9. LASSO mutual Regression plotting

It's trivial that Distance has an outlier. Omitting that amount, as the datasheet offers, due to medical protocol readings, distance difference isn't much variance in different cases. As obviously expected, aveOralF and aveOralM have a linear regression (please see cell [1,2] and [2,1]). Humidity and Temperature in this observation doesn't have a clear regression on the plot.

```
Call:
lm(formula = aveOralF ~ T_atm + Humidity + Distance, data = dataFLIR)

Residuals:
     Min       1Q   Median       3Q      Max
-1.19161 -0.20051 -0.06945  0.10027  2.64364

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.648e+01  2.251e-01 162.055   <2e-16 ***
T_atm       1.908e-02  9.380e-03   2.034   0.0422 *
Humidity    1.248e-03  9.461e-04   1.319   0.1874
Distance    2.115e-03  4.957e-03   0.427   0.6698
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3881 on 997 degrees of freedom
Multiple R-squared:  0.006737,  Adjusted R-squared:  0.003748
F-statistic: 2.254 on 3 and 997 DF,  p-value: 0.08055
```

Fig. 10. FLIR linear Model details output includes Residuals, Coefficients, and Interception

```
Call:
lm(formula = aveOralF ~ T_atm + Humidity + Distance, data = dataICI)

Residuals:
     Min       1Q   Median       3Q      Max
-1.18710 -0.19787 -0.06923  0.09908  2.64096

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.648e+01  2.244e-01 162.567  < 2e-16 ***
T_atm       2.011e-02  9.302e-03   2.162  0.03082 *
Humidity    3.603e-04  1.281e-04   2.812  0.00503 **
Distance    2.155e-03  4.942e-03   0.436  0.66293
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3869 on 997 degrees of freedom
Multiple R-squared:  0.01283,   Adjusted R-squared:  0.00986
F-statistic: 4.319 on 3 and 997 DF,  p-value: 0.004901
```

Fig. 11. ICI linear Model details output includes Residuals, Coefficients, and Interception

```
4 x 1 sparse Matrix of class "dgCMatrix"
                       s0
(Intercept) 36.483168447
T_atm        0.019078805
Humidity     0.001248110
Distance     0.002114517
```

Fig. 12. An example of a confusion matrix which includes Interception, T_atm, Humidity and Distance as a result of linear model and coefficient function in R

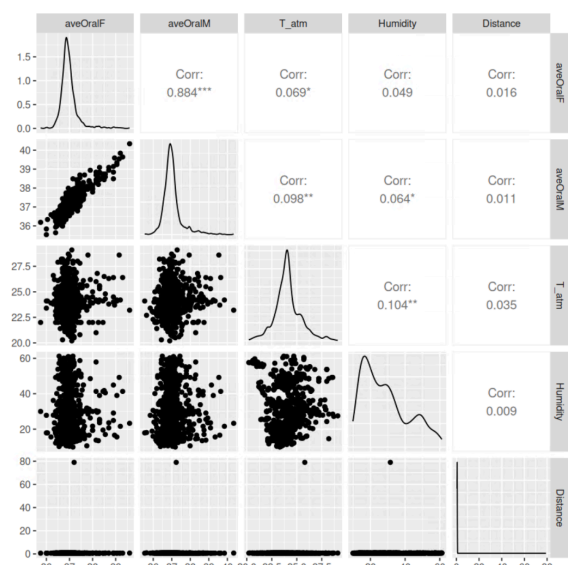

Fig. 13. Output of GGplot for Ridge Regression Model

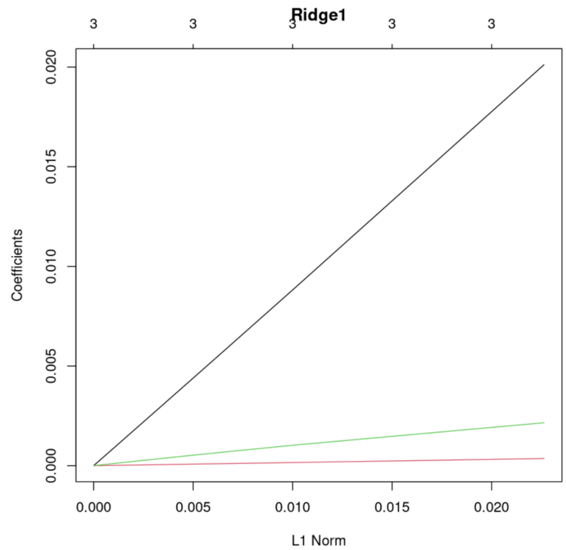

Fig. 14. A visual interpretation of coefficient vs. L1 Norm in a Ridge Regression Model

Lasso

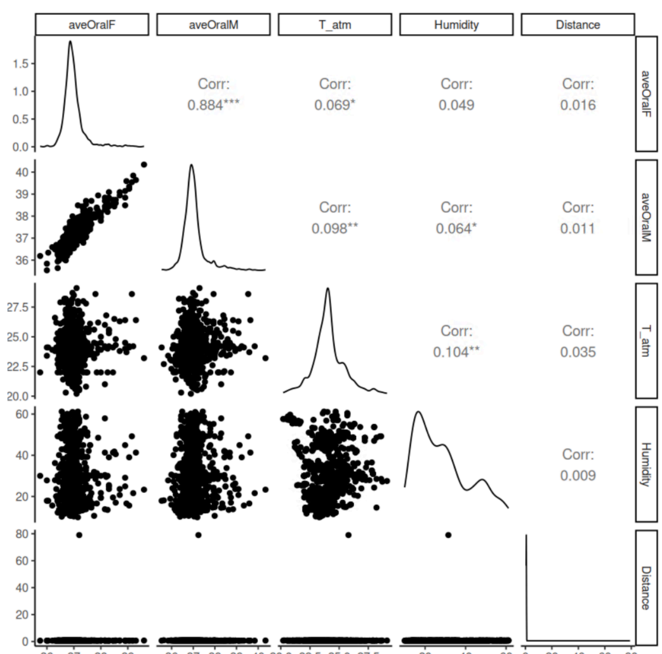

Fig. 15. Output of GGplot for LASSO Regression Model

This Figure in the diagonal cells reveals a more gaussian distribution mostly with one distinguishable global maximum except for Humidity which has multiple peaks.

Option 2.2: Neural Network

```
  target
   36 36.05 36.3 36.4 36.45 36.5 36.55 36.6 36.65 36.7 36.75 36.8 3
37  2     1    2    2     4    7     7    4    15   24    21   24
38  0     0    0    0     0    0     0    0     0    0     0    0
  target
   36.95 37 37.05 37.1 37.15 37.2 37.25 37.3 37.35 37.4 37.45 37.5
37    29 27    24   12    14    5     3    6     8    2     2    2
38     0  0     1    0     0    0     0    0     0    0     0    0
  target
   37.65 37.7 37.75 37.85 37.9 38 38.15 38.25 38.3 38.45 38.9 39.6
37     1    1     1     1    2  1     2     1    1     2    1    1
38     0    0     0     0    0  0     0     0    0     0    0    0

   35.75 36.2 36.35 36.4 36.45 36.5 36.55 36.6 36.65 36.7 36.75 36.
37     1    1     4    5     7    4    11   21    23   32    34   6
38     0    0     0    0     0    0     0    0     0    0     0

   36.9 36.95 37 37.05 37.1 37.15 37.2 37.25 37.3 37.35 37.4 37.45
37   60    56 50    41   34    19   32    18    5     7    8    10
38    0     1  0     0    0     1    1     0    0     0    0     0

   37.6 37.65 37.75 37.8 37.85 38 38.15 38.2 38.25 38.3 38.5 38.55
37    6     2     2    1     2  3     2    1     1    3    1     2
38    0     0     0    0     0  0     0    0     0    0    0     0
```

Fig. 16. Neural Network by Perceptron Method Outputs

As a quick review of numbers reveal, the algorithm modified prediction by omitting those labeled 0 by the weight * vector inner product.

One of the reasons for using Bootstrap is the potential of using to estimate the bias of an estimator by comparing the average of bootstrap estimates to the original estimate.

The bootstrap method behaves the empirical distribution of the sample as an approximation of the true population distribution, which looks proper for my research. (4)

My implementation worked more around two more common body temperatures which are 37 and 38.

My perceptron model had a single layer including the input features (T_scale, H_scale, D_scale) and the output (aveOralF). The perceptron is a simple neural network model with one layer that directly maps the input features (T_scale, H_scale, D_scale) to the output (aveOralF) through a set of weights and a bias = -threshold. My initial bias was set to 0, my final bias resulted in 36.92675 as you may also see on the NCC Jupiter page.

2.2.4. **Model Comparison** 20

I utilized Bootstrapping Validation because the size of datasets is low (1000 instances and a few features) also since multiple machine learning methods were implemented, this looks much satisfying in comparison with other types named as LOOC and Cross Validation models.

```
ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = dataNumFLIR, statistic = boot_fn, R = 1000)


Bootstrap Statistics :
        original        bias     std. error
t1* 36.483168468  1.323743e-01 0.3129133969
t2*  0.019078805 -2.187721e-03 0.0107327365
t3*  0.001248110  6.687018e-05 0.0008250446
t4*  0.002114517 -1.259982e-01 0.1937373942
```

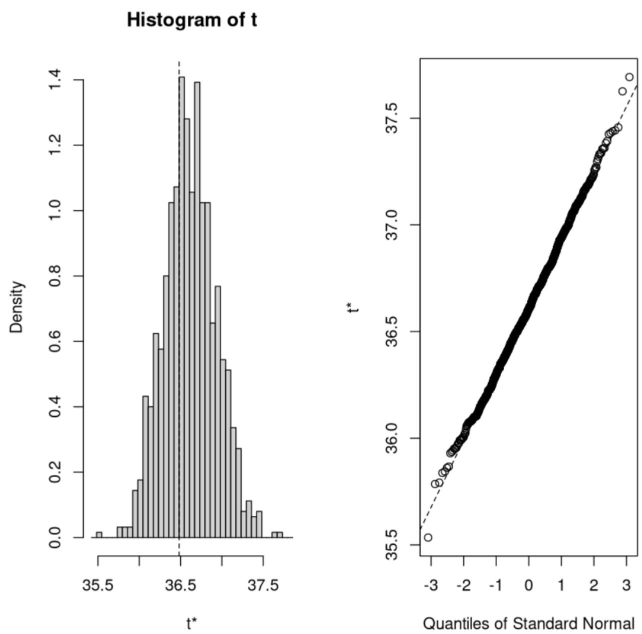

Fig. 17. Graphical representation of Perceptrons through histogram and Standard Norm Visualization which includes a spline to determine the trend

T_atm:

Original estimate: 0.0191
Bias: -0.0022, indicating a tiny(because absolute value of it is small) and underestimation (because of the negative sign).
Standard error:0.0107 declares the variability of this coefficient estimate which is still low and fine

Humidity:

Original estimate: 0.0012
Bias: 0.0001, indicating an incredibly ignorable(because absolute value of it is small) and overestimation (because of the positive sign).

Standard error:0.0008 declares the variability of this coefficient estimate which is still low and fine

Distance:
Original estimate: 0.0021
Bias: -0.1260, indicating a significant (in percentage 12.6%) and underestimation (because of the negative sign).
Standard error:0.1937 unfortunately declares the variability of this coefficient is high and unreliable.

What do these results tell us?
Coefficients for `T_atm` and `Humidity` have relatively small biases and standard errors, indicating stable estimates whilst coefficient under `Distance header` has a larger bias and standard error, suggesting less stability and higher variability in its estimate.

2.2.5. Results and Conclusion 10

I would like to make a quick review and some highlights of results to come up with an insightful conclusion.
Considering the nature of this project as a regression medical measurement of two methods of body temperature measuring enriched with multiple features starting from point by point measurement to personal major differences including but not limited to age range, ethnicity, gender, and some atmospheric numerical determiners including atmosphere temperature and humidity, we tried to define a small handleable problem to enjoy surfing a number of regression studies through machine learning methods utilizing narrow down datasets to be well-prepared for this case.
As the cropped datasets were numerical, option1.1. Which was the LASSO, RIDGE, and Elastic Net was chosen and a short perception was implemented on this numerical data. Afterwards, a Bootstrap Validation was conducted to check the level of suitableness of the crafted machine learning model.
Some features didn't play a significant role, including Distance, and we observed the goodness of model through the Bootstrap validation.
Although data preparation and data cleaning play a great role in data analysis and machine learning tasks,just small justifications including omitting missing mutual SubjectIDs, and missing 2 instances helped us to come up with a final version.
In the first ever scanning, cosmetic looked incredibly attractive to be in the neural network playing field,but due to some missing values, omitting that could save more space for other sort of model implementation.
The findings highlight the importance of considering both bias and variance in model evaluation. Both graphical and numerical interpretations provided in Figures.
The study provides a detailed comparison of different regression models, helping to identify the best approach for the given data.

Bibliography
Get assistance from [References]

1. *Facial and oral temperature data from a large set of human subject volunteers v1.0.0*. (2023, May 24). https://physionet.org/content/face-oral-temp-data/1.0.0/
2. UCI Machine Learning Repository. (n.d.). Infrared thermography temperature dataset. University of California, Irvine. Retrieved March 26, 2025, from https://archive.ics.uci.edu/dataset/925/infrared+thermography+temperature+dataset
3. Minitab. (n.d.). What are response and predictor variables? Minitab. Retrieved March 26, 2025, from https://support.minitab.com/en-us/minitab/help-and-how-to/statistical-modeling/regression/supporting-topics/basics/what-are-response-and-predictor-variables/
4. Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. Chapman & Hall/CRC. Retrieved from https://www.stat.cmu.edu/~larry/=sml/Boot.pdf