DSCI 5340

Predictive Analytics

Project 1

**PIMA Diabetes prediction**

**(Linear regression)**

By-

Suhail Bari

Lauren Brooks

## Introduction

Diabetes is a chronic disease in which a person’s pancreas is unable to properly respond or

produce the hormone, insulin. Over time, diabetes can cause harmful effects on the body such

as, cardiovascular problems, which can include chest pain, heart attack, stroke, and narrowing of

the arteries. Diabetes can also cause damage to other organs in the body such as the eyes, nerves,

feet and kidneys. This study was done on women at least twenty-one years of age and of Pima Indian

Heritage. Studying the sample size, we are trying to determine if these women do have diabetes based

on a few different factors such as, the number of times they have been pregnant, their plasma glucose

level over two hours, blood pressure, skin thickness, insulin level, body mass (weight), diabetes

pedigree Function, and age.

## Literature Review

1. A model for early prediction of diabetes

Informatics in Medicine Unlocked

Volume 16, 2019, 100204

The dataset used in this study, is originally taken from the National Institute of Diabetes and Digestive and Kidney Diseases (publicly available at: UCI ML Repository [29]). Our findings indicate a strong association of diabetes with body mass index (BMI) and with glucose level, which was extracted via the Apriori method. Artificial neural network (ANN), random forest (RF) and K-means clustering techniques were implemented for the prediction of diabetes. The methodology used was step wise. They followed the logical data preprocessing, cleaning, reduction and transformation steps.

They focused on binning the age, glucose and blood pressure variables into groups from young to old, low to high and starvation to obese. The latter used association rule mining by determining item sets and setting specific rules. To find the association between the item sets X and Y, they set a minimum support of that fraction of transactions which has both X and Y called minsupp. They defined rules as

Rule#1. If (BMI = Obesity) → Class = Yes

Rule#2. If (Glucose = Diabetes) → Class = Yes

Rule#3. If (Glucose = Diabetes BMI = Obesity) → Class = Yes

The models used were ANN, random forest, K-means clustering. The results were evaluated on the basis of accuracy of the AUROC curve. The ANN gave better results than the random forests model (even after increasing the number of branches) and k-means clustering. The AUROC curve of ANN gave a value of 0.816 and an accuracy of 75.7%.

1. Analysis of diabetes mellitus for early prediction using optimal features selection

Journal of Big Data volume 6

Article number: 13 (2019)

These researchers used information mining which is a procedure of removing valuable information from a gigantic set of data. The reason the sample is extracted is because the dataset is so large that any sort of operation to be performed on the data is very time taking and consumptive. The first step is creating an information step index which comprises predefined class names. This preparation set is utilized to assemble the new grouping model. In the next step, we have the informational indices consisting of information occasion without class names. The recently produced demonstrate connects to the test informational index to predict their class marks. The execution of the model is assessed through exactness rate, mistake rate and different measurements.

The implementation methods used were decision trees, naive bayesian, support vector machine, random forest and KNN. The accuracy of SVM was recorded at 77.73%. The accuracy of random forest was 75.39%. The accuracy of Naive Bayesian Classification was 73.48%. Decision tree classification showed an accuracy of 73.18% and the KNN was at 63%.

The decision tree algorithm and Random forest give the highest specificity of 98.20% and 98.00%. The Naive Bayesian Classification accuracy is 82.30%, hence it is able to map the features effectively from low dimensions to high dimensions. It gives the best fit to the data with respect to the diabetic and non-diabetic patients.

1. Prediction of diabetes mellitus in Chinese adults

Meinian Institute of Health, PKUHSC Meinian Public Health Research Institute

& Peking University Health Science Center (PKUHSC)

OP-0293

This study has an objective to establish a risk prediction model for 1-year incident diabetes in Chinese adults, based on 1,532,841 participants with remeasured

health checkup data.

The risk model was created with a random sample of 10% of the participants via stepwise regression model, and then validated using data from the remaining 10% of participants. Logistic regression model was used to estimate the risk of diabetes.

Results showed that 25,821 participants (1.71%) in the derivation cohort and 731 participants (3.9%) in the validation cohort developed diabetes at the second visit. Age, gender, body mass index, systolic blood pressure, fasting plasma glucose, total cholesterol, triglyceride, alanine aminotransferase, fatty liver were statistically significant predictors.

1. Diabetes in Older Adults

American Diabetes Association

Diabetes Care 2012 Dec; 35(12): 2650-2664.

This article dives into the fact that more than 25 percent of the U.S population age 65 and older have a larger percentage of those who have diabetes. This is problematic because it can cause a higher mortality rate, reduce their function, and a greater risk of institutionalization.

According to the study in the article, 22-33% of the 65 and older group range is type 2 diabetic, which is directly linked to overweight/obesity. In the study, it talks about having adults aged 45 and older screened every 1-3 years using an FPG test to treat early for type 2 diabetes.

In the results, there has been a table provided that dives into detail on a reasonable glucose level along with depending on where your level is at, if you should fast and for how long.

1. Pregnancy and Diabetes: How Women Handle the Challenges

[J Perinat Educ.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1595250/) 2005 Summer; 14(3): 23–32.

This article has two studies on a group of 14 swedish women and then the second group of 18 swedish women between ages 25 and 38 years old. Most of the women had higher risks and already had cardiovascular problems due to their diabetes. The women were split into three different categories which were, meaningfulness/meaninglessness, reconciliation/conflict, and shared control/unwillingly controlled.

The results showed that women who were in the meaningful/meaninglessness and were in the acceptant/ambivalent subcategory, knew they were pregnant, but continued to struggle even more so with their glucose levels. They usually had severe morning sickness or were separated from the child's father.

The women who were in the meaningfulness/meaninglessness with the subcategory hope/hopelessness were constantly worried about their child's health, but did note they were aware that they could have a healthy child, even with having diabetes.

The women who were in the meaningfulness/meaninglessness with the subcategory confirmed normality/fortified malaise showed normal signs of pregnancy and was able to have a normal delivery for their baby as well.

The women in the second category, reconciliation/conflict with the subcategory acceptance/opposition struggled to maintain their original lifestyles unless they accepted that some changes were going to need to be made in order to maintain their health and the baby's health. The women had a decrease in their HbA1c, which measures how well controlled the blood sugar is, with a mean of 6.8%.

In the category reconciliation/conflict with the subcategory self understanding/lack of self understanding, some of the women lacked a self understanding that their glucose levels would change periodically due to being pregnant and that caused a lot of issues. This issue had more to do with self understanding of their body and the changes that were being made.

The third category, which is Shared control/unwilling controlled with the subcategory Responsibility/Surrendered responsibility, the women express their need to have control over their bodies and the disease, but also wanted to give the responsibility to the healthcare professions to constantly monitor their glucose levels.

The women in the category Shared control/unwilling controlled with the subcategory Health-care staff as a supportive resource or as a controlling factor needed the constant supervision of the healthcare providers to make sure they were stable within their glucose levels.

The women in the category Shared control/unwilling control with the subcategory Relatives and employers who provide support or increase the pressure relied heavily on the child's father, and if the father was not in the picture, they relied heavily on other family members. These women ran the risk of being in shock and going into an insulin coma, therefore it was very important that their family members help them throughout their pregnancy.

Therefore, the overall success of women, in the study, who were pregnant, is to have constant support from family and the child's father, be consistent with having health care professionals monitor their glucose levels, and be acceptant of the chronic disease to make decisions to better themselves and their child.

## Hypothesis statement

Research question:

With all the data and variables in our dataset, what variables are better predictors for diabetes prediction.

Population: All the data with its variables.

Null Hypothesis: The variable glucose is a better predictor.

H0 : myu(u1) = myu(u2)

Alternate Hypothesis: The variable glucose is not a better predictor.

H1 : myu(u1) != myu(u2) != means not equal.

## Methodology

### Data Source

This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.This is a secondary dataset with 769 rows and 9 columns, which provides numerical information on levels of pregnancies, glucose, skin thickness etc.

#### Dimensions:

* Age – Age range of the population from 21 years to 81 years.
* Outcome - Class variable to check whether the patient has diabetes or not (0 or 1)

#### Measures:

* Pregnancies - Number of times pregnant
* Glucose - Plasma glucose concentration a 2 hours in an oral glucose tolerance test
* BloodPressure - Diastolic blood pressure (mm Hg)
* SkinThickness - Triceps skin fold thickness (mm)
* Insulin - 2-Hour serum insulin (mu U/ml)
* BMI - Body mass index (weight in kg/(height in m)^2)
* Diabetes pedigree function - A value from 0.08 to 2.42 which indicates family history of diabetes

### Handling Null values

There aren’t any null values to begin with in the dataset. However, there are significant zero values which will change the results. Logical reasoning defies the fact that variables such as

glucose or blood pressure or even body mass index will be 0 for any human being. Thus, we will remove the zeros from the dataset and will replace them with the mean of the values in the variable.

We will also explore the statistics without removing zeros from the dataset. SAS has the ability to skip rows which have missing values and provide statistics for only the rows which have data in them. This dataset is a random sample, so even if we reduce the sample, the rules of central limit theorem and random sampling apply to the dataset. The results will not be distorted.

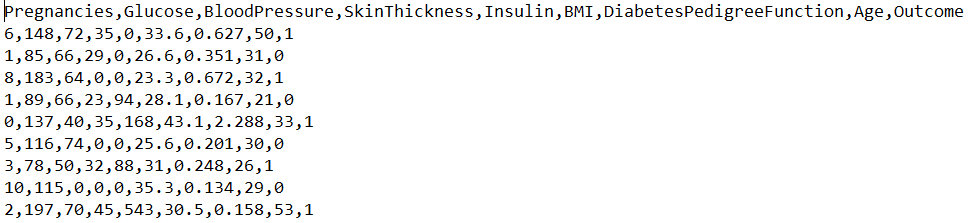


Figure 1: Data Snapshot (before)

Replace all the zeros with . (dot) values. SAS has the ability to read dot values and impute them.

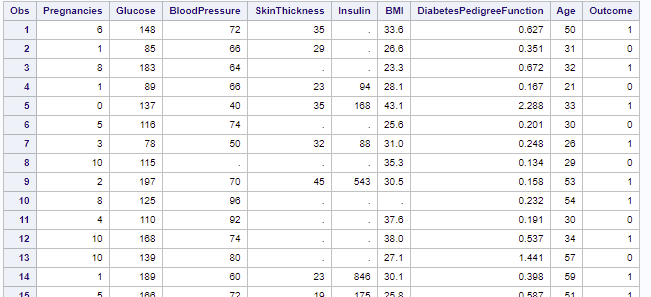


Figure 2: Data Snapshot (after)

#### Data Imputation

The dot values will them be imputed in SAS with the mean of all the values in a

Variable.

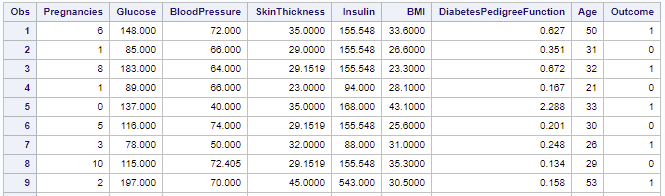


Figure 3: Data imputation (after)

### Descriptive Analytics

1. Analyzing the data overview

We can see that there are a total of 768 rows out of which 500 of the female candidates have an outcome of 0, which means, they do not have diabetes.

268 of the candidates have an outcome of 1, which means they have diabetes. The percentage of women having diabetes in this dataset is 34.9%.

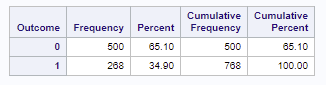


Figure 4: Cumulative

1. Data - Means step

When we look at the means of the data, we can see that the Insulin has the highest standard deviation among all the variables, followed by glucose.

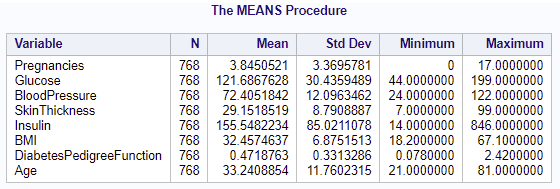
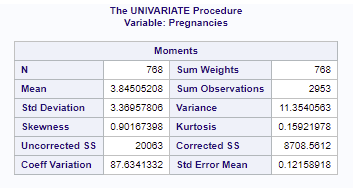
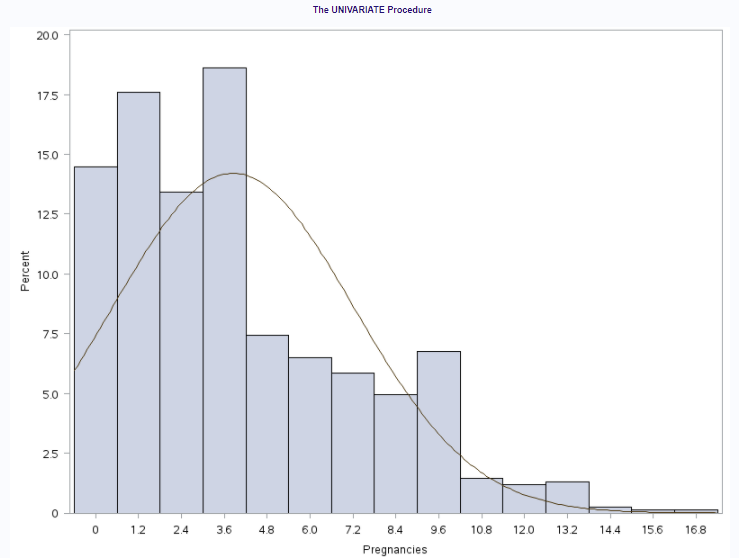


Figure 5: Means

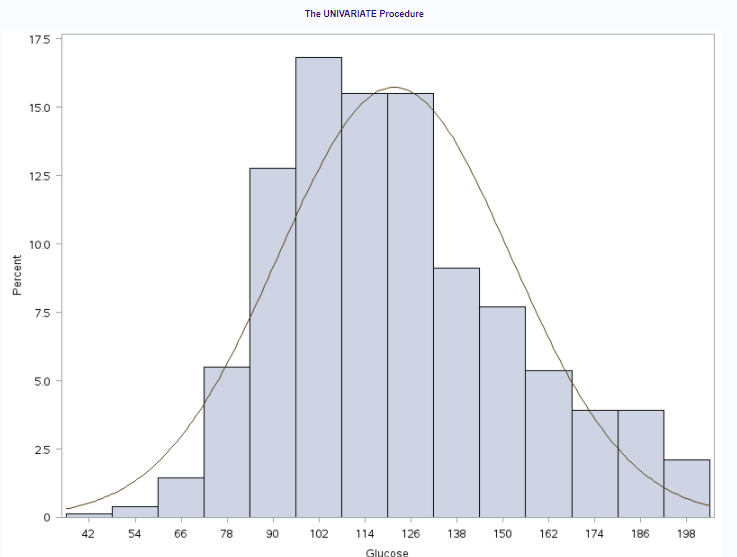
1. Data - Univariate step

Pregnancies

We can see the univariate stats for the variable pregnancies. We notice that the bar graph is not normally distributed with a variance of 11.35 and a standard deviation of 3.36. Since this variable has zero values as well, we will not be transforming this variable because the data after transformation becomes less normal.

Figure 5 & 6: Univariate (Pregnancies)

Glucose

This variable has a std dev of 30.43 and a very high variance of 926.34. However, the data seems to be uniformly distributed in the normal form. We will not be transforming this variable.

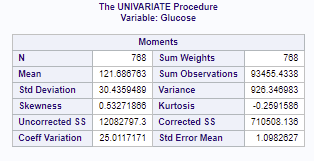
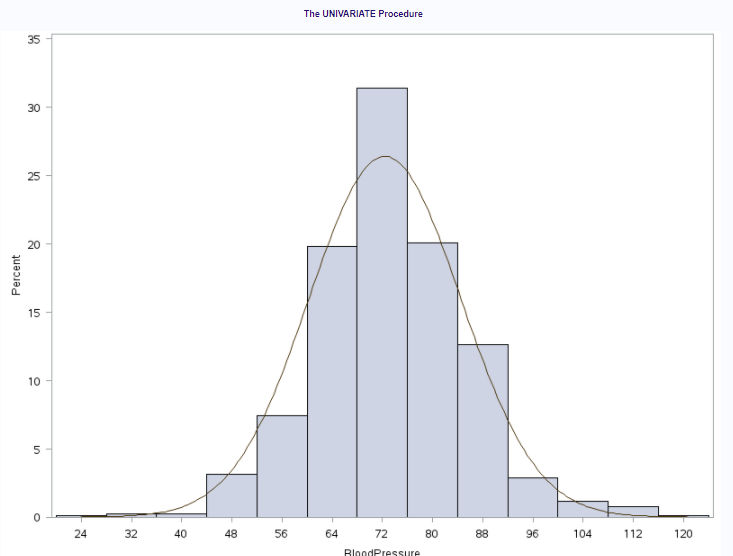


Figure 7 & 8: Univariate (Glucose)

BloodPressure

This variable has a std dev of 12.09 and a variance of 146.323. However, the data

seems to be uniformly distributed in the normal form. We will not be transforming this variable.

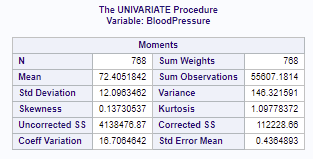
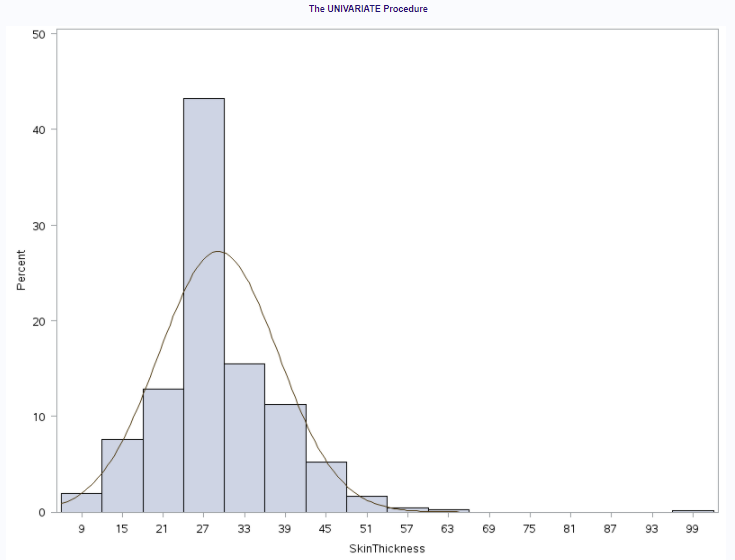


Figure 9 & 10: Univariate (BP)

SkinThickness

This variable has a standard deviation of 8.7 and a variance of 77.27. However, it has a very high kurtosis value of 5.41 and does not follow the principles of the normal distributions. We will be transforming this variable.



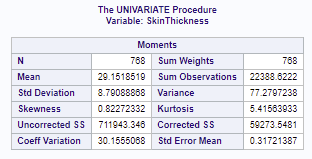
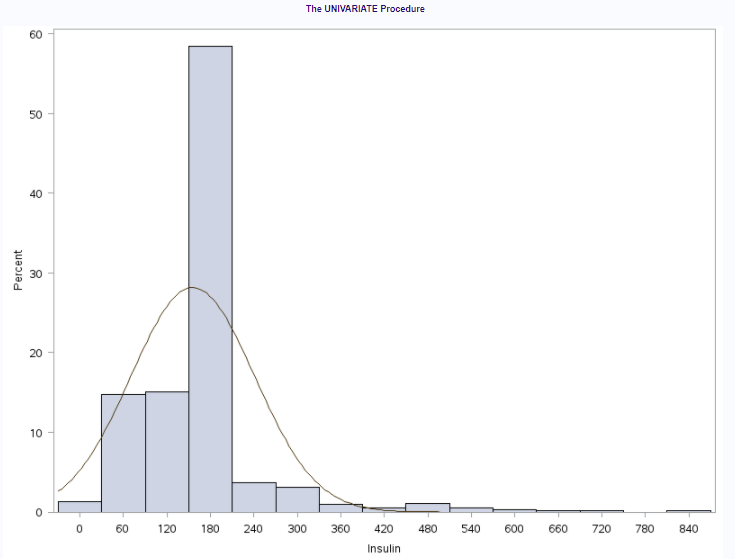


Figure 11 &12: Univariate (ST)

Insulin

This variable has a std dev of 85.02 and a very high variance of 7228.58. The skewness (positive) is 3.019 and the kurtosis is 15.185. This is a must transform variable in our dataset.



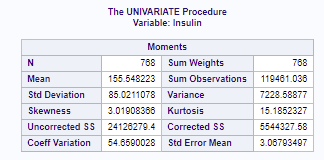
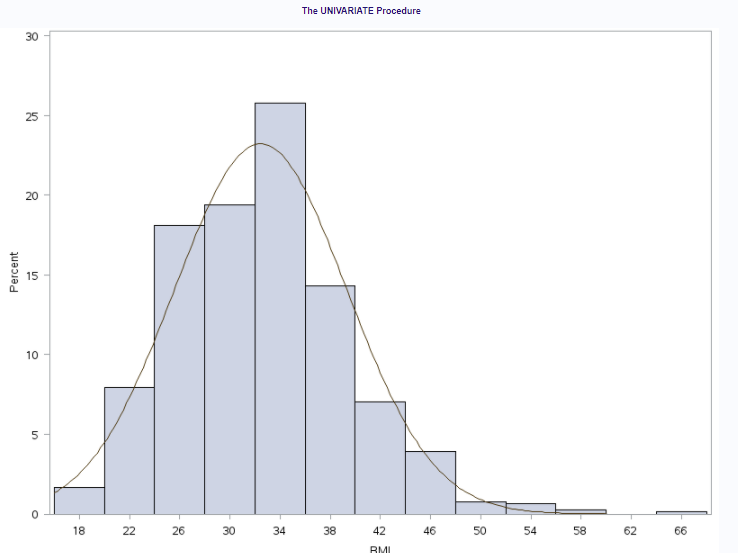
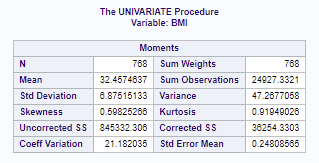


Figure 13 & 14: Univariate (Insulin)

BMI

This variable has a std dev of 6.87 and a variance of 47.26. However, the data seems to be uniformly distributed in the normal form. We will not be transforming this variable.

Figure 15 & 16: Univariate (BMI)



DiabetesPedigreeFunction

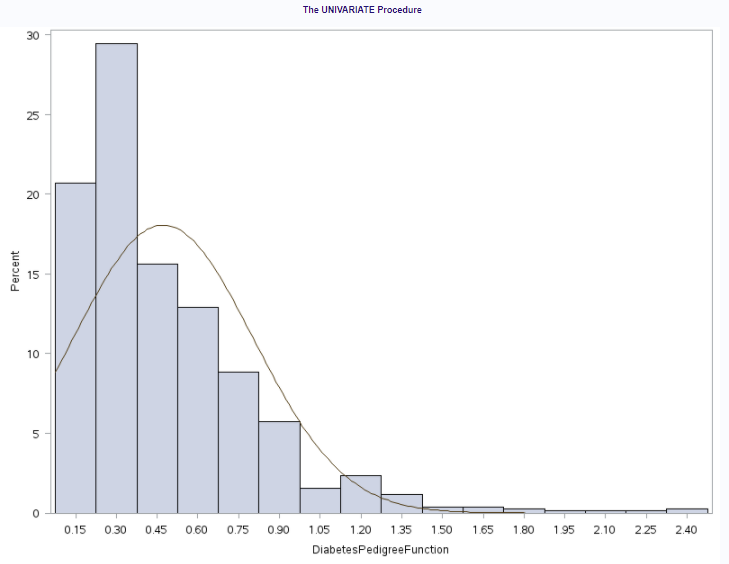
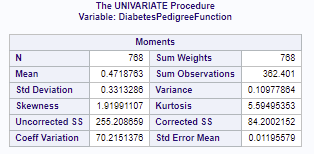
This variable has a std dev of 0.33 and a variance of 0.109. However, the data seems to be highly skewed positively with a kurtosis value of 5.59. We will be transforming this variable.

Figure 17 & 18: Univariate (DPF)

From the above procedure, we can finalize our variables for transformation:

* DiabetesPedigreeFunction
* Insulin
* SkinThickness

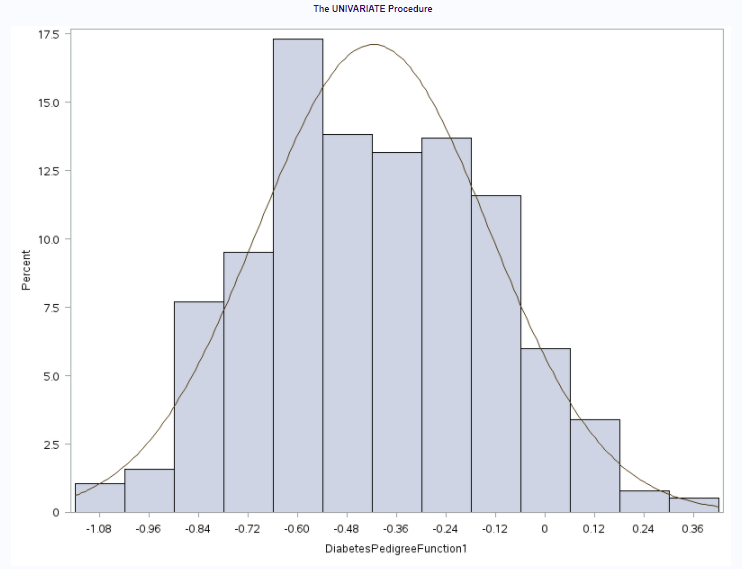


Figure 19: Transformed variable (DPF)

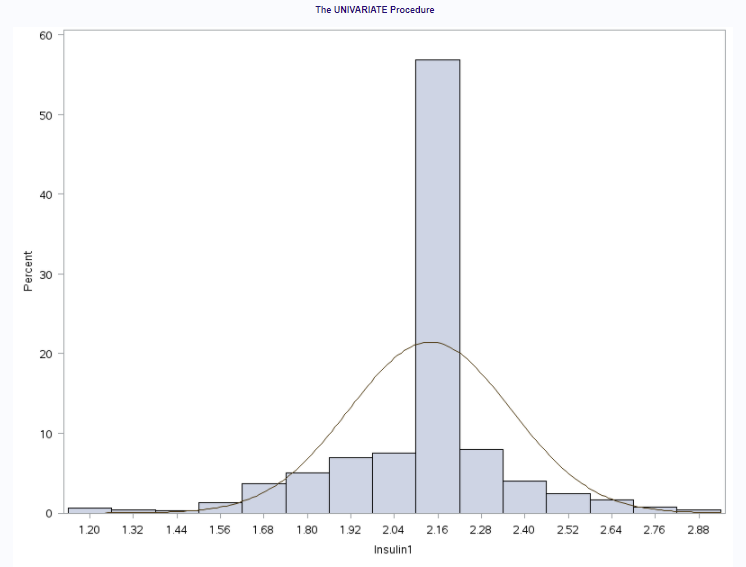


Figure 19: Transformed variable (Insulin)

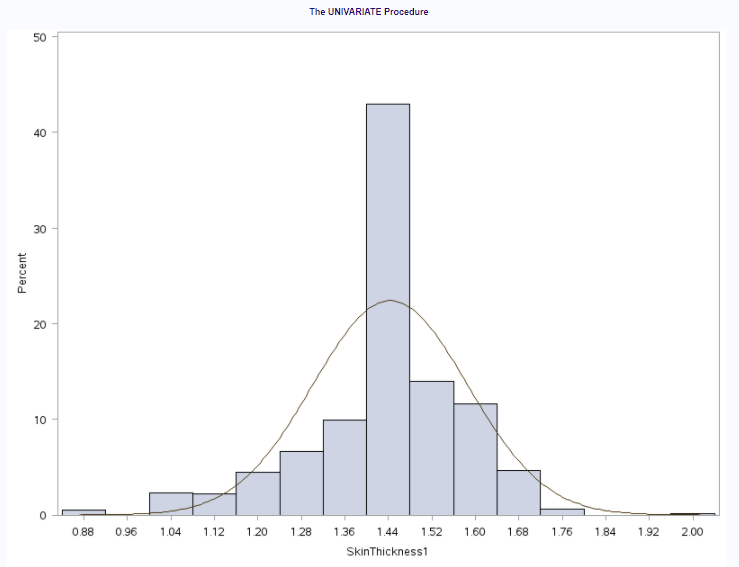


Figure 20: Transformed variable (ST)

The transformations will be concatenated with the original table. We need to drop the pre-transformed variable columns and proceed with correlation.

1. Data - Correlation step

As we pass the transformations, we get into correlating the data with the other variables. We need to see which variable correlated with the other highly. We can notice that Glucose and Insulin have a higher value. We can see that the highest correlation value is Age and Pregnancies but we will not explore that option. We also have Skin thickness and BMI.

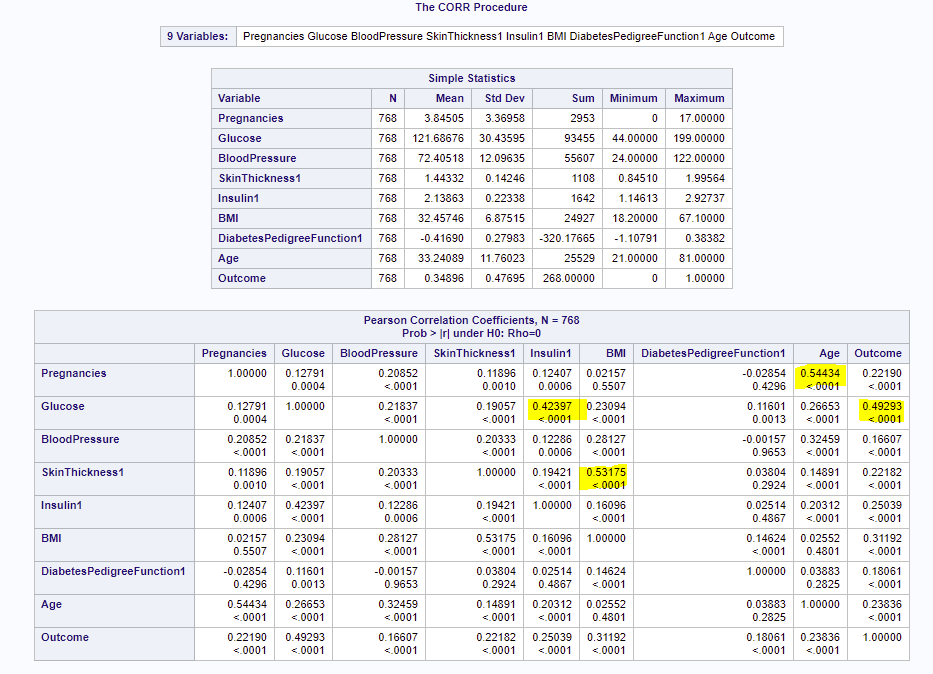


Figure 21: Correlation

1. Data - Plotting step

In the Insulin vs Glucose graph, we can see that there is a clear pattern on the data. As the glucose levels are increasing, the insulin is increasing. The both are positively correlated.

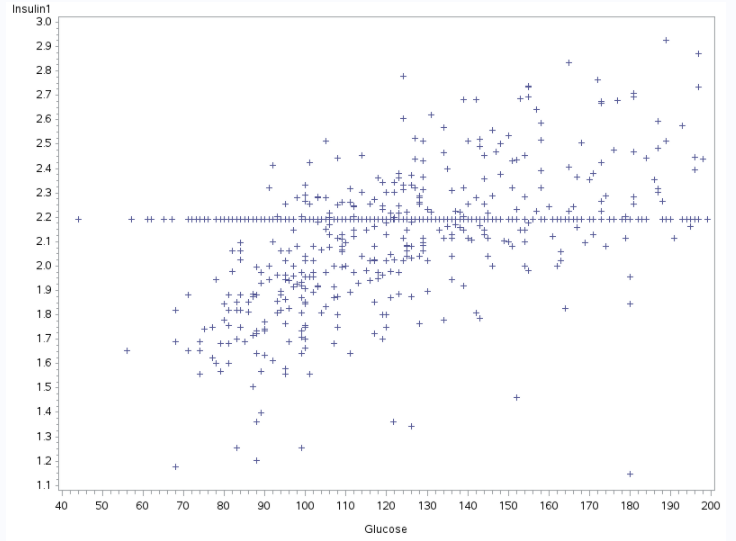


Figure 22: Glucose v/s Insulin

In the Skin thickness vs BMI graph, the variables seem to be positively correlated. In fact they appear to be bending in the form of a curve.

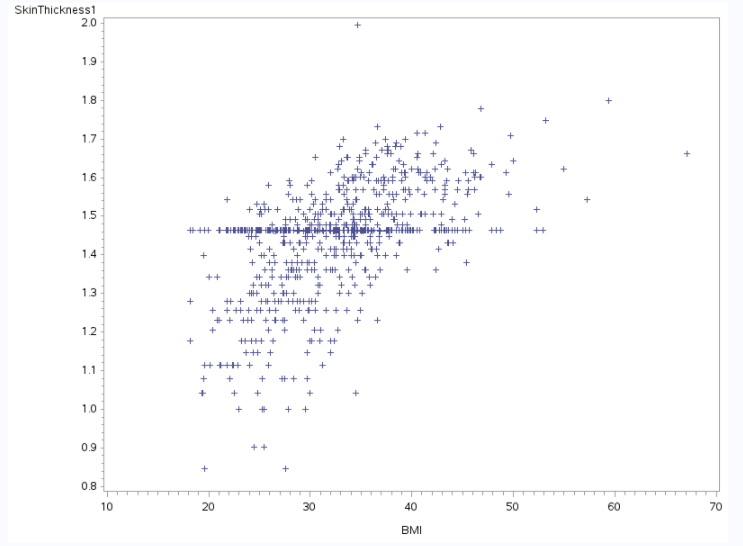


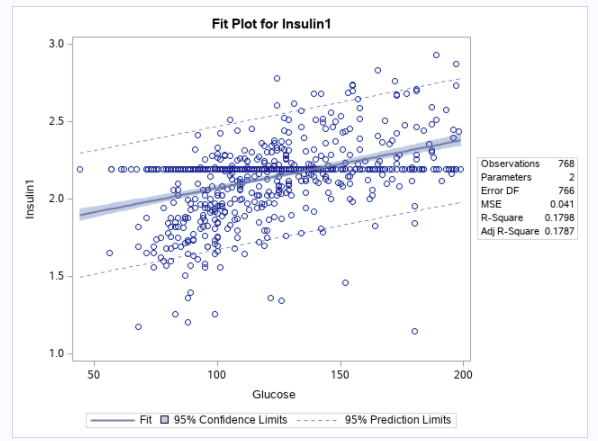
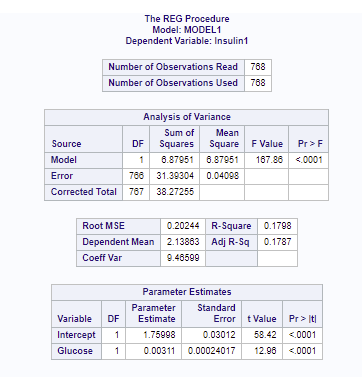
Figure 23: BMI v/s Skin thickness

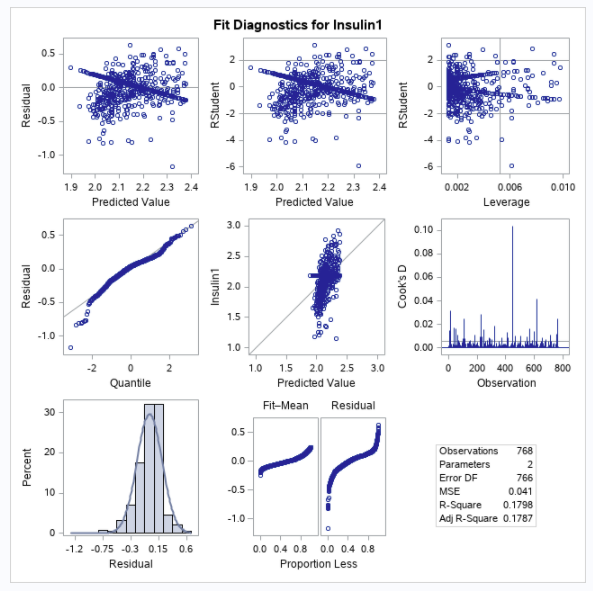
1. Data - regression step

Insulin vs Glucose

In the below step, we can see that the fit statistics give us an intercept value of 1.75998 and a standard error value of 0.3012. The adjusted R square value is 0.178 which is pretty good.

Figure 24,25,26: Insulin Glucose regression

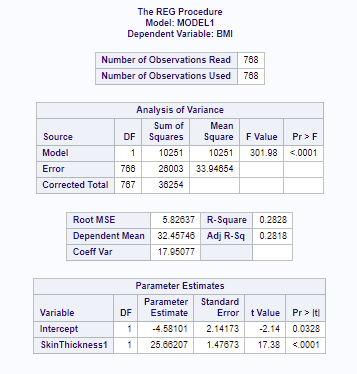


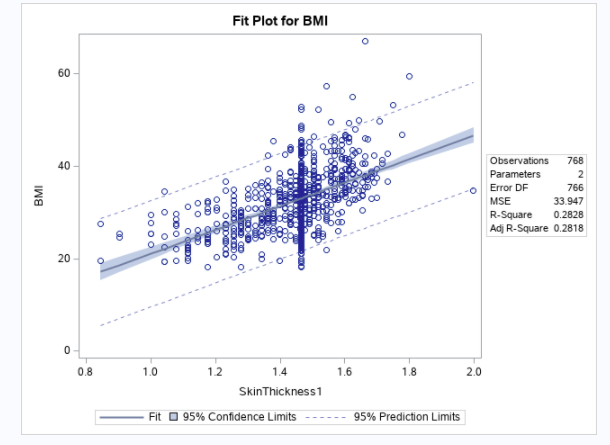
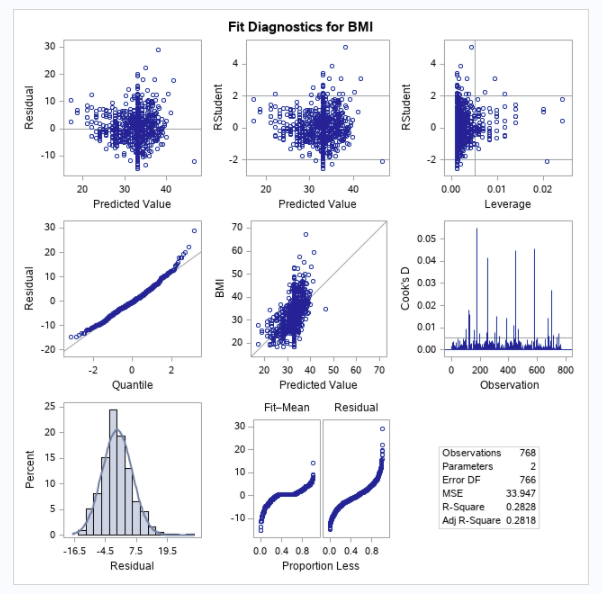


BMI vs Skin thickness

In the below step, we can see that the fit statistics give us an intercept value of -4.581 and a standard error value of 2.141. The adjusted R square value is 0.2818 which is lower than that of the Insulin vs Glucose.

Figure 27,28,29: BMI Skin thickness regression





## Code - SAS

**/\*Get the data in\*/**

data p1;

input Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age Outcome;

infile datalines delimiter=' ';

datalines;

6 148 72 35 . 33.6 0.627 50 1

1 85 66 29 . 26.6 0.351 31 0

8 183 64 . . 23.3 0.672 32 1

1 89 66 23 94 28.1 0.167 21 0

0 137 40 35 168 43.1 2.288 33 1

5 116 74 . . 25.6 0.201 30 0

3 78 50 32 88 31 0.248 26 1

10 115 . . . 35.3 0.134 29 0

2 197 70 45 543 30.5 0.158 53 1

8 125 96 . . . 0.232 54 1

4 110 92 . . 37.6 0.191 30 0

10 168 74 . . 38 0.537 34 1

10 139 80 . . 27.1 1.441 57 0

1 189 60 23 846 30.1 0.398 59 1

5 166 72 19 175 25.8 0.587 51 1

7 100 . . . 30 0.484 32 1

0 118 84 47 230 45.8 0.551 31 1

7 107 74 . . 29.6 0.254 31 1

1 103 30 38 83 43.3 0.183 33 0

1 115 70 30 96 34.6 0.529 32 1

3 126 88 41 235 39.3 0.704 27 0

8 99 84 . . 35.4 0.388 50 0

7 196 90 . . 39.8 0.451 41 1

9 119 80 35 . 29 0.263 29 1

11 143 94 33 146 36.6 0.254 51 1

10 125 70 26 115 31.1 0.205 41 1

7 147 76 . . 39.4 0.257 43 1

1 97 66 15 140 23.2 0.487 22 0

13 145 82 19 110 22.2 0.245 57 0

5 117 92 . . 34.1 0.337 38 0

5 109 75 26 . 36 0.546 60 0

3 158 76 36 245 31.6 0.851 28 1

3 88 58 11 54 24.8 0.267 22 0

6 92 92 . . 19.9 0.188 28 0

10 122 78 31 . 27.6 0.512 45 0

4 103 60 33 192 24 0.966 33 0

11 138 76 . . 33.2 0.42 35 0

9 102 76 37 . 32.9 0.665 46 1

2 90 68 42 . 38.2 0.503 27 1

4 111 72 47 207 37.1 1.39 56 1

3 180 64 25 70 34 0.271 26 0

7 133 84 . . 40.2 0.696 37 0

7 106 92 18 . 22.7 0.235 48 0

9 171 110 24 240 45.4 0.721 54 1

7 159 64 . . 27.4 0.294 40 0

0 180 66 39 . 42 1.893 25 1

1 146 56 . . 29.7 0.564 29 0

2 71 70 27 . 28 0.586 22 0

7 103 66 32 . 39.1 0.344 31 1

7 105 . . . . 0.305 24 0

1 103 80 11 82 19.4 0.491 22 0

1 101 50 15 36 24.2 0.526 26 0

5 88 66 21 23 24.4 0.342 30 0

8 176 90 34 300 33.7 0.467 58 1

7 150 66 42 342 34.7 0.718 42 0

1 73 50 10 . 23 0.248 21 0

7 187 68 39 304 37.7 0.254 41 1

0 100 88 60 110 46.8 0.962 31 0

0 146 82 . . 40.5 1.781 44 0

0 105 64 41 142 41.5 0.173 22 0

2 84 . . . . 0.304 21 0

8 133 72 . . 32.9 0.27 39 1

5 44 62 . . 25 0.587 36 0

2 141 58 34 128 25.4 0.699 24 0

7 114 66 . . 32.8 0.258 42 1

5 99 74 27 . 29 0.203 32 0

0 109 88 30 . 32.5 0.855 38 1

2 109 92 . . 42.7 0.845 54 0

1 95 66 13 38 19.6 0.334 25 0

4 146 85 27 100 28.9 0.189 27 0

2 100 66 20 90 32.9 0.867 28 1

5 139 64 35 140 28.6 0.411 26 0

13 126 90 . . 43.4 0.583 42 1

4 129 86 20 270 35.1 0.231 23 0

1 79 75 30 . 32 0.396 22 0

1 . 48 20 . 24.7 0.14 22 0

7 62 78 . . 32.6 0.391 41 0

5 95 72 33 . 37.7 0.37 27 0

0 131 . . . 43.2 0.27 26 1

2 112 66 22 . 25 0.307 24 0

3 113 44 13 . 22.4 0.14 22 0

2 74 . . . . 0.102 22 0

7 83 78 26 71 29.3 0.767 36 0

0 101 65 28 . 24.6 0.237 22 0

5 137 108 . . 48.8 0.227 37 1

2 110 74 29 125 32.4 0.698 27 0

13 106 72 54 . 36.6 0.178 45 0

2 100 68 25 71 38.5 0.324 26 0

15 136 70 32 110 37.1 0.153 43 1

1 107 68 19 . 26.5 0.165 24 0

1 80 55 . . 19.1 0.258 21 0

4 123 80 15 176 32 0.443 34 0

7 81 78 40 48 46.7 0.261 42 0

4 134 72 . . 23.8 0.277 60 1

2 142 82 18 64 24.7 0.761 21 0

6 144 72 27 228 33.9 0.255 40 0

2 92 62 28 . 31.6 0.13 24 0

1 71 48 18 76 20.4 0.323 22 0

6 93 50 30 64 28.7 0.356 23 0

1 122 90 51 220 49.7 0.325 31 1

1 163 72 . . 39 1.222 33 1

1 151 60 . . 26.1 0.179 22 0

0 125 96 . . 22.5 0.262 21 0

1 81 72 18 40 26.6 0.283 24 0

2 85 65 . . 39.6 0.93 27 0

1 126 56 29 152 28.7 0.801 21 0

1 96 122 . . 22.4 0.207 27 0

4 144 58 28 140 29.5 0.287 37 0

3 83 58 31 18 34.3 0.336 25 0

0 95 85 25 36 37.4 0.247 24 1

3 171 72 33 135 33.3 0.199 24 1

8 155 62 26 495 34 0.543 46 1

1 89 76 34 37 31.2 0.192 23 0

4 76 62 . . 34 0.391 25 0

7 160 54 32 175 30.5 0.588 39 1

4 146 92 . . 31.2 0.539 61 1

5 124 74 . . 34 0.22 38 1

5 78 48 . . 33.7 0.654 25 0

4 97 60 23 . 28.2 0.443 22 0

4 99 76 15 51 23.2 0.223 21 0

0 162 76 56 100 53.2 0.759 25 1

6 111 64 39 . 34.2 0.26 24 0

2 107 74 30 100 33.6 0.404 23 0

5 132 80 . . 26.8 0.186 69 0

0 113 76 . . 33.3 0.278 23 1

1 88 30 42 99 55 0.496 26 1

3 120 70 30 135 42.9 0.452 30 0

1 118 58 36 94 33.3 0.261 23 0

1 117 88 24 145 34.5 0.403 40 1

0 105 84 . . 27.9 0.741 62 1

4 173 70 14 168 29.7 0.361 33 1

9 122 56 . . 33.3 1.114 33 1

3 170 64 37 225 34.5 0.356 30 1

8 84 74 31 . 38.3 0.457 39 0

2 96 68 13 49 21.1 0.647 26 0

2 125 60 20 140 33.8 0.088 31 0

0 100 70 26 50 30.8 0.597 21 0

0 93 60 25 92 28.7 0.532 22 0

0 129 80 . . 31.2 0.703 29 0

5 105 72 29 325 36.9 0.159 28 0

3 128 78 . . 21.1 0.268 55 0

5 106 82 30 . 39.5 0.286 38 0

2 108 52 26 63 32.5 0.318 22 0

10 108 66 . . 32.4 0.272 42 1

4 154 62 31 284 32.8 0.237 23 0

0 102 75 23 . . 0.572 21 0

9 57 80 37 . 32.8 0.096 41 0

2 106 64 35 119 30.5 1.4 34 0

5 147 78 . . 33.7 0.218 65 0

2 90 70 17 . 27.3 0.085 22 0

1 136 74 50 204 37.4 0.399 24 0

4 114 65 . . 21.9 0.432 37 0

9 156 86 28 155 34.3 1.189 42 1

1 153 82 42 485 40.6 0.687 23 0

8 188 78 . . 47.9 0.137 43 1

7 152 88 44 . 50 0.337 36 1

2 99 52 15 94 24.6 0.637 21 0

1 109 56 21 135 25.2 0.833 23 0

2 88 74 19 53 29 0.229 22 0

17 163 72 41 114 40.9 0.817 47 1

4 151 90 38 . 29.7 0.294 36 0

7 102 74 40 105 37.2 0.204 45 0

0 114 80 34 285 44.2 0.167 27 0

2 100 64 23 . 29.7 0.368 21 0

0 131 88 . . 31.6 0.743 32 1

6 104 74 18 156 29.9 0.722 41 1

3 148 66 25 . 32.5 0.256 22 0

4 120 68 . . 29.6 0.709 34 0

4 110 66 . . 31.9 0.471 29 0

3 111 90 12 78 28.4 0.495 29 0

6 102 82 . . 30.8 0.18 36 1

6 134 70 23 130 35.4 0.542 29 1

2 87 . 23 . 28.9 0.773 25 0

1 79 60 42 48 43.5 0.678 23 0

2 75 64 24 55 29.7 0.37 33 0

8 179 72 42 130 32.7 0.719 36 1

6 85 78 . . 31.2 0.382 42 0

0 129 110 46 130 67.1 0.319 26 1

5 143 78 . . 45 0.19 47 0

5 130 82 . . 39.1 0.956 37 1

6 87 80 . . 23.2 0.084 32 0

0 119 64 18 92 34.9 0.725 23 0

1 . 74 20 23 27.7 0.299 21 0

5 73 60 . . 26.8 0.268 27 0

4 141 74 . . 27.6 0.244 40 0

7 194 68 28 . 35.9 0.745 41 1

8 181 68 36 495 30.1 0.615 60 1

1 128 98 41 58 32 1.321 33 1

8 109 76 39 114 27.9 0.64 31 1

5 139 80 35 160 31.6 0.361 25 1

3 111 62 . . 22.6 0.142 21 0

9 123 70 44 94 33.1 0.374 40 0

7 159 66 . . 30.4 0.383 36 1

11 135 . . . 52.3 0.578 40 1

8 85 55 20 . 24.4 0.136 42 0

5 158 84 41 210 39.4 0.395 29 1

1 105 58 . . 24.3 0.187 21 0

3 107 62 13 48 22.9 0.678 23 1

4 109 64 44 99 34.8 0.905 26 1

4 148 60 27 318 30.9 0.15 29 1

0 113 80 16 . 31 0.874 21 0

1 138 82 . . 40.1 0.236 28 0

0 108 68 20 . 27.3 0.787 32 0

2 99 70 16 44 20.4 0.235 27 0

6 103 72 32 190 37.7 0.324 55 0

5 111 72 28 . 23.9 0.407 27 0

8 196 76 29 280 37.5 0.605 57 1

5 162 104 . . 37.7 0.151 52 1

1 96 64 27 87 33.2 0.289 21 0

7 184 84 33 . 35.5 0.355 41 1

2 81 60 22 . 27.7 0.29 25 0

0 147 85 54 . 42.8 0.375 24 0

7 179 95 31 . 34.2 0.164 60 0

0 140 65 26 130 42.6 0.431 24 1

9 112 82 32 175 34.2 0.26 36 1

12 151 70 40 271 41.8 0.742 38 1

5 109 62 41 129 35.8 0.514 25 1

6 125 68 30 120 30 0.464 32 0

5 85 74 22 . 29 1.224 32 1

5 112 66 . . 37.8 0.261 41 1

0 177 60 29 478 34.6 1.072 21 1

2 158 90 . . 31.6 0.805 66 1

7 119 . . . 25.2 0.209 37 0

7 142 60 33 190 28.8 0.687 61 0

1 100 66 15 56 23.6 0.666 26 0

1 87 78 27 32 34.6 0.101 22 0

0 101 76 . . 35.7 0.198 26 0

3 162 52 38 . 37.2 0.652 24 1

4 197 70 39 744 36.7 2.329 31 0

0 117 80 31 53 45.2 0.089 24 0

4 142 86 . . 44 0.645 22 1

6 134 80 37 370 46.2 0.238 46 1

1 79 80 25 37 25.4 0.583 22 0

4 122 68 . . 35 0.394 29 0

3 74 68 28 45 29.7 0.293 23 0

4 171 72 . . 43.6 0.479 26 1

7 181 84 21 192 35.9 0.586 51 1

0 179 90 27 . 44.1 0.686 23 1

9 164 84 21 . 30.8 0.831 32 1

0 104 76 . . 18.4 0.582 27 0

1 91 64 24 . 29.2 0.192 21 0

4 91 70 32 88 33.1 0.446 22 0

3 139 54 . . 25.6 0.402 22 1

6 119 50 22 176 27.1 1.318 33 1

2 146 76 35 194 38.2 0.329 29 0

9 184 85 15 . 30 1.213 49 1

10 122 68 . . 31.2 0.258 41 0

0 165 90 33 680 52.3 0.427 23 0

9 124 70 33 402 35.4 0.282 34 0

1 111 86 19 . 30.1 0.143 23 0

9 106 52 . . 31.2 0.38 42 0

2 129 84 . . 28 0.284 27 0

2 90 80 14 55 24.4 0.249 24 0

0 86 68 32 . 35.8 0.238 25 0

12 92 62 7 258 27.6 0.926 44 1

1 113 64 35 . 33.6 0.543 21 1

3 111 56 39 . 30.1 0.557 30 0

2 114 68 22 . 28.7 0.092 25 0

1 193 50 16 375 25.9 0.655 24 0

11 155 76 28 150 33.3 1.353 51 1

3 191 68 15 130 30.9 0.299 34 0

3 141 . . . 30 0.761 27 1

4 95 70 32 . 32.1 0.612 24 0

3 142 80 15 . 32.4 0.2 63 0

4 123 62 . . 32 0.226 35 1

5 96 74 18 67 33.6 0.997 43 0

0 138 . . . 36.3 0.933 25 1

2 128 64 42 . 40 1.101 24 0

0 102 52 . . 25.1 0.078 21 0

2 146 . . . 27.5 0.24 28 1

10 101 86 37 . 45.6 1.136 38 1

2 108 62 32 56 25.2 0.128 21 0

3 122 78 . . 23 0.254 40 0

1 71 78 50 45 33.2 0.422 21 0

13 106 70 . . 34.2 0.251 52 0

2 100 70 52 57 40.5 0.677 25 0

7 106 60 24 . 26.5 0.296 29 1

0 104 64 23 116 27.8 0.454 23 0

5 114 74 . . 24.9 0.744 57 0

2 108 62 10 278 25.3 0.881 22 0

0 146 70 . . 37.9 0.334 28 1

10 129 76 28 122 35.9 0.28 39 0

7 133 88 15 155 32.4 0.262 37 0

7 161 86 . . 30.4 0.165 47 1

2 108 80 . . 27 0.259 52 1

7 136 74 26 135 26 0.647 51 0

5 155 84 44 545 38.7 0.619 34 0

1 119 86 39 220 45.6 0.808 29 1

4 96 56 17 49 20.8 0.34 26 0

5 108 72 43 75 36.1 0.263 33 0

0 78 88 29 40 36.9 0.434 21 0

0 107 62 30 74 36.6 0.757 25 1

2 128 78 37 182 43.3 1.224 31 1

1 128 48 45 194 40.5 0.613 24 1

0 161 50 . . 21.9 0.254 65 0

6 151 62 31 120 35.5 0.692 28 0

2 146 70 38 360 28 0.337 29 1

0 126 84 29 215 30.7 0.52 24 0

14 100 78 25 184 36.6 0.412 46 1

8 112 72 . . 23.6 0.84 58 0

0 167 . . . 32.3 0.839 30 1

2 144 58 33 135 31.6 0.422 25 1

5 77 82 41 42 35.8 0.156 35 0

5 115 98 . . 52.9 0.209 28 1

3 150 76 . . 21 0.207 37 0

2 120 76 37 105 39.7 0.215 29 0

10 161 68 23 132 25.5 0.326 47 1

0 137 68 14 148 24.8 0.143 21 0

0 128 68 19 180 30.5 1.391 25 1

2 124 68 28 205 32.9 0.875 30 1

6 80 66 30 . 26.2 0.313 41 0

0 106 70 37 148 39.4 0.605 22 0

2 155 74 17 96 26.6 0.433 27 1

3 113 50 10 85 29.5 0.626 25 0

7 109 80 31 . 35.9 1.127 43 1

2 112 68 22 94 34.1 0.315 26 0

3 99 80 11 64 19.3 0.284 30 0

3 182 74 . . 30.5 0.345 29 1

3 115 66 39 140 38.1 0.15 28 0

6 194 78 . . 23.5 0.129 59 1

4 129 60 12 231 27.5 0.527 31 0

3 112 74 30 . 31.6 0.197 25 1

0 124 70 20 . 27.4 0.254 36 1

13 152 90 33 29 26.8 0.731 43 1

2 112 75 32 . 35.7 0.148 21 0

1 157 72 21 168 25.6 0.123 24 0

1 122 64 32 156 35.1 0.692 30 1

10 179 70 . . 35.1 0.2 37 0

2 102 86 36 120 45.5 0.127 23 1

6 105 70 32 68 30.8 0.122 37 0

8 118 72 19 . 23.1 1.476 46 0

2 87 58 16 52 32.7 0.166 25 0

1 180 . . . 43.3 0.282 41 1

12 106 80 . . 23.6 0.137 44 0

1 95 60 18 58 23.9 0.26 22 0

0 165 76 43 255 47.9 0.259 26 0

0 117 . . . 33.8 0.932 44 0

5 115 76 . . 31.2 0.343 44 1

9 152 78 34 171 34.2 0.893 33 1

7 178 84 . . 39.9 0.331 41 1

1 130 70 13 105 25.9 0.472 22 0

1 95 74 21 73 25.9 0.673 36 0

1 . 68 35 . 32 0.389 22 0

5 122 86 . . 34.7 0.29 33 0

8 95 72 . . 36.8 0.485 57 0

8 126 88 36 108 38.5 0.349 49 0

1 139 46 19 83 28.7 0.654 22 0

3 116 . . . 23.5 0.187 23 0

3 99 62 19 74 21.8 0.279 26 0

5 . 80 32 . 41 0.346 37 1

4 92 80 . . 42.2 0.237 29 0

4 137 84 . . 31.2 0.252 30 0

3 61 82 28 . 34.4 0.243 46 0

1 90 62 12 43 27.2 0.58 24 0

3 90 78 . . 42.7 0.559 21 0

9 165 88 . . 30.4 0.302 49 1

1 125 50 40 167 33.3 0.962 28 1

13 129 . . . 39.9 0.569 44 1

12 88 74 40 54 35.3 0.378 48 0

1 196 76 36 249 36.5 0.875 29 1

5 189 64 33 325 31.2 0.583 29 1

5 158 70 . . 29.8 0.207 63 0

5 103 108 37 . 39.2 0.305 65 0

4 146 78 . . 38.5 0.52 67 1

4 147 74 25 293 34.9 0.385 30 0

5 99 54 28 83 34 0.499 30 0

6 124 72 . . 27.6 0.368 29 1

0 101 64 17 . 21 0.252 21 0

3 81 86 16 66 27.5 0.306 22 0

1 133 102 28 140 32.8 0.234 45 1

3 173 82 48 465 38.4 2.137 25 1

0 118 64 23 89 . 1.731 21 0

0 84 64 22 66 35.8 0.545 21 0

2 105 58 40 94 34.9 0.225 25 0

2 122 52 43 158 36.2 0.816 28 0

12 140 82 43 325 39.2 0.528 58 1

0 98 82 15 84 25.2 0.299 22 0

1 87 60 37 75 37.2 0.509 22 0

4 156 75 . . 48.3 0.238 32 1

0 93 100 39 72 43.4 1.021 35 0

1 107 72 30 82 30.8 0.821 24 0

0 105 68 22 . 20 0.236 22 0

1 109 60 8 182 25.4 0.947 21 0

1 90 62 18 59 25.1 1.268 25 0

1 125 70 24 110 24.3 0.221 25 0

1 119 54 13 50 22.3 0.205 24 0

5 116 74 29 . 32.3 0.66 35 1

8 105 100 36 . 43.3 0.239 45 1

5 144 82 26 285 32 0.452 58 1

3 100 68 23 81 31.6 0.949 28 0

1 100 66 29 196 32 0.444 42 0

5 166 76 . . 45.7 0.34 27 1

1 131 64 14 415 23.7 0.389 21 0

4 116 72 12 87 22.1 0.463 37 0

4 158 78 . . 32.9 0.803 31 1

2 127 58 24 275 27.7 1.6 25 0

3 96 56 34 115 24.7 0.944 39 0

0 131 66 40 . 34.3 0.196 22 1

3 82 70 . . 21.1 0.389 25 0

3 193 70 31 . 34.9 0.241 25 1

4 95 64 . . 32 0.161 31 1

6 137 61 . . 24.2 0.151 55 0

5 136 84 41 88 35 0.286 35 1

9 72 78 25 . 31.6 0.28 38 0

5 168 64 . . 32.9 0.135 41 1

2 123 48 32 165 42.1 0.52 26 0

4 115 72 . . 28.9 0.376 46 1

0 101 62 . . 21.9 0.336 25 0

8 197 74 . . 25.9 1.191 39 1

1 172 68 49 579 42.4 0.702 28 1

6 102 90 39 . 35.7 0.674 28 0

1 112 72 30 176 34.4 0.528 25 0

1 143 84 23 310 42.4 1.076 22 0

1 143 74 22 61 26.2 0.256 21 0

0 138 60 35 167 34.6 0.534 21 1

3 173 84 33 474 35.7 0.258 22 1

1 97 68 21 . 27.2 1.095 22 0

4 144 82 32 . 38.5 0.554 37 1

1 83 68 . . 18.2 0.624 27 0

3 129 64 29 115 26.4 0.219 28 1

1 119 88 41 170 45.3 0.507 26 0

2 94 68 18 76 26 0.561 21 0

0 102 64 46 78 40.6 0.496 21 0

2 115 64 22 . 30.8 0.421 21 0

8 151 78 32 210 42.9 0.516 36 1

4 184 78 39 277 37 0.264 31 1

0 94 . . . . 0.256 25 0

1 181 64 30 180 34.1 0.328 38 1

0 135 94 46 145 40.6 0.284 26 0

1 95 82 25 180 35 0.233 43 1

2 99 . . . 22.2 0.108 23 0

3 89 74 16 85 30.4 0.551 38 0

1 80 74 11 60 30 0.527 22 0

2 139 75 . . 25.6 0.167 29 0

1 90 68 8 . 24.5 1.138 36 0

0 141 . . . 42.4 0.205 29 1

12 140 85 33 . 37.4 0.244 41 0

5 147 75 . . 29.9 0.434 28 0

1 97 70 15 . 18.2 0.147 21 0

6 107 88 . . 36.8 0.727 31 0

0 189 104 25 . 34.3 0.435 41 1

2 83 66 23 50 32.2 0.497 22 0

4 117 64 27 120 33.2 0.23 24 0

8 108 70 . . 30.5 0.955 33 1

4 117 62 12 . 29.7 0.38 30 1

0 180 78 63 14 59.4 2.42 25 1

1 100 72 12 70 25.3 0.658 28 0

0 95 80 45 92 36.5 0.33 26 0

0 104 64 37 64 33.6 0.51 22 1

0 120 74 18 63 30.5 0.285 26 0

1 82 64 13 95 21.2 0.415 23 0

2 134 70 . . 28.9 0.542 23 1

0 91 68 32 210 39.9 0.381 25 0

2 119 . . . 19.6 0.832 72 0

2 100 54 28 105 37.8 0.498 24 0

14 175 62 30 . 33.6 0.212 38 1

1 135 54 . . 26.7 0.687 62 0

5 86 68 28 71 30.2 0.364 24 0

10 148 84 48 237 37.6 1.001 51 1

9 134 74 33 60 25.9 0.46 81 0

9 120 72 22 56 20.8 0.733 48 0

1 71 62 . . 21.8 0.416 26 0

8 74 70 40 49 35.3 0.705 39 0

5 88 78 30 . 27.6 0.258 37 0

10 115 98 . . 24 1.022 34 0

0 124 56 13 105 21.8 0.452 21 0

0 74 52 10 36 27.8 0.269 22 0

0 97 64 36 100 36.8 0.6 25 0

8 120 . . . 30 0.183 38 1

6 154 78 41 140 46.1 0.571 27 0

1 144 82 40 . 41.3 0.607 28 0

0 137 70 38 . 33.2 0.17 22 0

0 119 66 27 . 38.8 0.259 22 0

7 136 90 . . 29.9 0.21 50 0

4 114 64 . . 28.9 0.126 24 0

0 137 84 27 . 27.3 0.231 59 0

2 105 80 45 191 33.7 0.711 29 1

7 114 76 17 110 23.8 0.466 31 0

8 126 74 38 75 25.9 0.162 39 0

4 132 86 31 . 28 0.419 63 0

3 158 70 30 328 35.5 0.344 35 1

0 123 88 37 . 35.2 0.197 29 0

4 85 58 22 49 27.8 0.306 28 0

0 84 82 31 125 38.2 0.233 23 0

0 145 . . . 44.2 0.63 31 1

0 135 68 42 250 42.3 0.365 24 1

1 139 62 41 480 40.7 0.536 21 0

0 173 78 32 265 46.5 1.159 58 0

4 99 72 17 . 25.6 0.294 28 0

8 194 80 . . 26.1 0.551 67 0

2 83 65 28 66 36.8 0.629 24 0

2 89 90 30 . 33.5 0.292 42 0

4 99 68 38 . 32.8 0.145 33 0

4 125 70 18 122 28.9 1.144 45 1

3 80 . . . . 0.174 22 0

6 166 74 . . 26.6 0.304 66 0

5 110 68 . . 26 0.292 30 0

2 81 72 15 76 30.1 0.547 25 0

7 195 70 33 145 25.1 0.163 55 1

6 154 74 32 193 29.3 0.839 39 0

2 117 90 19 71 25.2 0.313 21 0

3 84 72 32 . 37.2 0.267 28 0

6 . 68 41 . 39 0.727 41 1

7 94 64 25 79 33.3 0.738 41 0

3 96 78 39 . 37.3 0.238 40 0

10 75 82 . . 33.3 0.263 38 0

0 180 90 26 90 36.5 0.314 35 1

1 130 60 23 170 28.6 0.692 21 0

2 84 50 23 76 30.4 0.968 21 0

8 120 78 . . 25 0.409 64 0

12 84 72 31 . 29.7 0.297 46 1

0 139 62 17 210 22.1 0.207 21 0

9 91 68 . . 24.2 0.2 58 0

2 91 62 . . 27.3 0.525 22 0

3 99 54 19 86 25.6 0.154 24 0

3 163 70 18 105 31.6 0.268 28 1

9 145 88 34 165 30.3 0.771 53 1

7 125 86 . . 37.6 0.304 51 0

13 76 60 . . 32.8 0.18 41 0

6 129 90 7 326 19.6 0.582 60 0

2 68 70 32 66 25 0.187 25 0

3 124 80 33 130 33.2 0.305 26 0

6 114 . . . . 0.189 26 0

9 130 70 . . 34.2 0.652 45 1

3 125 58 . . 31.6 0.151 24 0

3 87 60 18 . 21.8 0.444 21 0

1 97 64 19 82 18.2 0.299 21 0

3 116 74 15 105 26.3 0.107 24 0

0 117 66 31 188 30.8 0.493 22 0

0 111 65 . . 24.6 0.66 31 0

2 122 60 18 106 29.8 0.717 22 0

0 107 76 . . 45.3 0.686 24 0

1 86 66 52 65 41.3 0.917 29 0

6 91 . . . 29.8 0.501 31 0

1 77 56 30 56 33.3 1.251 24 0

4 132 . . . 32.9 0.302 23 1

0 105 90 . . 29.6 0.197 46 0

0 57 60 . . 21.7 0.735 67 0

0 127 80 37 210 36.3 0.804 23 0

3 129 92 49 155 36.4 0.968 32 1

8 100 74 40 215 39.4 0.661 43 1

3 128 72 25 190 32.4 0.549 27 1

10 90 85 32 . 34.9 0.825 56 1

4 84 90 23 56 39.5 0.159 25 0

1 88 78 29 76 32 0.365 29 0

8 186 90 35 225 34.5 0.423 37 1

5 187 76 27 207 43.6 1.034 53 1

4 131 68 21 166 33.1 0.16 28 0

1 164 82 43 67 32.8 0.341 50 0

4 189 110 31 . 28.5 0.68 37 0

1 116 70 28 . 27.4 0.204 21 0

3 84 68 30 106 31.9 0.591 25 0

6 114 88 . . 27.8 0.247 66 0

1 88 62 24 44 29.9 0.422 23 0

1 84 64 23 115 36.9 0.471 28 0

7 124 70 33 215 25.5 0.161 37 0

1 97 70 40 . 38.1 0.218 30 0

8 110 76 . . 27.8 0.237 58 0

11 103 68 40 . 46.2 0.126 42 0

11 85 74 . . 30.1 0.3 35 0

6 125 76 . . 33.8 0.121 54 1

0 198 66 32 274 41.3 0.502 28 1

1 87 68 34 77 37.6 0.401 24 0

6 99 60 19 54 26.9 0.497 32 0

0 91 80 . . 32.4 0.601 27 0

2 95 54 14 88 26.1 0.748 22 0

1 99 72 30 18 38.6 0.412 21 0

6 92 62 32 126 32 0.085 46 0

4 154 72 29 126 31.3 0.338 37 0

0 121 66 30 165 34.3 0.203 33 1

3 78 70 . . 32.5 0.27 39 0

2 130 96 . . 22.6 0.268 21 0

3 111 58 31 44 29.5 0.43 22 0

2 98 60 17 120 34.7 0.198 22 0

1 143 86 30 330 30.1 0.892 23 0

1 119 44 47 63 35.5 0.28 25 0

6 108 44 20 130 24 0.813 35 0

2 118 80 . . 42.9 0.693 21 1

10 133 68 . . 27 0.245 36 0

2 197 70 99 . 34.7 0.575 62 1

0 151 90 46 . 42.1 0.371 21 1

6 109 60 27 . 25 0.206 27 0

12 121 78 17 . 26.5 0.259 62 0

8 100 76 . . 38.7 0.19 42 0

8 124 76 24 600 28.7 0.687 52 1

1 93 56 11 . 22.5 0.417 22 0

8 143 66 . . 34.9 0.129 41 1

6 103 66 . . 24.3 0.249 29 0

3 176 86 27 156 33.3 1.154 52 1

0 73 . . . 21.1 0.342 25 0

11 111 84 40 . 46.8 0.925 45 1

2 112 78 50 140 39.4 0.175 24 0

3 132 80 . . 34.4 0.402 44 1

2 82 52 22 115 28.5 1.699 25 0

6 123 72 45 230 33.6 0.733 34 0

0 188 82 14 185 32 0.682 22 1

0 67 76 . . 45.3 0.194 46 0

1 89 24 19 25 27.8 0.559 21 0

1 173 74 . . 36.8 0.088 38 1

1 109 38 18 120 23.1 0.407 26 0

1 108 88 19 . 27.1 0.4 24 0

6 96 . . . 23.7 0.19 28 0

1 124 74 36 . 27.8 0.1 30 0

7 150 78 29 126 35.2 0.692 54 1

4 183 . . . 28.4 0.212 36 1

1 124 60 32 . 35.8 0.514 21 0

1 181 78 42 293 40 1.258 22 1

1 92 62 25 41 19.5 0.482 25 0

0 152 82 39 272 41.5 0.27 27 0

1 111 62 13 182 24 0.138 23 0

3 106 54 21 158 30.9 0.292 24 0

3 174 58 22 194 32.9 0.593 36 1

7 168 88 42 321 38.2 0.787 40 1

6 105 80 28 . 32.5 0.878 26 0

11 138 74 26 144 36.1 0.557 50 1

3 106 72 . . 25.8 0.207 27 0

6 117 96 . . 28.7 0.157 30 0

2 68 62 13 15 20.1 0.257 23 0

9 112 82 24 . 28.2 1.282 50 1

0 119 . . . 32.4 0.141 24 1

2 112 86 42 160 38.4 0.246 28 0

2 92 76 20 . 24.2 1.698 28 0

6 183 94 . . 40.8 1.461 45 0

0 94 70 27 115 43.5 0.347 21 0

2 108 64 . . 30.8 0.158 21 0

4 90 88 47 54 37.7 0.362 29 0

0 125 68 . . 24.7 0.206 21 0

0 132 78 . . 32.4 0.393 21 0

5 128 80 . . 34.6 0.144 45 0

4 94 65 22 . 24.7 0.148 21 0

7 114 64 . . 27.4 0.732 34 1

0 102 78 40 90 34.5 0.238 24 0

2 111 60 . . 26.2 0.343 23 0

1 128 82 17 183 27.5 0.115 22 0

10 92 62 . . 25.9 0.167 31 0

13 104 72 . . 31.2 0.465 38 1

5 104 74 . . 28.8 0.153 48 0

2 94 76 18 66 31.6 0.649 23 0

7 97 76 32 91 40.9 0.871 32 1

1 100 74 12 46 19.5 0.149 28 0

0 102 86 17 105 29.3 0.695 27 0

4 128 70 . . 34.3 0.303 24 0

6 147 80 . . 29.5 0.178 50 1

4 90 . . . 28 0.61 31 0

3 103 72 30 152 27.6 0.73 27 0

2 157 74 35 440 39.4 0.134 30 0

1 167 74 17 144 23.4 0.447 33 1

0 179 50 36 159 37.8 0.455 22 1

11 136 84 35 130 28.3 0.26 42 1

0 107 60 25 . 26.4 0.133 23 0

1 91 54 25 100 25.2 0.234 23 0

1 117 60 23 106 33.8 0.466 27 0

5 123 74 40 77 34.1 0.269 28 0

2 120 54 . . 26.8 0.455 27 0

1 106 70 28 135 34.2 0.142 22 0

2 155 52 27 540 38.7 0.24 25 1

2 101 58 35 90 21.8 0.155 22 0

1 120 80 48 200 38.9 1.162 41 0

11 127 106 . . 39 0.19 51 0

3 80 82 31 70 34.2 1.292 27 1

10 162 84 . . 27.7 0.182 54 0

1 199 76 43 . 42.9 1.394 22 1

8 167 106 46 231 37.6 0.165 43 1

9 145 80 46 130 37.9 0.637 40 1

6 115 60 39 . 33.7 0.245 40 1

1 112 80 45 132 34.8 0.217 24 0

4 145 82 18 . 32.5 0.235 70 1

10 111 70 27 . 27.5 0.141 40 1

6 98 58 33 190 34 0.43 43 0

9 154 78 30 100 30.9 0.164 45 0

6 165 68 26 168 33.6 0.631 49 0

1 99 58 10 . 25.4 0.551 21 0

10 68 106 23 49 35.5 0.285 47 0

3 123 100 35 240 57.3 0.88 22 0

8 91 82 . . 35.6 0.587 68 0

6 195 70 . . 30.9 0.328 31 1

9 156 86 . . 24.8 0.23 53 1

0 93 60 . . 35.3 0.263 25 0

3 121 52 . . 36 0.127 25 1

2 101 58 17 265 24.2 0.614 23 0

2 56 56 28 45 24.2 0.332 22 0

0 162 76 36 . 49.6 0.364 26 1

0 95 64 39 105 44.6 0.366 22 0

4 125 80 . . 32.3 0.536 27 1

5 136 82 . . . 0.64 69 0

2 129 74 26 205 33.2 0.591 25 0

3 130 64 . . 23.1 0.314 22 0

1 107 50 19 . 28.3 0.181 29 0

1 140 74 26 180 24.1 0.828 23 0

1 144 82 46 180 46.1 0.335 46 1

8 107 80 . . 24.6 0.856 34 0

13 158 114 . . 42.3 0.257 44 1

2 121 70 32 95 39.1 0.886 23 0

7 129 68 49 125 38.5 0.439 43 1

2 90 60 . . 23.5 0.191 25 0

7 142 90 24 480 30.4 0.128 43 1

3 169 74 19 125 29.9 0.268 31 1

0 99 . . . 25 0.253 22 0

4 127 88 11 155 34.5 0.598 28 0

4 118 70 . . 44.5 0.904 26 0

2 122 76 27 200 35.9 0.483 26 0

6 125 78 31 . 27.6 0.565 49 1

1 168 88 29 . 35 0.905 52 1

2 129 . . . 38.5 0.304 41 0

4 110 76 20 100 28.4 0.118 27 0

6 80 80 36 . 39.8 0.177 28 0

10 115 . . . . 0.261 30 1

2 127 46 21 335 34.4 0.176 22 0

9 164 78 . . 32.8 0.148 45 1

2 93 64 32 160 38 0.674 23 1

3 158 64 13 387 31.2 0.295 24 0

5 126 78 27 22 29.6 0.439 40 0

10 129 62 36 . 41.2 0.441 38 1

0 134 58 20 291 26.4 0.352 21 0

3 102 74 . . 29.5 0.121 32 0

7 187 50 33 392 33.9 0.826 34 1

3 173 78 39 185 33.8 0.97 31 1

10 94 72 18 . 23.1 0.595 56 0

1 108 60 46 178 35.5 0.415 24 0

5 97 76 27 . 35.6 0.378 52 1

4 83 86 19 . 29.3 0.317 34 0

1 114 66 36 200 38.1 0.289 21 0

1 149 68 29 127 29.3 0.349 42 1

5 117 86 30 105 39.1 0.251 42 0

1 111 94 . . 32.8 0.265 45 0

4 112 78 40 . 39.4 0.236 38 0

1 116 78 29 180 36.1 0.496 25 0

0 141 84 26 . 32.4 0.433 22 0

2 175 88 . . 22.9 0.326 22 0

2 92 52 . . 30.1 0.141 22 0

3 130 78 23 79 28.4 0.323 34 1

8 120 86 . . 28.4 0.259 22 1

2 174 88 37 120 44.5 0.646 24 1

2 106 56 27 165 29 0.426 22 0

2 105 75 . . 23.3 0.56 53 0

4 95 60 32 . 35.4 0.284 28 0

0 126 86 27 120 27.4 0.515 21 0

8 65 72 23 . 32 0.6 42 0

2 99 60 17 160 36.6 0.453 21 0

1 102 74 . . 39.5 0.293 42 1

11 120 80 37 150 42.3 0.785 48 1

3 102 44 20 94 30.8 0.4 26 0

1 109 58 18 116 28.5 0.219 22 0

9 140 94 . . 32.7 0.734 45 1

13 153 88 37 140 40.6 1.174 39 0

12 100 84 33 105 30 0.488 46 0

1 147 94 41 . 49.3 0.358 27 1

1 81 74 41 57 46.3 1.096 32 0

3 187 70 22 200 36.4 0.408 36 1

6 162 62 . . 24.3 0.178 50 1

4 136 70 . . 31.2 1.182 22 1

1 121 78 39 74 39 0.261 28 0

3 108 62 24 . 26 0.223 25 0

0 181 88 44 510 43.3 0.222 26 1

8 154 78 32 . 32.4 0.443 45 1

1 128 88 39 110 36.5 1.057 37 1

7 137 90 41 . 32 0.391 39 0

0 123 72 . . 36.3 0.258 52 1

1 106 76 . . 37.5 0.197 26 0

6 190 92 . . 35.5 0.278 66 1

2 88 58 26 16 28.4 0.766 22 0

9 170 74 31 . 44 0.403 43 1

9 89 62 . . 22.5 0.142 33 0

10 101 76 48 180 32.9 0.171 63 0

2 122 70 27 . 36.8 0.34 27 0

5 121 72 23 112 26.2 0.245 30 0

1 126 60 . . 30.1 0.349 47 1

1 93 70 31 . 30.4 0.315 23 0

;

run;

proc print data=p1;

run;

**/\*Missing values substitution - Mean\*/**

proc stdize data=work.p1 reponly method=mean out=p2 outstat=mean;

var Glucose BloodPressure SkinThickness Insulin BMI;

run;

proc print data=p2;

run;

**/\*Descriptive analytics\*/**

proc means data=work.p2;

var Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age;

run;

ods listing;

ods graphics off;

proc univariate data=work.p2 plots;

Histogram Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age / Normal;

VAR Pregnancies Glucose BloodPressure SkinThickness Insulin BMI DiabetesPedigreeFunction Age;

run;

**/\*Transform variables\*/**

DATA work.p3;

set work.p2;

SkinThickness1=log10(SkinThickness);

Insulin1=log10(Insulin);

DiabetesPedigreeFunction1=log10(DiabetesPedigreeFunction);

run;

DATA p4(DROP = SkinThickness Insulin DiabetesPedigreeFunction);

SET work.p3;

RUN;

**/\*Descriptive analytics\*/**

ods listing;

ods graphics off;

proc univariate data=p4 plots;

Histogram SkinThickness1 Insulin1 DiabetesPedigreeFunction1 / Normal;

VAR SkinThickness1 Insulin1 DiabetesPedigreeFunction1 ;

run;

**/\*Correlation analysis\*/**

proc corr data=p4;

var Pregnancies Glucose BloodPressure SkinThickness1 Insulin1 BMI DiabetesPedigreeFunction1 Age Outcome;

run;

**/\*Plotting the data\*/**

proc gplot data=p4;

plot Insulin1\*Glucose;

run;

proc gplot data=p4;

plot SkinThickness1\*BMI;

run;

proc gplot data=p4;

plot Age\*Pregnancies;

run;

**/\*Simple Linear Regression\*/**

proc reg data=p4;

model Insulin1=Glucose;

run;

proc reg data=p4;

model BMI=SkinThickness1;

run;

## 

## Conclusion

Regression is a good tool to predict values when it comes to simple datasets. Everytime, machine learning is not required to accomplish a task. The capability to predict diabetes early, assumes a vital role for the patient's appropriate treatment procedure. The results have shown that there is a strong association of BMI and glucose with diabetes.

Now, coming back to our business question. We can firmly say that the Hypothesis is true as we accept H0, we can reject H1. Glucose is a better predictor of diabetes than the other variables in the dataset.

## References

1. A model for early prediction of diabetes

Informatics in Medicine Unlocked

Volume 16, 2019, 100204

<https://doi.org/10.1016/j.imu.2019.100204>

1. Analysis of diabetes mellitus for early prediction using optimal features selection

Journal of Big Data volume 6

Article number: 13 (2019)

<https://doi.org/10.1186/s40537-019-0175-6>

1. Prediction of diabetes mellitus in Chinese adults

Meinian Institute of Health, PKUHSC Meinian Public Health Research Institute

& Peking University Health Science Center (PKUHSC)

OP-0293

<https://www.morressier.com/article/prediction-1year-incident-diabetes-mellitus-chinese-adults-prospective-cohort-study-1-million-people/5d9b622cea541d6ca8493bbb>

1. Diabetes in Older Adults

American Diabetes Association

Diabetes Care 2012 Dec; 35(12): 2650-2664.

<https://doi.org/10.2337/dc12-1801>

1. Pregnancy and Diabetes: How Women Handle the Challenges

[J Perinat Educ.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1595250/) 2005 Summer; 14(3): 23–32.

doi: [10.1624/105812405X57552](https://dx.doi.org/10.1624%2F105812405X57552)