BENIN JOURNAL OF STATISTICS
ISSN 2682-5767
Vol. 5, pp. 123– 136 (2022)

Implementation of Yeo-Johnson Transformation in Quantile Regression

M. T. Nwakuya[1 *], **and I. V. Anyaogu**[2]

[1,2]*Department of Mathematics and Statistics, University of Port Harcourt, Nigeria.*

(*Received: 23 November 2021; Accepted: 18 April 2022*)

Abstract. Power transformations offer a remedy for retaining the symmetry, homoscedasticity and normality assumptions that is typical of regression analysis. The Box-Cox power transformation is defined for a strictly positive response variable in regression model. In this work the Yeo-Johnson quantile regression model is investigated as a modification of Box-Cox quantile regression model to accommodate both negative and positive values of the response variable, while implementing the two-step method of estimation of the transformation parameter. The quantile regression without any transformation was also considered. The study examines the adequacy of the Yeo-Johnson quantile regression over the quantile regression. The results demonstrate that the Yeo-Johnson quantile regression model performed better that the quantile regression model at all quantiles considered in this paper. Comparison of both models was done using the Akaike Information criteria (AIC) and the mean square error (MSE) as the comparison criteria.

Keywords: Yeo-Johnson transformation, Box-Cox transformation, Yeo-Johnson Quantile Regression, Box-Cox quantile Regression and Transformation parameter.

Published by: Department of Statistics, University of Benin, Nigeria

1. Introduction

Data transformation is an essential tool used in statistical data analysis. Most often, raw data to be used for a given statistical analysis possess some features that violates basic assumption for which the inferential statistics to be carried out (e.g normality and linearity in the case of regression) are built upon. Their distribution is far from normal and often skewed. In order to handle such problem, data transformation is often used to preprocess the variable(s) in order to bring them closer to normality with linearity features.
The BoxCox transformation method is often used to handle this type of issue but its drawback is the inability to handle negative response values. Yeo and Johnson (2000) proposed the Yeo-Johnson transformation as an extension of the

* Corresponding author. Email: mauren.nwakuya@uniport.edu.ng

BoxCox transformation to handle both positive and negative response values. Powell (1991) applied the BoxCox quantile regression by adopting the maximum likelihood method as a means of estimating the transformation parameter. As an improvement to that, this paper aims to extend the BoxCox quantile regression analysis to accommodate both negative and positive response values through the implementation of the Yeo-Johnson transformation in a quantile regression analysis.

This paper adopted a different method of estimating the transformation parameter known as the Chamberlain Buchinsky Two Stage method (CBTS). The choice of CBTS is to allow the equi-variance property of the quantiles to be exploited, which to the best of our knowledge has not been applied in the Yeo-Johnson transformation and also to provide a method that is robust to outliers, unlike the maximum likelihood method that is very sensitive to outliers.

Varouchakis (2021) presented alternative methods for data transformation and revisited the applicability of a modified version of the well-known Box-Cox transformation technique. He discovered that this modified method competes successfully in data transformation and improves normalization of negative sign data or fluctuations, significantly. Othman and Ali (2021) proposed an algorithm to transform the positive original responses using Box-Cox transformation and the stationary responses using the Yeo-Johnson transformation. The results show that the functional time series predictions were more accurate using the Yeo-Johnson transformation on the stationary data set. Pallab (2021) proposed a semi-parametric estimation method, Box-Cox power transformation unconditioned quantile regression to estimate the impact of changes in the distribution of the response variable on the unconditional quantile of the outcome variable.

The results reveal that declining unionization can explain approximately 10% of the decline in the 50/10 wage gap in 1990-2000 and 23% in 2000-2010. Raymaekers and Rousseeuw (2021), in their work applied the Box-Cox and Yeo-Johnson transformation to achieve central normality using a reweighted maximum likelihood method that is robust to outliers to estimate the transformation parameter. The method compares favorably to existing techniques in an extensive simulation study and on real data. Yilin *et al.* (2020) proposed a robust alternative to simple linear regression for analysis of change in two repeated measurements of a continuous outcome, by applying a cprobit model that incorporates the Box-Cox transformation, comparing the cprobit and REM with and without transformation, they found out that the cprobit model was robust to skewed subject-specific intercept term. Onyegbuchulam *et al.* (2019), in their work applied different power transformations on a Weilbull distributed data in quantile regression and the results show that the square root of the response transformation gave a better result.

Nugroho (2019) proposed a new class of GARC models by applying Yeo-Johnson transformation to return series. The results show that the proposed model outperformed the initial model in fitting ten different international stock indices. Atkinson *et al.* (2021), in their work described an extended Yeo-Johnson transformation that allows positive and negative response values to have two different transformation parameters. This research paper provides an extension of BoxCox quantile regression that is called the Yeo-Johnson quantile regression. This Yeo-Johnson quantile regression method provides a means of transforming response variable with both negative and positive data points with outliers using a transformation parameter estimator known as CBTS that is robust to outliers, in a quantile regression analysis. The analysis was done using Qtools package in R.

2. Materials and methods

2.1 *Yeo-Johnson Transformation:*

Atkinson *et al.* (2021) stated that in choosing a transformation for a given data set, our choice of transformation should provide simple and more revealing analyses that lead to valid inferences. In order to transform positive variable to give it a more normal distribution one often resorts to power transformation Raymaekers and Rousseeuw, (2021). Box-Cox power transformation family is often chosen from the family of parametric transformations and it is equivalent to the power transformation for univariate data transformation. This transformation is defined as:

$$h_\lambda(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(y) & \text{if } \lambda = 0 \end{cases} \tag{1}$$

Where y stands for only values of positive observations or data points, which is then transformed to $h_\lambda(y)$ using λ as the transformation parameter. This parameter can be estimated using the method of maximum likelihood as proposed by Box and Cox (1964). The Box-Cox transformation possesses special features which makes it perhaps more preferable as the chosen response transformation method and or for transforming set of exogenous variables towards normality. At $\lambda = 1$, the Box-Cox transformation corresponds to no transformation as we can see in the (1) above.
When $\lambda = \frac{1}{2}$, this transformation will resolve to the square root transformation. When $\lambda = 0$, the logarithm transformation is chosen and with $\lambda = -1$ the inverse or reciprocal transformation is opted for. However, a major limitation associated with the family of Box-Cox transformations is that it can only be applied to positive response data points or observations. Yeo and Johnson (2000) proposed an alternative family of transformations that addresses this limitation and can deal with both positive and negative response data points. Another good feature of the Johnson's transformation is that it has many properties of the Box-Cox power family and it can simply be seen as an extension of the Box-Cox. The Yeo–Johnson transformation is described by the function below:

$$G_\lambda(y) = \begin{cases} \frac{(1+y)^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \text{ and } y \geq 0 \\ \log(y+1) & \text{if } \lambda = 0 \text{ and } y \geq 0 \\ \frac{-[(-y+1)^{2-\lambda} - 1]}{(2-\lambda)} & \text{if } \lambda \neq 2 \text{ and } y < 0 \\ -\log(-y+1) & \text{if } \lambda = 2 \text{ and } y < 0 \end{cases} \tag{2}$$

This transformation successfully deals with both positive and negative values of

the response. For y strictly positive numbers, the Yeo-Johnson's transformation resolves to the generalized Box-Cox transformation of $(y + 1)$ and $(-y + 1)$ but for y strictly negative numbers we now have it raised to the power of $2 - \lambda$. (2) shows the Yeo-Johnson transformation for a range of values of λ. In these transformations $\lambda = 1$ yields a linear relation while the right tail of the distribution of the response is compressed when the transformation is done with $\lambda < 1$ but it expands the left tail with $\lambda > 1$ making the transformation suitable for transforming both right-skewed and left-skewed distributions towards symmetry.
The traditional transformation parameter estimation by Yeo-Johnson and Box-Cox is usually done using the maximum likelihood method. The draw back in using MLE is the fact that it is very sensitive to outliers in the dataset; just one outlier can have arbitrary large effect on the estimate, hence the quest for an approach that is robust to outliers. This led to the idea of integrating the Yeo-Johnson transformation into the quantile regression, since the quantile regression is robust to outliers.

2.2 *Quantile Regression:*

Unlike regular linear regression which uses the method of least squares to estimate the conditional mean of the target object across different values of the predictor variable. Quantile regression on the other hand estimates the conditional quantiles of the target object across different values of the predictor variable. Quantile regression can be seen as an extension of linear regression that is used when the conditions of linear regression are not met, which includes the normality assumption and presence of outliers. In other words quantile regression can be used even when the response variable is not normal. The Traditional mean regression model is given by;

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, \ldots, n \tag{3}$$

Where p is equal to the number of predictor variable in the equation and n is the number of data points. The best linear regression line is found by minimizing the mean square error, which is given by;

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}))^2 \tag{4}$$

Dye (2020).
Quantile regression model takes similar structure to that of the linear regression model; the conditional quantile regression model equation for the τ^{th} quantile is given by;

$$Q_\tau(y_i) = \beta_0(\tau) + \beta_1(\tau) x_{i1} + \cdots + \beta_p(\tau) x_{ip}, \quad i = 1, \ldots, n \tag{5}$$

Koenker and Bassett (1978)
Where the quantile $\tau \in (0, 1)$, we can observe that the β's are not constant like in the traditional linear regression rather the β's are seen to be functions that

depend on the quantile (τ). Unlike the linear regression, the quantile regression even when a transformation does satisfy the assumptions for standard regression, the inverse transformation can be applied to the predicted quantiles of the transformed response;

$$Q_\tau(Y/X) = h^{-1}\left(Q_\tau(h(Y)/X)\right) \tag{6}$$

Hence $Q_\tau(Y/X)$ is preserved under transformation because of the equi-variance property of the quantile regression. But the inverse transformation of the linear regression does not predict the mean of the response when applied to the predicted mean of the transformed response;

$$E(Y/X) \neq h^{-1}\left(E(h(Y)/X)\right) \tag{7}$$

Quantile Regression helps to describe the entire conditional distribution of the response. It gives a principled alternative to the usual practice of stabilizing the variance of heteroscedastic data with a monotone transformation $h(Y)$ before fitting a standard regression model.
Estimation in Quantile Regression: Given the conditional quantile function in (5), estimation of $\beta_j(\tau)$ is obtained by solving the minimization problem;

$$\min_{\beta_j(\tau)} \sum_{i=1}^{n} \rho_\tau \left(y_i - \beta_0(\tau) - \sum_{i=1}^{p} X_{ij}\beta_j(\tau) \right),$$

where (check function)$\rho_\tau = \tau \max(r, 0) + (1 - \tau)\max(-r, 0)$
For each quantile level τ, the solution to the minimization problem yields distinct sets of regression coefficients at different levels of the quantile.

2.3 *Implementing Yeo-Johnson Transformation in Quantile Regression:*

To implement Yeo-Johnson transformation in quantile regression, the question of estimation of transformation parameter arises. In other to avoid the drawback of the maximum likelihood estimation, in this work we adopt the two step method (CBTS) by Chamberlain (1994) and Buchinsky (1995). The procedure is as follows:
First estimate $\beta_\tau(\lambda)$ conditional on λ by solving the minimization problem;

$$\hat{\beta}_\tau = \arg\min_{\beta} n^{-1} \sum_{i=1}^{n} \rho_\tau \left(y_{\lambda i} - x'\beta \right) \tag{8}$$

Secondly estimate λ_τ by solving the minimization problem;

$$\hat{\lambda}_\tau = \min_{\lambda} n^{-1} \sum_{i=1}^{n} \rho_\tau \left(y_i - \left(\lambda x'\hat{\beta}_\tau(\lambda) + 1 \right)^{-\frac{1}{\lambda}} \right) \tag{9}$$

Conditioned on the assumption that $\left(\lambda x'\hat{\beta}_{\tau}(\lambda)+1\right) > 0$, this procedure was implemented in R by Geraci (2016). In this research work, two quantile regression analysis were performed; one implementing the Yeo-Johnson transformation and the other without any form of transformation. The analysis considered 0.05, 0.25, 0.45, 0.5, 0.55, 0.75 and 0.95 quantiles. Results of both analyses were investigated and the resulting models were compared based on their mean square error (MSE) and Akaike Information criteria (AIC).

3. Results and Discussion

3.1 *Data Sets and Description:*

The data for this work is iris data obtained from R data base. The iris data set is a built-in data set in R that contains measurements on 4 different attributes (in centimeters) for 50 flowers from 3 different species. Also two sets of data were simulated with sample sizes 20 and 500 and a real life data from a survey carried out by team of researchers from Department of Physiology, University of Uyo. In the survey four hundred and ninety nine (499) adults between the age, 18- 65 years were recruited in the study.
Their ages were obtained, their Body Mass Indexes were calculated, and their pulse rate as well as their Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) were measured, (see Appendix A). The whole data were collated to form the data set for this research work. The data was analyzed using a new created function in QTOOLS package in R.

3.2 *Results and Discussion*

Table 1: Results for AIC and MSE (Iris Data)

τ	**Comparison Criteria**	**Untransformed**	**Yeo-Johnson Transformed**
	AIC	155.9925	**-169.4475**
0.05	MSE	0.3084184	**0.306363**
	AIC	114.3259	**-272.955**
0.25	MSE	0.1451838	**0.1436569**
	AIC	103.8186	**-344.3926**
0.45	MSE	0.0968846	**0.095437**
	AIC	100.965	**-344.655**
0.50	MSE	0.096485	**0.0952706**
	AIC	99.14373	**-341.9881**
0.55	MSE	0.100418	**0.0969796**
	AIC	113.5707	**-278.2643**
0.75	MSE	0.1522861	**0.148313**
	AIC	166.2496	**-138.9423**
0.95	MSE	0.3734668	**0.373455**

Table 1 above shows us that the Yeo-Johnson transformed data produced better results based on the AIC and Mean square error on all quantiles considered.

Table 2: Results for AIC and MSE (Simulated data for sample size 20)

τ	Comparison Criteria	Untransformed	Yeo-Johnson Transformed
	AIC	58.34604	**13.11372**
0.05	MSE	2.169321	**2.129071**
	AIC	54.20241	**7.885194**
0.25	MSE	1.017053	**0.9942761**
	AIC	51.11198	**-4.670824**
0.45	MSE	0.5315598	**0.5307091**
	AIC	50.39603	**-4.145197**
0.50	MSE	0.5435761	**0.5448418**
	AIC	49.64338	**-3.918325**
0.55	MSE	0.5509591	**0.5510574**
	AIC	50.35257	**5.49153**
0.75	MSE	0.9000754	**0.882133**
	AIC	44.129	**14.62694**
0.95	MSE	1.42068	**1.392843**

Table 2 above also shows us that the Yeo-Johnson transformed data produced better results based on the AIC and Mean square error on all quantiles considered.

Table 3: Results for AIC and MSE (Simulated data for sample size 500)

.

τ	Comparison Criteria	Untransformed	Yeo-Johnson Transformed
	AIC	1515.606	**446.1266**
0.05	MSE	2.169321	**2.401883**
	AIC	1302.12	**8.633267**
0.25	MSE	1.001341	**1.001267**
	AIC	1265.772	**-169.4629**
0.45	MSE	0.7013788	**0.7012255**
	AIC	1265.792	**-180.62337**
0.50	MSE	0.6857677	**0.6857469**
	AIC	1269.306	**-176.2**
0.55	MSE	0.6919367	**0.6918404**
	AIC	1353.177	**-17.54112**
0.75	MSE	0.9477509	**0.9502005**
	AIC	1666.518	**488.8919**
0.95	MSE	2.617625	**2.61636**

Table 3 above also shows us that the Yeo-Johnson transformed data produced better results based on the AIC and Mean square error on all quantiles considered.

Table 4: Results for AIC and MSE (Blood Pressure data)

τ	**Comparison Criteria**	**Untransformed**	**Yeo-Johnson Transformed**
	AIC	2899.314	**2281.376**
0.05	MSE	1269.852	**1268.808**
	AIC	2830.232	**2088.578**
0.25	MSE	700.3548	**694.157**
	AIC	2839.023	**1944.768**
0.45	MSE	438.975	**431.627**
	AIC	2845.402	**1933.0567**
0.50	MSE	425.609	**425.707**
	AIC	2854.415	**1929.096**
0.55	MSE	425.029	**420.390**
	AIC	2926.627	**2030.243**
0.75	MSE	574.488	**571.487**
	AIC	3098.348	**2400.318**
0.95	MSE	1815.874	**1805.874**

Table 4 above also shows us that the Yeo-Johnson transformed data produced better results based on the AIC and Mean square error on all quantiles considered.

4. Conclusion

This paper investigated the implementation of Yeo-Johnson transformation in quantile regression while using the two-step method of estimation proposed by Chamberlain (1994) and Buchinsky (1995) to estimate the transformation parameter. This transformed quantile regression method was then compared to the quantile regression without any transformation, using Iris data which is an inbuilt R data, a real data on blood pressure and two simulated data sets of sample sizes 20 and 500. The comparison was done using AIC and MSE as the comparison criteria. The results of this analysis revealed that at all quantiles (0.05, 0.25, 0.45, 0.5, 0.55, 0.75 and 0.95) considered, the Yeo-Johnson transformed quantile regression performed better than the quantile regression without transformation. The analysis was done in R using Qtools package. This study has been able to show:

- The estimation of the transformation parameter of Yeo-Johnson transformation using the CBTS method.
- That the mid-quantiles (0.45, 0.50 and 0.55) exhibit approximately same margin of error considering their MSE's, this confirms the literature that the mean regression and the median regression produce approximately same results.
- That the Yeo-Johnson transformed data sets produced better results in comparison to the untransformed data.
- That the tail ends of the distribution (0.05 and 0.95 quantiles) of the response variable have higher MSE.
- That the implemented CBTS method is robust to outliers, which gives it

an edge over existing methods like the MLE.

One of the limitations of this work is that, the method uses same transformation parameter for both the negative and the positive response value, but further research can be done to investigate the use of two different transformation parameters for the negative and positive data points in a quantile regression analysis using the CBTS estimation method. This work is limited to continuous response variable and investigations can be carried out to modify it to accommodate bounded response variable and count data.

Acknowledgement

We would like to express our deep gratitude to the anonymous reviewers of this paper, for their thorough commitment in ensuring that this paper gets improved on and published. Our thanks also go to the editorial team for their consideration of the publication of this paper.

References

Atkinson A. C., Riani M. and Corbellini, A. (2021). The Box-Cox Transformation: Review and Extensions. Statist. Sci., 36(2), 239-255, https://doi.org/10.1214/20-STS77-8

Box, G. and Cox, D. (1964). An Analysis of Transformation. Journal of the Royal Statistical Society B, 26(2), 211-252.

Buchinsky, M. (1995). Quantile regression, Box-Cox transformation model, and the U.S. wage structure, 1963-1987. Journal of Econometrics, 65(1), 109-154.

Chamberlain, G. (1994). Quantile Regression, Censoring, and the Structure of Wages. In: Sims, C. (ed.), Advances in Econometrics: Sixth World Congress, 1, Econometric Society Monograph.

Dye, S. (2020). Quantile Regression; www.https//towardsdatascience.com/quantile-regression-ff2343c4a03. Accessed 23rd Nov. 2021.

Gercia, M. (2016). Qtools: A Collection of Models and Tools for Quantile Inference. The R Journal, 8(2), 117-138.

Koenker, R. and Bassett, G. (1978). “Regression Quantiles,” Econometrica, 46, 33-50.

Nugroho, D. B. (2019). GARCH (1,1) Model with the Yeo-Johnson Transformed Returns. Journal of Physics: Conference Series, vol 1320 012013.

Onyegbuchulam, B. O., Nwakuya, M. T., Nwabueze, J. C. and Otu, A. O. (2019). Choice of Appropriate Power Transformation of Skewed Distribution for Quantile Regression Model. American Journal of Applied Mathematics, 7(3), 105-111.

Othman, S. A. and Ali, M. H. T. (2021). Improvement of the Non-parametric Estimation of Functional Stationary Time Series Using Yeo-Johnson Transformation with Application to Temperature Curve. Hindawi Advances in Mathematical Physics, 2021, 1-6, https://doi.org/10.1155/2021/6676400.

Pallab, K. G. (2021). Box-Cox Power Transformation Unconditional Quantile Regression with an Application on Wage Inequality. Journal of Applied Statistics, 48(16), 3086-3101.

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. In non-parametric And Semiparametric methods in Econometrics (W. Barnett, J. Powell, G. Tauchen, eds.), 357-384, Cambridge University Press, New York.

Raymaekers, J., Rousseeuw, P.J. (2021). Transforming variables to central normality. Machine Learning, https://doi.org/10.1007/s10994-021-05960-5

Varouchakis, E. A. (2021). Gaussian Transformation Methods for Spatial Data.Geosciences, 11(5), 196.

Yeo, I. and Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. Biometrika, 87(4).

Yilin, N., Nathelie, C. S., Joo, H. P., Kao, S. L., Yuan, N., Haokhoo, E. Y., Chinelee, S., Tai, E. S., Hartman, M., Reilly, M. and Tan, C. S. (2020). Robust Estimation of The Effect of an Exposure on The Change in a Continuous Outcome. BMC Med Res Methodol., 20, 145, https://doi.org/10.1186/s12874-020-01027-6.

Appendix

Blood Pressure Data

S/N	Age	BMI	PR	SBP	DBP	S/N	Age	BMI	PR	SBP	DBP
1	29	26.1	87	128	72	51	16	25.4	70	145	70
2	20	25.9	79	115	60	52	18	19.6	83	113	61
3	19	26.1	75	102	62	53	16	20.6	90	105	54
4	19	21.1	79	129	78	54	17	19.7	83	105	43
5	23	25.9	81	124	78	55	18	19.4	82	110	51
6	18	22.3	79	126	43	56	18	21.8	100	160	77
7	21	23	72	126	64	57	18	20.1	91	142	68
8	19	24	92	131	51	58	16	19.4	84	100	61
9	21	19.2	89	118	68	59	17	19.3	78	128	54
10	22	23.4	55	136	71	60	18	22.6	90	121	78
11	18	22.8	64	116	54	61	16	19.5	100	121	51
12	17	23.2	91	136	82	62	16	27	73	126	67
13	18	28	69	160	70	63	18	19.6	77	102	66
14	18	22.2	72	116	53	64	19	20.3	79	87	51
15	22	23.1	79	113	43	65	19	22.9	89	108	51
16	19	20.3	58	112	60	66	25	27.7	80	137	66
17	18	21.8	67	95	64	67	33	30.4	62	131	74
18	21	19.5	70	115	54	68	27	30.8	62	150	72
19	18	19.2	70	118	63	69	40	30.8	80	108	72
20	20	23.4	64	132	74	70	35	25.8	75	113	61
21	20	20.8	78	137	83	71	19	25.2	82	137	60
22	16	22.1	92	90	52	72	36	26	76	129	78
23	23	24.8	89	126	70	73	24	22	93	110	58
24	24	21.1	100	124	70	74	23	24.5	66	126	57
25	21	26.8	102	134	76	75	23	20.3	77	92	46
26	18	21.9	90	116	65	76	27	32.1	65	100	66
27	18	25.7	75	158	66	77	18	21.8	82	121	50
28	19	24.9	55	124	58	78	26	22.6	100	110	59
29	22	21.2	72	121	82	79	25	21.7	85	110	55
30	19	25.4	85	126	78	80	18	20.3	104	115	67
31	18	19.3	74	110	70	81	25	21.7	85	110	55
32	17	28.7	91	121	59	82	16	21.2	66	107	50
33	23	19.7	88	97	34	83	25	20	67	96	88
34	18	19.8	85	110	56	84	34	19.7	80	123	64
35	19	20.6	74	113	58	85	28	25.6	77	94	51
36	22	22	75	139	78	86	23	26.5	94	134	69
37	17	17.8	86	81	38	87	35	20.8	87	164	96
38	23	24.8	72	124	51	88	28	25	78	132	85
39	18	20.6	69	104	54	89	40	32.9	97	198	11
40	18	16.9	64	104	52	90	40	31.2	76	148	67
41	16	22.1	101	120	84	91	21	19.5	85	110	67
42	16	16.9	88	120	84	92	23	24.1	74	129	74
43	16	24.2	78	132	42	93	39	28.1	74	107	65
44	16	18.8	104	124	82	94	30	32	86	144	75
45	17	17.8	73	96	43	95	22	15.8	82	93	53
46	17	22.6	89	131	64	96	40	26.2	75	139	61
47	16	18.1	80	124	38	97	38	21.5	78	115	67
48	16	20.2	73	102	59	98	16	23.5	114	140	71
49	16	17.8	82	118	88	99	32	26.8	87	105	64
50	16	22.6	82	92	34	100	38	34.6	86	85	139

S/N	Age	BMI	PR	SBP	DBP
101	28	26.9	82	58	129
102	30	19.3	59	95	110
103	21	21.6	54	73	96
104	18	24	89	103	123
105	25	26	54	78	105
106	31	33.2	81	90	147
107	40	25.6	77	84	130
108	35	30.1	16	70	128
109	25	24.1	77	87	131
110	34	34.4	70	78	128
111	32	31.5	73	79	134
112	32	30.4	75	89	133
113	29	32.8	69	83	132
114	36	30.5	72	80	131
115	18	25.3	70	80	121
116	21	22.8	57	83	110
117	20	25.9	61	77	117
118	37	26.5	71	79	131
119	19	25.3	64	78	101
120	26	31.6	80	81	139
121	25	28.3	69	80	127
122	33	27.2	71	85	139
123	22	24.2	69	78	128
124	28	30.1	60	70	120
125	23	23.6	60	70	111
126	21	28.4	62	72	115
127	22	26	70	78	128
128	28	25.3	81	75	130
129	19	25.7	60	75	112
130	35	29.7	71	80	126
131	36	31.6	72	80	130
132	22	31.3	60	77	100
133	31	26.2	70	80	125
134	27	21.7	50	75	95
135	18	23.3	51	73	95
136	35	24.5	56	77	126
137	21	19.1	60	73	117
138	21	23.7	57	98	102
139	21	21.5	55	79	121
140	19	24.8	55	72	108
141	23	20.5	49	78	123
142	24	22.7	62	94	118
143	25	23.4	58	79	115
144	24	20.2	74	67	112
145	19	31.5	79	93	113
146	27	22.9	74	75	104
147	19	21.2	59	91	108
148	23	23	54	64	120
149	21	21.9	61	75	123
150	19	23.3	69	80	124
151	25	22.1	64	57	108
152	24	25	62	78	110
153	30	23.4	72	75	112
154	20	19.4	54	78	102
155	20	20	70	67	124
156	20	20.5	43	72	107
157	19	21.7	66	100	108
158	23	19.4	30	75	84
159	19	23.5	56	82	110
160	18	16	90	80	110
161	20	23.1	44	72	112
162	23	22.3	71	61	118
163	24	21.1	69	68	128
164	22	24.2	68	59	128

S/N	Age	BMI	PR	SBP	DBP
165	21	22.5	68	91	123
166	20	21.1	52	81	113
167	23	23.8	92	77	142
168	24	24.7	60	55	118
169	23	23.9	56	83	107
170	24	21.5	49	73	115
171	23	20.3	51	72	107
172	25	22.9	74	56	120
173	23	23.5	40	54	108
174	24	22.6	67	69	142
175	19	18.7	43	91	96
176	22	16.9	50	87	97
177	26	25.9	46	82	105
178	23	20	142	89	155
179	18	27.4	75	88	136
180	25	22.8	66	63	100
181	21	21.9	56	63	105
182	25	23	70	53	121
183	21	24.9	51	70	99
184	18	36.4	77	63	137
185	19	23.6	62	82	118
186	36	26	80	78	110
187	21	18.9	62	82	110
188	23	19.1	56	75	91
189	19	23.1	43	55	102
190	24	17.9	66	104	124
191	26	22.4	64	63	131
192	25	22	75	91	136
193	22	21	58	85	104
194	24	21.9	73	63	110
195	25	19.5	54	59	94
196	17	20.6	34	72	99
197	23	19.6	78	73	113
198	24	23.5	67	79	104
199	22	18.5	62	73	96
200	25	20.7	59	85	100
201	17	19.5	46	85	115
202	17	16.4	58	121	118
203	16	19.8	50	96	104
204	16	19.7	58	85	113
205	16	25.6	59	84	100
206	16	21	34	92	102
207	27	23.3	70	60	116
208	16	23.2	70	83	115
209	17	22.1	64	81	107
210	17	20.3	72	78	107
211	17	21.2	58	79	118
212	16	21.6	80	99	121
213	17	24.5	71	101	116
214	16	17.3	58	86	100
215	19	22.4	78	91	123
216	16	20.8	71	95	104
217	16	27.8	65	96	129
218	15	23.1	65	96	129
219	17	20.7	51	84	139
220	19	21.4	59	78	108
221	17	20.3	58	100	121
222	18	22.9	40	67	83
223	17	22	59	46	104
224	18	17.9	54	75	102
225	18	17.9	54	75	102
226	17	21.3	64	69	107
227	16	23.1	59	96	121
228	18	21.4	51	71	100

S/N	Age	BMI	PR	SBP	DBP
229	16	19.9	61	94	88
230	16	24.5	75	75	112
231	16	20.5	53	56	126
232	16	18.1	48	56	118
233	17	22.9	54	81	132
234	16	19.6	53	86	124
235	17	24.2	65	76	110
236	16	20.3	54	116	89
237	16	20.6	62	73	104
238	16	19.9	53	95	100
239	16	20.3	59	108	136
240	16	18.8	41	79	84
241	16	20.5	58	72	97
242	16	18.8	70	114	110
243	17	18.6	54	82	104
244	17	21.2	46	78	131
245	16	16.7	58	85	104
246	17	19.3	61	81	99
247	18	21.8	55	72	152
248	16	19.2	51	99	112
249	18	20	51	81	99
250	16	21	87	60	158
251	16	21.1	87	60	158
252	16	20.8	61	73	116
253	17	29.8	64	67	121
254	19	20.5	58	103	110
255	18	21	72	93	124
256	18	21.9	47	53	76
257	17	23.4	74	85	132
258	23	20	83	78	121
259	17	20.3	70	70	131
260	18	20.4	66	111	108
261	18	23.7	74	78	126
262	16	22.9	75	108	129
263	16	21.9	55	91	112
264	16	17.5	58	78	113
265	18	22.8	61	81	108
266	16	27	78	75	131
267	17	23.5	73	75	105
268	16	16.1	48	81	96
269	29	23.2	62	69	137
270	21	23.3	59	72	115
271	17	23.1	46	88	124
272	18	22.6	69	81	112
273	25	23.2	55	79	113
274	16	23.7	65	92	144
275	17	22.6	48	71	105
276	18	20.8	38	71	115
277	18	20.8	36	82	78
278	16	25.8	51	98	94
279	19	20	54	81	100
280	16	21.3	54	68	112
281	16	21.3	54	68	112
282	16	21.9	29	71	110
283	18	18.2	37	90	105
284	16	21.8	53	67	121
285	18	17.7	59	104	126
286	24	18.6	84	103	97
287	17	24.4	53	75	97
288	18	23.3	54	63	116
289	16	18.2	48	101	120
290	25	22.8	63	69	126
291	25	24.4	56	111	124
292	24	19.5	48	70	94

S/N	Age	BMI	PR	SBP	DBP
293	18	26.7	59	103	120
294	17	26.5	53	85	128
295	19	25	66	119	129
296	16	22.1	58	89	140
297	25	18.8	70	82	100
298	16	16.4	56	89	108
299	18	21.1	67	77	104
300	18	18.7	52	88	91
301	25	25	54	89	126
302	17	22.5	64	79	108
303	24	18.9	43	91	96
304	16	18.3	53	105	104
305	18	21.1	59	98	134
306	35	20.8	72	111	113
307	18	18.4	42	76	88
308	35	20.7	57	87	99
309	18	20.6	59	75	104
310	19	17.7	56	93	131
311	18	19.5	33	80	129
312	18	17.6	67	86	100
313	19	24.4	53	86	124
314	18	19.1	54	87	96
315	19	20.5	54	65	121
316	17	16.4	36	94	102
317	19	18.4	59	75	110
318	17	27.2	69	72	120
319	18	17.6	48	68	88
320	17	21.6	67	77	120
321	16	21.4	95	73	134
322	17	20.1	40	66	113
323	18	19.6	50	67	118
324	16	18.6	85	79	131
325	18	16.4	66	87	107
326	18	21.6	46	92	116
327	16	20.2	67	81	121
328	16	19	88	125	134
329	16	21.9	54	65	102
330	16	27.9	52	74	100
331	18	22.8	70	88	120
332	20	24.2	50	85	110
333	18	20.6	50	85	110
334	19	21.4	37	74	111
335	18	25.4	45	52	128
336	18	30.5	56	73	131
337	18	24.5	78	72	120
338	17	22.6	70	65	134
339	18	24.5	78	72	120
340	17	22.4	46	120	77
341	17	23.4	46	73	112
342	19	31.1	78	37	160
343	16	19.1	54	70	121
344	17	22.5	54	70	121
345	17	22.5	62	81	124
346	18	20.6	48	91	105
347	18	22.2	56	75	100
348	19	23.1	76	102	126
349	16	23.8	76	102	126
350	16	22.5	58	82	104
351	16	26.5	58	74	108
352	19	24	38	90	96
353	18	25.8	61	72	113
354	18	21.9	78	63	134
355	19	23.8	53	83	99
356	18	25.1	78	63	134

S/N	Age	BMI	PR	SBP	DBP
357	18	22.8	66	87	102
358	18	23.7	67	81	112
359	16	19.7	35	91	68
360	16	25.4	69	82	112
361	16	27	82	101	104
362	16	23.4	50	97	94
363	16	20.5	50	70	113
364	19	22.5	43	76	112
365	18	26.5	69	88	140
366	18	20.8	35	67	97
367	18	21.6	61	62	128
368	22	21.4	38	74	107
369	16	24.2	59	70	120
370	21	19	39	73	100
371	19	23.9	61	69	136
372	20	25.6	51	86	102
373	16	19.5	43	81	91
374	19	28.7	66	66	126
375	19	23.6	62	80	129
376	18	21.8	73	73	104
377	18	24.2	67	40	145
378	18	26.7	64	80	136
379	17	22.2	62	90	97
380	18	21.8	79	60	96
381	16	30.4	64	81	136
382	16	27.5	74	90	120
383	17	21.4	67	93	116
384	18	19.8	40	107	94
385	18	26	69	68	132
386	20	22.5	80	129	116
387	18	22.8	62	54	121
388	19	24.1	66	67	112
389	16	22.6	70	102	116
390	18	24	64	72	107
391	18	28.7	53	76	144
392	25	18.4	57	68	99
393	16	21.5	43	83	137
394	18	22.5	54	75	115
395	24	18.9	68	75	115
396	18	22.8	56	64	120
397	17	20.9	46	70	94
398	18	23.7	57	99	99
399	18	25	76	70	94
400	18	28	86	106	152
401	17	21.5	37	69	129
402	17	20.2	53	67	115
403	17	19.8	42	61	130
404	20	18	74	85	121
405	17	22	43	127	120
406	16	25.6	66	116	118
407	18	20.7	40	78	128
408	17	21.6	63	109	116
409	21	24.2	67	68	144
410	17	20	56	89	126
411	17	20.4	38	111	112
412	17	28.1	74	131	136
413	17	22.9	48	67	113
414	16	21.6	91	92	123
415	16	18.6	70	115	104
416	17	15.8	45	74	96
417	17	23.9	69	80	131
418	17	23	60	70	107
419	17	27.3	59	70	115
420	16	19.5	40	81	92

S/N	Age	BMI	PR	SBP	DBP
421	27	26.6	61	100	121
422	19	23.4	54	106	105
423	20	25.3	54	93	92
424	17	20.3	54	95	107
425	16	20.3	46	64	115
426	18	21.6	29	73	94
427	17	23.4	50	67	89
428	17	26.8	52	61	104
429	17	21.6	50	70	102
430	18	18.2	58	72	134
431	18	25	66	76	124
432	16	19.8	56	83	99
433	17	22.4	42	95	97
434	16	21.1	59	66	124
435	17	22.6	53	96	145
436	16	19.3	46	98	105
437	17	23.2	61	75	107
438	17	26.1	75	75	110
439	17	23.6	62	69	107
440	17	18.6	37	64	100
441	19	25.5	74	78	131
442	16	22.5	51	109	107
443	16	22.5	61	102	139
444	17	18.1	32	84	72
445	17	18.9	54	100	105
446	19	19.8	46	57	110
447	17	21.2	62	67	123
448	17	25.8	59	93	124
449	16	24.8	53	75	132
450	18	22.8	67	70	137
451	17	21.2	45	76	94
452	18	19.6	62	62	137
453	17	21.4	58	84	115
454	17	16.5	51	105	105
455	21	23	59	59	105
456	24	20.1	61	92	108
457	18	22.9	66	72	113
458	16	23.8	67	69	128
459	17	21.7	53	91	110
460	17	21.4	74	82	131
461	17	23.8	38	73	107
462	16	21	66	76	110
463	16	20.2	53	59	105
464	19	22.6	50	78	78
465	17	22.4	42	59	59
466	17	19.7	45	73	73
467	17	23.2	70	75	75
468	17	22.1	44	82	82
469	18	28.6	42	86	86
470	17	21.6	70	66	66
471	16	20.6	58	116	116
472	17	19	59	102	102
473	16	21.4	70	80	80
474	16	21.8	75	72	72
475	17	20.7	53	78	78
476	18	24.2	58	83	83
477	30	23.9	56	70	70
478	16	21.4	75	94	94
479	17	24.4	80	116	116
480	17	30	59	84	84
481	17	23.7	83	66	66
482	30	20.3	56	79	79
483	22	23.1	50	101	101
484	22	21	58	62	62

S/N	Age	BMI	PR	SBP	DBP
485	20	17.8	47	67	67
486	29	20.2	58	69	69
487	26	21.2	68	69	69
488	17	22	89	79	79
489	24	13.3	28	84	84
490	21	22.2	46	81	81
491	25	22.8	64	58	58
492	25	23.1	69	73	73
493	21	22.4	66	54	54
494	19	23.1	70	91	91
495	22	29.4	75	64	64
496	21	21.4	59	65	65
497	21	19.7	78	85	85
498	21	19.1	66	71	71
499	20	18.9	58	70	70