IE 7280 – Statistical Methods in Engineering

Prof. Nizar Zaarour

Regression Analysis: Relation of GDP Per Capita and Population of a Country to Per Capita Plastic Waste

Group-6

Naren Mohan, Rahul Gupta, Ritika Wadhwa

Table of Contents

Introduction

Motivation: Plastic is a synthetic organic polymer made from petroleum with properties ideally suited for a wide variety of applications including packaging, building and construction, household and sports equipment, vehicles, electronics, and agriculture. Over 300 million tons of plastic are produced every year, half of which is used to create single-use items such as shopping bags, cups, and straws. If discarded improperly, plastic waste can harm the environment and biodiversity.

Objective: With the ever-increasing usability of plastic, we start exploiting it, which has a highly negative effect on our environment, marine life, and wildlife. In this study, we are exploring how plastic waste generated per person per day and the GDP per capita of a country is related to each other. We aim to find if the high-income countries consume more plastic waste by performing a Simple Linear Regression Analysis with GDP per capita as a dependent variable and Per Capita Plastic Waste as the predictor.

Following that, we will introduce another dependent variable, Population of a Country, and perform a Multiple Linear Regression and see if it improves the model for prediction of Per Capita Plastic Waste.

Our objective is to find a regression model that best fits the variables GDP and Population to predict Per Capita Plastic Waste.

Once we understand factors that are related to generation of plastic waste, we can work on finding ways to reduce plastic generation or finding better ways of waste disposal and reusability of plastic.

Dataset: The dataset was found on https://ourworldindata.org/plastic-pollution. The dataset includes the following variables:

- Dependent Variable (Y) - Per Capita Plastic Waste in kg/person/per day,
- Predictor (X) - GDP of the country based on purchasing power parity (PPP) in constant 2017 International Dollars. PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates
- Predictor 2 (X2) - Population of the country

The Dataset is from 2010. The sample size is 60 and we have selected 60 countries randomly.

The complete dataset is provided in Appendix A.

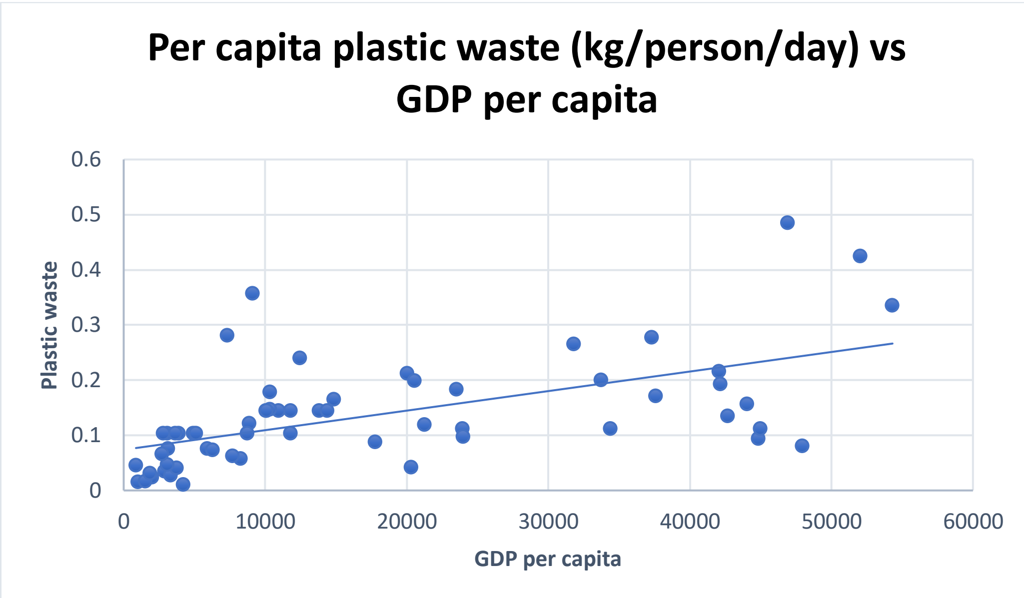
Limitations: The data represent total plastic waste generation and do not account for differences in waste management, recycling, or incineration. Therefore, we cannot say how much is disposed-off into the ocean or other waterways, and how much is recycled.

Software Used: We'll be using MS Excel and IBM SPSS for all calculations, tables, and charts.

Variables

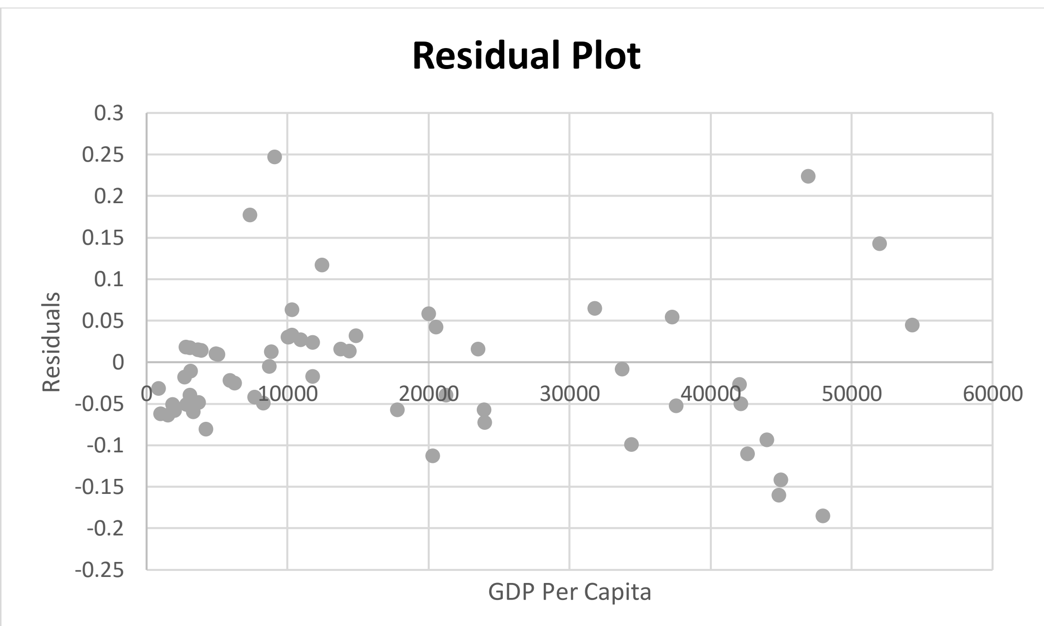
1. **Graphical relation between Per Capita Plastic Waste and GDP Per Capita**

Scatterplot



In the scatterplot, we can see a positive relationship between Per Capital Plastic Waste and GDP Per Capita. As the GDP increases, we can see an increase in Plastic Waste.

Residual Plot



From the Residual Plot, we can see an almost symmetrical distribution, with the tendency to cluster in the middle for the lower values of X-axis: GDP Per Capita. As we increase the GDP, we see an increase in outliers.

With the graphical representations, we can conclude that there is a positive correlation between X and Y.

2. **Descriptive Statistics for both variables**

	GDP Per Capita	Per Capita Plastic Waste
Mean	17590.01999	0.136183333
Median	10655.83833	0.112
Mode	#N/A	0.103
Variance	250953925.2	0.009413816
Standard Deviation	15841.52534	0.097024824
Skewness	0.87733291	1.446725175

Simple Linear Regression

Linear Regression with 1 variable

1. **ANOVA Output Tables**

Full Tables in Appendix B

Correlation Coefficient: Rxy

The Rxy value is 0.579 which shows a positive correlation between X and Y, i.e., GDP Per Capita and Plastic Waste Per Capita. However, according to the standards, even though there is a positive correlation, it is not a strong correlation since its <0.7

Relationship between the variables

Coefficient of variable is positive with a value of 3.543E-06 which indicates that the relationship between the dependent variable and the independent variable i.e., plastic consumption and GDP per capita of a country have a positive relationship

Regression Equation

b0 = 0.074

b1 = 3.544E-6

Linear Regression Equation: y = b0 + b1x

Y* = 0.074 + 3.544E-6 *x

Sum of Squares:

SSR = 0.189

SSE = 0.376

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-2 = 58)

MSR = 0.189 (d.f. = k = 1)

Coefficient of Determination: R^2 value

The value of R^2 which is 0.335 indicates that our model this model explains 33.5% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = 0.804897

Measure of much variability around the reference line (predicted values) from the actual values.

2. **Hypothesis Testing – Testing for Significance (α=0.05)**

<u>**F-Test**</u>

Ho: $\beta 1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta 1$ ≠ 0 (Alternative Claim: Slope = Relationship)

From the table, F-stat = 29.184, p-value= 0.000001

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

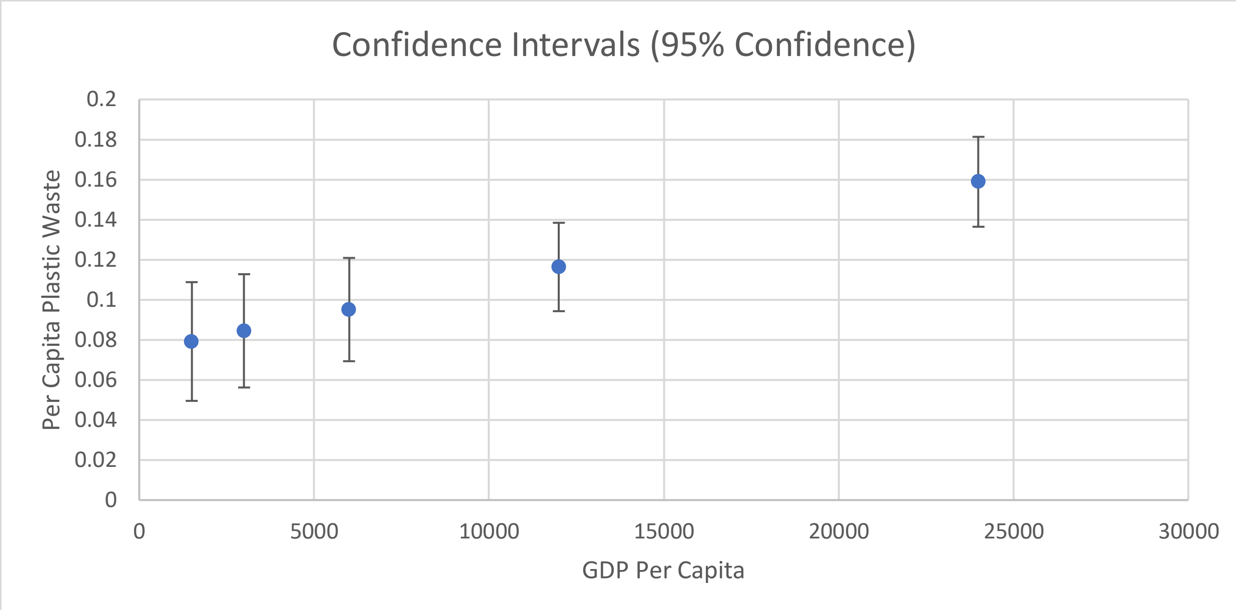
Since F-stat > F(α) -> It lies in the rejection area

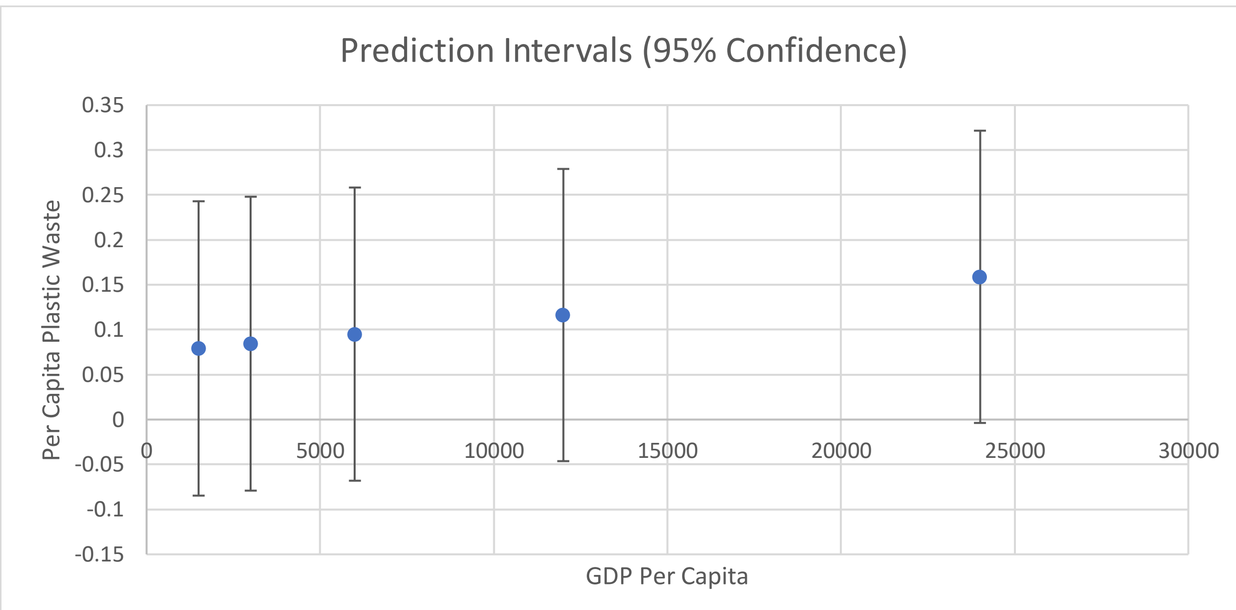
We have significant evidence to reject the Null Hypothesis Ho

So, $\beta 1$ ≠ 0, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between X and Y (GDP Per Capita and Plastic Waste Per Capita). We can say that the simple linear regression model (*Y* = 0.074 + 3.544E-6 *x)* is significant.

3. Confidence & Prediction Interval (*Full Table in Appendix B*)





Conclusions:

1. We found that there is a positive correlation between GDP and Plastic Waste with Correlation Coeff = 0.579
2. After building a Simple Linear Regression Model with equation ***Y* = 0.074 + 3.544E-6 *x*,** we found that Coeff of Determination to be 0. 335, so 33.5% of the data can be explained by the model.

3. To find if the model is significant, we performed hypothesis testing and found that with a significance level of 0.05, the linear regression model is significant and there is a relationship between X and Y.
4. The p-value from F-test is 0.000001, which is extremely small so we can conclude that the model will have evidence to show relationship between X and Y, even with a significance level of 0.01.
5. Further, we will build a model with log(X) and log(X^2) as a predictor by using Transformation to convert it to Linear Regression to see if there is an improvement in the model.

Transformed Logarithmic Regression

Full Tables in Appendix C

Building a Transformed Linear Regression Model with Predictor – Log(x) and Dependent Variable – Y

We used Backward Elimination from SPSS to choose the best predictor out of the two. The variables entered were log (GDP) and log (Population). We got the 3 different possible models (all outputs shown in Appendix C) and chose the 3rd model with the best results.

Correlation Coefficient: Rxy

The Rxy value is 0.605 which shows a positive correlation between log(X) and Y, i.e., log value of GDP Per Capita and Plastic Waste Per Capita. However, according to the standards, even though there is a positive correlation, it is not a strong correlation since its <0.7

Regression Equation

b0 = -.362

b1 = 0.124

Linear Regression Equation: y = b0 + b1* log(x)

Y* = -.362 + 0.124 log(*x)

Sum of Squares:

SSR = 0.207

SSE = 0.358

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-2 = 58)

MSR = 0.207 (d.f. = k = 1)

Coefficient of Determination: R^2 value

The adjusted value of R^2 which is 0.356 indicates that our model this model explains 35.6% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = .0785415

Measure of much variability around the reference line (predicted values) from the actual values.

Hypothesis Testing – Testing for Significance (α=0.05)

<u>**F-Test**</u>

Ho: $\beta1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta1 \neq 0$ (Alternative Claim: Slope = Relationship)

From the table, F-stat = 33.563, p-value= 2.98E-07

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

Since F-stat > F(α) -> It lies in the rejection area

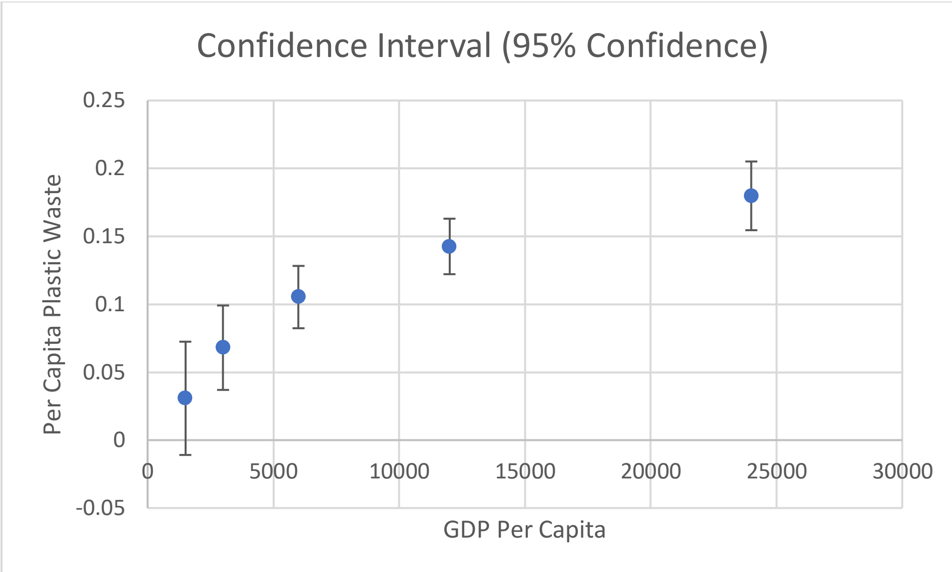
We have significant evidence to reject the Null Hypothesis Ho

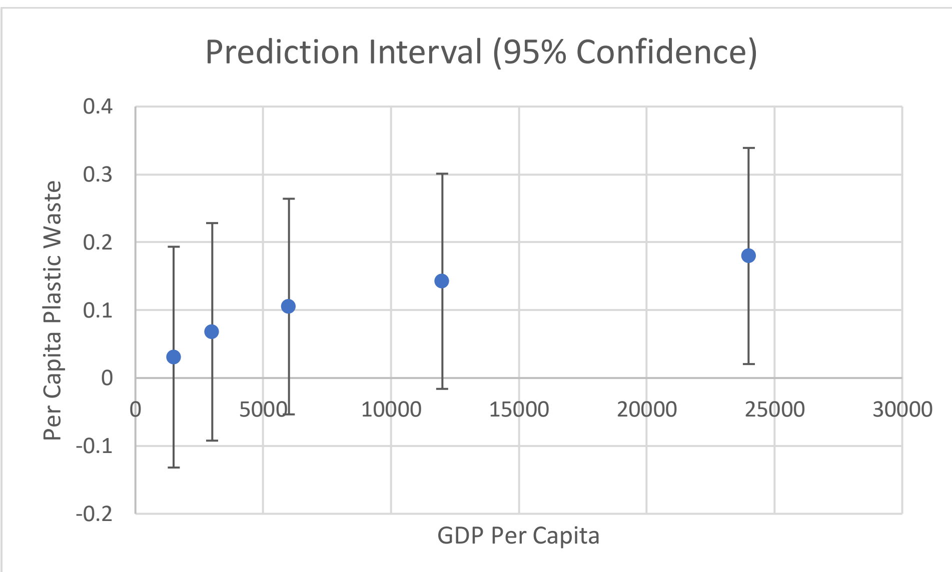
So, $\beta1 \neq 0$, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between X and Y (GDP Per Capita and Plastic Waste Per Capita).

Confidence and Prediction Intervals with 95% Confidence

Full Tables in Appendix C





Conclusion

1. The Transformed Logarithmic Regression Model shows improvement from the Simple Linear Regression Model
2. The Adjusted R-square value has improved, so more data can be explained by this model.
3. The Standard Error of Estimate is smaller than the Simple Linear Regression Model so there's lesser variability around the reference line.
4. The p-value is also much smaller than p-value obtained from the Simple Linear Regression.

5. We can conclude, the Transformed Logarithmic Regression Model performs better than the Simple Linear Regression.

Transformed Quadratic Regression

Full Tables in Appendix D

Building a Transformed Linear Regression Model with Predictor – Log(x^2) and Dependent Variable – Y

Correlation Coefficient: Rxy

The Rxy value is 0.607 which shows a positive correlation between log(X) and Y, i.e., log value of GDP Per Capita and Plastic Waste Per Capita.

Regression Equation

b0 = -.121

b1 = 0.016

Linear Regression Equation: y = b0 + b1* log(x^2)

***Y* = -.121 + 0.016 log(*x^2*)**

Sum of Squares:

SSR = 0.208

SSE = 0.357

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-2 = 58)

MSR = 0.208 (d.f. = k = 1)

Coefficient of Determination: R^2 value

The adjusted value of R^2 which is 0.357 indicates that our model this model explains 35.7% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = .0784563

Measure of much variability around the reference line (predicted values) from the actual values.

Hypothesis Testing – Testing for Significance (α=0.05)

<u>F-Test</u>

Ho: $\beta 1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta 1 \neq 0$ (Alternative Claim: Slope = Relationship)

From the table, F-stat = 33.762, p-value= 2.79E-07

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

Since F-stat > F(α) -> It lies in the rejection area

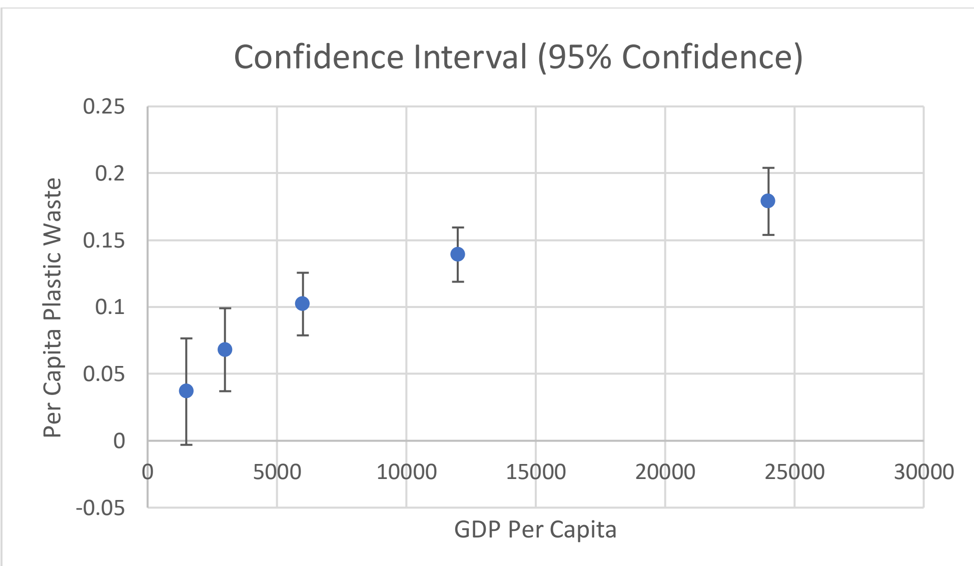
We have significant evidence to reject the Null Hypothesis Ho

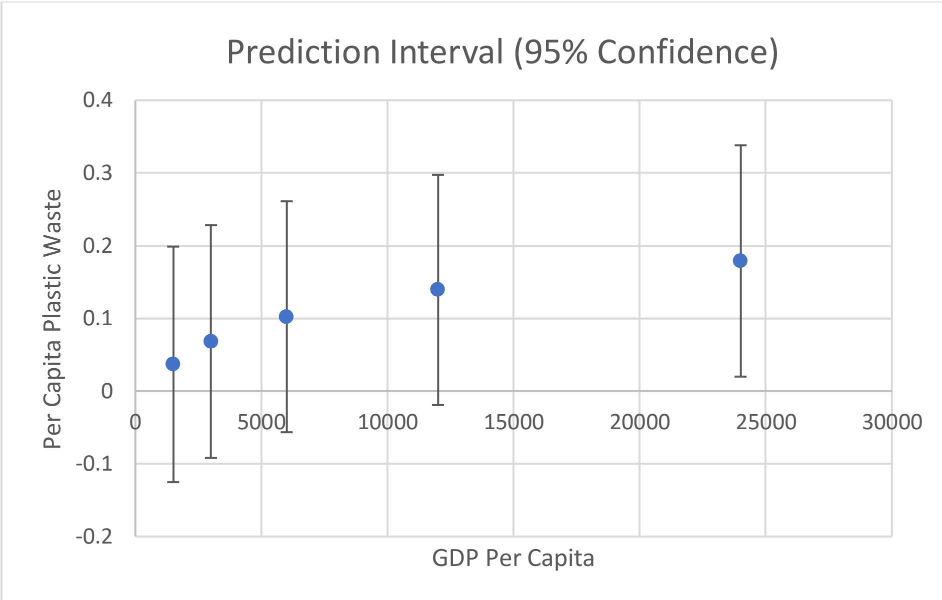
So, $\beta 1 \neq 0$, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between X and Y (GDP Per Capita and Plastic Waste Per Capita).

Confidence and Prediction Intervals with 95% Confidence

Full Tables in Appendix D





Conclusion

1. The Transformed Quadratic Regression Model shows slight improvement from the Logarithmic Regression Model
2. The Adjusted R-square value has slightly increased, and the Standard Error of Estimate is slightly smaller than the Logarithmic Regression Model so there's slightly lesser variability around the reference line.
3. The p-value is slightly smaller than p-value obtained from the Transformed Logarithmic Regression.
4. We can conclude, the Transformed Quadratic Regression Model performs better than the Simple Linear Regression and the Transformed Logarithmic Regression.
5. The Simple Linear Regression Model has shown good performance. Further, we will try adding another predictor and using Multiple Linear Regression to see if there is more improvement.

Multi Linear Regression

Multi Linear Regression – 3 Variables

Full Tables in Appendix E

Building a Multi Linear Regression model using 3 variables - GDPPerCapita, Population, GDP

Correlation Coefficient: Rxy

Rxy value is 0.579 which shows a positive correlation between GDPPerCapita and Y; is an indicator that GDPPerCapita will be a useful predictor.
Rxy value is -0.102 which shows a negative correlation with lower value between Population and Y; is an indicator that Population will not be a useful predictor.
Rxy value is 0.286 which shows a positive correlation between GDP and Y; is an indicator that GDP might or might not be a useful predictor.

Regression Equation

b0 = .083

b1 = 2.842E-6

b2 = -8.956E-11

b3 = 8.730E-15
Linear Regression Equation: y = b0 + b1* GDPPerCapita + b2 * Population + b3 * GDP

Ŷ = .083 + 2.842E-6 * GDPPerCapita + -8.956E-11 * Population + 8.730E-15 * GDP

Sum of Squares:

SSR = 0.203

SSE = 0.361

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-1-k = 56)

MSR = 0.068 (d.f. = k = 3)

Coefficient of Determination: R^2 value

The adjusted value of R^2 which is 0.326 indicates that our model this model explains 32.6% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = .0803263

Measure of much variability around the reference line (predicted values) from the actual values.

Hypothesis Testing – Testing for Significance (α=0.05)

<u>F-Test</u>

Ho: $\beta 1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta 1 \neq 0$ (Alternative Claim: Slope = Relationship)

From the table, F-stat = 10.513, p-value= 0.000014

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

Since F-stat > F(α) -> It lies in the rejection area

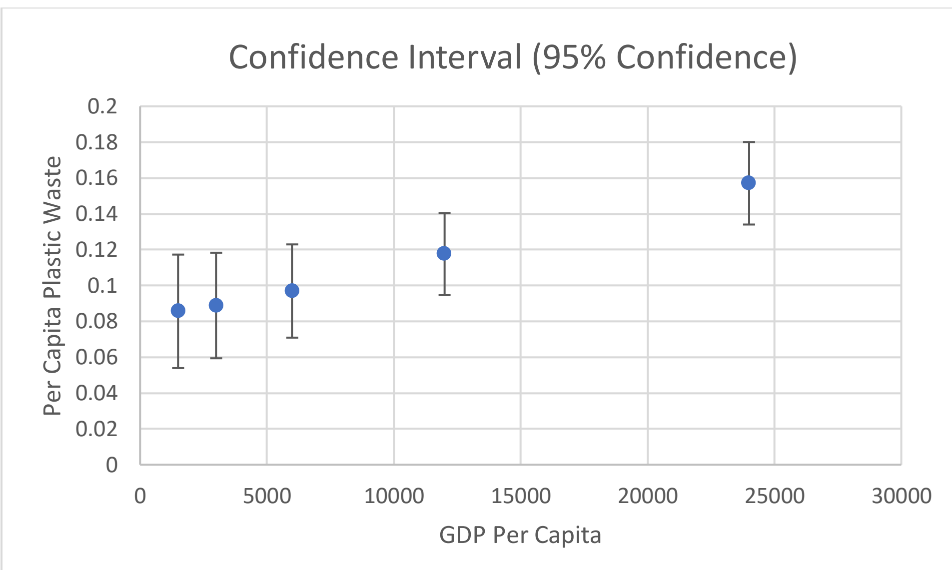
We have significant evidence to reject the Null Hypothesis Ho

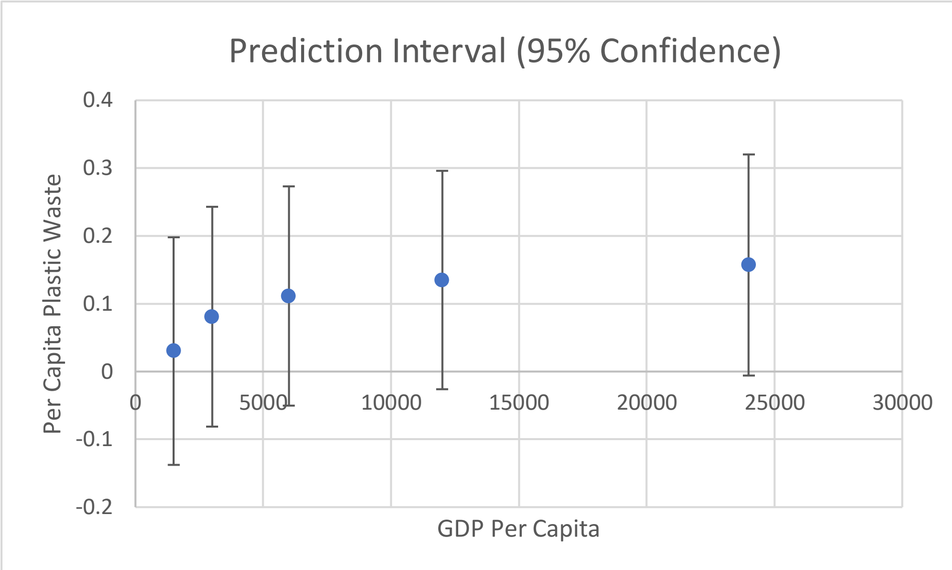
So, $\beta 1 \neq 0$, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between the 3 predictors and Y (GDP Per Capita, Population, GDP and Plastic Waste Per Capita).

Confidence and Prediction Intervals with 95% Confidence

Full Tables in Appendix E





Conclusion

1. The Multi Linear Regression Model with 3 variables do not generalize well as compared to some of the Linear Regression Model
2. The Adjusted R-square value has dropped, so even though more predictors has been added, less data can be explained by this model.
3. The Standard Error of Estimate is greater than the Simple Linear Regression Model so there's more variability as compared to previous model around the reference line.
4. The p-value is also much smaller than p-value obtained from the Simple Linear Regression which proves significance even with significance levels of 0.1, 0.05, 0.01.
5. We can conclude, the Multi Linear Regression Model with 3 variables do not perform better than the previous Linear Regression model.

Multi Regression - Transformed 4 Variables

Full Tables in Appendix F

Building a Multi Linear Regression model using Transformed 4 variables - GDPPerCapita, LogGPC, GPCSqrt, LogPopulation

Correlation Coefficient: Rxy

Rxy value is 0.579 which shows a positive correlation between GDPPerCapita and Y; is an indicator that GDPPerCapita will be a useful predictor.
Rxy value is 0.605 which shows a negative correlation with lower value between LogGPC and Y; is an indicator that LogGPC will be a useful predictor.
Rxy value is 0.599 which shows a positive correlation between GPCSqrt and Y; is an indicator

that GPCSqrt will be a useful predictor.
Rxy value is -0.042 which shows a positive correlation between LogPopulation and Y; is an indicator that LogPopulation might or might not be a useful predictor.

Regression Equation

b0 = -1.067

b1 = 1.482E-5

b2 = .448

b3 = -.006

b4 = -0.14

Linear Regression Equation: y = b0 + b1* GDPPerCapita + b2 * LogGPC + b3 * GPCSqrt + b4 * LogPopulation

Ŷ = -1.067 + 1.482E-5 * GDPPerCapita + .448 * LogGPC + -.006 * GPCSqrt + -0.14 * LogPopulation

Sum of Squares:

SSR = 0.222

SSE = 0.343

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-1-k = 55)

MSR = 0.056 (d.f. = k = 4)

Coefficient of Determination: R^2 value

The adjusted value of R^2 which is 0.349 indicates that our model this model explains 34.9% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = .0789178

Measure of much variability around the reference line (predicted values) from the actual values.

Hypothesis Testing – Testing for Significance (α=0.05)

<u>F-Test</u>

Ho: $\beta 1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta 1 \neq 0$ (Alternative Claim: Slope = Relationship)

From the table, F-stat = 8.923, p-value= 0.000013

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

Since F-stat > F(α) -> It lies in the rejection area

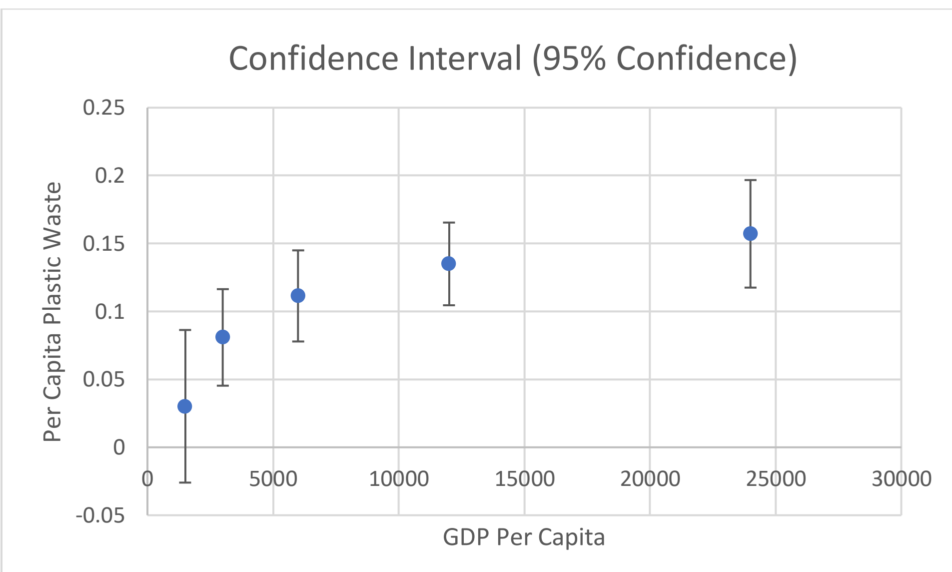
We have significant evidence to reject the Null Hypothesis Ho

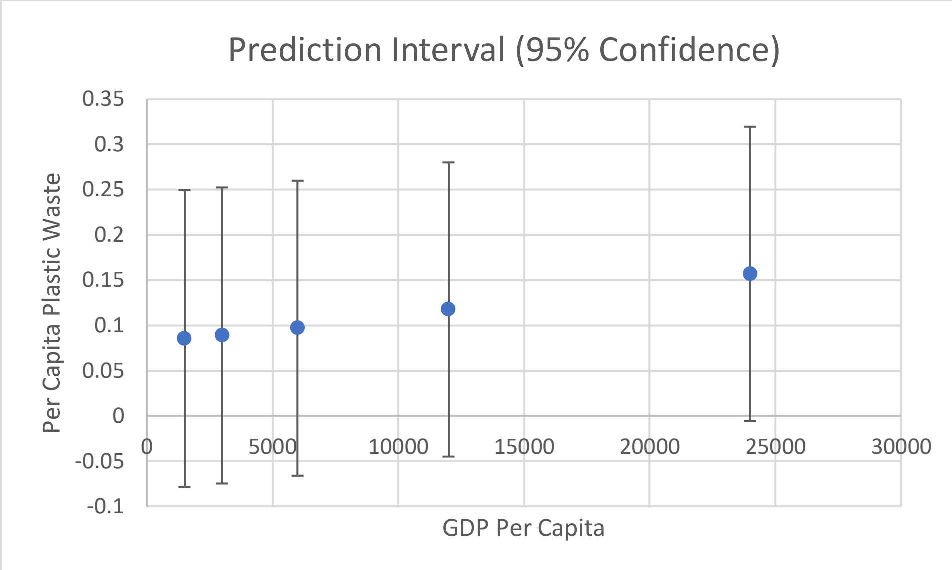
So, $\beta 1 \neq 0$, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between the 3 predictors and Y (GDPPerCapita, LogGPC, GPCSqrt, LogPopulation and Plastic Waste Per Capita).

Confidence and Prediction Intervals with 95% Confidence

Full Tables in Appendix F





Conclusion

1. The Multi Linear Regression Model with 4 Transformed variables generalize better than the previous multi linear regression model but does not better the performance when compared to some of the Linear Regression Model
2. The Adjusted R-square value has improved, so the transformed predictors has shown some significance towards fitting the data.
3. The Standard Error of Estimate is comparable to that of the Simple Linear Regression Model so there's lesser variability as compared to previous model around the reference line.
4. The p-value is also much smaller than p-value obtained from the Simple Linear Regression which proves significance even with significance levels of 0.1, 0.05, 0.01.
5. We can conclude, the Multi Linear Regression Model with 4 Transformed variables do perform better than the previous Multi Linear Regression model.

Multi Regression - Transformed 2 Variables

Full Tables in Appendix G

Building a Multi Linear Regression model using Transformed 4 variables - LogGPC, LogGPCSqrt

Correlation Coefficient: Rxy

Rxy value is 0.605 which shows a positive correlation with lower value between LogGPC and Y; is an indicator that LogGPC will be a useful predictor.
Rxy value is 0.604 which shows a positive correlation between LogGPCSqrt and Y; is an indicator that LogGPCSqrt will be a useful predictor.

Regression Equation

b0 = .504

b1 = .345

b2 = -.877

Linear Regression Equation: **Ŷ** = b0 + b1* LogGPC + b2 * LogGPCSqrt

Ŷ = .504 + .345 * LogGPC + -.877 * LogGPCSqrt

Sum of Squares:

SSR = 0.208

SSE = 0.357

SST = 0.565

Mean Square Error:

MSE = 0.006 (d.f. = n-1-k = 57)

MSR = 0.104 (d.f. = k = 2)

Coefficient of Determination: R^2 value

The adjusted value of R^2 which is 0.346 indicates that our model this model explains 34.6% of the data we have analyzed which had 60 observations.

Standard Error of Estimate

Se = .0789178

Measure of much variability around the reference line (predicted values) from the actual values.

Hypothesis Testing – Testing for Significance (α=0.05)

<u>F-Test</u>

Ho: $\beta 1$ = 0 (Initial Claim: No Slope = No Relationship)

Ha: $\beta 1$ ≠ 0 (Alternative Claim: Slope = Relationship)

From the table, F-stat = 16.577, p-value= 0.000002

Let α = 0.05 (one-tailed)

F(α) = 4.00687265

Since F-stat > F(α) -> It lies in the rejection area

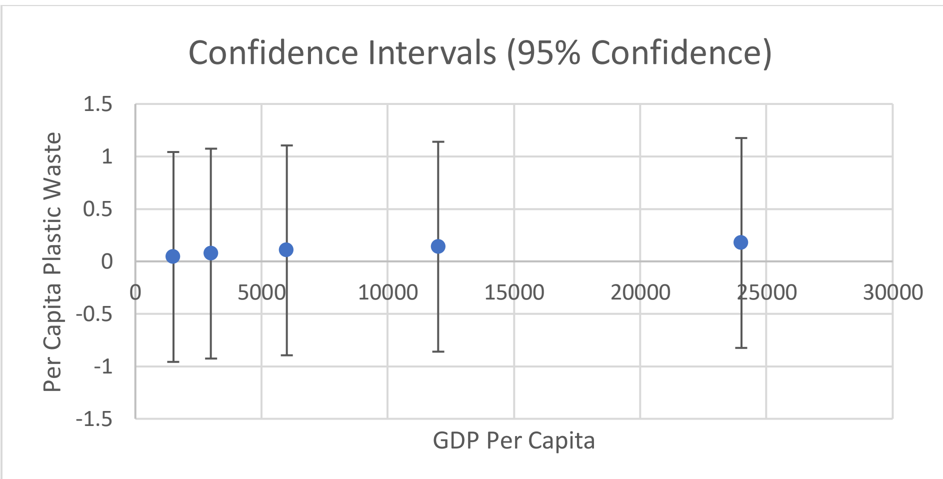
We have significant evidence to reject the Null Hypothesis Ho

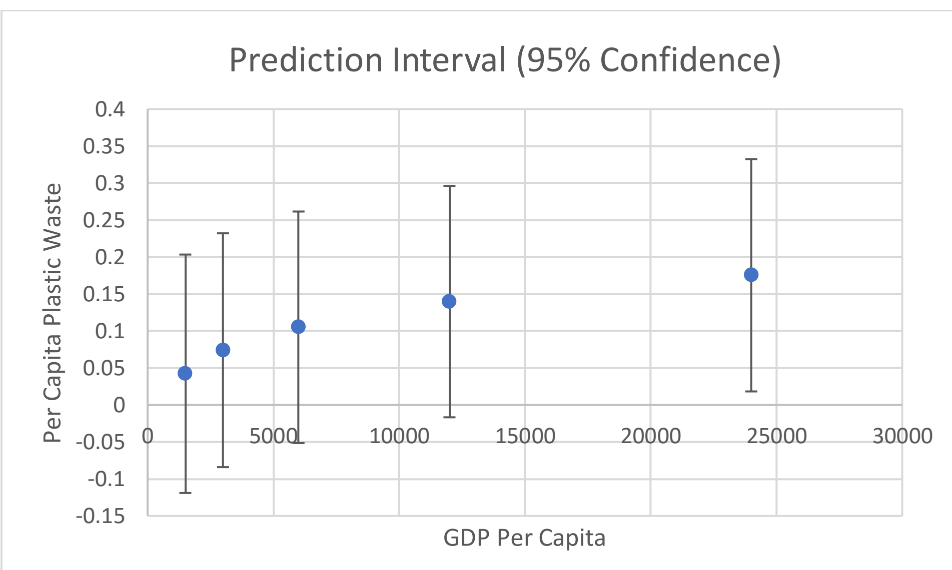
So, $\beta 1 \neq 0$, Slope is not 0

With a significance level of 0.05, we can conclude that we have evidence for a relationship between the 3 predictors and Y (LogGPC, LogGPCSqrt and Plastic Waste Per Capita).

Confidence and Prediction Intervals with 95% Confidence

Full Tables in Appendix G





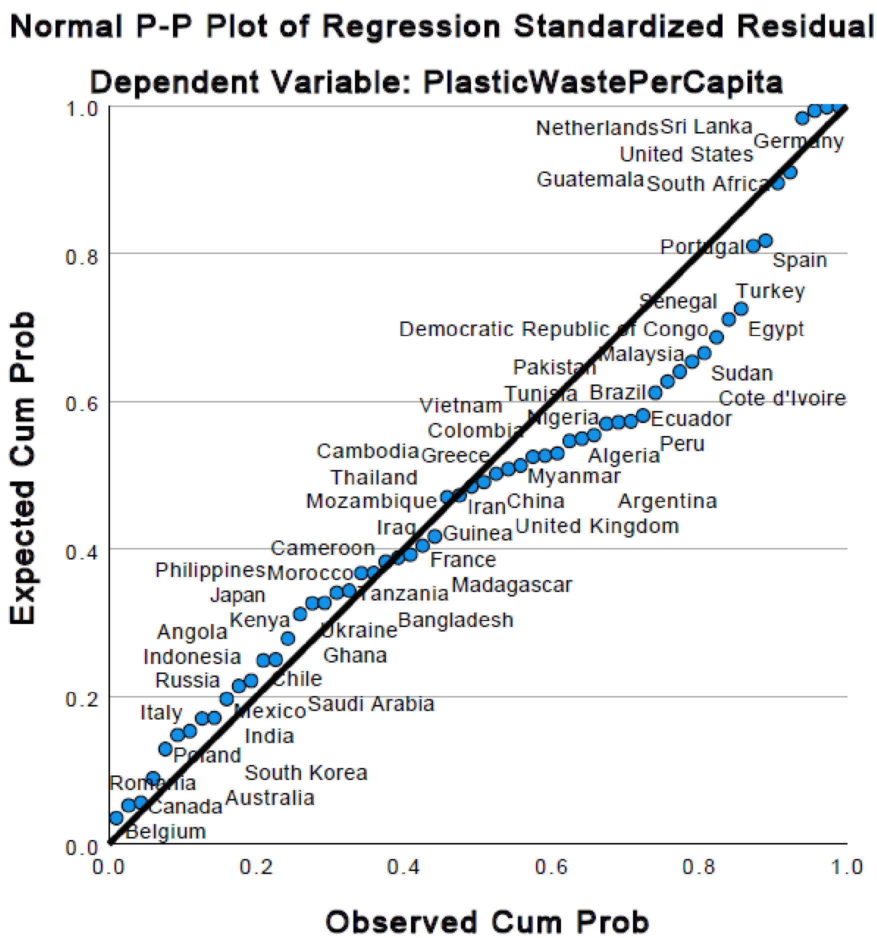
Conclusion

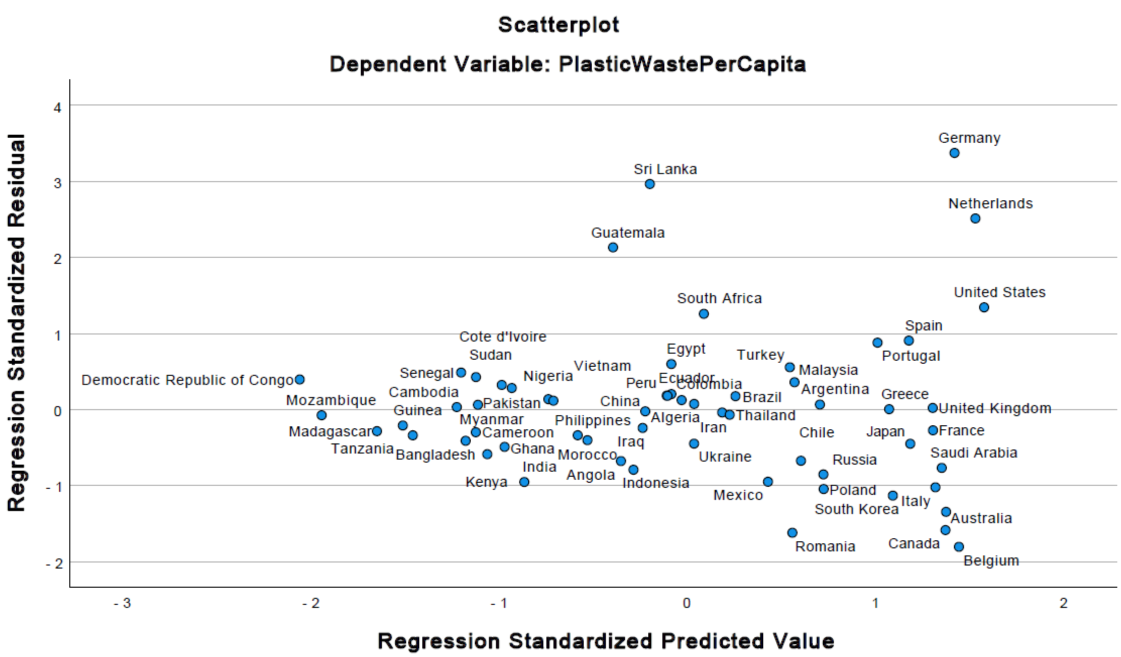
1. The Multi Linear Regression Model with 2 Transformed variables generalize better than the previous multi linear regression model but does not better the performance when compared to some of the Linear Regression Model
2. The Adjusted R-square value has improved, so the transformed predictors has shown some significance towards fitting the data.
3. The Standard Error of Estimate is comparable to that of the Simple Linear Regression Model so there's lesser variability as compared to previous model around the reference line.
4. The p-value is also much smaller than p-value obtained from the Simple Linear Regression which proves significance even with significance levels of 0.1, 0.05, 0.01.

We can conclude, the Multi Linear Regression Model with 4 Transformed variables do perform better than the previous Multi Linear Regression model.

Summary & Conclusions

1. We tested our dataset with 6 models – Simple Linear Regression, Transformed Logarithmic Regression, Transformed Quadratic Regression, Multi Linear Regression, Multi Linear Regression with 4 transformed variables, Multi Linear Regression with 2 transformed variables
2. We found that all models showed significant evidence to reject the Null Hypothesis (α = 0.05), so we can conclude that there is a relationship between both our predictors – GDP Per Capita and Population and our dependent variable – Per Capita Plastic Waste
3. **We can conclude that richer countries consume higher amounts of plastic**
4. Out of all the models, **Transformed Quadratic Regression** showed the best fit to the dataset.
5. With the **p-value of 2.79E-07** and **R^2 value of 0.357**, we concluded it as the best fit model
6. The final regression equation: **$Y = -.121 + 0.016\ (\log(x))^2$**
7. The final regression plot shows a good fit, and the residual plot shows an almost symmetrical distribution, with the tendency to cluster in the middle


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: PlasticWastePerCapita
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0
Netherlands
Sri Lanka
Germany
United States
Guatemala
South Africa
Portugal
Spain
Senegal
Turkey
Democratic Republic of Congo
Egypt
Malaysia
Sudan
Brazil
Cote d'Ivoire
Vietnam
Ecuador
Colombia
Peru
Cambodia
Greece
Algeria
Thailand
Myanmar
Mozambique
Iran
China
Argentina
Iraq
Guinea
United Kingdom
Cameroon
France
Philippines
Morocco
Madagascar
Japan
Kenya
Bangladesh
Angola
Ghana
Indonesia
Russia
Chile
Italy
Mexico
Saudi Arabia
India
South Korea
Australia
Canada
Belgium


Scatterplot
Dependent Variable: PlasticWastePerCapita
Regression Standardized Residual
Regression Standardized Predicted Value
4
3
2
1
0
- 1
- 2
- 3
- 2
- 1
0
1
2
Germany
Sri Lanka
Netherlands
Guatemala
South Africa
United States
Cote d'Ivoire
Spain
Sudan
Egypt
Turkey
Portugal
Vietnam
Senegal
Nigeria
Malaysia
Democratic Republic of Congo
Cambodia
Peru
Ecuador
Colombia
Argentina
Pakistan
Brazil
Greece
Mozambique
Guinea
China
United Kingdom
Myanmar
Philippines
Algeria
Thailand
Madagascar
Cameroon
Iran
Chile
Japan
France
Tanzania
Ghana
Iraq
Bangladesh
Morocco
Ukraine
Saudi Arabia
India
Russia
Angola
Kenya
Indonesia
Mexico
Poland
Italy
South Korea
Australia
Romania
Canada
Belgium

Limitations

1. The data represent total plastic waste generation and do not account for differences in waste management, recycling, or incineration. Therefore, we cannot say how much is disposed-off into the ocean or other waterways, and how much is recycled.
2. Our assumption is richer countries will have a better waste disposal and management system.
3. Data on the waste directly disposed-off in waterways will give a much more accurate picture of how much harm we're causing to the ecosystem.
4. The Dataset has a sample size of only 60. A greater sample size is suggested.
5. For Multiple Linear Regression, we could not add more samples while increasing the number of predictors.
6. For further and more in-depth study of this, we should increase the sample size. The larger the sample size the more information we have and so our uncertainty reduces. This will help us increase precision and narrow the confidence intervals.
7. Only Linear Regression Models were used in this study. We could use other models like Exponential Regression to find a better fit and narrower confidence interval for better prediction.

References

1. Dataset: https://ourworldindata.org/plastic-pollution
2. Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., ... & Law, K. L. (2015). Plastic waste inputs from land into the ocean. Science, 347(6223), 768-771.
3. K. L. Law, N. Starr, T. R. Siegler, J. R. Jambeck, N. J. Mallos, and George H. Leonard, "The United States’ contribution of plastic waste to land and ocean," Science Advances, vol. 6, no. 44, p. eabd0288, 2020, doi: 10.1126/sciadv.abd0288.

Appendix

Appendix A – Dataset

Entity	Per capita plastic waste (kg/person/day)	GDP per capita (USD)	Population (2010)
Germany	0.485	46929.99341	80827001
Netherlands	0.424	52032.9865	16682927
Sri Lanka	0.357	9126.865612	20261738
United States	0.335	54315.91248	309011469
Guatemala	0.28	7335.988974	14630420
Spain	0.277	37319.47622	46931011
Portugal	0.265	31798.15381	10596055
South Africa	0.24	12452.33753	51216967
United Kingdom	0.215	42089.01364	63459801
Turkey	0.212	20027.66683	72326992
Greece	0.2	33753.61775	10887640
Malaysia	0.198	20536.37251	28208028
France	0.192	42147.67168	62879535
Argentina	0.183	23521.27018	40895751
Egypt	0.178	10340.07323	82761244
Japan	0.171	37586.16715	128542349
Brazil	0.165	14873.20832	195713637
Saudi Arabia	0.156	44036.76735	27421468
Ecuador	0.147	10340.9703	15011114
Iran	0.144	13805.70382	73762519
Thailand	0.144	14399.04451	67195032
Colombia	0.144	11783.29991	45222699
Algeria	0.144	10970.70637	35977451
Peru	0.144	10066.46757	29027680

Tunisia	0.144	10113.37035	10635245
Italy	0.134	42664.35527	59325232
China	0.121	8884.588031	1368810604
Chile	0.119	21262.46699	17062531
Russia	0.112	23961.22029	143479273
South Korea	0.112	34394.49049	49545638
Australia	0.112	44991.78371	22154687
Pakistan	0.103	3906.873971	179424643
Nigeria	0.103	4932.3348	158503203
Vietnam	0.103	5089.411016	87967655
Ukraine	0.103	11778.3137	45792086
Sudan	0.103	3089.572476	34545014
Iraq	0.103	8748.508776	29741977
Cote d'Ivoire	0.103	3660.903847	20532944
Senegal	0.103	2797.11016	12678143
Poland	0.097	23996.13978	38329784
Canada	0.093	44861.52396	34147566
Mexico	0.087	17790.01192	114092961
Belgium	0.08	47971.63031	10938735
Philippines	0.075	5918.373229	93966784
Myanmar	0.075	3129.920255	50600827
Morocco	0.073	6281.464191	32343384
Cambodia	0.066	2716.699738	14312205
Angola	0.062	7692.434286	23356247
Indonesia	0.057	8286.73287	241834226
Cameroon	0.046	3086.221474	20341236
Democratic Republic of Congo	0.045	865.6840402	64563853
Romania	0.042	20303.1628	20471860
Ghana	0.04	3729.475903	24779614
Bangladesh	0.034	2883.466794	147575433
Guinea	0.03	1870.800051	10192168
Kenya	0.027	3329.853752	42030684
Tanzania	0.023	2006.971326	44346532
Madagascar	0.016	1553.404722	21151640
Mozambique	0.015	1027.208877	23531567
India	0.01	4234.979573	1234281163

Appendix B – ANOVA Output Tables & Confidence/Prediction Interval Table for SLR

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.579[a]	.335	.323	.0804897

a. Predictors: (Constant), GDPPerCapita

b. Dependent Variable: PlasticPerCapita

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.189	1	.189	29.184	<.001[b]
	Residual	.376	58	.006		
	Total	.565	59			

a. Dependent Variable: PlasticPerCapita

b. Predictors: (Constant), GDPPerCapita

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.074	.016		4.756	<.001
	GDPPerCapita	3.544E-6	.000	.579	5.402	<.001

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	.043	.105
	GDPPerCapita	.000	.000

a. Dependent Variable: PlasticPerCapita

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.076920	.266324	.136183	.0566091	60
Residual	-.1638424	.2508064	.0000000	.0798046	60
Std. Predicted Value	-1.047	2.299	.000	1.000	60
Std. Residual	-2.036	3.116	.000	.991	60

a. Dependent Variable: PlasticPerCapita

GDP Per Capita (X)	Predictions	Confidence Intervals over Mean		Prediction Intervals over Individual value	
1500	0.07917	0.04952	0.10881	-0.08466	0.24299
3000	0.08448	0.0562	0.11276	-0.0791	0.24806
6000	0.09511	0.06934	0.12089	-0.06805	0.25828
12000	0.11637	0.09432	0.13843	-0.04625	0.279
24000	0.1589	0.13646	0.18134	-0.00378	0.32157

Appendix C – ANOVA Output Tables for Transformed Logarithmic Regression

Descriptive Statistics

	Mean	Std. Deviation	N
PlasticWastePerCapita	.136183	.0978436	60
GDPPerCapita	17590.0200	15975.2115	60
LogGPC	4.027655	.4785812	60
LogPopulation	7.659646	.4593598	60

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.611[a]	.374	.340	.0794760	.374	11.141
2	.610[b]	.372	.350	.0789106	-.002	.192
3	.605[c]	.367	.356	.0785415	-.005	.459

Model Summary[d]

Model	Change Statistics		
	df1	df2	Sig. F Change
1	3	56	<.001
2	1	56	.663
3	1	57	.501

a. Predictors: (Constant), LogPopulation, LogGPC, GDPPerCapita

b. Predictors: (Constant), LogGPC, GDPPerCapita

c. Predictors: (Constant), LogGPC

d. Dependent Variable: PlasticWastePerCapita

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.211	3	.070	11.141	<.001[b]
	Residual	.354	56	.006		
	Total	.565	59			
2	Regression	.210	2	.105	16.854	<.001[c]
	Residual	.355	57	.006		
	Total	.565	59			
3	Regression	.207	1	.207	33.563	<.001[d]
	Residual	.358	58	.006		
	Total	.565	59			

a. Dependent Variable: PlasticWastePerCapita

b. Predictors: (Constant), LogPopulation, LogGPC, GDPPerCapita

c. Predictors: (Constant), LogGPC, GDPPerCapita

d. Predictors: (Constant), LogGPC

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.188	.238		-.791	.432
	GDPPerCapita	9.492E-7	.000	.155	.616	.540
	LogGPC	.095	.051	.466	1.854	.069
	LogPopulation	-.010	.023	-.047	-.438	.663
2	(Constant)	-.255	.181		-1.413	.163
	GDPPerCapita	1.028E-6	.000	.168	.677	.501
	LogGPC	.093	.051	.453	1.829	.073
3	(Constant)	-.362	.087		-4.182	<.001
	LogGPC	.124	.021	.605	5.793	<.001

Coefficients[a]

Model		95.0% Confidence Interval for B Lower Bound	Upper Bound	Correlations Zero-order	Partial	Part
1	(Constant)	-.664	.288			
	GDPPerCapita	.000	.000	.579	.082	.065
	LogGPC	-.008	.198	.605	.240	.196
	LogPopulation	-.055	.036	-.042	-.058	-.046
2	(Constant)	-.617	.106			
	GDPPerCapita	.000	.000	.579	.089	.071
	LogGPC	-.009	.194	.605	.235	.192
3	(Constant)	-.536	-.189			
	LogGPC	.081	.167	.605	.605	.605

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	(Constant)		
	GDPPerCapita	.177	5.652
	LogGPC	.177	5.649
	LogPopulation	.986	1.014
2	(Constant)		
	GDPPerCapita	.179	5.574
	LogGPC	.179	5.574
3	(Constant)		
	LogGPC	1.000	1.000

a. Dependent Variable: PlasticWastePerCapita

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.001228	.223729	.136183	.0592382	60
Std. Predicted Value	-2.278	1.478	.000	1.000	60
Standard Error of Predicted Value	.010	.025	.014	.003	60
Adjusted Predicted Value	-.003887	.224000	.136111	.0592939	60
Residual	-.1370517	.2691284	.0000000	.0778730	60
Std. Residual	-1.745	3.427	.000	.991	60
Stud. Residual	-1.789	3.511	.000	1.010	60
Deleted Residual	-.1440001	.2825015	.0000725	.0807992	60
Stud. Deleted Residual	-1.824	3.922	.014	1.057	60
Mahal. Distance	.001	5.190	.983	1.057	60
Cook's Distance	.000	.306	.019	.048	60
Centered Leverage Value	.000	.088	.017	.018	60

a. Dependent Variable: PlasticWastePerCapita

GDP Per Capita (X)	Predictions	Confidence Intervals over Mean		Prediction Intervals over Individual value	
1500	0.03078	-0.01092	0.07247	-0.13187	0.19343
3000	0.06804	0.03695	0.09912	-0.09222	0.2283
6000	0.1053	0.08237	0.12823	-0.05358	0.26418
12000	0.14256	0.12215	0.16298	-0.01598	0.3011
24000	0.17982	0.15454	0.20511	0.02058	0.33906

Appendix D – ANOVA Output Tables for Transformed Quadratic Regression

Descriptive Statistics

	Mean	Std. Deviation	N
PlasticWastePerCapita	.136183	.0978436	60
LogGPCSquared	16.447224	3.7966579	60

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	LogGPCSquared[b]	.	Enter

a. Dependent Variable: PlasticWastePerCapita

b. All requested variables entered.

Model Summary[b]

					Change Statistics	
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change
1	.607[a]	.368	.357	.0784563	.368	33.762

Model Summary[b]

	Change Statistics		
Model	df1	df2	Sig. F Change
1	1	58	<.001

a. Predictors: (Constant), LogGPCSquared

b. Dependent Variable: PlasticWastePerCapita

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.208	1	.208	33.762	<.001[b]
	Residual	.357	58	.006		
	Total	.565	59			

a. Dependent Variable: PlasticWastePerCapita

b. Predictors: (Constant), LogGPCSquared

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.121	.045		-2.664	.010
	LogGPCSquared	.016	.003	.607	5.810	<.001

GDP Per Capita (X)	Predictions	Confidence Intervals over Mean		Prediction Intervals over Individual value	
1500	0.03677	-0.00303	0.07657	-0.12524	0.19878
3000	0.06808	0.03707	0.09909	-0.092	0.22816
6000	0.10222	0.07881	0.12563	-0.05656	0.261
12000	0.13919	0.11889	0.15949	-0.01916	0.29755
24000	0.179	0.15393	0.20407	0.01996	0.33804

Appendix E – ANOVA Output Tables for Multi Linear Regression – 3 Variables

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.600[a]	.360	.326	.0803263

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.203	3	.068	10.513	<.001[b]
	Residual	.361	56	.006		
	Total	.565	59			

a. Dependent Variable: PlasticWastePerCapita

b. Predictors: (Constant), GDP, GDPPerCapita, Population

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.083	.017		4.834	<.001
	GDPPerCapita	2.842E-6	.000	.464	3.515	<.001
	Population	-8.956E-11	.000	-.213	-1.343	.185
	GDP	8.730E-15	.000	.245	1.434	.157

Coefficients[a]

Model		95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound	Correlations Zero-order	Correlations Partial	Correlations Part
1	(Constant)	.049	.117			
	GDPPerCapita	.000	.000	.579	.425	.376
	Population	.000	.000	-.102	-.177	-.144
	GDP	.000	.000	.286	.188	.153

Coefficients[a]

Model		Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)		
	GDPPerCapita	.655	1.526
	Population	.453	2.205
	GDP	.391	2.557

a. Dependent Variable: PlasticWastePerCapita

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.029979	.356068	.136183	.0587294	60
Std. Predicted Value	-1.808	3.744	.000	1.000	60
Standard Error of Predicted Value	.011	.072	.017	.011	60
Adjusted Predicted Value	.052182	.442571	.137566	.0649223	60
Residual	-.1427896	.2484075	.0000000	.0782575	60
Std. Residual	-1.778	3.092	.000	.974	60
Stud. Residual	-1.890	3.143	-.005	1.008	60
Deleted Residual	-.1613393	.2624228	-.0013827	.0846339	60
Stud. Deleted Residual	-1.935	3.432	.007	1.049	60
Mahal. Distance	.037	46.462	2.950	7.857	60
Cook's Distance	.000	.361	.023	.062	60
Centered Leverage Value	.001	.787	.050	.133	60

a. Dependent Variable: PlasticWastePerCapita

GDP Per Capita (X)	Predictions	Confidence Intervals over Mean		Prediction Intervals over Individual value	
1500	0.08559	0.0539	0.11727	-0.07841	0.24959
3000	0.08885	0.05938	0.11831	-0.07474	0.25243
6000	0.09693	0.07092	0.12294	-0.06607	0.25993
12000	0.11757	0.09466	0.14047	-0.04497	0.2801
24000	0.15706	0.13403	0.18009	-0.00549	0.31961

Appendix F – ANOVA Output Tables for Multi Regression - Transformed 4 Variables

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.627[a]	.394	.349	.0789178	.394	8.923

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.222	4	.056	8.923	<.001[b]
	Residual	.343	55	.006		
	Total	.565	59			

a. Dependent Variable: PlasticWastePerCapita

b. Predictors: (Constant), LogPopulation, GPCSqrt, LogGPC, GDPPerCapita

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.067	.697		-1.530	.132
	GDPPerCapita	1.482E-5	.000	2.420	1.416	.162
	LogGPC	.448	.268	2.193	1.670	.101
	GPCSqrt	-.006	.005	-3.902	-1.340	.186
	LogPopulation	-.014	.023	-.066	-.619	.538

Coefficients[a]

Model		95.0% Confidence Interval for B Lower Bound	Upper Bound	Correlations Zero-order	Partial	Part
1	(Constant)	-2.463	.330			
	GDPPerCapita	.000	.000	.579	.188	.149
	LogGPC	-.090	.987	.605	.220	.175
	GPCSqrt	-.016	.003	.599	-.178	-.141
	LogPopulation	-.060	.031	-.042	-.083	-.065

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	(Constant)		
	GDPPerCapita	.004	264.782
	LogGPC	.006	156.421
	GPCSqrt	.001	769.103
	LogPopulation	.968	1.033

a. Dependent Variable: PlasticWastePerCapita

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-.034081	.263300	.136183	.0613807	60
Std. Predicted Value	-2.774	2.071	.000	1.000	60
Standard Error of Predicted Value	.015	.051	.022	.007	60
Adjusted Predicted Value	-.092456	.273526	.134771	.0636474	60
Residual	-.1677835	.2539112	.0000000	.0761957	60
Std. Residual	-2.126	3.217	.000	.966	60
Stud. Residual	-2.283	3.387	.008	1.019	60
Deleted Residual	-.1935260	.2813257	.0014122	.0852253	60
Stud. Deleted Residual	-2.378	3.772	.018	1.063	60
Mahal. Distance	1.138	24.073	3.933	3.996	60
Cook's Distance	.000	.258	.025	.058	60
Centered Leverage Value	.019	.408	.067	.068	60

a. Dependent Variable: PlasticWastePerCapita

GDP Per Capita (X)	Predictions	Confidence Intervals over Mean		Prediction Intervals over Individual value	
1500	0.03019	-0.02597	0.08635	-0.13764	0.19802
3000	0.08088	0.04536	0.11641	-0.08121	0.24298
6000	0.11143	0.07789	0.14497	-0.05024	0.2731
12000	0.13499	0.10456	0.16543	-0.02606	0.29605
24000	0.15713	0.11754	0.19671	-0.00591	0.32016

Appendix G – ANOVA Output Tables for Multi Regression - Transformed 2 Variables

Model Summary [b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.606 [a]	.368	.346	.0791523	.368	16.577

ANOVA [a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.208	2	.104	16.577	<.001 [b]
	Residual	.357	57	.006		
	Total	.565	59			

a. Dependent Variable: PlasticWastePerCapita

b. Predictors: (Constant), LogGPCSqrt, LogGPC

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.504	2.633		.191	.849
	LogGPC	.345	.673	1.688	.513	.610
	LogGPCSqrt	-.877	2.666	-1.083	-.329	.743

Coefficients[a]

Model		95.0% Confidence Interval for B Lower Bound	Upper Bound	Correlations Zero-order	Partial	Part
1	(Constant)	-4.769	5.776			
	LogGPC	-1.002	1.692	.605	.068	.054
	LogGPCSqrt	-6.215	4.460	.604	-.044	-.035

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	(Constant)		
	LogGPC	.001	976.236
	LogGPCSqrt	.001	976.236

a. Dependent Variable: PlasticWastePerCapita

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.013868	.228776	.136183	.0593352	60
Std. Predicted Value	-2.061	1.561	.000	1.000	60
Standard Error of Predicted Value	.013	.046	.017	.006	60
Adjusted Predicted Value	-.002179	.232013	.135810	.0596019	60
Residual	-.1410663	.2652898	.0000000	.0777992	60
Std. Residual	-1.782	3.352	.000	.983	60
Stud. Residual	-1.850	3.474	.002	1.011	60
Deleted Residual	-.1520127	.2849689	.0003733	.0823022	60
Stud. Deleted Residual	-1.891	3.878	.015	1.059	60
Mahal. Distance	.499	19.084	1.967	2.929	60
Cook's Distance	.000	.298	.020	.049	60
Centered Leverage Value	.008	.323	.033	.050	60

a. Dependent Variable: PlasticWastePerCapita

GDP Per Capita (X)	**Predictions**	**Confidence Intervals over Mean**		**Prediction Intervals over Individual value**	
1500	**0.04216**	**-0.0024**	**0.08672**	**-0.119**	**0.20331**
3000	**0.07397**	**0.04247**	**0.10546**	**-0.08408**	**0.23201**
6000	**0.10497**	**0.08231**	**0.12763**	**-0.05155**	**0.26149**
12000	**0.13973**	**0.1179**	**0.16156**	**-0.01667**	**0.29613**
24000	**0.17535**	**0.14938**	**0.20132**	**0.01832**	**0.33239**