#1 Histogram and Q-Q plot of Personal Income to check the Normal Population Assumption without transformation.

Chart, histogram

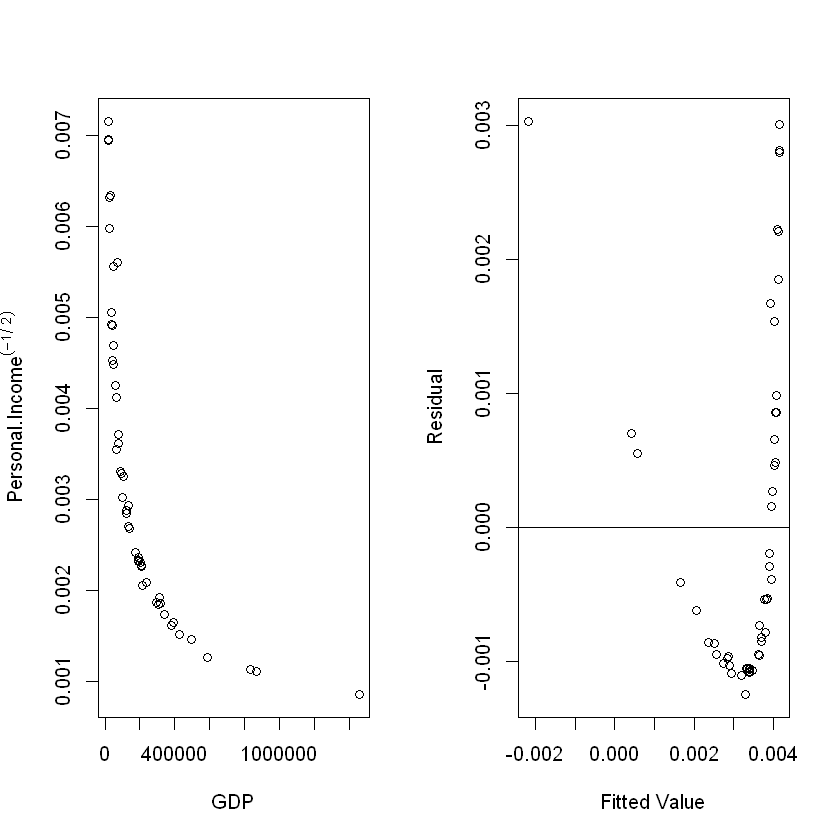
Description automatically generated

#1 Cont: Histogram and Q-Q plot of Personal Income to check the Normal Population Assumption with transformation of -1/2 on the Y-axis as this transformation make the plots more Normal and linear.

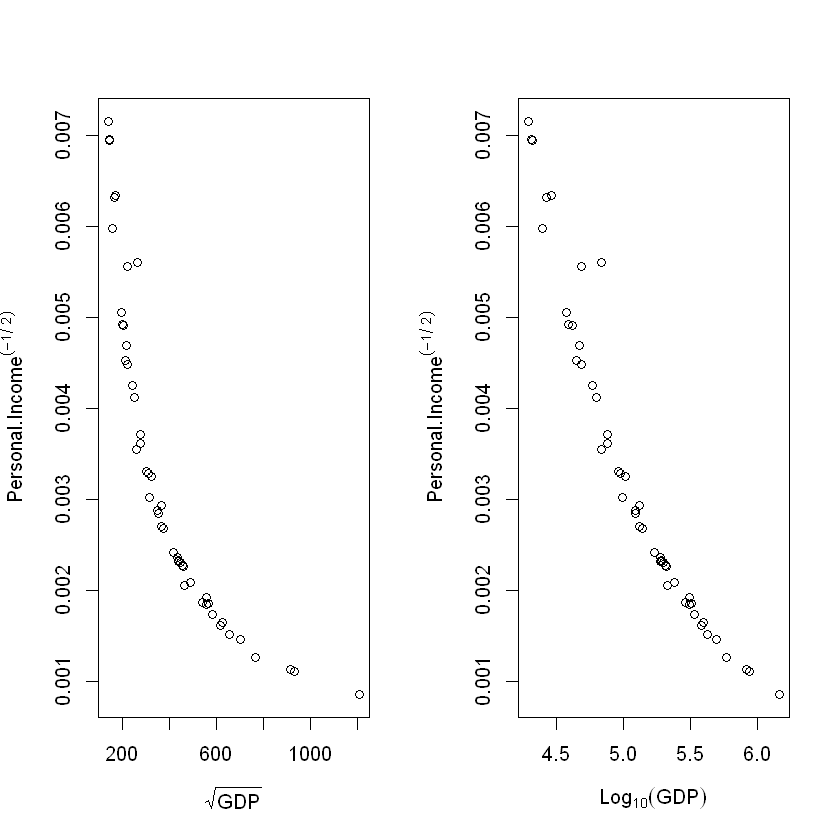
Chart, histogram

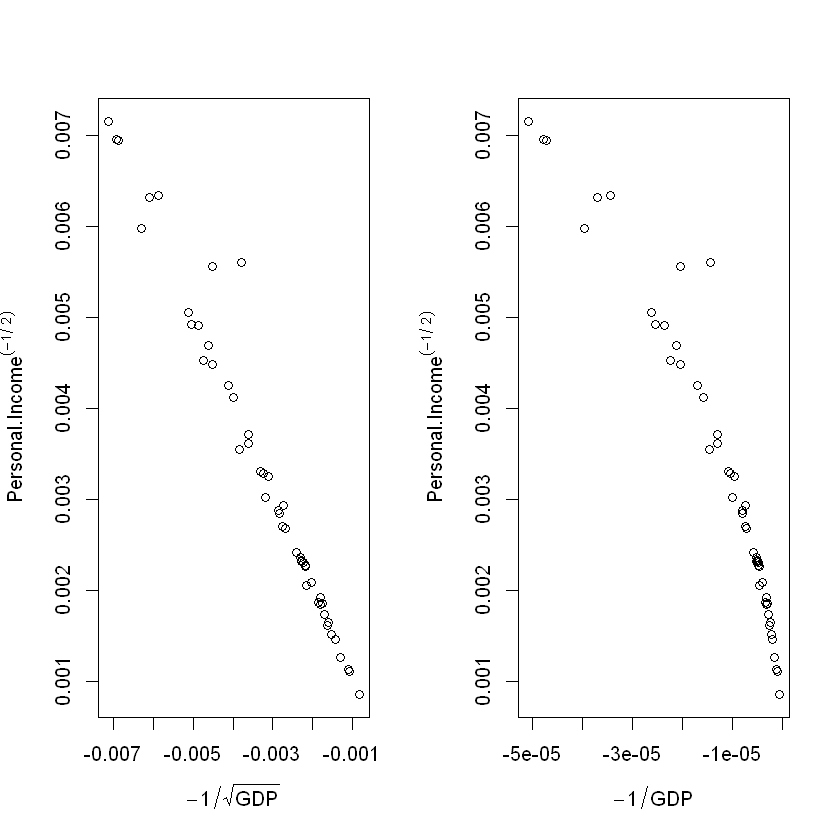
Description automatically generated

#2 Scatterplot of the transformed Personal Income against GDP along with Residuals to check linearity. None of these graphs have transformations and both are curved.

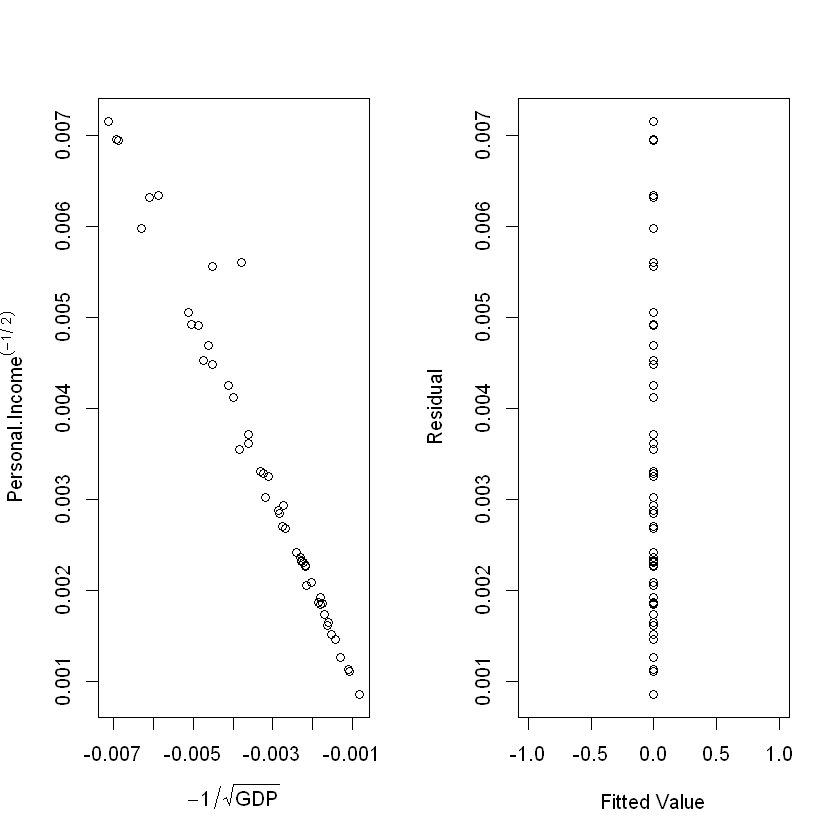


#2 Cont: Scatterplot of the transformed Personal Income against GDP along with Residuals to check linearity. These graphs have transformations as listed in the options. The best transformation which made the results linear was (-1/sqrt) on the x axis along with the prior mentioned transformation on the y-axis.





#3 Plot of Residuals and fitted values of most accurate transformation of plots.



#4 Base on the above graphs there seems to be no issues of leverage as the none of the points seems largely separated from the mean of the x-values. There are a few values towards the top portion of the plot which may seem like outliers but do not have large enough residuals to be considered. There are also no points which if taken away would incredibly influence the R^2 value of the plot.

#5 The 95% Prediction interval for a GDP of 300,000:

|  |  |
| --- | --- |
| (-0.007446221 | 0.007446221) |
|  |  |

Code:

#1:

Without Transformation

par(mfrow = c(1, 2))

hist(GDP1$Personal.Income, xlab = "GDP", main = "")

qqnorm(GDP1$Personal.Income)

qqline(GDP1$Personal.Income)

With Transformation

par(mfrow = c(1, 2))

hist(GDP1$Personal.Income^(-1/2), xlab = expression(Personal.Income^(-1/2)), main = "")

qqnorm(GDP1$Personal.Income^(-1/2))

qqline(GDP1$Personal.Income^(-1/2))

#2

Without Transformation

imod1 <- lm(Personal.Income^(-1/2) ~ GDP, data = GDP1)

par(mfrow = c(1,2))

plot(GDP1$GDP, GDP1$Personal.Income^(-1/2), xlab = 'GDP', ylab = expression(Personal.Income^(-1/2)))

plot(imod1$fitted.values, imod1$residuals, xlab = 'Fitted Value', ylab = 'Residual')

abline(0,0)

With Transformation

par(mfrow = c(1, 2))

plot(sqrt(GDP1$GDP), GDP1$Personal.Income^(-1/2), xlab = expression(sqrt(GDP)),

ylab = expression(Personal.Income^(-1/2)))

plot(log10(GDP1$GDP),GDP1$Personal.Income^(-1/2), xlab = expression(Log[10](GDP)),

ylab = expression(Personal.Income^(-1/2)))

par(mfrow = c(1, 2))

plot(-1/sqrt(GDP1$GDP),GDP1$Personal.Income^(-1/2), xlab = expression(-1/sqrt(GDP)),

ylab = expression(Personal.Income^(-1/2)))

plot(-1/GDP1$GDP, GDP1$Personal.Income^(-1/2), xlab = expression(-1/GDP),

ylab = expression(Personal.Income^(-1/2)))

#3 Desired Transformation and Line

imod2 <- lm(Personal.Income^(-1/2) ~ (-1/sqrt(GDP)), data = GDP1)

par(mfrow = c(1,2))

plot(-1/sqrt(GDP1$GDP), GDP1$Personal.Income^(-1/2), xlab = expression(-1/sqrt(GDP)), ylab = expression(Personal.Income^(-1/2)))

plot(imod2$fitted.values, imod2$residuals, xlab = 'Fitted Value', ylab = 'Residual')

abline(imod2)

#5 95% Prediction Interval

pred.data <- data.frame(GDP = 300000)

result.pred <- predict(imod2, newdata = pred.data, interval = 'prediction', level = 0.95)

result.pred