

Data Assignment 2

Dependent Variable Assigned: v46-Percentage of infant deaths due to Measles (to total reported infant deaths)

Q1) -

NOTE : All graphs, histograms, etc. can be compiled from the R source code, and for the sake of not wanting to be redundant, we are attaching a few graphs here.

(a)

v46	Percentage of infant deaths due to Measles (to total reported infant deaths)
v2	Pregnant women registered for Ante Natal Care within first trimester
v3	Pregnant women received 3 Ante Natal Care (ANC) check-ups
v6	Women having tested moderately anemic with hemoglobin(Hb)<11
v13	Institutional deliveries
v18	Women received postpartum checkup within 48 hours of delivery
v25	Percentage of reported live births (to reported births)
v29	New borns breastfed within 1 hour
v34	Fully immunized children in the age group of 9 to 11 months
v40	Percentage of infant deaths due to Sepsis (to total reported infant deaths)

v41	Percentage of infant deaths due to Asphyxia (to total reported infant deaths)
v42	Percentage of infant deaths due to Low Birth Weight (LBW) (to total reported infant deaths)
v43	Percentage of infant deaths due to Pneumonia (to total reported infant deaths)
v44	Percentage of infant deaths due to Diarrhea (to total reported infant deaths)
v45	Percentage of infant deaths due to Fever (to total reported infant deaths)
v47	Percentage of infant deaths due to other causes (to total reported infant deaths)

Our MLRM :- v46 ~ gdp +beds+ index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42 + v45

```

cat("
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
    v18 + v25 + v29 + v34 + v40 + v41 + v42 + v43 + v44 + v45 +
    v47, data = newData_kh)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.20961 -0.01018 -0.00493  0.00814  0.37897 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.690e+01  1.577e-01 614.497 < 2e-16 ***  
gdp        -1.225e-11  2.893e-11 -0.424   0.6719    
beds       -1.114e-08  1.185e-08 -0.940   0.3471    
index      3.122e-05  7.342e-05  0.425   0.6707    
v2         2.154e-07  9.439e-08  2.282   0.0225 *    
v3         -5.293e-10 8.230e-08 -0.006   0.9949    
v6         -1.678e-07 4.306e-08 -3.898  9.75e-05 ***  
v13        1.041e-07 8.870e-08  1.174   0.2404    
v18        -6.173e-08 8.543e-08 -0.723   0.4699    
v25        -1.609e-04 7.822e-04 -0.206   0.8371    
v29        7.135e-08 9.588e-08  0.744   0.4568    
v34        -9.350e-08 7.211e-08 -1.297   0.1948    
v40        -9.690e-01 1.423e-03 -680.743 < 2e-16 ***  
v41        -9.688e-01 1.424e-03 -680.298 < 2e-16 ***  
v42        -9.687e-01 1.423e-03 -680.738 < 2e-16 ***  
v43        -9.687e-01 1.424e-03 -680.337 < 2e-16 ***  
v44        -9.686e-01 1.427e-03 -678.754 < 2e-16 ***  
v45        -9.688e-01 1.426e-03 -679.239 < 2e-16 ***  
v47        -9.688e-01 1.423e-03 -680.601 < 2e-16 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0713 on 16183 degrees of freedom
Multiple R-squared:  0.9667,    Adjusted R-squared:  0.9667 
F-statistic: 2.612e+04 on 18 and 16183 DF,  p-value: < 2.2e-16

```

We explore many alternatives, the one we found most interesting is :- $v46 \sim gdp + beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v40 + v41 + v42 + v43 + v44 + v45 + v47$

```
> summary(LR)

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
    v18 + v25 + v29 + v34 + v42 + v45, data = newData_kh)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4174 -0.0850 -0.0624 -0.0322  9.0456 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.999e+00  4.158e-01   7.214 5.69e-13 ***  
gdp         1.666e-10  1.563e-10   1.066 0.286617    
beds        -1.411e-07  6.410e-08  -2.201 0.027782 *    
index       -5.636e-04  3.996e-04  -1.411 0.158389    
v2          2.963e-06  5.117e-07   5.791 7.12e-09 ***  
v3          -3.001e-06  4.464e-07  -6.722 1.86e-11 ***  
v6          -7.264e-07  2.309e-07  -3.145 0.001662 **    
v13         3.378e-06  4.801e-07   7.035 2.07e-12 ***  
v18         -1.238e-06  4.639e-07  -2.669 0.007621 **    
v25         -2.961e-02  4.232e-03  -6.997 2.73e-12 ***  
v29         -3.030e-06  5.184e-07  -5.845 5.16e-09 ***  
v34         1.421e-06  3.912e-07   3.633 0.000281 ***  
v42         -5.521e-04  2.355e-04  -2.345 0.019056 *    
v45         1.065e-03  3.534e-04   3.012 0.002600 **    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3882 on 16188 degrees of freedom
Multiple R-squared:  0.01334, Adjusted R-squared:  0.01255 
F-statistic: 16.84 on 13 and 16188 DF,  p-value: < 2.2e-16

> 9.689598e+01 |
```

```

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
    v18 + v25 + v29 + v34 + v40 + v41 + v42 + v43 + v44 + v45 +
    v47, data = newData_kh)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.20961 -0.01018 -0.00493  0.00814  0.37897 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.690e+01 1.577e-01 614.497 < 2e-16 ***  
gdp        -1.225e-11 2.893e-11 -0.424   0.6719    
beds       -1.114e-08 1.185e-08 -0.940   0.3471    
index      3.122e-05 7.342e-05  0.425   0.6707    
v2         2.154e-07 9.439e-08  2.282   0.0225 *    
v3         -5.293e-10 8.230e-08 -0.006   0.9949    
v6         -1.678e-07 4.306e-08 -3.898 9.75e-05 ***  
v13        1.041e-07 8.870e-08  1.174   0.2404    
v18        -6.173e-08 8.543e-08 -0.723   0.4699    
v25        -1.609e-04 7.822e-04 -0.206   0.8371    
v29        7.135e-08 9.588e-08  0.744   0.4568    
v34        -9.350e-08 7.211e-08 -1.297   0.1948    
v40        -9.690e-01 1.423e-03 -680.743 < 2e-16 ***  
v41        -9.688e-01 1.424e-03 -680.298 < 2e-16 ***  
v42        -9.687e-01 1.423e-03 -680.738 < 2e-16 ***  
v43        -9.687e-01 1.424e-03 -680.337 < 2e-16 ***  
v44        -9.686e-01 1.427e-03 -678.754 < 2e-16 ***  
v45        -9.688e-01 1.426e-03 -679.239 < 2e-16 ***  
v47        -9.688e-01 1.423e-03 -680.601 < 2e-16 ***  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0713 on 16183 degrees of freedom
Multiple R-squared:  0.9667,    Adjusted R-squared:  0.9667 
F-statistic: 2.612e+04 on 18 and 16183 DF,  p-value: < 2.2e-16

```

Its R-squared is 96.67%. The reason is that sum of v40 to v47 = 100% and we are regressing for v46 by including the variables v40 to v47 (except v46) in a linear combination. Effectively all these combined into a single variable that is 100% - v46, which is the reason for such a high “fit”.

Hence we discarded this model and only considered the variables that seemed to be more or less independent to each other in some sense (tending towards ceteris paribus).

(b)

{Answer to how the three plots are related}

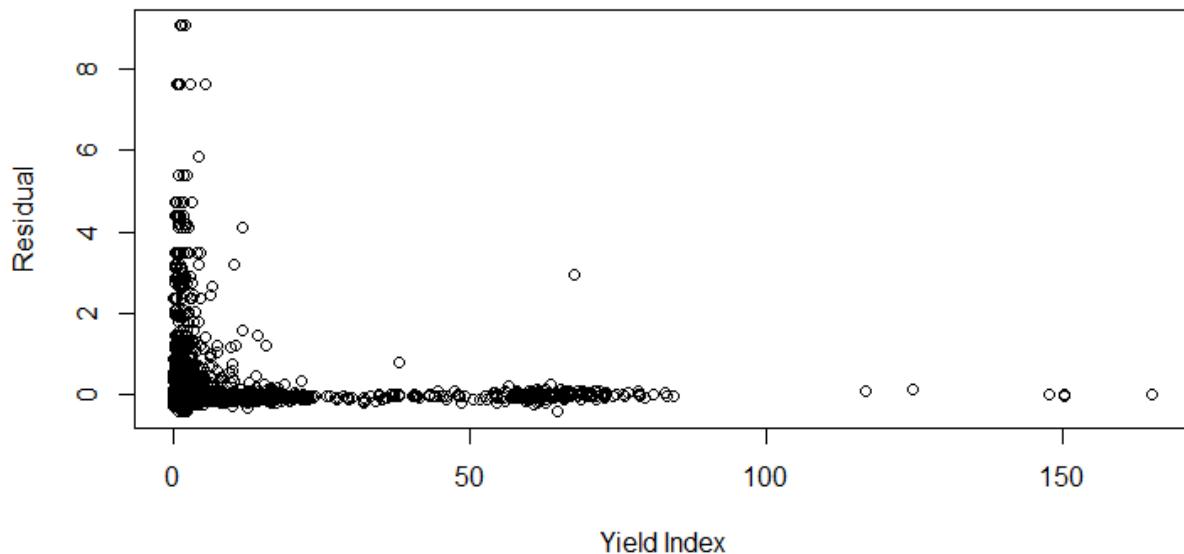
The plot with v46 on y-axis and index on x-axis (Y), the plot with predicted value of v46 on y-axis and index on x-axis (Y_hat), and the plot with residuals on y-axis and index on x-axis (U_hat) are co-related like so :-

$$Y = Y_{\text{hat}} + U_{\text{hat}}$$

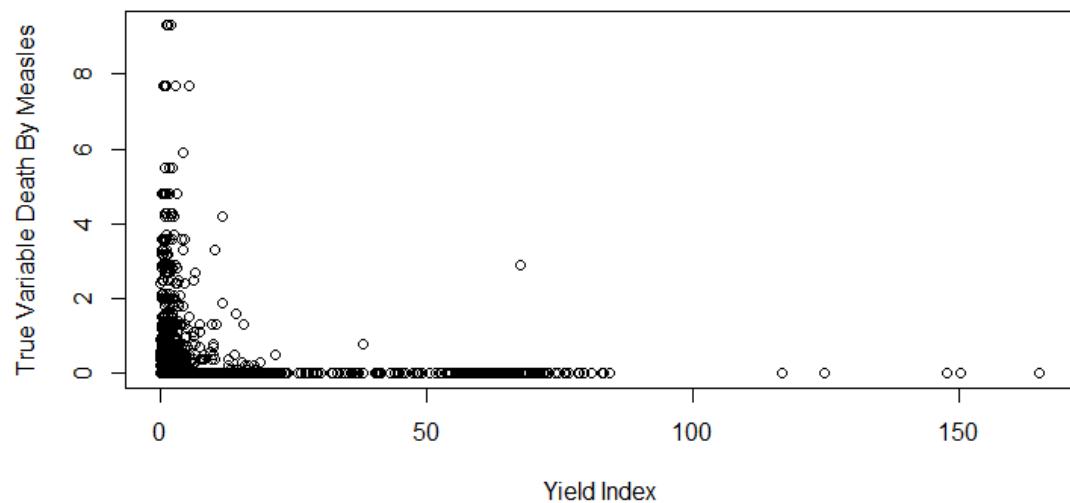
i.e, the superposition of the graphs Y_{hat} and U_{hat} gives us Y .

This is because these Linear Models already use $y = y_{\text{hat}} + u_{\text{hat}}$ as the basis for further analysis.

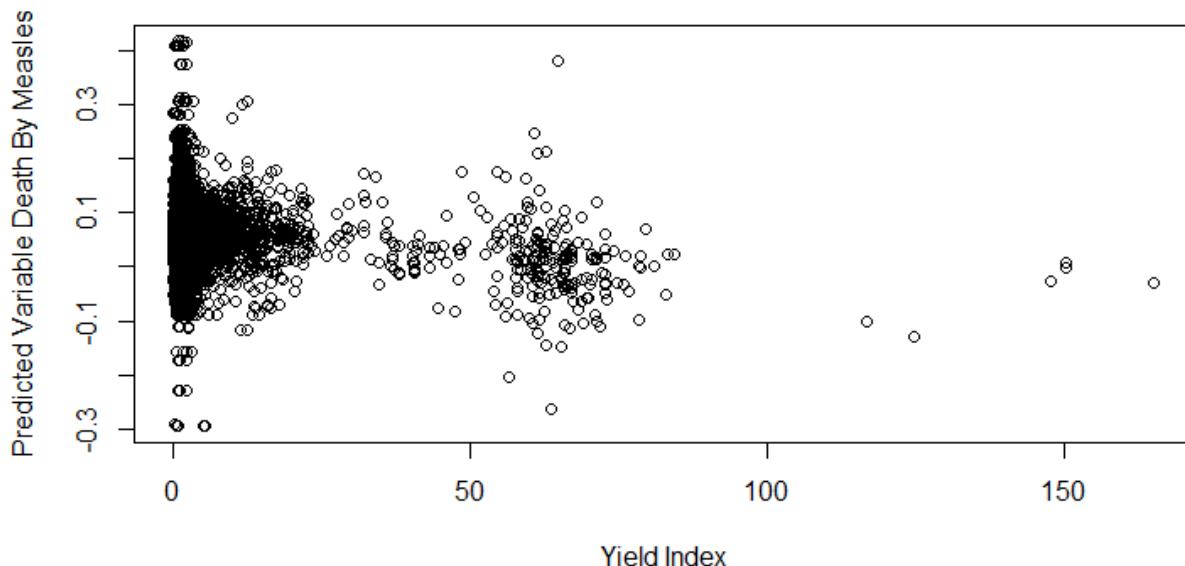
Kharif Plot Residuals VS Yield Index



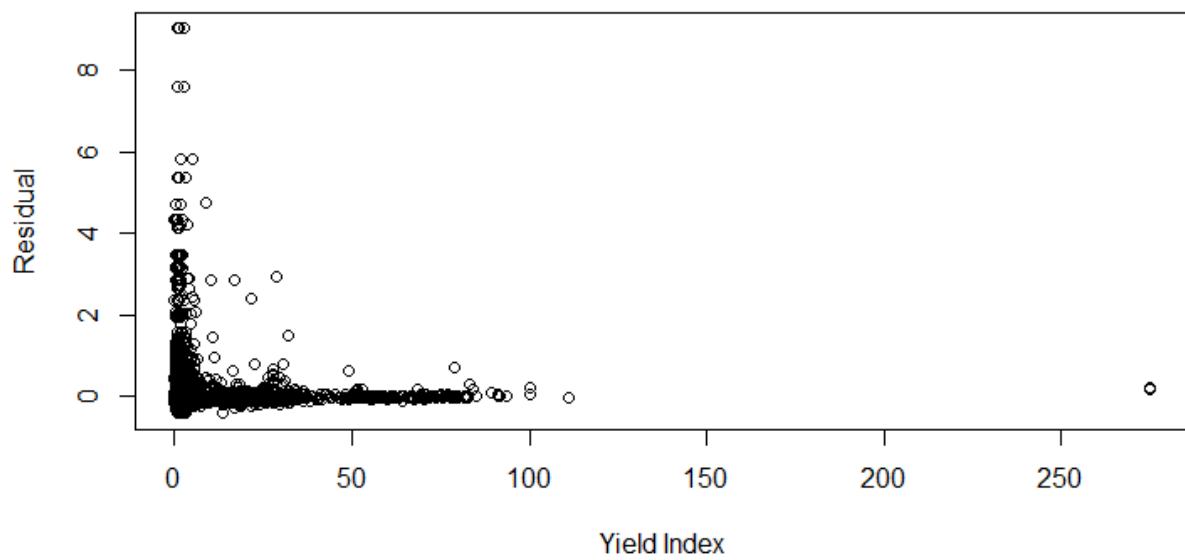
Kharif Plot True Variable Death By Measles VS Yield Index

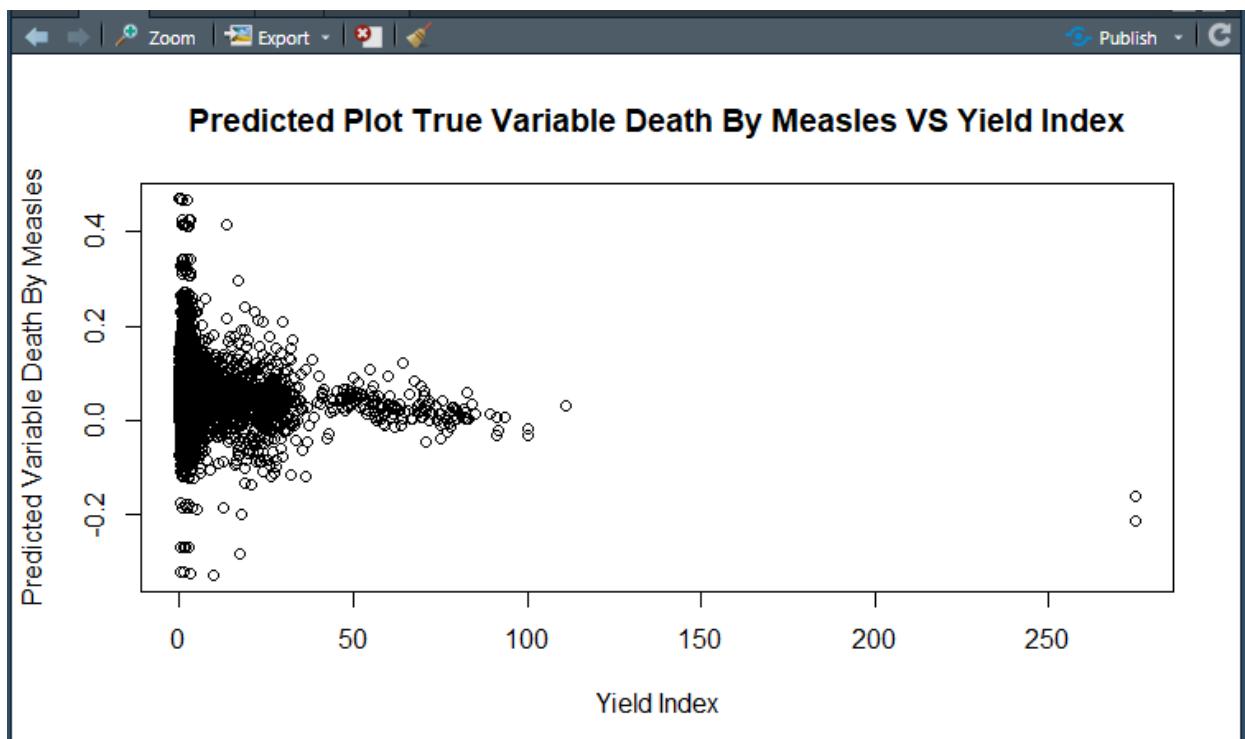
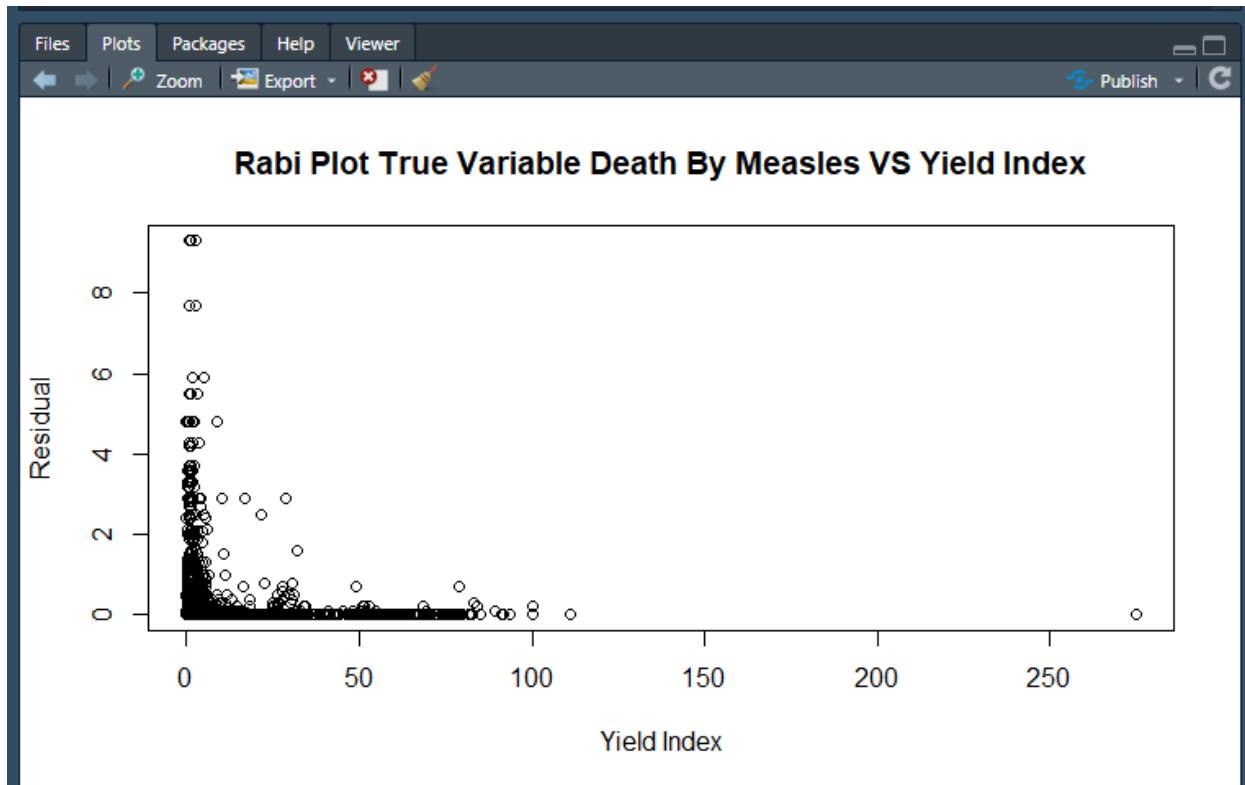


Predicted Plot True Variable Death By Measles VS Yield Index

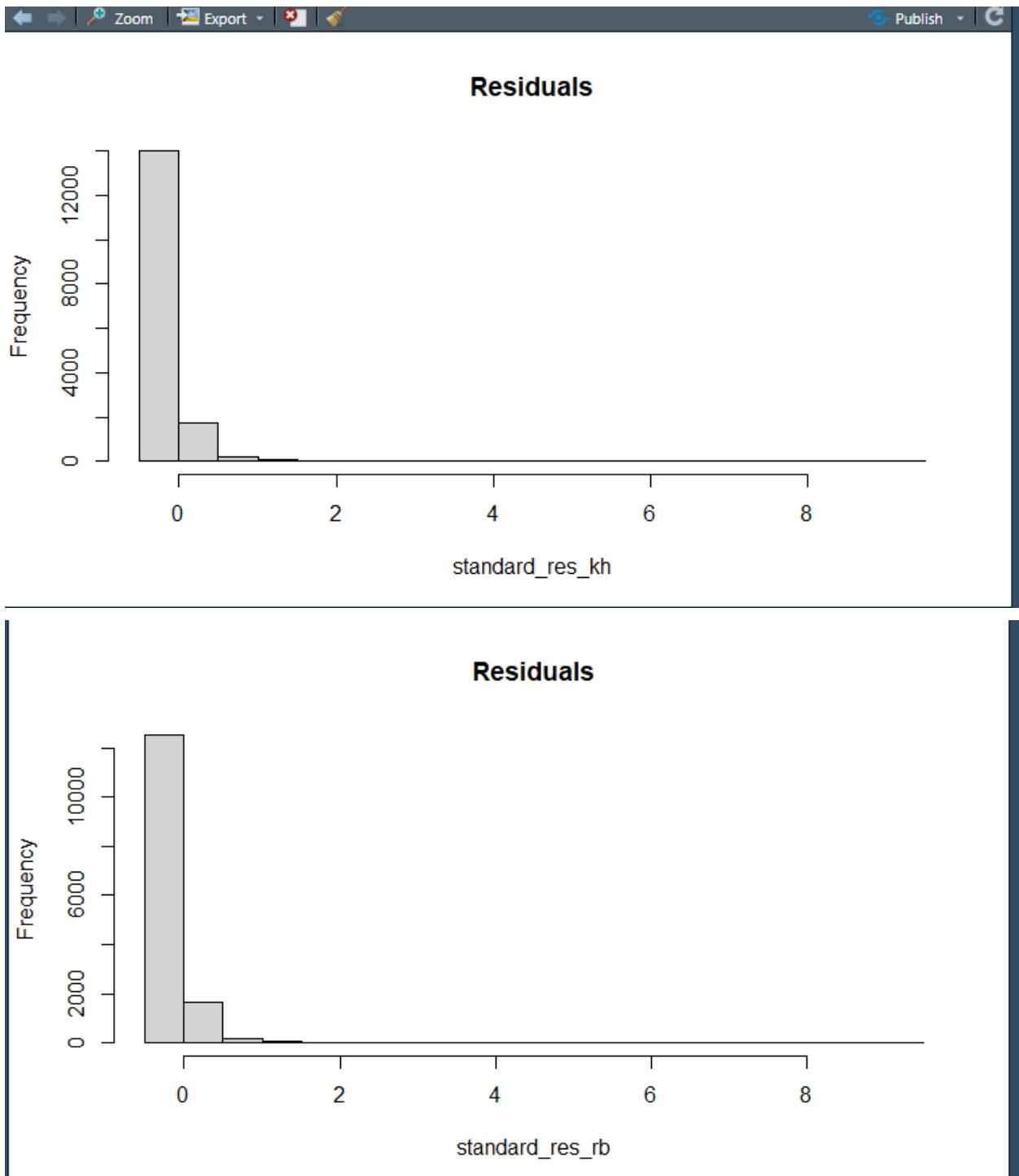


Rabi Plot Residuals VS Yield Index



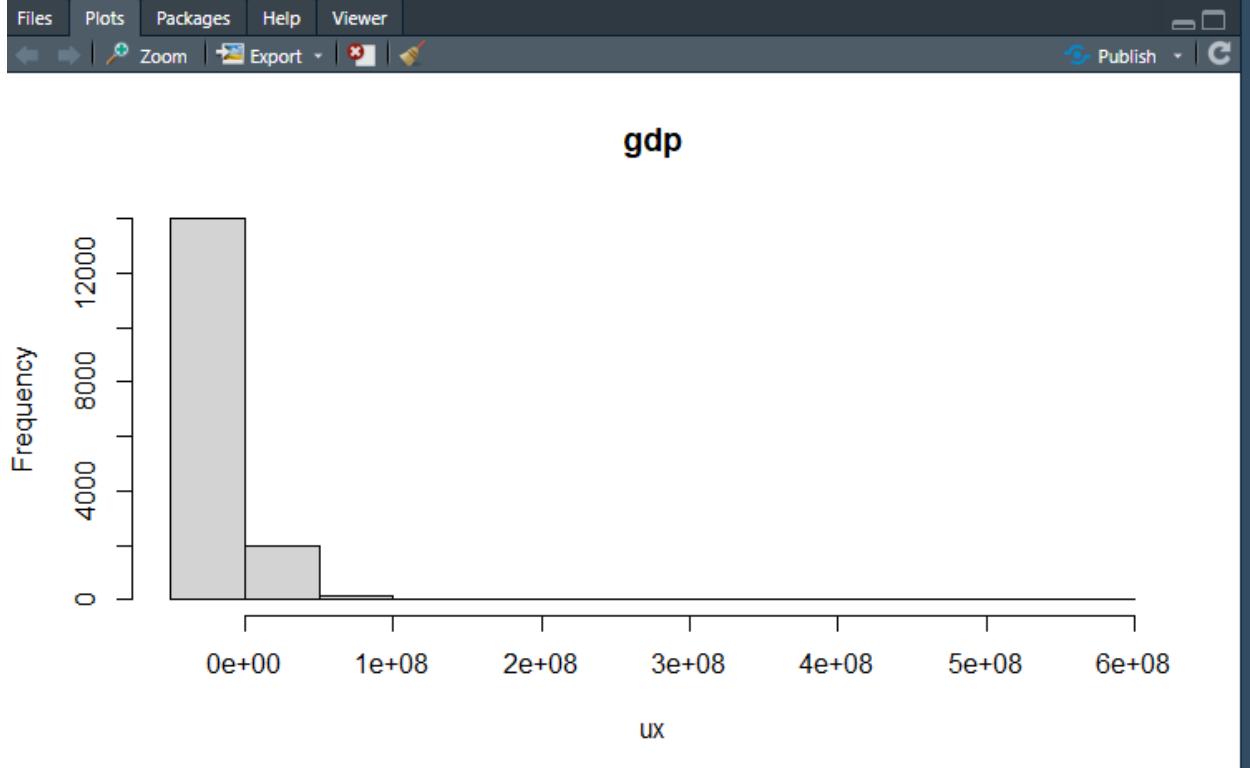


```
> print(sum(standard_res_kh))  
(c) summation u_it = [1] 2.141503e-13
```



(d) KHARIF

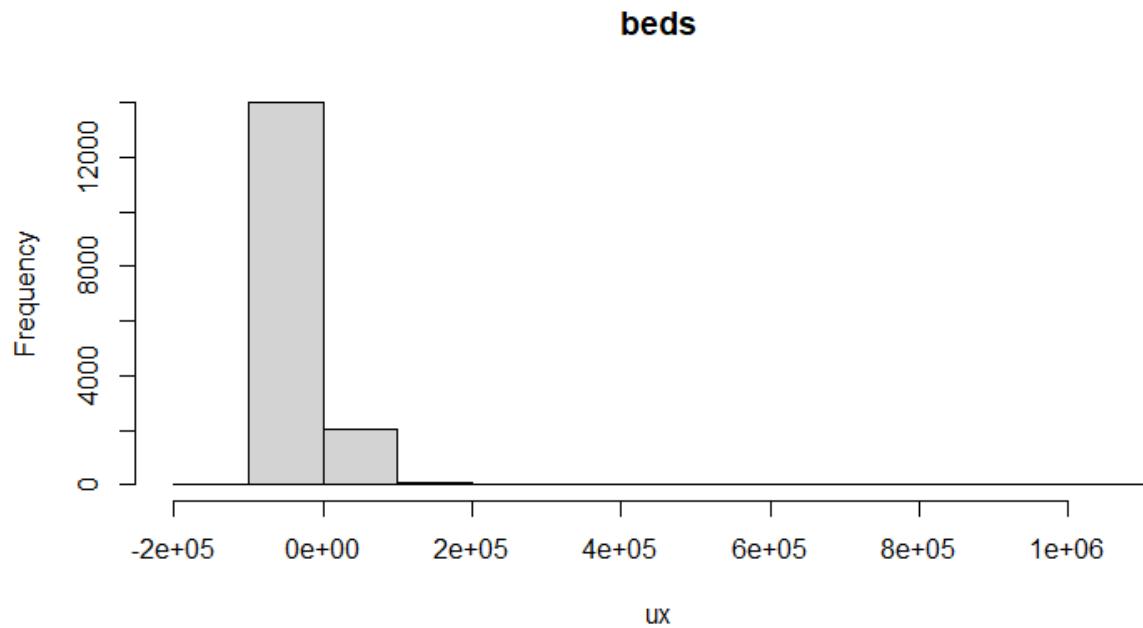
```
> hist(standard_res_kh,main = "Residuals")
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$gdp)
> print(sum(ux))
[1] -6.492244e-05
> hist(ux,main = "gdp")
> |
```



```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$beds)
> print(sum(ux))
[1] 8.488365e-08
> hist(ux,main = "beds")
> |
```

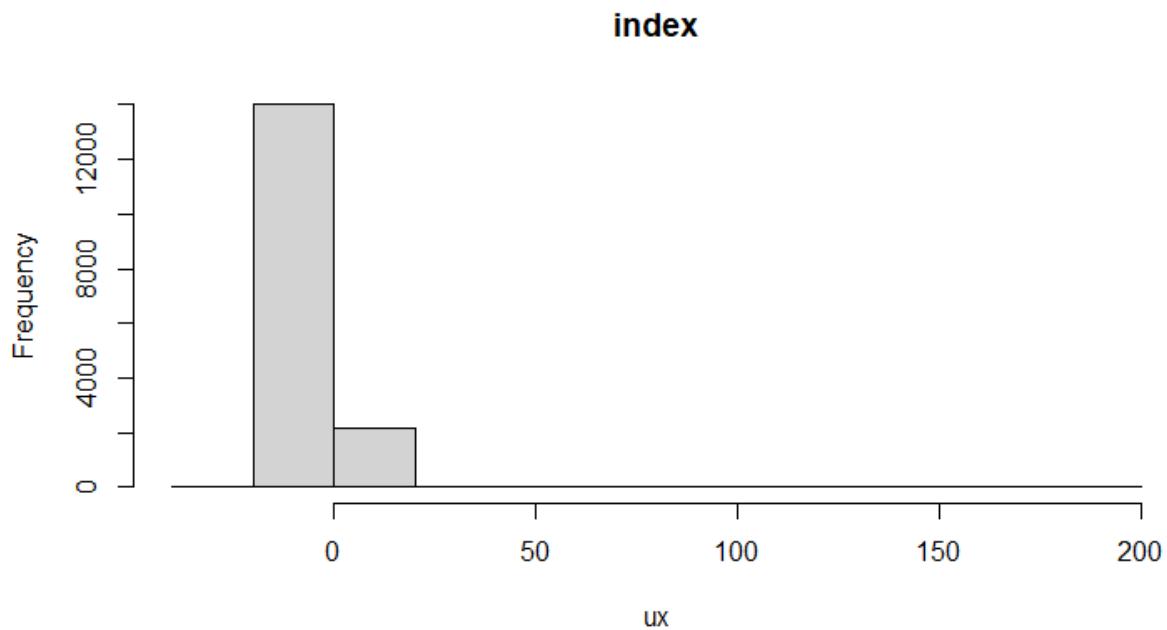
Files Plots Packages Help Viewer | Publish | C

Zoom Export |

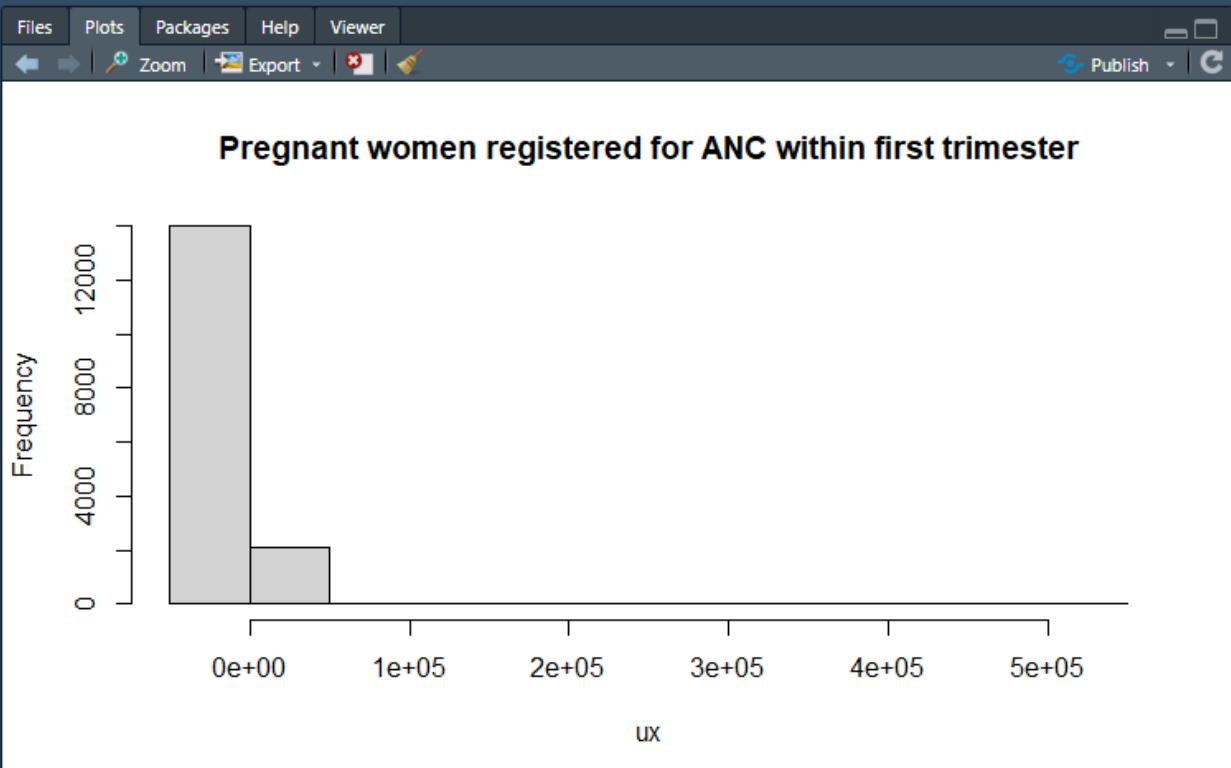


```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh)index  
> print(sum(ux))  
[1] 2.337085e-13  
> hist(ux,main = "index")  
> |
```

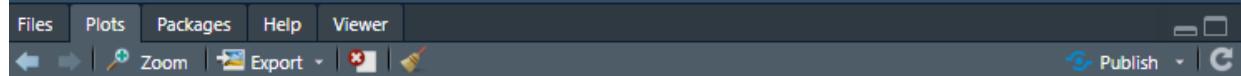
Files Plots Packages Help Viewer |



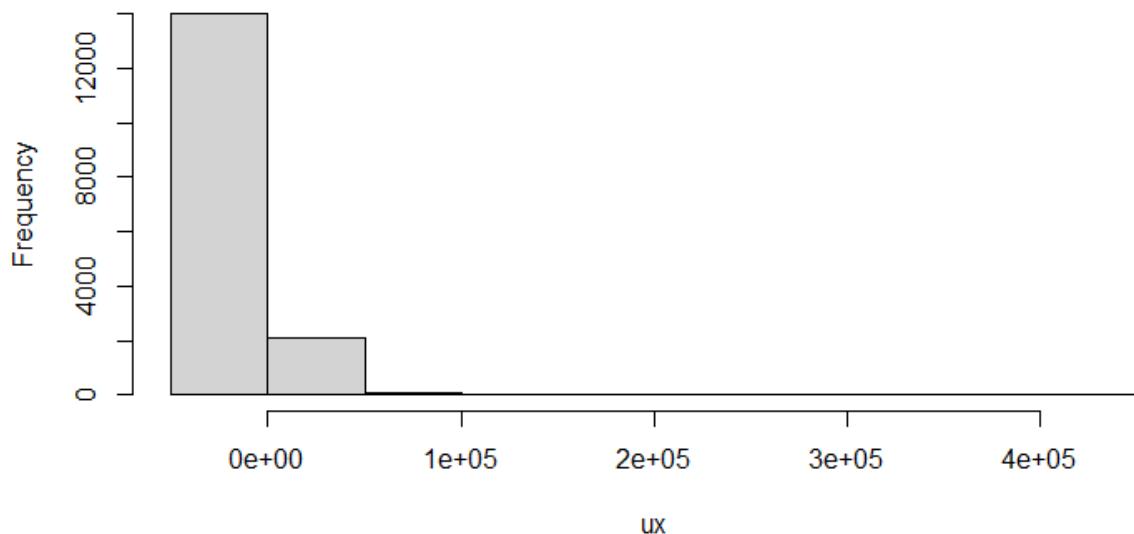
```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v2)
> print(sum(ux))
[1] 7.511801e-09
> hist(ux,main = "Pregnant women registered for ANC within first trimester")
> |
```



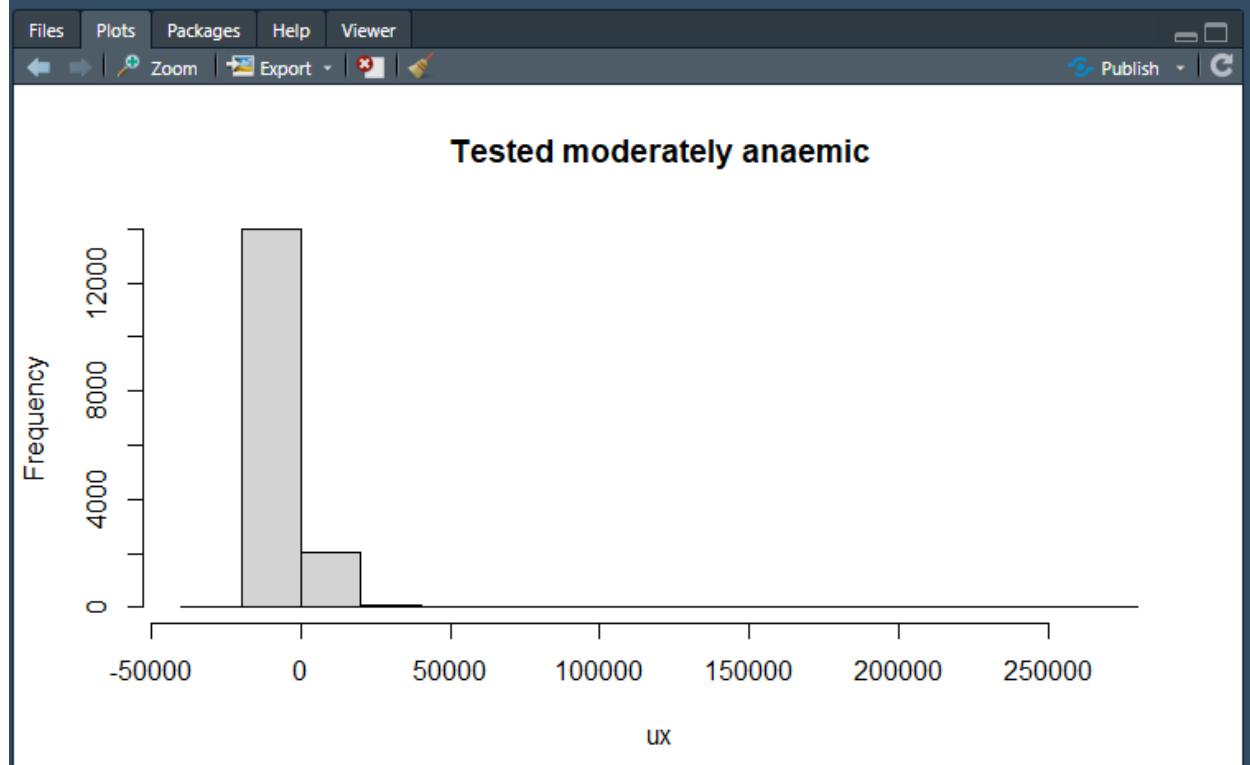
```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v3)
> print(sum(ux))
[1] 1.086423e-08
> hist(ux,main = "Pregnant women received 3 ANC check-ups")
> |
```



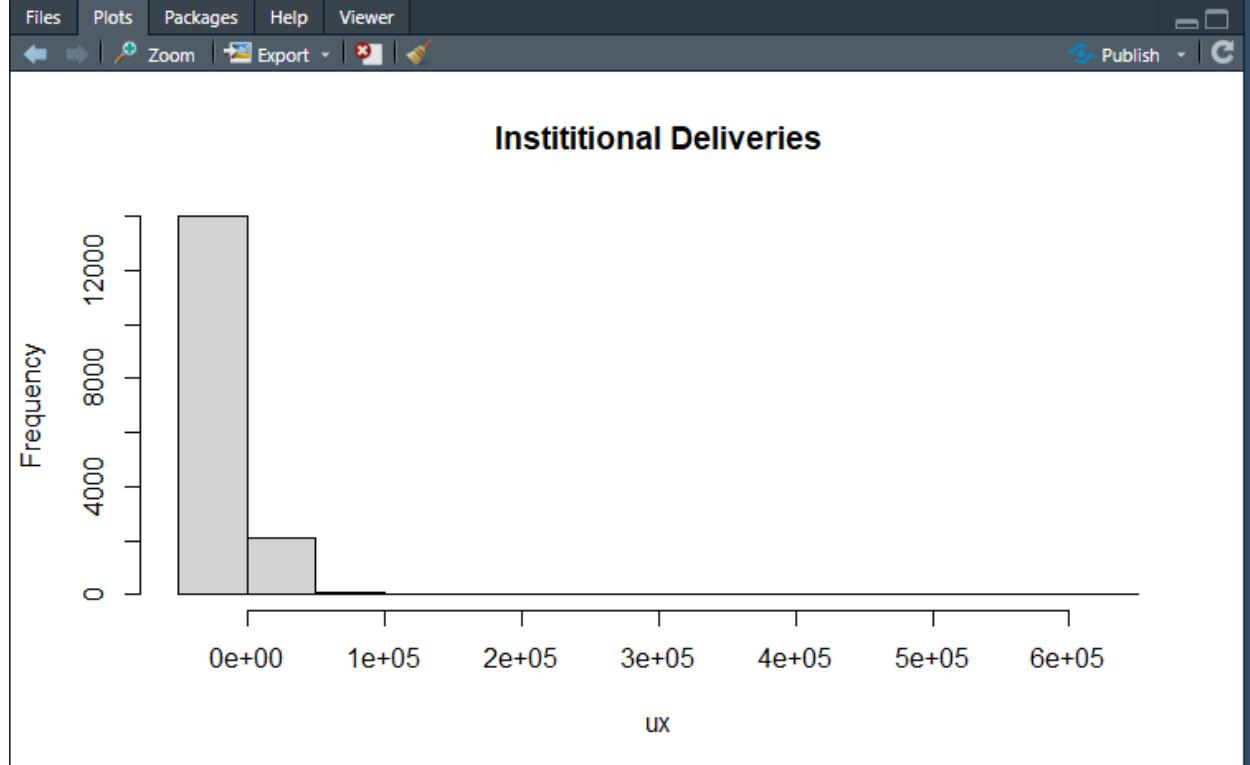
Pregnant women received 3 ANC check-ups



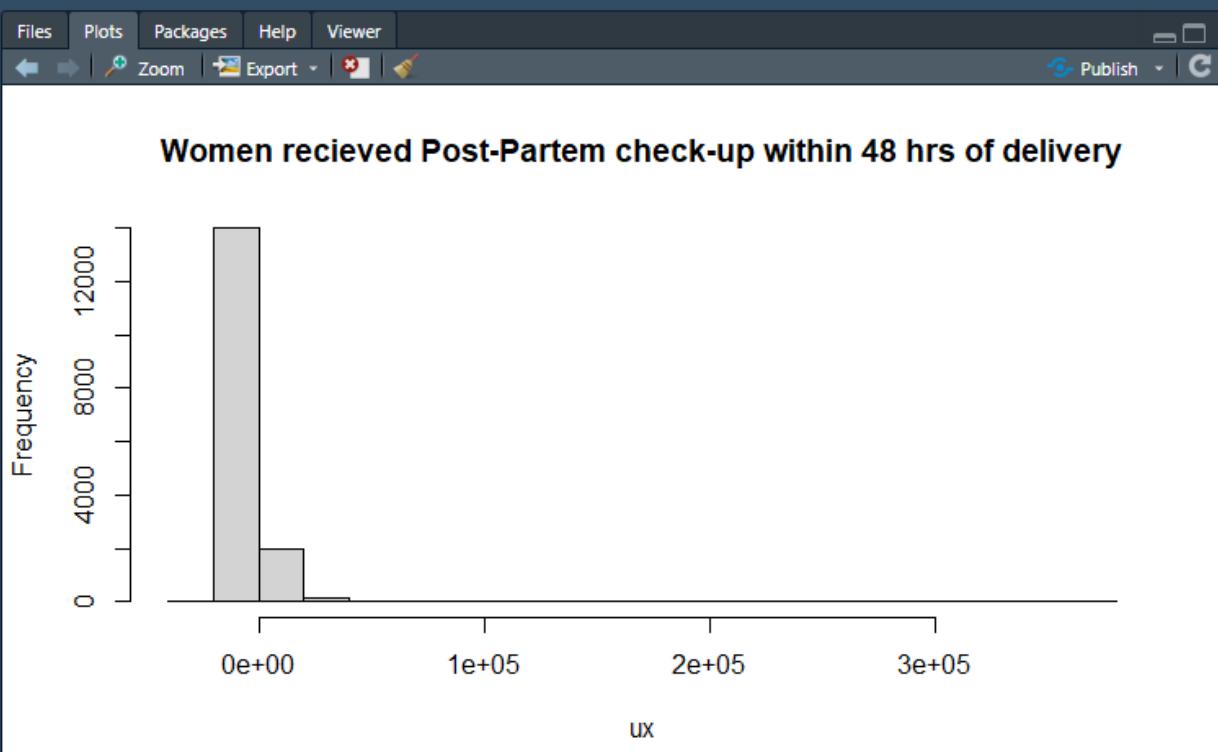
```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v6)
> print(sum(ux))
[1] 3.9729e-09
> hist(ux,main = "Tested moderately anaemic")
> |
```



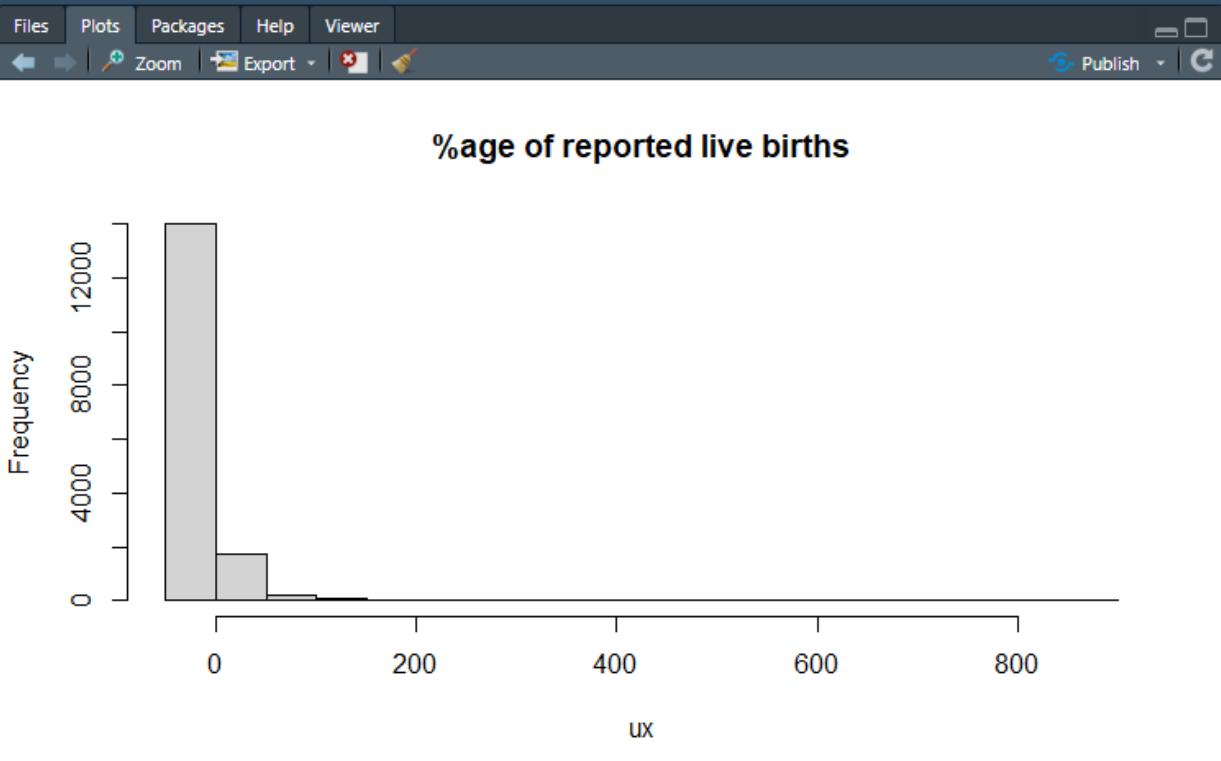
```
hist(ux,main = "Institutional Deliveries")
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v13)
> print(sum(ux))
[1] 6.117915e-09
> hist(ux,main = "Institutional Deliveries")
> |
```



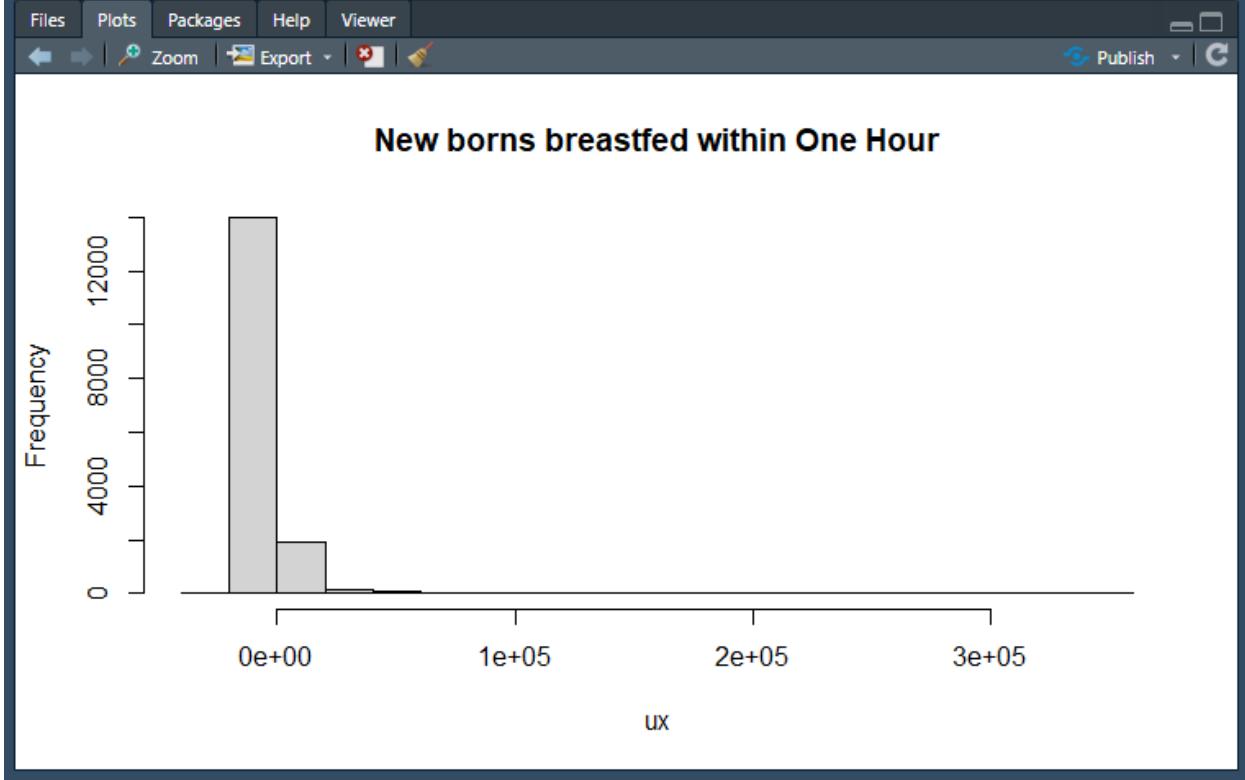
```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v18)
> print(sum(ux))
[1] -3.180645e-09
> hist(ux,main = "Women recieved Post-Partem check-up within 48 hrs of delivery")
> |
```



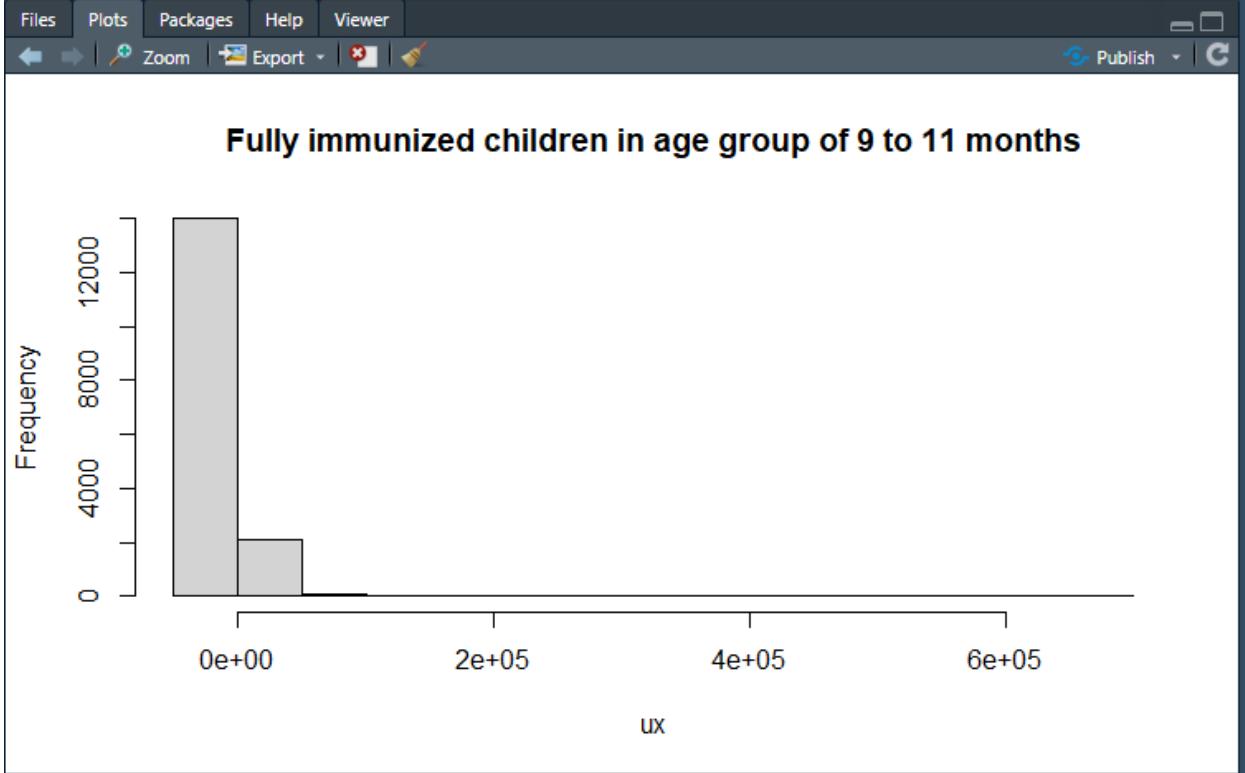
```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v25)
> print(sum(ux))
[1] 1.128916e-11
> hist(ux,main = "%age of reported live births")
> |
```



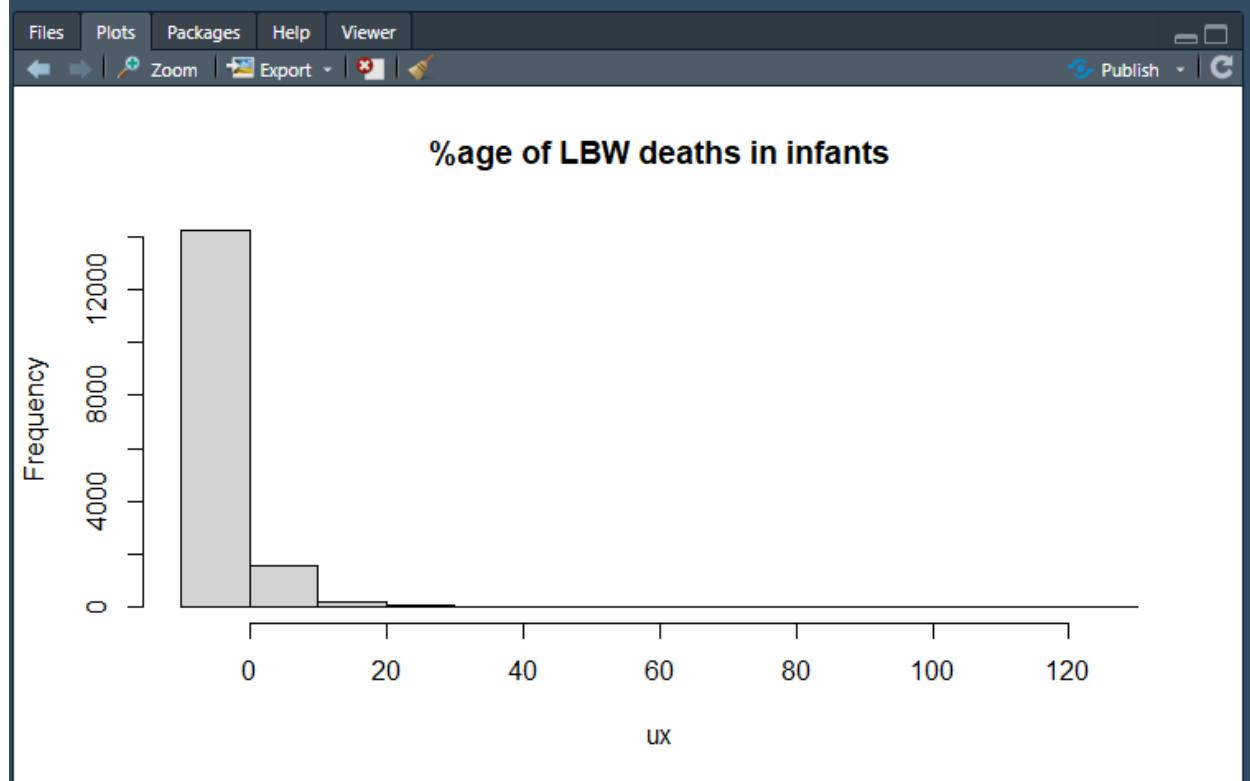
```
> hist(ux,main = "age of reported live births")
> ux <- c(res_ind_Data_kh$standard_res_kh$newData_kh$v29)
> print(sum(ux))
[1] 5.247784e-10
> hist(ux,main = "New borns breastfed within One Hour")
> |
```



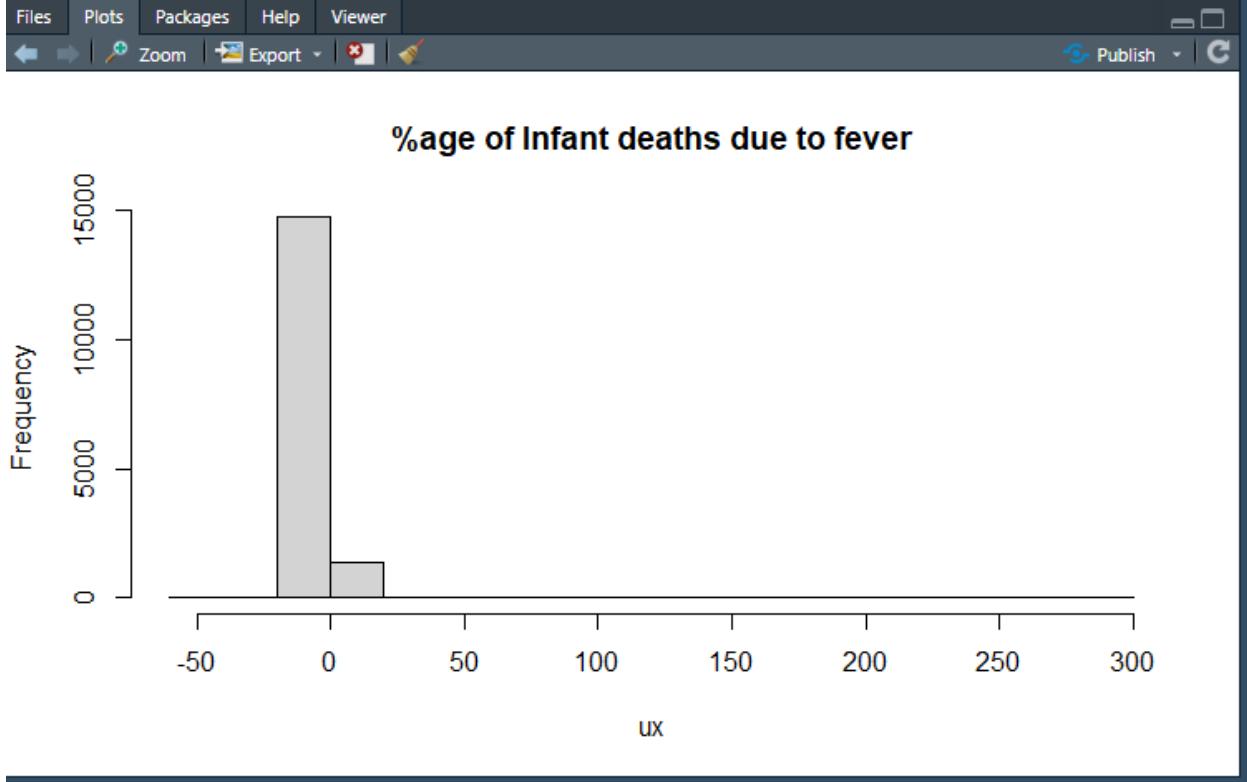
```
> hist(ux,main = "New borns breastfed within one hour")
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v34)
> print(sum(ux))
[1] 9.240523e-09
> hist(ux,main = "Fully immunized children in age group of 9 to 11 months")
> |
```



```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v42)
> print(sum(ux))
[1] -4.906402e-12
> hist(ux,main = "%age of LBW deaths in infants")
> |
```

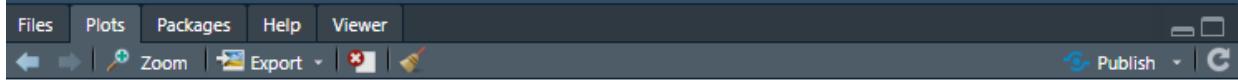


```
> ux <- c(res_ind_Data_kh$standard_res_kh*newData_kh$v45)
> print(sum(ux))
[1] -7.514604e-12
> hist(ux,main = "%age of Infant deaths due to fever")
> |
```

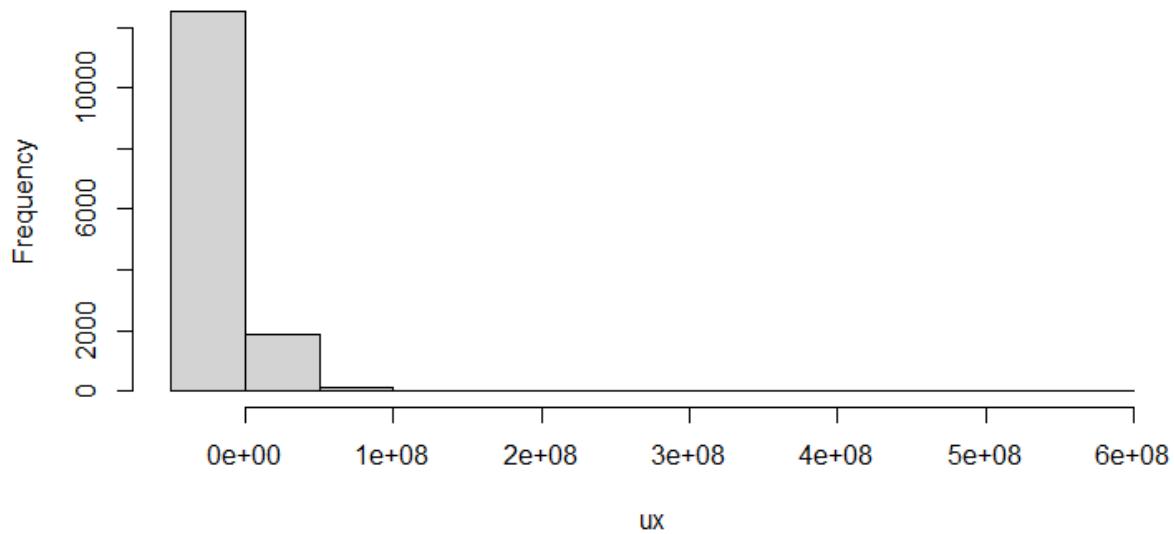


RABI

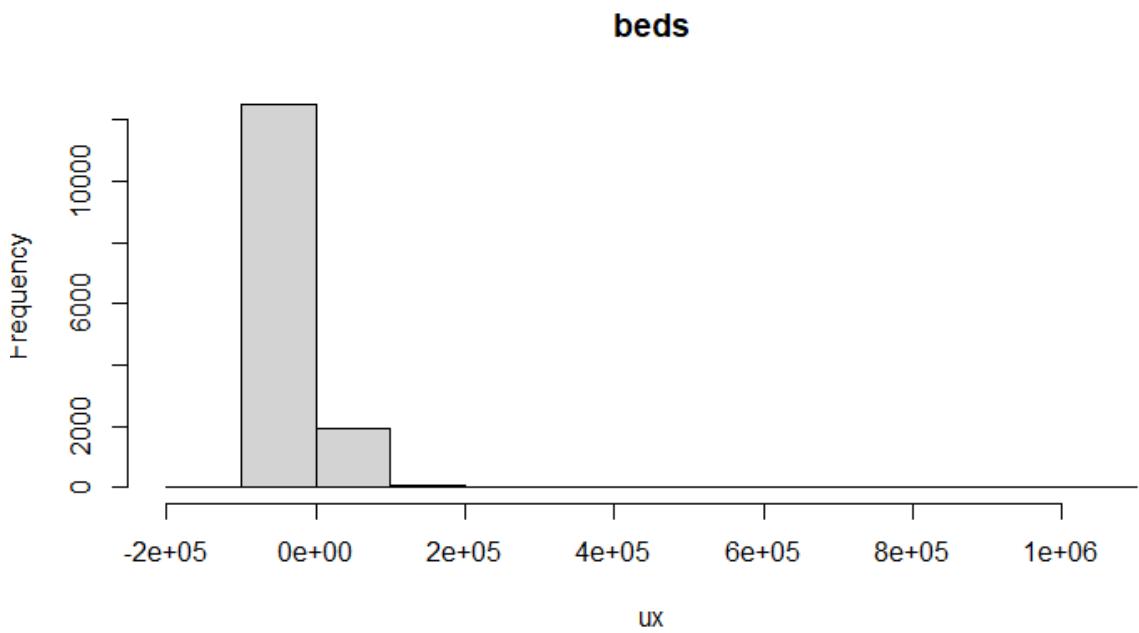
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$gdp)
> print(sum(ux))
[1] -1.607207e-05
> hist(ux, main = "gdp")
> |
```

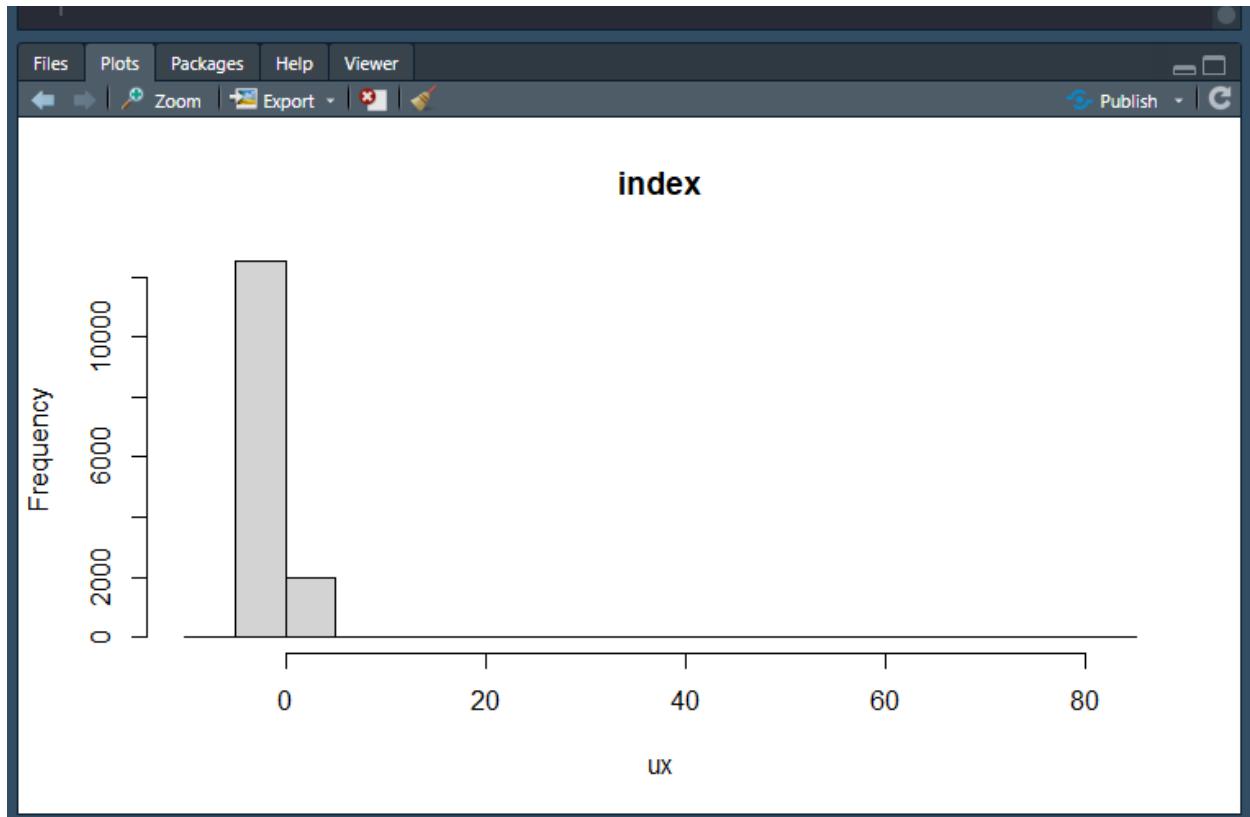


gdp

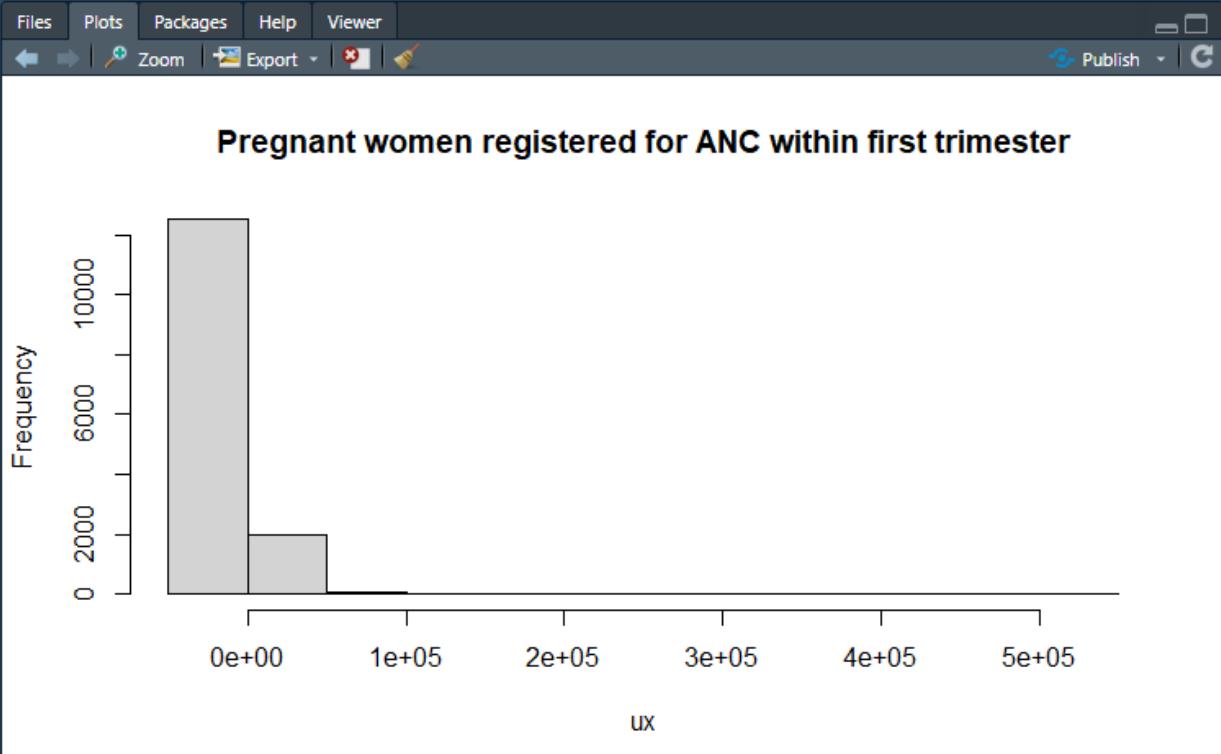


```
> hist(ux,main = "gap")
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$beds)
> print(sum(ux))
[1] -3.328898e-08
> hist(ux,main = "beds")
>
```

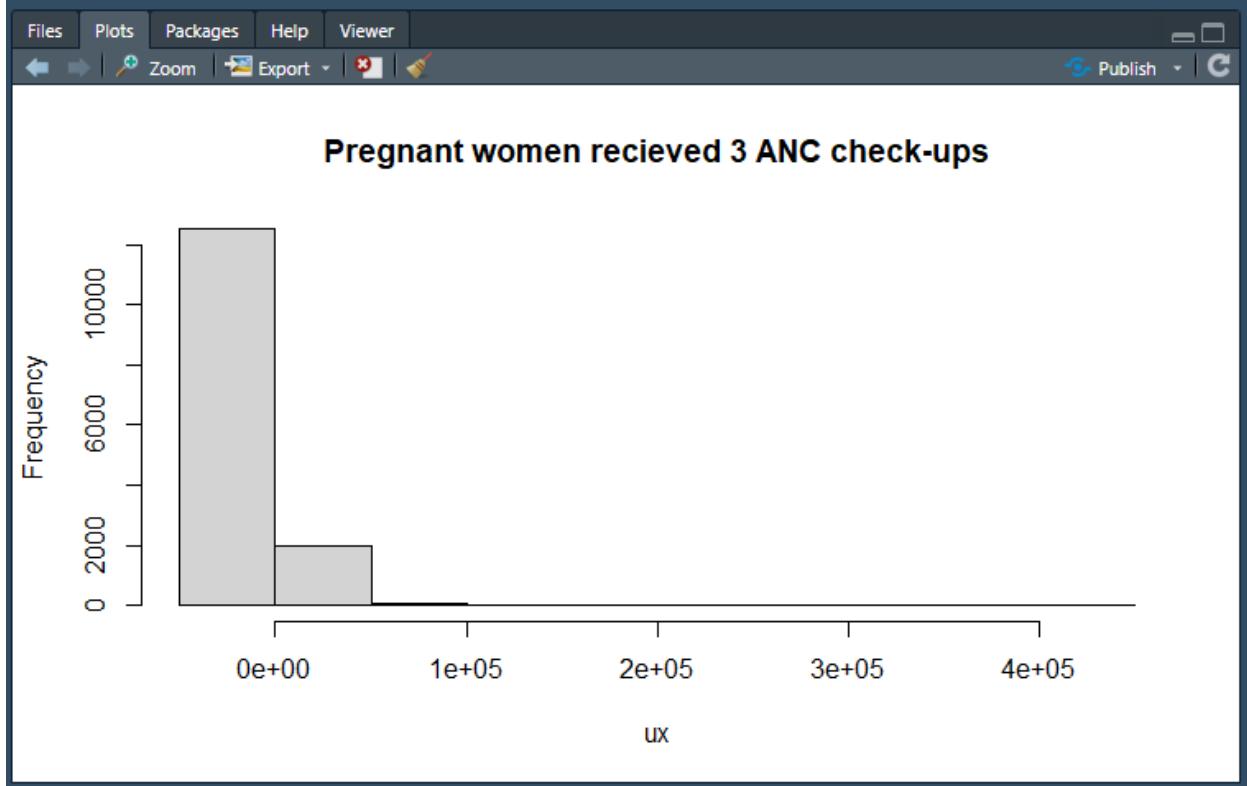




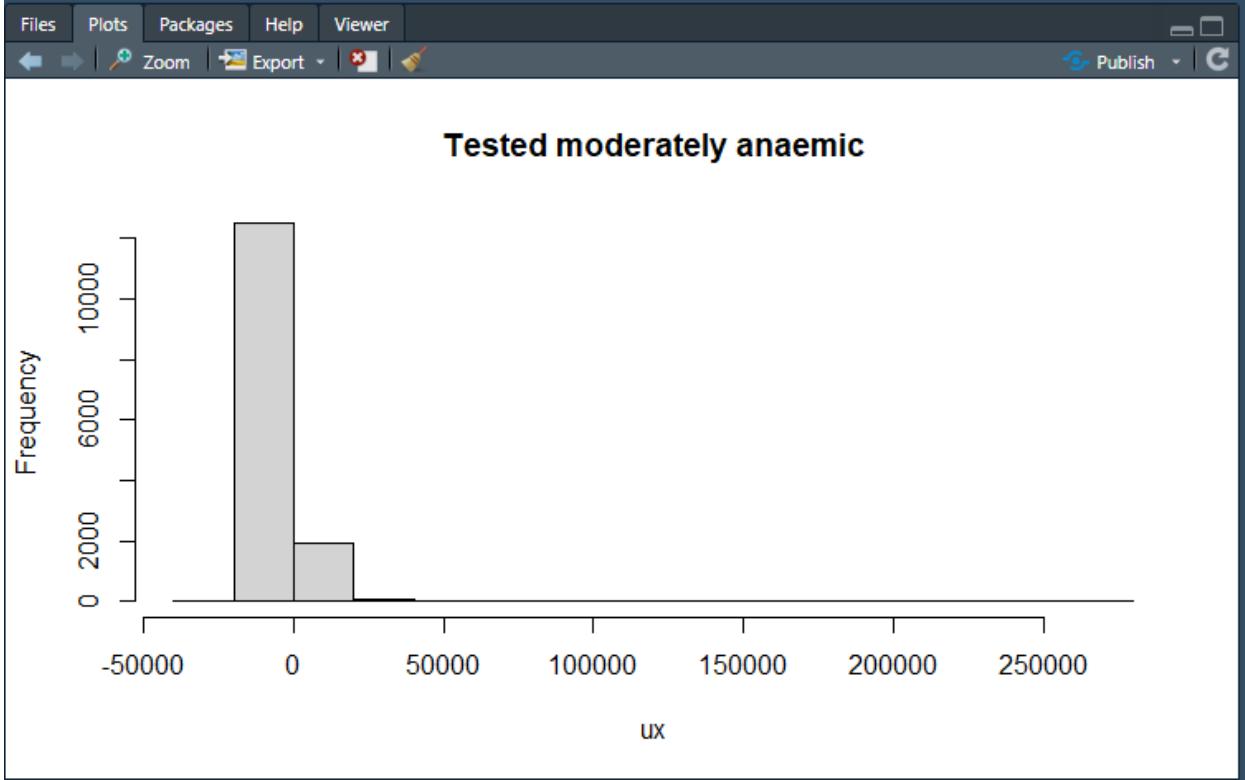
```
> hist(ux,main = "Index")
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v2)
> print(sum(ux))
[1] 1.231894e-08
> hist(ux,main = "Pregnant women registered for ANC within first trimester")
>
```



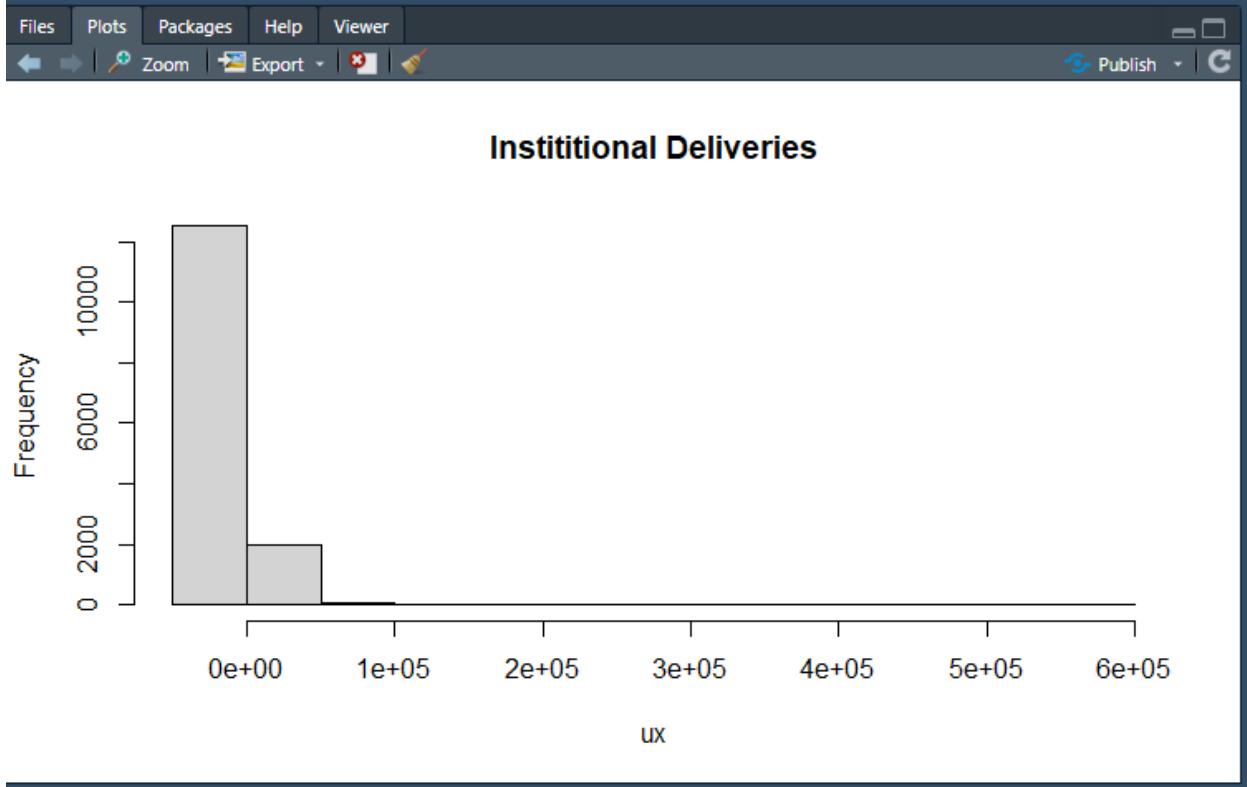
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v3)
> print(sum(ux))
[1] 9.937168e-09
> hist(ux,main = "Pregnant women received 3 ANC check-ups")
> |
```



```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v6)
> print(sum(ux))
[1] 4.314302e-09
> hist(ux,main = "Tested moderately anaemic")
> |
```



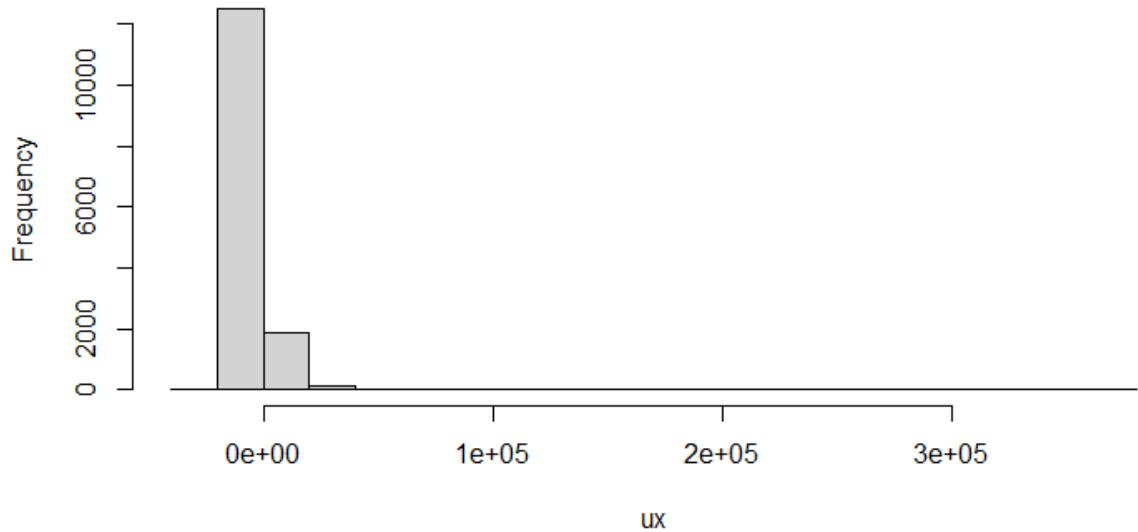
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v13)
> print(sum(ux))
[1] 1.3207e-08
> hist(ux,main = "Institutional Deliveries")
> |
```



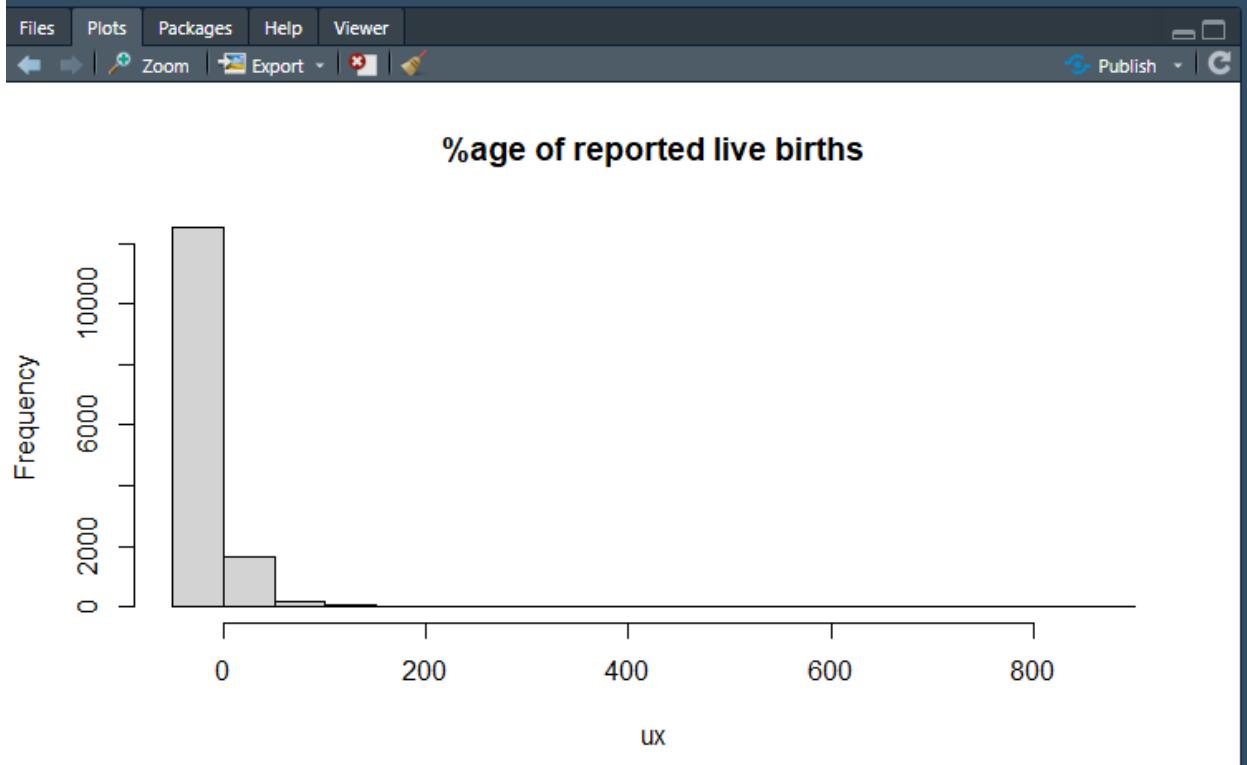
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v18)
> print(sum(ux))
[1] 1.234739e-08
> hist(ux,main = "Women received Post-Partem check-up within 48 hrs of delivery")
> |
```



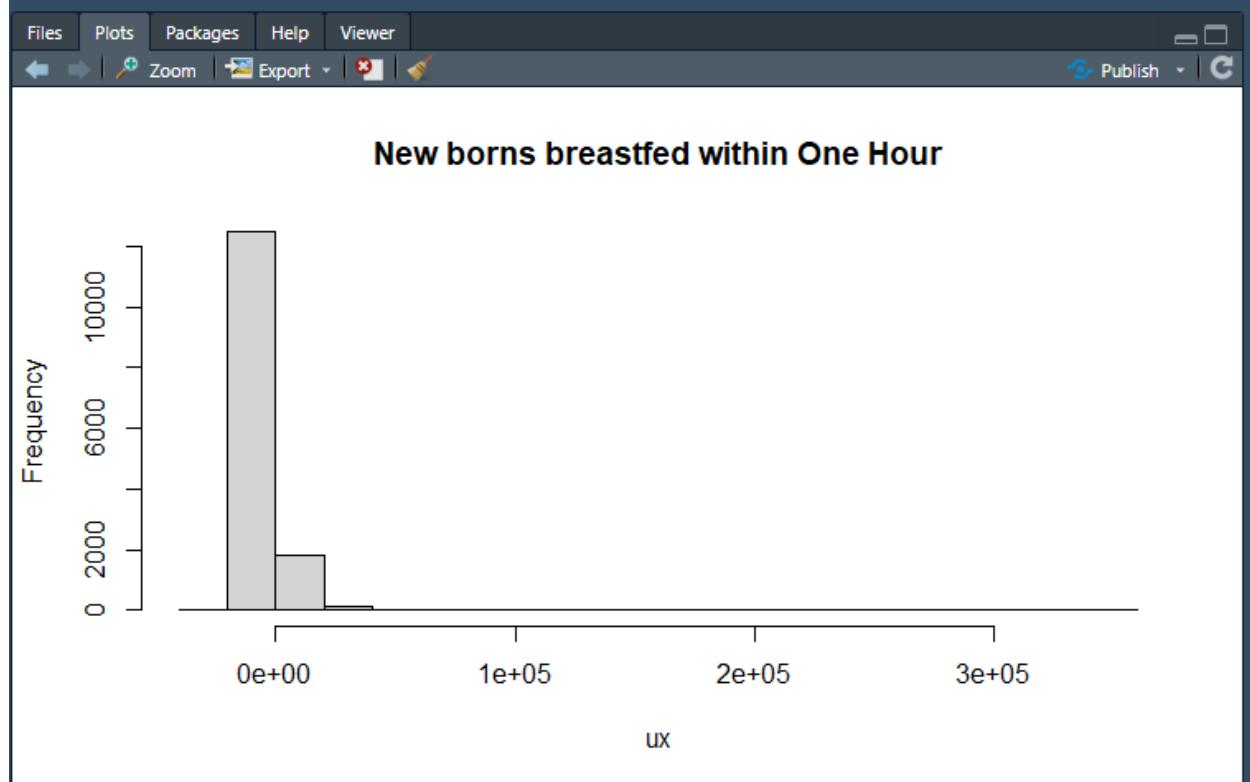
Women received Post-Partem check-up within 48 hrs of delivery



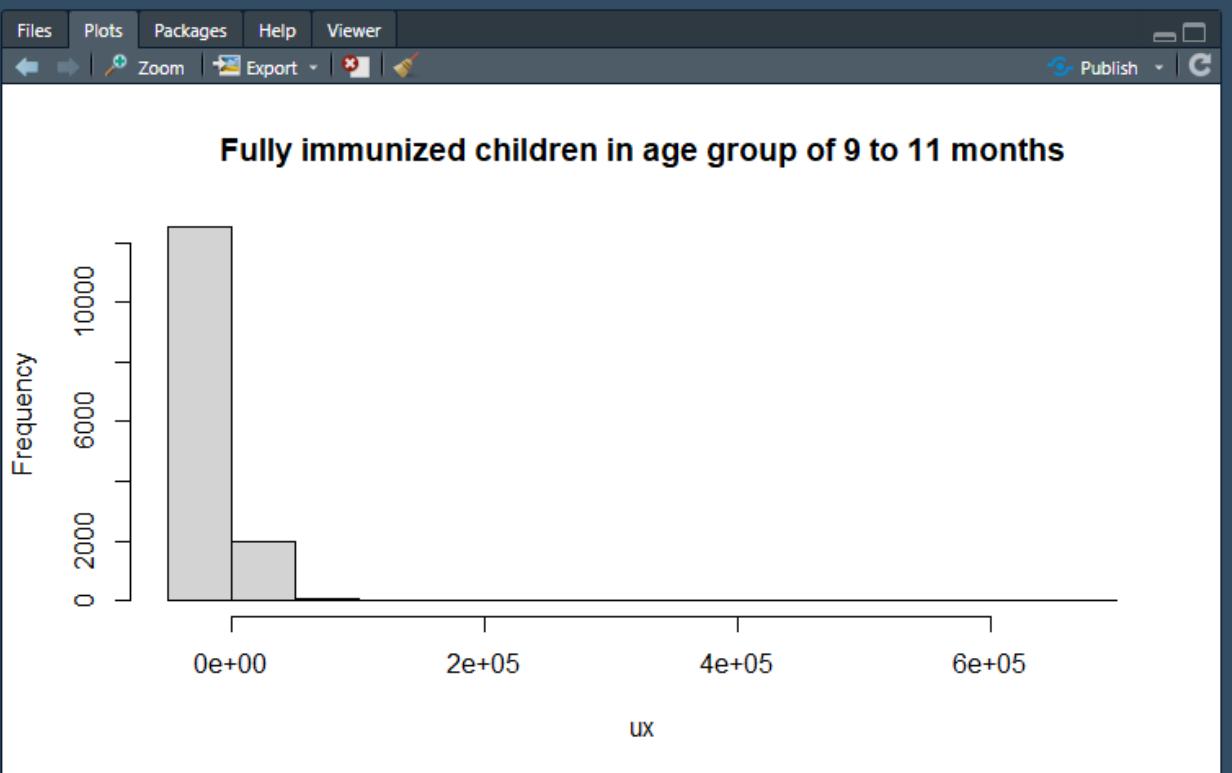
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v25)
> print(sum(ux))
[1] 1.409206e-11
> hist(ux,main = "%age of reported live births")
> |
```



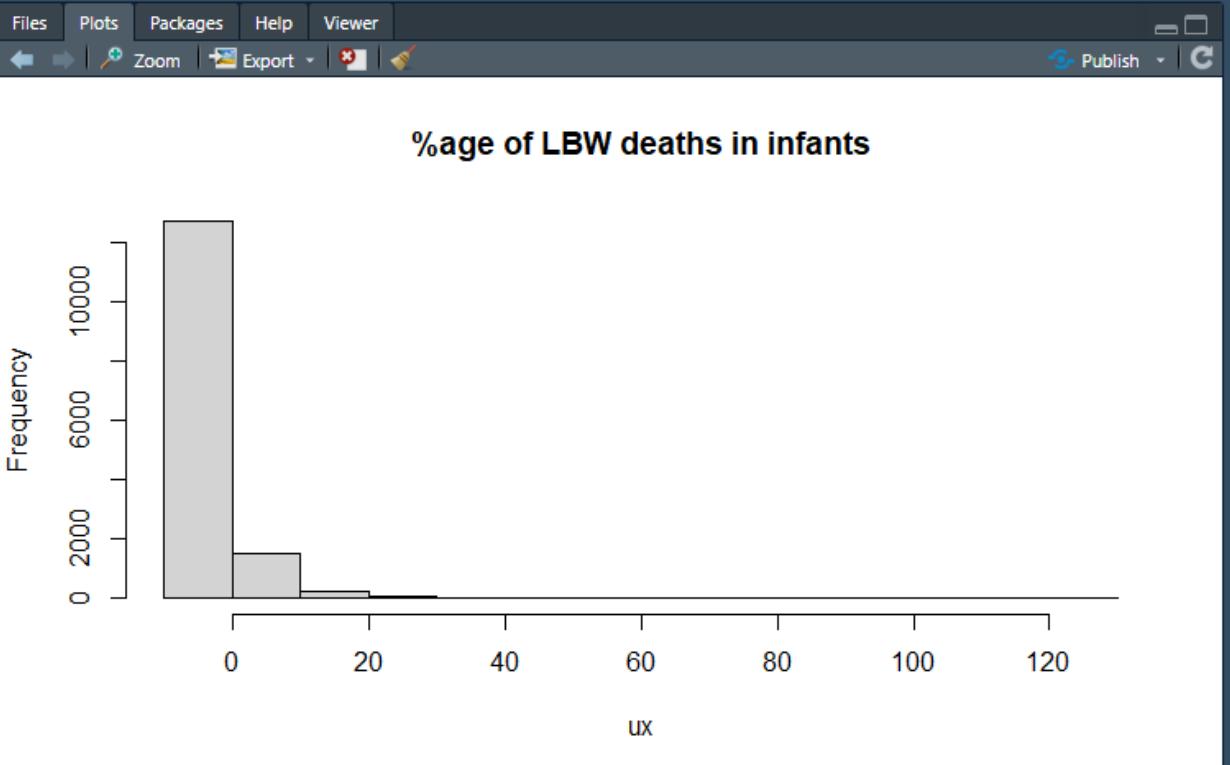
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v29)
> print(sum(ux))
[1] 1.119884e-08
> hist(ux,main = "New borns breastfed within One Hour")
> |
```



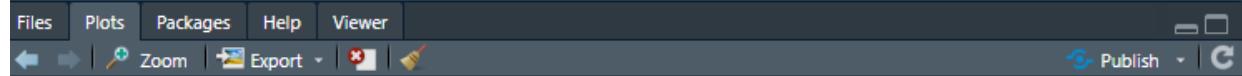
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v34)
> print(sum(ux))
[1] 7.498556e-09
> hist(ux,main = "Fully immunized children in age group of 9 to 11 months")
> |
```



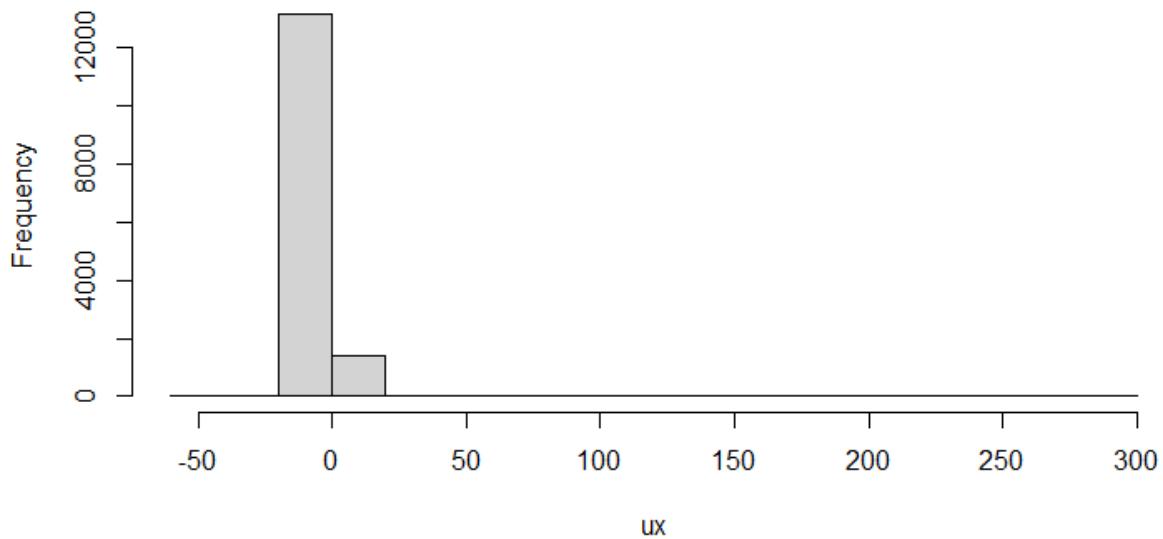
```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v42)
> print(sum(ux))
[1] 9.283435e-12
> hist(ux,main = "%age of LBW deaths in infants")
> |
```



```
> ux <- c(res_ind_Data_rb$standard_res_rb*newData_rb$v45)
> print(sum(ux))
[1] -5.17143e-12
> hist(ux,main = "%age of Infant deaths due to fever")
> |
```



%age of Infant deaths due to fever



Q2)

- a) After removing outliers for v46 for each season, Kharif and Rabi, we got the following intercepts and slope:
- i) Kharif
 - 1) B0: 0.0653645558042423
 - 2) B1: -0.000871357241736528
 - ii) Rabi
 - 1) B0: 0.0657506456804071
 - 2) B1: -0.00088251426953115
- b) **Monte Carlo Simulations:** In order to apply monte carlo simulations, we have assumed that the true values of B0 and B1 are those derived from the whole data as asked in the question.
In order to apply monte carlo simulations, we randomly take 80% data multiple times and then apply a regression model on this data. At the end, we take the average of B0 and B1 obtained from these random samples.

Thus we got,

- 1) Rabi
 - a) Avg B0: 0.0657659550755011 (Difference: -1.53e-5)
[True Value-Avg Value]
 - b) Avg B1: -0.000878243547802804 (Difference: -4.27e-6)
[True Value-Avg Value]
- 2) Kharif
 - a) Avg B0: 0.0654568128682429 (Difference: -9.22e-5)
[True Value-Avg Value]
 - b) Avg B1: -0.000878530687047374 (Difference: 7.17e-6)
[True Value-Avg Value]

Hence we see that the values obtained are very similar to that of true value. The difference is really small. From this, we can say that our OLS estimates are consistent.

Q3.

A. INTRODUCTION OF DUMMY VARIABLES

Model1)

$$v46 = gdp + beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42 + v45 + \text{South},$$

Code is provided in the r file.

B. (All screenshots attached below)

For Model 1) South

Rabi: We see that for the model having south as dummy variable, since $pr(t)>0.05$, our null hypothesis is correct.

Kharif:

We see that for the model having south as dummy variable, since $pr(t)>0.05$, our null hypothesis is correct.

For Model 2) North

Rabi: We see that for the model having north as dummy variable, since $pr(t)<0.05$, our null hypothesis is rejected.

Kharif:

We see that for the model having north as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

For Model 3) Central

Rabi: We see that for the model having central as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

Kharif:

We see that for the model having central as dummy variable, since $pr(t)>0.05$, our null hypothesis is correct.

For Model 4) East

Rabi: We see that for the model having east as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

Kharif:

We see that for the model having east as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

For Model 5) West

Rabi: We see that for the model having west as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

Kharif:

We see that for the model having west as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

For Model 6) North East

Rabi: We see that for the model having north east as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

Kharif:

We see that for the model having north east as dummy variable, since $pr(t)<0.05$, our null hypothesis is incorrect.

We see that we have checked 6 null hypothesis. As we see, that all null hypothesis should be simultaneously true to have no structural break. We see that that's not the case here. (Refer to sir's notes)

C. Our objective is to check structural breaks between southern and non-southern states.

UTILISING F TEST

We calculate SSE(U) and SSE(R). We want to see the change when the dummy variable south is added to the regression model. The basic concept is to include it(dummy var) in the unrestricted model and exclude it in the restricted model in order to check if it causes any effect. This will be determined using the F value.

$$F^* = \left(\frac{SSE(R) - SSE(F)}{df_R - df_F} \right) \div \left(\frac{SSE(F)}{df_F} \right)$$

For this formula

R-restricted

F- unrestricted

$$(1620.501-1620.363)/(13119-13118) / (1620.363/13118))$$

$$(1353.833-1353.531)/(11899-11898)/(1353.531/11898))-2.654$$

KHARIF

State	F* value
South	1.11

P-value restricted- 2.04e-11

P-value unrestricted- 3.226e - 11

```
> k_linear_model <- lm(formula = v46 ~ gdp + beds+ index + v2 + v3 + v6 + v13 +
+ v18 + v25 + v29 + v34 + v42 + v45 , data = kharif )
> summary(k_linear_model)

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
v18 + v25 + v29 + v34 + v42 + v45, data = kharif)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.2143 -0.0772 -0.0610 -0.0380  7.6132 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.794e+00  4.196e-01   4.276  1.92e-05 ***
gdp         1.238e-10  1.526e-10   0.811  0.41719    
beds        -1.354e-07  6.330e-08  -2.138  0.03252 *  
index       -3.313e-04  3.809e-04  -0.870  0.38444    
v2          1.372e-06  5.303e-07   2.587  0.00968 ** 
v3          -7.364e-07  4.601e-07  -1.601  0.10951    
v6          -2.220e-07  2.298e-07  -0.966  0.33414    
v13         7.628e-07  5.103e-07   1.495  0.13500    
v18         -6.845e-07  4.760e-07  -1.438  0.15039    
v25         -1.727e-02  4.271e-03  -4.044  5.29e-05 *** 
v29         -3.128e-07  5.524e-07  -0.566  0.57124    
v34         -5.830e-07  4.023e-07  -1.449  0.14729    
v42         -5.258e-04  2.359e-04  -2.229  0.02586 *  
v45         1.894e-04  3.502e-04   0.541  0.58867    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3515 on 13119 degrees of freedom
Multiple R-squared:  0.005969, Adjusted R-squared:  0.004984 
F-statistic:  6.06 on 13 and 13119 DF,  p-value: 2.04e-11
```

```
> k_dummy_linear_model <- lm(formula = v46 ~ gdp + beds+ index + v2 + v3 + v6 +
+ v13 + v18 + v25 + v29 + v34 + v42 + v45 + South , data = kharif )
> summary(k_dummy_linear_model)

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
v18 + v25 + v29 + v34 + v42 + v45 + South, data = kharif)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.2258 -0.0769 -0.0609 -0.0388  7.6147 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.908e+00  4.331e-01   4.404  1.07e-05 ***
gdp         1.693e-10  1.586e-10   1.068  0.2856    
beds        -1.645e-07  6.906e-08  -2.382  0.0172 *  
index       -2.828e-04  3.837e-04  -0.737  0.4611    
v2          1.326e-06  5.320e-07   2.493  0.0127 *  
v3          -7.913e-07  4.630e-07  -1.709  0.0875 .  
v6          -2.580e-07  2.323e-07  -1.110  0.2669    
v13         7.395e-07  5.108e-07   1.448  0.1477    
v18         -7.105e-07  4.766e-07  -1.491  0.1360    
v25         -1.844e-02  4.413e-03  -4.180  2.94e-05 *** 
v29         -1.920e-07  5.641e-07  -0.340  0.7335    
v34         -5.254e-07  4.060e-07  -1.294  0.1956    
v42         -5.414e-04  2.364e-04  -2.290  0.0220 *  
v45         2.301e-04  3.523e-04   0.653  0.5137    
South       1.018e-02  9.633e-03   1.057  0.2907    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3515 on 13118 degrees of freedom
Multiple R-squared:  0.006054, Adjusted R-squared:  0.004993 
F-statistic: 5.707 on 14 and 13118 DF,  p-value: 3.226e-11
```

RABI

State	F* value
South	2.654

P value restricted- 3.328e -08

P value unrestricted- 2.543e - 08

```

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
    v18 + v25 + v29 + v34 + v42 + v45 + South, data = rabi)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.1982 -0.0729 -0.0589 -0.0396  7.6199 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.666e+00  4.317e-01   3.858 0.000115 ***
gdp         1.096e-10  1.651e-10   0.664 0.506901  
beds        -9.867e-08 6.832e-08  -1.444 0.148719  
index       -6.338e-04 3.574e-04  -1.774 0.076156 .  
v2          1.206e-06 5.142e-07   2.346 0.018998 *  
v3          -8.769e-07 4.832e-07  -1.815 0.069610 .  
v6          -3.363e-07 2.305e-07  -1.459 0.144621  
v13         8.980e-07 5.100e-07   1.761 0.078288 .  
v18         -3.521e-07 4.724e-07  -0.745 0.456095  
v25         -1.604e-02 4.398e-03  -3.648 0.000266 *** 
v29         -3.516e-07 5.595e-07  -0.628 0.529731  
v34         -5.708e-07 4.025e-07  -1.418 0.156135  
v42         -4.520e-04 2.357e-04  -1.918 0.055159 .  
v45         1.302e-04 3.491e-04   0.373 0.709216  
South       1.565e-02 9.605e-03   1.630 0.103216  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3373 on 11898 degrees of freedom
Multiple R-squared:  0.005342, Adjusted R-squared:  0.004172 
F-statistic: 4.564 on 14 and 11898 DF, p-value: 2.543e-08

> sse_rd <- sum((fitted(rabi_dummy_linear_model) - rabi$v46)^2)
> print(sse_rd)
[1] 1353.531

```

```

Call:
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 +
    v18 + v25 + v29 + v34 + v42 + v45, data = rabi)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.1800 -0.0730 -0.0589 -0.0393  7.6175 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.495e+00 4.188e-01 3.569 0.000359 ***  
gdp         4.907e-11 1.609e-10 0.305 0.760365    
beds        -5.891e-08 6.382e-08 -0.923 0.356012    
index       -6.680e-04 3.568e-04 -1.872 0.061171 .    
v2          1.248e-06 5.136e-07 2.430 0.015120 *    
v3          -7.713e-07 4.789e-07 -1.611 0.107315    
v6          -2.793e-07 2.279e-07 -1.226 0.220245    
v13         9.239e-07 5.097e-07 1.812 0.069940 .    
v18         -3.140e-07 4.718e-07 -0.666 0.505723    
v25         -1.428e-02 4.263e-03 -3.349 0.000813 ***  
v29         -5.231e-07 5.495e-07 -0.952 0.341168    
v34         -6.613e-07 3.987e-07 -1.659 0.097168 .    
v42         -4.217e-04 2.350e-04 -1.795 0.072753 .    
v45         6.980e-05 3.472e-04 0.201 0.840659    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3373 on 11899 degrees of freedom
Multiple R-squared:  0.00512,   Adjusted R-squared:  0.004033 
F-statistic:  4.71 on 13 and 11899 DF,  p-value: 3.328e-08

> sse_ru <- sum((fitted(rabi_linear_model) - rabi$v46)^2)
> print(sse_ru)
[1] 1353.833

```

Hypotheses

HO- BETAsouth = 0 (i.& no structural breaks across the southern states)

H1- BETA!=0 (i.e.there is a presence of structural break)

KHARIF

As it can be observed, the F-stat value for both the regressions are greater than the F-value(1.11) we obtained using the sum of squared estimate of errors(sse). Since f stat is greater we can conclude that there is no structural break.

RABI

As it can be observed, the F-stat value for both the regressions are greater than the F-value(2.654) we obtained using the sum of squared estimate of errors(sse). Since f stat is greater we can conclude that there is no structural break.

3b (Screenshots)

<u>Coefficients:</u>				
	<u>Estimate</u>	<u>Std. Error</u>	<u>t value</u>	
(Intercept)	1.335e+00	4.227e-01	3.158	
gdp	5.684e-11	1.608e-10	0.353	
beds	-4.708e-08	6.395e-08	-0.736	
index	-5.844e-04	3.579e-04	-1.633	
v2	1.327e-06	5.142e-07	2.580	
v3	-8.973e-07	4.809e-07	-1.866	
v6	-2.390e-07	2.283e-07	-1.047	
v13	1.022e-06	5.108e-07	2.002	
v18	-4.136e-07	4.731e-07	-0.874	
v25	-1.270e-02	4.300e-03	-2.952	
v29	-5.219e-07	5.494e-07	-0.950	
v34	-6.182e-07	3.988e-07	-1.550	
v42	-5.259e-04	2.379e-04	-2.210	
v45	1.072e-04	3.473e-04	0.309	
Central	2.796e-02	1.013e-02	2.760	
<u>Pr(> t)</u>				
(Intercept)	0.00159	**		
gdp	0.72383			
beds	0.46165			
index	0.10255			
v2	0.00988	**		
v3	0.06209	.		

```

v6          0.29503
v13         0.04536 *
v18         0.38196
v25         0.00316 **
v29         0.34212
v34         0.12118
v42         0.02710 *
v45         0.75769
Central     0.00579 **

---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

```

Residual standard error: 0.3372 on 11898 degrees of freedom

Multiple R-squared: 0.005756, Adjusted R-squared: 0.004586

F-statistic: 4.92 on 14 and 11898 DF, p-value: 3.286e-09

```

> sse_rd <- sum((fitted(rabi_dummy_linear_model) -
rabi$v46)^2)
> print(sse_rd)
[1] 1352.967
>
> rabi_linear_model <- lm(formula = v46 ~ gdp +beds+
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34
+ v42 + v45, data = rabi )
> summary(rabi_linear_model)

```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1800	-0.0730	-0.0589	-0.0393	7.6175

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.495e+00	4.188e-01	3.569
gdp	4.907e-11	1.609e-10	0.305
beds	-5.891e-08	6.382e-08	-0.923
index	-6.680e-04	3.568e-04	-1.872
v2	1.248e-06	5.136e-07	2.430
v3	-7.713e-07	4.789e-07	-1.611
v6	-2.793e-07	2.279e-07	-1.226
v13	9.239e-07	5.097e-07	1.812
v18	-3.140e-07	4.718e-07	-0.666
v25	-1.428e-02	4.263e-03	-3.349
v29	-5.231e-07	5.495e-07	-0.952
v34	-6.613e-07	3.987e-07	-1.659
v42	-4.217e-04	2.350e-04	-1.795
v45	6.980e-05	3.472e-04	0.201
	Pr(> t)		
(Intercept)	0.000359	***	
gdp	0.760365		
beds	0.356012		
index	0.061171	.	
v2	0.015120	*	

```

v3          0.107315
v6          0.220245
v13         0.069940 .
v18         0.505723
v25         0.000813 ***
v29         0.341168
v34         0.097168 .
v42         0.072753 .
v45         0.840659
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

```

Residual standard error: 0.3373 on 11899 degrees of freedom

Multiple R-squared: 0.00512, Adjusted R-squared: 0.004033

F-statistic: 4.71 on 13 and 11899 DF, p-value: 3.328e-08

```

> sse_ru <- sum((fitted(rabi_linear_model) -
rabi$v46)^2)
> print(sse_ru)
[1] 1353.833
> k_dummy_linear_model <- lm(formula = v46 ~ gdp +beds+
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34
+ v42 + v45 + North , data = kharif )
>
> summary(k_dummy_linear_model)

```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45 + North, data =  
kharif)
```

Residuals:

Min	10	Median	30	Max
-0.2240	-0.0780	-0.0600	-0.0366	7.6173

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.815e+00	4.196e-01	4.325
gdp	1.802e-10	1.545e-10	1.166
beds	-1.802e-07	6.619e-08	-2.722
index	-4.686e-04	3.855e-04	-1.216
v2	1.455e-06	5.314e-07	2.738
v3	-6.326e-07	4.622e-07	-1.369
v6	-2.389e-07	2.299e-07	-1.039
v13	8.684e-07	5.123e-07	1.695
v18	-6.351e-07	4.764e-07	-1.333
v25	-1.752e-02	4.272e-03	-4.102
v29	-4.192e-07	5.542e-07	-0.756
v34	-7.581e-07	4.093e-07	-1.852
v42	-4.369e-04	2.390e-04	-1.828
v45	1.619e-04	3.503e-04	0.462
North	2.303e-02	9.960e-03	2.312
	Pr(> t)		
(Intercept)	1.54e-05	***	
gdp	0.24350		
beds	0.00650	**	

```
index          0.22410
v2            0.00619 **
v3            0.17109
v6            0.29876
v13           0.09009 .
v18           0.18247
v25           4.13e-05 ***
v29           0.44945
v34           0.06403 .
v42           0.06754 .
v45           0.64388
North         0.02077 *
```

Signif. codes:

```
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3514 on 13118 degrees of freedom

Multiple R-squared: 0.006374, Adjusted R-squared:

0.005314

F-statistic: 6.011 on 14 and 13118 DF, p-value:

5.258e-12

>

```
> sse_kd <- sum((fitted(k_dummy_linear_model) -  
kharif$v46)^2)
```

```
> print(sse_kd)
```

```
[1] 1619.841
```

>

```
> k_linear_model <- lm(formula = v46 ~ gdp +beds+ index  
+ v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42
```

```
+ v45 , data = kharif )  
> summary(k_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = kharif)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.2143	-0.0772	-0.0610	-0.0380	7.6132

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.794e+00	4.196e-01	4.276
gdp	1.238e-10	1.526e-10	0.811
beds	-1.354e-07	6.330e-08	-2.138
index	-3.313e-04	3.809e-04	-0.870
v2	1.372e-06	5.303e-07	2.587
v3	-7.364e-07	4.601e-07	-1.601
v6	-2.220e-07	2.298e-07	-0.966
v13	7.628e-07	5.103e-07	1.495
v18	-6.845e-07	4.760e-07	-1.438
v25	-1.727e-02	4.271e-03	-4.044
v29	-3.128e-07	5.524e-07	-0.566
v34	-5.830e-07	4.023e-07	-1.449
v42	-5.258e-04	2.359e-04	-2.229
v45	1.894e-04	3.502e-04	0.541
	Pr(> t)		
(Intercept)	1.92e-05	***	
gdp	0.41719		

<u>beds</u>	<u>0.03252 *</u>
<u>index</u>	<u>0.38444</u>
<u>v2</u>	<u>0.00968 **</u>
<u>v3</u>	<u>0.10951</u>
<u>v6</u>	<u>0.33414</u>
<u>v13</u>	<u>0.13500</u>
<u>v18</u>	<u>0.15039</u>
<u>v25</u>	<u>5.29e-05 ***</u>
<u>v29</u>	<u>0.57124</u>
<u>v34</u>	<u>0.14729</u>
<u>v42</u>	<u>0.02586 *</u>
<u>v45</u>	<u>0.58867</u>

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3515 on 13119 degrees of freedom

Multiple R-squared: 0.005969, Adjusted R-squared:

0.004984

F-statistic: 6.06 on 13 and 13119 DF, p-value:

2.04e-11

```
> sse_ku <- sum((fitted(k_linear_model) -  
kharif$v46)^2)  
> print(sse_ku)  
[1] 1620.501  
>  
> rabi_dummy_linear_model <- lm(formula = v46 ~ gdp  
+beds+ index + v2 + v3 + v6 + v13 + v18 + v25 + v29  
+ v34 + v42 + v45 + North , data = rabi )
```

```
> summary(rabi_dummy_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45 + North, data =  
rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1909	-0.0731	-0.0577	-0.0368	7.6224

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.532e+00	4.189e-01	3.657
gdp	1.358e-10	1.633e-10	0.831
beds	-1.293e-07	6.784e-08	-1.906
index	-7.238e-04	3.571e-04	-2.027
v2	1.323e-06	5.140e-07	2.574
v3	-6.186e-07	4.813e-07	-1.285
v6	-2.969e-07	2.278e-07	-1.303
v13	1.048e-06	5.112e-07	2.050
v18	-2.494e-07	4.721e-07	-0.528
v25	-1.470e-02	4.264e-03	-3.449
v29	-6.513e-07	5.510e-07	-1.182
v34	-8.736e-07	4.045e-07	-2.159
v42	-2.987e-04	2.383e-04	-1.254
v45	2.550e-05	3.473e-04	0.073
North	3.028e-02	9.925e-03	3.051
	Pr(> t)		
(Intercept)	0.000256	***	

<u>gdp</u>	<u>0.405769</u>
<u>beds</u>	<u>0.056717</u> .
<u>index</u>	<u>0.042701</u> *
<u>v2</u>	<u>0.010062</u> *
<u>v3</u>	<u>0.198759</u>
<u>v6</u>	<u>0.192634</u>
<u>v13</u>	<u>0.040343</u> *
<u>v18</u>	<u>0.597373</u>
<u>v25</u>	<u>0.000565</u> ***
<u>v29</u>	<u>0.237174</u>
<u>v34</u>	<u>0.030833</u> *
<u>v42</u>	<u>0.210040</u>
<u>v45</u>	<u>0.941467</u>
<u>North</u>	<u>0.002282</u> **

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3372 on 11898 degrees of freedom

Multiple R-squared: 0.005898, Adjusted R-squared: 0.004728

F-statistic: 5.042 on 14 and 11898 DF, p-value: 1.622e-09

```
> sse_rd <- sum((fitted(rabi_dummy_linear_model) -
rabi$v46)^2)
> print(sse_rd)
[1] 1352.774
>
> rabi_linear_model <- lm(formula = v46 ~ gdp + beds+
```

```
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34  
+ v42 + v45, data = rabi )  
> summary(rabi_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1800	-0.0730	-0.0589	-0.0393	7.6175

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.495e+00	4.188e-01	3.569
gdp	4.907e-11	1.609e-10	0.305
beds	-5.891e-08	6.382e-08	-0.923
index	-6.680e-04	3.568e-04	-1.872
v2	1.248e-06	5.136e-07	2.430
v3	-7.713e-07	4.789e-07	-1.611
v6	-2.793e-07	2.279e-07	-1.226
v13	9.239e-07	5.097e-07	1.812
v18	-3.140e-07	4.718e-07	-0.666
v25	-1.428e-02	4.263e-03	-3.349
v29	-5.231e-07	5.495e-07	-0.952
v34	-6.613e-07	3.987e-07	-1.659
v42	-4.217e-04	2.350e-04	-1.795
v45	6.980e-05	3.472e-04	0.201
	Pr(> t)		
(Intercept)	0.000359	***	

```
gdp      0.760365
beds    0.356012
index   0.061171 .
v2      0.015120 *
v3      0.107315
v6      0.220245
v13     0.069940 .
v18     0.505723
v25     0.000813 ***
v29     0.341168
v34     0.097168 .
v42     0.072753 .
v45     0.840659
```

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3373 on 11899 degrees of freedom

Multiple R-squared: 0.00512, Adjusted R-squared: 0.004033

F-statistic: 4.71 on 13 and 11899 DF, p-value: 3.328e-08

```
> sse_ru <- sum((fitted(rabi_linear_model) -
rabi$v46)^2)
> print(sse_ru)
[1] 1353.833
>
>
>
```

```

>
>
> k_dummy_linear_model <- lm(formula = v46 ~ gdp + beds+
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34
+ v42 + v45 + West , data = kharif )
>
> summary(k_dummy_linear_model)

```

Call:

```

lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
v18 + v25 + v29 + v34 + v42 + v45 + West, data =
kharif)

```

Residuals:

Min	1Q	Median	3Q	Max
-0.2351	-0.0782	-0.0585	-0.0371	7.6173

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	2.018e+00	4.256e-01	4.740
gdp	5.095e-10	1.969e-10	2.587
beds	-2.345e-07	7.091e-08	-3.307
index	-3.792e-04	3.811e-04	-0.995
v2	1.364e-06	5.301e-07	2.574
v3	-7.716e-07	4.601e-07	-1.677
v6	-3.648e-07	2.343e-07	-1.557
v13	1.062e-06	5.192e-07	2.045
v18	-6.579e-07	4.759e-07	-1.382
v25	-1.959e-02	4.335e-03	-4.519
v29	-3.198e-07	5.522e-07	-0.579

v34	-6.814e-07	4.034e-07	-1.689
v42	-5.148e-04	2.359e-04	-2.183
v45	2.013e-04	3.501e-04	0.575
West	-3.459e-02	1.117e-02	-3.098
	Pr(> t)		
(Intercept)	2.16e-06	***	
gdp	0.009679	**	
beds	0.000947	***	
index	0.319754		
v2	0.010067	*	
v3	0.093532	.	
v6	0.119533		
v13	0.040882	*	
v18	0.166849		
v25	6.27e-06	***	
v29	0.562572		
v34	0.091229	.	
v42	0.029090	*	
v45	0.565350		
West	0.001955	**	

<u>Signif. codes:</u>			
0	'***'	0.001	'**'
0.01	'*'	0.05	'. '
0.1	' '	1	

Residual standard error: 0.3513 on 13118 degrees of freedom

Multiple R-squared: 0.006696, Adjusted R-squared: 0.005636

F-statistic: 6.316 on 14 and 13118 DF, p-value: 8.393e-13

```

>
> sse_kd <- sum((fitted(k_dummy_linear_model) -
kharif$v46)^2)
> print(sse_kd)
[1] 1619.316
>
> k_linear_model <- lm(formula = v46 ~ gdp + beds+ index
+ v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42
+ v45 , data = kharif )
> summary(k_linear_model)

```

Call:

```

lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
v18 + v25 + v29 + v34 + v42 + v45, data = kharif)

```

Residuals:

Min	1Q	Median	3Q	Max
-0.2143	-0.0772	-0.0610	-0.0380	7.6132

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.794e+00	4.196e-01	4.276
gdp	1.238e-10	1.526e-10	0.811
beds	-1.354e-07	6.330e-08	-2.138
index	-3.313e-04	3.809e-04	-0.870
v2	1.372e-06	5.303e-07	2.587
v3	-7.364e-07	4.601e-07	-1.601
v6	-2.220e-07	2.298e-07	-0.966
v13	7.628e-07	5.103e-07	1.495
v18	-6.845e-07	4.760e-07	-1.438

v25	-1.727e-02	4.271e-03	-4.044
v29	-3.128e-07	5.524e-07	-0.566
v34	-5.830e-07	4.023e-07	-1.449
v42	-5.258e-04	2.359e-04	-2.229
v45	1.894e-04	3.502e-04	0.541
	Pr(> t)		
(Intercept)	1.92e-05	***	
gdp	0.41719		
beds	0.03252	*	
index	0.38444		
v2	0.00968	**	
v3	0.10951		
v6	0.33414		
v13	0.13500		
v18	0.15039		
v25	5.29e-05	***	
v29	0.57124		
v34	0.14729		
v42	0.02586	*	
v45	0.58867		

<u>Signif. codes:</u>			
0	'***'	0.001	'**'
0.01	'*'	0.05	'. '
0.1	' '	1	

Residual standard error: 0.3515 on 13119 degrees of freedom

Multiple R-squared: 0.005969, Adjusted R-squared: 0.004984

F-statistic: 6.06 on 13 and 13119 DF, p-value: 2.04e-11

```

> sse_ku <- sum((fitted(k_linear_model) -
kharif$v46)^2)
> print(sse_ku)
[1] 1620.501
>
> rabi_dummy_linear_model <- lm(formula = v46 ~ gdp
+ beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29
+ v34 + v42 + v45 + West, data = rabi )
> summary(rabi_dummy_linear_model)

```

Call:

```

lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
      v18 + v25 + v29 + v34 + v42 + v45 + West, data =
rabi)

```

Residuals:

Min	1Q	Median	3Q	Max
-0.1963	-0.0739	-0.0574	-0.0384	7.6214

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.668e+00	4.242e-01	3.933
gdp	3.946e-10	2.100e-10	1.879
beds	-1.463e-07	7.239e-08	-2.022
index	-7.058e-04	3.570e-04	-1.977
v2	1.262e-06	5.135e-07	2.458
v3	-7.965e-07	4.789e-07	-1.663
v6	-4.067e-07	2.332e-07	-1.744
v13	1.142e-06	5.167e-07	2.210
v18	-2.941e-07	4.718e-07	-0.623

v25	-1.608e-02	4.320e-03	-3.723
v29	-5.232e-07	5.494e-07	-0.952
v34	-7.502e-07	4.001e-07	-1.875
v42	-4.102e-04	2.350e-04	-1.746
v45	7.947e-05	3.471e-04	0.229
West	-3.088e-02	1.207e-02	-2.558
	Pr(> t)		
(Intercept)	8.45e-05	***	
gdp	0.060283	.	
beds	0.043218	*	
index	0.048040	*	
v2	0.013975	*	
v3	0.096291	.	
v6	0.081170	.	
v13	0.027152	*	
v18	0.533089		
v25	0.000198	***	
v29	0.341009		
v34	0.060801	.	
v42	0.080868	.	
v45	0.818915		
West	0.010536	*	

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3372 on 11898 degrees of freedom

Multiple R-squared: 0.005667, Adjusted R-squared:
0.004497

F-statistic: 4.843 on 14 and 11898 DF, p-value:

5.127e-09

```
> sse_rd <- sum((fitted(rabi_dummy_linear_model) -  
rabi$v46)^2)  
> print(sse_rd)  
[1] 1353.089  
>  
> rabi_linear_model <- lm(formula = v46 ~ gdp + beds +  
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34  
+ v42 + v45, data = rabi )  
> summary(rabi_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1800	-0.0730	-0.0589	-0.0393	7.6175

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.495e+00	4.188e-01	3.569
gdp	4.907e-11	1.609e-10	0.305
beds	-5.891e-08	6.382e-08	-0.923
index	-6.680e-04	3.568e-04	-1.872
v2	1.248e-06	5.136e-07	2.430
v3	-7.713e-07	4.789e-07	-1.611
v6	-2.793e-07	2.279e-07	-1.226
v13	9.239e-07	5.097e-07	1.812

v18	-3.140e-07	4.718e-07	-0.666
v25	-1.428e-02	4.263e-03	-3.349
v29	-5.231e-07	5.495e-07	-0.952
v34	-6.613e-07	3.987e-07	-1.659
v42	-4.217e-04	2.350e-04	-1.795
v45	6.980e-05	3.472e-04	0.201
	Pr(> t)		
(Intercept)	0.000359	***	
gdp	0.760365		
beds	0.356012		
index	0.061171	.	
v2	0.015120	*	
v3	0.107315		
v6	0.220245		
v13	0.069940	.	
v18	0.505723		
v25	0.000813	***	
v29	0.341168		
v34	0.097168	.	
v42	0.072753	.	
v45	0.840659		

<u>Signif. codes:</u>			
0	'***'	0.001	'**'
	0.01	'*'	0.05
	.	0.1	'
	'	1	

Residual standard error: 0.3373 on 11899 degrees of freedom

Multiple R-squared: 0.00512, Adjusted R-squared: 0.004033

F-statistic: 4.71 on 13 and 11899 DF, p-value: 3.328e-08

```

> sse_ru <- sum((fitted(rabi_linear_model) -
rabi$v46)^2)
> print(sse_ru)
[1] 1353.833
>
>
>
>
>
> k_dummy_linear_model <- lm(formula = v46 ~ gdp + beds +
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34
+ v42 + v45 + East , data = kharif )
>
> summary(k_dummy_linear_model)

```

Call:

```

lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
v18 + v25 + v29 + v34 + v42 + v45 + East, data =
kharif)

```

Residuals:

Min	1Q	Median	3Q	Max
-0.2124	-0.0794	-0.0619	-0.0339	7.6100

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.667e+00	4.216e-01	3.954
gdp	1.072e-10	1.527e-10	0.702
beds	-1.669e-07	6.414e-08	-2.603

<u>index</u>	-3.501e-04	3.808e-04	-0.919
v2	1.395e-06	5.302e-07	2.630
v3	-9.305e-07	4.644e-07	-2.004
v6	-2.567e-07	2.300e-07	-1.116
v13	6.374e-07	5.119e-07	1.245
v18	-6.317e-07	4.761e-07	-1.327
v25	-1.595e-02	4.292e-03	-3.717
v29	-1.290e-07	5.556e-07	-0.232
v34	-3.220e-07	4.113e-07	-0.783
v42	-5.332e-04	2.359e-04	-2.261
v45	1.752e-04	3.501e-04	0.500
East	-3.823e-02	1.260e-02	-3.035
<hr/>			
	Pr(> t)		
(Intercept)	7.73e-05	***	
gdp	0.482564		
beds	0.009250	**	
<u>index</u>	0.357939		
v2	0.008539	**	
v3	0.045109	*	
v6	0.264560		
v13	0.213046		
v18	0.184626		
v25	0.000202	***	
v29	0.816325		
v34	0.433738		
v42	0.023792	*	
v45	0.616747		
East	0.002410	**	
<hr/>			
<i>Signif. codes:</i>			
0	'***'	0.001	'**'
0.01	'*'	0.05	'.'
0.1	' '		' 1

Residual standard error: 0.3513 on 13118 degrees of freedom

Multiple R-squared: 0.006667, Adjusted R-squared: 0.005607

F-statistic: 6.289 on 14 and 13118 DF, p-value: 9.914e-13

>

```
> sse_kd <- sum((fitted(k_dummy_linear_model) -  
kharif$v46)^2)
```

```
> print(sse_kd)
```

[1] 1619.364

>

```
> k_linear_model <- lm(formula = v46 ~ gdp + beds + index  
+ v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42  
+ v45, data = kharif )
```

```
> summary(k_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = kharif)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.2143	-0.0772	-0.0610	-0.0380	7.6132

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.794e+00	4.196e-01	4.276

<u>gdp</u>	<u>1.238e-10</u>	<u>1.526e-10</u>	<u>0.811</u>		
<u>beds</u>	<u>-1.354e-07</u>	<u>6.330e-08</u>	<u>-2.138</u>		
<u>index</u>	<u>-3.313e-04</u>	<u>3.809e-04</u>	<u>-0.870</u>		
<u>v2</u>	<u>1.372e-06</u>	<u>5.303e-07</u>	<u>2.587</u>		
<u>v3</u>	<u>-7.364e-07</u>	<u>4.601e-07</u>	<u>-1.601</u>		
<u>v6</u>	<u>-2.220e-07</u>	<u>2.298e-07</u>	<u>-0.966</u>		
<u>v13</u>	<u>7.628e-07</u>	<u>5.103e-07</u>	<u>1.495</u>		
<u>v18</u>	<u>-6.845e-07</u>	<u>4.760e-07</u>	<u>-1.438</u>		
<u>v25</u>	<u>-1.727e-02</u>	<u>4.271e-03</u>	<u>-4.044</u>		
<u>v29</u>	<u>-3.128e-07</u>	<u>5.524e-07</u>	<u>-0.566</u>		
<u>v34</u>	<u>-5.830e-07</u>	<u>4.023e-07</u>	<u>-1.449</u>		
<u>v42</u>	<u>-5.258e-04</u>	<u>2.359e-04</u>	<u>-2.229</u>		
<u>v45</u>	<u>1.894e-04</u>	<u>3.502e-04</u>	<u>0.541</u>		
<hr/>					
$\Pr(> t)$					
<u>(Intercept)</u>	<u>1.92e-05</u>	***			
<u>gdp</u>	<u>0.41719</u>				
<u>beds</u>	<u>0.03252</u>	*			
<u>index</u>	<u>0.38444</u>				
<u>v2</u>	<u>0.00968</u>	**			
<u>v3</u>	<u>0.10951</u>				
<u>v6</u>	<u>0.33414</u>				
<u>v13</u>	<u>0.13500</u>				
<u>v18</u>	<u>0.15039</u>				
<u>v25</u>	<u>5.29e-05</u>	***			
<u>v29</u>	<u>0.57124</u>				
<u>v34</u>	<u>0.14729</u>				
<u>v42</u>	<u>0.02586</u>	*			
<u>v45</u>	<u>0.58867</u>				
<hr/>					
<i>Signif. codes:</i>					
<u>0</u>	<u>'***'</u>	<u>0.001</u>	<u>'**'</u>		
		<u>0.01</u>	<u>'*'</u>		
		<u>0.05</u>	<u>'. '</u>		
		<u>0.1</u>	<u>' '</u>		
			<u>1</u>		

Residual standard error: 0.3515 on 13119 degrees of freedom

Multiple R-squared: 0.005969, Adjusted R-squared: 0.004984

F-statistic: 6.06 on 13 and 13119 DF, p-value: 2.04e-11

```
> sse_ku <- sum((fitted(k_linear_model) -  
kharif$v46)^2)  
> print(sse_ku)  
[1] 1620.501  
≥  
> rabi_dummy_linear_model <- lm(formula = v46 ~ gdp  
+ beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29  
+ v34 + v42 + v45 + East , data = rabi )  
> summary(rabi_dummy_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45 + East, data =  
rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1762	-0.0755	-0.0583	-0.0359	7.6134

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.355e+00	4.215e-01	3.215

<u>gdp</u>	<u>2.778e-11</u>	<u>1.610e-10</u>	<u>0.173</u>
<u>beds</u>	<u>-9.222e-08</u>	<u>6.485e-08</u>	<u>-1.422</u>
<u>index</u>	<u>-6.934e-04</u>	<u>3.568e-04</u>	<u>-1.944</u>
<u>v2</u>	<u>1.190e-06</u>	<u>5.138e-07</u>	<u>2.316</u>
<u>v3</u>	<u>-9.370e-07</u>	<u>4.822e-07</u>	<u>-1.943</u>
<u>v6</u>	<u>-3.020e-07</u>	<u>2.279e-07</u>	<u>-1.325</u>
<u>v13</u>	<u>8.271e-07</u>	<u>5.107e-07</u>	<u>1.619</u>
<u>v18</u>	<u>-2.188e-07</u>	<u>4.729e-07</u>	<u>-0.463</u>
<u>v25</u>	<u>-1.283e-02</u>	<u>4.292e-03</u>	<u>-2.989</u>
<u>v29</u>	<u>-3.701e-07</u>	<u>5.520e-07</u>	<u>-0.671</u>
<u>v34</u>	<u>-4.030e-07</u>	<u>4.086e-07</u>	<u>-0.986</u>
<u>v42</u>	<u>-4.296e-04</u>	<u>2.349e-04</u>	<u>-1.829</u>
<u>v45</u>	<u>7.495e-05</u>	<u>3.471e-04</u>	<u>0.216</u>
<u>East</u>	<u>-3.108e-02</u>	<u>1.083e-02</u>	<u>-2.869</u>
<hr/>			
Pr(> t)			
(Intercept)	<u>0.00131</u>	**	
<u>gdp</u>	<u>0.86301</u>		
<u>beds</u>	<u>0.15502</u>		
<u>index</u>	<u>0.05197</u>	.	
<u>v2</u>	<u>0.02056</u>	*	
<u>v3</u>	<u>0.05203</u>	.	
<u>v6</u>	<u>0.18526</u>		
<u>v13</u>	<u>0.10537</u>		
<u>v18</u>	<u>0.64353</u>		
<u>v25</u>	<u>0.00281</u>	**	
<u>v29</u>	<u>0.50251</u>		
<u>v34</u>	<u>0.32403</u>		
<u>v42</u>	<u>0.06747</u>	.	
<u>v45</u>	<u>0.82902</u>		
<u>East</u>	<u>0.00412</u>	**	
<hr/>			

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3372 on 11898 degrees of freedom

Multiple R-squared: 0.005808, Adjusted R-squared: 0.004638

F-statistic: 4.965 on 14 and 11898 DF, p-value: 2.545e-09

```
> sse_rd <- sum((fitted(rabi_dummy_linear_model) -  
rabi$v46)^2)  
> print(sse_rd)  
[1] 1352.897  
>  
> rabi_linear_model <- lm(formula = v46 ~ gdp + beds +  
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34  
+ v42 + v45, data = rabi )  
> summary(rabi_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = rabi)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1800	-0.0730	-0.0589	-0.0393	7.6175

Coefficients:

Estimate	Std. Error	t value
----------	------------	---------

(Intercept)	1.495e+00	4.188e-01	3.569
gdp	4.907e-11	1.609e-10	0.305
beds	-5.891e-08	6.382e-08	-0.923
index	-6.680e-04	3.568e-04	-1.872
v2	1.248e-06	5.136e-07	2.430
v3	-7.713e-07	4.789e-07	-1.611
v6	-2.793e-07	2.279e-07	-1.226
v13	9.239e-07	5.097e-07	1.812
v18	-3.140e-07	4.718e-07	-0.666
v25	-1.428e-02	4.263e-03	-3.349
v29	-5.231e-07	5.495e-07	-0.952
v34	-6.613e-07	3.987e-07	-1.659
v42	-4.217e-04	2.350e-04	-1.795
v45	6.980e-05	3.472e-04	0.201
Pr(> t)			
(Intercept)	0.000359	***	
gdp	0.760365		
beds	0.356012		
index	0.061171	.	
v2	0.015120	*	
v3	0.107315		
v6	0.220245		
v13	0.069940	.	
v18	0.505723		
v25	0.000813	***	
v29	0.341168		
v34	0.097168	.	
v42	0.072753	.	
v45	0.840659		
<hr/>			
Signif. codes:			

0 **** 0.001 ** 0.01 * 0.05 . 0.1 ' ' 1

Residual standard error: 0.3373 on 11899 degrees of freedom

Multiple R-squared: 0.00512, Adjusted R-squared: 0.004033

F-statistic: 4.71 on 13 and 11899 DF, p-value: 3.328e-08

```
> sse_ru <- sum((fitted(rabi_linear_model) - rabi$v46)^2)
```

```
> print(sse_ru)
```

```
[1] 1353.833
```

```
>
```

```
>
```

```
>
```

```
>
```

```
> k_dummy_linear_model <- lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42 + v45 + Central , data = kharif )
```

```
>
```

```
> summary(k_dummy_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42 + v45 + Central, data = kharif)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.2058	-0.0774	-0.0612	-0.0375	7.6162

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.691e+00	4.233e-01	3.996
gdp	1.310e-10	1.526e-10	0.858
beds	-1.268e-07	6.347e-08	-1.998
index	-3.086e-04	3.811e-04	-0.810
v2	1.453e-06	5.321e-07	2.731
v3	-8.259e-07	4.627e-07	-1.785
v6	-1.979e-07	2.302e-07	-0.860
v13	8.372e-07	5.119e-07	1.635
v18	-7.599e-07	4.777e-07	-1.591
v25	-1.626e-02	4.307e-03	-3.775
v29	-3.189e-07	5.524e-07	-0.577
v34	-5.652e-07	4.024e-07	-1.405
v42	-5.979e-04	2.392e-04	-2.500
v45	2.147e-04	3.504e-04	0.613
Central	1.768e-02	9.699e-03	1.823
	Pr(> t)		
(Intercept)	6.49e-05	***	
gdp	0.390852		
beds	0.045791	*	
index	0.418042		
v2	0.006329	**	
v3	0.074247	.	
v6	0.389961		
v13	0.101998		
v18	0.111673		
v25	0.000161	***	

```
v29      0.563748  
v34      0.160176  
v42      0.012445 *  
v45      0.540069  
Central  0.068326 .
```

Signif. codes:

```
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3514 on 13118 degrees of freedom

Multiple R-squared: 0.006221, Adjusted R-squared: 0.005161

F-statistic: 5.866 on 14 and 13118 DF, p-value: 1.255e-11

>

```
> sse_kd <- sum((fitted(k_dummy_linear_model) -  
kharif$v46)^2)  
> print(sse_kd)  
[1] 1620.09
```

>

```
> k_linear_model <- lm(formula = v46 ~ gdp + beds+ index  
+ v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34 + v42  
+ v45 , data = kharif )  
> summary(k_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +  
v13 +  
v18 + v25 + v29 + v34 + v42 + v45, data = kharif)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.2143	-0.0772	-0.0610	-0.0380	7.6132

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.794e+00	4.196e-01	4.276
gdp	1.238e-10	1.526e-10	0.811
beds	-1.354e-07	6.330e-08	-2.138
index	-3.313e-04	3.809e-04	-0.870
v2	1.372e-06	5.303e-07	2.587
v3	-7.364e-07	4.601e-07	-1.601
v6	-2.220e-07	2.298e-07	-0.966
v13	7.628e-07	5.103e-07	1.495
v18	-6.845e-07	4.760e-07	-1.438
v25	-1.727e-02	4.271e-03	-4.044
v29	-3.128e-07	5.524e-07	-0.566
v34	-5.830e-07	4.023e-07	-1.449
v42	-5.258e-04	2.359e-04	-2.229
v45	1.894e-04	3.502e-04	0.541
	Pr(> t)		
(Intercept)	1.92e-05	***	
gdp	0.41719		
beds	0.03252	*	
index	0.38444		
v2	0.00968	**	
v3	0.10951		
v6	0.33414		
v13	0.13500		
v18	0.15039		

```
v25      5.29e-05 ***
v29      0.57124
v34      0.14729
v42      0.02586 *
v45      0.58867
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3515 on 13119 degrees of freedom

Multiple R-squared: 0.005969, Adjusted R-squared: 0.004984

F-statistic: 6.06 on 13 and 13119 DF, p-value: 2.04e-11

```
> sse_ku <- sum((fitted(k_linear_model) -
kharif$v46)^2)
> print(sse_ku)
[1] 1620.501
>
> rabi_dummy_linear_model <- lm(formula = v46 ~ gdp
+ beds + index + v2 + v3 + v6 + v13 + v18 + v25 + v29
+ v34 + v42 + v45 + Central , data = rabi )
> summary(rabi_dummy_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
v18 + v25 + v29 + v34 + v42 + v45 + Central, data =
rabi)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.1661	-0.0730	-0.0579	-0.0382	7.6217

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.335e+00	4.227e-01	3.158
gdp	5.684e-11	1.608e-10	0.353
beds	-4.708e-08	6.395e-08	-0.736
index	-5.844e-04	3.579e-04	-1.633
v2	1.327e-06	5.142e-07	2.580
v3	-8.973e-07	4.809e-07	-1.866
v6	-2.390e-07	2.283e-07	-1.047
v13	1.022e-06	5.108e-07	2.002
v18	-4.136e-07	4.731e-07	-0.874
v25	-1.270e-02	4.300e-03	-2.952
v29	-5.219e-07	5.494e-07	-0.950
v34	-6.182e-07	3.988e-07	-1.550
v42	-5.259e-04	2.379e-04	-2.210
v45	1.072e-04	3.473e-04	0.309
Central	2.796e-02	1.013e-02	2.760
	Pr(> t)		
(Intercept)	0.00159	**	
gdp	0.72383		
beds	0.46165		
index	0.10255		
v2	0.00988	**	
v3	0.06209	.	
v6	0.29503		
v13	0.04536	*	

```
v18      0.38196
v25      0.00316 **
v29      0.34212
v34      0.12118
v42      0.02710 *
v45      0.75769
Central  0.00579 **

---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3372 on 11898 degrees of freedom

Multiple R-squared: 0.005756, Adjusted R-squared: 0.004586

F-statistic: 4.92 on 14 and 11898 DF, p-value: 3.286e-09

```
> sse_rd <- sum((fitted(rabi_dummy_linear_model) -
rabi$v46)^2)
> print(sse_rd)
[1] 1352.967
>
> rabi_linear_model <- lm(formula = v46 ~ gdp +beds+
index + v2 + v3 + v6 + v13 + v18 + v25 + v29 + v34
+ v42 + v45, data = rabi )
> summary(rabi_linear_model)
```

Call:

```
lm(formula = v46 ~ gdp + beds + index + v2 + v3 + v6 +
v13 +
```

v18 + v25 + v29 + v34 + v42 + v45, data = rabi)

Residuals:

	Min	1Q	Median	3Q	Max
	-0.1800	-0.0730	-0.0589	-0.0393	7.6175

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.495e+00	4.188e-01	3.569
gdp	4.907e-11	1.609e-10	0.305
beds	-5.891e-08	6.382e-08	-0.923
index	-6.680e-04	3.568e-04	-1.872
v2	1.248e-06	5.136e-07	2.430
v3	-7.713e-07	4.789e-07	-1.611
v6	-2.793e-07	2.279e-07	-1.226
v13	9.239e-07	5.097e-07	1.812
v18	-3.140e-07	4.718e-07	-0.666
v25	-1.428e-02	4.263e-03	-3.349
v29	-5.231e-07	5.495e-07	-0.952
v34	-6.613e-07	3.987e-07	-1.659
v42	-4.217e-04	2.350e-04	-1.795
v45	6.980e-05	3.472e-04	0.201
	Pr(> t)		
(Intercept)	0.000359	***	
gdp	0.760365		
beds	0.356012		
index	0.061171	.	
v2	0.015120	*	
v3	0.107315		
v6	0.220245		
v13	0.069940	.	

```
v18      0.505723
v25      0.000813 ***
v29      0.341168
v34      0.097168 .
v42      0.072753 .
v45      0.840659

---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3373 on 11899 degrees of freedom

Multiple R-squared: 0.00512, Adjusted R-squared: 0.004033

F-statistic: 4.71 on 13 and 11899 DF, p-value: 3.328e-08

```
> sse_ru <- sum((fitted(rabi_linear_model) -
rabi$v46)^2)
> print(sse_ru)
[1] 1353.833
```