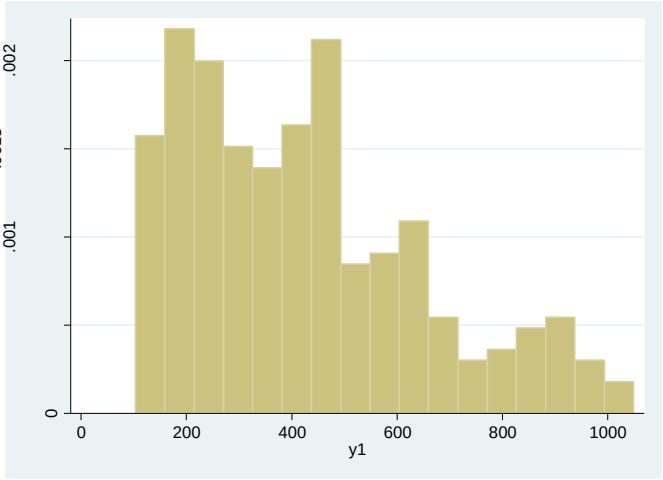
1 Plot histogram of y1i, y2i, y3i, compute descriptive statistics of these three variables, then determine limitation of these three dependent variables.

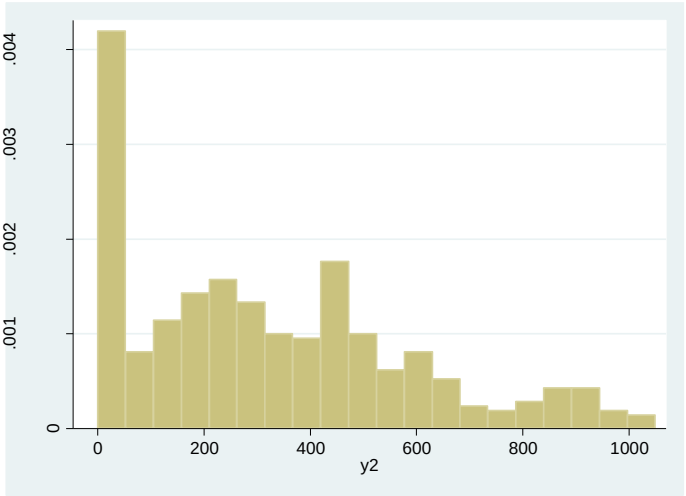
```
. sum y1 y2 y3

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y1 |        297    423.1683    227.8976   103.4663   1049.314
          y2 |        400    317.5914    266.3419          0   1049.314
          y3 |        400     10446.1    23186.57  -866.3348   99508.07
```

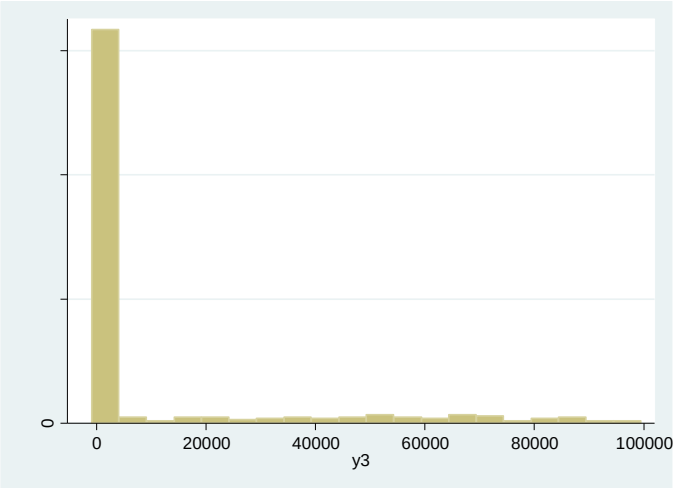
```
histogram y1
(bin=17, start=103.46626, width=55.638129)
```



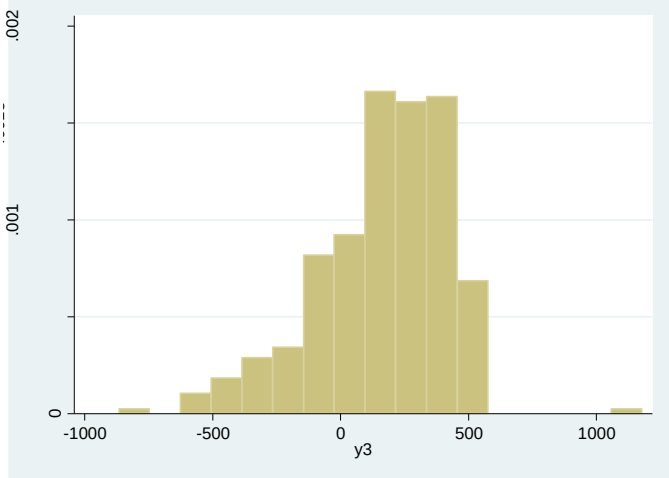

```
histogram y2
(bin=20, start=0, width=52.465723)
```



```
histogram y3
(bin=20, start=-866.33484, width=5018.7203)
```




```
histogram y3 if y3<2000
(bin=17, start=-866.33484, width=120.20064)
```



y1 has truncated problem

y2 has censored problem

y3 has outlier problem.

2 Estimate the model (1) for y1i, y2i, y3i using OLS, using truncated regression model, Tobit model, determine the most appropriated models for each y.

```
reg y1 x

      Source |       SS           df       MS      Number of obs   =       297
-------------+----------------------------------   F(1, 295)       =     30.30
```

```
       Model |  1431905.59         1  1431905.59   Prob > F        =    0.0000
    Residual |  13941534.7       295  47259.4396   R-squared       =    0.0931
-------------+----------------------------------   Adj R-squared   =    0.0901
       Total |  15373440.3       296  51937.2982   Root MSE        =    217.39

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   123.1693   22.37637     5.50   0.000     79.13176    167.2069
       _cons |   69.01677   65.56422     1.05   0.293    -60.01611    198.0496
------------------------------------------------------------------------------

. est store modely1

. truncreg y1 x, ll(103) nolog
(note: 0 obs. truncated)

Truncated regression
Limit:   lower =        103                     Number of obs     =        297
         upper =       +inf                     Wald chi2(1)      =      23.38
Log likelihood = -1973.9423                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   249.6062   51.61679     4.84   0.000     148.4392    350.7733
       _cons |  -456.6913   178.5338    -2.56   0.011    -806.6112   -106.7714
-------------+----------------------------------------------------------------
      /sigma |   306.4488   26.46552    11.58   0.000     254.5774    358.3203
------------------------------------------------------------------------------

. est store y1truncated

.  mfx compute, predict(e(0,.)) at(mean)

Marginal effects after truncreg
```

```
      y  = E(y1|y1>0) (predict, e(0,.))
         =  366.96443
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |   146.2602      26.102    5.60   0.000   95.1014  197.419   2.87532
------------------------------------------------------------------------------

.  mfx compute, predict(e(0,.)) at(median)

Marginal effects after truncreg
      y  = E(y1|y1>0) (predict, e(0,.))
         =  364.18865
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |   145.1256      25.703    5.65   0.000   94.7479  195.503   2.85627
------------------------------------------------------------------------------

. mfx compute, predict(e(0,.)) at(0)

Marginal effects after truncreg
      y  = E(y1|y1>0) (predict, e(0,.))
         =  134.87926
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |    37.5304     3.31249   11.33   0.000    31.038  44.0228         0
------------------------------------------------------------------------------

. lrtest modely1 y1truncated, force

Likelihood-ratio test                                 LR chi2(1)  =     89.69
(Assumption: modely1 nested in y1truncated)           Prob > chi2 =    0.0000
```

Since we reject H0 , so Truncated regression model is a more appropriated model in this case.

```
reg y2 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     76.58
       Model |  4567087.34         1  4567087.34   Prob > F        =    0.0000
    Residual |  23737174.2       398  59641.1413   R-squared       =    0.1614
-------------+----------------------------------   Adj R-squared   =    0.1592
       Total |  28304261.6       399  70937.9989   Root MSE        =    244.22

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   179.8325   20.55046     8.75   0.000     139.4315    220.2335
       _cons |  -179.3679   58.08821    -3.09   0.002     -293.566    -65.1698
------------------------------------------------------------------------------

. est store my2

. tobit y2 x, ll(0)

Tobit regression                                  Number of obs     =        400
                                                  LR chi2(1)        =      73.17
                                                  Prob > chi2       =     0.0000
Log likelihood = -2389.3077                       Pseudo R2         =     0.0151

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   219.1562   24.86468     8.81   0.000     170.2741    268.0383
       _cons |  -318.6265   70.93087    -4.49   0.000    -458.0714   -179.1816
-------------+----------------------------------------------------------------
      /sigma |   284.9073   11.56717                      262.1671    307.6475
------------------------------------------------------------------------------
            72  left-censored observations at y2 <= 0
           328     uncensored observations
```

```
             0 right-censored observations

. est store y2c

. mfx compute, predict(ystar(0,.)) at(mean)

Marginal effects after tobit
      y  = E(y2*|y2>0) (predict, ystar(0,.))
         =  310.40874
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |   184.7744      20.984    8.81   0.000   143.647  225.902   2.76346
------------------------------------------------------------------------------

. mfx compute, predict(ystar(0,.)) at(median)

Marginal effects after tobit
      y  = E(y2*|y2>0) (predict, ystar(0,.))
         =  308.36848
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |   184.3248      20.885    8.83   0.000    143.39  225.259    2.7524
------------------------------------------------------------------------------

. mfx compute, predict(ystar(0,.)) at(0)

Marginal effects after tobit
      y  = E(y2*|y2>0) (predict, ystar(0,.))
         =  18.851247
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
       x |   28.86472     8.50363    3.39   0.001   12.1979  45.5315         0
------------------------------------------------------------------------------
```

```
. lrtest my2 y2c, force

Likelihood-ratio test                                 LR chi2(1)  =    752.97
(Assumption: my2 nested in y2c)                       Prob > chi2 =    0.0000
```

Tobit regression model is a more appropriated model in this case.

```
.  reg y3 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     26.43
       Model |  1.3359e+10         1  1.3359e+10   Prob > F        =    0.0000
    Residual |  2.0115e+11       398   505402183   R-squared       =    0.0623
-------------+----------------------------------   Adj R-squared   =    0.0599
       Total |  2.1451e+11       399   537616802   Root MSE        =     22481

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   9726.042   1891.765     5.14   0.000     6006.941    13445.14
       _cons |  -16431.39   5347.288    -3.07   0.002    -26943.85   -5918.932
------------------------------------------------------------------------------

. est store my3

. reg y3 x if y3<=500

      Source |       SS           df       MS      Number of obs   =       314
-------------+----------------------------------   F(1, 312)       =     30.33
       Model |  1748440.94         1  1748440.94   Prob > F        =    0.0000
    Residual |  17984963.6       312  57644.1141   R-squared       =    0.0886
-------------+----------------------------------   Adj R-squared   =    0.0857
       Total |  19733404.5       313  63046.0209   Root MSE        =    240.09

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
           x |   130.1676   23.63496     5.51   0.000     83.66353    176.6717
       _cons |  -188.1423   64.52556    -2.92   0.004    -315.1025   -61.18199
------------------------------------------------------------------------------

. est store my3out

. tobit y3 x, ul(500) nolog

Tobit regression                                  Number of obs     =        400
                                                  LR chi2(1)        =      67.20
                                                  Prob > chi2       =     0.0000
Log likelihood = -2305.4404                       Pseudo R2         =     0.0144

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   216.5213   25.84755     8.38   0.000     165.7069    267.3357
       _cons |  -325.0374   72.09175    -4.51   0.000    -466.7645   -183.3102
-------------+----------------------------------------------------------------
      /sigma |   292.3574   12.15972                      268.4523    316.2625
------------------------------------------------------------------------------
             0  left-censored observations
           314     uncensored observations
            86 right-censored observations at y3 >= 500

. est store my3out_t

. lrtest my3 my3out_t, force

Likelihood-ratio test                                 LR chi2(1)  =   4538.61
(Assumption: my3 nested in my3out_t)                  Prob > chi2 =    0.0000
```

Tobit regression model is a more appropriated model in this case.