Analysis Panel Survey of Income Dynamics

Executive Summary

1. This data consists of age group between 30 - 50 but most people belong to 30 -35. Out of them
2. we can identify that there are lots of people who are having less salary levels (less than 50000). And there is a few people who are having higher earnings.
3. When the education level increases the salary also increases.
4. Initially when the hours of work are increasing, the earnings are increasing. When it reaches to some level (nearly 2000 hours), earnings are decreasing.
5. Earnings do not highly depend on the age level. People who are young can earn higher salaries or lower salaries and vise versa.
6. Individuals have averagely worked around 1600 hours annually
7. Average labor income of a person is $18894.
8. Divorced people shows high working hours compared to others.
9. Most of the individuals are form education level 11 - 13
10. There is no relationship between earnings and marital status

Git Link : https://github.com/jayasaigoutham/PSID-Analytics-

Group Member	Contribution
189101G - Kaumadi de Zoysa	Data Analysis/ Finding Insights/Documentation
189116G - Jayasai Goutheman	Data Analysis/ Finding Insights/Documentation
189109M - M.T.C.Jayasiri	Data Analysis/ Finding Insights/Documentation

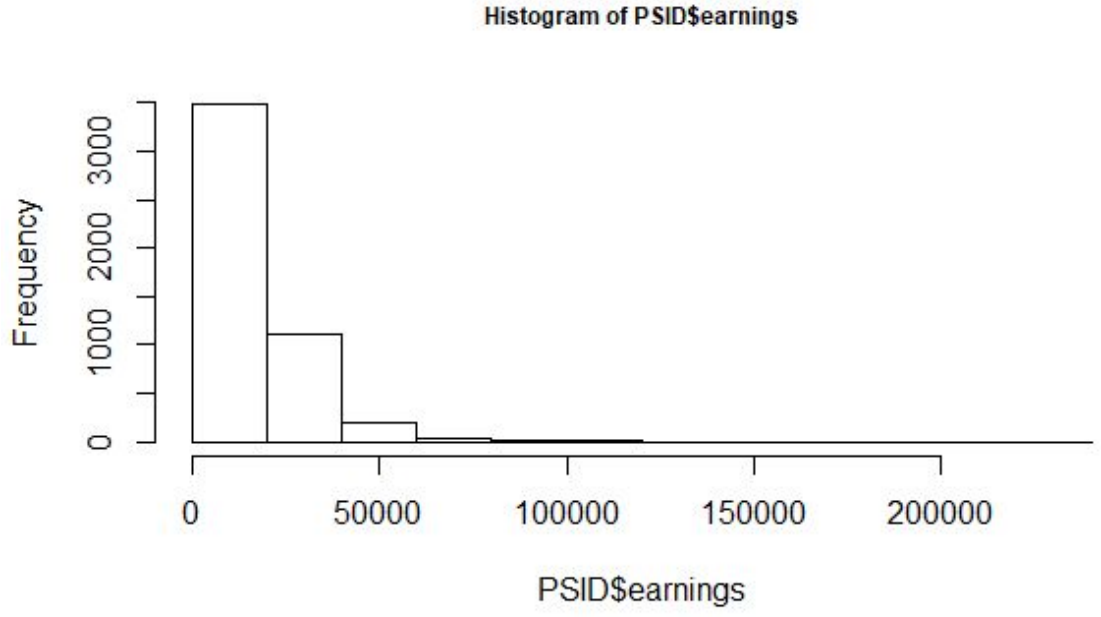
Insights

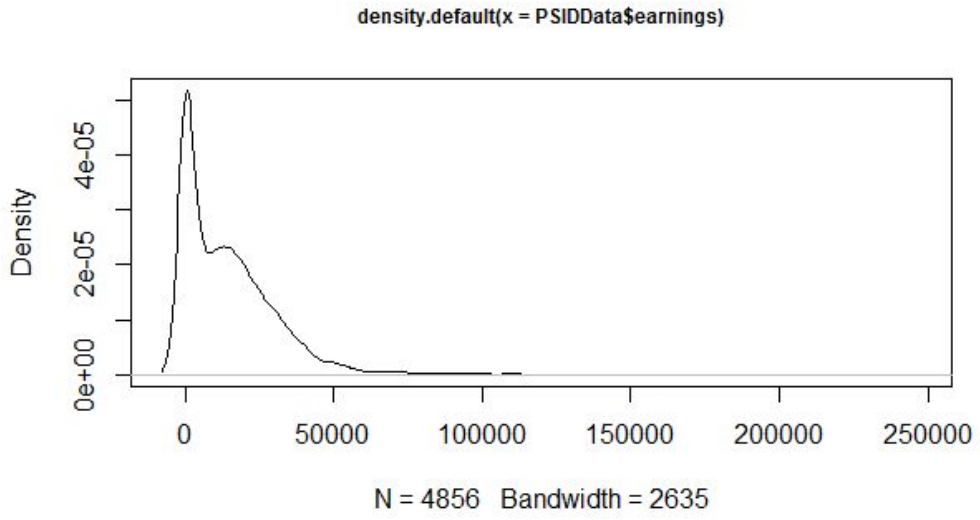
First we observed there are some of outliers and missing values in the dataset.
For instance Kids column contain values of 98 and 99 which are impossible.
And education levels also shown a large gap after level 17 then 98 and 99 values are shown. We have remove those data for first instance.

Summary of the data

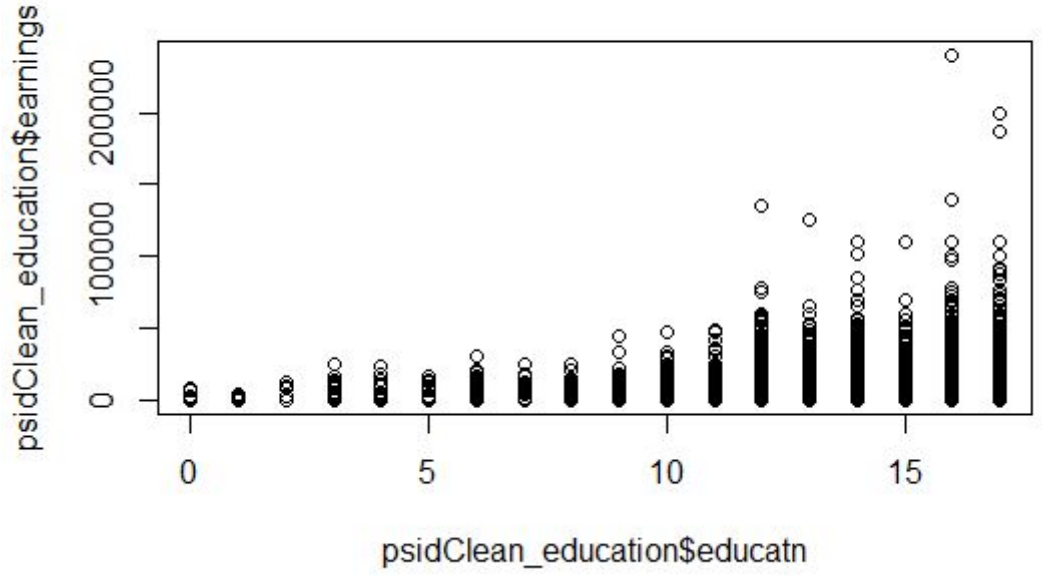
```
     Seq.No          intnum          persnum              age            educatn
 Min.   :   1   Min.   :   4   Min.   :  1.00   Min.   :30.0   Min.   : 0.00
 1st Qu.:1092   1st Qu.:1723   1st Qu.:  2.00   1st Qu.:34.0   1st Qu.:12.00
 Median :2220   Median :5235   Median :  5.00   Median :38.0   Median :12.00
 Mean   :2277   Mean   :4296   Mean   : 60.11   Mean   :38.5   Mean   :12.89
 3rd Qu.:3413   3rd Qu.:6413   3rd Qu.:170.00   3rd Qu.:43.0   3rd Qu.:14.00
 Max.   :4856   Max.   :9306   Max.   :199.00   Max.   :50.0   Max.   :17.00

    earnings          hours           kids                 married
 Min.   :     0   Min.   :   6   Min.   : 0.000   married      :2304
 1st Qu.:  8000   1st Qu.:1213   1st Qu.: 1.000   divorced     : 478
 Median : 16000   Median :1856   Median : 2.000   never married: 428
 Mean   : 18894   Mean   :1633   Mean   : 2.018   separated    : 203
 3rd Qu.: 26000   3rd Qu.:2040   3rd Qu.: 3.000   widowed      :  54
 Max.   :240000   Max.   :5025   Max.   :10.000   (Other)      :   5
```

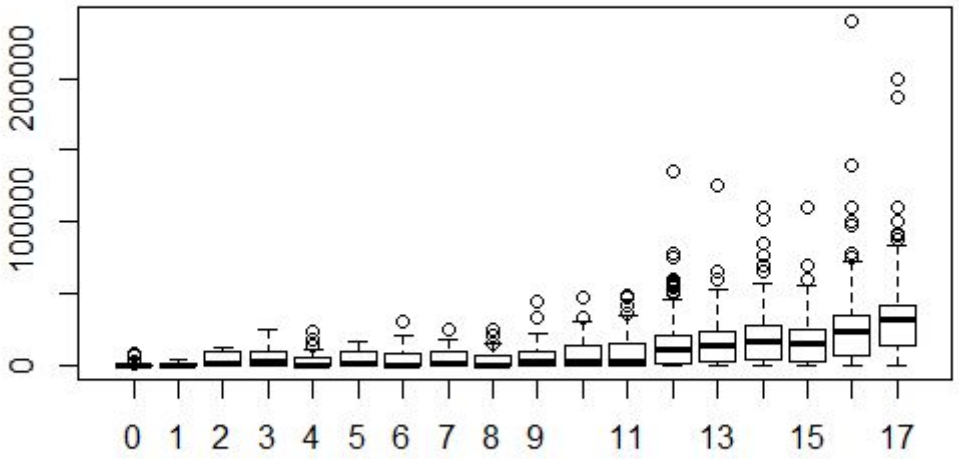




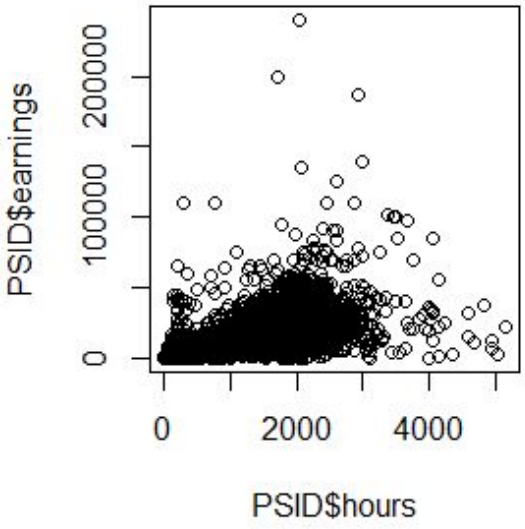
Above Histogram is drawn with the earnings data. From that we can identify that the there are lots of people who are having less salary (less than 50000). And there is a few people who are having higher earnings.



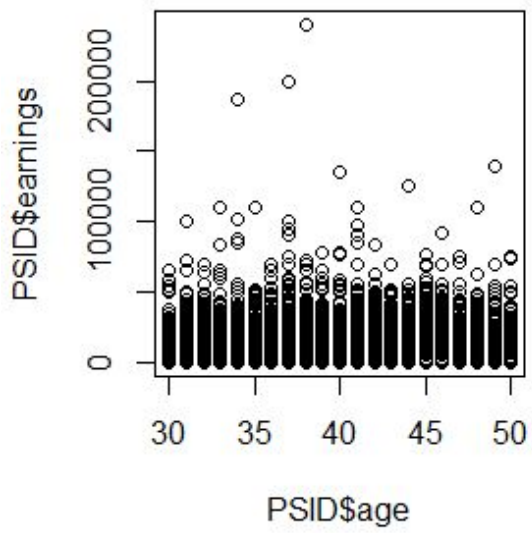
Above plot compares education level of the people in the data set with the earnings. We can say that when the education level increases the salary also increases. Education level which is within 90's range was considered as outliers.



This diagram depicts how earnings are varying in terms of education level. According to the data distribution, upto level 17 earnings are getting increased.



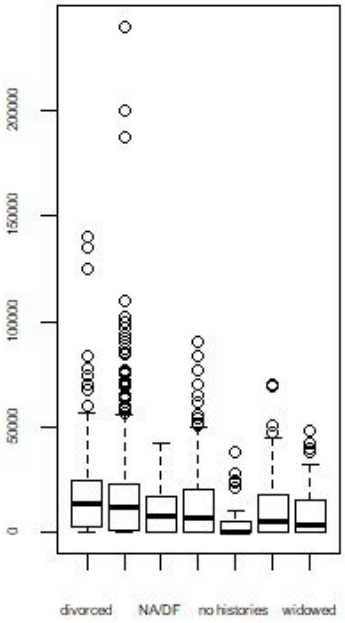
Above plot compares the hours people do work and the earnings. It shows that initially when the hours of work are increasing, the earnings are increasing. When it reaches to some level (nearly 2000 hours), earnings are decreasing.But these is a problem that since we exactly do not know the number of people who are shown in the black area.



Above plot compares the age and the earnings. Generally we think that when people grow old, their earnings will increase gradually, since people will get promotions and etc. But here we cannot see that trend. It is more or less the same. But these is a problem that since we exactly do not know the number of people who are shown in the black area.

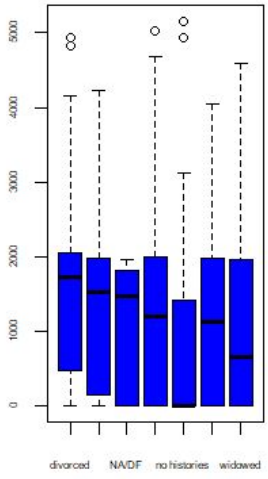
Below table shows how mean of earnings are varies according to their marital status

#	Marital Status	Mean
1	divorced	16486.755
2	married	14990.662
3	NA/DF	11577.778
4	never married	11698.154
5	no histories	4390.860
6	separated	10650.921
7	widowed	9613.878



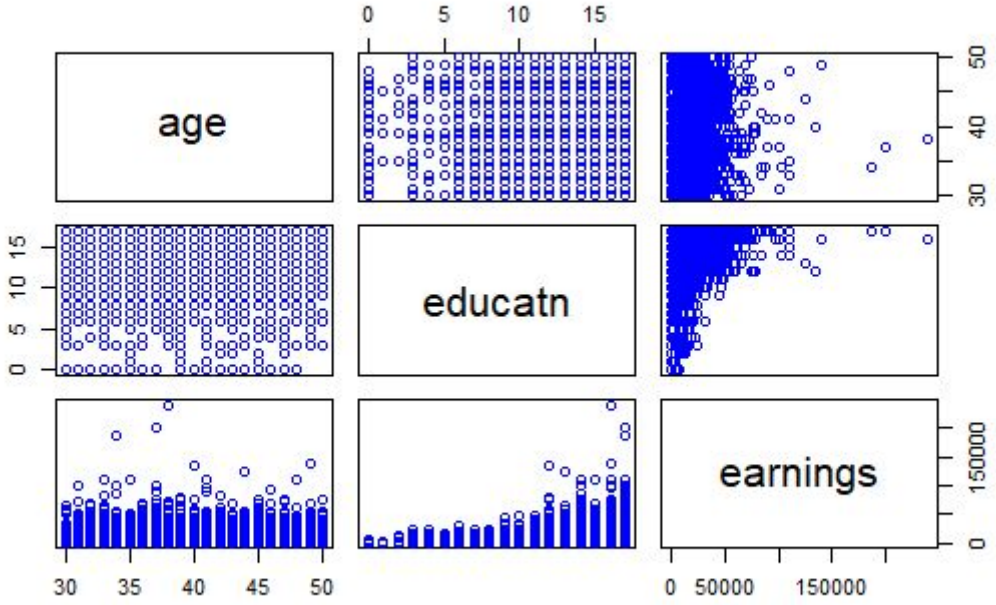
By looking at this we can say there is no a good relationship between marital status and the earnings.

Relationship between total working hours and marital status of person. According to this graph divorced person has high working hours.



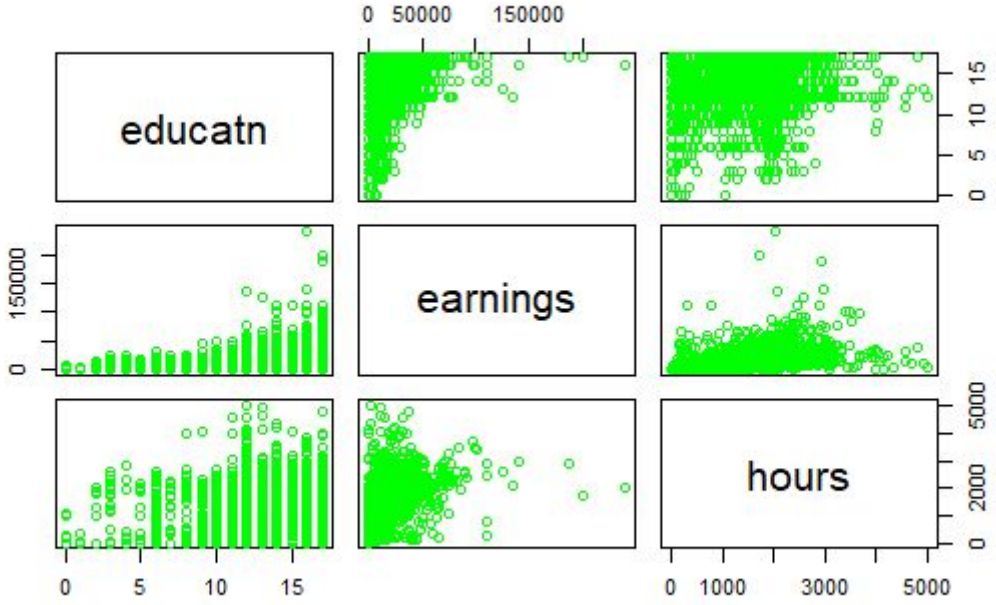
Age Vs Education Vs Earnings

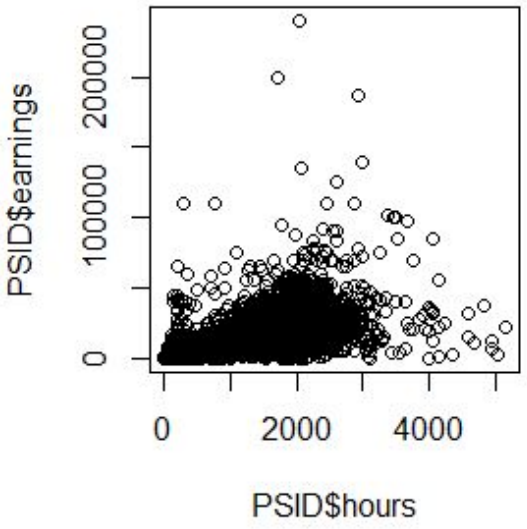
This is the comparison of age, education and earnings.

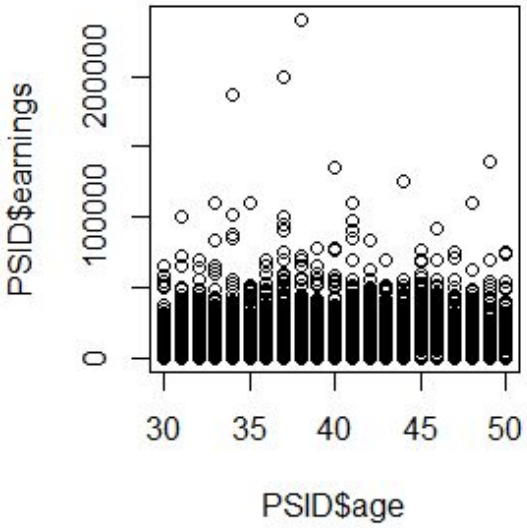


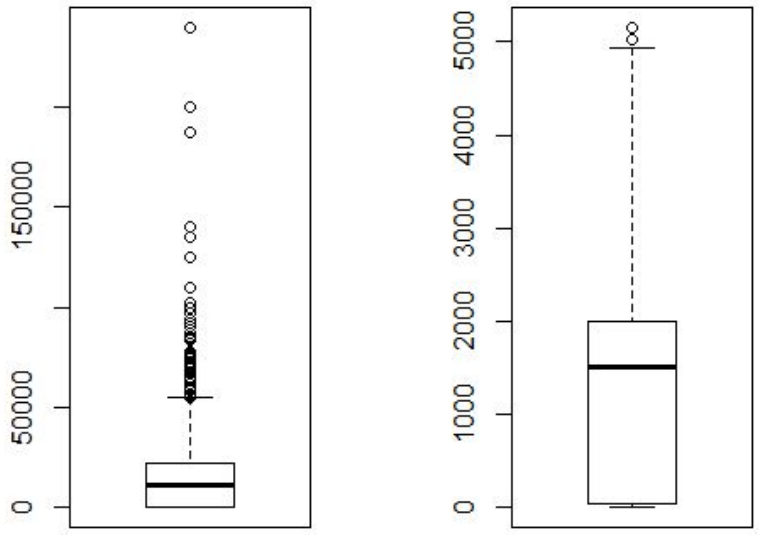
Education Vs Earnings Vs hours

This is the comparison of education, earnings and hours.

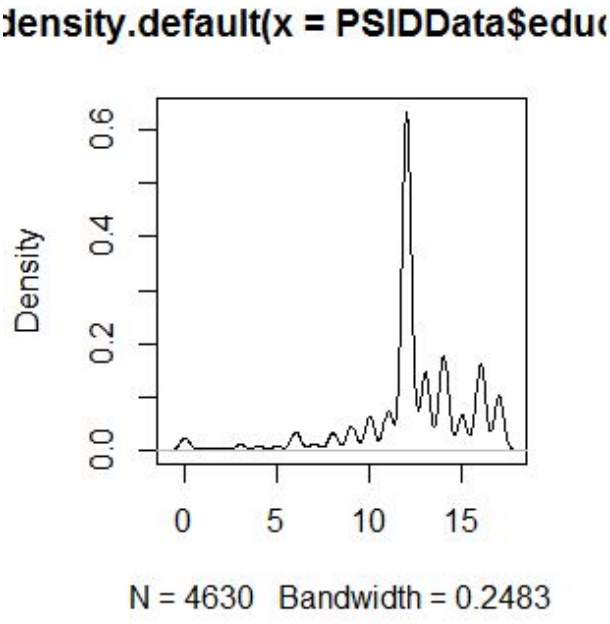



PSID$earnings
0
100000
200000
0
2000
4000
PSID$hours

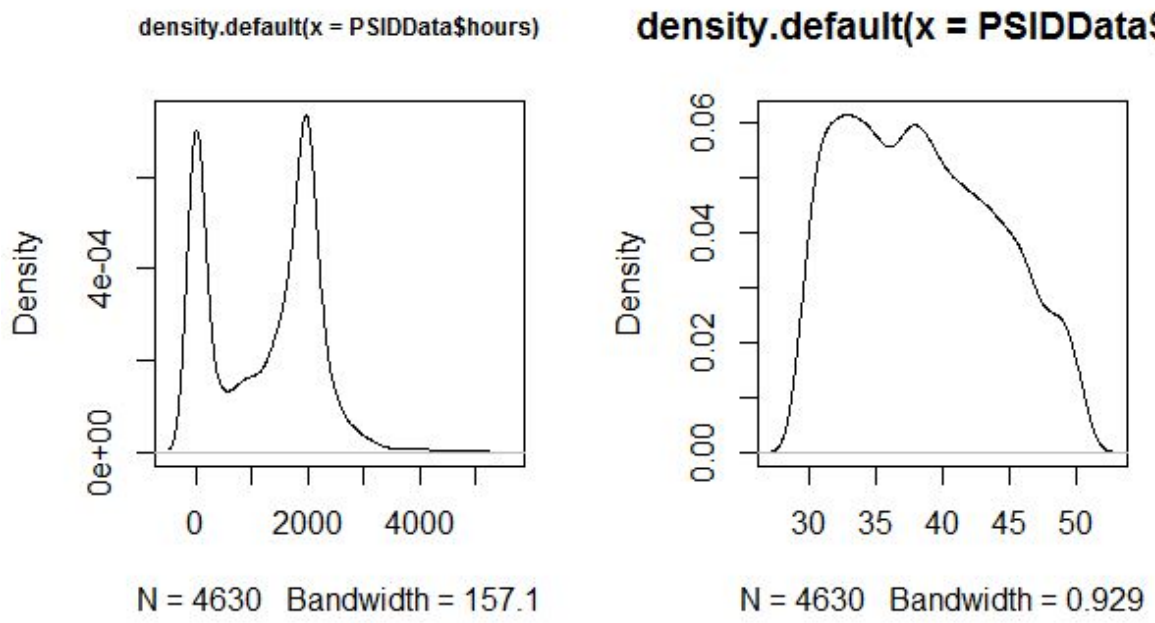

PSID$earnings
0
100000
200000
30
35
40
45
50
PSID$age



First diagram shows the average earnings of the panel members. The mean value is 14244.51
Next box plot shows average working hours of the members. It's 1235.335



When not considering outliers , Most of the individuals are form education level 11 - 13



When not considering outliers , the above graph shows that the most hours worked and individuals from age category