The Impact of Unconditional Cash Transfers on Child Labour in Indonesia

Data depicted information of 6000 households from the years 2005 and 2006, that either received UCT in 2006 or not. All analytical functions were performed in Python. Code used can be seen in **Appendix 1**.

The table below shows the code book created which was used throughout for further analysis and examining the impacts of UCT on incidents of child work.

**Table 1: Code book**

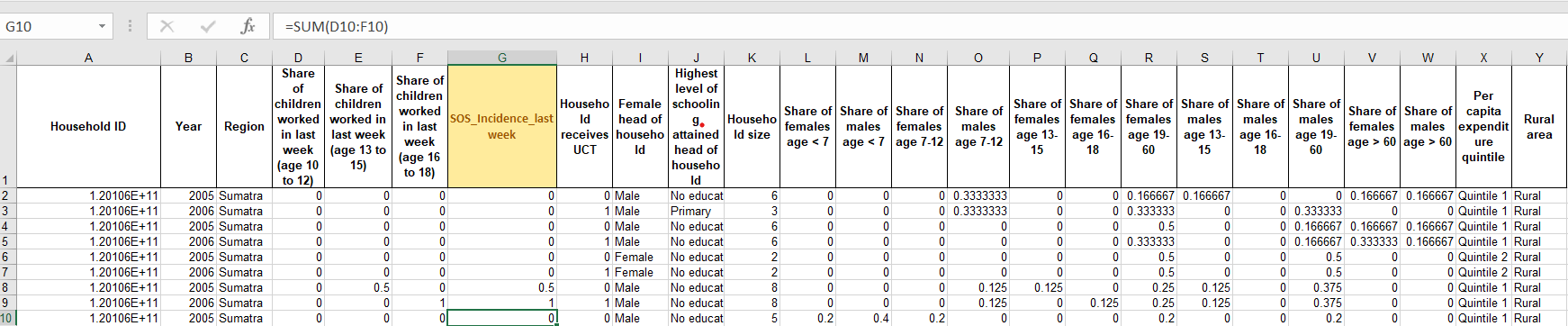
|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Count** |
| Region | 0=Sumatra  1=Java and Bali  2=Other Islands  3=Kalimantan  4=Sulawesi | **Region Distinct Counts** |
| 0= Sumatra=2410  1= Java and Bali=6408  2=Other islands=1210  3=Kalimantan=972  4=Sulawesi=1000 |
| Year | 0= 2005  1= 2006 | 0=6000  1=6000 |
| Female head of household | 0= Male  1= Female | 0 = 3905  1= 450 |
| Level of Schooling of HoH | 0= No education  1= Primary  2= Junior Sec  3= Higher  4= Senior Sec | 0= 864  1= 997  2= 587  3= 1153  4= 754 |
| Per capita expenditure quintile | 0= Quintile 1 (poorest)  1= Quintile 2  2= Quintile 3  3= Quintile 5 (richest)  4= Quintile 4 | 0= 864  1= 997  2= 587  3=1153  4=754 |
| Rural | 0= Rural  1= Urban | 0=2326  1= 2029 |
| Household receives UCT | 0= No UCT received in both years  1= UCT received in 2006 | 0= No UCT in both years = 10355 (86.29%)  1= received UCT in 2006 only = 1645 (13.71 %) |

**Data analysis**

MS Excel and Python were used to analyse the data, and examine the impacts of UCT on incidence of child work.

The share of children that worked in the last week of the survey were presented in the data in 3 groups of age for the years 2005 and 2006 for the same household. To show a cumulative value of all children that worked, the three categories were combined and named “SOS\_Incidence\_last week” (Sum of Share of Incidents of work of all children ranging from 10-18 years of age).

**Table 2: Sum of share of children that worked last week**



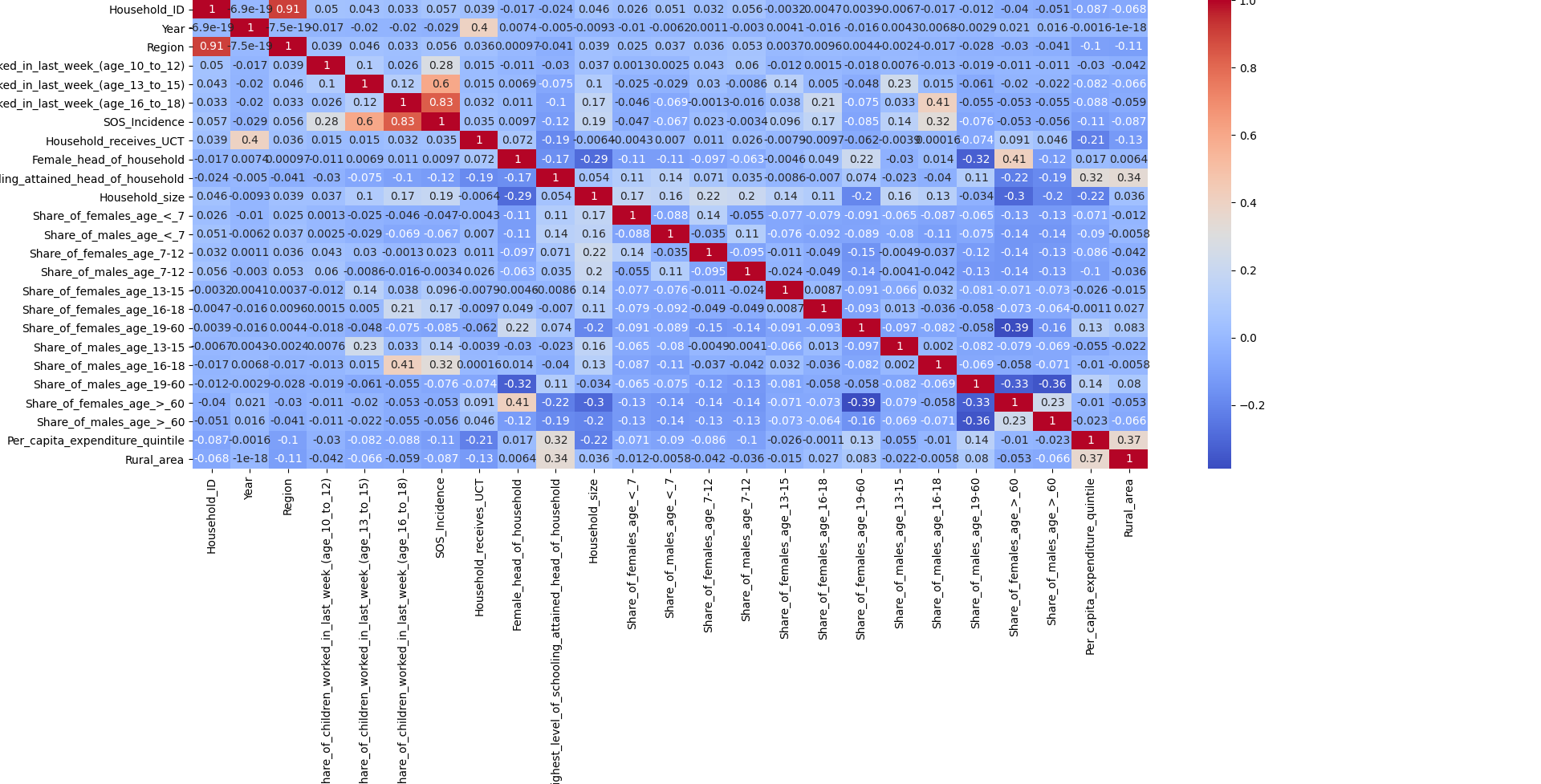
**Correlation**

A measure to depict how strongly two variables are interrelated, it was used to examine the relationship between share of child work and households that received UCT. A negative correlation

indicates a negative relation; a positive correction coefficient indicates a positive relation. The correlation (or correlation coefficient) ranges between 0 and 1; the higher the correlation, the stronger the variables are statistically related.

To inspect the relationship between the *numerical variable (share of children worked)* and a *nominal (categorical)* *variable (household receives UCT)*, *Point Biserial Correlation* test was run and the result was ‘**correlation = 0.035’.** Since the correlation, although positive, was low; it depicted that there is no significant relationship between share of child work done across the age groups of 10-18 years, and the households that received UCT that year. Further, a correlation matrix/correlation heat map was created which showed same correlation as the point biserial correlation described above.

**Table 3: Correlation Heatmap**

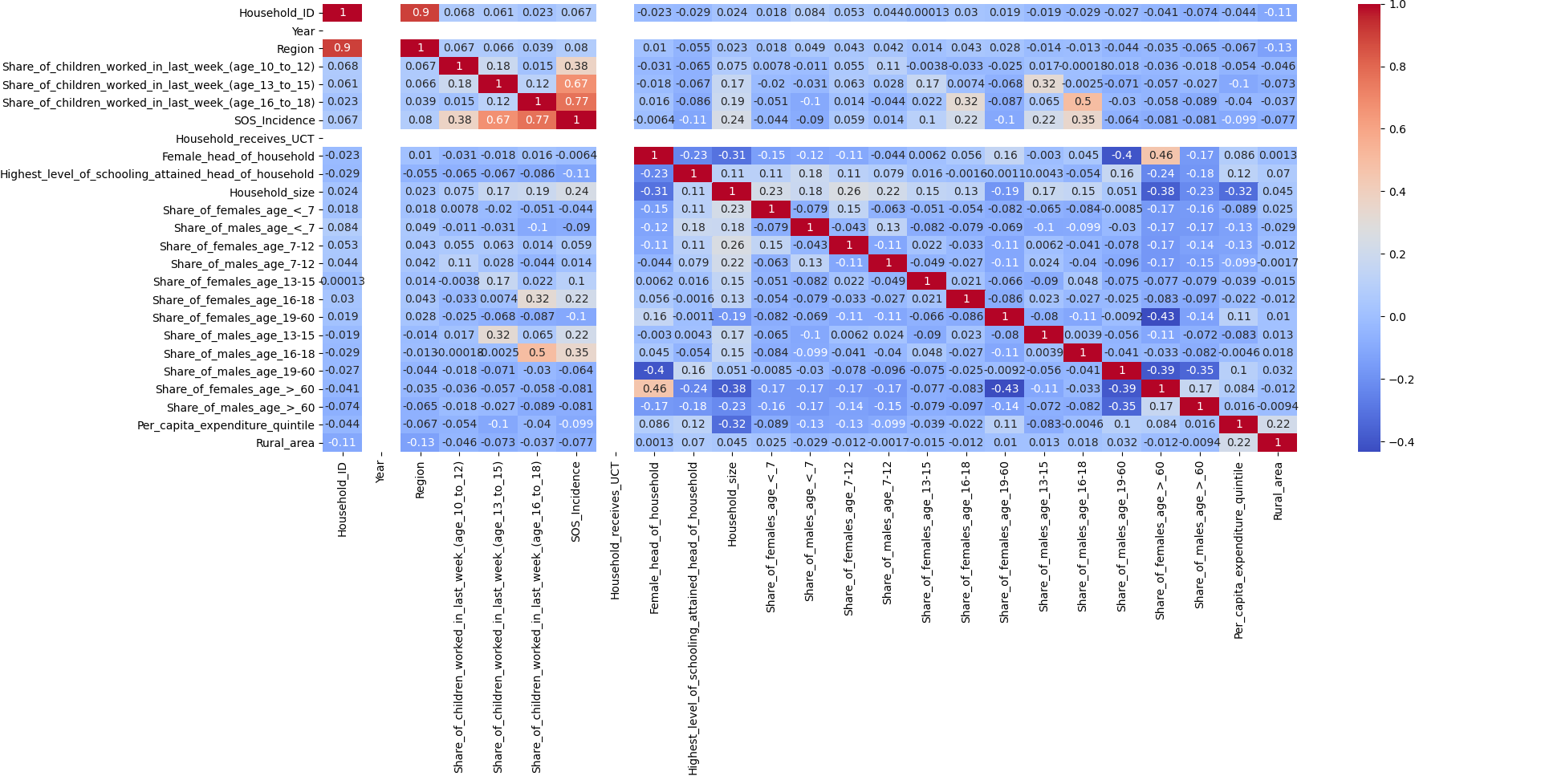
This heatmap confirms that there is no significant correlation between the combined values of child work incidences, or between the three categories of ages depicting share of children that worked in the last week. But it shows other correlations which can have significance, they are;

* there is a strong correlation between sum of share of children that worked and share of children under the category of 16-18 and 13-15 share of children that worked (corel=0.83 and correl=0.6 respectively). This means that a significant share of children that worked were between the ages of 13-18.
* There is a noticeable correlation between share of males of ages 16-18 and the share of work done by children in the same age category (correl=0.4), which tells us that more male children were involved in working in the last week of this survey.

**UCT recipient households**

What were the conditions before households received UCT? Not all households received UCT in the year 2006. To analyse and compare the impacts of the presence of UCT received in 2006 on incidence of child work; correlation among variables was *calculated for households in the year 2005 that RECEIVED unconditional cash transfer in 2006*.

**Table 4: 2005 data of households that RECEIVED UCT in 2006**

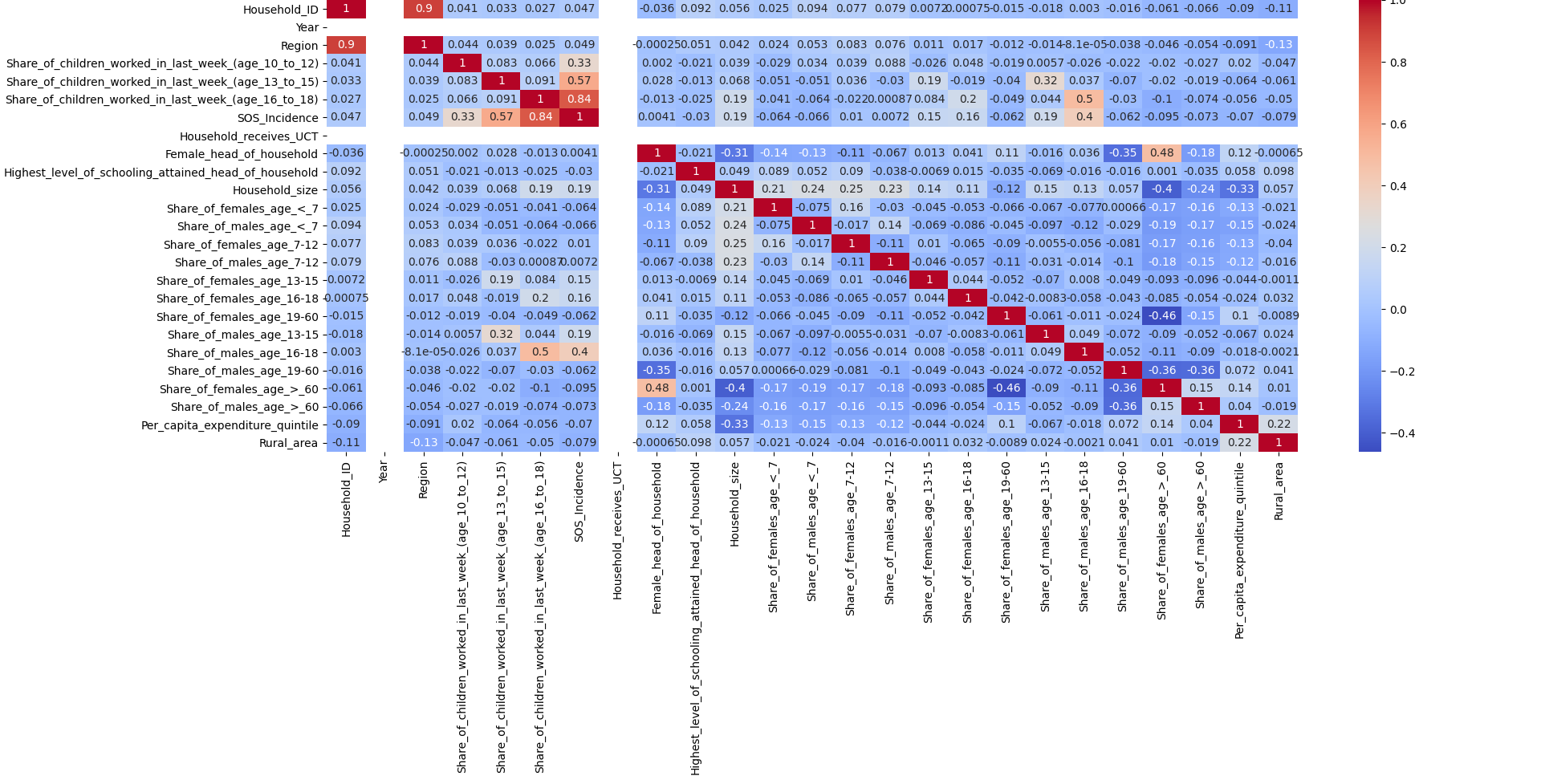


The above correlation matrix tells us that;

* Without the presence of unconditional cash transfer, there is a significant correlation between share of children that worked (SOS\_Incidence in image) and the share of children that worked between the ages of 16 to 18 (Correl=0.77). This correlation between work done by children and total number of share of children work incidences decreases with decrease in age category. It shows us that the maximum number of children that worked in that last week were between the ages of 16 to 18 and then between that ages of 13-15 (Correl=0.67).
* There is a durable correlation between *share of children that worked between the ages of 16-18* and *share of males present in the household of that same age* (Correl=0.5). This means that male children of the ages 16-18 in households comprised to a significant amount of the total working children numbers. Male children contributing to work between 13-15 years also showed correlation (=0.33), and so did female children between 16-18 that worked (Correl=0.32).
* There is a negative correlation (= -0.11) between the variables highest level of schooling attended by head of household and total share of children working across age groups. Which can mean education of head of household did not play any role in deciding of children of that household ended up working.

Comparing data of exactly these households, when they started receiving UCT in 2006 shows;

**Table 5: 2006 data of households that RECEIVED UCT in 2006, but not in 2005**

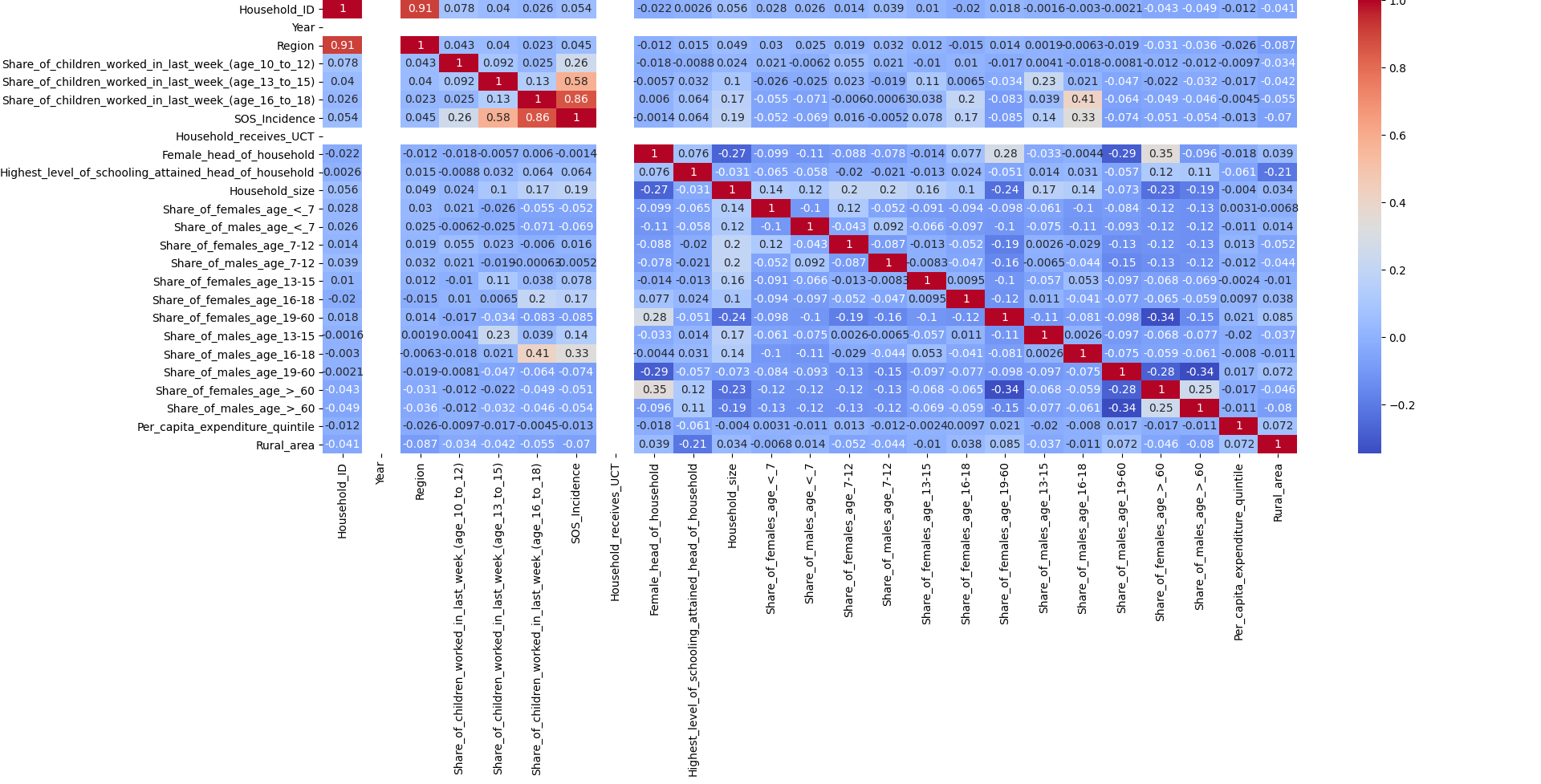


* Total share of children that worked across all ages (between ages 10-18) shows a greater correlation with share of children that worked in the 16-18 age bracket. It shows more children from this age group started working after receiving UCT from the government.
* The data about share of male children that worked (b/w ages 16-18) and the total number of males between this age group is the same, which means reception UCT had no effect on that. But the correlation between share of females in the (b/w age 16-18 and 13-15) and children that worked in these age brackets decreased in the year 2006 decreased as compared to the previous year. It can be concluded that UCT reception had a small effect on child work.
* The highest level of education of head of household showed no correlation with total share of children working after households received UCT, which was not the case in 2005. It can be said that receiving UCT

**Non-recipient of UCT households**

What happened to households that did not receive UCT in 2006? How were incidences of child work different for households that never received UCT? What does the comparison of these households look like with households that received UCT in 2006? To analyse and compare the impacts of the absence of UCT received in 2006 on incidence of child work, correlation among variables in the data was *calculated for households in the year 2005 that DID NOT RECEIVE UCT in 2006*.

**Table 6: 2005 data of households that DID NOT RECEIVE UCT in 2006**

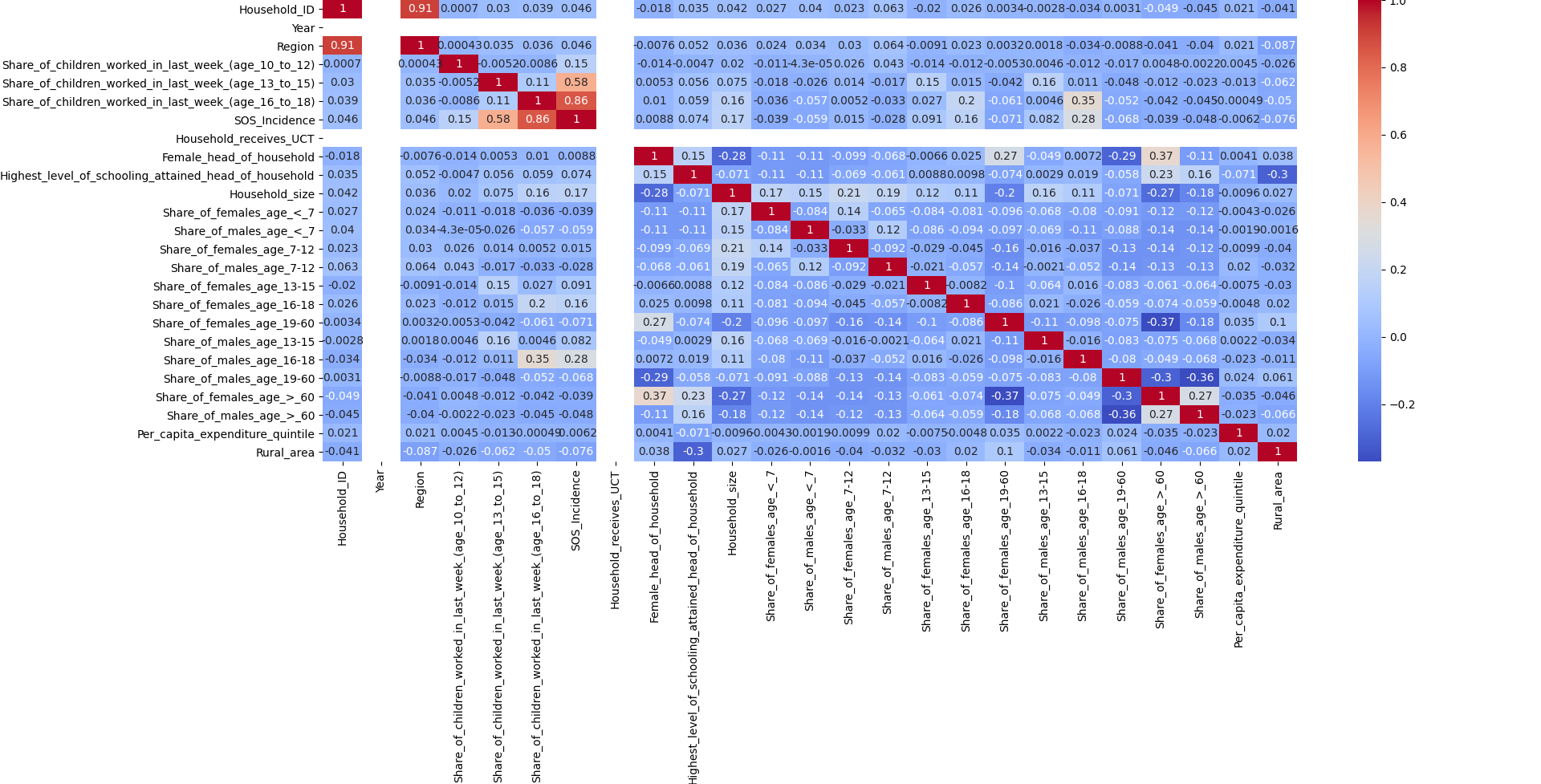


The above correlation heatmap shows that;

* There was a strong correlation between children working in the ages of 16-18 (Correl=0.86) and share of total incidences of child work. Which means children between these ages comprised maximum number of working children. Children between ages 13 to 15 also contributed to total share of working children (Correl=0.58).
* Out of the children working in the ages of 16-18, male children formed the highest number in this group (Correl=0.41), followed by male children of ages 13-15 (Correl=0.23). Female children of these age groups also show contribution in share of children that worked in these age groups (Correl=0.2 and 0.11 respectively).
* In this case, total share of children that worked across all age groups also shows correlation with household size. Which means, the bigger the house, the more children working in that household.

To study the effects of not receiving any UCT in the following year either, the correlation heat map showed

**Table 7: 2006 data of households that DID NOT RECEIVE UCT in 2006**



* There was a strong correlation between children working in the ages of 16-18 (Correl=0.86) and share of total incidences of child work. Which means children between these ages comprised maximum number of working children. Children between ages 13 to 15 also contributed to total share of working children (Correl=0.58).
* Out of the children working in the ages of 16-18, male children formed the highest number in this group (Correl=0.35), followed by male children of ages 13-15 (Correl=0.16). Female children of these age groups also show contribution in share of children that worked in these age groups (Correl=0.2 and 0.15 respectively).

Comparing households that **UCT recipient** and **UCT non-recipient** households in 2006 (Table 5 and Table 7)

* It shows that total share of child work incidences and share of children that worked in ages 10-12 was slightly *less strongly related* in the houses that did not receive UCT in 2006 as compared to recipient households. (Correl=0.15 and 0.33). Which means, receiving UCT made more children in the ages of 10-12 work for households.
* The share of male and female children in non-recipient households were *slightly less* correlated to share children that worked as compared to children of households that did receive UCT.

It should be noted that these correlations were not strong enough to imply causation, hence no regression analysis was conducted. Overall, it can be said that receiving UCT might have only slightly impacted incidence of child labour. There could be other factors also affecting the incidences of child work, as the comparison between recipient and non-recipient households does not show any major differences in impacts of UCT on child labour activities in these regions of Indonesia.

**Appendix 1: Python code**

**import** pandas **as** pd  
**import** seaborn  
**import** matplotlib.pyplot **as** plt  
**from** scipy **import** stats  
  
  
**def** analyze(data):  
 *###Start: Draw Heatmap for all data* plt.figure(figsize=(10, 10))  
 seaborn.heatmap(data.corr(), annot=**True**, cmap=**"coolwarm"**).set\_title(**'Correlation'**)  
 plt.show()  
 *###End: Draw Heatmap for all data  
  
 ###Start: Calculate PointBiSerial correlation* c1 = data.SOS\_Incidence  
 c2 = data.Household\_receives\_UCT  
 c3 = data.Female\_head\_of\_household  
 print(stats.pointbiserialr(c1, c3))  
 print(stats.pointbiserialr(c1, c2))  
 *###End: Calculate PointBiSerial correlation***def** read\_file():  
 *###Start: Read Data and reformat column names* dat = pd.read\_excel(**"Dataset.xls"**)  
 dat.columns = [col.replace(**' '**, **'\_'**) **for** col **in** dat.columns]  
 *###End: Read Data and reformat column names* **return** dat  
  
  
**def** factorize(fact\_data):  
 *###Start: Convert textual categorical values to numerical* fact\_data[**'Year'**] = pd.factorize(fact\_data.Year)[0]  
 fact\_data[**'Region'**] = pd.factorize(fact\_data.Region)[0]  
 fact\_data[**'Female\_head\_of\_household'**] = pd.factorize(fact\_data.Female\_head\_of\_household)[0]  
 fact\_data[**'Highest\_level\_of\_schooling\_attained\_head\_of\_household'**] = pd.factorize(fact\_data.Highest\_level\_of\_schooling\_attained\_head\_of\_household)[0]  
 fact\_data[**'Per\_capita\_expenditure\_quintile'**] = pd.factorize(fact\_data.Per\_capita\_expenditure\_quintile)[0]  
 fact\_data[**'Rural\_area'**] = pd.factorize(fact\_data.Rural\_area)[0]  
 *###End: Convert textual categorical values to numerical* **return** fact\_data  
  
  
**def** print\_stats(statData):  
 *###Start: Print Count of Values for categorical columns* print(statData.Year.value\_counts())  
 print(statData.Region.value\_counts())  
 print(statData.Female\_head\_of\_household.value\_counts())  
 print(statData.Highest\_level\_of\_schooling\_attained\_head\_of\_household.value\_counts())  
 print(statData.Per\_capita\_expenditure\_quintile.value\_counts())  
 print(statData.Rural\_area.value\_counts())  
 *###End: Print Count of Values for categorical columns*a = read\_file()  
print\_stats(a)  
a = factorize(a)  
*##Analyze all Data*analyze(a)  
  
*# Only 2005 data for households that received UCT in 2006 but did not in 2005*b = read\_file()  
b = b.drop\_duplicates(subset=[**'Household\_ID'**, **'Household\_receives\_UCT'**], keep=**False**).reset\_index(drop=**True**)  
b.drop(b[b.Year == 2006].index, inplace=**True**)  
b = factorize(b)  
analyze(b)  
  
*# Only 2005 data for households that did NOT receive UCT*c = read\_file()  
duplicates = c.duplicated(subset=[**'Household\_ID'**, **'Household\_receives\_UCT'**], keep=**False**).reset\_index(drop=**True**)  
c = c[duplicates]  
c.drop(c[c.Year == 2006].index, inplace=**True**)  
c = factorize(c)  
analyze(c)  
  
*# Only 2006 data for households that received UCT in 2006 but did not in 2005*d = read\_file()  
d = d.drop\_duplicates(subset=[**'Household\_ID'**, **'Household\_receives\_UCT'**], keep=**False**).reset\_index(drop=**True**)  
d.drop(d[d.Year == 2005].index, inplace=**True**)  
d = factorize(d)  
analyze(d)  
  
*# Only 2006 data for households that did NOT receive UCT*e = read\_file()  
duplicates = e.duplicated(subset=[**'Household\_ID'**, **'Household\_receives\_UCT'**], keep=**False**).reset\_index(drop=**True**)  
e = e[duplicates]  
e.drop(e[e.Year == 2005].index, inplace=**True**)  
e = factorize(e)  
analyze(e)