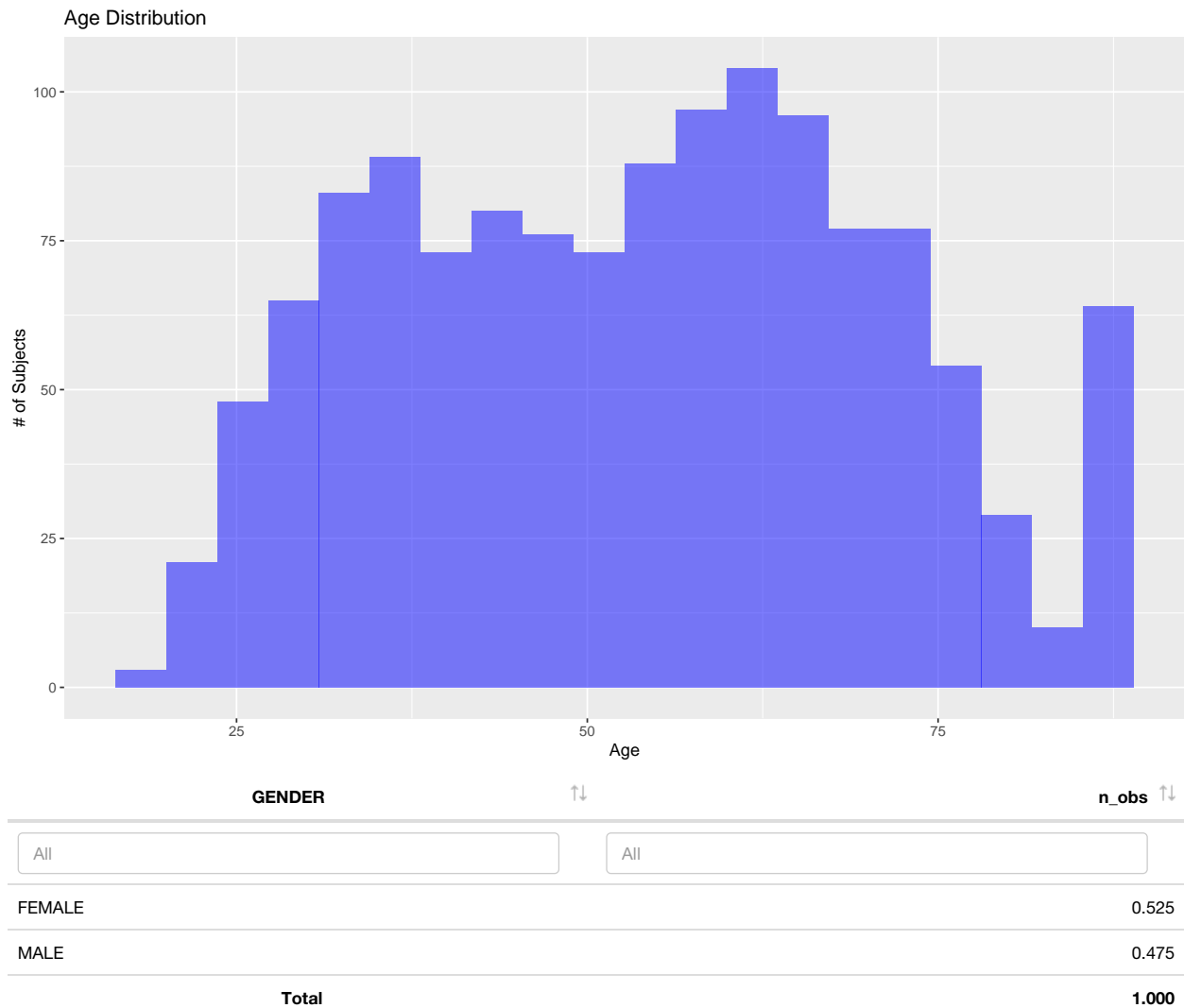


## Group 6 Project - Part B



## Exploratory Data Analysis

There is a moderate correlation between family income and overall expenditure

### Correlation

```
Data      : group6FilterIncome
Method     : Pearson
Variables  : FAM_Income, OVERALL_EXP
Null hyp.  : variables x and y are not correlated
Alt. hyp.  : variables x and y are correlated
```

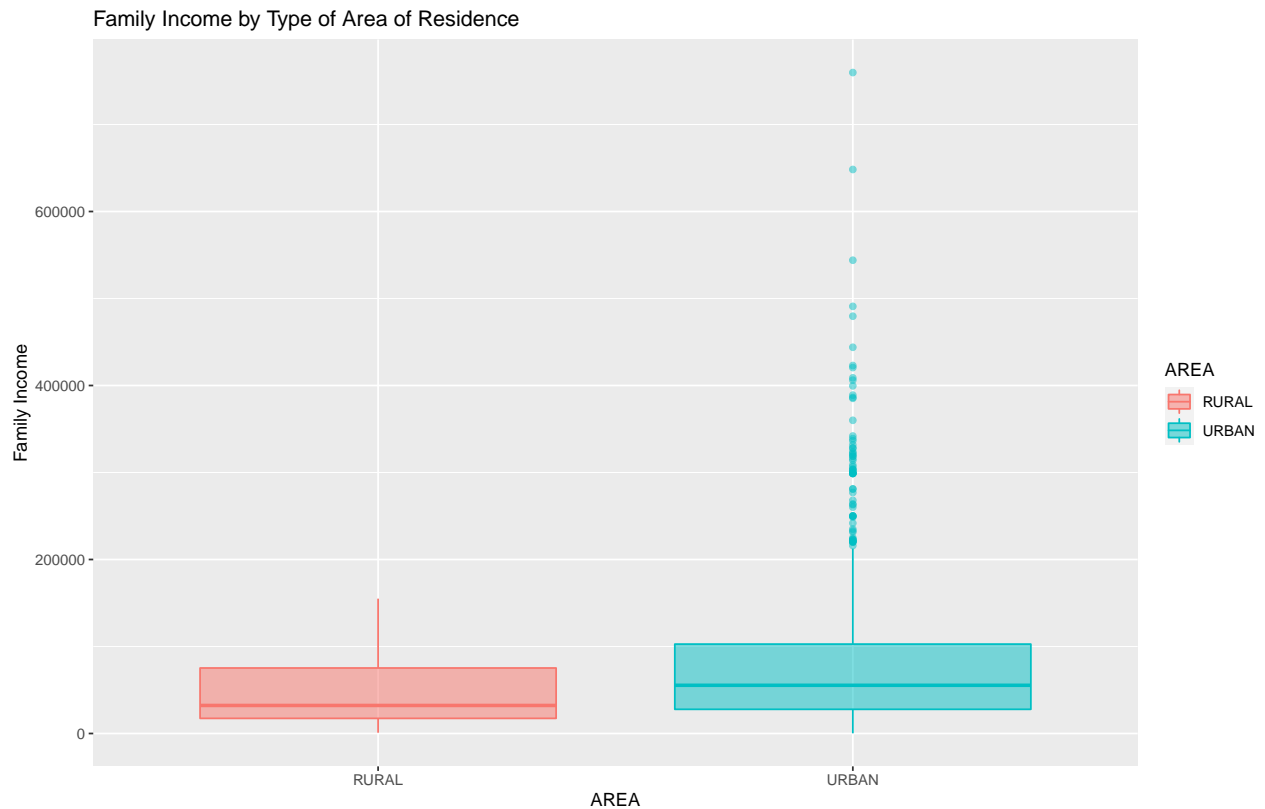
### Correlation matrix:

```
FAM_Income
OVERALL_EXP 0.46
```

### p.values:

```
FAM_Income
OVERALL_EXP 0.00
```

## Rural populations typically have lower household incomes than urban families



$$\text{rate}(\text{high exp} \mid \text{emp}) = 0.522 > 0.495 = \text{rate}(\text{high exp} \mid \sim \text{emp})$$

EMPLOYED			
ENT_EXP_IS_HIGH <span>↑↓</span>	FALSE. <span>↑↓</span>	TRUE. <span>↑↓</span>	Total <span>↑↓</span>
All	All	All	All
FALSE	0.505	0.478	<b>0.487</b>
TRUE	0.495	0.522	<b>0.513</b>
<b>Total</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>

## Correlation between income and Housing Expenditure is highest: 0.50

Correlation

Data : group6FilterIncome

Method : Pearson

Variables : FAM\_Income, ENTERTAINMENT\_EXP, ELCTRC\_EXP, FOOD\_EXP, WATER\_EXP, NAT\_GAS\_EXP, TRANS\_EXP, HEAL

Null hyp. : variables x and y are not correlated

Alt. hyp. : variables x and y are correlated

Correlation matrix:

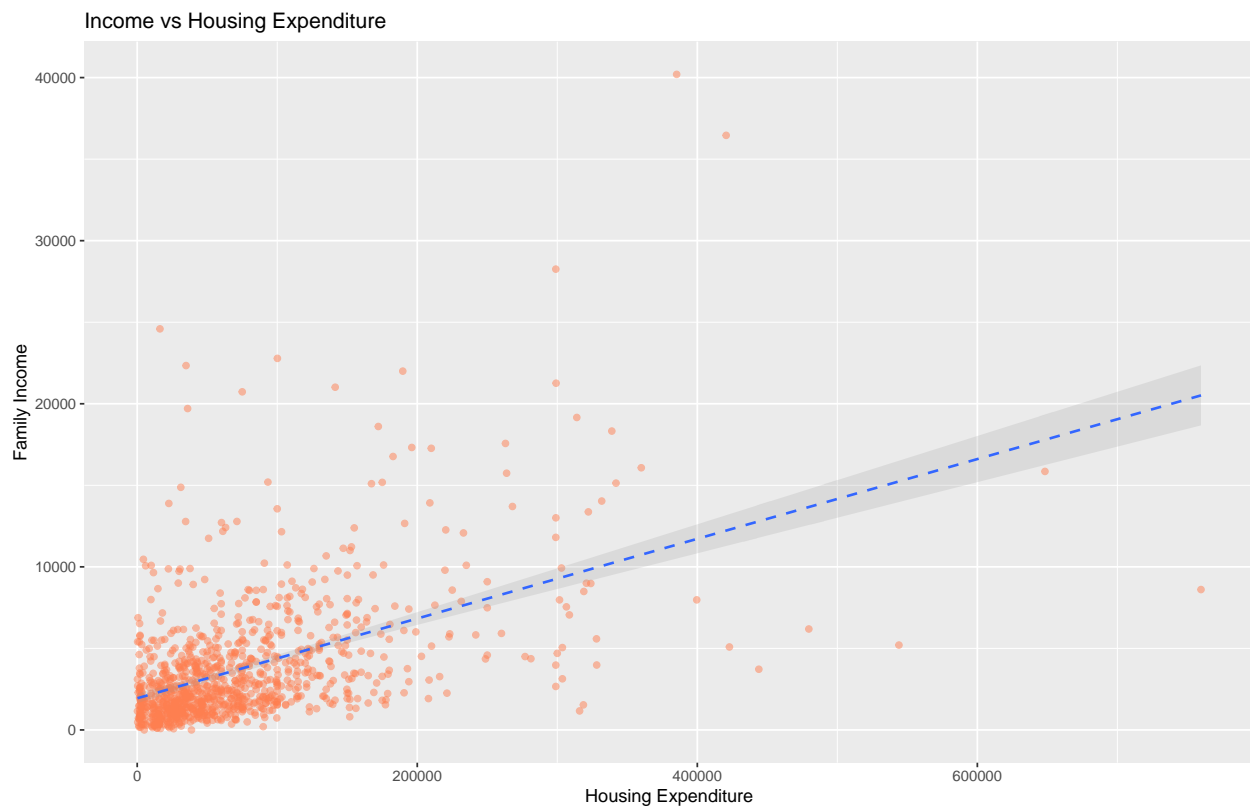
	FAM_Income	ENTERTAINMENT_EXP	ELCTRC_EXP	FOOD_EXP	WATER_EXP	NAT_GAS_EXP	TRANS_EXP	HEAL
ENTERTAINMENT_EXP	0.29							
ELCTRC_EXP	0.13	0.22						
FOOD_EXP	0.40	0.38	0.40					

WATER_EXP	0.24	0.23	0.32	0.35					
NAT_GAS_EXP	0.18	0.19	0.18	0.24	0.27				
TRANS_EXP	0.16	0.16	0.12	0.18	0.10	0.05			
HEALTH_EXP	0.17	0.20	0.19	0.23	0.18	0.18	0.10		
TELEPH_EXP	0.25	0.26	0.31	0.44	0.24	0.24	0.15	0.21	
EHOUS_EXP	0.50	0.41	0.38	0.49	0.36	0.26	0.13	0.23	

p.values:

	FAM_Income	ENTERTAINMENT_EXP	ELCTRC_EXP	FOOD_EXP	WATER_EXP	NAT_GAS_EXP	TRANS_EXP	HEALTH_EXP	TELEPH_EXP	EHOUS_EXP
ENTERTAINMENT_EXP	0.00									
ELCTRC_EXP	0.00	0.00								
FOOD_EXP	0.00	0.00	0.00							
WATER_EXP	0.00	0.00	0.00	0.00						
NAT_GAS_EXP	0.00	0.00	0.00	0.00	0.00					
TRANS_EXP	0.00	0.00	0.00	0.00	0.00	0.07				
HEALTH_EXP	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
TELEPH_EXP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
EHOUS_EXP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

$$\text{Income} = 9.821 * \text{Housing} + 39873.153$$



Linear regression (OLS)

Data : group6FilterIncome

Response variable : FAM\_Income

Explanatory variables: EHOUS\_EXP

Null hyp.: the effect of EHOUS\_EXP on FAM\_Income is zero

Alt. hyp.: the effect of EHOUS\_EXP on FAM\_Income is not zero

	coefficient	std.error	t.value	p.value	
(Intercept)	39873.153	2762.500	14.434	< .001	***
EHOUS_EXP	9.821	0.493	19.923	< .001	***

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

R-squared: 0.252, Adjusted R-squared: 0.251  
F-statistic: 396.932 df(1,1180), p.value < .001  
Nr obs: 1,182