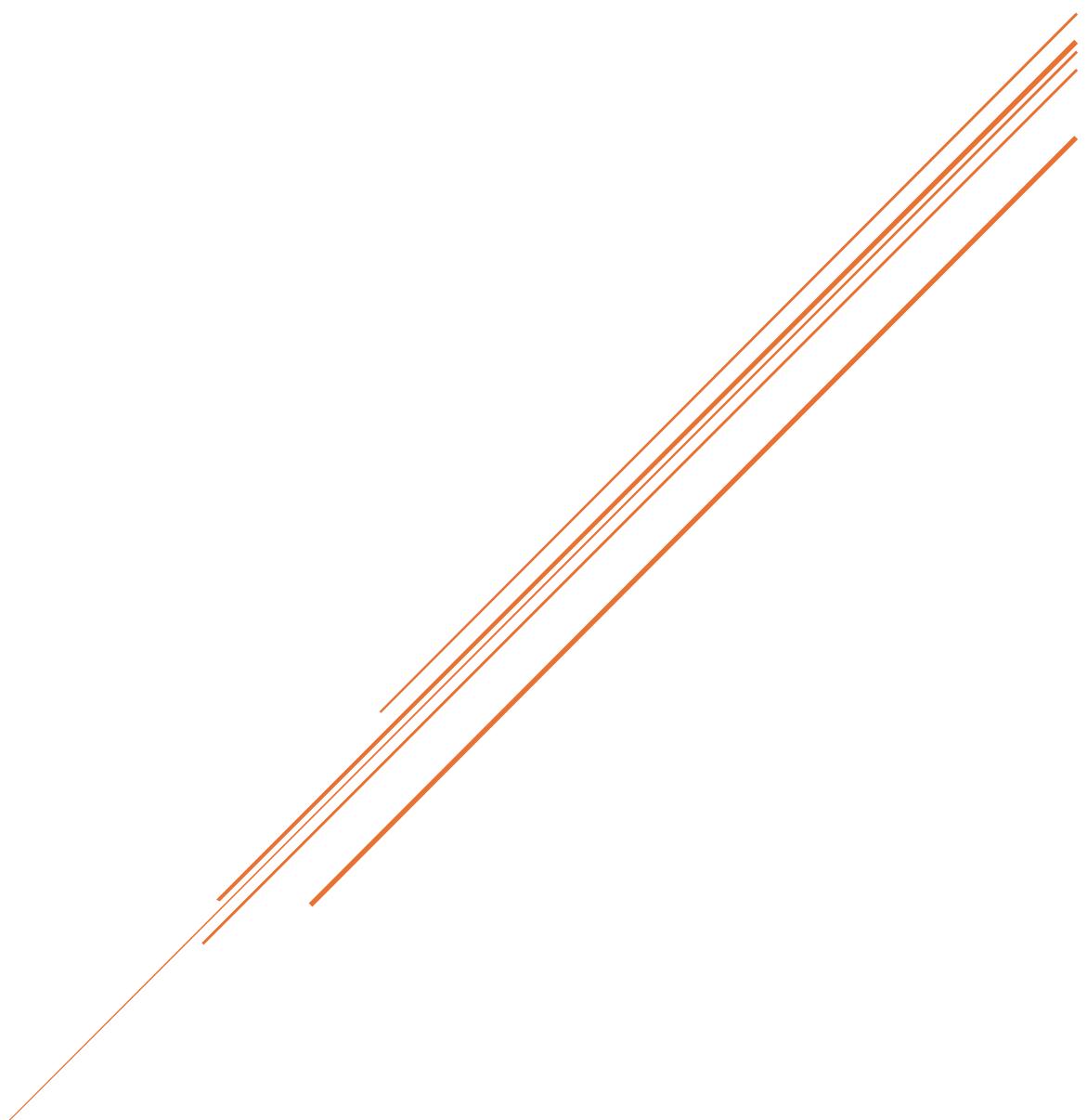


DATA MINING COURSEWORK

COP259



Contents

Part 4. Demographics of US elections	27
Plot the election results, mean income and mean educational attainment levels on the US map.....	27
Comment on the visual comparisons of the maps	32
Test the 2 following hypothesis:.....	32
Part 5. Your own data mining	35
<i>Hypothesis: The impact of weekly working hours on income is moderated by the type of occupation.....</i>	35
Results Analysis.....	41
Appendix.....	42

Part 1. PreprocessingS

Reduce to 5000 data points

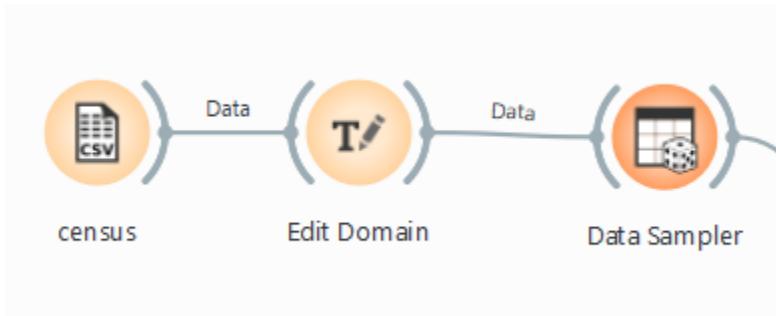
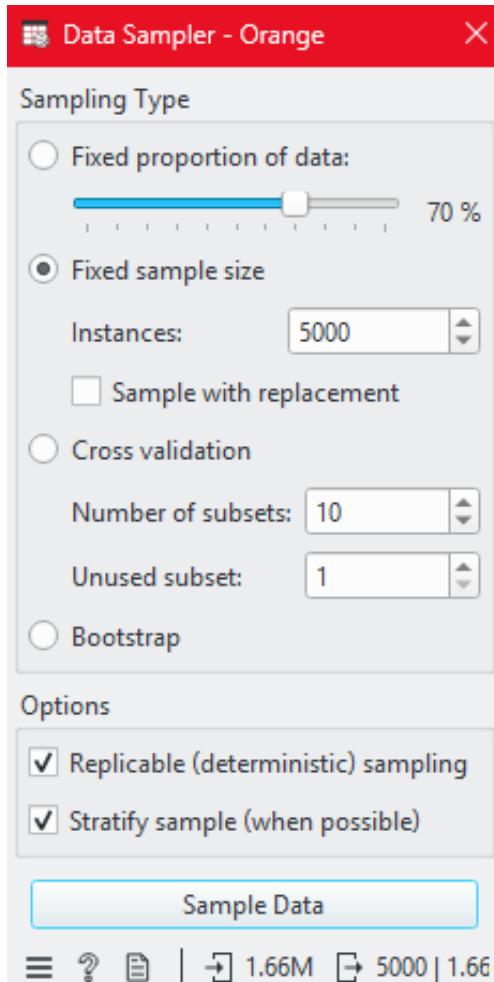


Figure 1 Data Sampler Node



The dataset has been reduced to 5000 points using Data Sampler to speed up computations.

Figure 2 Data Sampler Settings

Modify data structure

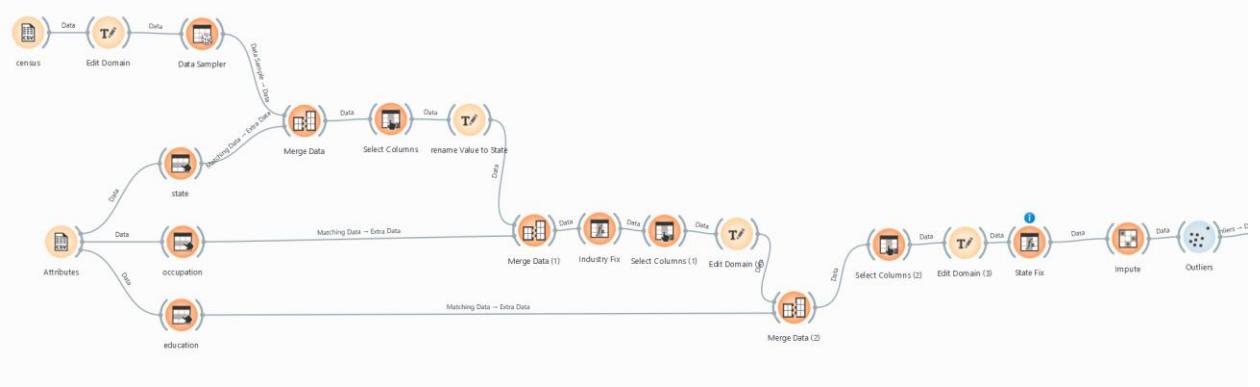


Figure 3 Preprocessing nodes

	income	State	Industry	education_lv1	age	CoW	education	marital	occupation	PoB	hours	sex	race
182	83500	Arkansas	CON	High-School	64	Government E...	16	Married	6720	US	40	Male	white
183	30000	New Hampshire	OFF	Post-High-School	28	Private Employee	19	Married	5140	US	40	Male	white
184	120000	Nevada	PRT	Post-High-School	45	Government E...	19	Married	3740	US	56	Male	non-white
185	44000	Ohio	RPR	High-School	34	Private Employee	17	Married	7100	US	40	Male	white
186	38000	Maryland	CLN	High-School	57	Government E...	16	Married	4220	US	40	Male	white
187	22000	California	OFF	High-School	20	Private Employee	16	Separated	5240	non-US	40	Female	non-white
188	19600	New York	EAT	No diploma	24	Private Employee	14	Married	4030	non-US	18	Female	non-white
189	69400	Nebraska	MGR	Post-High-School	68	Private Employee	19	Married	10	US	40	Female	white
190	35000	Pennsylvania	OFF	Post-High-School	62	Private Employee	19	Married	5120	US	40	Female	white
191	1500	Connecticut	TRN	Post-High-School	21	Private Employee	18	Single	9645	US	8	Female	white
192	95000	Florida	ENG	Bachelor's degree	49	Private Employee	21	Married	1450	US	40	Male	white
193	1400	Tennessee	EAT	No diploma	18	Private Employee	14	Single	4130	US	10	Male	white
194	50000	Nevada	MGR	Associate's degree	40	Self-Employed	20	Married	310	non-US	60	Male	non-white
195	161000	New Jersey	BUS	Professional degree beyond a ...	55	Private Employee	23	Married	650	US	45	Male	white
196	283000	North Carolina	EAT	No diploma	57	Private Employee	14	Married	4020	US	35	Male	non-white
197	15000	Massachusetts	MED	High-School	19	Private Employee	16	Single	3424	US	36	Female	white
198	18000	Illinois	EAT	Post-High-School	37	Private Employee	18	Married	4110	US	20	Female	white
199	30000	Massachusetts	EAT	No diploma	30	Private Employee	11	Married	4020	non-US	40	Male	non-white
200	22000	Louisiana	EAT	Post-High-School	25	Private Employee	19	Single	4110	non-US	40	Female	non-white
201	36000	Michigan	SAL	Associate's degree	59	Self-Employed	20	Married	4820	non-US	50	Male	non-white
202	103000	Missouri	MGR	Bachelor's degree	61	Private Employee	21	Married	10	US	45	Male	white
203	120000	New York	PRT	Bachelor's degree	43	Government E...	21	Married	3740	US	36	Male	white
204	50000	Washington	OFF	Post-High-School	24	Private Employee	19	Single	5000	US	40	Female	non-white
205	59000	Maine	PRD	High-School	38	Private Employee	17	Married	7700	US	48	Male	white
206	30000	Arizona	CON	High-School	59	Private Employee	16	Separated	6410	US	40	Male	white
207	35000	Minnesota	SAL	High-School	19	Private Employee	16	Married	4720	US	35	Female	white
208	150000	Wisconsin	MGR	Master's degree	51	Private Employee	22	Married	140	US	45	Male	white
209	192000	Maine	TRN	High-School	66	Self-Employed	16	Single	9130	US	60	Male	white
210	157000	New York	MGR	Bachelor's degree	65	Private Employee	21	Divorced	440	US	50	Male	white
211	60000	Georgia	TRN	Post-High-School	56	Private Employee	19	Divorced	9130	US	70	Male	non-white
212	17500	Mississippi	TRN	High-School	32	Private Employee	16	Single	9600	US	20	Male	white
213	48000	New Jersey	MED	Professional degree beyond a ...	64	Self-Employed	23	Married	3100	US	20	Male	white
214	6940	North Carolina	CON	Post-High-School	64	Self-Employed	18	Married	6230	US	35	Male	white
215	1500	Pennsylvania	TRN	No diploma	17	Private Employee	15	Single	9645	US	25	Male	white
216	50000	New Hampshire	BUS	Master's degree	30	Government E...	22	Married	710	US	38	Male	white
217	20400	New York	CON	No diploma	61	Private Employee	14	Divorced	6260	US	40	Male	white
218	55000	New York	OFF	Master's degree	31	Private Employee	22	Separated	5740	US	40	Female	non-white
219	32000	Florida	OFF	High-School	62	Private Employee	17	Married	5000	US	40	Female	white
220	1000	Georgia	OFF	Post-High-School	22	Private Employee	19	Married	5400	US	4	Female	white
221	50000	Texas	CMS	Bachelor's degree	35	Government E...	21	Single	2011	US	50	Female	white
222	25000	Maryland	MED	Bachelor's degree	22	Private Employee	21	Single	3500	US	30	Female	white
223	36000	Florida	CLN	High-School	50	Private Employee	16	Separated	4210	US	40	Male	white
224	80260	Indiana	PRD	Bachelor's degree	72	Private Employee	21	Married	8130	US	20	Male	white
225	15000	Maine	CLN	No diploma	33	Self-Employed	15	Married	4230	US	15	Female	white
226	20800	South Carolina	CLN	High-School	50	Government E...	17	Married	4220	US	40	Female	non-white
227	111000	North Carolina	MGR	Master's degree	45	Government E...	22	Divorced	230	US	40	Female	white
228	41000	Wisconsin	PRD	High-School	39	Private Employee	17	Married	8990	US	40	Male	white
229	2400	Texas	EAT	Post-High-School	19	Private Employee	18	Single	4055	US	15	Female	white
230	200000	Michigan	MGR	Bachelor's degree	56	Private Employee	21	Married	10	non-US	45	Male	non-white

Figure 4 Preprocessing final output

The data in the census file has been modified to display the data structure for:

- COW as private employee, government employee, self-employed, no pay. No data points exist for “unemployed”. (categorical)
- Education as no diploma, high-school, post-high-school, etc. while maintaining the numeric value associated with each data point. (categorical + numeric)

- Marital as single, married, widowed, divorced, separated. (categorical)
- Occupation as the first 3 digits of the industry title while maintaining the original numeric value. (text + numeric)
- PoB as US/ non-US. (categorical)
- Sex as male or female. (categorical)
- Race as white or non-white. (categorical)
- State as state's name. (categorical)

Handle missing data or Outliers

There are no missing data in datasets. However, there were a total of 430 datapoints which involved Outliers which were removed using Local Outlier Factor.

Inliers: Census Data: 4570 instances, 13 variables
 Features: 9 (5 categorical, 4 numeric) (no missing values)
 Target: numeric
 Metas: 3 (2 categorical, 1 string)

	income	State	Industry	education_lvl	age	CoW	education	marital	occupation	PoB	hours	sex	
1	40000	Illinois	SAL	Post-High...	28	Private Employee	18	Single	4700	US	50	Female	w
2	45200	Connecticut	CLN	Post-High...	56	Government ...	18	Divorced	4230	US	40	Male	n
3	58010	Illinois	ENG	Bachelor's degree	23	Private Employee	21	Single	1430	US	40	Male	w
4	68000	Illinois	SAL	Bachelor's degree	33	Private Employee	21	Married	4710	US	50	Female	w
5	75000	Connecticut	MED	Doctorate degree	38	Private Employee	24	Married	3250	US	40	Female	w
4													

Outliers: Census Data: 430 instances, 13 variables
 Features: 9 (5 categorical, 4 numeric) (no missing values)
 Target: numeric
 Metas: 3 (2 categorical, 1 string)

	income	State	Industry	education_lvl	age	CoW	education	marital	occupation	PoB	hours	sex	
1	45400	Michigan	TRN	High-School	83	Private Employee	17	Married	9610	US	16	Male	no
2	392000	Oregon	CMM	Doctorate degree	57	Government ...	24	Married	1010	non-US	99	Male	no
3	65500	Florida	MED	Bachelor's degree	62	Private Employee	21	Divorced	3255	US	24	Female	wf
4	3300	Rhode Island	OFF	Post-High...	19	Private Employee	18	Single	5240	US	10	Female	wf
5	11500	New Jersey	CLN	No diploma	73	Private Employee	1	Married	4220	non-US	40	Male	no
4													

Data: Census Data: 5000 instances, 14 variables
 Features: 9 (5 categorical, 4 numeric) (no missing values)
 Target: numeric
 Metas: 4 (3 categorical, 1 string)

	income	State	Industry	education_lvl	Outlier	age	CoW	education	marital	occupation	PoB	hours	
1	40000	Illinois	SAL	Post-High...	No	28	Private Employee	18	Single	4700	US	50	F
2	45200	Connecticut	CLN	Post-High...	No	56	Government ...	18	Divorced	4230	US	40	N
3	58010	Illinois	ENG	Bachelor's degree	No	23	Private Employee	21	Single	1430	US	40	N
4	68000	Illinois	SAL	Bachelor's degree	No	33	Private Employee	21	Married	4710	US	50	F
5	75000	Connecticut	MED	Doctorate degree	No	38	Private Employee	24	Married	3250	US	40	F
4													

Figure 5 Outliers node results

Exploratory Data Analysis

Feature Statistics								
	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.
N	income		56519.8	50000	39300	1.26636	150	830000
N	occupation		4146.93	2310	4150	0.639558	10	9830
N	age		42.898	52	43	0.350743	17	90
N	hours		38.3967	40	40	0.317445	1	96
N	education		18.5987	16	19	0.169695	1	24
C	State		California		3.5	0 (0 %)		
C	education_lv		High-School		1.82	0 (0 %)		
C	marital		Married		1.09	0 (0 %)		
C	CoW		Private Employee		0.77	0 (0 %)		
C	sex		Male		0.692	0 (0 %)		
C	race		white		0.526	0 (0 %)		
C	PoB		US		0.422	0 (0 %)		

Table 1 Feature Statistics results

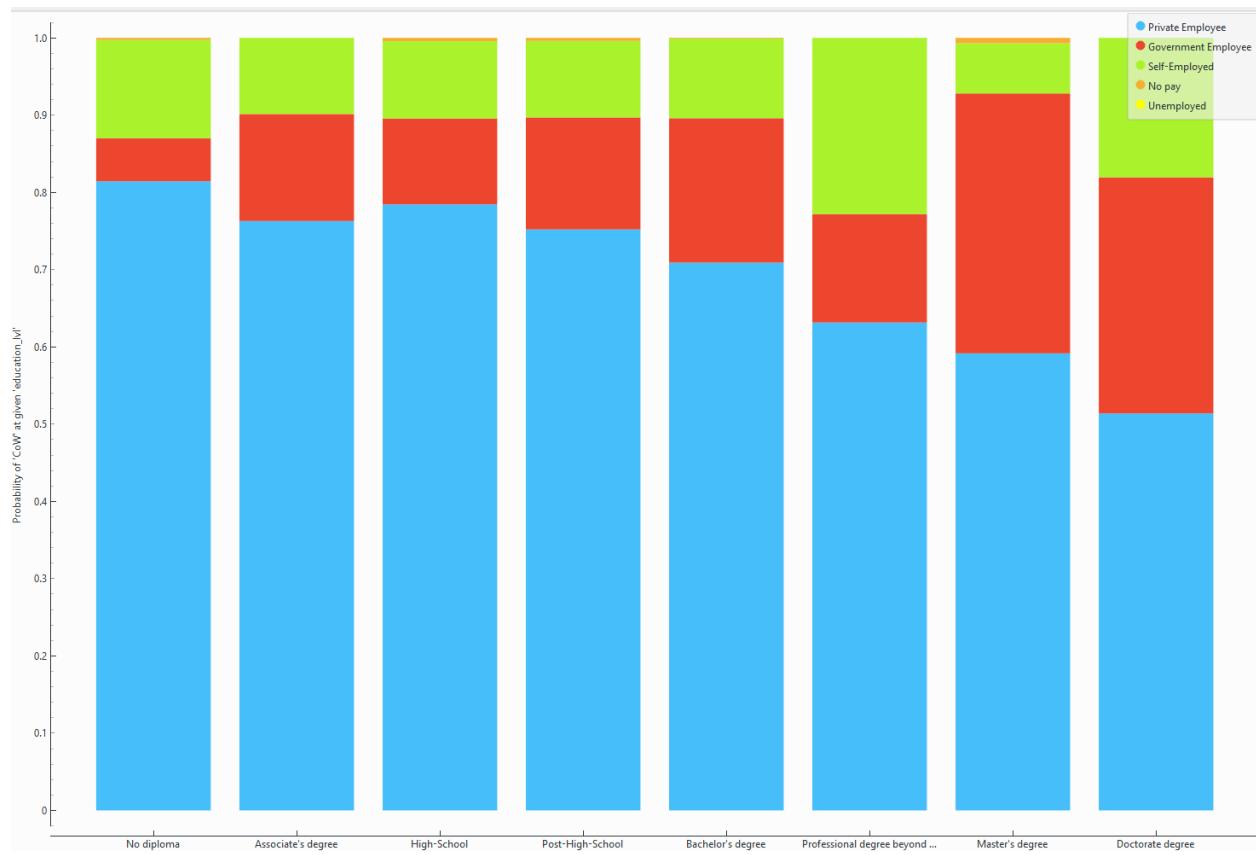


Table 2 CoW by education level

	#	Univar. reg.	RR...ff	
1	N occupation	NA	0.100	
2	N age	NA	0.089	
3	N hours	NA	0.079	
4	N education	NA	0.077	
5	C marital	5	NA	0.058
6	C CoW	5	NA	0.016
7	C PoB	2	NA	0.012
8	C race	2	NA	0.008
9	C sex	2	NA	0.007

Figure 6 Ranking variables importance

The distribution of the numerical variables of the data set along with some statistically relevant features are visualized above.

Part 2. Fairness in income distribution

Income Histogram

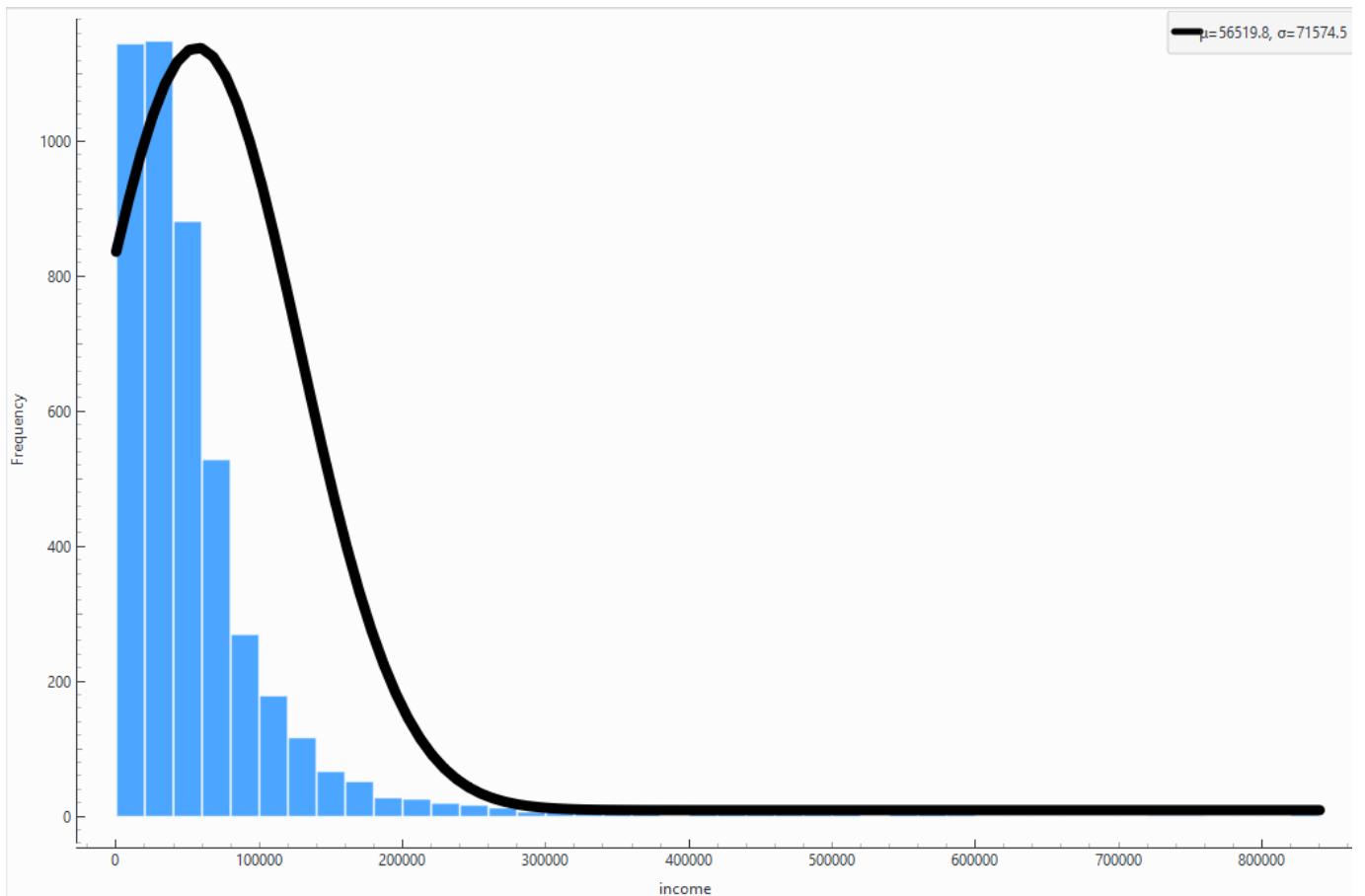


Table 3 Income Histogram

Log(Income) Histogram

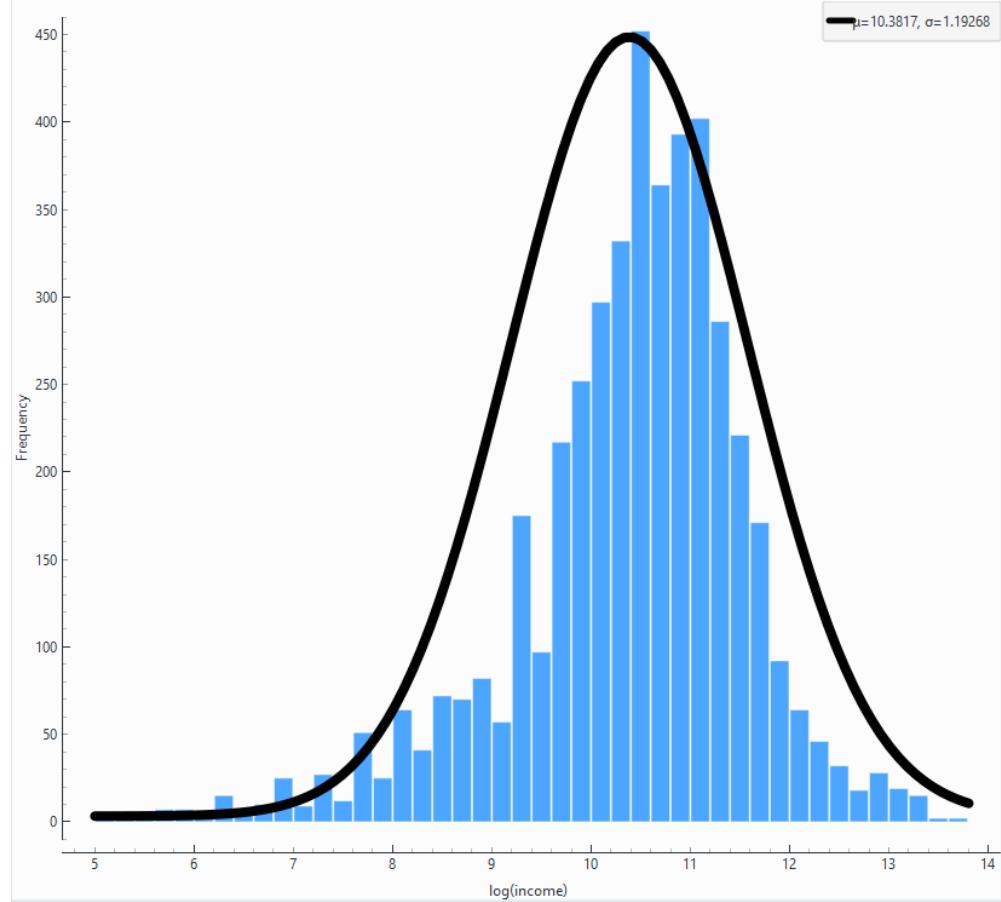


Table 4 Log(Income) Histogram

Zipf plot for income distribution

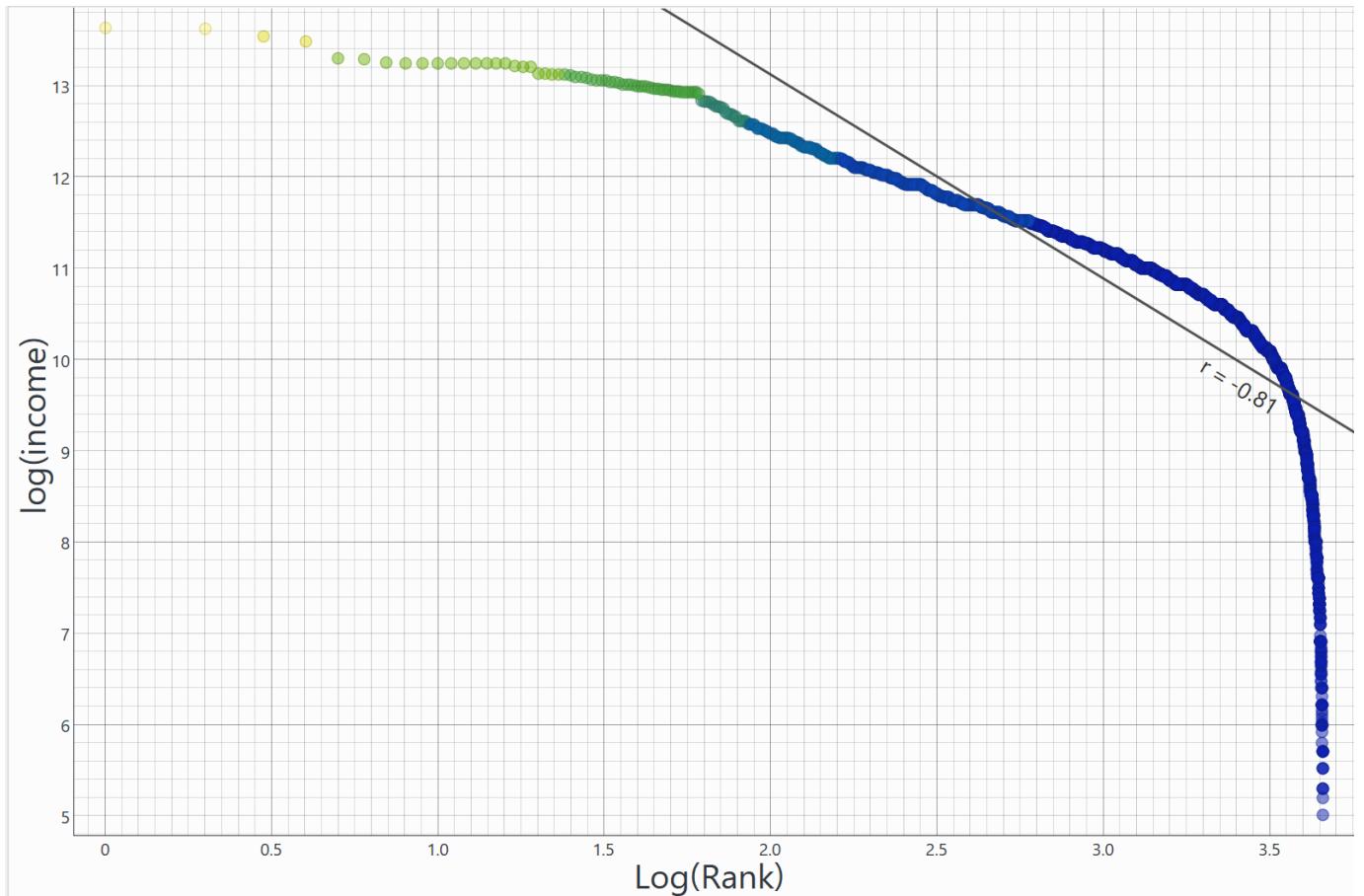


Table 5 Zipf Plot

Results interpretation

The log transformation has increased the correlation coefficients (data is more spread). This suggests a more linear relationship between the variables. The $\log(\text{income})$ graph is more useful as it helps identify patterns which may be less apparent in the regular income/frequency histogram. The Zipf plot graph checks if the income has followed a Pareto distribution and can be used to understand inequality in income (how top earners compare vs the rest). The results:

- A) Log(income) plot: data appears to be normally distributed with a mean $\log(\text{income})$ of 10.3817 and SD of 1.93.
- B) Zipf plot: the data in the upper tail appears to follow a power-law distribution (Pareto's principle) due to its linearity, indicating that a very small number of individuals hold significantly higher incomes.

Visualization of income distribution by sex, race and PoB

The three attributes vs income are visualized using box plot graphs.

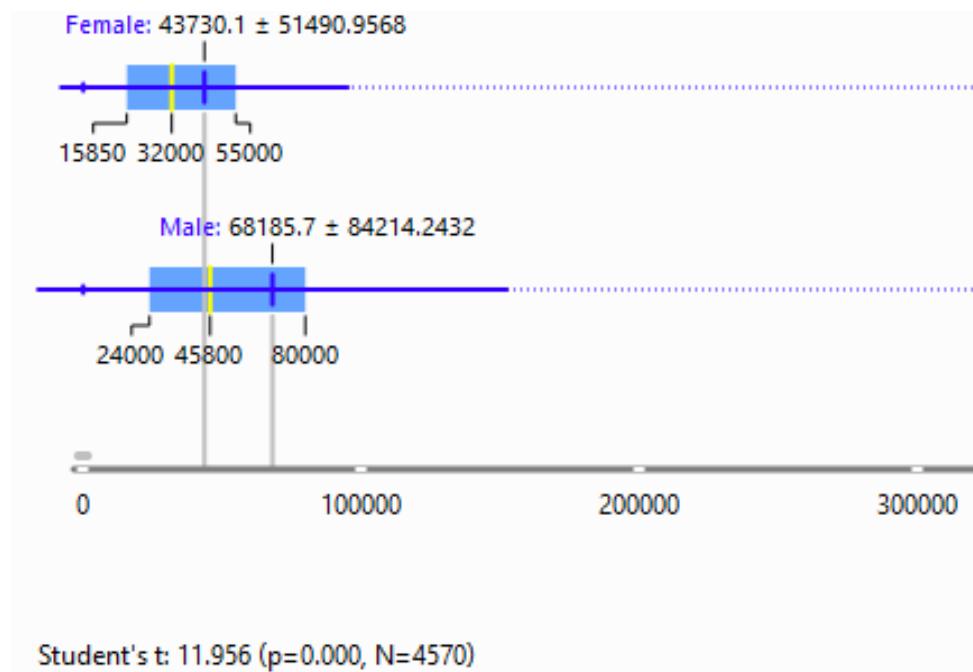


Table 6 Box Plot of income vs sex

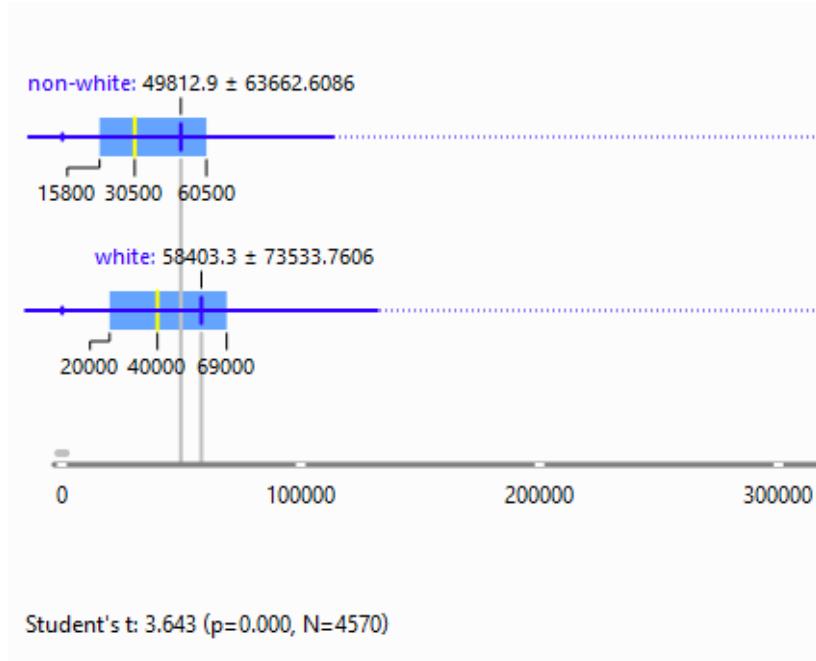


Table 7 Box Plot of income vs race

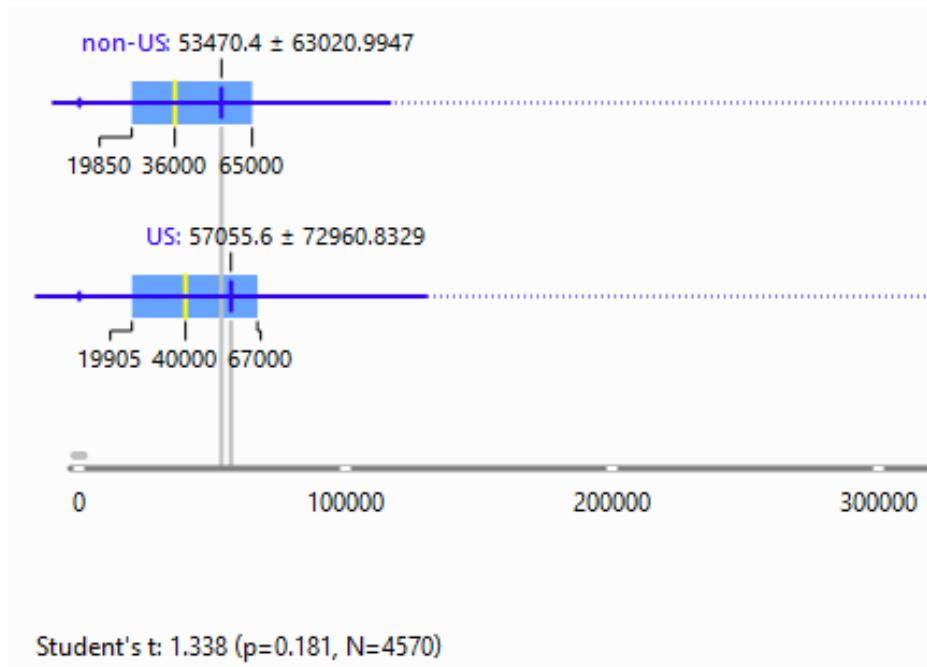


Table 8 Box Plot of income vs PoB

Compute t-test and comment on findings

The results are summarized in the table below:

Category	Mean Income	Standard Deviation	Median	Interquartile Range (IQR)	t-test values	p-value	Sample Size (N)
Gender							
Female	43,730.1	51,490.9568	32,000	15,850 - 55,000	11.956	0.000	4,570
Male	68,185.7	84,214.2432	45,800	24,000 - 80,000			
Race							
Non-White	49,812.9	63,662.6086	30,500	15,800 - 60,500	3.643	0.000	4,570
White	58,403.3	73,533.7606	44,000	20,000 - 69,000			
Nationality							
Non-US	53,470.4	63,020.9947	36,000	19,850 - 65,000	1.338	0.181	4,570
US	57,055.6	72,960.8329	40,000	19,905 - 67,000			

Table 9 Tabular t-test results

The t-test results show statistically significant differences between gender and race. The p-value of 0.0 indicates that the difference is highly unlikely to be caused by random chance for those two variables. The t-test result and p-value of 0.181 indicate that PoB is not a statistically significant variable.

Scatter plots of income vs age, weekly hours and education level

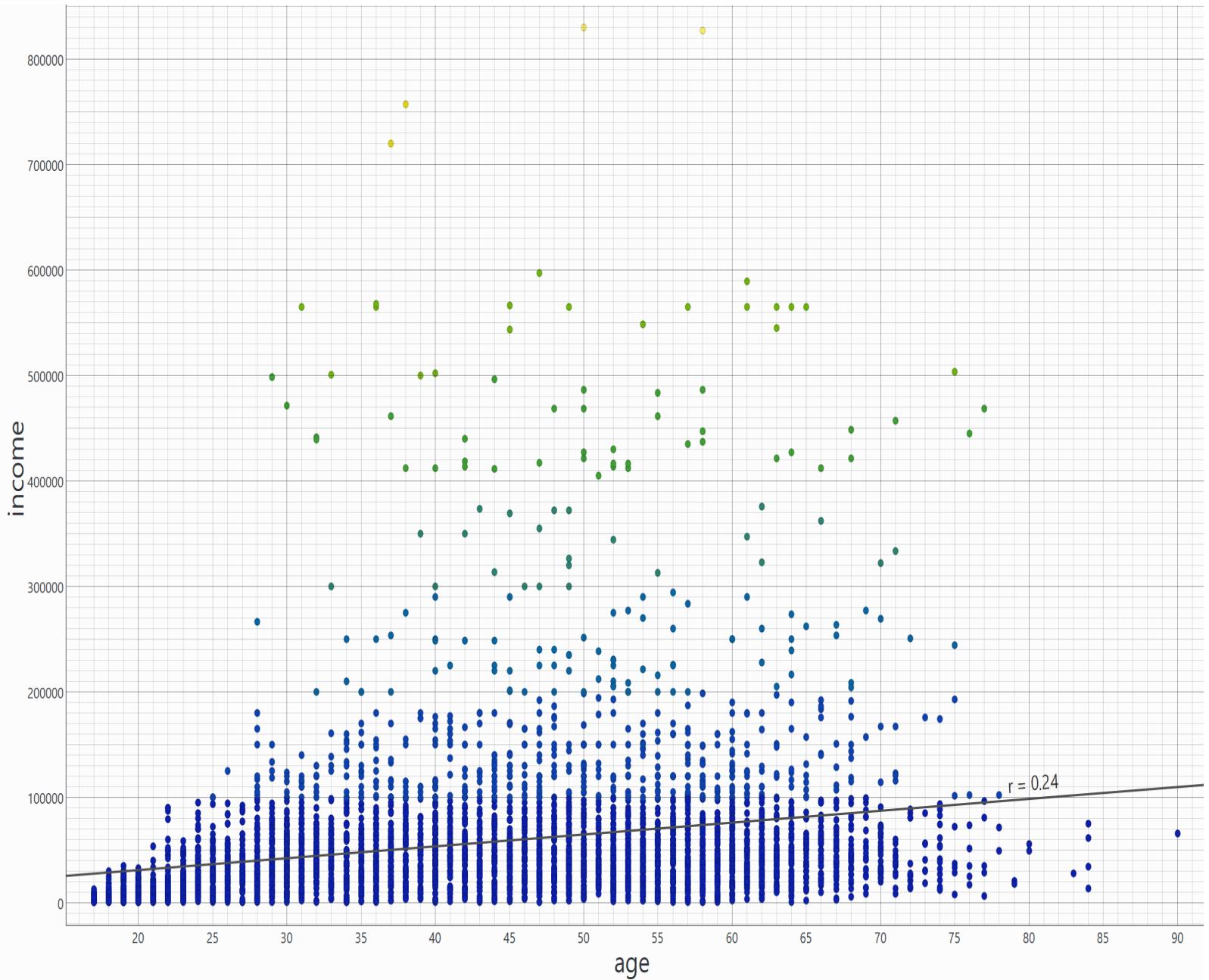


Figure 7 Income vs age

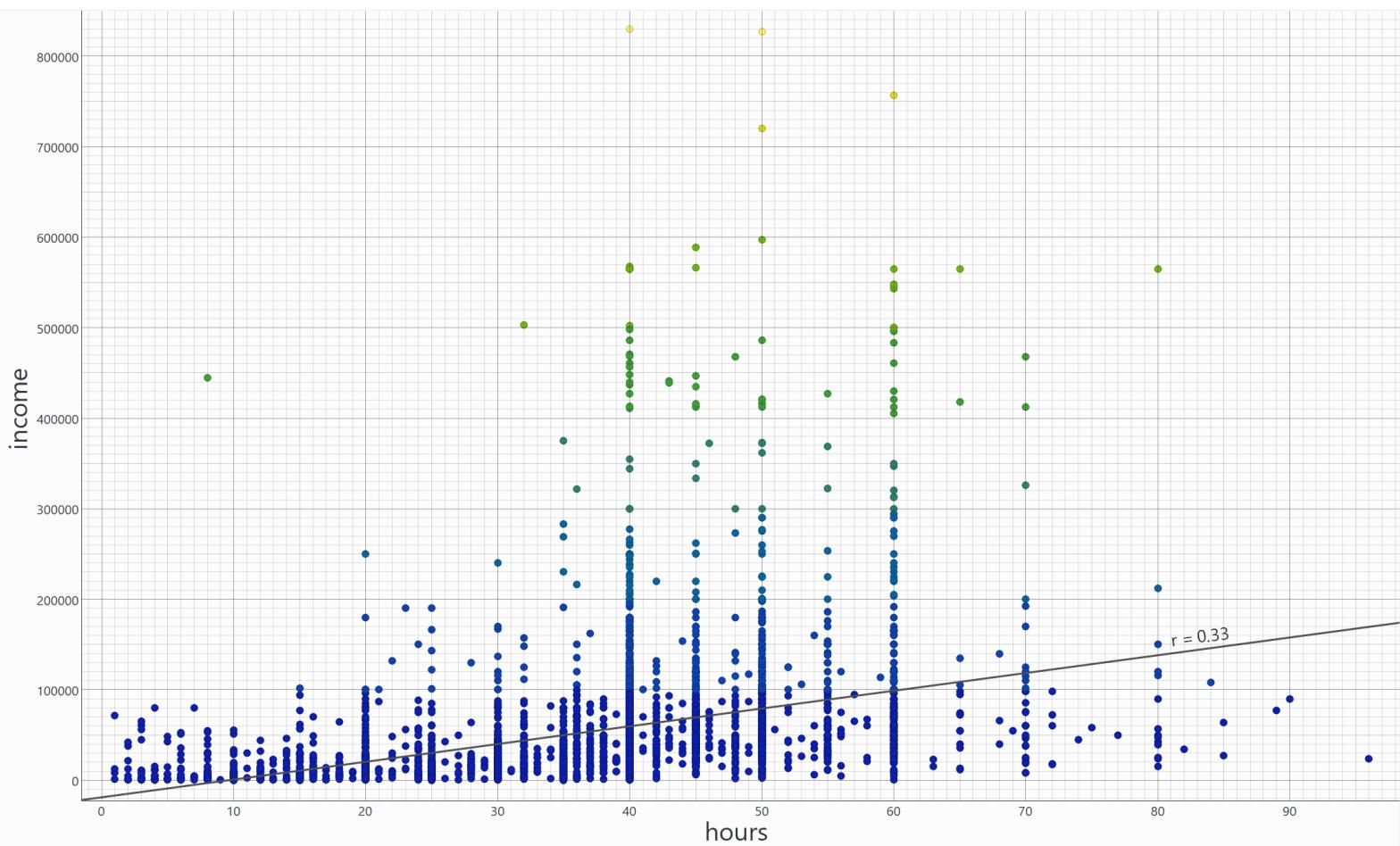


Figure 8 Income vs hours

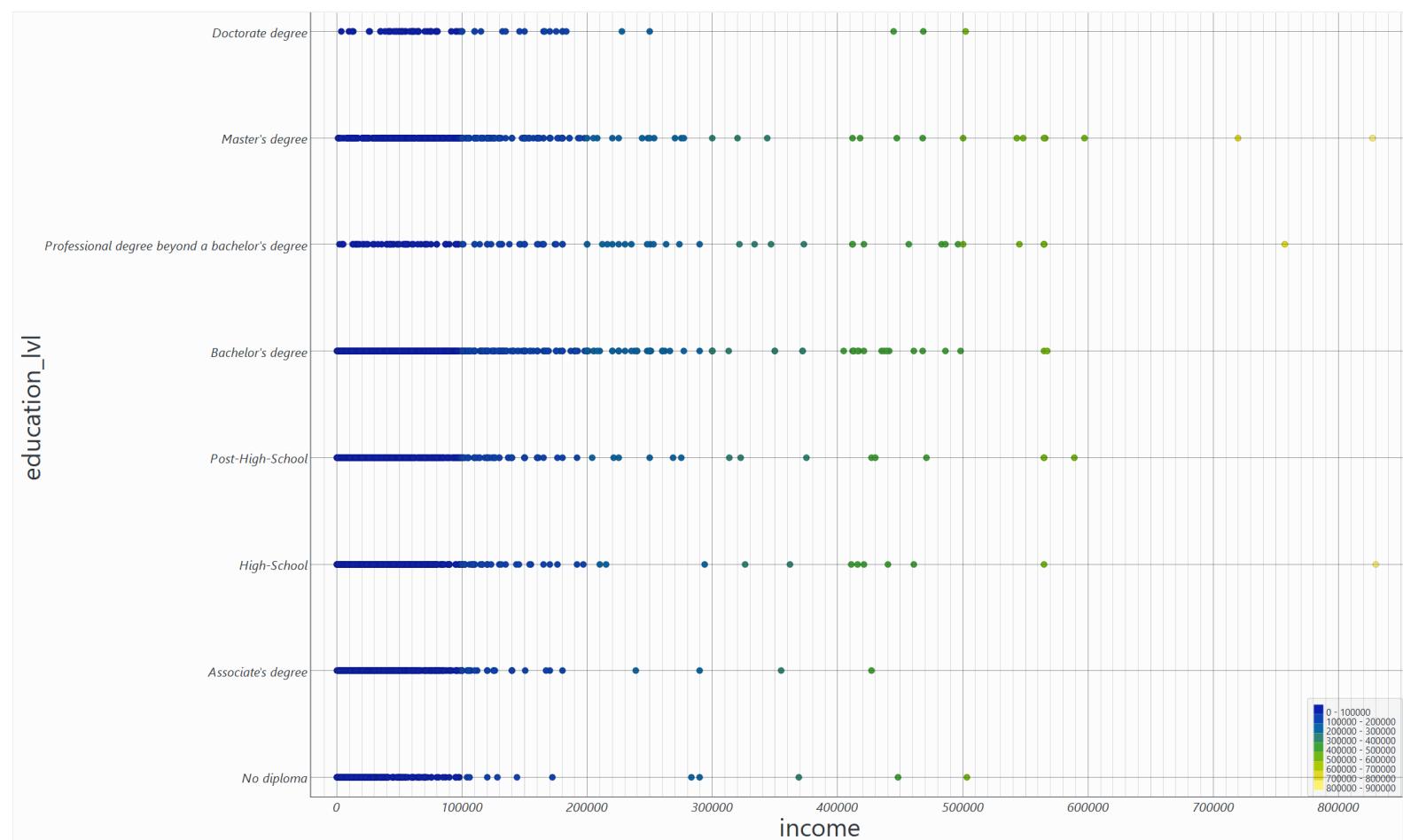


Figure 9 Income vs Edu_lvl

Pearson's Correlation for income

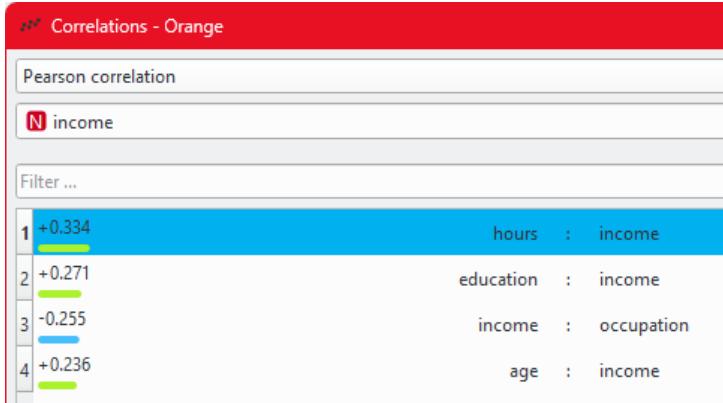


Figure 10 Pearson's Correlation for income

Log(income) scatter plots

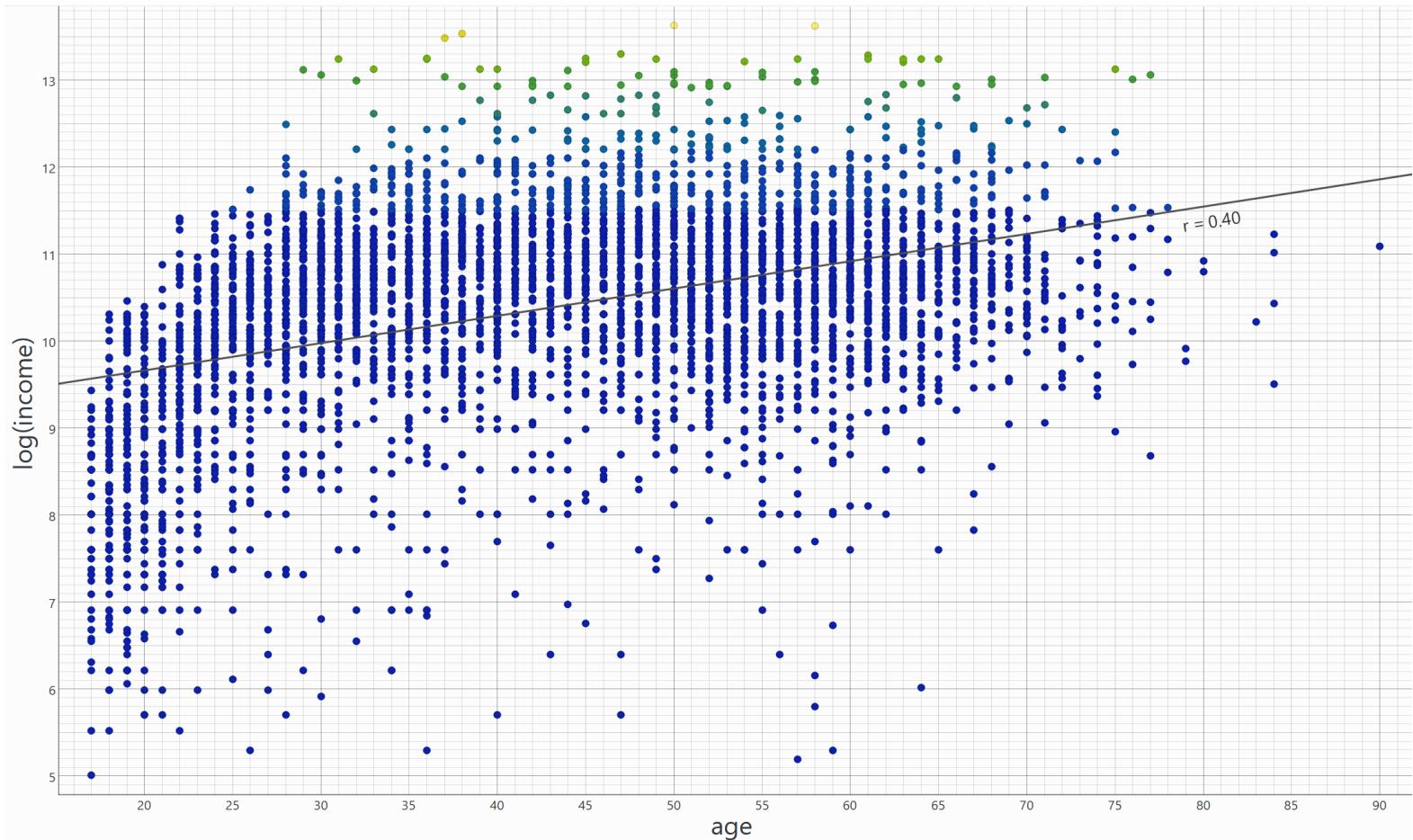


Figure 11 log(income) vs age

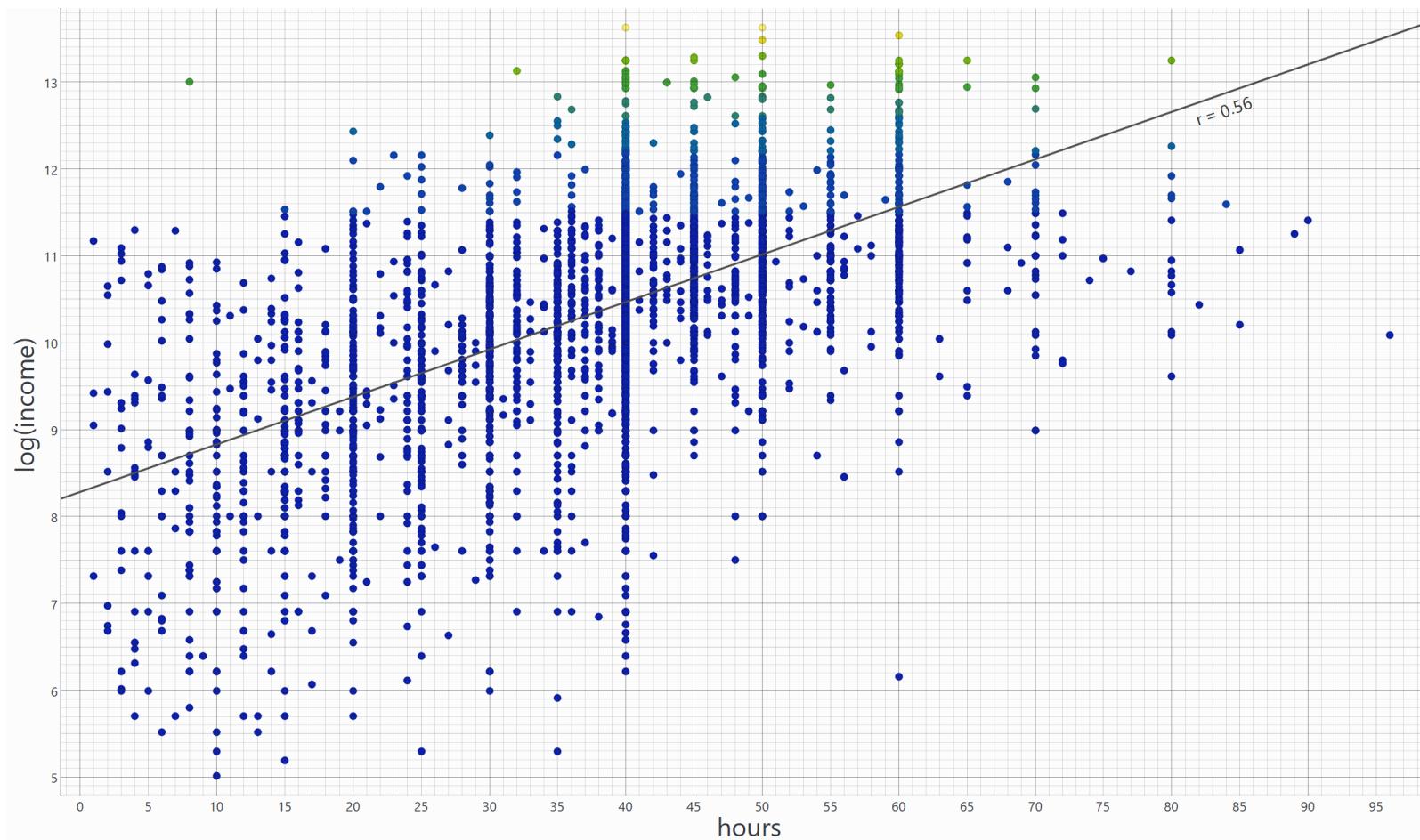


Figure 12 $\log(\text{income})$ vs hours

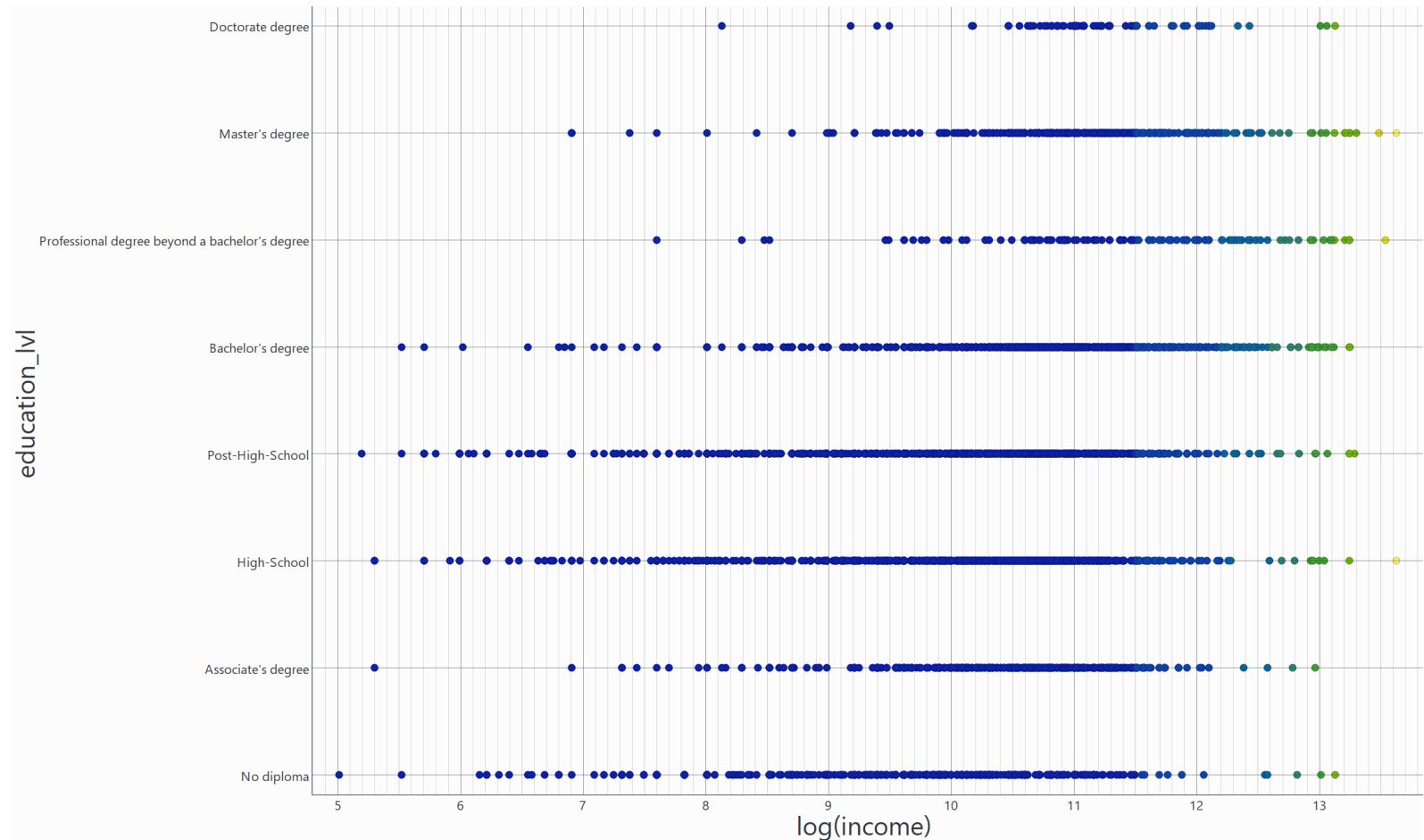


Figure 13 $\log(\text{income})$ vs edu_lvl

Pearson's Correlation for log(income)

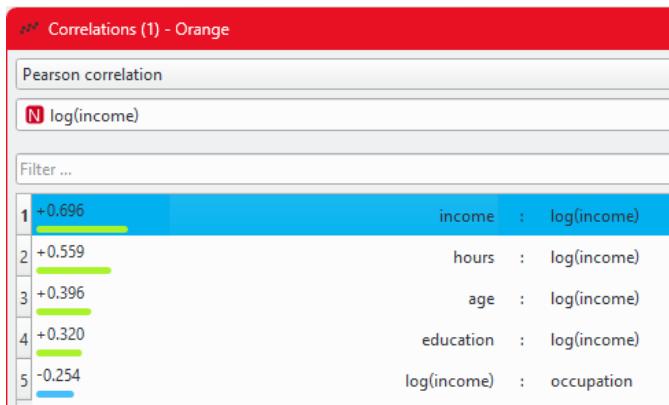


Figure 14 Pearson's Correlation for log(income)

Comment on findings

The usage of log transformations (log(income)) has reduced the influence of outliers and the impact of data variance. This makes the correlations of the variables more reliable and meaningful. The log(income)'s variable correlation was higher than those of income for the aforementioned reasons and the more meaningful variables are now in the order hours, age, education, occupation.

Part 3. Predicting Income

Mean income vs education level

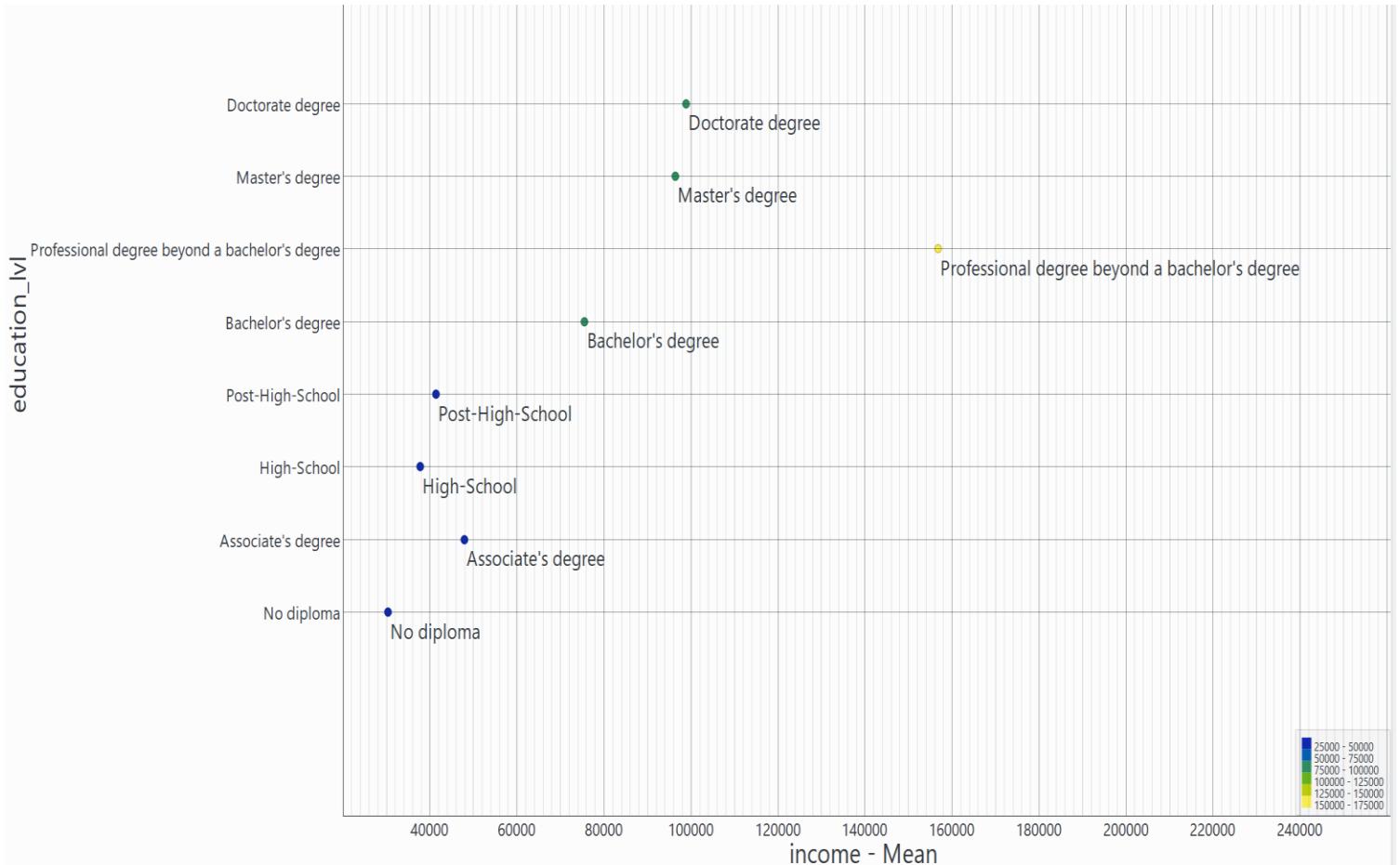


Figure 15 mean income vs edu_lvl

Estimate a monetary value for education

The mean income corresponding to the total years of education is visualized.

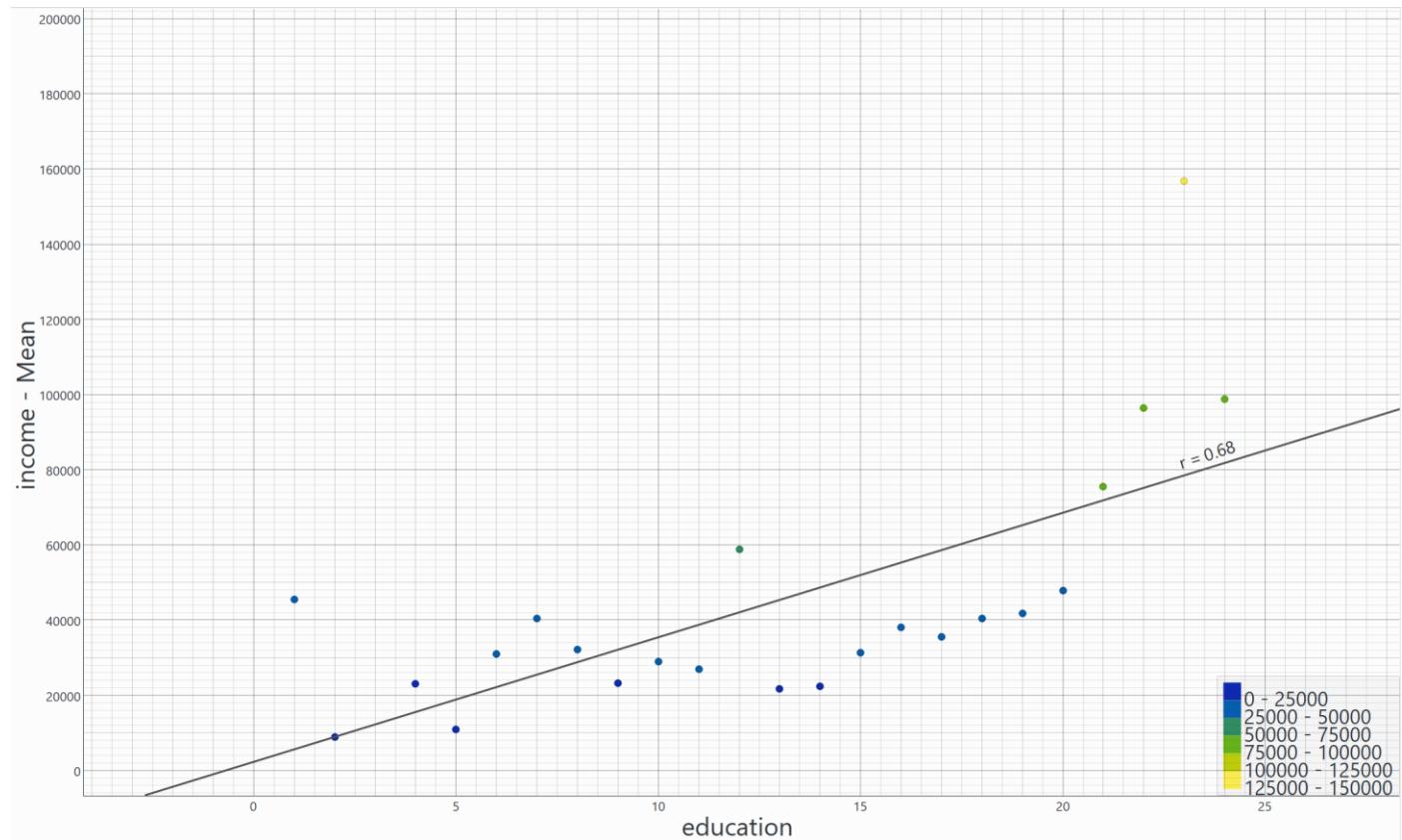


Figure 16 mean income vs years of education

Linear regression is used to calculate the monetary value of each year of education as well as the y-intercept.

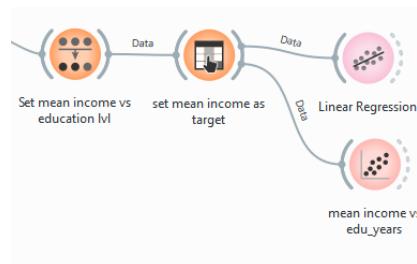


Figure 17 Linear Regression results

Coefficients: coefficients: 2 instances, 2 variables
Features: numeric (no missing values)
Metas: string

name	coef
1 intercept	2289.87
2 education	3312.65

Figure 18 linear regression Orange nodes

Analysis Limitation

This analysis does not factor in other important variables such as age, hours, industry type, CoW, experience, skills or the state. This omission can greatly imbalance the data.

The intercept assumes an annual income of \$2289.87 at 0 years of education.

Furthermore, the results of low years of education might not be useful or realistic.

By looking at education level vs mean-income (figure15), higher levels of education do not necessarily translate to a significant increase in mean-income, which suggests non-linearity.

Split population into high and low income

By connecting a formula node to the data set, a new categorical variable (income_class) is created to split the population into different income groups.

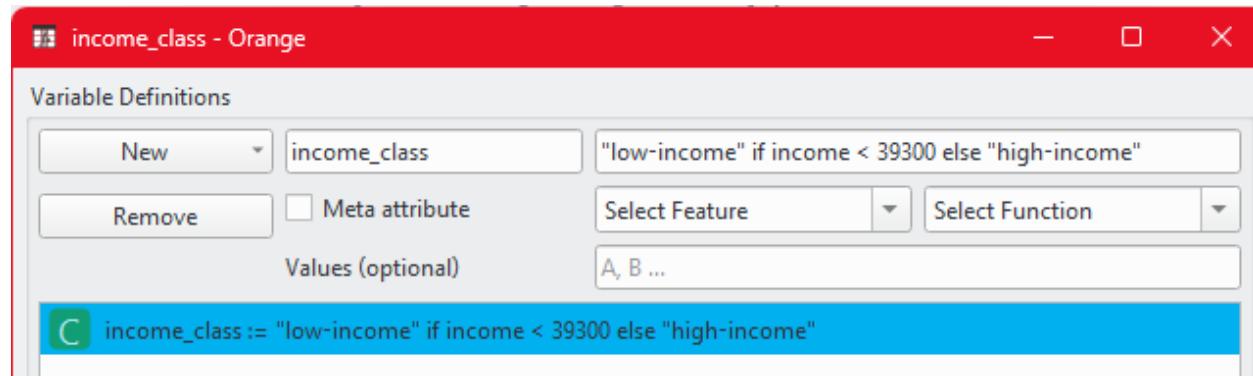


Figure 19 income_class node settings

Use different classifier models to predict which subgroup a person belongs to

After setting “income_class” as target and ignoring the variable “income”, 5 models were attached to a “test and score” node and the results were analyzed.

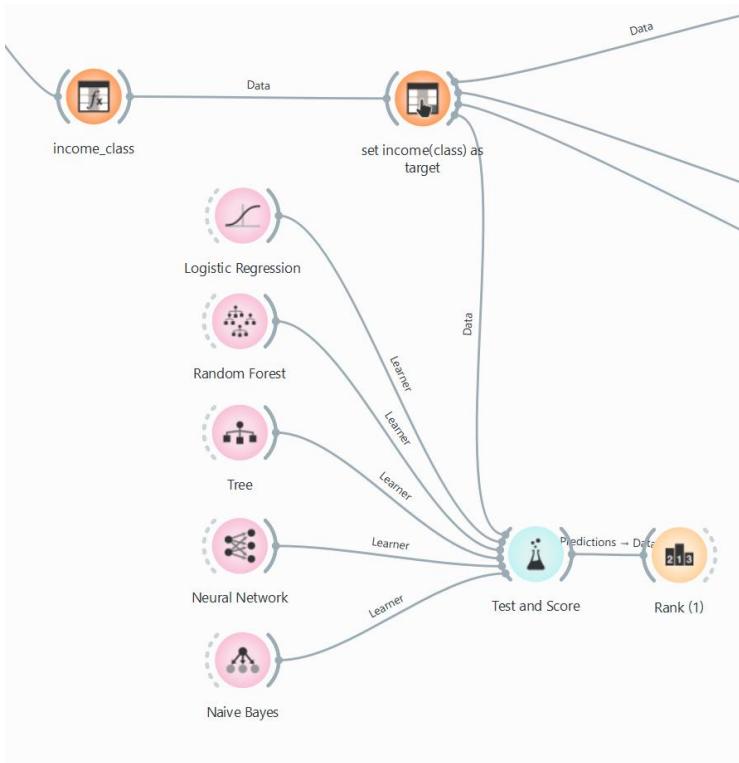


Figure 20 Models used in analysis

The results are displayed below:

Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.851	0.767	0.766	0.767	0.767	0.533
Random Forest	0.853	0.771	0.771	0.772	0.771	0.543
Tree	0.722	0.737	0.737	0.738	0.737	0.476
Neural Network	0.857	0.768	0.768	0.768	0.768	0.536
Naive Bayes	0.852	0.762	0.762	0.762	0.762	0.525

Compare models by:		Area under ROC curve	<input type="checkbox"/> Negligible diff.:	0.1
Logistic ...	Random ...	Tree	Neural N...	Naive Ba...
Logistic Regression		0.396	0.999	0.054
Random Forest	0.604		1.000	0.283
Tree	0.001	0.000		0.001
Neural Network	0.946	0.717	0.999	
Naive Bayes	0.601	0.420	1.000	0.187

Figure 21 Model results

Out all the models, Neural Network performed the best based on its high AUC score, F1 score and MCC. A side-by-side comparison shows its performance against the other models.

Collect feature ranking

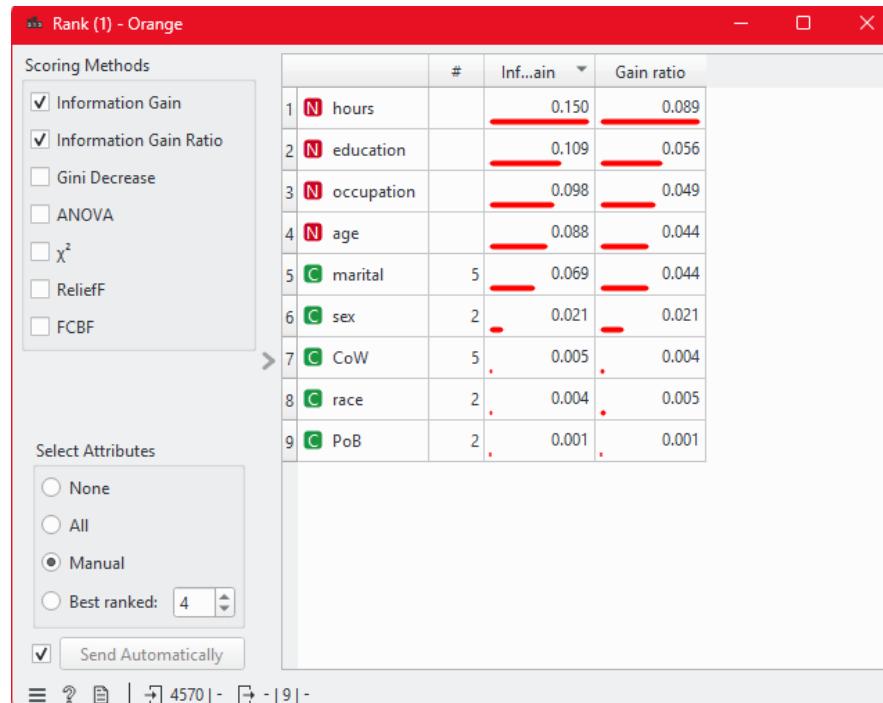


Figure 22 Feature ranking of data

The best predictors of income_class in order of importance appear to be

- Hours
- Education
- Occupation
- Age

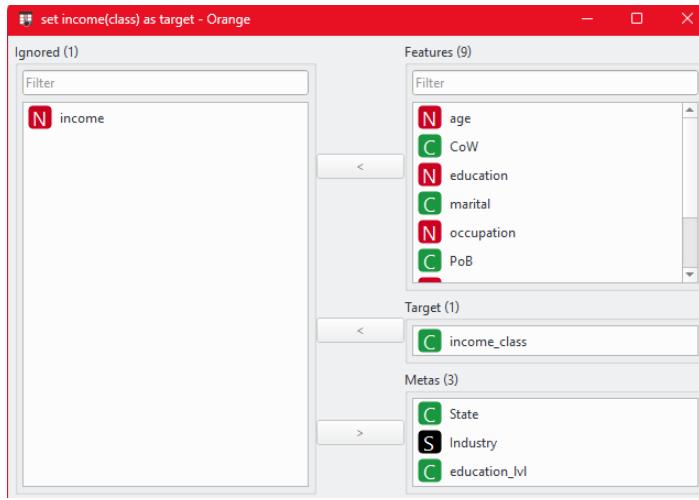


Figure 23 setting income_class as target

Income is ignored to avoid contamination of the results.

Train selected model on data

A “data sampler” node is used to split the data into training (67%) and testing data (33%).

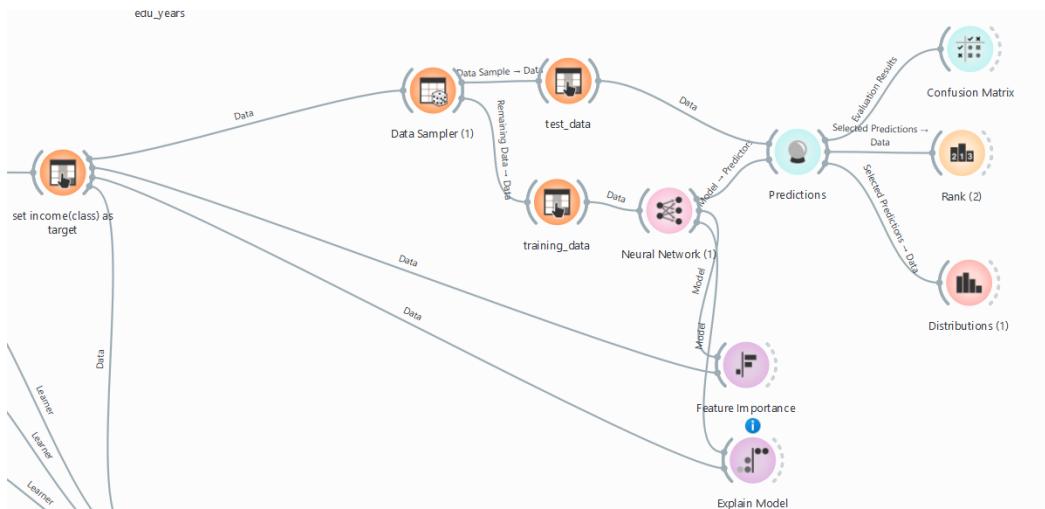


Figure 24 Node setup for data training

The results of the model and explainability techniques are displayed below.

		Predicted		
		high-income	low-income	Σ
Actual	high-income	79.8 %	20.2 %	732
	low-income	24.9 %	75.1 %	731
Σ		766	697	1463

Figure 25 Confusion Matrix results

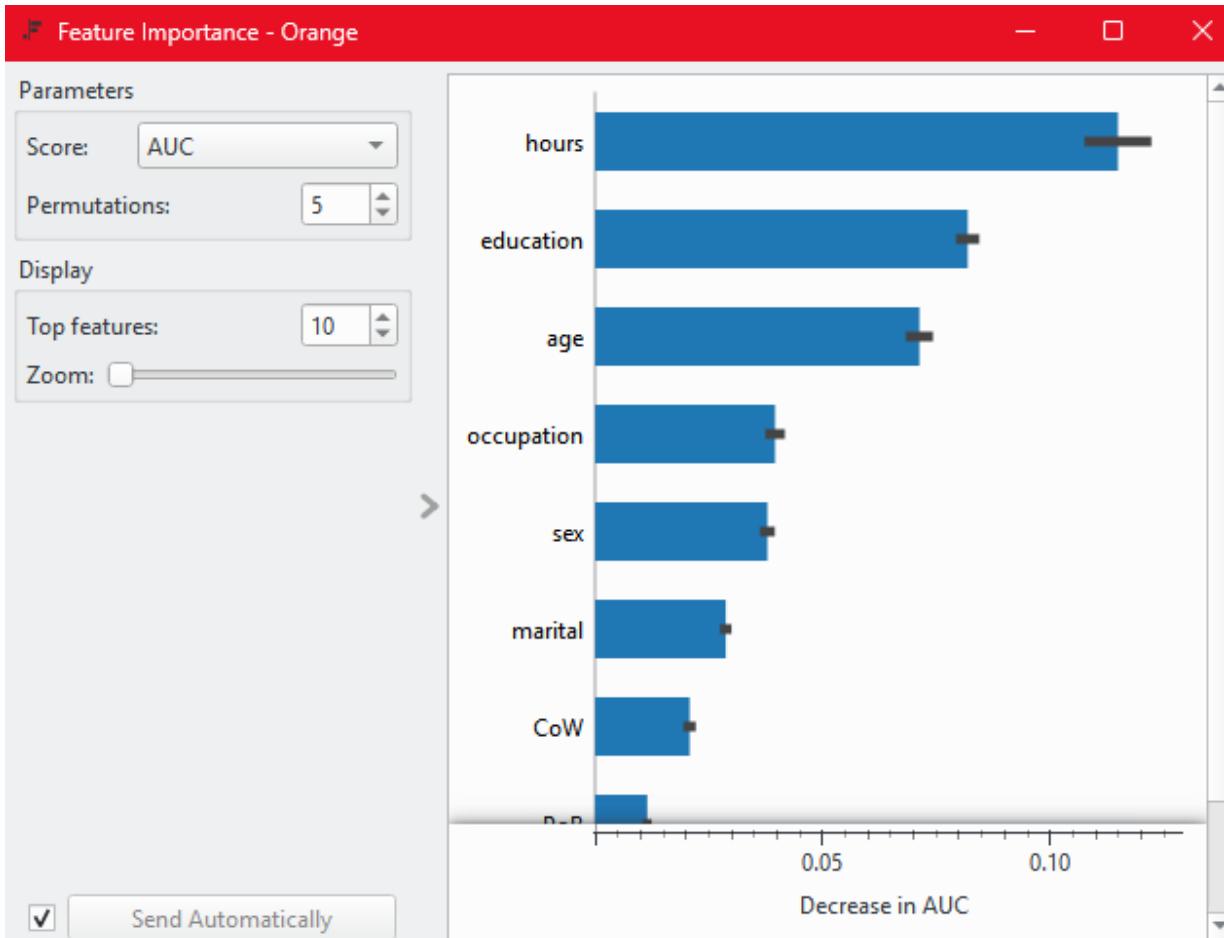


Figure 26 Feature Importance results

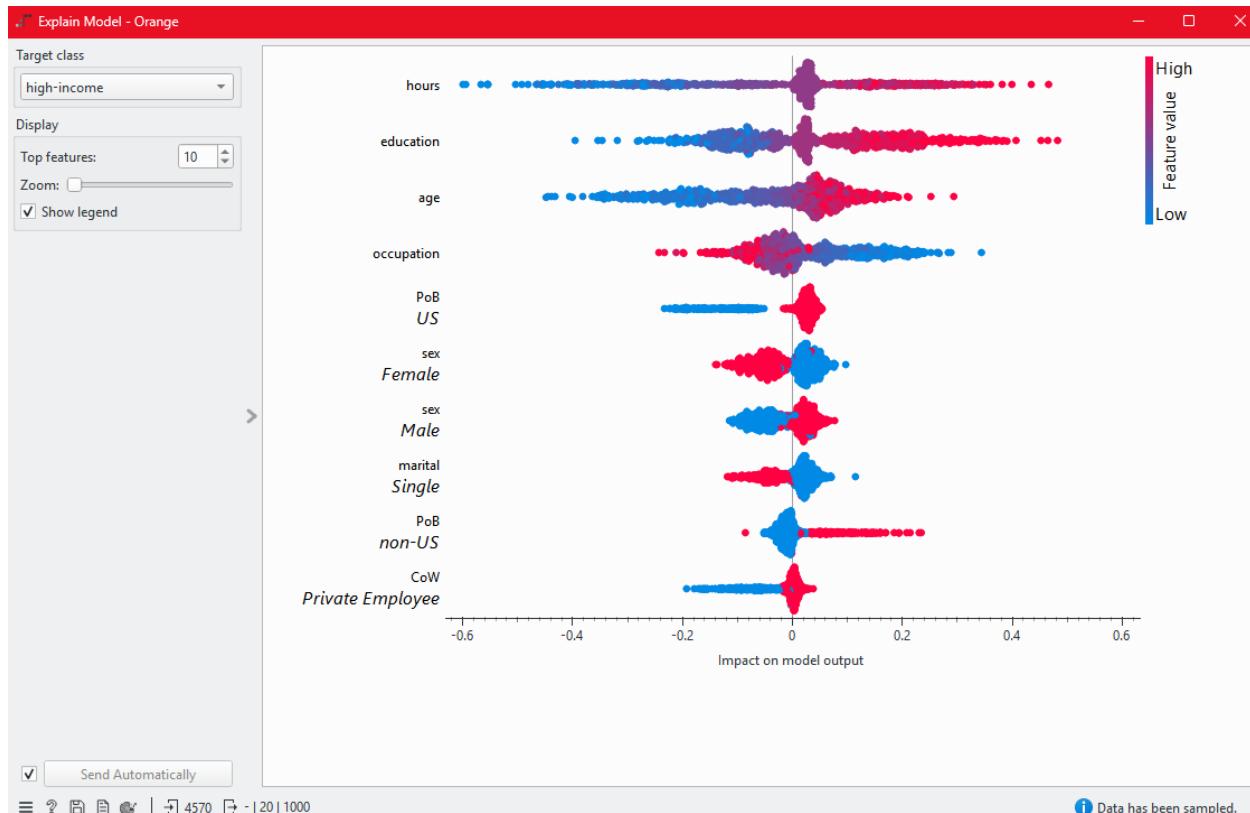


Figure 27 Explain Model results

Compare results and comment on findings

Both analyses highlight the same variables as having a strong impact on income levels. The explained model results shows a more nuanced impact of each variable while the feature importance shows a more general impact. The results mostly agree with each other; however the importance of variables slightly changes.

Part 4. Demographics of US elections

Plot the election results, mean income and mean educational attainment levels on the US map.

Data from the Census file is grouped by State, and Puerto Rico before merging it to Voting.csv. The data in Voting is cleaned to display one row for each State, showing who won only the winning candidate. The variables mean income, mean years in education and income level for each state are added to the dataset.

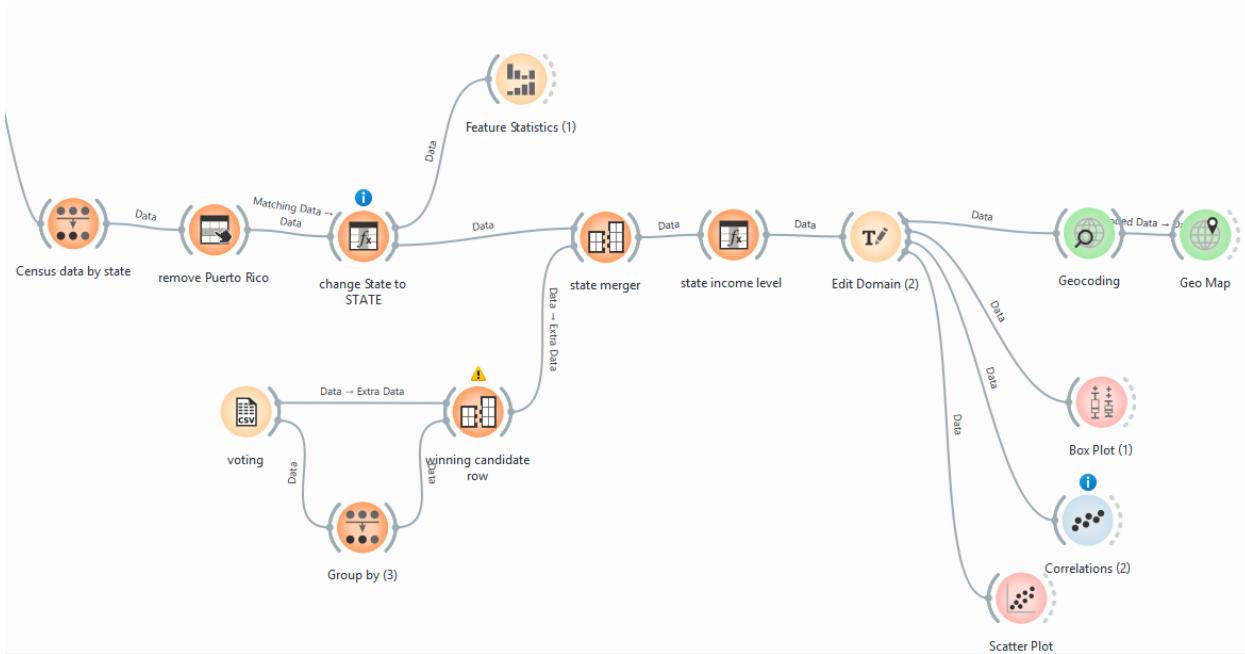


Figure 28 Part 4 Orange node setup

- Election Results (Red for Republican/ Blue for Democrat)



Figure 29 Election results mapped

- Income level per State

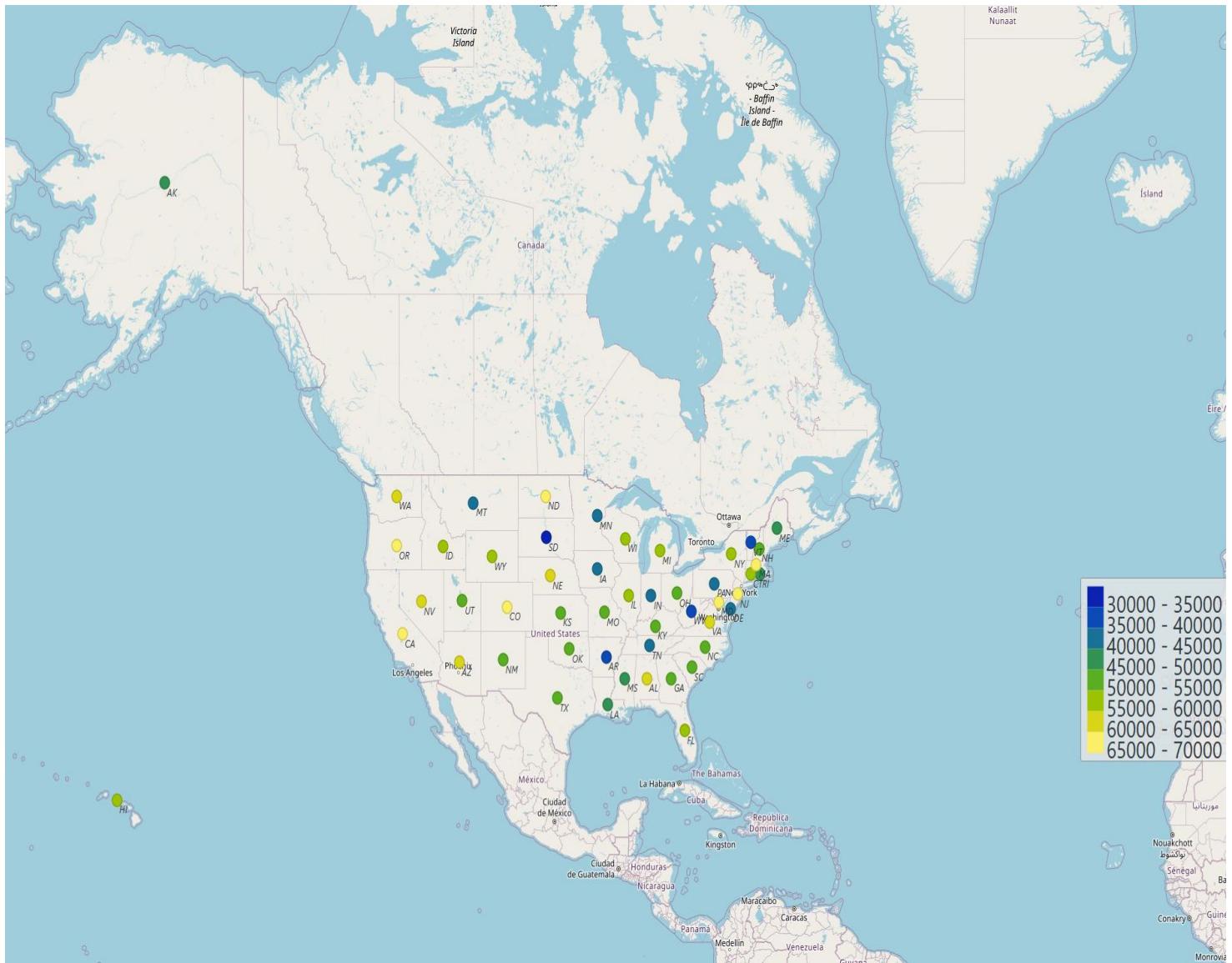


Figure 30 Income level mapped

- Years of education by State

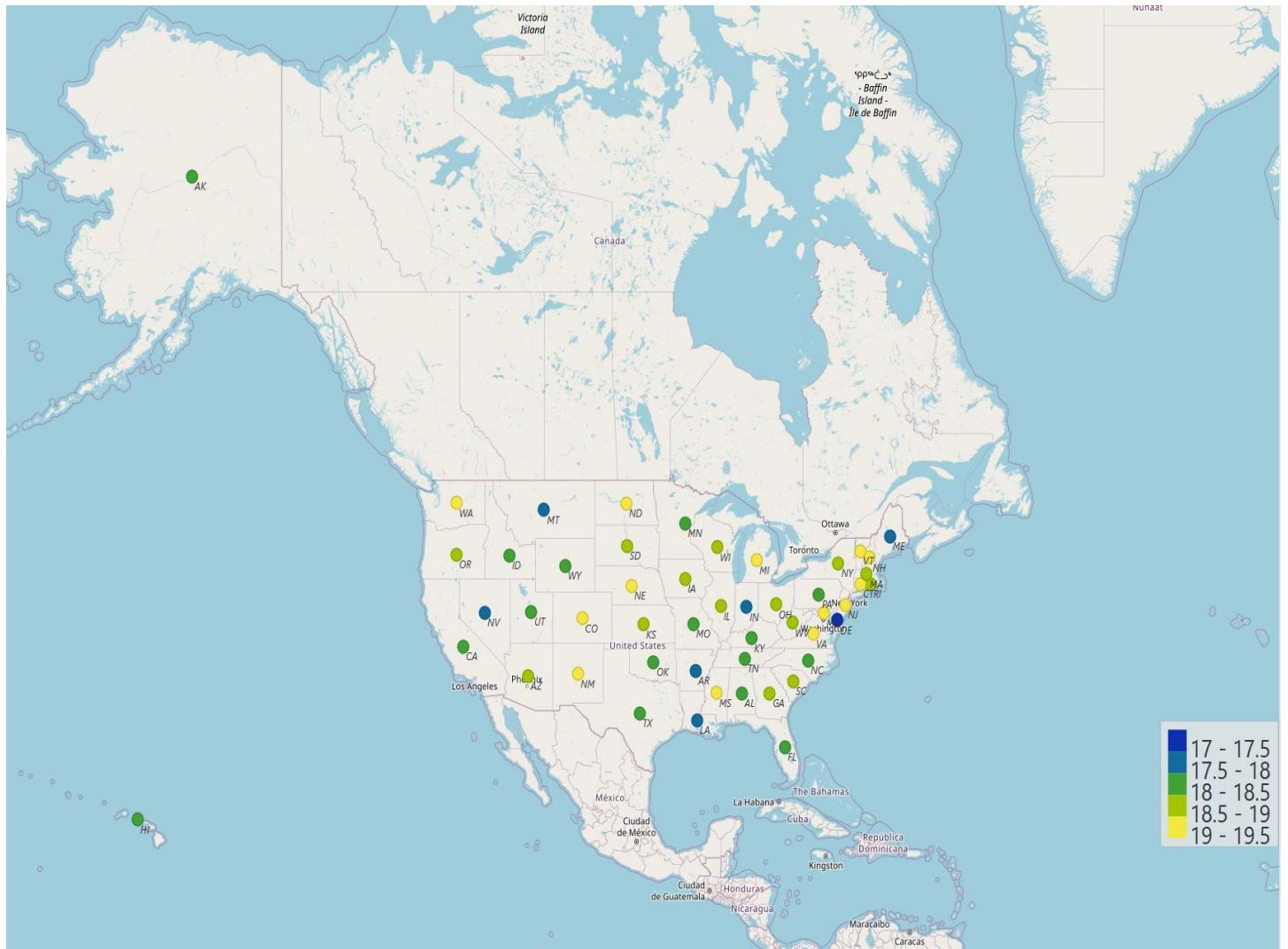


Figure 31 Average years in education mapped

- Combined graph showing winner, income (by size of dot) and years of education

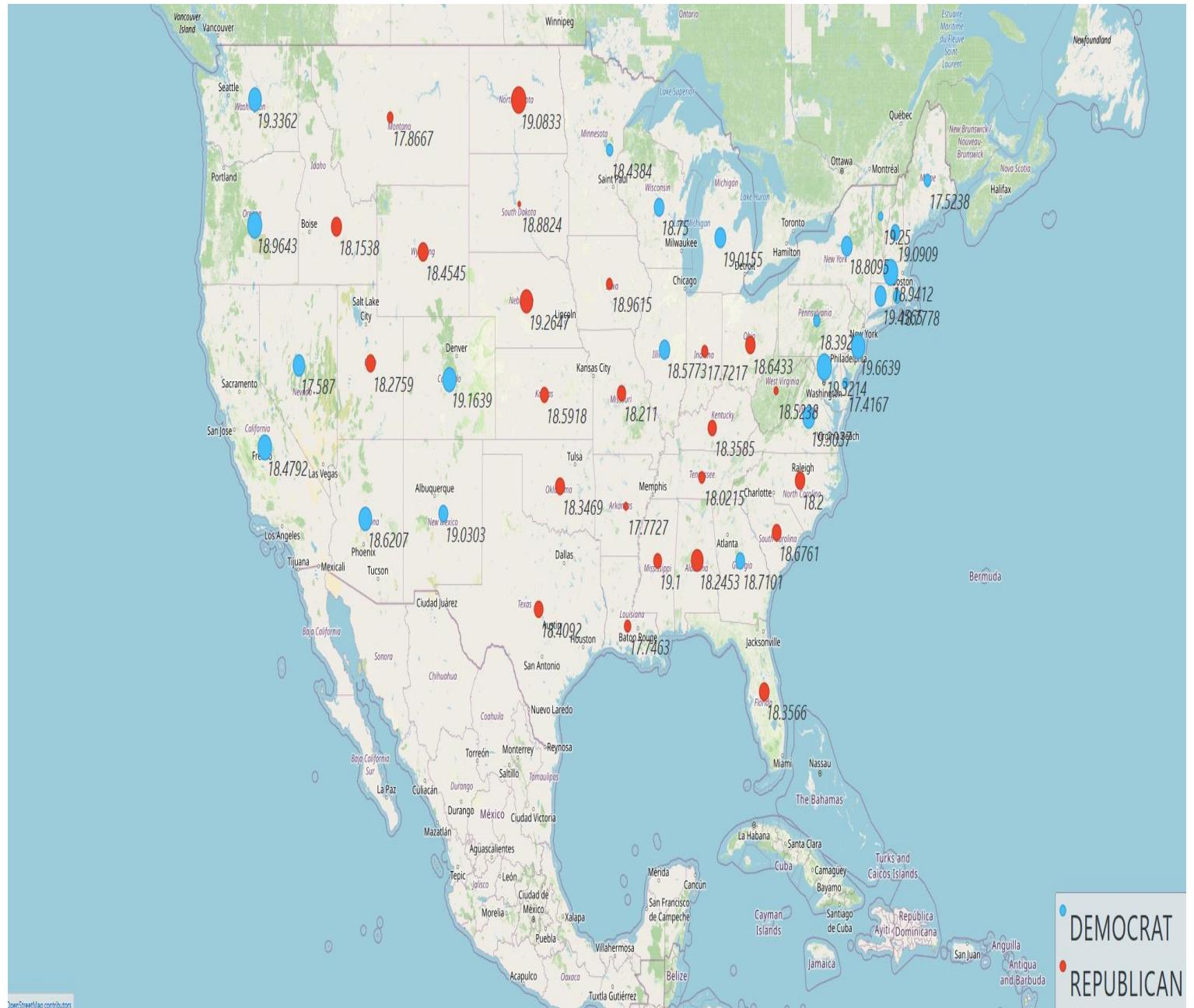


Figure 32 Map of multiple variables

Comment on the visual comparisons of the maps

- Regional results map: Southern America and the Midwest are more likely to vote for Trump whilst the North-eastern and West-coast states are likely to vote for Biden
- Income map: More urban and coastal states generally tend to have higher incomes
- Education levels map: More urban states appear to have higher educational attainment
- Combined map: There appears to be a correlation between education levels and income levels. Both these factors visually appear to correlate positively with voting for Biden, although more complex analysis might be required to establish that.

Test the 2 following hypothesis:

- *Low-income states voted for Trump*
- *States with high educational attainment voted for Biden*

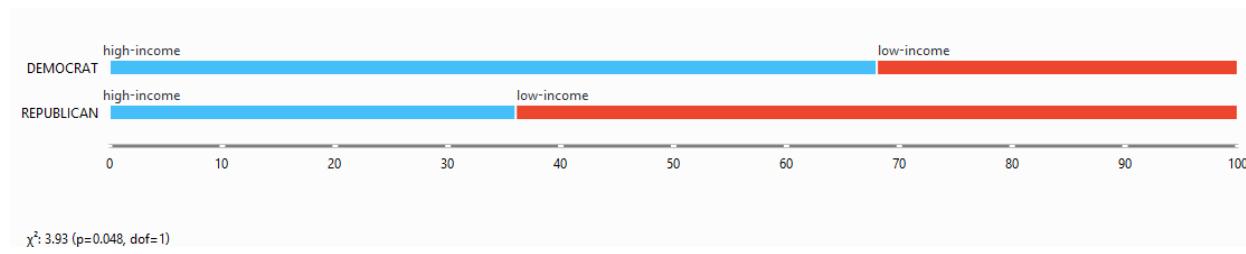


Figure 33 state_vote by state_income_level

The above box plot graph shows the election results of states based on their income levels. 65% of high-income states voted for the Democratic Party (Biden) whilst 67% of low-income states voted for the Republican Party (Trump).

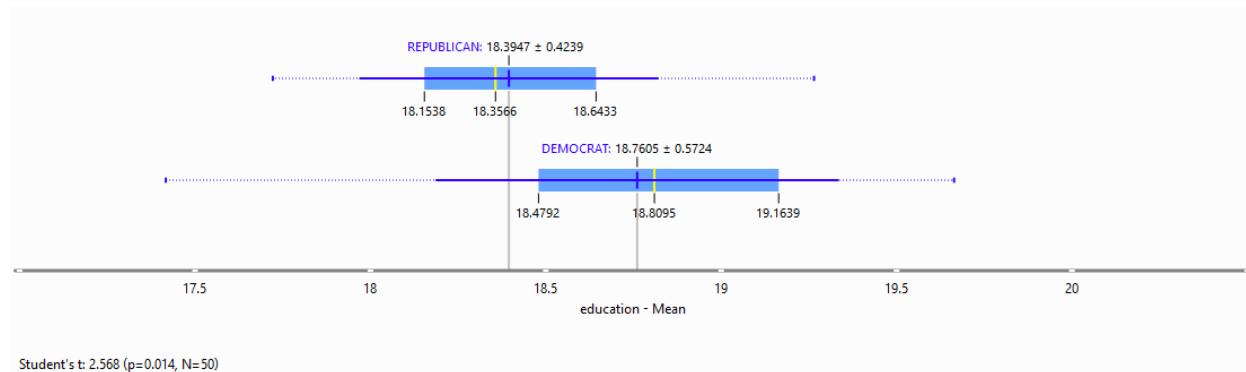


Figure 34 T-test results by Party

The above box plot shows the t-test results for each winning party by the distribution of years in education. The p-value (0.014) indicates that there is a statistically significant difference in mean education levels between the chosen political party. The spread of data for those who voted for Biden indicated a higher variability in education levels. The hypothesis is supported as Republic states have a lower mean education level (18.3947) vs Democrat winning states (18.8095).

Testing hypotheses using logistic regression:

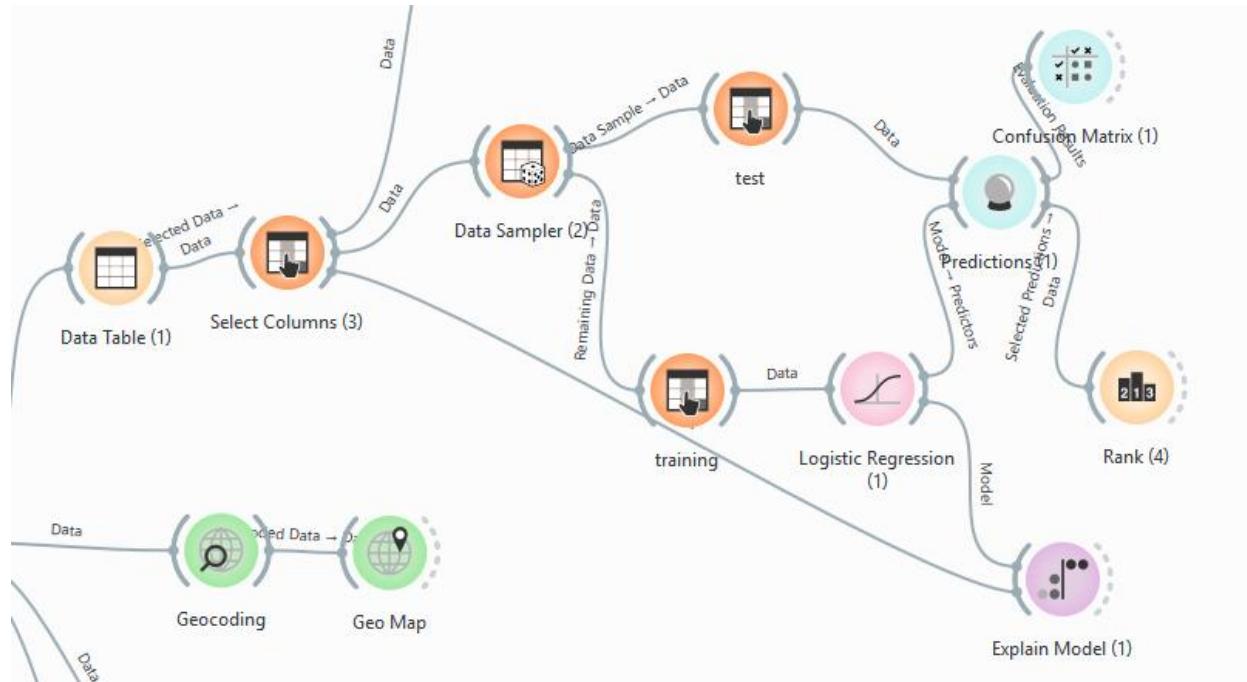


Figure 35 Logistic regression node setup

The state income level was set as target, and the data was sampled and ran through a Logistic Regression model to perform a regression analysis and analyze the impact of different variables.

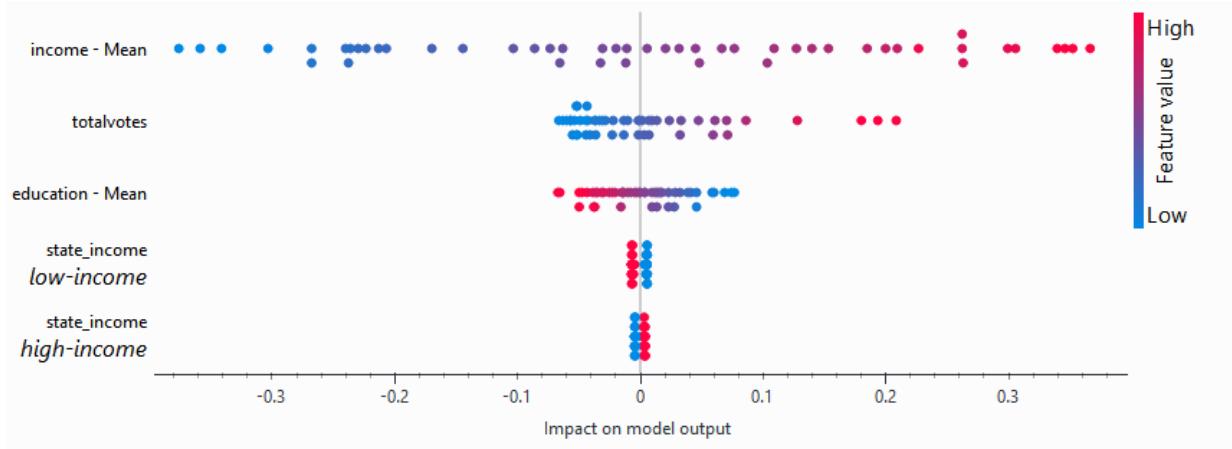


Figure 36 Democrat results

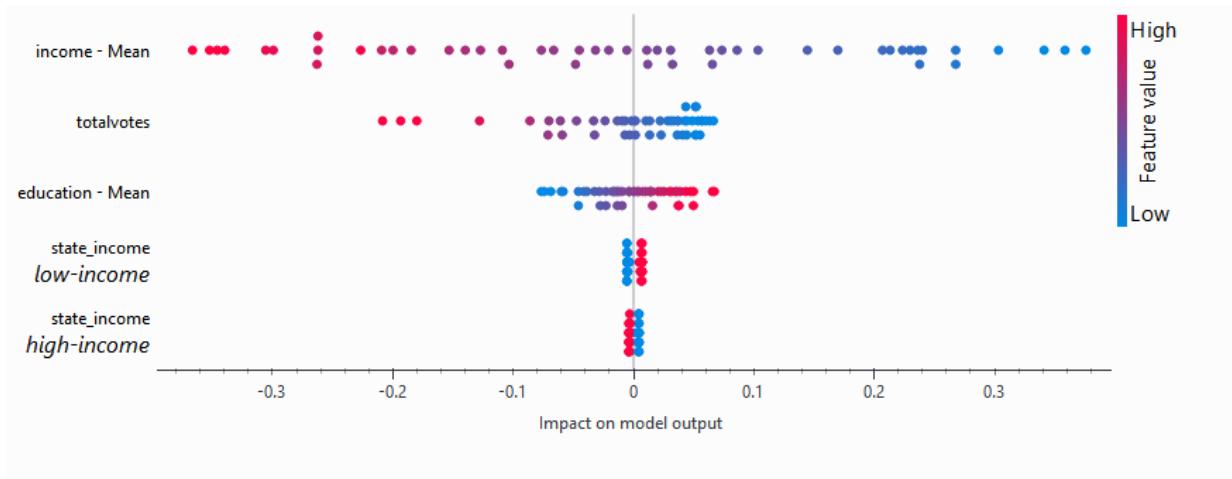


Figure 37 Republican results

By targeting the class “Democrats”, high-income states show a positive correlation with mean-income. A positive correlation is also observed with mean-education. For “Republicans” a positive correlation is shown with lower mean-income and mean-education. In both instances the mean-income results are more spread than the mean-education results, indicating that there is more diversity in income levels which might be influenced by type of industry, hours worked, and different costs-of-living associated with a State.

In summation, based on the statistical analysis and visualizations both hypotheses are supported.

Part 5. Your own data mining

Hypothesis: The impact of weekly working hours on income is moderated by the type of occupation

Exploratory analysis

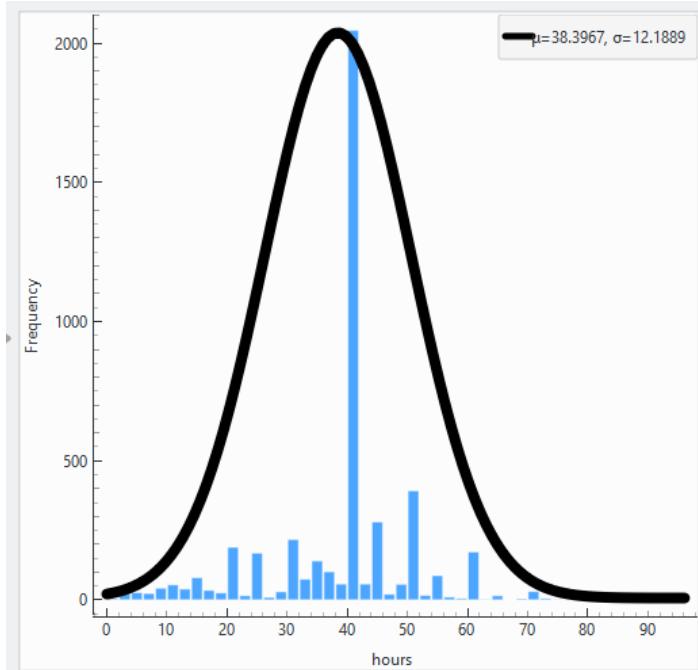


Figure 38 hours vs frequency graph

Figure 38 shows the distribution of weekly worked hours across the data set. The following graphs show the distribution of hours by industry, log(income) as well as the hourly rate data spread by each industry.

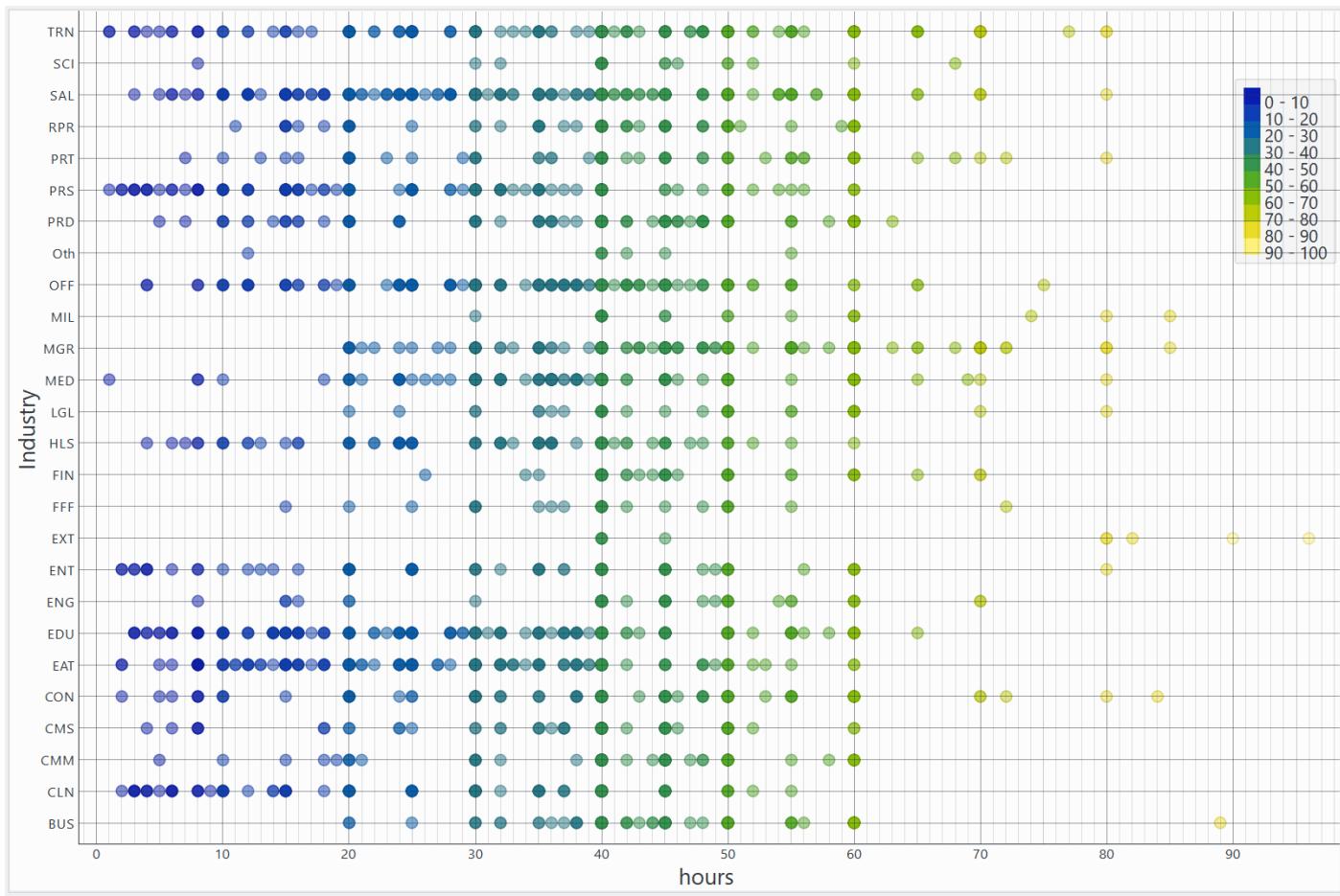


Figure 39 Weekly hours by industry Scatter Plot

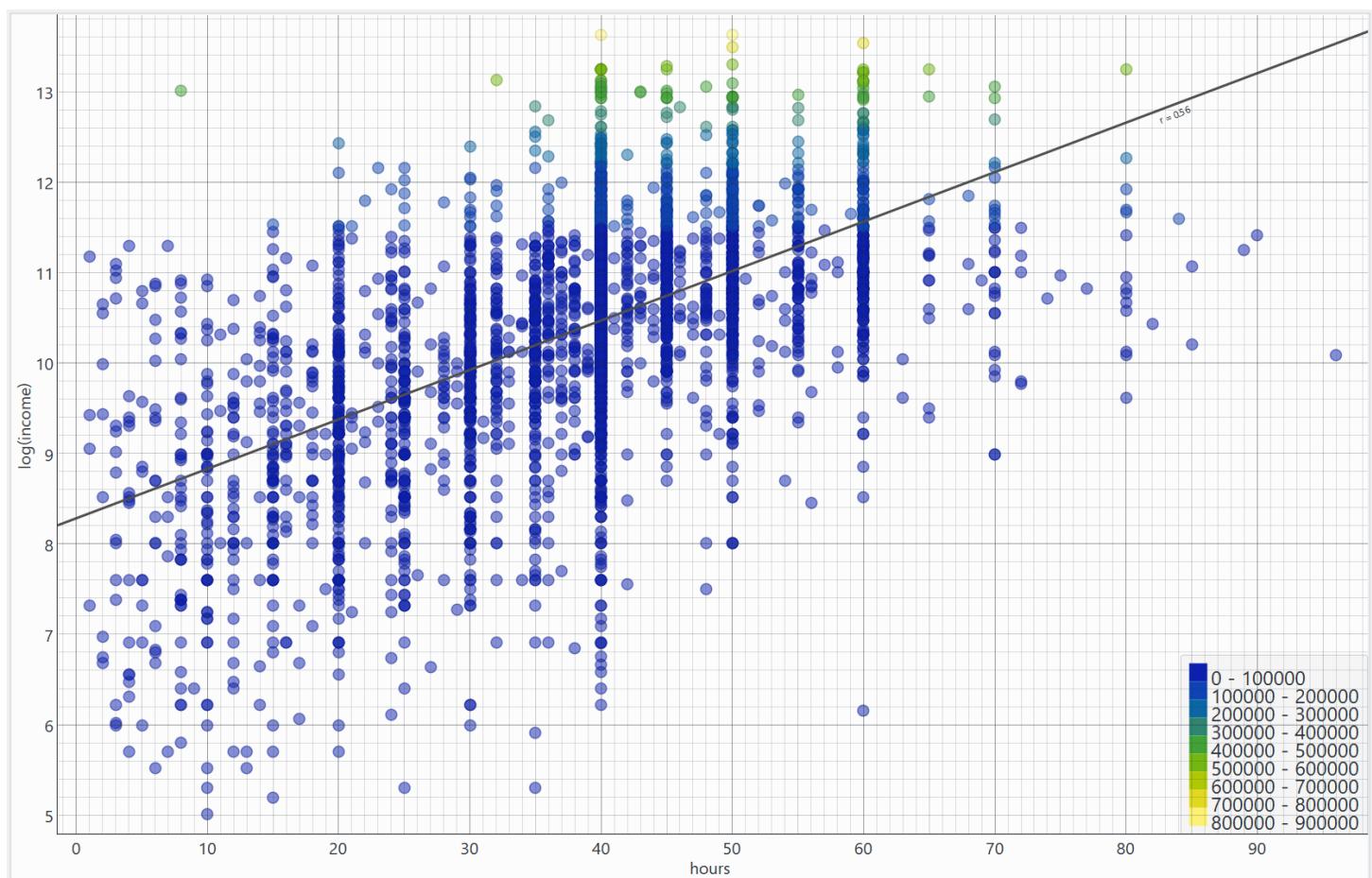


Figure 40 Weekly hours by log(income) Scatter Plot

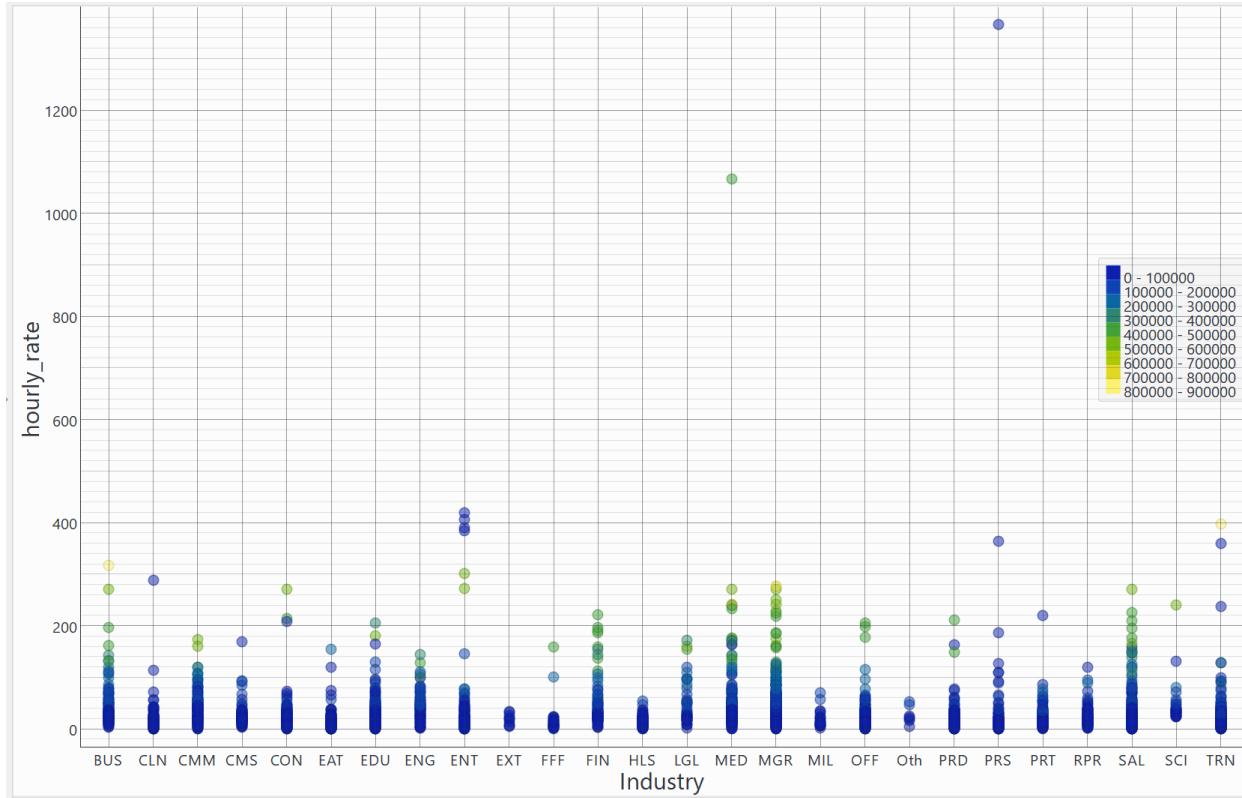


Figure 41 Hourly rate by Industry Scatter Plot

The t-test results (figure 42) show that industry type is statistically significant in determining the number of weekly hours worked as the p-value is 0.00.

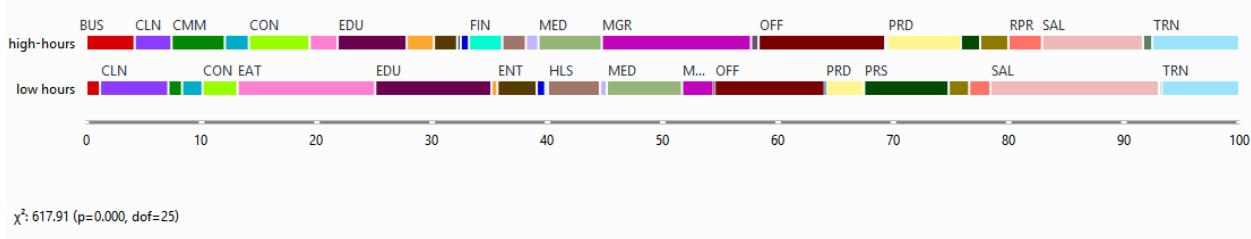


Figure 42 Box plot of Industry by number of hours

The summary statistics of the variables Hourly rate, hours and income are tabulated below:

	Hourly Rate	Hours	Income
Mean	28.1031	38.3967	56519.8
Mode	19.1781	40	50000
Median	19.1781	40	39300
Dispersion	1.50456	0.317445	1.26636
Min	0.109589	1	150
Max	1365.48	96	830000

Table 10

A new categorical category (hours_type) is created based off the median number of weekly hours worked to categorize the data into high and low-hours. Different classifier models were used to determine the hours_type. Those models are shown in Figure 43 and their results are displayed below.

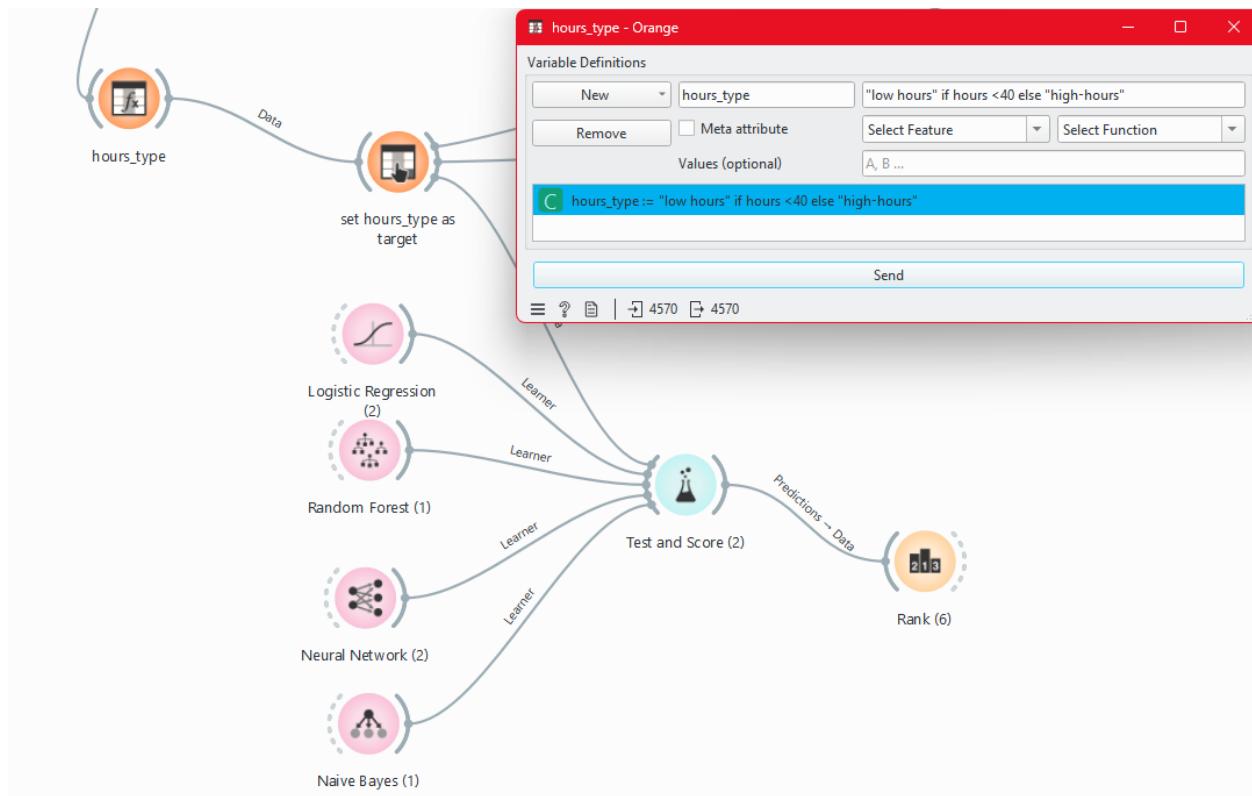


Figure 43 Orange setup showing models used

Evaluation results for target (None, show average over classes)						
Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression (2)	1.000	0.995	0.995	0.995	0.995	0.987
Random Forest (1)	0.891	0.850	0.845	0.847	0.850	0.625
Neural Network (2)	0.951	0.903	0.902	0.902	0.903	0.763
Naive Bayes (1)	0.800	0.764	0.766	0.770	0.764	0.445

Compare models by: Area under ROC curve					<input type="checkbox"/> Negligible diff.:	0.1
	Logistic Regression (2)	Random Forest (1)	Neural Network (2)	Naive Bayes (1)		
Logistic Regression (2)		1.000	1.000	1.000		
Random Forest (1)	0.000		0.002	0.999		
Neural Network (2)	0.000	0.998		1.000		
Naive Bayes (1)	0.000	0.001	0.000			

Figure 44 Test and Score results

Out of all the models, Logistic Regression performed best in the side-by-side comparison and by having the best scores across all parameters. This model was applied to training data and the results are displayed:

		Predicted		Σ
		high-hours	low hours	
Actual	high-hours	100.0 %	0.0 %	1031
	low hours	0.9 %	99.1 %	432
Σ		1035	428	1463

A rank node is attached to show the impact of each variables on weekly hours worked and an explain model is attached to the classifier.

	#	Inf...ain	Gain ratio	χ^2
N income		0.175	0.087	241.212
C Industry	26	0.105	0.025	2.414
N occupation		0.060	0.030	1.263
N hourly_rate		0.057	0.029	65.813
N age		0.042	0.021	12.569
C sex	2	0.027	0.027	28.449
C education_lvl	8	0.017	0.006	19.756
C marital	5	0.016	0.010	63.634
N education		0.012	0.006	14.141
C CoW	5	0.000	0.000	1.001
C race	2	0.000	0.000	0.462
C PoB	2	0.000	0.000	0.428

Figure 45 Rank node results

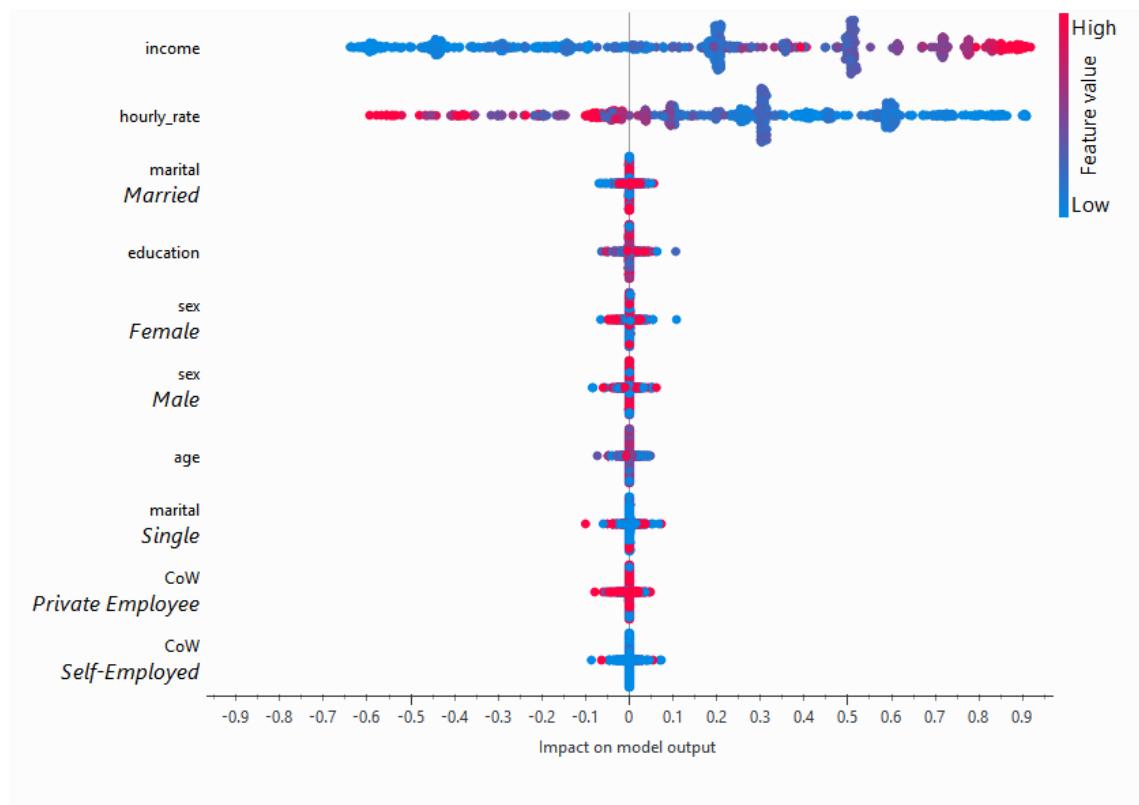


Figure 46 Explain Model node results

Results Analysis

The exploratory analysis shows a positive correlation between hours worked with income with a diverse variation across different types of industries. This suggests that the type of industry have a statistically significant impact on the number of weekly worked hours. The final Rank node and Explain model results establish that occupation type does moderate the number weekly hours worked supporting the hypothesis. However, due to the information gain, gain ratio and chi-square results, income has a higher influence on weekly hours over the type of industry.

Appendix

Part 1. Orange Nodes

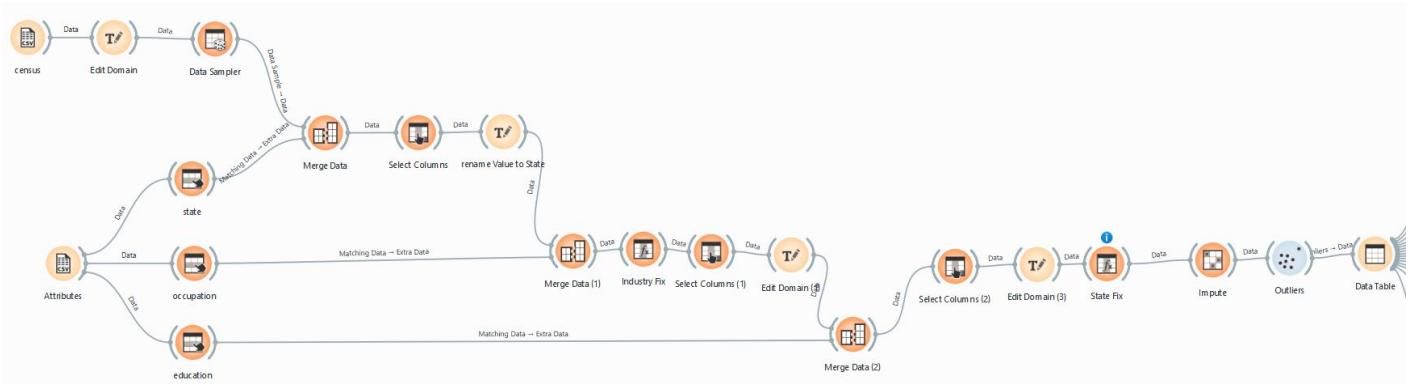


Figure 47 Part 1

Part 2. Orange Nodes

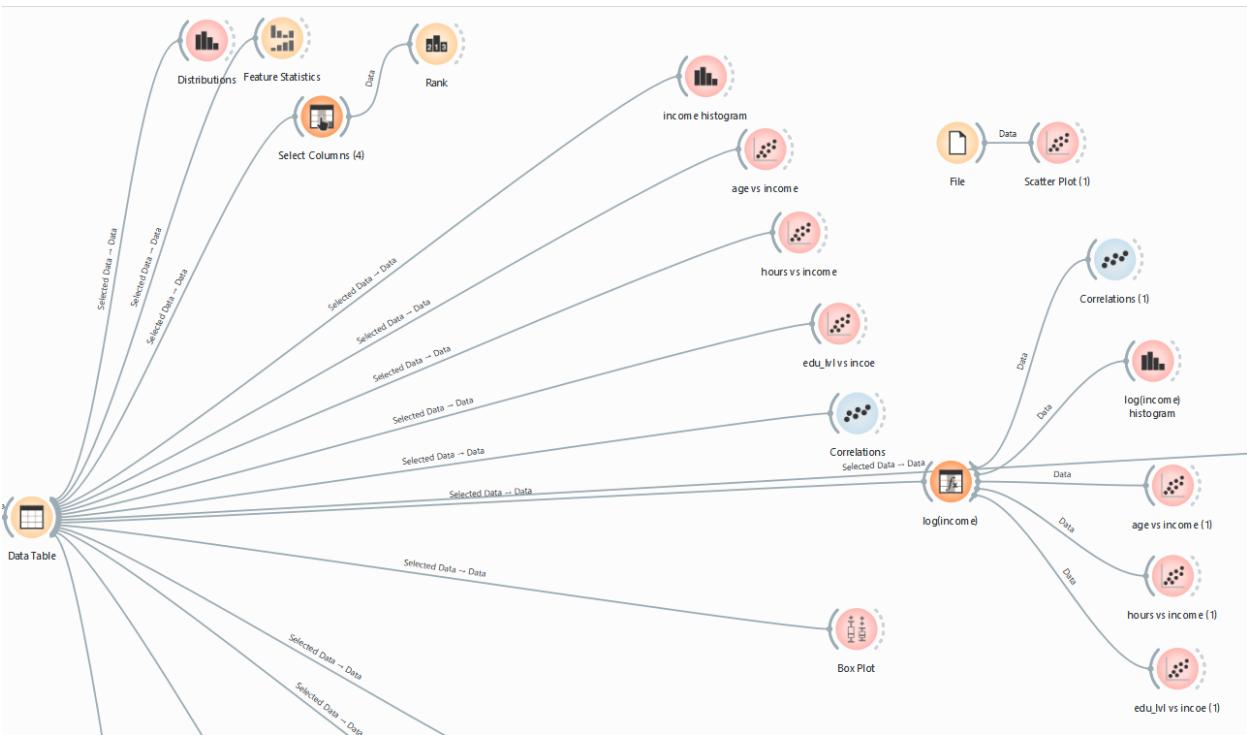


Figure 48 Part 2

Part 3. Orange Nodes

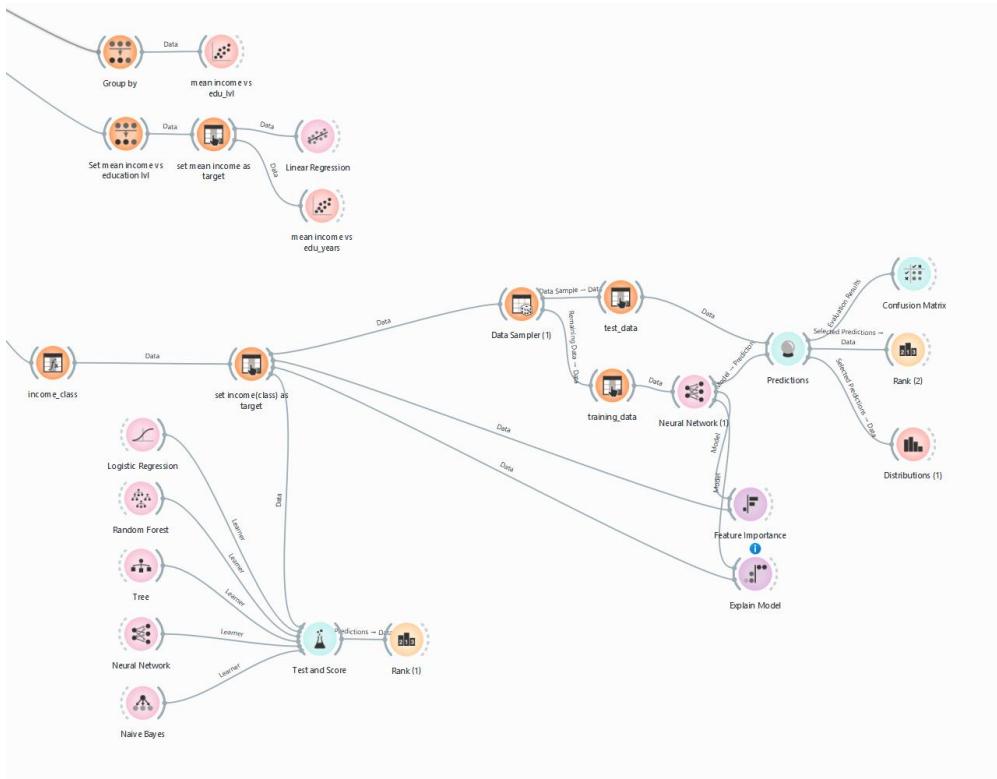


Figure 49 Part 3

Part 4. Orange Nodes

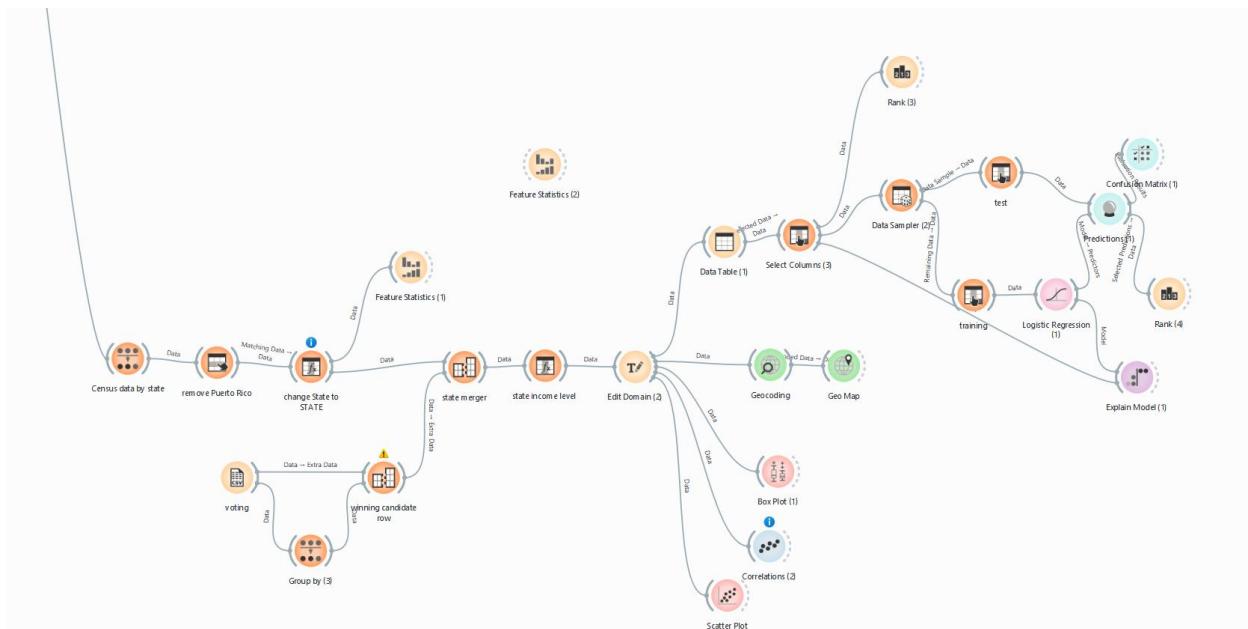


Figure 50 Part 4

Part 5. Orange Nodes

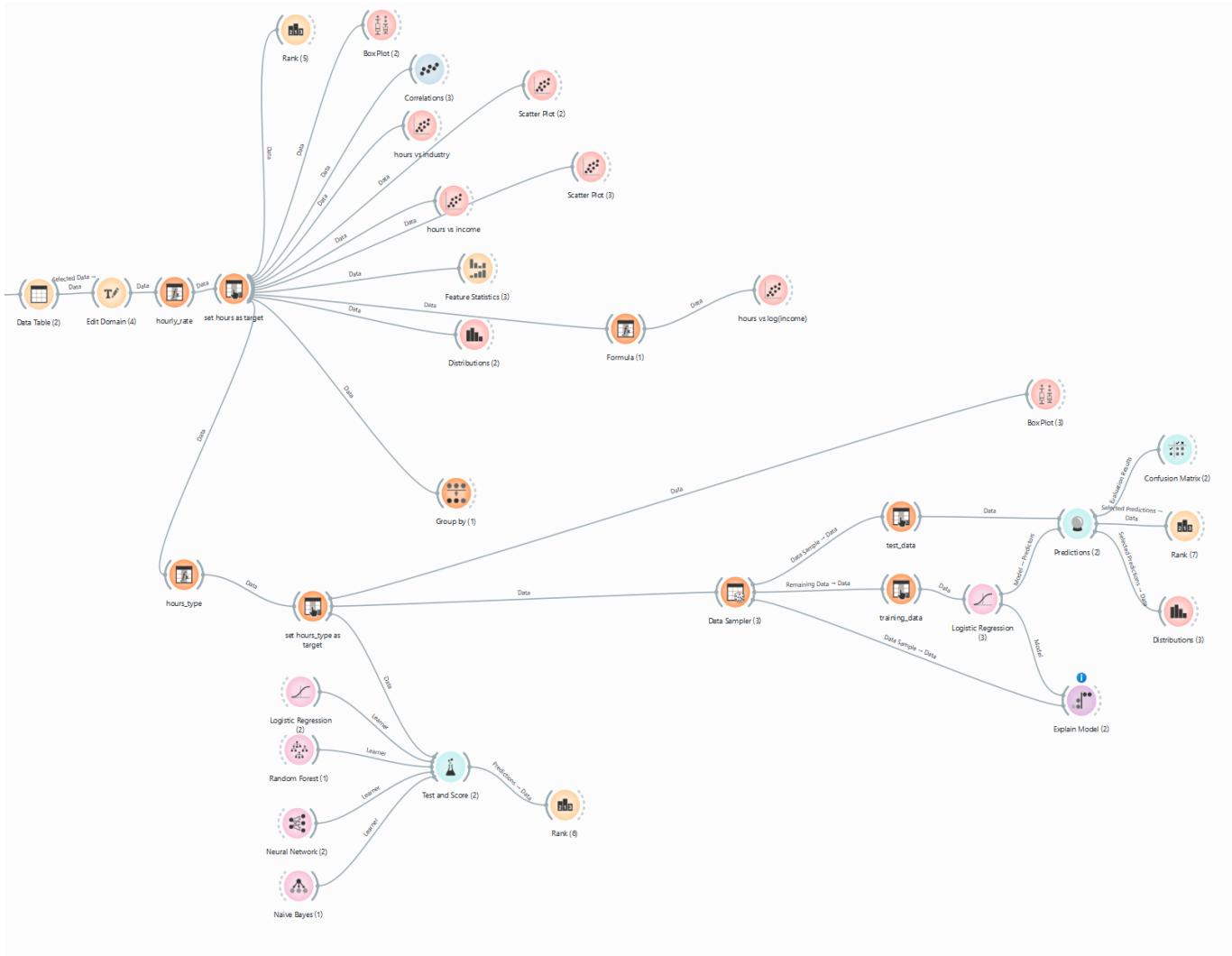


Figure 51 Part 5