

What Non-Work Related Factors Effect Time Spent Alone In the US?

Naima Abdirahman



Introduction

- The study explores factors predicting the total non-work-related alone time adults spend in the United States from 2003 to 2023.
- Data is sourced from the American Time Use Survey (ATUS).
- The aim is to understand how non-work-related variables influence alone time, including:
 - Household size
 - Age
 - Sex
 - Labor force status
 - Time spent with family
 - Time spent with spouse/partner exclusively
 - Presence of spouse/partner

Question: Are these variables good predictors for time spent alone?

Methodology

- Data is from the American Time Use Survey (ATUS) from 2003 to 2023.
- Data sourced from respondent and roster files, merged by the unique household identifier (TUCASEID).
- 10,000 unique individuals were randomly selected due to the large size of the full dataset.
- Dependent variable: Total time spent alone (excluding sleep).
- Subsetted by these predictors:
 - Age, Gender, Labor status, Household size, Presence of a spouse or partner, Time spent with family, Time spent with spouse/partner exclusively

Analysis Approach:

Exploratory data analysis (EDA) conducted using:

- Box plots
- Line plots
- Multicollinearity was assessed

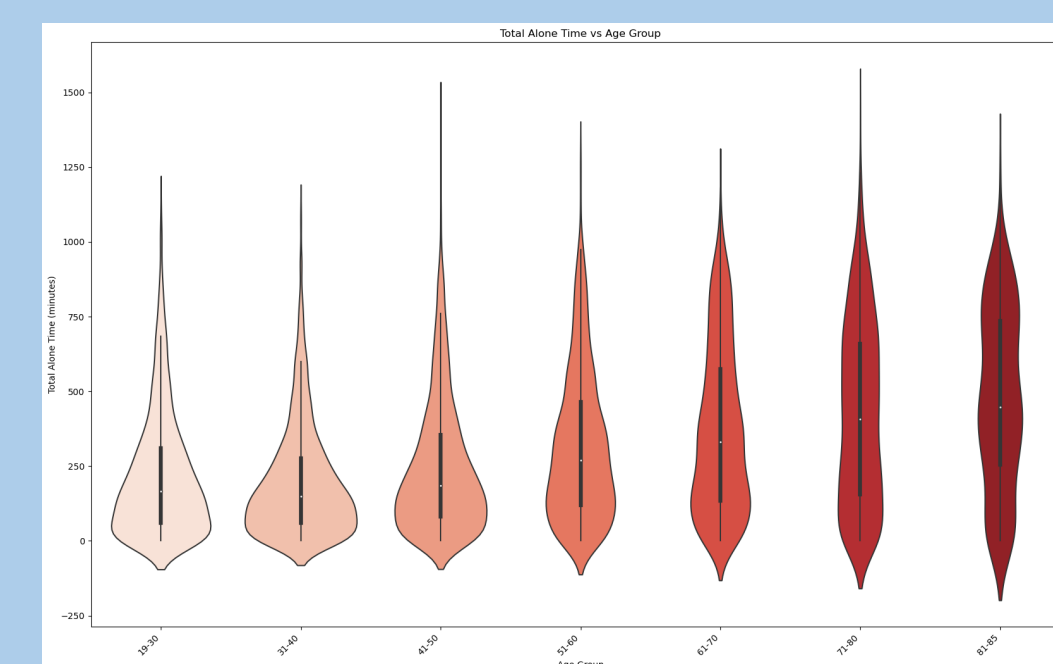
Modeling:

- OLS model
- Random Forest Model
- Model performance was assessed using Root Mean Squared Error (RMSE) after splitting data into training and testing sets.

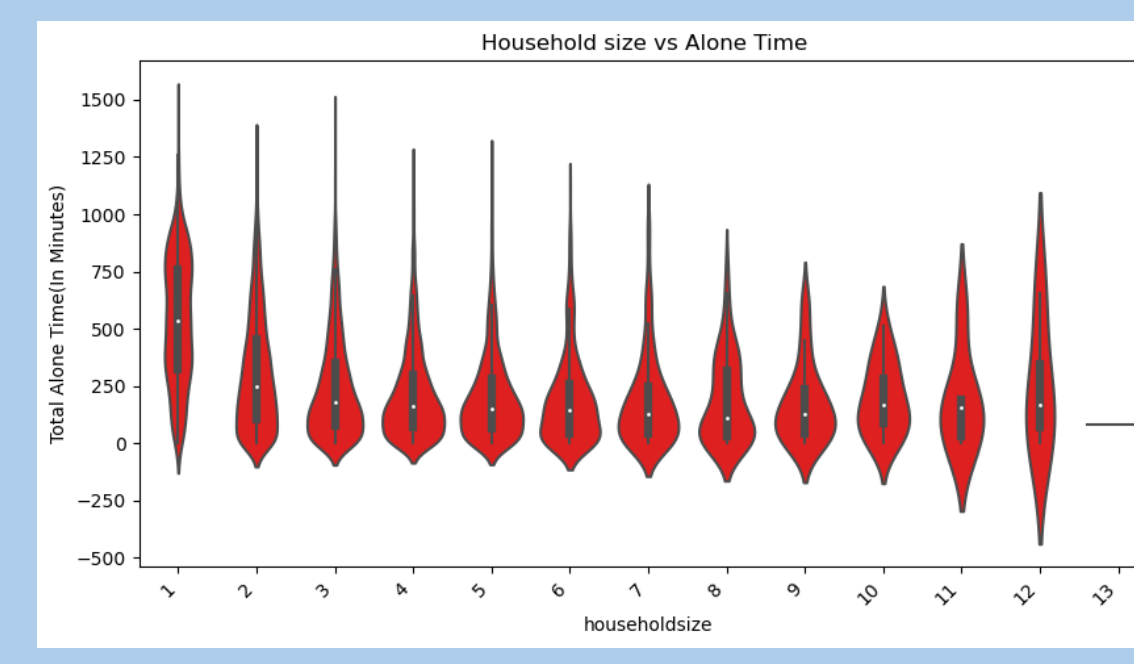
Tools and Libraries:

Python, with relevant statistical and machine learning libraries.

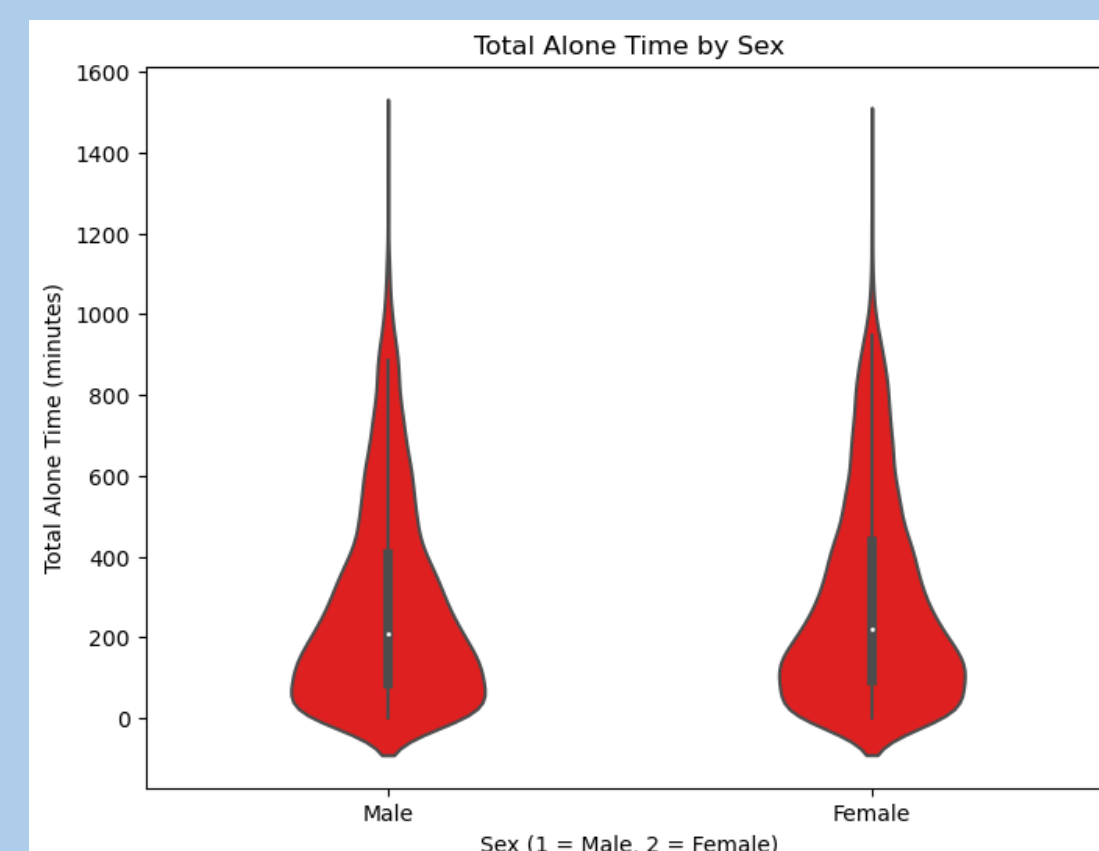
Data Exploration



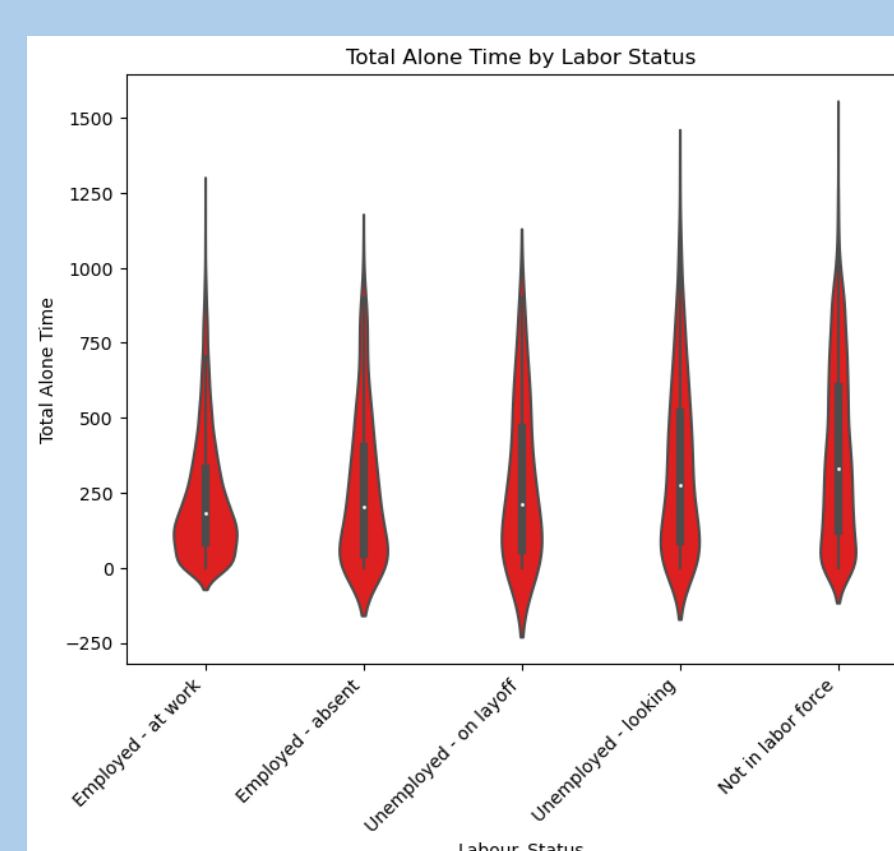
Alone time increases with age



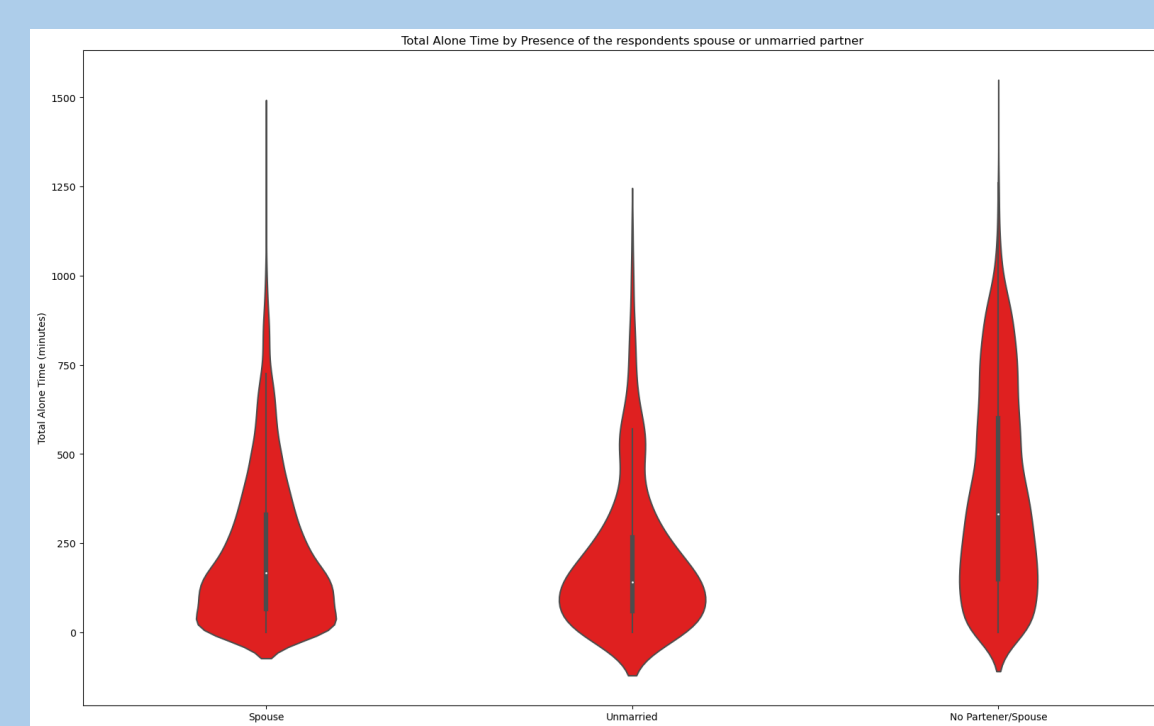
Alone time decreases with household size



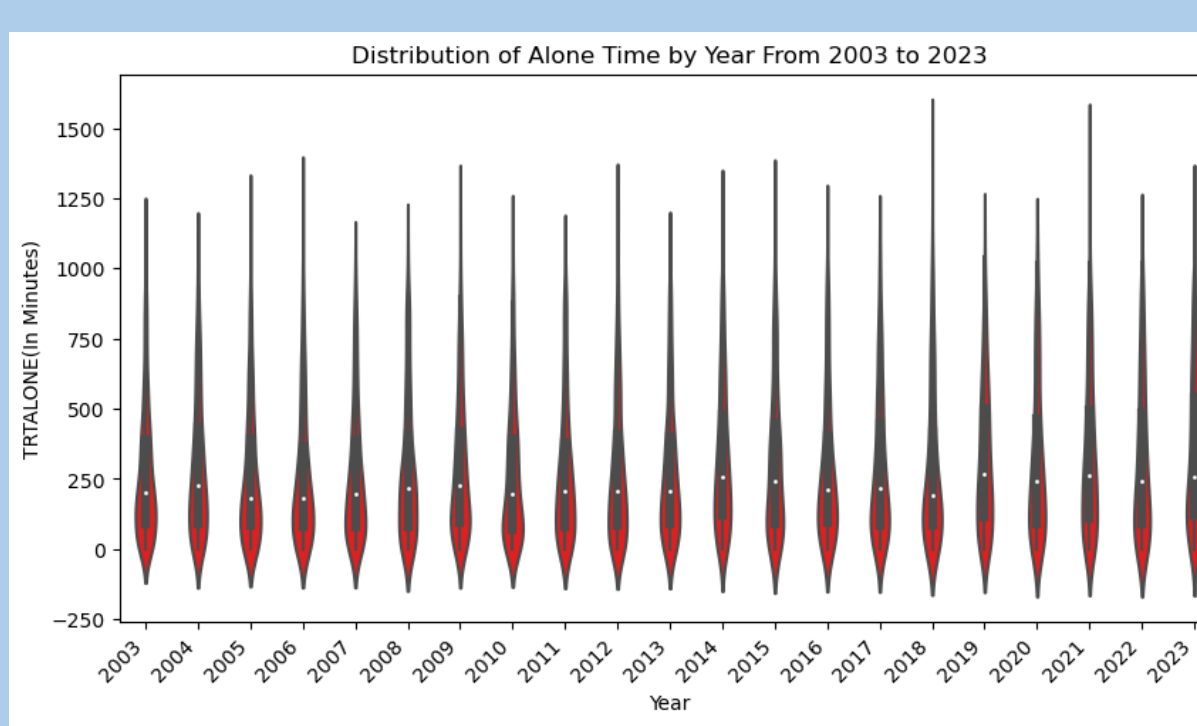
Men and Women spent similar amounts of alone time, but there are a few men who spent more alone time



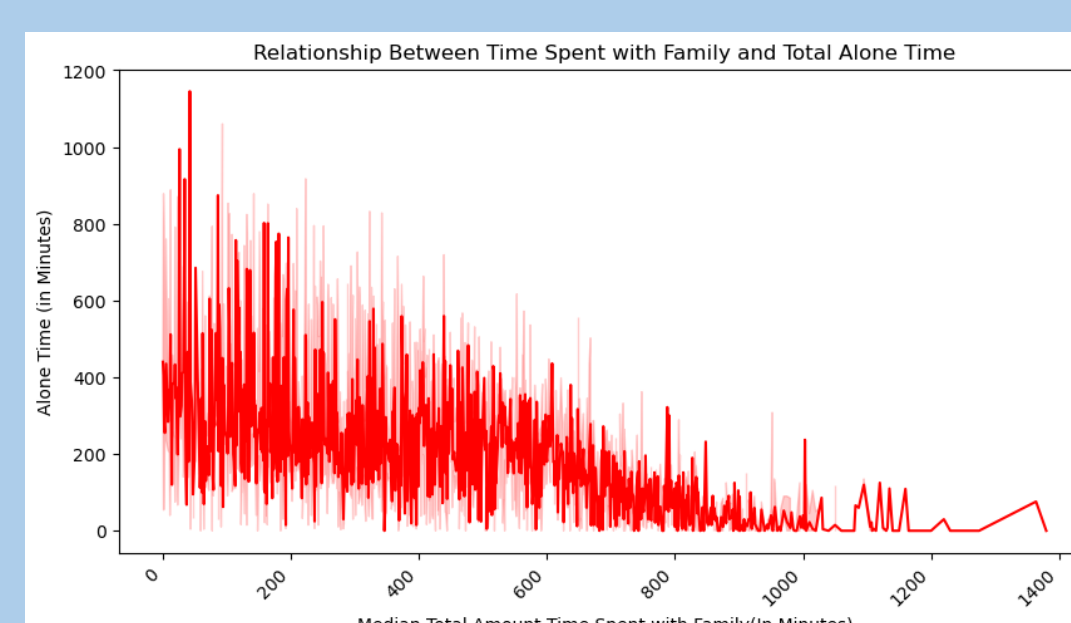
Unemployed- on layoff individuals spent the least amount of alone time



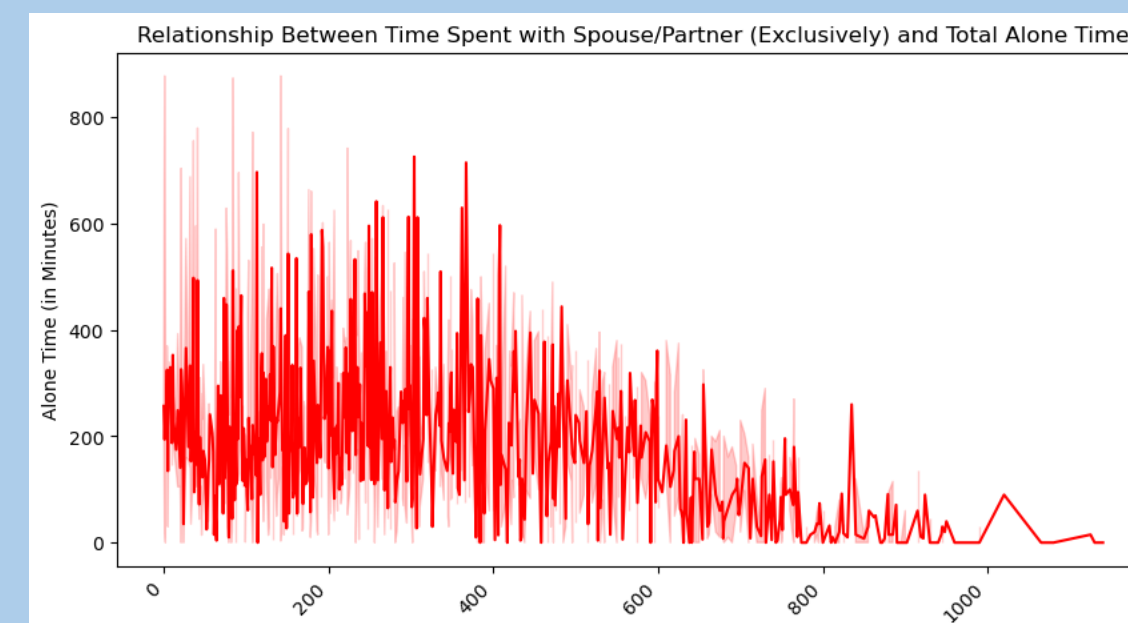
Those with a partner(married or not) spent the same amount of alone time. But, those with no partner/spouse spent more time alone time



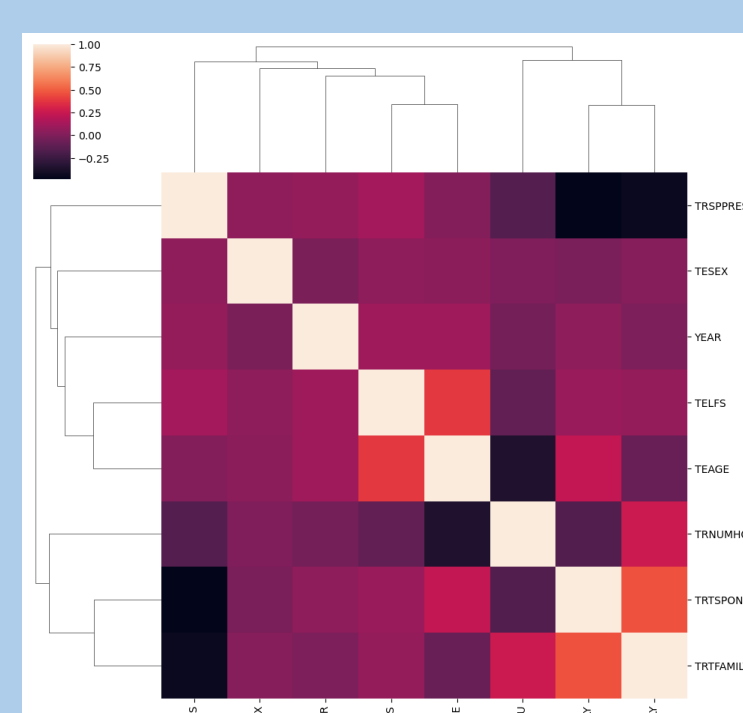
Total Alone Time spent by US individuals has fluctuated over the years increasing during the covid years



Alone Time relatively decreases with time spent with family



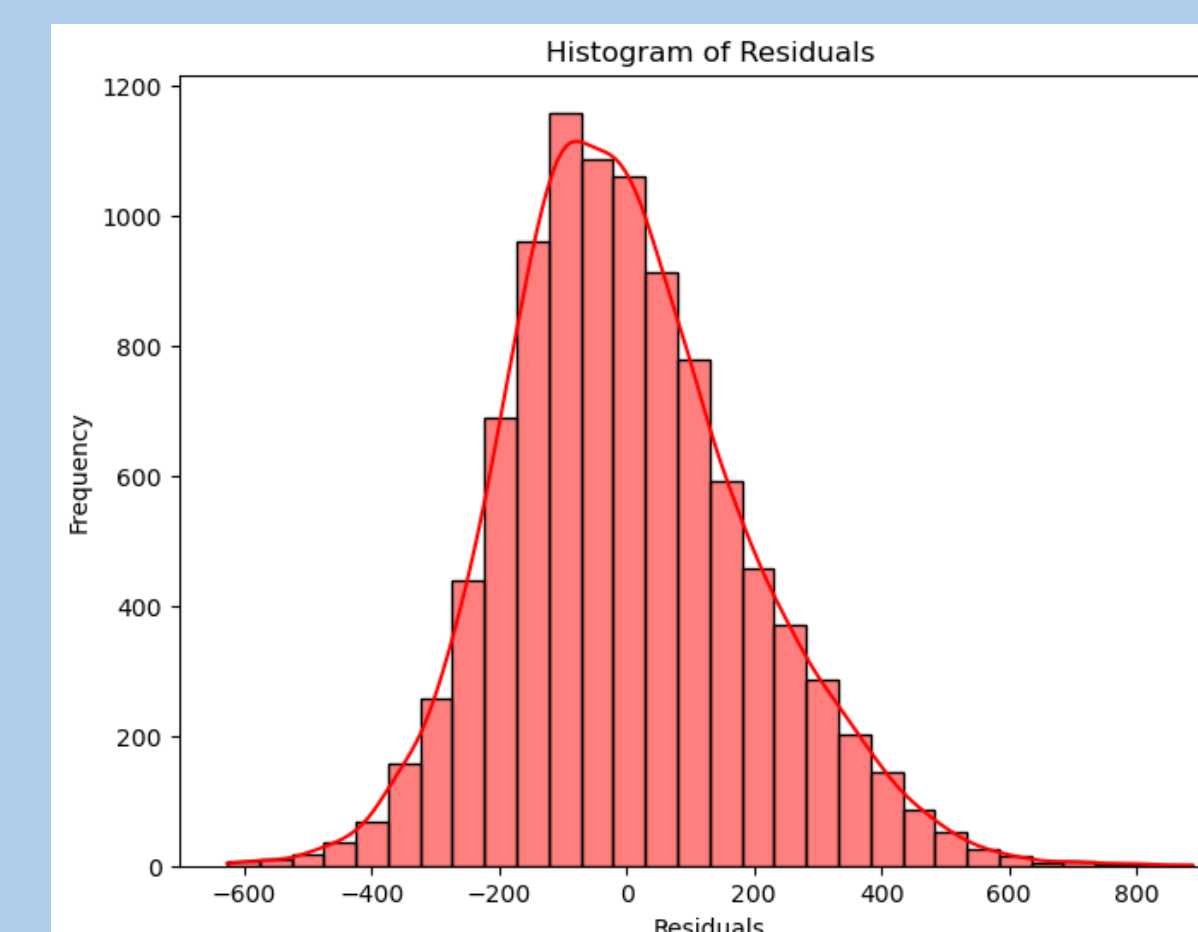
Alone Time relatively decreases with exclusive time spent with spouse/partner



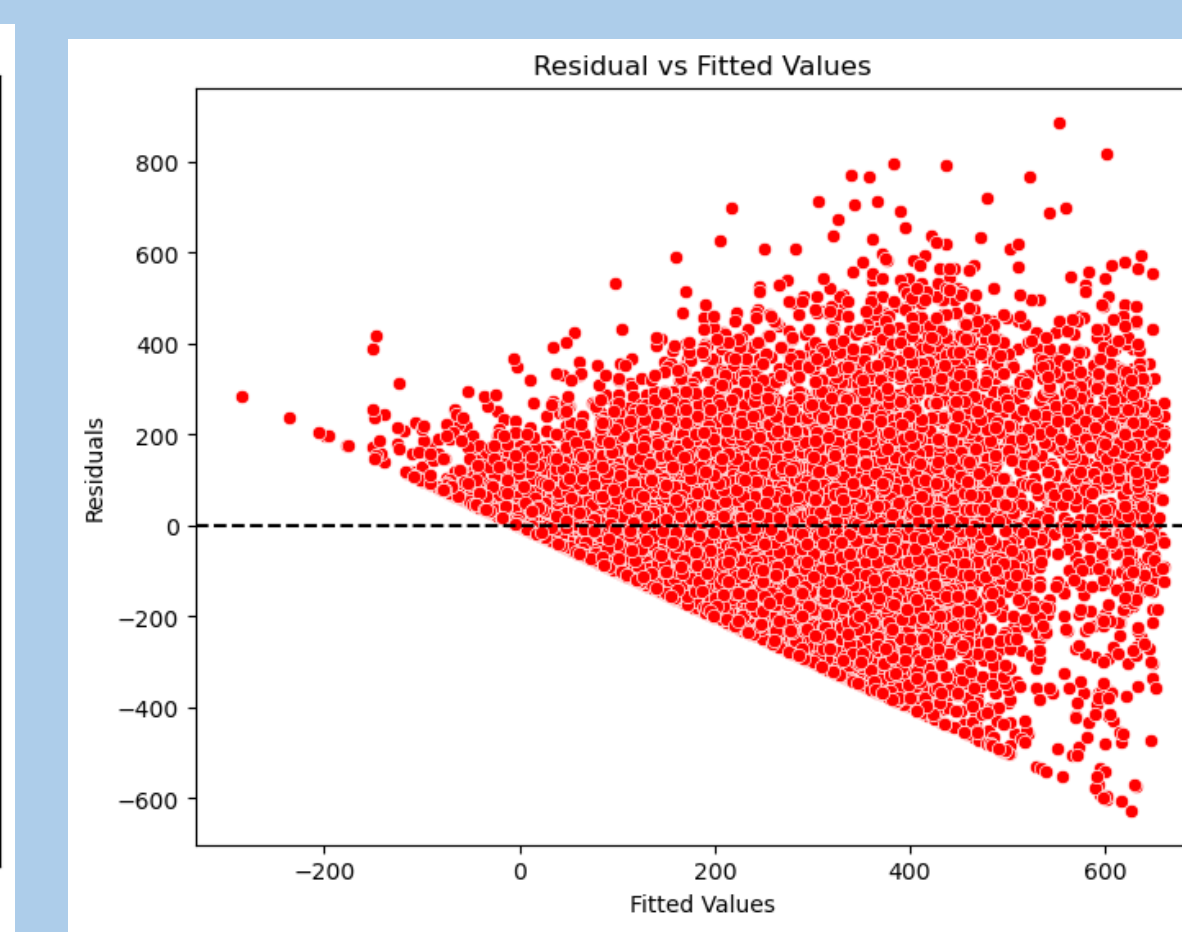
multicollinearity: none of the predictors highly correlate

Results

OLS Model



The histogram of the residuals is relatively normal so normality assumption is met



The residuals form a funnel shape Thus homoscedasticity is not met! Even after transforming the response variable using sqr transformation, the residuals still form a funnel-shape

Predictor	Coef	P-value
HH size	-20	<0.05
Age	2	<0.05
Sex	5.8	>0.05
Labour Status	34.9	<0.05
Time with family	-0.4	<0.05
Time spent with spouse/partner	-0.1	<0.05
Year	0.96	<0.05
Spouse/partner presence	-1.6	>0.05

Random Forest Model

	Feature	Importance
4	TimewithFamily(Minutes)	0.362423
5	Age	0.175492
2	Labour_Status	0.159631
7	YEAR	0.125606
0	HH_size	0.086059
3	TimewithSpouse/Partener(Minutes)	0.038851
1	Spouse_Partner_Presence	0.027328
6	Sex	0.024610

- The most important predictors identified by the RF model, are also statistically significant in the OLS model
- The 2 least important predictors, are not statistically significant in the OLS model

Model Performance Comparison

Model	RMSE	R^2
OLS	189.9	0.43
RF	186.6	-

Discussion

- Household size, labor status, time spent with spouse/partner, time spent with family, and age significantly predict time spent alone.
- Sex, and Spouse/Partner presence are not good predictors of time spent alone time
- Labor status and age are positively related to time spent alone
- Household size, time spent with spouse/partner, and time spent with family are negatively related to time spent alone.
- The OLS and the Random Forest models agree with the important predictors and have very close RMSEs indicating that these predictors have a strong and stable relationship with the response variable, regardless of the model type
- The RMSEs indicate that the model's error is approximately 13% of the maximum possible value, which suggests a moderate level of accuracy

Conclusion

- While comprehensive, the ATUS data is self-reported, which introduces potential biases due to recall or social desirability effects
- Unmeasured factors, such as mental health or lifestyle choices, may contribute to alone time.
- Future research can explore additional predictors or interactions to improve model accuracy.