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 demographic and socioeconomic variables influence alone time, including:
  - Household size
  - Age
  - Gender
  - Labor force status
  - Time spent with family
  - Time spent with spouse/partner exclusively

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).

Key Predictors:

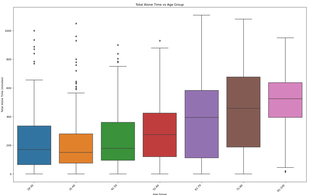
•Age, Gender, Labor status, Household size, Presence of a spouse or partner, Time spent with family, Time spent with spouse/ partner exclusively
Analysis Approach:
Exploratory data analysis (EDA) conducted using:
•Descriptive statistics
•Box plots
•Scatter plots
•Multicollinearity was assessed

Modeling:

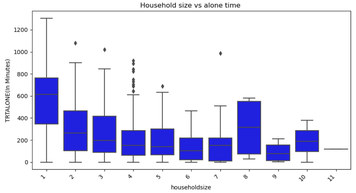
•Random Forest
•A generalized Linear Model was used for linear regression
•Model performance was assessed using Root Mean Squared Error (RMSE) after splitting data into training and testing sets.

Tools and Libraries:
Python, with relevant statistical and machine learning libraries.

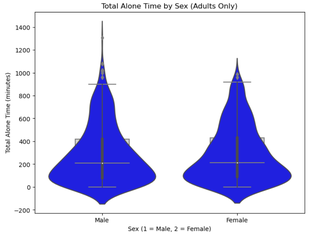
Data Exploration



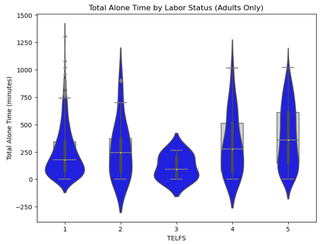
- Alone time increases with age



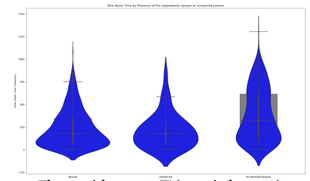
Alone time decreases with household size



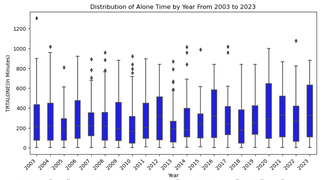
- Men and Women spent similar amounts of alone time, but there are a few men who spent more alone time



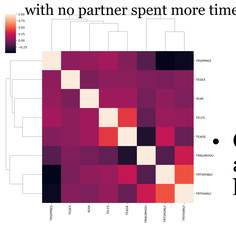
- Unemployed- on layoff individuals spent the least amount of alone time



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



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



- Checked for multicollinearity and none of the predictors highly correlate

Results

	Predictor	Levene Statistic	Levene P-Value
0	TRNUMHOU	8.333321	1.780966e-05
1	TRSPPRES	16.979487	5.608253e-08
2	TELFS	1.366243	2.436706e-01
3	TRTSPONLY	9.780898	1.814467e-03
4	TRTFAMILY	45.058240	2.631951e-27
5	TEAGE	5.382731	1.119837e-03
6	TESEX	0.071860	7.887031e-01
7	YEAR	0.811465	4.875870e-01

- The residuals do not have an equal variance for a Linear Regression to be used
- So, a Generalized Linear Model (the default) and a Random Forest Model were used instead

Random Forest Model Performance Metrics:

	Feature	Importance
3	TRTFAMILY	0.350179
6	TEAGE	0.174204
7	YEAR	0.115252
0	TRNUMHOU	0.092544
1	TELFS	0.091512
4	TRTSPONLY	0.031704
2	TRSPPRES	0.020982
5	TESEX	0.013277

- Root Mean Squared Error: 196.71
- R-squared: 0.368
  - On average, the model's predictions for the time spent alone are off by approximately 197 minutes
    - Thus, the model's predictions are not very accurate
  - The model explains 37% of the variance in the total alone time

GLM Model Performance Metrics:

- Root Mean Squared Error (RMSE): 192.79
- R-squared: 0.3932998610897971
- Significant Predictors (p-value < 0.05):
  - Household size (p-value = 0.001)
  - Labor Status(p-value = 0.000)
  - Time spent with spouse/partner (p-value = 0.009)
  - Time spent with family(p-value = 0.000)
  - AGE (p-value = 0.000)

Equation:

$$\text{TimeSpentAlone} = -581.2727 - 16.7944(\text{Householdsize}) + 31.0774(\text{Labor Status}) - 0.1252(\text{Time with spouse/partner}) - 0.3500(\text{Time with family}) + 2.8859(\text{AGE})$$

Conclusion

- Household size, labor status, time spent with spouse/partner, time spent with family, and age significantly predict time spent alone.
- 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.
- Future research can explore additional predictors or interactions to improve model accuracy.

Limitations

- 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.