Mobility Behavior of Mobile and Immobile People

G. Arnarson, F. Bartke, M. Van Laere, S. Nguyen

2025-01-17

1 Introduction

Discuss the research objectives and question:

- How are immobile people characterized?
- Which model can best predict their immobility?

---

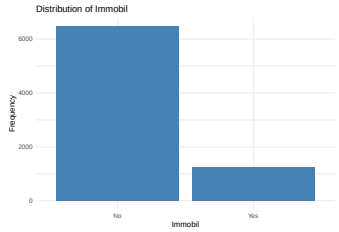
2 Literature Review

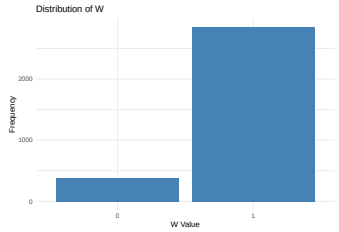
2.1 Variables Identified

- From literature: Age, working status, health, sex, income, possession of a car, etc.
- From common sense: Remote work, bike or motorcycle ownership, distance to public transport.

#Data manipulation

2.2 Dataset Preparation



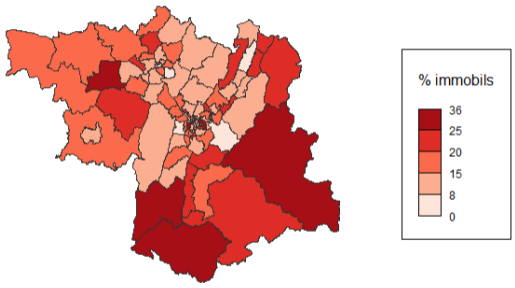



```
## 'data.frame':    7720 obs. of  17 variables:
##  $ id_pers     : num  1.01e+08 1.01e+08 1.01e+08 1.01e+08 1.01e+08 ...
##  $ immobil     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
##  $ sexe        : Factor w/ 2 levels "Male","Female": 1 1 2 1 2 1 2 1 2 1 ...
##  $ dispovp     : num  0 1 2 3 3 3 2 2 2 0 ...
##  $ age         : int  24 42 28 33 30 30 27 57 52 17 ...
##  $ age_group   : Factor w/ 3 levels "0-20","20-60",..: 2 2 2 2 2 2 2 2 2 1 ...
##  $ has_car     : Factor w/ 2 levels "No","Yes": 1 2 2 2 2 2 2 2 2 1 ...
##  $ cspgroup    : Factor w/ 4 levels "1","3","4","non_significant": 4 2 3 2 2 2 2 2 3 4 ...
##  $ W           : Factor w/ 2 levels "No","Yes": 2 2 NA 2 NA NA NA 2 NA NA ...
##  $ TYPE_HAB    : Factor w/ 2 levels "Groep 1","Groep 2": 2 2 NA 2 NA NA NA 2 NA NA ...
##  $ zoneres     : Factor w/ 32 levels "1","2","3","4",..: 3 3 NA 1 NA NA NA 2 NA NA ...
##  $ OCCU1       : Factor w/ 7 levels "1","2","3-5",..: 4 1 NA 2 NA NA NA 1 NA NA ...
##  $ travdom     : Factor w/ 2 levels "1","2": NA 2 NA 2 NA NA NA 2 NA NA ...
##  $ nb_pers     : Factor w/ 4 levels "1","2","3-4",..: 1 2 NA 3 NA NA NA 3 NA NA ...
##  $ parking_diff: Factor w/ 2 levels "No","Yes": 1 1 NA 2 NA NA NA 1 NA NA ...
##  $ retrait     : Factor w/ 2 levels "No","Yes": 1 1 NA 1 NA NA NA 1 NA NA ...
##  $ fullygrouped: Factor w/ 3 levels "City","Montagnes",..: 1 1 1 1 1 1 1 1 1 1 ...
```

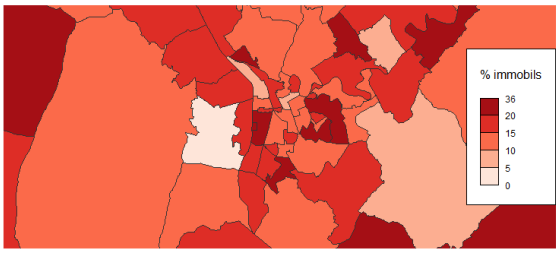
3 Data exploration

3.1 Mapping

Percentage immobile people per zone



Zoomed: Percentage immobile people per zone



4 Modeling

4.1 Logit model

```
##
## Call:
## glm(formula = immobil ~ W + dispovp + age_group, family = binomial,
##     data = allgreI_filtered)
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.3825     1.0614  -0.360    0.719
## W               -1.7194     0.1353 -12.707  < 2e-16 ***
## dispovp         -0.4772     0.1186  -4.023 5.73e-05 ***
## age_group20-60  -0.1243     1.0411  -0.119    0.905
## age_group60+     0.3523     1.0598   0.332    0.740
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 2524.6  on 3235  degrees of freedom
## Residual deviance: 2126.7  on 3231  degrees of freedom
##   (4484 observations deleted due to missingness)
## AIC: 2136.7
##
## Number of Fisher Scoring iterations: 6
```

	Odds_Ratio
(Intercept)	0.682
W	0.179
dispovp	0.621
age_group20-60	0.883
age_group60+	1.422

```
##
## Call:
## glm(formula = immobil ~ has_car + OCCU1 + W, family = binomial,
##     data = allgreI_filtered)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.6446     0.2645   2.437  0.01480 *
## has_carYes   -2.0670     0.2511  -8.232  < 2e-16 ***
## OCCU12        0.7037     0.2571   2.737  0.00620 **
## OCCU13      -11.4040   333.6457  -0.034  0.97273
## OCCU14        0.2661     0.3871   0.687  0.49179
## OCCU15      -11.4040   882.7434  -0.013  0.98969
## OCCU16       -0.9520     0.3564  -2.672  0.00755 **
## OCCU17       -0.8053     0.2461  -3.272  0.00107 **
## W            -1.7396     0.1494 -11.641  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2256.7  on 3059  degrees of freedom
## Residual deviance: 1877.4  on 3051  degrees of freedom
##   (4484 observations deleted due to missingness)
## AIC: 1895.4
##
## Number of Fisher Scoring iterations: 13
```

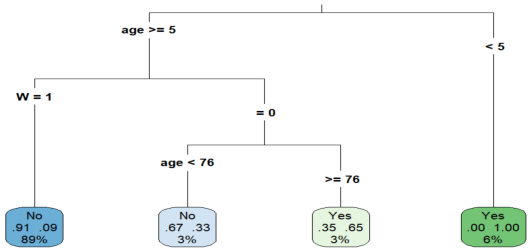
(Intercept)	1.905
has_carYes	0.127
OCCU12	2.021
OCCU13	0.000
OCCU14	1.305
OCCU15	0.000
OCCU16	0.386
OCCU17	0.447
W	0.176

```
## Warning: package 'caret' was built under R version 4.4.2
```

```
## Loading required package: lattice
```

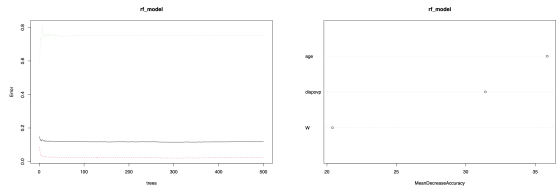
```
##              Metric Logit_Model
## Accuracy   Accuracy       0.884
## Precision Precision       0.525
## Recall       Recall       0.432
## F1         F1-Score       0.474
```

4.2 Decision tree




```
##            used.variables Accuracy Precision Recall F1.Score
## 1 age_group + dispovp + W     0.85      0.00    NaN      NaN
## 2          age + dispo + W     0.90      0.45   0.87     0.59
## 3     has_car + OCCU1 + W     0.85      0.00    NaN      NaN
```

4.3 Random forest



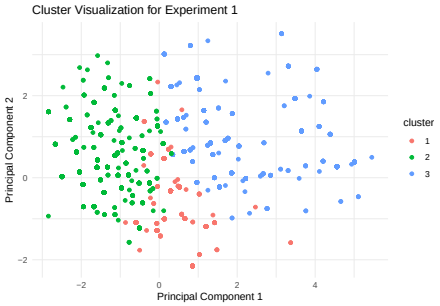

```
## Accuracy: 0.8813025
```

```
## Precision: 0.725
```

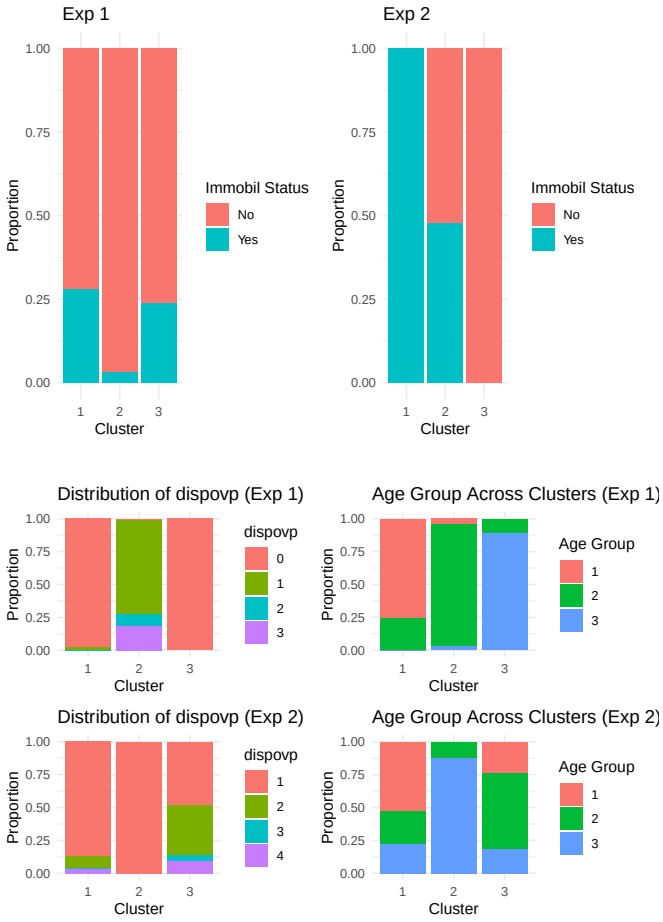
```
## Recall: 0.221374
```

```
## F1-Score: 0.3391813
```

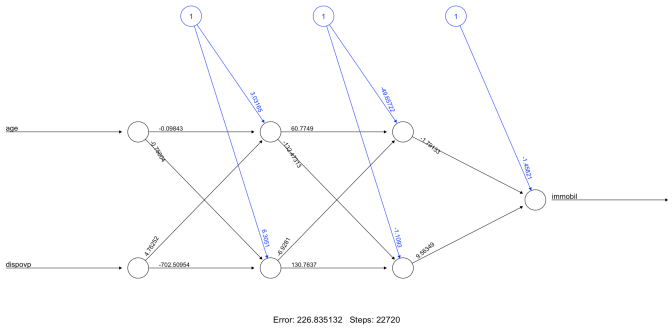
4.4 Clustering




```
## Warning: package 'patchwork' was built under R version 4.4.2
```



4.5 Neural network




```
## Comparison:
```

```
## NN1 Accuracy: 0.9  | NN2 Accuracy: 0.954
```

```
## NN1 Precision: 0.926  | NN2 Precision: 0.7
```

```
## NN1 Recall: 0.384  | NN2 Recall: 0.233
```

```
## NN1 F1 Score: 0.543  | NN2 F1 Score: 0.35
```

5 Results

```
##             Model Accuracy Precision Recall F_score
## 1     Logit Model    0.884     0.525  0.432   0.474
## 2   Decision tree    0.900     0.450  0.870   0.590
## 3   Random Forest    0.881     0.725  0.221   0.339
## 4 Neural Network1    0.900     0.926  0.384   0.543
## 5 Neural Network2    0.954     0.700  0.233   0.350
```