



# Understanding attitudes, experience and risk factors for domestic violence

International Team of Scientists working at JAMSTEC, Japan



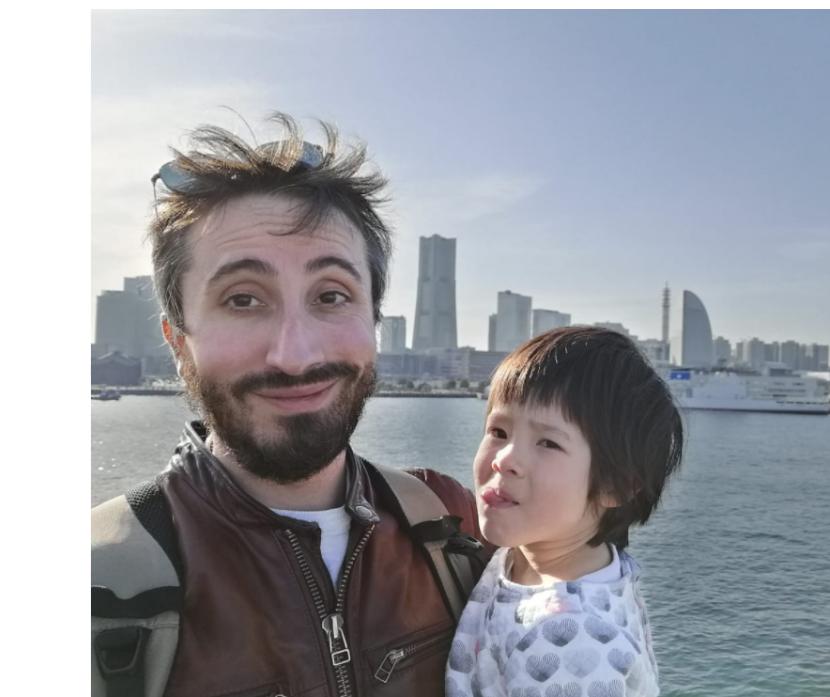
Prima Anugerahanti  
Indonesia



Gerlien Verhaegen  
Belgium



Heather Ritchie  
Scotland



Arthur Bauville  
France



Ettore Barbieri  
Italy

The Demographic and Health Surveys (DHS) Program has collected, analyzed, and disseminated accurate and representative data on population, health, HIV, and nutrition through more than **400 surveys in over 90 countries**.



## Using data provided by the DHS Program we wanted to:

- ▶ Investigate the **acceptance vs the experience** of domestic violence
- ▶ Determine the **controlling factors** of domestic violence
- ▶ Develop a tool that would **quickly and efficiently identify women who are at risk** of domestic violence

The Demographic and Health Surveys (DHS) Program has collected, analyzed, and disseminated accurate and representative data on population, health, HIV, and nutrition through more than 400 surveys in over 90 countries.



A mother and daughters in Jimma, Ethiopia work with coffee beans after their house has received Indoor Residual Spraying (IRS) to reduce malaria transmission. Photo Credit: ALRS Ethiopia PMI

## DHS Program data for acceptance and experience of domestic violence

► Acceptance: % of yes to at least one answer of:

It is the respondent's opinion that a husband is justified in hitting or beating his wife when:

**BASE:** For V744A to V744E is all women.

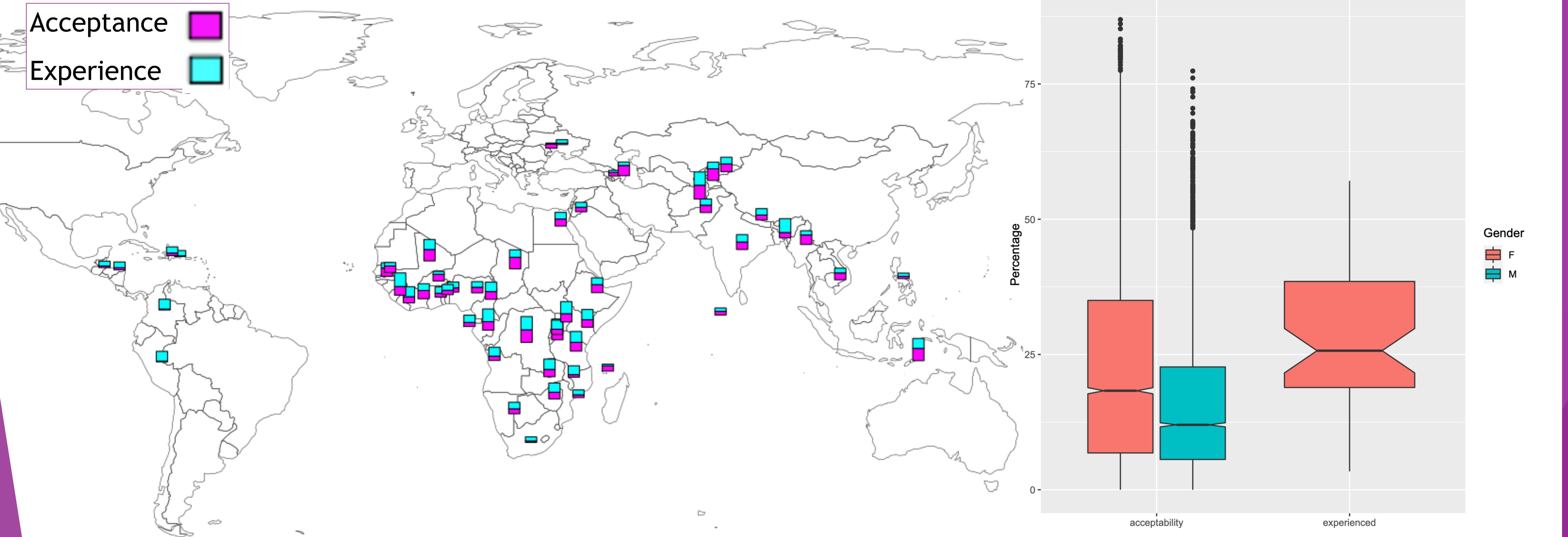
- V744A She goes out without telling him
- V744B She neglects the children
- V744C She argues with him
- V744D She refuses to have sex with him

► Experience: % of yes to at least one answer of:

### Physical violence

- BASE:** Women currently in union ( $V502 = 1$ ).  
Spouse ever pushed, shook or threw something.  
Spouse ever slapped.  
Spouse ever punched with fist or something harmful.  
Spouse ever kicked or dragged.  
Spouse ever tried to strangle or burn.  
Spouse ever threatened with knife/gun or other weapon.  
Spouse ever attacked with knife/gun or other weapon.  
Spouse ever physically forced sex when not wanted.  
Spouse ever forced other sexual acts when not wanted.  
Spouse ever twisted her arm or pulled her hair.
- D105A Physical violence
- D105B Physical violence
- D105C Physical violence
- D105D Physical violence
- D105E Physical violence
- D105F Physical violence
- D105G Physical violence
- D105H Physical violence
- D105I Physical violence
- D105J Physical violence
- D105K Physical violence
- D105L Physical violence
- D105M Physical violence
- D105N Physical violence
- D106 Experienced any less severe violence.
- D107 Experienced any severe violence.
- D108 Experienced any sexual violence.

# Acceptance vs experience of domestic violence

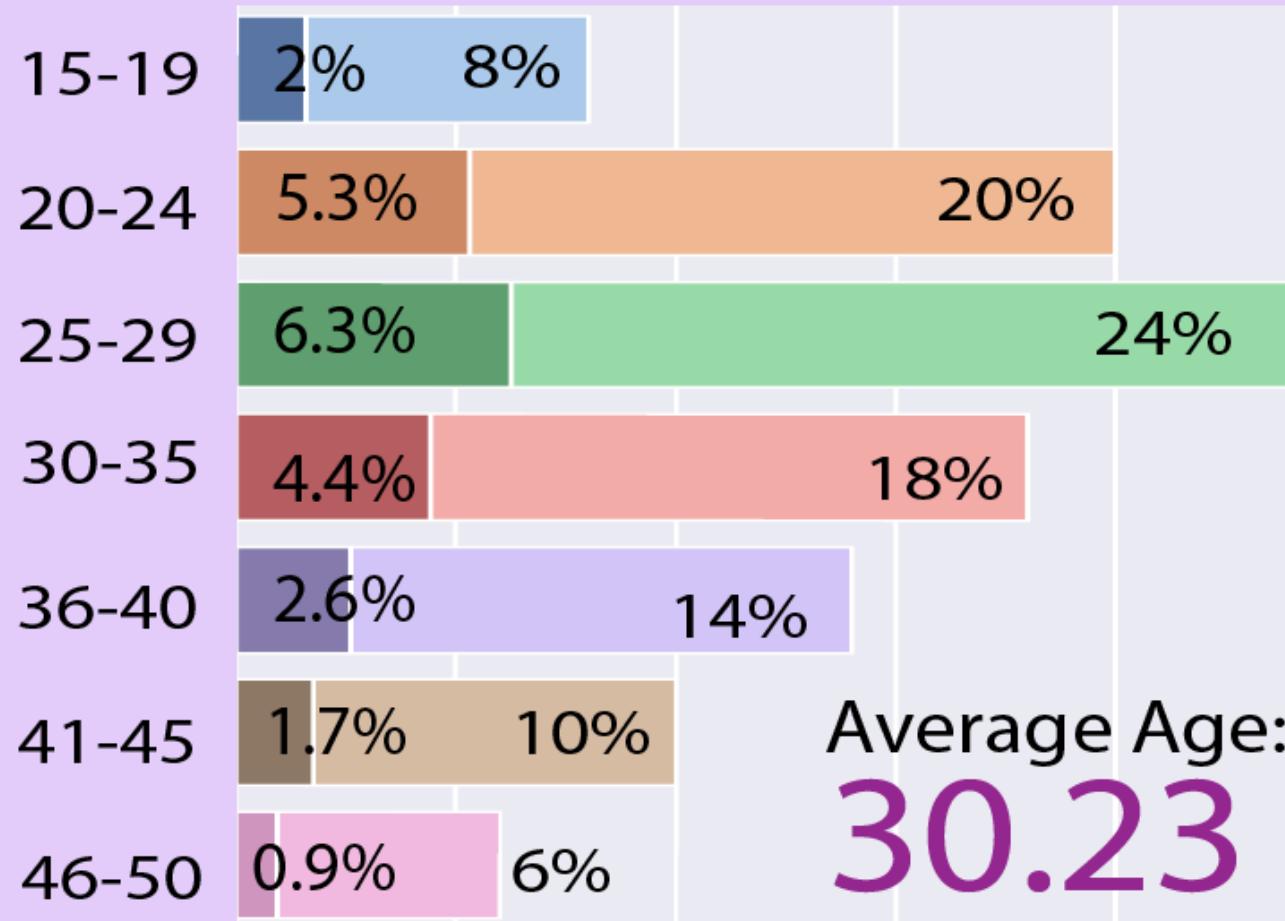


- ▶ Domestic violence is **widespread** amongst the countries surveyed by the DHS
- ▶ On average across all countries **26% of women** surveyed have experienced physical and/or sexual domestic violence
- ▶ Overall **more women than men** agreed that domestic violence was acceptable in certain circumstance

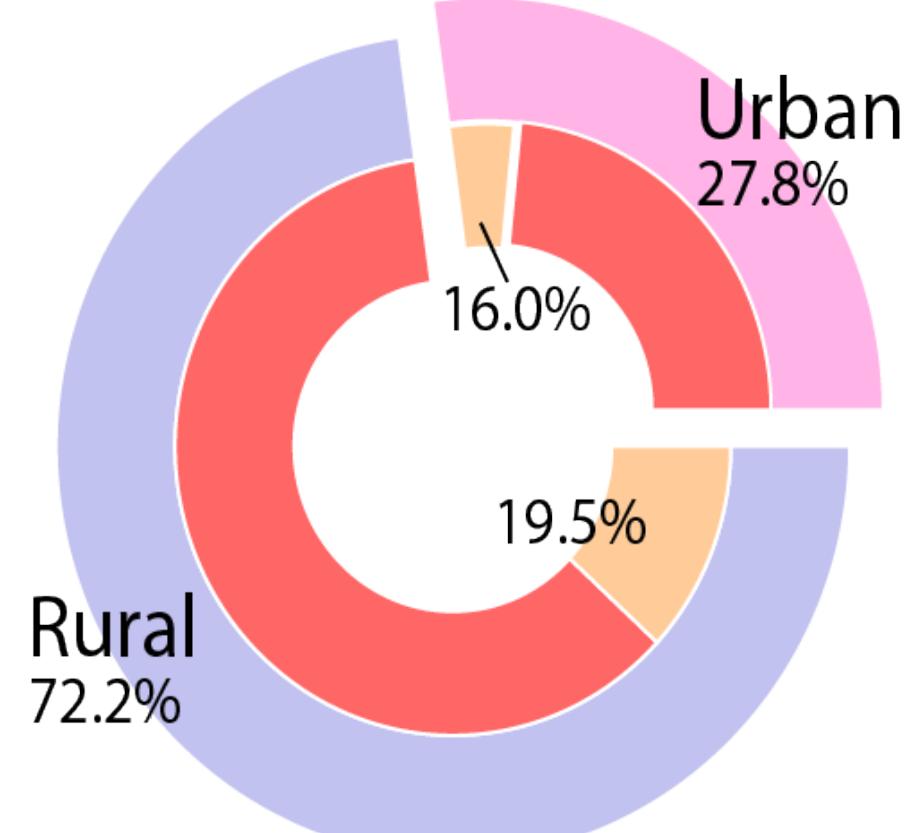
# Demographic of Respondents

5120 female and married respondents in Congo between 2013-2014

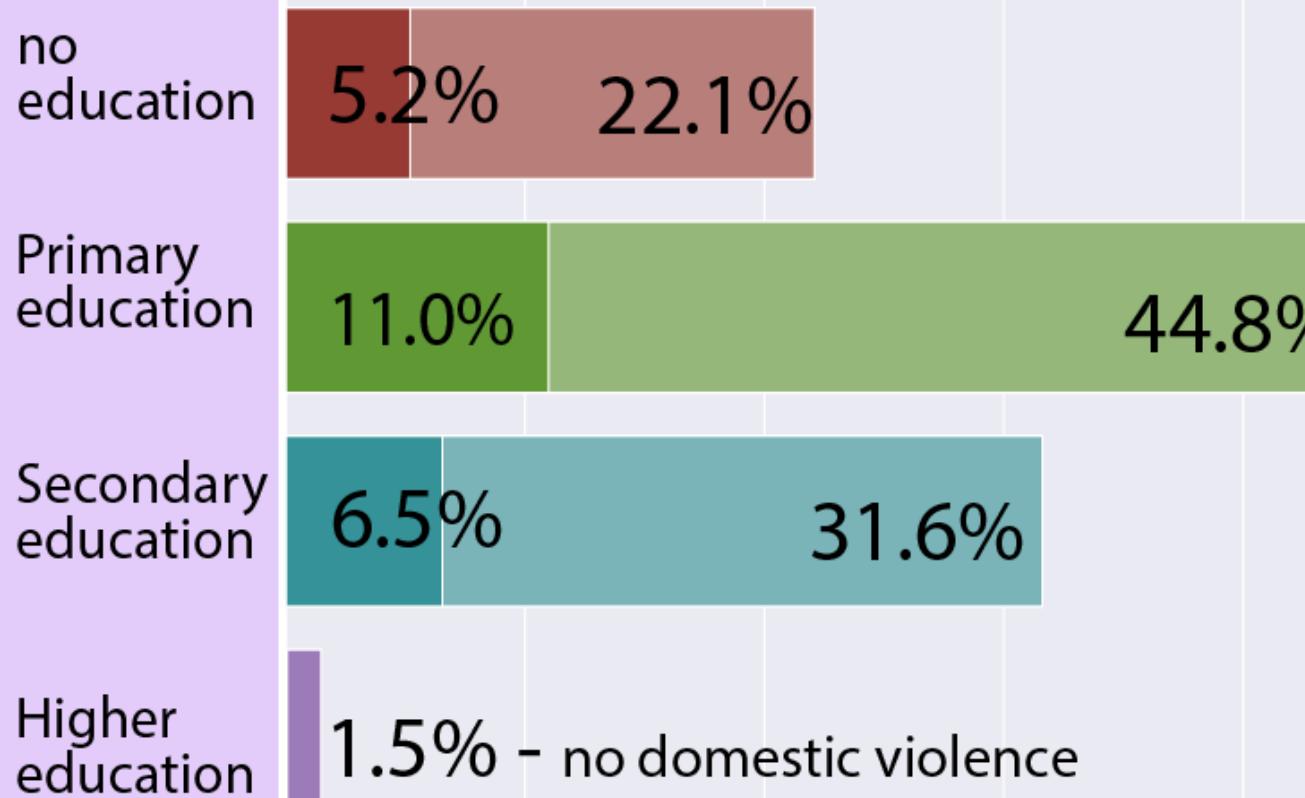
## Age



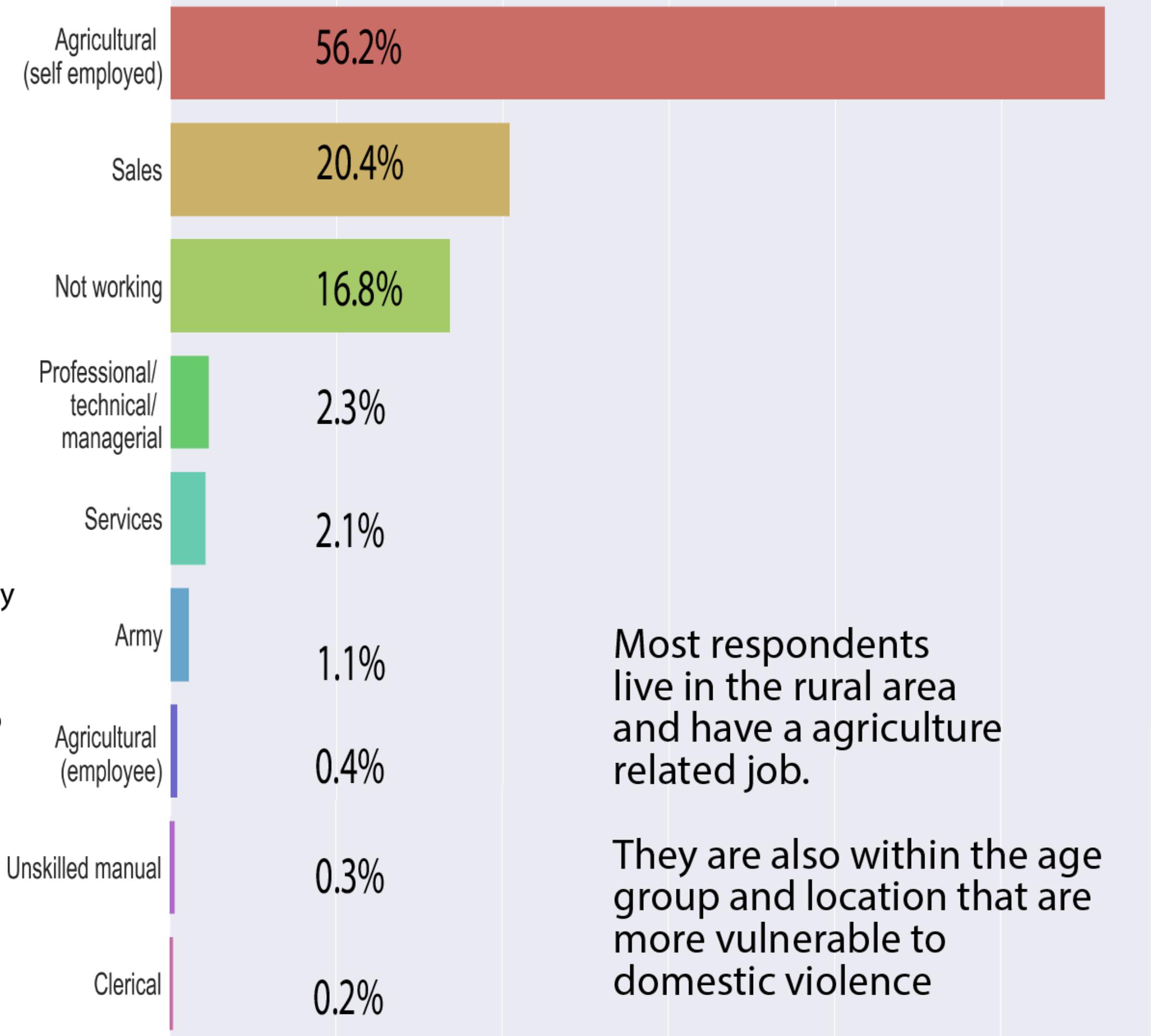
## Location



## Education

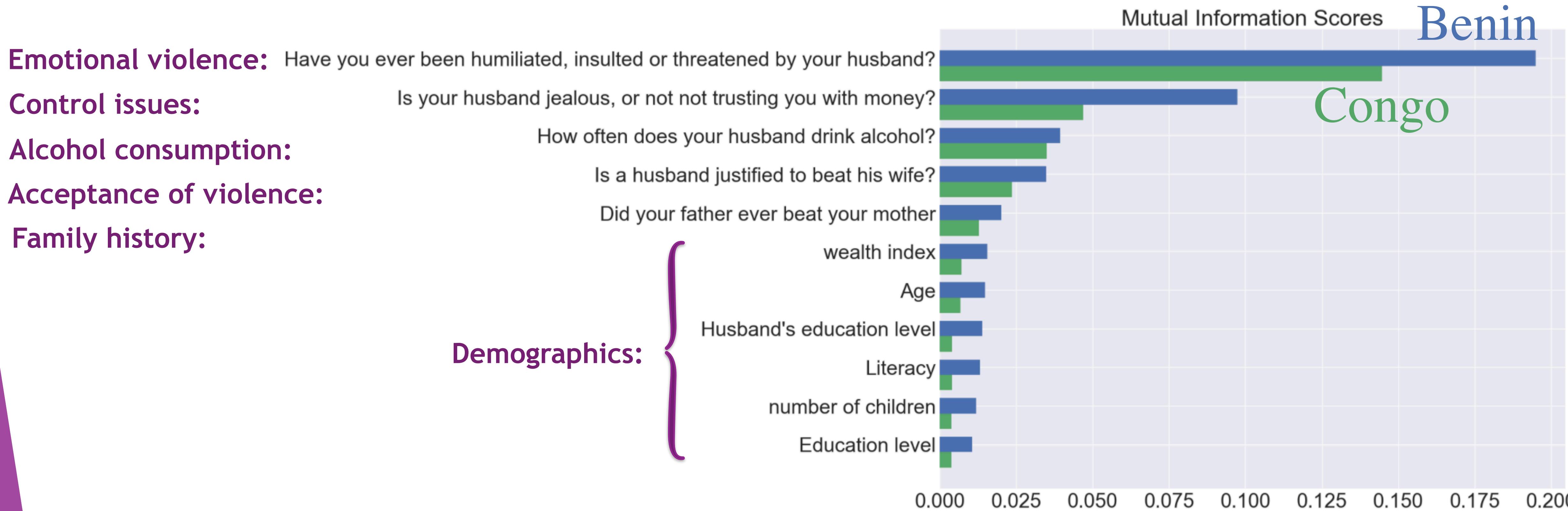


## Occupation



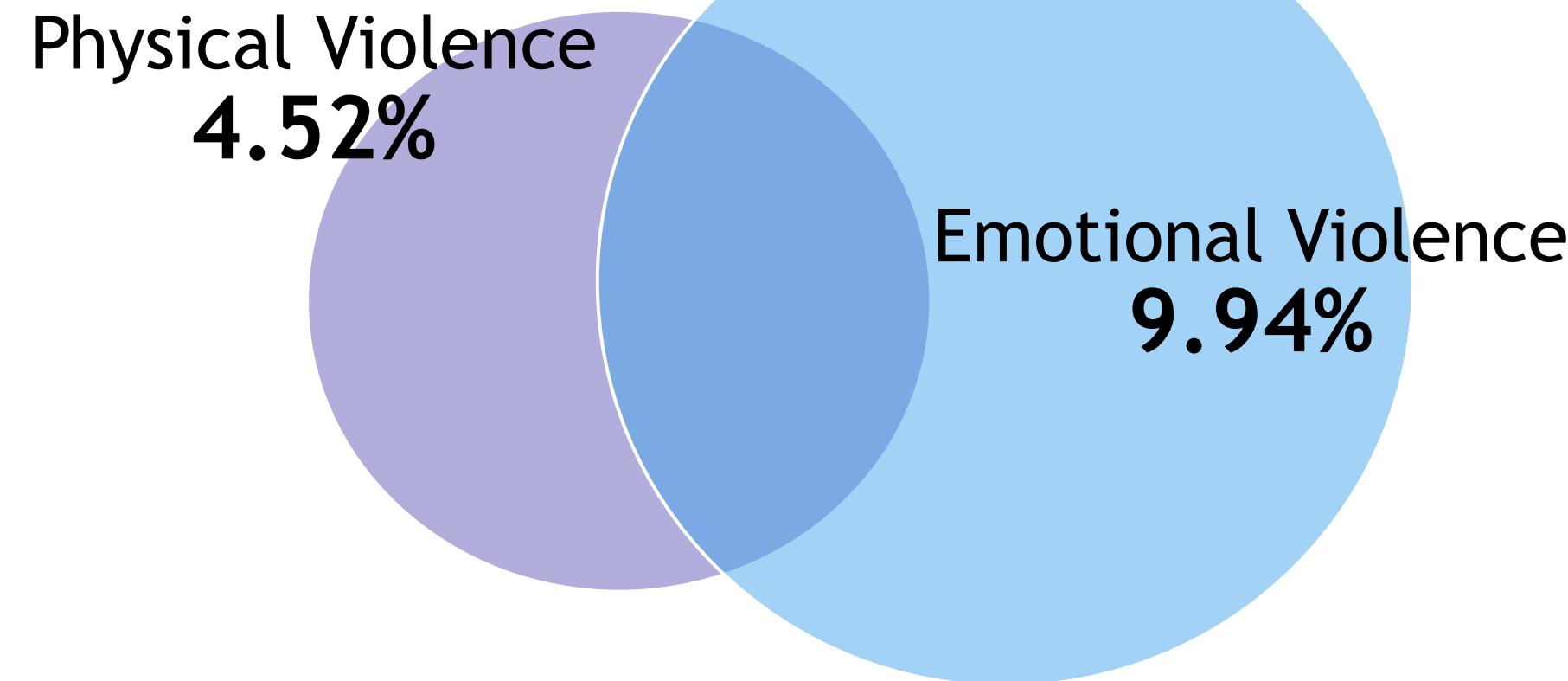
# From which information can we predict answers to the question:

► “Have you experienced physical or domestic violence in the last 12 months?”

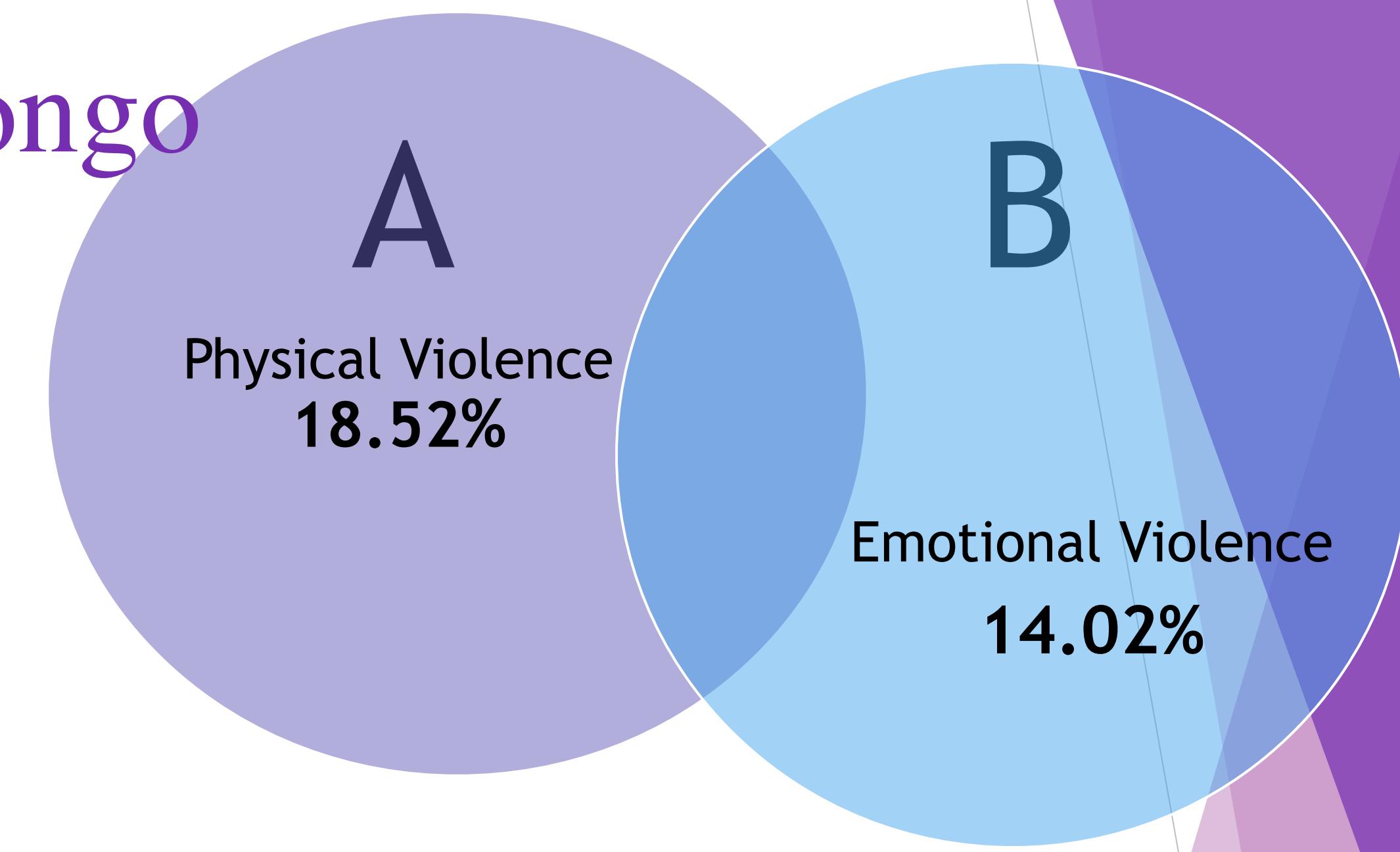


# Is emotional violence predictive of physical/sexual violence

Benin



Congo



$$P(B | A)$$

$$\begin{cases} \text{(Benin)} & 63\% \\ \text{(Congo)} & 51\% \end{cases}$$

} of women experiencing **physical violence** also experience **emotional violence**

$$P(A | B)$$

$$\begin{cases} \text{(Benin)} & 29\% \\ \text{(Congo)} & 67\% \end{cases}$$

} of women experiencing **emotional violence** also experience **physical violence**

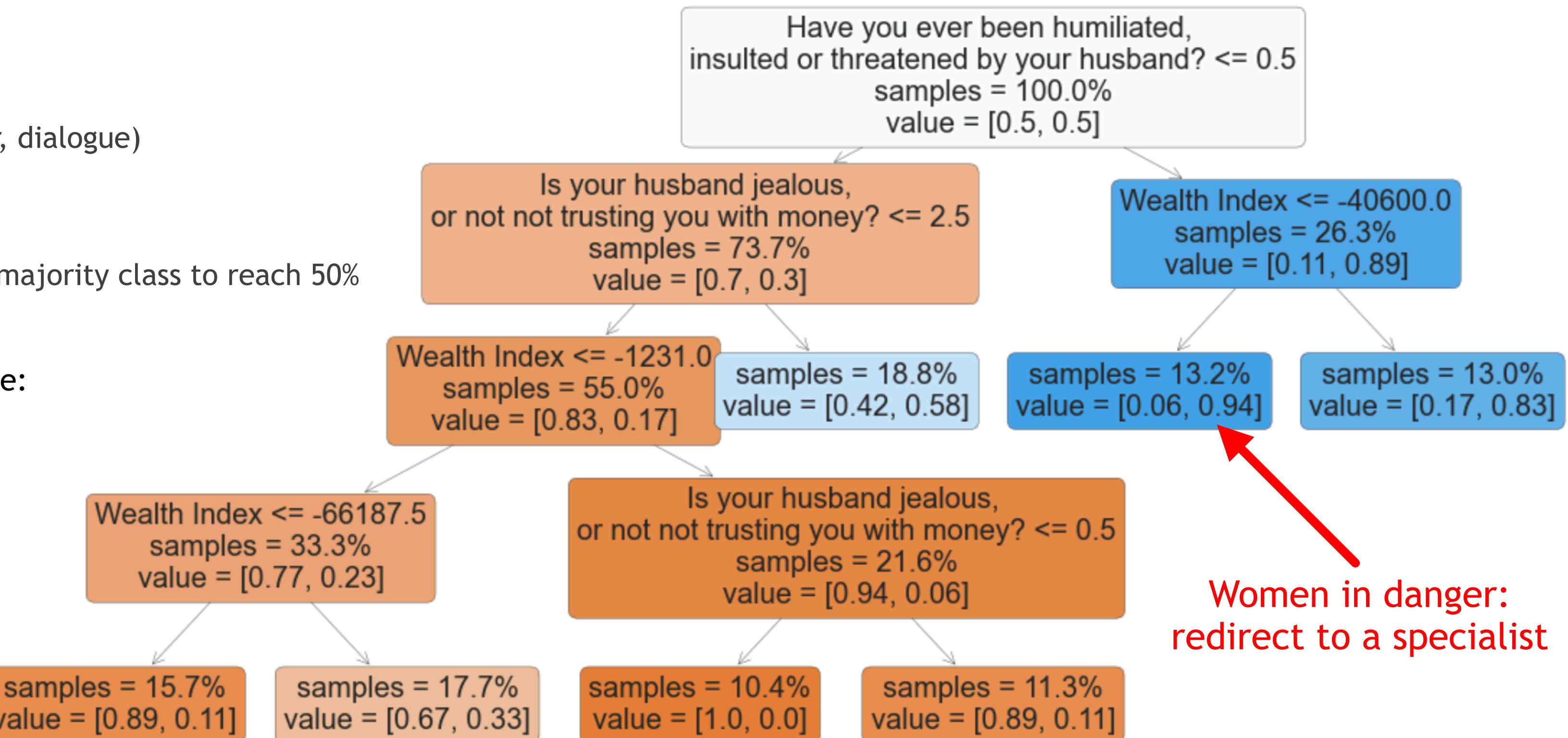
In both countries emotional and physical violence are correlated but emotional violence seems to be a precursor of physical/sexual violence only in Benin

# How well can we identify women at risk of physical violence?

- ▶ Binary classification tree: doesn't require a computer (paper, dialogue)
- ▶ True: this woman is experiencing violence
- ▶ Since the dataset is initially imbalanced we subsampled the majority class to reach 50% False, 50% True
- ▶ Hyperparameters were tuned to optimize **recall** because:

$$\text{recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

- ▶ True positives are women in danger (experiencing physical or sexual violence)
- ▶ False negatives are women in danger but not detected (**should be minimized**)
- ▶ False positives are women who are not experiencing physical violence but may in the future



Women in danger:  
redirect to a specialist

- ▶ The model can be used to:
  - ▶ design a personalized questionnaire
  - ▶ quantitatively interpret the results of a general questionnaire
  - ▶ Raise awareness in the general public





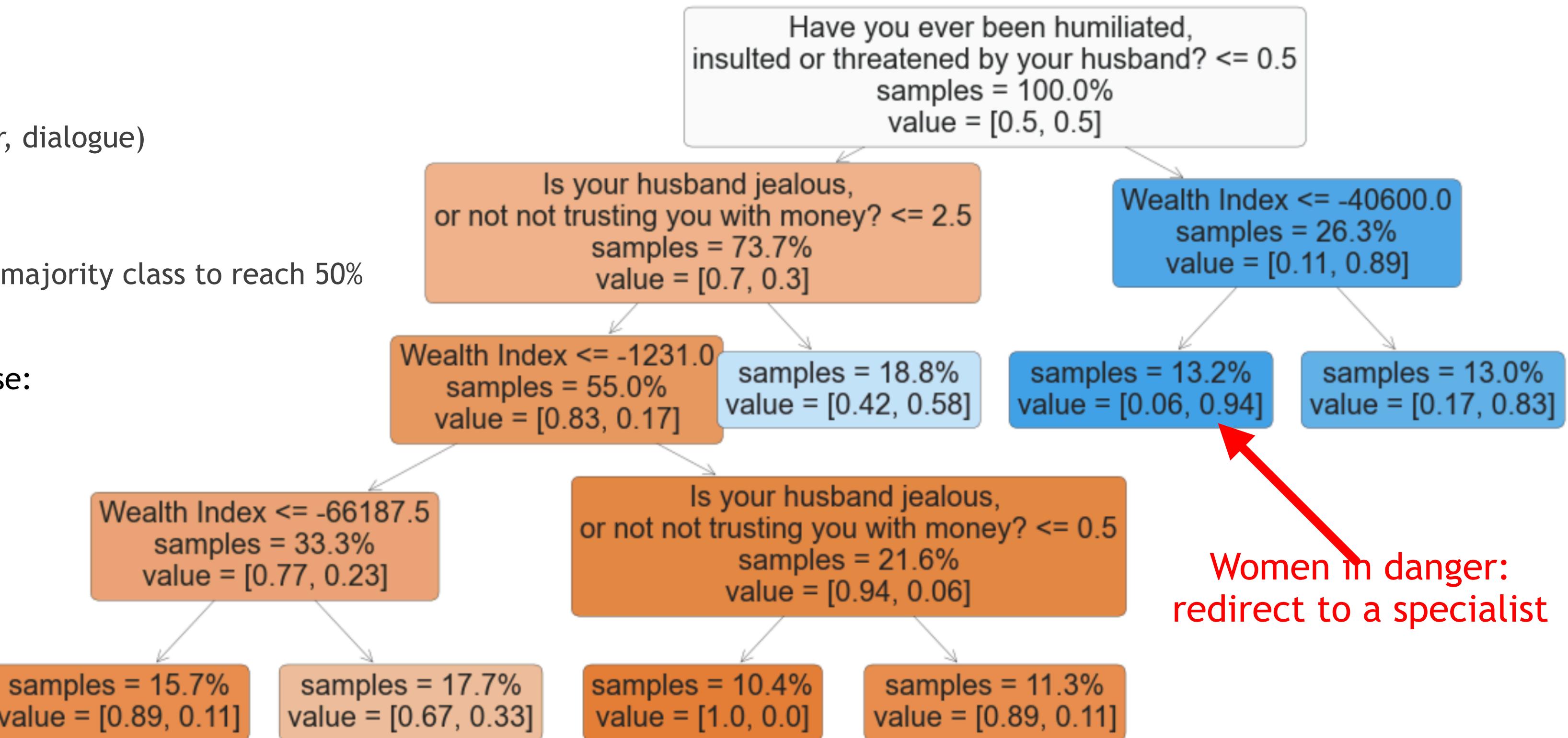
# Additional information

# How well can we identify women at risk of physical violence?

- ▶ Binary classification tree: doesn't require a computer (paper, dialogue)
- ▶ True: this woman is experiencing violence
- ▶ Since the dataset is initially imbalanced we subsampled the majority class to reach 50% False, 50% True
- ▶ Hyperparameters were tuned to optimize **recall** because:

$$\text{recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

- ▶ True positives are women in danger (experiencing physical or sexual violence)
- ▶ False negatives are women in danger but not detected (**should be minimized**)
- ▶ False positives are women who are not experiencing physical violence but may in the future



## Model's performance

Actual	Predicted label		Metrics				
	False	True	precision	recall	f1-score	support	
False	38%	10%					
True	13%	38%	False True	0.74 0.78	0.78 0.74	0.76 0.76	37 39

- ▶ The model can be used to:

- ▶ design a personalized questionnaire
- ▶ quantitatively interpret the results of a general questionnaire
- ▶ Raise awareness in the general public

```

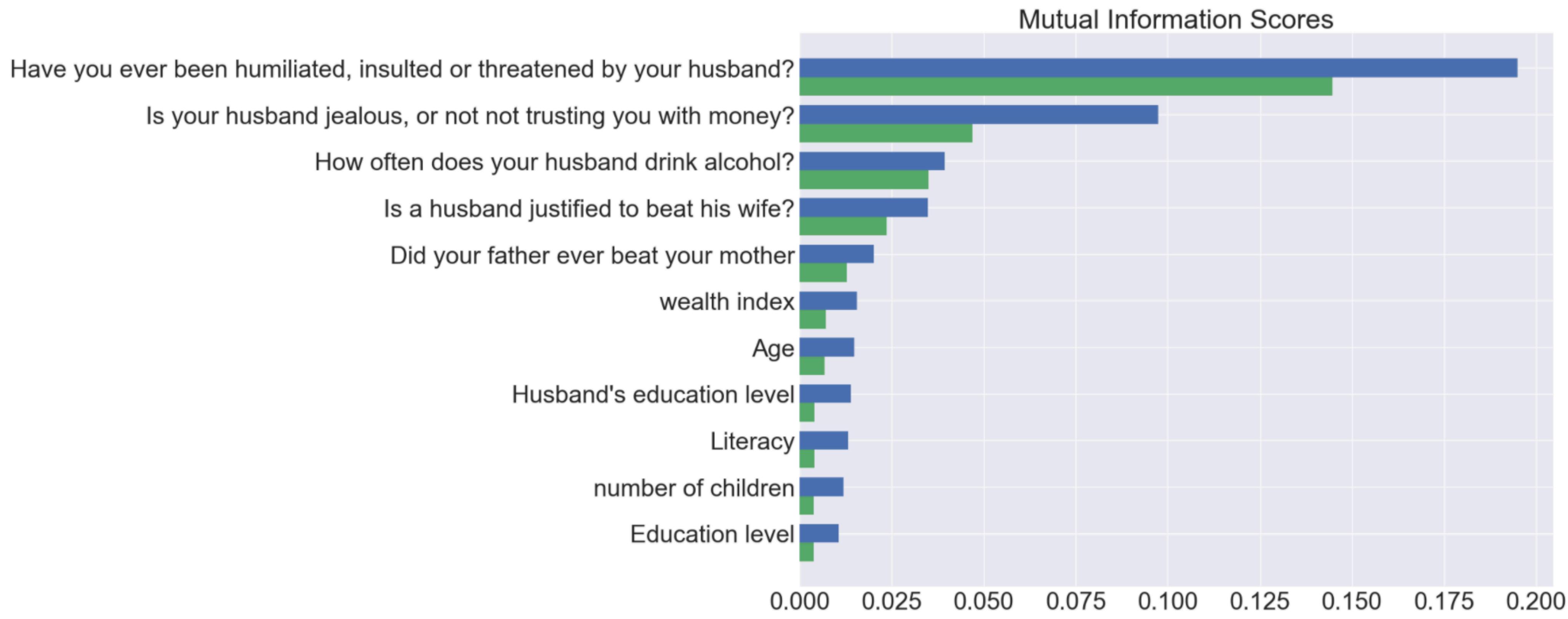
for itarget, target_feature in enumerate([physical_violence]):
    y = X[target_feature + '_sum']>0

    plt.subplot(1,2,itarget+1)
    this_X = X.drop(columns=[target_feature + "_sum"] )
    discrete_features = this_X.dtypes == 'int8'
    mi_scores = pd.Series(np.zeros(this_X.columns.shape), index= this_X.columns)
    n = 20
    for i in range(n):
        X_sub, y_sub = subsample(this_X,y)

        mi_scores += 1.0/n*make_mi_scores(X_sub, y_sub, discrete_features)

mi_scores.rename(dict(zip(X.columns,explicit_feature_names)), inplace=True)
plot_mi_scores(mi_scores)

```



# Model code

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import GradientBoostingClassifier, RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn import tree
from sklearn.metrics import confusion_matrix, classification_report
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import f1_score, roc_auc_score
y = X[physical_violence + "_sum"]>0
this_X = X.drop(columns=[physical_violence + "_sum"])
this_X.columns
discrete_features = this_X.dtypes == 'int8'
X_model = pd.get_dummies(this_X, columns=this_X.columns[discrete_features])
imbalance_fac = 1
X_model, y = subsample(X_model,y, imbalance_fac=imbalance_fac)
X_train, X_test, y_train, y_test = train_test_split(X_model.values,y.values, test_size=0.2)

param_grid = dict(max_depth=[9],
                  min_samples_leaf=[0.05, 0.07, 0.1, 0.15, 0.2, 0.25, .3, 0.5],
                  )

clf = DecisionTreeClassifier(min_impurity_decrease=0.0001,
                             class_weight={0:1, 1:imbalance_fac},
                             )

clf_cv = GridSearchCV(clf, param_grid, cv=3, scoring='recall')

clf_cv.fit(X_train, y_train)
display(clf_cv.best_params_)
y_pred = clf_cv.best_estimator_.predict(X_test)

print(confusion_matrix(y_test, y_pred)/y_test.shape[0])
classif_dict = classification_report(y_test,y_pred, output_dict=True)
print("f1_score:", classif_dict['True']['f1-score'])
print(classification_report(y_test,y_pred))
```

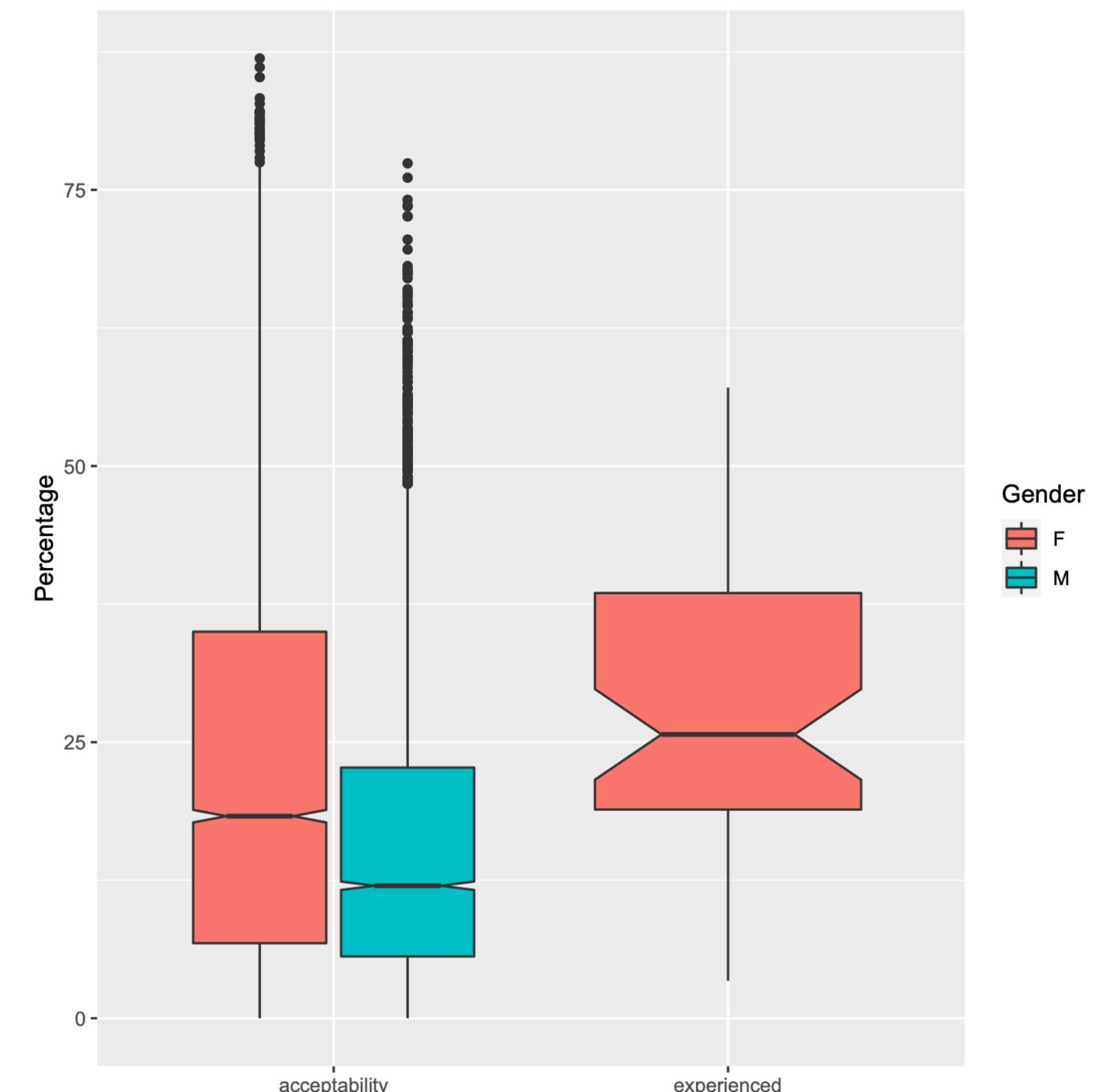
# Can we use these data to predict which women are at most risk of future violence?



# Code

```
dat <- read.csv("acceptability_vs_experienced_vs_gender.csv")
library(ggplot2)

ggplot(dat, aes(x = Category, y = Value))+
  geom_boxplot(aes(fill = Gender),notch=TRUE,las=2)+
  labs(y="Percentage", x="")
```



# Code

```
library(rworldmap)  
  
mapBars( dat, nameX="lon", nameY="lat",  
         nameZs=c('acceptability','experienced'), barWidth = 3,  
         borderCol= "grey17", zColours = c("magenta","cyan"))
```

