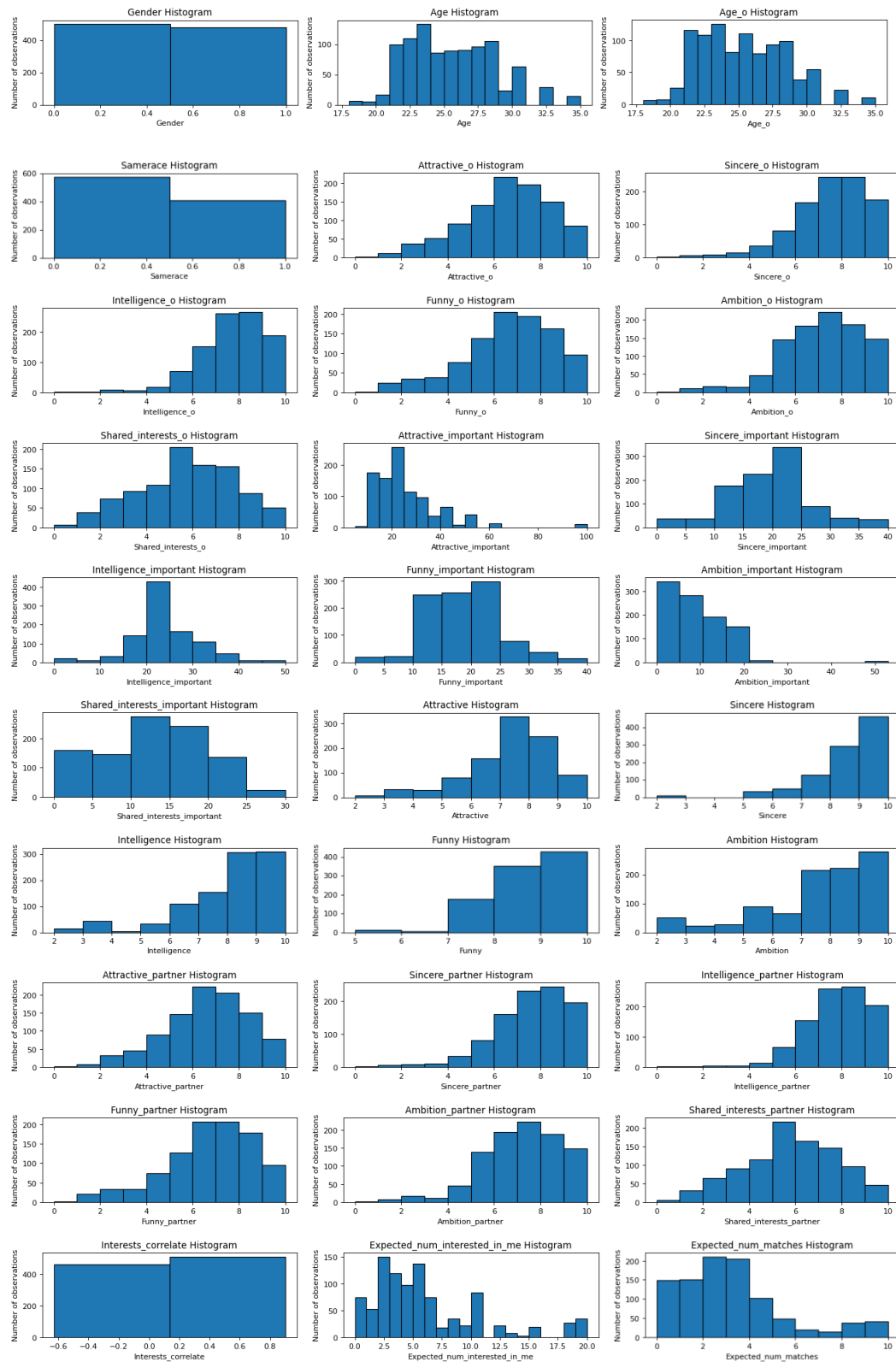
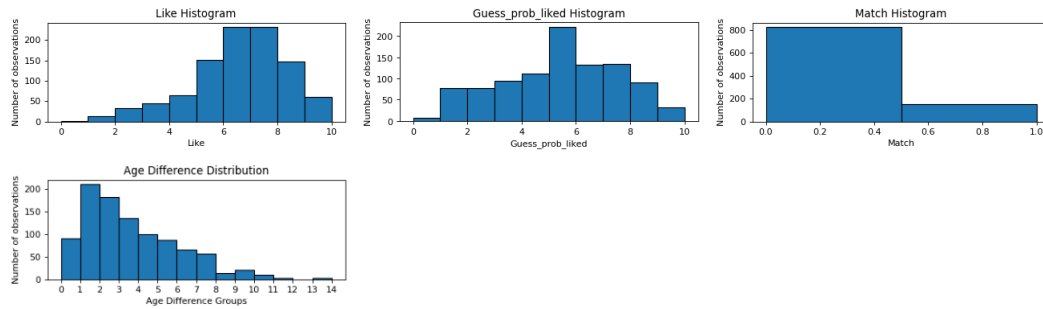


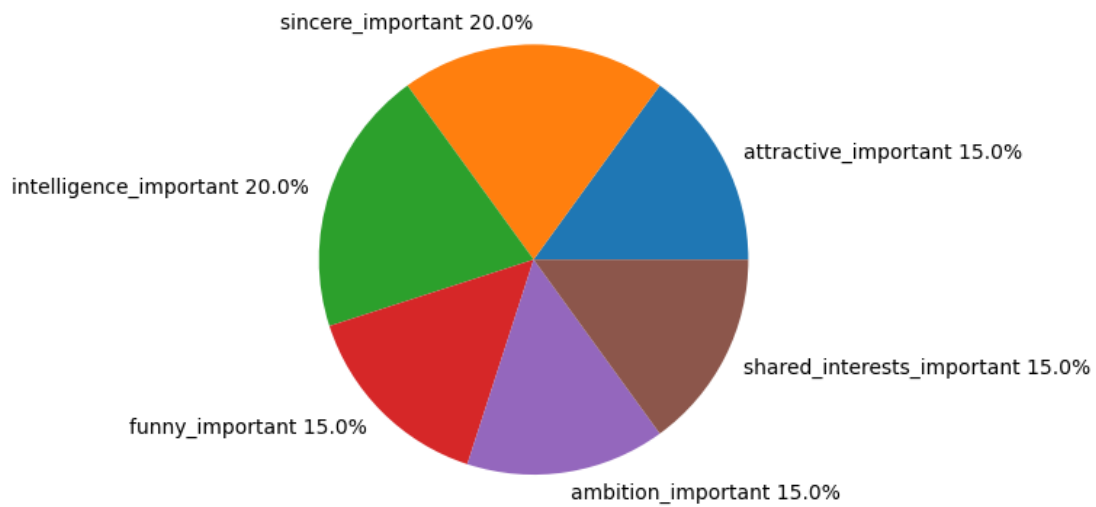
# 1. Histograms of numerical variables





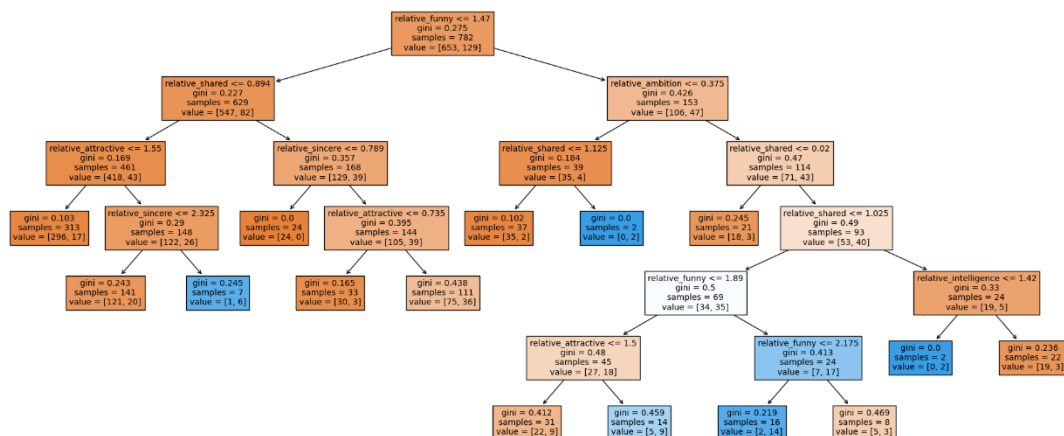
There are 34 numeric variables.

## 2. The pie chart for the importance of each quality of the first participant



The first participant wants her partner has 6 characteristics, especially in sincerity, intelligence, occupying 20% per characteristic. Other 4 characteristics occupy 15% individually.

## 3. Prediction model (decision tree) using relative quality scores.



The root node at the top starts by checking if the "relative\_funny" feature is greater than or equal to 1.47. Each internal node represents a decision point, where the algorithm evaluates a specific feature and determines the next step based on the outcome.

#### 4. Evaluation of the decision tree:

Accuracy: 0.86

Confusion Matrix:

```
[[165  5]
```

```
[ 23  3]]
```

Classification Report:

	precision	recall	f1-score	support
False	0.88	0.97	0.92	170
True	0.38	0.12	0.18	26
accuracy			0.86	196
macro avg	0.63	0.54	0.55	196
weighted avg	0.81	0.86	0.82	196

True positive: 165, False positive: 5, False negative: 23, True Negative: 3

The accuracy of this decision tree is 0.86.