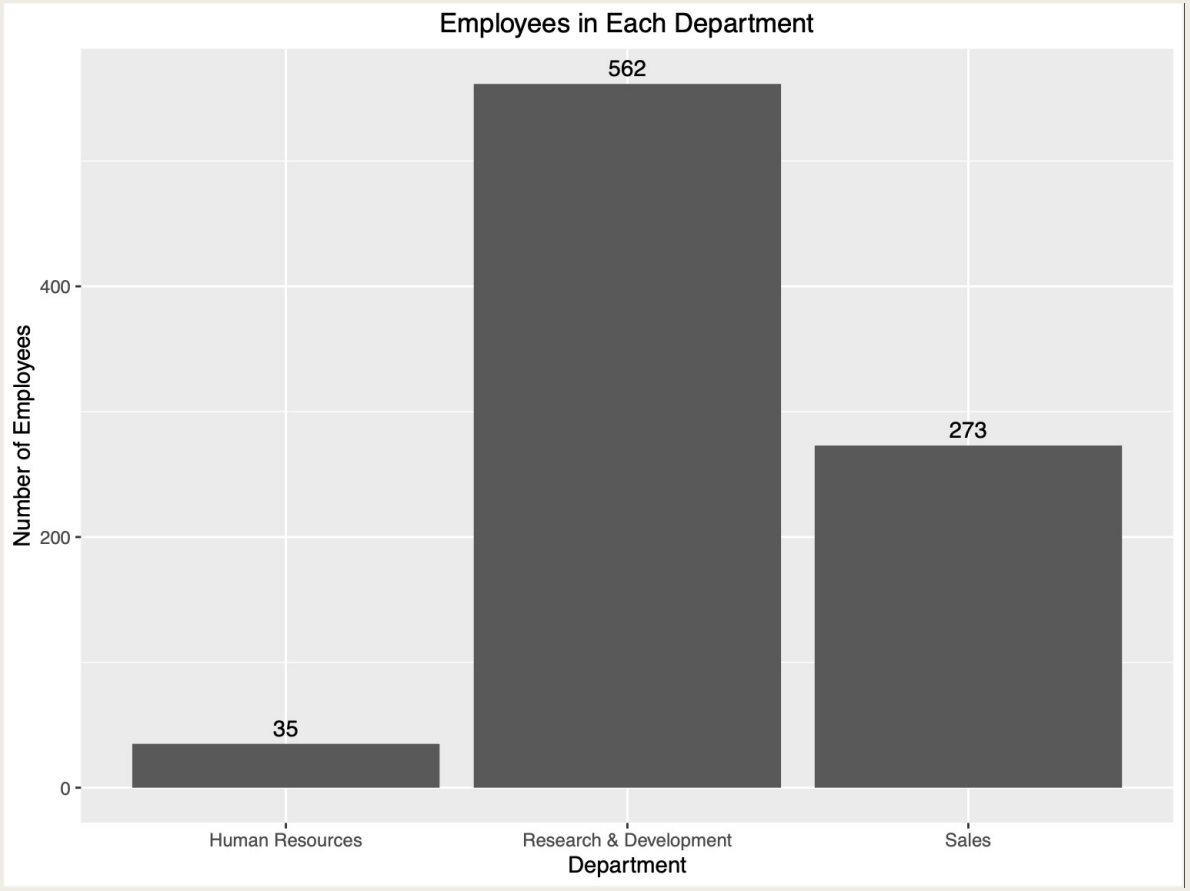
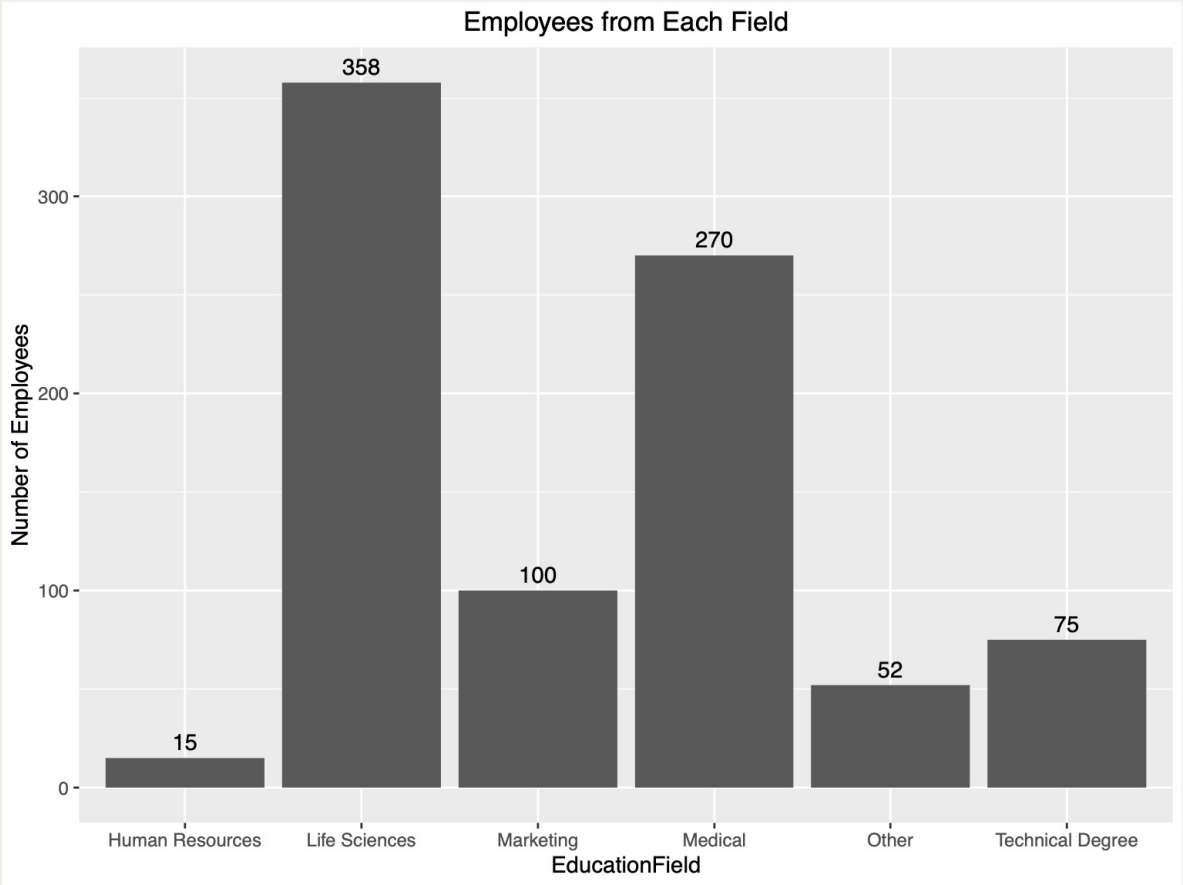
CASE STUDY 2

Gowtham Katta

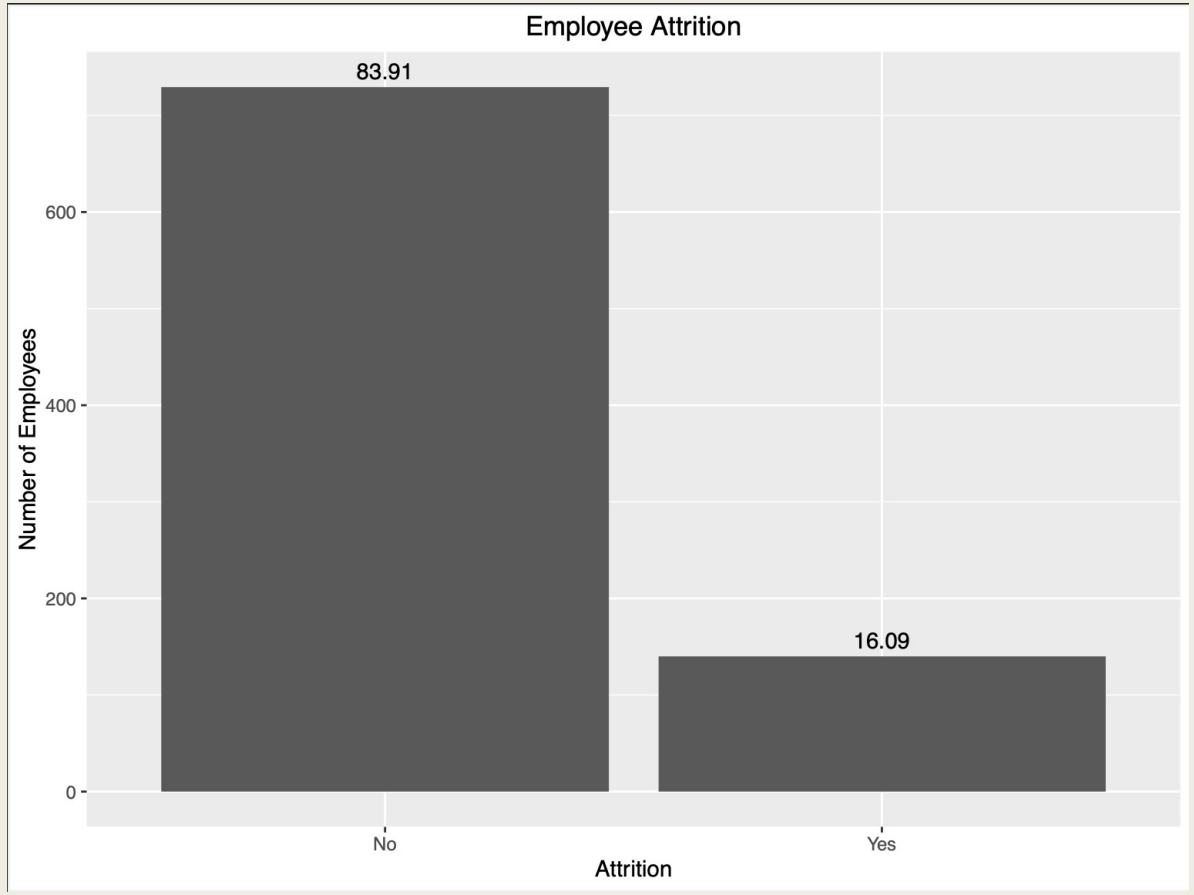
About the Company



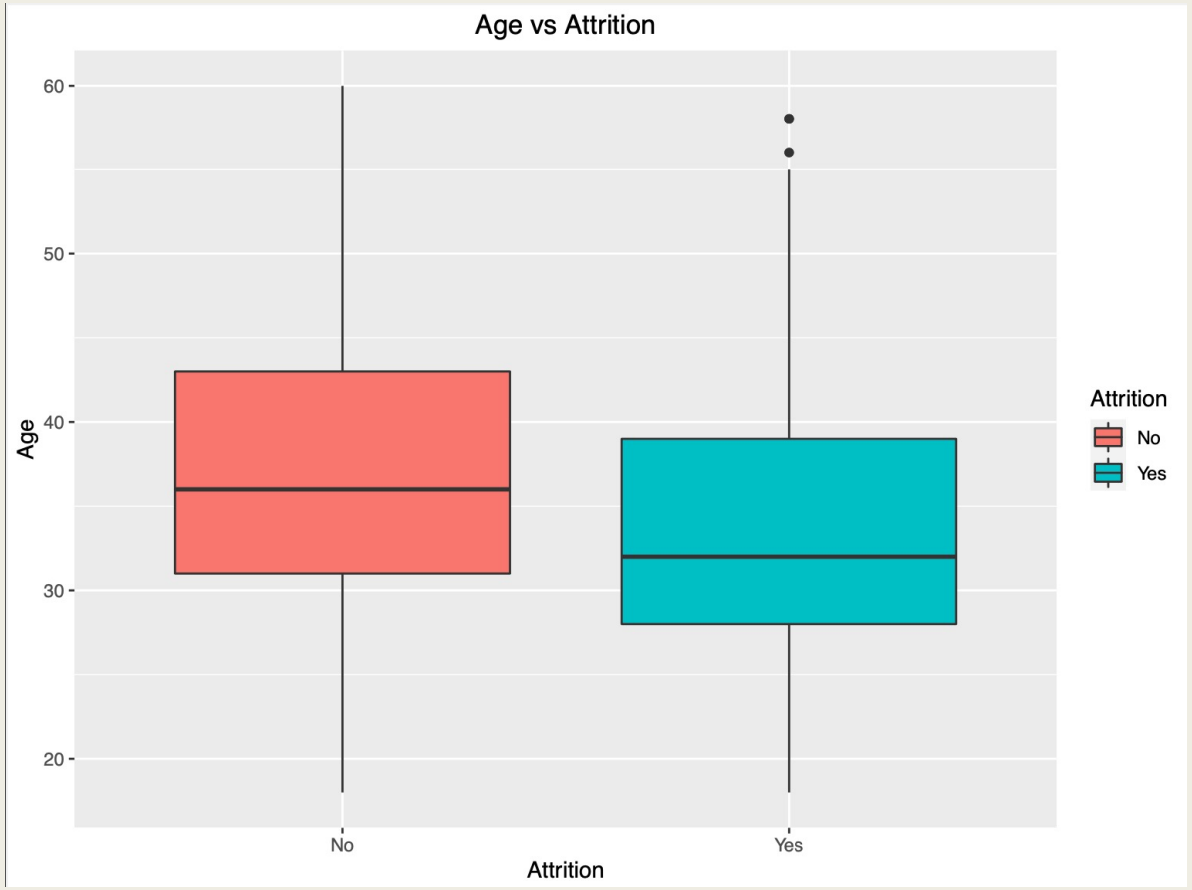
About the Company (cont.)



What factors affect employee attrition?



Factor 1: Age



Age of No Attrition

- Median: 36

Vs.

Age of Yes Attrition

- Median: 32

Factor 2: Monthly Income

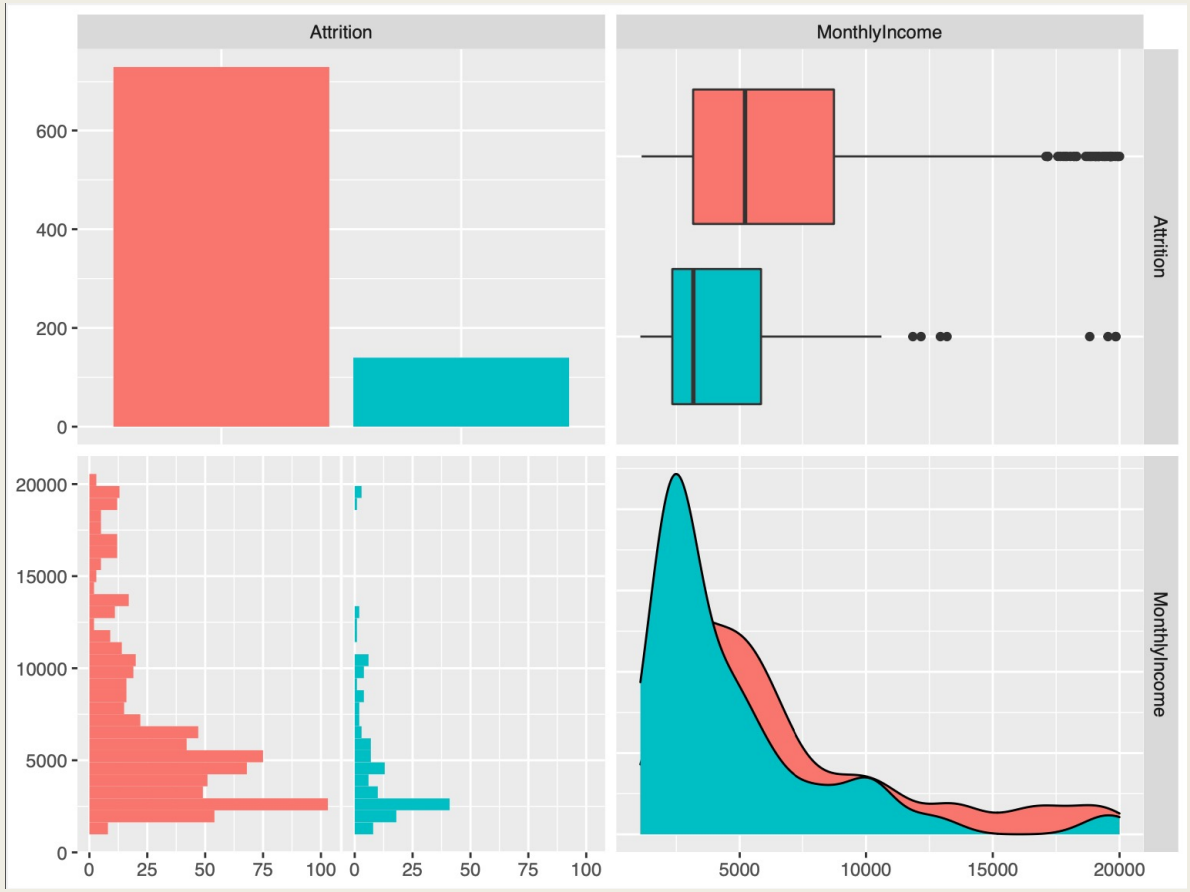
Monthly Income of No Attrition:

- Median: 5208
- Average: 6702

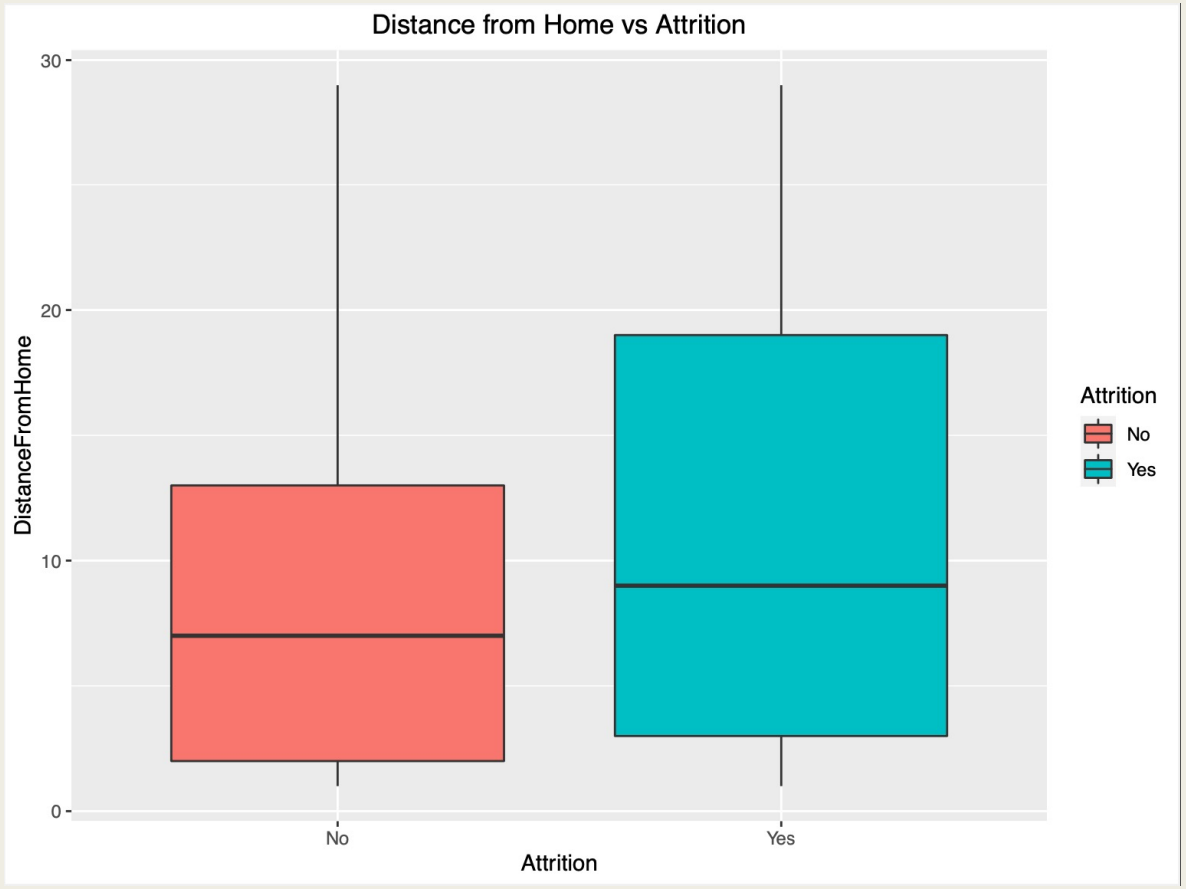
Vs.

Monthly Income of Yes Attrition:

- Median: 3171
- Average: 4765



Factor 3: Distance from Home



Distance from Home of No Attrition

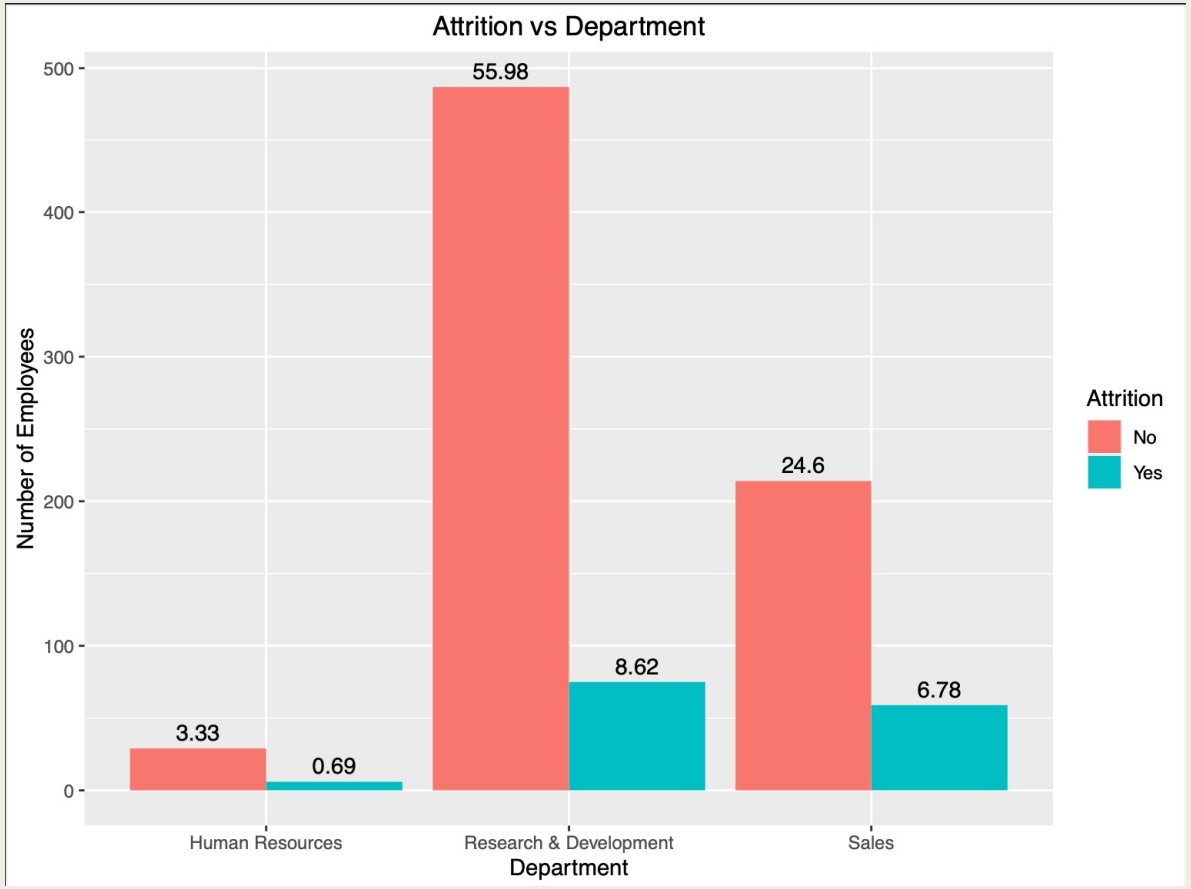
- Median: 9

Vs.

Distance from Home of Yes Attrition

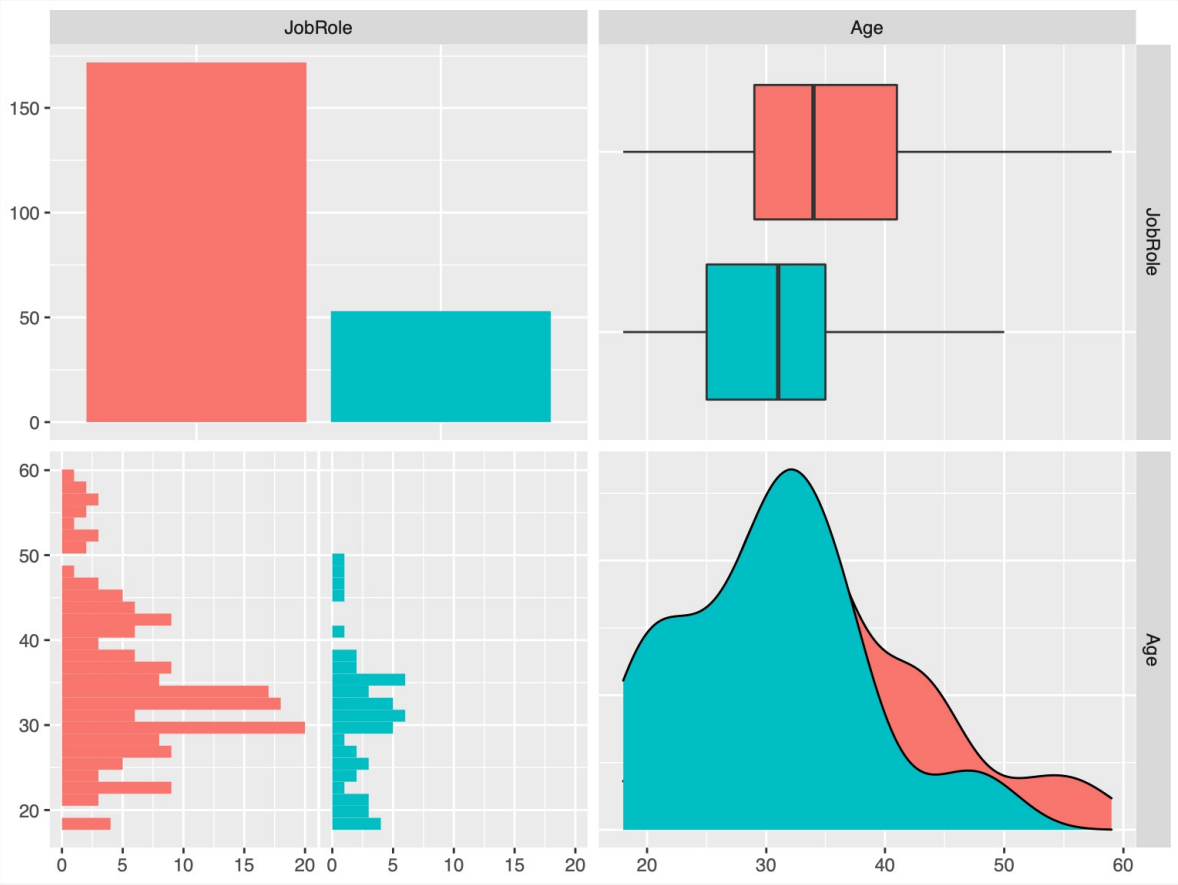
- Median: 7

Another Factor could be...



When comparing departments, the Sales department has the highest attrition rate!

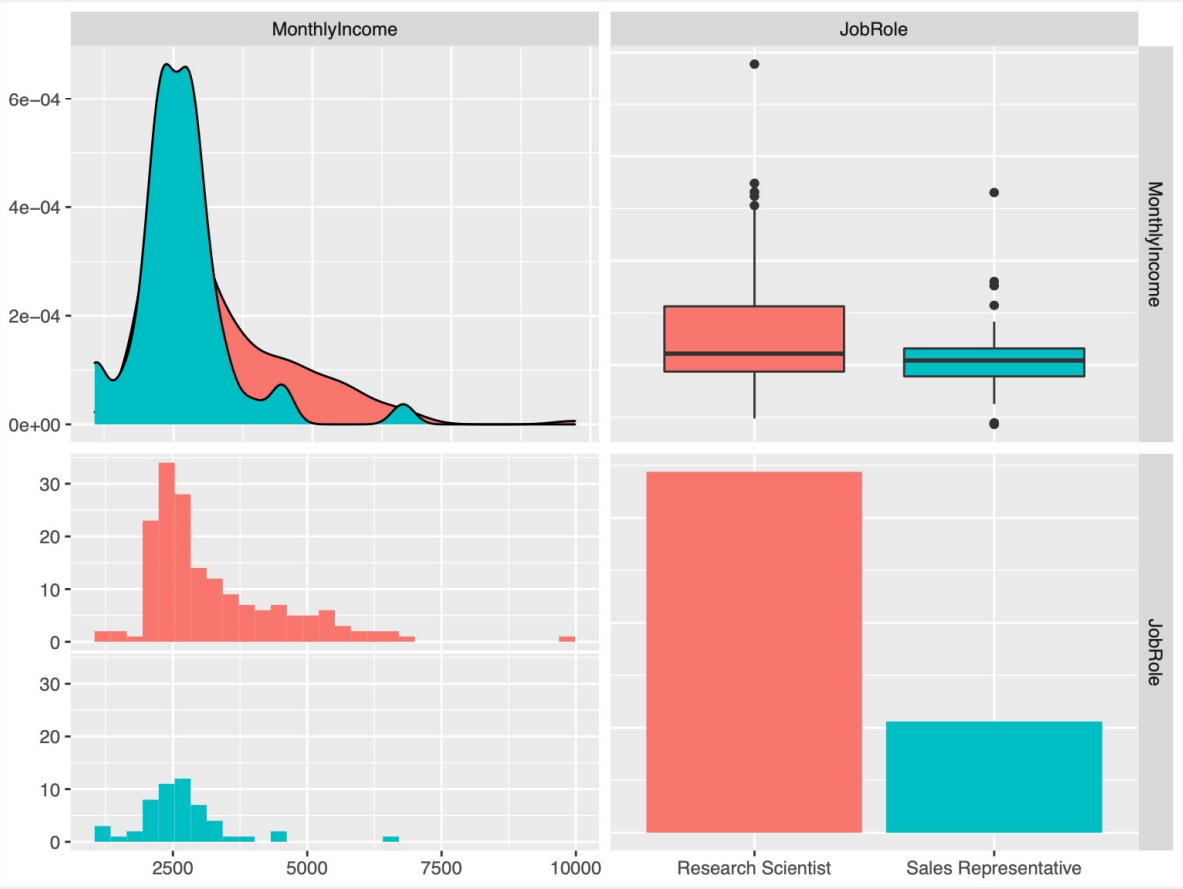
Why does Sales have higher attrition?



Sales Representative vs Research Scientist

Employees in Sales Representatives are younger!

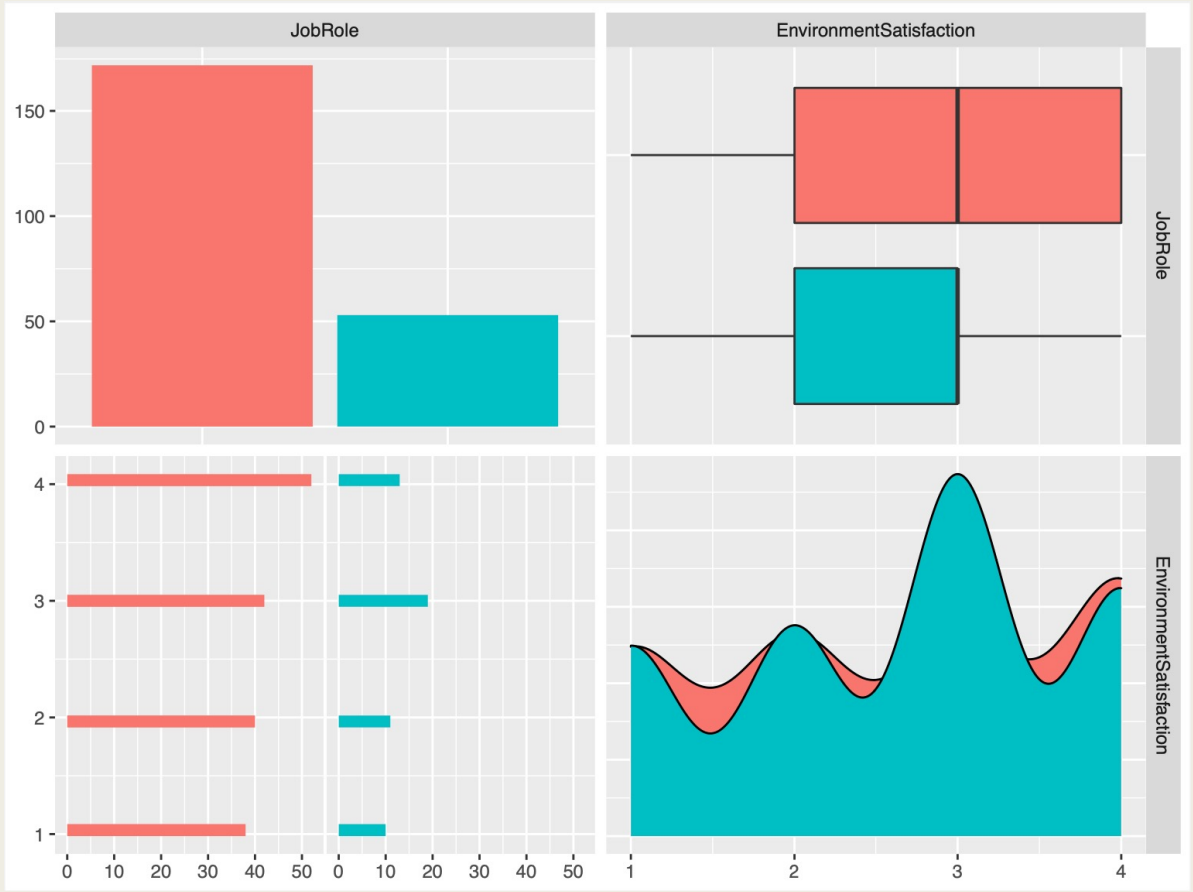
Analysis (cont.)



Sales Representatives also have a lower Monthly Income.

Analysis (cont.)

Sales Representatives also have a lower Environment Satisfaction.



Predicting Attrition

```
               Accuracy : 0.6092
                 95% CI : (0.5471, 0.6688)
    No Information Rate : 0.8429
    P-Value [Acc > NIR] : 1

                  Kappa : 0.1291

 Mcnemar's Test P-Value : 8.373e-12

            Sensitivity : 0.6091
            Specificity : 0.6098
         Pos Pred Value : 0.8933
         Neg Pred Value : 0.2252
             Prevalence : 0.8429
         Detection Rate : 0.5134
   Detection Prevalence : 0.5747
      Balanced Accuracy : 0.6094

       'Positive' Class : No
```

We can predict employee attrition with a 60.92% accuracy.

We can predict those who stay (no attrition) with a 60.91% sensitivity.

We can predict those who leave (yes attrition) with a 60.98% specificity.

Predicting Monthly Income

```
lm(formula = MonthlyIncome ~ JobLevel + TotalWorkingYears + YearsWithCurrManager,
    data = TrainSalary)

Residuals:
    Min      1Q  Median      3Q     Max
-4781.7  -918.3    43.2   741.6  3886.0

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          -1737.06     123.86 -14.024  < 2e-16 ***
JobLevel              3727.41      80.05  46.561  < 2e-16 ***
TotalWorkingYears       71.73      12.21   5.873 7.07e-09 ***
YearsWithCurrManager   -56.72      18.58  -3.052  0.00237 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1389 on 605 degrees of freedom
Multiple R-squared:  0.9098, Adjusted R-squared:  0.9093
F-statistic:  2034 on 3 and 605 DF,  p-value: < 2.2e-16
```

Model →

Predicted Monthly Income = Job Level + Total Working Years + Years with Current Manager

RMSE →

$1354.42

The End