



Análise de Attrition

Valesca Souza



Análise de Dados

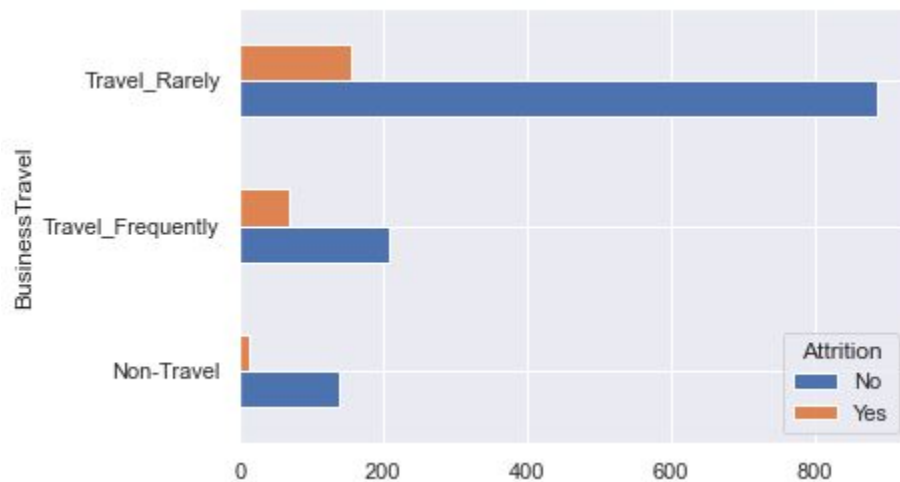
O primeiro passo foi descobrir qual dos atributos categóricos tem uma relação estatisticamente significativa com Attrition, que é o atributo que queremos avaliar. Para isto, foi realizado o teste Qui-quadrado que mede a relação de dependência entre duas variáveis categóricas, verificando como os valores esperados desviam dos valores observados.

Quando temos um alto valor de Qui-quadrado (nosso p-value será baixo), significa que temos evidência estatística para inferir que os valores observados e esperados não são os mesmos, portanto possuem dependência entre si. Quanto mais alto o Qui-quadrado, maior a dependência entre as variáveis.

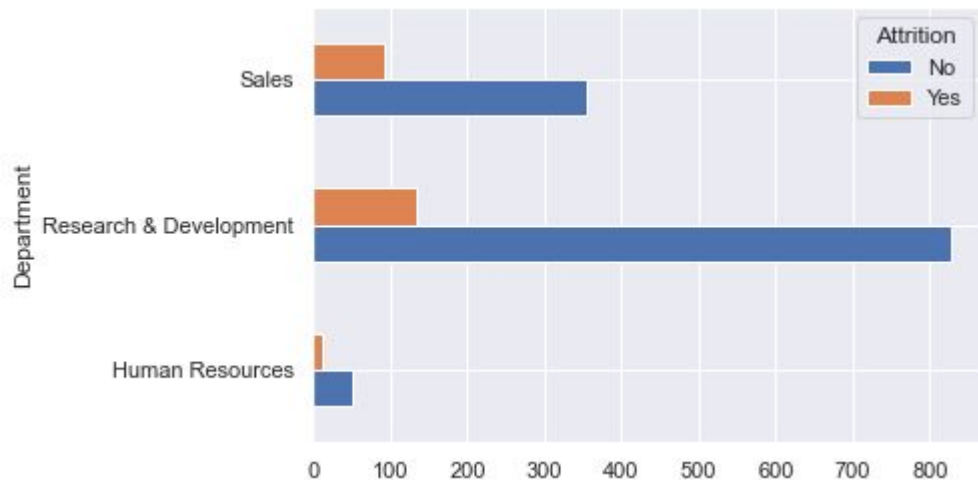


| | Attrition | p < 0.05 |
|--------------------------|-----------|----------|
| Over18 | 1.0 | False |
| PerformanceRating | 0.990075 | False |
| Education | 0.545525 | False |
| Gender | 0.290572 | False |
| RelationshipSatisfaction | 0.154972 | False |
| EducationField | 0.006774 | True |
| Department | 0.004526 | True |
| WorkLifeBalance | 0.000973 | True |
| JobSatisfaction | 0.000556 | True |
| EnvironmentSatisfaction | 0.000051 | True |
| BusinessTravel | 0.000006 | True |
| JobInvolvement | 0.000003 | True |
| MaritalStatus | 0.0 | True |
| StockOptionLevel | 0.0 | True |
| JobLevel | 0.0 | True |
| JobRole | 0.0 | True |
| OverTime | 0.0 | True |

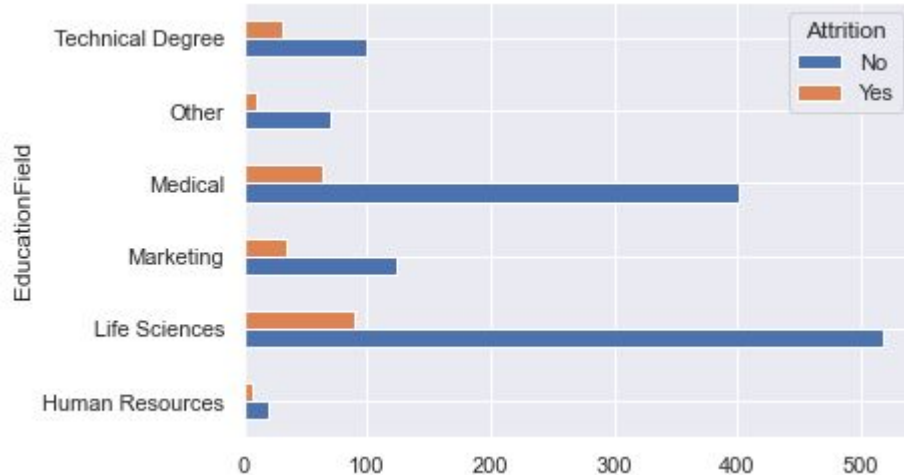
A maioria dos funcionários com desgaste (Attrition) raramente viaja



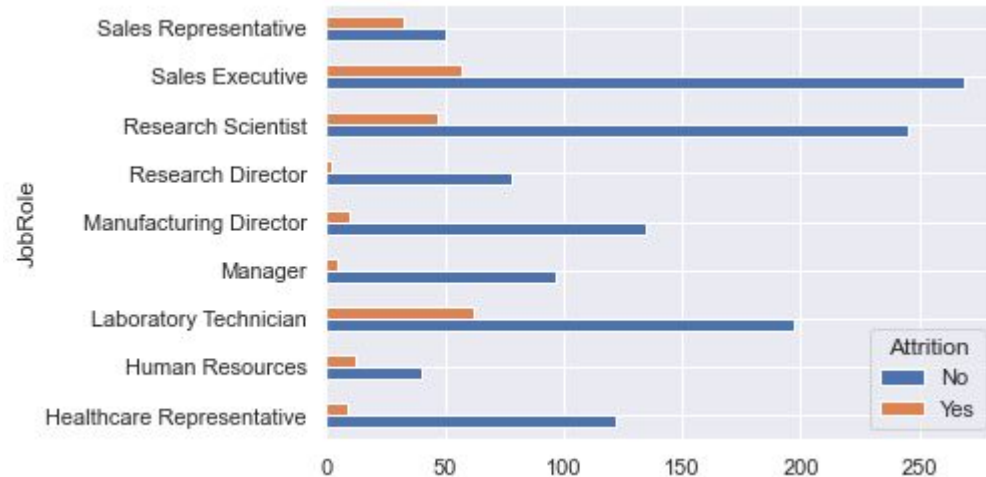
Os departamentos Research & Development e Sales são responsáveis por cerca de 95% dos funcionários com desgaste



O campo da educação de mais da metade dos funcionários com desgaste é Life Sciences e Medical

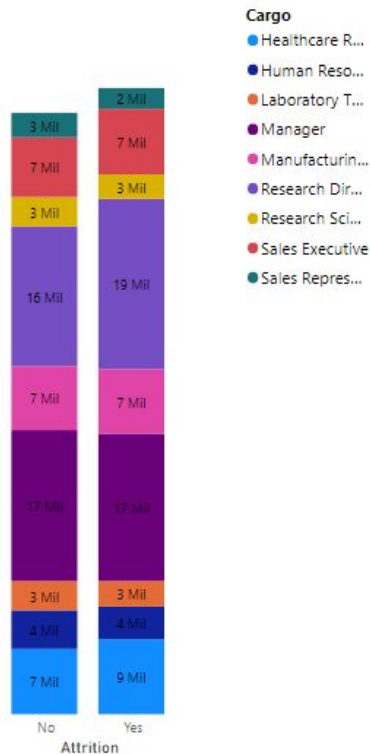


Cerca de 70% dos funcionários com desgaste possuem os seguintes cargos laboratory technicians, sale executives, and research scientists

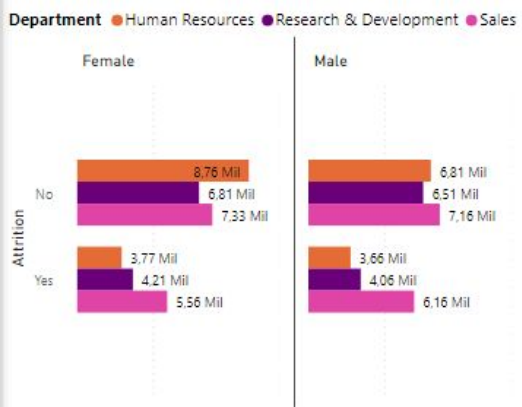


Análise de Renda Mensal

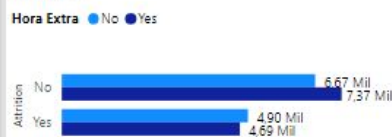
Média de da Renda Mensal por Cargo e Attrition



Média de Renda Mensal por Attrition, Departamento e Gênero



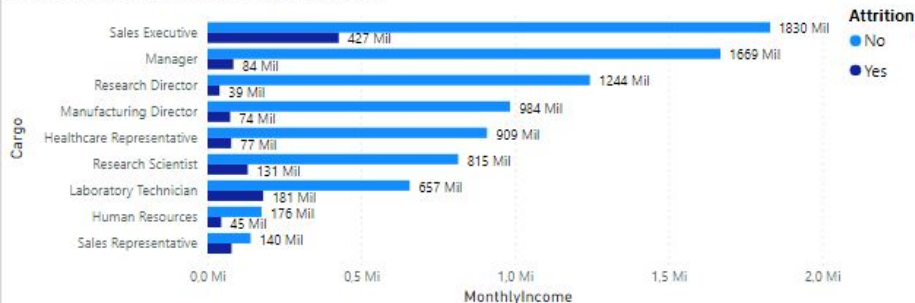
Média de Renda Mensal por Horas Extras e Attrition



Média de Renda Mensal por Nível no Trabalho



Média da Renda Mensal por Cargo e Attrition





Previendo o desgaste

- Transformação de dados categóricos em numéricos;
- Aplicando o SMOTE (Synthetic Minority Over-sampling Technique) para gerar dados sintéticos da classe minoritária;
- Separar dados de treino e teste, em 75% para treino e 25% para teste;
- Aplicado o Random Forest ($n_estimators = 900$) para prever o Attrition:

| Relatório de classificação: modelo Random Forest | | | | |
|--|-----------|--------|----------|---------|
| | precision | recall | f1-score | support |
| No | 0.9026 | 0.9719 | 0.9360 | 925 |
| Yes | 0.9695 | 0.8950 | 0.9308 | 924 |



Conclusão

Neste trabalho foi feita algumas análises dos dados para entender o que pode afetar o desgaste de funcionários. Bem como a implementação de um algoritmos de machine learn, o Random Forest que obteve um recall de 97,5% para funcionários sem desgaste e 89,5% para funcionários com desgaste.