Professional Report

Employee Attrition Analysis

Author: Paulo Henrique Pereira da Cunha

1. Executive Summary

This report analyzes employee attrition in a dataset of 1,470 records. The objective was to understand **who leaves, why, and the financial impact**, using **Python** for data cleaning and exploratory analysis, and **Power BI** for visualization.

The overall attrition rate was **16.1%**, concentrated in the **Sales and HR departments**. High-risk profiles include employees **under 25 years old**, those with **less than 1 year at the company**, **low-income employees**, and those working **excessive overtime**.

The estimated financial impact is **USD 6.8 million**, with potential savings of **USD 1 million** if retention programs reduce turnover by up to 15% (benchmarks: SHRM, Gallup, Deloitte).

Recommendations include implementing **structured onboarding, salary review, and career development plans**.

2. Report Objective

Identify the main factors associated with employee attrition and propose actions to reduce turnover costs and improve talent retention.

3. Data Description

- **Source**: HR dataset (simulated, Kaggle/IBM HR Attrition).
- **Size**: 1,470 records, 237 attrited employees.
- **Key variables**: age, monthly income, years at company, department, job role, overtime, satisfaction (job, environment, relationships, work-life balance).
- **Notes**:
    - No significant missing values.
    - Outliers analyzed via IQR and kept, as they represent legitimate variations (high salaries, long careers).

4. Methodology

1. **ETL and Cleaning**
    - Standardization of column names into Portuguese.
    - Creation of derived variables (age bands, tenure bands, income quartiles, attrition flag).
2. **EDA (Exploratory Data Analysis)**
    - Histograms for age, income, years at company, and commuting distance.
    - Attrition rates segmented by department, role, income, age, and overtime.
    - Point-biserial correlations to assess associations with attrition.
3. **Comparative Statistics**
    - Comparison of average salary and satisfaction between attrited and active employees.
4. **Financial Calculation**
    - Estimated cost = annual salary of attrited × replacement rate (30–70%, default 50%).
    - Potential savings of 15% (based on benchmarks SHRM, Gallup, Deloitte).

Tools used: Python (Pandas, Matplotlib, Seaborn), Power BI.

5. Analysis and Results

- Overall attrition rate: 16.1% (237 attrited).
- By department: Sales (20.6%), HR (19%), R&D (13.8%).
- By age: ≤25 years (35.7%) vs 36–45 years (9.1%).
- By tenure: ≤1 year (34.9%) vs 10+ years (8.1%).
- Overtime: 30.5% (Yes) vs 10.4% (No).
- Income: low (29.2%) vs high (10.3%).
- Average salary: USD 4,787 (attrited) vs USD 6,832 (retained).
- Satisfaction: consistently lower among attrited employees.
- Estimated turnover cost: USD 6.8 million.
- Potential savings: USD 1.02 million with 15% reduction.

6. Insights and Conclusions

- Attrition is **not random**: concentrated among **young, low-income, short-tenure employees**, and those working overtime.
- **Low satisfaction and low pay** reinforce attrition risk.
- The **financial impact is significant**, justifying investment in retention policies.
- **Sales and HR** are critical departments to prioritize.

7. Recommendations

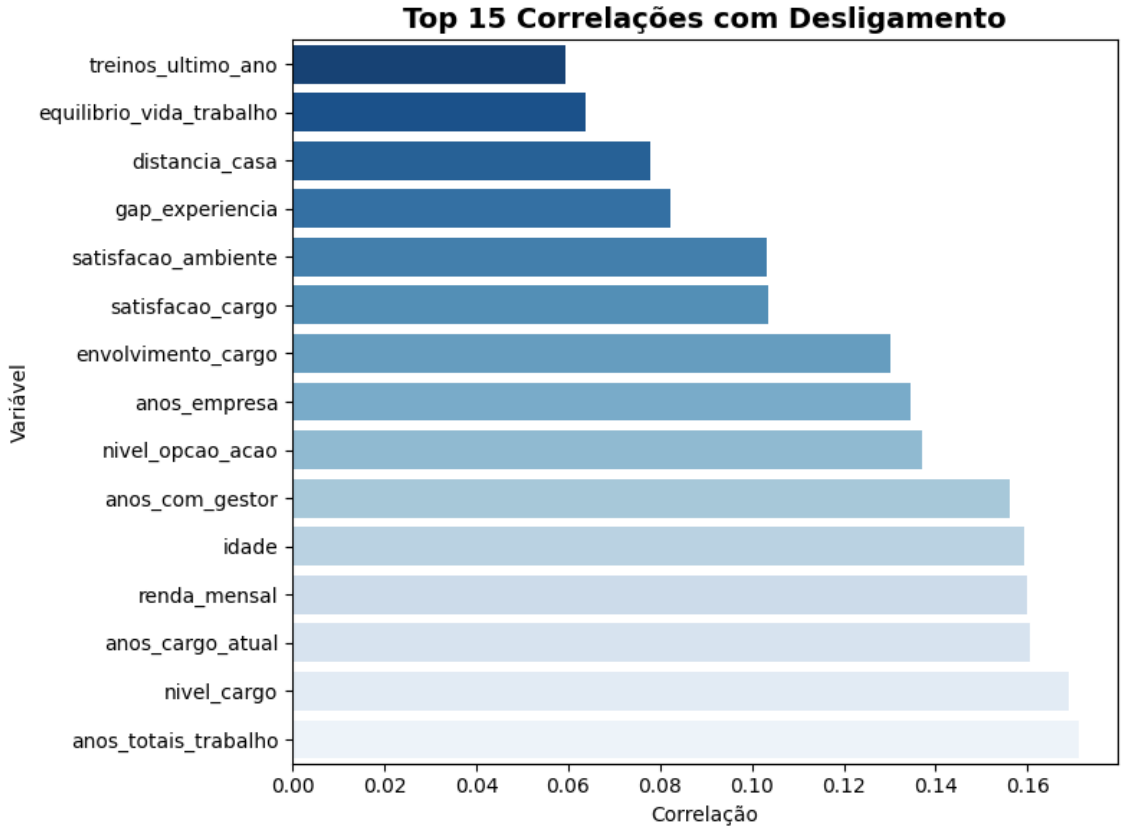
1. Structured onboarding to reduce exits within the first 12 months.
2. Career development and promotion plans to reduce stagnation.
3. Salary review to address income inequality.
4. Overtime management to avoid employee overload.
5. Engagement programs focused on work environment, feedback, and recognition.

8. Study Limitations

- Dataset is simulated (not real).
- Some qualitative variables may not capture real behavioral factors.
- Correlations found are weak (max ≈ 0.17), indicating attrition is multifactorial and not explained by single variables.

Appendix A - Complementary Charts (Python)

1. Top 10 Correlations with Attrition

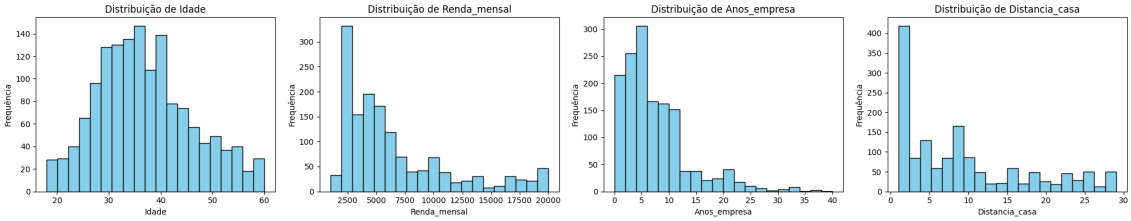


Horizontal bar chart showing correlation coefficients between numerical variables and employee attrition.

Insights:

- Correlations are weak (max ≈0.17), showing that no single variable explains attrition.
- Tenure and years in role show slight association with exits → stagnation risk.
- Younger employees and lower salaries are associated with higher attrition.
- Job satisfaction and work-life balance have weak negative correlation, highlighting the role of workplace experience.

2. Profile Distributions (Age, Income, Tenure, Commuting Distance)



Histograms showing workforce distribution in four dimensions: age, monthly income, years at company, commuting distance.

Insights:

- **Age:** majority between 30–40 → young-adult workforce.
- **Income:** concentrated in lower ranges (up to USD 5,000) → marked salary inequality.
- **Tenure:** majority with ≤5 years → relatively new employees.
- **Distance:** most live close (1–5 km), but a minority commute long distances (up to 30 km).

Appendix B - Summary of Methodological Flow

1. ETL (Extract, Transform, Load)
    - Import dataset
    - Standardize columns
    - Create derived variables
2. EDA (Exploratory Data Analysis)
    - Descriptive statistics
    - Histograms and distributions
    - Segmentations by age, tenure, income, overtime
3. Statistical Analysis
    - Point-biserial correlations
    - Mean comparisons (attrited vs active)
4. Financial Impact
    - Estimation of turnover cost
    - Calculation of potential savings using benchmarks
5. Conclusions and Recommendations
    - Risk profiles
    - Associated factors
    - Practical retention suggestions

Appendix C - References

- SHRM (Society for Human Resource Management). *Onboarding New Employees: Maximizing Success*.
- Gallup. *State of the Global Workplace Report*.
- Deloitte Insights. *Human Capital Trends*.
- Mercer. *Global Talent Trends*.