Employee Turnover Report

Sonia Rus Gelo

The Human Capital Department - 25 January 2022

Report of the Managing Director

Index

1. Background

1.1 Report Basis

1. The following report has been prepared to provide an up to date **analysis of employee turnover** as part of the Board planning considerations.
2. The specific aims of the report regard the establishment of the departments having the highest and lowest employee turnover rate, as well as **identifying** the variables which seem to be better **predictors** of employee departure from the company. Finally, some recommendations will be delivered in order to prevent potential issues.
3. The incidence of a **high turnover has significant resource implication**s and places constraints on the ability to deliver a high quality service provision. Therefore, the need to retain skilled and experienced staff is important as the company responds to ongoing financial pressures.
4. The data presented in this report is based upon an analysis on **10,000 employees**.
5. The information was collected from **exit interviews**, **performance reviews**, and **employee records**.
6. Employee records include benefits, eligibility, training history, performance reviews, disciplinary actions, job experience and compensation history.

1.2 Objectives

1. Which department has the **highest** employee turnover? Which one has the **lowest**?
2. Investigate which **variables** seem to be **better predictors** of employee departure.
3. What **recommendations** would you make regarding ways to **reduce employee turnover**?

2. Data Overview

In [1]:
```
# Libraries
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns
from scipy import stats
```

2.1 The Data Set

In [2]:
```
# Read Data
df = pd.read_csv('./data/employee_churn_data.csv')
df.head()
```

Out[2]:

	department	promoted	review	projects	salary	tenure	satisfaction	bonus	avg_hrs_month	left
0	operations	0	0.577569	3	low	5.0	0.626759	0	180.866070	no
1	operations	0	0.751900	3	medium	6.0	0.443679	0	182.708149	no
2	support	0	0.722548	3	medium	6.0	0.446823	0	184.416084	no
3	logistics	0	0.675158	4	high	8.0	0.440139	0	188.707545	no
4	sales	0	0.676203	3	high	5.0	0.577607	1	179.821083	no

2.2 Variables explanation

- "department" - the department the employee belongs to.
- "promoted" - 1 if the employee was promoted in the previous 24 months, 0 otherwise.
- "review" - the composite score the employee received in their last evaluation.
- "projects" - how many projects the employee is involved in.
- "salary" - for confidentiality reasons, salary comes in three tiers: low, medium, high.
- "tenure" - how many years the employee has been at the company.
- "satisfaction" - a measure of employee satisfaction from surveys.
- "avg_hrs_month" - the average hours the employee worked in a month.
- "left" - "yes" if the employee ended up leaving, "no" otherwise.

3. Relevant Issues

Employee turnover has been assessed on the basis of the **number of leaving employees per department as a percentage** of the total number of staff employed **by each department** based on:

$$\frac{a}{b} = \frac{x}{100}$$

a = employees that left | b = total employees by department | x = turnover ratio per department

In [3]:
```
# Hands-on
proportion_table = pd.crosstab(df.department, df.left).apply(lambda r:np.around(r/r.sum() * 100,2),axis=1)
proportion_table.sort_values('yes', ascending = False, inplace = True)
proportion_table.index = [*map(lambda x:x.capitalize(), proportion_table.index)]

# Print
print("Turnover ratio (%) per department: \n" )
proportion_table
```

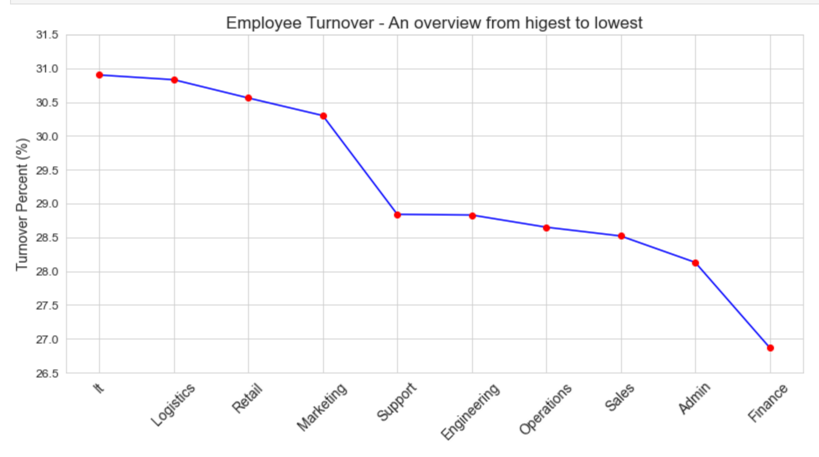
```
Turnover ratio (%) per department:
```

Out[3]:

left	no	yes
It	69.10	30.90
Logistics	69.17	30.83
Retail	69.44	30.56
Marketing	69.70	30.30
Support	71.16	28.84
Engineering	71.17	28.83
Operations	71.35	28.65
Sales	71.48	28.52
Admin	71.87	28.13
Finance	73.13	26.87

A table showing the overall turnover rate within each department is set out below:

In [4]:
```
# Plot 'proportion_table' from high to low
sns.set_style("whitegrid")

bins = np.arange(26.5 ,32, 0.5)

figg = plt.figure(figsize = (11, 6), constrained_layout = True)
ax = figg.subplots()

plt.scatter(proportion_table.index, proportion_table['yes'], color = 'r', zorder = 2)
plt.plot(proportion_table.index, proportion_table['yes'], color='b', zorder = 1)

# Details
plt.title("Employee Turnover - An overview from higest to lowest", fontsize = 17)
plt.yticks(bins,fontsize = 12)
plt.ylabel("Turnover Percent (%)", fontsize = 14)
plt.xticks(rotation = 45, fontsize = 14)

# Show
plt.show()
```



The figures indicate that the IT department has the highest employee turnover rate at 30.90%. On the other hand, the Finance department has the lowest employee turnover rate, at 26.87%.

4. Turnover predictors

4.1 Variables correlation

Employee turnover predictors has been assessed through he **Pearson Correlation Coefficient**, a quantitative value of the relationship between two or more variables. The correlation coefficient can vary **from -1.00** (negative correlation) **to 1.00** (positive correlation). Based on:

$$\frac{N\sum xy - \sum x \sum y}{\sqrt{[\sum x^2 - (\sum x)^2] * [\sum y^2 - (\sum y)^2]}}$$

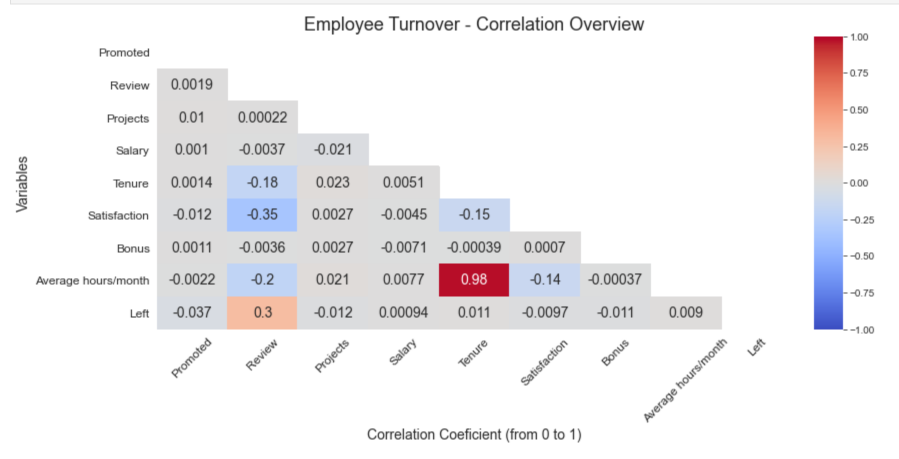
N = Total sample | x = independent variable | y = dependent variable

In [5]:
```
# map the categorical variables
df['left'] = df['left'].map({'yes':1, 'no':0})
df['salary'] = df['salary'].map({'low':0, 'medium':1, 'high':2})

# Correlation matrix
corr_matrix = df.corr()

# print
print("Turnover correlation: \n")
print(corr_matrix["left"].sort_values(ascending=False))
```

```
Turnover correlation:

left             1.000000
review           0.304294
tenure           0.010521
avg_hrs_month    0.009008
salary           0.000943
satisfaction    -0.009721
bonus           -0.011485
projects        -0.012408
promoted        -0.036777
Name: left, dtype: float64
```

According to this data, the only variable which seems to be a valuable turnover predictor is "Review" (0.3), defined as "The composite score which the employee received in their last evaluation".

In [6]:
```
# Heatmap
figg = plt.figure(figsize = (12,6), constrained_layout = True)
ax = figg.subplots()

# For cleaner index and ticks
df_heat = df.copy()
df_heat.columns = ['Department', 'Promoted', 'Review',
                   'Projects', 'Salary', 'Tenure', 'Satisfaction',
                   'Bonus', 'Average hours/month', 'Left']

# Plot Corpus
upp_mat = np.triu(df.corr())

sns.heatmap(df_heat.corr(), vmin = -1, vmax = +1, annot = True,
            cmap = 'coolwarm',annot_kws = {'size':14}, mask = upp_mat)

# Appareance
plt.title("Employee Turnover - Correlation Overview", fontsize = 18)
plt.ylabel("Variables", fontsize = 14)
plt.yticks(fontsize = 12)
plt.xlabel("Correlation Coeficient (from 0 to 1)", fontsize = 14)
plt.xticks(rotation = 45, fontsize = 12)

# Show
plt.show()
```



Other variables considered, such as "Tenure" (0.01), defined as "How many years the employee has been at the company", or "Average hours/month" (0.009), defined as "The average hours the employee worked in a month", have minimal values. Therefore, they are not to be considered when designing intervention strategies.

4.2 p-value Signification

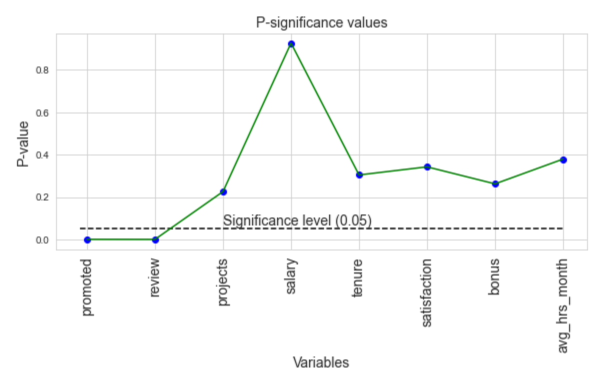
In technical terms, a p-value is the **probability** of obtaining an effect at least as extreme as that of the sample data, assuming that the null hypothesis is true.

In [7]:
```
# p-value
p_list = []

def p_value_wchair(cols):
    for col in cols:
        r, p = stats.pearsonr(df['left'], df[col])
        p_list.append(p)
        print('Variable', col, 'p value: ', p)

p_value_wchair(df.columns[1:-1])
```

```
Variable promoted p value:  0.0003270460442308153
Variable review p value:  1.522017428119139e-203
Variable projects p value:  0.22559419297868366
Variable salary p value:  0.9266091135901224
Variable tenure p value:  0.3041673502064808
Variable satisfaction p value:  0.3424439867190504
Variable bonus p value:  0.26201637368964104
Variable avg_hrs_month p value:  0.37900768131613305
```

In [8]:
```
# Declare axes
figg = plt.figure(figsize = (8, 5), constrained_layout = True)
ax = figg.subplots()

# corpus
plt.plot(p_list, color='g')
plt.scatter(df.columns[1:-1], p_list, color = 'b', zorder = 2)


# lines
plt.hlines(0.05, -0.1, 7, linestyles = 'dashed', color = 'black')
plt.annotate('Significance level (0.05)',(2,0.065), size = 14)

# params
plt.title("P-significance values", fontsize = 14)
plt.xticks(range(0,len(df.columns[1:-1])), df.columns[1:-1], rotation = 90, fontsize = 14)
plt.ylabel("P-value", fontsize = 14)
plt.xlabel("Variables", fontsize = 14)

# plot
plt.plot()
```

Out[8]: []



P-significance level tells that with more than a 99,9% probability, **review** (0.3) is the most correlated variable with turnover, as **promotion** (-0.03), although has a weak correlation, could be in the scope afterwards.

5. Recommendations

5.1 Intervention implementation

In order to guarantee the department employees engagement and retention, a **change of the review criteria** used in previous evaluations is deemed necessary in order to minimize the impact of this variable on employee turnover. Implementing this strategy will entail the following action points:

A. Designing a **questionnaire** for the IT department employees where they may be able to assign a specific value to the review criteria used in previous evaluations. Thus they will choose themselves the review criteria applied in further evaluations and in doing so they will be more likely to accept the validity of the results.

B. First of all, the different **review criteria will be broken down in detail** and **transformed** into questionnaire items.

C. Second, the department **employees will respond to the items** according to a Likert scale by selecting the number which they consider to reflect the **perceived quality** of the specific review criteria.

D. The **format** of each item will have five levels as follows:

```
1. Strongly disagree

2. Disagree

3. Neither agree nor disagree

4. Agree

5. Strongly agree
```

E. A specific **Staff App** will be developed to **extend the channels of communication** during the term the employees are provided to complete the questionnaire. This will provide them the possibility to ask any question concerning the content of the questionnaire items as well as the interpretation of the scale levels.

F. The questionnaire will be launched as soon as possible in order to efficiently **avoid the employee turnover** provoked by the negative impact of the review variable.

5.2 Effectiveness measurement

The impact of this strategy will be measured by the same indicators and variables analyzed for this report over the current financial year. The same sort of sources, such as exit interviews, performance reviews, and employee records will be taken into account.

7. References

Aburumman, Omar & Arabiat, Khuzama. (2021). Determinants of Employee Turnover Intention. 10.9734/bpi/mpebm/v1/11061D.

Allen, E.; Seaman, C. (2007). Likert Scales and Data Analyses. Quality Progress, 64–65.

Babbie, Earl R. (2005). The Basics of Social Research. Belmont, CA: Thomson Wadsworth, p. 174.

Burrell, S. (2014). IT Staff Turnover Intentions, Job Modification, and the Effects of Work Recognition at Large Public Higher Education Institutions.

C. A. Al Mamun and M. N. Hasan (2017). Factors affecting employee turnover and sound retention strategies in business organization: A conceptual view. Problem Perspectives Management, 15(1), 63–71. doi: 10.21511/ppm.15(1).2017.06.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.).

Hom, Peter & Allen, David & Griffeth, Rodger. (2019). Theories of Employee Turnover. 10.4324/9781315145587-4.

Iqbal, Shuja & Hongyun, Tian & Akhtar, Shamim & Ahmad, Usama & Ankomah, Fred. (2020). Impacts of Supervisor Support on Turnover Intentions: Mediating Role of Job Satisfaction. Asian Journal of Education and Social Studies. 1-9. 10.9734/ajess/2020/v6i330174.

Kendall, M. G., Stuart, A. (1973) The Advanced Theory of Statistics, Volume 2: Inference and Relationship, Griffin.

Latham, Gary P. (2006). Work Motivation: History, Theory, Research, And Practice. Thousand Oaks, Calif.: Sage Publications. p. 15.

Lim, Paul & Parker, Andrew. (2020). Employee Turnover. 10.1108/978-1-78973-483-620201006.

Mitrofanova, A. (2015). Managing Employee Turnover. Management of the Personnel and Intellectual Resources in Russia. 4. pp. 47-51. 10.12737/13240.

Moro, S.; Ramos, F.; Rita, P. (2020) What drives job satisfaction in IT companies?. International Journal of Productivity and Performance Management 70(2), 391-407.

Mobley, W. H. (1983). Employee Turnover: Causes, Consequences, and Control, Price Industrial Labor Relations Review. 36(3), 506–507.

Rusch, T.; Lowry, P.; B.; Mair, P.; Treiblmaier, H. (2017). Breaking free from the limitations of classical test theory: Developing and measuring information systems scales using item response theory. Information & Management, 54(2),189–203. doi:10.1016/j.im.2016.06.005.

Srivastava, S. & Kanpur, R. (2014). A Study On Quality Of Work Life: Key Elements & Its Implications. IOSR Journal of Business and Management. 19(5), 411–423. https://doi. org/10.9790/487x-16315459.