People Analytics

Project Report

at the Faculty of Business, Economics, and Law

Friedrich-Alexander-Universität Erlangen-Nürnberg

Schöller Endowed Chair for Information Systems

Prof. Dr. Sven Laumer

Presented by: Muhammad Fahad Ali

Rothenburger Str 145/106

90439, Nurnberg

+49 017674768100

mfahad.ali@fau.de

22995110

Final Date of Submission: 2024-02-01

Content

1 Well-Being

1.1 Importance of Monitoring Employee Well-being

Monitoring employee well-being is crucial for FAU Airlines for several reasons:

1. **Employee Performance and Productivity:** Well-being directly impacts employee performance and productivity. Satisfied and healthy employees are likely to be more engaged, focused, and efficient in their tasks, ultimately contributing to the overall success of the company.

2. **Employee Morale and Satisfaction:** Monitoring well-being helps in understanding the level of job satisfaction and morale within the workforce. High levels of well-being are associated with greater job satisfaction, which, in turn, contributes to higher retention rates and reduced turnover.

3. **Organizational Reputation:** A company that prioritizes employee well-being is likely to build a positive reputation in the industry. This can attract top talent and create a positive image for customers and stakeholders.

4. **Workplace Culture:** Well-being is closely tied to the organizational culture. Monitoring well-being allows the company to assess whether its culture is supportive, inclusive, and conducive to employee happiness and fulfillment.

5. **Healthcare Costs:** Unhealthy employees may lead to increased healthcare costs for the company. Monitoring well-being can help identify and address health issues early, potentially reducing long-term healthcare expenses.

A) Factors Leading to Low Well-being and Unhealthy Work-Life Balance

Several factors can contribute to low levels of well-being and an unhealthy work-life balance:

1. **Excessive Workload:** Employees facing consistently high workloads and tight deadlines may experience stress and burnout, negatively impacting their well-being.

2. **Lack of Job Security:** Insecurity about job stability can lead to heightened stress levels and decreased well-being among employees.

3. **Poor Work-Life Balance:** Employees struggling to balance work and personal life may experience fatigue, impacting their overall well-being.

4. **Limited Autonomy and Control:** Employees with limited control over their work or decision-making processes may feel disengaged and stressed.

5. **Inadequate Social Support:** Lack of supportive relationships and a sense of isolation in the workplace can contribute to low well-being.

B) Practical Approach to Monitor Well-being

FAU Airlines can employ a practical approach to monitor well-being by:

1. **Well-being Surveys:** Conduct regular surveys to gather feedback on various dimensions of well-being. Include questions related to workload, job satisfaction, work-life balance, and overall happiness.

2. **Performance Indicators:** Establish key performance indicators (KPIs) related to well-being, such as absenteeism rates, turnover rates, and employee satisfaction scores.

3. **Well-being Dashboard:** Develop a well-being dashboard that consolidates data from surveys and performance indicators. This provides a visual representation of well-being trends and allows for real-time monitoring.

4. **Alarming System:** Implement an alarming system within the well-being dashboard to alert HR and management when significant deviations from established KPIs occur.

5. **Root Cause Analysis:** Conduct thorough analyses to identify the root causes of any issues affecting well-being. This may involve investigating workload distribution, interpersonal dynamics, or organizational policies.

6. **Support Programs:** Use insights from the monitoring process to design and implement targeted support programs, such as stress management workshops, flexible work arrangements, and mental health resources.

By adopting such a practical approach, FAU Airlines can proactively address issues affecting employee well-being, creating a healthier and more productive work environment.

1.2 Stress induced by technology

A) Potential Influence on Employees' Stress and Consequences

1. **Technostress Types:** The introduction of a new Computer Reservation System (CRS) at FAU Airlines may induce technostress among employees. Depending on their perceptions, employees may experience either Challenge Technostress or Hindrance Technostress.

2. **Employee Perceptions:**

- **Challenge Technostress:** Employees viewing the new CRS as an opportunity to enhance their skills and improve efficiency may experience positive outcomes.

- **Hindrance Technostress:** Those perceiving the system as a threat hindering their daily work routine might face negative consequences.

3. **Control Levels:** According to Beaudry & Pinsonneault (2005), the control level employees feel over the new system will impact outcomes. High control may lead to positive consequences, while low control may result in negative outcomes.

4. **Positive Consequences:** Increased efficiency, better psychological well-being, higher engagement levels, improved job performance, and a technical mastery of the new system.

5. **Negative Consequences:** Dissatisfaction with jobs, decreased loyalty, hindered performance, exhaustion, burnout, and reduced overall well-being.

B) Coping Strategies and Recommendations

1. **Training and Support:**

Provide comprehensive training to check-in agents on using the new CRS.

Offer ongoing support through workshops and resources to enhance their technological skills.

2. **Clear Communication:**

- Ensure transparent communication about the purpose, benefits, and potential challenges of the new system.
- Set realistic expectations and goals regarding technology use and workload management.

3. **Employee Involvement:**

- Involve employees in the decision-making process regarding the CRS implementation.
- Seek feedback, address concerns, and consider their input to enhance their sense of control and ownership.

4. **Workload Management:**

- Implement strategies to manage workload effectively, considering the impact of the new CRS on job demands.
- Provide tools and resources to streamline processes and prevent role overload.

5. **Well-being Support:**

- Offer resources and programs to support employee well-being, such as stress management workshops and mental health support.
- Encourage work-life balance and provide opportunities for relaxation and rejuvenation.

6. **Personalized Communications:**

- Engage in articulated personal level communications with employees to understand their perceptions of the new CRS and address concerns.
- By implementing these strategies, FAU Airlines can mitigate potential technostress, ensuring a smoother transition to the new CRS and fostering positive outcomes for employees. This approach aligns with literature emphasizing the importance of training, communication, involvement, workload management, and well-being support in managing technostress.

1.3 Employee well-being dataset

A) Data Understanding and Preprocessing (Refer to employee_wellbeing.ipynb)

The dataset "wellbeing_dataset.csv" is used to determine the main factors that influence the WORK_LIFE_BALANCE_SCORE (WLB) of employees. Before analyzing the data, the following preprocessing steps are performed:

- **Convert string datatypes to numeric:** Age and Gender columns are converted to numeric using LabelEncoder.
- **Drop irrelevant columns:** Columns that are considered irrelevant for the analysis are dropped.

B) Data Analysis Steps

- **Import Libraries:** The necessary libraries (pandas, numpy, seaborn, and matplotlib) are imported Figure 1.

```
# import necessary libraries

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
```

Figure 1

- **Read the Dataset:** The CSV file is read into a pandas Data Frame Figure 2.

```python
# reading the CSV file into a pandas dataframe
df = pd.read_csv("employee_wellbeing.csv")

# displaying all column names
print(df.columns)
```
✓ 0.0s

```
Index(['EMP_ID', 'AGE', 'GENDER', 'STATUS', 'EMPLOYMENT', 'SUFFICIENT_INCOME',
       'SALARY', 'TO_DO_COMPLETED', 'DAILY_STRESS', 'CORE_CIRCLE',
       'SUPPORTING_OTHERS', 'SOCIAL_NETWORK', 'ACHIEVEMENT', 'FLOW',
       'DAILY_STEPS', 'SLEEP_HOURS', 'LOST_VACATION', 'PERSONAL_AWARDS',
       'TIME_FOR_HOBBY', 'HEALTHY_DIET', 'WORK_LIFE_BALANCE_SCORE'],
      dtype='object')
```

Figure 2

- **Data Preprocessing**

1. **Encoding Categorical Variables:** The Age and Gender columns are converted from strings to numeric using LabelEncoder.

2. **Dropping Irrelevant Columns:** After analyzing the data, columns that are irrelevant to the analysis are dropped Figure 3.

```python
# check for any missing values in the data and drop 'EMP_ID' column adi t is irrelevant for analysis
df.isnull().values.any()
df= df.drop(columns=['EMP_ID'], axis = 1)

df_encoder = LabelEncoder()
for i in range(20):
    if df.iloc[:,i].dtype == 'object':
        df.iloc[:,i] = df_encoder.fit_transform(df.iloc[:,i])
df.head()
```
✓ 0.0s

	AGE	GENDER	STATUS	EMPLOYMENT	SUFFICIENT_INCOME	SALARY	TO_DO_COMPLETED	DAILY_STRESS	CORE_CIRCLE
0	1	0	0	1	1	1	6	2	5
1	1	0	2	0	2	0	5	3	3
2	1	0	2	0	2	0	2	3	4
3	2	0	2	0	1	2	3	3	3
4	2	0	1	0	1	1	5	1	3

Figure 3

C) Findings

- **Daily Stress Analysis:**

1. Daily Stress by Gender: The data analysis shows that males have a higher average level of daily stress than females. The average daily stress score for females is 2.55, while the average daily stress score for males is 2.92. This means that males experience stress more frequently and intensely than females Figure 4.

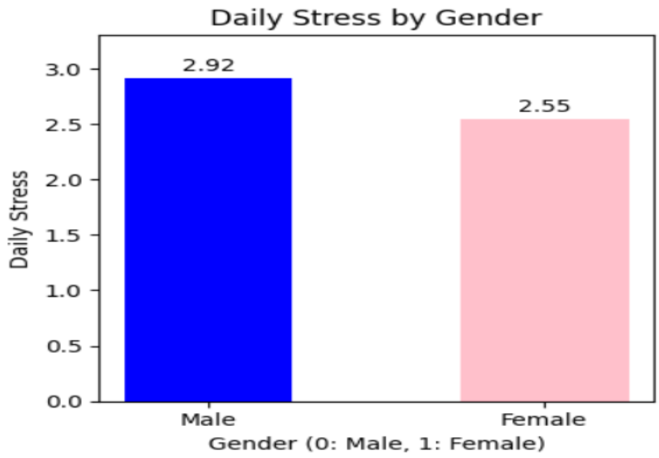

Figure 4

2. **Daily Stress by Jobe Role:** The data analysis shows that the average daily stress level of check-in agents is 2.79, while the average daily stress level of flight attendants is 2.76; both values are similar to each other Figure 5.

Figure 5

3. **Hobby Time Analysis by Gender:** The data analysis shows that males spend an average of 3.2 hours per week on hobbies, while females spend an average of 3.3 hours per week on hobbies. This means that males spend about 3.7% less time on hobbies than females Figure 6.

```
Mean Time for Hobby by Gender:
GENDER
0    3.202676
1    3.334009
Name: TIME_FOR_HOBBY, dtype: float64
```

Figure 6

4. **Correlation Values and Heatmap Generation**: Using Pearson Correlation, and checking the correlation value of each variable with the WORK_LIFE_BALANCE_SCORE, any variable with a correlation of greater than +0.09 or less than -0.09 is considered to have a strong impact on the WLB score and this can also be confirmed with the heatmap below. The 19 variables that are crucial for affecting the score are as follows in Figure 7 & Figure 8 .

```
Columns with the highest impact on WORK_LIFE_BALANCE_SCORE:
AGE = 0.09199385597641332
SUFFICIENT_INCOME = 0.29481329713164744
SALARY = -0.26493993026980167
TO_DO_COMPLETED = 0.552218439271271
DAILY_STRESS = -0.36463487395974803
CORE_CIRCLE = 0.50479649677039
SUPPORTING_OTHERS = 0.5505493474781307
SOCIAL_NETWORK = 0.4120603490297315
ACHIEVEMENT = 0.5773849696677533
FLOW = 0.4837366506391596
DAILY_STEPS = 0.42868079914573903
SLEEP_HOURS = 0.19578563132024046
LOST_VACATION = -0.27029497357851584
PERSONAL_AWARDS = 0.5133261768718717
TIME_FOR_HOBBY = 0.5168612314679437
HEALTHY_DIET = 0.4425562997481686
```

Figure 7

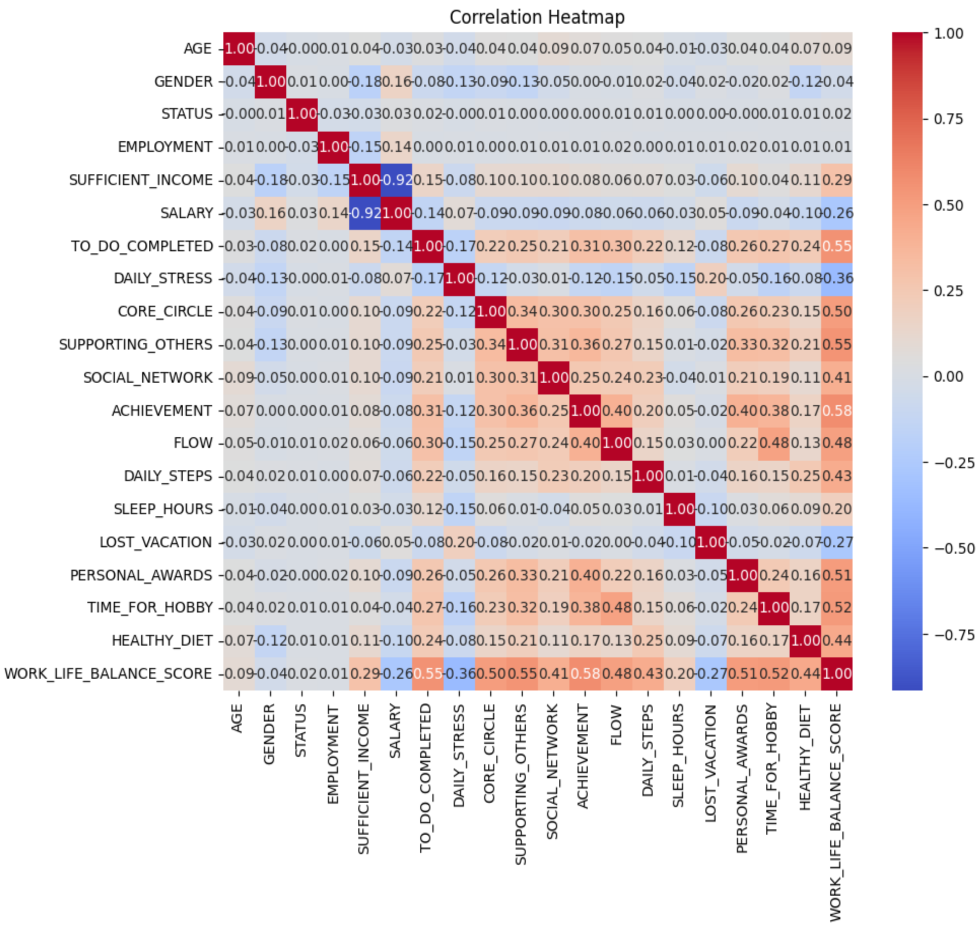

Figure 8

In summary, the examination of data and preprocessing procedures enables the identification of key factors impacting employees' WORK_LIFE_BALANCE_SCORE. The correlation heatmap offers valuable insights into the interplay of attributes, informing decision-making and suggesting potential strategies to enhance work-life balance for FAU Airline employees.

1.4 Predictive well-being algorithm

A) FAU Airline Predicted Model

- **Objective:** The goal of this analysis is to develop a Linear Regression model to predict the WORK_LIFE_BALANCE_SCORE of employees based on relevant features, and subsequently, evaluate its performance.
- **Data Preparation:** To prepare the dataset for modeling, unnecessary columns including 'AGE,' 'GENDER,' 'STATUS,' and 'EMPLOYMENT' were dropped, as they were deemed not directly contributing to the predictive task. The remaining features were utilized to predict the target variable 'WORK_LIFE_BALANCE_SCORE.'

- **Data Splitting:** The dataset was split into training and testing sets using an 80-20 split, with 80% of the data allocated for training and 20% for testing. The random state parameter was set to 42 to ensure reproducibility.
- **Model Selection and Training:** Linear Regression was chosen as the predictive model due to its simplicity and interpretability. The model was trained on the training set using the features (X_train) and the target variable (Y_train) Figure 9.

```
# Evaluate the model
r2 = r2_score(y_test, y_pred)
print(f"R2 Score: {r2}")
```

✓ 0.0s

```
R2 Score: 0.8767169465412897
```

Figure 9

- **Model Evaluation:** The trained model was used to make predictions on the test set (X_test), and the R2 score, a common metric for regression models, was computed for evaluation Figure 10 .

```
# Train the Linear Regression model
model = LinearRegression()
model.fit(X_train, y_train)
```

Figure 10

B) Finding

Model Accuracy

1. **R2 Score:** The R2 score serves as a metric indicating the proportion of variability in the dependent variable (Work-Life Balance scores) that can be accounted for by the independent variables (other factors). Ranging from 0 to 1, a higher R2 score signifies a more effective model fit. In our case, the printed R2 score of 0.87 denotes the degree to which our model explains the variation in Work-Life Balance scores. This elevated R2 score suggests strong predictive performance, affirming the model's effectiveness in capturing and explaining the variability in Work-Life Balance scores.

2. **Scatter Plot:** A scatter plot has been generated to visually depict the association between the observed WLB scores and the model's predicted WLB scores. This graphical representation

aids in assessing the degree of alignment between the model's predictions and the true values in Figure 11.

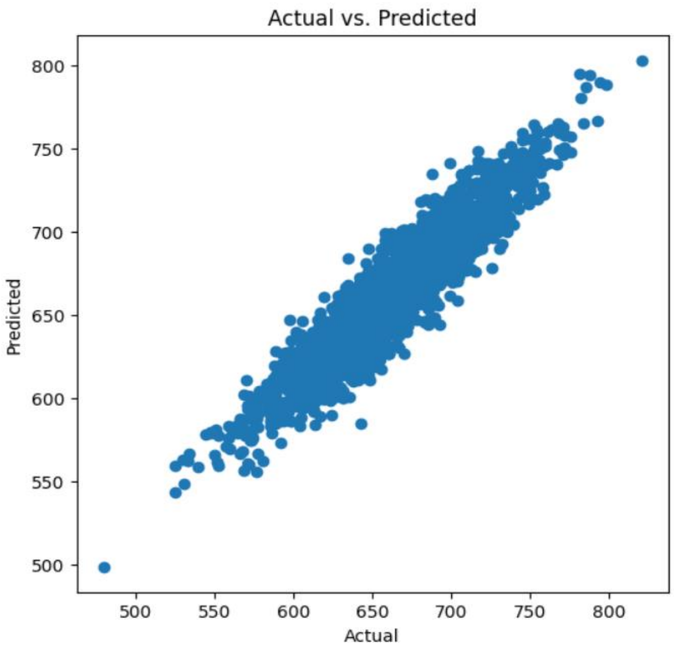

Figure 11

3. **Model Summary:** Below is the representation of the difference between the actual and predicted values generated by the model Figure 12 .

Tabular Form of Differences:

	Actual	Predicted	Difference
6252	600.6	596.045964	4.554036
4684	645.6	654.767473	-9.167473
1731	671.7	666.093179	5.606821
4742	707.4	696.148910	11.251090
4521	723.4	720.746745	2.653255

Figure 12

The summary of the model highlights that certain variables significantly impact the Work-Life Balance (WLB) Score.

- GENDER
- EMPLOYMENT
- SUFFICIENT_INCOME
- SALARY
- SLEEP_HOURS

The substantial "Coef" values associated with each of these variables suggest a notable influence on the model. This implies that these variables are likely pivotal factors contributing significantly to the determination of an individual's WLB Score.

C) Prediction for a new employee

Add a new employee with the following attributes for prediction of a new employee Figure 13.

```
# Create a new record for prediction
new_record = pd.DataFrame({
    'SUFFICIENT_INCOME': [1],
    'SALARY': [2],
    'TO_DO_COMPLETED': [6],
    'DAILY_STRESS': [5],
    'CORE_CIRCLE': [7],
    'SUPPORTING_OTHERS': [4],
    'SOCIAL_NETWORK': [4],
    'ACHIEVEMENT': [3],
    'FLOW': [6],
    'DAILY_STEPS': [7],
    'SLEEP_HOURS': [7],
    'LOST_VACATION': [0],
    'PERSONAL_AWARDS': [3],
    'TIME_FOR_HOBBY': [10],
    'HEALTHY_DIET': [1]
})
```

Figure 13

Predicted WLB Score for the newly added employee Figure 14.

```
# Make predictions for the new record
new_record_pred = model.predict(new_record)

# Display the predicted WLB score for the new record
print(f"Predicted WLB Score for the New Record: {new_record_pred[0]:.2f}")

✓ 0.0s

Predicted WLB Score for the New Record: 653.51
```

Figure 14

D) Recommendations for Improving WLB Score at FAU Airlines

Recommendations for Enhancing Work-Life Balance (WLB) Scores at FAU Airlines: To elevate the Work-Life Balance (WLB) scores of FAU Airlines' workforce, proposed strategies are drawn from insights derived from the Linear Regression model.

1. **Employee Engagement:** Encourage regular feedback from employees regarding their work-life balance, job satisfaction, and factors influencing their WLB ratings. Regular surveys and

fostering open communication channels can provide valuable insights into employee preferences and concerns.

2. **Flexible Work Arrangements:** Introduce policies that support flexible work arrangements, such as remote work options, flexible schedules, or compressed workweeks, allowing individuals to better manage their professional and personal commitments.

3. **Wellness Initiatives:** Implement wellness programs focused on supporting both the physical and mental well-being of employees. Initiatives may include stress management workshops, mindfulness sessions, and access to health and fitness resources.

4. **Skill Development:** Provide training and resources to enhance employees' time management skills, prioritize tasks effectively, and cultivate a healthy balance between work and personal life.

5. **Recognition and Incentives:** Recognize and reward employees who exhibit positive work-life balance behaviors. The implementation of these measures can foster a positive work environment, elevate work-life balance, and contribute to improved WLB scores, ultimately enhancing overall employee well-being.

2 Turnover

2.1 Employee Turnover

A) Voluntary Employee Turnover

Voluntary employee turnover refers to the intentional and independent decision made by an employee to resign or terminate their employment contract with an organization. This decision is driven by the employee's personal choice, reflecting a desire to discontinue their association with the company. It contrasts with involuntary turnover, where the employer initiates the termination of employment.

B) Significance and Concerns for FAU Airlines

Voluntary employee turnover is of high concern for FAU Airlines due to several reasons:

1. **Cost Implications:**

- Employee turnover is costly, involving expenses related to recruitment, onboarding, and training of new employees.
- Replacing a departing employee can cost between 30% and 400% of their annual salary, leading to a substantial financial burden for the company.

2. **Knowledge Loss:**

- Each departing employee takes with them valuable knowledge, skills, and experience, resulting in a loss of intellectual capital for the organization.
- The departure of high-performing employees may have a particularly significant impact on the airline's operations.

3. **Disruption to Operations:**

- Frequent turnover can disrupt the normal functioning of the organization, leading to potential inefficiencies and gaps in workforce continuity.
- It may affect team dynamics and collaboration, impacting overall productivity and performance.

C) Organization Equilibrium Theory and Causes of Employee Turnover

The Organization Equilibrium Theory posits that turnover occurs when individuals perceive that their contributions to the organization exceed the incentives they receive. In the context of FAU Airlines, potential causes for employees deciding to quit voluntarily include:

1. **Perceived Inequity:**

- Employees may feel that their contributions, such as skills, effort, and dedication, are not adequately recognized or rewarded by the organization.
- If employees believe that their efforts outweigh the benefits they receive, they may decide to leave in search of better opportunities.

2. **Lack of Job Satisfaction:**

- Dissatisfaction with work-related factors, such as job responsibilities, work environment, or organizational culture, can lead to employees seeking alternative employment.
- Unfulfilled expectations and unhappiness in the workplace contribute to a higher likelihood of voluntary turnover.

3. **Limited Growth Opportunities:**

- Employees may leave if they perceive a lack of opportunities for career advancement, skill development, or personal growth within the organization.
- Organizations that fail to provide a clear path for career progression may struggle to retain ambitious and career-oriented employees.

4. **Ineffective Employee Motivation:** If the organization fails to motivate and engage employees through recognition, rewards, and a positive work environment, employees may lose motivation and decide to leave.

5. **Poor Work-Life Balance:** In cases where employees feel overworked or experience challenges in maintaining a healthy work-life balance, they may seek employment elsewhere for improved well-being.

Understanding these causes allows FAU Airlines to develop targeted strategies and interventions to address the root issues contributing to voluntary employee turnover. This can involve enhancing recognition and rewards, improving job satisfaction, providing growth opportunities, and fostering a positive work environment.

2.2 Employee Turnover Theories

Case 1: Sarah - Job Embeddedness Theory

Sarah's decision to leave the small, innovative software company after its acquisition by FAU Airlines suggests a fit with the Job Embeddedness Theory.

Explanation of Job Embeddedness Theory

The Job Embeddedness Theory posits that individuals are less likely to leave an organization when they feel deeply embedded or connected to various aspects of their job and community. The dimensions of community fit, fit to the organization, community connection, connection to the organization, community-related waiver, and organization-related waiver provide insights into an individual's attachment to the workplace and surrounding community.

Application to Case 1

- **Community Fit:** Sarah may have felt a strong fit within the small, innovative software company's community.
- **Fit to the Organization:** The autonomy and casual work environment in the software company could have provided a perfect fit for Sarah's skills and talents.
- **Community Connection:** Her connection to the community might have been strong, making it difficult for her to leave.
- **Connection to the Organization:** The sudden acquisition and shift to a large, hierarchical corporation disrupted her connection to the organization.
- **Community-Related Waiver:** Sarah may have found it hard to leave the tight-knit community of the small software company.
- **Organization-Related Waiver:** The change in the company's culture and structure may have diminished Sarah's opportunities for advancement.

Conclusion

Sarah's decision aligns with the Job Embeddedness Theory, where her deep connection to the community, organization, and reluctance to give up her position played a significant role in her decision to leave.

Case 2: John - Turnover Event Theory

John's contemplation of quitting FAU Airlines despite aligning with its customer service priorities suggests an application of the Turnover Event Theory.

Explanation of Turnover Event Theory

The Turnover Event Theory explains how employees who remain in a company after turnover develop positive or negative attitudes towards their own turnover, leading to various consequences. The theory considers factors influencing employees' post-turnover attitudes.

Application to Case 2

- **Turnback Consideration:** John is contemplating quitting despite working for almost two years.
- **Alignment with Company Values:** John aligns with FAU Airlines customer service priorities, reflecting a positive attitude towards the organization.
- **Limited Social Connections:** John's limited social connections and meaningful relationships within the organization contribute to a negative attitude.
- **Consequences:** Negative attitudes may lead to increased thoughts of quitting and potential turnover.

Retention Strategies for FAU Airlines

- Encourage Social Connections: Implement initiatives to foster social connections, such as team-building activities, mentorship programs, and regular employee events.
- Enhance Communication: Improve communication channels to ensure employees feel connected and valued.

Recognition Programs: Implement employee recognition programs to acknowledge and reward contributions, enhancing job satisfaction.

- **Career Development Opportunities:** Provide opportunities for professional growth and career advancement to enhance employees' connection to the organization.
- **Feedback Mechanisms:** Establish feedback mechanisms to address concerns and gather insights for continuous improvement.

Conclusion

John's situation aligns with the Turnover Event Theory, emphasizing the need for FAU Airlines to address factors influencing his decision to quit, focusing on improving social connections and overall job satisfaction.

2.3 Employee Turnover Dataset

A) Data Understanding and Preprocessing (Refer to employee_turnover.ipynb)

Reads the " turnover_dataset.csv" CSV file into a panda Data Frame and checks for missing values in the data. Any rows with missing values are dropped from the Data Frame. Categorical data types are converted to numeric using LabelEncoder, and the Data Frame is updated accordingly in Figure 15 & Figure 16.

```
# Read the CSV file into a DataFrame
df = pd.read_csv("employee_turnover.csv")

# Check for any missing values in the DataFrame
df.isnull().values.any()

# Print the DataFrame and check for missing values
df.head()
```

✓ 0.0s

	job_satisfaction_level	salary_satisfaction_level	satisfaction_with_team_leader
0	0.38	0.36	2
1	0.80	0.91	4
2	0.11	0.13	2
3	0.72	0.86	4
4	0.37	0.32	1

Figure 15

```
# Convert categorical data types into numeric equivalents
df['Department'] = pd.Categorical(df['Department']).codes
df['salary'] = pd.Categorical(df['salary']).codes

df.dtypes
```

✓ 0.0s

```
job_satisfaction_level           float64
salary_satisfaction_level        float64
satisfaction_with_team_leader      int64
score_last_evaluation            float64
overtime_in_a_week                 int64
average_working_hours_monthly      int64
time_spend_company                 int64
work_accident                      int64
left                               int64
promotion_last_5years              int64
Department                          int8
salary                              int8
dtype: object
```

Figure 16

B) Finding

- **Average Job Satisfaction Level of Employees Who Left:** The average job satisfaction level of employees who left FAU Airlines is 0.44, which is slightly lower than the average job satisfaction level of employees who stayed 0.67 Figure 17.

```
Average Job Satisfaction Level of Employees Who Left: 0.44
```

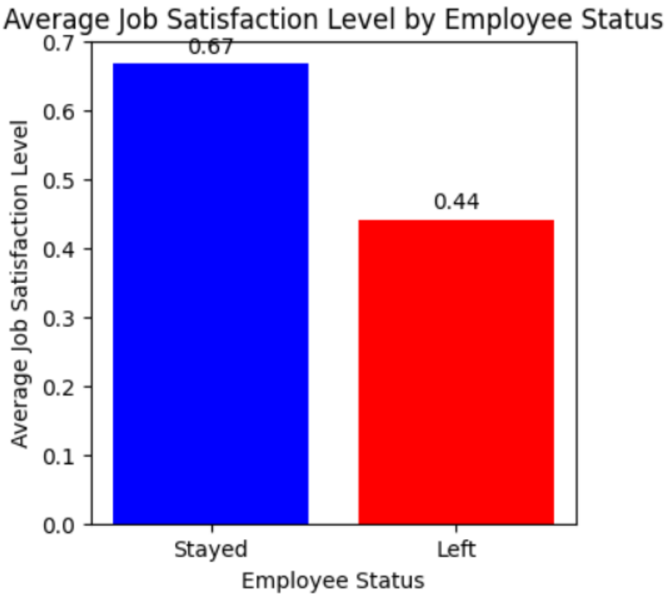

Figure 17

- **Average Salary Satisfaction Level of Employees Who Left:** The average salary satisfaction level of employees who left FAU Airlines is 0.5, which is also slightly lower than the average salary satisfaction level of employees who stayed 0.63.

```
Average Salary Satisfaction Level of Employees Who Left: 0.50
```

Figure 18

- **Average Time Spent with FAU Airlines for Employees Who Left:** Employees who left FAU Airlines on average spent 3.88 hours per week on average with the company, while employees who stayed on average spent 3.38 hours on average per week with the company Figure 19.

```
Average Time Spent with FAU Airlines for Employees Who Left: 3.88 years
```

Figure 19

- **Does Salary Play a Role for People Who Decide to Quit Their Job**?: Yes, salary does play a role in employee turnover. Employees who left FAU Airlines had a lower average salary than employees who stayed Figure 20.

```
Salary Effect on Turnover: 1.3458415009801177
```

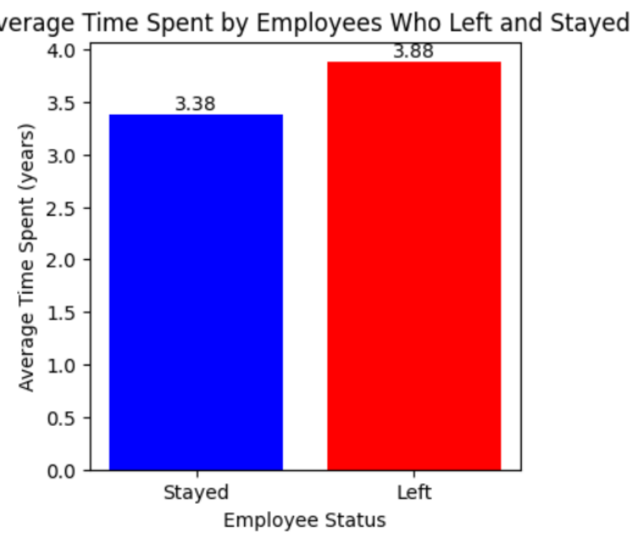

Figure 20

- **Correlation Matrix:** The correlation matrix in Figure 21 shows that the most correlated attributes to the left label are salary_satisfaction_level (-0.44), time_spend_company (-0.36), years_of_experience (-0.34), and promotion_last_5_years (-0.32). This means that employees who are less satisfied with their salary, have worked at FAU Airlines for a shorter period of time, have less work experience, and have not been promoted in the past five years are more likely to leave the company.

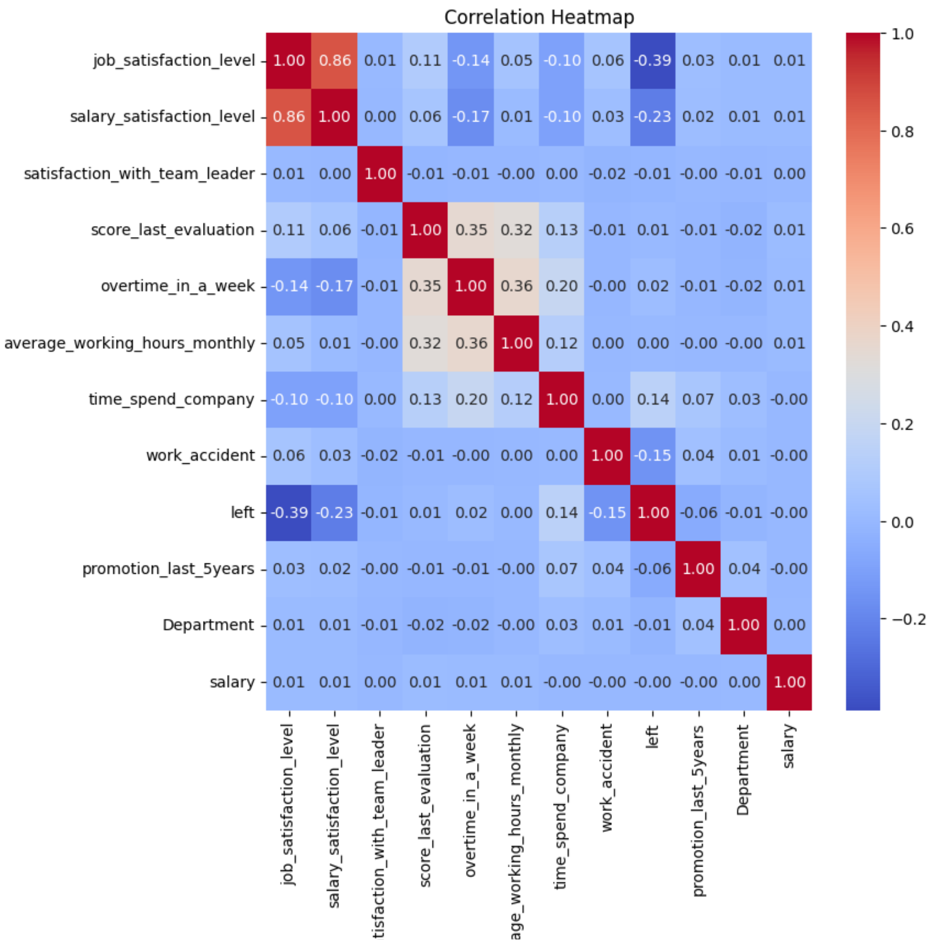

Figure 21

C) Conclusion

The analysis suggests that several factors, including job satisfaction, salary, time spent at the company, work experience, and promotion history, can contribute to employee turnover at FAU Airlines. To reduce turnover, the company may want to focus on improving job satisfaction, providing competitive salaries, and creating a culture of recognition and promotion.

2.4 Turnover prediction

To develop a predictive model for employee turnover, the choice of a machine learning algorithm is crucial. For this particular scenario, Logistic Regression, a widely used algorithm for binary classification tasks, is selected. Logistic Regression is well-suited for predicting binary outcomes, specifically discerning whether an employee will either leave (1) or stay (0).

In the process of predicting employee turnover at FAU Airlines, the following technical steps were undertaken:

A) Data Preparation

- Dividing the data into dependent and independent variables.
- Independent variables include all columns except the 'left' column.
- The 'left' column serves as the dependent variable Figure 22.

```
# Assume X as target column, and other columns Y as features
X = df.drop('left', axis=1)
y = df['left']
```

Figure 22

B) Data Splitting

- Splitting the data into training and testing sets (80-20 split) Figure 23.
- This allows for model training on a subset of data and evaluating its performance on unseen data.

```
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 23

C) Model Selection

- A Logistic Regression model was chosen for this analysis Figure 24.
- Parameters were tuned for increased max_iter and 'sag' solver to enhance convergence.

```python
# Train the Logistic Regression model with increased max_iter and 'liblinear' solver
model = LogisticRegression(max_iter=1000, solver='liblinear', random_state=42)
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)
```

Figure 24

D) Model Evaluation

The model was evaluated using key metrics shown in Figure 25.

- Accuracy Score
- Confusion Matrix
- Classification Report (precision, recall, f1-score)

```python
# Evaluate the model's performance
accuracy = accuracy_score(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred)
```

Figure 25

E) Results

- **Accuracy:** The model achieved an accuracy of 80% on the testing set.
- **Confusion Matrix:** Breakdown of true positives, true negatives, false positives, and false negatives.
- **Classification Report:** Detailed precision, recall, and f1-score for both classes (0: Stayed, 1: Left) in Figure 26.

```
Accuracy: 0.80
Confusion Matrix:
[[2131  163]
 [ 445  261]]
Classification Report:
              precision    recall  f1-score   support

           0       0.83      0.93      0.88      2294
           1       0.62      0.37      0.46       706

    accuracy                           0.80      3000
   macro avg       0.72      0.65      0.67      3000
weighted avg       0.78      0.80      0.78      3000
```

Figure 26

Discussion

- The Logistic Regression model exhibits an accuracy of 80%, demonstrating its ability to correctly classify instances.
- The precision and recall scores highlight the model's effectiveness in identifying both stayed and left cases.

Implications and Recommendations for FAU Airlines

- Factors such as Employee Satisfaction Level and Work Accidents significantly impact turnover.
- FAU Airlines should prioritize maintaining high job satisfaction levels and providing support to employees following accidents or shocks.
- Regular employee feedback surveys and flexible work arrangements can contribute to a healthier work environment and potentially reduce turnover.

In conclusion, the Logistic Regression model provides valuable insights into potential turnover factors, allowing FAU Airlines to take proactive measures to retain employees and improve overall organizational well-being.

References

The information and valuable insights provided throughout the lectures and notes have been instrumental in aiding the development and formulation of this report.

Declaration of Academic Integrity at the Schöller Endowed Chair for Information Systems

I hereby certify that I have prepared the submitted work independently, and without the unauthorized assistance of third parties, as well as without the use of unauthorized aids. The work has not been submitted in the same or similar form to any other examination authority, nor has it been accepted by any other examination authority as part of an examination.

The passages in the work, which have been taken from other sources in terms of wording or meaning, are identified by indicating the origin. This also applies to drawings, sketches, picture representations and sources from the Internet.

I am aware that the use of artificial intelligence is permitted for work at the Schöller Endowed Chair of Information Systems, Digitalization in Business and Society (esp. to improve the text written by myself). However, the intellectual core of the respective work has been developed by me, and the scientific methods that are part of the work have been carried out by myself. Furthermore, I have transparently communicated the aids used in the work.

Violations of the above-mentioned rules are to be qualified as deception or attempted deception and lead to an assessment of the examination with "failed". Further sanctions are possible in the case of multiple or particularly drastic violations of the rules by the examination board.

Muhammad Fahad Ali | Nurnberg 02.04.2024