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 17674768100

mfahad.ali@fau.de

22995110

Final Date of Submission: 2023-12-17

Contents

1 Personnel Planning

1.1 Future workforce needs

The HR Staffing Model could be generally used by any organization including FAU Airlines to determine the number of future employees needed.

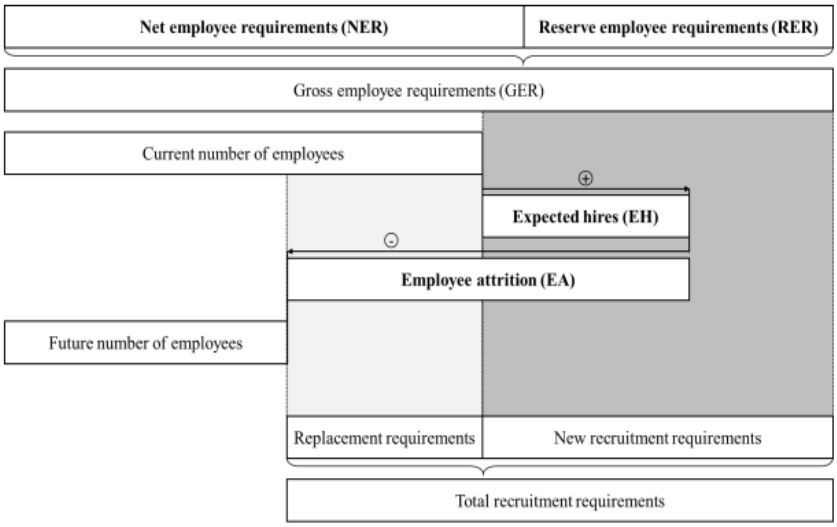

Figure 1.1

Figure 1.1 above cites the HR staffing model in detail where an organization can realize the number of employees needed in the future. The model takes into account the net employee requirements (NER) and reserve employee requirements (RER) to determine the gross employee requirements (GER). The model takes into account variables such as the current number of employed employees, expected hires (EH), and employee attrition (EA) to calculate the number of employees needed.

The organization needs to define further aspects to further operationalize this in practice. These include extracting all the possible variables of the environment the organization operates in and realizing the industry that the organization is a part of. The question to ask here is “How many and which employees do I need in my daily business and when?”. These variables are dependent on industry, individual organization, and whether it is an office or a store and further different aspects. For example, at the FAU Airline level, some variables that we need to extract in order to achieve our goal are as follows :

1. **Strategic Direction:** Define primary strategies aligned with FAU Airline's goals. Consider the internal and external environment, including social, political, economic, and technological factors. Assess the characteristics and capabilities of the current workforce.

2. **Assess Future Needs:** Analyze the demands and distribution of airline staff. Identify the specific skills and qualifications required for check-in agents and flight attendants roles at FAU Airline. Identify any gaps between the current workforce and future requirements.

3. **Gap-Closing Strategies:** Implement strategies to bridge the identified gaps. Forecast the demand for check-in agents and flight attendants based on FAU Airline's goals and growth plans. Implement recruitment strategies to acquire new check-in agents and flight attendants with the required skills.

4. **Development:** Address skill needs within the check-in agents and flight attendants through targeted development programs. Provide training and development opportunities for existing staff to meet future requirements.

5. **Determine Staff Needs:** Identify the specific number of check-in agents and flight attendants required based on patient care demands and service level. Consider factors such as the availability of staff, their working hours, and fluctuations in patient volume. Take into account seasonal variations or changes in tourist demand that may impact airline staffing levels.

By integrating these steps and variables, you can estimate the future number of check-in agents and flight attendants needed for FAU Airline. It's important to customize the model and adapt it according to FAU Airline's specific requirements, industry norms, and historical data.

1.2 FAU Airlines Staffing model.

Personnel planning is a critical aspect of organizational management, and the HR Staffing Model provides a systematic approach to calculate the future need for new employees. One key element of this model is the Gross Employee Requirements (GER), a metric derived from the Net Employee Requirements (NER) and Reserve Requirements (RER). The steps involved in calculating the Future Number of Employees for an airline that plans to hire new staff are:

Defining Gross Employee Requirements (GER):

Gross Employee Requirements represent the total number of new employees an organization needs to meet its future staffing demands. It is calculated by summing the Net Employee Requirements (NER) and the Reserve Requirements (RER). NER is the difference between the desired workforce size and the current workforce, while RER accounts for additional staff required to cover unexpected events or fluctuations in demand.

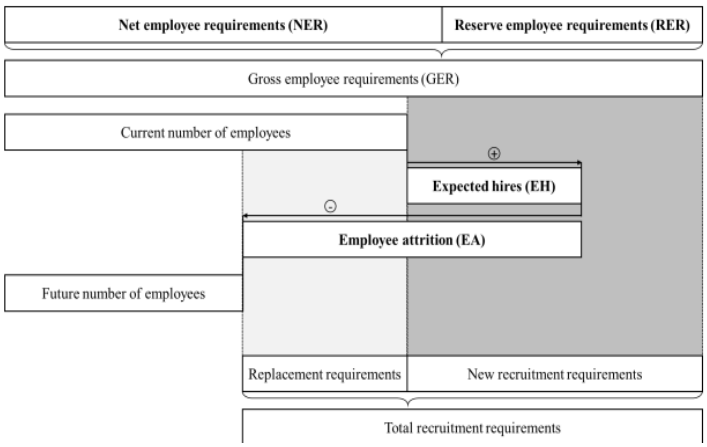

Figure 1.2

Determining GER using HR Staffing Model:

The HR Staffing Model involves several steps to calculate GER:

1. **Calculate Net Employee Requirements (NER):** Subtract the current number of employees from the desired workforce size. NER = Desired Workforce Size - Current Workforce.

2. **Calculate Reserve Requirements (RER):** Anticipate any unforeseen events, attrition, or fluctuations in demand and determine the additional staff needed to cover these situations.

3. **Calculate Gross Employee Requirements (GER):** Add NER and RER to obtain the total number of new employees required for future staffing.

Calculating Future Number of Employees for FAU Airline:

In the case of the airline planning to hire new check-in agents and flight attendants, the steps would be as follows:

1. Check-in Agents:

- Current Check-in Agents: 4
- Expected Hires (EH) for Check-in Agents: 20
- Attrition (EA) for Check-in Agents: 1
- Net Employee Requirements (NER) for Check-in Agents: C+EH = F+EA
  20 + 4 = F + 1
  After solving for F: **F=23**
- Reserve Requirements (RER) for Check-in Agents: Anticipate any additional needs.
- Gross Employee Requirements (GER) for Check-in Agents: NER + RER.

Therefore future number of Check-In agents are 23.

2. Flight Attendants:

- Current Flight Attendants: 6
- Expected Hires (EH) for Flight Attendants: 16
- Attrition (EA) for Flight Attendants: 2
- Net Employee Requirements (NER) for Flight Attendants: C+EH = F+EA
  16 + 6 = F + 2
  After solving for F: F=20
- Reserve Requirements (RER) for Flight Attendants: Anticipate any additional needs.
- Gross Employee Requirements (GER) for Flight Attendants: NER + RER.

The HR Staffing Model, through the calculation of Gross Employee Requirements, provides organizations with a systematic approach to determine the total number of new employees needed for future staffing. FAU airline can apply this model involves considering the current workforce, expected hires, and attrition to strategically plan and allocate resources for an optimal and efficient workforce.

Therefore future number of Flight Attendents agents are 20.

1.3 Recruiting new hires

Recruitment is a critical aspect of personnel management, and the choice between passive recruiting and active sourcing plays a pivotal role in shaping an organization's workforce. According to the concepts of passive recruiting and active sourcing we provide recommendations for FAU Airlines in their pursuit of hiring check-in agents and flight attendants Figure 1.3.

Figure 1.3

Passive Recruiting vs. Active Sourcing:

Passive recruiting, involves a more indirect marketing approach, targeting intrinsically motivated applicants with strong negotiating positions. On the other hand, active sourcing is characterized by a proactive recruitment strategy, where motivation relies on persuasion, albeit with potential challenges in negotiating positions Figure 1.3.

Recommendation for FAU Airlines:

Considering the dynamic and customer-oriented nature of the airline industry, an active sourcing approach is recommended for FAU Airlines. Proactively seeking and recruiting candidates allows the company to tailor its workforce to meet specific requirements and respond swiftly to changing demands.

Recruitment Channels for FAU Airlines:

To implement active sourcing successfully, FAU Airlines should leverage a mix of online and offline channels. Recommended channels include:

1. **Online Job Portals:** Utilize popular job portals to reach a wide audience of potential candidates actively seeking employment in the aviation sector.

2. **Social Media Platforms:** Leverage social media platforms to showcase the company culture, highlight employee experiences, and engage with potential candidates directly.

3. **Aviation Industry Events:** Attend and participate in industry-specific events and job fairs to connect with individuals passionate about aviation careers.

4. **Employee Referral Programs:** Encourage current employees to refer qualified candidates, tapping into their networks and promoting a sense of community within the organization.

5. **Collaboration with Training Institutes:** Partner with aviation training institutes to identify and recruit individuals who have received relevant training and education.

Adopting an active sourcing strategy aligns with the dynamic needs of the airline industry and positions FAU Airlines to build a skilled and adaptable workforce. Leveraging a diverse range of recruitment channels will further enhance the company's ability to attract top talent for check-in agents and flight attendants, ensuring a positive and efficient passenger experience.

1.4 Opened check-in desks per shift

A) To solve this problem we must describe the objective and the approach to solve the problem.

- **Objective:** Determine the optimal number of check-in agent required for each shift at the FAU airlines, in a way that satisfies the constraints.
- **Solution Approach:** This is an optimization problem that requires Linear Programming (LP). We will use PuLP, which is a Python library used for solving tasks related to Linear Programming.

B) Now we consider the input parameters and the decision variable

- **Input Parameters:** Number of shifts, Average number of customer per hour, Time to entertain easch customer 2.4 minutes, Wage rate is assumed 1, Service level.
- **Decision variable:** Optimal number of check-in agent for each shift in a day.

C) Now, Getting to know the dataset tells us

- Three 6 hours shifts for every day.
- Dataset contain average number of customer per hour.
- **Service level** determines the number of employees required for certain customer.

In our case, a check-in agent take 2.4 minutes per customer.

Service level, therefore, becomes 25 customer per hour, which has been added to the dataset.

D) Translation of the problem into a Linear Problem

- **Decision Variable:** The number of check-in agents required for each shift.
- **Objective Function:** Minimize the cost of wages paid to all employees by finding the optimal number.
- **Constraints:** Service level and the average number of customers in each time window.

E) Programming Steps (Refer to file “fau_airline_shifts.ipynb”) :

- Import libraries such as pandas and pulp, and then import the dataset into a data Frame Figure 1.4.

```
1  #importing necessary libraries.
2  import pandas as pd
3  from pulp import *
4  import math as mt
```

Figure 1.4

- Read and pre process data Figure 1.5.

```
# initalize and load dataset and remove unwanted rows and replace the Nan values

week = ['Mon', 'Tue', 'Wen', 'Thu', 'Fri', 'Sat','Sun']
time_pp = 2.4
df = pd.read_csv('fau_airlines_shifts.csv')
df = df.iloc[:18]
df = df.fillna(0).applymap(lambda x: 1 if x == "X" else x)
```

Figure 1.5

- Calculate service level, extract pessenger and shifts details Figure 1.6.

```
# calculate service level, extract pessenger and shifts details
weeks_days = df.columns[4:]

# convert column headers to an array
weeks_days = weeks_days.to_numpy()

passengers= df[weeks_days].values

shifts_columns= df.columns[1:4]
shifts = df[shifts_columns]
shifts_pd = len(shifts_columns)

shift_slots = len(df[df[shifts_columns] == 1])
service_level = int(60 / time_pp)

print("Shift Per Day: ",shifts_pd)
print("Service Level: ",service_level)
print("Total Shift Slots: ",shift_slots)
```

✓ 0.0s

```
Shift Per Day:  3
Service Level:  25
Total Shift Slots:  18
```

Figure 1.6

- Creating dictionary using LpVariable.dicts. {'0_0': num_workers_0_0, '0_1': num_workers_0_1, '0_2': num_workers_0_2, '1_0': num_workers_1_0, '1_1': num_workers_1_1, '1_2': num_workers_1_2, '2_0': num_workers_2_0, '2_1': num_workers_2_1, '2_2': num_workers_2_2, '3_0': num_workers_3_0, '3_1': num_workers_3_1, '3_2': num_workers_3_2, '4_0': num_workers_4_0, '4_1': num_workers_4_1, '4_2': num_workers_4_2, '5_0': num_workers_5_0, '5_1': num_workers_5_1, '5_2': num_workers_5_2, '6_0': num_workers_6_0, '6_1': num_workers_6_1, '6_2': num_workers_6_2}.

- Agent must entertain average no of pessengers Figure 1.7

```
# agent must entertain average no of pessengers
for day_of_week in range (0, 7) :
    for t in range(shift_slots) :
        prob += lpSum(
            [shifts.iloc[t, j] * num_workers[f'{day_of_week}_{j}']
             for j in range(shifts_pd)]) >= mt.ceil(passengers[t][day_of_week] / service_level)

print(prob)
```

Figure 1.7

- `MINIMIZE`
  1*num_workers_0_0 + 1*num_workers_0_1 + 1*num_workers_0_2 + 1*num_workers_1_0 + 1*num_workers_1_1 + 1*num_workers_1_2 + 1*num_workers_2_0 + 1*num_workers_2_1 + 1*num_workers_2_2 + 1*num_workers_3_0 + 1*num_workers_3_1 + 1*num_workers_3_2 + 1*num_workers_4_0 + 1*num_workers_4_1 + 1*num_workers_4_2 + 1*num_workers_5_0 + 1*num_workers_5_1 + 1*num_workers_5_2 + 1*num_workers_6_0 + 1*num_workers_6_1 + 1*num_workers_6_2 + 0

F) Findings:

- The linear programming model yields the optimal number of check-in agents needed for each shift at FAU airline Figure 1.8.

```
#formulating results
for index in num_workers:
    index_parts = index.split('_')
    day_of_week = int(index_parts[0])
    day = week[day_of_week]
    shift = index_parts[1]
    print (f"The number of Agents needed for day {day} shift {shift} is {int(num_workers[index].value())} workers")
```

```
The number of Agents needed for day Mon shift 0 is 10 workers
The number of Agents needed for day Mon shift 1 is 5 workers
The number of Agents needed for day Mon shift 2 is 11 workers
The number of Agents needed for day Tue shift 0 is 1 workers
The number of Agents needed for day Tue shift 1 is 4 workers
The number of Agents needed for day Tue shift 2 is 6 workers
The number of Agents needed for day Wen shift 0 is 2 workers
The number of Agents needed for day Wen shift 1 is 2 workers
The number of Agents needed for day Wen shift 2 is 5 workers
The number of Agents needed for day Thu shift 0 is 11 workers
The number of Agents needed for day Thu shift 1 is 11 workers
The number of Agents needed for day Thu shift 2 is 8 workers
The number of Agents needed for day Fri shift 0 is 12 workers
The number of Agents needed for day Fri shift 1 is 11 workers
The number of Agents needed for day Fri shift 2 is 3 workers
The number of Agents needed for day Sat shift 0 is 10 workers
The number of Agents needed for day Sat shift 1 is 8 workers
The number of Agents needed for day Sat shift 2 is 3 workers
The number of Agents needed for day Sun shift 0 is 2 workers
The number of Agents needed for day Sun shift 1 is 2 workers
The number of Agents needed for day Sun shift 2 is 5 workers
```

Figure 1.8

G) **Conclusion:** The ultimate outcome illustrates the necessary number of Agents needed for each shift daily, determined by the average customer count and service level.

A) To solve second part of the problem:

- Made changes in datset in file “fau_airline_shifts.ipynb” Figure 1.9.

```
#formulating results
for index in num_workers:
    index_parts = index.split('_')
    day_of_week = int(index_parts[0])
    day = week[day_of_week]
    shift = index_parts[1]
    print (f"The number of Agents needed for day {day} shift {shift} is {int(num_workers[index]. value())} workers")
```

```
The number of Agents needed for day Mon shift 0 is 16 workers
The number of Agents needed for day Mon shift 1 is 8 workers
The number of Agents needed for day Mon shift 2 is 17 workers
The number of Agents needed for day Tue shift 0 is 1 workers
The number of Agents needed for day Tue shift 1 is 5 workers
The number of Agents needed for day Tue shift 2 is 8 workers
The number of Agents needed for day Wen shift 0 is 2 workers
The number of Agents needed for day Wen shift 1 is 3 workers
The number of Agents needed for day Wen shift 2 is 8 workers
The number of Agents needed for day Thu shift 0 is 17 workers
The number of Agents needed for day Thu shift 1 is 17 workers
The number of Agents needed for day Thu shift 2 is 12 workers
The number of Agents needed for day Fri shift 0 is 18 workers
The number of Agents needed for day Fri shift 1 is 17 workers
The number of Agents needed for day Fri shift 2 is 4 workers
The number of Agents needed for day Sat shift 0 is 15 workers
The number of Agents needed for day Sat shift 1 is 13 workers
The number of Agents needed for day Sat shift 2 is 5 workers
The number of Agents needed for day Sun shift 0 is 2 workers
The number of Agents needed for day Sun shift 1 is 3 workers
The number of Agents needed for day Sun shift 2 is 8 workers
```

Figure 1.9

B) Findings:

- The linear programming model yields the optimal number of check-in agents needed for each shift at FAU airline Figure 1.10.

```
#task 02
# make changes in dataset

df['Avg_Passenger_No_3'] = df['Avg_Passenger_No_1']
df['Avg_Passenger_No_7'] = df['Avg_Passenger_No_6']

passengers *= 2
service_level = 32
```

Figure 1.10

C) Conclusion: Similarly, the ultimate outcome illustrates the necessary number of Agents needed for each shift daily, determined by the average customer count and service level.

2 Onboarding and Performance

2.1 A New Hire

A well-managed onboarding process acts as the building block for the new employee to perform well in their job and not end up as a potential turnover by long-term retaining. According to Bauer (2010), an organization should provide the employee with valuable connections a clarified agenda, an understanding of the culture, and the required policy-related material. This outlines that if an employee is clear about their role expectations, and if they have a sense of self-confidence, then they will perform well with concise goals in mind and feel welcomed. The theory by Van Maanen and Schein (1979) discusses the six-dimensional model for the structure of onboarding. The theory, on further work by Jones (1986) categorizes the six dimensions into factors such as context, content, and social aspects as a whole of newcomer information seeking. Bauer et al (2007) further expanded the three factors into role clarity, selfefficacy, and social acceptance. These models combined highlight the importance of the onboarding process by outlining that if the six-dimensional organizational socialization tactics and 3 factors of newcomer information seeking are managed efficiently, then the outcome by the new employee is increased performance, job satisfaction, commitment to the organization, the intention to remain, and a decrease in turnover probability Figure 2.1.

Dimension	Tasks	
Collective vs. Individual	Group learning	Individual learning
Formal vs. Informal	Formal approach (socialization event is separate from actual job)	Informal approach (on the job training)
Sequential vs. Random	Sequential learning events	Random events (order is unknown)
Fixed vs. Variable	Fixed timetable	Variable timeline of events
Serial vs. Disjunctive	Serial approach offers a mentor	Disjunctive approach leaves new hires to navigate on their own
Investiture vs. Divestiture	The investiture approach provides new hires with positive feedback from experienced peers.	Divestiture approach provides a newcomer with negative feedback until full adaptation.

Figure 2.1

Importance of Socialization During Onboarding:

Effective socialization during onboarding is crucial for several reasons:

1. Cultural Integration: Socialization fosters an understanding of the organization's culture, values, and norms, aiding in the smooth integration of new employees.

2. Team Building: It promotes team building and collaboration, creating a positive work environment where employees feel connected and supported.

3. Role Clarity: Socialization helps clarify roles and responsibilities, ensuring that new hires understand their positions within the organization.

4. Job Satisfaction: Employees who undergo successful socialization tend to experience higher job satisfaction, contributing to overall organizational success.

Factors Contributing to Successful Onboarding:

Based on scientific literature, key factors for successful onboarding include:

1. Clear Communication: Transparent and consistent communication helps new hires understand their roles, expectations, and the organization's values.

2. Mentorship Programs: Implementing mentorship programs provides new employees with guidance and support, aiding in their professional development.

3. Structured Training: Well-structured training programs ensure that new hires are equipped with the necessary skills and knowledge for their roles.

4. Integration of Technology: Utilizing technology enhances the onboarding experience, making information easily accessible and improving efficiency.

Recommended Socialization Approach for FAU Airlines:

FAU Airlines should adopt a comprehensive socialization approach that includes:

1. Orientation Sessions: Conduct orientation sessions to introduce new hires to the company's mission, values, and organizational structure.

2. Buddy System: Implement a buddy system pairing new hires with experienced employees to provide guidance and foster camaraderie.

3. Interactive Training Modules: Develop interactive training modules incorporating multimedia elements for an engaging onboarding experience.

4. Cultural Immersion Activities: Organize activities that immerse new employees in the company's culture, such as team-building exercises and cultural awareness sessions.

Recommended Welcome Events:

To welcome everyone on board, FAU Airlines can organize events such as:

1. Welcome Luncheon: Host a welcome luncheon where new hires can meet key team members and establish initial connections.

2. Team-Building Workshops: Conduct team-building workshops to facilitate collaboration and communication among new and existing employees.

3. Company Tour: Arrange a company tour to familiarize new hires with different departments and facilities.

4. **Welcome Packages:** Provide welcome packages with essential information, company merchandise, and a personalized welcome letter.

Prioritizing socialization during onboarding is pivotal for FAU Airlines to ensure the seamless integration and long-term success of new check-in agents and flight attendants. Adopting the recommended socialization approach and organizing welcoming events will contribute to a positive onboarding experience, ultimately benefiting both the company and its employees.

2.2 Recommender System

This problem requires to be solved through a content-based recommender system that will help our new hire, emp_100, to identify the colleagues with whom they share similarities so that their performance could be increased and a smooth onboarding process is provided to them.

A) Programming Steps (Refer to file “fau_airline_onboarding.ipynb”) :

- Import libraries such as pandas and sklearn, then import the dataset into a data frame Figure 2.2.

```
#importing necessary libraries.
import pandas as pd
import numpy as np

# import TfidfVector from sklearn.
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import linear_kernel
```

✓ 0.0s

Figure 2.2

- Read and pre process dataset Figure 2.3.

```
#load dataset and prints columns
employees = pd.read_csv("fau_airlines_onboarding.csv")
print(employees.columns)
```

✓ 0.0s

```
Index(['id', 'study degree', 'previous_experience', 'civil_status',
       'personality_traits', 'hobbies', 'favourite sport'],
      dtype='object')
```

Figure 2.3

- Join columns, and return a new column with all the charactersFigure 2.4.

```
#joined column and a new column results from them
def create_soup(x):
    return ''.join(x['study degree']) + ' ' + ''.join(x['previous_experience']) + ' ' + ''.join(x['civil_status']) + ' '
employees['soup'] = employees.apply(create_soup, axis=1)
```

Figure 2.4

- We now calculate the Term-Frequency-Inverse Document Frequency Figure 2.5.

```
#TfidfVectorizer class that willbe used to creat TF-IDF vector
tfidf = TfidfVectorizer(stop_words='english')
tfidf_matrix = tfidf.fit_transform(employees['soup'])
tfidf_matrix.shape
```

✓ 0.0s

```
(100, 107)
```

Figure 2.5

- Calculate the cosine similarity between all pairs Figure 2.6.

```
# Calculate the cosine similarity between all pairs
cosine_sim = linear_kernel(tfidf_matrix, tfidf_matrix)
```
✓ 0.0s

Figure 2.6

- Construct a reverse map of indices and employee Ids Figure 2.7.

```
# construct a reverse map of indices and employee IDs
indices = pd.Series(employees.index, index=employees['id']).drop_duplicates()
```
✓ 0.0s

Figure 2.7

- The function get_recommendations will be used to compute the similarity score of the newcomer to the existing employees Figure 2.8.

```
def get_recommendations(ID, cosine_sim=cosine_sim):

    # get the index of the employee that matches the employee ID
    IDx = indices[ID]

    # get the pairwise similarity scores of all employees with the specified employee ID
    sim_scores = list(enumerate(cosine_sim[IDx]))

    # sort employees based on the similarity scores
    sim_scores = sorted(sim_scores, key=lambda x: x[1], reverse=True)

    # get the scores of the three most similar employees
    sim_scores = sim_scores[1:4]

    # get employee indices
    employees_indices = [i[0] for i in sim_scores]

    # return the top three most similar employees
    return employees['id'].iloc[employees_indices]
```

Figure 2.8

B) Findings:

- The algorithm returns the following results after a successful run Figure 2.9.

```
recomm=get_recommendations('emp_100', cosine_sim)
print(recomm)
```
✓ 0.0s

```
41    emp_042
88    emp_089
75    emp_076
Name: id, dtype: object
```

Figure 2.9

- The above results indicate that employees emp_042, emp_089, and emp_076 are the top three recommendations for the newcomer to interact with Figure 2.10.

```
1 employees[employees["id"].isin(['emp_042','emp_089','emp_076'])]
```

✓ 0.0s Python

	id	study degree	previous_experience	civil_status	personality_traits	hobbies	favourite sport	soup
41	emp_042	media	Advanced beginner.	in a relationship	Extrovert, expresses themselves easily, outgoi...	Interest centered around learning, reading and...	tennis	mediaAdvanced beginner.in a relationshipExtrov...
75	emp_076	psychology	Novice.	in a relationship	Extrovert, expresses themselves easily, outgoi...	Interest centered around learning, reading and...	tennis	psychologyNovice.in a relationshipExtrovert, e...
88	emp_089	economics	Advanced beginner.	in a relationship	Extrovert, expresses themselves easily, outgoi...	Interest centered around learning, reading and...	swimming	economicsAdvanced beginner.in a relationshipEx...

Figure 2.10

C) Programming Steps for Task 02 (Refer to file "fau_airline_onboarding.ipynb") :

- Add new employee (emp_101) to the dataset Figure 2.11.
- Then, perform all steps listed above.

```
#task 02
#add new employee

new_employee = {
    'id': 'emp_101',
    'study degree': 'communication',
    'previous_experience': 'Competent.',
    'civil_status': 'single',
    'personality_traits': 'Helpful, emotionally intelligent and gets along well with others',
    'hobbies': 'Interest centered around fitness, health, and wellbeing',
    'favourite sport': 'swimming'
}

# Convert the dictionary to a DataFrame and append it to the original DataFrame
new_employee_df = pd.DataFrame([new_employee])
employees = pd.concat([employees, new_employee_df], ignore_index=True)
employees
```

Figure 2.11

D) Findings:

- The algorithm returns the following results after a successful run Figure 2.12.

```
1 recomm=get_recommendations('emp_101', cosine_sim)
2 print(recomm)
```

✓ 0.0s

```
38    emp_039
21    emp_022
62    emp_063
Name: id, dtype: object
```

Figure 2.12

- The above results indicate that employees emp_042, emp_089, and emp_076 are the top three recommendations for the newcomer to interac with Figure 2.13.

```
1 employees[employees["id"].isin(['emp_039','emp_022','emp_063'])]
```

✓ 0.0s Python

	id	study degree	previous_experience	civil_status	personality_traits	hobbies	favourite sport	soup
21	emp_022	business	Advanced beginner.	single	Helpful, emotionally intelligent and gets alon...	Interest centered around fitness, health, and ...	swimming	businessAdvanced beginner.singleHelpful, emoti...
38	emp_039	economics	Expert.	single	Helpful, emotionally intelligent and gets alon...	Interest centered around fitness, health, and ...	swimming	economicsExpert.singleHelpful, emotionally int...
62	emp_063	economics	Advanced beginner.	single	Helpful, emotionally intelligent and gets alon...	Interest centered around fitness, health, and ...	tennis	economicsAdvanced beginner.singleHelpful, emot...

Figure 2.13

2.3 Factors that affect employee performance

Factors Affecting Employee Performance:

Employee performance is a multifaceted aspect influenced by various factors. Diamantidis & Chatzoglou's model and leadership style, provides insights into the factors with the highest impact on employee performance in the airline industry and offers recommendations for FAU Airlines to enhance the performance of their employees Figure 2.14.

1. Diamantidis & Chatzoglou's Model: The proposed model suggests that job-related factors have the strongest impact on employee performance, followed by employee-related factors and firm/environment-related factors. Additionally, firm/environment-related factors significantly affect both job-related and employee-related factors, while job-related factors strongly influence employee-related factors.

2. Leadership Style: Leadership style is a crucial factor influencing employee performance. Transformational leadership, characterized by idealized influence, intellectual stimulation, and inspirational motivation, has been associated with positive employee performance. Leaders who can create a shared vision, articulate clear goals, empower employees, and exhibit ethical behavior contribute to enhanced performance.

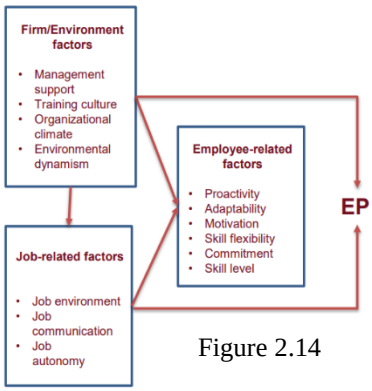

Figure 2.14

Based on my opinion and considering the dynamic nature of the airline industry, job-related factors such as effective training, clear communication, and job design are likely to have the highest impact on employee performance. Ensuring that employees are well-trained, understand their roles clearly, and have the necessary tools for their tasks is crucial in this fast-paced environment.

Recommendations for FAU Airlines:

1. Comprehensive Training Programs: Invest in comprehensive training programs that not only cover technical skills but also emphasize the company's values, service

standards, and safety protocols. This will empower employees to perform at their best from the start.

2. Clear Communication Channels: Establish clear communication channels to ensure that information flows seamlessly across all levels. This includes regular updates on industry trends, company news, and feedback mechanisms for employees to express their concerns or suggestions.

3. Job Redesign for Efficiency: Continuously evaluate and redesign job roles to enhance efficiency and align with industry best practices. This could involve optimizing workflows, incorporating new technologies, and ensuring that employees have the necessary resources to perform their tasks effectively.

4. Transformational Leadership Training: Provide leadership training programs that focus on transformational leadership behaviors. This can empower leaders within the organization to inspire and motivate employees, creating a positive and high-performance work culture.

5. Recognition and Rewards Programs: Implement recognition and rewards programs to acknowledge and appreciate outstanding employee performance. This not only boosts morale but also reinforces a culture of excellence.

Understanding and addressing the factors that impact employee performance is crucial for the success of any organization, especially in the demanding environment of the airline industry. By implementing the recommended strategies, FAU Airlines can foster an environment that promotes high employee performance, contributing to the overall success and competitiveness of the company.

2.4 Employee performance analysis

The FAU Airline management aims to initiate a data science project to analyze the current employee data and determine the core underlying causes of performance issues. The project aims to achieve three main insights.

- Employee job role performance analysis.
- Identification of top important factors affecting employee performance.
- Development of a supervised machine learning model to predict employee performance.

A) Expolatory Data Analysis (Refer to file “fau_airline_performance.ipynb”) :

- Import libraries such as pandas, matplotlib and sklearn, then import the dataset into a data frame Figure 2.15.

```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
```

Figure 2.15

- Read and pre process dataset Figure 2.16.

```
# Load the dataset
df = pd.read_csv('fau_airlines_performance.csv')
df.columns
```

```
Index(['EmpNumber', 'Age', 'Gender', 'MaritalStatus', 'EmpJobRole',
       'DistanceFromHomeKm', 'EmpJobSatisfaction',
       'EmpEnvironmentSatisfaction', 'EmpRelationshipSatisfaction',
       'EmpJobInvolvement', 'EmpJobExperience',
       'EmpHourlyRate -> change according to the age & experience', 'OverTime',
       'EmpLastSalaryHikePercent', 'EmpWorkLifeBalance',
       'YearsSinceLastPromotion', 'Attrition', 'PerformanceRating'],
      dtype='object')
```

Figure 2.16

- Employee Job Role Performance Analysis Figure 2.17 and Figure 2.18.

```
# Assuming df is your DataFrame with the necessary data
# Perform Exploratory Data Analysis (EDA)
# Print mean performance of each job role in a column chart
mean_performance_by_role = df.groupby('EmpJobRole')['PerformanceRating'].mean()

# Plotting the horizontal bar chart
plt.figure(figsize=(10, 6))
bars = mean_performance_by_role.sort_values().plot(kind='barh', color='skyblue')

# Annotating the bars with their respective values
for bar in bars.patches:
    plt.text(bar.get_width(), bar.get_y() + bar.get_height() / 2,
             f'{bar.get_width():.2f}',  # Format the value to two decimal places
             va='center', ha='left', color='black', fontsize=10)

plt.title('Mean Performance by Job Role')
plt.xlabel('Mean Performance Rating')
plt.ylabel('Job Role')
plt.show()
```

Figure 2.17

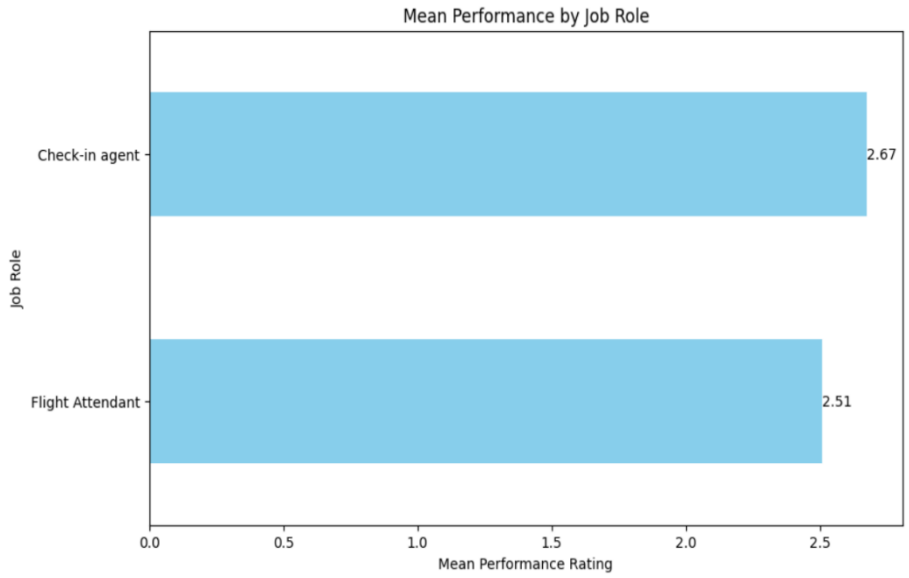

Figure 2.18

- Check null values Figure 2.19.

```
# Check for missing values
missing_values = df.isnull().sum()
```

Figure 2.19

- Calculate the correlation coefficient of the variables to the performance rating on Figure 2.20 and Figure 2.21.

```
# Create and print a correlation matrix
correlation_matrix = df.corr()
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=False, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Matrix')
```

Figure 2.20

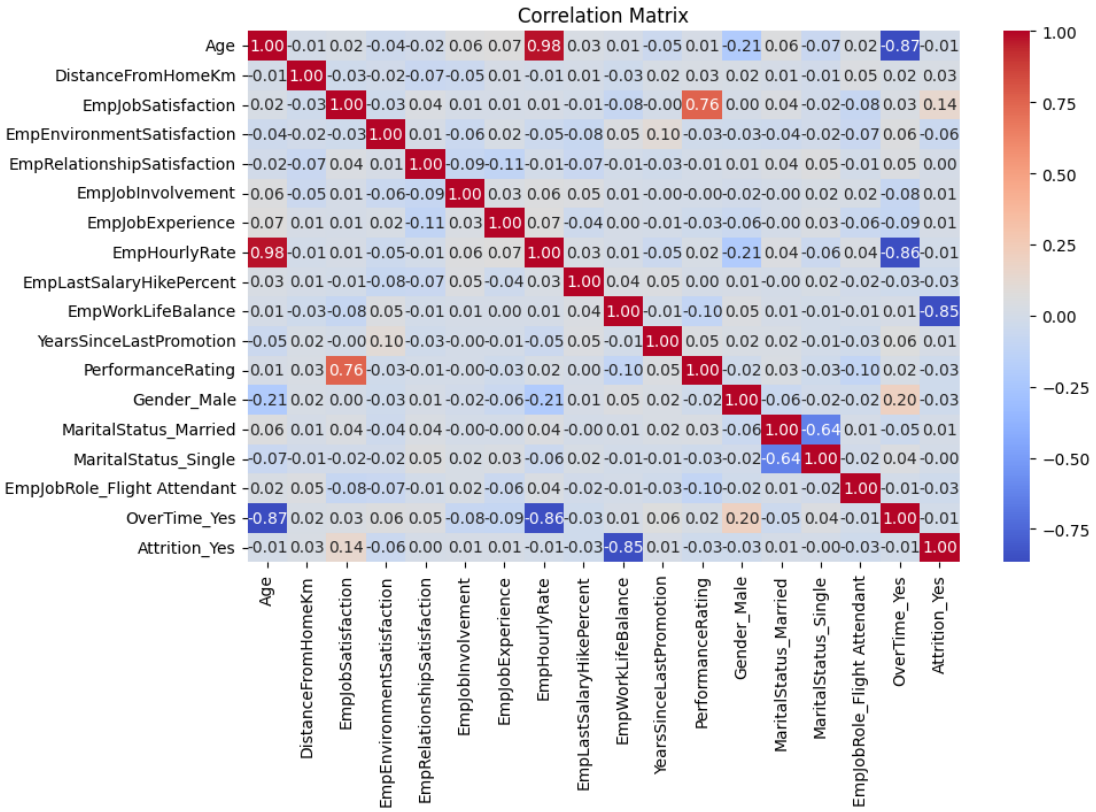

Figure 2.21

- Determine the variables highly correlated with PerformanceRating Figure 2.22

```
# Determine the variables highly correlated with PerformanceRating

# Assuming correlation_matrix is your DataFrame with the correlation matrix
# performance_correlation is the Series containing correlations with PerformanceRating

# Specify the variables of interest
performance_correlation = correlation_matrix['PerformanceRating'].sort_values(ascending=False)
variables_list = ['EmpJobSatisfaction', 'EmpEnvironmentSatisfaction', 'EmpRelationshipSatisfaction']

# Filter the correlation values for the specified variables
selected_correlations = performance_correlation[variables_list]

# Create a DataFrame with the selected correlations
correlation_table = pd.DataFrame({'Variable': selected_correlations.index, 'Correlation': selected_correlations.values})

# Display the table
print("Correlation with Performance Rating:")
print(correlation_table)

```

```
Correlation with Performance Rating:
                    Variable  Correlation
0          EmpJobSatisfaction     0.758983
1   EmpEnvironmentSatisfaction    -0.025695
2  EmpRelationshipSatisfaction    -0.005870
```

Figure 2.22

- Cloumns with the heighest impact on Performance Figure 2.23

```
#cloumns with the heighest impact on Performance
pos = []
print("cloumns with the heighest impact on Performance")

for column in correlation_matrix.columns:
    if column != 'PerformanceRating':
        if(correlation_matrix[column]['PerformanceRating'] >= 0.1
           ) or (
            correlation_matrix[column]['PerformanceRating'] <= -0.09
           ):
               pos.append(column)
               print(column, '=', correlation_matrix[column]['PerformanceRating'])

```

```
cloumns with the heighest impact on Performance
EmpJobSatisfaction = 0.7589828944507343
EmpWorkLifeBalance = -0.10161238394247978
EmpJobRole_Flight Attendant = -0.10469536063631464
```

Figure 2.23

B) Supervised machine learning model:

The trained MLP classifier can predict employee performance ratings based on the identified significant factors. The model's accuracy score provides an estimate of its performance in predicting performance ratings.

- Creating machine learning model using random forest classifier Figure 2.24.

```
# Machine Learning Model (Random Forest Classifier)
# Assuming 'PerformanceRating' is the target variable
X = df.drop(['PerformanceRating'], axis=1)
y = df['PerformanceRating']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
```

✓ 0.0s

Figure 2.24

- Train the random forest classifier Figure 2.25.

```
# Train the Random Forest Classifier
rf_classifier = RandomForestClassifier(random_state=42)
rf_classifier.fit(X_train, y_train)

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

✓ 0.2s

Figure 2.25

- Generate a classification report Figure 2.26.

```
# Generate a classification report
class_report = classification_report(y_test, y_pred, labels=[2, 3, 4, 5], target_names=[str(i) for i in range(2, 6)])
print("\nClassification Report:\n", class_report)
```

✓ 0.0s

```
Classification Report:
              precision    recall  f1-score   support

           2       0.72      0.68      0.70       105
           3       0.49      0.55      0.51        66
           4       0.40      0.33      0.36         6
           5       1.00      1.00      1.00        13

    accuracy                           0.64       190
   macro avg       0.65      0.64      0.64       190
weighted avg       0.65      0.64      0.65       190
```

Figure 2.26

C) Conclusion:

The model yields an accuracy of 68%. Therefore the model could be used for predicting the performance rating of a new employee. It will yield positive predictions for the employees with a probability of 68%. Overall, this project will help FAU Airline gain valuable insights into employee performance, identify factors influencing performance, and predict future employee performance.

3 References

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

4 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 12. 17.2023