Tegevuskäik – mida täpsem ja selgem, seda täpsem tulemus

1. Could you make me a spreadsheet named - Contract Table, containing information about 400 employees in an IT company. The columns should be:

Employee ID

Contract ID

Gender

Full name

Category

Position

Level

Workload

Period Date

Can you give me a CSV file for that information?

1. Now based on this information could you make me another spreadsheet named – Salary Table and the columns should be :

Employee ID

Contract ID

Salary

Benefits

Period Date

Can you give me a CSV file for that information?

**ESIMENE PYTHONI KOOD**

import pandas as pd

import random

from faker import Faker

from datetime import datetime, timedelta

# Initialize Faker

fake = Faker()

random.seed(42)

Faker.seed(42)

# Number of employees

num\_employees = 400

# Sample data pools

genders = ['Male', 'Female', 'Other']

positions = ['Software Engineer', 'Data Scientist', 'System Administrator', 'Product Manager', 'UX Designer', 'Business Analyst']

levels = ['Junior', 'Mid', 'Senior', 'Lead']

workloads = ['40h/week', '20h/week', '30h/week']

# ISCO-08 detailed unit groups relevant to IT

isco\_detailed = {

"1330": "ICT Services Managers",

"2152": "Electronics Engineers",

"2511": "Systems Analysts",

"2512": "Software Developers",

"2513": "Web and Multimedia Developers",

"2514": "Applications Programmers",

"2519": "Software & App Developers Not Elsewhere Classified",

"2521": "Database Designers and Administrators",

"2522": "Systems Administrators",

"2523": "Computer Network Professionals",

"2529": "ICT Professionals Not Elsewhere Classified",

"3511": "ICT Operations Technicians",

"3512": "ICT User Support Technicians"

}

# Function to generate contract period and renewal status

def generate\_contract\_period\_and\_status():

start\_date = fake.date\_between(start\_date='-3y', end\_date='today')

months\_duration = random.randint(6, 36)

end\_date = start\_date + timedelta(days=30 \* months\_duration)

start\_str = start\_date.strftime('%d-%m-%Y')

end\_str = end\_date.strftime('%d-%m-%Y')

period\_str = f"{start\_str} to {end\_str}"

today = datetime.today().date()

if end\_date.date() > today:

renewal\_status = random.choice(['Active', 'Pending', 'Renewed'])

else:

renewal\_status = random.choice(['Expired', 'Renewed', 'Not Renewed'])

return period\_str, renewal\_status

# Generate data

data = []

for i in range(1, num\_employees + 1):

employee\_id = f"E{i:04d}"

contract\_id = f"C{i:04d}"

gender = random.choice(genders)

full\_name = fake.name\_male() if gender == 'Male' else fake.name\_female() if gender == 'Female' else fake.name()

isco\_code, isco\_title = random.choice(list(isco\_detailed.items()))

position = random.choice(positions)

level = random.choice(levels)

workload = random.choice(workloads)

period\_date, renewal\_status = generate\_contract\_period\_and\_status()

data.append([

employee\_id,

contract\_id,

gender,

full\_name,

f"{isco\_code} - {isco\_title}",

position,

level,

workload,

period\_date,

renewal\_status

])

# Create DataFrame and save to CSV

columns = ['Employee ID', 'Contract ID', 'Gender', 'Full name', 'ISCO Group',

'Position', 'Level', 'Workload', 'Period Date', 'Renewal Status']

contract\_table = pd.DataFrame(data, columns=columns)

contract\_table.to\_csv("Contract\_Table.csv", index=False)

print("✅ Contract\_Table.csv has been created successfully.")

1. Siis sai seatud tingimuseks, et viimane kuupäev peaks lõppema aastaga 2024
2. Benefits tulbal oli tekstiline väärtus, see sai muudetud numbriliseks (nt. boonused, ületunnitöö vms)
3. Gender tulbas oli antud – female, male ja other. Palusime other eemaldada
4. Seejärel muutsime veel tulpasid:

* change workload data to mainly full time workers and the values should be workload=percent from full position (mainly 1, sometimes 0.5, rarely 0.75 or other)
* divide Period Date column into 2 columns: Start Date and End Date
* women and men salary should have small pay gap within the same job position category
* the table should also have employees with changed contractual conditions (position, category, level of position) and some employees who have started to work in 2024 and some employees who have left the company in year 2024

1. Veel mõned muudatused:

* start date may be between 2013 to 2024 and please do not end contract on the last day of the year and start date cannot be later than end date
* End date should be mostly empty
* We want to have employees in the company who leave the company in 2024.
* Workload values should only be: 1; 0,25; 0,5 and 0,75