exercise-module-40

October 23, 2024

1 Module 41: Importing and most common queries in SQL using Python

Author: Juliho David Castillo Colmenares

[81]:
```
import sqlite3
import pandas
```

[82]:
```
# Create a datafile for the database
data = pandas.read_csv("recursos_humanos.csv")
data.head()
```

[82]:

	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	left	promotion_last_5years	sales	salary
0	0.38	0.53	2	157	3	0	1	0	sales	low
1	0.80	0.86	5	262	6	0	1	0	sales	medium
2	0.11	0.88	7	272	4	0	1	0	sales	medium
3	0.72	0.87	5	223	5	0	1	0	sales	low
4	0.37	0.52	2	159	3	0	1	0	sales	low

[83]:
```
# Create a new SQLite database and establish a connection
conn = sqlite3.connect("data.db")
cursor = conn.cursor()
```

[84]:
```
# Create a table named 'Details' that matches the structure of the CSV file
cursor.execute(
    """
CREATE TABLE IF NOT EXISTS Details (
    satisfaction_level REAL,
    last_evaluation REAL,
    number_project INTEGER,
    average_monthly_hours INTEGER,
    time_spend_company INTEGER,
    work_accident INTEGER,
    left INTEGER,
    promotion_last_5years INTEGER,
    sales TEXT,
    salary TEXT
)
"""
)
```

[84]:
```
<sqlite3.Cursor at 0x226ffccd640>
```

[85]:
```
# Insert the CSV data into the 'Details' table
data.columns = [
    col.replace("average_montly_hours", "average_monthly_hours").lower()
    for col in data.columns
]
data.to_sql("Details", conn, if_exists="replace", index=False)
```

[85]:
```
14999
```

[86]:
```
# 1. Calculate the average satisfaction level of all employees
cursor.execute("SELECT AVG(satisfaction_level) FROM Details")
avg_satisfaction = cursor.fetchone()[0]
avg_satisfaction
```

[86]:
```
0.6128335222348156
```

[87]:
```
# 2. Compare satisfaction levels between employees who stayed and those who left
cursor.execute("SELECT AVG(satisfaction_level) FROM Details WHERE left = 1")
avg_satisfaction_left = cursor.fetchone()[0]
avg_satisfaction_left
```

[87]:
```
0.44009801176141133
```

[88]:
```
cursor.execute("SELECT AVG(satisfaction_level) FROM Details WHERE left = 0")
avg_satisfaction_stayed = cursor.fetchone()[0]
avg_satisfaction_stayed
```

[88]:
```
0.666809590479524
```

[89]:
```
# 3. Calculate the average monthly hours for employees with low or medium salary
cursor.execute(
    "SELECT AVG(average_monthly_hours) FROM Details WHERE salary IN ('low',␣
 ↪'medium')"
)
avg_hours_low_medium_salary = cursor.fetchone()[0]
avg_hours_low_medium_salary
```

[89]:
```
201.15666327568667
```

[90]:
```
# 4. Extract records of employees promoted in the last 5 years who left the␣
 ↪company
cursor.execute("SELECT * FROM Details WHERE promotion_last_5years = 1 AND left␣
 ↪= 1")
promoted_left = cursor.fetchall()
promoted_left[:5]
```

[90]:
```
[(0.45, 0.51, 2, 160, 3, 1, 1, 1, 'sales', 'low'),
 (0.79, 0.59, 4, 139, 3, 0, 1, 1, 'management', 'low'),
 (0.41, 0.46, 2, 160, 3, 0, 1, 1, 'sales', 'low'),
 (0.11, 0.79, 6, 292, 4, 0, 1, 1, 'technical', 'low'),
 (0.41, 0.56, 2, 154, 3, 0, 1, 1, 'support', 'medium')]
```

[91]:
```
# 5. Extract records of employees with a last evaluation of 0.9 or higher
cursor.execute("SELECT * FROM Details WHERE last_evaluation >= 0.9")
high_evaluation = cursor.fetchall()
high_evaluation[:5]
```

[91]:
```
[(0.89, 1.0, 5, 224, 5, 0, 1, 0, 'sales', 'low'),
 (0.84, 0.92, 4, 234, 5, 0, 1, 0, 'sales', 'low'),
 (0.78, 0.99, 4, 255, 6, 0, 1, 0, 'sales', 'low'),
 (0.09, 0.95, 6, 304, 4, 0, 1, 0, 'sales', 'low'),
 (0.89, 0.92, 5, 242, 5, 0, 1, 0, 'sales', 'low')]
```

[92]:
```
# Close the connection
conn.close()
```

[93]:
```
# Display the results to the user
print(f"Average Satisfaction Level: {avg_satisfaction:.2f}")
print(f"Average Satisfaction Level (Left): {avg_satisfaction_left:.2f}")
print(f"Average Satisfaction Level (Stayed): {avg_satisfaction_stayed:.2f}")
print(f"Average Monthly Hours (Low/Medium Salary): {avg_hours_low_medium_salary:
 ↪.2f}")
print(f"Number of Promoted Employees Who Left: {len(promoted_left)}")
print(f"Number of Employees with High Evaluation: {len(high_evaluation)}")
```

```
Average Satisfaction Level: 0.61
Average Satisfaction Level (Left): 0.44
```

```
Average Satisfaction Level (Stayed): 0.67
Average Monthly Hours (Low/Medium Salary): 201.16
Number of Promoted Employees Who Left: 19
Number of Employees with High Evaluation: 2988
```