Pandas SQL

Assignment 1 Question

Pandas SQL: Assignment 1

Table of Contents

1.Introduction

This assignment will help you to consolidate the concepts learnt in the session.

2.Problem Statement

Problem statement 1:

Read the following data set:

https://archive.ics.uci.edu/ml/machine-learning-databases/adult/

Rename the columns as per the description from this file:

https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.names

Task: Create a sql db from adult dataset and name it sqladb

1. Select 10 records from the adult sqladb

2. Show me the average hours per week of all men who are working in private sector

3. Show me the frequency table for education, occupation and relationship, separately

4. Are there any people who are married, working in private sector and having a master's degree

5. What is the average, minimum and maximum age group for people working in different sectors

6. Calculate age distribution by country

7. Compute a new column as 'Net-Capital-Gain' from the two columns 'capital-gain' and 'capital-loss'

Problem statement 2:

Read the following data set:

https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data

Task:

1. Create an sqlalchemy engine using a sample from the data set

2. Write two basic update queries

3. Write two delete queries

4. Write two filter queries

5. Write two function queries

Note: Solution submitted via github must contain all the source code and output.

3. Expected Output

This assignment consists of 200 marks and the solution file needs to be submitted in the upcoming session in .ipynb format for evaluation.