1. Tell me about yourself

**SQL Questions:**

1. What is the order of Execution in SQL query?
2. Suppose we have 2 Tables with 10 rows each, so what will be the maximum and minimum number of rows we get after joining these two tables.
3. Write a SQL query to extract the id’s who haven’t transacted in the month of ‘jan’.

|  |  |  |
| --- | --- | --- |
| Id | Month | amt |
| 1 | jan | 23 |
| 1 | Feb | 3 |
| 1 | March | 14 |
| 2 | jan | 23 |
| 3 | jan | 4 |
| 4 | March | 5 |
| 5 | jan | 12 |
| 2 | April | 34 |
| 3 | April | 2 |
| 4 | Feb | 1 |
| 5 | jan | 20 |

**Sales Table:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Retailer | Store # | SKU | Week | Year | Units | Dollars |
| 101 | 727 | 599506 | 16 | 2021 | 3 | 545 |
| 102 | 2686 | 1004823827 | 19 | 2021 | 2 | 43.94 |
| 101 | 6173 | 681059 | 18 | 2021 | 10 | 65.55 |
| 101 | 2578 | 1002878652 | 21 | 2021 | 2 | 16.4 |
| 102 | 1419 | 625356 | 19 | 2021 | 0 | 55 |
| 101 | 727 | 599506 | 16 | 2021 | 3 | 545 |

**Store Table:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Store # | Store Name | Store Address | Store City | State |
| 362 | HARAHAN | 5151 CITRUS BLVD | HARAHAN | LA |
| 6220 | GLASTONBURY | 115 PUTNAM BLVD | GLASTONBURY | LM |
| 2844 | MAPLE GROVE | GROVE CIRCLE | MAPLE GROVE | CT |
| 485 | LOWE'S OF FT. OGLETHORPE | 2215 BATTLEFIELD PKWY | FORT OGLETHORPE | GA |
| …………….. | ……….. |  |  |  |

**SKU Table:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SKU | |  | | --- | | **SKU\_DESCRIPTION** | | Category | Sub Category |
| 1000168869 | RED QUAD ADVANCED FO | CAULKS | SOLVENT |
| 1000169079 | BEIGE QUAD ADVANCED | CAULKS | SOLVENT |
| 1348292 | SERTA QUEEN HYBRID 10-IN MATTRESS | Hybrid | NULL |
| 1000169445 | RED QUAD ADVANCED FO | CAULKS | SOLVENT |
| 319762 | |  | | --- | | 4-IN PRO SPPUPSPRY W 15-FT QTRNOZ | | SPRAYS | PRO-S |
| ………. |  |  |  |

**Fiscal\_Yr\_Wk Table:**

|  |
| --- |
| Week\_ending\_Dt |
| 2017-02-12 00:00:00.000 |
| 2017-02-19 00:00:00.000 |
| 2017-02-26 00:00:00.000 |
| 2017-03-05 00:00:00.000 |
| 2017-03-12 00:00:00.000 |

1. Write a query to get top 4 highest selling states.
2. Print the number of Stores having sales less than 50 in each week for Year 2021.
3. Write a query to get Store, Store Name, SKU, Category and total dollars and Avg units for each.
4. From Fiscal\_Yr\_Wk table get Month, Year from Week\_ending\_Dt. Give alias to each column.
5. Write a query to get total number of units and sales (SALES TABLE) of each store (pull only top 10), in year 2020?
6. Write a SQL query to change the structure of Table A into Table B (wide format to long format)

Table A

|  |  |  |  |
| --- | --- | --- | --- |
| Id | dt\_falg | nb\_flag | ws\_flag |
| 1 | Best | Rest | Rest |
| 2 | Rest | Rest | Best |
| 3 | Best | Best | Rest |

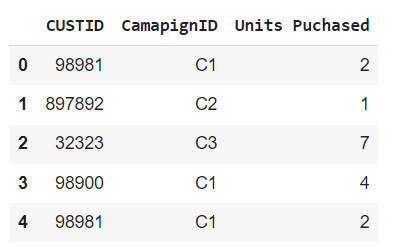
Table B

|  |  |  |
| --- | --- | --- |
| Id | lob | flag |
| 1 | dt | Best |
| 1 | nb | Rest |
| 1 | ws | Rest |
| 2 | dt | Rest |
| 2 | nb | Rest |
| 2 | ws | Best |
| 3 | dt | Best |
| 3 | nb | Best |
| 3 | ws | Rest |

**Python Question**

import pandas as pd  
import numpy as np  
  
df\_dict = {'CUSTID': ['98981', '897892','32323','98900','98981'], 'CamapignID': ['C1', 'C2','C3','C1', 'C1'],'Units Puchased':[2,1,7,4,2]}  
campaign\_df = pd.DataFrame(data=df\_dict)  
campaign\_df

* After executing the above program output displayed below :



* Follow the above table write python program given below

1. Calculate the total Units purchased per campaign

2. Calculate the total Units purchased, excluding those Units sold from campaign 2 (C2)

3.return / find the custid with the highest units purcahsed across all campaigns aggregations not required.