SQL Assignment

Guidelines:

i. ***This assignment is mandatory for everyone.***
ii. ***Use the below datasets or Tables to solve the below queries.***
iii. ***It is mandatory to submit the answer with the screenshot of the output you have received.***
iv. *There will only be a single attempt for each exam and* ***no deadline extension in case of assignments****.*
v. *Any case of unfair means or* ***plagiarism would lead to debarring*** *in final placements without any further consideration.*

Problem Statement 1:

Retail Depart in Bank accepts the application for new account opening and sends it to the backend operations team to process. There are three steps involved in processing the application. Each step is processed by an Agent.

Step1: Maker: He is doing the data entry job, enter the information into the system.

Step 2: Checker: He validates the data entered by Maker, and can accept or reject it.

If It is rejected, then it is sent back to Maker and it is reverified else if it's successful it is sent to Approver.

Step3: Approver: He approves the application and account is

created. Below are the two tables,

Table1: Agent Data: where agent details are stored

Agent Table	
Agent_Id	**Name**
1	Vijay
2	Rajesh
3	Satish
4	Anji

Table2: Case_transaction_details: Each step in the Backend operation is stored.

Case_transaction_details					
Case_Id	**Stage**	**Login_Time**	**Logout_time**	**Agent_id**	**Status**
101	Maker	5/11/2019 10:20	10:30	2	Success
102	Maker	10:25	10:35	1	Success
103	Maker	10:40	10:50	2	Success
101	Checker	10:45	11:00	3	Success
101	Approver	11:15	11:30	2	Success
102	Checker	10:50	11:00	1	Reject
102	Maker	11:15	11:45	4	Reverify
103	Checker	11:30	11:40	2	Reject

Please Write SQL queries for below,

1. How many unique cases per day?
2. Case Id which is rejected by checker but still not reverified?
3. Top Agent names with who processed more applications?

Problem 2:
Dataset:

https://drive.google.com/file/d/15XT-eZRgFDNaCbpvZheax4oLNQglhoAW/view?usp=sharing

Instructions:-

1. Download the data files for this Assignment. Your first task is to create tables from the data files. In order to do so, please follow the steps given below:
2. Open MySQL Workbench
3. Connect to your database using the connection you have created
4. Create a database named Superstores
5. In the “Navigator” pane on the left-hand side, you will find the created database.
6. Right-click on the superstores
7. You will see the option called “Table Data Import Wizard”. Click on it.
8. Follow the wizard to create tables by providing the .csv data files that you have downloaded.
9. You need to follow the “Table Data Import Wizard” for each data file given for this assignment.

Task 1:- Understanding the Data

1. Describe the data in hand in your own words.
2. Identify and list the Primary Keys and Foreign Keys for this dataset provided to you(In case you don’t find either primary or foreign key, then specially mention this in your answer)

Task 2:- Basic & Advanced Analysis

1. Write a query to display the Customer_Name and Customer Segment using alias name “Customer Name", "Customer Segment" from table Cust_dimen.
2. Write a query to find all the details of the customer from the table cust_dimen order by desc.

3. Write a query to get the Order ID, Order date from table orders_dimen where 'Order Priority' is high.
4. Find the total and the average sales (display total_sales and avg_sales)
5. Write a query to get the maximum and minimum sales from maket_fact table.
6. Display the number of customers in each region in decreasing order of no_of_customers. The result should contain columns Region, no_of_customers.
7. Find the region having maximum customers (display the region name and max(no_of_customers)
8. Find all the customers from Atlantic region who have ever purchased 'TABLES' and the number of tables purchased (display the customer name, no_of_tables purchased)
9. Find all the customers from Ontario province who own Small Business. (display the customer name, no of small business owners)
10. Find the number and id of products sold in decreasing order of products sold (display product id, no_of_products sold)
11. Display product Id and product subcategory whose product category belongs to Furniture and TechnologyTechnlogy. The result should contain columns of product id, product subcategory.
12. Display the product categories in descending order of profits (display the product category wise profits i.e. product_category, profits)?
13. Display the product category, product sub-category and the profit within each subcategory in three columns.
14. Display the order date, order quantity and the sales for the order.
15. Display the names of the customers whose name contains the
    i) Second letter as 'R'
    ii) Fourth letter as 'D'
16. Write a SQL query to make a list with Cust_Id, Sales, Customer Name and their region where sales are between 1000 and 5000.
17. Write a SQL query to find the 3rd highest sales.
18. Where is the least profitable product-subcategory shipped the most? For the least profitable product sub-category, display the region-wise no_of_shipments and the profit made in each region in decreasing order of profits (i.e. region, no_of_shipments, profit_in_each_region)
    - Note: You can hardcode the name of the least profitable product subcategory

Problem 3:

Dataset:

https://drive.google.com/drive/folders/15WrWbPKWD6BJQdm1_DQ17IGs0ql4_qs2?usp=sharing

1. List the names of all left-handed batsmen from England. Order the results alphabetically.
2. List the names and age (in years, should be an integer) as on 2018-12-02 (12th Feb 2018) of all bowlers with skill "Legbreak googly" who are 28 or more in age. Order the result in decreasing order of their ages. Resolve ties alphabetically.
3. List the match ids and toss winning team IDs where the toss winner of a match decided to bat first. Order results in increasing order of match ids.
4. In the match with match id 335987, list the over ids and runs scored where at most 7 runs were scored. Order the over ids in decreasing order of runs scored. Resolve ties by listing the over ids in increasing order.
5. List the names of those batsmen who were bowled at least once in alphabetical order of their names.
6. Find the bowler who has the best average overall. Bowling average is calculated using the following formula:
   bowling average = Number of runs given/ Number of wickets taken

   Calculate the average up to 3 decimal places and return the bowler with the lowest average runs per wicket. In case of a tie, return the results in alphabetical order.

7. List the players and the corresponding teams where the player played as "CaptainKeeper" and won the match. Order results alphabetically on the player's name.
8. List the names of all players and their runs scored (who have scored at least 50 runs in any match). Order results in decreasing order of runs scored. Resolve ties alphabetically.
9. List the player names who scored a century but their teams lost the match. Order results alphabetically.
10. List match ids and venues where KKR has lost the game. Order results in increasing order of match id.
11. List the names of top 10 players who have the best batting average in season 5. Batting average can be calculated according to the following formula:

    batting average(player) = Number of runs scored by player/ Number of matches player has batted in

    The output should contain exactly 10 rows. Report results up to 3 decimal places. Resolve ties alphabetically.