**SQL Mini Project**

**Part - A**

**ICC Test Cricket**

**Dataset: ICC Test Batting Figures.csv**

Test cricket is the form of the sport of cricket with the longest match duration and is considered the game's highest standard. Test matches are played between national representative teams that have been granted ‘Test status’, as determined and conferred by the International Cricket Council (ICC). The term Test stems from the fact that the long, gruelling matches are mentally and physically testing. Two teams of 11 players each play a four-innings match, which may last up to five days (or longer in some historical cases). It is generally considered the most complete examination of a team's endurance and ability.

The Data consists of runs scored by the batsmen from 1877 to 2019 December.

**Data Dictionary:**

|  |  |  |
| --- | --- | --- |
| **Sl No** | **Column Name** | **Column Description** |
| 1 | Player | Name of the player and country the player belongs to |
| 2 | Span | The duration of years between which the player was active |
| 3 | Mat | No of matches played by the player |
| 4 | Inn | No of innings played by the player |
| 5 | NO | No of matches the player was NOT OUT by the end of the match. |
| 6 | Runs | Total number of runs scored by the player |
| 7 | HS | Highest Score of the player |
| 8 | Avg | Average runs scored by the player in all the matches |
| 9 | 100 | No of centuries scored by the player |
| 10 | 50 | No of fifties scored by the player |
| 11 | 0 | No of Duck outs of the player |
| 12 | Player Profile | Link to the profiles of the players |

**Tasks to be performed:**

1. Import the csv file to a table in the database.
2. Remove the column 'Player Profile' from the table.
3. Extract the country name and player names from the given data and store it in separate columns for further usage.
4. From the column 'Span' extract the start\_year and end\_year and store them in separate columns for further usage.
5. The column 'HS' has the highest score scored by the player so far in any given match. The column also has details if the player had completed the match in a NOT OUT status. Extract the data and store the highest runs and the NOT OUT status in different columns.
6. Using the data given, considering the players who were active in the year of 2019, create a set of batting order of best 6 players using the selection criteria of those who have a good average score across all matches for India.
7. Using the data given, considering the players who were active in the year of 2019, create a set of batting order of best 6 players using the selection criteria of those who have the highest number of 100s across all matches for India.
8. Using the data given, considering the players who were active in the year of 2019, create a set of batting order of best 6 players using 2 selection criteria of your own for India.
9. Create a View named ‘Batting\_Order\_GoodAvgScorers\_SA’ using the data given, considering the players who were active in the year of 2019, create a set of batting order of best 6 players using the selection criteria of those who have a good average score across all matches for South Africa.
10. Create a View named ‘Batting\_Order\_HighestCenturyScorers\_SA’ Using the data given, considering the players who were active in the year of 2019, create a set of batting order of best 6 players using the selection criteria of those who have highest number of 100s across all matches for South Africa.
11. Using the data given, Give the number of player\_played for each country.
12. Using the data given, Give the number of player\_played for Asian and Non-Asian continent

**Part – B**

“Richard’s Supply” is a company which deals with different food products. The company is associated with a pool of suppliers. Every Supplier supplies different types of food products to Richard’s supply. This company also receives orders for the food products from various customers. Each order may have multiple products mentioned along with the quantity. The company has been maintaining the database for 2 years.

Refer to the following Entity-Relationship diagram of the database.

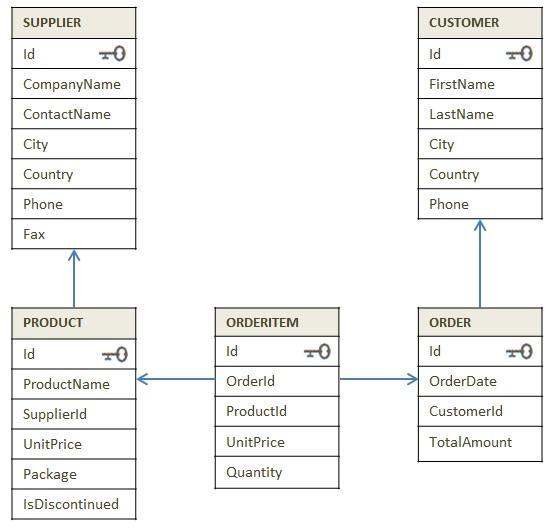


Diagram: E-R Diagram of Supply\_chain database

Try to get insight of business through the dataset

**Instruction:** Execute the SQL files in the sequence given below.

1. 1\_DDL\_Case Study
2. 2\_Data
3. 3\_Data Constraints
4. Company sells the product at different discounted rates. Refer actual product price in product table and selling price in the order item table. Write a query to find out total amount saved in each order then display the orders from highest to lowest amount saved.
5. Mr. Kavin want to become a supplier. He got the database of "Richard's Supply" for reference. Help him to pick:

a. List few products that he should choose based on demand.

b. Who will be the competitors for him for the products suggested in above questions.

1. Create a combined list to display customers and suppliers details considering the following criteria

* Both customer and supplier belong to the same country
* Customer who does not have supplier in their country
* Supplier who does not have customer in their country

1. Every supplier supplies specific products to the customers. Create a view of suppliers and total sales made by their products and write a query on this view to find out top 2 suppliers (using windows function) in each country by total sales done by the products.
2. Find out for which products, UK is dependent on other countries for the supply. List the countries which are supplying these products in the same list.