

UBISOFT INDIA DATA ANALYST TEST

The main goal of this test is to understand your technical skills and problem solving abilities. Please use Graphs to support your answers. Also provide SQL query for each of the below questions.

PART A]

Please refer to the dataset attached in the mail to answer these questions.

1. From the given Dataset, determine and analyze Day 1, Day 3, Day 7, Day 15 and Day 30 Retention. (Refer to **Session start** tab)
2. Create a Funnel of level progression and determine if any areas of the game are acting as blockers. (Refer to **Level Start** tab). Also, identify how many times levels are replayed.
3. What percentage of users make first time purchases between levels 1 to 7? (Refer to **Spenders** tab)
4. What percentage of users make a repeat purchase? At what level maximum repeat purchases happen? (Refer to **Spenders** tab)
5. Which packs generate the maximum revenue? (Refer to **Spenders** tab)

PART B]

1. Difference between Normalization and Denormalization (Single line difference) ? Types of Normalization and Denormalization?
Which would you like to prefer for the Query optimization?
2. Write SQL query to display the data in following manner:
Table sample:

User	Salary
A	2000
B	3000
C	8000
D	4000
E	9000
F	5500
G	6000

Required Output:

User	Salary	Total2
A	2000	2000
B	3000	5000
C	8000	13000
D	4000	17000
E	9000	26000
F	5500	31500
G	6000	37500

3. Write a procedure to Calculate daily active users in last 30 days.

Ex : If you are calculating daily active users on 1st of march 2011 then it should calculate that from last 30 days it means it starts from 1st Feb 2011 to 1st March 2011 and shows the result on 1st March 2011 likewise it shows the result date wise but for last 30 days. A user may or may not login for give date.

Schema:

UserID	Login Date
--------	------------

4. Which database table is used to store all procedure in Mysql and their details?

Write a query to show all the procedure?

5. Write the SQL query to display the data in user specified order:

User	Salary
A	2000
D	8540
G	7540
C	7000

Required output:

User	Salary
C	2000
A	8540
D	7540
G	7000

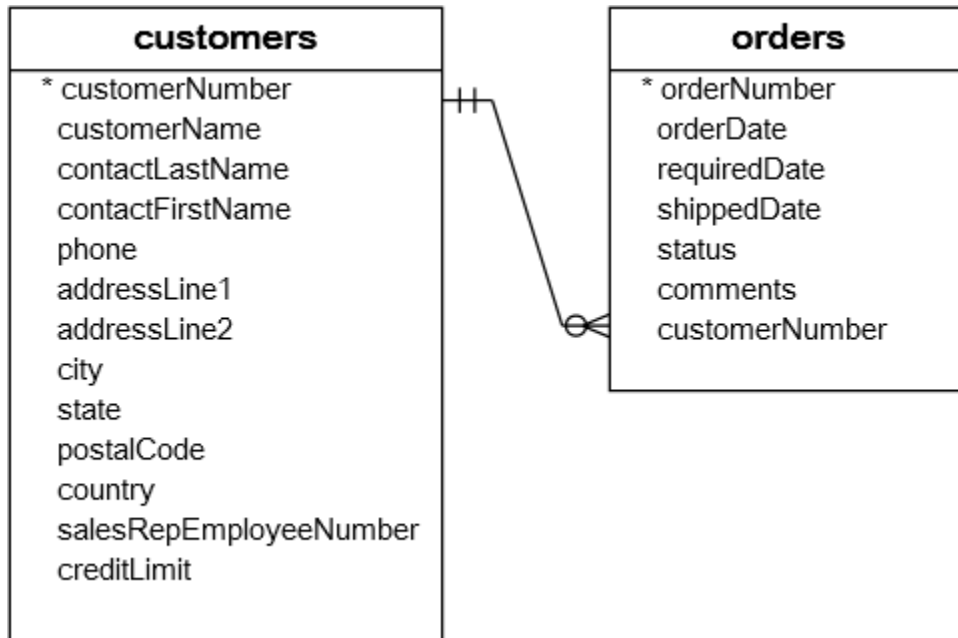
6. Write the SQL query for following output?

User	Salary
A	01/01/2011
A	02/01/2011
B	04/01/2011
B	08/01/2011
C	10/01/2011

Required Output:

User	Count	User Details
A	2	A,A
B	2	B,B
C	1	C

7. Based on the table below write a query to find out customers who have not yet placed at least one order.



PART C] Refer to the attached dataset (CA.csv (is) present in the Test file). This database has the spend information of all the players. Identify different group of players present in the data and profile them based on their spend behavior/pattern. Use a statistical technique for this exercise.