

Case Study

on

DW Concept, SQL, Advance SQL

for Hexaware – Batch 02

Contents

1	Instructions	3
2	Case Study – I	4
2.1	Introduction	4
2.2	Available Data	4
2.3	Entity Relationship Diagram	5
2.4	Table Description and Data:	6
3	Case Study Question	10
3.1	Initial Understanding of Case Study	10
3.2	Customer Journey	10
3.3	Data Analysis Question	10

learn from vinit...

1 Instructions

This case study is prepared for Hexaware – Batch who have undergone training from Vinit in Feb 2023 & carries 70 marks. Please go through the requirement and problem statement in different sections of this document.

Expectation from participants is that they will use SQL Server Database for performing the exercise.

Take the screenshot of SQL Query and its output and paste in Microsoft Word Document and upload in Examly.

DDL and required data is mentioned in section 2.4 of this document. Expectation from participants is they will respect the naming convention used and will create their own insert statement as mentioned in data table.

Each answer sheet will go through plagiarism, if found that anyone has copied or have taken reference from Internet the assignment submission will be rejected and will be reported to Hexaware HR team including Hexaware learning SPOCs.

Each participant must attend these case studies without seeking any help from anyone not limited to trainer, colleagues, friends, seniors, internet etc.

2 Case Study – I

2.1 Introduction

After advancement of Internet and 4G, subscription-based series are popular on Television and Mobile App(s) and have reached to common man irrespective of his/her geographical location. Vinit realised that there was a medium to large gap in the market – he wanted to create a new streaming service that will only have beverage related content – something like Amazon Prime but with only beverage shows!

Vinit along with his few friends launched his new start-up BTech-Chai-Wala in year 2019 and started selling weekly, monthly and annual subscriptions, giving his customers unlimited on-demand access to exclusive beverages videos from around the world!

Vinit created BTech-Chai-Wala with a data driven mindset as he has played different roles in his IT Career like Data Analyst, Enterprise Data Architect, Principal Architect etc. and wanted to ensure all future investment decisions and new features were decided using data available in Data Warehouse and Data Lake built for BTech-Chai-Wala. This case study focuses on using subscription style digital data to answer important business questions for BTech-Chai-Wala...

2.2 Available Data

Vinit has shared the data design for BTech-Chai-Wala and also short descriptions on each of the database tables – this case study focuses on only 3 tables but there will be a challenge to create a new table for the BTech-Chai-Wala team.

All datasets exist within the BTech-Chai-Wala database schema – be sure to include this reference within your SQL scripts as you start exploring the data and answering the Case Study Questions.

2.3 Entity Relationship Diagram

Following diagram depicts the relationship between the tables



Hint:

1. Create Database:

a. `CREATE DATABASE [btech-chai-wala];`

2. Create Tables:

a. Table - plans:

```

use [btech-chai-wala];
create table [dbo].[plans]
(
    plan_id INT NOT NULL PRIMARY KEY IDENTITY(1,1),
    plan_name VARCHAR(100),
    plan_price DECIMAL
);
  
```

b. Table – subscriptions:

```

use [btech-chai-wala];
create table [dbo].[subscriptions]
(
    subscrip_id INT NOT NULL PRIMARY KEY IDENTITY(1,1),
    cust_id INT,
    plan_id INT,
    start_date DATE,
    is_active VARCHAR(1)
);
  
```

c. Table – customer:

```

use [btech-chai-wala];
create table [dbo].[customer]
(
    cust_id INT NOT NULL PRIMARY KEY IDENTITY(1,1),
    cust_name varchar(100),
    cust_city varchar(100),
    cust_dob DATE
);
  
```

2.4 Table Description and Data:

1. Table – plans:

- i. Customers can choose with which plans to join BTech-Chai-Wala when they first sign-up.
- ii. Basic plan customers have limited access and can only stream their videos and is only available monthly at INR 499 + GST.
- iii. Pro plan customers have no watch time limits and are able to download videos for offline viewing. Pro plans start at INR 999 + GST a month or INR 11,988 + GST for an annual subscription.
- iv. Customers can sign up to an initial 7-day free trial will automatically continue with the pro monthly subscription plan unless they cancel, downgrade to basic or upgrade to an annual pro plan at any point during the trial.
- v. When customers cancel their BTech-Chai-Wala service - they will have a churn plan record with a null price, but their plan will continue until the end of the billing period.

plan_id	plan_name	plan_price (INR)
1	Trail	0.00
2	Basic_Monthly	499.00
3	Pro_Monthly	999.00
4	Pro_Annual	11,998.00
5	Chrun	Null

* GST is calculated while paying and here we can exclude and consider plan price only.

2. Table – subscription:

- i. Customer subscriptions show the exact date where their specific plan_id starts.

- ii. If customers downgrade from a pro plan or cancel their subscription – the higher plan will remain in place until the period is over – the start_date in the subscriptions table will reflect the date that the actual plan changes.
- iii. When customers upgrade (from lower plan to higher plan) their account from a basic plan to a pro or annual pro plan – the higher plan will take effect straightaway.
- iv. When customers churn (leaves) – they will keep their access until the end of their current billing period but the start_date will be technically the day they decided to cancel their service.
- v. Column is_active? represents: 1 – Active Subscription, 0 – Inactive Subscriptions.

subscrip_id	cust_id	plan_id	start_date (yyyy-mm-dd)	is_active?
1	1	2	2019-08-01	0
2	1	3	2019-08-08	0
3	5	1	2020-01-17	0
4	5	1	2020-03-17	0
5	5	3	2021-03-24	0
6	5	3	2021-03-24	0
7	5	5	2021-04-29	0
8	5	5	2022-04-30	1
9	6	1	2020-05-31	0
10	6	2	2020-06-07	0
11	9	1	2020-06-22	0
12	9	3	2020-06-29	0
13	8	1	2020-07-06	0
14	8	3	2020-07-13	0
15	7	1	2020-08-01	0

subscrip_id	cust_id	plan_id	start_date (yyyy-mm-dd)	is_active?
16	1	1	2020-08-07	1
17	7	2	2020-08-08	0
18	9	4	2020-08-29	0
19	2	1	2020-09-20	0
20	2	1	2020-09-20	0
21	2	4	2020-09-27	0
22	2	4	2020-09-27	0
23	6	4	2020-10-21	0
24	2	1	2020-11-19	0
25	10	1	2020-11-19	0
26	2	5	2020-11-26	1
27	10	5	2021-11-26	1
28	3	1	2020-12-15	0
29	3	1	2020-12-15	0
30	3	2	2020-12-22	0
31	4	2	2020-12-22	0
32	4	3	2021-03-29	1
33	3	3	2021-03-29	1
34	6	1	2021-05-31	0
35	6	2	2021-06-07	0
36	9	1	2021-06-22	0
37	9	3	2021-06-29	0
38	8	1	2021-07-06	0
39	8	3	2021-07-13	1
40	7	1	2021-08-01	1
41	9	4	2022-08-29	1

subscrip_id	cust_id	plan_id	start_date (yyyy-mm-dd)	is_active?
42	6	4	2021-10-21	1

3. Table – customer

- Customers are from different cities as mentioned below in the table.
- Customer Data of Birth (dob) is in format YYYY-MM-DD (year-month-date).

cust_id	cust_name	cust_city	cust_dob (yyyy-mm-dd)
1	Chaitanya	Bangalore	2001-02-02
2	Harini	Mumbai	2001-07-19
3	Indu	Chennai	2001-01-21
4	Jameela	Hyderabad	2002-05-01
5	Neeha	Bangalore	2000-03-15
6	Reshma	Mumbai	2001-02-02
7	Sahithi	Chennai	2000-11-02
8	Sharmila	Hyderabad	2001-04-22
9	Sreelakshmi	Bangalore	2000-06-17
10	Subasri	Mumbai	2002-01-30

3 Case Study Question

This case study is split into an initial data understanding question before diving straight into data analysis questions.

3.1 Initial Understanding of Case Study

1. Summarize your understanding of Section 01 of this document is less than 60 words.

3.2 Customer Journey

2. Based on the 10 customers provided in the section 2.4 of this document refer subscriptions & customer table, customers having age 21 as of date get their onboarding dates.

Hint:

- a. Try to keep it as short as possible - you may also want to run some sort of join to make your explanations a bit easier!
- b. SQL query output should include cust_id, cust_name, cust_city, plan_name, plan_price and start_date of subscription.
- c. It will include both active and inactive subscriptions.

3.3 Data Analysis Question

3. How many customers has BTech-Chai-Wala with activate subscription as on 5th May 2021?
4. What plan start_date values occur after the year 2020 for above dataset (refer section 2.4)?
5. What is the customer count and percentage breakdown of all 5 plan_name values at 2020-12-31?
6. How many customers have upgraded to an annual plan between 2020 and 2021?
7. How many customers downgraded from a pro monthly to a basic monthly plan in 2020 or 2021?