

SQL CASE STUDY

Nikita Rathi

BDAP batch 03 FT

SP JAIN SCHOOL OF GLOBAL MANAGEMENT

INDEX

<u>REQUIREMENT DOCUMENT.....</u>	<u>3</u>
<u>E-R DIAGRAM</u>	<u>5</u>
<u>SCHEMA DIAGRAM.....</u>	<u>6</u>
<u>TABLE Creation</u>	
<u>Queries.....</u>	
<u>.....7</u>	
<u>BUSINESS QUERIES ALONG WITH BUSINESS</u>	
<u>BENEFITS AND QUERY DEFINITION:</u>	<u>9</u>

Requirement document

13 July 2016 at 12:28 am

Boo box is an advertising social media and technology company. It is an advertisement network which connect publishers to advertisers.

Business Objective:

Business objective is to extract maximum details about advertisers, ads preferred, most profitable ads and their categories and sub categories, publisher preferences, format preferred, preferred duration under each category in order to do sentiment analysis on customers and compare commodity preferences in the market under each category.

Following is the data requirement for the detailed client analysis of Boo Box.com :

- Data requirements include details of advertiser, advertisement offered, preferred publishers or editors.
- Boo bOx provides ads across various countries
- Request collected are under 4 major categories health, entertainment, energy and education.
- Each ad category has further sub categories on the basis of ad type. For e.g. Under health cat_A represent beverages, cat_B represent Food commodity, cat_c represent medical aid, etc
- Details to be pinned includes:
 - **Advertiser details:**
 - advertiser company name*, company branch, country
 - contact details, official address, tin no
 - **Advertisement details:**
 - Product name, billing address,
 - **Format, duration, size, display, date**
 - **Publisher details**

- Publisher name
- Country
- Feedback for ads published under different publishers is to be noted.

Following assumptions are taken before framing E-R diagram and schema:

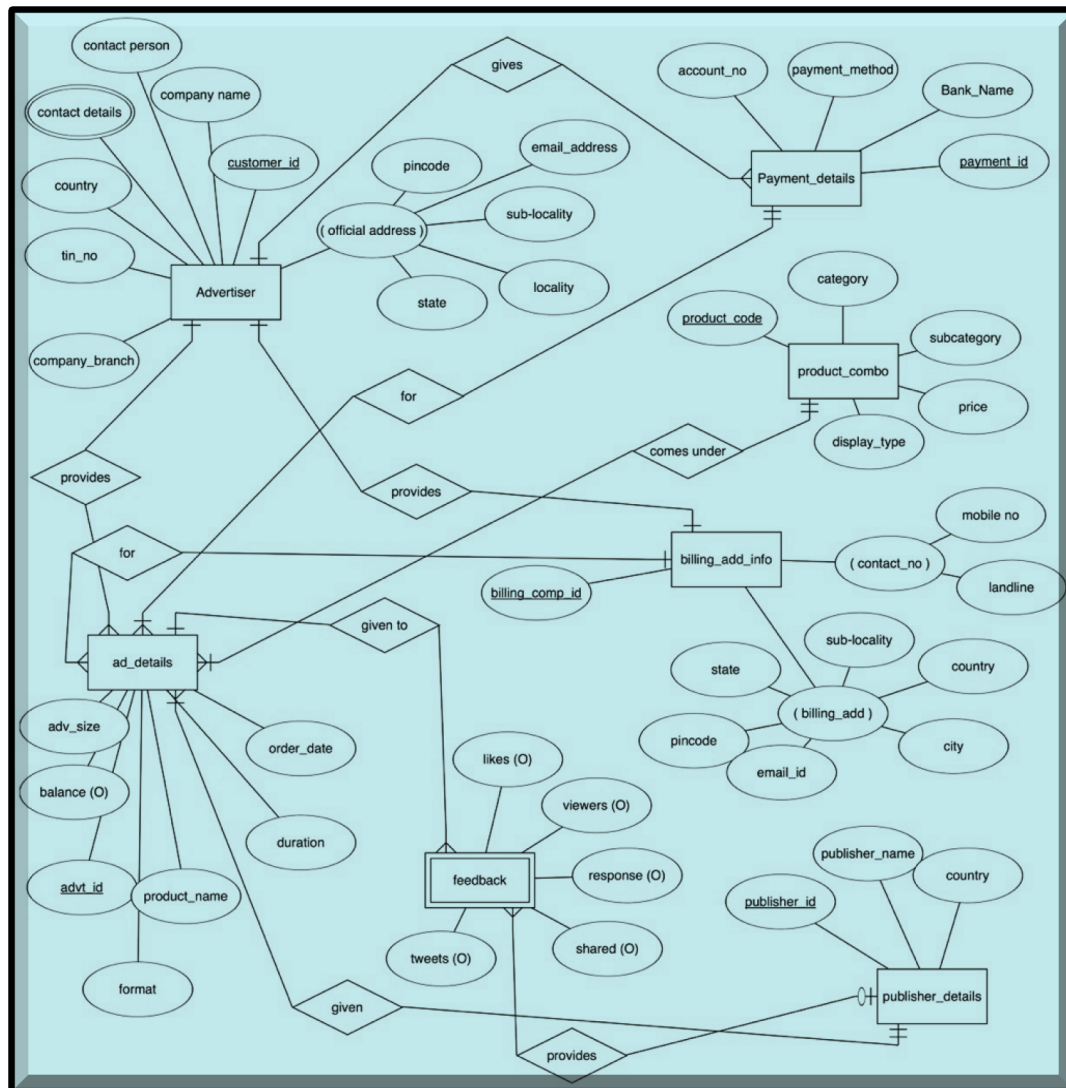
- Billing address will be provided by customer while enrolling for ad
And will be stored for ad transactions.
- Payment details are also stored in separate table. Each time with reference to customer id details are fetch in to transaction table.
- Attribute response in feedback is valued with the response of clients which customers get on advertisements. These response are noted internally and stored as a count in response attribute for further analysis.
- There are 7 entities and under each entity several attributes are described according to the requirement.
- Advertiser, Product_combo, Payment_details , billing_add_info and Publisher_details are master tables.
- Ad_details is a transaction table which fetch details from biling_add_info,advertiser,payment_details and publisher_details using references mentioned.

Case Study

15 July 2016 at 8:00 pm

E-R diagram

Nikita Rathi



Case Study 01

15 July 2016 at 9:30 pm

Schema diagram

Nikita Rathi

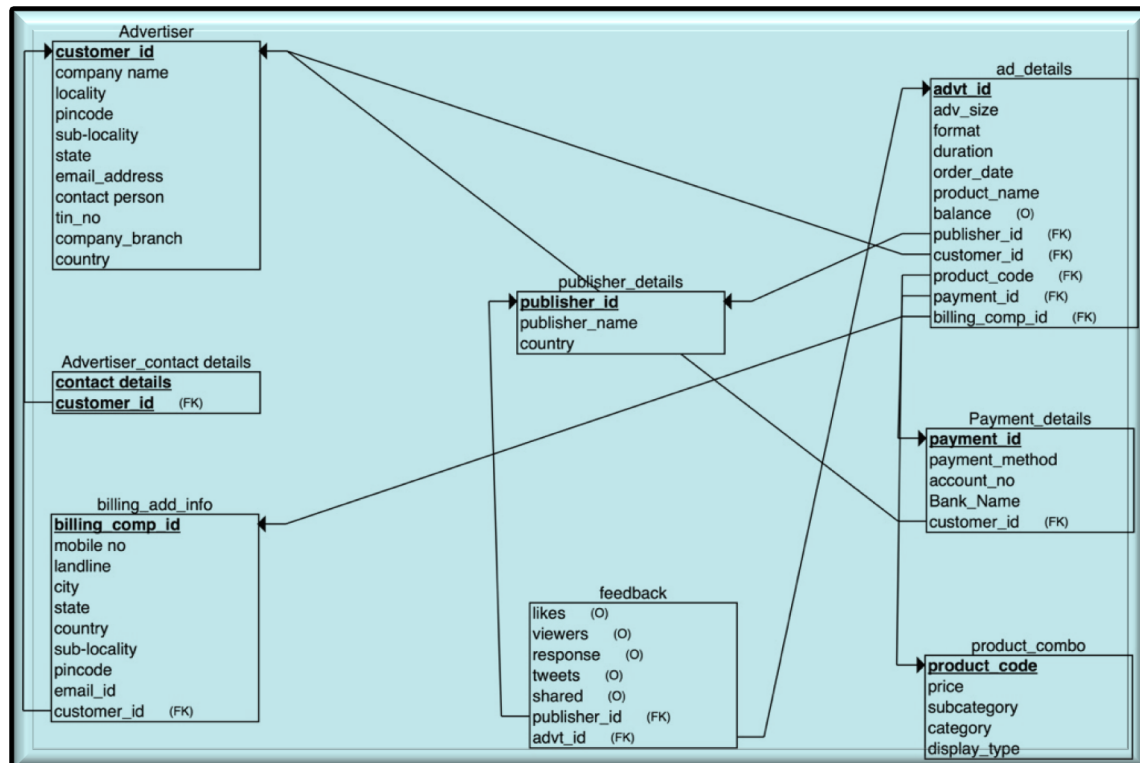


Table Creation Queries

```
create database case_study_db_01;  
use case_study_db_01;
```

```
CREATE TABLE Advertiser  
(  
    company_name INT NOT NULL,  
    customer_id INT NOT NULL,  
    locality INT NOT NULL,  
    pincode INT NOT NULL,  
    sub_locality INT NOT NULL,  
    state INT NOT NULL,  
    email_address INT NOT NULL,  
    contact_person INT NOT NULL,  
    tin_no INT NOT NULL,  
    company_branch INT NOT NULL,  
    country INT NOT NULL,  
    PRIMARY KEY (customer_id)  
);
```

```
CREATE TABLE product_combo  
(  
    price INT NOT NULL,  
    subcategory INT NOT NULL,  
    category INT NOT NULL,  
    product_code INT NOT NULL,  
    display_type INT NOT NULL,  
    PRIMARY KEY (product_code)  
);
```

```
CREATE TABLE publisher_details  
(  
    publisher_id INT NOT NULL,  
    publisher_name INT NOT NULL,  
    country INT NOT NULL,  
    PRIMARY KEY (publisher_id)  
);
```

```
CREATE TABLE Payment_details  
(  
    payment_method INT NOT NULL,  
    account_no INT NOT NULL,  
    payment_id INT NOT NULL,  
    Bank_Name INT NOT NULL,  
    customer_id INT NOT NULL,  
    PRIMARY KEY (payment_id),  
    FOREIGN KEY (customer_id) REFERENCES Advertiser(customer_id)  
);
```

```
CREATE TABLE billing_add_info  
(  
    mobile_no INT NOT NULL,
```

```

landline INT NOT NULL,
city INT NOT NULL,
state INT NOT NULL,
country INT NOT NULL,
sub_locality INT NOT NULL,
pincode INT NOT NULL,
email_id INT NOT NULL,
billing_comp_id INT NOT NULL,
customer_id INT NOT NULL,
PRIMARY KEY (billing_comp_id),
FOREIGN KEY (customer_id) REFERENCES Advertiser(customer_id)
);

```

```

CREATE TABLE Advertiser_contact_details
(
contact_details INT NOT NULL,
customer_id INT NOT NULL,
PRIMARY KEY (contact_details, customer_id),
FOREIGN KEY (customer_id) REFERENCES Advertiser(customer_id)
);

```

```

CREATE TABLE ad_details
(
adv_size INT NOT NULL,
format INT NOT NULL,
advt_id INT NOT NULL,
duration INT NOT NULL,
order_date INT NOT NULL,
product_name INT NOT NULL,
balance INT,
publisher_id INT NOT NULL,
customer_id INT NOT NULL,
product_code INT NOT NULL,
payment_id INT NOT NULL,
billing_comp_id INT NOT NULL,
PRIMARY KEY (advt_id),
FOREIGN KEY (publisher_id) REFERENCES publisher_details(publisher_id),
FOREIGN KEY (customer_id) REFERENCES Advertiser(customer_id),
FOREIGN KEY (product_code) REFERENCES product_combo(product_code),
FOREIGN KEY (payment_id) REFERENCES Payment_details(payment_id),
FOREIGN KEY (billing_comp_id) REFERENCES billing_add_info(billing_comp_id)
);

```

```

CREATE TABLE feedback
(
likes INT,
viewers INT,
response INT,
tweets INT,
shared INT,
publisher_id INT NOT NULL,
advt_id INT NOT NULL,
FOREIGN KEY (publisher_id) REFERENCES publisher_details(publisher_id),
FOREIGN KEY (advt_id) REFERENCES ad_details(advt_id)
);

```

```

show tables;

```


Business queries along with business benefits and query definition:

16 July 2016 at 10:30 pm

1. Which companies are our frequent customers and how many orders they give each month?

- **Business benefit:** It will help in analyzing frequent customers and framing attractive schemes for them. Also in advising future clients for suitable advertisement months.
- **Query definition:**

```
SELECT
    a.customer_id'Company_id',
    adv.company_name'Company_Name', adv.company_branch 'branch',adv.country'country',
    monthname(a.order_date) 'month_name',COUNT(a.advt_id)'ads_printed'
FROM
    Advertiser adv,
    Ad_details a
WHERE
    adv.customer_id = a.customer_id
AND monthname(a.Order_date) = (SELECT
    monthname(od.Order_date)
    FROM
        Ad_details od where od.customer_id=a.customer_id)
GROUP BY a.customer_id , 5
having count(a.advt_id) > 3 ;
```

2. Profit earnings month wise with each company including count of orders?

Business benefit: To track our sales monthly and which month we need to bring attractive schemes for new customers.

Query definition:

```
SELECT
    adv.customer_id 'customer-id',
    adv.company_name 'company_name',
    adv.company_branch 'branch',
    adv.country 'country',
    MONTHNAME(a.order_date) 'month',
    COUNT(*) 'advertisemnets_undertaken',
    SUM(p.price * a.duration) 'total price'
FROM
    advertiser adv,
```

```

ad_details a,
product_combo p
WHERE
    adv.customer_id = a.customer_id
    AND p.product_code = a.product_code
    AND MONTHNAME(a.order_date) = (SELECT
        MONTHNAME(av.order_date)
    FROM
        ad_details av
    WHERE
        av.customer_id = a.customer_id)
GROUP BY adv.customer_id , 5
order by 7 desc;

```

3. Which is the most opted combination of sub category and category by companies?

Business benefit: To analyze which combination is most favored and to work on it with publishers

to bring new innovations in it.

Query definition:

```

SELECT
    p.subcategory 'sub-category',
    p.category 'category',
    p.product_code 'product-code',
    COUNT(a.advt_id)
FROM
    product_combo p,
    ad_details a
WHERE
    a.product_code = p.product_code
GROUP BY p.product_code
ORDER BY 4 DESC
LIMIT 0 , 1;

```

4. Which companies are frequent ad takers in sub category 'food'?

Business benefit: To analyze which companies' food products are favored and bring new ideas for business benefit in food domain advertisements.

Query Definition:

```

SELECT
    ad.company_name 'company name',
    a.customer_id 'company id',
    ad.company_branch 'company branch',
    ad.country 'country',
    p.subcategory 'sub_category_name',
    COUNT(a.advt_id) 'count'
FROM
    Advertiser ad,
    ad_details a,
    product_combo p
WHERE

```

```

ad.customer_id = a.customer_id
AND a.product_code = p.product_code
AND p.subcategory = 'FOOD'
GROUP BY a.customer_id
HAVING 5 > 4;

```

5. Which sub category is most taken under given combos?

Business benefit: To analyze sub categories and bring new combinations with them.

Business query:

```

SELECT
p.subcategory 'subcategory',
COUNT(p.product_code) 'product_code'
FROM
product_combo p,
ad_details a
WHERE
a.product_code = p.product_code
GROUP BY p.subcategory
ORDER BY 2 DESC
LIMIT 0 , 4;

```

6. Which are the top 3 most favorable country for our business?

Business benefit: To analyze where should we open new ventures and extend our advertisement business.

Business query:

```

SELECT
'Most favourable country for our business is ',
b.country,
COUNT(a.advt_id)'no of advertisements offered'
FROM
advertiser adv,
ad_details a,
billing_add_info b
WHERE
adv.country = b.country
AND adv.customer_id = a.customer_id
GROUP BY b.country
ORDER BY 2 DESC
LIMIT 0 , 3;

```

7. Which publisher is best suited under each Subcategory?

Business benefit: Which publisher gets the best rating. To suggest for new clients which publisher could suit best for there advertisement under several categories.

Query Definition:

```

SELECT
    pr.subcategory,
    p.publisher_name,
    COUNT(pr.product_code) 'advertisements undertaken',
    SUM(likes) 'likes',
    SUM(response) 'response',
    SUM(viewers) 'viewers',
    SUM(shared) 'shared',
    SUM(tweets) 'tweets'
FROM
    publisher_details p,
    product_combo pr,
    ad_details a,
    feedback f
WHERE
    p.publisher_id = a.publisher_id
    AND a.publisher_id = f.publisher_id
    AND pr.product_code = a.product_code
    AND f.advt_id = a.advt_id
GROUP BY pr.subcategory , p.publisher_name
ORDER BY 3 DESC
LIMIT 0 , 10;

```

8. Which product combo is most responded?

Business benefit: Market study of categories and sub categories filtered for each product. To analyze consumers demand of each product and with what tagline they prefer. Eg. Is Bournvita more responded as energy tagline or as a milk flavor.

Query definition:

```

SELECT
    p.product_code 'product code',
    p.category 'product category',
    p.subcategory 'product sub category',
    SUM(likes),
    SUM(tweets),
    SUM(shared),
    SUM(response),
    SUM(viewers),
    pub.publisher_name 'publisher name',
    COUNT(a.advt_id)
FROM
    publisher_details pub,
    product_combo p,
    ad_details a,
    feedback f
WHERE
    p.product_code = a.product_code
    AND pub.publisher_id = a.publisher_id
    AND f.publisher_id = a.publisher_id
    AND f.advt_id = a.advt_id
GROUP BY p.product_code , pub.publisher_name
ORDER BY 4 , 5 , 6 , 7 , 8 DESC

```

```
LIMIT 0 , 4;
```

_9. Which category is not responded by viewers?

Business benefit: To suggest advertisers before they advertise for the products. To bring up new publishers with new taglines for that category to bring up sale.

Query definition:

```
SELECT
  p.product_code 'product_code',
  p.category 'product category',
  p.subcategory 'product sub category',
  SUM(viewers) 'viewers'
FROM
  product_combo p,
  ad_details a,
  feedback f,
  publisher_details pub
WHERE
  p.product_code = a.product_code
  AND f.publisher_id = a.publisher_id
  AND f.advt_id = a.advt_id
GROUP BY p.product_code
HAVING 4 < 50;
```

10. How many companies did not return after one order and under which advertiser they published there add?

Business benefit: To track sale with respect to each company and to analyze where our published ads for there products failed

Business query:

```
SELECT
  a.customer_id 'company id',
  adv.company_name 'company_name',
  adv.company_branch 'company_branch',
  COUNT(a.advt_id) 'no.of advertisements',
  p.publisher_name 'publisher-name'
FROM
  advertiser adv,
  ad_details a,
  publisher_details p
WHERE
  a.customer_id = adv.customer_id
  AND p.publisher_id = a.publisher_id
GROUP BY a.customer_id , p.publisher_name
HAVING 4 < 1;
```

11. which companies opt for small advertisement but are frequent visitors?

Business benefit: To analyze the numbers of small benefit giving companies and to come up with new ideas to attract multiple companies at a time.

Query definition:

```
SELECT
    a.customer_id 'company id',
    a.company_name 'company-name',
    a.company_branch 'company branch',
    COUNT(aa.advt_id)
FROM
    advertiser a,
    ad_details aa
WHERE
    a.customer_id = aa.customer_id
    AND aa.adv_size = 'small'
    AND aa.duration < 5
GROUP BY a.customer_id
ORDER BY 4 DESC
LIMIT 0 , 3;
```

12. which bank is mostly opted for transactions by our clients?

Business benefit: To tie up with banks and attract clients with discounts in alignment with those banks. Also where Our company could get benefits in banking services.

Query definition:

```
SELECT
    pa.bank_name 'Bank Name',
    COUNT(DISTINCT a.customer_id) 'no.of companies transacting'
FROM
    Payment_details pa,
    ad_details a,
    Advertiser aa
WHERE
    a.customer_id = aa.customer_id
    AND pa.payment_id = a.payment_id
GROUP BY pa.bank_name
ORDER BY 2 DESC
LIMIT 0;
```

13. Which duration is most favored by the companies under different categories?

Business benefit: To align the prices accordingly seeing the duration favored by companies in market for various categories.

Query definition:

```
SELECT
    a.duration 'preferred no.of days',
    COUNT(DISTINCT a.customer_id) 'Count of companies',
    p.product_code 'product combo',
    p.category 'category',
    p.subcategory 'subcategory'
FROM
    ad_details a,
    product_combo p
WHERE
```

```

    p.product_code = a.product_code
GROUP BY p.product_code , a.duration
ORDER BY 3 DESC
LIMIT 0;

```

14. Which category is least opted by customers?

Business benefit: To do market study for that particular category and take decision for dropping that category or customizing it with others.

Query definition:

```

SELECT
    p.product_code 'least opted product code',
    p.category 'product category',
    p.subcategory 'product sub category',
    COUNT(a.advt_id) 'advertisements taken'
FROM
    product_combo p,
    ad_details a,
    advertiser adv
WHERE
    p.product_code = a.product_code
    AND a.customer_id = adv.customer_id
GROUP BY p.product_code
HAVING 5 < 3
LIMIT 0 , 4;

```