**Case Study #6 - Clique Bait**

**A. Digital Analysis**

**1 - How many users are there?**

SELECT

COUNT(DISTINCT user\_id) AS user\_count

FROM users

**2 - How many cookies does each user have on average?**

SELECT

ROUND(AVG(cookie\_count)) AS avg\_cookie

FROM

(SELECT

user\_id,

COUNT(cookie\_id) AS cookie\_count

FROM users

GROUP BY user\_id) AS sub

**3 - What is the unique number of visits by all users per month?**

SELECT

MONTH(event\_time) as calendar\_month,

COUNT(DISTINCT visit\_id) as visit\_count

FROM events

GROUP BY calendar\_month

**4 - What is the number of events for each event type?**

SELECT

event\_type,

COUNT(\*) as event\_count

FROM events

GROUP BY event\_type

**5 - What is the percentage of visits which have a purchase event?**

SELECT

COUNT(DISTINCT visit\_id) / (SELECT COUNT(DISTINCT visit\_id) FROM events) \* 100 as percentage

FROM events AS e

JOIN event\_identifier AS ei ON e.event\_type = ei.event\_type

WHERE event\_name = 'Purchase'

**6 - What is the percentage of visits which view the checkout page but do not have a purchase event?**

SELECT ROUND((1 - SUM(purchase\_count)/SUM(view\_checkout\_count) ) \* 100,2) AS percentage

FROM

(SELECT

visit\_id,

SUM(IF(event\_name = 'Page View' AND page\_name = 'Checkout',1,0)) AS view\_checkout\_count,

SUM(IF(event\_name = 'Purchase',1,0)) AS purchase\_count

FROM events AS e

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

GROUP BY visit\_id

) as x

**7 - What are the top 3 pages by number of views?**

SELECT

page\_name,

COUNT(\*) AS view\_count

FROM events AS e

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

WHERE event\_name = 'Page View'

GROUP BY page\_name

ORDER BY view\_count DESC

LIMIT 3

**8 - What is the number of views and cart adds for each product category?**

SELECT

product\_category,

SUM(IF(event\_name = 'Page View',1,0)) AS view\_count,

SUM(IF(event\_name = 'Add to Cart',1,0)) AS cart\_add

FROM events AS e

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

WHERE product\_category IS NOT NULL

GROUP BY product\_category

ORDER BY view\_count DESC

**9 - What are the top 3 products by purchases?**

WITH check\_purchase AS

(SELECT

DISTINCT visit\_id

FROM events

WHERE event\_type=3)

SELECT

page\_name,

SUM(IF(event\_name = 'Add to Cart',1,0)) AS purchase\_count

FROM events as e

LEFT JOIN check\_purchase AS cp ON cp.visit\_id = e.visit\_id

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

WHERE product\_id IS NOT NULL

GROUP BY page\_name

ORDER BY purchase\_count DESC

LIMIT 3

**B. Product Funnel Analysis**

Using a single SQL query - create a new output table which has the following details:

1. How many times was each product viewed?
2. How many times was each product added to cart?
3. How many times was each product added to a cart but not purchased (abandoned)?
4. How many times was each product purchased?

CREATE VIEW product\_analysis AS

WITH check\_purchase AS

(SELECT

visit\_id,

1 AS is\_purchase

FROM events

WHERE event\_type=3)

SELECT

product\_id,

product\_category,

page\_name,

SUM(IF(event\_name = 'Page View',1,0)) AS views,

SUM(IF(event\_name = 'Add to Cart',1,0)) AS cart\_adds,

SUM(IF(event\_name = 'Add to Cart' and is\_purchase IS NULL,1,0)) AS abandoned,

SUM(IF(event\_name = 'Add to Cart' and is\_purchase = 1,1,0)) AS purchased

FROM events as e

LEFT JOIN check\_purchase AS cp ON cp.visit\_id = e.visit\_id

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

WHERE product\_id IS NOT NULL

GROUP BY product\_id,page\_name,product\_category

Additionally, create another table which further aggregates the data for the above points but this time for each product category instead of individual products.

WITH check\_purchase AS

(SELECT

visit\_id,

1 AS is\_purchase

FROM events

WHERE event\_type=3)

SELECT

product\_category,

SUM(IF(event\_name = 'Page View',1,0)) AS views,

SUM(IF(event\_name = 'Add to Cart',1,0)) AS cart\_adds,

SUM(IF(event\_name = 'Add to Cart' and is\_purchase IS NULL,1,0)) AS abandoned,

SUM(IF(event\_name = 'Add to Cart' and is\_purchase = 1,1,0)) AS purchased

FROM events as e

LEFT JOIN check\_purchase AS cp ON cp.visit\_id = e.visit\_id

JOIN event\_identifier AS ei ON ei.event\_type = e.event\_type

JOIN page\_hierarchy AS ph ON ph.page\_id = e.page\_id

WHERE product\_id IS NOT NULL

GROUP BY product\_category

**1 - Which product had the most views, cart adds and purchases?**

SELECT

\*

FROM product\_analysis

ORDER BY views DESC,cart\_adds DESC,purchased DESC

**2 - Which product was most likely to be abandoned?**

SELECT

\*

FROM product\_analysis

ORDER BY abandoned DESC

LIMIT 1

**3 - Which product had the highest view to purchase percentage?**

SELECT

product\_id,

product\_category,

page\_name,

ROUND(purchased/views \* 100,2) AS percentage

FROM product\_analysis

ORDER BY percentage DESC

LIMIT 1

**4 - What is the average conversion rate from view to cart add?**

SELECT

ROUND(AVG(cart\_adds/views \* 100),2)AS percentage

FROM product\_analysis

ORDER BY percentage

**5 - What is the average conversion rate from cart add to purchase?**

SELECT

ROUND(AVG(purchased/cart\_adds \* 100),2)AS percentage

FROM product\_analysis

ORDER BY percentage

**C. Campaigns Analysis**

Generate a table that has 1 single row for every unique visit\_id record and has the following columns:

* user\_id
* visit\_id
* visit\_start\_time: the earliest event\_time for each visit
* page\_views: count of page views for each visit
* cart\_adds: count of product cart add events for each visit
* purchase: 1/0 flag if a purchase event exists for each visit
* campaign\_name: map the visit to a campaign if the visit\_start\_time falls between the start\_date and end\_date
* impression: count of ad impressions for each visit
* click: count of ad clicks for each visit
* (Optional column) cart\_products: a comma separated text value with products added to the cart sorted by the order they were added to the cart (hint: use the sequence\_number)

WITH subtable AS

(SELECT

\*,

FIRST\_VALUE(event\_time) OVER (PARTITION BY visit\_id) AS visit\_start\_time,

LAST\_VALUE(event\_type) OVER (PARTITION BY visit\_id) AS last\_event

FROM events)

SELECT

visit\_id,

user\_id,

visit\_start\_time,

SUM(IF(event\_type = 1,1,0)) AS page\_views,

SUM(IF(event\_type = 2,1,0)) AS cart\_adds,

IF(last\_event = 3,1,0) AS purchase,

campaign\_name,

SUM(IF(event\_type = 4,1,0)) AS impression,

SUM(IF(event\_type = 5,1,0)) AS click,

GROUP\_CONCAT(IF(event\_type = 2 and product\_id IS NOT NULL,page\_name,NULL)) as cart\_products

FROM subtable AS s

JOIN users AS u ON u.cookie\_id = s.cookie\_id

JOIN page\_hierarchy AS ph ON ph.page\_id = s.page\_id

LEFT JOIN campaign\_identifier AS ci on s.event\_time BETWEEN ci.start\_date AND ci.end\_date

GROUP BY

visit\_id,

user\_id,

visit\_start\_time,

purchase,

campaign\_name