#Staging data in cloud storage using the google cloud shell

echo "Creating bucket: gs://$DEVSHELL\_PROJECT\_ID"

gsutil mb gs://$DEVSHELL\_PROJECT\_ID

echo "Copying data to our storage from public dataset"

gsutil cp gs://cloud-training/bdml/v2.0/data/data-to-insights.csv gs://$DEVSHELL\_PROJECT\_ID

gsutil cp gs://cloud-training/bdml/v2.0/data/google-analytics-360.csv gs://$DEVSHELL\_PROJECT\_ID

echo "Show the files in our bucket"

gsutil ls gs://$DEVSHELL\_PROJECT\_ID

echo "View some sample data"

gsutil cat gs://$DEVSHELL\_PROJECT\_ID/data-to-insights.csv

gsutil cat gs://$DEVSHELL\_PROJECT\_ID/ google-analytics-360.csv

#Exploring the ecommerce data

# Out of the total visitors who visited our website, what % made a purchase?

WITH visitors AS(

SELECT

COUNT(DISTINCT fullVisitorId) AS total\_visitors

FROM `data-to-insights.ecommerce.web\_analytics`

),

purchasers AS(

SELECT

COUNT(DISTINCT fullVisitorId) AS total\_purchasers

FROM `data-to-insights.ecommerce.web\_analytics`

WHERE totals.transactions IS NOT NULL

)

SELECT

total\_visitors,

total\_purchasers,

total\_purchasers / total\_visitors AS conversion\_rate

FROM visitors, purchasers

# What are the top 5 selling products?

SELECT

p.v2ProductName,

p.v2ProductCategory,

SUM(p.productQuantity) AS units\_sold,

ROUND(SUM(p.localProductRevenue/1000000),2) AS revenue

FROM `data-to-insights.ecommerce.web\_analytics`,

UNNEST(hits) AS h,

UNNEST(h.product) AS p

GROUP BY 1, 2

ORDER BY revenue DESC

LIMIT 5;

# How many visitors bought on subsequent visits to the website?

# visitors who bought on a return visit (could have bought on first as well

WITH all\_visitor\_stats AS (

SELECT

fullvisitorid, # 741,721 unique visitors

IF(COUNTIF(totals.transactions > 0 AND totals.newVisits IS NULL) > 0, 1, 0) AS will\_buy\_on\_return\_visit

FROM `data-to-insights.ecommerce.web\_analytics`

GROUP BY fullvisitorid

)

SELECT

COUNT(DISTINCT fullvisitorid) AS total\_visitors,

will\_buy\_on\_return\_visit

FROM all\_visitor\_stats

GROUP BY will\_buy\_on\_return\_visit

#What is the total number of transactions generated per device browser in July 2017?

SELECT

device.browser,

SUM ( totals.transactions ) AS total\_transactions

FROM `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_\*`

WHERE

\_TABLE\_SUFFIX BETWEEN '20170701' AND '20170731'

GROUP BY

device.browser

ORDER BY

total\_transactions DESC

# What was the real bounce rate per traffic source?

SELECT

source,

total\_visits,

total\_no\_of\_bounces,

( ( total\_no\_of\_bounces / total\_visits ) \* 100 ) AS bounce\_rate

FROM (

SELECT

trafficSource.source AS source,

COUNT ( trafficSource.source ) AS total\_visits,

SUM ( totals.bounces ) AS total\_no\_of\_bounces

FROM `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_\*`

WHERE

\_TABLE\_SUFFIX BETWEEN '20170701' AND '20170731'

GROUP BY

source )

ORDER BY

total\_visits DESC

# What was the average number of product pageviews for users who made a purchase in July 2017?

SELECT

( SUM(total\_pagesviews\_per\_user) / COUNT(users) ) AS avg\_pageviews\_per\_user

FROM

SELECT

fullVisitorId AS users,

SUM(totals.pageviews) AS total\_pagesviews\_per\_user

FROM`bigquery-public-data.google\_analytics\_sample.ga\_sessions\_\*`

WHERE

\_TABLE\_SUFFIX BETWEEN '20170701' AND '20170731'

AND

totals.transactions >=1

GROUP BY

users )