SQL Project – Google Store Visitor Data

BUAN 6320.006

My Pham, Jiaxing Ji, Yang Chen

Group #: 23

Contents

Data Model

Assumptions/Notes About Data Entities and Relationships

Visitors are determined by FullVisitorID. Visitors also have information of SocialEngagementType.

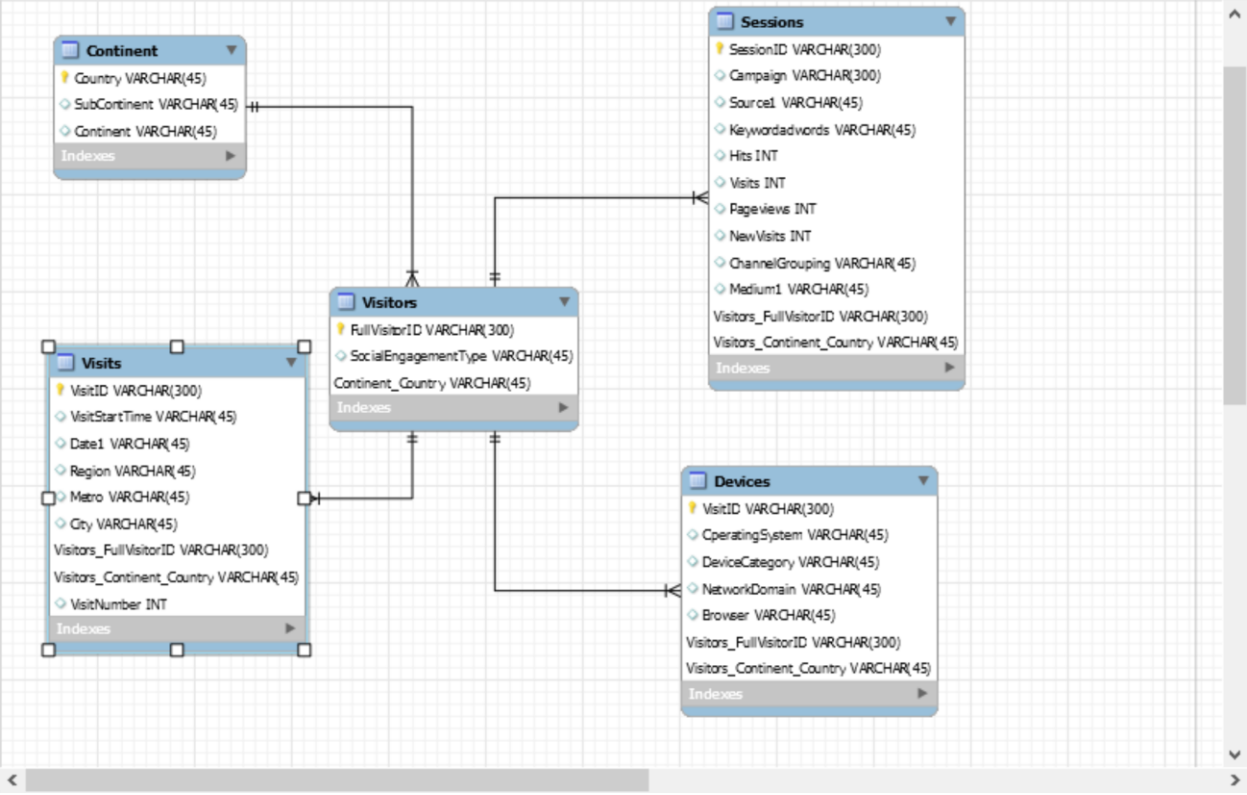
Each continent contains different subcontinents, which include different countries. Each continent can have multiple visitors (users), and each visitor is from only one continent.

Each visit to the store is determined by VisitID. A visit has other information such as VisitNumber, VisitStartTime, Date, and location information such as Region, Metro and City. A visitor can have more than one visit, and each visit is specified by only one visitor.

Each visitor can have different sessions, determined by SessionID (a combination of FullVisitorID and VisitID). A session includes information of Campaign, Source, Keyword, Hits, Visits, Pageviews, NewVisits, ChannelGrouping and Medium. Each session may have only one user.

Devices, determined by VisitID, include information of Browser, OperatingSystem, DeviceCategory and NetworkDomain. A user may have different devices, and each device can only be used by one user.

Entity-Relationship Diagram

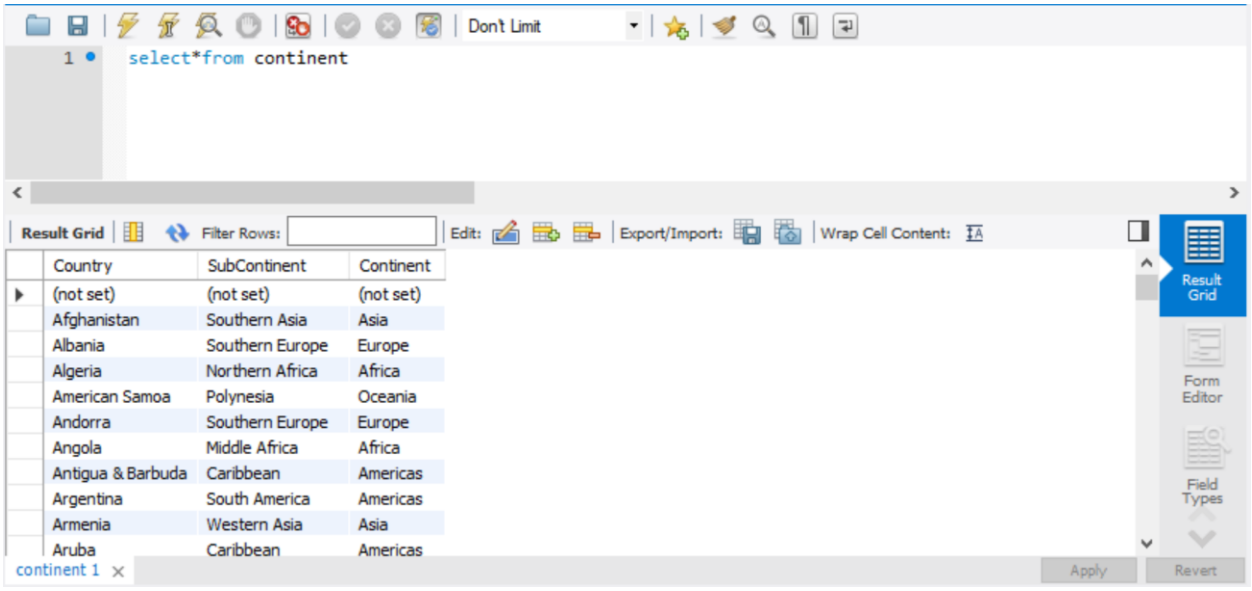


Physical Database

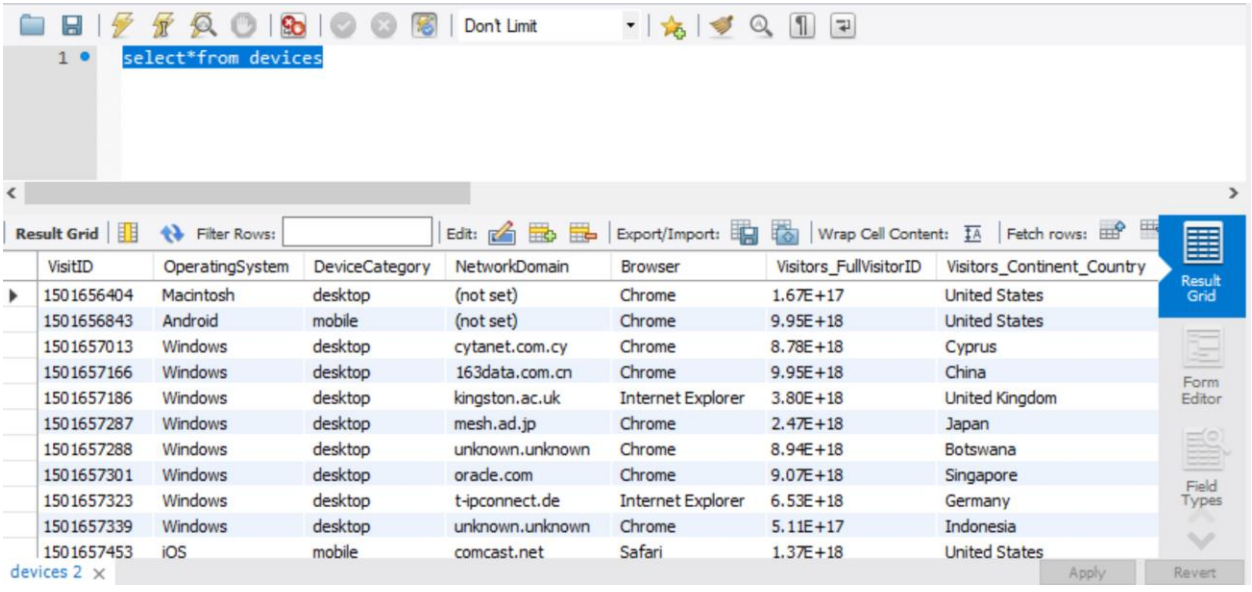
Assumptions/Notes About Data Set

We imported whole database into one table, project1, and insert the columns into different tables

Screen shot of Physical Database objects



Country	SubContinent	Continent
(not set)	(not set)	(not set)
Afghanistan	Southern Asia	Asia
Albania	Southern Europe	Europe
Algeria	Northern Africa	Africa
American Samoa	Polynesia	Oceania
Andorra	Southern Europe	Europe
Angola	Middle Africa	Africa
Antigua & Barbuda	Caribbean	Americas
Argentina	South America	Americas
Armenia	Western Asia	Asia
Aruba	Caribbean	Americas



VisitID	OperatingSystem	DeviceCategory	NetworkDomain	Browser	Visitors_FullVisitorID	Visitors_Continent_Country
1501656404	Macintosh	desktop	(not set)	Chrome	1.67E+17	United States
1501656843	Android	mobile	(not set)	Chrome	9.95E+18	United States
1501657013	Windows	desktop	cytanet.com.cy	Chrome	8.78E+18	Cyprus
1501657166	Windows	desktop	163data.com.cn	Chrome	9.95E+18	China
1501657186	Windows	desktop	kingston.ac.uk	Internet Explorer	3.80E+18	United Kingdom
1501657287	Windows	desktop	mesh.ad.jp	Chrome	2.47E+18	Japan
1501657288	Windows	desktop	unknown.unknown	Chrome	8.94E+18	Botswana
1501657301	Windows	desktop	oracle.com	Chrome	9.07E+18	Singapore
1501657323	Windows	desktop	t-ipconnect.de	Internet Explorer	6.53E+18	Germany
1501657339	Windows	desktop	unknown.unknown	Chrome	5.11E+17	Indonesia
1501657453	iOS	mobile	comcast.net	Safari	1.37E+18	United States

```
1 • select*from sessions
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows:

SessionID	Campaign	Source1	Keywordadwords	Hits	Visits	Pageviews	NewVisits	ChannelGrou
0000000259678714014_1511912001	(not set)	google	keyword	3	1	3	1	Organic Sear
0000000259678714014_1511914780	(not set)	google	keyword	16	1	10	0	Organic Sear
0000049363351866189_1505779414	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000049363351866189_1505796426	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000049363351866189_1505863874	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000053049821714864_1517178742	Data Share Promo	Partners	adwordsClickInfo	1	1	1	1	Affiliates
0000059488412965267_1519200377	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000059488412965267_1519772893	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000059488412965267_1519787699	(not set)	(direct)	adwordsClickInfo	1	1	1	1	Direct
0000085840370633780_1504918786	(not set)	google	keyword	2	1	2	1	Organic Sear

Result Grid | Form Editor | Field Types

sessions 3 ×

Apply | Revert

```
1 • select*from visitors
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows:

FullVisitorID	SocialEngagementType	Continent_Country
"desktop}"""	Not Socially Engaged	United States
1.00E+15	Not Socially Engaged	Philippines
1.00E+16	Not Socially Engaged	United States
1.00E+17	Not Socially Engaged	United States
1.00E+18	Not Socially Engaged	United States
1.00E+19	Not Socially Engaged	United States
1.01E+15	Not Socially Engaged	Turkey
1.01E+16	Not Socially Engaged	China
1.01E+17	Not Socially Engaged	Argentina
1.01E+18	Not Socially Engaged	United States
1.02E+16	Not Socially Engaged	United States

Result Grid | Form Editor | Field Types

visitors 4 ×

Apply | Revert

```
1 • select*from visits
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows:

VisitID	VisitStartTime	Date1	Region	Metro	City
1501656404	1501657280	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501656843	1501657213	20170802	California	San Francisco-Oakland-San Jose CA	Mountain View
1501657013	1501657239	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501657166	1501657216	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501657186	1501657203	20170802	England	London	London
1501657287	1501657287	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501657288	1501657288	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501657301	1501657301	20170802	(not set)	(not set)	(not set)
1501657323	1501657323	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset
1501657339	1501657339	20170802	not available in demo dataset	not available in demo dataset	not available in demo dataset

Result Grid | Form Editor | Field Types

visits 5 ×

Apply | Revert

Data in the Database

Table Name	Primary Key	Foreign Key	# of Rows in Table
Visitors	FullVisitorId	Continent_Country	3827
Visits	VisitID	Visitors_FullVisitorID,Visitors_Continent_Country	779,504
Continents	Country		219
Devices	VisitID	Visitors_FullVisitorID,Visitors_Continent_Country	3,827
Sessions	SessionID	Visitors_FullVisitorID,Visitors_Continent_Country	803,863

SQL Queries

Query 1

Question

Which user had the maximum number of visits and when?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

VisitorID: 1957458976293870000 on 2018-04-16

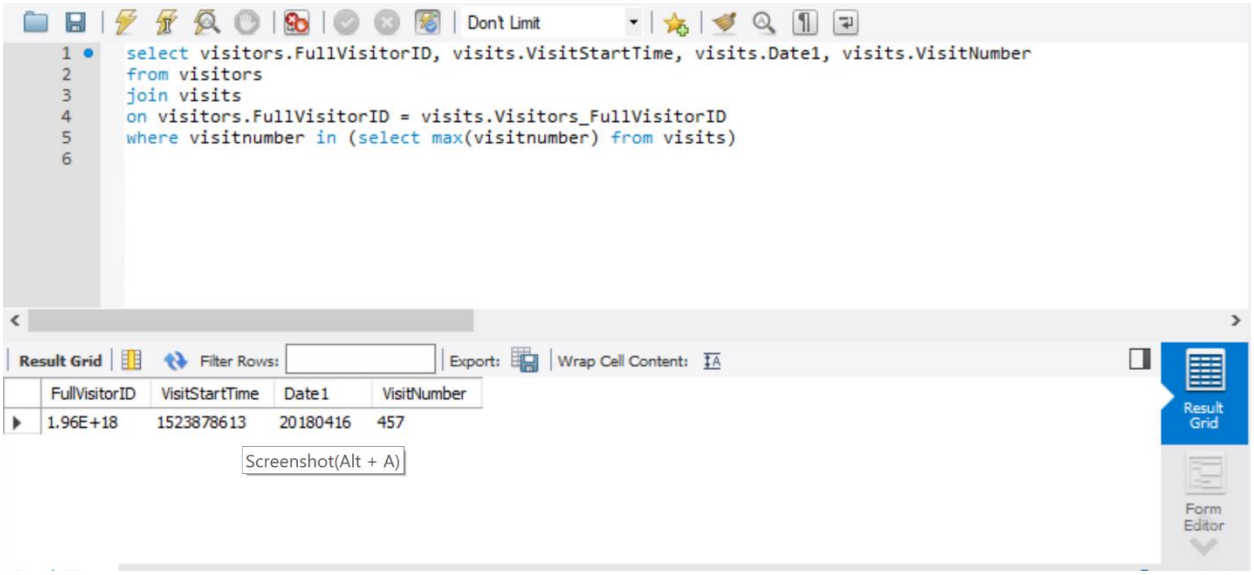
18 rows

Translation

select visitorid, visitstarttime,date,visitnumber from the visitors table joined with the visits table on visitorid where visitnumber is maximum.

Screen Shot of SQL Query and Results

select visitors.FullVisitorID, visits.VisitStartTime, visits.Date1, visits.VisitNumber
from visitors join visits on visitors.FullVisitorID = visits.Visitors_FullVisitorID
where visitnumber in (select max(visitnumber) from visits);



Query 2

Question

Which operating system (devices) was the most popular amongst store visitors?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

Windows

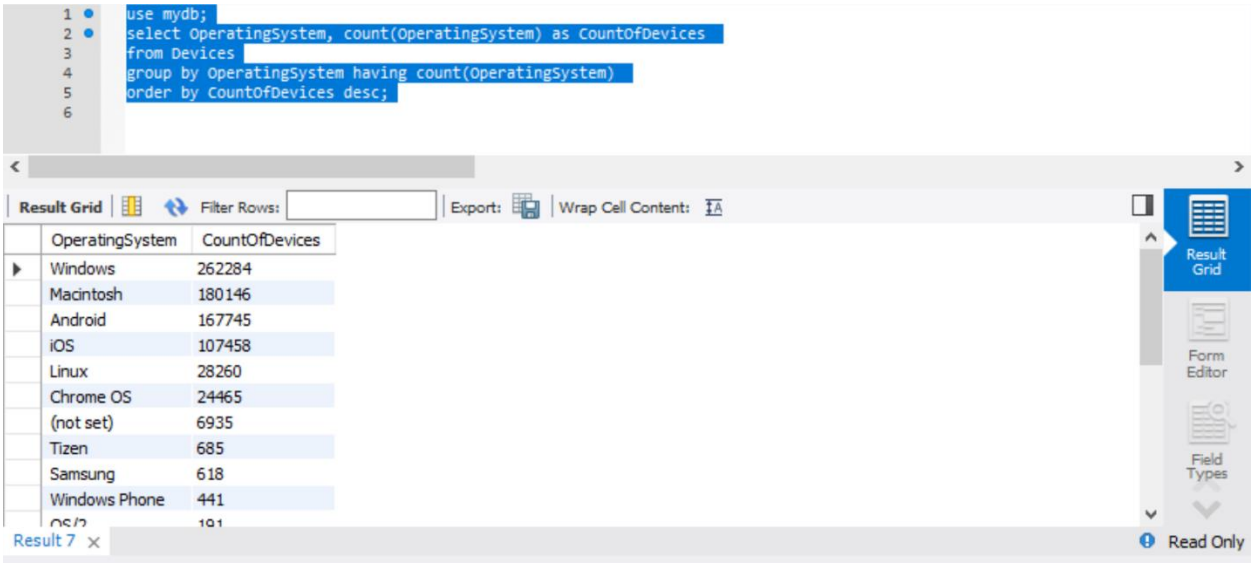
23 rows

Translation

select operatingsystem and count of operatingsystem as countofdevices from the devices table, which group by operatingsystem and having count of operatingsystem and order by descending order of count of devices

Screen Shot of SQL Query and Results

select OperatingSystem, count(OperatingSystem) as CountOfDevices
from Devices
group by OperatingSystem having count(OperatingSystem)
order by CountOfDevices desc;



OperatingSystem	CountOfDevices
Windows	262284
Macintosh	180146
Android	167745
iOS	107458
Linux	28260
Chrome OS	24465
(not set)	6935
Tizen	685
Samsung	618
Windows Phone	441
OS/2	191

Query 3

Question

Which date had the least and most number of visits?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

2017-08-05, 2017-12-13

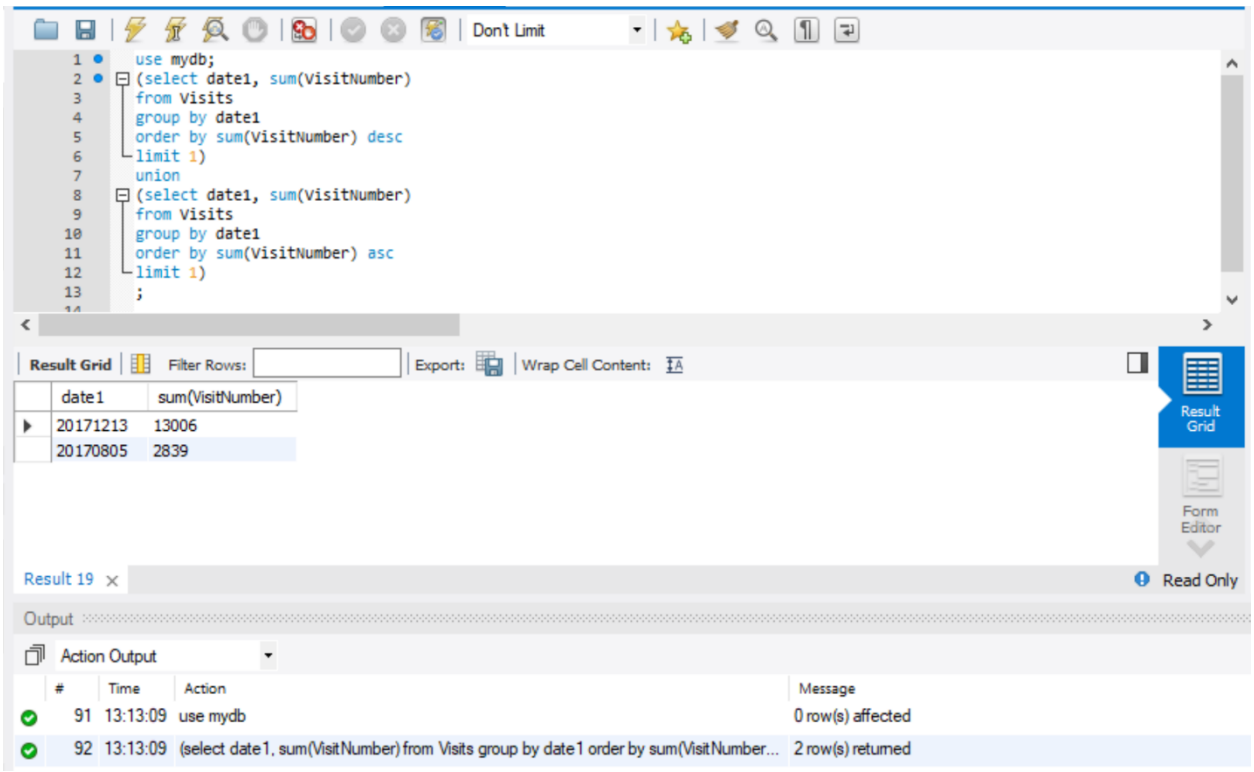
2 rows

Translation

select only 1 row of date and sum of visit number from the visits table, group by date and order by descending order of sum of visit number union select only 1 row of date and sum of visit number from the visits table, group by date and order by ascending order of sum of visit number

Screen Shot of SQL Query and Results

use mydb;
(select date1, sum(VisitNumber) from Visits group by date1 order by sum(VisitNumber) desc
limit 1)
union
(select date1, sum(VisitNumber) from Visits group by date1 order by sum(VisitNumber) asc
limit 1);



Query 4

Question

Which country had the most number of iOS users who were socially engaged?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

No country

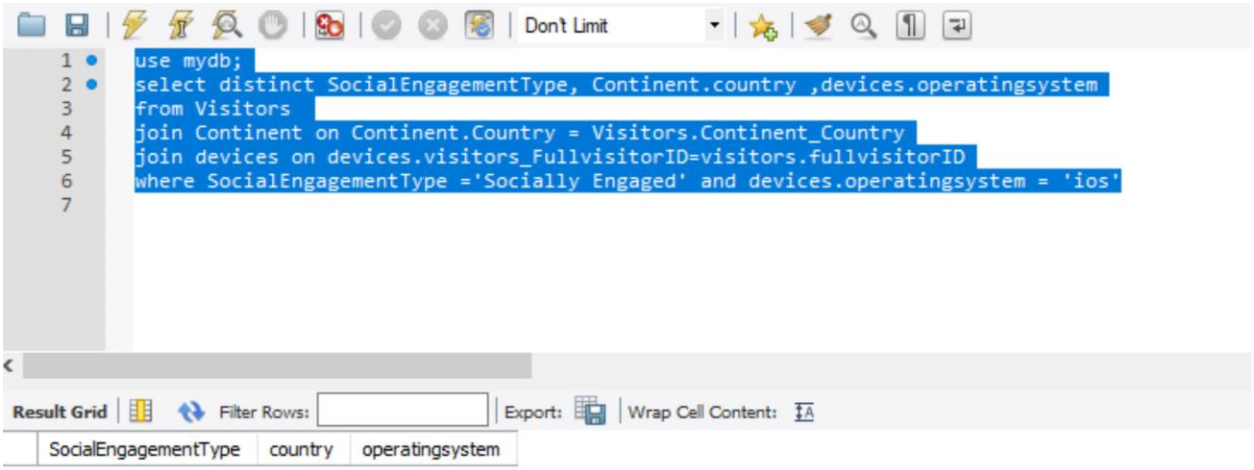
0 Row

Translation

select country, social engagement type, operatingsystem from the visitors table joined with the continent table on country, joined with the devices table on visitorid, where social engagement type is socially engaged and operatingsystem is iOS.

Screen Shot of SQL Query and Results

use mydb;
select distinct SocialEngagementType, Continent.country ,devices.operatingsystem from Visitors
join Continent on Continent.Country = Visitors.Continent_Country
join devices on devices.visitors_FullvisitorID=visitors.fullvisitorID
where SocialEngagementType ='Socially Engaged' and devices.operatingsystem = 'ios';



Query 5

Question

Provide a breakdown of unique visitors by operating system type

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

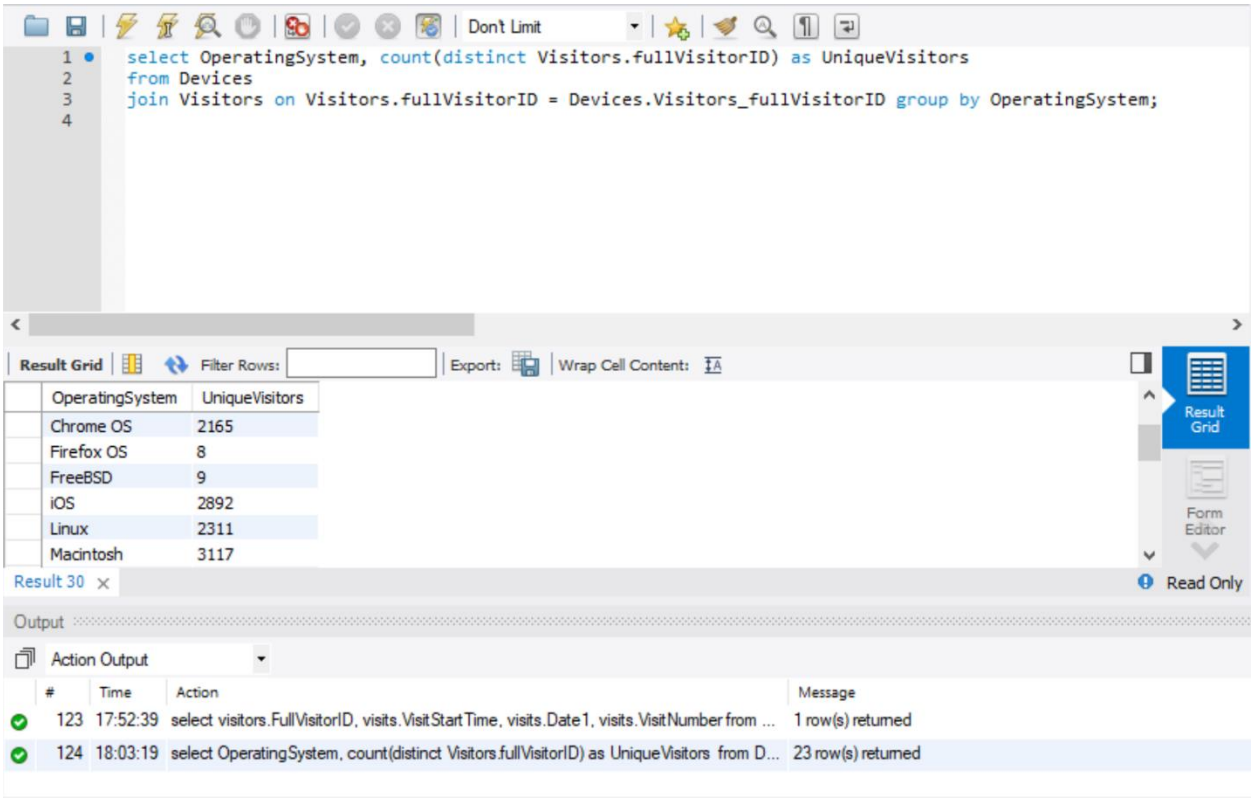
23 rows

Translation

select operatingsystem and count of each visitorid of its operatingsystem from the devices table joined with the visitors table on visitorid, group by operatingsystem

Screen Shot of SQL Query and Results

select OperatingSystem, count(distinct Visitors.fullVisitorID) as UniqueVisitors from Devices
join Visitors on Visitors.fullVisitorID = Devices.Visitors_fullVisitorID group by OperatingSystem;



Query 6

Question

Which country had the least number of hits higher than zero?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

6 countries.

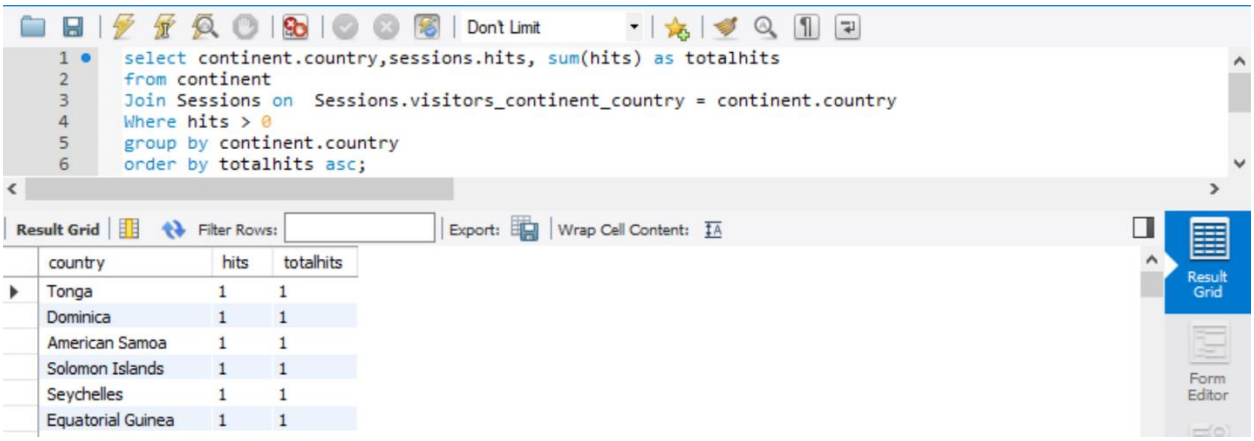
219 rows

Translation

select country, hits, and sum of hits as totalhits from the continent table joined with the sessions table on country where hits is greater than 0, group by country and order by ascending order of totalhits

Screen Shot of SQL Query and Results

select continent.country,sessions.hits, sum(hits) as totalhits from continent
Join Sessions on Sessions.visitors_continent_country = continent.country
Where hits > 0
group by continent.country
order by totalhits asc;



Query 7

Question

Which region had more blackberry users than iOS users?

Notes/Comments About SQL Query and Results (Include # of Rows in Result)

11 regions

11 rows

Translation

select region, BBnum, operatingsystem, Inumber from the table of select visit.region, operatingsystem, count of operatingsystem as BBnumer from the devices table joined with visits on visitid where the operatingsystem is blackberry and group by visits region as A
Joined with
The table of select visit.region, operatingsystem, count of operatingsystem as Inum from the devices table joined with visits on visitid where the operatingsystem is ios and group by visits region as B
Where BBnum large than Inum
Group by region

Screen Shot of SQL Query and Results

select a.region, a.operatingsystem,sum(BBnum),b.operatingsystem,sum(Inum) from
(select visits.region, operatingsystem, count(operatingsystem) as BBnum from devices
join visits on visits.visitID=devices.visitId
where operatingsystem like '%BlackBerry%' group by visits.region) as A
join
(select visits.region, operatingsystem, count(operatingsystem) as Inum from devices
join visits on visits.visitID=devices.visitId
where operatingsystem like '%ios%'
group by visits.region) as B on A.region=A.region
where a.BBnum>b.Inum group by region;