



SQL Project – Google Store Visitor Data

BUAN 6320.003

Group Members

Sai Pratheek Banda

Prithvi Gude

Nikhil Kumar Kasham

Sriharshith Reddy Nimmala

Group #: 10

Contents

Data Model	3
Assumptions/Notes About Data Entities and Relationships	3
Entity-Relationship Diagram	3
Physical Database	4
Assumptions/Notes About Data Set	4
Screen shot of Physical Database objects	4
Data in the Database	7
SQL Queries	Error! Bookmark not defined.
Query 1	Error! Bookmark not defined.
Question	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result) ...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results	Error! Bookmark not defined.
Query 2	Error! Bookmark not defined.
Question	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result) ...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results	Error! Bookmark not defined.
Query 3	Error! Bookmark not defined.
Question	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result) ...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results	Error! Bookmark not defined.
Query 4	Error! Bookmark not defined.
Question	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result) ...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results	Error! Bookmark not defined.
Query 5	Error! Bookmark not defined.

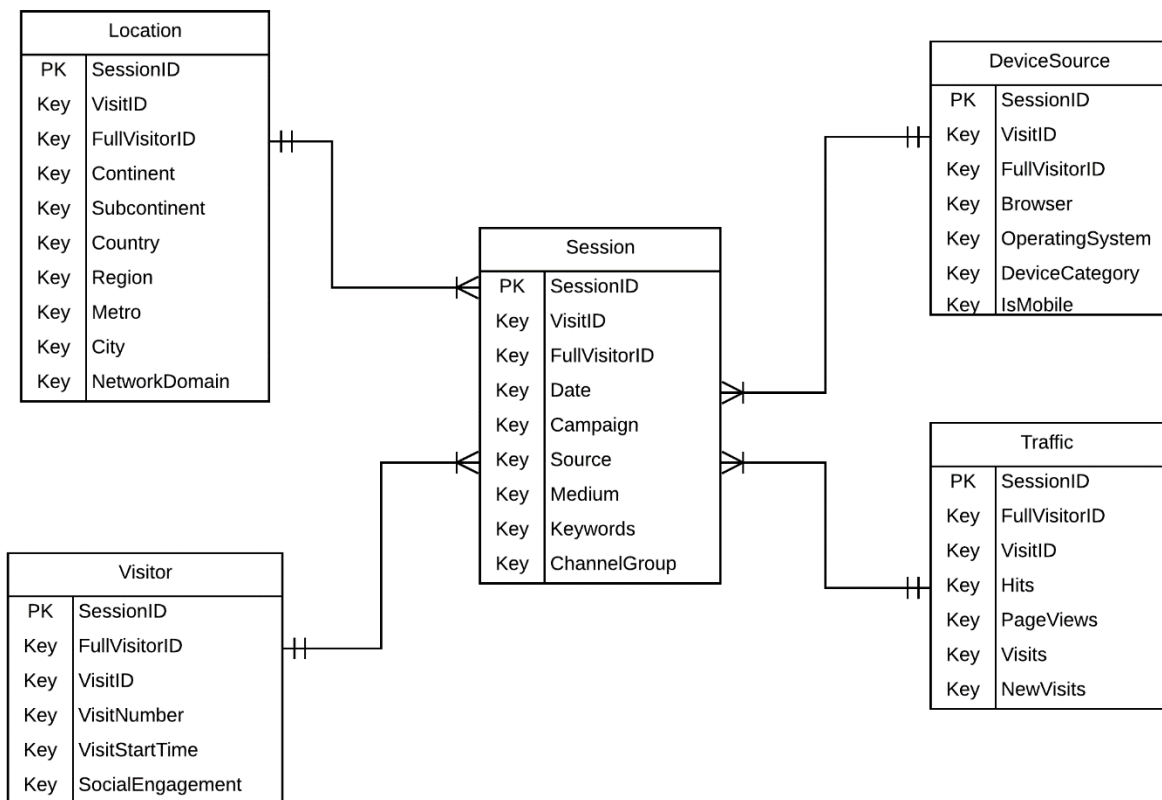
Question.....	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result)...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results.....	Error! Bookmark not defined.
Query 6.....	Error! Bookmark not defined.
Question.....	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result)...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results.....	Error! Bookmark not defined.
Query 7.....	Error! Bookmark not defined.
Question.....	Error! Bookmark not defined.
Notes/Comments About SQL Query and Results (Include # of Rows in Result)...	Error! Bookmark not defined.
Translation	Error! Bookmark not defined.
Screen Shot of SQL Query and Results.....	Error! Bookmark not defined.

Data Model

Assumptions/Notes About Data Entities and Relationships

- In the given dataset, the data is not in 1NF in the first place because, the columns “device”, “geoNetwork”, “totals” & “trafficSource” have multiple data entries, which stops us from going forward with the data model.
- As soon as those clustered columns are parsed and the data set is expanded, the data is transformed into 1 NF as every column in the data set is dependent on the sessionID
- The data set is in 2NF, because no composite primary key in our dataset. Instead all the columns are dependent solely on the primary key i.e., sessionID
- The data set is in 3NF, because every column in the dataset is dependent solely on the primary key, and there is no dependency between any other columns except for the primary key

Entity-Relationship Diagram



Physical Database

Assumptions/Notes About Data Set

The tables created are Session, DeviceSource, Location, Traffic, Visitor.

- The table “Location” contains SessionID, VisitID, FullVisitorID, Continent, Subcontinent, Country, Region, Metro, City, NetworkDomain where SessionID is the primary key. These columns have been put together in this table because all the columns give out the location details of the user and their session details.
- The table “Session” contains SessionID, VisitID, FullVisitorID, Date, Campaign, Source, Medium, Keywords and Channelgroup where SessionID is the primary key. These columns have been put together in this table because all the columns give out the session details of the user and their medium of entry, channel through which they made use of keywords and made a search, and through which campaign they got to know about the product and made the purchase
- The table “Device Source” contains SessionID, VisitID, FullVisitorID, Browser, OperatingSystem, DeviceCategory and isMobile where SessionID is the primary key. These columns have been put together in this table because all the columns give out the Device details of the user and whether they have used a mobile or desktop to make a session and the browser and OS used to run the session
- The table “Visitor” contains SessionID, VisitID, FullVisitorID, VisitNumber, VisitStarttime and SocialEngagement where SessionID is the primary key. These columns have been put together in this table because all the columns give out the respective visit details of the user and their social engagement status, also the time their session starts

MySQL Workbench

Local instance MySQL Router

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

bowlingleagueexample
entertainmentagencyexample
googlecustomerevenue
projectdb
recipesexample
sakila
salesorderexample
schoolschedulingexample
sys
world

Administration Schemas

Information

Schema: googlecustomerevenue

Query 1 Administration - Data Import/Res... devicesource location session traffic visitor

Limit to 1000 rows

1

Select FullVisitorID, VisitStartTime from visitor where VisitNumber=(Select max(VisitNumber) from Visitor);

Result Grid

FullVisitorID	VisitStartTime
7.28E+18	1508200018
4.60E+17	1511026542
7.12E+18	1516377477
8.28E+17	1516401822
3.04E+18	1515372752
3.88E+18	1507812699
7.06E+17	1514228053
9.25E+18	1519653360
1.24E+17	1511012939
7.48E+18	1507219710
8.96E+18	1519936835
3.94E+18	1502397038
8.84E+18	1511189632
5.21E+18	1516648562

Visitor 2 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
8	23:41:18	SELECT count(*)	2 row(s) returned	9.391 sec / 0.000 sec
9	23:45:25	SELECT count(*)	22 row(s) returned	0.734 sec / 0.000 sec
10	23:47:05	SELECT count(*)	2 row(s) returned	0.766 sec / 0.000 sec
11	23:47:40	SELECT count(*)	1 row(s) returned	0.469 sec / 0.000 sec
12	23:54:07	Select FullVisitorID from Visitor where VisitNumber=(Select max(VisitNumber) from Visitor) LIMIT 0, 1000	19 row(s) returned	0.953 sec / 0.000 sec
13	23:55:37	Select FullVisitorID, VisitStartTime from Visitor where VisitNumber=(Select max(VisitNumber) from Visitor) LIMIT 0, 1000	19 row(s) returned	0.922 sec / 0.000 sec

Object Info Session

Type here to search

11:55 PM 10/28/2018

Screen shot of Physical Database objects

Session

MySQL Workbench

Local instance MySQL Router

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

bowlingleagueexample
entertainmentagencyexample
googlecustomerevenue
projectdb
recipesexample
sakila
salesorderexample
schoolschedulingexample
sys
world

Administration Schemas

Information

No object selected

Query 1 Administration - Data Import/Res... devicesource location session traffic visitor

Limit to 1000 rows

1

SELECT * FROM googlecustomerevenue.session;

Result Grid

sessionid	fullvisitorid	visitid	vdate	campaign	vsources	vmedium	keyword	channelGrouping
6.17E+28	6.17E+18	1508151024	20171016	(not set)	google	organic	(not provided)	Organic Search
6.44E+27	6.44E+17	1508175522	20171016	(not set)	google	organic	(not provided)	Organic Search
6.06E+28	6.06E+18	1508143220	20171016	(not set)	google	organic	(not provided)	Organic Search
2.38E+28	2.38E+18	1508193530	20171016	(not set)	google	organic	(not provided)	Organic Search
2.31E+28	2.31E+18	1508217442	20171016	(not set)	google	organic	(not provided)	Organic Search
4.13E+28	4.13E+18	1508186358	20171016	(not set)	google	organic	(not provided)	Organic Search
4.32E+28	4.32E+18	1508203650	20171016	(not set)	google	organic	(not provided)	Organic Search
5.88E+28	5.88E+18	1508184397	20171016	(not set)	google	organic	(not provided)	Organic Search
5.15E+27	5.15E+17	1508189652	20171016	(not set)	google	organic	(not provided)	Organic Search
6.43E+28	6.43E+18	1508175502	20171016	(not set)	google	organic	(not provided)	Organic Search
7.03E+28	7.03E+18	1508190324	20171016	(not set)	google	organic	(not provided)	Organic Search
2.86E+28	2.86E+18	1508198751	20171016	AW - Dr...	google	cpc	6qHsCsdK...	Paid Search
7.91E+28	7.91E+18	1508186327	20171016	(not set)	google	cpc	NA	Display
4.45E+28	4.45E+18	1508184708	20171016	(not set)	google	organic	(not provided)	Organic Search

session1 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	23:30:03	SELECT * FROM googlecustomerevenue.devicesource LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec
2	23:30:20	SELECT * FROM googlecustomerevenue.location LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
3	23:30:38	SELECT * FROM googlecustomerevenue.session LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec

Object Info Session

Type here to search

11:30 PM 10/28/2018

Traffic

MySQL Workbench

Local instance MySQL Router

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- entertainmentagencyexample
- googlecustomerevenue
 - Tables
 - devicesource
 - location
 - session
 - traffic
 - visitor
 - Views
 - Stored Procedures
 - Functions
- projectdb
- recipesexample
- sakila
- salesorderexample
- schoolsschedulingexample
- sys
- world

Administration Schemas

Information

No object selected

Query 1 Administration - Data Import/Res... devicesource location session traffic

1. `SELECT * FROM googlecustomerevenue.traffic;`

Result Grid

sessionid	fullvisitorid	visid	hits	pageviews	visits	newVisits
61678713306171123631508151024	6167871330617112363	1508151024	4	4	1	NA
06436976409779156181508175522	0643697640977915618	1508175522	5	5	1	1
6059383810968229466143220	6059383810968229466	1508143220	7	7	1	1
23767200785634236311508193530	2376720078563423631	1508193530	8	4	1	1
23145445207954400381508217442	2314544520795440038	1508217442	9	4	1	1
4133039884103923671508186358	413303988410392367	1508186358	11	5	1	1
43204788502073975571508203650	4320478850207397557	1508203650	37	15	1	1
58764382479901571311508184397	5876438247990157131	1508184397	52	22	1	1
051491268737029441508189652	05149126873702944	1508189652	5	5	1	NA
64305670315316772121508175502	6430567031531677212	1508175502	6	6	1	1
70263740701572406531508190324	7026374070157240653	1508190324	10	10	1	1
28617243041343537791508196731	2861724304134353779	1508196731	3	3	1	1
79082471172896303661508186327	7908247117289630366	1508186327	3	3	1	1
44521279523516640461508184708	4452127952351664046	1508184708	3	3	1	1

traffic 1 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	23:30:03	SELECT * FROM googlecustomerevenue.devicesource LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.000 sec
2	23:30:20	SELECT * FROM googlecustomerevenue.location LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
3	23:30:38	SELECT * FROM googlecustomerevenue.session LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec
4	23:30:47	SELECT * FROM googlecustomerevenue.traffic LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.000 sec

Object Info Session

Type here to search

11:30 PM 10/28/2018

Visitor

MySQL Workbench

Local instance MySQL Router

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- bowlingleagueexample
- entertainmentagencyexample
- googlecustomerevenue
 - Tables
 - devicesource
 - location
 - session
 - traffic
 - visitor
 - Views
 - Stored Procedures
 - Functions
- projectdb
- recipesexample
- sakila
- salesorderexample
- schoolsschedulingexample
- sys
- world

Administration Schemas

Information

Table: visitor

Columns:

- sessionid varchar(25)
- fullvisitorid varchar(25)
- visid varchar(25)
- visNumber varchar(25)
- visStartTime varchar(25)
- socialEngagementType varchar(25)

Query 1 Administration - Data Import/Res... devicesource location session traffic visitor

1. `SELECT * FROM googlecustomerevenue.visitor;`

Result Grid

sessionid	fullvisitorid	visid	visNumber	visStartTime	socialEngagementType
6.16787E+28	6.17E+18	1508151024	2	1508151024	Not Socially Engaged
6.43698E+27	6.44E+17	1508175522	1	1508175522	Not Socially Engaged
6.05938E+28	6.06E+18	1508143220	1	1508143220	Not Socially Engaged
2.37672E+28	2.38E+18	1508193530	1	1508193530	Not Socially Engaged
2.31454E+28	2.31E+18	1508217442	1	1508217442	Not Socially Engaged
4.13304E+28	4.13E+18	1508186358	1	1508186358	Not Socially Engaged
4.32048E+28	4.32E+18	1508203650	1	1508203650	Not Socially Engaged
5.8764E+28	5.88E+18	1508184397	1	1508184397	Not Socially Engaged
5.14991E+27	5.15E+17	1508189652	6	1508189652	Not Socially Engaged
6.43057E+28	6.43E+18	1508175502	1	1508175502	Not Socially Engaged
7.02637E+28	7.03E+18	1508190324	1	1508190324	Not Socially Engaged
2.86172E+28	2.86E+18	1508196731	1	1508196731	Not Socially Engaged
7.90823E+28	7.91E+18	1508186327	1	1508186327	Not Socially Engaged
4.45212E+28	4.45E+18	1508184708	1	1508184708	Not Socially Engaged

visitor 1 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	23:30:03	SELECT * FROM googlecustomerevenue.devicesource LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.000 sec
2	23:30:20	SELECT * FROM googlecustomerevenue.location LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
3	23:30:38	SELECT * FROM googlecustomerevenue.session LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec
4	23:30:47	SELECT * FROM googlecustomerevenue.traffic LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.000 sec
5	23:30:55	SELECT * FROM googlecustomerevenue.visitor LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

Type here to search

11:30 PM 10/28/2018

DeviceSource

The screenshot shows the MySQL Workbench interface. The 'Schemas' pane on the left lists various databases, including 'googlecustomerrevenue'. The 'Query' pane in the center shows a SQL query: `SELECT * FROM googlecustomerrevenue.devicesource;`. The 'Result Grid' displays the query results, which include columns for sessionid, fullvisitorid, visitid, browser, operating system, ismobile, and DeviceCategory. The 'Output' pane at the bottom shows the execution details, including the time taken (23:30:03) and the number of rows returned (1000).

sessionid	fullvisitorid	visitid	browser	operatingsystem	ismobile	DeviceCategory
6.16787E+28	6167871330617112263	1508151024	Chrome	Macintosh	FALSE	desktop
6.43698E+27	0643697640977915618	1508175522	Chrome	Windows	FALSE	desktop
6.05938E+28	6059383810968229466	1508143220	Chrome	Macintosh	FALSE	desktop
2.37672E+28	2376720078963423631	1508193530	Safari	iOS	TRUE	mobile
2.31454E+28	2314544030795440038	1508211942	Safari	Macintosh	FALSE	desktop
4.13304E+28	4133039884103392367	1508186358	Chrome	Linux	FALSE	desktop
4.32048E+28	4320478850207397557	1508203650	Chrome	Macintosh	FALSE	desktop
5.87644E+28	5876438247980157131	1508184397	Chrome	Windows	FALSE	desktop
5.14591E+27	0514591268737702944	1508189652	Chrome	Macintosh	FALSE	desktop
6.43057E+28	6430567031531677212	1508175502	Chrome	Windows	FALSE	desktop
7.02637E+28	7026374070157240653	1508190324	Chrome	Windows	FALSE	desktop
2.86172E+28	2861724304134353779	1508196731	Chrome	Android	TRUE	mobile
7.90823E+28	7908247117289630366	1508186327	Chrome	Linux	FALSE	desktop
4.45213E+28	4452127952351664046	1508184708	Chrome	Windows	FALSE	desktop

Location

Data in the Database

Table Name	Primary Key	Foreign Key	# of Rows in Table
Session	SessionID		804684
Traffic	SessionID		804684
Visits	SessionID		804684
Location	SessionID		804684
DeviceSource	SessionID		804684