Q: What To Do with NULLs?
A: NULL
(Understanding the dataset)

Etienne Toussaint
University of Edinburgh
etienne.toussaint@ed.ac.uk

Paolo Guagliardo
University of Edinburgh
paolo.guagliardo@ed.ac.uk

Leonid Libkin
University of Edinburgh & ENS Paris
l@libk.in

Juan Sequeda
data.world
juan@data.world

Most of the data we are gathering through the questionnaire is going to be made public for other investigation. The only part which will remain private until further are the custom answers each of the respondent may share. To be more specific, the public data set would have the knowledge that such custom answer exists, but the content of it would not appear yet. The reason is that we want to share the data set as soon as possible, and we do not have an automatic way to ensure that the custom answers do not leak any private information about our respondent. Therefore such data will not be used in our first analysis either, another more exploratory approach may come later and we will make a pseudonym version of the complete data-set at this time.

0.1 Demographic questions

Figure 1: Question QID219

Question. What best describes your domain according to ISIC specification?

Options. The choices corresponds to the Level 1 from ISIC Rev. 4 ie. the International Standard Industrial Classification of All Economic Activities. The answer is stored as the text value among option selected by the respondent.

Figure 2: Question QID224

Question. What best describes your role according to ISIC specification?

Options. The choices corresponds to the Level 4 for networking and informatics roles from ISIC Rev. 4 ie. the International Standard Industrial Classification of All Economic Activities. The answer is stored as the text value among option selected by the respondent. This question is displayed if and only if the respondent did not answer Education on the question QID219.

Figure 3: Question QID225

Question. What best describes your position?

Options. The answer is stored as the text value among option selected by the respondent:

- Assistant / Associate Professor
- Engineer
- Lecturer / Senior Lecturer
- Management / Leadership
- PhD Student
- MSc Student
- Bachelor Student
- Postdoc
- Professor
- Research assistant
- Researcher
- Tenure Track

This question is displayed if and only if the respondent did answer Education on the question QID219.

Figure 4: Question QID171

Question. How often do you manipulate SQL code?

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly

0.2 Data management questions

Figure 5: Question QID58

Question. What type of data do you use? [TICK ALL THAT AP-PLY]

Options. The answer is stored as an array of text value among options selected by the respondent.
- Text documents
- Relational databases
- Graph databases
- Column-store
- CSV files
- Spreadsheets
- XML
- RDF
- JSON
- other.

The other option is bind to a text box so the respondent can give a custom answer.

+last question

Figure 6: Question QID80

Question. What are the relational databases you interact with? [TICK ALL THAT APPLY]

Options. The answer is stored as an array of text value among options selected by the respondent.
- Oracle database
- Microsoft SQL Server
- PostgreSQL
- MySQL
- Snowflake
- DB2
- SQLite
- Teradata
- Vertica
- IBM DB2
- Other

The other option is bind to a text box so the respondent can give a custom answer.

Figure 7: Question QID57

Question. How would you describe the size of data that you interact with?

Options. The answer is stored as the text value among option selected by the respondent.
Very Small (<1Gb) - Small (several Gb) - Medium (tens of Gb) - Large (hundreds of Gb) - Very large (>1Tb)

Figure 8: Question QID64

Question. According to you, what could be the meaning of the NULL placeholder appearing in a database? [TICK ALL THAT AP-PLY]

Options. The answer is stored as an array of text value among options selected by the respondent.
- The value exists and could be anything
- The value does not exist
- The value exists and is equal to a known constant (for example 0 or empty string)
- The value exists and is equal to an unknown constant
- The data is dirty
- There is a bug
- Nothing is known about the value
- It denotes a non-applicable field
- other.

The other option is bind to a text box so the respondent can give a custom answer.

Figure 9: Question QID60

Question. How often do encounter NULL values?

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly

Figure 10: Question QID82

Question. How often do you explicitly specify a column as NOT NULL?

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly

Figure 11: Question QID62

Question. How often do you explicitly add NOT NULL to the following constraints? [PRIMARY KEY (1), FOREIGN KEY (2), UNIQUE (3)].

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly

Figure 12: Question QID62

Question. How often do you use the following SQL operators/functions to deal with NULL values? [IFNULL(), ISNULL(), IS [NOT] NULL, COALESCE(), IS [NOT] DISTINCT, Other].

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly
The other option is bind to a text box so the respondent can give a custom answer.

Figure 13: Question QID120

Question. How often do you use the following JOIN operators on columns with NULLs? [[INNER] JOIN, LEFT | RIGHT [OUTER] JOIN, FULL [OUTER] JOIN, Other].

Options. The answer is stored as the text value among option selected by the respondent.
Never - Infrequently - Occasionally - Often - Regularly
The other option is bind to a text box so the respondent can give a custom answer.
From our perspective the only sensible data we might collect are email address, and customs answers. The questions we are asking are generic enough so it is not possible to extract someone identity from them. In order to ensure anonymity, we will release the 2 different data-set to the public, the first one will be striped of the respondent email and custom answers. While the second one will contain the custom answers if and only if they are not threatening anonymity. The custom answers are going to be check manually beforehand.

0.3 Relational Algebra questions

Most questions in this section follow the same format. Given the following Customers, Orders and OrdersPaidBy tables 14 and a specific query, we ask the respondent to **give[s] the rows [he] would like to be in the answer** with a drag and drop process. For each row in the answers pool there is a corresponding encoding given for each question, and we associate to each encoding an array of numbers. The number in the array corresponds to the multiplicity of the given row in the answer built by the respondent. We store such number each time the respondent request to see the SQL answers and at submission.

Example 0.1. Consider the XML tree below for question QID110.

```
<QID110>
<c1>[0,0,4]</c1>
<c2>[0,2,3]</c2>
<c3>[0,1,0]</c3>
<c4>[0,0,0]</c4>
<c5>[0,1,0]</c5>
<Other>[0,1,2]</Other>
<natLanguage>0</natLanguage>
<sqlasked>2</sqlasked>
</QID110>
<QID180_FIRST_CLICK>0</QID180_FIRST_CLICK>
<QID180_LAST_CLICK>0</QID180_LAST_CLICK>
<QID180_PAGE_SUBMIT>118.72</QID180_PAGE_SUBMIT>
<QID180_CLICK_COUNT>0</QID180_CLICK_COUNT>
```

The tree means that the respondant checked the SQL answer 2 times (sqlasked). When he first check SQL his answer was empty; after checking a second time his answer contained 2 times the row which encoding is c2 and 1 times the rows encoded as c3, c6 and Other. Finally his final answer contains 4 times the row which encoding is c1, 3 times the row encoded as c2 and 2 times the row encoded as Other. Finally the respondent never asked for the natural language interpretation of the query. The question QID132 is an hidden to the respondent and correponds to a Timing question. The time values are stored in seconds.

Figure 14: Caption

CUSTOMERS				ORDERS				ORDERSPAIDBY	
cid	name	alias	oid	cid	price	taxes	oid	cid	
c1	Etienne	Etienne	o1	c1	30	NULL	o1	c3	
c2	Leonid	NULL	o2	c1	NULL	0	o2	c3	
c3	Paolo	Juan	o3	c2	NULL	NULL	o3	NULL	
c4	NULL		o4	c2	NULL	NULL	o4	NULL	
c5	NULL	NULL	o5	NULL	10	0	o5	NULL	
			o6	NULL	20	10	o6	NULL	

Figure 15: Query and encoding of answers for QID110

```sql
SELECT cid , name
  FROM CUSTOMERS
  WHERE name = alias ;
```

ROWS			
cid	name	encoding	SQL
c1	Etienne	c1	1
c2	Leonid	c2	0
c3	Paolo	c3	0
c4	NULL	c4	0
c5	NULL	c5	0
other	other	Other	0

NATURAL LANGUAGE. *The query returns the attributes cid and name of all CUSTOMERS rows for which its attribute name is equal to its attribute alias.*

Figure 16: Query and encoding of answers for QID179

```sql
SELECT cid , name
  FROM CUSTOMERS
  WHERE name <> alias ;
```

ROWS			
cid	name	encoding	SQL
c1	Etienne	c1	0
c2	Leonid	c2	0
c3	Paolo	c3	1
c4	NULL	c4	0
c5	NULL	c5	0
other	other	Other	0

NATURAL LANGUAGE. *The query returns the attributes cid and name of all CUSTOMERS rows for which its attribute name is different from its attribute alias.*

Figure 17: Query and encoding of answers for QID172

```sql
SELECT oid
  FROM ORDERS
  WHERE price = taxes ;
```

ROWS		
oid	encoding	SQL
o1	o1	0
o2	o2	0
o3	o3	0
o4	o4	0
o5	o5	0
o6	o6	0
other	Other	0

NATURAL LANGUAGE. *The query returns the attribute oid of all ORDERS rows for which its attribute price is equal to its attribute taxes.*

Figure 18: Query and encoding of answers for QID175

```
SELECT oid
  FROM ORDERS
 WHERE price  <>  taxes ;
```

ROWS

oid		encoding	SQL
o1		o1	0
o2		o2	0
o3		o3	0
o4		o4	0
o5		o5	1
o6		o6	1
other		Other	0

NATURAL LANGUAGE. *The query returns the attribute oid of all OR-DERS rows for which its attribute price is different from its attribute taxes.*

Figure 19: Query and encoding of answers for QID185;QID186;QID187

```
(1) SELECT DISTINCT c.cid
    FROM CUSTOMERS AS c , ORDERS AS o
    WHERE c.cid = o.cid;

(2) SELECT c.cid
    FROM CUSTOMERS AS c
    WHERE c.cid IN
        (SELECT o.cid
        FROM ORDERS AS o );

(3) SELECT c.cid
    FROM CUSTOMERS AS c
    WHERE EXISTS
        (SELECT o.cid
        FROM ORDERS AS o
        WHERE c.cid=o.cid);
```

ROWS

cid		encoding	SQL
c1		c1	1
c2		c2	1
c3		c3	0
c4		c4	0
c5		c5	0
NULL		NULL_Value	0
other		Other	0

NATURAL LANGUAGE. *The queries return the attribute cid of all CUS-TOMERS rows for which the attribute cid also appears as the attribute cid of an ORDERS row.*

Explanation. This question is a special case aimed to reduce the cognitive load of the survey. In SQL the 3 queries presented above compute the them result, and we present this answer to the respondent. Then the question QID181 ask him whether if the respondent is satisfied with the SQL result for each of the query. If the respondent is satisfied (encoded as 1), we do not ask for further detail. If he is not (encoded as 2), an additional question is prompted asking the the respondent to build the answer if wanted. The matching between question and query is the following: [QID185;(1)] [QID186;(2)] and [QID187;(3)].

Figure 20: Query and encoding of answers for QID20

```sql
SELECT DISTINCT c.cid
FROM CUSTOMERS AS c
WHERE c.cid NOT IN
    (SELECT o.cid
    FROM ORDERS AS o );
```

cid	ROWS encoding	SQL
c1	c1	0
c2	c2	0
c3	c3	0
c4	c4	0
c5	c5	0
NULL	NULL_Value	0
other	Other	0

NATURAL LANGUAGE. *The query returns the attribute cid of all CUS-TOMERS rows for which the attribute cid does not appear as the attribute cid of an ORDERS row.*

Figure 21: Query and encoding of answers for QID19

```sql
SELECT c.cid
FROM CUSTOMERS AS c
WHERE NOT EXISTS
    (SELECT o.cid
    FROM ORDERS AS o
    WHERE c.cid=o.cid);
```

cid	ROWS encoding	SQL
c1	c1	0
c2	c2	0
c3	c3	1
c4	c4	1
c5	c5	1
NULL	NULL_Value	0
other	Other	0

NATURAL LANGUAGE. *The query returns the attribute cid of all CUS-TOMERS rows for which the attribute cid does not appear as the attribute cid of an ORDERS row.*

Figure 22: Query and encoding of answers for QID21, QID226

```sql
(1) (SELECT opb.cid
    FROM ORDERSPAIDBY AS opb)
        MINUS
    (SELECT o.cid
    FROM ORDERS AS o);


(2) (SELECT opb.cid
    FROM ORDERSPAIDBY AS opb)
        EXCEPT
    (SELECT o.cid
    FROM ORDERS AS o);
```

cid	ROWS encoding	SQL
c1	c1	0
c2	c2	0
c3	c3	1
c4	c4	0
c5	c5	0
NULL	NULL_Value	0
other	Other	0

NATURAL LANGUAGE. *The query returns the attribute cid of all CUS-TOMERS rows for which the attribute cid does not appear as the attribute cid of an ORDERS row.*

Explanation. Due to the difference between PSQL and ORACLE languages, those 2 questions are under a mutually exclusive display logic. The first question (1) is displayed if and only if the respondent has claimed familiarity with either Oracle or MySQL DBMS and not with PSQL in question ??. Otherwise question 2 is given. The matching between question and query is the following: [QID21;(1)] and [QID226;(2)].

Figure 23: Query and encoding of answers for QID86

```
(SELECT opb.cid
   FROM ORDERSPAIDBY AS opb)
      EXCEPT ALL
   (SELECT o.cid
   FROM ORDERS AS o);
```

ROWS

cid		encoding	SQL
c1		c1	0
c2		c2	0
c3		c3	2
c4		c4	0
c5		c5	0
NULL		NULL_Value	2
other		Other	0

NATURAL LANGUAGE. *The query returns some cid attribute with a multiplicity equal to the difference between the number of rows in ORDERSPAIDBY and the number of rows in ORDERS with this cid attribute*

Explanation. Please note that this question is displayed to every respondent despite the fact that it is not valid on many DBMS. Moreover immediate equivalent does not exist on MySQL or Oracle.

Figure 24: Query, database and encoding of answers for QID88

```
SELECT o.oid ,
     o.price + o.taxes AS result
FROM ORDERS AS o
```

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**

NATURAL LANGUAGE. *The query returns the attribute cid and the addition of the attributes price and taxes of all ORDERS rows.*

ANSWERS

Options		encoding	SQL
oid | result o1 | 30		QID88~1~track	FALSE
oid | result o1 | 0		QID88~6~track	FALSE
oid | result o1 | **NULL**		QID88~3~track	TRUE
Other		QID88~5~track	FALSE

Figure 25: Query, database and encoding of answers for QID93

```
SELECT o.cid ,
     SUM(price) AS result
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**
o2	c1	**NULL**	0

NATURAL LANGUAGE. *The query returns the attribute cid and the sum of the attribute price of all ORDERS rows which share the same attribute cid.*

ANSWERS

Options		encoding	SQL
cid | result c1 | 30		QID93~1~track	TRUE
cid | result c1 | 0		QID93~5~track	FALSE
cid | result c1 | **NULL**		QID93~3~track	FALSE
Other		QID93~6~track	FALSE

Figure 26: Query, database and encoding of answers for QID119

```
SELECT o.cid,
     MIN(price) AS minimum
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**
o2	c1	**NULL**	0

NATURAL LANGUAGE. *The query returns the attribute cid and the minimum of the attribute price among all ORDERS rows which share the same attribute cid.*

ANSWERS

Options		encoding	SQL
cid | minimum c1 | 30		QID119~1~track	TRUE
cid | minimum c1 | 0		QID119~5~track	FALSE
cid | minimum c1 | **NULL**		QID119~3~track	FALSE
Other		QID119~6~track	FALSE

0.4 Aggregation and Arithmetic questions

Most questions in this section follow the same format. Given and a specific database and query, we propose some table each corresponding to a potential answers. We ask the respondent to **specify how much [he] like[s] each of this table as an answer.** For each table in the answers pool there is a corresponding encoding given for each question, and we associate to each encoding an array of numbers. The number in the array corresponds to the rating from 0 to 5 the respondent give to each possibility. We store such number each time the respondent request to see the SQL answers and at submission. Not that we offer the respondent the possibility to give a custom answer, by default this answer is given a 5 stars rating in order to incentive the respondent to give a better alternative to the one already offered, in the public data-set those 5 are becoming 0 if the user didn't give a custom answers, and keep their actual value otherwise.

Example 0.2. Consider the XML tree below for question QID110.

```
<QID88>
<QID88~1~track>[0,0,4]</QID88~1~track>
```

Figure 29: Query, database and encoding of answers for QID105

```
SELECT o.cid,
    MIN(price) AS minimum
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o3	c2	**NULL**	**NULL**
o4	c2	**NULL**	**NULL**

NATURAL LANGUAGE. *The query returns the attribute cid and the minimum of the attribute price among all ORDERS rows which share the same attribute cid.*

ANSWERS

Options			encoding	SQL
cid c2 / minimum 0			QID105~1~track	FALSE
cid c2 / minimum **NULL**			QID105~3~track	TRUE
Other			QID105~5~track	FALSE

Figure 30: Query, database and encoding of answers for QID98

```
SELECT o.cid ,
    SUM(price) AS result
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o5	**NULL**	10	0
o6	**NULL**	20	10

NATURAL LANGUAGE. *The query returns the attribute cid and the sum of the attribute price of all ORDERS rows which share the same attribute cid.*

ANSWERS

Options			encoding	SQL
cid **NULL** / result 0			QID98~6~track	FALSE
cid **NULL** / result 30			QID98~1~track	TRUE
cid **NULL** 10, **NULL** 20 / result			QID98~3~track	FALSE
Other			QID98~7~track	FALSE

Figure 31: Query, database and encoding of answers for QID106

```
SELECT o.cid,
    MIN(price) AS minimum
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o5	**NULL**	10	0
o6	**NULL**	20	10

NATURAL LANGUAGE. *The query returns the attribute cid and the minimum of the attribute price among all ORDERS rows which share the same attribute cid.*

ANSWERS

Options			encoding	SQL
cid **NULL** / minimum 10			QID106~1~track	TRUE
cid **NULL** / minimum 0			QID106~6~track	FALSE
cid **NULL** 10, **NULL** 20 / minimum			QID106~3~track	FALSE
Other			QID106~7~track	FALSE

Figure 27: Query, database and encoding of answers for QID96

```
SELECT o.oid ,
    o.price + o.taxes AS result
FROM ORDERS AS o
```

ORDERS

oid	cid	price	taxes
o3	c2	**NULL**	**NULL**

NATURAL LANGUAGE. *The query returns the attribute cid and the addition of the attributes price and taxes of all ORDERS rows.*

ANSWERS

Options			encoding	SQL
oid o3 / result 0			QID96~1~track	FALSE
oid o3 / result **NULL**			QID96~3~track	TRUE
Other			QID96~5~track	FALSE

9

Figure 33: Query, database and encoding of answers for QID118

```
SELECT SUM(o.price) AS result
FROM CUSTOMERS AS c
    LEFT JOIN ORDERS
    AS o ON c.cid = o.cid
```

CUSTOMERS

cid	name	alias
c1	Etienne	Etienne
c3	Paolo	Juan

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**

NATURAL LANGUAGE. *The query returns the sum of the attributes price of all rows in the following database:*

cid	name	alias	oid	price	taxes
c1	*Etienne*	*Etienne*	*o1*	*30*	***NULL***
c3	*Paolo*	*Juan*	***NULL***	***NULL***	***NULL***

ANSWERS

Options		encoding	SQL
result 0		QID118~1~track	FALSE
result 30		QID118~3~track	TRUE
result **NULL**		QID118~9~track	FALSE
Other		QID118~10~track	FALSE

Figure 34: Query, database and encoding of answers for QID113

```
SELECT MIN(o.price) AS minimum
FROM CUSTOMERS AS c
    LEFT JOIN ORDERS
    AS o ON c.cid = o.cid
```

CUSTOMERS

cid	name	alias
c1	Etienne	Etienne
c3	Paolo	Juan

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**

NATURAL LANGUAGE. *The query returns the minimum of the attribute price among all rows in the following database:*

cid	name	alias	oid	price	taxes
c1	*Etienne*	*Etienne*	*o1*	*30*	***NULL***
c3	*Paolo*	*Juan*	***NULL***	***NULL***	***NULL***

ANSWERS

Options		encoding	SQL
minimum 0		QID113~1~track	FALSE
minimum 30		QID113~3~track	TRUE
minimum			

Figure 35: Query, database and encoding of answers for QID200

```
SELECT o.cid,
    SUM(price) AS result
FROM CUSTOMERS AS c
    LEFT JOIN ORDERS
    AS o ON c.cid = o.cid
GROUP BY o.cid
```

CUSTOMERS

cid	name	alias
c1	Etienne	Etienne
c3	Paolo	Juan

ORDERS

oid	cid	price	taxes

NATURAL LANGUAGE. *The query returns the attribute cid and the sum of the attribute price of all rows in the following database which share the same attribute cid.*

cid	name	alias	oid	price	taxes
c1	*Etienne*	*Etienne*	***NULL***	***NULL***	***NULL***
c3	*Paolo*	*Juan*	***NULL***	***NULL***	***NULL***

ANSWERS

Options			encoding	SQL
cid **NULL**	result 0		QID200~6~track	FALSE
cid **NULL**	result **NULL**		QID200~1~track	TRUE
cid **NULL** **NULL**	result 0 0		QID200~3~track	FALSE
cid **NULL** **NULL**	result **NULL** **NULL**		QID200~9~track	FALSE
Other			QID200~7~track	FALSE

Figure 32: Query, database and encoding of answers for QID109

```
SELECT c.cid,
    o.price + o.taxes AS result
FROM CUSTOMERS AS c
    LEFT JOIN ORDERS AS o
    ON c.cid = o.cid;
```

CUSTOMERS

cid	name	alias
c3	Paolo	Juan

ORDERS

oid	cid	price	taxes
o1	c1	30	**NULL**

NATURAL LANGUAGE. *The query returns the attribute cid and the addition of the attributes price and taxes of all rows in the following database:*

cid	name	alias	oid	price	taxes
c3	*Paolo*	*Juan*	***NULL***	***NULL***	***NULL***

ANSWERS

Options		encoding	SQL

10

Figure 36: Query, database and encoding of answers for QID201

```
SELECT o.cid,
    MIN(price) AS minimum
FROM CUSTOMERS AS c
    LEFT JOIN ORDERS
    AS o ON c.cid = o.cid
GROUP BY o.cid
```

CUSTOMERS

cid	name	alias
c1	Etienne	Etienne
c3	Paolo	Juan

ORDERS

oid	cid	price	taxes

NATURAL LANGUAGE. *The query returns the attribute cid and the minimum of the attribute price among all rows in the following database which share the same attribute cid.*

cid	name	alias	oid	price	taxes
c1	*Etienne*	*Etienne*	*NULL*	*NULL*	*NULL*
c3	*Paolo*	*Juan*	*NULL*	*NULL*	*NULL*

ANSWERS

Options		encoding	SQL
cid / minimum **NULL** / 0		QID201~6~track	FALSE
cid / minimum **NULL** / **NULL**		QID201~1~track	TRUE
cid / minimum **NULL** / 0 **NULL** / 0		QID201~3~track	FALSE
cid / minimum **NULL** / **NULL** **NULL** / **NULL**		QID201~9~track	FALSE
Other		QID201~7~track	FALSE

Figure 28: Query, database and encoding of answers for QID97

```
SELECT o.cid ,
    SUM(price) AS result
FROM ORDERS AS o
GROUP BY o.cid
```

ORDERS

oid	cid	price	taxes
o3	c2	**NULL**	**NULL**
o4	c2	**NULL**	**NULL**

NATURAL LANGUAGE. *The query returns the attribute cid and the sum of the attribute price of all ORDERS rows which share the same attribute cid.*

ANSWERS

Options		encoding	SQL
cid / result c2 / 0		QID97~1~track	FALSE
cid / result c2 / **NULL**		QID97~3~track	TRUE
Other		QID97~5~track	FALSE

```
<QID88~2~track>[0,2,3]</QID88~2~track>
<QID88~3~track>[0,1,0]</QID88~3~track>
<QID88~4~track>[0,0,0]</QID88~4~track>
<natLanguage>0</natLanguage>
<sqlasked>2</sqlasked>
</QID88>
<QID146_FIRST_CLICK>0</QID146_FIRST_CLICK>
<QID146_LAST_CLICK>0</QID146_LAST_CLICK>
<QID146_PAGE_SUBMIT>118.72</QID146_PAGE_SUBMIT>
<QID146_CLICK_COUNT>0</QID146_CLICK_COUNT>
```

The tree means that the respondent checked the SQL answer 2 times (sqlasked). When he first check SQL the star ranking of each table was 0. After checking a second time he gave 2 stars amung 5 available to the potential answer encoded as QID88 2 track and 1 star to the one encoded as QID88 3 track. Finally his final answer gives 4 stars to the potential answer encoded as QID88 1 track and 3 stars to the one encoded as QID88 2 track. Finally the respondent never asked for the natural language interpretation of the query. The question QID146 is hidden to the respondent and corresponds to a Timing question. The time values are stored in seconds.