

#	Table	Field	Check Type	Check status	Result value	Expected value	Description	SQL
1	CALLS	Sales_Person_ID	FK consistency	OK	0	0	Are there any invalid Sales_Person_ID values? (not in the ID list (PK) in the table SALES_PEOPLE or null)	SELECT * FROM calls LEFT join sales_people ON calls.Sales_Person_ID=sales_people.Sales_Person_ID WHERE sales_people.Sales_Person_ID IS NULL;
2	CALLS	Customer_ID	FK consistency	OK	0	0	Are there any invalid Customer_ID values? (not in the ID list (PK) in the table CUSTOMER or null)	SELECT * FROM calls LEFT join customer ON calls.Customer_ID=customer.Customer_ID WHERE customer.Customer_ID IS NULL;
3	CALLS	Call_Date_UTC	Historization check	OK	01-08-21 0:10	-	How much history entity stores?	SELECT min(calls.Call_Date_UTC) FROM calls;
4	CALLS	Call_Duration_Sec	Analysis Range	OK	2.2773	<10%	What is the percentage of dropped calls (time <5 sec)	SELECT (SELECT count(Call_Duration_Sec)*100 from calls WHERE Call_Duration_Sec <5)/count(*) from calls;
5	SALES_PEOPLE	Sales_Person_ID	PK unique	OK	0	0	Are there any duplicates of Sales_Person_ID in SALES_PEOPLE table?	SELECT Sales_Person_ID, count(*) FROM sales_people GROUP BY Sales_Person_ID HAVING count(*)>1;
6	SALES_PEOPLE	Sales_Person_Name	Pattern	ERROR	4	0	Do all the names make sense?	SELECT Sales_Person_Name from sales_people WHERE Sales_Person_Name LIKE '%0%' OR Sales_Person_Name LIKE '%1%' OR Sales_Person_Name LIKE '%2%' OR Sales_Person_Name LIKE '%\%%' OR Sales_Person_Name LIKE '%\&%' OR Sales_Person_Name LIKE '%@%';

7	SALES_PERSON	Sales_Person_Country	Pattern	ERROR	<table><tr><th>Sales_Person_Country</th><th>count</th></tr><tr><td>France</td><td>44</td></tr><tr><td>Germany</td><td>4</td></tr><tr><td>Germany</td><td>21</td></tr><tr><td>UK</td><td>30</td></tr><tr><td>United States</td><td>49</td></tr><tr><td>United States</td><td>1</td></tr><tr><td>US</td><td>1</td></tr></table>	Sales_Person_Country	count	France	44	Germany	4	Germany	21	UK	30	United States	49	United States	1	US	1	according to policies	Which countries are represented in the table? Are there any inconsistencies in country names?	SELECT Sales_Person_Country, count(Sales_Person_Country) from sales_people Group by (Sales_Person_Country);			
Sales_Person_Country	count																										
France	44																										
Germany	4																										
Germany	21																										
UK	30																										
United States	49																										
United States	1																										
US	1																										
8	SALES_PERSON	Sales_Person_City	Completeness	ERROR	38	0	How many sales people do not have a city assigned to them?	SELECT count(Sales_Person_City) from sales_people WHERE Sales_Person_City="" OR Sales_Person_City IS NULL;																			
9	SALES_PERSON	Sales_Person_Address	Analysis Range/Pattern	OK	70.70%	>50	What is the percentage of valid address? (not empty, starting with numbers)	SELECT (SELECT count(Sales_Person_Address)*100 from sales_people where Sales_Person_Address REGEXP '[0-9][0-9][0-9]')/count(*) from sales_people;																			
10	CUSTOMER	Customer_ID	PK unique	OK	1229	1229	Are there any duplicates among customer ID in the customer table?	Select count(*), count(distinct `Customer_ID`) from Customer;																			
11	CUSTOMER	Customer_Category	Conformity Analysis	ERROR	<table><tr><th>Customer_Category</th><th>count</th></tr><tr><td>Customer_Category</td><td>194</td></tr><tr><td>Customer_Category</td><td>1</td></tr><tr><td>Large</td><td>334</td></tr><tr><td>M</td><td>6</td></tr><tr><td>Medium</td><td>256</td></tr><tr><td>Small</td><td>407</td></tr><tr><td>SX</td><td>1</td></tr><tr><td>XS</td><td>30</td></tr></table>	Customer_Category	count	Customer_Category	194	Customer_Category	1	Large	334	M	6	Medium	256	Small	407	SX	1	XS	30	categories are ambiguous; exist empty categories	categories are not ambiguous; no empty categories	Can we split all customers by customer category?	SELECT `Customer_Category`, count(`Customer_Category`) FROM `customer` GROUP BY `Customer_Category`;
Customer_Category	count																										
Customer_Category	194																										
Customer_Category	1																										
Large	334																										
M	6																										
Medium	256																										
Small	407																										
SX	1																										
XS	30																										
12	CUSTOMER	Customer_Name	Completeness	ERROR	1221	1229	Is the customer name provided?	SELECT count(*) AS all_clients, (SELECT count(customer_name) from customer WHERE NOT customer_name="" OR																			

								customer_name IS NULL) AS have_names from customer;												
13	CUSTOMER	Customer_Phone	Pattern/Completeness	ERROR	874	1229	Do all the phone numbers correspond to mask "ddd-ddd-dddd"?	SELECT count(`Customer_Phone`) FROM `customer` WHERE `Customer_Phone` REGEXP '^...-...-....\$';												
14	CUSTOMER	Customer_Address	Conformity	Ok	no	no	Are there cases when several phone numbers are recorded to one address?	SELECT `Customer_Address`, count(Customer_Phone) FROM `customer` WHERE `Customer_Address` REGEXP '[0-9][a-z][A-Z]' GROUP BY `Customer_Address` HAVING count(Customer_Phone)>1;												
15	CUSTOMER	Customer_City	Pattern	ERROR	<table><thead><tr><th>Customer_City</th><th>count</th></tr></thead><tbody><tr><td>Customer_City</td><td>1</td></tr><tr><td>NUUUULL</td><td>1</td></tr><tr><td>CITRUS HEIGHTS</td><td>9</td></tr><tr><td>WEAVERVILLE</td><td>10</td></tr><tr><td>MILL VALLEY</td><td>13</td></tr></tbody></table>	Customer_City	count	Customer_City	1	NUUUULL	1	CITRUS HEIGHTS	9	WEAVERVILLE	10	MILL VALLEY	13	no	Are there unexpected city names? It is difficult to say for sure that some city does not exist. But grouping and sorting in ascending order helps identify those of minimum number - for further investigation with business (or open sources)	SELECT `Customer_City`, count(`Customer_City`) FROM `customer` WHERE `Customer_City` REGEXP '[a-z]+' GROUP BY `Customer_City` ORDER BY count(`Customer_City`) ASC;
Customer_City	count																			
Customer_City	1																			
NUUUULL	1																			
CITRUS HEIGHTS	9																			
WEAVERVILLE	10																			
MILL VALLEY	13																			