Universitat Oberta de Catalunya
uoc.edu

Name: Estela Martinez Moran

OLAP Design with Tableau Desktop to perform the OLAP schema of the DWH database using Tableau Desktop.

Explanation and justification about the steps performed and it is included the final file dwh-olap.twb

For the creation of an OLAP structure, what will be done is to map of design's database (table of facts and dimensions) with our design. So it is possible to create a schema with fewer elements, the same elements or more elements than those existing in the database.

1) Then we introduce the clue to connect to MySQL ;

- ● Port: 3306
  ● Username: root
  ● Password: xxxx

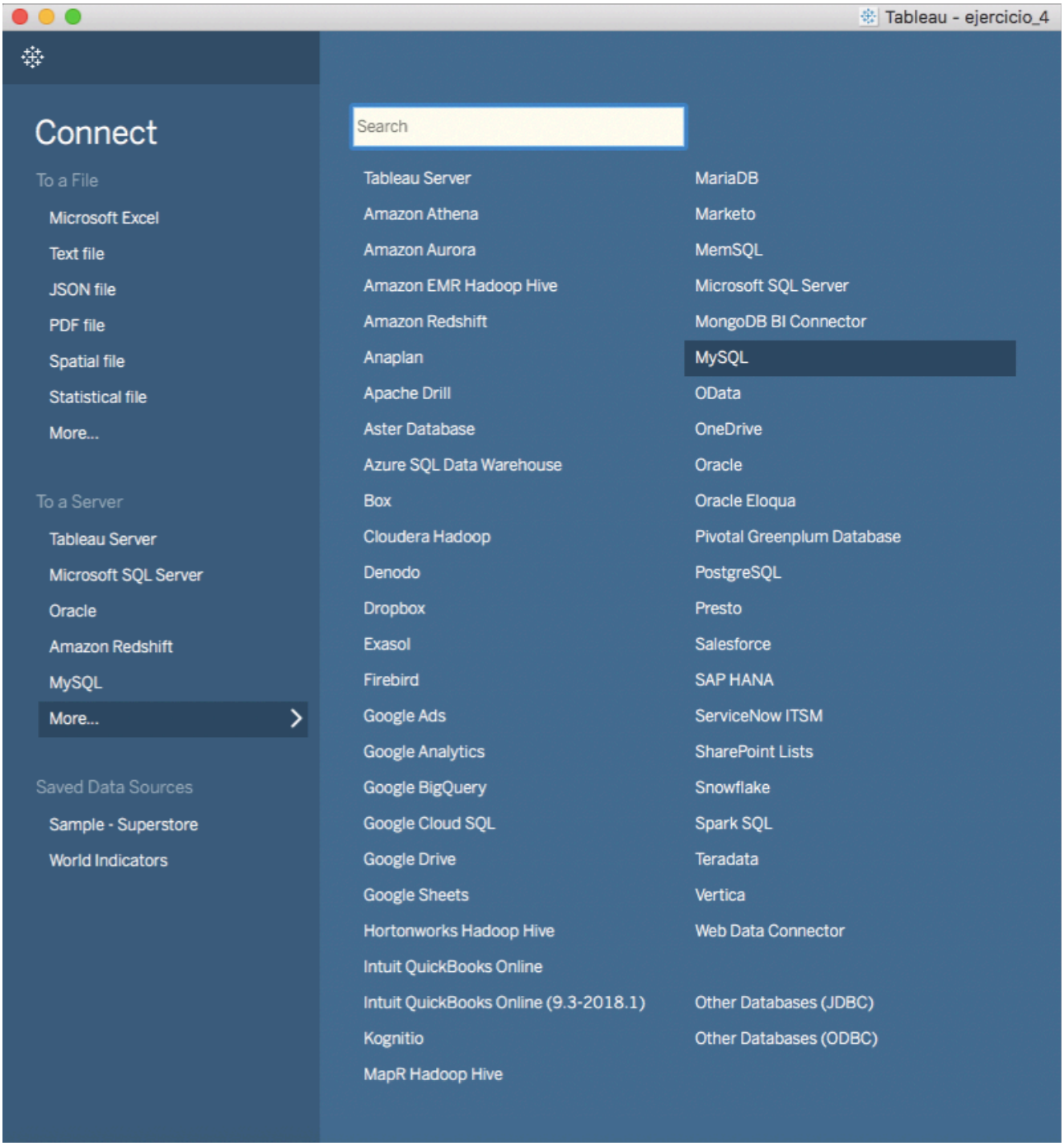
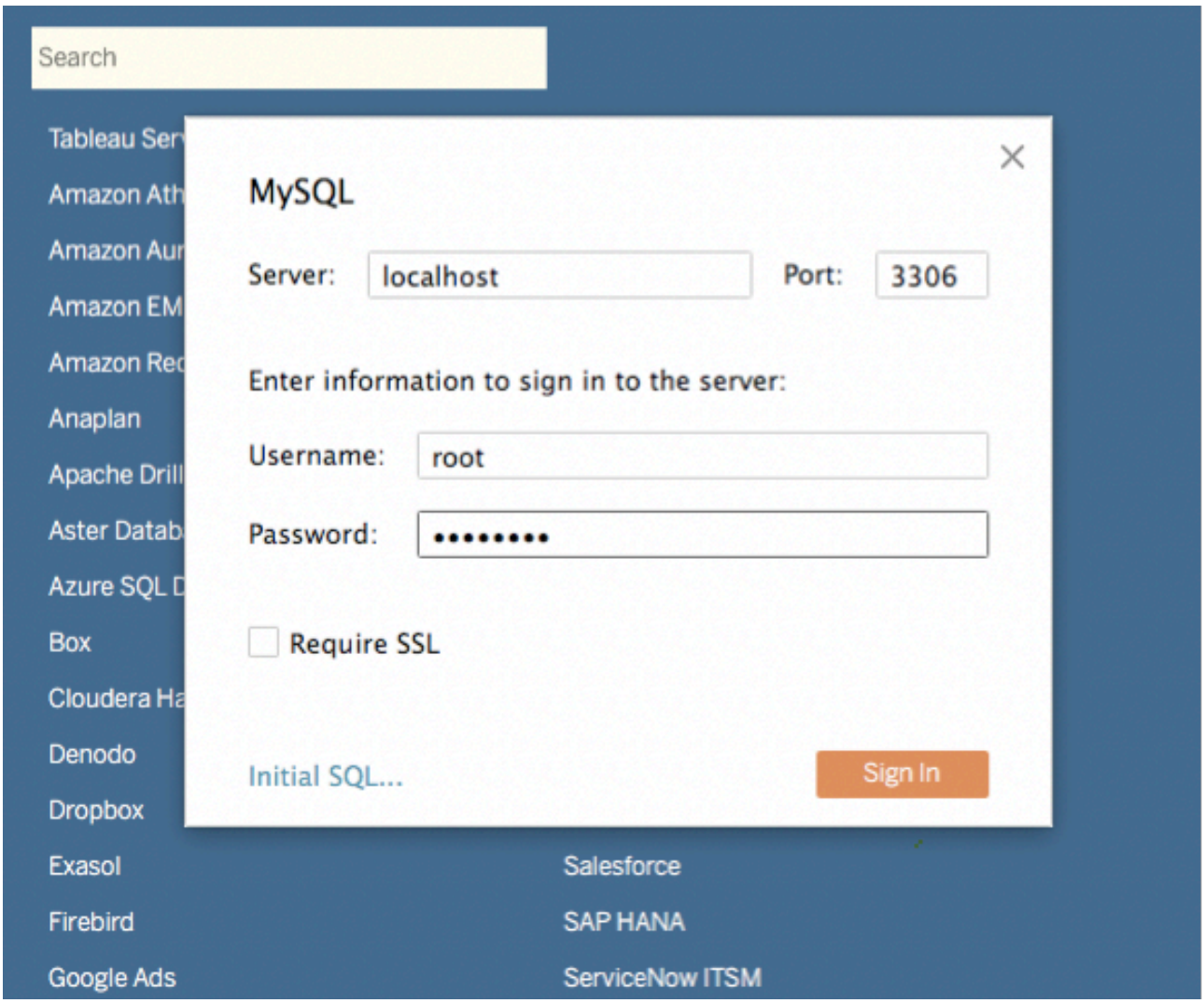


Figura 1. Selection of MySQL as data base in Tableau



2) Figura 2. Configuration of acces to MySQL

Next we drag all the dimension tables to the same canvas. Because the fields that refer to the same identifier are called the same in the dimension table and in the fact table, Tableau recognizes them and suggests the join to perform.

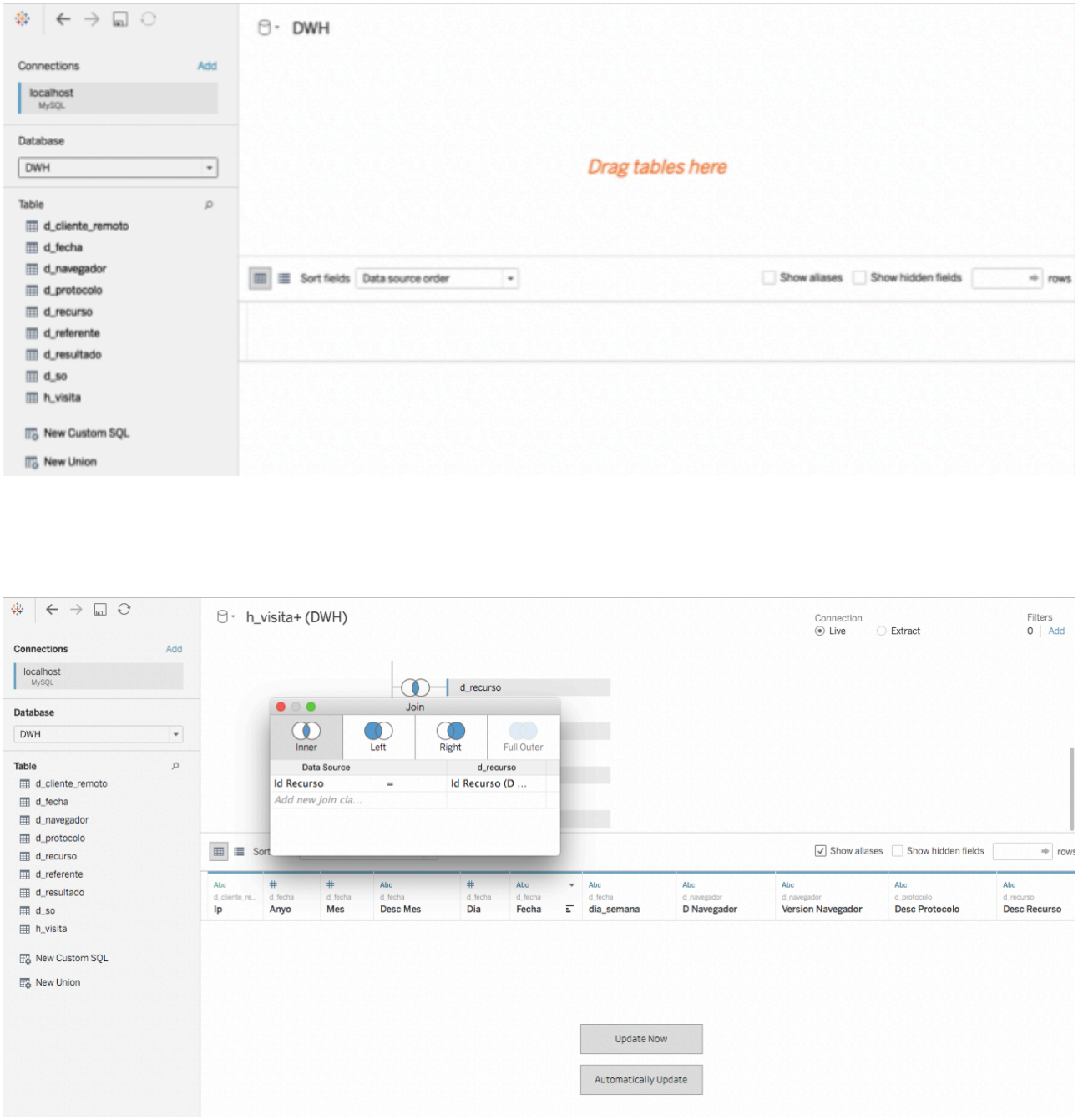


Figura 4. Union of h_visita and d_recurso

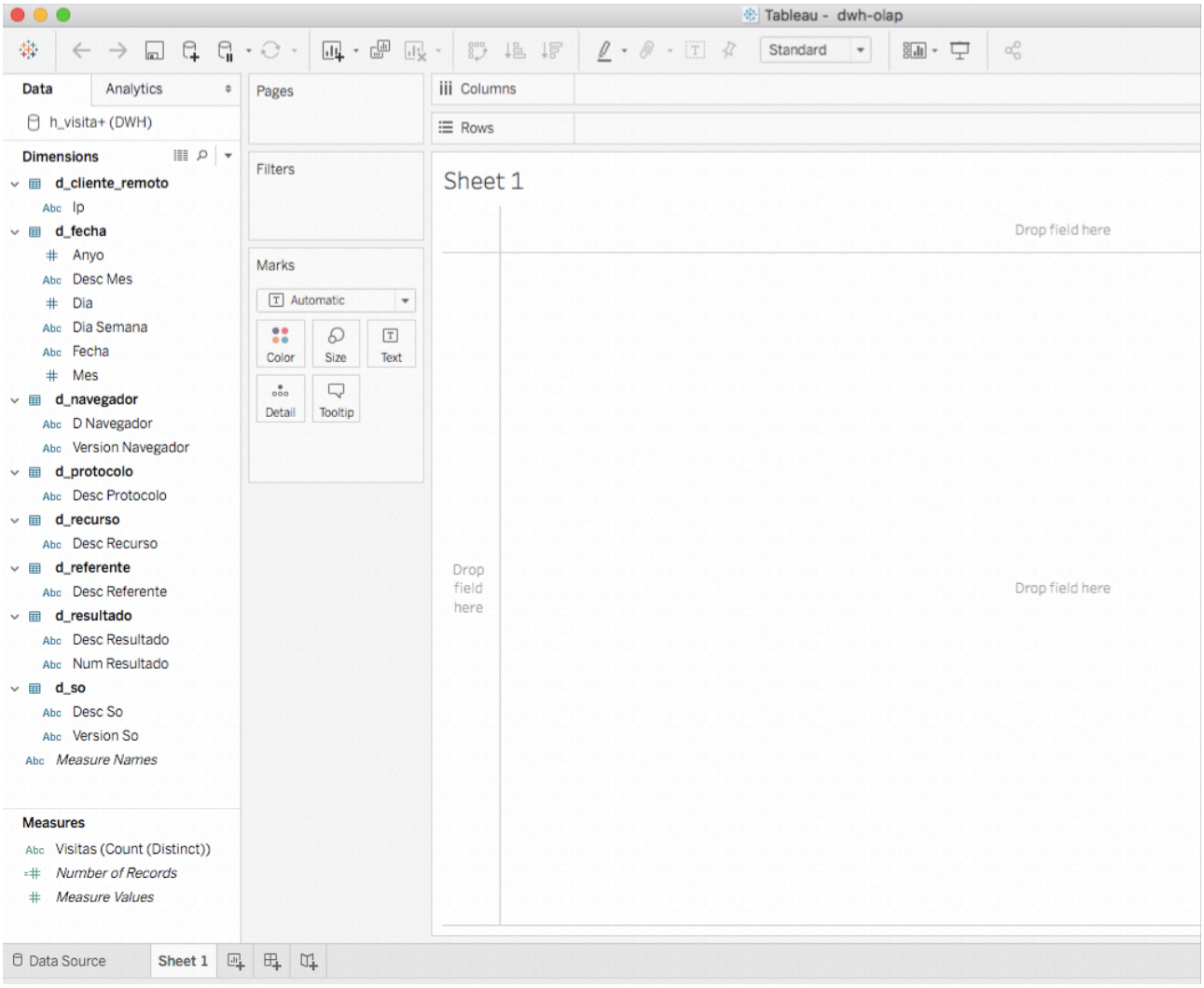


Figura 5.Dimension and metrics of DWH

The steps to follow to obtain the previous figure are;

1**) Sort the metrics section.** In our DWH we have a single metric that is the number of visits and that is calculated from the id_visita field of the h_visita fact table.

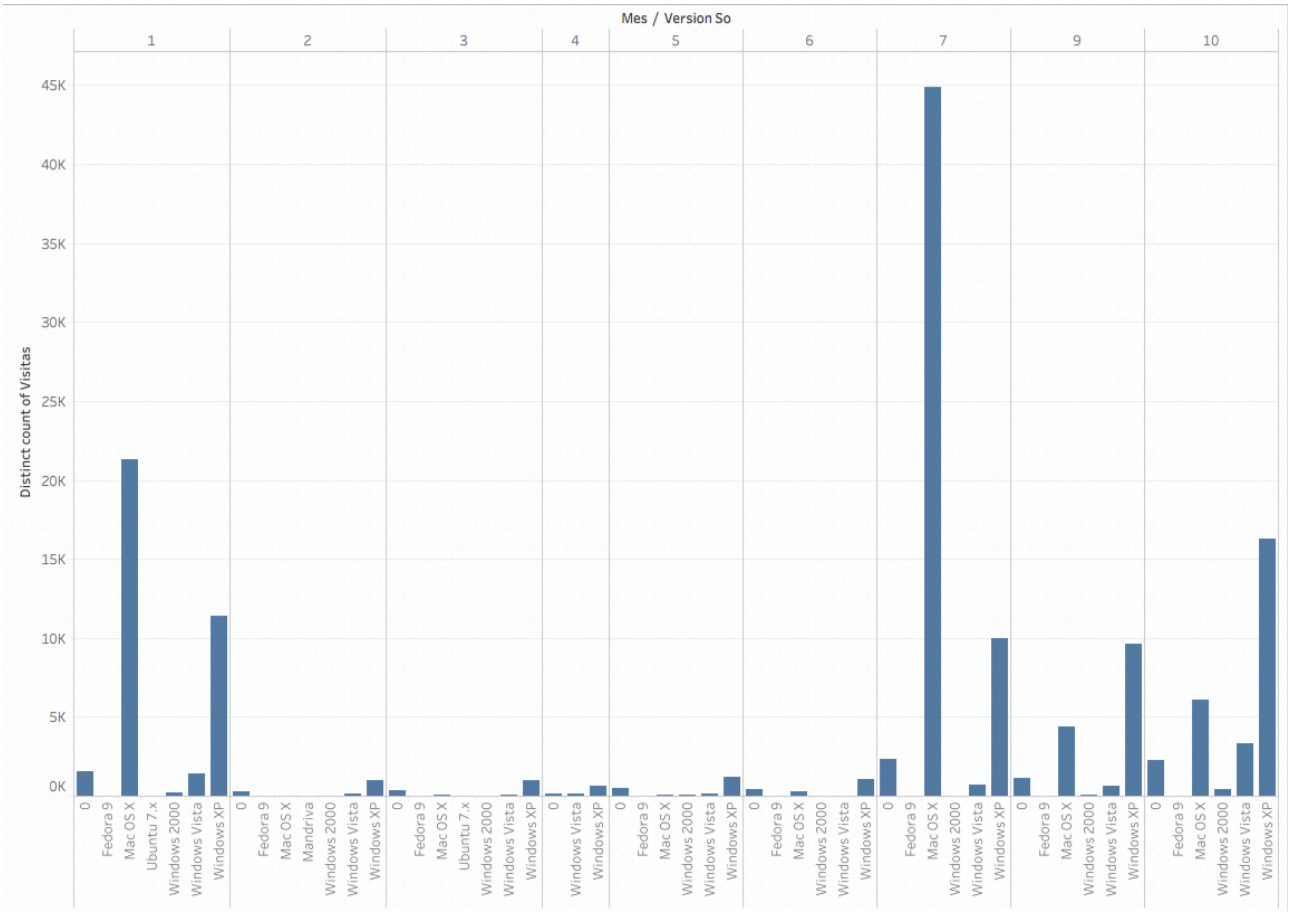
2) Within the list we can see that this field is in the list of dimensions. We drag it to the metrics section. We renamed it to “Visits” and we note that Tableau is going to add the suffix “(Count (Distinct))” to highlight that it is a measure that will use the count distinct aggregation function to be calculated.

3) The year, month and day fields are not considered metrics, in our case we are going to do the reverse procedure and move them to the dimensions section.

4) In this file **dwh-olap.twb** we can see the final version.

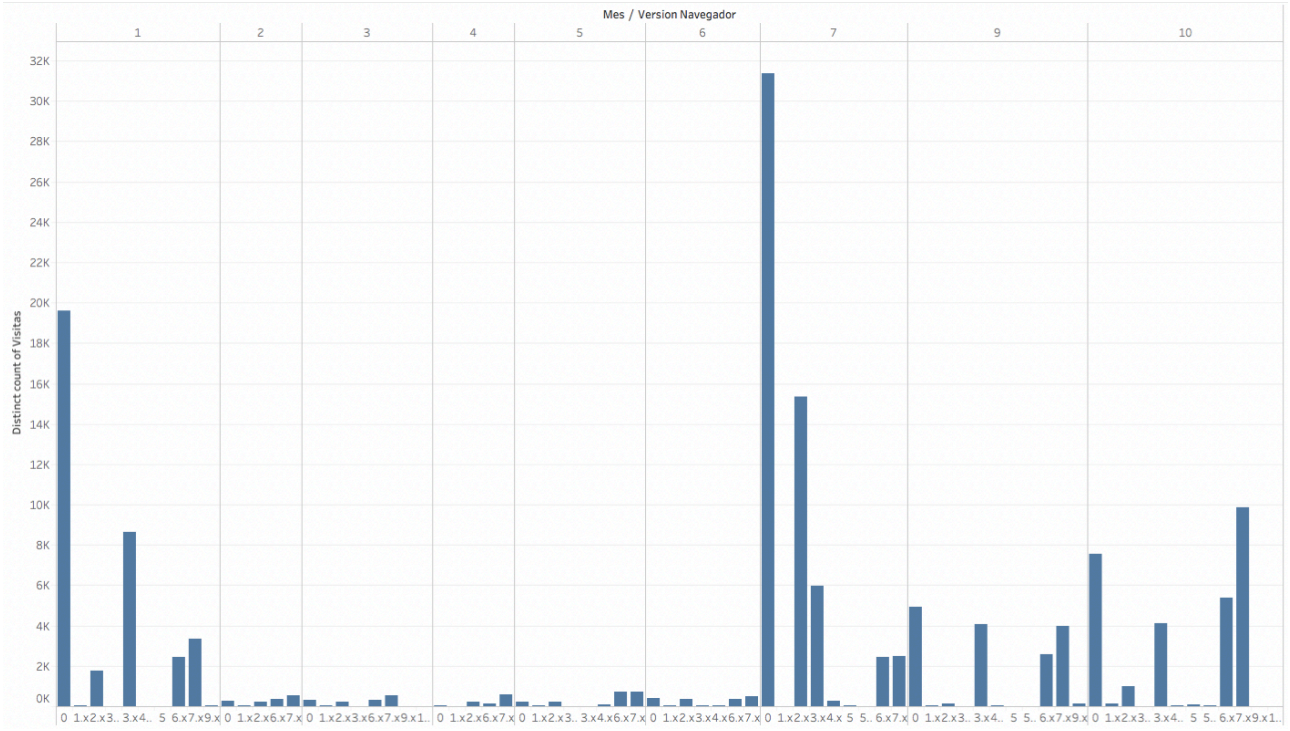
Business questions

¿A Which is the distribution of visits according to Operative System and month of the year?



	Desc Mes								
Version So	January	February	March	April	May	June	July	Septem..	October
0	1.550	318	366	154	490	424	2.368	1.158	2.270
Fedora 9	26	22	22		22	22	10	26	28
Mac OS X	21.340	16	60		40	256	44.908	4.422	6.068
Mandriva		8							
Ubuntu 7.x	6		12						
Windows 2000	222	10	14		106	4	28	98	404
Windows Vista	1.404	110	48	162	130	24	682	610	3.296
Windows XP	11.388	964	956	666	1.214	1.060	10.012	9.628	16.266

¿And according to navigate and month of the year?



Y la tabla pivotante;

	Version Navegador												
Desc Mes	0	1.x	2.x	3.01	3.x	4.01	4.x	5	5.5	6.x	7.x	9.x	11.x
January	19.582	34	1.788	12	8.658	4		8		2.444	3.358	48	
February	266	34	206							378	564		
March	324	32	244		8					298	554	14	4
April	54	8	206							144	570		
May	218	30	206	2	14		78			744	710		
June	416	48	366		52		24			382	502		
July	31.386	18	15.358		5.982		294	30	2	2.458	2.480		
September	4.918	26	154	10	4.076	24		18	10	2.574	4.006	126	
October	7.572	154	1.010	18	4.102	44		112	24	5.400	9.872	22	2

Pregunta 3 (25 puntos). ¿Cuáles son las tres semanas con más visitas en el periodo enero-junio 2017?

The weeks with the highest number of visits correspond to the first, second and fourth of the month of January 2017. These days that stand out the most from those weeks are; Saturday 7th as the first week of January, Saturday 14th as the second week of January and Thursday 26th as the fourth week of January.

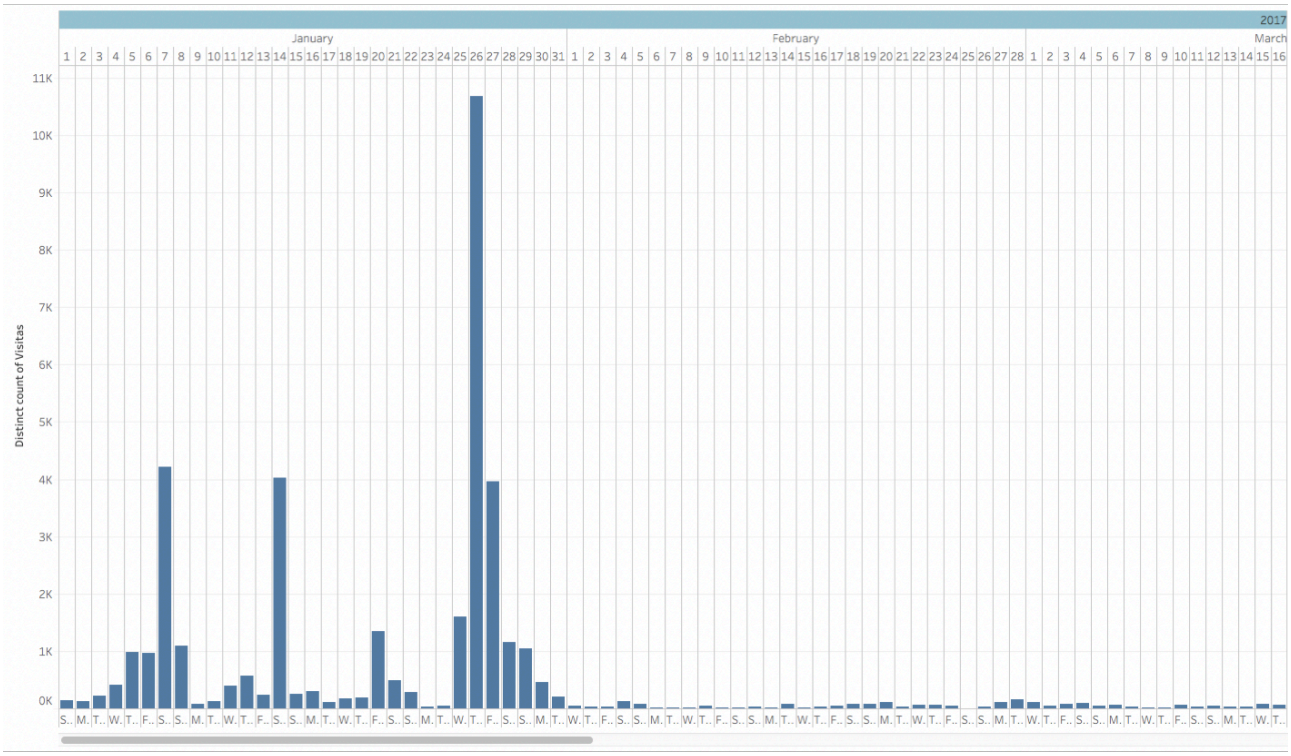
To obtain the above results, the first thing we will do is filter by the months of January to June of the year 2017 (in the "Filters" section of Tableau and select the months mentioned). We then draw a graph. To achieve this, we go to the “Show me” section and select the horizontal bar graph as explained in previous sections.

The output of the graph is as follows (we only show the section where the three weeks with the most visits appear);

How many users (ips) have visited us every month from January to June 2017?

There are a total of 619 users who have visited us every month from January to June.

For the calculation, as we had already filtered in the previous question for the months of January to June 2017, the filter of the MES variable is maintained and we count the number of users (IP). To do this, within the main Tableau screen, we drag this variable and it returns it directly (using the code IP(COUNT(Distinct))) with the following output;



Anyo	
2017	619

What percentage of increase/decrease in visits has there been by browser and version between February and March 2017?

We are interested in comparing the percentage of increase/decrease in visits between the months of February and March according to the version of the operating system. To do this, we create a pivot table with the columns "Month" and "Year" and in the rows OS Version and Browser Version. It gives us the following output as a result;

The idea would be that for browser version 0 of operating system 0, the increase between February and March is 0.72% (The difference between the column for March and February). To get the total percentages to appear by column in each cell of the pivot table, we do the following;

		2017	
Version So	Version Na..	February	March
0	0	17,68%	18,40%
	1.x	0,55%	0,54%
	2.x		0,41%
	6.x	0,55%	2,30%
	7.x	3,18%	3,11%
Fedora 9	1.x	1,52%	1,49%
Mac OS X	0	0,69%	3,52%
	2.x	0,41%	
	3.x		0,54%
Mandriva	2.x	0,55%	
Ubuntu 7.x	2.x		0,81%
Windows 2000	1.x	0,28%	
	2.x	0,41%	
	6.x		0,95%
Windows Vista	7.x	7,60%	3,25%
Windows XP	1.x		0,14%
	2.x	12,85%	15,29%
	6.x	25,55%	16,91%
	7.x	28,18%	31,12%
	9.x		0,95%
	11.x		0,27%

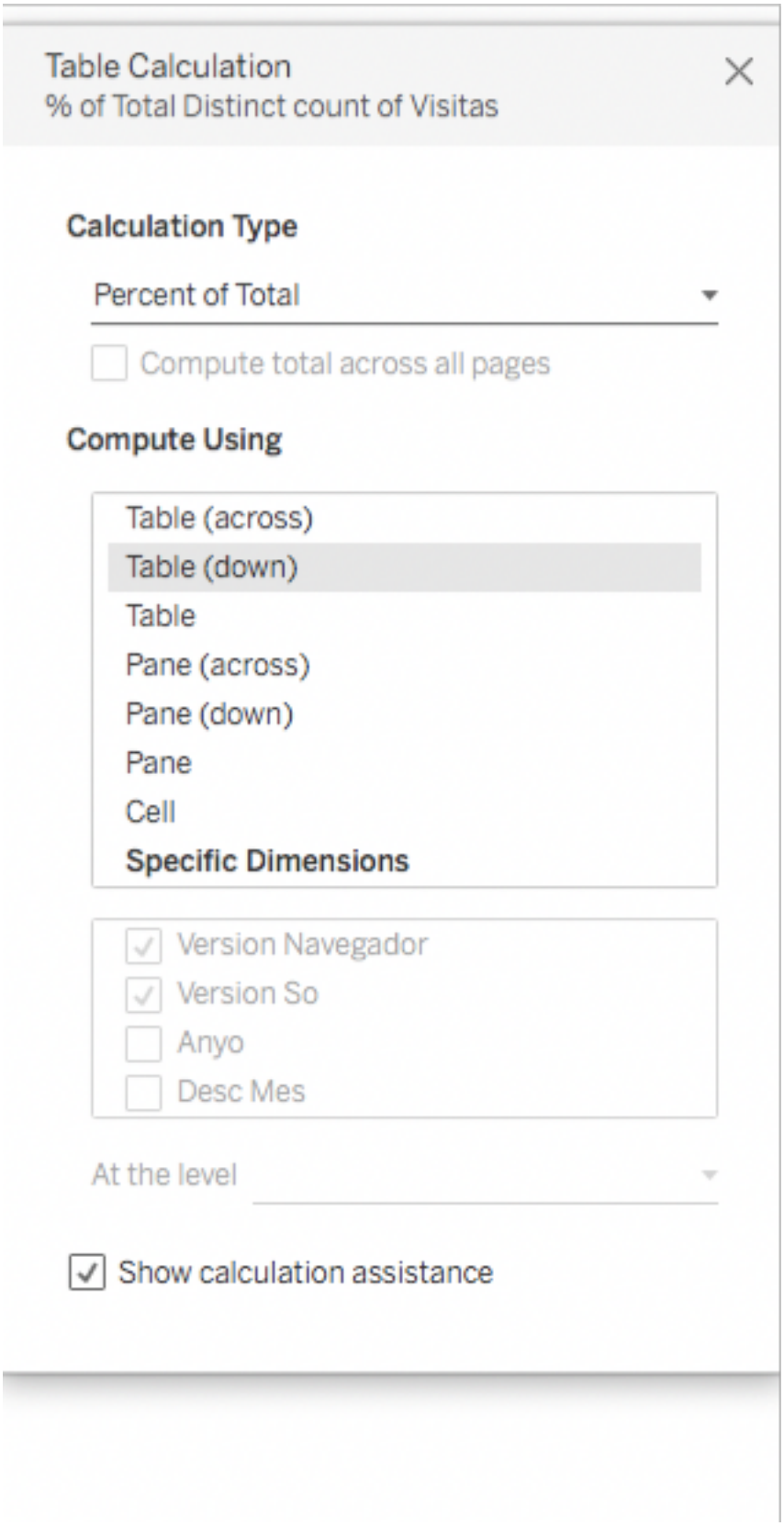


Table Calculation
% of Total Distinct count of Visitas
Calculation Type
Percent of Total
Compute total across all pages
Compute Using
Table (across)
Table (down)
Table
Pane (across)
Pane (down)
Pane
Cell
Specific Dimensions
Version Navegador
Version So
Anyo
Desc Mes
At the level
Show calculation assistance