Tableau + TPC-DS

Braulio Blanco
Gonçalo Moreira

+ableau

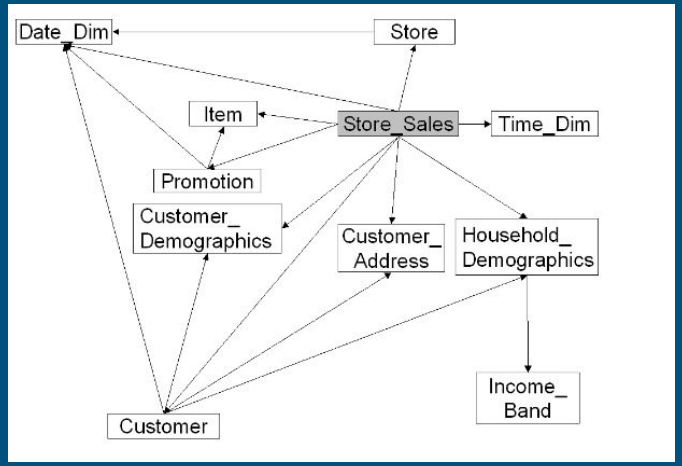

Date_Dim
Store
Item
Store_Sales
Time_Dim
Promotion
Customer_ Demographics
Customer_ Address
Household_ Demographics
Income_ Band
Customer

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms
CHALLENGERS
LEADERS
Microsoft
Tableau
MicroStrategy
ThoughtSpot
Qlik
Looker
Sisense
Salesforce
TIBCO Software
SAS
GoodData
Domo
BOARD International
SAP
Yellowfin
Logi Analytics
Oracle
Information Builders
IBM
Pyramid Analytics
Birst
NICHE PLAYERS
VISIONARIES
ABILITY TO EXECUTE
COMPLETENESS OF VISION
As of January 2019
© Gartner, Inc
Source: Gartner (February 2019)

Tableau suite

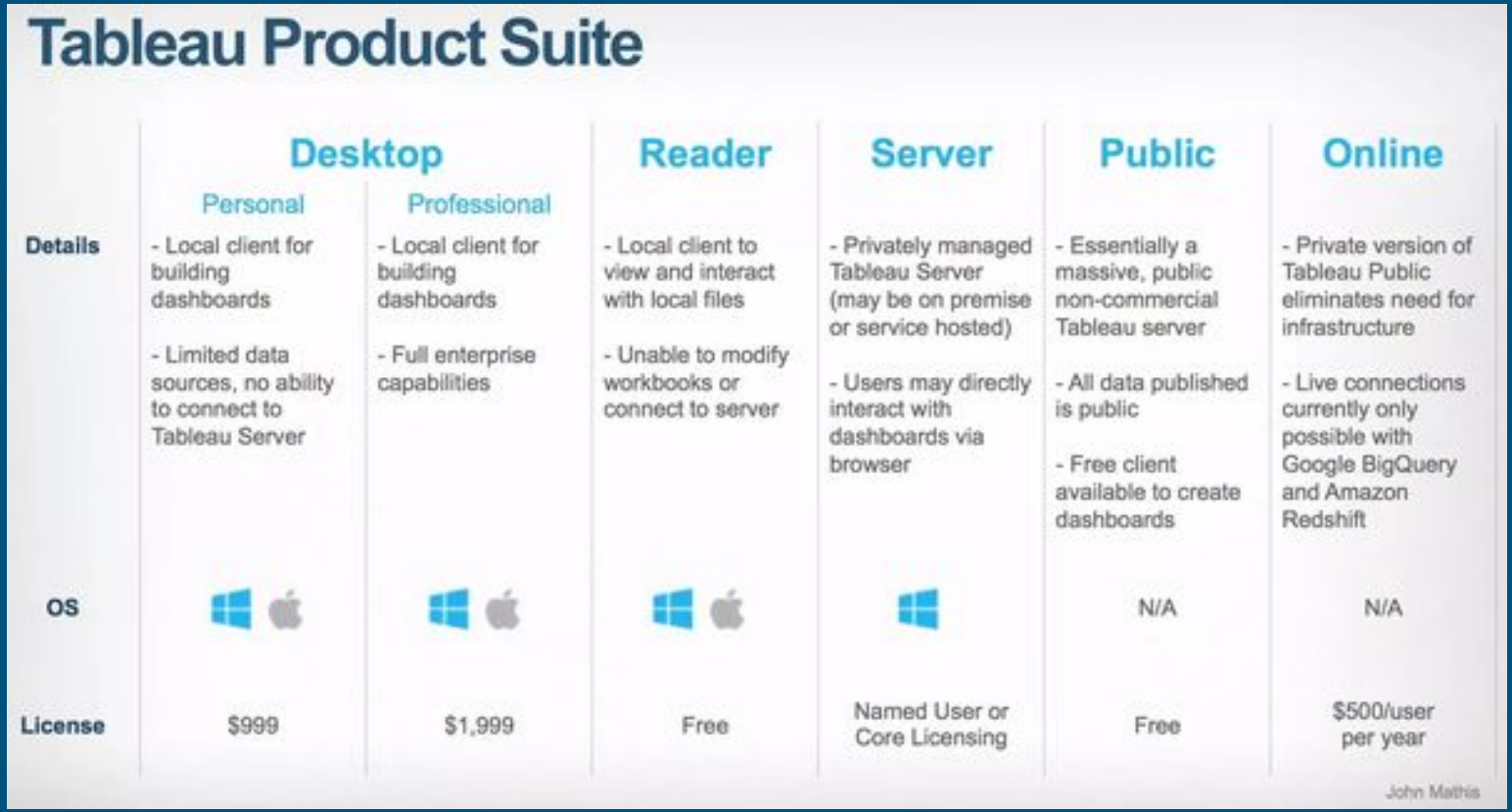


Tableau Desktop

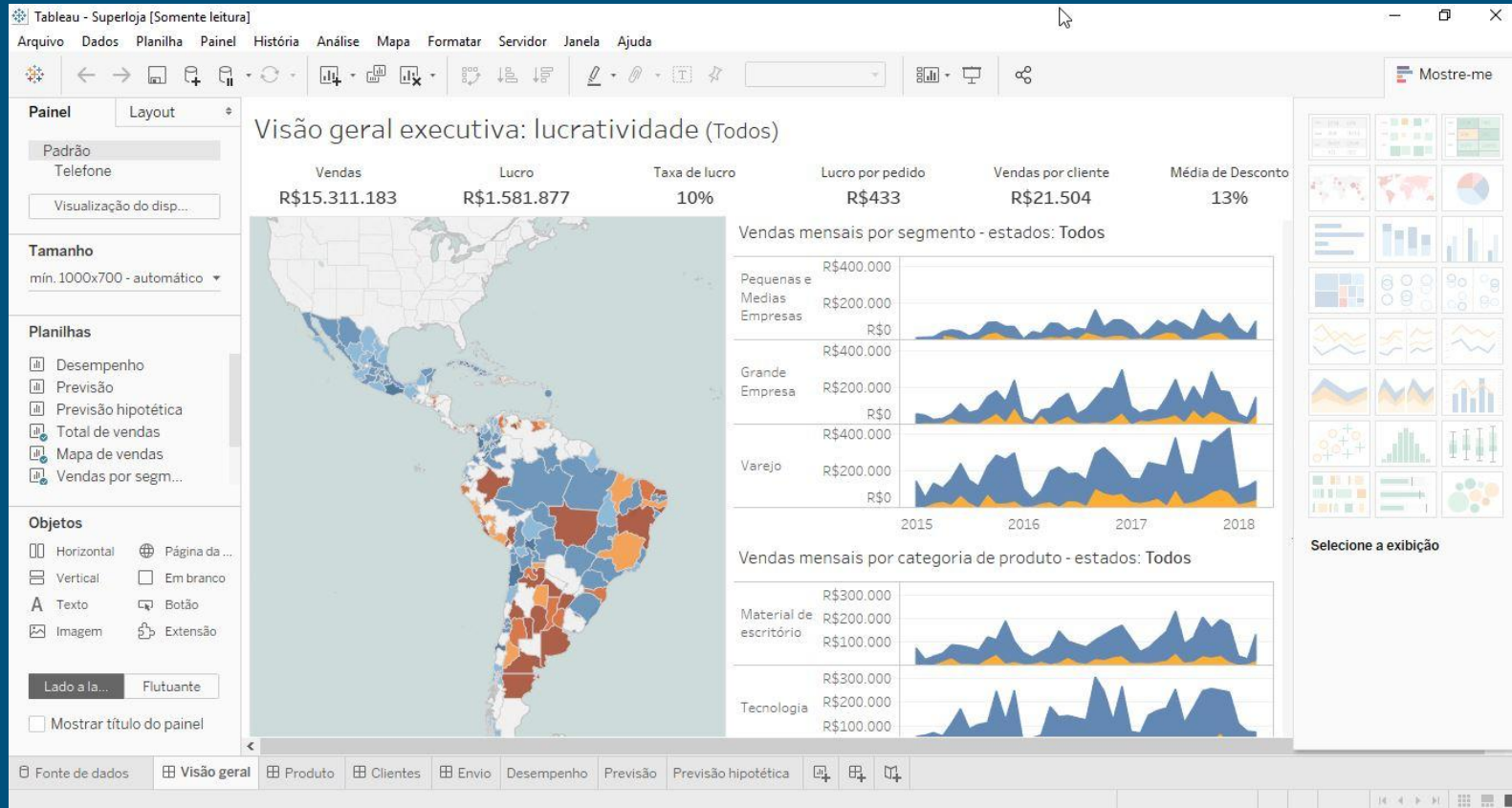


Tableau Prep Builder

Tableau Server on cloud

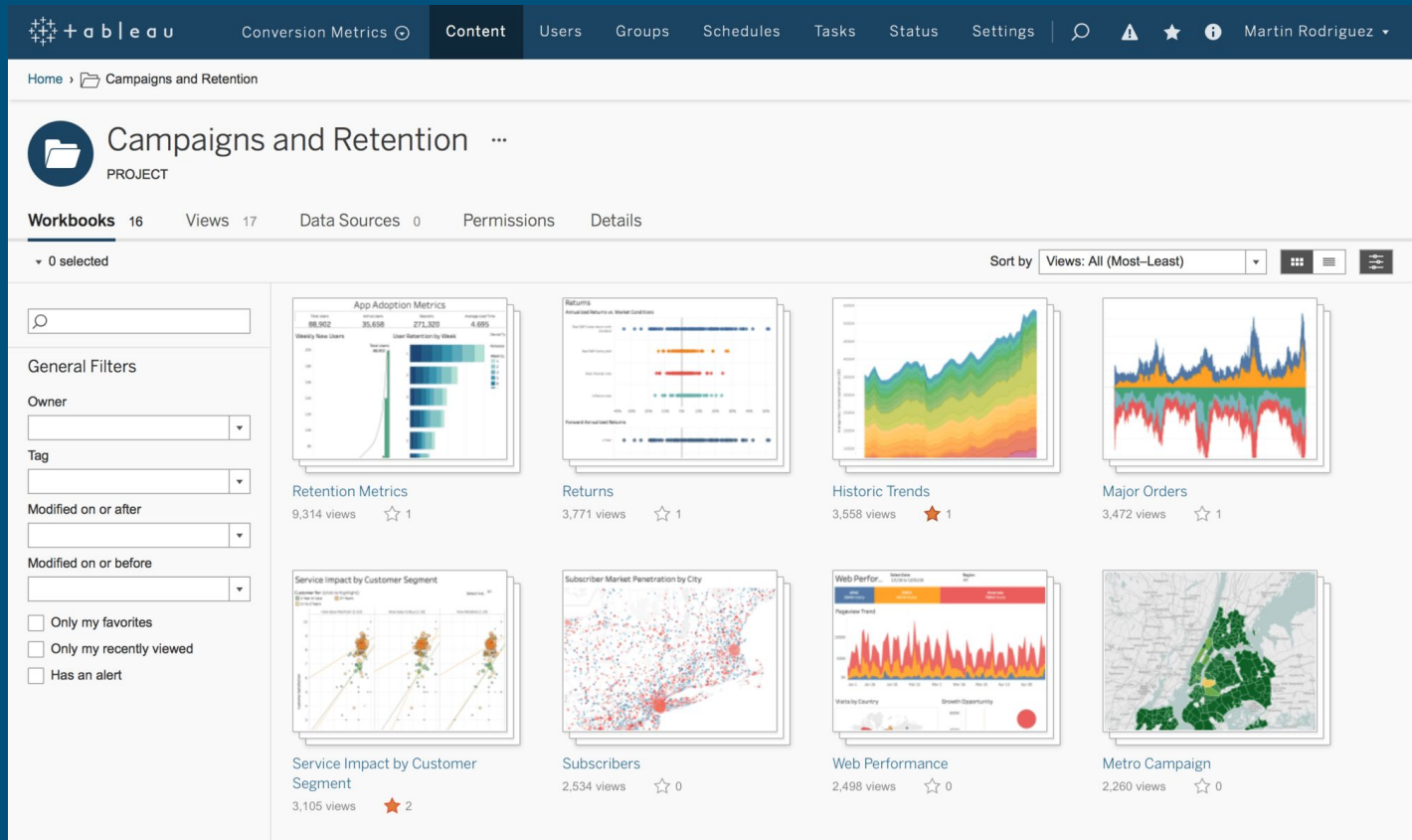
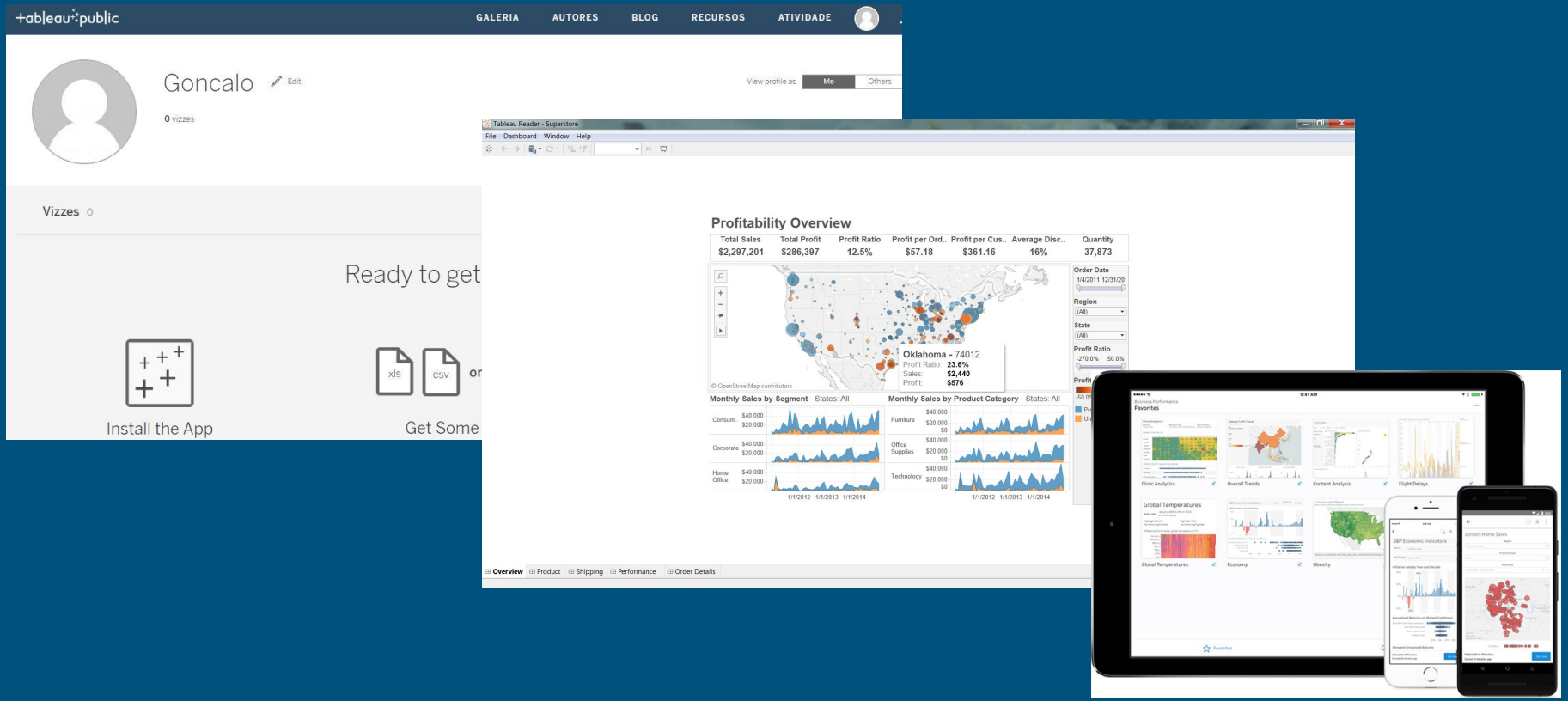


Tableau Public, Mobile and Reader

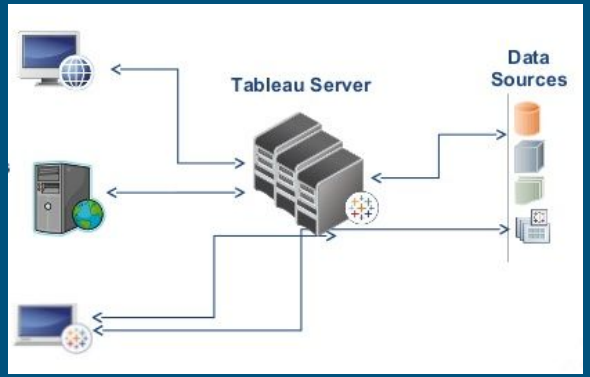


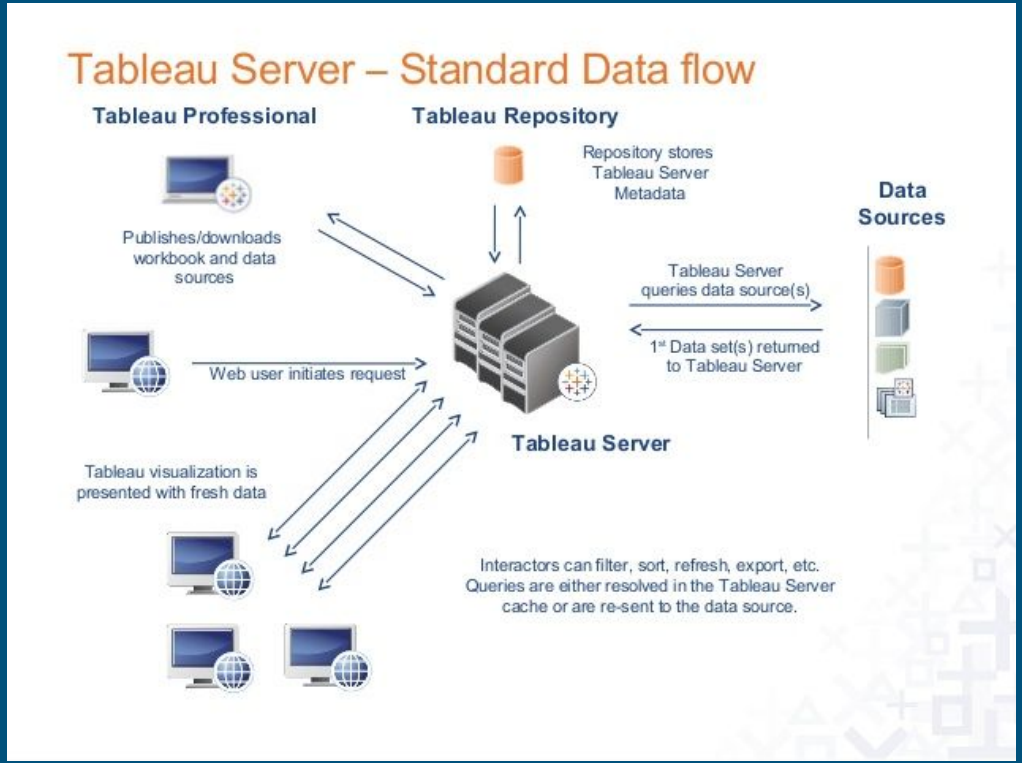
Data sources

- Tableau Server
- Actian Matrix
- Actian Vector
- Amazon Athena
- Amazon Aurora
- Amazon EMR Hadoop Hive
- Amazon Redshift
- Anaplan
- Apache Drill
- Banco de dados Aster
- Azure SQL Data Warehouse
- Box
- Cloudera Hadoop
- Denodo
- Dropbox
- Exasol
- Firebird
- Google Ads
- Google Analytics
- Google BigQuery
- Google Cloud SQL
- Google Drive
- Planilhas Google
- Hortonworks Hadoop Hive
- IBM BigInsights
- IBM DB2
- IBM PDA (Netezza)
- Intuit QuickBooks Online
- Intuit QuickBooks Online (9.3-2018.1)
- Kognitio
- MapR Hadoop Hive
- MariaDB
- Marketo
- MarkLogic
- MemSQL
- Microsoft Analysis Services
- Microsoft PowerPivot
- Microsoft SQL Server
- MonetDB
- Conector do MongoDB BI
- MySQL
- OData
- OneDrive
- Oracle
- Oracle Eloqua
- Oracle Essbase
- Pivotal Greenplum Database
- PostgreSQL
- Presto
- Progress OpenEdge
- Salesforce
- SAP HANA
- SAP NetWeaver Business Warehouse
- SAP Sybase ASE
- SAP Sybase IQ
- ServiceNow ITSM
- Listas do SharePoint
- Snowflake
- Spark SQL
- Splunk
- Teradata
- Teradata OLAP Connector
- Virtualização de dados TIBCO
- Vertica
- Conector de dados da web

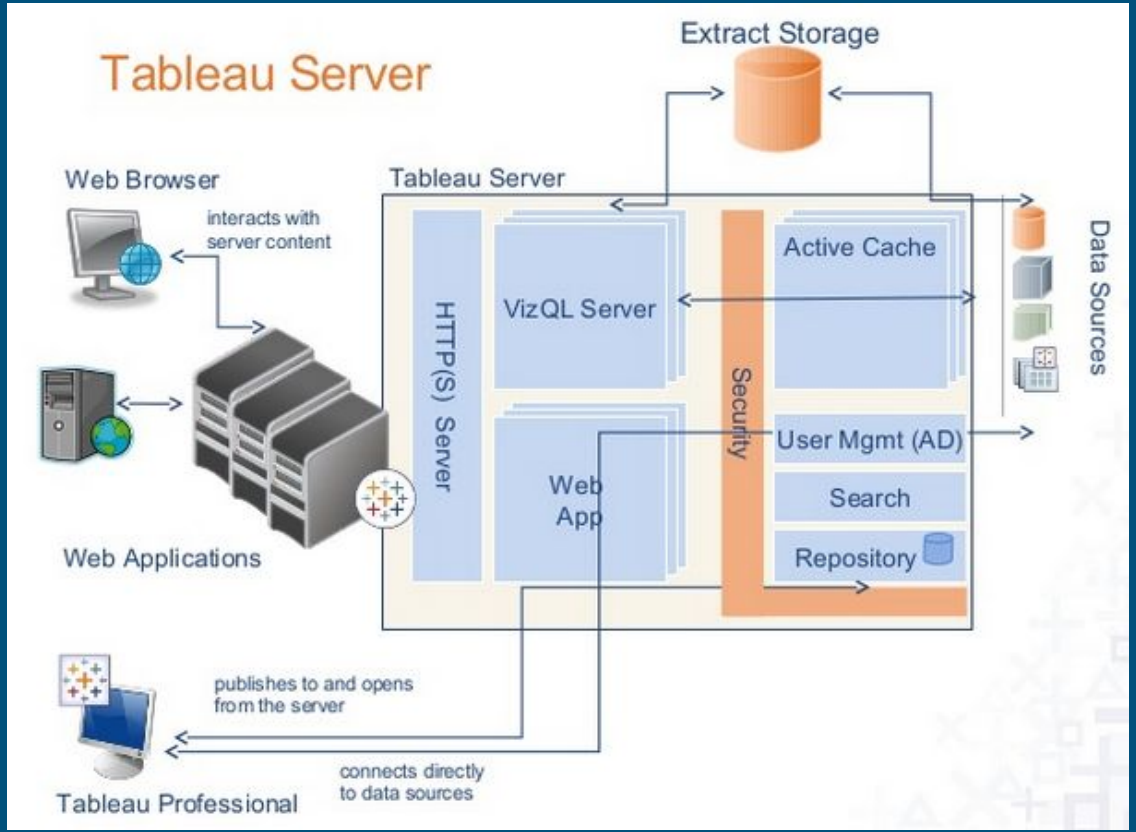
- Outros bancos de dados (JDBC)
- Outros bancos de dados (ODBC)

Basic Architecture



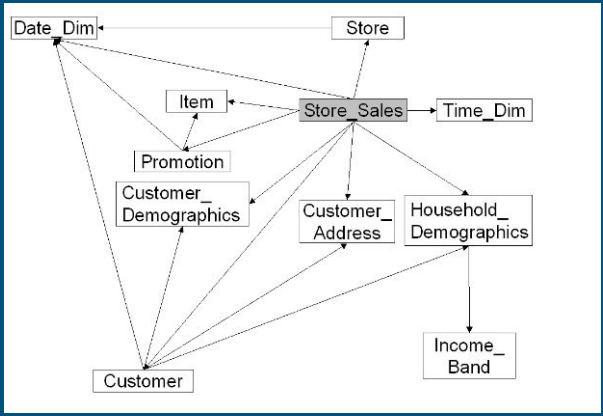


Basic Architecture

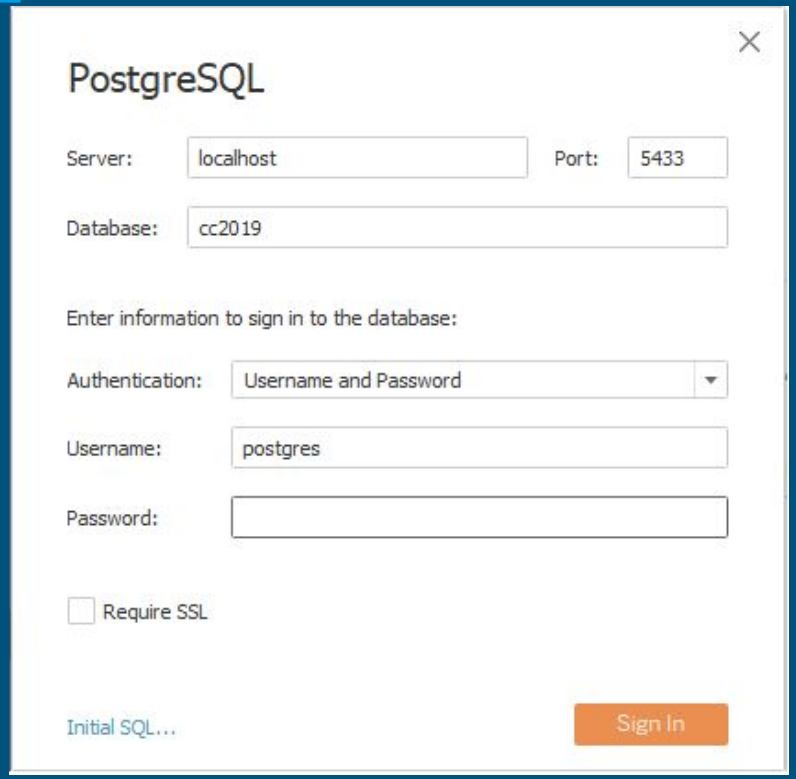


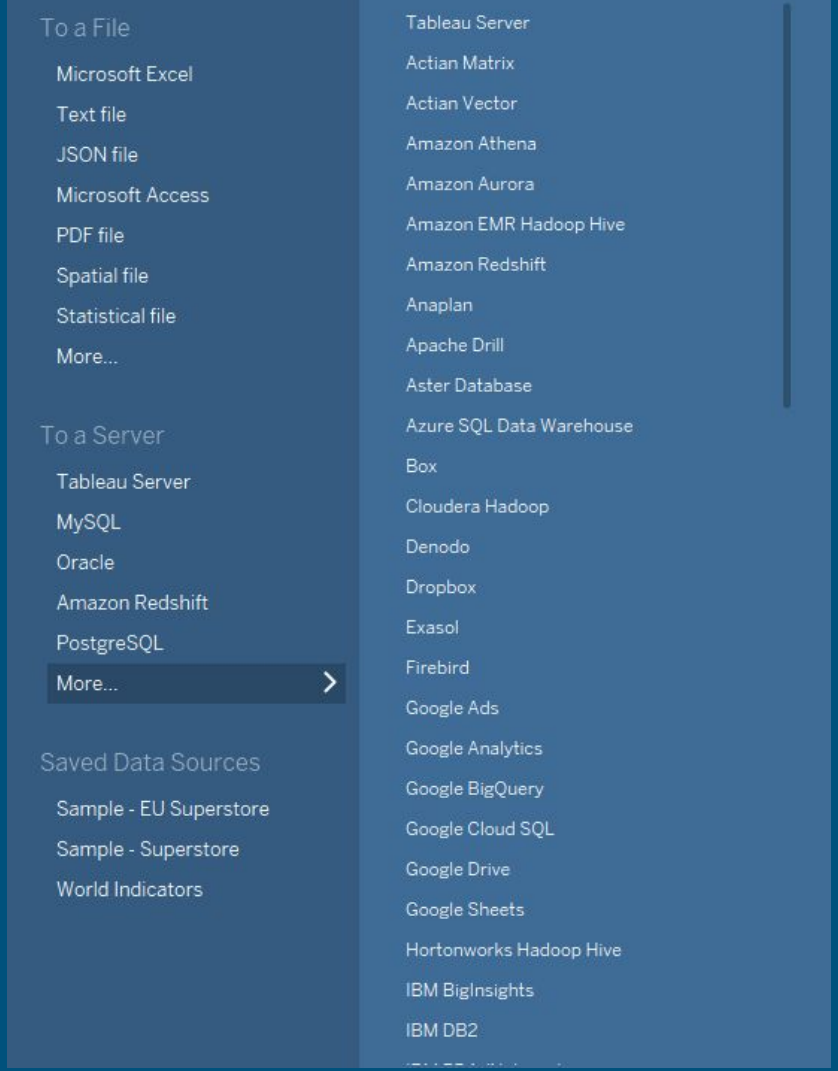
Practical case using TPC-DS

Table	Avr Row Size in bytes	Sample Row Counts. Number of rows are within 1/100th Percent of these numbers							
		1GB	100GB	300GB	1TB	3TB	10TB	30TB	100TB
call_center	305	6	30	36	42	48	54	60	60
catalog_page	139	6	30	36	42	48	54	60	60
catalog_returns	166	11,718	20,400	26,000	30,000	36,000	40,000	46,000	50,000
catalog_sales	226	144,067	14,404,374	43,193,472	143,996,756	432,018,033	1,440,033,112	4,319,925,093	14,400,509,482
customer	132	1,441,548	143,997,065	431,969,836	1,439,980,416	4,320,078,880	14,399,964,710	43,200,404,822	144,001,292,896
customer_address	110	100,000	2,000,000	5,000,000	12,000,000	30,000,000	65,000,000	80,000,000	100,000,000
customer_ demographics	42	50,000	1,000,000	2,500,000	6,000,000	15,000,000	32,500,000	40,000,000	50,000,000
date_dim	141	1,920,800	1,920,800	1,920,800	1,920,800	1,920,800	1,920,800	1,920,800	1,920,800
household_ demographics	21	73,049	73,049	73,049	73,049	73,049	73,049	73,049	73,049
income_band	16	7,200	7,200	7,200	7,200	7,200	7,200	7,200	7,200
inventory	16	20	20	20	20	20	20	20	20
item	281	11,745,000	399,330,000	585,684,000	783,000,000	1,033,560,000	1,311,525,000	1,627,857,000	1,965,337,830
promotions	124	18,000	204,000	264,000	300,000	360,000	402,000	462,000	502,000
reason	38	300	1,000	1,300	1,500	1,800	2,000	2,300	2,500
ship_mode	56	35	55	60	65	67	70	72	75
store	263	20	20	20	20	20	20	20	20
store_returns	134	12	402	804	1,002	1,350	1,500	1,704	1,902
store_sales	164	287,514	28,795,080	86,393,244	287,999,764	863,989,652	2,879,970,104	8,639,952,111	28,799,941,488
time_dim	59	2,880,404	287,997,024	864,001,869	2,879,987,999	8,639,936,081	28,799,983,563	86,399,341,874	287,998,696,432
warehouse	117	86,400	86,400	86,400	86,400	86,400	86,400	86,400	86,400
web_page	96	5	15	17	20	22	25	27	30
web_returns	162	60	2,040	2,604	3,000	3,600	4,002	4,602	5,004
web_sales	226	71,763	7,197,670	21,599,377	71,997,522	216,003,761	720,020,485	2,160,007,345	7,200,085,924
web_site	292	719,384	72,001,237	216,009,853	720,000,376	2,159,968,881	7,199,963,324	21,600,036,511	71,999,537,298

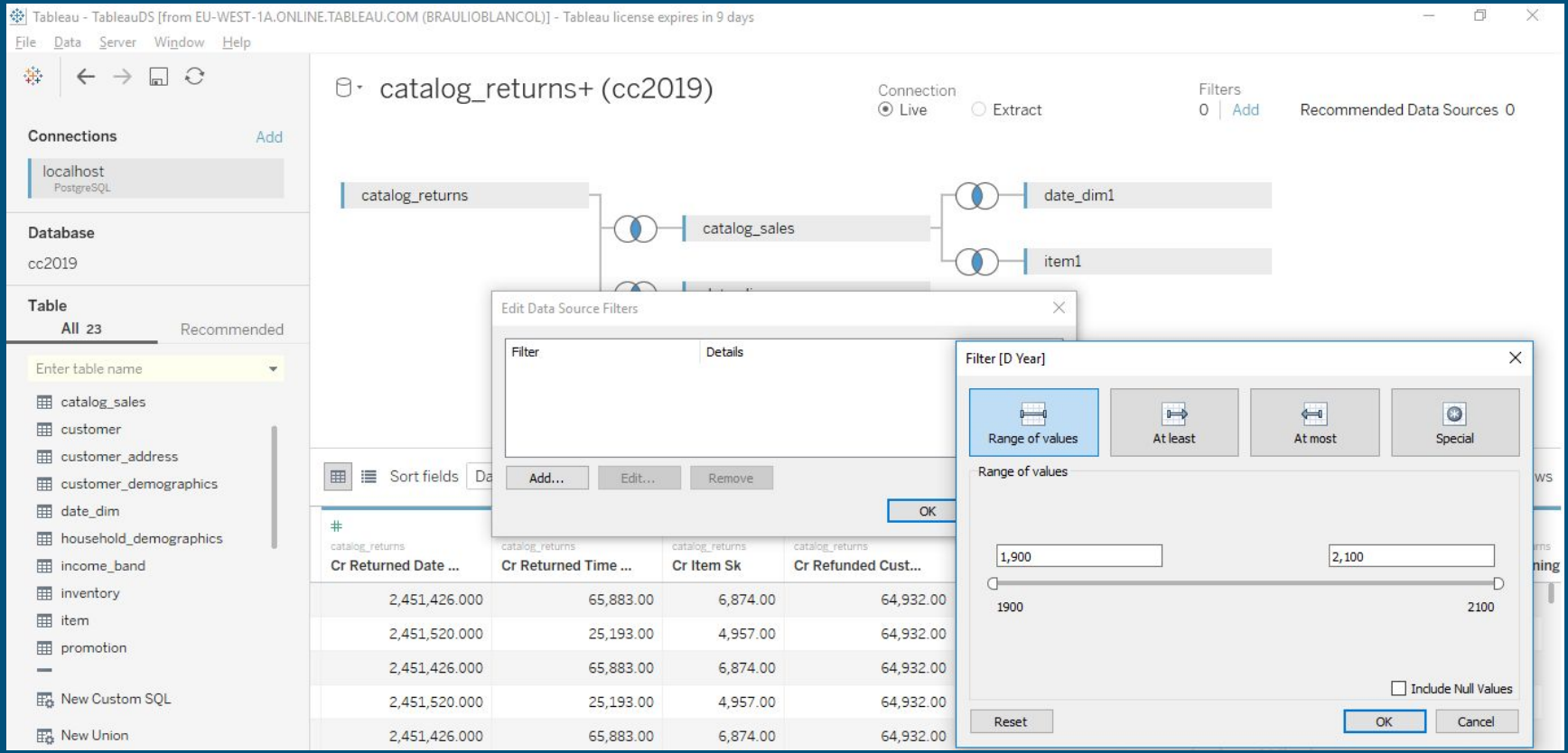


Using Tableau Desktop

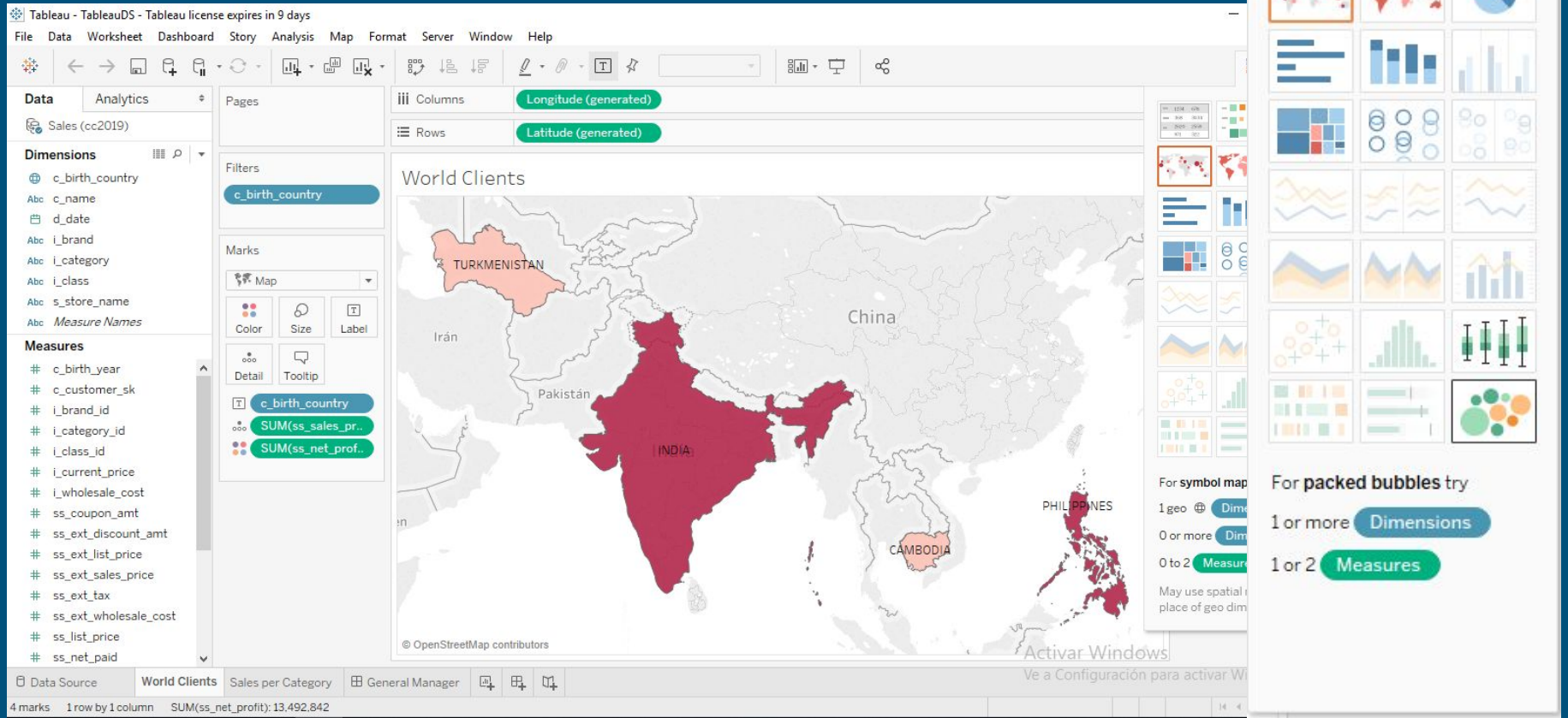




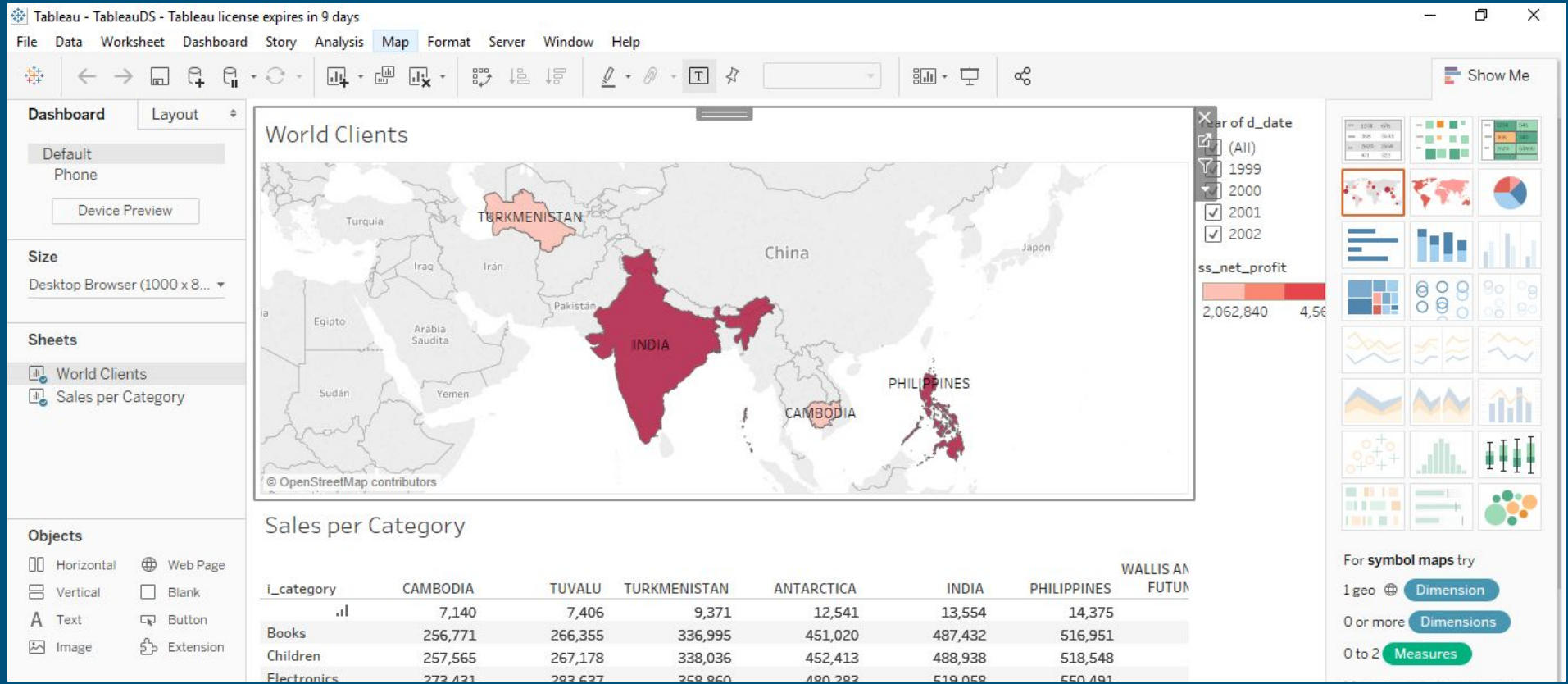
Creating data sources using Tableau Tools



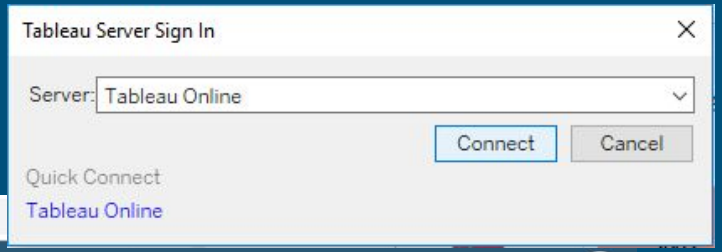
Creating worksheets

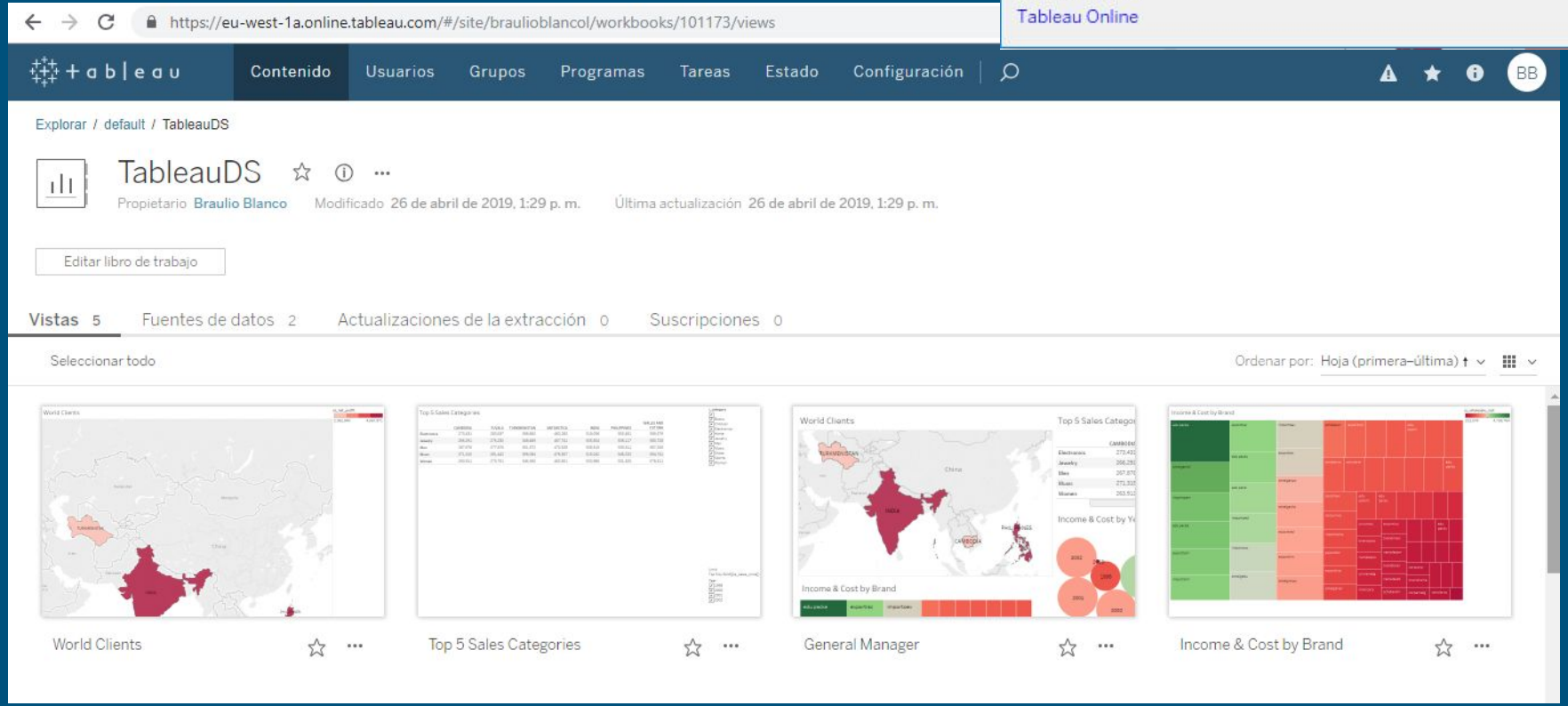


Creating dashboards

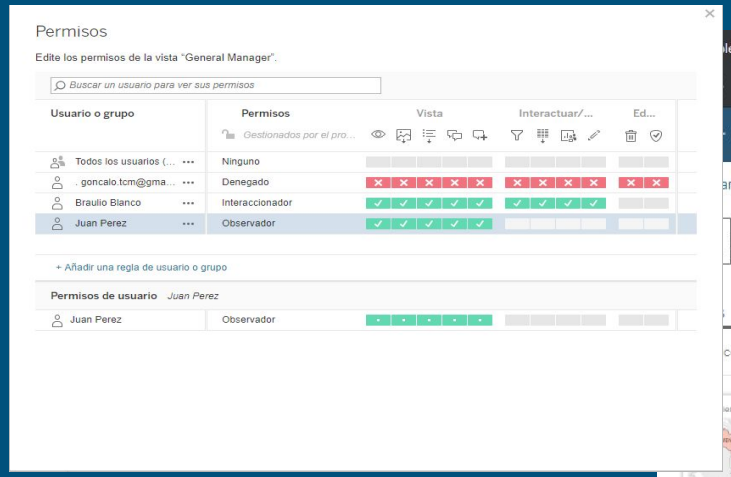


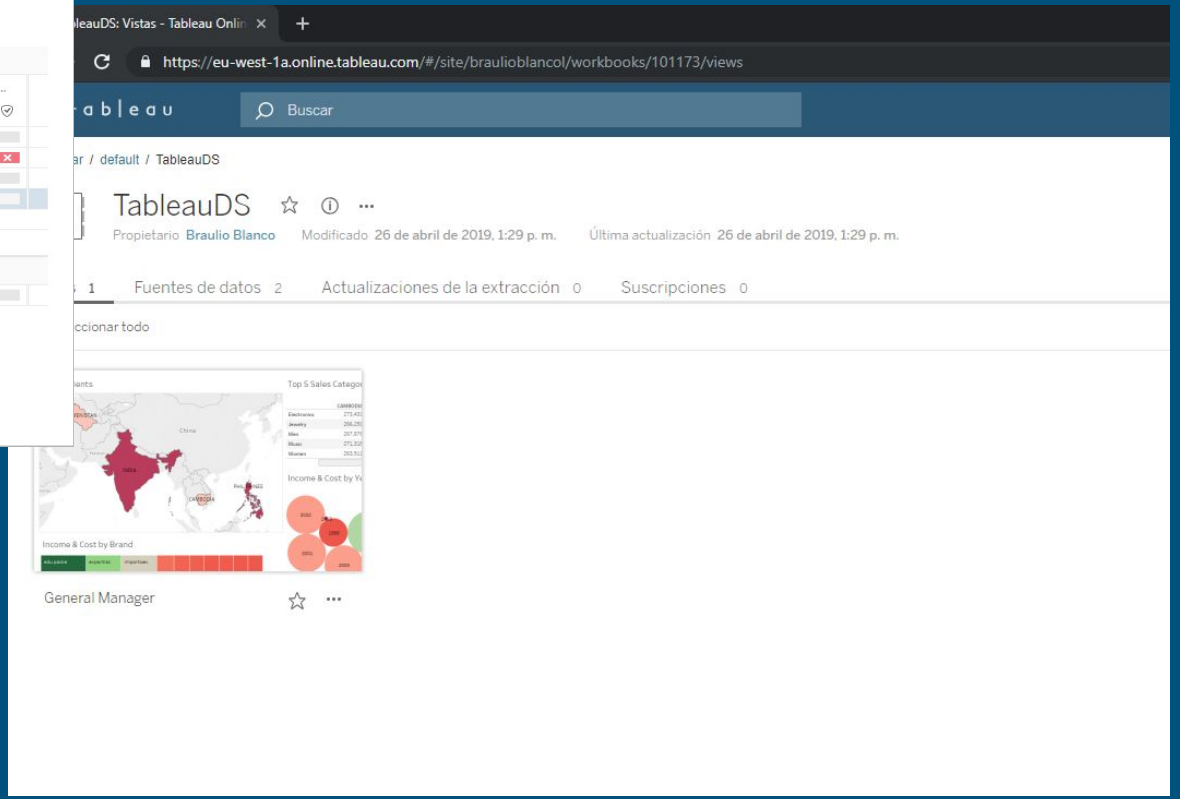
Publishing to the cloud



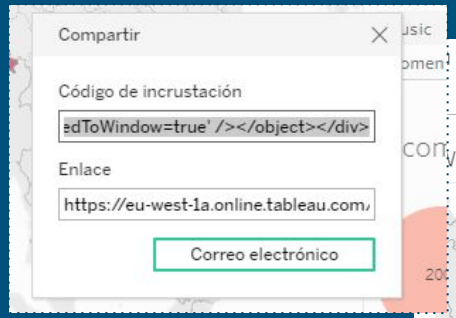


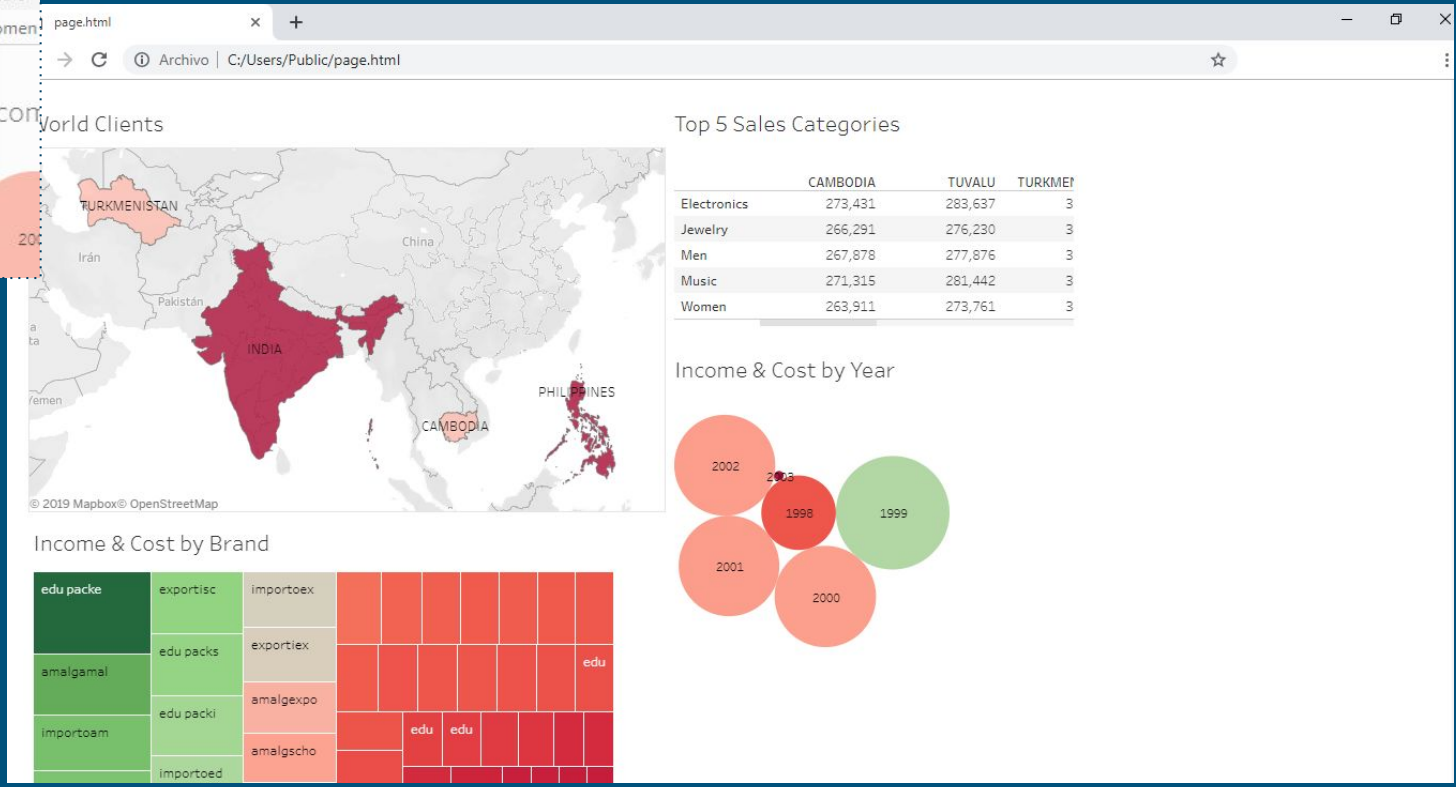
Security: Authentication and Authorization





Integration





CONCLUSIONS

- Tableau has designer tools that are very intuitive even for final users with low support of IT, this could allow the companies to obtain a direct participation of the business users in the Business Intelligence project development process.
- The architecture of Tableau is quite simple but allows to configure a basic distributed architecture and the disconnected data using the extract feature allow improve the performance.
- The integration feature can be useful if Single Sign On is implemented and the size of the charts are not heavy to be passed through the network.

Learn more...

→ Start free trial
→ Tableau Help
→ Tableau Blog
→ Tableau Community
→ Tableau Certification
→ Tableau E-learning
→ YouTube channel
→ Classroom training (in-person + virtual)
→ Tableau Architecture