

Comparativa de tecnologies de BI

Martí Rius Ollé



1 INTRODUCCIÓ - CONTEXT DEL TREBALL

1.1 Motivació

He decidit fer el meu Treball de Fi de Grau (TFG) sobre Business Intelligence, ja que és un camp que m'agrada i vull aprofundir-me més en les seves diferents eines i possibilitats que ofereix. També s'ha plantejat la idea de realitzar el treball amb l'empresa on treballo, que és Mercanza, una empresa especialitzada en Business Intelligence que es centra en l'ús de Qlik.

A aquest punt, us preguntareu, ¿què és exactament el Business Intelligence? Bé, el BI és un conjunt d'eines i tecnologies que permeten a les organitzacions des de la recopilació de dades fins a l'anàlisi d'aquestes dades i la generació d'informació per a prendre les millors decisions possibles. En definitiva, el BI ajuda les empreses a comprendre millor el seu entorn, a identificar les tendències tant en el mercat com dins de la pròpia empresa, a avaluar el rendiment dels seus productes i empleats, i, en definitiva, a optimitzar les seves decisions en totes les àrees de l'empresa.

I quines eines hi ha? Us preguntareu, doncs bé, hi ha moltes eines diferents, cadascuna una amb les seves aventatges i inconvenients. No obstant això, les tres més populars són:

- Power BI: Desenvolupada per Microsoft, destaca per la seva integració amb altres productes de la mateixa companyia.
- Qlik View/Sense: Plataforma líder del mercat fins que va arribar Power BI, té una gran comunitat.
- Tableau: És una eina sòlida, tot i que no destaca particularment en cap àrea.

Aquestes són actualment les tres tecnologies més estandaritzades. Pel meu TFG, he decidit realitzar una comparativa entre Power BI i Qlik Sense. Personalment, crec que aquestes dues són les millors opcions i m'agradaria aprofundir en les seves capacitats. Cal destacar que utilitzo Qlik Sense a diari en la meva feina, així que el que més m'interessa del treball és explorar les capacitats de Power BI, ja que és una eina en expansió i comparar-la amb la que considero la millor.

1.2 Objectius

Com he mencionat anteriorment, el meu objectiu principal és avaluar les capacitats de cadascuna de les eines, però a més d'això, hi ha altres objectius que he de complir per arribar al meu objectiu final. Aquests objectius inclouen:

1. **Creació del model de dades:** En cada una de les eines, s'ha de desenvolupar el model de dades. Cada eina utilitza un llenguatge de programació diferent. Qlik Sense utilitza el seu propi llenguatge anomenat Qlik Sense Expression Language, que es basa en SQL. D'altra banda, Power BI utilitza DAX, un llenguatge desenvolupat per Microsoft. En aquesta fase, el meu interès és comprendre les capacitats i la velocitat d'execució de cada llenguatge.
2. **Creació de les visualitzacions:** Un cop s'ha establert el model de dades, cal crear les visualitzacions per mostrar els resultats de l'anàlisi. En aquesta etapa, vull explorar les capacitats de generació de gràfics de cadascuna de les eines, així com les opcions de personalització i la interactivitat que ofereixen.
3. **Realitzar la comparativa final:** En aquesta fase, aglutinaré tots els aspectes rellevants per a la comparació final. Això inclourà els llenguatges de programació utilitzats tant pel model com per les visualitzacions, les capacitats de visualització, la velocitat d'execució, el rendiment de les aplicacions, les funcionalitats específiques de cada eina, la portabilitat i altres aspectes rellevants.

A través d'aquests objectius, pretenc aprofundir en la comprensió de les diferències i similituds entre Power BI i Qlik Sense per a poder realitzar una comparativa exhaustiva i arribar a una conclusió fonamentada.

1.3 Metodologia

Seguiré una metodologia àgil mitjançant Jira, ja que crec que és la millor manera de planificar-me i mantenir-me en el camí correcte. Personalment, trobo que tenir un calendari amb tasques, dates i estimacions de temps és molt beneficiós. També vull destacar la facilitat amb què Jira permet crear tasques segons les necessitats i gestionar-les mitjançant un tauler, assignant-les com a "pendents", "en curs", "en revisió" o "finalitzades". Això em permet mantenir el control del progrés del projecte i el seguiment d'aquest.

Pel que fa al repositori, utilitzaré GitHub, ja que considero que és la millor opció per fer seguiment dels documents del projecte i accedir-hi des de qualsevol lloc. A més, és la plataforma amb la qual em sento més còmode, ja que he utilitzat GitHub anteriorment en altres projectes. Personalment, sempre he preferit GitHub a altres opcions com GitLab o Bitbucket.

1.4 Planificació

Per poder desenvolupar el projecte correctament hauria de seguir la següent seqüència d'activitats:

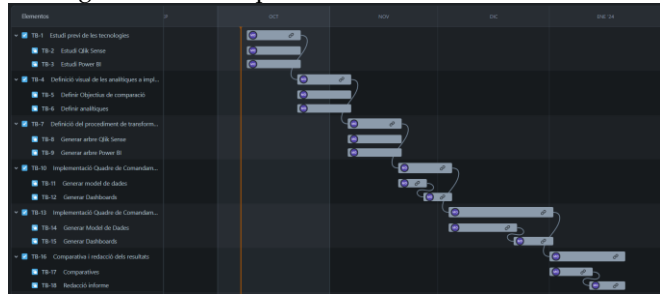
- Estudi previ de les dues tecnologies – 20 hores
- Definició visual de les analítiques a implementar i objectius de comparació – 30 hores
- Definició del procediment de transformació de dades per construir l'ETL – 40 hores
- Implementació Quadre de Comandament de Vendes en Qlik – 70 hores
- Implementació Quadre de Comandament de Vendes en Power BI – 70 hores
- Estudi de les proves comparatives definides i redacció dels resultats - 30 hores
- Reunions i documentació – 40 hores.

Utilitzant Jira quedaria tal que:

Tipo	#	Clave	Resumen
✓	✓	TB-1	Estudi previ de les tecnologies
	✓	TB-2	Estudi Qlik Sense
	✓	TB-3	Estudi Power BI
✓	✓	TB-4	Definició visual de les analítiques a implementar i objectius de comparació
	✓	TB-5	Definir Objectius de comparació
	✓	TB-6	Definir analítiques
✓	✓	TB-7	Definició del procediment de transformació de dades per construir l'ETL
	✓	TB-8	Generar arbre Qlik Sense
	✓	TB-9	Generar arbre Power BI
✓	✓	TB-10	Implementació Quadre de Comandament de Vendes en Qlik Sense
	✓	TB-11	Generar model de dades
	✓	TB-12	Generar Dashboards
✓	✓	TB-13	Implementació Quadre de Comandament de Vendes en Power BI
	✓	TB-14	Generar Model de Dades
	✓	TB-15	Generar Dashboards
✓	✓	TB-16	Comparativa i redacció dels resultats
	✓	TB-17	Comparatives
	✓	TB-18	Redacció informe

Il·lustració 1 Llista Jira

I el diagrama de Gantt quedaria:



Il·lustració 2: Cronograma Jira

Es pot observar que tot i que l'implementació dels quadres de comandament d'ambes eines els he assignat unes 70 hores orientatives de treball, en el diagrama he otorgat un major pes a Power BI que a Qlik. Això es deu al fet que ja tinc experiència prèvia amb Qlik i, per tant, considero que podré completar l'implementació més ràpidament, per aquest motiu prefereixo assignar-li un major temps a Power BI; per poder desenvolupar fins al mateix punt tot i que la quantitat d'hores en cada eina sigui diferent.

2 MODEL DE DADES – QLIK SENSE

2.1 Conexions Utilitzades

Per tal de realitzar el model de dades en Qlik he utilitzat les següents connexions:

1. **OLE DB:** Per fer la connexió a la base de dades (Certification.mdb); en ser una base de dades en local (desktop) únicament s'havia d'afegir la ruta com a source, sense usuari ni password.

Il·lustració 3: Connexió OLE DB

2. **Carpeta:** Pels altres documents (.PNG, .xlsx, .txt) els tenia en una carpeta en local i únicament havia d'afegir la ruta de la carpeta a la connexió.

Il·lustració 4: Connexió Carpeta

2.2 Funcions Utilitzades

- **LET i SET:** Per declarar variables; LET s'utilitza si es vol avaluar el valor, mentre que SET si es vol prendre el valor literal; per exemple si afegeixes una data com a SET obtindràs la data; per exemple 07/11/2023; però si l'afegeixes com a LET obtindràs $3.14 \cdot e^{-4}$; degut a que et calcularà 7 dividit entre 11 entre 2023.
- **LOAD:** Per carregar dades d'una taula; per extreure valor d'una taula s'utilitza el següent format: LOAD, [Camps que vols carregar],FROM,[Taula a la qual pertanyen] (d'una connexió). Exemple:

```
LOAD
SpecialOfferID,
Description,
DiscountPct,
"Type",
Category
FROM [lib://Data/Certification Data.xlsx]
(ooxml, embedded labels, table is [Special Offer]);
```

Il·lustració 5: Exemple LOAD

- **RESIDENT:** Igual que un LOAD, però en comptes de carregar una taula des d'una connexió, carregues els camps que vols d'un altra taula anteriorment creada en l'script.
- **QUALIFY i UNQUALIFY:** S'utilitzen quan vols tenir el nom de la taula com a prefix pels noms dels camps. L'he hagut d'utilitzar, ja que he duplicat la taula Adress en dues, perquè tenia BillAdress i ShipAdress; Qlik sense no permet unions múltiples (unir una taula amb un altre amb més d'un camp) així que s'ha de duplicar la taula Adress per poder apuntar una cap a Ship Adress i l'altre cap a Bill Adress. Qualify et permet que aquestes dues taules noves anomenades ShipAdress i BillAdress no s'uneixin entre elles, ja que us podeu imaginar que en duplicar la taula tots els camps dins d'elles són idèntics, i en comptes d'anar renombrant un a un perquè no s'uneixin entre ells és molt més còmode utilitzar el qualify perquè tots els camps de Ship Adress comencin amb Ship per exemple; Així AdressID, City, PostalCode, TerritoryID etc.. quedaran com a ShipAdressID, ShipCity, ShipPostalCode, ShipTerritoryID etc.. i no s'unirien ni amb la taula de Sales ni amb la taula de Bill Adress.
- **CROSSTABLE:** S'usa per convertir una taula creuada en una taula simple, s'entén millor amb l'exemple. Tenia una taula de Quotes la qual estava de la següent manera:

	275	276	277	279	280	282
01/07/2020	200.000,00	850.000,00	400.000,00	550.000,00	700.000,00	700.000,00
01/10/2020	190.000,00	808.000,00	380.000,00	525.000,00	665.000,00	665.000,00
01/01/2021	190.000,00	808.000,00	380.000,00	525.000,00	665.000,00	665.000,00
01/04/2021	220.000,00	935.000,00	440.000,00	605.000,00	770.000,00	770.000,00
01/07/2021	580.000,00	1.600.000,00	775.000,00	630.000,00	805.000,00	450.000,00
01/10/2021	550.000,00	1.520.000,00	735.000,00	600.000,00	765.000,00	425.000,00
01/01/2022	550.000,00	1.520.000,00	735.000,00	600.000,00		425.000,00
01/04/2022	638.000,00	1.750.000,00	853.000,00	693.000,00		495.000,00
01/07/2022	667.000,00	1.850.000,00	850.000,00	725.000,00		660.000,00
01/10/2022	635.000,00	1.750.000,00	800.000,00	680.000,00		627.000,00
01/01/2023	635.000,00	1.750.000,00	800.000,00	680.000,00		627.000,00
01/04/2023	735.000,00	2.035.000,00	935.000,00	795.000,00		726.000,00
01/07/2023	700.000,00	1.940.000,00	890.000,00	725.000,00		693.000,00

Il·lustració 6: Taula CROSSTABLE inicial

Tenint el número d'empleat adalt (vertical) i la data a l'esquerra (horitzontal) i les quotes pertinents en cada un. El que busquem és obtenir la següent taula:

Data	EmployeeID	Quota
01/07/2020	275	200.000,00
01/07/2020	276	850.000,00
01/07/2020	277	400.000,00
01/07/2020	279	550.000,00
01/07/2020	280	700.000,00
01/07/2020	282	700.000,00
01/10/2020	275	190.000,00
01/10/2020	276	808.000,00
01/10/2020	277	380.000,00
01/10/2020	279	525.000,00
01/10/2020	280	665.000,00
01/10/2020	282	665.000,00

Il·lustració 7: Taula resultat CROSSTABLE

Tenint el camp Data i l'id d'empleat com a columnes per poder iterar-les.

- **CONCATENATE i NOCONCATENATE:** Concatenate serveix tal com diu el seu nom per concatenar dades a una taula; sense juntarles amb les anteriorment afegides; i noconcatenate just el contrari, perquè no es concatenin les dades si no existeixen.
- **MAPPING:** Serveix per crear una taula de correspondències. Consta de dos valors: el valor clau (valor que té tant la taula MAPPING com a la taula on s'aplica) i el valor que es vol portar a la nova taula. Mitjançant el valor clau s'afegeixen els valors pertinents a la nova taula 1-1.
- **DROP TABLE:** Com el seu nom indica serveix per eliminar una taula de memòria.
- **RENAME:** S'utilitza per canviar de nom les taules.
- **JOIN:** Tal com el seu nom indica serveix per unir una taula amb un altra. Hi ha quatre tipus diferents:
 - Left-Join: Es carrega primer la taula de l'esquerra i seguidament se li afegeixen els camps de l'altra taula mitjançant una referència. (si no hi ha referència no s'afegeixen)
 - Inner-Join: Només s'afegeixen els camps amb el mateix ID a les dues taules.
 - Outer-Join: S'afegeixen tots els camps; hi hagi o no relació.
 - Contrari al Left-Join: Es carrega primer la taula de la dreta i se li afegeixen únicament els camps amb els quals té referència.

Exemple:

Examples

DATASETS

Target Table

Trade ID	Asset Class
101533	Fixed Income
606601	Commodities

Incoming Dataset

Trade ID	Exchange
101533	LSE
79052	Hong Kong

OPERATION

LEFT JOIN

INNER JOIN

OUTER JOIN

RIGHT JOIN

OUTPUT

LEFT JOIN

Trade ID	Asset Class	
101533	Fixed Income	LSE
606601	Commodities	

INNER JOIN

Trade ID	Asset Class	
101533	Fixed Income	LSE

OUTER JOIN

Trade ID	Asset Class	
101533	Fixed Income	LSE
606601	Commodities	
79052		Hong Kong

RIGHT JOIN

Trade ID	Asset Class	
101533	Fixed Income	LSE
79052		Hong Kong

Il·lustració 8: Exemples Joins

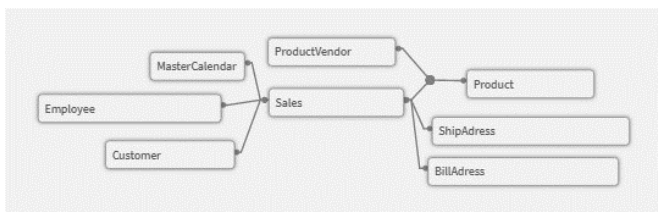
També hi ha altres funcions menors com: **Num** (interpreta els valors d'una cadena com a números), **Text**(com a text), **Date** (com a dates) etc..

2.3 Model Final

Inicialment, el model constava d'una base de dades amb 15 taules diferents, dos excels i un txt. En finalitzar totes les modificacions he acabat amb vuit taules.

En Qlik Sense és molt més òptim quant a rendiment tenir una taula central amb molts registres i pocs salts en altres taules, ja que al programa li costa molt més passar d'una taula a un altre que llegir molts registres de la mateixa taula. Per aquest motiu el meu model es basa en la taula de Sales; on s'emmagatzemen tots els moviments del negoci. Aquesta taula es connecta amb el calendari, la taula Employee (Empleats), que conté tots els noms i dades dels empleats de l'empresa, la taula Customers (Clients) que conté tota la informació dels clients, ShipAddress i BillAddress, que tal com he comentat amb anterioritat és on es guarden tant les adreces d'enviament amb totes les seves dades com les direccions de pagament i finalment les taules Product Vendor i Product que contenen els venedors de productes i les seves dades i els productes en vendes respectivament. Aquestes dues taules també estan unides entre elles òbviament, mitjançant els productes (Un producte és venut per X venedor i comprat per Y).

Aquesta és la imatge final del model simplificat:



Il·lustració 9: Model Final Simplificat

3 ESTAT DE L'ART

En l'actualitat hi ha moltes eines de business intelligence, les més rellevants són:

- Power BI: Aplicació gratuïta desenvolupada per Microsoft la qual ha entrat fa poc en el mercat i l'està revolucionant, gràcies a la seva portabilitat amb altres eines de Microsoft i el fet que sigui gratis.
- Qlik Sense: Plataforma líder del mercat desenvolupada per Qlik (successió de Qlik View), és l'eina més robusta i professional que hi ha.
- Tableau: Eina sòlida de BI, destaca per la seva quantitat de gràfics diferents, i la seva portabilitat per a dispositius tablets i mòbils.
- Looker: Eina nova de BI desenvolupada per Google. Se sol utilitzar en cloud, tot i que la seva potència no és comparable amb les anteriors, és força nova i li falta polir detalls.

Tenint aquests punts en compte, he decidit fer la comparativa entre Qlik Sense (plataforma líder del mercat des de fa més de 20 anys) amb Power BI (nova eina desenvolupada per una multinacional que està revolucionant el mercat). La principal diferència que es veu a primera vista és el preu; mentre que Power BI és una eina gratuïta, Qlik Sense és força car, i vull veure si val la pena el cost respecte a les millores que pot suposar respecte a Power BI.

4 BIBLIOGRAFIA

Ayuda de Qlik Inicio / Qlik Help. Available at: <https://help.qlik.com/es-ES/> (Accessed: 10 November 2023).

Consiga que sus Datos Tengan un efecto inmediato Visualización de datos / Microsoft Power BI. Available at: <https://powerbi.microsoft.com/es-es/> (Accessed: 11 November 2023).

Looker studio: Business insights visualizations / google cloud Google. Available at: https://cloud.google.com/looker-studio?gad_source=1&gclid=Cj0KCQiAjMKqBhCgARIsAP-DgWlweZfZWqCwEdrGqxR6YJAooccl8dH4-2T8wJUHg3PuWArG62Gr2E5IaAgZi-EALw_wcB&gclsrc=aw.ds (Accessed: 11 November 2023).

Software de análisis e inteligencia de negocios Tableau. Available at: <https://www.tableau.com/es-es> (Accessed: 11 November 2023).

Support: Qlik community Qlik. Available at: <https://community.qlik.com/t5/Support/ct-p/qlikSupport> (Accessed: 10 November 2023).