
This is the **published version** of the bachelor thesis:

Camprubí Pujol, Arnau; Oropesa Física, Ana, dir. Atmira Pharma Visualization.
2022. (1394 Enginyeria de Dades)

This version is available at <https://ddd.uab.cat/record/264623>

under the terms of the  license

Atmira Pharma Visualization

Arnau Camprubí Pujol

Resum— Atmira Pharma Visualization és un repte que presenta Cajamar UniversityHack 2022. Aquest repte consisteix en crear la millor aplicació i/o visualització de dades anonimitzades facilitades per Atida Mifarma, una empresa de comerç electrònic líder en Espanya i Portugal en la venda de productes de parafarmàcia i farmàcia. Així doncs, l'objectiu general del TFG és donar una visualització útil i completa per l'empresa Atida Mifarma que li permeti conèixer de forma visual dades rellevants pel seu dia a dia, a més d'ajudar-la en la seva missió, visió i valors. A més, s'han realitzat dues recomanacions a l'empresa per tal que els seus clients siguin més proactius i ajudar al seu creixement.

Paraules clau— Visualització, parafarmàcia, clients, empresa, creixement, comerç electrònic, aplicació.

Abstract— Atmira Pharma Visualization is a challenge presented by Cajamar UniversityHack 2022. This challenge consists in creating the best application and/or visualization of anonymized data provided by Atida Mifarma, a leading e-commerce company in Spain and Portugal in the sale of parapharmacy products and pharmacy. Therefore, the general objective of this Final Degree Project is to provide a useful and complete visualization for the company Atida Mifarma that allows them to visualize relevant data for their day to day, as well as help them in their mission, vision and values. In addition, two recommendations have been made to the company to make its customers more proactive and help them grow.

Index Terms— Visualization, parapharmacy, customers, business, growth, e-commerce, application.

1 INTRODUCCIÓ - CONTEXT DEL TREBALL

En aquest apartat es parla del context del treball, dels objectius i de la metodologia utilitzada.

1.1 Context del treball

Cajamar UniversityHack 2022 és un esdeveniment dirigit específicament als alumnes dels millors centres formatius de *Data Science* d'Espanya [1]. Dels dos reptes que presenten aquest any, s'ha escollit participar en el repte Atmira Pharma Visualization, per, a més a més de participar en l'esdeveniment, fer-lo servir com a projecte de Treball de Fi de Grau (TFG).

El repte Atmira Pharma Visualization té com a *Data Partner* Atida Mifarma, una empresa de comerç electrònic líder en Espanya i Portugal en la venda de productes de parafarmàcia i farmàcia [2] [3]. Amb seu a Albacete i Madrid, més de 10 anys d'experiència al sector i un equip de més de 120 professionals, ofereix una experiència fiable i personalitzada abans, durant i després del procés de compra.

L'any 2019 Mifarma es va unir a Atida amb l'objectiu de convertir-se en la plataforma de salut holística *online* més gran a Europa. D'aquesta manera, treballen dia a dia per construir un ecosistema en línia, l'objectiu del qual és transformar el panorama de salut i benestar, convertint-se

en la farmàcia en línia més gran d'Europa.

En resum, Atmira Pharma Visualization consisteix en crear la millor aplicació i/o visualització de dades anonimitzades facilitades per Atida Mifarma, amb la informació de negoci del període transcorregut entre el gener de 2017 i el desembre de 2018.

1.2 Objectius

L'objectiu general del meu TFG és donar una visualització útil i completa per l'empresa Atida Mifarma que li permeti conèixer de forma visual dades rellevants pel seu dia a dia, a més d'ajudar-la en la seva missió, visió i valors.

Com a objectius específics es plantegen els següents:

1. Realitzar una investigació sobre l'empresa Atida Mifarma per a conèixer a què es dedica, quin són els seus interessos i valors, quin perfil de clients té i quin és el seu objectiu i missió a llarg termini.
2. Analitzar els 3 *datasets* que Cajamar UniversityHack 2022 posa a la meua disposició per tal de conèixer de què estan compostats i quins són els atributs útils pels interessos de l'empresa.
3. Localitzar els punts estratègics de l'empresa. Per exemple, amb quins productes guanyen més diners, quins productes tenen més rotació de stock, des de quina zona compren més els clients, quins són els descomptes que més funcionen, etc.
4. Mostrar aquests punts estratègics i crear tendències en relació a ells per analitzar-los i crear futures accions.
5. Proposar suggerències per tal que els clients

- E-mail de contacte: arnaucamprubipujol@gmail.com
- Treball tutoritzat per: Ana Oropesa (DEIC)
- Curs 2021/22

d'Atida Mifarma siguin més proaciuts.

6. Proposar suggerències per tal d'ajudar a Atida Mifarma a ser la farmàcia online més gran d'Europa.
7. Avaluar si les dades mostrades són útils pels interessos de l'empresa.

1.3 Metodologia i planificació

La metodologia utilitzada és Kanban, un mètode visual de gestió de projectes que permet aconseguir un procés organitzat, productiu i eficient a l'hora de realitzar les diferents tasques [4].

L'eina que s'ha utilitzat per a planificar totes les tasques és un diagrama de Gantt [5] [6]. El diagrama de Gantt es pot veure referenciat a la secció d'apèndix 1. Per a desenvolupar-lo s'ha utilitzat l'aplicació Lucid App [7].

2 ESTAT DE L'ART

En aquest apartat es parlarà de la investigació sobre Atida Mifarma i de les eines utilitzades.

2.1 Investigació sobre Atida Mifarma

Atida Mifarma és una empresa de comerç electrònic líder en Espanya i Portugal en la venda de productes de parafarmàcia i farmàcia. La companyia ofereix un ampli catàleg de productes per a la cura i el benestar a través d'un servei ràpid i personalitzat recolzat per experts farmacèutics.

Amb més de 15.000 comandes, més de 4.000 clients actius, més de 700 treballadors, 16 anys d'experiència, 7 oficines i operant en més de 9 països, diferents, Atida Mifarma ofereix una experiència fiable i personalitzada abans, durant i després del procés de compra [8].

Els valors que defineixen a Atida Mifarma són: ètica, empatia, humilitat, honestat, curiositat, valentia i diversió.

Atida Mifarma vol ajudar als seus clients a aconseguir ser la millor versió d'ells mateixos, ajudant-los a prendre les millors decisions respecte a la salut. Creuen en un futur on la pròpia salut i benestar s'han de convertir en una experiència personal i de confiança. El seu somni és aconseguir que el benestar passi d'un sentit més curatiu i reactiu a una cosa més proactiva, on s'inclouï la prevenció.

L'any 2019 Mifarma es va unir a Atida amb l'objectiu de convertir-se en la plataforma de salut holística online més gran a Europa. D'aquesta manera, treballen dia a dia per construir un ecosistema en línia, l'objectiu del qual és transformar el panorama de salut i benestar, convertint-se en la farmàcia en línia més gran d'Europa.

2.2 Eines utilitzades

Les eines utilitzades es divideixen en les eines per a fer *Data Cleaning* i les eines per a fer les visualitzacions.

2.2.1 Eines utilitzades per a fer *Data Cleaning*

Com que les bases de dades no tenen el format que interessa, cal que realitzem la neteja de dades o *Data Cleaning*.

Tableau Prep Builder és una eina de Tableau dissenyada per a preparar dades de forma fàcil i intuïtiva que permet combinar dades, donar-los forma i netejar-los per a l'anàlisi [9] [10]. S'utilitzarà aquest software per el procés de *Data Cleaning*.

2.2.2 Eines utilitzades per a fer les visualitzacions

Aprofitant que es té experiència prèvia amb Tableau, es pensa que és una bona idea fer-lo servir per a fer les visualitzacions.

Tableau és un software de *Business Intelligence (BI)* que serveix per analitzar i visualitzar dades. Permet a les persones i les organitzacions treure el màxim partit de les seves dades [11] [12] [13] [14].

Power BI també és un software potent per a proporcionar visualitzacions interactives de dades, però el fet de ja tenir unes bases amb Tableau és el que es fa decantar per aquesta eina [15].

3 DESENVOLUPAMENT

En aquesta secció es parlarà del desenvolupament del treball.

3.1 Anàlisi dels 3 *datasets* inicials

A continuació es troba la descripció de cada *dataset*:

1. 'items_ordered_2years.txt', al que es referirà com a *dataset* 'sales'.
Dataset en format TXT separat per barres verticals ("|") que fa servir el punt com indicador decimal en les variables numèriques.
Inclou la informació de les ventes realitzades a través de la web des del gener de 2017 al desembre del 2018 a Espanya.
2. 'products.csv', al que es referirà com a *dataset* 'products'.
Dataset en format CSV (separats per comes) que fa servir el punt com indicador decimal en les variables numèriques.
Presenta la informació del catàleg de productes disponible a la web.
3. 'products_cat.csv', al que es referirà com a *dataset* 'products_cat'.
Dataset en format CSV (separats per comes) que fa servir el punt com indicador decimal en les variables numèriques.
Conté la classificació dels productes a diferents nivells.

Per a relacionar els tres *datasets* s'ha creat relacions, que descriuen com es relacionen dues taules (o *datasets*) entre elles, bastant-se en els camps que tenen en comú i sense fusionar-les.

Al tenir només tres *datasets*, es creen dues connexions:

- Connexió entre *datasets* 'sales' i 'products' relacionats pel camp 'product_id'.
- Connexió entre *datasets* 'products' i 'products_cat' relacionats pel camp 'sku'.

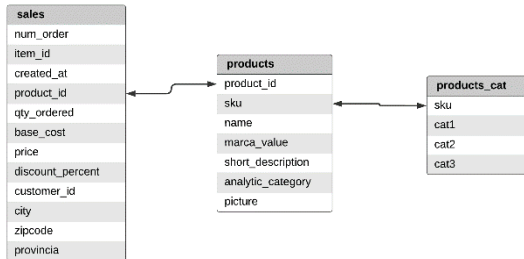


Figura 1 – Diagrama entitat-relació dels 3 datasets

3.2 Data Cleaning – Tableau Prep Builder

Per a fer la neteja de dades s'han utilitzat diferents operacions repetides vegades. Aquestes són les operacions més rellevants:

- Canviar el tipus de dades d'alguns camps d'enter a cadena, com per exemple en el cas de 'num_order', ja que en algunes ocasions és una cadena de caràcters, i en canvi interessa tenir tots els camps numèrics per a fer futures comparacions correctament.
- Agrupar valors com per exemple agrupar les categoritzacions de primer nivell 'nutrición' i 'Nutrición', ja que tot i tractar-se del mateix, una comença amb majúscula i l'altre en minúscula.
- Segmentar les files del *dataset* 'products_cat' on alguns 'SKU' tenen més informació de la necessària. Per exemple, s'ha segmentat el 'SKU' = [001268,Salud,"Huesos, articulaciones y músculos",Colágeno], de forma que el 'SKU' sigui 001268 i la resta les 3 categoritzacions corresponents.
- Eliminar columnes.

Ex: eliminar les columnes no necessàries del *dataset* 'provincias'.

El diagrama complet que s'ha realitzat amb Tableau Prep Builder pel *Data Cleaning* es pot veure referenciat a la secció d'apèndix 2.

3.3 Utilització de *datasets* externs

S'ha utilitzat dos *datasets* externs, és a dir, *datasets* que tot i no formar part dels inicials, es pensa que aporten informació rellevant pel desenvolupament del treball.

3.3.1 *Dataset* extern 1 – Provincies

El fet de que hi hagi més de 8.000 ciutats a Espanya dificulta totes les visualitzacions relacionades amb la localitat. Es pensa que si es tingués una nova columna indicant la província corresponent de cada ciutat es podrien controlar més fàcilment les zones més populars des d'on es realitzen compres, ja que de províncies a Espanya n'hi ha només 52.

És per això que s'ha utilitzat un *dataset* de INE que re-

laciona els dos dígitos inicials dels codis postals d'Espanya amb la província corresponent [16].

Per a poder ajuntar els codis postals del *dataset* 'sales' amb els codis postals del *dataset* 'provincias', s'ha hagut de crear una nova columna de 'sales' que inclou només els dos primers dígitos del codi postal corresponent. D'aquesta manera es poden ajuntar ambdós *datasets* sense problemes.

3.3.2 *Dataset* extern 2 – Malalties cròniques

Es pensa que les malalties cròniques són malalties de llarga duració i progressió generalment lenta. Les persones que pateixen aquestes malalties acostumen a utilitzar diferents productes farmacèutics constantment, per tant, es pensa que seria interessant poder oferir productes relacionats amb aquestes malalties per a augmentar la proactivitat dels clients.

S'ha utilitzat un *dataset* per a consultar les malalties cròniques més comunes a Espanya durant el 2020 [17]. Simplement s'ha ordenat els percentatges de major a menor, per a veure les malalties més comunes.

D'aquest *dataset* s'extreuen les 6 malalties cròniques més comunes a Espanya durant el 2020 (ordenades de més a menys comú):

- Tensió alta.
- Colesterol alt.
- Artrosi (excloent artritis).
- Mal d'esquena crònic (lumbar).
- Mal d'esquena crònic (cervical).
- Al·lèrgia crònica com rinitis, conjuntivitis o dermatitis al·lèrgica, al·lèrgia alimentària o d'un altre tipus.

3.4 Estat de l'empresa. Primer contacte amb les dades

S'ha creat diferents gràfics dels *datasets* per a familiaritzar-se amb ells i conèixer la informació bàsica de l'empresa Atida Mifarma. A continuació s'expliquen tant quins són els gràfics creats de cada *dataset*, com els resultats que s'extreuen d'ells.

3.4.1 *Dataset* 'sales'

Mapa de calor que mostra les províncies que s'han gastat més diners al llarg dels dos anys.

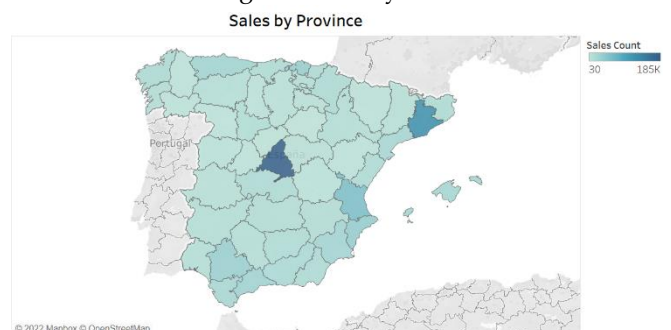


Figura 2 – Sales by Province 1

Es veu clarament que les províncies que s'han gastat més diners són Madrid i Barcelona.

Mapa de calor que mostra l'evolució dels diners que s'ha gastat cada província mes a mes. Es pot seleccionar un mes en concret per tal de veure'l en detall.

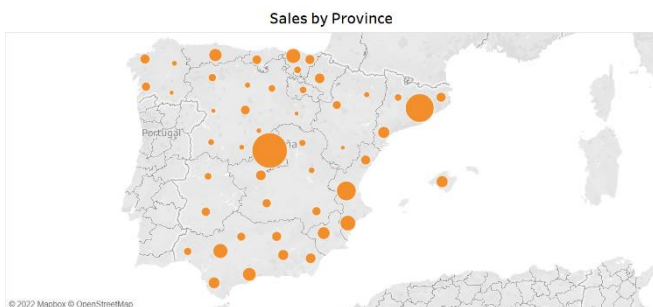


Figura 3 – Sales by Province 2

Es pot veure que tot i en els mesos més baixos de facturació, Madrid i Barcelona segueixen sent les províncies que més diners es gasten.

Gràfic de línies que mostra la facturació mensual de l'empresa respecte el cost base des del primer a l'últim mes.

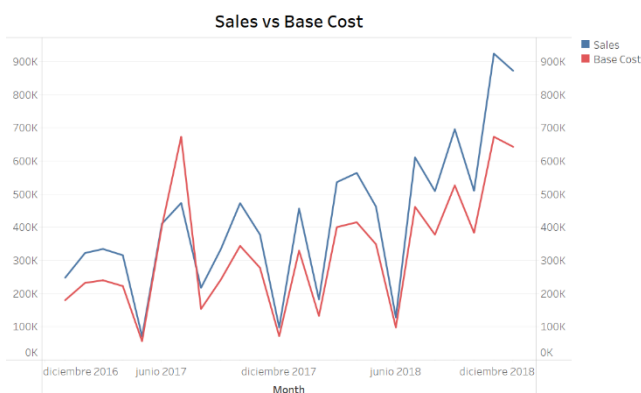


Figura 4 – Sales vs Base Cost

És d'utilitat per a veure ràpidament que el juliol de 2017 l'empresa va tenir un benefici negatiu. També es veu que els dos mesos on el benefici ha sigut major han estat els dos darrers mesos (novembre i desembre de 2018).

Gràfic de línies que mostra el benefici mensual de l'empresa des del primer a l'últim mes.



Figura 5 – Profit by Month

Destaca que mentre el juliol del 2017 el benefici va ser de gairebé -200.000€, el juliol de l'any següent el benefici va ser d'aproximadament 150.000€.

Gràfic de línies múltiples que mostra el benefici de cada categoria de productes des del primer a l'últim mes. Es permet filtrar el gràfic per mesos, de manera que es pugui veure el benefici de cada categoria més clarament de qualsevol mes en concret.

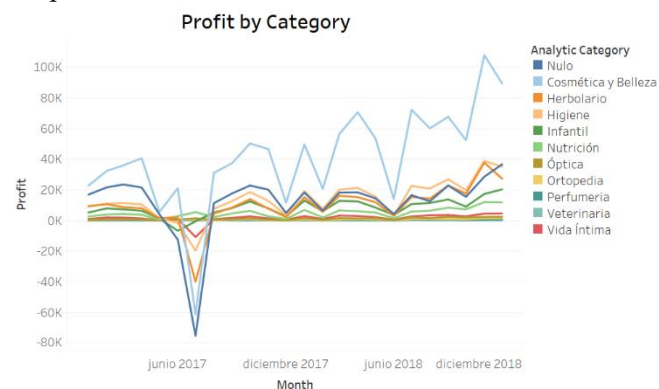


Figura 6 – Profit by Category

Es pot observar que el juliol de 2017, el pitjor mes en relació a la facturació, la categoria que més va generar beneficis va ser la de 'Nutrición'. En canvi, el juliol del 2018, la categoria més beneficiosa va ser sense cap dubte la de 'Cosmética y Belleza'.

Gràfic de línies múltiples que mostra el benefici de l'any 2017 respecte al del 2018.

Es permet filtrar el gràfic per data, per tant, es pot comparar fàcilment el benefici d'un any respecte l'altre en un període de temps en concret.

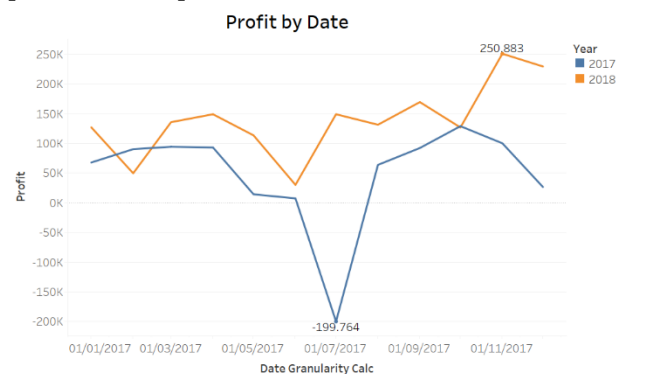


Figura 7 – Profit by Date

S'observa que el 2018 en general va ser un any amb més beneficis que el 2017.

3.4.2 Dataset 'products'

Gràfic de barres horitzontals que mostra les marques que inclouen més productes. Es mostren el top 5% de les marques que inclouen més productes.

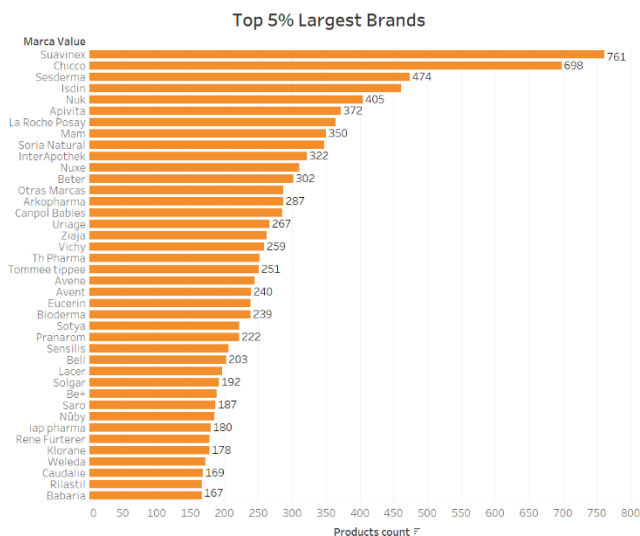


Figura 8 – Top 5% Largest Brands

Es pot veure que la marca més popular de l'empresa és Suavinex.

Gràfic circular que mostra les categories que contenen més productes.

Top Analytic Categories

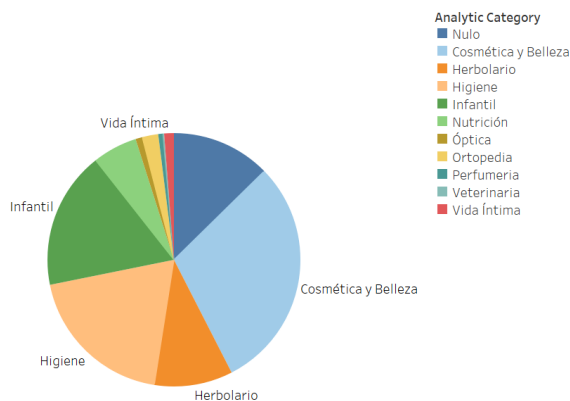


Figura 9 – Top Analytic Categories

Es pot veure que la categoria que inclou més productes és 'Cosmética y Belleza', seguida de 'Higiene'.

Gràfic de barres apilades que mostra el benefici respecte la facturació de cada categoria de productes en un mes en concret.

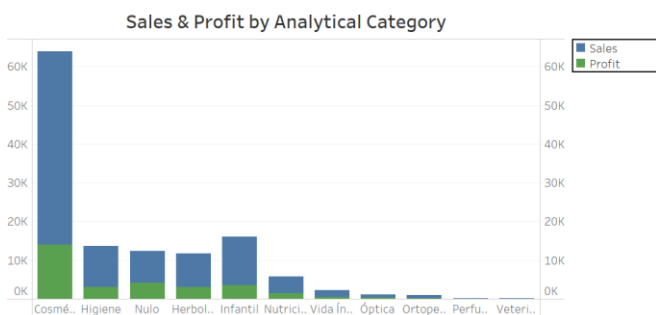


Figura 10 – Profit by Category

Es permet filtrar el gràfic per mesos, de manera que es pugui veure el benefici respecte la facturació de cada categoria de qualsevol mes en concret.

3.4.3 Dashboards

Combinació dels gràfics 'Figura 7 - Profit by Date' i 'Figura 10 - Profit by Category'.

Es permet obtenir informació sobre les categories de productes que més benefici han generat simplement fent un click al mes que interessa. En aquest exemple, el mes seleccionat es desembre de 2018.



Sales & Profit by Analytical Category

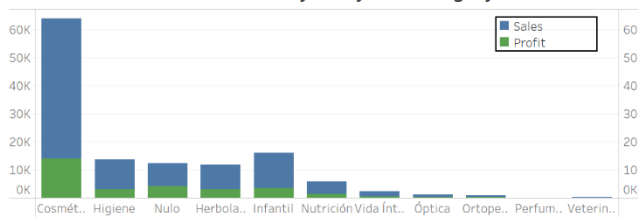


Figura 11 – Dashboard (Profit by Date & Profit by Category)

Es pot veure que el desembre de 2018, la categoria que més benefici va provocar a l'empresa va ser 'Cosmética y Belleza'.

3.5 Resum de l'estat de l'empresa

Com a resum, de tots els gràfics realitzats en aquest apartat n'extreiem la següent informació:

- Les províncies des de les que es recauden més diners són Madrid, Barcelona, València i Alicante.
- El juliol de 2017 l'empresa va tenir un benefici negatiu.
- Els dos mesos on el benefici ha sigut major han estat novembre i desembre de 2018.
- El juliol del 2017 el benefici va ser de gairebé -200.000€ i la categoria que més va generar beneficis va ser la de 'Nutrición'. En canvi, el juliol de l'any següent el benefici va ser d'aproximadament 150.000€ i la categoria més beneficiosa va ser sense cap dubte la de 'Cosmética y Belleza'.
- 'Cosmética y Belleza' és la categoria que més es ven, la que més productes conté i la que més beneficis genera.
- En general, el 2018 va ser un any amb més beneficis que el 2017.

3.6 Gràfics avançats. Aprofundiment en l'empresa

S'ha creat diferents gràfics avançats que aprofundixen en l'empresa Atida Mifarma. A continuació

s'expliquen tant quins són els gràfics creats, com els resultats que s'extreuen d'ells.

Gràfic de línies múltiples que mostra la comparació de les vendes i el cost base respecte el promig dels descomptes de cada mes.

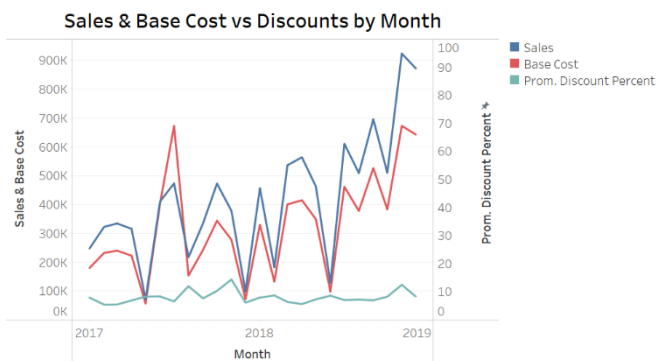


Figura 12 – Sales & Base Cost vs Discounts by Month

Es pot veure que generalment, quan hi ha pics d'ofertes les vendes no augmenten. Això significa que no hi ha un aparent relació entre els descomptes i les vendes.

Gràfic compost relacionat amb el creixement anual. En el primer d'ells, s'observa l'acomulació de les vendes del 2018 respecte el 2017 per mesos. En el segon, es pot consultar quina és la diferència acumulada de vendes per mesos.

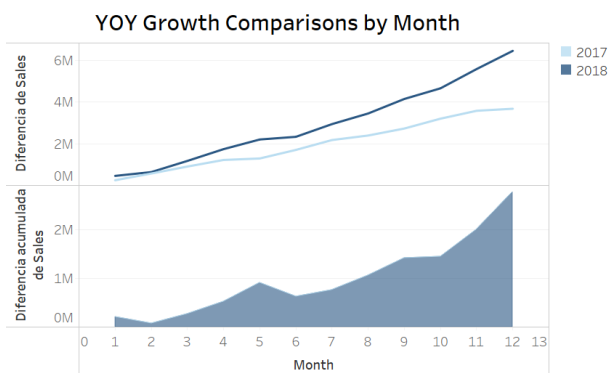


Figura 13 – YOY Growth Comparisons by Month

Analitzant els dos gràfics, es veu que durant 2018 hi va haver més vendes que durant 2017. És a partir de juny del 2018 quan l'acomulació de vendes del 2018 comença a ser més gran, augmentant ràpidament en els mesos següents.

Gràfic de línies per a veure el creixement anual de les vendes per mesos.

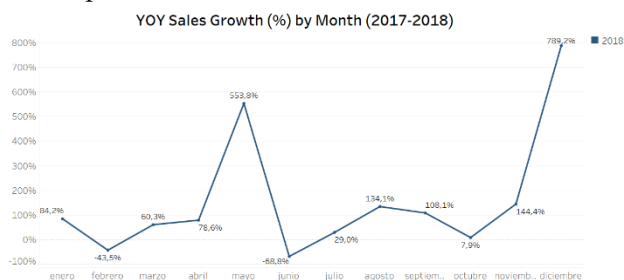


Figura 14 – YOY Sales Growth (%) by Month

D'aquest gràfic n'extreiem informació molt valuosa: que maig de 2018 i desembre de 2018 són dos mesos amb un creixement de vendes superior al 500% respecte l'any anterior.

Per a veure si el benefici de l'empresa també ha augmentat en aquests dos mesos, s'ha creat aquest gràfic de línies, que és igual que l'anterior, però utilitza el benefici com a mesura a comparar respecte el 2017.

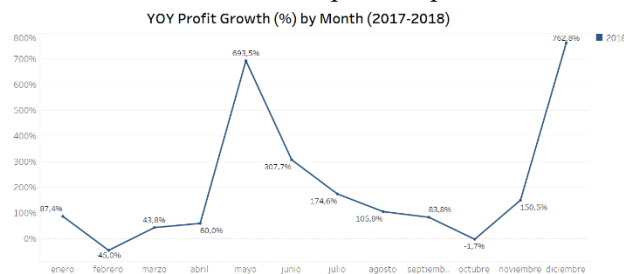


Figura 15 – YOY Profit Growth (%) by Month

Es veu que segueixen destacant tant maig de 2018 i desembre de 2018 com a mesos amb més creixement de benefici d'Atida Mifarma.

3.7 Resum dels gràfics avançats

Com a resum, de tots els gràfics realitzats en aquest apartat n'extreiem la següent informació:

- No hi ha un aparent relació entre els descomptes i les vendes.
- És a partir de juny del 2018 el mes en el que l'acomulació de vendes del 2018 comença a ser més gran, augmentant ràpidament en els mesos següents.
- Maig de 2018 i desembre de 2018 són dos mesos amb un creixement de vendes superior al 500% respecte l'any anterior.

4 SOLUCIÓ

En aquesta secció es parlarà de la solució dels dos objectius plantejats.

4.1 Objectiu 1 – Clients més proactius

En aquest apartat es descriu tot el procés realitzat per a proposar suggerències per tal que els clients d'Atida Mifarma siguin més proactius.

Tal i com s'ha comentat a l'apartat 3.3.2, s'ha extret d'un dataset extern, les 6 malalties cròniques més comunes a Espanya durant el 2020:

- Tensió alta.
- Colesterol alt.
- Artrosi (excloent artritis).
- Mal d'esquena crònic (lumbar).
- Mal d'esquena crònic (cervical).
- Al·lèrgia crònica com rinitis, conjuntivitis o dermatitis al·lèrgica, al·lèrgia alimentària o d'un altre tipus.

Seguidament, s’han relacionat les 6 malalties cròniques amb paraules clau, que serveixen per a buscar tots els productes relacionats amb cada malaltia.

A continuació, s’ha realitzat la cerca en el camp ‘short_description’ del dataset products, de manera que cada malaltia crònica queda associada amb tots aquells productes que en la seva descripció inclouen alguna de les paraules clau identificades anteriorment. Per exemple, la malaltia ‘alergia’ s’ha associat amb les paraules clau ‘rinitis’, ‘conjuntivitis’ i ‘dermatitis’.

Seguidament, s’ha realitzat un dashboard combinant dos gràfics. El primer és un gràfic de barres que mostra els 10 productes que més s’han venut per cada una de les malalties. El segon mostra informació relacionada amb productes, de manera que cada vegada que es fa click en un dels productes del primer gràfic, el segon en mostra la seva categoria i descripció.

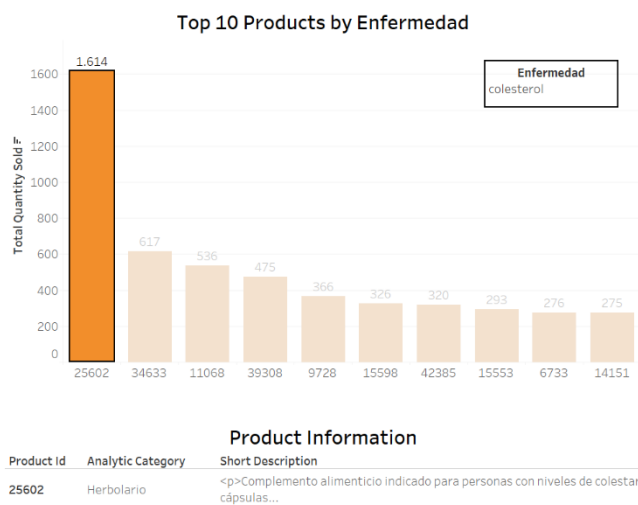


Figura 16 – Dashboard (Top 10 Products by Enfermedad & Product Information)

Com a proposta per tal que els clients d’Atida Mifarma siguin més proactius, es suggereix que l’empresa segueixi els següents passos:

1. Quan es detecta que un client compra 5 productes relacionats amb una malaltia crònica, s’associa que el client la pateix.
2. D’aquesta manera, es pot afegir un sistema de recomanacions a la pàgina web en el que es facin, a cada client associats a una malaltia, suggerències dels 10 productes associats a aquesta.

D’aquesta manera, els clients rebrien suggerències de productes del seu interès i s’aconseguiria que fossin més proactius.

4.2 Objectiu 2 – Farmàcia en línea més gran d’Europa

En aquest apartat es descriu tot el procés realitzat per a proposar suggerències per tal d’ajudar a Atida Mifarma a ser la farmàcia online més gran d’Europa.

A l’apartat 3.6 s’ha vist que els dos mesos de més crei-

xement d’Atida Mifarma respecte el 2017 van ser maig i desembre de 2018. Aquest és el motiu pel que a continuació, es realitza una investigació per a consultar les ventes i el profit d’ambdós mesos i els productes venguts respectivament.

4.2.1 Sales i profit

En aquest apartat s’analitzen les ventes i el profit de maig de 2017 i 2018.

4.2.1.1 Anàlisis de ventes i profit de maig

Dashboard en el que es poden comprar en 2 gràfics circulars, quines categories de productes es van vendre més en maig de 2017 respecte en maig de 2018.

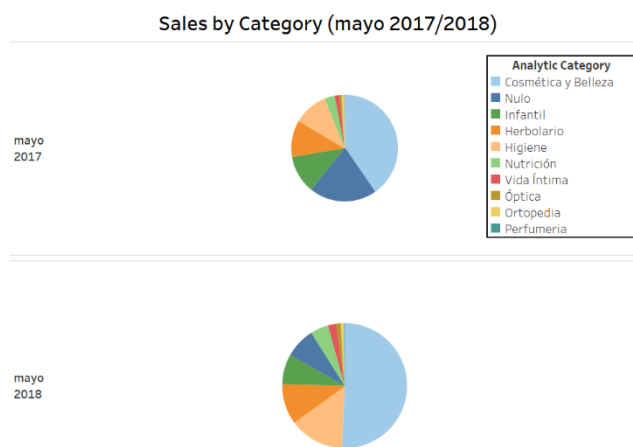


Figura 17 – Sales by Category (mayo 2017/2018)

Es pot veure que ‘Cosmética y Belleza’ és la categoria més venguda en ambdós mesos. Tot i que ‘Higiene’, ‘Herbolario’ i ‘Infantil’ ocupin aproximadament el mateix percentatge, el maig de 2017 es ven vendre bastants productes de categoria ‘Nula’.

Dashboard en el que es poden comparar en 2 gràfics circulars, quines categories de productes van tenir més benefici en maig de 2017 respecte en maig de 2018.

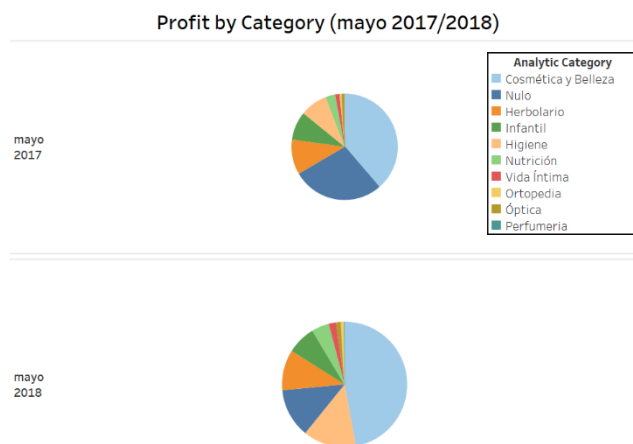


Figura 18 – Profit by Category (mayo 2017/2018)

tes l'empresa hi guanya diners. A més, dels 15 productes que més es van vendre en desembre de 2018, n'hi ha 5 que no es van vendre en maig de 2017.

És per això que seria interessant veure els productes venuts en desembre de 2018 i no en desembre de 2017.

4.2.3 Comparació de preus de venda unitaris i costos base unitaris

En aquest apartat es comparen els preus de venda unitaris i el cost base unitari de cadascun dels productes que es van vendre en maig i desembre de 2018 respecte 2017.

4.2.3.1 Comparació de preus de venda unitaris i costos base unitaris durant maig

Gràfic de barres múltiples que compara el preu de venda unitari i el cost base unitari de cadascun dels productes que es van vendre en maig de 2018 respecte a maig de 2017.

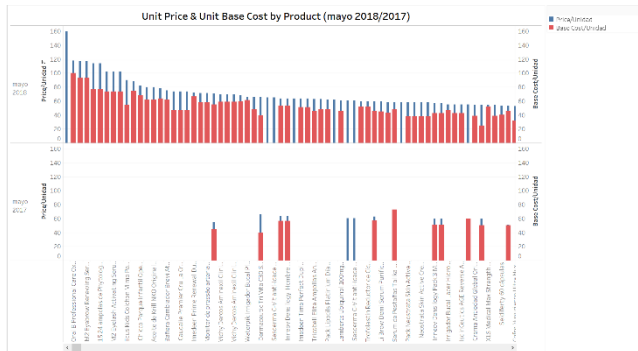


Figura 23 – Unit Price & Unit Base Cost by Product (mayo 2018/2017)

Es veu que en maig de 2018 es van vendre molts més productes diferents que en maig de 2017. A més, veiem que no hi ha hagut cap reducció de costos base, és a dir, que els productes segueixen costant el mateix, però ara es venen en més quantitat.

4.2.3.2 Comparació de preus de venda unitaris i costos base unitaris durant desembre

Gràfic de barres múltiples que compara el preu de venda unitari i el cost base unitari de cadascun dels productes que es van vendre en desembre de 2018 respecte a desembre de 2017.

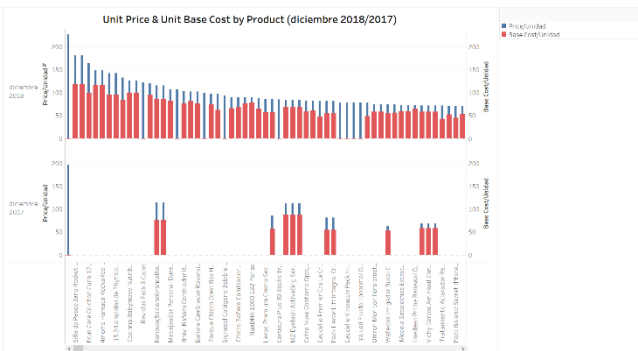


Figura 24 – Unit Price & Unit Base Cost by Product (diciembre 2018/2017)

Es pot veure clarament que en desembre de 2018 es van vendre molts més productes diferents que en desembre de 2017. A més, veiem que no hi ha hagut cap reducció de costos base, és a dir, que els productes segueixen costant el mateix, però ara es venen en més quantitat.

4.2.4 Recompte de productes diferents per mesos

En aquest apartat s'analitzen els productes diferents que s'han venut per mesos.

Gràfic de línies que mostra el recompte de productes diferents que hi ha hagut en cada mes.

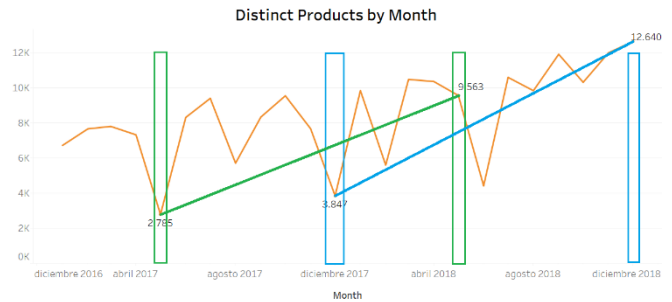


Figura 25 – Distinct Products by Month, donde los rectángulos verdes equivalen a mayo y los azules a diciembre

Tal i com es pot veure, tant maig de 2017 com desembre de 2017 son els dos mesos en els que el catàleg de productes era menys ampli. En canvi, en els seus respectius mesos del any següent, el catàleg de productes es va ampliar més de tres vegades.

4.2.5 Resolució de l'objectiu 2

En aquest apartat s'explica la solució de l'objectiu 2, és a dir, proposar suggerències per tal d'ajudar a Atida Mifarma a ser la farmàcia online més gran d'Europa.

S'ha vist que l'empresa no s'ha diversificat en categories, és a dir, que ha seguit vendent productes de les mateixes categories, però en molta més quantitat i varietat.

La raó per la qual Atida Mifarma va créixer tant en maig i desembre de 2018 és perquè els productes disponibles en la seva pàgina web van augmentar dràsticament d'un any a l'altre, oferint així als seus clients més productes.

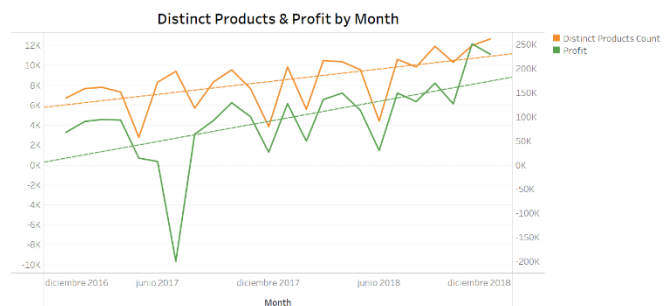


Figura 26 – Distinct Products & Profit by Month

Tot i això, en aquest gràfic podem observar que el benefici no creix al mateix temps que els productes diferents venuts. Això significa que s'ha d'anar en compte ja que

vendre més productes diferents, tot i ser el factor diferencial que fa créixer més a l'empresa en ventes, no és una cosa que pugui perdurar en el temps. És per això que Atida Mifarma haurà de prendre altres accions en un futur com les que es comten a continuació.

Per tal d'ajudar a Atida Mifarma a ser la farmàcia online més gran d'Europa, es suggereix que l'empresa segueixi les següents propostes:

1. Eliminar del catàleg els productes que no es venen en certa mesura.
2. Analitzar la rotació que hi ha en el mercat per a trobar futurs productes a afegir al catàleg.
3. Créixer en la quantitat de productes venguts diferents sempre i quan no s'inclouguin en productes sense rotació.

5 CONCLUSIONS

En aquest apartat s'expliquen les conclusions extretes després de tot el treball realitzat.

Primer de tot, s'ha assolit tots els objectius plantejats, tant els principals com els específics.

S'ha vist que és important fer *Data Cleaning*, ja que permet identificar dades incompletes, incorrectes, inexactes, no pertinents, etc. i després substituir, modificar o eliminar aquestes dades. Per exemple, s'ha agrupat valors de les categoritzacions de primer nivell 'nutrició' i 'Nutrició'.

També s'ha descobert que els gràfics relacionats amb el creixement anual són gràfics empresarials que la majoria d'empreses entenen, i per tant, que són útils en el món empresarial. S'han de fer servir perquè l'empresa ho sap interpretar sense cap explicació.

Conèixer al client per qui es treballa és molt important perquè sinó no s'hagués pogut respondre a les necessitats o les voluntats de l'empresa.

És clau saber que una empresa no té sempre tot el coneixement. Permetre la utilització de *datasets* externs ha donat moltes possibilitats d'explorar noves dades i poder fer recomanacions com l'associació de malalties cròniques amb els clients.

A l'hora de realitzar qualsevol treball sempre s'ha de reservar temps pels imprevistos. Sense aquest temps no es podria a ver fet un anàlisi tan profund en l'estat inicial.

6 TREBALL FUTUR

En aquest apartat es proposen les futures millores que es podrien dur a terme.

Primer de tot, analitzar les malalties estacionals i crear un altre sistema de recomanació de productes que canvia-

ran per cada estació de l'any.

Seria interessant comparar les dades de venda amb les d'altres empreses de farmàcia Europees.

Finalment, analitzar dies de ventes puntuals com per exemple *Black Friday*, per tal de comprovar si són dies on es ven més o menys.

AGRAÏMENTS

M'agradaria agrair a la meva tutora Ana Oropesa, a la meva família i als meus amics que han suposat en tot moment un suport moral molt important.

BIBLIOGRAFIA

- [1] Esdeveniment Cajamar UniversityHack 2022 | <https://www.cajamardatalab.com/datathon-cajamar-universityhack-2022/>
- [2] Repte Atmira Pharma Visualization | <https://www.cajamardatalab.com/datathon-cajamar-universityhack-2022/retos/atmira-pharma-visualization/>
- [3] Empresa Atida Mifarma | <https://www.atida.com/es-es>
- [4] Informació sobre la metodologia Kanban | <https://www.esan.edu.pe/conexion-esan/los-principios-de-kanban-para-la-gestion-de-proyec- tos#:~:text=El%20m%C3%A9todo%20Kanban%20se%20enfoca, un%20C3%A1rea%20de%20la%20empresa>
- [5] Informació sobre els diagrames de Gantt | <https://saasrank.es/diagrama-de-gantt-ventajas-y-desventajas-planificacion-de-proyectos/>
- [6] Informació sobre els diagrames de Gantt | <https://economipedia.com/definiciones/diagrama-de-gantt.html#:~:text=El%20diagrama%20de%20Gantt%20es,inde pendiente%20dentro%20de%20un%20proceso>
- [7] Software per a crear el diagrama de Gantt | <https://lucid.app/>
- [8] Investigació sobre Atida Mifarma | <https://about.atida.com/es/>
- [9] Tableau Prep Builder | <https://www.tableau.com/es-es/products/prep>
- [10] Informació sobre Tableau Prep Builder | https://help.tableau.com/current/prep/es-es/prep_about.htm#:~:text=Tableau%20Prep%20Builder%20es%20una,para%20su%20an%C3%A1lisis%20en%20Tableau
- [11] Tableau | <https://www.tableau.com/es-es>
- [12] Informació sobre Tableau | https://www.youtube.com/watch?v=TPMIzRRaBQ&ab_cha nnel=freeCodeCamp.org
- [13] Informació sobre Tableau | https://www.youtube.com/watch?v=LRgupZ6B_Zg
- [14] Unir *datasets* en Tableau | https://www.youtube.com/watch?v=Xbg2WabUIK&ab_cha nnel=PenguinAnalytics
- [15] Power BI | <https://powerbi.microsoft.com/es-es/>
- [16] *Dataset* províncies | https://www.ine.es/daco/daco42/codmun/cod_provincia.htm
- [17] *Dataset* malalties cròniques | https://www.ine.es/jaxi/Tabla.htm?path=/t00/mujeres_hombres/tablas_1/&file=d03005.px

A2. SECCIÓ D'APÈNDIX

