



ÁRUFÁVAROZÓ CÉG ÜZLETI FOLYAMATAINAK TÁMOGATÁSA

SUPPORTING THE BUSINESS PROCESSES OF A FREIGHT FORWARDING COMPANY

Óbudai Egyetem
Neumann János Informatikai Kar
Kiberfizikai Rendszerek Intézet

SZAKDOLGOZAT FELADATLAP

Hallgató neve:	Horváth Dávid
Törzskönyvi száma:	T/005660/FI12904/N
Neptun kódja:	

A dolgozat címe:

Árufuvarozó cég üzleti folyamatainak támogatása
Supporting the business processes of a freight forwarding company

Intézményi konzulens:	Dr. Holyinka Péter
Külső konzulens:	

Beadási határidő:	2022. május 15.
-------------------	-----------------

A záróvizsga tárgyai:	Számítógép architektúrák Big Data és üzleti intelligencia specializáció
-----------------------	---

A feladat

Ismerje meg egy árufuvarozással foglalkozó cég működését, a jövőbeli döntések támogatási igényeit. Vizsgálja meg a jövőbeli döntések megkönnyítésének-, valamint az ügyfelek számára könnyebb elérés biztosításának lehetőségeit. Határozza meg a célokat kielégítő rendszerrel szemben támasztott követelményeket, és a fejlesztéshez szükséges eszközöket. Készítse el a rendszertervet, és fejlessze ki a rendszert. Tervezzon meg egy weboldalt, mely kellően kiszolgálja a megrendelők igényeit.

A dolgozatnak tartalmaznia kell:

- a vállalkozás bemutatását és aktuális helyzetét,
- a feladat pontos ismertetését és a követelmény specifikációt,
- a szükséges fejlesztői eszközök ismertetését és választásuk indoklását,
- a megvalósítandó rendszer tervét, és a tesztelési tervet,
- az eredményeket és a továbbfejlesztési lehetőségeket.



Dr. Fleiner Rita

Dr. Fleiner Rita
intézetigazgató

A szakdolgozat elévülésének határideje: **2024. május 15.**
(OE TVSz 55.§ szerint)

A dolgozatot beadásra alkalmasnak tartom:

.....
külső konzulens

Dr. Fleiner Rita

intézményi konzulens

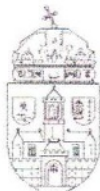


HALLGATÓI NYILATKOZAT

Alulírott hallgató kijelentem, hogy a szakdolgozat / diplomamunka saját munkám eredménye, a felhasznált szakirodalmat és eszközöket azonosíthatóan közöltem. Az elkészült szakdolgozatomban / diplomamunkámban található eredményeket az egyetem és a feladatot kiíró intézmény saját céljára térítés nélkül felhasználhatja.

Budapest, 2022.05.13.

Horváth Péter
hallgató aláírása



KONZULTÁCIÓS NAPLÓ

Hallgató neve: Neptun kód: Tagozat:
HORVÁTH DÁVID..... [REDACTED]..... NAPPALI.....
Telefon: Levelezési cím (pl: lakcím):
[REDACTED]

Szakdolgozat / Diplomamunka¹ címe magyarul:

ÁRUFUVAROZÓ CÉG ÜZLETI FOLYAMATAINAK TÁMOGATÁSA.....

Szakdolgozat / Diplomamunka² címe angolul:

SUPPORTING THE BUSINESS PROCESSES OF A FREIGHT FORWARDING
COMPANY

Intézményi konzulens:

Külső konzulens:

DR. HOLYINKA PÉTER.....

Kérjük, hogy az adatokat nyomtatott nagybetűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	2021.09.23.	KEZDETI DOLGOG MEGBESZÉLÉSE	Holyi
2.	2021.09.25.	TÉMAVÁLASZTÁS ÁTGONDOLÁSA	Holyi
3.	2021.10.26.	CÉGBESZÁMOLÓ ÉS AKTUÁLIS ÁLLAPOT	Holyi
4.	2021.12.10.	HIÁNYZÓ DOLGOK ÁTBESZÉLÉSE, VÉGLEGESÍTÉS	Holyi

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplomamunka 2 / Diplomamunka 3 / Diplomamunka 4³ tantárgy követelményét teljesítette, beszámolóra / védésre⁴ bocsátható.

Budapest, 2021.12.11.....

intézményi konzulens

¹ Megfelelő aláhúzendő!

² Megfelelő aláhúzendő!

³ Megfelelő aláhúzendő!

⁴ Megfelelő aláhúzendő!



KONZULTÁCIÓS NAPLÓ

Hallgató neve: Neptun kód: Tagozat:
HORVÁTH DÁVID..... [REDACTED]..... NAPPALI.....
Telefon: Levelezési cím (pl: lakcím):
[REDACTED]

Szakdolgozat / Diplomamunka¹ címe magyarul:

ÁRUFUVAROZÓ CÉG ÜZLETI FOLYAMATAINAK TÁMOGATÁSA.....

Szakdolgozat / Diplomamunka² címe angolul:

SUPPORTING THE BUSINESS PROCESSES OF A FREIGHT FORWARDING
COMPANY

Intézményi konzulens: Külső konzulens:
DR. HOLYINKA PÉTER

Kérjük, hogy az adatokat nyomtatott nagybetűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	2022.03.09.	ADATBÁZIS TERVÉNEK MEGBESZÉLÉSE	[Signature]
2.	2022.04.06.	WEBOLDAL	[Signature]
3.	2022.04.25.	AKTUÁLIS ÁLLAPOT BEMUTATÁSA	[Signature]
4.	2022.05.06.	TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK	[Signature]

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplomamunka 2 / Diplomamunka 3 / Diplomamunka 4³ tantárgy követelményét teljesítette, beszámolóra / védésre⁴ bocsátható.

Budapest, 2022.05.09.....

[Signature]
.....
intézményi konzulens

¹ Megfelelő aláhúzendő!

² Megfelelő aláhúzendő!

³ Megfelelő aláhúzendő!

⁴ Megfelelő aláhúzendő!

Tartalom

1. BEVEZETÉS.....	9
2. CÉLOK ÉS KÖVETELMÉNYEK.....	10
3. SZEMÉLYES KÖTŐDÉS ÉS A VÁLLALAT BEMUTATÁSA.....	12
3.1. Személyes érintettség	12
3.2. Vállalat bemutatása	12
4. ADATBÁZIS-KEZELŐ RENDSZEREK BEMUTATÁSA.....	13
4.1. Relációs adatbázis-kezelő rendszerek	13
4.1.1. MS SQL Server	13
4.1.2. PostgreSQL	13
4.2. NoSQL adatbázis-kezelő rendszerek	13
4.2.1. Redis	14
4.2.2. Apache Cassandra	14
4.3. A kiválasztott adatbázis-kezelő rendszer	14
5. ADATTÁRHÁZAK ELMÉLETI ÖSSZEFOGLALÓJA	15
5.1. Az adattárház fogalma	15
5.1.1. Ralph Kimball féle adattárház	15
5.1.2. William H. Inmon féle adattárház	16
5.2. Adattárház séma Relációs OLAP esetén.....	16
5.2.1. Csillag séma	16
5.2.2. Galaxis séma	17
5.2.3. Hópehely séma	18
6. ADATINTEGRÁCIÓ, ETL FOLYAMAT.....	19
6.1. Az ETL részei.....	19
6.1.1. Extract (adatkinyerés) fázis.....	19
6.1.2. Transform (átalakítás) fázis	19
6.1.3. Load (betöltés) fázis	19
6.2. ETL eszközök.....	20
6.2.1. Microsoft SQL Server Integration Services (SSIS)	20
6.2.2. SAP Data Integrator	20
6.2.3. Kettle Pentaho.....	20
6.2.4. iCEDQ Standard Edition	21
6.3. A kiválasztott ETL eszköz	21

7. ADATVIZUALIZÁCIÓ ISMERTETÉSE, ESZKÖZEINEK BEMUTATÁSA.....	22
7.1. Az adatvizualizáció tulajdonságai	22
7.2. Adatvizualizációs eszközök.....	23
7.2.1. Microsoft Power BI.....	23
7.2.2. Matplotlib.....	23
7.2.3. SAP Lumira	24
7.2.4. Google Charts	24
7.3. A kiválasztott adatvizualizációs eszköz.....	25
8. ADATBÁZIS KIALAKÍTÁSA.....	26
8.1. Adatbázis terv	26
8.2. Az adatbázis táblái és adatai	27
9. ADATTÁRHÁZ TERVEZÉSE ÉS MEGVALÓSÍTÁSA.....	30
9.1. Az ETL folyamat	30
9.1.1. Extract.....	30
9.1.2. Transform	32
9.1.3. Load.....	33
10. RIPORTERKÉSZÍTÉS	35
10.1. Melyik időszakban volt a legtöbb megrendelés?	35
10.2. Merre találhatóak a megrendelők boltjai, illetve melyik partner rendelte a legtöbb mennyiséget?	36
10.3. Mennyi árut szállított egy-egy autó összesen?	37
10.4. Ki mennyi terméket rendel átlagosan?	38
11. WEBOLDAL TERVEZÉSE ÉS MEGVALÓSÍTÁSA.....	39
11.1. Tervezés.....	39
11.2. Megvalósítás.....	39
12. TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK.....	44
13. ÖSSZEFOGLALÁS	45
14. SUMMARY	46
ÁBRAJEGYZÉK.....	47
TÁBLÁZATJEGYZÉK.....	47
RÖVIDÍTÉSEK:.....	47
IRODALOMJEGYZÉK.....	48
MELLÉKLET.....	50

1. BEVEZETÉS

A szakdolgozatom témája az áru fuvarozás témaköréhez kapcsolódik. Mit is értünk pontosan fuvarozás alatt? A fuvarozás egy bizonyos díj, fizetség ellenében folytatott árutovábbítás.

A mai világban rengeteg cég, vállalat, magánszemély foglalkozik különféle áruk fuvarozásával, mely történhet vízen, szárazföldön, de akár levegőben is. Első körben nagyon különbözőknek tűnhetnek, hiszen azt gondolhatjuk, hogy például a légi szállítás és a szárazföldi szállítás teljesen eltér egymástól. Azonban ez nem teljesen van így. Lényegében van egy megrendelő személy, a megrendelt áru és van egy szállító fél. Így már mindjárt jobban össze lehet kapcsolni az imént felsorolt lehetőségeket. A mai modern világban mi sem egyszerűbb, ha szükségünk van valamire megrendeljük az interneten és rövid időn belül meg is kapjuk a kívánt árut. Ezáltal már kapcsolatba is lépünk valamilyen fuvarozó, szállító céggel, aki a megrendelt árut kiszállítja nekünk.

Manapság nem egyszerű elkezdni egy ilyen vállalkozást, hiszen rengeteg konkurencia van a piacon. Nehéz egyből felvenni velük a versenyt. Emellett rengeteg mindenre oda kell figyelni, nagyon sok szabály és törvény van, amit kötelezően be kell tartani, mind a munkáltató, mind a munkavállaló részéről.

A szakdolgozatom célja, hogy a megkeresett vállalat által biztosított adatok alapján készítsek egy megfelelő adattárházat, mely segítségével könnyen készíthetők majd többféle riportot valamilyen adatvizualizációs eszköz támogatásával, ami elősegíti a cég jövőbeli működését, fejlődését vagy akár különböző újítások bevezetését.

2. CÉLOK ÉS KÖVETELMÉNYEK

A legelső feladatom az volt, hogy találjak egy olyan vállalkozást, amely szívesen segítene a terveim megvalósításában. Ez szerencsére gyorsan sikerült is, ugyanis ismertem olyan vállalkozót, akiről tudtam, hogy örömmel értesül róla, hogy az ő segítségével szeretném megvalósítani a szakdolgozatomat. A sikeres együttműködést követően rendkívül barátságosan és segítőkészen tette fel a kérdést, hogy ő miben tudna segíteni. Ekkor vázoltam fel neki az elképzeléseimet, hogy miről is lenne szó, milyen terveim vannak ehhez kapcsolódóan. Az első kérésem felé az volt, ha tud biztosítson számomra forrásadatokat, amiknek segítségével el tudom indítani a folyamatot. Erre a kérdésemre gyorsabban kaptam választ, mint ahogy gondoltam, ugyanis ők egy az i-Cell Kft. által nyújtott szoftvert használnak, ami tárolja a számomra szükséges adatokat. Ezt követően közösen ábeszteltük, hogy neki, mint tulajdonos, milyen információkkal tudnék segíteni a jövőre nézve és így megtárgyaltuk a rendszer felé támasztott követelményeket. Első körben bizonyos kérdésekre szeretne volna tudni a választ, melyek a következők.

- Egyes autókra lebontva, merre közlekedtek
- Mekkora mennyiséget rendelnek átlagban a vevők
- Melyik időszak volt a legjobb szállítás szempontjából
- Melyik megrendelőhöz történt a legtöbb szállítás

Ezeknek az információknak a tudatában számos dolgot le lehet majd szűrni, amelyből a cég profitálhat. Ennek érdekében a forrásadatokból egy olyan adattárház megvalósítása volt a feladatom, ami elősegíti ezen kérdések megválaszolását. Majd a végleges adatokból elkészült vizualizációk segítségével választ kaphatunk olyan dolgokra, hogy jelenleg elég munkaerő dolgozik-e a vállalatnál, a munkások mennyire vannak leterhelve, megfelelő számú gépjármű van-e munkában, esetleg lehetne-e bővíteni a flottát. Ennek következtében több megrendelést tudna a cég vállalni, ezzel nagyobb bevételt és profitot szerezni. Mely időszakok a nehezebbek, amikor esetleg elkélne még plusz munkaerő.

A vállalkozás folyamatos növekedése és bővülése érdekében a cég szeretne nyitni az online tér felé, ezért feladatként kaptam egy olyan kezdetleges weboldal elkészítését, mely kielégíti a már meglévő vagy az újonnan érkező megrendelőket. Ezt az oldalt kis bővítés és fejlesztést követően, akár az a nyilvánossággal is meg lehet osztani. Maga a weboldal HTML alapra fog épülni. Az oldalon megtalálható lesz a cégről egy bemutatkozás, hogy milyen munkakörökkel foglalkoznak, kik és milyen gépjárművel végzik a munkájukat, hogyan lehet felvenni a céggel a kapcsolatot és még pár általános információ legyen szó közösségi média elérésről vagy a szolgáltatások áráról. Ezáltal a vezetők is láthatnak a cég fontosabb dolgairól egy áttekinthető platformot, illetve későbbiekben a megrendelők számára is fontos lehet, hiszen a mai modern világban honnan is informálódna az ember, ha nem az

internetről. Jelenleg a piacon lévő legtöbb cégnek van fellelhető weboldala és nekem is az a célom ezzel a projekttel, hogy hozzásegítsem a céget egy alap weboldalhoz.

3. SZEMÉLYES KÖTŐDÉS ÉS A VÁLLALAT BEMUTATÁSA

3.1. Személyes érintettség

Az ember napi szinten találkozik fuvarozást végző emberekkel. Bárhol sétálhatunk az utcán, nagy valószínűséggel megpillanthatunk egy olyan furgont, amiben valamilyen árut vélhetünk felfedezni. Ki ne találkozott volna már egy csomagszállító futárral vagy a strandon, amikor épp mellette tölték el a rengeteg megrendelt üdítőt. A legfőbb szempont azonban mégsem ez volt a témaválasztás kapcsán, hanem a személyes tapasztalat a fuvarozásban végzett munkám során. Több nyáron keresztül volt lehetőségem bepillantani a különböző munkafolyamatokba. Egy üdítőitalokat forgalmazó cég kötelékében voltam, ahol magtapasztaltam mennyi minden kell ahhoz, hogy a megkapott rendelést a futár sikeresen felvegye és ki tudja szállítani a megrendelőnek. Mennyi precizitás és figyelem kell ahhoz, hogy pontosan az legyen kikészítve, amit a vevő megrendelt és se több, se kevesebb ne szerepeljen a listán. Egy-két év elteltével már egy vállalkozónak köszönhetően én magam szállíthattam ki az összekészített árukat. Nekem kellett találkoznom a megrendelőkkel, az én kezemből vették át a megrendelt árut. Saját magam voltam a felelős azért, hogy épségben odaérjen a megrendelt árucikk. Ehhez rengetek mindenre oda kellett figyelnem az utak során, hogy ne történjen semmi probléma és a megrendelő elégedett legyen.

3.2. Vállalat bemutatása

Az általam választott vállalkozás jelenleg zalaszentgróti telephellyel rendelkezik. De tekintsünk vissza a legelejére. Maga a vállalkozás 2013-ban indult. Az ötlet onnét származott, hogy a tulajdonos megkeresett egy vállalkozót, hogy másodállásban dolgozhasson nála, mint áruterítő. Az ott eltöltött bő fél év során tapasztaltak alapján elhatározásra jutott, hogy indít egy saját vállalkozást. Vásárolt egy Transit típusú gépjárművet és szezonális, 3 hónapra szóló szerződést kapott a Coca-Cola HBC Magyarország Kft-től. Ezt követően a nyárra szóló szerződésből éves megállapodás lett a két fél között, így évről évre tudott fejlődni és bővülni a vállalkozás. A hatalmas fejlődést az is mutatja, hogy kezdetben egy gépjármű és egy személy volt maga a vállalkozás, ami csak áruterítéssel foglalkozott. Jelenleg ott tart, hogy 14 gépjárművel rendelkezik és több mint tíz főt foglalkoztat. Legfőbb tevékenységük még mindig az áruterítés, hiszen 10 gépjármű végzi ezt a munkafolyamatot a Coca-Cola HBC Magyarország Kft-nek. Emellett foglalkoznak még veszélyes áru fuvarozással, amit 2 teherautóval végeznek. Rendelkezésre áll egy darus autó is, amit leginkább a fahordás során használnak a munkások, illetve van egy autó az egyéb fuvaroknak is fenntartva, legyen szó költöztetésről vagy valamilyen magánfuvarról. A munkákat jelenleg 11 fix ember végzi, akik főállásban dolgoznak. Sokszor előfordul, hogy szükség van beugrós emberekre. Ők általában 2-3 napot dolgoznak egy héten és napi bejelentéssel vannak foglalkoztatva.

4. ADATBÁZIS-KEZELŐ RENDSZEREK BEMUTATÁSA

4.1. Relációs adatbázis-kezelő rendszerek

A relációs modell az egyik legelterjedtebb adatmodell. Erre a modellre épül többek között a relációs adatbázis-kezelő rendszerek adatmodellje is. Gyakorlatban az 1980-as években kezdték el használni annak ellenére, hogy magát a relációs adatmodellt már 1970-ben definiálta egy bizonyos E. F. Codd, aki az IBM kutatója volt. A modell a logikai adatbázis kereteit szabja meg, nem foglalkozik az adatok tárolásával, módosításával és betöltésével. A relációs adatbázis kétdimenziós táblákból áll, ahol a tábla sorai a rekordok, oszlopai pedig a tulajdonságok.

4.1.1. MS SQL Server

A Microsoft SQL Server egy relációs adatbázis-kezelő rendszer (RDBMS), melyet a Microsoft fejlesztett és forgalmaz és amely a tranzakciófeldolgozás, az üzleti intelligencia és az analitikai alkalmazások széles skáláját támogatja vállalati informatikai környezetben. Az SQL Server elsősorban egy soralapú táblastruktúra köré épül, amely különböző táblákban lévő kapcsolódó adatelemeket összekapcsolja egymással, elkerülve az adatok redundáns tárolásának szükségességét az adatbázison belül több helyen. ⁽¹¹⁾

A többi RDBMS szoftverhez hasonlóan a MSSQL Server is az SQL-re épül, egy szabványos programozási nyelvre, amelyet az adatbázis adminisztrátorok és más informatikai szakemberek használnak adatbázisok kezelésére és a bennük lévő adatok lekérdezésére. ⁽¹¹⁾

4.1.2. PostgreSQL

A PostgreSQL egy hatékony, nyílt forráskódú objektum-relációs adatbázis-rendszer, amely az SQL nyelvet használja. Minden nagyobb operációs rendszeren fut. A PostgreSQL nagyon elterjedt a megbízhatóságának, adatintegritásának, bővíthetőségének és sok más tulajdonságának köszönhetően, így nem meglepő, hogy sok ember és vállalat választja ezt a nyílt forráskódú adatbázis-kezelő rendszert. ⁽¹²⁾

A PostgreSQL több olyan funkcióval rendelkezik, melyek azzal a céllal jöttek létre, hogy a legnagyobb segítséget nyújtsa a fejlesztőknek az alkalmazások készítésében, a rendszergazdáknak az adatok integritásának védelmében, valamint az adatok kezelésében. Legfontosabb pozitív tulajdonságai közé tartozik, hogy ingyenesen elérhető, nyílt forráskódú és számtalan bővítménnyel rendelkezik. ⁽¹²⁾

4.2. NoSQL adatbázis-kezelő rendszerek

Az NoSQL adatbázisok nem relációs adatbázisok és másképpen tárolják el az adatokat, mint a relációs táblák. Sok különféle típusú adatbázis gyűjtőneve. Ezeknek az adatbázisoknak négy típusát különböztetjük meg. Elsőként érdemes megemlíteni a kulcs-érték típusút. Ez a legegyszerűbb típus. Simán eltárolja a kulcs-érték párokat,

lényegében egy asszociatív tömbről van szó. Beszélhetünk dokumentum adatbázisokról, amik tartalmazhatnak kulcs-érték párokat, tömböket vagy beágyazott dokumentumokat is. Ezek leginkább JSON vagy XML formátumban vannak tárolva. Az oszlopalapú adatbázisok hasonlítanak leginkább a hagyományos tábla szerkezetre. Előnyös lehet ez a fajta, ha az adatbázis oszlopain végzünk bizonyos lekérdezéseket. A gráf típusú adatbázisoknál az adatok értelemszerűen csomópontok és élek formájában jelennek meg. Képes komplex, hierarchikus szerkezetű adatok tárolására is.

4.2.1. Redis

A Redis egy nyílt forráskódú, memórián belüli adatstruktúra tároló, melyet adatbázisként is használnak. A Redis több adatstruktúrát biztosít, legyen szó karakterláncokról, listákról, bittérképekről vagy akár adatfolyamokról. Továbbá beépített replikációval is rendelkezik. ^([3])

A maximális teljesítmény érdekében a Redis egy memórián belüli adatkészlettel dolgozik. Támogatja az aszinkron replikációt, a nagyon gyors, nem blokkoló első szinkronizálással és az automatikus újracsatlakozással. ^([3])

A Redis ANSI C nyelven íródott és küldő függőségek nélkül működik a legtöbb POSIX rendszerben, mint például a Linux vagy az OS X. A Linux és OS X az a két operációs rendszer, ahol a Redis leggyakrabban fejlesztik és tesztelik. A Windows build-ekhez nincs hivatalos támogatás. ^([3])

4.2.2. Apache Cassandra

Az Apache Cassandra egy elosztott NoSQL adatbázis. Tervezésüknél fogva a NoSQL adatbázisok könnyűek, nyílt forráskódúak, nem relációsak és nagyrészt elosztottak. Ezek az adatbázisok lehetővé teszik a rendkívül nagy mennyiségű, eltérő adattípusok gyors, ad hoc rendszerezését és elemzését. A Cassandra azon NoSQL adatbázisok közé tartozik, melyek kezelték a korábbi adatkezelési technológiák korlátait. ^([4])

Az Apache Cassandra az egyetlen elosztott NoSQL adatbázis, amely állandó rendelkezésre állást, rendkívül gyors olvasási-írási teljesítményt és korlátlan lineáris méretezhetőséget biztosít, amelyek lehetővé teszik a sikeres modern alkalmazások igényeinek kielégítését. ^([4])

4.3. A kiválasztott adatbázis-kezelő rendszer

Fejlesztőkörnyezetnek a Microsoft SQL Servert választottam. Mindenképp relációs adatbázis-kezelő rendszerre volt szükségem és mivel ez Microsoft termék, könnyebb volt megtanulni a kezelését és megjelenésében is tapasztaltam már ismert dolgokat. Az is a rendszer mellett szólt, hogy volt már szerencsém használni, így nem kellett teljesen nulláról elsajátítanom a működéshez szükséges folyamatokat. Nem utolsósorban könnyen kezelhető és a számomra szükséges dolgok tökéletesen megvalósíthatók vele.

5. ADATTÁRHÁZAK ELMÉLETI ÖSSZEFOGLALÓJA

5.1. Az adattárház fogalma

Az adattárház definícióra az évek során többféle megfogalmazás is történt, melyeket kisebb nagyobb sikerrel fogadtak el a szakemberek. Manapság két olyan változat is létezik, melyek a mai napig megállják a helyüket és elfogadott definícióként tartják számon. A két definíció az adattárház témakörében jártas, legismertebb személyektől származik, akiket sokan az adattárház atyjaként ismernek el. Ők pedig név szerint Ralph Kimball és Bill Inmon.

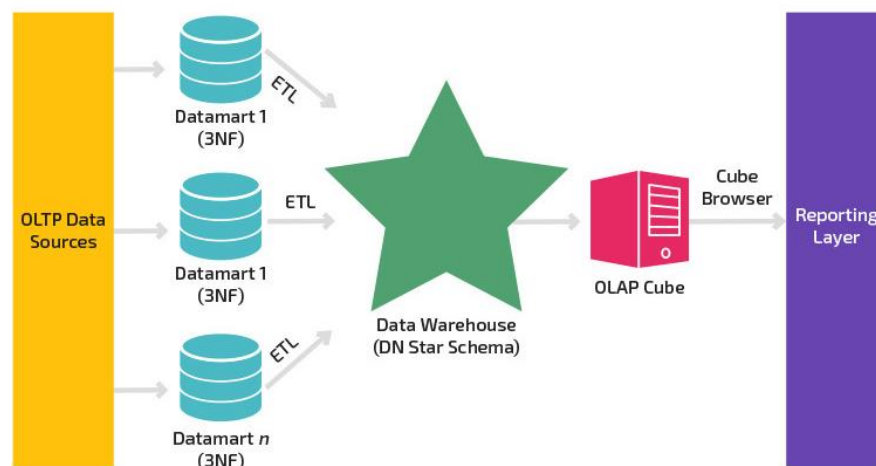
5.1.1. Ralph Kimball féle adattárház

„The conglomeration of an organization’s data warehouse staging and presentation areas, where operational data is specifically structured for query and analysis performance and ease-of-use.” ([5])

Ralph Kimball így definiálta az adattárház fogalmát, ami magyarul azt jelenti, hogy egy szervezeti adattárház átmeneti és bemutatási területeknek az összessége, ahol az operációs adatok kifejezetten strukturáltak a lekérdezésre, az elemzési teljesítményre és az egyszerű használatra.

Kimball szerint a tervezés folyamata a legfontosabb üzleti folyamatokkal kezdődik. A forrásokból adatpiacokat hozunk létre, melyek tárgyspecifikus területeken összegyűjtik az adatokat. Az ő szemléletmódja, amit az 1. ábra mutat be a bottom-up, vagyis az alulról felfelé irányuló megközelítést alkalmazza.

Az ő nevéhez fűződik az a megfogalmazás, hogy „the data warehouse is nothing more than the union of all data marts” ([6]), azaz az adattárház nem más, mint az összes adatpiac egyesülése.



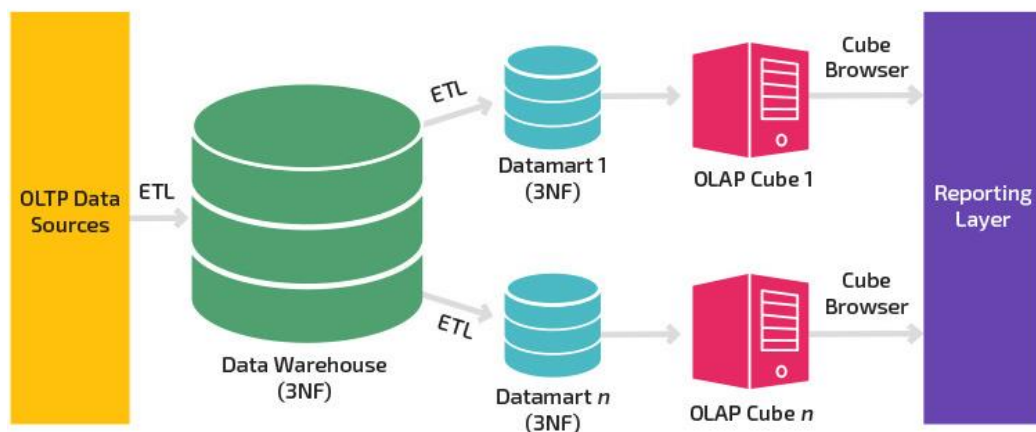
1. ábra Kimball modell⁽¹⁸⁾

5.1.2. William H. Inmon féle adattárház

„A data warehouse is a subject oriented, integrated, nonvolatile, and time variant collection of data in support of management’s decisions.”⁽¹⁷⁾

Inmon szerint az adattárház nem más, mint témakör orientált, integrált, időfüggő, de időben nem változó adatok gyűjteménye, amelyet a cég stratégiai döntéshozatalának támogatására használnak.

A 2. ábra Bill Inmon szemléletmódját mutatja, amely a top-down, azaz a felülről lefelé irányuló megközelítést alkalmazza, amelyben az adattárház a központi adattár és a szervezet adatrendszereinek a legfontosabb összetevője. Ez a megközelítés először a központosított vállalati adatmodellt építi fel. Az adattárház pedig ugyanennek a modellnek a fizikai ábrázolása. Ebben modellben az adattárházban lévő adatok integrálva vannak, vagyis az adattárház a forrása azoknak az adatoknak, amelyek a különböző adatpiacokba kerülnek.



2. ábra Inmon modell⁽¹⁸⁾

5.2. Adattárház séma Relációs OLAP esetén

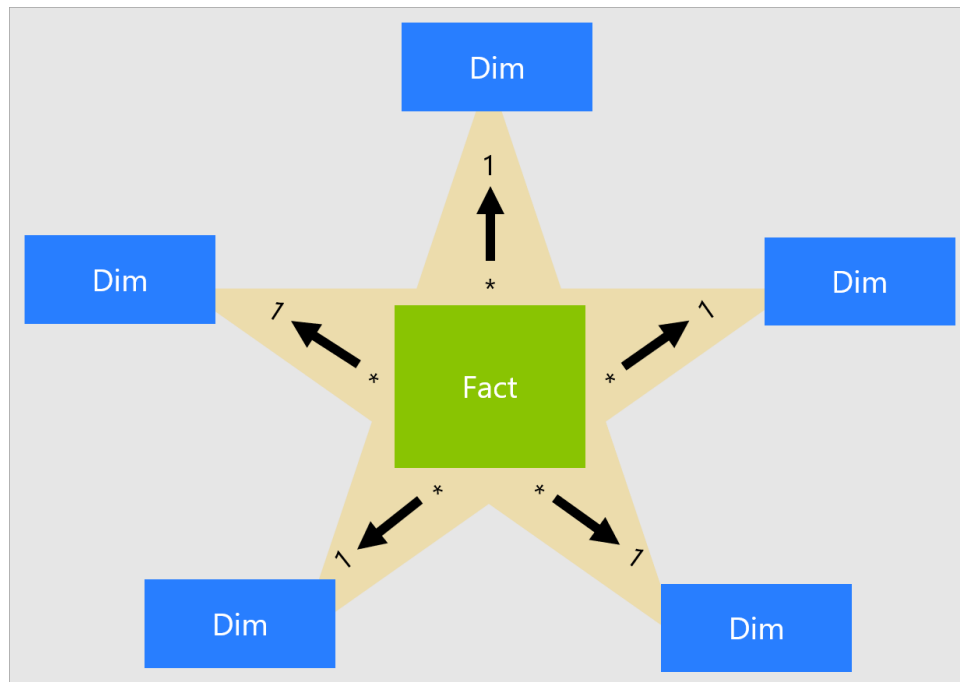
A ROLAP vagyis a Relational Online Analytical Processing, magyarul relációs online analitikus feldolgozást jelent. Ilyen feldolgozás esetén az adatok eltárolását tradicionális relációs adatbáziskezelővel végezzük.

ROLAP architektúrájú adatbázis esetén nagyon fontos kiemelni a ténytábla és a dimenziótábla fogalmát, hiszen ezeknek az egymással való kapcsolata határozza meg azt, hogy milyen sémájú adatbázisról van szó. Ezeknek a sémáknak három különböző fajtáját tudjuk megkülönböztetni. Beszélhetünk csillagsémáról, galaxis modellről és hópehely sémáról.

5.2.1. Csillag séma

A csillagséma, amelyet a 3. ábra szemléltet a leginkább kedvelt séma a három közül. Ennek a felépítése a legegyszerűbb, hiszen egyetlen ténytáblából és pár

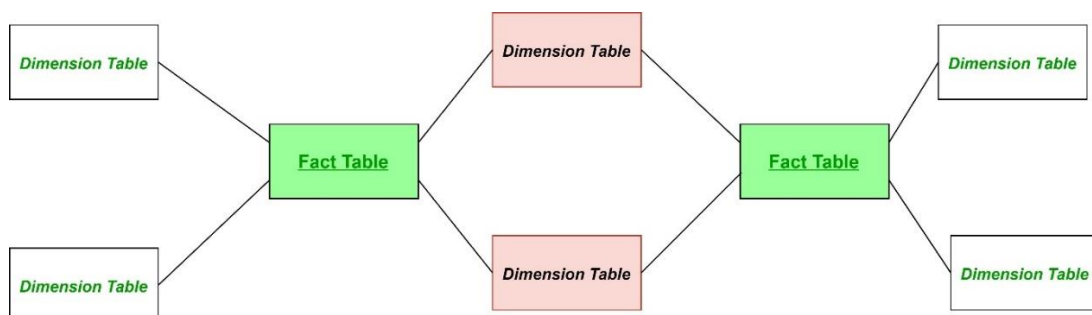
dimenziótáblából áll. Ebből adódóan a ténytábla mindegyik dimenziótáblával közvetlen kapcsolatban áll. Nevét onnét kapta, hogy úgy rendezzük el a táblákat, hogy középen van mindig a ténytábla körülötte pedig a dimenziótáblák, ami pont úgy néz ki, mint egy csillag. Amellett, hogy kevés táblából áll, előnye még, hogy könnyű belőle megérteni és kiolvasni az adatokat, valamint rendkívül gyorsan hozzá is férhetünk ezekhez az adatokhoz.



3. ábra Csillag séma⁽¹⁹⁾

5.2.2. Galaxis séma

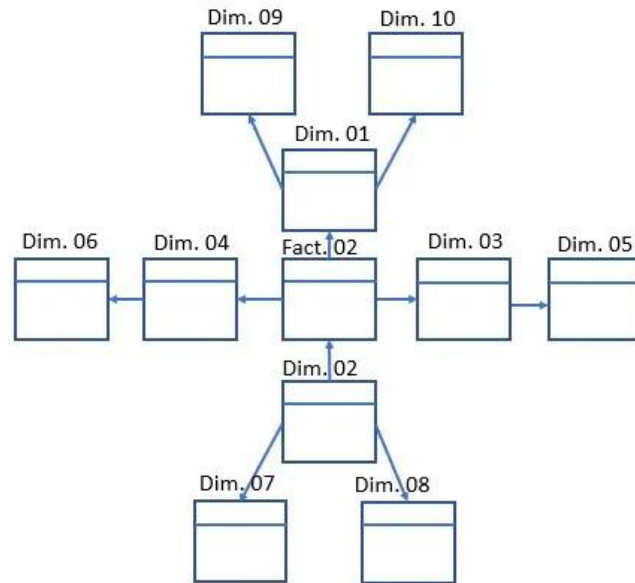
A 4. ábrán a galaxis séma látható, amely a csillag sémának egy kibővített változata. Ebben az esetben nem egy, hanem több ténytáblánk van és ezek köré épülnek fel a dimenziótáblák. Két ténytábla közötti kapcsolatot a közös dimenziótáblák biztosítják. Nincsenek redundáns adatok, ugyanis megvan a ténytáblák közötti átjárhatóságnak a lehetősége. Nem szerepel duplikáltan egy dimenziótábla sem.



4. ábra Galaxis séma⁽²⁰⁾

5.2.3. Hópehely séma

A hópehely séma esetében, amelyet az 5. ábrán figyelhetünk meg, akár csak a csillagséma esetén egy ténytábláról beszélhetünk, amelyhez több dimenziótábla kapcsolódik, azonban a hópehely séma egyik különlegessége, hogy a dimenziótáblák egymással is össze lehetnek kapcsolva. A dimenziótáblákat elkezdjük normalizálni, mivel a hierarchikus adatokat nem szeretnénk redundánsan eltárolni.



5. ábra Hópehely séma^([21])

6. ADATINTEGRÁCIÓ, ETL FOLYAMAT

A 6. ábrán az ETL (Extract, Transform, Load) folyamat figyelhető meg, amely nagyon fontos funkciót tölt be az adatintegrációs stratégiákban. Adattárházak tervezése során felmerülhet az a kérdés, miért is van szükség az ETL folyamatra? Tudjuk, hogy vannak különböző források és ezek valamilyen módon mindig bekerülnek az adattárházba. Ezeknek a forrásoknak a felépítése általában mindig más. Az ETL folyamatnak az a célja, hogy a különféle adatokat be tudjuk integrálni az adattárházba. Tehát az ETL egy folyamat, mely különböző forrásokból kinyeri az adatokat, azokat átalakítja, majd betölti egy adattárházba vagy egy adattárba.

6.1. Az ETL részei

6.1.1. Extract (adatkinyerés) fázis

Ez a rész az ETL folyamat első fázisa, amely az adatok kinyeréséről szól a különböző forrásrendszerekből. Az a célja, hogy ez az adatkinyerés minél gyorsabban megtörténjen és az adatokat a különféle forrásokból bele tudjuk tenni egy adatbázisba. Minél kevesebb ideig foglaljuk le a performanciát, azaz a teljesítményt, annál jobb, ugyanis így csak rövid ideig lassítjuk a többi folyamatot. Nagyon lényeges aspektusa ez az egész folyamatnak, hiszen az, hogy helyesen nyerjük ki az adatokat, nagymértékben befolyásolja a következő lépések sikerességét. ⁽¹⁸⁾

6.1.2. Transform (átalakítás) fázis

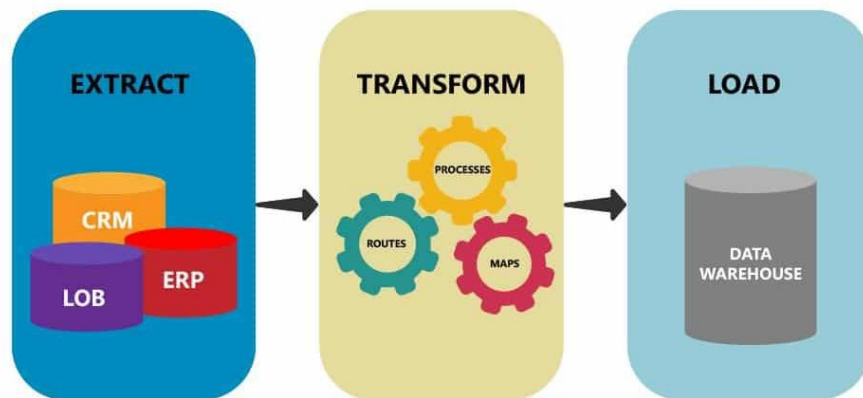
A forrás adatokat át kell alakítani olyan formára, hogy azok megfeleljenek az adattárházban meghatározott követelményeknek. Ez a folyamat olyan szabályokat és előírásokat alkalmazhat, amelyek biztosítják az adatok minőségét és elérhetőségét. A transform folyamat több alfolyamatból áll. Ilyen például a tisztítás, ami megoldja az adatban lévő ellentmondásos és hiányos értékek problémáját. A hitelesítés, mely formázási szabályt alkalmaz az adatkészletre. A deduplikáció a redundás adatok eltávolítását oldja meg. Az ellenőrzés a használhatatlan adatokat eltávolítja és jelöli a rendellenességet. A rendezés pedig az adatokat típus szerint rendezi.

A transform folyamat szintén nagyon fontos része az ETL-nek, hiszen ez a rész javítja az adatok integritását és segít abban, hogy az adatok megfelelően és használható állapotban kerüljenek a következő helyre. ⁽¹⁸⁾

6.1.3. Load (betöltés) fázis

Az adatok betöltése az ETL folyamat utolsó lépése. Ebben a részben az átalakított adatok az átmeneti helyről átkerülnek a céladatbázisba. A betöltés több módon is történhet. Lehet például úgynevezett teljes betöltés, ami azt jelenti, hogy minden adat, ami az előző fázisból érkezik, új, egyedi rekordként kerül bele az adattárházba. Emellett beszélhetünk differenciális adatbetöltésről is. Ez egy kevésbé átfogó, de jobban kezelhető megoldás. Ennek a betöltésnek az a lényege, hogy végez egy összehasonlítást

a meglévő adatok és az éppen bejövő adatok között és csak azokat az adatokat tölti be, amelyek még nem szerepelnek az adattárházban. ⁽¹⁸⁾



6. ábra ETL folyamat⁽¹²²⁾

6.2. ETL eszközök

6.2.1. Microsoft SQL Server Integration Services (SSIS)

Az SQL Server Integration Services egy platform a vállalati szintű adatintegrációs és adatátalakítási megoldások kiépítésére. Ennek segítségével különböző problémákra találhat megoldást a fájlok másolásával vagy letöltésével, esetleg adattárházak betöltésével. Rendkívül sok forrásból tudunk kinyerni, átalakítani és betölteni adatokat az eszköz segítségével egy vagy akár több helyre is. Van lehetőség kód nélkül is megoldani a feladatot, ebben segítenek a grafikus eszközök. ⁽¹⁷⁾

6.2.2. SAP Data Integrator

A SAP Data Integrator ETL, ELT és szöveges adatelemzési lehetőséget nyújt egyetlen megoldásban. Ez az eszköz segít abban, hogy a különböző forrásrendszerből származó adatokat egy környezetbe helyezzük. A SAP Data Integrator segítségével bármilyen forrásból kinyerhetjük az adatokat, azokat át is tudjuk alakítani és formázni tetszőleges adatbázisba. A szoftver könnyebbé teszi az adatok mozgatásához és átalakításához szükséges műveletek fejlesztését, ugyanis széleskörű könyvtárral rendelkezik e témakörben. ⁽¹⁵⁾

6.2.3. Kettle Pentaho

A Pentaho egy vezető üzleti intelligencia platform, amely lehetővé teszi a szervezet számára, hogy könnyen hozzáférjen az adatokhoz, előkészítse és elemezze a könnyen használható és intuitív felületeken keresztül. A Pentaho biztosítja, hogy az adatok függetlenül a méretüktől és egyéb tulajdonságaiktól megfelelően legyenek irányítva. Hatékony ETL képességeket nyújt, amelyek magukba foglalják az adatok gyűjtését,

akár több forrásból. Az adatok egy helyre való összevonását, valamint azt, hogy az adatok egységes formátumban jelenjenek meg.^[(16)]

6.2.4. iCEDQ Standard Edition

Az iCEDQ egy Java-ba épített, memórián belüli motor, amely a java adatbázis kapcsoló segítségével tud csatlakozni többféle adatforráshoz. Ez az eszköz támogatja az adattárház tesztelését és az üzleti intelligencia tesztelését is. A szabálmotor csatlakozik a forrás- és céladatrendszerhez. Az iCEDQ bármilyen típusú fájlból és adatbázisból kiolvassa az adatokat. A motor mikro kötegként olvassa be az adatokat a memóriába. Az összehasonlítás és az értékelés alapján képes kiszűrni a helytelen információkat. Ha a motor problémát talál, azokat egy úgynevezett kivétel riportba írja. Az eszköz képes arra, hogy rengeteg adatsort hasonlítson össze egyszerre a memóriába. A végén a teljes kivétel riport elérhető a felhasználó számára, abból a célból, hogy a rossz adatokat elemezni tudják.^[(14)]

6.3. A kiválasztott ETL eszköz

Az ETL folyamat megtervezése és megvalósítása során a Microsoft SQL Server Integration Services (SSIS) programot fogom használni. Ez a program illeszkedik legjobban az adatbázis kezelőhöz, hiszen ugyanahhoz a gyártócsoporthoz tartoznak. Emellett teljes mértékben kivitelezhető vele a számomra szükséges folyamatok elkészítése. Az egyszerű használat mellett rendkívül sok objektum közül lehet választani a toolbox résznél, ezáltal színesebben és magasabb szinten tudom majd végrehajtani a folyamatokat. Könnyen kezelhető és felhasználóbarát.

7. ADATVIZUALIZÁCIÓ ISMERTETÉSE, ESZKÖZEINEK BEMUTATÁSA

7.1. Az adatvizualizáció tulajdonságai

Az adatvizualizáció alatt azokat a módszereket értjük, amelyek segítségével a birtokunkban lévő információkat vizuálisan tudjuk megjeleníteni. Ez segít abban, hogy az emberi agy könnyebben feldolgozza a kapott információt és betekintést nyerjen a részletekbe. Az adatvizualizációnak egyik fő célja, hogy a sablonok és trendek beazonosítását megkönnyítse elég nagy adathalmazokban. ^([9])

Az adatvizualizáció az adattudományi folyamatnak egyik lépése. Ez a lépés kimondja, hogy az adatok összegyűjtése, feldolgozása és modellezése után muszáj őket valamilyen módon vizualizálni, hogy az elkészült diagramokból, képekből különféle következtetéseket lehessen levonni. Az adatvizualizáció tágabb értelemben az adatmegjelenítési architektúra tudományágának is az eleme. Ennek a legfőbb célja, hogy az adatok azonosítása, helyük meghatározása, formázása és továbbítása minél hatékonyabb módon történjen. ^([9])

Az adatvizualizáció gyors és hatékony módot biztosít az információk univerzális módon történő közlésére. A gyakorlat segít a vállalkozásoknak azonosítani, hogy mik azok a tényezők, melyek hatással vannak az ügyfelek magatartására. Megállapítja azokat a területeket, amelyeken javítani kellene, esetleg nagyobb odafigyelést igényelne. Ezen felül is rengeteg előnnyel bír az adatvizualizáció. Segíti a fontos információk gyors befogadását, általa gyorsabban meg tudjuk hozni a szükséges döntéseket, könnyebben megértjük a következő lépéseket, amelyeket a cég vagy vállalat fejlődése érdekében érdemes megtenni. Biztosítja az információk értelemszerű elosztását, így mindenki számára növeli a betekintések megosztásának lehetőségét. ^([9])

Az adatvizualizáció tudománya annak megértésében rejlik, hogy az emberek hogyan gyűjtik össze az információkat és dolgozzák fel azokat. Ehhez kapcsolódóan Daniel Kahneman és Amos Tversky egy olyan kutatásban dolgoztak együtt, amely két különböző módszert definiál az információgyűjtésre és az információfeldolgozásra. A közösen kidolgozott elméletért, amit kilátáselméletnek nevezünk mindketten Nobel-díjat kaptak. Az egyik módszer a gyors, automatikus és tudattalan gondolatfeldolgozást vetíti a középpontba. Ezzel a módszerrel gyakran találkozhatunk a hétköznapi életben is. Ide tartozik egy táblán vagy papíron lévő szöveg elolvasása, egy hang beazonosítása, hogy honnan is halljuk., vagy éppen legyen szó színek megkülönböztetéséről, vagy csak egy könnyű matematikai feladat megoldásáról. Míg a másik módszer a lassú, logikus és ritka gondolatfeldolgozásra fókuszál. Olyanokra, mint például a telefonszám visszamondása, összetett matematikai feladatok megoldása vagy az összetett társadalmi jelzések megértése. ^([9])

7.2. Adatvizualizációs eszközök

A mai világban az adatvizualizációs eszközöket leginkább üzleti intelligencia jelentéskészítő eszközként használják. Ebben a részben pár ilyen termék bemutatása fog történni.

7.2.1. Microsoft Power BI

A jelenlegi piacon talán az egyik legismertebb adatvizualizációs eszköz a Microsoft Power BI. A Microsoft célja ezzel az eszközzel az üzleti intelligencia és vizualizáció támogatása. A felhasználói felülete meglehetősen intuitív, így, ha valaki használt már többféle Microsoft terméket biztosan kiigazodik ezen is probléma nélkül. ⁽¹⁰⁾

A Power BI segít a különböző adatkészletek összekapcsolásában, az adatok adatmodellé történő átalakításában és tisztításában, illetve a diagramok vagy grafikonok létrehozásában vizuális módon. Az elkészült munkákat tetszés szerint meg lehet osztani a többi felhasználóval a szervezetben belül. A Power BI-ből létrehozott adatmodellek többféle módon használhatók fel a szervezetek számára. A jelentések valós időben tudnak válaszolni a feltett kérdésekre és segítséget nyújthat az előrejelzésekben, hogy a részlegek megfelelő üzleti mutatókkal rendelkezzenek. A Power BI rendelkezik egy olyan lehetőséggel is, hogy az adminisztrátorok és a vezetők számára egyfajta vezetői irányítópultot biztosít, ezzel jobban betekinhetnek a részlegek folyamataiba. ⁽¹⁰⁾

Tudjuk, hogy a Power BI egy önkiszolgáló eszköz, amely elemzéseket juttat el a dolgozóknak, azonban mégis leginkább adatelemző és üzleti intelligencia szakemberek használják abból a célból, hogy különféle adatmodellt készítsenek, mielőtt a riportokat továbbítják a szervezetben belül. Ezenkívül a Power BI egy adminisztrátori portált kínál a rendszergazdák számára monitorozáshoz és a licencek konfigurálásához. ⁽¹⁰⁾

A Microsoft folyamatosan bővíti az eszköz funkcióit. Manapság már olyan fontos funkciókat tartalmaz, mint a mesterséges intelligencia, amivel a felhasználók hozzáférhetnek a képfelismeréshez vagy a szövegelemzéshez. Ide tartozik még a közös adatmodell támogatás, amivel lehetővé teszik az adatsémák számos használatát. Továbbá személyre szabhatjuk az alapértelmezett eszközöket és még újakat is tudunk importálni a platformra és ami talán a legérdekesebbnek tűnhet az a Cortana integráció. Ez a funkció lehetővé teszi a felhasználók számára, hogy a mobiltelefonjukon szóban lekérdezzenek adatokat és hozzáférjenek az eredményekhez, mindezt a Microsoft digitális asszisztense, Cortana segítségével. ⁽¹⁰⁾

7.2.2. Matplotlib

A Matplotlib az egyik legerősebb és legnépszerűbb adatvizualizációs könyvtár a Pythonhoz, amely NumPy tömbökre épül. Lehetővé teszi statikus, animált és interaktív vizualizációk létrehozását Pythonban. Tervezése és megírása John D. Hunter nevéhez fűződik, aki 2002-ben kezdett el foglalkozni vele. Számos tudományos publikáció ábrái készülnek a Matplotlib eszköz segítségével. ⁽¹¹⁾

A Matplotlib egyik talán legfontosabb tulajdonsága, hogy jól tud együttműködni több operációs rendszerrel és grafikus háttérprogrammal. Számtalan háttérprogramot és kimeneti típust támogat, emiatt a felhasználó számíthat arra, hogy függetlenül attól, milyen operációs rendszert használ vagy milyen kimeneti formátumot szeretne, a rendszer működni fog. Mindent egybevetve talán ez a Matplotlib egyik legnagyobb erőssége. Ezáltal sokan elkezdtek használni és ez nagy felhasználói bázishoz vezetett, amiből adódott a fejlesztői bázis növekedése is. Végül oda jutott a dolog, hogy a Python világában mindenütt jelen van. ⁽¹¹¹⁾

A könyvtárát rengeteg területen előszeretettel használják. Ilyenek például a képfeldolgozás, azon belül is a képmódosítás vagy a képek megértése különböző ábrázolással. Megtalálható a jelfeldolgozásnál is, mint a zajelemzés és a mélységbecslés. Továbbá a gépi tanulás és az adattudomány területén is találkozhatunk vele csakúgy, mint az akadémiai kutatások során. ⁽¹¹¹⁾

7.2.3. SAP Lumira

A SAP Lumira vizuális intelligencia eszközként ismert, melyet adatok megjelenítésére és történetek létrehozására használnak. Ezek grafikus részleteket biztosítanak az adatokról. Az adatok adatkészletként kerülnek be a Lumirába. Ezt követően van lehetőség szűrőket, hierarchiakat és oszlopokat alkalmazni a dokumentum elkészítéséhez. Az adatok hatékony megjelenítéséhez különféle diagramokat választhat. ⁽¹²¹⁾

A SAP Lumira az üzleti elemzők számára készült, akik tetszés szerint módosíthatják az adatstruktúrákat és a korrelációkat. Adatvizualizációkat és történeteket hozhatnak létre több adatforrásból. Ez az eszköz segít az adatokat a szervezeti igényekhez igazítani vizualizációk segítségével. ⁽¹²¹⁾

A Lumira lehetővé teszi az eredmények kimutatását és előrejelzést biztosít a változó piaci helyzetnek megfelelően. Több adatforrásból is létrehozhat adatvizualizációkat és történeteket. Lehetőséget nyújt megosztani az elkészített vizualizációkat különböző platformokon. ⁽¹²¹⁾

7.2.4. Google Charts

A Google Charts tökéletes módot biztosít a webhelyen lévő adatok megjelenítésére. Az egyszerű vonaldiagramoktól a bonyolult hierarchikus fatérképekig a diagramgaléria számos használatra kész diagramot ajánl a felhasználó számára.

A Google Charts egy infografikus szoftver, amely interaktív diagramokat biztosít mobileszközökhöz és böngészőkhöz. A szoftver olyan lehetőséget nyújt az eszközökkel, melyek lehetővé teszik, hogy élő adatokat jelenítsen meg az oldalán. Rendkívül könnyen használható, emellett hatékony megoldást biztosít a feladatok elvégzésére, valamint ingyenesen elérhető. Tulajdonságai közé tartozik még a gazdag galéria, a testre

szabható lehetőségek, a HTML5 tulajdonságok, az ingyenes diagrameszközök és a dinamikus adatok is. ⁽¹³⁾

7.3. A kiválasztott adatvizualizációs eszköz

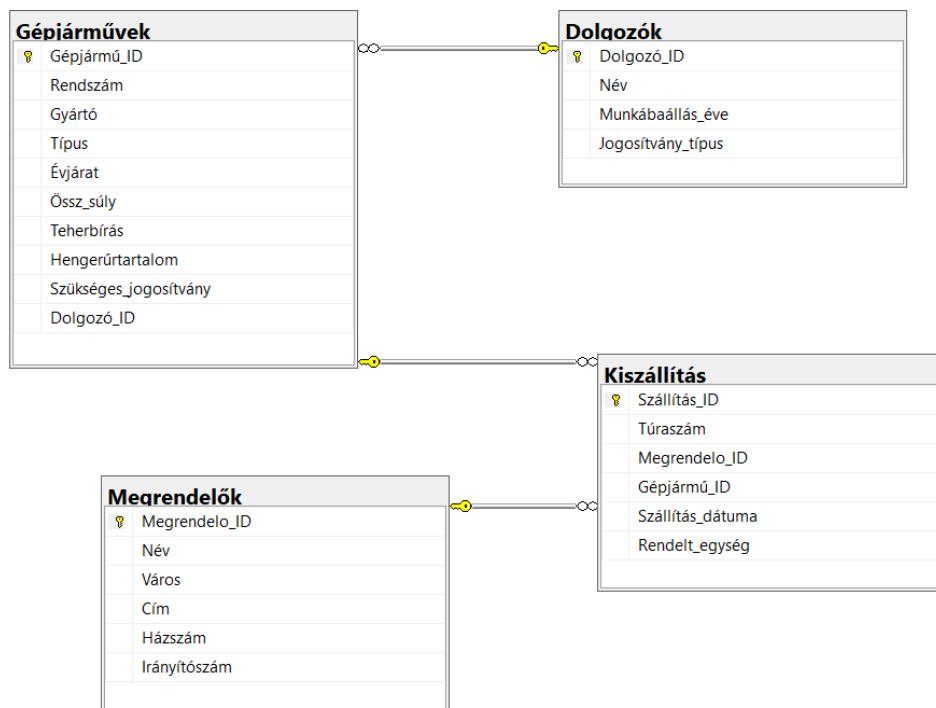
A munkám során az adatok vizualizálására, illetve a riportok elkészítéséhez a Microsoft Power BI eszközt fogom használni. Tanulmányaim során már megtapasztalhattam, hogy miért is ez az egyik legnépszerűbb vizualizációs eszköz. Ahol lehetséges Microsoft terméket használok, ezért már a Power BI első használatakor ismerős volt a megjelenés, ugyanis nagyban hasonlít a többi Microsoft termék megjelenéséhez. Emiatt gyorsan megbarátkoztam vele és könnyen megismerkedtem a lényegesebb dolgaival. Könnyen lehet importálni az adatokat és erre sok lehetőséget is biztosít, ami számomra fontos volt. Nagyon sok diagram közül lehet választani, ezáltal jobbnál jobb megoldásokat tudok majd alkalmazni a munkám során. Célom ugyanis, hogy érdekes, színes és érthető vizualizációkat tudjak készíteni. Könnyen kezelhető és rövid idő alatt el lehet sajátítani a szükséges dolgokat.

8. ADATBÁZIS KIALAKÍTÁSA

A cég által biztosított adatok Excel formátumban kerültek hozzám. Elégé rendszertelen és strukturálatlan állapotban voltak, ugyanis a cégnél nincs olyan ember, aki rendszeresen karbantartaná és odafigyelne az adatok megfelelő tárolására. A rengeteg adatból ki kellett válogatnom azokat, melyek a munkám során hasznosak és fontosak lehetnek. Emellett figyelnem kellett arra is, hogy a cég számára kialakítsak egy olyan adatbázist, amivel a későbbiekben megkönnyíthetem a munkájukat és átláthatóbbá tehetem az adataik tárolását.

8.1. Adatbázis terv

Az adatbázist SQL Server segítségével hoztam létre, amelyet a 7. ábra is mutat és ahova az általam megalkotott Excel fájlokat csv-be való konvertálást követően töltöttem be.



7. ábra Adatbázis terv

8.2. Az adatbázis táblái és adatai

Gépjárművek tábla

Oszlopnév	Típus	Magyarázat
Gépjármű_ID	int	A gépjárművekhez tartozó egyedi azonosító
Rendszám	varchar(10)	Az adott gépjárműnek a rendszáma.
Gyártó	varchar(50)	Milyen gyártmányú a gépkocsi.
Típus	varchar(50)	A gépjármű típusa.
Évjárat	int	A gépjármű gyártásának az éve.
Össz_súly	float	A gépjármű teljes súlya.
Teherbírás	float	Mekkora súlyt bír el a gépjármű.
Hengerűrtartalom	int	A hengerek össz űrtartalma cm ³ -ben.
Szükséges_jogosítvány	varchar(5)	Milyen jogosítvánnyal vezethető az gépjármű.
Dolgozó_ID	int	Melyik dolgozóhoz tartozik az autó.

1. táblázat Gépjárművek tábla

Az 1. táblázat tartalmazza azon teherautók adatait, melyekkel a cég a mindennapos szállításokat végzi. Ezek az adatok nem voltak benne a kapott forrásban, így nekem kellett összegyűjteni az egyes autókra vonatkozó adatokat, de szerencsére segítségemre volt a tulajdonos, ezáltal nem tartott sok időbe.

Dolgozók tábla

Oszlopnév	Típus	Magyarázat
Dolgozó_ID	int	A dolgozókhoz tartozó egyedi azonosító.
Név	varchar(50)	A dolgozó neve.
Munkábaállás_éve	int	Mióta dolgozik valaki a cégnél.
Jogosítvány_típus	varchar(10)	Az illető milyen jogosítvánnyal rendelkezik.

2. táblázat Dolgozók tábla

A 2. táblázat szemlélteti a vállalatnál dolgozó sofőrök neveit és a hozzájuk tartozó adatokat.

Megrendelők tábla

Oszlopnév	Típus	Magyarázat
Megrendelő_ID	int	A megrendelőhöz tartozó egyedi azonosító
Név	varchar(100)	A megrendelő személy vagy üzlet neve.
Város	varchar(100)	Melyik városban található a megrendelő.
Cím	varchar(100)	A megrendelő boltjának, házának a címe.
Házzszám	int	A megrendelő boltjának, házának a házzsáma.
Irányítószám	int	A város irányítószáma.

3. táblázat Megrendelők tábla

A 3. táblázatban megfigyelhetők azon vevők adatai, akik az aktuális időszakban igénybe vették a cég által nyújtott szolgáltatást.

Kiszállítás tábla

Oszlopnév	Típus	Magyarázat
Szállítás_ID	int	A szállításhoz tartozó egyedi azonosító.
Túraszám	int	Melyik szállításhoz tartozik az adott elem.
Megrendelő_ID	int	A szállítás megrendelőjének azonosítója.
Gépjármű_ID	int	Annak a gépjárműnek az azonosítója, amellyel végzik a szállítást.
Szállítás_dátuma	date	A szállítás dátuma.
Rendelt_egység	int	A rendelés mennyisége.

4. táblázat Kiszállítás tábla

A 4. táblázat azokat az adatokat tartalmazza, amelyek egy adott megrendelés kiszállításához tartoznak.

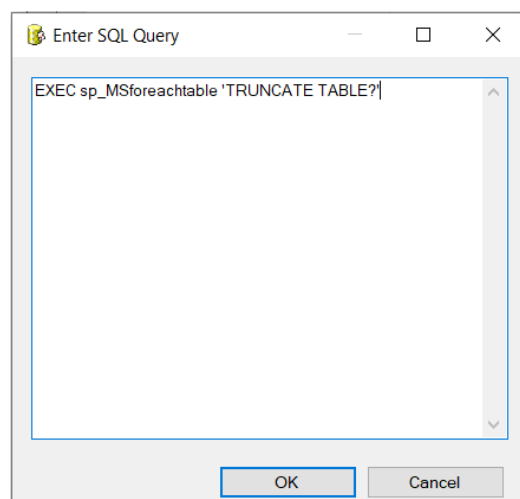
9. ADATTÁRHÁZ TERVEZÉSE ÉS MEGVALÓSÍTÁSA

9.1. Az ETL folyamat

Az SQL Server Management Studio segítségével létrehoztam három új adatbázist a fentebb említett adatbázisból, melyeket az ETL folyamat során fogok használni. Legelső lépésként egy STAGE nevű réteget készítettem. Ennél a résznél még minden adatot betöltünk, nem foglalkozva azzal, hogy a későbbiekben nem lesz szükségünk rá. Fontos dolog még, hogy típusként varchar, esetleg nvarchar adjunk meg. Célunk ennél a résznél a gyors és zökkenőmentes adatbetöltés. A következő rétegünk a DW, azaz a DataWarehouse lesz. Ennél a folyamatnál már hozzárendeljük a megfelelő típusokat az egyes értékekhez és úgy töltjük be az adatbázisba. Ebben a rétegben történik a transzformálás. A harmadik réteg pedig a DM, vagyis a DataMart. Itt történik a csillagséma megvalósítása és betöltése. Ezeket a folyamatokat részletezem a következő fejezetekben.

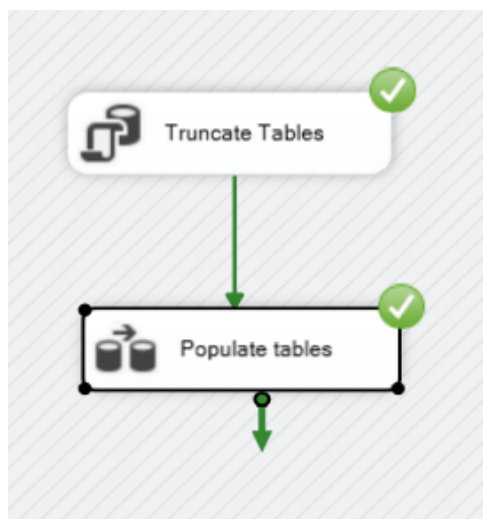
9.1.1. Extract

Első lépésként egy Execute SQL Taskot futtatunk le. Ennek a feladata, hogy ha rákapcsolódunk a STAGE rétegre mindig kiürítjük az összes tábla tartalmát, amelynek kódját a 8. ábra mutatja.



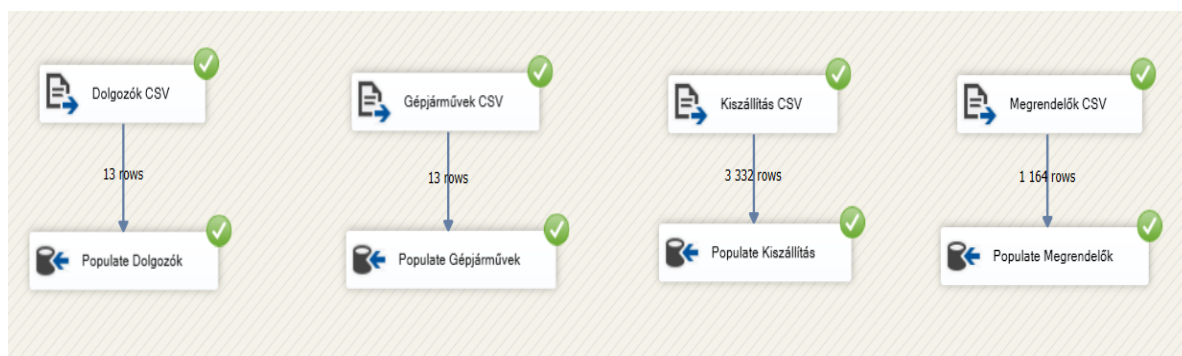
8. ábra Truncate table

A táblák kiürítését követően elkezdődhet az adatok betöltése, amelyet a 9. ábrán figyelhetünk meg. Ebben a rétegben az adatok típuskonverzió nélkül, szöveges formában kerülnek betöltésre. Mivel az adataink csv formátumban vannak, ezért a Flat File Source segítségével tudjuk beolvasni őket majd egy OLE DB Destination használatával töltjük be a megfelelő helyre, amely a 10. ábrán látható.



9. ábra Extract folyamat

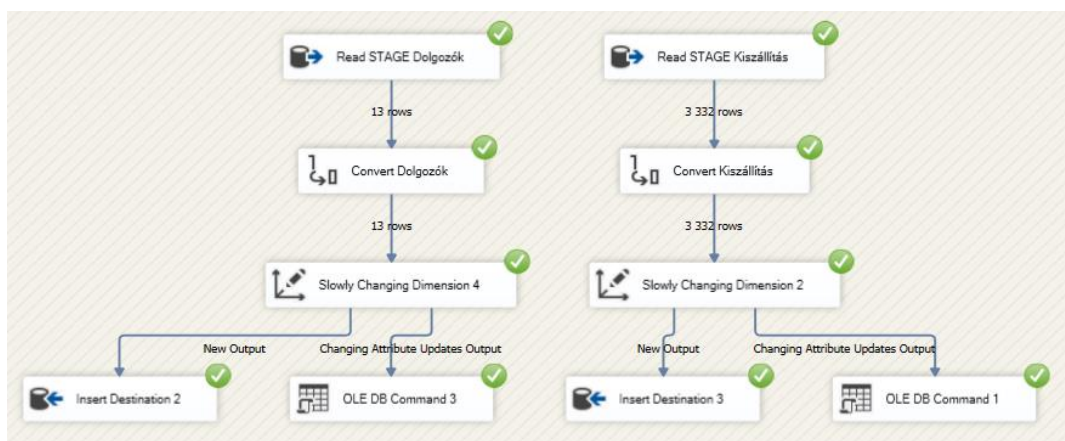
Az adatbázis nem rendelkezik semmilyen megszorítással, így lehetővé válik a folyamatok párhuzamos futása, ami csökkenti a program futási idejét. Ezen folyamatok során a fájlból áttöltöttük az adatokat a STAGE adatbázisba.



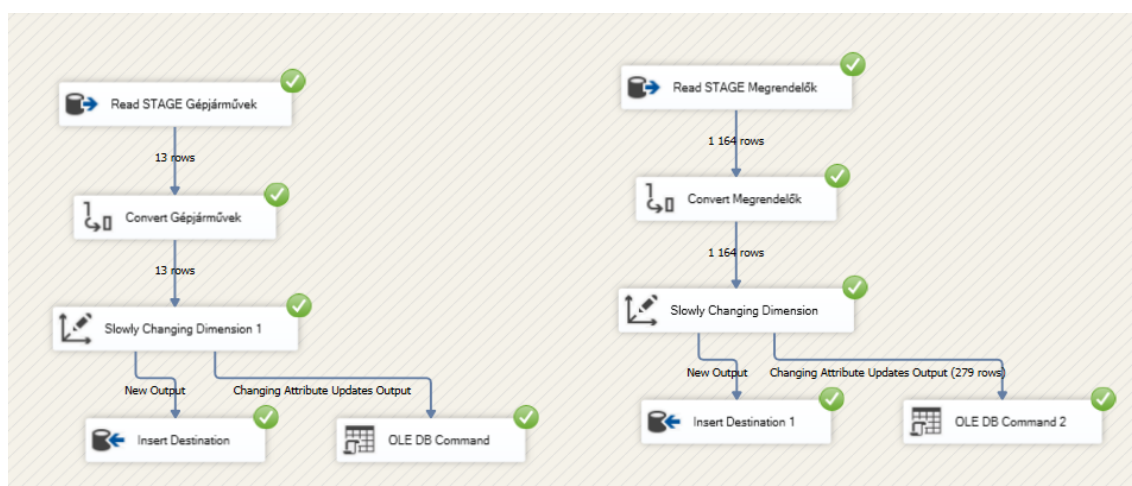
10. ábra Extract réteg betöltés

9.1.2. Transform

Ebben a rétegben történik a STAGE adattábláinak az áttöltése az adattárházba, amelyeket a 11. és a 12. ábrán láthatunk. Itt foglalkozunk azzal, hogy az adatok megfelelőképpen legyenek konvertálva. A típusokat a forrásadatbázisnak megfelelően szükséges kiválasztani a problémák elkerülése érdekében. Egy Data Flow Task segítségével lehetséges feltölteni a DW adatbázis tábláit.



11. ábra Transform I.

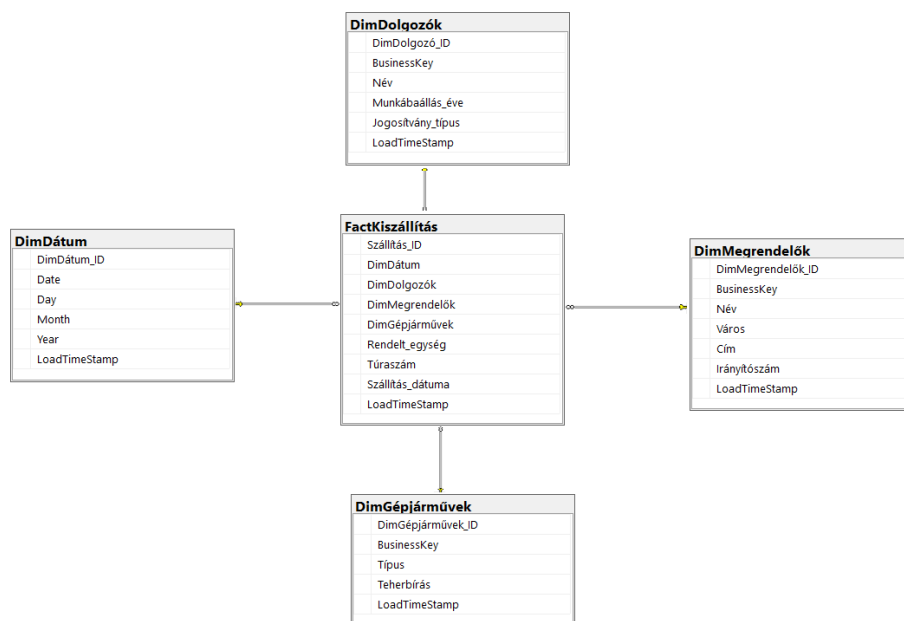


12. ábra Transform II.

Az ábrákon megfigyelhető egy Slowly Changing Dimension(SCD) metódus, amiről szükségeszerű néhány információt megosztani. Ennek a tool-nak az a lényege, hogy nyomon követi a változásokat. Több fajtát is megkülönböztetünk belőle. Én az úgynevezett Type-1 változatot használtam, ami azt jelenti, hogy egy esetleges változást követően egyszerűen felülírja a meglévő adatot az új adattal. Így az adattárház továbbra is naprakész állapotban van.

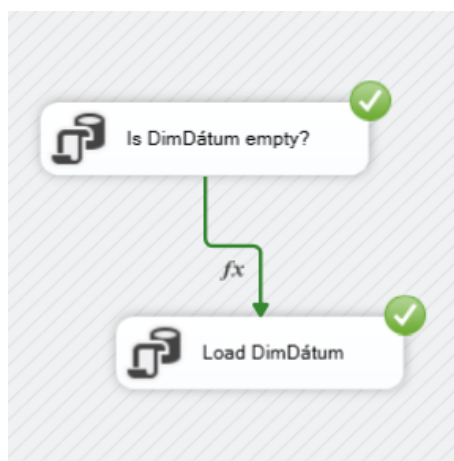
9.1.3. Load

A Load rétegben történik meg a csillag séma feltöltése, amelyet a 13. ábrán figyelhetünk meg.



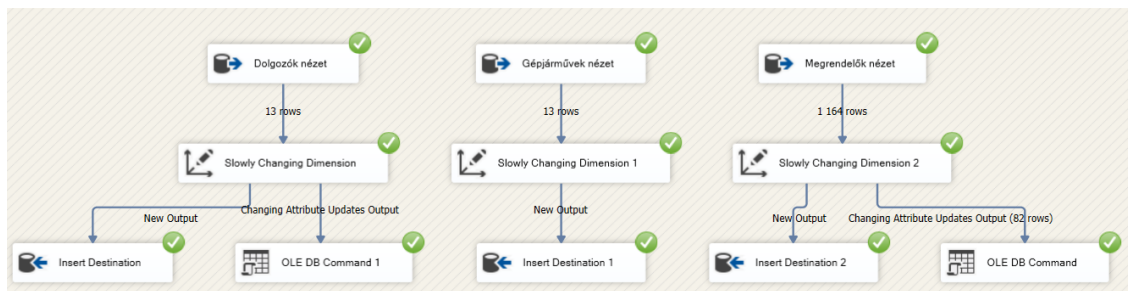
13. ábra Csillag séma adatai

Ehhez először létre kell hozni a DW adatbázisban a dimenzió táblák és a ténytábla feltöltéséhez szükséges nézeteket. Ezt követően első lépésként megvizsgáljuk, hogy a dátum dimenzió fel van-e töltve a szükséges adatokkal. Ezt a 14. ábra mutatja be. Ezt SSIS-ben lehet ellenőrizni. Amennyiben a dimenzió üres, úgy egy tárolt eljárás segítségével fel kell tölteni.



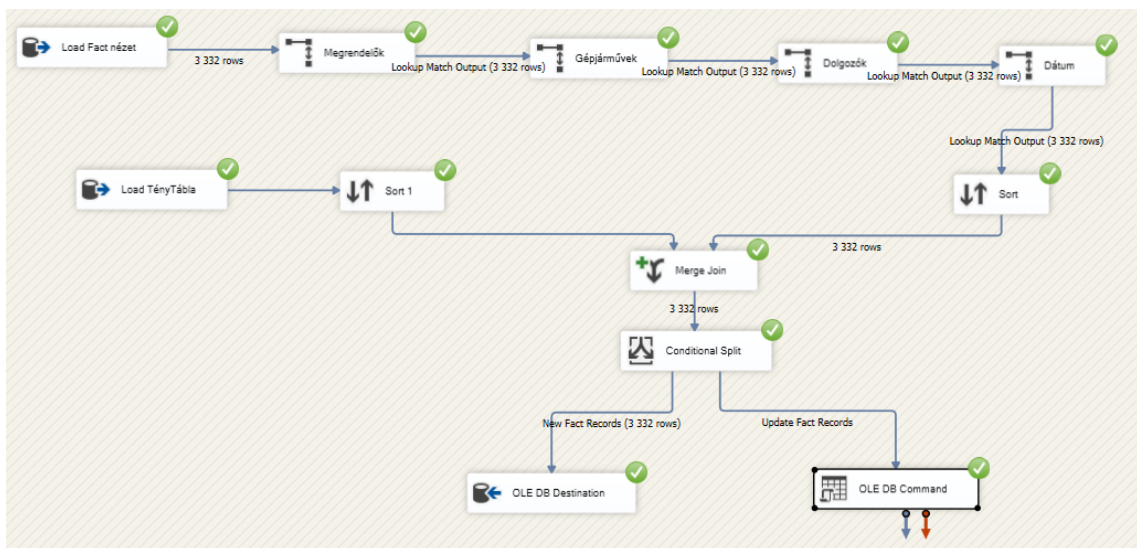
14. ábra Dátum dimenzió feltöltése

Ezt követően az elkészült nézetek segítségével elkezdjük feltölteni a dimenziókat adatokkal, amit a 15. ábrán láthatunk. Az adatok forrása a létrehozott nézetek lesznek és egy SCD modul használatával tudjuk a legfrissebb adatokkal feltölteni a tábláinkat.



15. ábra Dimenzió táblák feltöltése

A dimenzió táblákon való sikeres adatbetöltés után már csak a ténytábla feltöltése van hátra az ETL folyamatból, amely a 16. ábrán figyelhető meg. Ez talán az egyik legnehezebb és legösszetettebb feladat az egész folyamat során. Először a ténytábla nézetéből beolvassuk az adatokat, majd Lookup elemek segítségével minden tény rekordhoz meg kell keresni a hozzá tartozó dimenziók elsődleges kulcsát. Ha ezzel megvagyunk, meg kell vizsgálni, hogy a ténytáblában már megtalálhatóak-e az adatok. Amennyiben nem szerepelnek a táblában, úgy be kell őket szűrni, viszont szereplés esetén, úgy, hogy csak módosítás történt egy OLE DB Command modul használatával le kell futtatni egy frissítés utasítást, ezáltal a rekord oszlopai frissülnek.

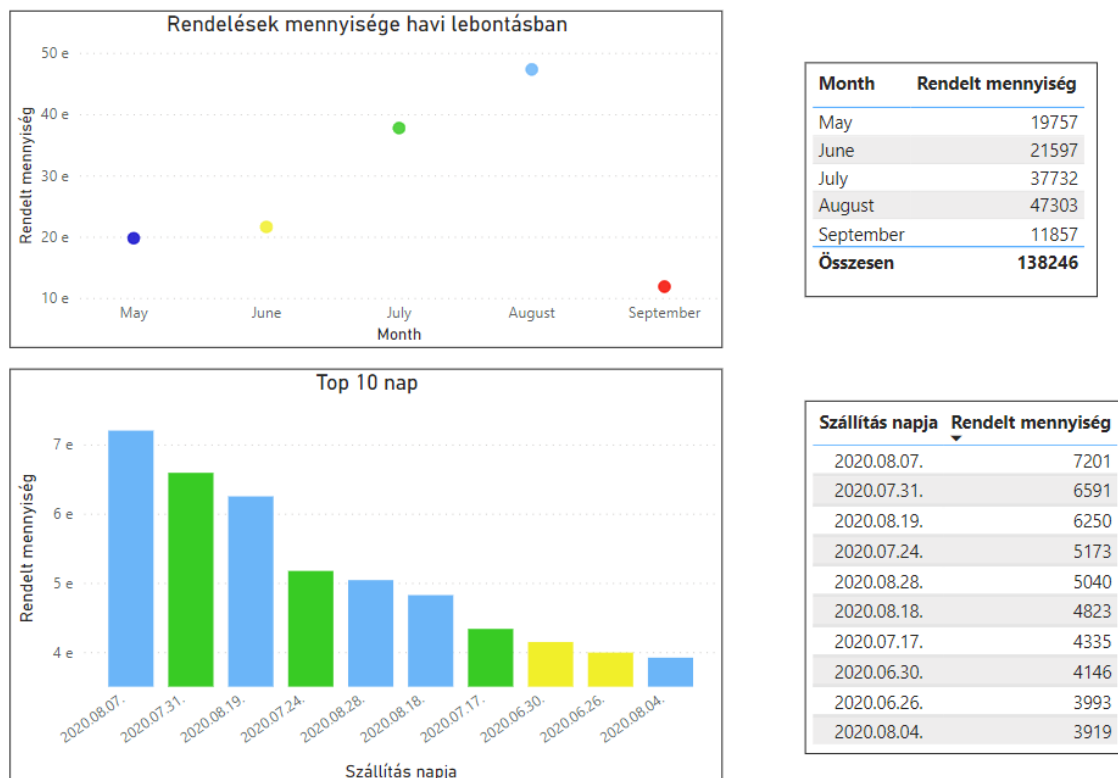


16. ábra Ténytábla feltöltése

10.RIPORTERKÉSZÍTÉS

A követelmények és a megbeszéltek alapján a cég számára fontos vizualizációk elkészítése volt a feladatom az adattárház megtervezése és megvalósítása után.

10.1. Melyik időszakban volt a legtöbb megrendelés?



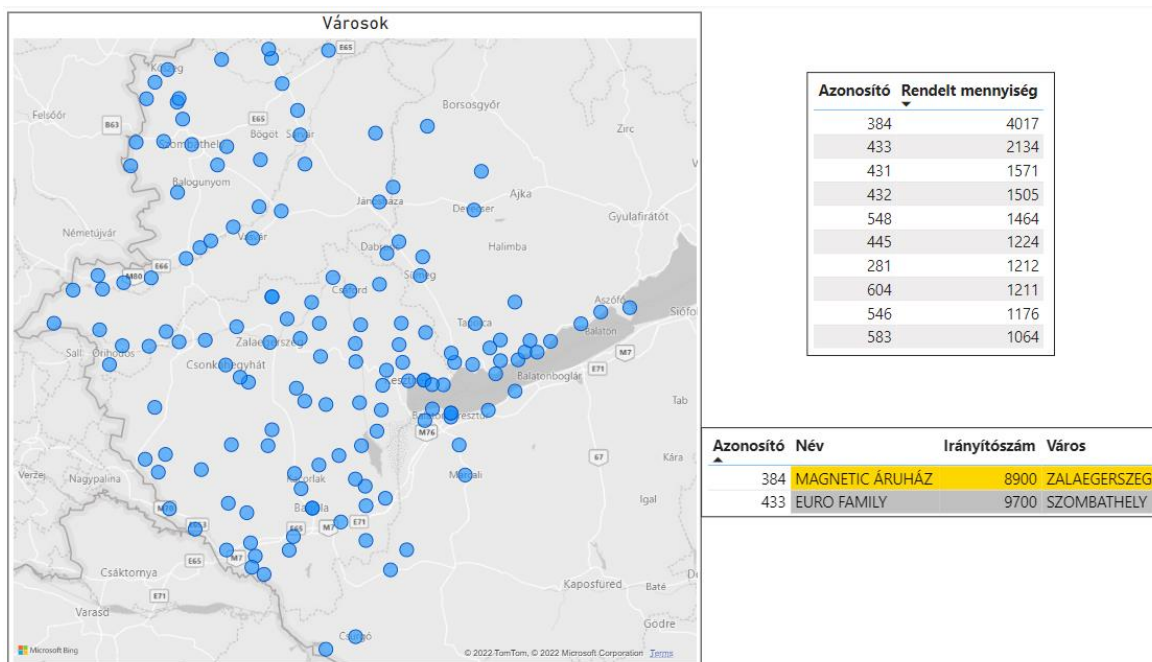
17. ábra Melyik időszakban volt a legtöbb megrendelés?

A vállalkozás számára lényeges volt, hogy az adatok feldolgozását követően bizonyos kérdésekre választ kapjon. Fontosnak tartották, hogy egy bővített nyári időszakban melyik az a hónap vagy esetleg nap amikor a legtöbb szállítás történt. Ezeket az információkat a 17. ábrán láthatjuk.

A felső részben hónapokra lebontva látható az, hogy mennyit rendeltek a partnerek. Nem nagy meglepetésre az augusztus hónap lett a legeredményesebb, hiszen általában akkor van a legmelegebb és akkor van szükség leginkább a frissítő italokra. Szépen kivehető a diagramból, hogy a nyár érkezésével elége megnőtt a rendelések mennyisége és meglehetősen jelentős növekedés tapasztalható a hónapok során. Ezt követően a nyár végeztével elég nagyot csökkent az a szám, ami a rendelt mennyiséget mutatja.

Az alsó oszlopdiagramon annak a tíz napnak az adatai láthatóak, amikor a legtöbb rendelt áru került kiszállításra. A színek segítségével könnyen kivehető, hogy nem véletlen az augusztusi eredmény, hiszen a top 10-ben is öt augusztusi dátum szerepel.

10.2. Merre találhatóak a megrendelők boltjai, illetve melyik partner rendelte a legtöbb mennyiséget?

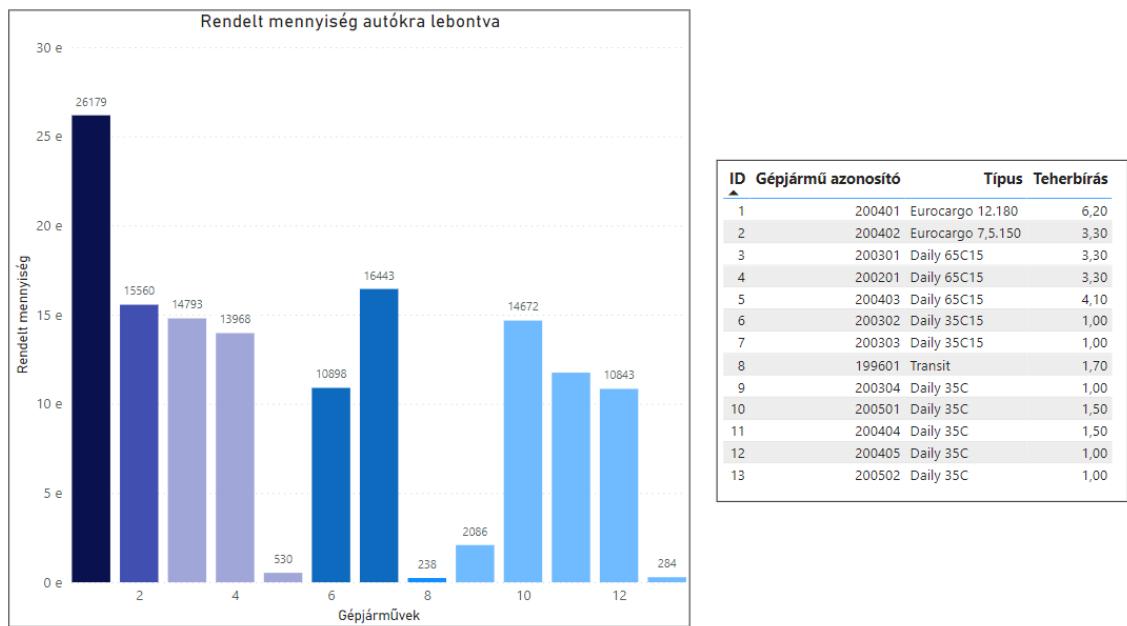


18. ábra A boltok elhelyezkedése és a legtöbb terméket rendelt üzlet

A tulajdonos egyik kérése az volt, hogy hiába rendelkeznek az autók GPS rendszerrel jobban átlátná a dolgokat, ha valamilyen eszköz segítségével megnézné, hogy merre szállítanak árut az emberei. Ez okból készült el a 18. ábrán látható térkép, ami megmutatja az egyes boltoknak az elhelyezkedését. Fontos kiemelni, hogy mind a cég, mind a lerakó, ahonnan az árukat szállítják zalaszentgróti telephellyel rendelkezik. Ennek tudatában jobban érthető, hogy miért is az ország nyugati részébe történnek a kiszállítások. Zala megyét igazából teljesen lefedik a sofőrök. Emellett Vas megye területeire is rendszeresen járnak az autók, valamint főként a nyári időszakban a Balaton környéki településeken is feltűnnek.

A kép jobb oldalán azoknak a boltoknak az azonosítóit lehet látni, ahol a rendelt áruk mennyiségének a száma a legjobbak között van. Azonosító alapján készítettem egy táblázatot is, melyben az első két helyezett üzlet adatai találhatóak. Ezt a listát két jól ismert megyeszékhelyi üzletlánc vezeti és korán sem meglepő, hiszen egy büfé vagy kiskereskedés nem tudja felvenni ekkora áruházakkal a versenyt, de nyilván nem is az a céljuk nekik.

10.3. Mennyi árut szállított egy-egy autó összesen?



19. ábra Szállított áruk mennyisége gépjárművekre lebontva

Árufuvarozással foglalkozó cég számára mindig is elsőfokú figyelmet kapnak a gépjárművek, hiszen az a munkaeszköz. Ha nincs mivel szállítani az árut, akkor nem lesz bevétel és nem lesz nyereség, amit egy cég sem szeretne. Ebből kifolyólag arra kértek, hogy az adatokból derítsem ki a cég rendelkezésére álló autóknek a leterheltségét, illetve azt is, hogy mennyire vannak egyensúlyban az egyes autók használatuk szerint, esetleg szükséges lenne-e változtatni az autók beosztásán.

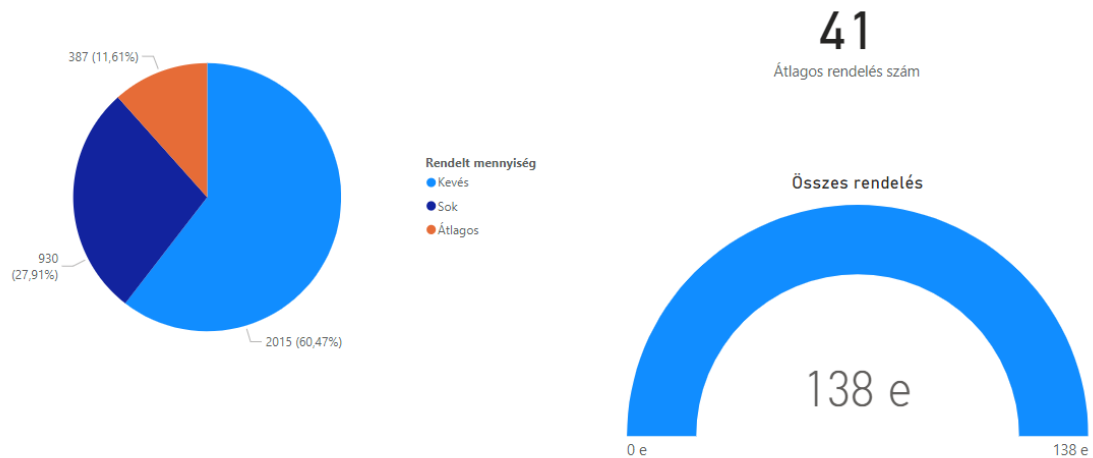
A 19. ábrán már az első pillantásra szembetűnik, hogy nem a legjobb elosztásban történnek a fuvarok kiosztása. Nyilvánvaló, hogy előfordulnak különbségek, ugyanis az autók terhelhetősége és teljesítménye között is vannak eltérések. Azonban, ha részletesebben megnézzük a jobb oldali kis táblázatot és csak az utolsó két sorra tekintünk rá, látjuk, hogy azonos autókról beszélhetünk. A diagramot látva mégis óriási különbséget vélhetünk felfedezni az autók használata kapcsán.

Az cseppet sem meglepő, hogy a legnagyobb autó szállította főként a legtöbb árut, viszont, ha megnézzük a második legmagasabb számot és a hozzá tartozó teherautót, akkor igencsak meglepődhetünk, ugyanis egy, a legkisebb teherbírással rendelkező autó oszlopához tartozik az az érték.

Összességében nem a legjobb elosztásban történtek a szállítások. Ezen a területen lehet még bőven fejlődni és láthatjuk, hogy mekkora segítséget tud nyújtani egy egyszerű ábra.

10.4. Ki mennyi terméket rendel átlagosan?

Megrendelt mennyiségek aránya a kiszállításokhoz képest



20. ábra Ki mennyi terméket rendel átlagosan?

A 20. ábrán látható diagramm arra ad választ, hogy az egyes megrendelők átlagosan mennyit rendelnek és az a szám csoportosítva hova is tartozik. A két egyszerű diagramról leolvasható, hogy összesen nagyjából 138 ezer termékrendelés történt és a rendelések során átlagosan 41 terméket választanak ki a partnerek.

A kördiagram elkészítését megelőzően csoportosítottam a vevőket, abból a szempontból, hogy ki mennyi terméket vesz egy-egy rendelés során. Jól látható, hogy a partnerek kétharmad része minimális mennyiségű terméket szokott rendelni. Ez számokban azt jelenti, hogy kevesebb, mint 35 terméket kér egy rendelés során. Jóval kevesebben választották az átlagos mennyiségű megrendelést. Ebbe a rétegbe azok a vevők tartoznak, akik egy vásárlás között átlagban 35 és 45 terméket választanak. A maradék közel 28% pedig az átlagnál több terméket tesz a kosárba.

11.WEBOLDAL TERVEZÉSE ÉS MEGVALÓSÍTÁSA

Az általam megkeresett cég viszonylag rövid időn belül érdemleges növekedésen ment keresztül, hiszen néhány éve egyszemélyes vállalkozásként indult, mára pedig több, mint 15 főt foglalkoztat. A folyamatos fejlődés és növekedés kapcsán jutottak arra az elhatározásra, hogy a közeljövőben szeretnének jobban nyitni az online világra. A közösségi felületeken való megjelenés mellett szeretnének egy saját weboldalt, ahol bővebb és gyorsabb tájékoztatást tudnának adni az érdeklődő partnerek, megrendelők számára. Ez ügyben meg is kértek engem, hogy ha időm és energiám engedi tervezzek meg egy kezdetleges weboldalt, amiből ki lehet majd indulni és egy jó alapot biztosít arra az időre, amikor megtörténik az elhatározás a saját oldal iránt.

11.1. Tervezés

Legelső lépésként megnéztem pár hasonló, áru fuvarozással foglalkozó cég oldalát és próbáltam onnan gyűjteni néhány ötletet, amit esetleg én is fel tudok használni a későbbiekben. Ezt követően megkérdeztem a tulajdonost, hogy milyen elvárásai vannak a weboldallal kapcsolatban. Elmondta az ő elképzelését, hogy mit szeretne látni, majd pedig megmutattam neki az addigra elkészült terveimet. Ezután kértem, hogy biztosítson kellő tartalmat, gondolok itt mind szövegre, mind képekre.

11.2. Megvalósítás

A weboldal elkészítéséhez a Microsoftnak a Visual Studio alkalmazását használtam. Azért választottam ezt az eszközt, mivel rendkívül egyszerű a használata, emellett felhasználóbarát, több kisegítő lehetőséggel is rendelkezik, valamint valós időben látom a változtatásokat.

Magát a weboldalt HTML (HyperText Markup Language) nyelven írtam, ami egy leíró nyelv és előszeretettel használják weboldalak készítésekor. Emellett még CSS (Cascading Style Sheets) kóddal tettem szebbé a kinézetet.

Kezdőoldal

OSBÁTH FUVAR KEZDŐOLDAL GÉPJÁRMŰVEK TÖRTÉNETÜNK KAPCSOLAT

**Megbízható, pontos, gyors fuvarozó céget keres?
Megtalálta!**

Köszönti Önt az Osbáth fuvar csapata!

Munkatársainkkal azon dolgozunk, hogy a megrendelő kérését 100%-ban teljesítsük.
Ha kérdése van forduljon hozzánk bizalommal.

+36 30 7574421
balazs.fuvar@gmail.com

Ömlesztett áru-szállítás
Kavics, homok,
kő, murva,
hulladék.
Gabona,
takarmány.

Silós szállítás
Cement, mészkő,
gipsz, vakolat.
Takarmány, liszt.

Raklapos áruk
Építőanyagok,
önrakodós és
fixes autókkal.
Takarmány, egyéb
darabáru.

Egyéb szállítás
Faszállítás,
veszélyes anyag
szállítás
Költöztetés.

21. ábra Kezdőoldal

A 21. ábrán a vállalkozás kezdőoldala látható. Ennek létrehozásakor figyelembe kellett vennem, hogy az érdeklődő ezzel a felülettel találkozik először, amikor megnyitja majd a weboldalt. Úgy kellett megalkotnom, hogy barátságkeltő, mégis kellően határozott megjelenést keltsen. Az első benyomás nagyon fontos mindenhol, természetesen itt is. Sok minden eldől abban a pillanatban, amikor valaki megnyitja az oldalt. Ha negatív élménnyel gazdagodik a felhasználó, könnyen lehet többet meg se látogatja az oldalunkat.

Minden oldal tetején megtalálható egy menüsáv és ezzel egy sorban egy logó is. A fő oldalon helyet kapott egy köszöntő szöveg, ami üdvözli a felhasználót és felhívja a figyelmét, hogy többet nem is kell keresgélennie, megtalálta a legjobb megoldást a problémájára. Ezen felül megtalálhatók a cég elérését biztosító adatok, valamint az első oldalon látható a cég szolgáltatásainak listája, hogy milyen lehetőségeket tud nyújtani a megrendelők számára, szépen rendszerezve, kategorizálva.

Gépjárművek



OSBÁTH
FUVAR

KEZDŐOLDAL GÉPJÁRMŰVEK TÖRTÉNETUNK KAPCSOLAT

 Iveco Daily 65c15 Össztömeg: 6,5t Terhelhetőség: 3,3t Évjárat: 2004 Hengerűrtartalom: 2800 cm³	 Ford Transit Össztömeg: 3,5t Terhelhetőség: 1,7t Évjárat: 1996 Hengerűrtartalom: 2500 cm³	 Pediküre Össztömeg: 3,7t Terhelhetőség: 1,5t Évjárat: 2002 Hengerűrtartalom: 2500 cm³
 Iveco Daily 65c15 Össztömeg: 6,5t Terhelhetőség: 3,3t Évjárat: 2005 Hengerűrtartalom: 2800 cm³	 Iveco Daily 65c15 Össztömeg: 6,5t Terhelhetőség: 3,3t Évjárat: 2003 Hengerűrtartalom: 2800 cm³	 Iveco Eurocargo 7,5.150 Össztömeg: 7,5t Terhelhetőség: 3,3t Évjárat: 2004 Hengerűrtartalom: 4000 cm³

22. ábra Gépjárművek oldal

A 22. ábrán a cég tulajdonában álló teherautókról kaphatunk bővebb tájékoztatást. Az érdeklődő képek segítségével könnyebben eligazodik az autók sokaságában, emellett olyan információkat megtudhat a cég autóiról, mint például milyen teljesítményt tud nyújtani, mekkora súlyt tud szállítani, valamint maga az autók típusa is olvasható a felületen.

Történetünk



OSBÁTH
FÚVAR

KEZDŐOLDAL GÉPJÁRMŰVEK TÖRTÉNETÜNK KAPCSOLAT

Cégtörténet



Kezdetek

Tulajdonos Osbáth Balázs

A vállalkozás 2013 nyarán indult egy Ford Transit típusú gépjármű vásárlásával. Ekkor még csak szezonális, 3 hónapra szóló szerződés volt a Coca-Cola HBC Magyarország Kft-vel. Itt még csak egy egyszemélyes vállalkozásról beszélhettünk.



Telephely

Zalaszentgrót

Pár év alatt rengeteget változott a vállalkozás. Évről évre egyre csak bővült a autóflootta, ezért 2016-ban szükséges volt egy nagyobb telephely birtokba vétele. Itt kényelmesen elfértek már a meglévő gépjárművek és a szereléshez is biztosítva voltak a körülmények.



Jelen

2022

Jelenleg a vállalkozás 14 gépjárművel rendelkezik és több, mint tíz főt foglalkoztat. Legfőbb tevékenység még mindig az áruiterítés, azonban már a szolgáltatások között megtalálható veszélyes áru szállítás is. Emellett van egy darus autó, ami leginkább fahordások során van használatban. Foglalkozunk még természetesen költöztetéssel és vállalunk egyéb magánfuvarokat is.

© 2022 Osbáth fuvar

23. ábra Történetünk oldal

A harmadik oldal a vállalkozás történetet foglalja össze röviden, amelyet a 23. ábra szemléltet a kezdetektől egészen a jelenig. Fénykép segítségével a felhasználó megtudhatja, hogy melyik járművel is indult el maga a vállalkozás, amikor még csak az az egy autó rótt az utakat. Azt követően a telephely képe tárulkozik a vendég szeme elé, ami szintén bizonyítja, hogy szépen fejlődött és most is fejlődik a cég. Az utolsó részben pedig az aktualitásokról olvashat a kedves látogató, hogy jelenleg mennyi gépjárművel rendelkezik a cég és azokat mire használja.

Kapcsolat



OSBÁTH
FUVAR

KEZDŐOLDAL GÉPJÁRMŰVEK TÖRTÉNETÜNK KAPCSOLAT



Telephely
Május 1. utca 4.
8790 Zalaszentgrót

Telefon
 +36 30 7574421

Email
 balazs.fuvar@gmail.com

Kapcsolat

Figyelem! Ne felejtse el a kitölteni a megjegyzés mezőt!

A választ: ☐ e-mailben kérem ☐ telefonon kérem.



Keressen minket bizalommal a hét
bármelyik napján, hétvégén is.

© 2022 Osbáth fuvar

24. ábra Kapcsolat oldal

Az utolsó oldal nem más, mint a kapcsolatok, amely a 24. ábrán látható. Itt található bővebb információ arról, hogyan is lehet a céggel kapcsolatba lépni, legyen szó telefonról, emailről vagy akár a kapcsolat form kitöltésével. Az oldal tetején egy beágyazott térkép található, ezzel is segítve az érdeklődőt a pontosabb helymeghatározáshoz.

12.TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK

- Weboldal éles rendszerbe helyezése
- Weboldal backend részének megírása
- Riportok megjelenítése a weboldalon mind a munkatársak, mind az érdeklődők számára
- Az adatbázis rendszeres karbantartására szolgáló szoftver bevezetése

13.ÖSSZEFOGLALÁS

A szakdolgozatom célja, hogy a munkámmal segítsek az általam választott árufuvarozással foglalkozó vállalkozás üzleti folyamataiban. Annak érdekében, hogy a lexikális tudásom is megfelelő legyen, első lépésként a szükséges irodalomkutatást végeztem el, ahol sok érdekes és új dolgot ismerhettem meg a különböző témákkal kapcsolatban. Amikor már úgy éreztem, hogy megvan a megfelelő háttér, elkezdtem a fejlesztéseket, amiket az előre megbeszélt követelmények és célok alapján készítettem el. Ennél a résznél az első feladatom az volt, hogy a forrásadatok alapján megtervezzek és megvalósítsak egy adatbázist. Ezt követően egy adattárház létrehozása várt rám, mint következő feladat. A sikeres adatkinyerést követően, az volt a feladatom, hogy ezeket az adatokat a megfelelő típusú és formátumú alakítsam át. Végül jöhetett a már kész és megfelelő adatok adattárházba történő betöltése. Ezzel az első fő feladatom be is fejeződött, de természetesen vártak még rám megoldandó dolgok. Az adattárház sikeres megvalósítását követte a különböző vizualizációk elkészítése. A munkám ezen szakaszában is a kapott követelmények alapján készítettem el a szükséges riportokat. Annak érdekében, hogy minél jobban és könnyebben lehessen az információk birtokába jutni, színes, egyszerű és egyértelmű diagramok elkészítése történt a részemről. Végül az utolsó fejlesztéshez érkeztem. Egy weboldal készítése várt rám, mint utolsó, megvalósítandó feladat. Azért kért meg ezzel a feladattal a tulajdonos, mert a közeljövőben szeretne jobban belépni az online világba és az érdeklődők részére is könnyebb megtalálást biztosítani. Ennek érdekében hoztam létre az oldalt, amit színes képekkel és megfelelő információkkal láttam el.

14.SUMMARY

The purpose of my thesis is to provide help for the business processes of the chosen freight forwarding company. In order to ensure that my lexical knowledge was adequate, I first did the necessary research, during which I was able to learn many interesting and new things about different topics. Once I felt that I had the proper background, I started the development process, based on the requirements and objectives that I had discussed beforehand. At this point, my first task was to design and implement a database based on the source data. After that, the next step was to set up a data warehouse. After successful data gathering, my task was to convert this data into the appropriate type and format. Finally, it was time to load the prepared and correct data into the data warehouse. This was the end of my first main task, but of course there was still more to be done. The successful implementation of the data warehouse was followed by the creation of the various data visualisations. During this stage I also prepared the necessary reports based on the received requirements. In order to make the information as accessible and easy to understand as possible, I created colourful, simple and clear diagrams. Finally, I reached the last part of the development, which was creating a website. The owner of the company requested this project because in the near future , he wants to enter the online world to reach more customers and to gain more visibility. To achieve these goals, I have created a website, containing lots of meaningful information and colorful images.

ÁBRAJEGYZÉK

1. ábra Kimball modell ⁽¹⁸⁾	15
2. ábra Inmon modell ⁽¹⁸⁾	16
3. ábra Csillag séma ⁽¹⁹⁾	17
4. ábra Galaxis séma ⁽²⁰⁾	17
5. ábra Hópehely séma ⁽²¹⁾	18
6. ábra ETL folyamat ⁽²²⁾	20
7. ábra Adatbázis terv	26
8. ábra Truncate table	30
9. ábra Extract folyamat	31
10. ábra Extract réteg betöltés	31
11. ábra Transform I.	32
12. ábra Transform II.	32
13. ábra Csillag séma adatai	33
14. ábra Dátum dimenzió feltöltése	33
15. ábra Dimenzió táblák feltöltése	34
16. ábra Ténytábla feltöltése	34
17. ábra Melyik időszakban volt a legtöbb megrendelés?	35
18. ábra A boltok elhelyezkedése és a legtöbb terméket rendelt üzlet	36
19. ábra Szállított áruk mennyisége gépjárművekre lebontva	37
20. ábra Ki mennyi terméket rendel átlagosan?	38
21. ábra Kezdőoldal	40
22. ábra Gépjárművek oldal	41
23. ábra Történetünk oldal	42
24. ábra Kapcsolat oldal	43

TÁBLÁZATJEGYZÉK

1. táblázat Gépjárművek tábla	27
2. táblázat Dolgozók tábla	28
3. táblázat Megrendelők tábla	28
4. táblázat Kiszállítás tábla	29

RÖVIDÍTÉSEK:

HTML: HyperText Markup Language
CSS: Cascading Style Sheets
ETL: Extract, Transform, Load
IBM: International Business Machines
RDBMS: Relational Database Management System
SQL: Structured Query Language
NoSQL: Not only Structured Query Language
JSON: JavaScript Object Notation
XML: Extensible Markup Language
POSIX: Portable Operating System Interface for uniX
ROLAP: Relational Online Analytical Processing
SSIS: SQL Server Integration Services
SCD: Slowly Changing Dimension

IRODALOMJEGYZÉK

- [1] Adam Hughes, Craig Stedman: Microsoft SQL Server
(<https://searchdatamanagement.techtarget.com/definition/SQL-Server>), utoljára megtekintve: 2021. 11. 19.
- [2] Magnus Hagender, Jonathan S. Katz: About PostgreSQL
(<https://www.postgresql.org/about/>), utoljára megtekintve: 2021.12.08.
- [3] (ismeretlen szerző) Introduction to Redis
(<https://redis.io/topics/introduction>), utoljára megtekintve: 2021.12.08.
- [4] (ismeretlen szerző) What is Apache Cassandra?
(https://cassandra.apache.org/_/cassandra-basics.html), utoljára megtekintve: 2021.12.08.
- [5] Ralph Kimball, Margy Ross: The Data Warehouse Toolkit - Second Edition. John Wiley & Sons, Inc., 2002
- [6] Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker: The Data Warehouse Lifecycle Toolkit. John Wiley & Sons, Inc., 2008
- [7] W. H. Inmon: Building the Data Warehouse, John Wiley & Sons, Inc., 2002
- [8] (ismeretlen szerző) What is Extract, Transform, Load? Definition, Process and Tools
(<https://www.talend.com/resources/what-is-etl/>), utoljára megtekintve: 2021.12.08.
- [9] Kate Brush, Ed Burns: Data visualization
(<https://searchbusinessanalytics.techtarget.com/definition/data-visualization>), utoljára megtekintve: 2021.12.09.
- [10] Jesse Scardina, Lauren Horwitz: Microsoft Power BI
(<https://searchcontentmanagement.techtarget.com/definition/Microsoft-Power-BI>), utoljára megtekintve: 2021.12.09.
- [11] Jake VanderPlas: Visualization with Matplotlib
(<https://jakevdp.github.io/PythonDataScienceHandbook/04.00-introduction-to-matplotlib.html>), utoljára megtekintve: 2021.12.10.
- [12] (ismeretlen szerző) SAP Lumire Tutorial
(https://www.tutorialspoint.com/sap_lumira/index.htm), utoljára megtekintve: 2021.12.10.
- [13] (ismeretlen szerző) Google Charts
(<https://www.predictiveanalyticstoday.com/google-charts/>), utoljára megtekintve: 2021.12.10.

- [14] (ismeretlen szerző) iCEDQ Standard Edition
(<https://icedq.com/standard-edition>),
utoljára megtekintve: 2021.12.10.
- [15] (ismeretlen szerző) SAP Data Integrator
(<https://www.sapstore.com/solutions/46899/SAP-Data-Integrator---premium-edition>),
utoljára megtekintve: 2021.12.10.
- [16] (ismeretlen szerző) What Is Pentaho Data Integration (Kettle): All There Is To Know(<https://www.spec-india.com/blog/pentaho-data-integration-kettle>)
utoljára megtekintve: 2021.12.10.
- [17] (ismeretlen szerző) SQL Server Integration Services
(<https://docs.microsoft.com/hu-hu/sql/integration-services/sql-server-integration-services?view=sql-server-ver15>), utoljára megtekintve: 2021.12.10.
- [18] (ismeretlen szerző) Kimball Model
(<https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/>), utoljára megtekintve: 2021.12.10.
- [19] (ismeretlen szerző) A csillagséma és a Power BI-ban játszott szerepének a bemutatása
(<https://docs.microsoft.com/hu-hu/power-bi/guidance/star-schema>), utoljára megtekintve: 2021.12.10.
- [20] (ismeretlen szerző) Fact Constellation in Data Warehouse Modelling
(<https://www.geeksforgeeks.org/fact-constellation-in-data-warehouse-modelling/>),
utoljára megtekintve: 2021.12.10.
- [21] (ismeretlen szerző) Introduction to Snowflake Schema
(<https://www.educba.com/snowflake-schema/>),
utoljára megtekintve: 2021.12.10.
- [22] (ismeretlen szerző) How to Automate the ETL Process For Data From Magento & Google Analytics
(<https://medium.datadriveninvestor.com/how-to-automate-the-etl-process-for-data-from-magento-google-analytics-72a6f8f0f6f5>),
utoljára megtekintve: 2021.12.10.

MELLÉKLET

```
CREATE DATABASE STAGE COLLATE SQL_Latin1_General_CP1_CI_AS
GO
USE STAGE
GO
CREATE TABLE STAGE.dbo.STA_Dolgozók(
    Dolgozó_ID varchar(256),
    Név varchar(256),
    Munkábaállás_éve varchar(256),
    Jogosítvány_típus varchar(256)
);
CREATE TABLE STAGE.dbo.STA_Gépjárművek(
    Gépjármű_ID varchar(256),
    Rendszám varchar(256),
    Gyártó varchar(256),
    Típus varchar(256),
    Évjárat varchar(256),
    Össz_súly varchar(256),
    Teherbírás varchar(256),
    Hengerűrtartalom varchar(256),
    Szükséges_jogosítvány varchar(256),
    Dolgozó_ID varchar(256)
);
CREATE TABLE STAGE.dbo.STA_Megrendelők(
    Megrendelo_ID varchar(256),
    Név varchar(256),
    Város varchar(256),
    Cím varchar(256),
    Házszám varchar(256),
    Irányítószám varchar(256)
);
CREATE TABLE STAGE.dbo.STA_Kiszállítás(
    Szállítás_ID varchar(256),
    Túraszám varchar(256),
    Megrendelo_ID varchar(256),
    Gépjármű_ID varchar(256),
    Szállítás_dátuma varchar(256),
    Rendelt_egység varchar(256)
);

CREATE DATABASE DW COLLATE SQL_Latin1_General_CP1_CI_AS
GO
USE DW
GO
CREATE TABLE DW.dbo.DW_Dolgozók(
    Dolgozó_ID int,
    Név varchar(50),
```

```

        Munkábaállítás_éve int,
        Jogosítvány_típus varchar(10),
        LoadTimeStamp TIMESTAMP
    );
CREATE TABLE DW.dbo.DW_Gépjárművek(
    Gépjármű_ID int,
    Rendszám varchar(10),
    Gyártó varchar(50),
    Típus varchar(50),
    Évjárat int,
    Össz_súly float,
    Teherbírás float,
    Hengerűrtartalom int,
    Szükséges_jogosítvány varchar(5),
    Dolgozó_ID int,
    LoadTimeStamp TIMESTAMP
);
CREATE TABLE DW.dbo.DW_Megrendelők(
    Megrendelo_ID int,
    Név varchar(100),
    Város varchar(100),
    Cím varchar(100),
    Házszám int,
    Irányítószám int,
    LoadTimeStamp TIMESTAMP
);
CREATE TABLE DW.dbo.DW_Kiszállítás(
    Szállítás_ID int,
    Túraszám int,
    Megrendelo_ID int,
    Gépjármű_ID int,
    Szállítás_dátuma date,
    Rendelt_egység int,
    LoadTimeStamp TIMESTAMP
);

CREATE DATABASE DM COLLATE SQL_Latin1_General_CP1_CI_AS
GO
USE DM
GO
CREATE TABLE DM.dbo.FactKiszállítás(
    Szállítás_ID BIGINT IDENTITY(1,1),
    DimDátum int,
    DimDolgozók int,
    DimMegrendelők int,
    DimGépjárművek int,
    Rendelt_egység int,

```

```

Túraszám int,
Szállítás_dátuma int,
LoadTimeStamp TIMESTAMP
);
CREATE TABLE DM.dbo.DimDolgozók(
    DimDolgozó_ID INT IDENTITY(1,1),
    BusinessKey int,
    Név varchar(50),
    Munkábaállás_éve int,
    Jogosítvány_típus varchar(10),
    LoadTimeStamp TIMESTAMP
);
CREATE TABLE DM.dbo.DimMegrendelők(
    DimMegrendelők_ID INT IDENTITY(1,1),
    BusinessKey int,
    Név varchar(100),
    Város varchar(100),
    Cím varchar(100),
    Irányítószám int,
    LoadTimeStamp TIMESTAMP
);
CREATE TABLE DM.dbo.DimGépjárművek(
    DimGépjárművek_ID INT IDENTITY(1,1),
    BusinessKey int,
    Típus varchar(50),
    Teherbírás float,
    LoadTimeStamp TIMESTAMP
);
CREATE TABLE DM.dbo.DimDátum(
    DimDátum_ID int,
    Date date,
    Day int,
    Month int,
    Year int,
    LoadTimeStamp TIMESTAMP
);

CREATE OR ALTER VIEW dbo.DimDolgozókVW AS
SELECT
    Dolgozó_ID AS [BusinessKey],
    Dolgozó_ID AS [Dolgozó_ID],
    Név AS [Név],
    Munkábaállás_éve AS [Munkábaállás_éve],
    Jogosítvány_típus AS [Jogosítvány_típus]
FROM [DW_Dolgozók];

CREATE OR ALTER VIEW dbo.DimGépjárművekVW AS

```

```

SELECT
    Gépjármű_ID AS [BusinessKey],
    Gépjármű_ID AS [Dolgozó_ID],
    Típus AS [Típus],
    Teherbírás AS [Teherbírás]
FROM [DW_Gépjárművek];

CREATE OR ALTER VIEW dbo.DimMegrendelőkVW AS
SELECT
    Megrendelo_ID AS [BusinessKey],
    Megrendelo_ID AS [Dolgozó_ID],
    Név AS [Név],
    Város AS [Város],
    Cím AS [Cím],
    Irányítószám AS [Irányítószám]
FROM [DW_Megrendelők];

CREATE OR ALTER VIEW dbo.FactKiszállításVW AS
SELECT
    ROW_NUMBER() OVER(ORDER BY Szállítás_ID) AS ID,
    K.Túraszám AS [Túraszám],
    K.Szállítás_dátuma AS [Szállítás_dátuma],
    K.Rendelt_egység AS [Rendelt_egység],
    M.Megrendelo_ID AS [Megrendelo_ID],
    G.Gépjármű_ID AS [Gépjármű_ID],
    D.Dolgozó_ID AS [Dolgozó_ID]
FROM [DW].[dbo].[DW_Megrendelők] M inner join
[DW].[dbo].[DW_Kiszállítás] K ON K.Megrendelo_ID = M.Megrendelo_ID inner join
[DW].[dbo].[DW_Gépjárművek] G ON G.Gépjármű_ID = K.Gépjármű_ID inner join
[DW].[dbo].[DW_Dolgozók] D ON D.Dolgozó_ID = G.Dolgozó_ID;

CREATE OR ALTER PROCEDURE LoadDimDátum
@StartDate DATE ='2020-01-01',
@EndDate DATE ='2020-12-31'
AS
BEGIN
    SET XACT_ABORT ON;
    TRUNCATE TABLE dbo.DimDátum;
    BEGIN TRAN;
    WHILE @StartDate < EndDate
    BEGIN
        INSERT INTO dbo.DimDátum(
            [DimDátum_ID],
            [Date],
            [Day],
            [Month],
            [Year]

```

```

    )
    SELECT [DimDateID] =
YEAR(@StartDate)*1000+MONTH(@StartDate)*100+DAY(@StartDate),
        [Date]=@StartDate,
        [Day]=DAY(@StartDate),
        [Month]=MONTH(@StartDate),
        [Year]=YEAR(@StartDate)
    SET @StartDate=DATEADD(DD,1,@StartDate)
END
SELECT*FROM dbo.DimDátum
COMMIT;
END;

```