Adattárházak és üzleti intelligencia csoportmunka

**Készítette:** Kun Szabolcs István – Data Modeler XOUEFI

Bódi István Márk – Scrum Master NEPTUNKOD

Jákli Viktor – SQL Developer NEPTUNKOD

Torontáli Máté Barnabás – BI Developer NEPTUNKOD

Varga Márk András – ETL Developer GG04UL

**Szak:** Mérnök Informatikus BSc

**Félév:** 2021/2022/1

Tartalomjegyzék

[1. Bevezetés 3](#_Toc84547957)

[1.1. Use-case 3](#_Toc84547958)

[1.2. Adathalmaz 3](#_Toc84547959)

1. Bevezetés

A csapatunk kihívás elé állítja magát, így arra vállalkozott, hogy egy olyan adattárházat építsen ki, mely segítségével nyomon követhetjük, és ezáltal elemezhetjük az Adventure Works Cycles fiktív, multinacionális kerékpár gyártó céget.

Habár a munka során a [Microsoft](https://docs.microsoft.com/en-us/sql/samples/adventureworks-install-configure?view=sql-server-ver15&tabs=ssms) példa adatbázisával dolgozunk, de ettől függetlenül törekszünk arra, hogy egy, olyan rendszert hozzunk létre, amely képes támogatni a vállalatot üzletkritikus döntéshozatalaikban és betekintést gyújtson a különböző értékesítési folyamatokba. A fejlesztés során az adatok tárolásához Docker konténerben hostolt Microsoft SQL Server adott otthon, míg az ETL folyamatok megvalósításához a Microsoft SSIS szolgáltatását vettük igénybe. A Microsoft Power BI segítségével készítettük el a riport réteget, mely egy jól átlátható, különböző személyre szabási és szűrésilehetőséget biztosító felület létrehozásában könnyítette meg a munkánkat.

## Use-case

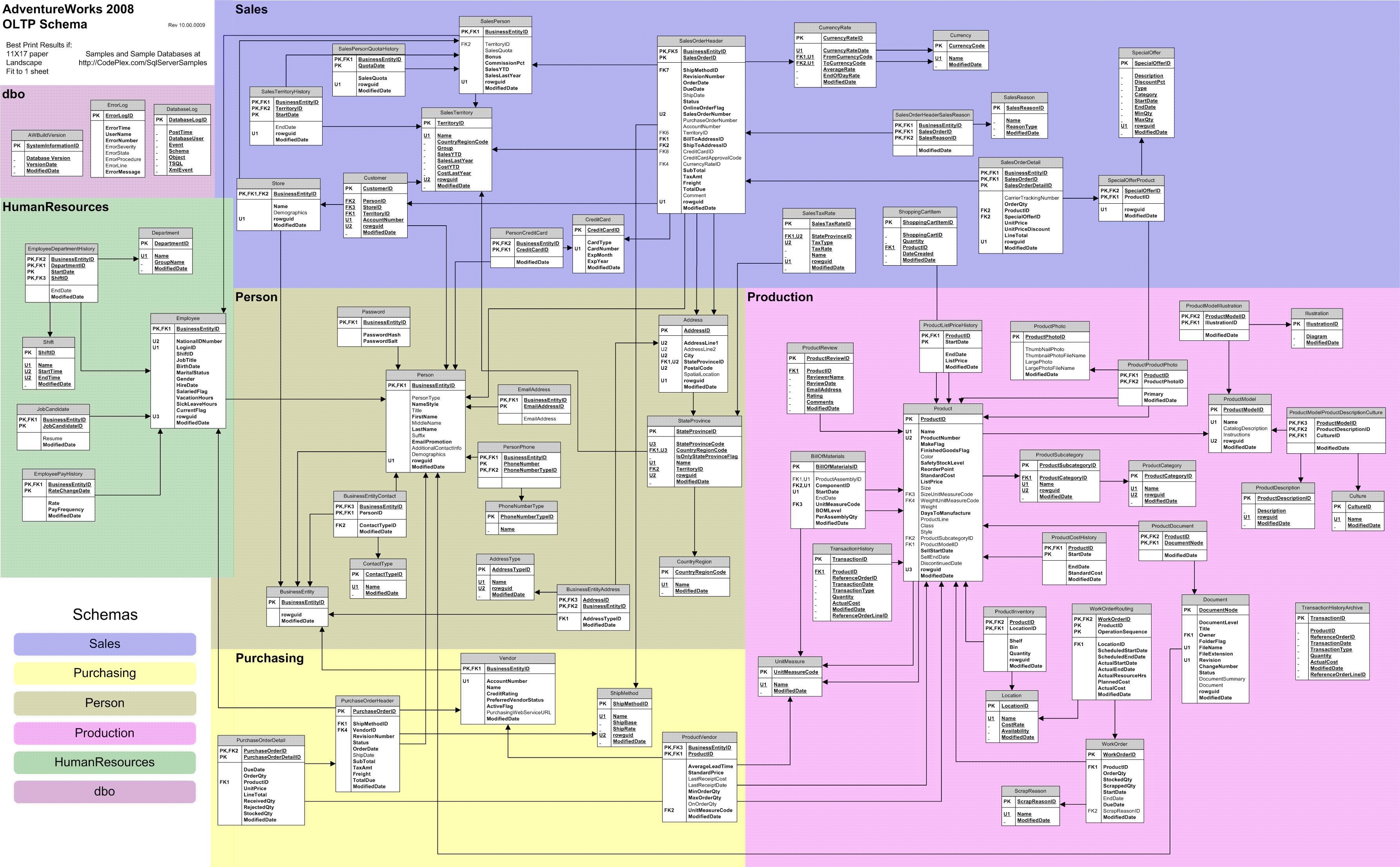
A rendszer segítségével a felhasználó képes lesz a termék, értékesítési kolléga, régió és idő dimenziók mentén az értékesítés adatait (tényeket) kiértékelni.

Példák use-casekre a teljesség igénye nélkül:

* Kimutatás az alkalmazottak értékesítéseiről régiókra bontva.
* Termék kategóriák mentén kimutatni, hogy mely régiókban mennyi az eladott mennyiség.
* Éves kimutatás kollégákra lebontva, hogy ki hogyan teljesített.

## Adathalmaz

Az AdventureWorks példa adatbázis Production, Sales, Person, Purchasing és HumanResources sémáiból használtunk fel táblákat.



1. ábra AdventureWorks adatbázis felépítése

A töltéshez és továbbá a dimenzió táblák és tény tábla létrehozásához az alábbi táblákra van szükség a forrásadatbázisból:

* Product: legyártott termékekkel kapcsolatos információkat tartalmaz, mint például termék tulajdonságok, árak és stb.
* ProductModel: a termék modelleket tartalmazza, a termék főtörzse.
* ProductSubcategory: a termékek alkategóriáit tartalmazza.
* ProductCategory: a termékek főkategóriáit tartalmazza.
* Person: személy entitásokat tartalmaz, így egyben alkalmazottak és vásárlók adatait is, hierarchiában áll az Employee és Customer táblával.
* Employee: alkalmazott specifikus adatokat tartalmaz.
* Customer: ügyfélspecifikus adatokat tartalmaz.
* Store: üzlet információk.
* SalesTerritory: az eladási régiókat tartalmazza.
* SalesPerson: eladókat tartalmazza, akik régiókhoz kötődnek, nem minden értékesítésnél van eladó megjelölve, hiszen rögzítünk online rendeléseket is.
* SalesOrderHeader: a rendelés törzs adatait tartalmazza. (Például: vevő ID, értékesítő ID, rendelés dátuma, státusz, stb.)
* SalesOrderDetail: a rendelés részletesebb adatait tartalmazza, szám mennyiségeket. Egy rendeléshez több is tartozhat, hiszen több tételt is rendelhetünk egyszerre különböző termékekből. (Például: egységárak, kedvezmények, stb.)
* ShipMethod: a konténer szállító cég adatait tartalmazza.



2. ábra Felhasznált táblák a forrás adatbázisból

A csillagséma létrehozása során 1 tény táblát és 6 dimenzó táblát alkottunk meg:

* DimProduct: Product, ProductSubcategory, ProductModel és ProductCategory adatait tartalmazza.
* DimTerritory: SalesTerritory tábla adatait tartalmazza.
* DimShipMethod: ShipMetod tábla adatait tartalmazza.
* DimSalesPerson: SalesPerson, SalesTerritory és Empolyee táblák adatait tartalmazza.
* DimCustomer: Customer, Store, CustomerTerritory és Person táblák adatait tartalmazza.
* DimData: generált idő dimenzió tábla.
* FactSalesOrder: SalesOrderHeader és SalesOrderDetail táblák adatait ölelifel.

Page-1

3. ábra Csillagséma

1. Planning calls

## 21/10/08

* Csapat onboardolasa az adatmodellel kapcsolatban
* Adatmodell átadása SQL fejlesztonek
* Kérdések:
  + Hol történjen az aggregáció?
  + Kellenek-e atributom megszoritasok?