**Universiteti i Prishtinës “Hasan Prishtina”**

**Fakulteti Inxhinierisë Elektrike dhe Kompjuterike**



|  |  |  |
| --- | --- | --- |
| Punoi : | 1. Gentrit Ademi | [gentrit.ademi2@student.uni-pr.edu](mailto:gentrit.ademi2@student.uni-pr.edu) |

Prishtinë, 2024

Përmbajtja

[**Abstrakti** 3](#_Toc171237067)

[I. Hyrje 4](#_Toc171237068)

[**II.** **Qëllimi i punimit** 6](#_Toc171237069)

[**III.** **Pjesa kryesore** 8](#_Toc171237070)

[**IV.** **Konkluzione** 10](#_Toc171237071)

[Referencat 11](#_Toc171237072)

# **Abstrakti**

Ky punim trajton ndërtimin dhe implementimin e një Data Warehouse (DWH) për një website e-commerce, duke përfshirë automatizimin e procesit të transferimit të të dhënave dhe vizualizimin e tyre përmes mjeteve të Business Intelligence (BI). Tema është e rëndësishme për shkak të rritjes së vazhdueshme të tregtisë elektronike dhe nevojës për analizë të thelluar të të dhënave për të përmirësuar vendimmarrjen. Në literaturën ekzistuese, janë propozuar zgjidhje të ndryshme për ndërtimin e DWH-ve, të cilat përfshijnë metoda të ndryshme të organizimit të skemave (si Star, Snowflake, etj.) dhe teknika të ndryshme ETL (Extract, Transform, Load). Megjithatë, shumica e këtyre zgjidhjeve kërkojnë konfigurim dhe mirëmbajtje komplekse, dhe mund të jenë të shtrenjta dhe të vështira për t'u zbatuar për bizneset e vogla dhe të mesme. Zgjidhja e propozuar në këtë punim përfshin një skemë Star për ndërtimin e DWH-së dhe përdorimin e mjeteve ETL për të automatizuar procesin e transferimit të të dhënave nga baza e të dhënave burimore në DWH çdo 1 minut. Përveç kësaj, përdorimi i BI tools, si PowerBI, mundëson krijimin e grafëve dhe dashboards të ndryshme për analizë të avancuar të të dhënave. Avantazhi kryesor i kësaj zgjidhjeje është thjeshtësia dhe kostoja e ulët e implementimit dhe mirëmbajtjes, duke e bërë atë të përshtatshme për bizneset e vogla dhe të mesme. Gjithashtu, automatizimi i procesit të transferimit të të dhënave dhe vizualizimi i të dhënave me BI tools ofrojnë një platformë të fuqishme për analizë dhe vendimmarrje të bazuar në të dhëna, duke përmirësuar performancën dhe efikasitetin e biznesit.

# I. Hyrje

#### Pse është interesante të diskutohet për problemin në fjalë?

Në epokën e digjitalizimit, e-commerce është bërë një nga sektorët më të rëndësishëm të ekonomisë globale. Rritja e vazhdueshme e tregtisë elektronike ka sjellë me vete nevojën për analiza më të thelluara të të dhënave për të përmirësuar vendimmarrjen dhe për të rritur konkurrencën. Data Warehouse dhe Business Intelligence janë dy teknologji kyçe që mund të ndihmojnë bizneset e e-commerce të kuptojnë më mirë klientët e tyre, të optimizojnë proceset e tyre operacionale dhe të rrisin të ardhurat.

#### Përshkruaje motivimin tuaj: A është e ndërlidhur puna e juaj me ndonjë projekt të caktuar?

Motivimi për këtë punim rrjedh nga nevoja për të përmirësuar analizën dhe menaxhimin e të dhënave në një projekt konkret: një website për shitje të produkteve (e-commerce). Ky projekt kërkon një sistem të integruar që të mbledhë, ruajë dhe analizojë të dhënat e shitjeve dhe klientëve për të marrë vendime më të informuara dhe për të përmirësuar performancën e përgjithshme të biznesit.

#### Pse ja vlen të punohet në ketë fushë e në veçanti në temën tuaj?

Punimi në këtë fushë dhe temë është thelbësor për bizneset e e-commerce, pasi u ofron atyre mjetet dhe teknikat e nevojshme për të përmirësuar efikasitetin dhe efektivitetin e tyre. Përmirësimi i mënyrës se si të dhënat mblidhen, përpunohen dhe analizohen mund të ketë një ndikim të madh në rritjen e të ardhurave dhe përmirësimin e përvojës së klientit.

#### Cila është sfida, Çka është e reja në këtë temë?

Sfida kryesore në këtë temë është integrimi i të dhënave nga burime të ndryshme, krijimi i një strukture të organizuar për ruajtjen e të dhënave dhe automatizimi i procesit të transferimit të të dhënave. Ajo që është e re në këtë temë është përdorimi i metodave të përparuara të ETL dhe BI tools për të krijuar një sistem të plotë dhe të automatizuar për analizën e të dhënave.

#### Cili është kontributi i juaj në këtë fushë?

Kontributi ynë konsiston në krijimin dhe implementimin e një DWH për një website e-commerce, duke përfshirë automatizimin e procesit të transferimit të të dhënave dhe vizualizimin e tyre përmes mjeteve të BI. Kjo ofron një qasje të re dhe më efikase për menaxhimin dhe analizën e të dhënave në sektorin e e-commerce.

#### Cili është problemi i juaj që duhet zgjedhur?

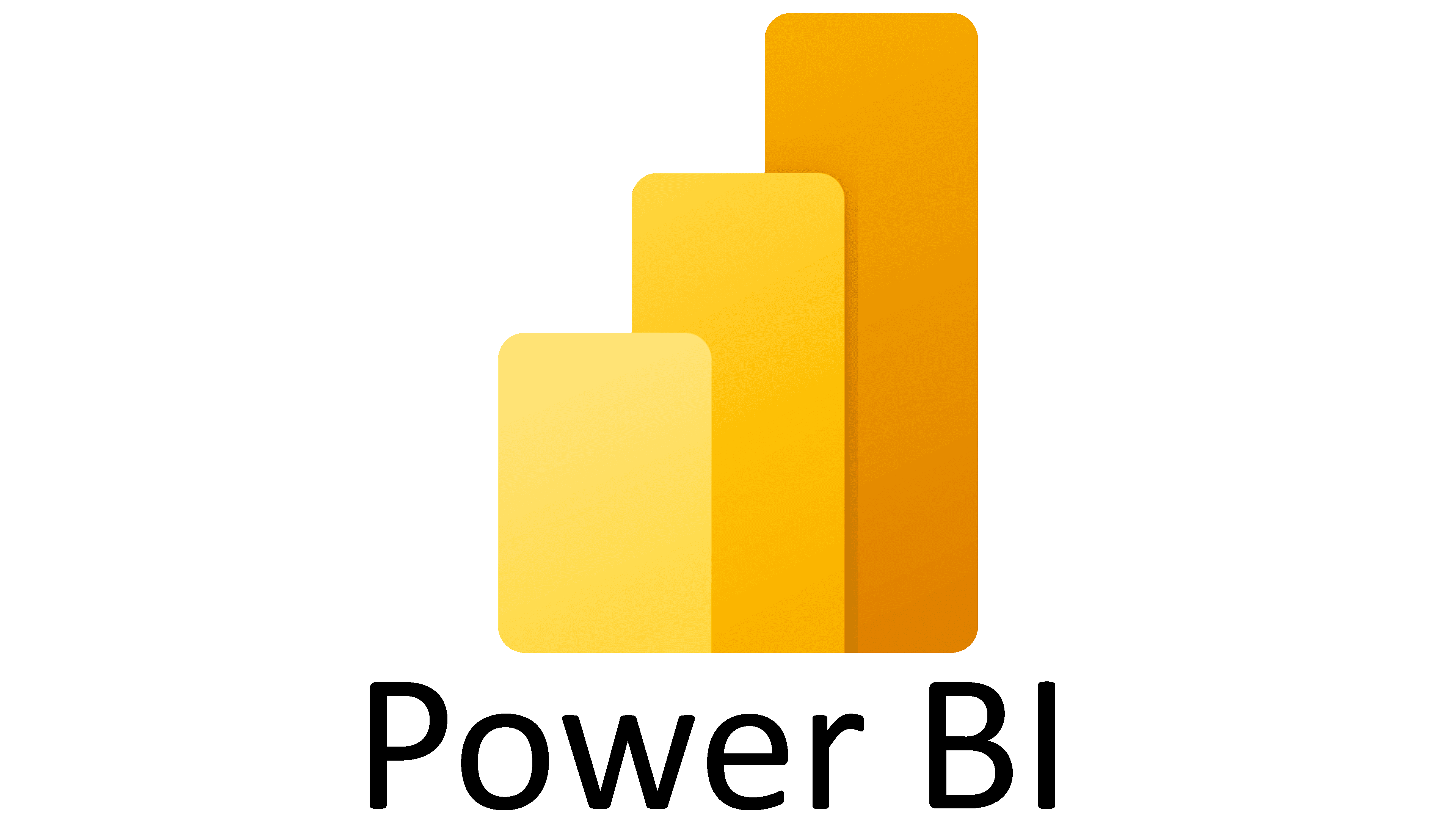
Problemi kryesor është menaxhimi dhe analizimi i një sasie të madhe të dhënash nga burime të ndryshme për të marrë vendime të informuara dhe për të përmirësuar performancën e biznesit.

#### Si duhet zgjedhur problemin?

Problemi duhet të zgjidhet duke krijuar një Data Warehouse të strukturuar dhe të automatizuar që mbledh dhe ruan të dhënat nga burime të ndryshme. Kjo përfshin përdorimin e metodave ETL për të transferuar të dhënat në mënyrë efikase dhe BI tools për të analizuar dhe vizualizuar të dhënat.

#### Figura 1: Arkitektura e përgjithshme e Data Warehouse dhe Data Flow

Në Figurën 1 është paraqitur arkitektura e përgjithshme e Data Warehouse dhe Data Flow, duke treguar procesin e mbledhjes dhe transferimit të të dhënave nga burimet e ndryshme në DWH përmes ETL dhe përdorimin e BI tools për vizualizimin e të dhënave.



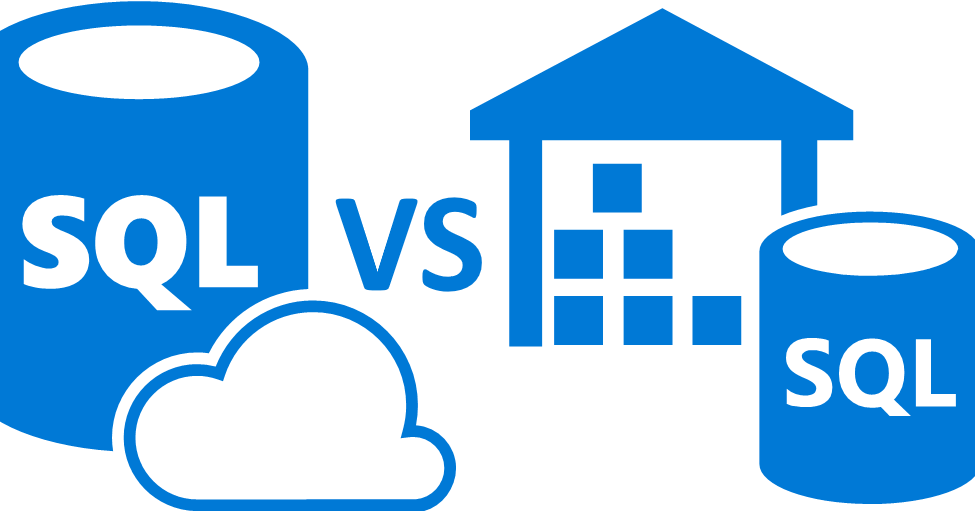


Figura 1: (logo, 2024)

# **Qëllimi i punimit**

#### Cilat janë arsyet që keni përzgjedhur tools-in që përdorni?

Në këtë projekt, kemi përzgjedhur disa mjete kryesore: një sistem DBMS (Database Management System) për ndërtimin e Data Warehouse (DWH), një mjet ETL (Extract, Transform, Load) për automatizimin e transferimit të të dhënave dhe një mjet BI (Business Intelligence) për vizualizimin e të dhënave. DBMS-ja e zgjedhur ofron mbështetje të gjerë për menaxhimin e bazave të të dhënave dhe është e përshtatshme për ndërtimin e DWH-së. Mjeti ETL siguron një proces të automatizuar dhe efikas për transferimin dhe transformimin e të dhënave, ndërsa mjeti BI, si PowerBI, ofron aftësi të fuqishme për analizën dhe vizualizimin e të dhënave.

#### Çka ka të veçantë përdorimi i tools-it që keni marrë?

Përdorimi i këtyre mjeteve në kombinim ofron një qasje të integruar dhe të fuqishme për menaxhimin dhe analizën e të dhënave. DBMS-ja e zgjedhur lejon një strukturë të organizuar dhe të shkallëzuar për ruajtjen e të dhënave. Mjeti ETL mundëson automatizimin e procesit të mbledhjes dhe transformimit të të dhënave, duke reduktuar kohën dhe përpjekjet e nevojshme për mirëmbajtjen e DWH-së. PowerBI ofron mundësi të avancuara për krijimin e grafëve dhe dashboards, duke e bërë analizën e të dhënave të qasshme dhe të kuptueshme për përdoruesit fundorë.

#### Cilat janë përparësitë dhe mangësitë nga veglat tjera në qëllimin që përdorni tools-in tuaj?

**Përparësitë:**

* **DBMS:** Lehtësia në menaxhimin e bazave të të dhënave të mëdha dhe të ndërlikuara, mbështetje për operacione të avancuara SQL.
* **ETL Tool:** Automatizimi i proceseve të transferimit dhe transformimit të të dhënave, integrim me burime të ndryshme të të dhënave.
* **PowerBI:** Aftësi të avancuara për vizualizimin dhe analizën e të dhënave, lehtësia në përdorim dhe ndërfaqe miqësore për përdoruesit.

**Mangësitë:**

* **DBMS:** Mund të jetë kompleks dhe kërkon aftësi të veçanta për administrim dhe optimizim.
* **ETL Tool:** Mund të kërkojë kohë për konfigurim fillestar dhe mirëmbajtje të vazhdueshme.
* **PowerBI:** Kërkon një licencë të paguar për përdorimin e plotë të funksionaliteteve të tij.

#### Cilat janë zgjidhjet tuaja më të mira që i ofroni me përdorimin e tools-it të zgjedhur në projektin tuaj?

Zgjidhja jonë më e mirë përfshin krijimin e një Data Warehouse të strukturuar dhe të automatizuar që mbledh dhe ruan të dhënat nga burime të ndryshme të një website e-commerce. Kjo do të mundësojë:

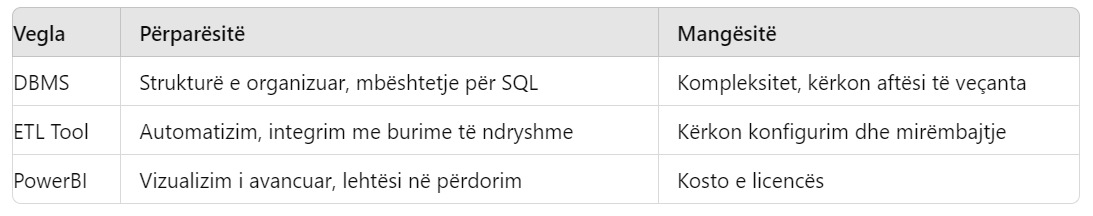
* **Automatizimin e DataFlow:** Procesi i mbledhjes dhe transferimit të të dhënave nga burimet e ndryshme do të jetë plotësisht i automatizuar, duke reduktuar gabimet njerëzore dhe duke kursyer kohë.
* **Analizën e Avancuar të Të Dhënave:** Përdorimi i PowerBI për të krijuar grafë dhe dashboards do të lejojë një analizë të thelluar dhe të kuptueshme të të dhënave, duke ndihmuar në vendimmarrje të informuar.
* **Efikasitet dhe Shkallëzueshmëri:** Zgjidhja jonë është e shkallëzuar dhe mund të zgjerohet lehtësisht për të përballuar rritjen e të dhënave dhe nevojat e biznesit.

#### Metodat e Mbledhjes së Informatave dhe Hulumtimi

Për mbledhjen e informatave të duhura, kemi përdorur metoda të ndryshme:

* **Analiza e Literaturës:** Hulumtimi i artikujve shkencorë dhe literaturës ekzistuese për të kuptuar praktikat më të mira dhe teknikat e avancuara në ndërtimin dhe menaxhimin e DWH-së dhe përdorimin e BI tools.
* **Intervista dhe Konsultime:** Biseda me ekspertë të fushës për të mbledhur njohuri praktike dhe rekomandime.
* **Eksperimentim dhe Testim:** Zbatimi i prototipeve dhe testimi i tyre për të vlerësuar efektivitetin dhe efikasitetin e zgjidhjeve të propozuara.

Tabela 1: Krahasimi i Veglave për Data Warehouse dhe BI Tools



Në Tabelën 1 është paraqitur krahasimi i veglave të përdorura për ndërtimin e DWH-së dhe përdorimin e BI tools, duke theksuar përparësitë dhe mangësitë e secilës vegël.

1. **Customers** (CustomerID, Name, Email, Phone, Address, City, Country)
2. **Products** (ProductID, ProductName, Category, Price, StockQuantity)
3. **Sales** (SaleID, SaleDate, CustomerID, ProductID, Quantity, TotalPrice)
4. **Suppliers** (SupplierID, SupplierName, ContactName, ContactEmail, Phone, Address, City, Country)
5. **Inventory** (InventoryID, ProductID, SupplierID, Quantity, ReceivedDate)
6. **Orders** (OrderID, CustomerID, OrderDate, Status)
7. **OrderDetails** (OrderDetailID, OrderID, ProductID, Quantity, UnitPrice)
8. **Employees** (EmployeeID, FirstName, LastName, Email, Phone, HireDate, DepartmentID)
9. **Departments** (DepartmentID, DepartmentName)
10. **Shippers** (ShipperID, ShipperName, Phone)
11. **Shipments** (ShipmentID, OrderID, ShipperID, ShipDate)
12. **Payments** (PaymentID, OrderID, PaymentDate, Amount, PaymentMethod)
13. **Reviews** (ReviewID, ProductID, CustomerID, ReviewDate, Rating, Comment)
14. **Categories** (CategoryID, CategoryName, Description)
15. **ProductCategories** (ProductCategoryID, ProductID, CategoryID)
16. **Promotions** (PromotionID, PromotionName, StartDate, EndDate, DiscountPercentage)
17. **ProductPromotions** (ProductPromotionID, ProductID, PromotionID)
18. **Locations** (LocationID, LocationName, Address, City, Country)
19. **InventoryLocations** (InventoryLocationID, InventoryID, LocationID, Quantity)
20. **CustomerSupport** (SupportID, CustomerID, SupportDate, IssueDescription, ResolutionDescription)

Kjo listë përmban emrat e tabelave të ndryshme dhe kolonat për secilën prej tyre, të shkruar në një rresht për lehtësi referimi.

CREATE EVENT IF NOT EXISTS update\_dim\_customers

ON SCHEDULE EVERY 1 MINUTE

STARTS CURRENT\_TIMESTAMP

ON COMPLETION PRESERVE

DO

INSERT INTO DWH\_big\_data.DimCustomers (CustomerID, Name, Email, Phone, Address, City, Country)

SELECT CustomerID, Name, Email, Phone, Address, City, Country

FROM big\_data.Customers

ON DUPLICATE KEY UPDATE

Name = VALUES(Name),

Email = VALUES(Email),

Phone = VALUES(Phone),

Address = VALUES(Address),

City = VALUES(City),

Country = VALUES(Country);

-- Eventi që pas 1 minuti bënë kalimin e të dhënave ngadatabaza big\_data në dwh\_big\_data.

# **Pjesa kryesore**

#### Arkitektura e Data Warehouse

Krijimi i një Data Warehouse (DWH) të strukturuar është thelbësor për menaxhimin dhe analizën efikase të të dhënave. Për këtë projekt, kemi përzgjedhur skemën Star, e cila është e njohur për thjeshtësinë dhe performancën e saj të lartë në query të mëdha dhe komplekse. Kjo skemë përbëhet nga një tabelë faktike qendrore e cila lidhet me disa tabela dimensionale.

##### Star Schema

* **Tabelat Faktike:** Facts.Sales
  + Kjo tabelë përmban të gjitha të dhënat kryesore të transaksioneve të shitjeve, duke përfshirë ID-në e produktit, ID-në e klientit, sasinë e shitur, çmimin total, dhe datën e shitjes.
* **Tabelat Dimensionale:** Dimensions.Products, Dimensions.Customers
  + Dimensions.Products përmban informacione të detajuara për produktet, si emri i produktit, kategoria, dhe furnizuesi.
  + Dimensions.Customers përmban informacione për klientët, si emri, adresa, dhe detajet e kontaktit.

#### Metodologjia ETL (Extract, Transform, Load)

Për të automatizuar procesin e transferimit të të dhënave nga burimi në DWH, kemi përdorur një mjet ETL. Procesi ETL përfshin tre hapa kryesorë:

1. **Extract:** Mbledhja e të dhënave nga burimet e ndryshme, përfshirë bazën e të dhënave të e-commerce (big\_data).
2. **Transform:** Transformimi i të dhënave në formatin e duhur për ruajtje në DWH. Kjo përfshin pastrimin e të dhënave, agregimin, dhe normalizimin.
3. **Load:** Ngarkimi i të dhënave të transformuara në tabelat e DWH-së (dwh\_big\_data).

#### Automatizimi i DataFlow

Automatizimi i procesit të ETL është një komponent kyç për të siguruar që DWH të jetë gjithmonë i përditësuar. Kemi konfiguruar procesin ETL të ekzekutohet çdo 1 minuta për të mbledhur dhe përpunuar të dhënat e reja. Ky automatizim është realizuar përmes një sistemi të planifikimit të detyrave (scheduler) që ekzekuton skriptet ETL në intervale të rregullta.

#### Përdorimi i Business Intelligence (BI) Tools

Për vizualizimin dhe analizën e të dhënave, kemi përdorur PowerBI, një nga mjetet më të fuqishme dhe të përdorura për BI. PowerBI ofron mundësi të shumta për krijimin e grafëve dhe dashboards që ndihmojnë në analizën e të dhënave në mënyrë të kuptueshme dhe të lehtë për përdoruesit.

##### Vizualizimet Kryesore

* **Chart Column Vertical:** Paraqet shitjet mujore të produkteve për të analizuar trendet e shitjeve.
* **Chart Column Horizontal:** Tregon shitjet sipas kategorive të produkteve, duke ndihmuar në identifikimin e kategorive më të suksesshme.
* **Pie/Donut Chart:** Përqindja e shitjeve nga secili klient, duke ndihmuar në identifikimin e klientëve më të rëndësishëm.
* **Linear Graph:** Trendet e shitjeve gjatë një periudhe të caktuar, duke ofruar një pamje të përgjithshme të performancës së biznesit.

#### Mbledhja e Informatave dhe Hulumtimi

Për të mbledhur informatat e duhura dhe për të hartuar zgjidhjen, kemi ndjekur një qasje të kombinuar:

1. **Analiza e Literaturës:** Hulumtuam artikuj shkencorë dhe libra për të kuptuar praktikat më të mira dhe teknikat e avancuara në ndërtimin e DWH-së dhe përdorimin e BI tools.
2. **Intervista dhe Konsultime:** Realizuam intervista me ekspertë të fushës për të mbledhur njohuri praktike dhe rekomandime të vlefshme.
3. **Eksperimentim dhe Testim:** Zbatuam dhe testuam prototipet për të vlerësuar efektivitetin dhe efikasitetin e zgjidhjeve të propozuara.

**Shënim:** Pjesa e paraqitjeve vizule është e lidhur me **ANEX\_1**, që gjendet si pjesë e dokumentacionit të këtij projekti.

# **Konkluzione**

#### Teknikat dhe Puna e Realizuar

Në këtë projekt, kemi përdorur teknikat dhe mjetet e mëposhtme:

* **Skema Star:** Struktura efikase për organizimin e tabelave në DWH.
* **ETL Tool:** Automatizimi i procesit të mbledhjes dhe transformimit të të dhënave çdo 1 minut.
* **PowerBI:** Krijimi i grafëve dhe dashboards për vizualizimin e të dhënave.

#### Rezultatet e Arritura

* **Struktura e DWH-së:** Një Data Warehouse e optimizuar për ruajtjen e të dhënave të e-commerce.
* **Automatizimi i DataFlow:** Proces ETL i automatizuar që përditëson DWH çdo 1 minut.
* **Vizualizimet e Të Dhënave:** Grafë dhe dashboards në PowerBI për analizën e të dhënave të shitjeve.

#### Arritja e Qëllimit

Kemi krijuar një sistem të integruar dhe të automatizuar për menaxhimin dhe analizën e të dhënave të një website e-commerce, duke përfshirë një DWH të strukturuar, proces ETL të automatizuar, dhe vizualizime të avancuara të të dhënave.

Në përmbledhje, projekti ka ofruar një zgjidhje të fuqishme për menaxhimin dhe analizën e të dhënave të një website e-commerce. Përmirësimet e ardhshme do të ndihmojnë në zgjerimin dhe optimizimin e sistemit, duke ofruar më shumë vlerë për përdoruesit dhe vendimmarrësit.

# Referencat

(n.d.)., L. i. (2024, 06 29). *https://www.facebook.com/flx/warn/*. Gjetur në /?u=https%3A%2F%2Fwww.microsoft.com%2Fen-us%2Fpower-platform%2Fproducts%2Fpower-bi&h=AT3yKjwfUtFHJbKRQXu1t8R5TySUVNrPUt0O4J9sLT-33BexR2xdGJ-zJTc3-zq5q0TTttAXArYdQ31n2PzrxCDI3GllAhGs-a6OidMKU5BQmBtmJic0NKyQSZ\_\_\_JF4DSONC-mzeJRhF40: Log in or sign up to view. (n.d.). https://www.facebook.com/flx/warn/?u=https%3A%2F%2Fwww.microsoft.com%2Fen-us%2Fpower-platform%2Fproducts%2Fpower-bi&h=AT3yKjwfUtFHJbKRQXu1t8R5TySUVNrPUt0O4J9sLT-33BexR2xdGJ-zJTc3-zq5q0TTttAXArYdQ31n2PzrxCDI3GllAhGs-a6Oid

logo, D. s. (2024, 06 29). *- Google Search. (n.d.). https://www.google.com/*. Gjetur në search?q=DWH+sql+logo&sca\_esv=0f377be67df6cd70&udm=2&biw=1920&bih=904&ei=VUGKZtrEDoiMxc8PrJ272AI&ved=0ahUKEwiaw5vbsZSHAxUIRvEDHazODisQ4dUDCBA&uact=5&oq=DWH+sql+logo&gs\_lp=Egxnd3Mtd2l6LXNlcnAiDERXSCBzcWwgbG9nb0jLGFDsBlizFnACeACQAQCYAWKgAcwDqgEBNbgBA8gBAPgB: search?q=DWH+sql+logo&sca\_esv=0f377be67df6cd70&udm=2&biw=1920&bih=904&ei=VUGKZtrEDoiMxc8PrJ272AI&ved=0ahUKEwiaw5vbsZSHAxUIRvEDHazODisQ4dUDCBA&uact=5&oq=DWH+sql+logo&gs\_lp=Egxnd3Mtd2l6LXNlcnAiDERXSCBzcWwgbG9nb0jLGFDsBlizFnACeACQAQCYAWKgAcwDqgEBNbgBA8gBAPgB

OpenData. (2024, 06 29). *faqja kryesore. (n.d.).* Gjetur në OpenData - Faqja Kryesore.: https://opendata.gov.al/