**TUGAS AKHIR SISTEM KECERDASAN BISNIS**

**PROYEK MEMBANGUN SISTEM KECERDASAN BISNIS SEDERHANA**

**SEBUAH STUDI PADA DATASET ADVENTURE WORKS (SALES)**

****

**KELOMPOK: 04**

**DISUSUN OLEH:**

**6182001033 / Alexander Bleuvito Fevrier**

**6182001039 / Wilson**

**6182001051 / Emeralda Checia Pitaloka**

**KELAS: A**

**Dosen Kelas: Lydia Mutiara Dewi, MA**

**UNIVERSITAS KATOLIK PARAHYANGAN**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI DAN SAINS**

**BANDUNG**

**2024**

1. **PROFIL SINGKAT PERUSAHAAN**

Adventure Works menciptakan suasana bisnis imajiner dengan fokus pada produk dan layanan yang mencakup berbagai aspek seperti petualangan, barang konsumen, dan kebutuhan pelanggan. Dalam penggunaannya, perusahaan ini memberikan data fiktif yang mencakup informasi tentang pelanggan, produk, dan transaksi untuk tujuan pelatihan dan demonstrasi. Adventure Works berupa *database* yang memberikan dataset untuk memungkinkan penggunanya mengasah dan berlatih mengenai pemahaman konsep. AdventureWorks dirancang untuk memberikan penggunaan praktis dan contoh konkret saat melibatkan diri dalam pengembangan dan manajemen database.

Selain melakukan penjualan sepeda, Adventure Works juga menyediakan beberapa aksesoris, pakaian, dan juga komponen. Beberapa contoh aksesoris yang disediakan adalah botol minuman, rem sepeda, dll. Sedangkan di bagian pakaian, Adventure Works menyediakan topi, sarung tangan, dan juga seragam untuk berolahraga sepeda. Dan untuk di kategori komponen, Adventure Works menyediakan rantai, ban, dan masih banyak lainnya.

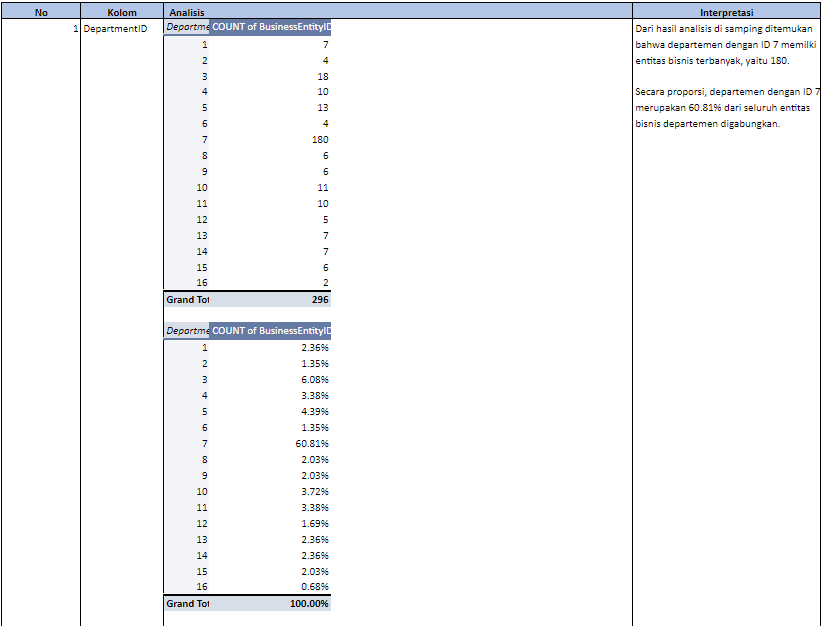
Adventure Works melayani pelanggan secara global, seperti Australia, Kanada, Perancis, Jerman, Inggris, dan juga Amerika Serikat. Meskipun Adventure Works merupakan sebuah perusahaan fiktif, perusahaan ini menyediakan database yang sangat lengkap mencakup penjualan, pembelian, dll.

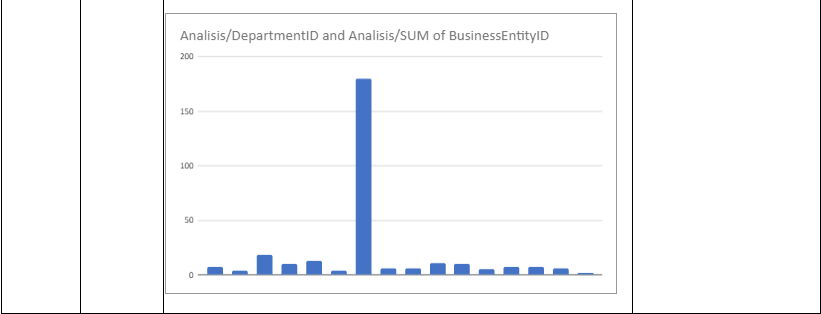
1. **EDA dan IDENTIFIKASI KEBUTUHAN PERUSAHAAN**

Berikut adalah hasil EDA dan kebutuhan perusahaan untuk dataset adventure works:

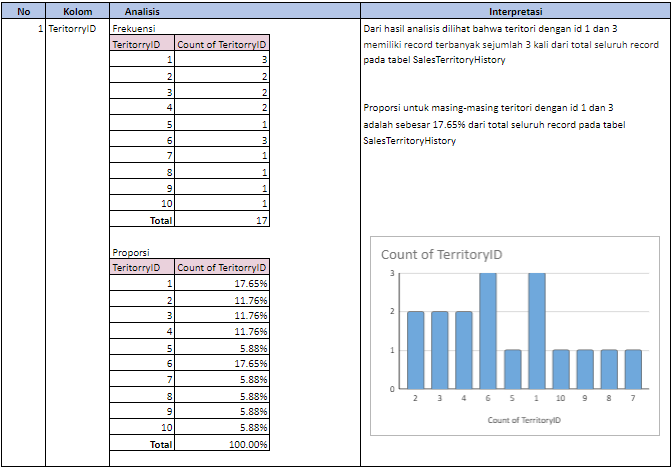
1. Melakukan analisis pada tabel Customer dan SalesOrderHeader untuk mensegmentasi konsumen berdasarkan perilaku dan lokasi pembelian, sehingga perusahaan dapat mengimplementasikan sistem personalisasi untuk meningkatkan penjualan.
2. Mengoptimalkani performa sales pada tabel SalesOrderHeader, SalesTeritorry, dan Product untuk mengoptimalkan performa sales
3. Memanfaatkan data dari tabel SalesOrderHeader, Product, dan Sales Territory untuk mencari produk-produk yang paling laku, serta wilayah yang paling menguntungkan. Kemudian dari situ melakukan penyesuaian harga, penawaran produk, dan berujung pada peningkatan pendapatan.
4. Memantau inventori produk dan memprediksi permintaan customer akan produk dengan menggunakan data yang ada di tabel *SalesOrderDetails*. Ditindaklanjuti lagi dengan mengimplementasikan strategi optimasi inventori untuk mengurangi biaya kirim dan memastikan produk selalu dapat dibeli customer

* **EDA Tabel EmployeeDepartmentHistory**

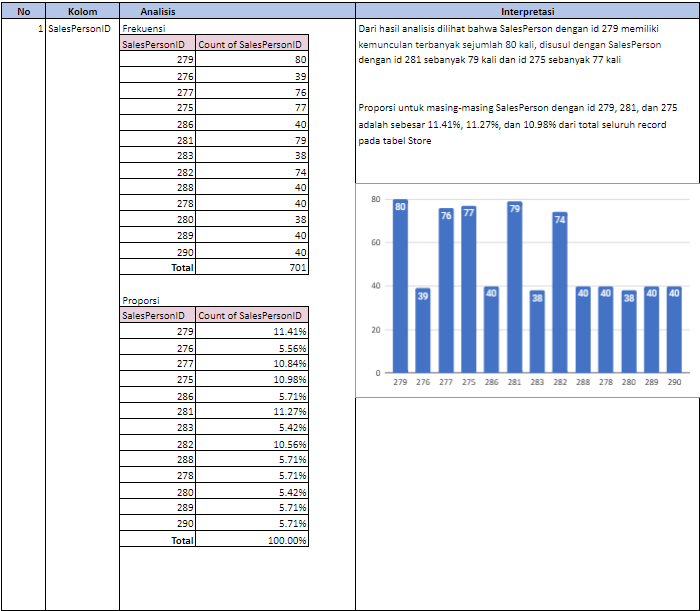
****

****

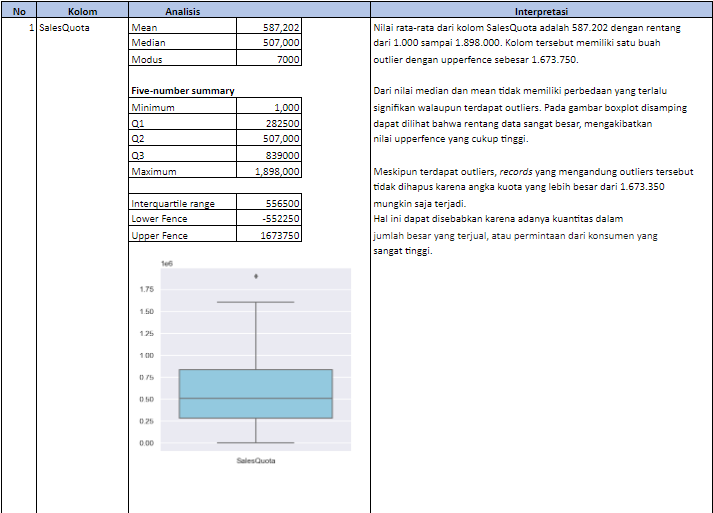
* **EDA Tabel SalesTerritoryHistory**

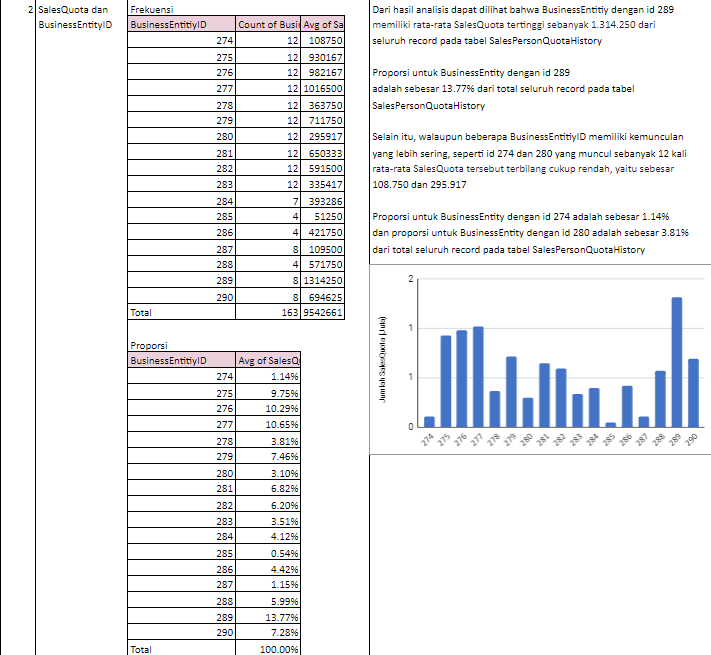
****

* **EDA Tabel Store**

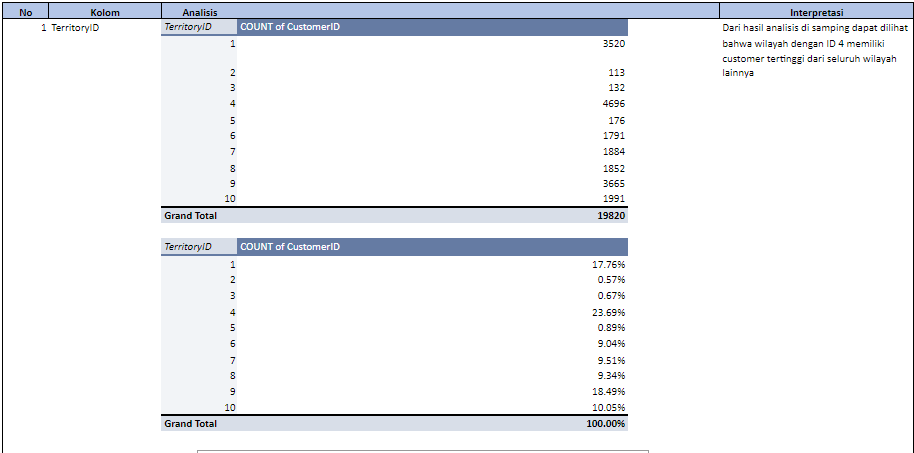
****

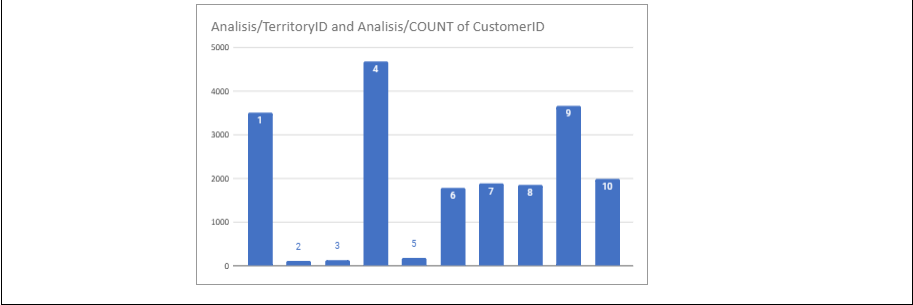
* **EDA Tabel SalesPersonQuotaHistory**

****

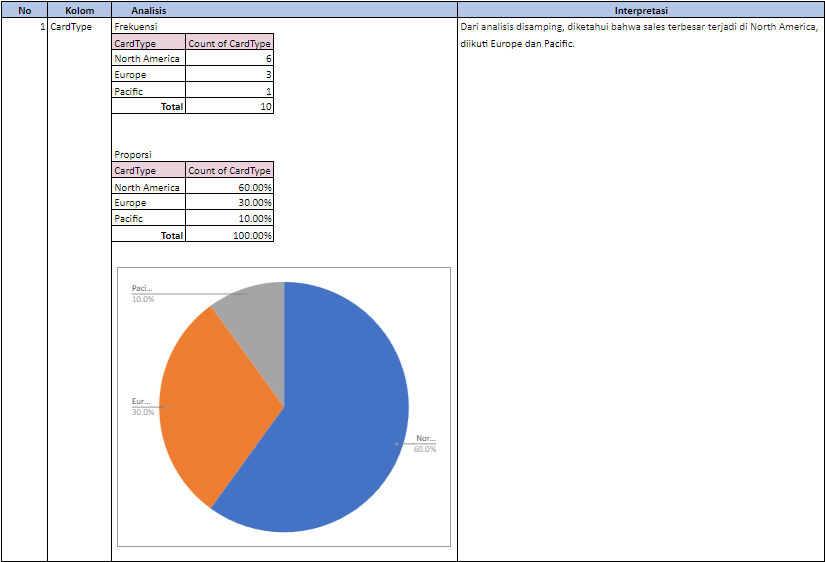
****

* **EDA Tabel Customer**

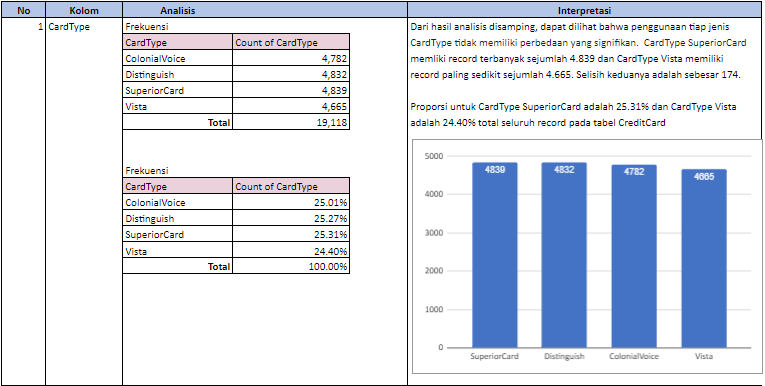
****

****

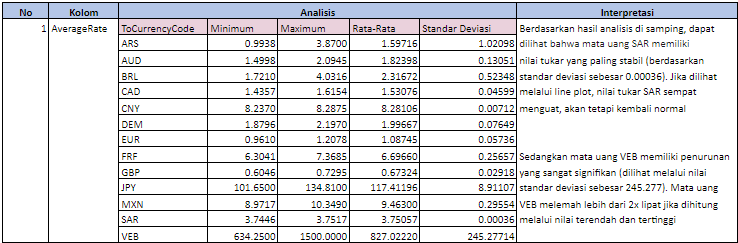
* **EDA Tabel SalesTerritory**

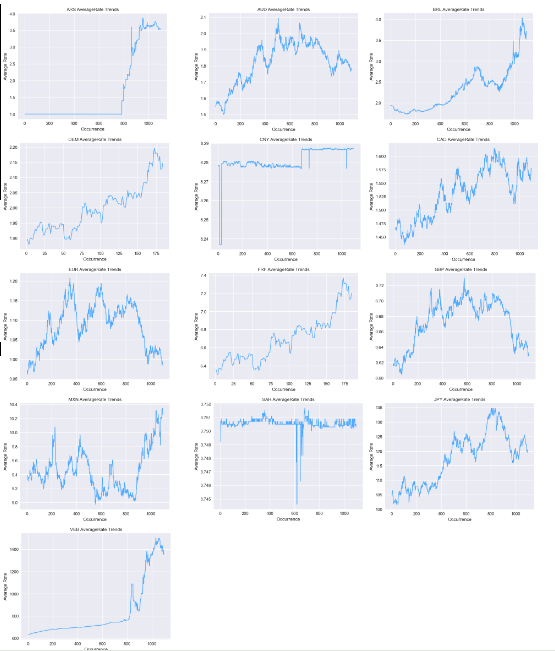
****

* **EDA Tabel CreditCard**

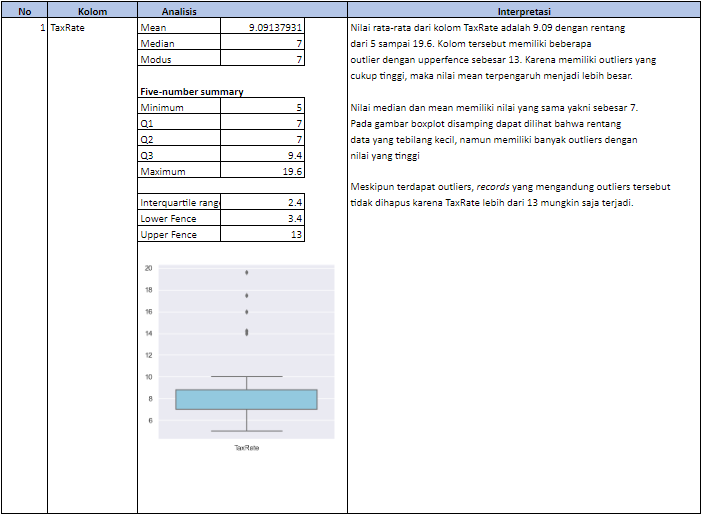
****

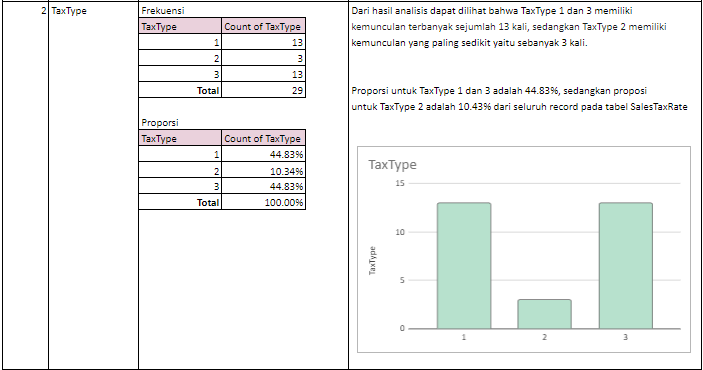
* **EDA Tabel CurrencyRate**

****

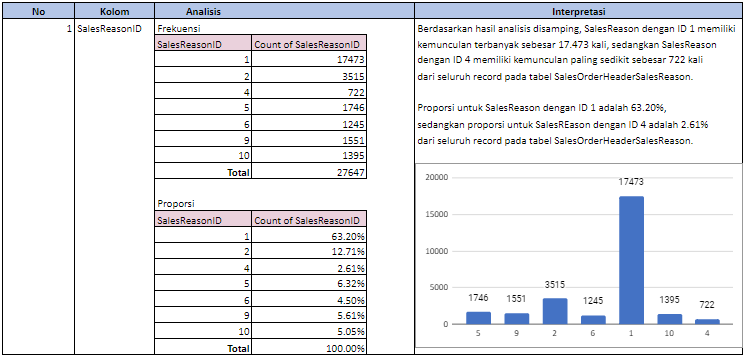
****

* **EDA Tabel SalesTaxRate**

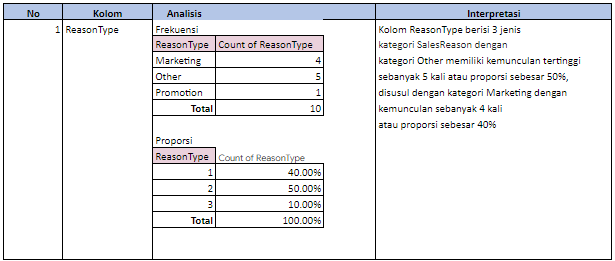
****

****

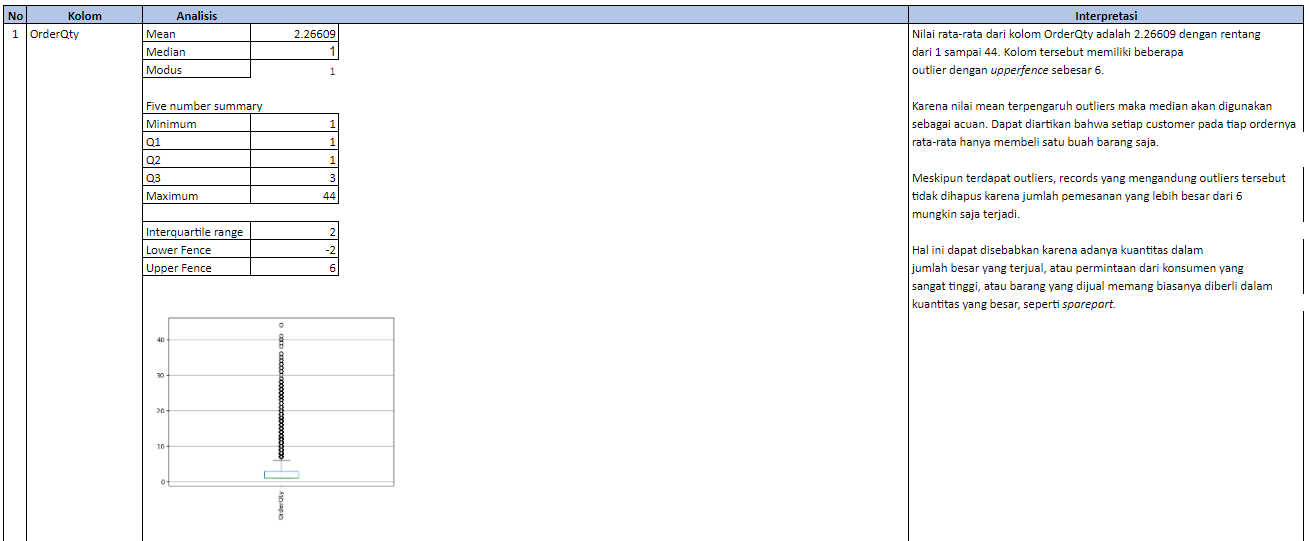
* **EDA Tabel SalesOrderHeaderSalesReason**

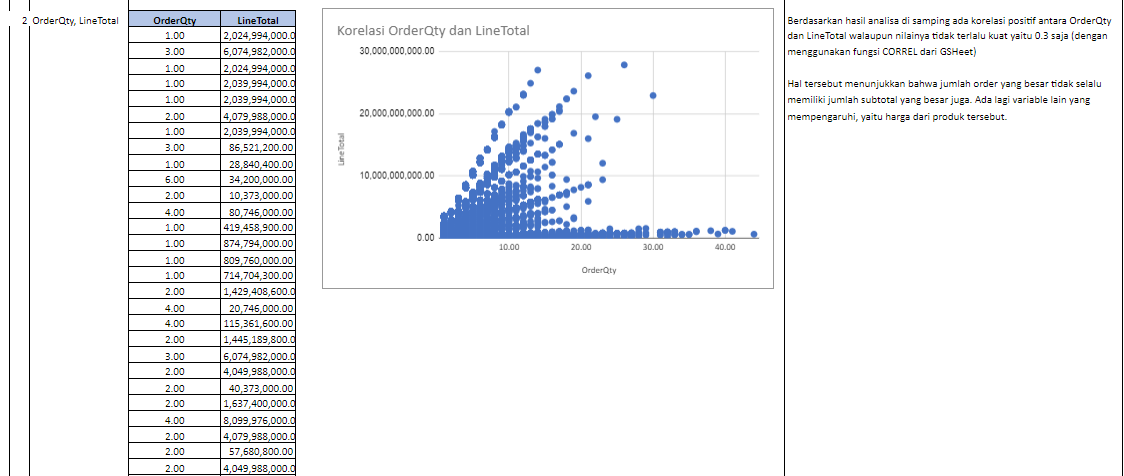
****

* **EDA Tabel SalesReason**

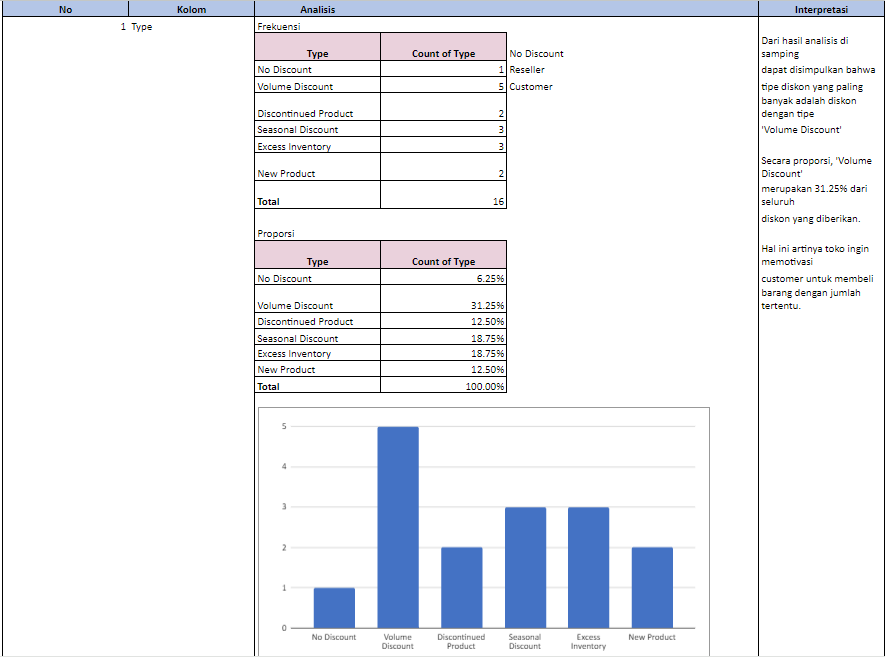
****

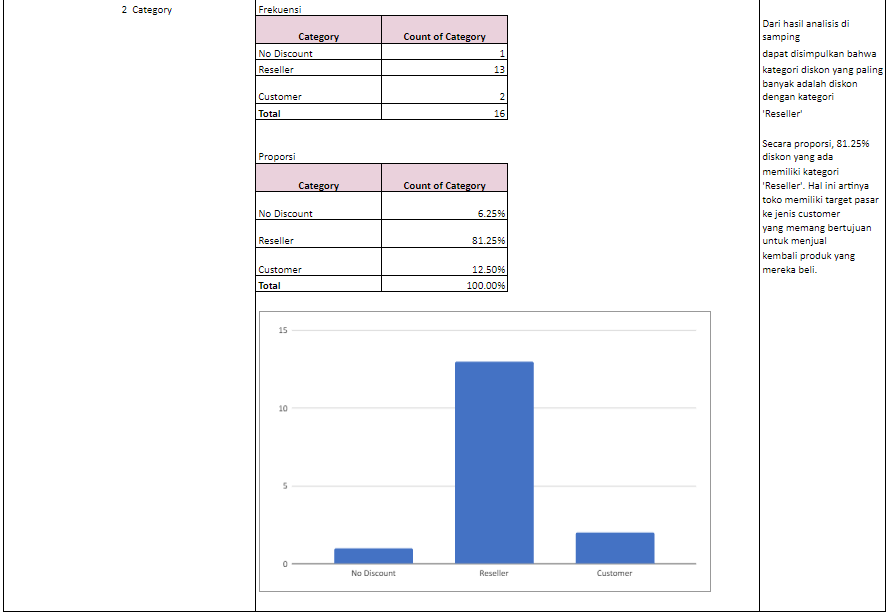
* **EDA Tabel SalesOrderDetail**

****

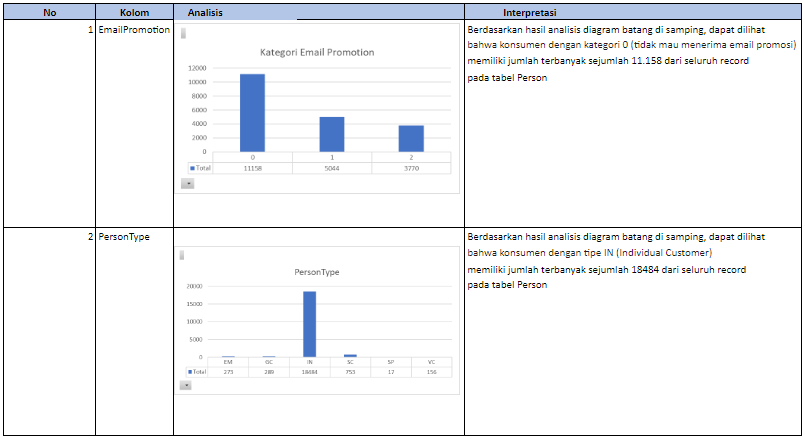
****

* **EDA Tabel SpecialOffer**

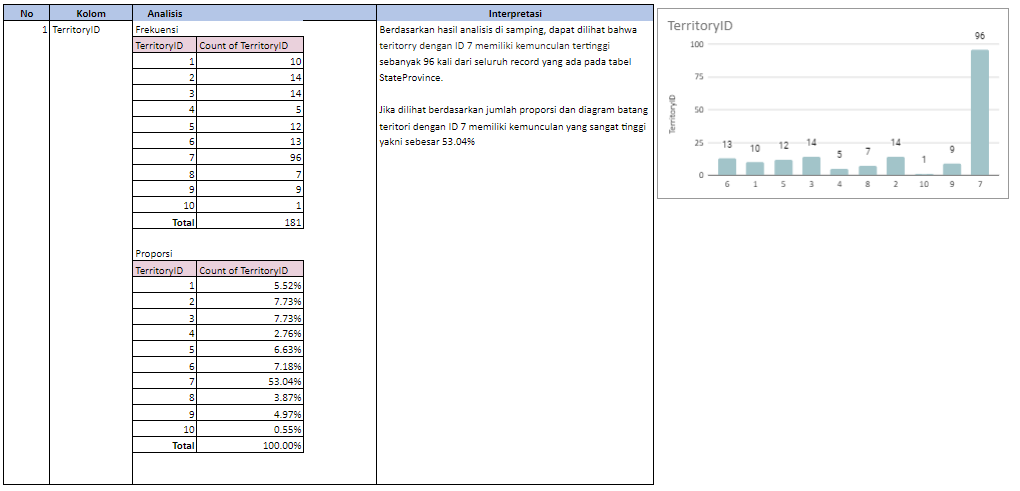


****

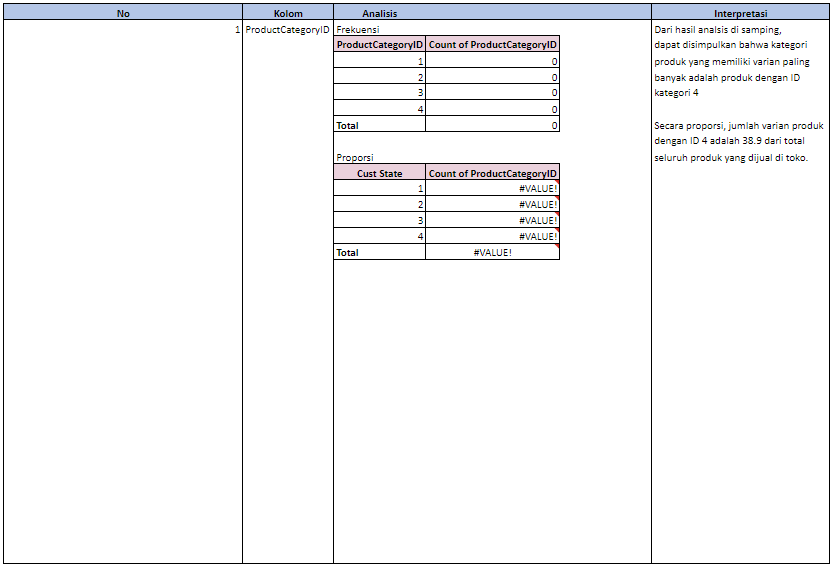
* **EDA Tabel Person**

****

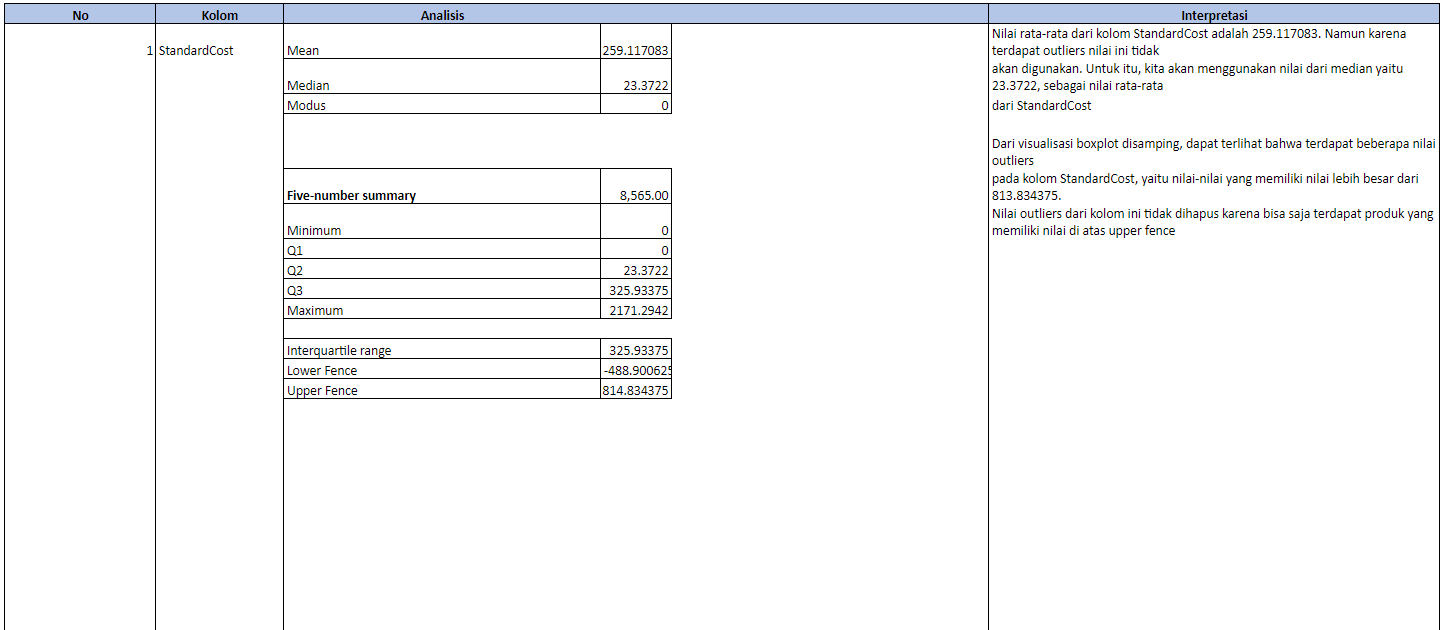
* **EDA Tabel StateProvince**

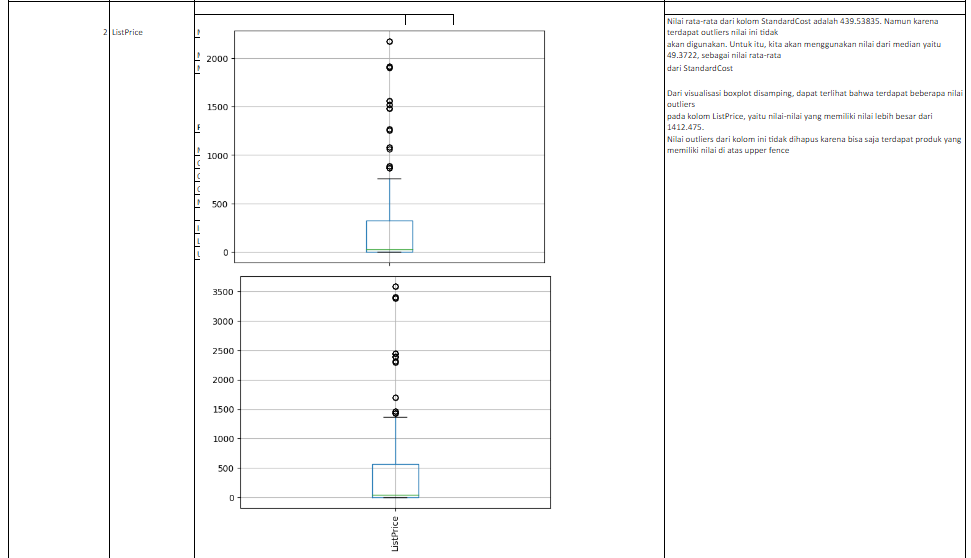
****

* **EDA Tabel ProductListPriceHistory**

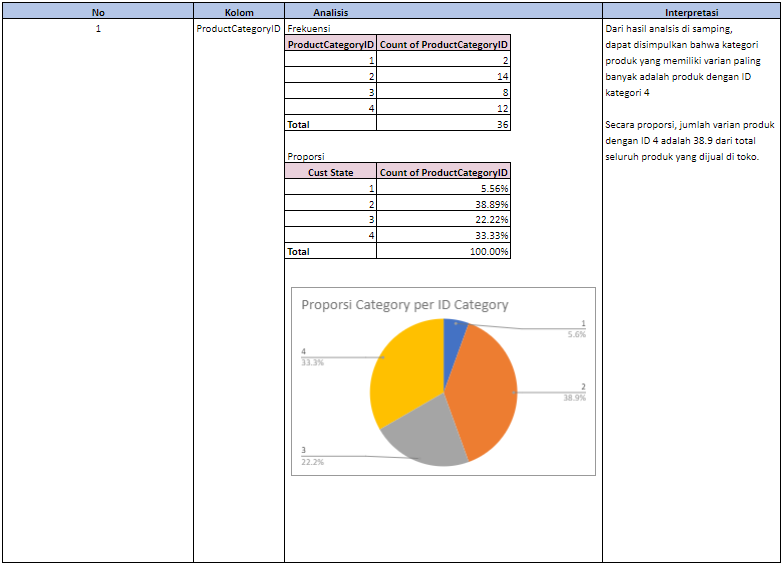
****

* **EDA Tabel Product**

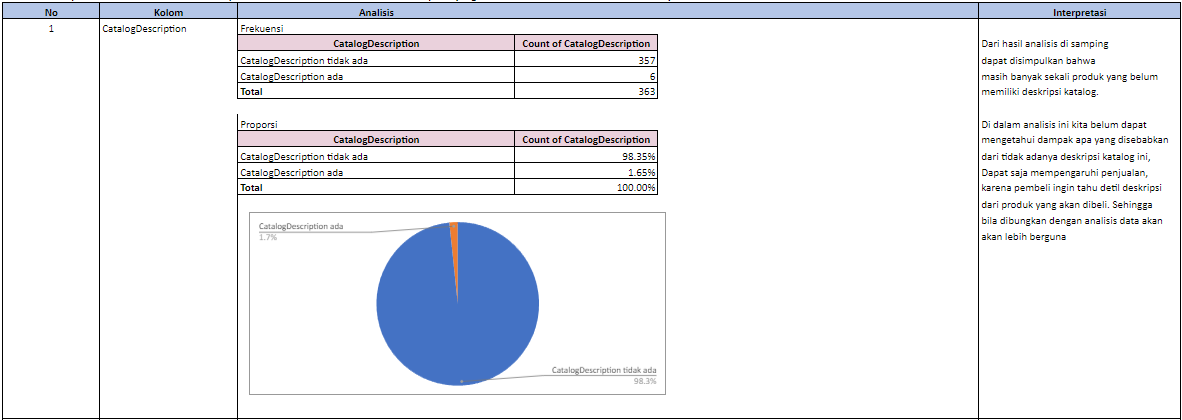
****

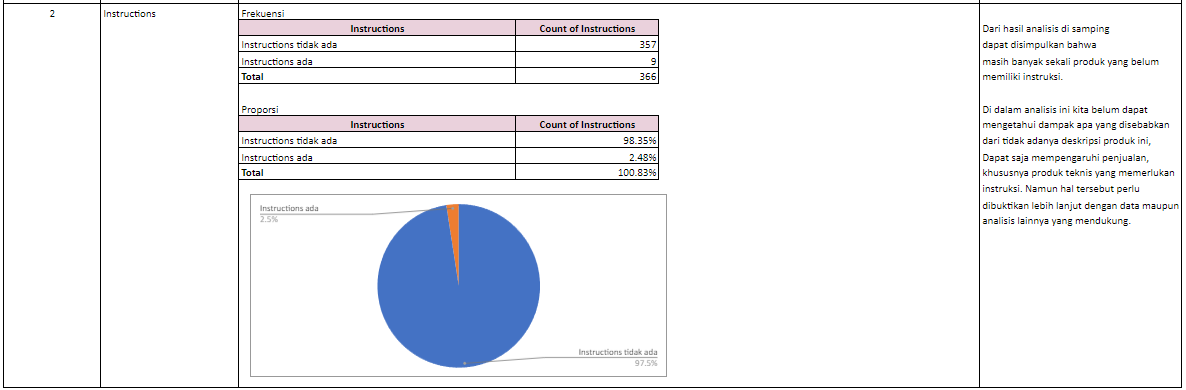
****

* **EDA Tabel ProductSubCategory**



* **EDA Tabel ProductModel**

****

****

1. **DATA PREPARATION [[1]](#footnote-1)**

Pada tahap penyiapan data (*data preparation*) dilakukan serangkaian proses yang terdiri dari tahap transformasi data dari bentuk mentah menjadi data yang dapat langsung digunakan dalam model, tahap deteksi dan perbaikan berbagai jenis kesalahan dalam data agar sesuai dengan kebutuhan analisis, serta tahap penggabungan data dari berbagai sumber data yang berbeda. Data yang telah disiapkan dapat disimpan ke dalam tempat penyimpanan (*data warehouse*) untuk kemudian dianalisis sehingga didapatkan *insights* yang bermanfaat bagi organisasi atau perusahaan. Data yang diterima dari tahap pengumpulan data kemungkinan besar adalah "*a diamond in the rough*". Tugas pada tahap penyiapan data adalah membersihkan dan mempersiapkan data untuk digunakan dalam fase pemodelan dan pelaporan. Tahapan ini sangatlah penting karena model akan berkinerja lebih baik dan tidak perlu membuang waktu untuk memperbaiki hasil yang buruk atau yang dikenal dengan istilah “*Garbage in, Garbage out*”.

**III.1 *Data Cleansing***

Pembersihan data adalah subproses dari proses *data science* yang berfokus pada penghapusan kesalahan dalam data sehingga data menjadi representasi yang benar dan konsisten. Terdapat setidaknya ada dua jenis kesalahan ada:

* *Interpretation error*, ketika data langsung diambil tanpa di cek terlebih dahulu kebenarannya.

Contoh: Usia seseorang lebih dari 300 tahun.

* *Inconsistencies*, ketidakkonsistenan antara sumber data terhadap nilai standar.

Contoh: "Female” di satu tabel dan "F" di tabel lain, “Pounds” di satu tabel dan “Dollars” di tabel lainnya.

**III.2 *Data Transformation [[2]](#footnote-2)***

Setiap model memerlukan bentuk data tertentu sehingga dibutuhkan transformasi data agar data sesuai untuk kebutuhan analisis dan pembuatan model. Beberapa contoh bentuk transformasi data[[3]](#footnote-3):

* Concatenate (mis: menggabungkan nama depan dan nama belakang seseorang yang berada di 2 kolom terpisah menjadi nama lengkap di yang berada di 1 kolom)
* Customized calculation (mis: menghitung BMI berdasarkan tinggi badan dan berat badan)
* Categorize(mis: mengkategorikan jenis umur seseorang (1 –100 tahun) yang berjenis numerik menjadi kategori umur seperti: batita, balita, remaja, dewasa, lansia)
* Pengkodean: (mis: mengkodekan jenis kelamin, pria == 1 dan Wanita == 0)

**III.3 *Combining Data***

Kumpulan data dapat berasal dari beberapa tempat yang berbeda, sehingga dibutuhkan pengintegrasian berbagai sumber data. Data bervariasi dalam ukuran, jenis, dan struktur, mulai dari *database*, file Excel, hingga dokumen teks. Terdapat dua operasi untuk menggabungkan informasi dari kumpulan data yang berbeda. Operasi pertama adalah *joining*: memperkaya pengamatan dari satu tabel dengan informasi dari tabel lain. Operasi kedua adalah *appending* atau *stacking*: menambahkan pengamatan dari satu tabel ke pengamatan dari tabel lain.

**Daftar Aktivitas Data Preparation**

|  |  |  |
| --- | --- | --- |
| No | Aktivitas | Tabel |
| 1 | Menghapus kolom ModifiedDate | ProductModelProductDescriptionCulture |
|  | Menambah kolom Culture untuk mengetahui arti dari CultureID tabel tersebut. Nilai dari CultureName ini diambil dari tabel Production.Culture, kemudian dijoin dengan tabel Production.ProductModelProductDescriptionCulture berdasarkan CultureID. | ProductModelProductDescriptionCulture |
|  | Menghapus kolom CultureID. Hal ini dilakukan karena sudah ada kolom yang menjelaskan CultureID, sehingga kolom ini tidak lagi diperlukan. | ProductModelProductDescriptionCulture |
| 2 | Mengahpus kolom ModifiedDate | ProductModel |
|  | Menghapus kolom rowguid | ProductModel |
| 3 | Menghapus kolom ModifiedDate | ProductCategory |
|  | Menghapus kolom rowguide | ProductCategory |
| 4 | Menghapus kolom ModifiedDate | ProductSubcategory |
|  | Menghapus kolom rowguide | ProductSubcategory |
| 5 | Menghapus kolom ModifiedDate | EmployeeDepartmentHistory |
|  | Mengganti format tanggal EndDate menjadi MM/DD/YYYY | EmployeeDepartmentHistory |
|  | Mengisi seluruh sel yang kosong dengan "-" | EmployeeDepartmentHistory |
| 6 | Menghapus kolom ModifiedDate | Department |
| 7 | Menghapus kolom rowguide | Employee |
|  | Menghapus kolom ModifiedDate | Employee |
|  | Mengganti format tanggal menjadi MM/DD/YYYY | Employee |
|  | Mengubah format MaritalStatus menjadi Single/Married | Employee |
|  | Mengubah format Gender menjadi Male/Female | Employee |
| 8 | Menghapus kolom ModifiedDate | Customer |
|  | Menghapus kolom rowguide | Customer |
|  | Mengisi seluruh sel yang kosong dengan "-" | Customer |
| 9 | Menghapus kolom ModifiedDate | SalesTerritory |
|  | Menghapus kolom rowguide | SalesTerritory |
| 10 | Menghapus kolom ModifiedDate | PersonCreditCard |
| 11 | Menghapus kolom ModifiedDate | StateProvince |
|  | Menghapus kolom rowguide | StateProvince |
| 12 | Menghapus kolom ModifiedDate | CountryRegion |
| 13 | Menghapus kolom ModifiedDate | SalesTerritoryHistory |
|  | Menghapus kolom rowguide | SalesTerritoryHistory |
|  | Mengubah format tanggal menjadi MM/DD/YYYY | SalesTerritoryHistory |
| 14 | Menghapus kolom ModifiedDate | Store |
|  | Menghapus kolom rowguide | Store |
| 15 | Menghapus kolom ModifiedDate | SalesPersonQuotaHistory |
|  | Menghapus kolom rowguide | SalesPersonQuotaHistory |
|  | Mengubah format tanggal menjadi MM/DD/YYYY | SalesPersonQuotaHistory |
| 16 | Menghapus kolom ModifiedDate | BusinessEntity |
|  | Menghapus kolom rowguide | BusinessEntity |
| 17 | Menghapus kolom ModifiedDate | CreditCard |
| 18 | Menghapus kolom ModifiedDate | CurrencyRate |
|  | Mengubah format tanggal menjadi MM/DD/YYYY | CurrencyRate |
|  | Menambahkan kolom RateDifferences | CurrencyRate |
| 19 | Menghapus kolom ModifiedDate | SalesTaxRate |
|  | Menghapus kolom rowguide | SalesTaxRate |
| 20 | Menghapus kolom ModifiedDate | Currency |
| 21 | Menghapus kolom ModifiedDate | SalesOrderHeaderSalesReason |
| 22 | Menghapus kolom ModifiedDate | SalesReason |
| 23 | Mengganti nilai 'W' menjadi 'Womens', 'M' menjadi 'Mens', 'U' menjadi 'Universal', dan 'NULL' menjadi '-' di kolom Style. | Product |
|  | Mengganti nilai 'H' menjadi 'High', 'M' menjadi 'Mendium', 'L' menjadi 'Low', dan 'NULL' menjadi '-' di kolom Class. | Product |
|  | Mengganti nilai 'R' menjadi 'Road', 'M' menjadi 'Mountain', 'T' menjadi 'Touring', 'S' menjadi 'Standard', dan 'NULL' menjadi '-' di kolom ProductLine. | Product |
|  | Menghapus kolom rowguid karena sepertinya tidak akan digunakan dalam analisis. | Product |
|  | Menghapus kolom ModifiedDate karena sepertinya tidak akan digunakan dalam analisis. | Product |
|  | Menghapus kolom DiscontinuedDate karena semua nilainya NULL yang artinya semua produk masih diproduksi dan membuat kolom tidak memiliki nilai. | Product |
|  | Menambah kolom AvailableForSale untuk menunjukkan produk masih dijual atau tidak. Nilai 'Yes' artinya produk masih dijual, nilai 'No' artinya produk sudah tidak dijual. | Product |
|  | Mengganti nilai 'NULL' pada kolom Color. | Product |
|  | Mengganti nilai '0' menjadi 'No' dan '1' menjadi 'Yes' pada kolom 'FinishedGoodFlag' | Product |
|  | Menggani nama Kolom 'FinishedGoodFlag' menjadi 'IsFinishedGood'. | Product |
|  | Menggani nilai '1' menjadi 'Manufactured', dan '0' menjadi 'Purchased' di kolom MakeFlag. | Product |
|  | Mengganti nama kolom 'MakeFlag' menjadi 'ObtainMethod'. | Product |
|  | Mengganti format nilai 'StandardCost' menjadi currency dengan menggunakan $. | Product |
|  | Mengganti format nilai 'ListPrice' menjadi currency dengan menggunakan $. | Product |
|  | Mengganti format tanggal kolom SellStartDate menjadi MM/DD/YYYY, | Product |
|  | Mengganti format tanggal kolom SellEndDate menjadi MM/DD/YYYY, | Product |
|  | Mengganti nilai kosong kolom dengan '-'. | Product |
|  | Mengganti nilai kosong kolom ProductSubcategoryID dengan '0' karena tidak ada record yang bernilai 0. | Product |
|  | Mengganti nilai kosong kolom ProductModelID dengan '0' karena tidak ada record yang bernilai 0. | Product |
|  | Menghapus kolom ProductNumber karena semua nilainya unique dan tidak akan dipakai dalam analisis. | Product |
| 24 | Mengganti format tanggal kolom StartDate menjadi MM/DD/YYYY. | ProductListPriceHistory |
|  | Mengganti format tanggal kolom EndDate menjadi MM/DD/YYYY. | ProductListPriceHistory |
|  | Menghapus kolom ModifiedDate | ProductListPriceHistory |
|  | Mengganti format nilai kolom ListPrice menjadi currency dengan $. | ProductListPriceHistory |
|  | Menambah kolom IsCurrentlyUsed untuk menandakan harga sedang digunakan atau tidak. 'Yes' artinya harga sedang digunakan, 'No' sebaliknya. | ProductListPriceHistory |
| 25 | Menghapus kolom rowguid. | SpecialOfferProduct |
|  | Menghapus kolom ModifiedDate. | SpecialOfferProduct |
| 26 | Menghapus kolom rowguid. | SpecialOffer |
|  | Menghapus kolom ModifiedDate. | SpecialOffer |
|  | Mengganti format tanggal kolom StartDate menjadi MM/DD/YYYY. | SpecialOffer |
|  | Mengganti format tanggal kolom EndDate menjadi MM/DD/YYYY. | SpecialOffer |
| 27 | Menghapus kolom rowguid. | SalesOrderDetail |
|  | Menghapus kolom ModifiedDate | SalesOrderDetail |
|  | Mengubah format nilai kolom UnitPrice menjadi currency dengan $. | SalesOrderDetail |
|  | Mengubah format nilai kolom LineTotal menjadi currency dengan $. | SalesOrderDetail |
|  | Mengganti nilai kosong kolom CarrierTrackingNumber dengan '-'. | SalesOrderDetail |
|  | Menambah kolom HasBeenShipped untuk menunjukkan pesanan sudah dikirim atau belum. 'Yes' berarti pesanan sudah dikirim, dan 'No' berarti pesanan belum dikirim. | SalesOrderDetail |
| 28 | Menhapus kolom DateCreated. | ShoppingCartItem |
|  | Menghapus kolom ModifiedDate. | ShoppingCartItem |
| 29 | Menghapus kolom ModifiedDate | Address |
|  | Menghapus kolom rowguide | Address |
|  | Mengganti kolom kosong dengan "-" | Address |
| 30 | Menghapus kolom ModifiedDate | SalesOrderHeader |
|  | Menghapus kolom rowguide | SalesOrderHeader |
|  | Mengganti kolom kosong dengan "-" | SalesOrderHeader |
|  | Menambah kolom Season untuk mengetahui pemesanan dibuat di musim apa. | SalesOrderHeader |
| 31 | Menghapus kolom ModifiedDate | Person |
|  | Menghapus kolom rowguide | Person |
|  | Mengganti kolom kosong dengan "-" | Person |
|  | Mengganti 'EM' dengan 'Non-Sales Employee', 'SC' dengan 'Sales Contact', 'GC' dengan 'General Contact', 'VC' dengan 'Vendor Contact, dan 'SP' dengan 'Sales Person' pada kolom PersonType. | Person |
|  | Menggani '0' dengan 'Western' dan '1' dengan 'Eastern' pada kolom NameStyle. | Person |
|  | Menambah kolom Fullname untuk menggabungkan nama orang yang terdapat pada record. | Person |

**Tabel ProductModelProductDescriptionCulture**

|  |  |  |
| --- | --- | --- |
| **ProductModelID** | **ProductDescriptionID** | **Culture** |
| 1 | 1199 | English |
| 1 | 1467 | Arabic |
| 1 | 1589 | French |
| 1 | 1712 | Thai |
| 1 | 1838 | Hebrew |
| 127 | 2006 | Arabic |
| 127 | 2007 | French |
| 127 | 2008 | Thai |
| 127 | 2009 | Hebrew |
| 127 | 2010 | Chinese |

**Tabel ProductModel**

|  |  |  |  |
| --- | --- | --- | --- |
| **ProductModelID** | **Name** | **CatalogDescription** | **Instructions** |
| 1 | Classic Vest | - | - |
| 2 | Cycling Cap | - | - |
| 3 | Full-Finger Gloves | - | - |
| 4 | Half-Finger Gloves | - | - |
| 5 | HL Mountain Frame | - | - |
| 124 | ML Mountain Rear Wheel | - | - |
| 125 | HL Mountain Rear Wheel | - | - |
| 126 | LL Road Rear Wheel | - | - |
| 127 | Rear Derailleur | - | - |
| 128 | Rear Brakes | - | - |

**Tabel ProductCategory**

|  |  |
| --- | --- |
| **ProductCategoryID** | **Name** |
| 1 | Bikes |
| 2 | Components |
| 3 | Clothing |
| 4 | Accessories |

**Tabel ProductSubcategory**

|  |  |  |
| --- | --- | --- |
| **ProductSubcategoryID** | **ProductCategoryID** | **Name** |
| 1 | 1 | Mountain Bikes |
| 2 | 1 | Road Bikes |
| 3 | 1 | Touring Bikes |
| 4 | 2 | Handlebars |
| 5 | 2 | Bottom Brackets |
| 33 | 4 | Lights |
| 34 | 4 | Locks |
| 35 | 4 | Panniers |
| 36 | 4 | Pumps |
| 37 | 4 | Tires and Tubes |

**Tabel EmployeeDepartmentHistory**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BusinessEntityID** | **DepartmentID** | **ShiftID** | **14-01-2009** | **EndDate** |
| 1 | 16 | 1 | 14/01/2009 | NULL |
| 2 | 1 | 1 | 31/01/2008 | NULL |
| 3 | 1 | 1 | 11/11/2007 | NULL |
| 4 | 1 | 1 | 5/12/2007 | NULL |
| 4 | 2 | 1 | 31/05/2010 | NULL |
| 286 | 3 | 1 | 30/05/2013 |  |
| 287 | 3 | 1 | 16/04/2012 |  |
| 288 | 3 | 1 | 30/05/2013 |  |
| 289 | 3 | 1 | 30/05/2012 |  |
| 290 | 3 | 1 | 30/05/2012 |  |

**Tabel Department**

|  |  |  |
| --- | --- | --- |
| **DepartmentID** | **Name** | **GroupName** |
| 1 | Engineering | Research and Development |
| 2 | Tool Design | Research and Development |
| 3 | Sales | Sales and Marketing |
| 4 | Marketing | Sales and Marketing |
| 5 | Purchasing | Inventory Management |
| 12 | Document Control | Quality Assurance |
| 13 | Quality Assurance | Quality Assurance |
| 14 | Facilities and Maintenance | Executive General and Administration |
| 15 | Shipping and Receiving | Inventory Management |
| 16 | Executive | Executive General and Administration |

**Tabel Employee**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BusinessEntityID** | **NationalIDNumber** | **LoginID** | **OrganizationNode** | **OrganizationLevel** |
| 1 | 295847284 | adventure-works\ken0 |  | 0 |
| 2 | 245797967 | adventure-works\terri0 | 58 | 1 |
| 3 | 509647174 | adventure-works\roberto0 | 5AC0 | 2 |
| 4 | 112457891 | adventure-works\rob0 | 5AD6 | 3 |
| 5 | 695256908 | adventure-works\gail0 | 5ADA | 3 |
| 286 | 758596752 | adventure-works\lynn0 | 95AB | 3 |
| 287 | 982310417 | adventure-works\amy0 | 9,50E+01 | 2 |
| 288 | 954276278 | adventure-works\rachel0 | 95EB | 3 |
| 289 | 668991357 | adventure-works\jae0 | 95ED | 3 |
| 290 | 134219713 | adventure-works\ranjit0 | 95EF | 3 |

**Tabel Customer**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CustomerID** | **PersonID** | **StoreID** | **TerritoryID** | **AccountNumber** |
| 1 | NULL | 934 | 1 | AW00000001 |
| 2 | NULL | 1028 | 1 | AW00000002 |
| 3 | NULL | 642 | 4 | AW00000003 |
| 4 | NULL | 932 | 4 | AW00000004 |
| 5 | NULL | 1026 | 4 | AW00000005 |
| 30114 | 1985 | 1986 | 7 | AW00030114 |
| 30115 | 1987 | 1988 | 6 | AW00030115 |
| 30116 | 1989 | 1990 | 4 | AW00030116 |
| 30117 | 1991 | 1992 | 4 | AW00030117 |
| 30118 | 1993 | 1994 | 3 | AW00030118 |

**Tabel SalesTeritorry**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **TerritoryID** | **Name** | **CountryRegionCode** | **Group** | **SalesYTD** | **SalesLastYear** | **CostYTD** | **CostLastYear** |
| 1 | Northwest | US | North America | 78.871.867.882 | 32.986.944.938 | 0 | 0 |
| 2 | Northeast | US | North America | 24.021.768.476 | 36.071.489.371 | 0 | 0 |
| 3 | Central | US | North America | 3.072.175.118 | 32.050.140.767 | 0 | 0 |
| 4 | Southwest | US | North America | 105.108.538.739 | 53.665.757.098 | 0 | 0 |
| 5 | Southeast | US | North America | 25.386.672.515 | 39.250.714.318 | 0 | 0 |
| 7 | France | FR | Europe | 47.723.983.078 | 23.965.397.601 | 0 | 0 |
| 8 | Germany | DE | Europe | 38.052.023.478 | 13.079.497.917 | 0 | 0 |
| 9 | Australia | AU | Pacific | 59.778.149.154 | 22.785.489.776 | 0 | 0 |
| 10 | United Kingdom | GB | Europe | 50.129.053.656 | 16.358.233.967 | 0 | 0 |

**Tabel PersonCreditCard**

|  |  |
| --- | --- |
| **BusinessEntityID** | **CreditCardID** |
| 293 | 17038 |
| 295 | 15369 |
| 297 | 8010 |
| 299 | 5316 |
| 301 | 6653 |
| 20773 | 16305 |
| 20774 | 16309 |
| 20775 | 16317 |
| 20776 | 16347 |
| 20777 | 16363 |

**Tabel StateProvince**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **StateProvinceID** | **StateProvinceCode** | **CountryRegionCode** | **IsOnlyStateProvinceFlag** | **Name** |
| 1 | AB | CA | 0 | Alberta |
| 2 | AK | US | 0 | Alaska |
| 3 | AL | US | 0 | Alabama |
| 4 | AR | US | 0 | Arkansas |
| 5 | AS | AS | 1 | American Samoa |
| 177 | 91 | FR | 0 | Essonne |
| 178 | 92 | FR | 0 | Hauts de Seine |
| 179 | 93 | FR | 0 | Seine Saint Denis |
| 180 | 94 | FR | 0 | Val de Marne |
| 181 | 95 | FR | 0 | Val d'Oise |

**Tabel CountryRegion**

|  |  |
| --- | --- |
| **CountryRegionCode** | **Name** |
| AD | Andorra |
| AE | United Arab Emirates |
| AF | Afghanistan |
| AG | Antigua and Barbuda |
| AI | Anguilla |
| YE | Yemen |
| YT | Mayotte |
| ZA | South Africa |
| ZM | Zambia |
| ZW | Zimbabwe |

**Tabel SalesTerritoryHistory**

|  |  |  |  |
| --- | --- | --- | --- |
| **SalesTeritorryID** | **TerritoryID** | **StartDate** | **EndDate** |
| 275 | 2 | 5/31/2011 | 11/29/2012 |
| 275 | 3 | 11/30/2012 | NULL |
| 276 | 4 | 5/31/2011 | NULL |
| 277 | 3 | 5/31/2011 | 11/29/2012 |
| 277 | 2 | 11/30/2012 | NULL |
| 284 | 1 | 9/30/2012 | NULL |
| 286 | 9 | 5/30/2013 | NULL |
| 288 | 8 | 5/30/2013 | NULL |
| 289 | 6 | 5/30/2012 | NULL |
| 290 | 7 | 5/30/2012 | NULL |

**Tabel Store**

|  |  |  |
| --- | --- | --- |
| BusinessEntityID | Name | SalesPersonID |
| 292 | Next-Door Bike Store | 279 |
| 294 | Professional Sales and Service | 276 |
| 296 | Riders Company | 277 |
| 298 | The Bike Mechanics | 275 |
| 300 | Nationwide Supply | 286 |
| 1988 | Retreat Inn | 282 |
| 1990 | Technical Parts Manufacturing | 281 |
| 1992 | Totes & Baskets Company | 277 |
| 1994 | World of Bikes | 277 |
| 2051 | A Bicycle Association | 275 |

**Tabel SalesPersonQuotaHistory**

|  |  |  |
| --- | --- | --- |
| **BusinessEntityID** | **QuotaDate** | **SalesQuota** |
| 274 | 5/31/2011 | 28,000 |
| 274 | 8/31/2011 | 7,000 |
| 274 | 12/1/2011 | 91,000 |
| 274 | 2/29/2012 | 140,000 |
| 274 | 5/30/2012 | 70,000 |
| 290 | 2/28/2013 | 404,000 |
| 290 | 5/30/2013 | 1,262,000 |
| 290 | 8/30/2013 | 1,057,000 |
| 290 | 11/30/2013 | 707,000 |
| 290 | 3/1/2014 | 908,000 |

**Tabel CreditCard**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CreditCardID | CardType | CardNumber | ExpMonth | ExpYear |
| 1 | SuperiorCard | 33332664695310 | 11 | 2006 |
| 2 | Distinguish | 55552127249722 | 8 | 2005 |
| 3 | ColonialVoice | 77778344838353 | 7 | 2005 |
| 4 | ColonialVoice | 77774915718248 | 7 | 2006 |
| 5 | Vista | 11114404600042 | 4 | 2005 |
| 19233 | SuperiorCard | 33335458414079 | 1 | 2005 |
| 19234 | Vista | 11114074915665 | 1 | 2007 |
| 19235 | ColonialVoice | 77774511327745 | 1 | 2006 |
| 19236 | Vista | 11112645284978 | 10 | 2008 |
| 19237 | SuperiorCard | 33336254511031 | 5 | 2005 |

**Tabel CurrencyRate**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CurrencyRateID | CurrencyRateDate | FromCurrencyCode | ToCurrencyCode | AverageRate |
| 1 | 5/31/2011 | USD | ARS | 1.0000 |
| 2 | 5/31/2011 | USD | AUD | 1.5491 |
| 3 | 5/31/2011 | USD | BRL | 1.9379 |
| 4 | 5/31/2011 | USD | CAD | 1.4641 |
| 5 | 5/31/2011 | USD | CNY | 8.2781 |
| 13528 | 5/27/2014 | USD | USD | 1.0000 |
| 13529 | 5/28/2014 | USD | USD | 1.0000 |
| 13530 | 5/29/2014 | USD | USD | 1.0000 |
| 13531 | 5/30/2014 | USD | USD | 1.0000 |
| 13532 | 5/31/2014 | USD | USD | 1.0000 |

**Tabel SalesTaxRate**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SalesTaxRateID | StateProvinceID | TaxType | TaxRate | Name |
| 1 | 1 | 1 | 14 | Canadian GST + Alberta Provincial Tax |
| 2 | 57 | 1 | 14.25 | Canadian GST + Ontario Provincial Tax |
| 3 | 63 | 1 | 14.25 | Canadian GST + Quebec Provincial Tax |
| 4 | 1 | 2 | 7 | Canadian GST |
| 5 | 57 | 2 | 7 | Canadian GST |
| 27 | 79 | 1 | 8.8 | Washington State Sales Tax |
| 28 | 77 | 3 | 10 | Taxable Supply |
| 29 | 20 | 3 | 16 | Germany Output Tax |
| 30 | 84 | 3 | 19.6 | France Output Tax |
| 31 | 14 | 3 | 17.5 | United Kingdom Output Tax |

**Tabel Currency**

|  |  |
| --- | --- |
| CurrencyCode | Name |
| AED | Emirati Dirham |
| AFA | Afghani |
| ALL | Lek |
| AMD | Armenian Dram |
| ANG | Netherlands Antillian Guilder |
| VEB | Bolivar |
| VND | Dong |
| XOF | CFA Franc BCEAO |
| ZAR | Rand |
| ZWD | Zimbabwe Dollar |

**Tabel SalesOrderHeaderSalesReason**

|  |  |
| --- | --- |
| SalesOrderID | SalesReasonID |
| 43697 | 5 |
| 43697 | 9 |
| 43702 | 5 |
| 43702 | 9 |
| 43703 | 5 |
| 75119 | 1 |
| 75120 | 1 |
| 75121 | 1 |
| 75122 | 1 |
| 75123 | 1 |

**Tabel SalesReason**

|  |  |  |
| --- | --- | --- |
| SalesReasonID | Name | ReasonType |
| 1 | Price | Other |
| 2 | On Promotion | Promotion |
| 3 | Magazine Advertisement | Marketing |
| 4 | Television Advertisement | Marketing |
| 5 | Manufacturer | Other |
| 6 | Review | Other |
| 7 | Demo Event | Marketing |
| 8 | Sponsorship | Marketing |
| 9 | Quality | Other |
| 10 | Other | Other |

**Tabel Product**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ProductID** | **Name** | **MakeFlag** | **IsFinishedGood** | **Color** |
| 1 | Adjustable Race | Purchased | No | - |
| 2 | Bearing Ball | Purchased | No | - |
| 3 | BB Ball Bearing | Manufactured | No | - |
| 4 | Headset Ball Bearings | Purchased | No | - |
| 316 | Blade | Manufactured | No | - |
| 995 | ML Bottom Bracket | Manufactured | Yes | - |
| 996 | HL Bottom Bracket | Manufactured | Yes | - |
| 997 | Road-750 Black, 44 | Manufactured | Yes | Black |
| 998 | Road-750 Black, 48 | Manufactured | Yes | Black |
| 999 | Road-750 Black, 52 | Manufactured | Yes | Black |

**Tabel ProductListPriceHistory**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ProductID** | **StartDate** | **EndDate** | **ListPrice** | **IsCurrentlyUsed** |
| 707 | 5/31/2011 | 5/29/2012 | $336.442.00 | No |
| 707 | 5/30/2012 | 5/29/2013 | $336.442.00 | No |
| 707 | 5/30/2013 | NULL | $34.99 | No |
| 708 | 5/31/2011 | 5/29/2012 | $336.442.00 | No |
| 708 | 5/30/2012 | 5/29/2013 | $336.442.00 | No |
| 995 | 5/30/2013 | NULL | $101.24 | No |
| 996 | 5/30/2013 | NULL | $121.49 | No |
| 997 | 5/30/2013 | NULL | $539.99 | No |
| 998 | 5/30/2013 | NULL | $539.99 | No |
| 999 | 5/30/2013 | NULL | $539.99 | No |

**Tabel SpecialOfferProduct**

|  |  |
| --- | --- |
| **SpecialOfferID** | **ProductID** |
| 1 | 680 |
| 1 | 706 |
| 1 | 707 |
| 1 | 708 |
| 1 | 709 |
| 16 | 984 |
| 16 | 985 |
| 16 | 986 |
| 16 | 987 |
| 16 | 988 |

**Tabel SpecialOffer**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SpecialOfferID** | **Description** | **DiscountPct** | **Type** | **Category** |
| 1 | No Discount | 0,00 | No Discount | No Discount |
| 2 | Volume Discount 11 to 14 | 0,02 | Volume Discount | Reseller |
| 3 | Volume Discount 15 to 24 | 0,05 | Volume Discount | Reseller |
| 4 | Volume Discount 25 to 40 | 0,10 | Volume Discount | Reseller |
| 5 | Volume Discount 41 to 60 | 0,15 | Volume Discount | Reseller |
| 12 | LL Road Frame Sale | 0,35 | Excess Inventory | Reseller |
| 13 | Touring-3000 Promotion | 0,15 | New Product | Reseller |
| 14 | Touring-1000 Promotion | 0,20 | New Product | Reseller |
| 15 | Half-Price Pedal Sale | 0,50 | Seasonal Discount | Customer |
| 16 | Mountain-500 Silver Clearance Sale | 0,40 | Discontinued Product | Reseller |

**Tabel SalesOrderDetail**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SalesOrderID** | **SalesOrderDetailID** | **CarrierTrackingNumber** | **OrderQty** | **ProductID** |
| 43659 | 1 | 4911-403C-98 | 1 | 776 |
| 43659 | 2 | 4911-403C-98 | 3 | 777 |
| 43659 | 3 | 4911-403C-98 | 1 | 778 |
| 43659 | 4 | 4911-403C-98 | 1 | 771 |
| 43659 | 5 | 4911-403C-98 | 1 | 772 |
| 75122 | 121313 | - | 1 | 878 |
| 75122 | 121314 | - | 1 | 712 |
| 75123 | 121315 | - | 1 | 878 |
| 75123 | 121316 | - | 1 | 879 |
| 75123 | 121317 | - | 1 | 712 |

**Tabel Address**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **AddressID** | **AddressLine1** | **AddressLine2** | **City** | **StateProvinceID** |
| 1 | 1970 Napa Ct. | - | Bothell | 79 |
| 2 | 9833 Mt. Dias Blv. | - | Bothell | 79 |
| 3 | 7484 Roundtree Drive | - | Bothell | 79 |
| 4 | 9539 Glenside Dr | - | Bothell | 79 |
| 5 | 1226 Shoe St. | - | Bothell | 79 |
| 32517 | 177 11th Ave | - | Sammamish | 79 |
| 32518 | 8040 Hill Ct | - | Redmond | 79 |
| 32519 | 137 Mazatlan | - | Seattle | 79 |
| 32520 | 5863 Sierra | - | Bellevue | 79 |
| 32521 | 7145 Matchstick Drive | - | Sammamish | 79 |

**Tabel SalesOrderHeader**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SalesOrderID** | **RevisionNumber** | **OrderDate** | **DueDate** | **ShipDate** |
| 43659 | 8 | 5/31/2011 | 6/12/2011 | 6/7/2011 |
| 43660 | 8 | 5/31/2011 | 6/12/2011 | 6/7/2011 |
| 43661 | 8 | 5/31/2011 | 6/12/2011 | 6/7/2011 |
| 43662 | 8 | 5/31/2011 | 6/12/2011 | 6/7/2011 |
| 43663 | 8 | 5/31/2011 | 6/12/2011 | 6/7/2011 |
| 75119 | 8 | 6/30/2014 | 7/12/2014 | 7/7/2014 |
| 75120 | 8 | 6/30/2014 | 7/12/2014 | 7/7/2014 |
| 75121 | 8 | 6/30/2014 | 7/12/2014 | 7/7/2014 |
| 75122 | 8 | 6/30/2014 | 7/12/2014 | 7/7/2014 |
| 75123 | 8 | 6/30/2014 | 7/12/2014 | 7/7/2014 |

**Tabel Person**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BusinessEntityID** | **PersonType** | **NameStyle** | **Title** | **FirstName** |
| 1 | Non-Sales Employee | Western | - | Ken |
| 2 | Non-Sales Employee | Western | - | Terri |
| 3 | Non-Sales Employee | Western | - | Roberto |
| 4 | Non-Sales Employee | Western | - | Rob |
| 5 | Non-Sales Employee | Western | Ms. | Gail |
| 20773 | Individual (retail) | Western | - | Crystal |
| 20774 | Individual (retail) | Western | - | Isabella |
| 20775 | Individual (retail) | Western | - | Crystal |
| 20776 | Individual (retail) | Western | - | Crystal |
| 20777 | Individual (retail) | Western | - | Crystal |

**Tabel ShoppingCartItem**

|  |  |  |  |
| --- | --- | --- | --- |
| **ShoppingCartItemID** | **ShoppingCartID** | **Quantity** | **ProductID** |
| 2 | 14951 | 3 | 862 |
| 4 | 20621 | 4 | 881 |
| 5 | 20621 | 7 | 874 |

1. **DATA ANALYSIS [[4]](#footnote-4)**

Analisis data merupakan proses yang berkaitan dengan pemeriksaan rangkaian data secara menyeluruh guna menghasilkan *data driven decision making* dan informasi yang berguna untuk beberapa hal seperti pengambilan keputusan, peningkatan pengetahuan, mitigasi risiko, dan menghindari bias dalam pengambilan keputusan.

Terdapat lima proses inti dari analisis data:

1. Stating and refining the question
2. Exploring the data
3. Building formal statistical models
4. Interpreting the results
5. Communicating the results

Pada masing-masing proses inti tersebut, terdapat tiga buah tahap yang disebut “*epicycle of data analysis*” yang dilakukan secara iteratif, yaitu:

1. *Setting Expectations*

Mengembangkan ekspektasi adalah proses memikirkan dengan sengaja apa yang diharapkan sebelum melakukan sesuatu. Contoh: Menentukan *upper* dan *lower bound*.

1. *Collecting Information*

Langkah ini meliputi pengumpulan informasi tentang pertanyaan pada data. Pertanyaan dapat didasarkan pada mengumpulkan informasi dengan melakukan pencarian literatur atau bertanya kepada para ahli untuk memastikan bahwa pertanyaan Anda adalah pertanyaan yang baik.

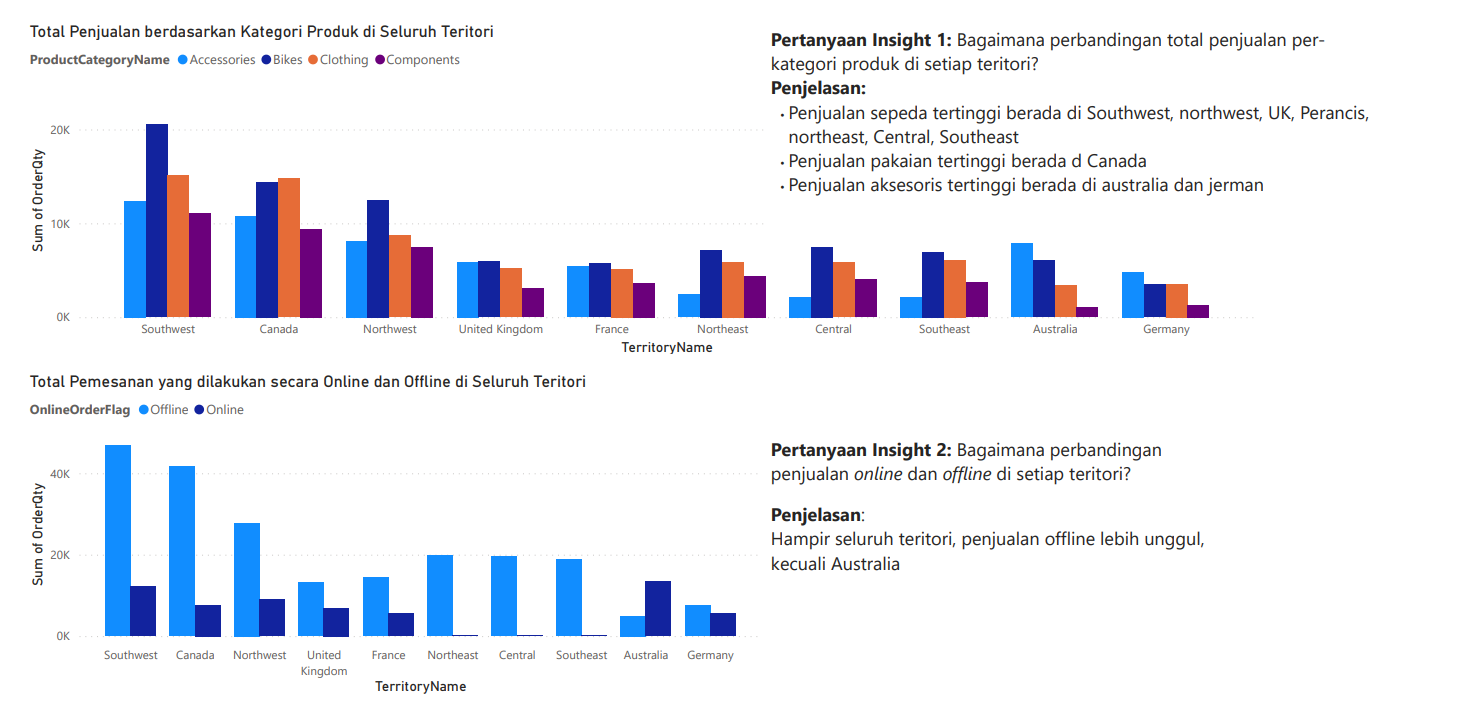
1. *Stating and Refining the Question*

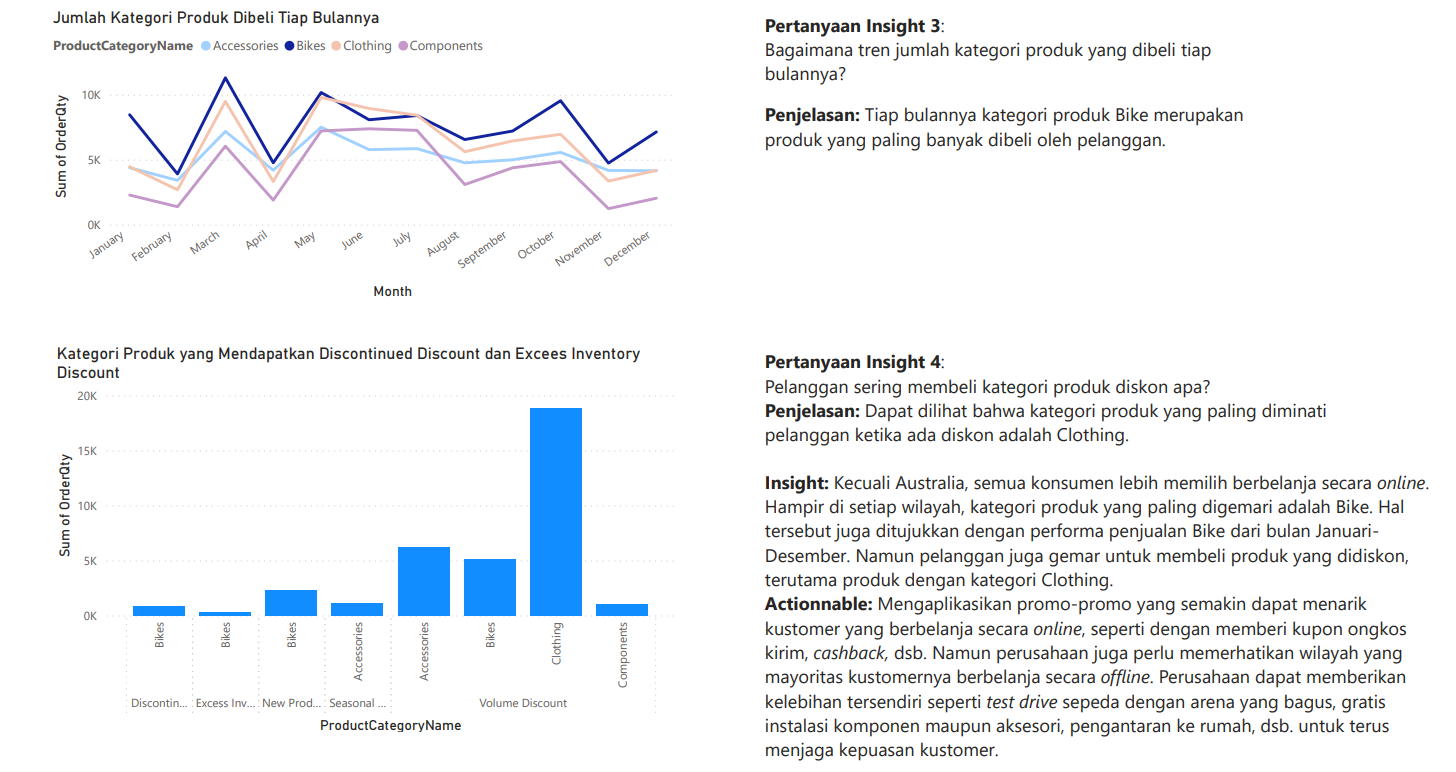
Memahami jenis pertanyaan yang ditanyakan merupakan langkah paling mendasar yang dapat dilakukan untuk memastikan bahwa interpretasi terhadap hasil adalah benar. Sebelum mempelajari cara mengajukan pertanyaan, terdapat beberapa pertimbangan mengenai jenis-jenis pertanyaan. Ada enam jenis pertanyaan dasar yang didasarkan pada makalah yang diterbitkan di *Science* oleh Roger dan Jeff Leek. Berikut enam jenis pertanyaan tersebut:

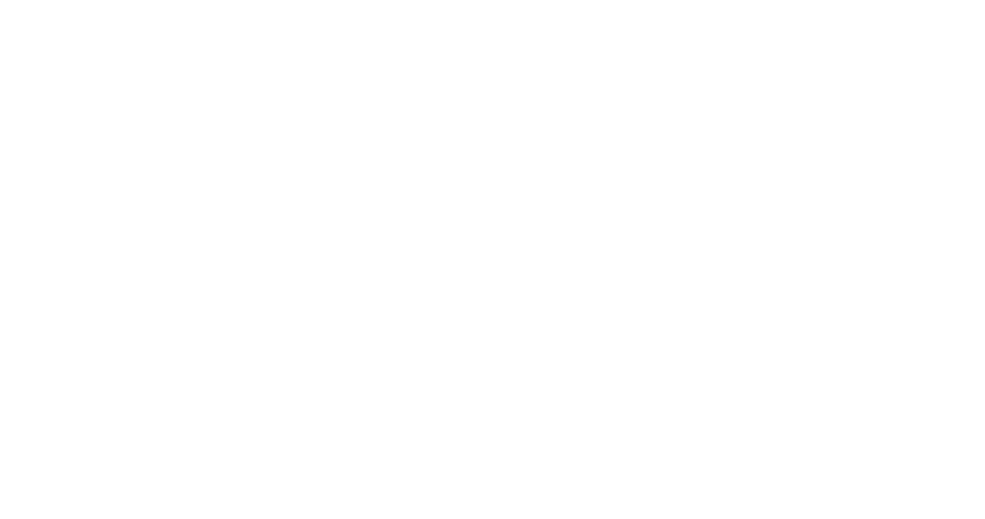
1. *Descriptive*: Pertanyaan yang berusaha meringkas karakteristik dari sekumpulan data.
2. *Exploratory (Hypothesis-Generating):* Menganalisis data untuk melihat apakah terdapat pola, tren, atau hubungan antar variabel.
3. *Inferential*: Pernyataan ulang hipotesis yang diajukan sebagai sebuah pertanyaan dan akan dijawab dengan menganalisis serangkaian data yang berbeda.
4. *Predictive*: Untuk mencari faktor-faktor apa saja yang dapat digunakan untuk memprediksi suatu perilaku/kejadian tertentu.
5. *Causal*: Menanyakan apakah mengubah satu faktor akan mengubah faktor lain, secara rata-rata, dalam suatu populasi.
6. *Mechanistic*: Pertanyaan untuk mencari tahu bagaimana perubahan satu variabel dapat memengaruhi perubahan variabel lainnya.

**Kebutuhan Pertama**

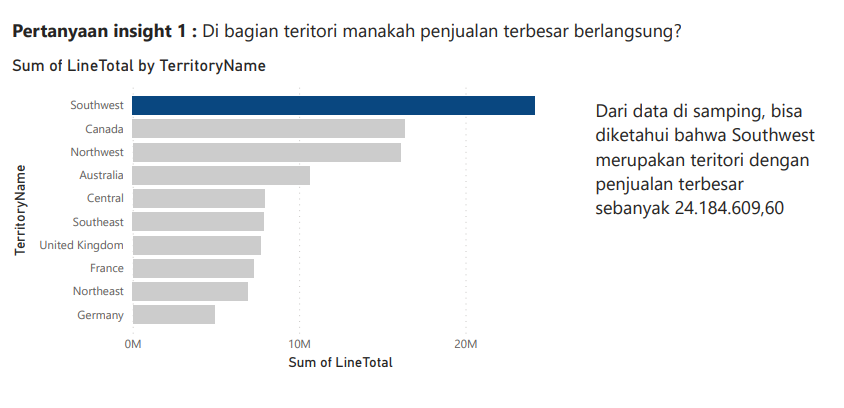
Melakukan analisis pada tabel Customer dan SalesOrderHeader untuk mensegmentasi konsumen berdasarkan perilaku dan lokasi pembelian, sehingga perusahaan dapat mengimplementasikan sistem personalisasi untuk meningkatkan penjualan.

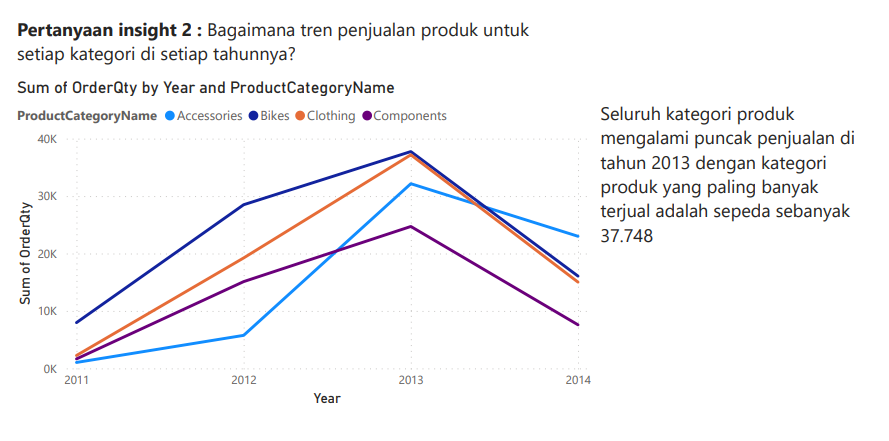


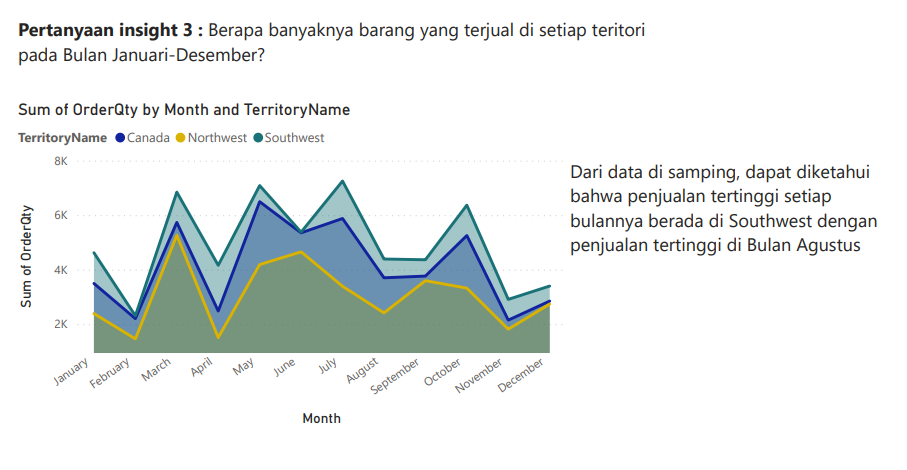


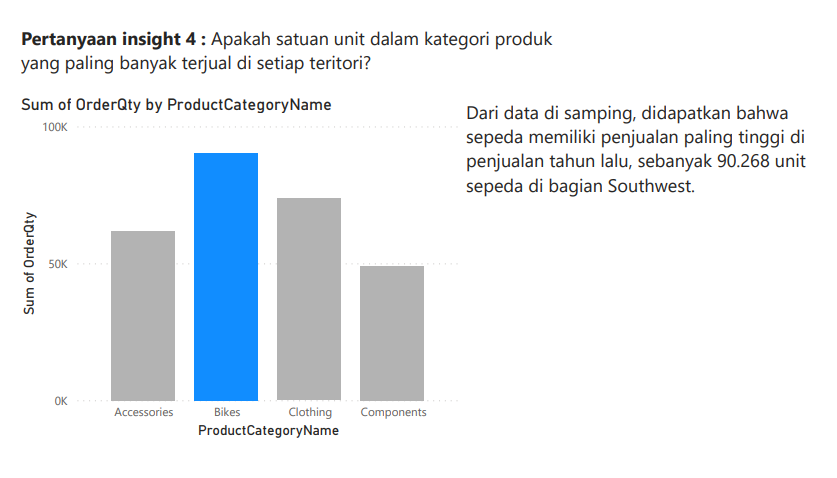


**Kebutuhan Kedua**

Mengevaluasi performa sales pada tabel SalesOrderHeader, SalesTeritorry, dan Product untuk mengoptimalkan performa sales

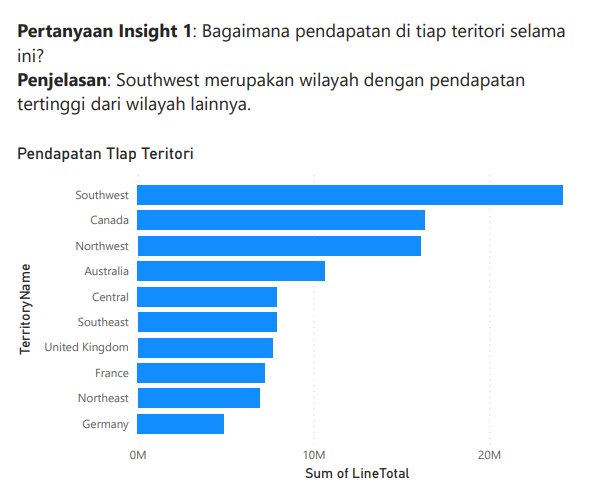


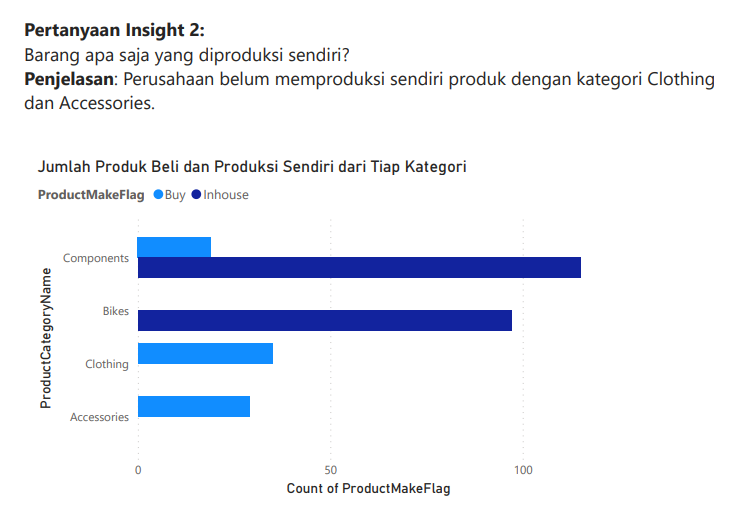


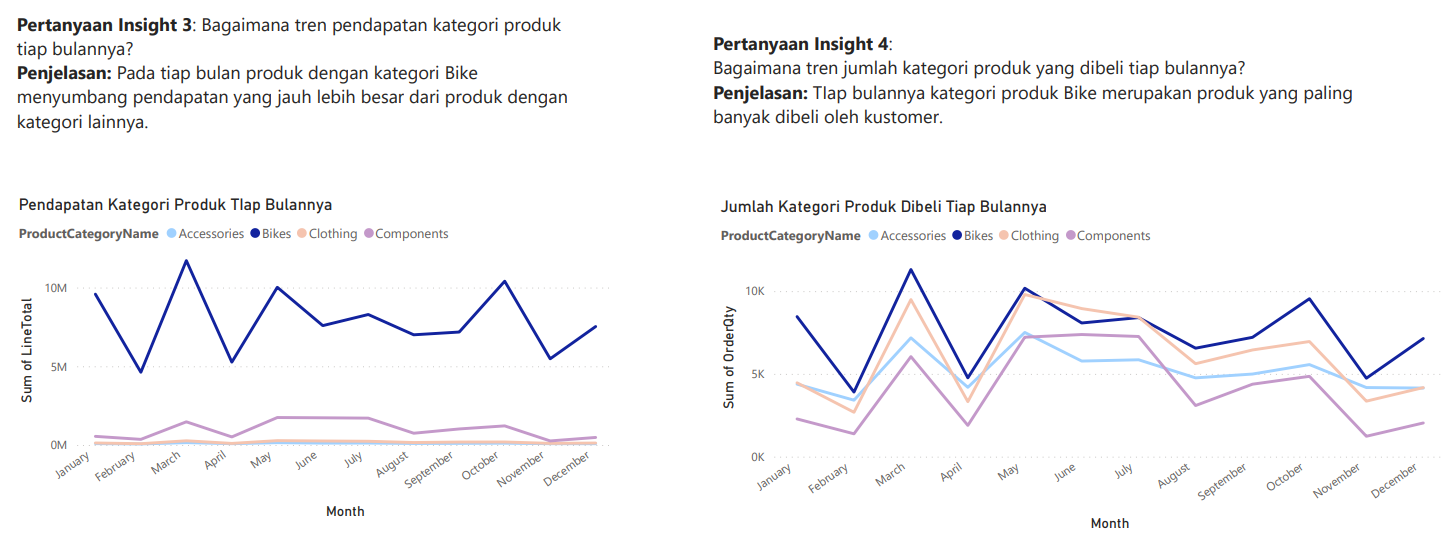


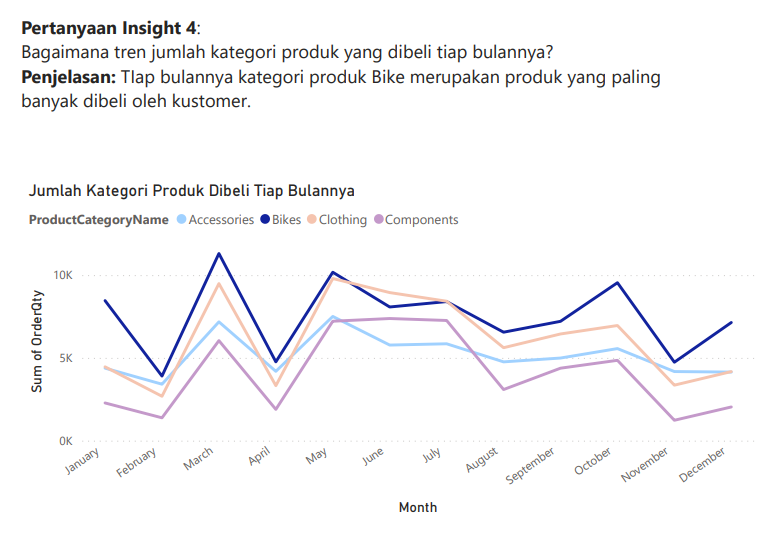
**Kebutuhan Ketiga**

Memanfaatkan data dari tabel SalesOrderHeader, Product, dan Sales Territory untuk mengenali produk-produk yang paling laris, serta wilayah yang paling menguntungkan. Kemudian dilakukan penyesuaian harga, penawaran produk, dan peningkatan pendapatan.



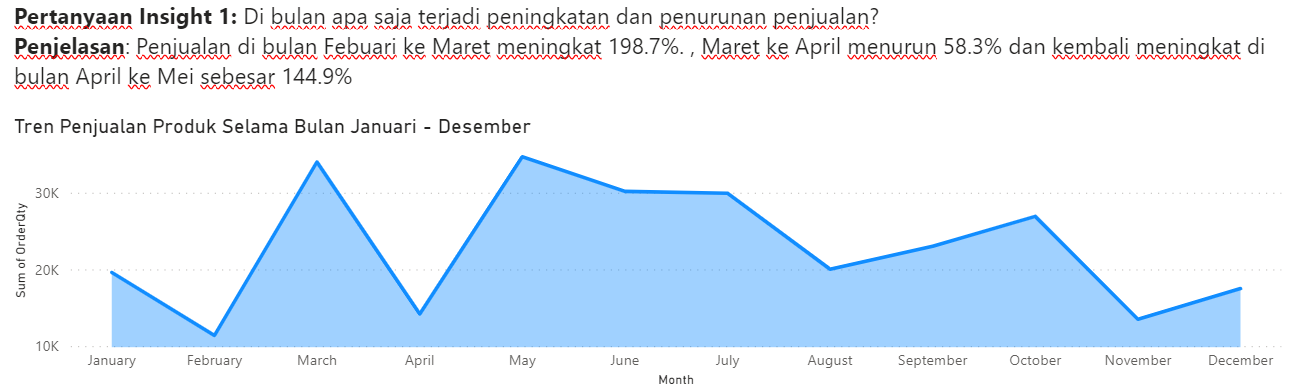




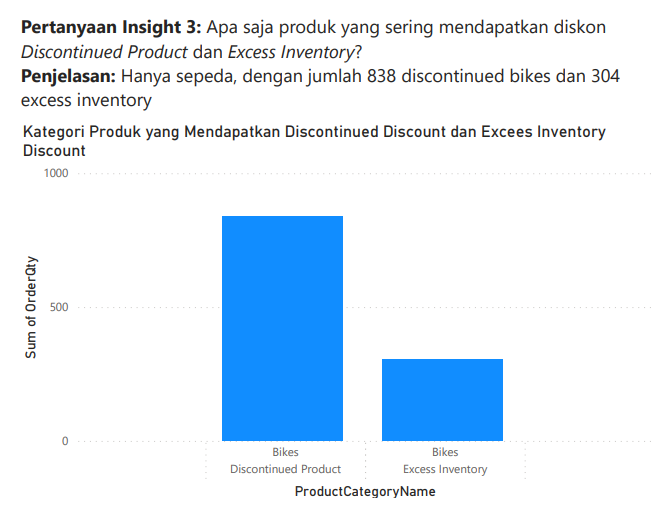


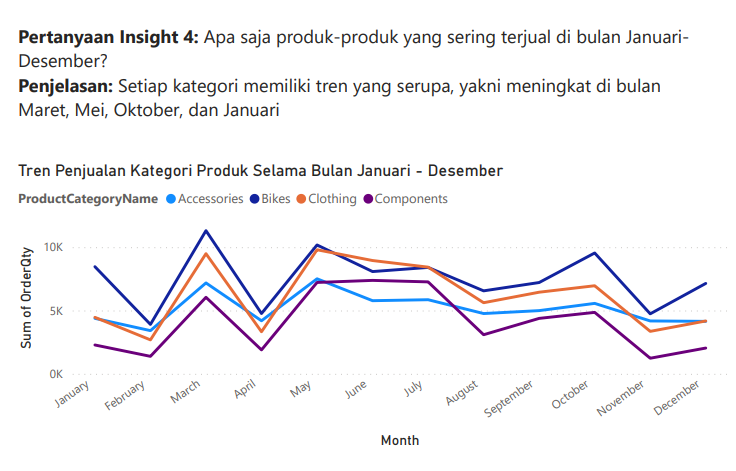
**Kebutuhan Keempat**

Memantau inventori produk dan memprediksi permintaan customer akan produk dengan menggunakan data-data yang ada di table SalesOrderDetails. Ditindaklanjuti lagi dengan mengimplementasikan strategi optimasi inventori untuk mengurangi biaya kirim dan memastikan produk selalu dapat dibeli customer.









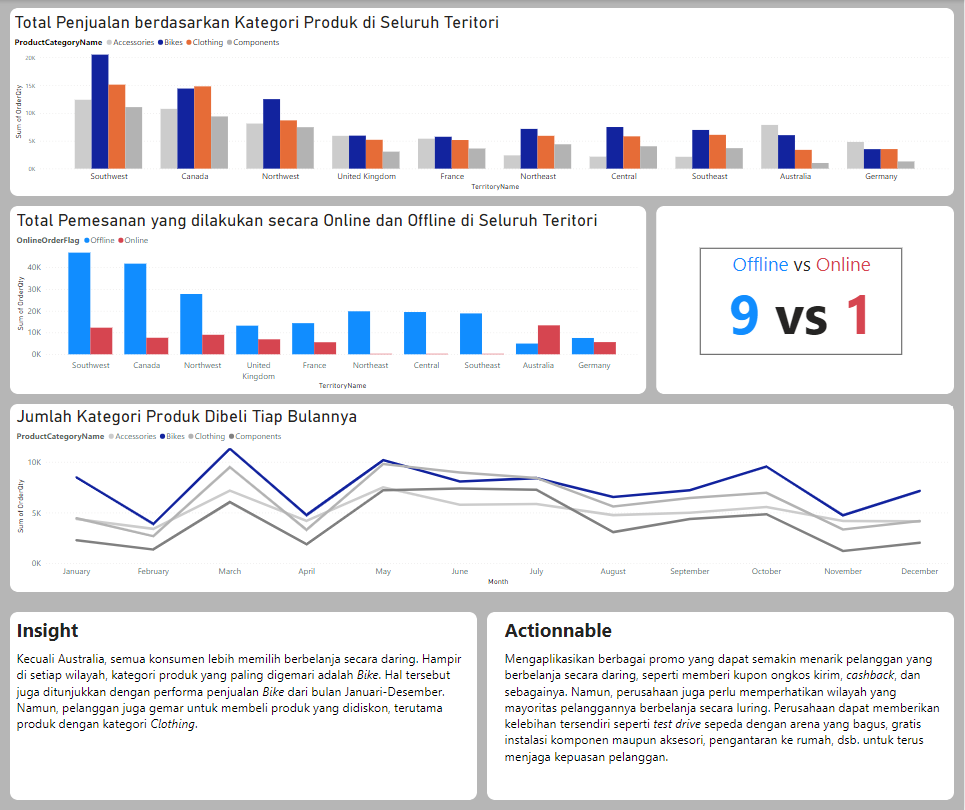
1. **DASHPORT**

**Definisi Dashport:**

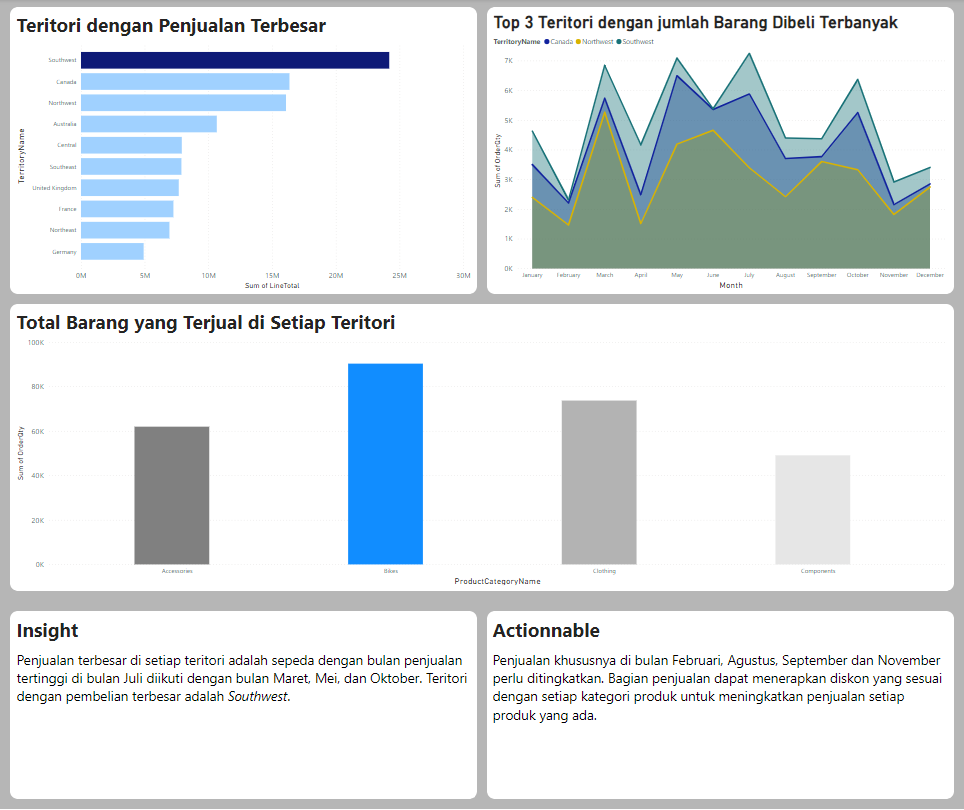
*Dashport* merupakan akronim yang berasal dari kata *dashboard* dan *report*. Dalam konteks organisasi, *dashboard* merupakan kumpulan berbagai visualisasi yang memberikan pandangan menyeluruh atas organisasi tersebut, sedangkan *report* merupakan visualisasi berhubungan yang menampilkan analisis mendalam atas suatu topik. *Report* hanya berfokus pada area tertentu organisasi saja, digunakan untuk menjawab pertanyaan dengan menampilkan data yang detail, dan informasi yang disajikan bersifat *snapshot* (tidak *realtime*). Maka dari itu, *dashport* menyajikan kumpulan visualisasi berhubungan mengenai suatu area tertentu organisasi yang disajikan dalam suatu *platform* tertentu untuk memberikan gambaran mengenai area organisasi tersebut

**Link dashport:**

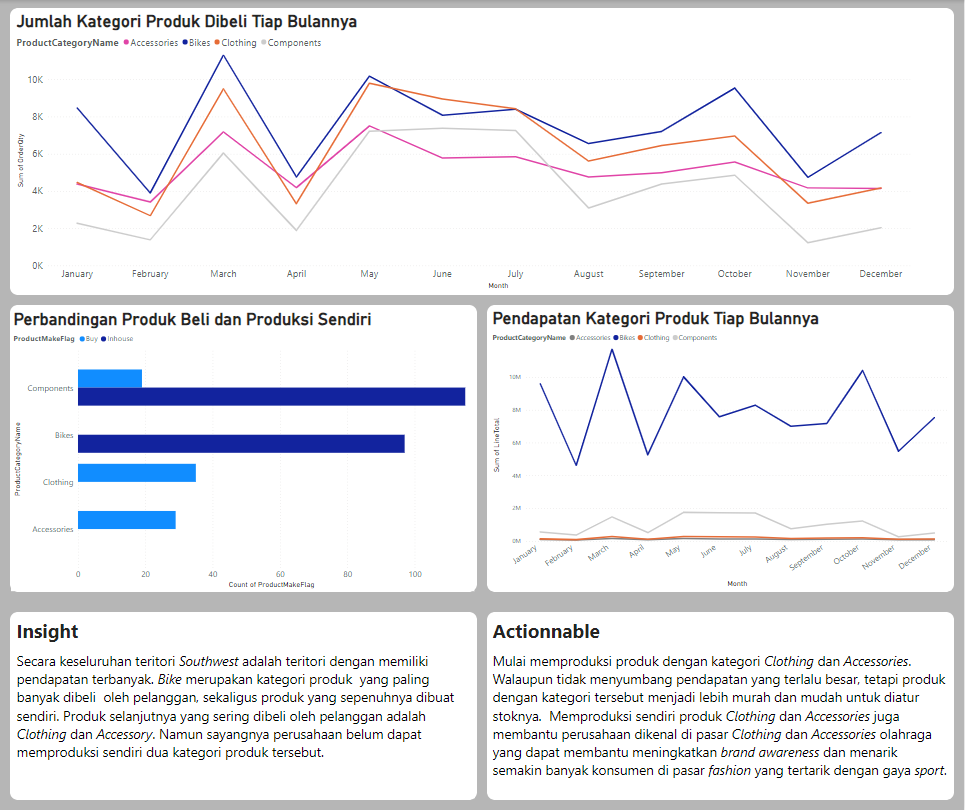
|  |
| --- |
| [**https://app.powerbi.com/links/2CjueyNCQm?ctid=8850c7df-a3c9-4f0f-bef5-dd85775a44bd&pbi\_source=linkShare**](https://app.powerbi.com/links/2CjueyNCQm?ctid=8850c7df-a3c9-4f0f-bef5-dd85775a44bd&pbi_source=linkShare) |

****

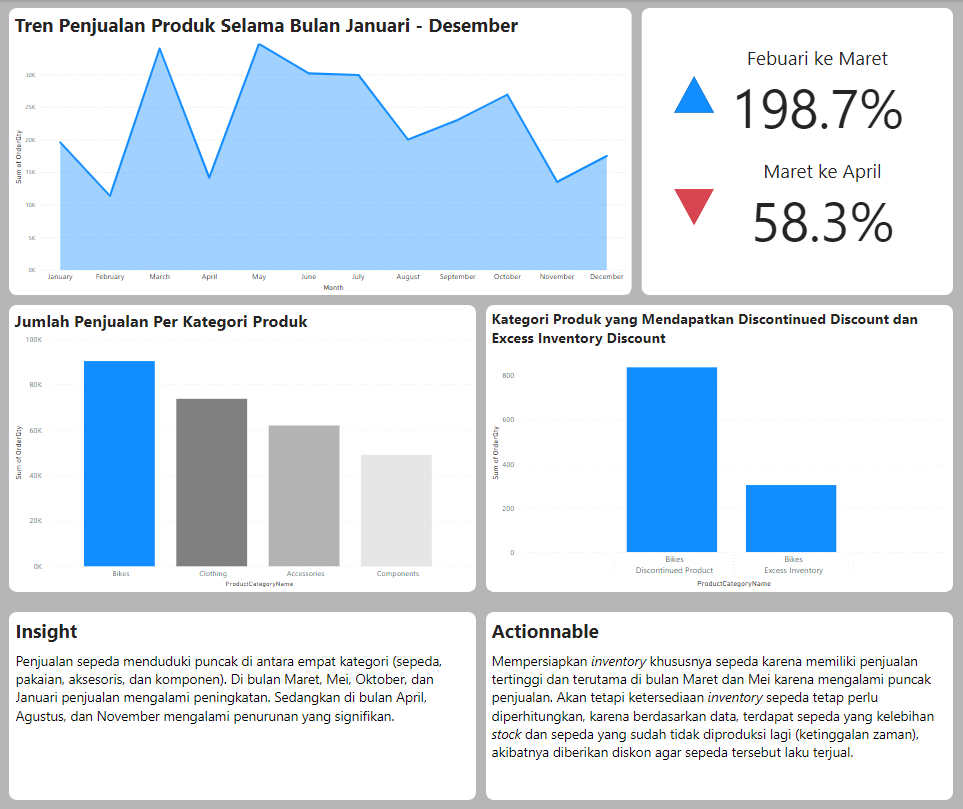
**Gambar 1. Dashport 1**

****

**Gambar 2. Dashport 2**

****

**Gambar 3. Dashport 3**



**Gambar 4. Dashport 4**

1. **ACTIONABLE INSIGHTS**

Berikut adalah insights yang didapat dari kegiatan analisis data:

1. Kecuali Australia, semua konsumen lebih memilih berbelanja secara daring. Hampir di setiap wilayah, kategori produk yang paling digemari adalah *Bike*. Hal tersebut juga ditunjukkan dengan performa penjualan *Bike* dari bulan Januari-Desember. Namun, pelanggan juga gemar untuk membeli produk yang didiskon, terutama produk dengan kategori *Clothing*.
2. Penjualan terbesar di setiap teritori adalah sepeda dengan bulan penjualan tertinggi di bulan Juli diikuti dengan bulan Maret, Mei, dan Oktober. Teritori dengan pembelian terbesar adalah *Southwest*.
3. Secara keseluruhan teritori *Southwest* adalah teritori dengan memiliki pendapatan terbanyak. Sepeda merupakan kategori produk yang paling banyak diincar dibeli oleh pelanggan, sekaligus produk yang sepenuhnya dibuat sendiri. Produk selanjutnya yang sering dibeli oleh pelanggan adalah *Clothing* dan *Accessory*. Namun sayangnya perusahaan belum dapat memproduksi sendiri dua kategori produk tersebut.
4. Penjualan sepeda menduduki puncak di antara empat kategori (sepeda, pakaian, aksesoris, dan komponen). Di bulan Maret, Mei, Oktober, dan Januari penjualan mengalami peningkatan. Sedangkan di bulan April, Agustus, dan November mengalami penurunan yang signifikan.

Berikut adalah actionable insights yang didapat dari kegiatan analisis data:

1. Mengaplikasikan berbagai promo yang dapat semakin menarik pelanggan yang berbelanja secara daring, seperti memberi kupon ongkos kirim, *cashback*, dan sebagainya. Namun, perusahaan juga perlu memperhatikan wilayah yang mayoritas pelanggannya berbelanja secara luring. Perusahaan dapat memberikan kelebihan tersendiri seperti *test drive* sepeda dengan arena yang bagus, gratis instalasi komponen maupun aksesori, pengantaran ke rumah, dsb. untuk terus menjaga kepuasan pelanggan.
2. Penjualan khususnya di bulan Februari, Agustus, September dan November perlu ditingkatkan. Bagian penjualan dapat menerapkan diskon yang sesuai dengan setiap kategori produk untuk meningkatkan penjualan setiap produk yang ada.
3. Mulai memproduksi produk dengan kategori *Clothing* dan *Accessories*. Walaupun tidak menyumbang pendapatan yang terlalu besar, tetapi produk dengan kategori tersebut menjadi lebih murah dan mudah untuk diatur stoknya. Memproduksi sendiri produk *Clothing* dan *Accessories* juga membantu perusahaan dikenal di pasar *Clothing* dan *Accessories* olahraga yang dapat membantu meningkatkan *brand awareness* dan menarik semakin banyak konsumen di pasar *fashion* yang tertarik dengan gaya *sport*.
4. Mempersiapkan *inventory* khususnya sepeda karena memiliki penjualan tertinggi dan terutama di bulan Maret dan Mei karena mengalami puncak penjualan. Akan tetapi ketersediaan *inventory* sepeda tetap perlu diperhitungkan, karena berdasarkan data, terdapat sepeda yang kelebihan *stock* dan sepeda yang sudah tidak diproduksi lagi (ketinggalan zaman), akibatnya diberikan diskon agar sepeda tersebut laku terjual.
5. **LAMPIRAN**

***Import Libraries***

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

**Membaca File CSV**

columns = ['SalesTaxRateID','StateProvinceID','TaxType','TaxRate', 'Name', 'rowguid', 'ModifiedDate' ]

df = pd.read\_csv('../AdventureWorks-oltp-install-script/SalesTaxRate.csv', delimiter='\t', header=None, names=columns)

df.shape **#Melihat Jumlah baris dan Kolom**

df.head() **#Melihat record 5 baris pertama**  
  
df.tail() **#Melihat record 5 baris terakhir**

df.info() **#Cek tipe data dan *null***

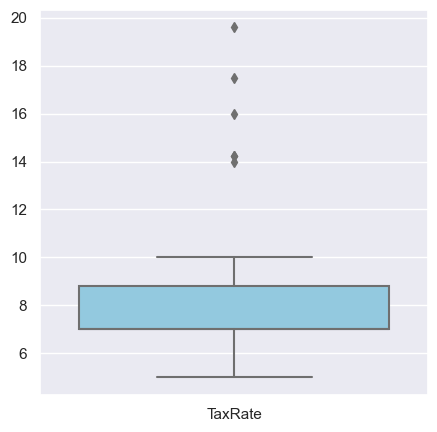
df.describe() **#Deskripsi tiap kolom (nilai min, max, std, mean)**  
  
df.isna().any() **#Cek kolom yang memiliki nilai *null***

**Membuat visualisasi dengan *boxplot***  
  
sns.color\_palette("husl", 9)

sns.set(rc={'figure.figsize':(5, 5)})

sns.boxplot(data=df[['TaxRate']], color='skyblue')

plt.show()



Lampiran 1. Kode program untuk melakukan EDA pada tabel-tabel AdventureWorks

SELECT

StateProvinceID, StateProvince, CountryRegion, [Group], SalesOrderID, OrderDate, SpecialOfferID, OnlineOrderFlag, ProductID, OrderQty, LineTotal

FROM

(

SELECT

T1.[TerritoryID], T1.SalesOrderID, T1.[OrderDate], T2.[SpecialOfferID], T1.[OnlineOrderFlag], T2.[ProductID], T2.OrderQty, T2.LineTotal

FROM

Sales.SalesOrderHeader AS T1 INNER JOIN Sales.SalesOrderDetail AS T2 ON

T1.SalesOrderID=T2.SalesOrderID

) AS T1

INNER JOIN

(

SELECT

T1.[TerritoryID], T1.StateProvinceID, T1.Name AS 'StateProvince', T2.Name AS 'CountryRegion', T2.[Group] AS 'Group'

FROM

Person.StateProvince AS T1 INNER JOIN Sales.SalesTerritory AS T2

ON T1.TerritoryID=T2.TerritoryID

) AS T2

ON T1.TerritoryID=T2.TerritoryID

Lampiran 2. Query untuk membuat Star Schema. Akan dilakukan transformasi lagi pada hasil dari query ini sebelum benar-benar dapat digunakan.

1. Cielen, D., B., M. A. D., & Ali, M. (2016). *Introducing data science: Big data, Machine Learning, and more, using python tools*. Manning Publications. [↑](#footnote-ref-1)
2. Sherman, R. (2015). *Business intelligence guidebook: From Data Integration to Analytics*. Elsevier/Morgan Kaufmann. [↑](#footnote-ref-2)
3. Slide Kuliah Sistem Kecerdasan Bisnis 06A\_DataPreparation [↑](#footnote-ref-3)
4. Peng, R. D., & Matsui, E. (2015). *The Art of Data Science: A Guide for anyone who works with data*. Leanpub. [↑](#footnote-ref-4)