|  |  |  |
| --- | --- | --- |
| Indeks | 258996, 260338 | Marcin Blicharski, Adam Bednarski |
| Grupa | E | |

W miarę możliwości/potrzeby proszę o dodatkowy komentarz do kodu – w szczególności, warto zaznaczyć zrealizowane dodatkowe elementy, czy skomentować przyjętą interpretację pewnych elementów składowych.

# Zadanie 1

Jako rozwiązanie zadania prześlij zrzut ekranu pokazujący daną wizualizację, opisz wybrane składowe wizualizacji (np. zastosowane atrybuty oraz ich umiejscowienie) i dla każdego punktu zastanów się (i krótko przedstaw) na jakiego rodzaju pytania/potrzeby użytkowników (decyzje) może odpowiadać dana wizualizacja.

## Macierz

Obraz zawierający stół

Opis wygenerowany automatycznie

## Wykres kolumnowy

Obraz zawierający wykres

Opis wygenerowany automatycznie

## Wykres KOŁOWY

Obraz zawierający wykres, diagram kołowy

Opis wygenerowany automatycznie

## Wykres punktowy

Obraz zawierający wykres

Opis wygenerowany automatycznie

## Konkluzje dot. zadaniA 1

Macierz – w wizualizacji do pola wiersze został użyty atrybut „ProductCategory.Name” oznaczający kategorię produktu, do pola kolumn został wstawiony atrybut „Product.Color” oznaczający kolor produktu, natomiast do pola wartości zagregowany do liczby wystąpień (Count) atrybut „Product.ProductID” (klucz główny tabeli). Wizualizacja może być przydatna dla użytkowników, których przykładowo interesuje, ile produktów o danym kolorze znajduje się w danej kategorii. Można ją też wykorzystać do porównania, ile mamy produktów w ramach danej kategorii i/lub koloru.

Wykres kolumnowy - w wizualizacji do pola osi X został wstawiony atrybut „SSDYEAR” oznaczający rok zakończenia sprzedaży produktu, do pola osi Y został wstawiony zagregowany do liczby wystąpień (Count) atrybut „Product.ProductID” (klucz główny tabeli). Wizualizacja może być przydatna dla użytkowników, których przykładowo interesuje, ile sztuk produktu zostało sprzedanych w danym roku. Pozwoli to np. na analizę sytuacji w której znajduje się firma – czy występuje sprzedażowy wzrost lub spadek na przestrzeni ostatnich lat.

Wykres kołowy - w wizualizacji do pola legenda został wstawiony atrybut „ProductCategory.Category”, do pola wartości został wstawiony zagregowany do liczby wystąpień (Count) atrybut „Product.ProductID” (klucz główny tabeli). Wizualizacja może być przydatna dla użytkowników, których interesuje procentowy udział produktów o poszczególnych kategoriach. Sam wykres kołowy jest bardzo przejrzysty i czytelny jako forma reprezentacji graficznej udziału procentowego poszczególnych wartości w zbiorze.

Wykres punktowy - w wizualizacji do osi X został wstawiony atrybut „Product.ListPrice”, do pola osi Y został wstawiony atrybut „Product.Profit”. Wizualizacja może być przydatna dla użytkowników, których przykładowo interesuje zależność zysku ze sprzedaży produktów od ich ceny. Wykres dostarcza też informacji o rozkładzie wartości zysku (np. widoczne jest duże zagęszczenie punktów blisko punktu (0,0), co prawdopodobnie oznacza dużo produktów zapewniających niski zysk w stosunku do swojej też niskiej ceny. Do wykresu dodaliśmy „Zoom slider” z uwagi na duże odchylenie standardowe, by móc sprawdzić szczegółowo jak rozkładają się wartości w różnych przedziałach np. dla ceny w przedziale od 0 do 1000.

# Zadanie 2

Jako rozwiązanie zadania prześlij kod z edytora zaawansowanego i krótko opisz rozwiązanie. Dodatkowo odpowiedz na dod. pytanie

## KOD – sprzedawca

Table.NestedJoin(#"Sales vSalesPerson", {"CountryRegionName"}, #"Person CountryRegion", {"Name"}, "Person CountryRegion", JoinKind.LeftOuter)

Table.NestedJoin(Source, {"TerritoryName"}, #"Sales SalesTerritory", {"Name"}, "Sales SalesTerritory", JoinKind.LeftOuter)

Table.ExpandTableColumn(#"Merged Queries", "Person CountryRegion", {"CountryRegionCode", "Name"}, {"Person CountryRegion.CountryRegionCode", "Person CountryRegion.Name"})

Table.ExpandTableColumn(#"Expanded Person CountryRegion", "Sales SalesTerritory", {"TerritoryID", "Name", "Group", "SalesYTD", "SalesLastYear"}, {"Sales SalesTerritory.TerritoryID", "Sales SalesTerritory.Name", "Sales SalesTerritory.Group", "Sales SalesTerritory.SalesYTD", "Sales SalesTerritory.SalesLastYear"})

Table.RemoveColumns(#"Expanded Sales SalesTerritory",{"Title", "PhoneNumber", "PhoneNumberType", "EmailAddress", "EmailPromotion", "AddressLine1", "AddressLine2", "PostalCode"})

## Kod – region sprzedaży

Table.ExpandRecordColumn(#"Removed Columns", "Sales.SalesOrderHeader", {"SalesOrderID", "OrderDate", "SalesOrderNumber", "CustomerID", "SalesPersonID", "TerritoryID"}, {"Sales.SalesOrderHeader.SalesOrderID", "Sales.SalesOrderHeader.OrderDate", "Sales.SalesOrderHeader.SalesOrderNumber", "Sales.SalesOrderHeader.CustomerID", "Sales.SalesOrderHeader.SalesPersonID", "Sales.SalesOrderHeader.TerritoryID"})

Table.ExpandTableColumn(#"Expanded Sales.SalesOrderHeader", "Sales SalesOrderHeader", {"Sales.SalesTerritory"}, {"Sales SalesOrderHeader.Sales.SalesTerritory"})

Table.ExpandRecordColumn(#"Expanded Sales SalesOrderHeader", "Sales SalesOrderHeader.Sales.SalesTerritory", {"TerritoryID", "Name", "CountryRegionCode", "Group", "Person.CountryRegion"}, {"Sales SalesOrderHeader.Sales.SalesTerritory.TerritoryID", "Sales SalesOrderHeader.Sales.SalesTerritory.Name", "Sales SalesOrderHeader.Sales.SalesTerritory.CountryRegionCode", "Sales SalesOrderHeader.Sales.SalesTerritory.Group", "Sales SalesOrderHeader.Sales.SalesTerritory.Person.CountryRegion"})

## Kod – sprzedaż

Table.NestedJoin(#"Sales SalesOrderDetail", {"SalesOrderID"}, #"Sales SalesOrderHeader", {"SalesOrderID"}, "Sales SalesOrderHeader", JoinKind.LeftOuter)

Table.RemoveColumns(Source,{"CarrierTrackingNumber", "SpecialOfferID", "rowguid", "ModifiedDate"})

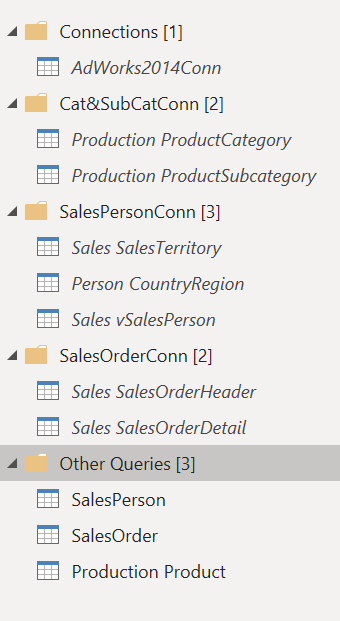
## Zastanów się, czy możliwe jest dołączenie do modelu informacji o czasie realizacji zamówienia (w dniach), wartości podatku czy kosztach przesyłki. Odpowiedź uzasadnij (nie implementuj)

Możliwe jest dołączenie do modelu informacji o czasie realizacji zamówienia (w dniach). Z tabeli SalesOrderHeader należy obliczyć liczbę dni miedzy datą z atrybutu ShipDate (oznaczającą przybliżoną datę dostarczenia produktu) i datą z atrybutu OrderDate (oznaczającą datę zamówienia produktu). Oczywiście będzie to wartość przybliżona, gdyż ShipDate jest **estymowaną** datą dostarczenia przesyłki.

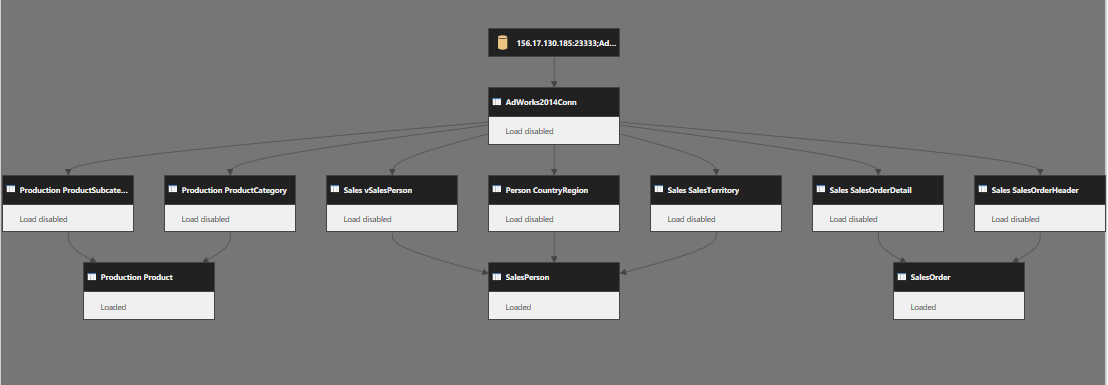
Wartość podatku zawarta jest w atrybucie TaxAmt w tabeli SalesOrderHeader, więc możemy dołączyć ją do modelu informacji. Należy tutaj pamiętać, że jest to podatek od całości zamówienia, a nie poszczególnych produktów w zamówieniu.

Koszty przesyłki zawarte są w atrybucie Freight. Tutaj również jest to koszt całego zamówienia, jednak wpływ na niego ma prawdopodobnie więcej parametrów (np. waga, rozmiar itd.). Możemy tę informację zawrzeć w modelu informacji.

## Konkluzje dot. zadaniA 2



Zadanie zrealizowaliśmy na podstawie instrukcji z polecenia, czyli utworzyliśmy dwa zapytania – **SalesPerson** (składa się na nie widok vSalesPerson oraz informacje o regionie sprzedaży – tabele CountryRegion oraz SalesTerritory) dla informacji o sprzedawcy oraz **SalesOrder** (składają się na nie tabele SalesOrderHeader oraz SalesOrderDetail) dla informacji o zamówieniach. Same źródła zostały, wzorem z poprzedniej listy, odseparowane od logiki przetwarzania. Z SalesPerson usunęliśmy nieistotne według nas atrybuty, np. typ numeru telefonu lub adresy email. Są to ważne informacje w bazie danych, ale dla nas pod kątem analizy danych sprzedażowych lub produktowych te informacje nie są w żaden sposób potrzebne. Z SalesOrder również usunęliśmy niepotrzebne atrybuty, np. rowguid czy ModifiedDate (te atrybuty usuwaliśmy już w przypadku innych tabel, na innej liście zadań). Z uwagi na strukturę, na jaką się zdecydowliśmy kody źródłowe wklejone powyżej nie odnoszą się bezpośrednio do nagłówków przygotowanych przez Pana tj. „kod sprzedaży”, „kod region sprzedaży”. Końcowy widok połączeń jest następujący:

****

# Zadanie 3

Jako rozwiązanie zadania prześlij zrzut ekranu gotowego raportu (wizualizacje + filtry) i przygotuj krótki opis (zgodnie z treścią zadania).

## Raport – zrzut ekranu

Obraz zawierający wykres

Opis wygenerowany automatycznie

Widok strony raportu

Obraz zawierający tekst

Opis wygenerowany automatycznie

Filtry podpunkty e) i f)

Obraz zawierający tekst

Opis wygenerowany automatycznie

Filtr podpunkt g)

## Raport – opis

1. Wizualizacja zwraca poprawne wyniki i nie było żadnych problemów – problemem może być tylko przejrzystość takiej wizualizacji (bardzo dużo rekordów, brak agregacji danych).
2. Wizualizacja zwraca poprawne wyniki, brak problemów.
3. Macierz wyświetla poprawne wyniki, jednak taki sposób wizualizacji nie jest optymalny dla tych danych, gdyż występuje dużo pól bez wartości, a filtrowanie nie poprawia w tym przypadku wyglądu.
4. Wizualizacja zwraca poprawne wyniki.
5. Hierarchia zdefiniowana w lewym dolnym rogu, zwraca poprawne wyniki.
6. Wykres kolumnowy zwraca poprawne wyniki.
7. Wybierając jeden z regionów pozostaje on wyszczególniony zarówno na tej wizualizacji, jak i na wszystkich innych. Wartości prezentowane na innych wizualizacjach zostają ograniczone tylko do tych, które odnoszą się do wybranego regionu. Pozostałe wartości są wyszarzane lub zerowane.
8. Wybierając jeden z elementów tabeli mogę następnie wybrać tylko te elementy z wykresu kolumnowego, które odnoszą się do elementu tabeli. Największy sens ma to np. dla USA.
9. Sterowanie góra-dół strzałkami pozwala na przechodzenie między poziomami w hierarchii, w sposób zdefiniowany przez użytkownika. W tym przypadku najwyżej w hierarchii jest TerritoryGroup. Po przejściu na jego poziom wyświetlane są wyniki sprzedażowe w podziale tylko na TerritoryGroup (3 kolumny). Po przejściu poziom niżej wyświetlane są wyniki sprzedażowe w podziale na Region (6 kolumn) itd.
10. Po wybraniu konkretnego koloru w filtrze (np. Black) wartości wyświetlane na wizualizacjach (wszystkich) ograniczają się tylko do tych odnoszących się do produktów o kolorze czarnym. Pozostałe są zerowane lub wyszarzane.
11. Filtrowania możemy dokonać za pomocą wybrania konkretnego zakresu daty, a wartości wyświetlane na wizualizacjach (wszystkich) odnoszą tylko do tych, które mieszczą się w wybranym zakresie daty.
12. Filtrując w taki sposób, by wyświetlał się tylko jeden region o największej sprzedaży (TOP N = 1) uzyskujemy informację o tym, że są to Stany Zjednoczone. Po kliknięciu w tę kolumnę wykres kołowy wyszarza się dla wszystkich wartości oprócz United States. Podobnie zachowują się inne wizualizacje.
13. Wizualizacja zwraca poprawne wyniki.
14. Wizualizacja zwraca poprawne wyniki.
15. Wizualizacja zwraca poprawne wyniki.

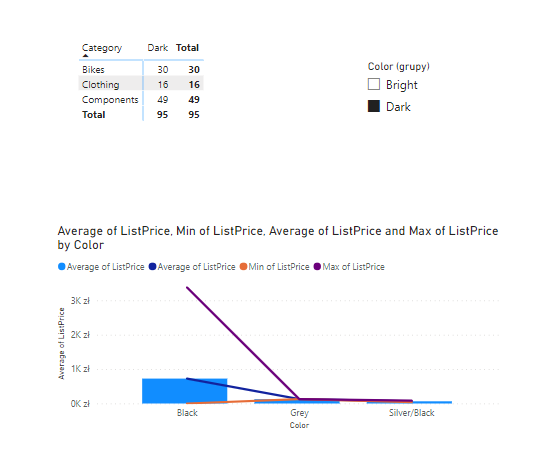
## Konkluzje dot. zadaniA 3

Zadanie 3 w bardzo dobry sposób rozjaśniło nam działanie filtrów na różnych poziomach oraz zastosowanie poszczególnych wizualizacji. Szczególnie przydatną funkcją jest zaznaczanie konkretnych pól, kolumn lub fragmentów wykresów, co pozwala na ograniczenie wyników odnoszących się do tego pola na całej stronie raportu.

# Zadanie 4

Jako rozwiązanie zadania prześlij zrzut ekranu gotowego raportu (wizualizacje + filtry) i przygotuj krótki opis (zgodnie z treścią zadania).

## Raport ½ – zrzut ekranu



Widok strony raportu

Obraz zawierający tekst

Opis wygenerowany automatycznie

Filtr kolorów

Obraz zawierający tekst

Opis wygenerowany automatycznie

Filtr zadanie 1a

## Raport ½ – opis

Wykonany został podpunkt 1.

1. Do zidentyfikowania kategorii produktów, w podziale na kolory, o więcej niż 5 produktach została użyta macierz przy użyciu atrybutów Category oraz Color w kolejno wierszach i kolumnach, a dla pola wartości „liczba elementów ProductID” – na tym atrybucie został zastosowany filtr dla wartości większych lub równych 5 (zinterpretowaliśmy to jako więcej niż 5 produktów w ramach jednej kategorii, a nie koloru). Ponadto zdecydowaliśmy się na zastosowanie hierarchii kolorów, aby użytkownik mógł wybrać czy chce wyświetlać wartości w podziale na Bright/Dark, czy w podziale na poszczególne kolory.
2. Na wykresie kolumnowym do osi X wybrany został atrybut Color, na osi y – średnia z atrybutu „ListPrice” jako średnia cena sprzedaży. Dodatkowo jako linie osi y wybraliśmy kolejno wartość minimalną z ListPrice, średnią z ListPrice oraz maksymalną z ListPrice.
3. Aby dodać fragmentator z podziałem na grupy kolorów został dodany nowy atrybut - Color(grupy) i następnie użyty w polu wizualizacji.

## Raport 3 – zrzut ekranu

Obraz zawierający wykres

Opis wygenerowany automatycznie

Widok strony raportu

## Raport 3 – opis

## Konkluzje dot. zadaniA 3

### PODSUMOWANIE – Konkluzje do całej listy zadań

Zadanie 1 z listy było dla nas można powiedzieć rozgrzewkowe i nie wymagało większego nakładu pracy. Jesteśmy usatysfakcjonowani z uzyskanych wyników.

Zadanie 2 może nie przysporzyło nam większych problemów, ale wymagało dogłębnej analizy i interpretacji tego, co chcemy osiągnąć. Ostatnie zdecydowaliśmy się na pewną interpretację polecenia, która według nas dała najbardziej efektywne wyniki (mogliśmy zrealizować następne zadania, a sama struktura, którą nadaliśmy połączeniom w zadaniu 2 jest bardzo przejrzysta i klarowna. Widoczne są poszczególne zależności między tabelami oraz zachowana została logika w rozumowaniu.

Zadanie 3 potraktowaliśmy jako sprawdzian naszej wiedzy w kontekście wizualizacji i raportowania danych, a żaden podpunkt nie sprawił nam kłopotów. Jesteśmy usatysfakcjonowani z uzyskanych wyników.

Zadanie 4 podobnie jak zadanie 3 nie przysporzyło nam problemów, a uzyskane wyniki pozwalają na analizę danych w sposób określony w poleceniu.