

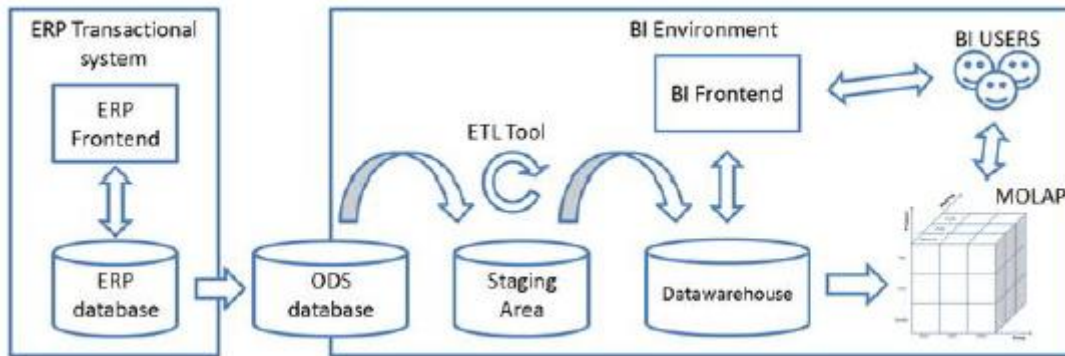
1. Business Intelligence dla każdego

W końcu biznes wymaga inteligencji. Taka jest prawda, zwłaszcza jeśli udajesz, że prowadzisz swój biznes poprawnie, ponieważ bez inteligencji nie da się prowadzić dochodowego biznesu. Mówiąc z perspektywy książki informatycznej Business Intelligence (BI) należy do świata Analytics. Business Intelligence to zestaw narzędzi i procesów, które pomagają podejmować decyzje w oparciu o dokładne dane, oszczędzając czas i wysiłek. Główną ideą narzędzia BI jest możliwość łatwej analizy danych w oparciu o koncepcje biznesowe bez posiadania wiedzy technicznej na temat narzędzi bazodanowych lub innych źródeł zawierających dane. Narzędzia BI udają, że wydobywają wiedzę z naszych przechowywanych danych w oparciu o trzy główne filary: niezawodność, dostępność i atrakcyjne doświadczenie użytkownika. Wyobraź sobie, że jesteś dyrektorem generalnym małej firmy zajmującej się produkcją ciastek i na podstawie raportów sprzedaży według analizowanych produktów stwierdzasz, że z każdym miesiącem sprzedaż ciasteczek Cream Chocolate spada. Liczby, które widzisz, stanowią obecnie około połowy kwoty, którą sprzedawano na początku roku. Jako dyrektor generalny masz różne możliwości: usunąć ciastko Cream Chocolate z katalogu; zmienić formułę Cream Chocolate; ustawić premię dla swojego działu handlowego, jeśli sprzedadzą ten produkt; lub zwolnij menedżera marki, ponieważ jej dział przynosi Ci straty, co znajduje odzwierciedlenie w analizie zysków i strat. Ale co się stanie, jeśli prawdziwym problemem jest to, że ten produkt zmienił swój wewnętrzny kod z powodu zmiany formuły kremu, a katalog produktów nie jest poprawnie zaktualizowany w Twoim systemie BI - i nie odzwierciedla prawidłowo sprzedaży z nowym kodem? Twoje poprzednie decyzje byłyby błędne, ponieważ opierasz je na błędnej analizie danych. Dlatego niezawodność jest podstawowym wymogiem we wszystkich projektach IT, ale jest szczególnie istotna w narzędziach BI, ponieważ można ich używać do podejmowania głównych decyzji, od strategicznego zarządzania firmą po podstawowe działania operacyjne. Na tej podstawie obowiązkowe jest, aby dane oferowane przez nasz system BI były spójne, a każdy wymiar analizy musi zapewniać prawidłowe wyniki w oparciu o jakość danych. Teraz wyobraź sobie, że pracujesz na linii montażowej samochodu; jest piątek o 20:00 i musisz złożyć zamówienie, aby uzupełnić swój magazyn różnymi elementami przed powrotem do domu. Uruchamiasz pulpit nawigacyjny magazynu, który sugeruje ilość każdej sztuki, o którą chcesz zapytać, a źródłem tej informacji są informacje ze środy po południu, ponieważ proces codziennego odświeżania jeszcze się nie zakończył. W przyszłym tygodniu zatrzymasz się na linii montażowej z powodu brakujących kół lub będziesz miał całkowicie zalany magazyn, ponieważ poprosiłeś o 100 zderzaków, które dotarły w zeszły czwartek. Podobny powód może spowodować ten sam rezultat, jeśli nie możesz uzyskać dostępu do systemu w wymaganym czasie z powodu pewnych prac konserwacyjnych na platformie i musisz oszacować zamówienie na podstawie tego, czego Twoim zdaniem brakuje. Nasz system musi być dostępny dla naszych użytkowników, kiedy nasi użytkownicy muszą z niego korzystać. Wydaje się to oczywistym warunkiem, ale istnieją dwa główne czynniki, które mogą spowodować, że nie uda nam się osiągnąć tego celu. Nasz system musi być stabilny, działać poprawnie w godzinach pracy, a dane muszą być aktualizowane odpowiednio do naszych docelowych konsumentów i ich wymagań. Ostatnią główną cechą systemu, który zamierzamy zbudować, jest to, że nasz dostęp do dostępnych danych musi być przyjazny dla użytkownika i dostosowany do oczekiwań i możliwości konsumentów. Nie możesz zapewnić eksperckiej analizy danych za pomocą narzędzia, które nie może wchodzić w interakcję z informacjami, a z drugiej strony możesz mieć kłopoty, jeśli Twój personel handlowy, który nie ma pojęcia o komputerach, będzie musiał opublikować zapytanie SQL, aby uzyskać dane, które wymagają analizy. Aby móc zapewnić narzędzie przyjazne dla użytkownika, musisz najpierw poznać swojego użytkownika i zgodzić się z jego wymaganiami w oparciu o jego potrzeby. Również dostarczone rozwiązanie musi być dostosowane do możliwości użytkownika.

Osiągnięcie tych trzech punktów nie jest łatwym zadaniem, ale są one podstawą do dostarczenia rentownego i trwałego rozwiązania BI wewnątrz Twojej firmy lub dla Twoich klientów.

Co to jest analiza biznesowa?

Oprócz krótkiego wprowadzenia, które omówiliśmy powyżej, i biorąc pod uwagę, że będziemy mówić głównie o Business Intelligence, chcielibyśmy bardziej szczegółowo przeanalizować, jak odpowiedzieć na te dwa proste pytania: Co oznacza BI? Jak właściwie zrozumieć koncepcję BI? BI ma wiele definicji w wielu publikacjach. Definicja BI w Wikipedii jest następująca: „Inteligencja biznesowa to zestaw teorii, metodologii, architektur i technologii, które przekształcają surowe dane w znaczące i przydatne informacje do celów biznesowych”. Naszym zdaniem jest to bardzo ciekawa definicja, ponieważ pokazuje pełny obraz rozwiązań BI, a nie tylko zwykłe skupienie się na narzędziach front-endowych, o których wspominają niektóre definicje. Ponieważ posiadanie rozwiązania BI oznacza podążanie za pewnymi teoriami w definiowaniu procesu, tak że określony model danych stosuje metodologie, które pomagają osiągnąć wydajność podczas projektu wdrożenia, a następnie późniejszej konserwacji, które określają poprawną architekturę, która zapewnia odpowiedni zwrot z inwestycji w oparciu o korzyści, jakie uzyskasz z projektu BI, a na koniec wybierz zestaw technologii, który spełni Twoje wymagania, specyfikacje i możliwości ekonomiczne. Na rysunku



możesz zobaczyć schemat głównych komponentów BI. Zapamiętaj to, ponieważ zrozumienie tego będzie przydatne. Na rysunku widać, że głównym źródłem informacji całego systemu jest ERP (pomimo tego, jak zobaczymy, może być wiele innych źródeł), wtedy mamy bazę danych ODS, która zawiera bezpośrednią ekstrakcję z ERP; może to być baza danych lub niektóre tabele w naszej bazie danych, ale ta koncepcja zwykle istnieje, wykorzystując bezpośrednie ekstrakcje z ERP, aby nie przeciążać systemu źródłowego. Za pomocą naszego narzędzia ETL przeniesiemy informacje z ODS do bazy danych Staging Area, gdzie będziemy je przetwarzać, a na koniec umieścimy je w hurtowni danych, do której uzyskamy dostęp za pomocą naszego narzędzia front-end BI. Jest całkiem możliwe, że mamy tylko bazę danych, a rozróżnienie między ODS, Staging Area i Datawarehouse to tylko używane przez nas tabele lub różne schematy w bazie danych. Wreszcie możemy mieć system MOLAP, który pomoże nam zrealizować budżet na kolejny rok. Zobaczymy szczegóły każdego elementu. Przeanalizujemy niektóre teorie podczas tego wprowadzenia. Istnieją również inne interesujące koncepcje definiujące BI. Jednym z nich jest skupienie się na korzyściach, jakie firma może odnieść z wdrożenia tego rodzaju systemu, o ile możesz być bardziej wydajny w zadaniach administracyjnych związanych ze zbieraniem informacji i wykorzystaniem swego czasu na analizę informacji w celu wyciągnięcia wniosków. Należy również zauważyć, że informacje, którymi zarządzamy, mogą pochodzić ze źródeł wewnętrznych i zewnętrznych, aby móc analizować, w jaki sposób wykonujemy nasze działania, ale także porównywać się z naszymi konkurentami, jeśli publikują informacje lub próbują analizować dane z naszych rynków docelowych do którego będziemy mieli dostęp. Inną ciekawą koncepcją jest możliwość przewidywania przyszłości. Nie mówimy o czarownicach ani wróżbitach; mówimy o znalezieniu prawidłowych wzorców, które

pozwolą nam przewidzieć, jaka może być nasza sprzedaż, jeśli warunki pozostaną takie same. Co może się stać, jeśli nowy konkurent wejdzie na nasz główny rynek i ukradnie nam 20% udziału w rynku? Lub co może być wynikiem zwiększenia o 25% naszego zespołu sprzedaży o 25%? W tej koncepcji kluczową cechą jest uzyskanie umiejętności wykrywania, które zmienne są ze sobą skorelowane, a które z nich są prawie niezależne. Interaktywność jest również jednym z elementów, które mogą dać ci wyobrażenie o tym, czym jest BI. To naprawdę interesujące, że analitycy biznesowi mogą badać i przeglądać dane, aby móc odkryć te ukryte wzorce, które mogą dać ci wgląd w najbliższą przyszłość. Kolejnym elementem, o którym warto wspomnieć przy różnych definicjach BI, jest wiedza, która jest wynikiem zastosowania technik BI do dużych ilości danych przechowywanych w naszych bazach danych lub uproszczenia formuły: jeśli łączysz Dane + Analiza, otrzymujesz Wiedzę.

Ewolucja BI

Koncepcja Business Intelligence, o której mowa tu, pojawiła się po raz pierwszy i została opisana przez Howarda Dresnera w 1989 roku. Opisał Business Intelligence jako „koncepcje i metody usprawniające podejmowanie decyzji biznesowych za pomocą systemów wsparcia opartych na faktach”. Od późnych lat 90. użycie tego terminu zostało uogólnione i można znaleźć niezliczone odniesienia do BI w książkach technicznych i artykułach internetowych. Pojawienie się BI jest bezpośrednio związane z konsolidacją systemów transakcyjnych na całym świecie. Zanim wszędzie zainstalowano usługi transakcyjne, narzędzia komputerowe były wykorzystywane głównie do analizy wysokiego poziomu; ilość informacji zapisywanych w systemach była na tyle mała, że można je było analizować bezpośrednio, bez potrzeby stosowania dodatkowego narzędzia. Gdy w scenariuszach biznesowych pojawiły się systemy transakcyjne, ilość danych do zarządzania wzrosła wykładniczo. Pomyśl o firmie detalicznej, która miała comiesięczne informacje o zakupionych jednostkach danego produktu i zapasie, który pozostał w sklepie; teraz ma informacje o każdym bilecie dowolnego klienta, ze szczegółowymi produktami, które kupili. Mogą uzyskać relacje między produktami; mogą analizować metody płatności; jeśli ich klienci płacą kartami, mogą uzyskać nazwę klientów i mogą przeanalizować, ile razy dany klient odwiedza nasz sklep, jaki rodzaj produktów kupuje i wiele innych analiz. A to tylko przykład; możesz przełożyć ten przykład na swoją firmę i zrozumieć, dlaczego potrzebujesz BI we własnym biznesie.

Uwaga: większość ostatnich odniesień dotyczy specjalistycznego oprogramowania obsługującego funkcje BI, a wiele firm konsultingowych ma dedykowane zespoły i projekty BI tylko po to, by spełniać wymagania programistyczne dotyczące narzędzi BI, traktując resztę rozwiązania jako pomocnicze komponenty samo narzędzie BI. Bądź więc ostrożny, jeśli myślisz o zatrudnieniu wsparcia doradczego, upewniając się, że ich oszacowanie kosztów projektu zawiera wszystkie wymagane elementy do Twojej prośby.

Od strategii do taktyki

BI doświadczyło również pewnych zmian w zakresie swoich projektów, rozprzestrzeniając się w różnych organizacjach, od raportów najwyższego kierownictwa i pulpitów nawigacyjnych po codzienne analizy operacyjne. Korzyści płynące z BI zostały udowodnione przez menedżerów najwyższego szczebla, którzy zauważyli, że BI oferuje ich organizacjom wiele możliwości czerpania zysków z narzędzi BI wdrażających projekty BI od dołu do góry ich firm. Na przestrzeni lat możemy zauważyć, że wdrożenia odeszły od wdrożeń strategicznych, które pomagają top managerom w decyzjach, które muszą podjąć, aby właściwie kierować firmami. Obejmuje to wszystkie środowiska wewnątrz organizacji, w tym te najniższe, aby ułatwić pracownikom podejmowanie decyzji, np. które produkty

muszę poprosić o uzupełnienie do magazynu lub jaki kolor najlepiej ubrać manekina w sklepie, w którym pracuję . BI przeszło od wspomagania decyzji strategicznych do taktycznych.

Uwaga: początkowe implementacje BI były znane pod różnymi akronimami, które ujawniają charakter docelowych użytkowników tych początkowych wdrożeń. Jednym z początkowych akronimów określających tego rodzaju systemy był DSS, czyli Decisional Support System, co pokazuje nam, że naszą grupą docelową będą decydenci. Z pewnością każda osoba w naszej firmie będzie podejmować decyzje, ale najważniejsze decyzje są zwykle podejmowane przez menedżerów, liderów i dyrektorów. Innym ciekawym akronimem, który wyraża to samo, jest EIS (Executive Information System), który w tym przypadku zawiera bezpośrednio nazwiska docelowych użytkowników platformy BI: kadry kierowniczej firmy.

Big Data

Obecnie najważniejszy nurt BI odnosi się do koncepcji Big Data i Data Lake. Sam Big Data opiera się na możliwości wykorzystania narzędzi BI i możliwości analitycznych do wydobywania informacji z niewiarygodnie ogromnej ilości danych generowanych każdego dnia przez naszych pracowników, klientów i użytkowników platform na wielu różnych platformach, takich jak portale społecznościowe, ogłoszenia o pracę w sieci, fora, blogi, aplikacje i zasoby mobilne, urządzenia mobilne, informacje GPS itp., które są zapisane w nieustrukturyzowanych systemach i których nie można zaatakować za pomocą standardowego uzbrojenia hurtowni danych; a wynika to z natury hurtowni danych, co przeanalizujemy w następujących sekcjach. DWH opiera się na ustrukturyzowanej bazie danych, która zawiera jednorodną informację ładowaną do naszego systemu wykorzystując procesy ETL, które zapewniają integralność danych, wielokrotne kontrole i walidacje; a tego rodzaju procesy są zbyt złożone, aby można je było odczytać ze źródeł Big Data, ponieważ moc obliczeniowa wymagana do przeprowadzenia tego rodzaju analizy jest zbyt wysoka. W systemie Big Data dokładność nie jest tak krytyczna jak w scenariuszu DWH. Wykorzystanie systemu Big Data do analizy logów Facebooka i pominięcie komentarza, który mógłby dostarczyć informacji od potencjalnego klienta wśród 1 miliarda użytkowników to coś, co możemy zaakceptować, ale jeśli pominiemy wiersz w swoim systemie księgowym, wygeneruje to niezgodność księgową. Big Data może służyć jako źródło zaopatrzenia naszego systemu BI: jest to dodatkowy komponent w warstwie infrastruktury i nie zastąpi naszej analizy sprzedaży, finansów ani operacji, którą możemy mieć na miejscu. Aby móc wspierać wymagania Big Data, wymagana jest nowa koncepcja inna niż DWH. W tym scenariuszu możemy zlokalizować Data Lake. Ideą tej koncepcji jest to, że nie musisz przetwarzać całej ilości dostępnych danych, aby stworzyć ustrukturyzowane źródło danych dla swojego systemu BI. Zamiast tego powinieneś uzyskać bezpośredni dostęp do swojego źródła danych, aby wyłowić potrzebne informacje (pomysł na jezioro pochodzi z wędkowania; pomysłowy, prawda?).

Internet przedmiotów

Można go traktować również jako źródło do analizy Big Data, jednak Internet of Things chciałbym omówić osobno ze względu na możliwości, jakie może on zaoferować projektom BI. Internet rzeczy jest związany z niewiarygodną ilością informacji, które można wydobyć z przychodzących urządzeń elektronicznych, które będą używane w wielu elementach na całym świecie. Teraz mamy samochody z dostępem do Internetu, lodówki, które mogą nam powiedzieć, czego brakuje w środku, maszyny sprzątające, które mogą wysłać nam SMS-a, gdy skończą, czy roboty sprzątające, które można zaprogramować do uruchamiania ze smartfona. Wyobraź sobie, ile informacji mogłoby to dostarczyć do analizy w środowiskach Big Data. Ten Internet Rzeczy zapewnia nieskończone możliwości badań i rozwoju, w których wiedza wydobyta z informacji dostarcza danych wejściowych, które będą niezwykle interesujące do analizy.

Charakterystyka BI

W ramach cech, które już skomentowaliśmy, istnieją dodatkowe i tematy związane z BI, na których chcielibyśmy skupić się bardziej szczegółowo, i które wymagają osobnych części.

Ukryte istotne informacje

Wykonując ponownie ćwiczenie wyobraźni, wyobraź sobie, że jesteś kierownikiem Sales Force w swoim kraju w swojej firmie i że masz pod swoją kontrolą pięciu przedstawicieli handlowych różnych narodowości: jeden z nich jest Chińczykiem, a drugi Włochem. Pozostali to rodowici mieszkańcy kraju. Mają przypisane regiony na podstawie liczby klientów, których muszą odwiedzić, oraz wielkości regionu w Twoim kraju. Robią regularne wizyty, aby odwiedzić wszystkich klientów z minimalną częstotliwością jednej wizyty na trzy miesiące. W tym scenariuszu przygotowujesz skonsolidowany raport dla wszystkich krajów i widzisz, że co trzy miesiące przynosi szczyt całkowitej sumy sprzedaży dla całego kraju. Wstępny wniosek, jaki możesz wyciągnąć, jest taki, że pod koniec kwartału Twój przedstawiciel handlowy sprzedaje więcej ze względu na jakiś cel, który sobie przyświeca lub że Twoi klienci mają kwartalną częstotliwość zakupów. Postanawiasz zmienić obiektywny proces na comiesięczny, aby zmotywować swoich pracowników do rozłożenia sprzedaży na cały rok, ale może to wyrzucić zbyt dużą presję i jeden z nich może zachorować. Zamiast wstępnej konkluzji i reakcji, aby potwierdzić, że podstawową przyczyną wahań sprzedaży jest okresowość, przechodzisz do analizy regionów i stwierdzasz, że są dwa regiony po pięć, które mają wyższy szczyt na wykresie, a pozostałe regiony są stabilne. Postanawiasz wyszukać przedstawiciela handlowego każdego regionu i widzisz, że są to regiony z pracownikami zagranicznymi. Porównując między nimi widać, że przez dwa miesiące w kwartale są poniżej średniej sprzedaży, ale w trzecim już nad resztą. Tak więc zmienność nie wynika tylko z okresowości, ale zależy również od pracownika. W tym momencie chcesz przeprowadzić głębszą analizę, aby zbadać, dlaczego są one bardziej zmienne w porównaniu z resztą. Następnie możesz przejrzeć miasta odwiedzane miesięcznie i docenić, że w trzecim miesiącu kwartału zawsze odwiedzają te same miasta na podstawie kwartalnej częstotliwości sprzedaży. Następnie możesz użyć analizy pochodzącej z Narodowego Instytutu Statystyki dotyczącej imigracji w poszczególnych miastach i skorelować z miastami o najwyższej sprzedaży dla tych dwóch pracowników, wykrywając, że głównym krajem pochodzenia imigracji w miastach, w których chiński pracownik ma lepsze wyniki sprzedaży, jest Chiny, a drugim krajem źródłowym w miastach, w których Włoch jest liderem sprzedaży, są Włochy. Z tymi wynikami mogłoby być interesujące przetestowanie reorganizacji przypisania klienta poprzez przypisanie, jeśli to możliwe, klienta do kogoś w Twojej organizacji z tej samej narodowości, ponieważ lepiej się rozumieją w negocjacjach, które prowadzą, zamiast robić to jak teraz przez geografii. Z tego przykładu można się dowiedzieć, że wiedza oprócz informacji może być ukryta wśród miliardów wierszy danych, a aby z tych danych wydobyć ważne informacje, trzeba wiedzieć nie tylko technicznie, jak się nimi bawić i umieć w pełni wykorzystać funkcjonalności Twojej platformy BI, ale także wymagana jest wiedza o tym, czego szukasz, w połączeniu z odrobiną intuicji.

Precyzja

Kiedy zarządzasz tysiącami, milionami lub miliardami wierszy w celu przeprowadzenia analizy, możesz pomyśleć, że jeśli stracisz kilka danych, nie wpłynie to na ogólną kwotę, ponieważ bardziej interesują Cię trendy i ewolucja danych niż posiadanie dokładnych danych do momentu trzecia liczba dziesiątna. Być może masz rację, ale bardzo ważne jest również dopasowanie danych do źródła informacji, ponieważ musisz oferować klientom niezawodność. Wyjaśnijmy przykład, aby pokazać znaczenie małych danych. W firmie produkującej produkty konsumenckie, takie jak środki czystości, masz zdefiniowany z klientami proces zarządzania zamówieniami zwrotu w przypadku produktu, który jest w złym stanie lub bezpośrednio uszkodzony. Może to spowodować małe zlecenia zwrotu kwot, które

wpływają do systemu transakcyjnego i są przekazywane do DWH. Możesz zignorować te polecenia w swoim DWH, ponieważ zarządzanie danymi głównymi jest zupełnie inne i przeprowadzając ogólną analizę, nie zauważyłbyś żadnej różnicy; pokazujesz sprzedaż brutto w wysokości 2 234 555,34 USD zamiast rzeczywistych 2 232 434,96 USD. Ale to 0,09% sprzedaży brutto, które zostało pominięte, wiąże się z wysokimi kosztami zarządzania zamówieniami zwrotów, jeśli nie weźmiesz ich pod uwagę, możesz przegapić interesujące dane dla Twojej organizacji, które mogą przełożyć się na Twoje przychody netto z dużą marżą, ponieważ niektóre zwroty od ważnych klientów wiążą się z pewnymi karami, które są dodawane do kosztów zarządzania.

Istotne KPI

Kluczowy wskaźnik wydajności (KPI) jest powiązany z miarą danych, które są przydatne do mierzenia ewolucji Twojej firmy. Określenie, które KPI musisz monitorować, aby mieć pewność, że nie przegapię żadnej istotnej miary mojej firmy, jest jednym z krytycznych kroków we wdrażaniu projektu BI. Podczas ich definiowania należy wziąć pod uwagę wiele kwestii, ale najważniejszą z nich jest trafność metryki. Oczywiście trafność miernika będzie zależała od grupy docelowej; jeśli definiujesz je ze swoim działem sprzedaży, najbardziej odpowiednie metryki będą dotyczyły sprzedaży, sprzedanych jednostek lub wolumenu, głównie sprzedaży brutto i sprzedaży netto, ponieważ jest całkiem możliwe, że te metryki są bezpośrednio powiązane do Twoich wyników i celów sprzedażowych, ale jeśli jesteś z działu finansowego, będziesz bardziej zainteresowany kosztami i przychodami netto. Jeśli pracujesz w dziale obsługi klienta, interesują Cię zwroty ze sprzedaży, jakość sprzedawanych towarów itp. Ale to nie jedyny ważny parametr, gdy chcesz mieć pokaźny zestaw KPI. W przypadku działu sprzedaży wyobraź sobie, że zostałeś przeniesiony od jednego klienta do drugiego. Twój stary klient kupował 3 miliony dolarów rocznie, a nowy kupuje 8 milionów dolarów. Czy ten numer jest dobry czy zły? Czy ta liczba sama w sobie może powiedzieć, jaka jest wydajność na tego klienta? Na pewno nie. Najlepsze analizy BI, gdy mierzysz swoją ewolucję, opierają się na wartościach procentowych; trzeba to z czymś porównać. Jak te dane mają się do sprzedaży z zeszłego roku? Jak porównuje się te dane z moimi celami rocznymi? Takim rodzajem analizy interesuje się większość odbiorców informacji; chcą otrzymywać informacje, że można łatwo zinterpretować.

Na czas

Ile razy słyszałeś typowe zdanie typu „Proszę dostarczyć raport sprzedaży na mój pulpit do jutra rano przed godziną 09:00” lub „Ta informacja powinna być zostać dostarczona naszemu kierownikowi wczoraj”? Wspominaliśmy już o tym, jak ważne jest dostarczanie informacji we właściwym czasie, kiedy jest to wymagane, ale chcielibyśmy rozwinąć ten pomysł. Aby informacje były dostarczane na czas, musimy najpierw dowiedzieć się, co oznacza dla naszych klientów terminowość dla każdego zestawu informacji. W definicji zakresu projektu, oś tymczasowa jest jednym z głównych tematów do oceny w celu sprawdzenia, czy zakres projektu jest realistyczny, czy nie. Należy wziąć pod uwagę różne elementy, które mogą mieć na to wpływ, takie jak wolumetria do przeniesienia, pojemność sprzętu, złożoność procesu i zależności w zakresie ładowania informacji. W tej definicji istotne jest również rozważenie przyszłej ewolucji wszystkich tych tematów: w jaki sposób można zwiększyć wolumetrię do zarządzania i wolumetrię docelową (w zależności od strategii procesu może to zależeć nie tylko od informacji do załadowania do bazy danych, ale także od rozmiaru tabel docelowych, zwłaszcza jeśli używasz strategii aktualizacji lub scalania); jak może zwiększyć wydajność sprzętu, jeśli mamy skalowalny system lub jeśli mamy najnowocześniejszy sprzęt i nie możemy się rozwijać bez przebudowy pełnej platformy z nowymi serwerami, aby kontrolować wzrost złożoności procesów i sprawdzać, czy wcześniejsze procesy nie będą cierpieć z powodu opóźnień, które mogłyby wpłynąć na nasze procesy ETL. Wszystkie te rozważania są szczególnie istotne, jeśli mówimy o procesach dziennych lub śróddziennych, ale nie są tak istotne, gdy harmonogram ETL jest tygodniowy lub miesięczny

Analiza firmy : Cykl życia i ciągłość

Poprawa

Cóż, jesteśmy w szóstym miesiącu po uruchomieniu naszego projektu BI i co miesiąc analizujemy, które są bestsellerami w naszym katalogu produktów, a które sprzedają się najgorzej. Widzimy, że podobnie jak w zeszłym miesiącu dany produkt pojawia się na dole listy i dzieje się to miesiąc po miesiącu dla tego samego produktu, ale naszym zdefiniowanym działaniem jest podążanie za ewolucją. W przyszłym miesiącu jest całkiem możliwe, że ten sam produkt pojawi się na dole listy sprzedaży. Czego brakuje w naszym systemie BI? Wydaje się jasne … działanie. Musimy działać zgodnie z wynikami, które uzyskaliśmy z naszej analizy, a następnie sprawdzić wyniki naszych działań. Znajdujemy się w cyklu życia analityki firmy, gdzie nasz system BI może pomóc nam zrozumieć cały proces.

* **Analizuj dane:** Kiedy mówimy o analizie danych, nie możemy poprzestać na wstępnej analizie. W poprzednim przykładzie nie możemy poprzestać na określeniu, które produkty mają najgorszy wskaźnik sprzedaży, musimy zbadać, co jest tego główną przyczyną lub czy istnieje więcej niż jeden czynnik, który może wpłynąć na wyniki, które zaobserwowaliśmy. Musimy spróbować skorelować informacje z inną analizą, spróbować zobaczyć inne wizualizacje informacji, przeanalizować informacje w różnych wymiarach, próbując wyizolować, na czym polega problem tego produktu. Być może jest to kwestia złego zaprojektowania produktu, wykorzystania niewłaściwego kanału sprzedaży, zastosowanej reklamy nie jest wystarczająco dobra lub jest skierowana do niewłaściwego targetu; może przedstawiciel handlowy tego produktu ma złe wyniki dla wszystkich produktów, które ma w portfolio lub jest to produkt tymczasowy, który staramy się wyprzedać w jego głównym sezonie – próbujemy sprzedawać lody zimą w naszym lokalnym biurze na biegunie północnym.

* **Zdefiniuj plan działania:** Po określeniu, co jest podstawową przyczyną wykrytego problemu lub co jest naszą mocną stroną w naszym bestsellerze, musimy zdefiniować zestaw działań, które pomogą nam poprawić/ulepszyć to, co jest nie tak. Jeśli zobaczymy, że mamy w katalogu naprawdę zły produkt, zdecydujemy się usunąć go z katalogu; jeśli region nie będzie właściwy, zdecydujemy się przenieść przedstawiciela handlowego do lepszego regionu; jeśli sprzedajemy lody na biegunie północnym, lepiej zamknąć naszą firmę. Nasz plan działania musi zawierać przede wszystkim działania, odpowiedzialność i terminy.

* **Działaj:** Plan działania nie ma żadnej użyteczności, jeśli go nie przestrzegasz. Musimy więc postępować zgodnie z nim w ramach zaplanowanego kalendarza.

* **Sprawdź wyniki:** Musimy zmierzyć, w jaki sposób zmieniamy KPI, które oceniamy, aby zobaczyć, czy nasze działania przynoszą pożądane rezultaty: czy powodują odwrotny skutek niż oczekiwano lub czy moje działania nie mają żadnego wpływu na wynik.

Uwaga: Głównym wynikiem dobrej analizy jest dobry plan działania. Musimy spróbować dostarczyć narzędzie BI, które pozwoli Twoim analitykom łatwo opracować plan działania i zmierzyć wyniki naszych działań w celu wdrożenia planu działania.

Korzyści z BI

Korzystanie z BI może przynieść Twojej firmie wiele korzyści. Jeśli Twoje narzędzie BI zostało poprawnie wdrożone i osiągnęłaś poprawną wydajność w swoim systemie BI, uzyskasz korzyści, które klasyfikujemy jako bezpośrednie i pośrednie.

Korzyści bezpośrednie

Za bezpośrednie korzyści uważamy wszystkie te korzyści, które daje korzystanie z narzędzia BI, które są bezpośrednio związane z wdrożeniem BI. Z pewnością uzyskasz te korzyści, jeśli będziesz w stanie wdrożyć rozwiązanie BI z dobrą strategią i wydajnością. Jednym ze sposobów, aby to umożliwić, jest przestrzeganie zaleceń, które staramy się przekazać w tej książce. Jeśli zrobisz to w ten sposób, możesz uzyskać:

Optymalizacja zasobów: jeśli próbujesz wdrożyć rozwiązanie BI, jest całkiem możliwe, że obecnie Twój system analityczny jest czymś w rodzaju ponownej kompilacji różnych zespołów w Excelu, skonsolidowanej w jednym skoroszycie, z którego wyodrębniasz różne wykresy, tabele przestawne i wysokie -stosunki poziomów. Istnieje możliwość, że bardzo często znajdziesz jakiś błąd w formacie źródłowych arkuszy Excela i będziesz musiał zmodyfikować dziesiątki stron, ponieważ musisz dodać nową kolumnę w swojej analizie lub popełniłeś błąd w agregacji, co wymaga ponownego wykonania całej ekstrakcji. Rozwiązanie BI powinno bezpośrednio dostarczać informacje w wymaganym formacie; powinieneś być w stanie zapisywać raporty w formacie, którego potrzebujesz, aby otrzymywać najlepsze produkty lub porównywać z ostatnim okresem, i powinieneś być w stanie zaplanować zadanie, aby otrzymywać te raporty codziennie, co tydzień lub co miesiąc.<br

Oszczędność kosztów: wynikająca z wcześniej skomentowanej optymalizacji zasobów, uzyskasz oszczędności, szczególnie w zasobach ludzkich, które są inwestowane w bieżący proces analizy ręcznej. Projekt BI pozwoli Ci również zaoszczędzić wiele innych kosztów dzięki wdrożeniu raportowania analizy kosztów w Twoim rozwiązaniu BI, które zapewni Ci wiele punktów działania w celu optymalizacji procesów i operacji.

Uwaga: Spróbuj zmierzyć ilość czasu, jaki Twoja firma inwestuje w gromadzenie, agregowanie i formatowanie informacji, które pozwalają poznać wynik bieżącego ćwiczenia, okres zamknięcia miesiąca lub dzienną sprzedaż; zsumuj cały ten czas, mnożąc koszt jednostkowy zasobów ludzkich, a następnie porównaj go z całkowitym kosztem projektu BI. W ten sposób będziesz mógł przeanalizować bezpośredni zwrot z inwestycji (ROI), jaki uzyskasz z wdrożenia projektu BI.

SVOT: akronim od Single Version of Truth; jedną z korzyści płynących z posiadania systemu referencyjnego, który zawiera wszystkie dane wymagane do analizy, jest to, że wszyscy będą korzystać z tego samego systemu, dzięki czemu wszystkie działy będą uzyskiwać te same dane, zamiast wielu arkuszy Excel z ręcznymi ekstrakcjami, ręcznymi modyfikacjami i personalnymi rozważaniami. Wszystkie działy uzgodnią, jaką łączną wielkość sprzedaży osiągniesz w każdym miesiącu lub jaki był ubiegłoroczny przychód netto.

Jedna osoba odpowiedzialna za informacje: W ramach wdrażania projektu BI zdefiniujesz osobę odpowiedzialną za BI lub dział, który będzie Twoim pojedynczym punktem kontaktowym w zakresie wszystkich informacji. Nie będziesz już zależny od informacji napływających z różnych kanałów, plików Excela wypełnianych i formatowanych przez wielu pracowników, działów czy wymiany e-maili pomiędzy wszystkimi działami kontrolingu sprzedaży. Dział IT będzie centralizował proces dostarczania informacji.

Analiza samoobsługowa: jeśli jesteś informatykiem w małej firmie i odpowiadasz za wyodrębnianie i formatowanie danych dla użytkowników biznesowych, jest to jedna z najbardziej interesujących korzyści, jakie może zapewnić wdrożenie rozwiązania BI. Będą mogli samodzielnie generować raporty. Stworzysz infrastrukturę, która pozwoli im przeprowadzać wszelkie analizy, o których mogą pomyśleć, pod warunkiem posiadania dostępnych informacji, których potrzebują. Ale będą w stanie samodzielnie formatować dane, drażnić dostępne informacje, obracać dane, segmentować dane, filtrować informacje, wprowadzać dowolne modyfikacje, na które chcesz im zezwolić.

Możliwości szczegółowej analizy: Jeśli jesteś w stanie wdrożyć solidny model danych z integralnością danych, zweryfikowanymi tabelami wyszukiwania i pełnymi wymiarami, otrzymasz solidny system, który pozwoli ci analizować informacje na najbardziej szczegółowym poziomie, jaki możesz sobie wyobrazić. Z pewnością, aby móc to zrobić, musisz użyć poziomu sprzętu, który odpowiada wymaganiom dotyczącym oczekiwanego czasu odpowiedzi, ilości danych i równoległości użytkownika. Ale jeśli sprzęt jest wystarczająco wydajny, będziesz mieć możliwość dostarczenia systemu analitycznego, który zapewni zagregowaną analizę na wysokim poziomie najbardziej szczegółowych dostępnych informacji, zawsze wykorzystując funkcje BI do filtrowania ilości informacji, które system zwraca użytkownikowi.

Korzyści pośrednie

Wywodzące się z poprzedniej grupy korzyści, ale także wynikające z konfiguracji samego rozwiązania BI, będziesz mieć możliwość uzyskania innych korzyści niematerialnych, które nie są bezpośrednie, ale które można uznać za konsekwencję Twojego projektu BI. Projekt BI da Ci narzędzia do podejmowania właściwych decyzji i działania zgodnie z tymi decyzjami; jeśli to zrobisz, będziesz w stanie osiągnąć zestaw korzyści. Na poniższej liście pokazujemy tylko kilka przykładów, które uważamy za ogólnie istotne, ale może to być dla Ciebie podstawa do określenia, na czym skupiają się Twoje działania wynikające z analizy BI. Wzrost sprzedaży: Będziesz w stanie przeanalizować, którzy klienci kupują Twój produkt i przeanalizować, czy znajdziesz wśród nich wspólne wzorce, aby skoncentrować swoje strategie marketingowe na tych klientach, którzy spełniają te wzorce. Dzięki temu będziesz mieć możliwość analizy swoich produktów wiele cech produktu, które mogą dać kombinację atrybutów maksymalizujących lub minimalizujących wyniki sprzedaży, a zobaczysz, które promocje sprzedaży są bardziej skuteczne, porównując je ze sprzedażą w pozostałej części roku. Podsumowując, będziesz mieć w zasięgu ręki narzędzia, które pozwolą Ci w prosty sposób przeprowadzić złożone analizy

Redukcja kosztów: Opierając się na projekcie BI, możesz zredukować koszty swojej firmy w oparciu o wiele perspektyw, o ile posiadanie potężnego narzędzia analitycznego pomoże ci usprawnić proces kontroli kosztów. Gdy opracujesz projekt BI z wymaganymi informacjami, będziesz mógł przeanalizować wszystkie dostępne koszty w Twojej organizacji, koszty operacyjne, zasoby ludzkie, wynajem i leasingu, koszty finansowe, wydatki pracownicze, koszty wydziałowe itp.

Powiernictwo klientów : Możesz analizować reakcje klientów na kampanię marketingową. Możesz również analizować wykorzystanie kart lojalnościowych, będziesz w stanie zweryfikować ewolucję pozyskiwania klientów, a BI pozwoli Ci połączyć informacje pochodzące od Twojego operatora transakcyjnego i logistycznego, aby zweryfikować, czy dostarczają Twoje produkty na czas, wśród wielu

inna analiza, o której możesz pomyśleć.

Powinowactwo produktów: jeśli skoncentrujesz się na wymiarze produktu, będziesz w stanie przeanalizować, które produkty są ze sobą powiązane na podstawie preferencji klienta. Będziesz mógł analizować, które pary produktów kupują Twoi klienci, jakie relacje nie są oczywiste, stosując podstawowe logiczne myślenie. Można sobie wyobrazić, że jeśli klient kupuje monitor komputerowy, to całkiem możliwe, że interesuje go mysz i klawiatura, ale może trudniej jest określić, które tytuły książek lub filmy mogą mieć związek, jeśli nie jest to bezpośrednio związane z kategorią, autorem lub reżyserem.

Segmentacja klientów: Segmentacja umożliwi grupowanie klientów na podstawie ich cech. Istnieją pewne cechy bezpośrednio z nimi związane, które zwykle pochodzą z danych podstawowych klienta w Twoim systemie transakcyjnym lub CRM (moduł relacji z klientami), takie jak wiek, miasto, adres, dochód netto, liczba dzieci lub wszelkie preferencje, które możesz uzyskać od dowolnego narzędzia do

zbierania informacji. Ale możesz też zastosować segmentację opartą na metrykach, takich jak liczba nabytych produktów, jakie są jego preferowane kategorie, jaka jest jego metoda płatności, jakie są jego preferowane kanały dystrybucji, pochodząca od niego sprzedaż brutto, okres zakupu lub dowolne inne pole który wymaga wcześniejszego obliczenia wyprowadzonego z podstawowego modelu, który wdrażasz.

Analiza demograficzna : W związku z segmentacją klientów, szczególnie istotna jest analiza demografii interesującego Cię regionu. Możesz zacząć od informacji pochodzących z wewnętrznych źródeł zebranych za pomocą jakiegoś narzędzia CRM związanego z Twoimi klientami lub możesz przeanalizować informacje dostarczane przez zewnętrzne źródła, takie jak publiczne instytucje statystyczne, które publikują ogólne informacje o krajach, regionach, miastach lub okręgach. Możesz połączyć analizę demograficzną z wcześniejszą segmentacją klientów i skupić się na tych dzielnicach, w których okolica może być bardziej zainteresowana Twoim produktem. Również na podstawie demografii możesz zdecydować, gdzie otworzyć sklep lub gdzie zlokalizować punkt informacji dla klientów.

Analiza geograficzna: Można to uznać za segmentację klientów lub analizę demograficzną, ale wolimy zachować ją jako oddzielną korzyść ze względu na nowe potężne wizualizacje map, które pozwalają zlokalizować na mapie liczbę klientów, sprzedaż lub inne dane które chcesz analizować.

Analiza procesu produkcyjnego: Analizując proces produkcyjny, linię montażową lub łańcuch dystrybucji, możesz uniknąć nadprodukcji i nadwyżki zapasów, stosując techniki just-in-time za pomocą rozwiązania BI. Możesz także przeanalizować stracony czas, czas oczekiwania, wąskie gardła, ilość zasobów zużywanych przez wszystkie etapy procesu produkcyjnego lub jakie zależności można zoptymalizować/usunąć z przepływu pracy produkcyjnej. **Analiza jakości:** Możesz także skoncentrować swoją analizę na jakości produktów lub usług, które oferujesz, analizując liczbę incydentów według typologii, źródło incydentów według etapów linii produkcyjnej, analizuj opinie klientów lub sprawdzaj terminowość rozwiązywania i dostarczania projektów.

Analiza produktywności pracowników: Twoja firma będzie łatwiej definiować cele i analizować ich realizację, korzystając z funkcji BI, aby uzyskać dane dotyczące zrealizowanych wizyt przedstawicieli handlowych, efektywności tych wizyt, liczby zamówień na pracownika, wartości sprzedaży brutto, liczby wyprodukowanych sztuk , liczba spowodowanych usterek na pracownika, zużyte zasoby na pracownika, zrealizowane szkolenia czy dni robocze.

Cele

Możesz połączyć i połączyć wszystkie te korzyści i analizy wskaźników KPI, aby zdefiniować i realizować różne cele strategiczne dla swojej organizacji, definiując w ten sposób, która strategia poprowadzi Twoją firmę do sukcesu. Poniżej znajdziesz kilka typowych celów, ale tak jak na liście korzyści, jest to tylko zestaw przykładów; powinieneś być w stanie zdefiniować własne cele w oparciu o misję firmy.

Zoptymalizuj wydajność firmy: Zwiększając sprzedaż, zmniejszając koszty lub łącząc oba te czynniki, możesz zmaksymalizować wydajność swojej firmy. Ostatecznie dla wyników naprawdę ważny jest dochód netto, więc wszelkie wskaźniki uwzględnione w obliczeniach dochodu netto mogą rosnąć lub maleć, aby zmaksymalizować wydajność. Aby działać na podstawie któregośkolwiek z tych wskaźników, możesz użyć swojego rozwiązania BI, które pomoże Ci wykryć sposób, w jaki to robisz i śledzić Twoje działania w oparciu o dowolne konkretne wskaźniki.

Wykrywanie rynków niszowych: jest to cel, który można zdefiniować w szczególności na podstawie analizy danych zewnętrznych, takich jak dane demograficzne i analizy konkurencji. Dzięki połączeniu

analizy zewnętrznej z wiedzą o tym, co Twoja firma może zaoferować rynkowi, możesz znaleźć nisze rynkowe dla konkretnego produktu, który znajduje się w Twoim katalogu lub który można opracować dla tego rynku.

Popraw zadowolenie klientów : Wydaje się dość oczywiste, że monitorując KPI, które pozwalają mierzyć jakość usług oferowanych klientom, będziesz w stanie je poprawić. Również Twoje narzędzie BI może służyć do raportowania klientom, w jaki sposób wykonujesz tę jakość usług lub inne raporty, które mogą być dla nich przydatne. Może również poprawić relacje z twoimi dostawcami, o ile możesz również dostarczać dostawcy informacji przydatnych do rozwoju jego produktów i marketingu. Przykładem tego celu może być producent odzieży, który sprzedaje swoje produkty do różnych sieci hipermarketów; jeśli hipermarket wysyła informacje o stanie sprzedaży produktów producenta, producent może ocenić, czy projektuje zgodnie z fajną linią, czy też jego produkty zostaną zwrócone, ponieważ hipermarket nie jest w stanie sprzedać żadnych sukienek.

Popraw zadowolenie pracowników: niektóre z przedstawionych wskaźników KPI i korzyści dotyczą kontroli pracowników, na przykład produktywności, sprzedaży na pracownika lub kontroli kosztów, co może powodować u pracownika poczucie, że jest on wysoce kontrolowany i monitorowany, co nie zawsze jest przyjemne dla nich. Ale z drugiej strony możesz zaoferować wiele innych korzyści płynących z BI; poprawa produktywności może dać pracownikowi czas do poświęcenia się ciekawszymi zadaniami.

Skrócenie czasu poświęcanego na powtarzalne zadania, takie jak zbieranie i formatowanie informacji, pozwoli pracownikom Twojej firmy skupić się na analizie danych; wszystkie opcje oferowane przez platformę BI ułatwią im pracę; a także mogą dowiedzieć się o możliwościach narzędzi BI, temat, który poprawi ich programy nauczania.

Kto może skorzystać z BI?

Na podstawie wszystkich korzyści, które analizowaliśmy w poprzedniej sekcji, wydaje się oczywiste, że prawie każda firma może w różnym stopniu skorzystać z BI, od podstawowego operatora, który może podejmować decyzje w oparciu o przyjazny interfejs użytkownika, po wypchnięcie zdefiniowanej przycisk do dyrektora generalnego, który może zdecydować, na którym rynku inwestować w ciągu najbliższych czterech lat; w końcu każda osoba w naszej organizacji może w pewnym momencie zostać poproszona o podjęcie decyzji i może użyć BI, aby pomóc mu zdecydować, które opcje wybrać. W każdym razie podajmy kilka przykładów według działów:

* Ogólne kierownictwo: To był pierwotny zespół docelowy dla narzędzi BI, więc są oni członkami organizacji, którzy mogą czerpać większe korzyści z narzędzi BI. Dobre narzędzie BI musi być nastawione na łatwe pokazywanie istotnych informacji, na które należy zwrócić uwagę, szybkie skupienie się na celu oraz wyciąganie wniosków i podejmowanie decyzji w krótkim czasie. Podsumowując sposób na życie zespołu wykonawczego.

* Dział sprzedaży: Sprzedaż to najważniejsze zadanie w każdej firmie. Możesz wyprodukować najlepszy produkt w swojej kategorii, ale jeśli go nie sprzedasz, Twoja firma zbankrutuje. Tak więc historycznie te zespoły były drugim krokiem we wdrażaniu BI. Analizuj trendy sprzedaży, wykorzystuj BI jako wsparcie do definiowania celów sprzedażowych, a także do śledzenia, jaki jest status na rzeczywistych vs obiektywnych KPI to główne analizy, na których się skupisz.

* Zespół finansowy: Sprzedaż prawie nieskończona, ale z kosztem jednostkowym wyższym niż cena, spowodowałaby nieskończone straty. Dlatego kolejnym krokiem we wdrożeniach BI jest zwykle analiza finansowa firmy, aby skupić się na poprawie wydajności firmy.

* Dział zakupów: wywodzący się z analizy finansowej i kosztowej, a zwłaszcza dla firm produkcyjnych lub odsprzedających, jest również dość istotny w celu obniżenia kosztów nabycia surowców i towarów. Aby pomóc Ci w tym trudnym zadaniu, możesz również skorzystać z funkcji BI, ponieważ zawsze potrzebujesz go do analizy informacji w celu wyciągnięcia wniosków.

* Dział kadr: Inne ważne koszty wynikające z analizy kosztów finansowych to wynagrodzenia, diety, wydatki i inne koszty osobowe naszych pracowników. Gdy dział finansowy naciska na obniżenie kosztów zasobów ludzkich, wsparcie narzędzia BI może pomóc w analizie danych pracowników.

* Operatorzy: Każda osoba w dowolnym zespole operacyjnym w firmie może mieć możliwość podjęcia decyzji: od prośby o uzupełnienie zapasów, po zaoferowanie klientowi produktu pasującego do reszty przedmiotów, które ma w koszyku. Ponownie, posiadanie dobrego rozwiązania BI może ułatwić te zadania każdemu, kto bierze udział w procesie decyzyjnym.

Komponenty Platformy BI

W tej sekcji przeanalizujemy, jakie są główne komponenty całej platformy BI, co posłuży jako teoretyczne wprowadzenie do kolejnych rozdziałów, w których dogłębnie dowiesz się, jak zainstalować i wdrożyć rozwiązanie BI oparte na głównym oprogramowaniu open source narzędzia.

Źródło ERP: Najczęstszym źródłem informacji w rozwiązaniu BI i zwykle pierwszym, które należy wykorzystać, jest Planowanie Zasobów Przedsiębiorstwa. Tego rodzaju narzędzie transakcyjne służy do kontrolowania i zapisywania głównych operacji zachodzących wewnątrz firmy. Ten komponent sam w sobie nie może być uważany za komponent BI; ale jak na razie jest źródło projektu, więc o tym wspomnieliśmy. W zależności od włączonych funkcji ERP w Twojej firmie będziemy w stanie wydobyć główne informacje o sprzedaży, finansach, operacjach, zasobach ludzkich czy zapasach. Aby wydobyć informacje mamy dwie możliwości: dostęp przez jakiś interfejs API lub dostęp bezpośrednio do bazy danych zawierającej informacje ERP. Nasze preferowane podejście to pierwsze, jeśli narzędzie na to pozwala; w ten sposób nie będziesz tak zależny od zmiany struktury tabel w bazie danych. Interfejs komunikacyjny, który umożliwia ekstrakcję danych, prawdopodobnie będzie trzymał się ewolucji ERP, podczas gdy nikt nie zapewni, że struktura tabel pozostanie taka sama. Z drugiej strony struktura tabeli może być poufna, więc musisz zbadać, które tabele i kolumny zawierają informacje, które chcesz wyodrębnić podczas korzystania z interfejsu API, o których będziesz mieć informacje z funkcjami interfejsu i parametrami do wyodrębnienia wymaganych danych. Głównym problemem, jaki można znaleźć w użyciu API, jest wydajność, jaką oferuje ten interfejs, ponieważ wydajność ERP jest zwykle zoptymalizowana pod kątem operacji transakcyjnych o niskim wolumenie, ale uruchomienie ekstrakcji ETL może zarządzać tysiącami wierszy w jednej ekstrakcji. Będziesz musiał ocenić możliwości i wydajność obu metod ekstrakcji, aby móc zdecydować, która opcja jest najlepsza.

Baza danych

Do przechowywania danych. To jest główny cel Bazy Danych. To może być wystarczające wprowadzenie do baz danych, ale przejdźmy do rozwinięcia tego pomysłu nieco bardziej. Zapewne wiesz, czym jest baza danych, więc nie ma sensu zaczynać tutaj od definicji z Wikipedii. Skupmy się na wykorzystaniu bazy danych w BI rozwiązaniu. Baza danych to ogólna nazwa, której używamy w odniesieniu do technologii przechowywania danych, ale w przypadku środowiska BI zwykle mówimy o niej jako Datawarehouse. Z pewnością jest to podstawowy składnik tej architektury; bez bazy danych nie byłoby nic. Zawsze możesz zrobić własny ETL z procedurami ładowania, własnym rozwiązaniem do raportowania z jakąś aplikacją internetową pobierającą dane z bazy danych; ale bez bazy danych nie byłbyś w stanie nic zrobić. Głównym celem bazy danych będzie zawieranie informacji, które będą dostępne z narzędzia BI, ale także będzie zawierała informacje pomocnicze do ładowania danych, które

w zależności od narzędzi, których używamy do ETL i front-endu, będą wymagały od Ciebie zapisywanie obiektów wewnętrznych do bazy danych, która może być taka sama jak ta zawierająca dane główne lub inna. Podczas definiowania opcji konfiguracji bazy danych należy wziąć pod uwagę kilka kwestii:

- * Istnieją parametry do optymalizacji dla hurtowni danych, które muszą być ustawione w inny sposób niż transakcyjna baza danych.

- * Musisz określić, jaka jest wymagana dostępność dla użytkowników.

- * Będziesz musiał określić, jaka jest częstotliwość ładowania i które okno ładowania jest dostępne do odświeżenia informacji.

- * Będziesz musiał narysować definicję środowiska; możesz mieć pośredniczącą bazę danych pomiędzy transakcyjną a hurtownią danych, która może służyć jako ODS (Operational Data Storage).

- * Będziesz musiał zdefiniować zasady tworzenia kopii zapasowych zgodnie z częstotliwością ładowania. Nie ma sensu zapisywać codziennych kopii zapasowych, jeśli dane zmieniają się raz w miesiącu.

ETL: Komponent ETL, będący akronimem słów Extraction, Transformation i Load, ma dość opisową nazwę. Jak możesz sobie wyobrazić, jego główne funkcje to Extract, Transform i Load. Co to znaczy? Nie będziesz używać swoich tabel transakcyjnych do opracowywania analizy modelu za pomocą narzędzia BI; wykorzystasz w tym celu swoją hurtownię danych. Będziesz więc musiał wyodrębnić informacje ze źródłowego systemu ERP i załadować je do hurtowni danych. W połowie tego procesu będziesz musiał dostosować informacje do wymaganej struktury hurtowni danych, tworząc tabele faktów z pożądanymi polami, tabele relacji, które muszą respektować jedną do wielu relacji między wymiarami, zapewniając, że wszystkie możliwe wartości łączenia pola są obecne w tabeli i że pola, które łączą się z resztą tabel, nie mają wartości pustych lub że tabele przeglądowe zawierają wszystkie możliwe wartości faktów i wartości tabel relacji. W ramach tych podstawowych operacji jest całkiem możliwe, że masz przekształcenia, takie jak rachunek zagregowanych tabel, proces robienia codziennego zdjęcia jakiejś tabeli, aby zachować historię jej ewolucji, obciążenie tabel bezpieczeństwa, które pozwalają na zabezpieczenie na poziomie wiersza dostępu lub wszelkie inne przekształcenia, których możesz potrzebować w procesie uzupełniania hurtowni danych. W rozdziale 5 tej książki przeanalizujemy ten proces i niektóre narzędzia open source.

Narzędzie front-end: czasami traktowane jako samo rozwiązanie BI, będziesz potrzebować rozwiązania front-end, które umożliwi użytkownikom interakcję z danymi zapisanymi w hurtowni danych. To narzędzie BI będzie głównym kanałem komunikacji między użytkownikami a danymi, dlatego zdecydowanie zaleca się, aby oba połączenia działały poprawnie. Twoi użytkownicy będą wymagać wystarczającej wiedzy na temat narzędzia BI, a narzędzie BI musi być przyjazne dla użytkownika. Z drugiej strony narzędzie BI musi być w pełni kompatybilne do pracy z Twoją bazą danych; musisz zapewnić certyfikowaną interoperacyjność między obydwojema komponentami zgodnie ze specyfikacjami dostawcy. W dalszej części tego rozdziału przeanalizujemy, jakie są główne możliwości narzędzia BI, a także w rozdziale 8 tej książki zobaczymy kilka rozwiązań open source, które można wykorzystać do rozpoczęcia projektu BI.

Narzędzie do budżetowania: Na podstawie decyzji, które podejmiesz z wykorzystaniem danych z narzędzia front-end, zdefiniujesz działania, które powinny znaleźć odzwierciedlenie w budżecie na nadchodzące okresy. Czasami tego rodzaju narzędzia nie są brane pod uwagę jako część BI, ale uważamy, że są one ważnym narzędziem, które finalizuje cykl życia analityki firmy i zawiera również funkcje analizy danych, dlatego zdecydowaliśmy się uwzględnić je w naszej definicji platformy jako komponent BI. Jako narzędzie do budżetowania możesz po prostu mieć arkusz zawierający cel na

przyszły rok, ale my przeanalizujemy bardziej zaawansowane narzędzia, które zapewnią Ci dodatkowe funkcje do tego celu, takie jak analiza WhatIf lub udostępnianie danych. Więcej informacji zobaczmy później w sekcji MOLAP

Lokalizacja platformy BI

Kiedy już wiemy, które komponenty będą częścią naszej platformy BI, musimy zdecydować, gdzie je ulokować. Aby podjąć taką decyzję, musimy zastanowić się, jaka jest polityka naszej firmy wobec pozostałych serwerów, które posiadamy; w przypadku wyboru rozwiązania licencjonowanego musimy wziąć pod uwagę różne ceny licencji dla wersji onpremise i cloud, biorąc pod uwagę zarówno licencje na system operacyjny, jak i licencje na narzędzia. Ponadto nasza firma może nałożyć pewne ograniczenia bezpieczeństwa dotyczące przesyłania poufnych danych do chmury oraz możliwe zmiany na naszej platformie. Głównie będziemy mieli trzy opcje:

On-Premise: będziesz mieć serwery zlokalizowane w Twoim CPD, w sieci firmowej i całkowicie dedykowane Tobie. Będziesz musiał zadbać o konserwację i aktualizację wersji zainstalowanego systemu operacyjnego i oprogramowania.

Chmura: będziesz korzystać ze współdzielonej infrastruktury serwerów wirtualnych kupując tylko pojemność i zapominając o utrzymaniu systemu operacyjnego i oprogramowania; robi to firma działająca w chmurze. Możesz mieć zawsze zaktualizowaną platformę do najnowszej wersji bez dbania o procesy aktualizacji.

Hybrydowy: Możesz myśleć w rozwiązaniu hybrydowym, umieszczając niektóre serwery w swoim CPD, a inne w chmurze. W tym przypadku najczęstszym podejściem jest posiadanie interfejsu BI w chmurze, o ile zwykłe ograniczenia bezpieczeństwa są silniejsze dla baz danych, więc być może twoja polityka bezpieczeństwa nie pozwala na posiadanie bazy danych w chmurze, ale o ile twoje Narzędzie BI front-end nie zapisuje żadnych danych, może znajdować się w dowolnym miejscu.

Koncepcje BI

W świecie BI istnieje wiele koncepcji, które mogą mieć różne nazwy w różnych narzędziach BI, ale ostatecznie odnoszą się do tego samego tematu. W tej części przyjrzymy się różnym koncepcjom, zaczynając od koncepcji hurtowni danych, poprzez model logiczny zestawu tabel z hurtowni danych, które są wykorzystywane bezpośrednio do raportowania, następnie zobaczymy model fizyczny i jego powiązania z logicznym oraz jakie elementy modelu są odwzorowywane w narzędziach BI i jak możemy je wykorzystać.

Hurtownia danych

Jedną z najważniejszych koncepcji związanych z rozwiązaniem BI jest Datawarehouse (DWH). DWH jest podstawą tradycyjnych systemów BI, ponieważ jest to miejsce, w którym znajdują się dane, które chcesz przeanalizować, zwykle obsługiwane przez system baz danych. Idea poniżej DWH polega na tym, że możesz zbierać dane z wielu źródeł, oczyszczać je, zapewniać ich integralność, zapewniać spójność i kompletność, aby mieć niezawodną implementację BI. Możesz mieć wiele źródeł ze swojej firmy, dane pochodzące z systemu transakcyjnego do standardowego przepływu Zamówienia/Dostawy/Faktury, dane pochodzące z narzędzia magazynowego, które kontroluje zapasy, które masz w środku twojego magazynu, dane pochodzące od firmy logistycznej dostarczającej Twoje produkty, ręczne grupowanie klientów lub produktów wykonane przez Twoich analityków, dane pochodzące od Twoich klientów detalicznych pokazujące, jak sprzedają Twoje produkty itp. Możesz mieć heterogeniczne dane w tych różnych źródłach, ale podstawą technik DWH jest powiązanie wszystkich tych danych w jakiś sposób z kompletnymi tabelami przeglądowymi, w których masz

wszystkie możliwe wartości dla swoich wymiarów, bez przerw w hierarchii i z unikalnymi relacjami między tabelami. Aby to osiągnąć, będziesz potrzebować wdrożenia procesów integralności, które mogą wymagać pewnych ręcznych działań wstecznych w celu poprawienia niektórych danych transakcyjnych lub głównego źródła klienta/produktu. Oczyszczanie danych jest wymagane, ponieważ dość często system transakcyjny pozwala na wprowadzanie informacji bez zbytej kontroli. Ostatecznie jest to kwestia braku koordynacji między systemem transakcyjnym a narzędziem BI. Można to łatwo wyjaśnić na prawdziwym przykładzie, który znaleźliśmy u jednego z naszych klientów. W tej firmie dość elastycznie wykorzystują bardzo popularne narzędzie transakcyjne, w którym można je ustawić wśród wielu innych cech, jakim jest przypisany do danego klienta członek zespołu sprzedaży. Ta relacja, oparta na projektowaniu procesu przez reguły biznesowe, jest wyjątkowa; nie powinno być klienta z przypisanym więcej niż jednym pracownikiem działu sprzedaży, ale system transakcyjny nie ma tego ograniczenia – pozwala ustawić więcej niż jednego. Kiedy otrzymujemy w systemie BI osobę przypisaną do klienta, widzimy dwie różne możliwości, co może powodować powielanie danych podczas analizy zysków klientów w hierarchii sił sprzedaży. To, co zrobiliśmy, to wdrożenie procesu, który gwarantuje, że masz tylko jedno zadanie; wybieramy jeden losowo, ale informujemy osobę odpowiedzialną za zarządzanie danymi podstawowymi o poprawieniu przypisania w transakcji, aby móc opublikować wiarygodną informację dla tego klienta przy kolejnym ładowaniu informacji. Inny przykład z życia wzięty, który widzieliśmy w przypadku różnych klientów reguły biznesowej, która jest ustalana przez odpowiedzialną firmę, ale czasami nie jest przestrzegana przez operatorów biznesowych, dotyczy przepływu Zamówienie-Dostawa-Faktura lub z powodu jakiegoś problemu technicznego, który może mieć problemy z integralnością. W tym przypadku główny problem, który znaleźliśmy, jest natury technicznej. U tego klienta istnieje automatyczny system zapisywania danych w bazie DWH. Za każdym razem, gdy ktoś wprowadza zamówienie, dostawę lub fakturę do systemu transakcyjnego, jest to automatycznie przesyłane do DWH. Reguła biznesowa mówi, że nie możesz mieć faktury bez powiązania z dostawą lub dostawy bez powiązania z zamówieniem. Ale jeśli z powodu jakiegoś problemu technicznego otrzymasz dostawę, ale nie otrzymałeś wcześniej powiązanego zamówienia, system BI odrzuca dostawę na stół kwarantanny, dopóki nie otrzymamy zamówienia w naszym systemie, a także ostrzega nas na koniec miesiąca o odrzuconych fakturach lub dostawach. Te dwa przykłady to tylko niektóre z niemal nieskończonych możliwości, jakie można znaleźć w naturze różnych środowisk biznesowych, które można znaleźć u swojego operatora.

Uwaga: Wdrażając system DWH, musisz dokładnie zrozumieć, jakie reguły biznesowe mają zastosowanie do Twojego modelu, i wprowadzić ograniczenia techniczne, aby wymusić realizację tych reguł, określając również działania naprawcze, które należy wykonać.

Innym pomysłem związanym z DWH jest możliwość przechowywania szerszej historii danych w stosunku do danych, do których masz dostęp w systemie źródłowym, o ile jednym z podstawowych wymagań stawianych systemowi BI jest możliwość analizy ewolucji Twoich danych w przeszłości oszacować, jaka może być przyszła ewolucja, a jeśli chcesz to zrobić konsekwentnie, będziesz potrzebować jak najwięcej informacji. DWH jest zwykle wypełniany dziennymi, tygodniowymi lub miesięcznymi procesami, więc zwykle informacje, o które można zapytać, mają pewne opóźnienie wynoszące co najmniej kilka godzin, a wynika to ze zwykłych procesów ETL, które są używane do wypełniania danych, a także ograniczenia i kontrole, które muszą wprowadzić, aby zapewnić wiarygodność danych. Istnieje również tendencja do posiadania intraday BI, ale jest to możliwe tylko w zależności od relacji między dostępnym sprzętem, optymalizacją procesów ETL i ilością danych, które chcesz uwzględnić w swoim DWH. Wewnątrz hurtowni danych znajdują się różne grupy tabel:

Tabele wejściowe: zawierają informacje bezpośrednio ze źródła danych, które próbujesz przeanalizować, czasami nazywane tabelami ODS, ponieważ zawierają informacje pobrane bezpośrednio z bazy danych ODS; Przechowywanie danych operacyjnych; który zwykle jest klonem, pełnym lub częściowym, systemu transakcyjnego, czasami nazywane są tabelami wejściowymi.

Tabele tymczasowe: Zawierają informacje tylko podczas procesu ETL, ponieważ są zwykle używane do rozwiązywania niektórych procesów obliczeniowych, których ze względu na problemy z wydajnością, złożoność lub inne przyczyny techniczne nie można rozwiązać w jednym kroku ETL. To są tabele, które będziemy uważać za obszar przejściowy. Mogą znajdować się w osobnej bazie danych lub w hurtowni danych.

Tabele końcowe: Są to zestawy tabel, które zostaną opublikowane w narzędziu BI do analizy. Ta grupa tabel będzie bezpośrednio związana z modelami logicznymi i fizycznymi, które przeanalizujemy w następujących sekcjach.

DataMart

Ideą DataMartu jest wyodrębnienie informacji z konkretnego obszaru wewnątrz firmy. Podczas gdy DWH przechowuje całe informacje, DataMart będzie zawierał informacje wydziałowe. W zależności od strategii możesz zdefiniować DataMart jako część DWH lub Twój DataMart może być umieszczony w oddzielnej bazie danych, a DWH jest poprzednim etapem Twojego DataMart. Możesz odizolować swój DataMart od reszty środowiska, umieszczając go na innym serwerze, w innej instancji bazy danych, w innej bazie danych, w innym schemacie bazy danych w tej samej bazie danych lub po prostu oddzielając go w sposób logiczny (przez nazwy tabel z przedrostkami lub sufiksami). Nie ma ogólnej rekomendacji, aby przeprowadzić to oddzielenie, będzie to zależęć od ilości danych, którymi zarządzasz w DataMart w porównaniu z całym środowiskiem, budżetu na wdrożenie DataMart, poziomu izolacji, na który możesz sobie pozwolić, oraz parametryzację, którą umożliwiała implementacja bazy danych. Wszystkie poprzednie rozważania, które wyjaśniliśmy w odniesieniu do DWH, można zastosować do DataMart. Jedyna różnica polega na tym, że Datamart zawiera tylko podzbiór danych. W dalszej części tego rozdziału omówimy projekty wewnątrz obiektów BI; DataMart może być bezpośrednio powiązany z projektem.

Model logiczny

Całkiem możliwe, że jeśli jesteś częścią zespołu IT w swojej firmie i chcesz zacząć od zdefiniowania struktury DWH wymaganej do obsługi analizy BI, którą chcesz wykonać, całkiem możliwe, że zaczniesz myśleć o tabelach, pola, typy danych, klucze podstawowe, klucze obce, widoki i inne techniczne rzeczy, zaczynając od utworzenia przykładowej tabeli, która na końcu jest konwertowana na ostateczną tabelę. Ale potem możesz zobaczyć brakujące pola, nie wiesz, jak je wypełnić, jak są powiązane z tym samym polem w innych tabelach itp. Raczej nie zalecamy rozpoczynania od logicznej definicji modelu, który chcesz wdrażać w oparciu o byty i relacje między nimi, zamiast bezpośrednio zaczynać od definicji technicznej. Dzięki modelowi logicznemu możesz zobaczyć, które tabele są powiązane z innymi, jakie pola można wykorzystać do łączenia tych tabel, a także możesz łatwo sprawdzić, czy wymagania analizy biznesowej odpowiadają proponowanemu modelowi i jego głównemu celowi ; służy do interakcji i tłumaczenia tych wymagań biznesowych na strukturę bazy danych. Wewnątrz modelu logicznego, a zatem odzwierciedlonego w modelu fizycznym, możemy znaleźć głównie trzy rodzaje tabel:

Tabele faktów : są to zazwyczaj największe tabele w modelu, ponieważ zawierają dane biznesowe do analizy, podsumowania i agregacji na podstawie pożądaných pól. Zawierają szczegółowe informacje o sprzedaży, kosztach, krokach operacyjnych, ruchach księgowych, danych dotyczących zasobów ludzkich, wizytach przedstawicieli handlowych lub innych danych, które chcesz przeanalizować. Nie

jest wymagane posiadanie maksymalnego poziomu szczegółowości we wszystkich tabelach faktów, w rzeczywistości zaleca się wstępne obliczenie niektórych tabel zbiorczych, aby uzyskać lepszy czas odpowiedzi z interfejsu raportowania.

Tabele relacji: Te tabele służą do łączenia wielu pojęć między nimi. Mogą one opierać się na bezpośrednich relacjach służących do definiowania hierarchii logicznych, jako harmonogram zawierający relacje między dniem, tygodniem, miesiącem, kwartałem i rokiem, lub mogą być wykorzystywane do wiązania niezależnych pojęć, takich jak produkt i klient, w oparciu o źródłową tabelę faktów, jak np. produkt i klienci, którzy mają jakikolwiek rejestr sprzedaży w całej historii DWH.

Tabele przeglądowe : Nazywane również tabelami głównymi, zawierają głównie identyfikator pojęcia i opis tego pojęcia, a także mogą zawierać identyfikatory wyższej hierarchii w przypadku powiązanych ze sobą różnych atrybutów. W poprzednim przykładzie dotyczącym czasu można mieć powiązaną tabelę wyszukiwania dla dnia zawierającą identyfikator dnia, format daty i identyfikator miesiąca, a następnie można znaleźć tabelę miesiąca zawierającą identyfikator miesiąca, opis miesiąca i powiązany identyfikator roku. Czasami do zdefiniowania relacji można użyć tabeli przeglądowej.

Model relacyjny

Modele logiczne są oparte na modelu relacyjnym, o ile są zwykle zlokalizowane w relacyjnej bazie danych. Nazywa się relacyjnym, ponieważ podstawą modelu są relacje między danymi, zwykle zapisywane w tabelach. Tabela jest zdefiniowana przez kolumny, a każdy wiersz danych jest relacją istniejącą między różnymi polami. Definiując prosty przykład, możesz rozważyć tabelę sprzedaży, w której masz kod produktu, kod klienta i kwotę sprzedaży sprzedanej w tym miesiącu. Z tej tabeli możesz pobrać wszystkie relacje między produktem a klientem, które mają jakiś związek ze sprzedażą. Możesz także łączyć tabele za pomocą instrukcji wyboru SQL (Structured Query Language), aby łączyć informacje z więcej niż jednej tabeli i uzyskiwać relacje pochodne. Ale więcej informacji zobaczymy w rozdziale 3 poświęconym SQL. Istnieje wiele klasyfikacji modeli, ale w tej książce omówimy główne typy modeli używane w hurtowniach danych, modele znormalizowane i zdenormalizowane, modele gwiazdy i płotka śniegu.

Model znormalizowany

Model znormalizowany ma na celu zredukowanie do minimum redundancji danych, optymalizację kosztów przechowywania poprzez uniknięcie wielokrotnego powtarzania tych samych danych. Ten rodzaj modelu jest wysoce zalecany do rozwiązań transakcyjnych i może być również używany w modelach DWH, ale należy zdawać sobie sprawę z jego ograniczeń. Ponieważ będziesz mieć związek tylko raz, nie możesz pozwolić na wiele zadań. Jeśli zmienisz relację, wpłynie to na wszystkie wiersze w DWH.

Model zdenormalizowany

Model zdenormalizowany chce poprawić wydajność zapytań, unikając łączenia w czasie wykonywania. Aby to zrobić, wymaga powtarzania danych wzdłuż tabel, aby zminimalizować liczbę połączeń wymaganych do rozwiązania zapytania użytkownika.

Uwaga: Najbardziej typową sytuacją w DWH jest sytuacja pośrednia między wysoką normalizacją a wysoką denormalizacją. W zależności od charakteru hierarchii i atrybutów użyjesz strategii znormalizowanej, strategii zdenormalizowanej lub strategii pośredniej.

Model gwiazdy

Model gwiazdny jest rodzajem modelu relacyjnego szeroko stosowanego w hurtowniach danych, szczególnie w przypadku małych DataMartów, gdzie masz tabele faktów (omówimy to w następnych sekcjach), które zawierają informacje o Twoich danych sprzedażowych, operacyjnych lub finansowych, a następnie mieć również kilka tabel przeglądowych połączonych z tabelą faktów przez niektóre kluczowe pola. Te tabele przeglądowe będą zawierać opisy lub inne pojęcia związane z polami kluczowymi.

Model płatka śniegu

Pomyśl teraz o płatku śniegu. Zwizualizujesz płatek z dużym rdzeniem, a następnie gałęzie, które dzielą się na mniejsze gałęzie. W podobny sposób możesz pomyśleć o modelu danych płatka śniegu, w którym pośrodku znajdują się duże tabele faktów, połączone z nimi przez niektóre kluczowe tabele, które zawierają podstawę do wyodrębnienia hierarchii głównych atrybutów, a następnie możesz znaleźć bezpośrednio połączoną tabelę wzorcową wyszukiwania do tych tabel kluczowych lub mniejszych tabel kluczowych, które wiążą dane z innymi tabelami głównymi

Model fizyczny

Po zdefiniowaniu modelu logicznego przejdziesz do spraw technicznych i będziesz musiał określić, które pola zawierają twoje tabele; które będą bezpośrednio powiązane z podmiotami modelu logicznego; jaki jest typ danych dla wszystkich tych pól; które pola będą używane do łączenia tabel; które będą unikalnymi kluczami tabel; jeśli Twój model będzie wymuszał integralność danych przy użyciu kluczy obcych lub jeśli wolisz unikać ograniczeń fizycznych w swoim modelu i opracować logiczne kontrole w celu ułatwienia procesu ETL.

Uwaga: Ważne jest zdefiniowanie nomenklatury dla tabel i pól. Ta nomenklatura może być bardzo ścisła, ale zalecamy posiadanie pośredniego poziomu nomenklatury między bardzo opisowymi nazwami a nazwami tylko technicznymi. W pełni otwarte nazwy mogą prowadzić do nazw tabel, takich jak SALES_AT_CUSTOMER_AND_PRODUCT_LEVEL lub SALES_AT_CUSTOMER_AND_PRODUCT_LEVEL_INCLUDING_PLANT. Jeśli spróbujesz naprawić bardzo ścisłą nomenklaturę, możesz znaleźć nazwy takie jak XP55GES001FT, które nie dają ci pojęcia, co może zawierać ta tabela. Ale z opcją pośrednią możesz naprawić, że twoje tabele faktów zaczynają się od F_, tabele relacji od R_, tabele wyszukiwania od L_, a następnie pozwalają niektórym znakom spróbować ustawić pewne znaczące nazwy, takie jak F_SALES_CUS_PRD lub F_SALES_CUS_PRD_PLT.

Jeśli chodzi o nazewnictwo pól, zaleca się przestrzeganie tych dwóch prostych zasad:

1. Nie ustawiaj tej samej nazwy pola dla różnych pojęć: zwykle narzędzia BI rozpoznają, kiedy masz tę samą nazwę pola i próbują połączyć informacje przez to pole, powodując utratę części danych. Musisz wziąć pod uwagę, że coś, co wydaje się być podobne, jak Region, może powodować niezgodność danych, jeśli mówisz o regionie klienta lub regionie pracownika, ponieważ w niektórych przypadkach mogą się one różnić.
2. Ustaw tę samą nazwę pola i typ dla tej samej koncepcji: podstawowa przyczyna jest taka sama, zwykle narzędzia BI rozpoznają to samo pole jako to samo pojęcie, ale ściśle przestrzeganie tego zmniejszy złożoność modelu, ponieważ nie musisz zapamiętać różne nazwy pól w różnych tabelach, aby dołączyć, jeśli mówisz o tych samych informacjach.

Chcemy zwrócić uwagę na znaczenie tych dwóch pomysłów z kilkoma dodanymi prawdziwymi przykładami. Wewnątrz klienta, który analizował wizyty swoich pracowników w sklepach, w których sprzedają swoje produkty, mieliśmy informację o tym, kto odwiedził i nazwaliśmy ją EMPLOYEE_ID. Wszyscy klienci powinni mieć przydzielonego jednego pracownika, ale tylko jednego. Wydaje się jasne,

że można mieć tabelę relacji między klientem a pracownikiem, a ponieważ byliśmy wtedy bardzo oryginalni, nazwaliśmy EMPLOYEE_ID pole związane z pracownikiem, który przypisał klienta. Ale co się stało, gdy próbowaliśmy zrobić analizę wizyt na pracownika? Narzędzie wykryło, że EMPLOYEE_ID jest dostępne w obu tabelach i zinterpretowało, że są takie same, więc połączyło obie tabele przez to pole i traciliśmy informację, czy wizytę wykonał inny pracownik, czy zmieniło się przypisanie klienta.

Główne obiekty BI

W różnych narzędziach BI możemy znaleźć podobne obiekty, które mogą mieć różne nazwy, ale są powiązane z podobną ideą. W tej sekcji postaramy się przeanalizować, które to obiekty, ale powinniśmy skupić się na idei stojącej za obiektem, a nie na samej nazwie obiektu.

Projekt

Jako projekt traktujemy zestaw informacji, który ma wystarczająco dużo podobieństw, aby można je było wspólnie analizować w oparciu o niektóre kluczowe pola relacji. W różnych narzędziach usłyszysz o projektach, wszechświatach, środowiskach, ksiązkach, wdrożeniach, ale wszystkie odnoszą się do tej samej koncepcji, grupy tabel, które mają sens razem analizować, ponieważ mają jakiś związek. Koncepcja projektu w BI jest bezpośrednio powiązana z koncepcją DataMart w bazie danych, którą już zdefiniowaliśmy jako zestaw powiązanych tabel, więc przeanalizujemy w projekcie BI bazę danych DataMart. Projekty są zwykle odizolowane od siebie i często narzędzia BI nie pozwalają na transport obiektów lub łączenie informacji pochodzących z różnych projektów, więc musisz być ostrożny przy podejmowaniu decyzji, czy umieszczasz nową analizę w tym samym projekcie, czy też wolisz utworzyć osobną, aby nie wymagać ponownej analizy, gdy zdasz sobie sprawę, że informacje muszą być analizowane w połączeniu z istniejącą.

Tabela

Najczęściej stosowana strategia polega na tym, że tabele bazy danych są odwzorowywane na pewnego rodzaju tabelę logiczną lub definicję tabeli, która jest własnym obiektem narzędzia BI, a następnie są wykorzystywane jako podstawa do projektowania pozostałych obiektów. Większość narzędzi BI będzie miała tego rodzaju logiczne definicje tabel wstawianych do swoich katalogów obiektów. To mapowanie może pochodzić z tabeli lub z widoku, a czasami samo narzędzie umożliwia definiowanie własnych widoków (lub selekcji) na poziomie narzędzia BI, zamiast robić to na poziomie bazy danych. Wszystkie tabele mają co najmniej jedno pole, które w kolejnych sekcjach podzielimy na dwa typy.

Fakt

Fakty to jeden z rodzajów pól, które możemy znaleźć w naszych tabelach. Są to zazwyczaj pola liczbowe, a ich główną cechą jest to, że można je agregować. Podstawą analizy są fakty i dlatego zwykle nazywane są danymi pomiarowymi. Łącząc fakt z formułą, którą chcemy zastosować do podsumowania danych (jakiś przykład zobaczymy, gdy będziemy mówić o możliwościach grupowania), otrzymamy miary, które również w zależności od narzędzia można znaleźć jako metryki, wskaźniki lub KPI. W zależności od narzędzia sam obiekt faktu nie istnieje i odwzorowujesz bezpośrednio pole za pomocą formuły, której chcesz użyć do jego agregacji. Przykładami faktów mogą być sprzedaż brutto, koszt, przychód, poświęcony czas, wielkość zapasów lub wykorzystana przestrzeń.

Wymiar

Jeśli niektóre pola są używane do agregacji, pozostałe pola podają poziom, na którym chcesz przeprowadzić tę agregację. Pola te są również nazywane danymi opisowymi, ponieważ opisują sposób dystrybucji danych i ilość analizowanych danych. Nazewnictwo różnych narzędzi to Oś, Atrybut,

Wymiar i można je pogrupować w Hierarchie, które zdefiniują relacje między tymi atrybutami lub wymiarami. W tym przypadku możemy myśleć o dacie, kliencie, produkcie, zakładzie, regionie, kraju, biurze sprzedaży lub jakimkolwiek innym polu, które pozwala podzielić informacje.

Raporty

Na końcu drogi głównym efektem działania narzędzia BI będą raporty na poziomie umożliwiającym użytkownikom końcowym analizę informacji. W zależności od narzędzia BI, ale także od charakterystyki i złożoności raportów, możemy znaleźć różne koncepcje, takie jak Raport, Dokument, Pulpit nawigacyjny, Widok, Analiza, Książka, Skoroszyt itp., ale wszystkie odnoszą się do końcowego interfejsu, który będzie służyć do analizy informacji. Mogą przejść od prostej siatki sprzedaży według kategorii do zestawu interaktywnych wizualizacji, które można przeglądać, przegłądać, obracać, drukować lub wysyłać e-mailem.

Podejścia BI

Istnieją różne podejścia BI, które zapewniają różne funkcje. Niektóre narzędzia BI łączą więcej niż jedno podejście w jedną platformę, ale zwykle nie obejmują wszystkich funkcjonalności. Musisz zdefiniować zakres swojego projektu, zanim będziesz mógł właściwie wybrać, które części rozwiązania będą potrzebne.

Zapytanie i raportowanie

Query and Reporting (Q&R) to początkowy etap typowych wdrożeń BI, o ile głównym celem systemu BI jest dostarczanie informacji w jakims kompleksowym formacie do ich analizy. Te informacje pochodzą z bazy danych, więc aby uzyskać te informacje, użyjemy Zapytania, które jest żądaniem do hurtowni danych w celu pobrania stamtąd danych. Z drugiej strony zwracane informacje są sformatowane w formacie przyjaznym dla użytkownika, aby były czytelne dla analityków. Mówiąc o Q&R mamy na myśli nie tylko możliwość dostępu do hurtowni danych w celu wydobycia z niej informacji, ale także narzędzie, które pozwala je analizować, zagłębiać się w szczegóły, sprawdzać czy wyróżniają się liczby, sprawdzać nietypowe zachowania odbiegające od normy trendy, filtruj otrzymane informacje i formatuj je zgodnie ze standardami firmy. Główną zaletą narzędzia Q&R jest to, że nie musisz znać języka SQL, aby móc wyodrębnić informacje z bazy danych; ten SQL jest generowany automatycznie przez narzędzie BI. W niektórych przypadkach, jeśli masz do tego umiejętności techniczne, narzędzie może dać ci możliwość modyfikowania, dostrajania lub bezpośredniego pisania od zera uruchamianego SQL, na wypadek gdyby model w narzędziu BI nie był w pełni dostosowane do poniższej bazy danych lub że potrzebujesz uzyskać informacje z innego zestawu tabel, który nie jest odwzorowany w katalogu BI. Ta możliwość będzie zależała również od używanego narzędzia BI; nie wszystkie pozwalają na tworzenie i modyfikowanie zapytań w SQL. W zależności od możliwości narzędzia BI i modelu Twojej organizacji, będą również tworzyć własne metryki, niestandardowe wymiary lub złożone filtry i grupowania. Możliwości BI przeanalizujemy w następnej sekcji. To rozwiązanie Q&R jest szczególnie przeznaczone dla analityków, którzy będą w stanie uzyskać większe korzyści z tej elastyczności tworzenia raportów ad hoc niż większość komponentów w organizacji.

Udostępnianie informacji

Analitycy w Twojej firmie będą bardzo zadowoleni z możliwości korzystania z raportów i zapytań ad hoc oraz przeprowadzania skomplikowanych analiz tworząc złożone metryki, ale zdecydowana większość pracowników w Twojej firmie będzie bardzo zadowolona, jeśli będzie miała bezpośredni dostęp do już utworzonych raportów lub uzyska swoje raporty w swoich skrzynkach pocztowych lub w

udostępnionej lokalizacji. Standardowym podejściem w środowisku BI jest posiadanie niewielkiego podzbioru pracowników z umiejętnościami analityka, którzy będą czerpać zyski z rozwiązania Q&R, ale następnie posiadanie grupy programistów (mogą to być te same grupy analityków), którzy będą tworzyć raporty do udostępniania z resztą organizacji albo za pośrednictwem tego samego narzędzia BI, aby użytkownicy mogli łączyć się z narzędziem BI w celu generowania raportów również z włączonymi funkcjami BI, albo za pośrednictwem jakiejś usługi dystrybucyjnej, takiej jak poczta e-mail lub folder udostępniony. Ta dystrybucja może być zwykle wykonana za pomocą narzędzia BI, ponieważ większość z nich pozwala na automatyczne dostarczanie wiadomości e-mail.

Pulpit nawigacyjny

Mówiąc o kokpitach BI, możesz pomyśleć o desce rozdzielczej samolotu. Można tam znaleźć wiele wskaźników pochodzących z różnych źródeł, które dają przegląd osiągnięć całego samolotu, od głównych silników po klapy skrzydłowe. Podobnie jak w desce rozdzielczej samolotu, deska rozdzielcza BI jest szczególnie istotna, aby mieć informacje o alertach, o ile trzeba zwrócić szczególną uwagę na te metryki, które są poza standardowym zakresem roboczym. Jak skomentowaliśmy w poprzednich sekcjach mówiących o istotności KPI, musimy skoncentrować nasze pulpity nawigacyjne na tych wskaźnikach, które są naprawdę istotne dla przeprowadzanej przez nas analizy, a także muszą być naprawdę znaczące w porównaniu z poprzednimi latami lub celami. Pulpity nawigacyjne BI mogą również oferować pewne funkcje, takie jak dynamiczne selektory, panele informacyjne, zależne siatki lub wykresy, w których kliknięcie części wykresu powoduje filtrowanie zależnej wizualizacji; przejść do szczegółowych informacji; przejść do powiązanego pulpitu nawigacyjnego; wszystkie opcje formatu, których możesz wymagać kolorów, czcionki, stylu, obrazu i kształtu; wiele układów informacji; podpowiedzi; lub osadzanie multimedii wśród innych funkcji.

Import danych

Jednym z głównych trendów w platformach BI jest możliwość szybkiej analizy informacji poprzez umożliwienie użytkownikowi dostępu do kanału importu własnych plików danych do interfejsu BI w celu wdrożenia szybkich dashboardów z wieloma wizualizacjami, dając użytkownikowi samoobsługę BI. Ta możliwość skraca również czas opracowywania projektów BI, które w przeszłości były dużymi projektami z długimi terminami realizacji. W zależności od narzędzia będziesz mógł importować pliki w różnych formatach (Excel, CSV, tekst), łączyć się z własnymi źródłami danych, łączyć się z interfejsami usług internetowych, takich jak Xquer, lub korzystać z plików znajdujących się na współdzielonych platformach chmurowych.

Wykrywanie danych

Ściśle powiązane z Importem danych jest podejście Data Discovery, które polega na zestawie wizualizacji specjalnie przeznaczonych do wyszukiwania trendów i wyjątków w bardzo łatwy sposób za pomocą bardzo intuicyjnego interfejsu. Ten rodzaj interfejsu jest bezpośrednią ewolucją pulpitu nawigacyjnego poprzez uproszczenie elementów sterujących i menu, ograniczenie dozwolonych opcji, ale skupienie się na mocniejszych. Być może użytkownik nie dysponuje bardzo precyzyjną siatką wyrównania, aby tworzyć idealne wykresy, ale może z łatwością powiązać jeden wykres z drugim i filtrować oba za pomocą prostych paneli filtrujących. Główną ideą Data Discovery jest umożliwienie użytkownikowi tworzenia własnych dashboardów bez znajomości narzędzi BI, tylko z bardzo intuicyjnym interfejsem z głównymi funkcjonalnościami.

MOLAP

Bazy danych wielowymiarowego przetwarzania analitycznego OnLine (MOLAP) są zorientowane na lepszą wydajność procesu zapytania. Zwykle są postrzegane jako sześcian, o ile mają wiele wymiarów. Tak, ściśle mówiąc, sześcian ma tylko trzy wymiary, ale jest to ilustracyjny sposób ich narysowania, aby odróżnić je od relacyjnych tabel relacyjnej bazy danych. Zamiast wielu tabel z faktami, relacjami i informacjami wyszukiwania, bazy danych MOLAP zawierają wszystkie informacje razem, wstępnie obliczone dla wszystkich wymiarów, na wszystkich poziomach różnych wymiarów, biorąc pod uwagę przecięcie ze wszystkimi poziomami pozostałych wymiarów. To wstępne obliczenie ma dwa główne implikacje dotyczące wydajności. Pobieranie informacji z bazy danych jest bardzo sprawne, ale z drugiej strony ładowanie informacji do bazy danych MOLAP może zająć bardzo dużo czasu; innymi słowy, jeśli chcesz mieć rozsądny czas ładowania, będziesz musiał zmoderować poziom szczegółowości swoich kostek MOLAP. Trudno znaleźć bazę danych MOLAP o takim samym, ani podobnym poziomie szczegółowości jak hurtownia danych. Bazy danych MOLAP zazwyczaj mają również możliwość zapisywania w nich danych przez użytkowników końcowych, a ta funkcja otwiera możliwość wykorzystania ich jako narzędzi do planowania i budżetowania. Dość często można znaleźć bazę danych MOLAP z wymiarem o nazwie Scenariusz lub Wersja, który może być używany do przechowywania rzeczywistych danych z bieżącego roku i prognozy na następny rok, a także z różnymi wersjami przyszłej prognozy, aby móc przeprowadzić analizę WhatIF, ponieważ na przykład, co by się stało, gdybym w przyszłym roku zwiększył sprzedaż o 15%, inwestując 10% więcej w reklamę? Teraz spróbujmy zwiększyć 5% w tym produkcie, ale 15% w innym produkcie. Teraz zobaczymy, czy obniżę koszty produkcji o 11% itd. Możesz przeprowadzić i porównać wiele symulacji różnych scenariuszy przed opublikowaniem celów firmy na następny rok. Czasami MOLAP nie jest uważany za klasyczną funkcjonalność BI, ponieważ BI miał być narzędziem analitycznym rzeczywistości bez możliwości pisania, ale obecnie uważamy, że dość ważna jest możliwość rysowania scenariuszy bawiących się różnymi zmiennymi, a kiedy już to zrobisz zdefiniowałeś, jaki jest twój cel, zamknij cykl życia firmy, próbując go osiągnąć.

Eksploracja danych

Jeśli widzieliśmy ostatnio, że MOLAP pozwala wyobrazić sobie, jaka może być przyszłość, zobaczymy, że eksploracja danych pozwoli ci ją przewidzieć. Eksploracja danych pomoże Ci to przewidzieć na podstawie możliwości wykrywania ukrytych trendów i wzorców. Jak możesz sobie wyobrazić, te możliwości predykcyjne opierają się na przeszłych informacjach i formułach prognozowania, więc chociaż możesz wykonać bardzo dobrą robotę, definiując je, istnieje możliwość, że nieoczekiwany fakt może całkowicie zmienić scenariusz, powodując zupełnie inne wyniki niż twoje planowanie. Dzięki eksploracji danych próbujesz znaleźć model, który uzasadnia relację między wejściem a wyjściem twojego procesu, dzięki czemu możesz lepiej zrozumieć swój biznes. Wyobraź sobie, że prowadzisz firmę zajmującą się parkowaniem samochodów. Aby wymodelować Twój biznes, zacznijmy od relacji pomiędzy posiadanymi parkingami a miesięczną kwotą faktury. Zaczynasz z 10 miejscami parkingowymi, a Twoja miesięczna kwota faktury to 2000 \$. Po roku możesz dokupić 10 dodatkowych miejsc parkingowych, a Twoja miesięczna faktura wzrośnie do 4000 USD. Na podstawie tych danych zależność między liczbą miejsc parkingowych a miesięczną kwotą faktury wydaje się być liniowa. Decydujesz się więc na budowę pełnego parkingu budowlanego na 500 miejsc parkingowych, ale potem widzisz, że nie jesteś w stanie wynająć więcej niż 40, a także cena wynajmu tych parkingów spada. Twój model nie był poprawny, ponieważ nie wzięłeś pod uwagę gęstości zaludnienia na swoim obszarze, liczby istniejących parkingów poza twoim itp. Szczególnie interesujące jest znalezienie tych wzorców, aby zwrócić uwagę na nietypowe wartości, wyniki zerowe i rozproszenie danych, ponieważ mogą udzielić ci informacji zwrotnej na temat poprawności twojego modelu. Więc jeśli chcesz uzyskać jak najdokładniejsze wyniki, musisz być bardzo wyczerpujący podczas definiowania modelu eksploracji danych, uwzględniając jak najwięcej zmiennych, aby uzyskać model biznesowy, który pozwoli ci

przewidzieć przyszłe wyniki. Podstawowe obliczenia prognozy opierają się na dwóch głównych składnikach: poprzednich wynikach i trendzie. Następnie musisz rozpocząć analizę tego, jak różnią się one w zależności od różnych wymiarów. Zwykle możesz zacząć od czasu, aby sprawdzić, czy występuje sezonowość, cykl w ciągu miesiąca lub wiele lat. Następnie możesz spróbować sprawdzić, czy produkt wpływa w jakiś sposób ze względu na cykl życia produktu, który jest korelacją między inwestycjami w reklamę a przychodami netto, lub jakąkolwiek inną zmienną, którą chcesz wziąć pod uwagę.

Przychodzące podejścia

Trudno powiedzieć, dokąd możemy dojść w ciągu najbliższych dziesięciu lat w podejściach i opcjach BI. Jak skomentowano w sekcji Ewolucja BI, obecnie głównym celem jest opracowywanie rozwiązań Big Data, aby móc uzyskać dostęp do trendów w nieskończonej ilości informacji, które są generowane wszędzie przez każde urządzenie elektroniczne. Będzie to wymagało dostosowania narzędzi BI, aby miały wystarczającą moc do analizy tej ilości danych, a także zmiany filozofii klasycznego Datawarehouse, aby móc zarządzać taką ilością informacji bez rygorów reguł Datawarehouse.

Możliwości BI

W zależności od narzędzia, którego używasz, zobaczysz te możliwości z różnymi szczegółami, na przykład możesz znaleźć narzędzie, które pozwala drążyć dane, wystarczy dwukrotnie kliknąć nagłówek pola, inne, które pozwalają również wybrać sposób drążenia lub inne, że wiercenie jest ustawiane przez administratorów i nie można wyjść z określonych ścieżek, ale mają one głównie pewne cechy, które postaramy się przeanalizować w tej sekcji. Zaraz skończysz to nudne wprowadzenie, aby zacząć od bardziej technicznych rozdziałów, zwłaszcza od rozdziału 4 do rozdziału 10, które będą znacznie bardziej interaktywne. Daj spokój!

Drążenie

Widzieliśmy już kilka przykładów, które wykorzystują możliwości drążenia, aby dojść do pewnych wniosków, ale chcielibyśmy zdefiniować, jakie jest znaczenie drążenia w środowisku BI. Drill to możliwość poruszania się po danych w modelu w celu uzyskania bardziej szczegółowych informacji o niektórych pojedynczych wierszach. Jest to szczególnie przydatne, gdy próbujesz odkryć, co powoduje niezwykłą ilość danych, która wyróżnia się na tle pozostałych wartości, które widzisz. Kiedy analizujesz raport według pracownika i jeden z pracowników sprzedaje trzy razy więcej jednostek niż średnia lub inny pracownik sprzedaje mniej niż połowę, możesz przejść do hierarchii produktów, aby zobaczyć, co jest pierwotną przyczyną różnicy, lub przejść do szczegółów czas, aby sprawdzić, czy pracownik był poza biurem przez dwa tygodnie z powodu urlopu, lub przeszukać kraje lub regiony, aby sprawdzić, czy bestseller obejmuje jakiś zamożny obszar itp. Przykład użycia arkusza programu Excel z tabelą przestawną do zilustruj, że może to być następujący, w którym możesz drążyć z projektu RUN, aby przeanalizować, które zadania są wykonywane w tym projekcie.

Obracanie

Dane przestawne odnoszą się do możliwości przenoszenia pojęć z wierszy do kolumn i odwrotnie, przenoszenia niektórych pojęć do przestrzeni filtrów, przenoszenia niektórych pojęć w celu generowania różnych stron informacyjnych, ogólnie możliwości gry polami w celu uporządkowania danych które wyświetlasz w przyjaznym formacie. Na przykład łatwiej jest zrozumieć raport, który analizuje trendy, jeśli masz koncepcję czasu w kolumnach, jeśli masz ją w wierszach lub w przypadku wykresu liniowego, znacznie łatwiej jest zobaczyć ewolucję, jeśli czas jest na X oś. Ale może w połączeniu z drążeniem wykryjesz miesiąc, w którym wszyscy Twoi klienci mają wzrost sprzedaży, a następnie przeniesiesz atrybut czasu do obszaru filtrowania, wybierzesz miesiąc ze wzrostem i

przeniesiesz atrybut produktu ze strony do obszar wierszy, aby je porównać, przenosząc klienta do kolumn, aby zobaczyć relację między klientem a produktem w danym miesiącu.

Wizualizacje

W wielu przypadkach graficzne przedstawienie danych ułatwia prawidłowe zrozumienie danych ukrytych za zwykłą siatką informacji. Dlatego wszystkie narzędzia BI oferują graficzne wizualizacje danych, które w zależności od narzędzia mogą być bardziej zaawansowane i ewoluować niż inne. Większość narzędzi, a może wszystkie, oferuje zestaw wykresów, takich jak grafika słupkowa, grafika liniowa lub grafika powierzchniowa, wszystkie pionowe i poziome; wykresy kołowe, rozproszone lub pierścienie są również powszechne w większości narzędzi. Następnie możesz znaleźć konkretną zaawansowaną wizualizację, która jest dostarczana tylko przez niektóre narzędzia, takie jak chmura słów, mapa cieplna, wykresy relacji, wykresy kaskadowe, sunburst, ułożone słupki, wykres strumieniowy lub wiele innych wizualizacji, które można wykorzystać do lepszego zrozumienia raport.

Uwaga: niektóre narzędzia umożliwiają korzystanie z niestandardowych wizualizacji, które można opracowywać i importować samodzielnie. Jeśli masz jakieś specjalne możliwości wizualizacji lub jakiegokolwiek rodzaj dostosowania, powinieneś sprawdzić dokumentację swojego dostawcy, aby zobaczyć, czy zezwalają na tego rodzaju dodatkową wizualizację.

Wizualizacje map

We wszystkich możliwych opcjach wizualizacji istnieje rodzaj wizualizacji, który naszym zdaniem wymaga dla nich dodatkowej sekcji. Pochodzące z uniwersalizacji możliwości GPS w urządzeniach mobilnych, które są wykorzystywane do gromadzenia danych w połączeniu z możliwościami integracji map z platformami takimi jak Google Maps, ESRI Maps lub CartoDB wśród wielu innych opcji komercyjnych; możesz mieć w swoim narzędziu BI wizualizację mapy, która ułatwia poprawność zrozumienia, w jaki sposób dystrybuowana jest Twoja sprzedaż w całym kraju lub w jaki sposób Twoje siły sprzedaży obejmują określony obszar.

Sortowanie

Inną funkcją, której będziesz potrzebować w swojej analizie, są opcje sortowania. Jest to również oferowane przez większość narzędzi, ale mogą również występować pewne różnice w opcjach, na które pozwalają. Możesz sortować rosnąco lub malejąco, według jednego lub wielu pól, na podstawie pól numerycznych, alfanumerycznych lub dat, na podstawie wymiarów lub metryk, a także z możliwością sortowania wyświetlanych stron.

Grupowanie

Istnieją dwa rodzaje grupowania w świecie BI. Pierwsza jest podstawowa dla analizy BI; nie możesz szczegółowo przeanalizować setek milionów wierszy, więc musisz zebrać te informacje w coś czytelnego. Aby to zrobić, wszystkie narzędzia BI zapewniają możliwości grupowania przy użyciu języka SQL, który jest używany do wysyłania zapytań do bazy danych. Aby uzyskać pogrupowane informacje, możesz wybrać szczegółowe pola, które chcesz przeanalizować, i użyć funkcji agregacji, takich jak suma, średnia, maksimum, minimum, ostatnia, mediana itp. - zwykle w odniesieniu do pól liczbowych. Otrzymujesz więc wynik formuły zastosowanej na wszystkich dostępnych informacjach na pożądanym poziomie wyjściowym. Ten pierwszy typ grupowania zostanie przeanalizowany z dalszymi szczegółami w rozdziale SQL. Ale istnieje również inne grupowanie uważane za grupowanie elementów, ponieważ nie jest to grupowanie techniczne oparte na wartościach pola; zamiast tego jest to grupa tych wartości do niestandardowej agregacji. Wyjaśnijmy to na przykładzie, który pokaże ci, o czym mówimy. W swoim DWH masz pole, które klasyfikuje region twojego klienta, oraz drugie, które pokazuje, gdzie

znajduje się kraj. Z drugiej strony w twojej strukturze odpowiedzialnej za sprzedaż masz zespół odpowiedzialny za regiony północne i inny zespół odpowiedzialny za regiony południowe. Możesz mieć raport sprzedaży według kraju lub sprzedaży według regionu, ale być może chcesz pogrupować niektóre regiony, aby pokazać zagregowaną wartość ich zestawu i porównać ją z resztą, a pole kraju nie jest dla ciebie przydatne, ponieważ potrzebujesz dodatkowych szczegółów. Dzięki funkcjom grupowania możesz grupować regiony południowe i północne, aby wyświetlać informacje tylko na tym poziomie szczegółowości, ani w regionie, ani na poziomie kraju. Aby wykonać tego rodzaju grupowanie elementów, możesz znaleźć różne opcje w różnych narzędziach. Niektóre z nich umożliwiają grupowanie informacji w locie na podstawie wyników raportu za pośrednictwem tego samego interfejsu raportu, a inne narzędzia umożliwiają tworzenie atrybutów pochodnych lub wymiarów w celu wyświetlenia informacji pogrupowanych na tym poziomie podczas korzystania z tych obiektów ; czasami te narzędzia umożliwiają grupowanie informacji na podstawie wyników metryki, a także prostsze narzędzia BI będą wymagały przeniesienia tej możliwości do silnika bazy danych, więc musisz opracować dodatkowe tabele/widoki w bazie danych, aby powiązać różne elementy .

Filtracja

Kiedy Twoje dane bazowe składają się z setek milionów wierszy, setek pól w tysiącach tabel, jest więcej niż możliwe, że nie wykorzystasz tak dużej ilości informacji w jednym raporcie. Zwykle raporty są filtrowane według jednego lub kilku pól. Ogólnie rzecz biorąc, chcesz przeanalizować informacje dzienne, zamknięcie miesiąca finansowego, porównać je z tym samym miesiącem poprzedniego roku, zobaczyć ewolucję ostatnich dwóch lat, ale nie uzyskać w raporcie całej historii hurtowni danych. Będziesz filtrować według dnia, miesiąca lub roku. Możesz także wymagać filtrowania według kategorii produktów, za które odpowiadasz, usuwając z danych źródłowych informacje związane z wewnętrznym kodem klienta, którego Twój zespół używa do przydzielania bezpłatnych próbek lub ograniczając informacje do dziesięciu najlepszych klientów w sprzedaży dla każdego regionu. Ponownie, możliwości filtrowania będą się różnić w zależności od platformy. W zależności od platformy będziesz mógł zakwalifikować się na podstawie wymiaru lub atrybutu (równy, większy, niższy, pomiędzy, na liście, nie na liście, jest pusty, nie jest pusty, zawiera, zaczyna się itp.), wybrać spośród różnych wartości tego wymiaru lub przefiltruj wynik a obliczenia, które mogą być wynikiem na poziomie raportu lub na innym poziomie (możesz przefiltrować dziesięciu najlepszych klientów, aby przeanalizować ich sprzedaż, lub przefiltrować dziesięciu najlepszych klientów i przeprowadzić analizę według produktu, aby zobaczyć, które są ich ulubionymi produktami). W zależności od narzędzia BI będziesz mieć możliwość dynamicznego filtrowania w czasie wykonywania. Możesz mieć stały raport ze zmiennym filtrem, co pozwoli ci, używając tego samego raportu, wyodrębnić informacje o różnych klientach lub regionach.

Wyrażenia warunkowe

Wewnątrz wyrażeń warunkowych możemy znaleźć wiele formuł, takich jak przypadek, a jeśli połączy się je z logicznymi, takimi jak i, lub, i nie, to da to możliwość tworzenia złożonych obliczeń, które można wykorzystać do segmentacji klientów lub przypisania celów pracownikom. Wyobraź sobie, że wychodząc z Twojej transakcji, otrzymujesz informacje finansowe. Otrzymujesz różne wartości kosztów, a znak pochodzący z transakcji ma różne znaczenia w zależności od kosztów. Kiedy przetwarzasz zachętę do sprzedaży, jest to wartość ujemna, ale gdy otrzymujesz koszty operacyjne, jest to wartość dodatnia. Jeśli chcesz mieć metrykę, która sumuje wszystkie koszty, będziesz potrzebować zdefiniowania takiej formuły:

```
Case when cost_type = 'SI' then - Cost_value else +  
cost_value
```

Uwaga: Aby opracować złożone obliczenia, będziesz potrzebować wyrażeń warunkowych. Oczywiście tego typu warunki można zastosować w bazie danych, ale jeśli pozwala na to Twoje narzędzie BI, będziesz mieć możliwość zrobienia tego na poziomie BI. Zawsze istnieje równowaga między złożonością rozwiązana w bazie danych a złożonością rozwiązana w narzędziu BI. W większości przypadków będzie to zależać od preferencji programisty; można to również zdefiniować w wytycznych dla programistów i książce najlepszych praktyk.

Sumy częściowe

Gdy uruchamiasz raporty płaszczyzny z więcej niż jednym atrybutem lub wymiarem w obszarze podziału, ponieważ wymagasz przeprowadzenia połączonej analizy różnych koncepcji, możliwe, że chcesz poznać część szczegółów każdej kombinacji atrybutów, które są sumami częściowymi każdej wartości jednego lub więcej atrybutów. Również na przykładzie pochodzącym z Excela możesz zobaczyć, że będziesz mógł wiedzieć, jaka jest łączna liczba godzin dziennie, wliczając wszystkich współpracowników w projekcie.

Administracja

Wszystkie platformy BI muszą być zarządzane przez jakiś zespół administracyjny. Ponownie funkcjonalności i możliwości platformy będą się znacznie różnić w zależności od narzędzia. Wszystkie muszą mieć w taki czy inny sposób definicję bezpieczeństwa, zarządzanie użytkownikami, zarządzanie grupami, zarządzanie rolami, które określają, jaki dostęp jest dozwolony dla każdego użytkownika, z jakich uprawnień i funkcjonalności może korzystać oraz z jakich danych użytkownik jest pozwolono zobaczyć. Ale oprócz podstawowego zarządzania bezpieczeństwem, będziesz miał możliwość ustawienia pewnych limitów dostrajających wydajność serwera, parametrów konfiguracyjnych wpływających na sposób dostępu do bazy danych, parametrów konfiguracji webowej w przypadku narzędzia z komponentem webowym, konfiguracji klastra w przypadku wysokie wymagania wydajnościowe, integracja bezpieczeństwa z uwierzytelnianiem LDAP lub Single Sign On. Będziesz także musiał ocenić, jakie są Twoje wymagania administracyjne, aby upewnić się, że narzędzie BI, o którym myślisz, jest właściwe.

Interfejs użytkownika

Innym tematem, który należy ocenić, są interfejsy oferowane przez narzędzie BI w celu uzyskania dostępu i doświadczenia użytkownika końcowego. W celu uzyskania dostępu do informacji znajdziesz narzędzia, które oferują Twoim użytkownikom możliwość dostępu poprzez instalację klient-serwer, instalację jednostanowiskową, dostęp sieciowy do połączenia z systemem lub oferują otrzymywanie informacji z niektóre narzędzia dystrybucyjne, głównie e-mail lub współdzielony zasób sieciowy. Oczywiście istnieją narzędzia, które oferują je wszystkie i inne narzędzia, do których dostęp jest bardziej ograniczony. Ten temat powinien być również uwzględniony w twoim panelu decyzyjnym, ponieważ może to spowodować odrzucenie użycia jakiegoś narzędzia ze względu na twoje wytyczne dotyczące bezpieczeństwa lub ponieważ zwiększa złożoność projektu, na przykład w przypadku konieczności opracowania instalacji pakietu w porozumieniu z twoimi Najlepsze praktyki instalacyjne.

Uwaga: Podsumowując, najlepszą strategią znalezienia najodpowiedniejszej opcji jest utworzenie panelu decyzyjnego w arkuszu lub slajdzie w celu oceny wszystkich możliwości różnych platform, które są obowiązkowe dla Twojego projektu, które warto mieć i które są bez znaczenia; ustaw ocenę dla każdej sekcji, a następnie wybierz najlepsze możliwe narzędzie, które odpowiada Twojemu budżetowi.

Wniosek

Jeśli dotarłeś tutaj, powinieneś mieć teraz ogólny zarys tego, czym jest rozwiązanie BI, jakie są jego główne komponenty i funkcjonalności, jakie są jego najważniejsze możliwości oraz główne kamienie milowe jego ewolucji; powinieneś znać również pewne pojęcia dotyczące BI i jego głównego komponentu, jakim jest hurtownia danych. Przejdźmy teraz do szczegółowej analizy, jaka jest nasza zalecana metodologia wdrażania platformy BI, i przeanalizujemy bardziej szczegółowo, jak zainstalować i wdrożyć główne części BI w Twojej organizacji. Ciesz się!