Analiza danych sprzedażowych z wykorzystaniem Snowflake i Power BI w chmurze Azure

Usługi sieciowe w biznesie

Małgorzata Radomska 173203 Inżynieria i analiza danych

Spis treści

1.	Wstę	p	3
	1.1.	Cel projektu	3
	1.2.	Wykorzystywane technologie	3
	1.2.1	Snowflake	3
	1.2.2	Microsoft Azure	4
	1.2.3	B. Power BI	5
	1.2.4	l. Dlaczego wybrano te technologie?	6
2.	Imple	ementacja projektu	8
	2.1. Utv	worzenie konta Snowflake	8
	2.2.	Rozpoczęcie pracy w Snowflake	0
	2.2.1	Tworzenie nowej bazy danych	0
	2.2.2	lmport danych	.1
	2.2.3	3. Tworzenie zapytań	4
	2.2.4	l. Tworzenie widoków	5
	2.3.	Instalacja Power BI Desktop1	7
	2.4.	Integracja Snowflake z Power BI	9
	2.5.	Tworzenie raportu	.3
	2.5.1	Tworzenie wizualizacji2	4
3.	Pods	umowanie	7

1. Wstęp

1.1. Cel projektu

Celem projektu jest zaprojektowanie i wdrożenie procesu analizy danych sprzedażowych z wykorzystaniem platformy Snowflake działającej w chmurze Microsoft Azure oraz integracja z narzędziem wizualizacji danych Power BI. Projekt ma na celu zademonstrowanie możliwości nowoczesnych rozwiązań chmurowych w zakresie magazynowania danych, przetwarzania zapytań SQL oraz prezentacji wyników w formie interaktywnych dashboardów, wspierających podejmowanie decyzji biznesowych.

1.2. Wykorzystywane technologie

1.2.1. Snowflake



Co to jest?

Snowflake to platforma do przechowywania i analizy danych, działająca całkowicie w chmurze. W przeciwieństwie do tradycyjnych hurtowni danych, które wymagają własnej infrastruktury (serwerów, baz danych, administratorów), Snowflake jest usługą typu Data Warehouse as a Service (DWaaS) – działa w pełni online, bez potrzeby instalowania czegokolwiek na komputerze użytkownika.

Jak działa?

Snowflake to platforma, która w pełni działa w chmurze, co oznacza, że wszystkie dane przechowywane są na zdalnych serwerach dostawców chmurowych, takich jak Microsoft Azure, Amazon Web Services (AWS) czy Google Cloud. Użytkownik nie musi instalować żadnego dodatkowego oprogramowania – wystarczy dostęp do przeglądarki internetowej i konto w Snowflake, aby móc rozpocząć pracę z danymi.

Kluczowym elementem działania Snowflake jest jego unikalna architektura, która rozdziela trzy główne funkcje systemu: przechowywanie danych, przetwarzanie danych i zarządzanie usługami. Dzięki temu rozwiązaniu Snowflake umożliwia niezależne skalowanie każdego z tych elementów, co w praktyce oznacza dużą elastyczność i wydajność.

Dane w Snowflake przechowywane są w formacie zoptymalizowanym do szybkiego odczytu i analizy, a użytkownicy mogą łączyć w jednym miejscu zarówno dane strukturalne (na przykład tabele w formacie CSV), jak i dane półstrukturalne, takie jak JSON. Nad warstwą przechowywania znajduje się warstwa przetwarzania, która odpowiada za wykonywanie zapytań i analiz. W Snowflake przyjmuje ona formę tzw. virtual warehouses, czyli niezależnych zasobów obliczeniowych przypisanych do konkretnego użytkownika lub zespołu. Każdy zespół może korzystać z własnego magazynu bez obawy, że jego zapytania spowolnią pracę innych użytkowników systemu.

Trzecią warstwą, spajającą całą platformę, jest system zarządzania, który odpowiada za bezpieczeństwo, uwierzytelnianie, optymalizację zapytań SQL i wiele innych procesów. Snowflake

automatycznie dobiera najlepszy sposób wykonania zapytania na podstawie dostępnych zasobów i charakteru danych, optymalizuje zasoby oraz dba o integralność danych, co sprawia, że korzystanie z tej platformy jest wygodne nawet dla osób bez zaawansowanej wiedzy technicznej.

Dzięki rozdzieleniu warstwy przechowywania danych od warstwy obliczeniowej, użytkownik może w dowolnym momencie zwiększyć moc przetwarzania bez potrzeby przenoszenia czy reorganizacji danych, lub – odwrotnie – zwiększyć pojemność magazynu danych bez wpływu na wydajność zapytań. Jest to ogromna przewaga nad tradycyjnymi hurtowniami danych, gdzie pamięć i moc obliczeniowa są ze sobą nierozerwalnie związane i wymagają kosztownej rozbudowy całej infrastruktury. W Snowflake płacimy tylko za to, z czego faktycznie korzystamy, co czyni tę platformę idealnym rozwiązaniem dla firm o dynamicznie zmieniających się potrzebach analitycznych.

Główne zalety:

- Działa w chmurze nie trzeba martwić się o serwery ani oprogramowanie.
- Skalowalność automatyczne dopasowanie mocy obliczeniowej do potrzeb użytkownika.
- Wydajność bardzo szybkie przetwarzanie dużych zbiorów danych.
- Bezpieczeństwo dane są szyfrowane, a dostęp do nich może być dokładnie kontrolowany.
- Integracja z narzędziami można łatwo połączyć Snowflake z Power BI, Tableau, Excel, itp.
- Obsługa różnych typów danych zarówno tradycyjnych (np. tabele), jak i półstrukturalnych (JSON, XML).
- Praca równoległa wiele zespołów może pracować jednocześnie bez wzajemnego spowalniania się.

Dla kogo?

Z Snowflake korzystają zarówno małe firmy, jak i duże korporacje (np. Disney, Siemens, Zoom, Żabka). To rozwiązanie świetnie sprawdza się tam, gdzie analiza danych ma wspierać podejmowanie decyzji biznesowych, marketingowych, finansowych czy logistycznych.

1.2.2. Microsoft Azure



Co to jest?

Microsoft Azure to rozbudowana platforma chmurowa stworzona przez firmę Microsoft, która umożliwia tworzenie, uruchamianie i zarządzanie aplikacjami oraz usługami w globalnej sieci centrów danych. Azure oferuje setki gotowych usług – od przechowywania danych, przez bazy danych i sztuczną inteligencję, aż po analizę danych, automatyzację procesów i cyberbezpieczeństwo. Wszystko to dostępne jest w modelu chmurowym, co oznacza, że użytkownik nie musi posiadać własnej infrastruktury – płaci tylko za zasoby, z których rzeczywiście korzysta.

Jak działa?

Azure działa w oparciu o infrastrukturę rozproszoną na całym świecie – centra danych znajdują się w wielu krajach, co zapewnia wysoką dostępność usług oraz niskie opóźnienia. Gdy firma chce przechowywać dane, stworzyć aplikację webową, wdrożyć usługę analityczną lub zbudować środowisko testowe, może to zrobić bez kupowania i utrzymywania własnych serwerów – zasoby zostają przydzielone w chmurze.

Platforma pozwala użytkownikom zarządzać zasobami przez intuicyjny interfejs przeglądarkowy (portal Azure), z poziomu aplikacji mobilnej lub za pomocą narzędzi programistycznych takich jak Visual Studio. Co ważne – Azure wspiera wiele języków programowania, systemów operacyjnych i baz danych, dzięki czemu jest elastyczny i otwarty na różne potrzeby użytkowników.

W kontekście analizy danych – Azure oferuje usługi takie jak Azure Synapse Analytics, Azure Data Lake, Azure SQL Database, a także obsługuje zewnętrzne platformy, np. Snowflake. Integracja z Power BI jest płynna i szybka, ponieważ wszystkie te narzędzia należą do ekosystemu Microsoft. Dodatkowo, dzięki Azure Active Directory, zarządzanie dostępem i bezpieczeństwem użytkowników odbywa się w sposób centralny i zautomatyzowany.

Główne zalety:

- Szeroki zakres usług od prostego przechowywania danych po zaawansowane rozwiązania
 Al i big data
- Wysoka dostępność centra danych Microsoft rozmieszczone są globalnie
- Elastyczność i skalowalność zasoby można dostosować do aktualnych potrzeb firmy
- Integracja z ekosystemem Microsoft łatwe połączenie z Power BI, Excel, Teams czy Windows Server
- Zaawansowane zabezpieczenia szyfrowanie, kontrola dostępu, certyfikaty zgodności
- Obsługa różnych technologii działa z Linuxem, Pythonem, JavaScript, bazami SQL i NoSQL.

Dla kogo?

Microsoft Azure jest przeznaczony zarówno dla małych i średnich firm, które chcą przejść do chmury bez dużych inwestycji w infrastrukturę, jak i dla dużych korporacji, które potrzebują globalnie dostępnych, skalowalnych i bezpiecznych rozwiązań IT. Jest także popularnym wyborem wśród deweloperów tworzących aplikacje internetowe i mobilne, specjalistów od danych oraz analityków, którzy korzystają z usług takich jak Snowflake, Power BI czy Azure Data Lake.

1.2.3. Power BI



Co to jest?

Power BI to narzędzie stworzone przez firmę Microsoft, które służy do analizy danych i tworzenia interaktywnych raportów oraz dashboardów. Umożliwia łączenie się z różnymi źródłami danych, ich

przekształcanie, modelowanie i prezentację w czytelnej formie wizualnej – za pomocą wykresów, map, tabel i wskaźników. Dzięki Power BI użytkownicy mogą podejmować lepsze decyzje biznesowe, bazując na aktualnych, przystępnie przedstawionych danych.

Jak działa?

Power BI działa w oparciu o trójskładnikowy ekosystem: Power BI Desktop, Power BI Service i Power BI Mobile. Użytkownik rozpoczyna pracę zwykle od aplikacji Power BI Desktop, gdzie może pobrać dane z różnych źródeł (np. pliki Excel, bazy SQL, chmury typu Snowflake), oczyścić je i przygotować za pomocą wbudowanego narzędzia Power Query, a następnie stworzyć model danych i raporty wizualne. Gotowe raporty można następnie opublikować w Power BI Service, czyli chmurowej platformie, gdzie można nimi zarządzać, udostępniać je innym użytkownikom i automatycznie odświeżać dane. Power BI Mobile umożliwia przeglądanie raportów na smartfonach i tabletach, co sprawdza się szczególnie w pracy zdalnej lub podczas spotkań.

Power BI wyróżnia się prostotą obsługi – tworzenie wizualizacji odbywa się metodą przeciągnij i upuść, a gotowe raporty są w pełni interaktywne. Użytkownik może filtrować dane, zaznaczać elementy wykresów i analizować szczegóły bez potrzeby znajomości języków programowania. Co więcej, Power BI integruje się bezproblemowo z innymi produktami Microsoft, takimi jak Excel, Teams, SharePoint czy Azure, a także z wieloma zewnętrznymi bazami danych i usługami.

Główne zalety:

- Intuicyjny interfejs i łatwe tworzenie wizualizacji
- Szeroka gama źródeł danych lokalnych, chmurowych, baz danych, plików, API
- Interaktywne raporty i dashboardy, dostępne także z poziomu przeglądarki i aplikacji mobilnej
- Automatyczne odświeżanie danych i harmonogramy aktualizacji
- Wysoki poziom integracji z Microsoft Azure, Snowflake i innymi platformami analitycznymi
- Zaawansowane funkcje modelowania danych i język DAX do obliczeń
- Możliwość współpracy zespołowej i dzielenia się raportami w organizacji

Dla kogo?

Power BI jest narzędziem dla analityków danych, menedżerów, pracowników działów sprzedaży, marketingu i finansów, a także wszystkich, którzy chcą w prosty sposób analizować i prezentować dane. Jest idealny zarówno dla małych firm, które chcą zacząć pracować z danymi, jak i dla dużych organizacji, które potrzebują rozbudowanych i zautomatyzowanych raportów biznesowych. Dzięki dostępności bezpłatnej wersji Power BI Desktop oraz chmurowej wersji z elastycznymi planami, narzędzie to jest szeroko dostępne i skalowalne w zależności od potrzeb.

1.2.4. Dlaczego wybrano te technologie?

W projekcie wykorzystywane są Snowflake, Microsoft Azure oraz Power BI, ponieważ razem tworzą elastyczne i nowoczesne środowisko do analizy danych sprzedażowych. Snowflake to chmurowa hurtownia danych, która pozwala na szybkie i skalowalne przetwarzanie dużych zbiorów danych za pomocą SQL. Dzięki rozdzieleniu przechowywania i mocy obliczeniowej jest bardziej elastyczna niż tradycyjne bazy danych.

Microsoft Azure jest wybierany jako platforma chmurowa, ponieważ oferuje stabilne środowisko, wysoką dostępność i natywną integrację z Power BI, co usprawnia połączenie między narzędziami.

Power BI to intuicyjne narzędzie do wizualizacji danych, które umożliwia tworzenie interaktywnych raportów bez potrzeby programowania. Jako produkt Microsoftu, bardzo dobrze współpracuje z Azure i Snowflake, co upraszcza proces analizy i prezentacji danych.

Zestaw tych technologii jest powszechnie wykorzystywany w nowoczesnych organizacjach. Coraz więcej przedsiębiorstw przenosi swoje rozwiązania analityczne do chmury, dlatego znajomość tych narzędzi staje się coraz bardziej ceniona na rynku pracy. Warto zdobywać doświadczenie w ich obsłudze, ponieważ mogą stanowić istotny element pracy analityka.

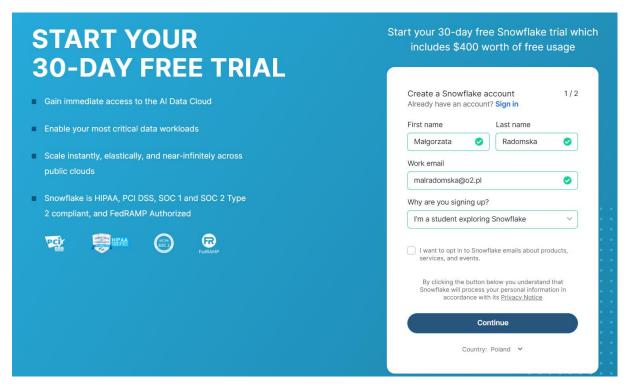
2. Implementacja projektu

2.1. Utworzenie konta Snowflake

Na pierwszym etapie rejestracji użytkownik zakłada darmowe konto w Snowflake (https://www.snowflake.com/en/), które umożliwia korzystanie z platformy przez 30 dni w ramach bezpłatnego trialu. W tym czasie dostępne jest do wykorzystania darmowych zasobów obliczeniowych o wartości 400 dolarów.

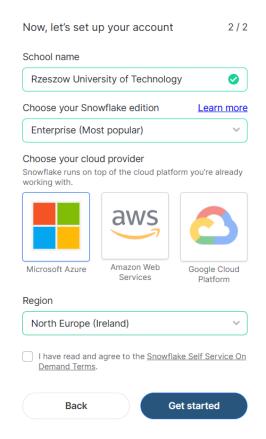
Po upływie 30 dni lub wyczerpaniu limitu kredytów konto zostaje zablokowane do czasu wykupienia płatnego planu, jednak dane przechowywane w Snowflake nie są usuwane – użytkownik może je odzyskać po przejściu na wersję płatną. Snowflake działa w modelu pay-as-you-go, co oznacza, że płacimy tylko za realnie wykorzystane zasoby – np. za uruchomienie zapytania SQL, przetwarzanie danych czy pracę wirtualnego magazynu (warehouse).

Na formularzu należy podać imię i nazwisko, adres e-mail (na który zostanie wysłany link aktywacyjny) oraz odpowiedzieć na pytanie, w jakim celu zakładane jest konto – w tym przypadku wybrano opcję, która odzwierciedla akademicki charakter projektu.

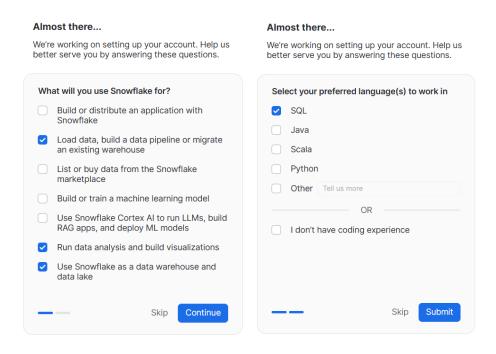


W kolejnym etapie użytkownik podaje nazwę uczelni oraz wybiera edycję. Tryb Enterprise oferuje m.in. możliwość cofania się do wersji danych do 90 dni wstecz, możliwość pracy na wielu magazynach danych jednocześnie oraz wsparcie dla materializowanych widoków, co czyni go idealnym do testów i analizy danych.

Jako chmurę wybrano Microsoft Azure, ponieważ Power BI – narzędzie używane w projekcie do wizualizacji – również pochodzi od Microsoftu. Dzięki temu integracja Snowflake z Power BI jest szybka i bezproblemowa. Region ustawiono na North Europe (Ireland), aby zapewnić dobrą wydajność i zgodność z europejskimi regulacjami (RODO).



Na poniższym etapie użytkownik określa, do czego będzie wykorzystywać Snowflake. W projekcie wybrano opcje związane z ładowaniem danych, analizą oraz wykorzystaniem platformy jako hurtowni danych. W kolejnym kroku zaznaczono SQL jako preferowany język pracy – jest to podstawowy język zapytań wykorzystywany w Snowflake. Po kliknięciu Submit konto jest gotowe do użycia.



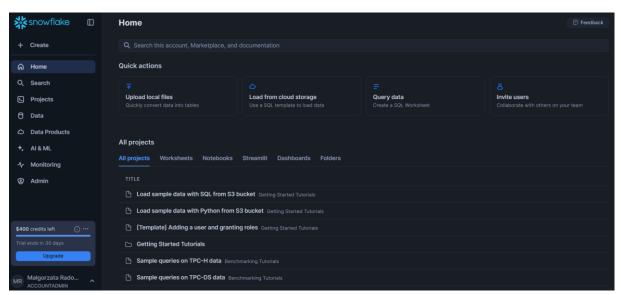
Po zalogowaniu do konta Snowflake użytkownik trafia na ekran główny, który umożliwia szybki dostęp do najważniejszych funkcji platformy. W centralnej części widoczne są tzw. Quick actions, czyli skróty do najczęstszych działań:

- Upload local files wgrywanie danych z pliku (np. CSV),
- Load from cloud storage ładowanie danych z chmury,
- Query data uruchamianie zapytań SQL,
- Invite users zapraszanie innych użytkowników do współpracy.

Po lewej stronie znajduje się menu nawigacyjne, z którego można przejść m.in. do:

- Create tworzenie zasobów,
- Projects zarządzanie projektami i notatkami SQL,
- Data przegląd i zarządzanie bazami danych,
- Monitoring śledzenie wykorzystania zasobów,
- Admin zarządzanie kontem i uprawnieniami.

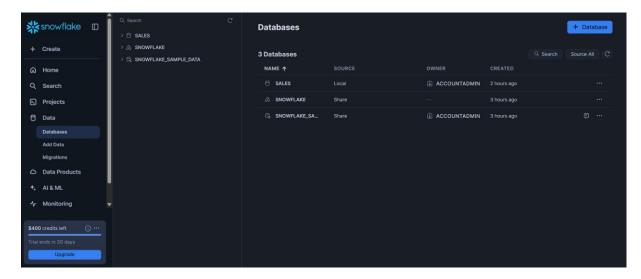
Na dole widać też ilość pozostałych darmowych kredytów (\$400) oraz czas pozostały do końca triala (30 dni).



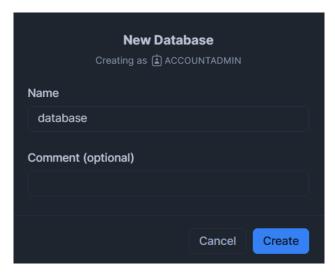
2.2. Rozpoczęcie pracy w Snowflake

2.2.1. Tworzenie nowej bazy danych

Tworzenie nowej bazy danych w Snowflake jest bardzo proste. Wystarczy przejść do zakładki Data > Databases, a następnie kliknąć przycisk + Database w prawym górnym rogu.



Na ekranie tworzenia bazy użytkownik nadaje jej nazwę i opcjonalnie może dodać komentarz. Po kliknięciu Create baza zostaje utworzona i jest gotowa do dalszej pracy.

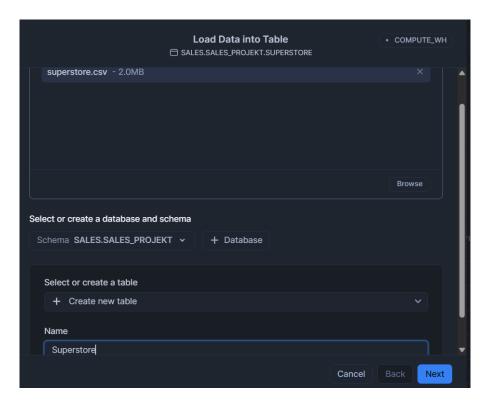


Analogicznie, po kliknięciu utworzonej bazy danych, można dodać nowy schemat, który porządkuje strukturę danych w ramach jednej bazy.

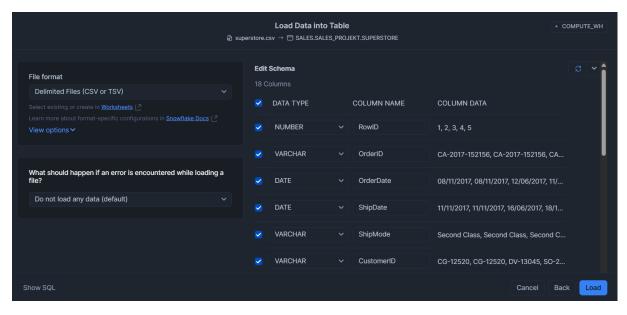
Dzięki temu etapowi środowisko Snowflake jest przygotowane do importu danych z pliku CSV i dalszej analizy SQL.

2.2.2. Import danych

Po wybraniu opcji Upload local files z ekranu głównego otwiera się okno, w którym możemy załadować plik z danymi bezpośrednio z komputera. W kolejnym kroku przypisujemy dane do konkretnej bazy danych i schematu, a następnie wybieramy istniejącą tabelę lub tworzymy nową, nadając jej nazwę.



Snowflake automatycznie rozpoznaje typy danych na podstawie zawartości pliku, ale użytkownik ma możliwość ich modyfikacji. Dodatkowo dostępne są zaawansowane ustawienia (po rozwinięciu listy), takie jak ustawienie formatu daty, separatora kolumn czy sposób obsługi błędów.



Przed zatwierdzeniem importu można kliknąć Show SQL (lewy dolny róg), aby podejrzeć kod SQL, który zostanie wykonany – zawiera on wszystkie wybrane ręcznie ustawienia, w tym strukturę tabeli, format pliku oraz komendę COPY INTO.

```
Load Data into Table

1 CREATE TABLE "SALES", "SALES, PROJEKT", "Superstore" ( RowID NUMBER(38, 0) , OrderID VARCHAR , OrderDate DATE , ShipDate DATE , ShipMode VARCHAR , CustomerID VARCHAR , CustomerID VARCHAR , CustomerID VARCHAR , Segment VARCHAR , Country VARCHAR , City VARCHAR , ProductIName VARCHAR , Segment VARCHAR , SubCategory VARCHAR , ProductIName VARCHAR , ProductIName VARCHAR , SubCategory VARCHAR , ProductIName VARCHAR , Sales NUMBER(38, 4) );

2 CREATE TEMP FILE FORMAT "SALES". "SALES_PROJEKT". "temp_file_format_2025-05-05118:20:46.7972"

4 TYPE=CSV

5 SKIP_HEADER-1

6 FIELD_DELIMITER",

7 TRIM_SPACE_TRUE

8 FIELD_OPTIONALIT_ENCLOSED_BY="""

9 REPLACE_INVALID_CHARACTERS=TRUE

10 DATE_FORMAT-AUITO

11 TIME_FORMAT-AUITO

12 TIMESTAMP_FORMAT=AUITO

13 FANCESS" "SALES_PROJEKT" "Superstore"

14 FORM (SELECT SI. 52, 53, 54, 55, 57, 58, 59, 510, 511, 512, 513, 514, 515, 516, 517, $18

15 FORM (SELECT SI. 52, 53, 54, 55, 57, 58, 59, 510, 511, 512, 513, 514, 515, 516, 517, $18

16 FORM "SALES" "SALES_PROJEKT" "superstore"

17 FILES ("2025-05-05118:1242-7322/Superstore.csv")

18 FILE_FORMAT = ""SALES" "SALES_PROJEKT". "temp_file_format_2025-05-05118:20:46.7972"

19 ON_ERROR-ABORT_STATEMENT

10 MERROR-ABORT_STATEMENT

10 MERROR-ABORT_STATEMENT

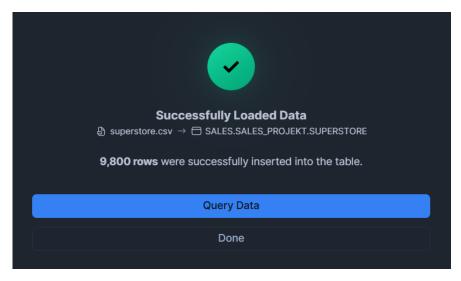
11 SALES "SALES_PROJEKT". "temp_file_format_2025-05-05118:20:46.7972"

10 NERROR-ABORT_STATEMENT

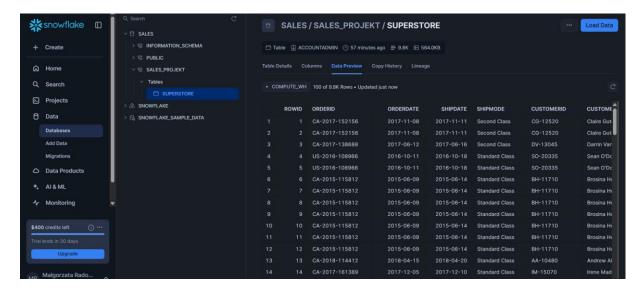
10 MERROR-ABORT_STATEMENT

10 MERROR-ABORT_STATEMENT
```

Po kliknięciu Load, Snowflake informuje o sukcesie lub ewentualnych błędach importu. Jeśli wszystko przebiegło pomyślnie, użytkownik widzi komunikat z liczbą załadowanych wierszy.



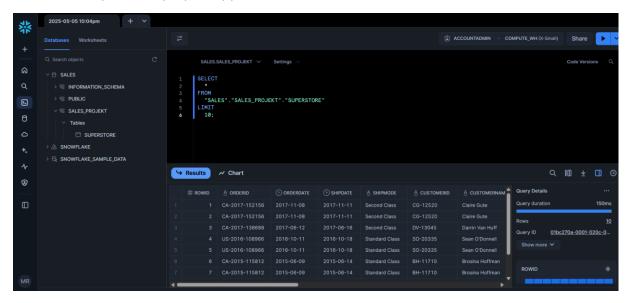
Po wybraniu opcji Done zostajemy automatycznie przeniesieni do zakładki Data > Databases, gdzie w panelu po lewej stronie można rozwinąć listę baz danych, schematów i tabel. Dla każdej tabeli dostępny jest podgląd danych oraz szczegóły techniczne.



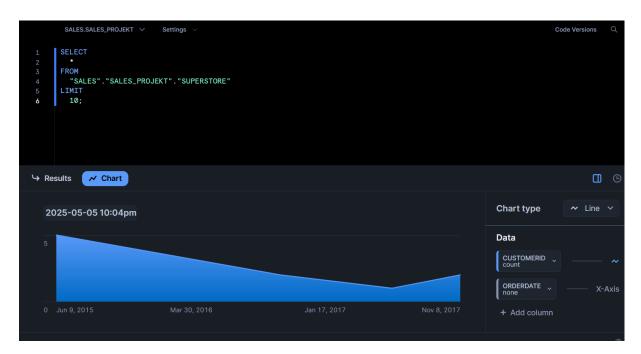
2.2.3. Tworzenie zapytań

Aby rozpocząć pracę z językiem SQL w Snowflake, należy przejść do zakładki Projects > Worksheets. To właśnie tutaj możemy tworzyć nowe skoroszyty robocze, w których zapisujemy i uruchamiamy zapytania SQL. Na górze wybieramy bazę danych, schemat oraz warehouse, z którego mają korzystać zapytania.

W skoroszycie wpisujemy dowolne kwerendy, a po ich uruchomieniu otrzymujemy wyniki w dolnym panelu. Dodatkowo, po prawej stronie wyświetlane są statystyki zapytania, takie jak czas wykonania, liczba zwróconych wierszy czy ID zapytania.



Snowflake umożliwia również tworzenie wykresów bezpośrednio z wyników – wystarczy kliknąć zakładkę Chart, a następnie wybrać kolumny, które chcemy zobrazować (np. liczbę klientów w czasie). To szybki sposób na wizualną analizę danych bez konieczności korzystania z zewnętrznych narzędzi.



Wyniki zapytań można zapisać w historii, pobrać jako plik CSV, skopiować do schowka, a także udostępnić innym użytkownikom Snowflake w ramach współpracy.

2.2.4. Tworzenie widoków

Aby ułatwić i uporządkować późniejszą integrację Snowflake z Power BI, w projekcie zostały utworzone dedykowane widoki SQL (views). Widoki te pełnią rolę wstępnie przefiltrowanych i zagregowanych zestawów danych, które można traktować jak gotowe tabele. Dzięki nim Power BI nie musi samodzielnie przetwarzać dużych ilości danych, co znacząco poprawia wydajność, przejrzystość i elastyczność dashboardu.

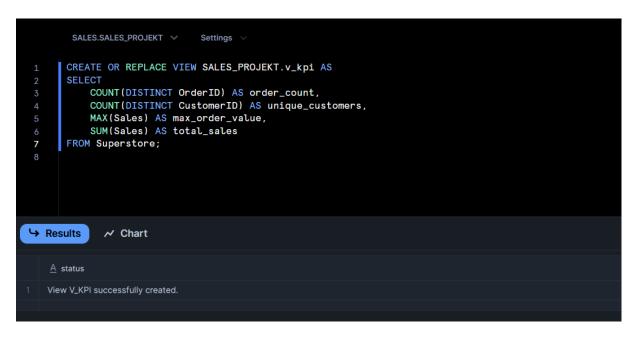
Dzięki wykorzystaniu widoków:

- zmniejsza się obciążenie Power BI, ponieważ przetwarzanie danych odbywa się po stronie Snowflake,
- łatwiej utrzymać spójność danych, ponieważ widoki można aktualizować centralnie,
- użytkownik końcowy otrzymuje przejrzyste i gotowe źródła danych, bez potrzeby tworzenia skomplikowanych zapytań w Power BI.

Aby utworzyć widok w Snowflake, należy przejść do zakładki Worksheets i upewnić się, że wybrana jest odpowiednia baza danych, schemat oraz warehouse. Następnie w edytorze zapytań wpisujemy polecenie CREATE OR REPLACE VIEW, które tworzy widok o określonej nazwie i zawartości.

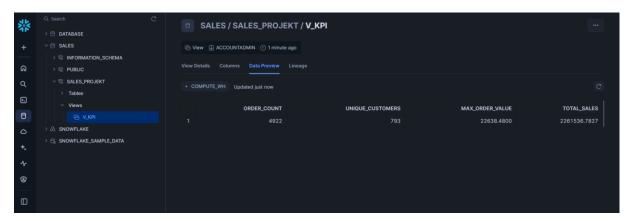
Na przykład w przedstawionym przypadku utworzono widok SALES_PROJEKT.v_kpi, który zapisuje najważniejsze informacje biznesowe, takie jak liczba zamówień, liczba klientów, maksymalna wartość zamówienia i całkowita sprzedaż. Po uruchomieniu zapytania w dolnym panelu pojawia się informacja o poprawnym utworzeniu widoku: "View V_KPI successfully created".

Taki widok działa jak tabela i może być wykorzystywany w kolejnych zapytaniach lub eksportowany do Power BI.



Aby potwierdzić, że widok został prawidłowo utworzony, możemy przejść do zakładki Data w panelu po lewej stronie. Po rozwinięciu odpowiedniej bazy danych i schematu pojawią się dwie osobne sekcje: Tables i Views. W sekcji Views znajdziemy utworzony wcześniej widok – w tym przypadku V_KPI.

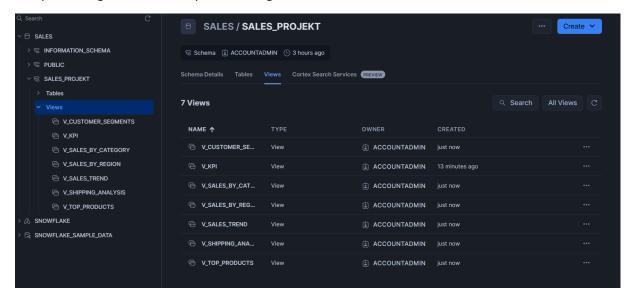
Po kliknięciu na nazwę widoku mamy dostęp do zakładki Data Preview, która działa analogicznie jak w przypadku tabel. Umożliwia ona szybki podgląd danych zwracanych przez widok, co pozwala zweryfikować poprawność zapytania SQL bez konieczności pisania dodatkowych kwerend.



Utworzono kolejne widoki, które możemy zauważyć w sekcji Views w ramach schematu SALES_PROJEKT. Są to:

- V_KPI zestawienie kluczowych wskaźników (liczba zamówień, klienci, suma i maksymalna wartość sprzedaży),
- V_SALES_TREND sprzedaż zagregowana miesięcznie, umożliwiająca analizę trendów,
- V_SALES_BY_REGION zestawienie sprzedaży oraz liczby zamówień według regionów,
- V SALES BY CATEGORY sprzedaż podzielona na kategorie i podkategorie produktów,
- V TOP PRODUCTS dziesięć najlepiej sprzedających się produktów,
- V CUSTOMER SEGMENTS analiza segmentów klientów oraz średnich wartości zamówień,
- V_SHIPPING_ANALYSIS analiza metod dostawy i średniego czasu wysyłki.

Widoki te posłużą jako źródła danych do dalszej analizy wizualnej w Power BI, umożliwiając budowę kompleksowego dashboardu sprzedażowego.



Snowflake to elastyczna i wydajna platforma do przechowywania oraz przetwarzania danych, która oferuje szerokie możliwości analityczne, m.in. dzięki obsłudze języka SQL, tworzeniu widoków i łatwej integracji z narzędziami BI. Dotychczasowe działania pozwoliły przygotować struktury danych i widoki potrzebne do dalszej analizy.

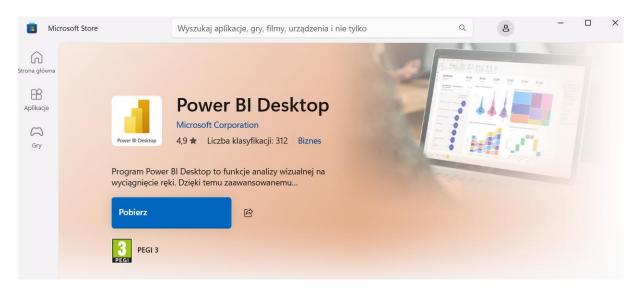
W kolejnej części projektu zajmiemy się integracją Snowflake z Power BI oraz budową interaktywnego raportu z wizualizacją kluczowych informacji sprzedażowych.

2.3. Instalacja Power BI Desktop

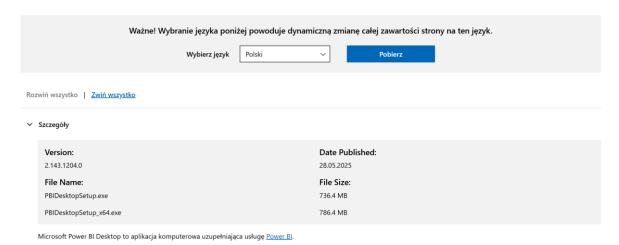
W celu instalacji Power BI Desktop, czyli darmowej wersji oprogramowania, należy wejść na stronę Microsoftu pod linkiem https://www.microsoft.com/pl-pl/power-platform/products/power-bi/desktop.



Wybranie opcji *Pobierz teraz* skutkuje przeniesieniem do Microsoft Store, skąd można łatwo pobrać aplikację na komputer.



Natomiast po wyborze przycisku *Zaawansowane opcje pobierania* pojawia się możliwość pobrania pliku z instalatorem.

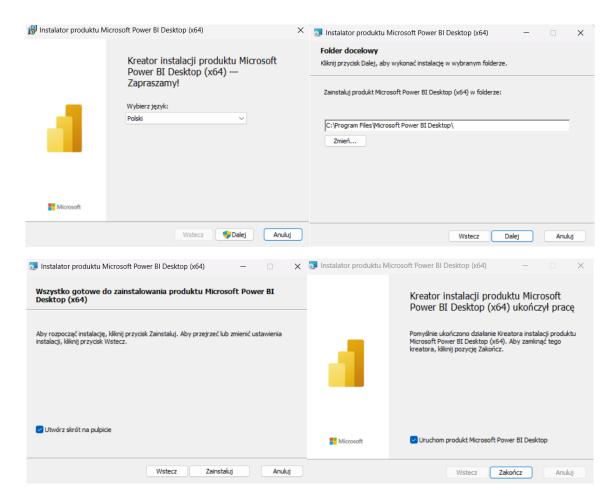


Po wyborze języka aplikacji i kliknięciu przycisku Pobierz należy wybrać plik instalacyjny.

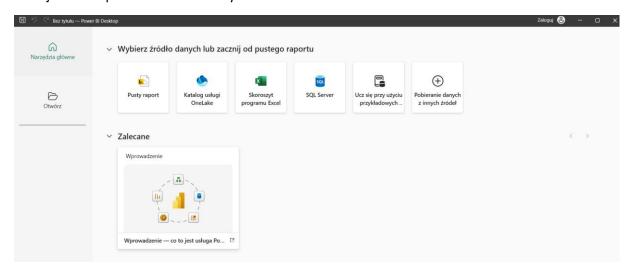


W przypadku, gdy komputer obsługuje aplikacje 64-bitowe, należy wybrać zaznaczony powyżej plik.

Po pobraniu pliku otwiera się instalator, gdzie w kolejnych krokach aplikacja jest dostosowywana pod użytkownika. Wybierany jest język, ścieżka oraz czy ma zostać utworzony skrót na pulpicie.

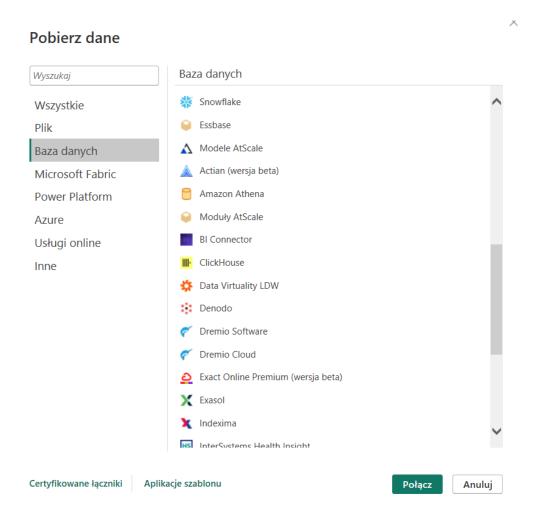


Po pomyślnej instalacji Power BI Desktop otwiera się strona aplikacji, gdzie możemy rozpocząć tworzenie nowego raportu, połączyć się ze źródłem danych lub otworzyć istniejący już plik. Microsoft oferuje także wprowadzenie do korzystania z Power BI.



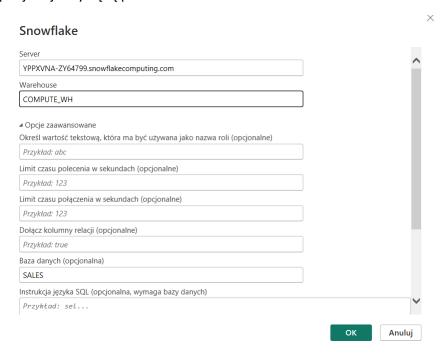
2.4. Integracja Snowflake z Power BI

Power BI umożliwia łączenie się z różnymi źródłami danych, np. OneLake, skoroszyt Excela czy SQL Server. W celu pobrania danych ze Snowflake wybieramy opcję *Pobieranie danych z innych źródeł*, po czym otwiera się okno z szerokim wyborem źródeł różnych kategorii.

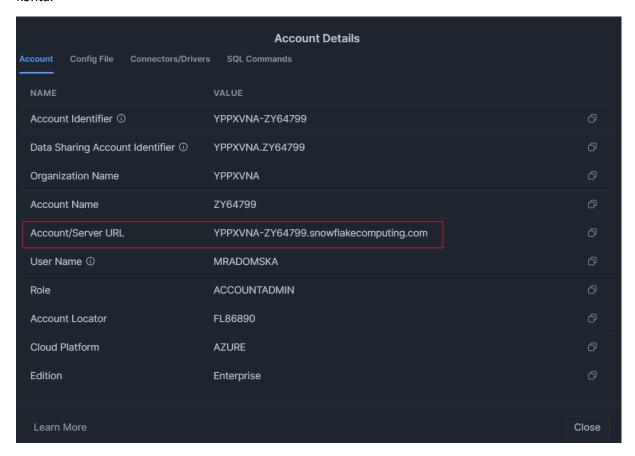


Snowflake można odnaleźć w kategorii Baza danych. Wybieramy tę opcję i klikamy Połącz.

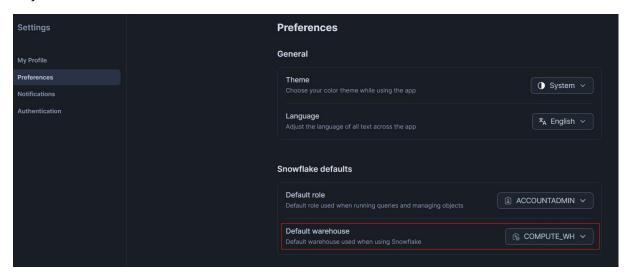
W kolejnym kroku Power BI prosi o podanie nazwy serwera i hurtowni danych oraz opcjonalne informacje, np. z jakiej bazy będą pobierane dane.



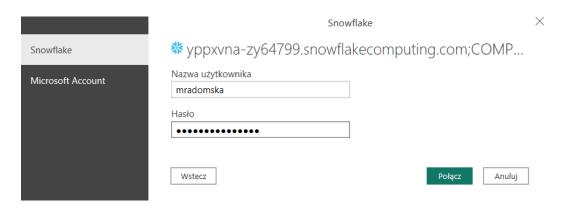
Potrzebne informacje można znaleźć na koncie Snowflake. Nazwę serwera można znaleźć w szczegółach konta.



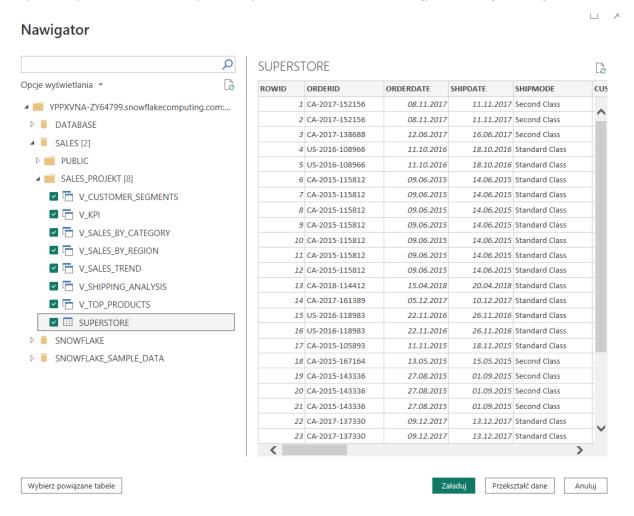
Natomiast informacja o wykorzystywanej hurtowni znajduje się w ustawieniach w zakładce *Preferences*.



W następnym etapie należy zalogować się danymi do konta Snowflake.



Po pomyślnym procesie logowania pojawia się okno, w którym możemy wybrać bazę danych oraz konkretne tabele, które chcemy załadować do Power BI. Można zauważyć, że oprócz tabeli SUPERSTORE ze wszystkimi danymi, wyświetlają się też widoki utworzone wcześniej w Snowflake. Aby móc stworzyć raport w oparciu o dane, należy zaznaczyć tabele oraz widoki, a następnie wcisnąć *Załaduj*.



Następnie Power BI prosi o wybór trybu pracy ze źródłem danych. Przy wyborze *Importuj* dane są pobierane i zapisywane w pliku, co przyspiesza działanie raportu, jednak przy ewentualnych zmianach dane należy odświeżać ręcznie. Natomiast wybór DirectQuery zapewnia wysyłanie zapytań do Snowflake, co oznacza, że dane zawsze są aktualne. Może to jednak spowalniać pracę. W przypadku gdy tworzymy raport na zamkniętym zestawie danych najlepszą opcją będzie wybranie *Importuj*. Jeśli jednak raport powinien uwzględniać na żywo zmiany w bazie danych, lepszym wyborem będzie *DirectQuery*. Ponieważ projekt opiera się na gotowym zestawie danych, wybrano pierwszą opcję.

Ustawienia połączenia

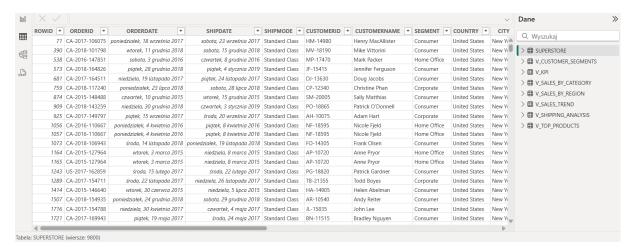
Możesz wybrać sposób łączenia się z tym źródłem danych. Importowanie pozwala załadować kopię danych do usługi Power BI. Zapytanie DirectQuery będzie łączyć się na żywo z tym źródłem danych.

ImportujDirectQuery

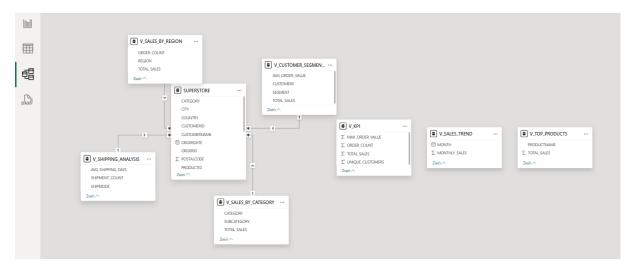
Dowiedz się więcej o zapytaniach DirectQuery



Po pomyślnym załadowaniu danych mamy możliwość podejrzenia widoku tabelarycznego dla tabeli Superstore oraz każdego z widoków.

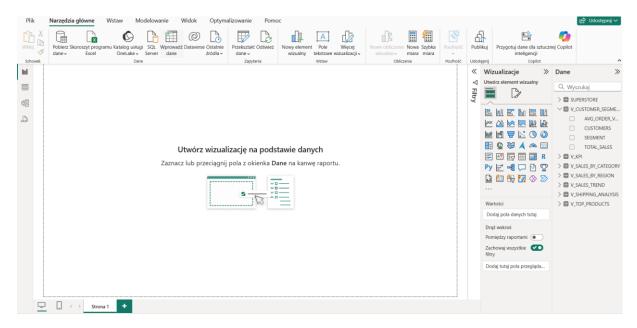


W zakładce poniżej można z kolei zobaczyć model danych. Tabele mogą być ze sobą połączone relacjami, jeśli są w jakiś sposób od siebie zależne. Mogą też występować osobno.



2.5. Tworzenie raportu

Aby stworzyć raport mamy do dyspozycji pustą planszę, na której można dowolnie ustawiać elementy.



W bocznym pasku mamy do wyboru różnego rodzaju wizualizacje, np.:

- Wykresy słupkowe,
- Wykres liniowy,
- Wykresy warstwowe,
- Wykres punktowy,
- Wykresy kołowe,
- Mapę drzewa,
- Mapę,
- Kartogram,
- Miernik,
- Kluczowy wskaźnik wydajności,
- Tabele,
- Macierz i inne.

Każda z wizualizacji opiera się na danych widocznych po prawej stronie. Wybrane kolumny możemy przeciągać na odpowiednie pola wizualizacji.

2.5.1. Tworzenie wizualizacji

W aplikacji Power BI stworzono interaktywny dashboard prezentujący kluczowe informacje sprzedażowe. Wykorzystano różne typy wizualizacji, odpowiednio dopasowane do charakteru danych:

Karty

- o Widok: V KPI
- Przedstawiają najważniejsze wskaźniki (KPI): całkowita sprzedaż, liczba zamówień, liczba unikalnych klientów oraz największa wartość pojedynczego zamówienia.

Wykres kolumnowy

- Widok: V_SALES_BY_REGION
- o Pokazuje łączną wartość sprzedaży w podziale na regiony (West, East, Central, South).

Wykres kolumnowy poziomy

Widok: V_TOP_PRODUCTS

o Prezentuje najlepiej sprzedające się produkty według wartości sprzedaży.

Wykres pierścieniowy

Widok: V_SALES_BY_CATEGORY

o Pokazuje: udział procentowy sprzedaży dla kategorii produktów (Technology, Furniture, Office Supplies), z wyświetleniem konkretnych wartości w tysiącach.

Wykres kołowy

Widok: V_CUSTOMER_SEGMENTS

 Przedstawia strukturę sprzedaży według segmentu klienta (Consumer, Corporate, Home Office), wraz z udziałem procentowym i wartościami.

Wykres liniowy

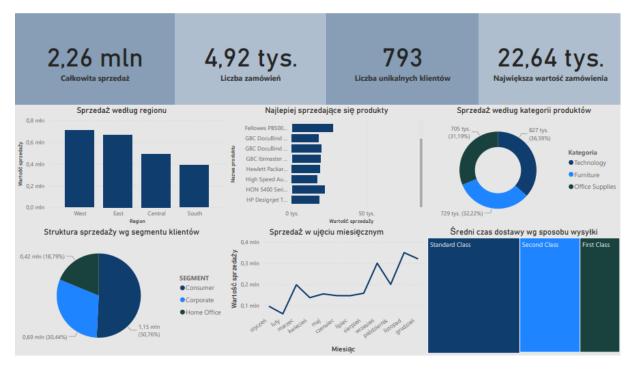
Widok: V_SALES_TREND

 Ilustruje zmiany wartości sprzedaży w ujęciu miesięcznym, umożliwiając identyfikację sezonowości i trendów.

• Wizualizacja typu "Treemap"

Widok: V_SHIPPING_ANALYSIS

 Przedstawia średni czas dostawy w zależności od wybranego sposobu wysyłki (Standard Class, Second Class, First Class).



Tworząc dashboard, wykorzystano także funkcje Power BI pozwalające na interaktywność między wykresami oraz przejrzyste formatowanie danych. Wybór odpowiednich typów wizualizacji,

zastosowanie relacji między widokami oraz możliwość dynamicznego filtrowania informacji pozwoliły na stworzenie czytelnego i funkcjonalnego raportu analitycznego.

Gotowy dashboard stworzony w Power BI można w łatwy sposób udostępniać i eksportować. W wersji Power BI Desktop dostępne są następujące opcje:

- Eksport do pliku PDF umożliwia zapisanie statycznego raportu w formie dokumentu, który można dołączyć do prezentacji lub dokumentacji.
- Zapis jako plik .pbix pozwala przechować raport lokalnie na komputerze i otworzyć go ponownie w Power BI Desktop.
- Publikacja online (wymaga konta Power BI w chmurze Microsoft) umożliwia publikację raportu w Power BI Service, gdzie można go udostępniać współpracownikom, osadzać na stronie internetowej lub utworzyć dashboard interaktywny dostępny online.
- Tworzenie subskrypcji lub alertów (w Power BI Service) pozwala automatycznie wysyłać raporty e-mailem lub ustawiać powiadomienia w zależności od wartości danych.

Dzięki tym funkcjom raport może być nie tylko narzędziem do analizy, ale także skutecznym sposobem komunikacji wyników biznesowych z innymi użytkownikami – zarówno technicznymi, jak i nietechnicznymi.

3. Podsumowanie

W ramach projektu przeprowadzono kompleksową analizę danych sprzedażowych z wykorzystaniem platformy Snowflake oraz narzędzia Power BI. Cały proces rozpoczął się od utworzenia struktury bazy danych w Snowflake, do której następnie zaimportowano dane sprzedażowe w formacie CSV. Dane zostały uporządkowane i przygotowane do dalszej analizy poprzez tworzenie widoków SQL, które w przejrzysty sposób grupowały istotne informacje, takie jak wartości sprzedaży według regionu, kategorii, segmentu klienta czy czasu.

Kolejnym etapem była integracja Snowflake z Power BI, gdzie utworzono interaktywny dashboard prezentujący kluczowe wskaźniki KPI, m.in. całkowitą wartość sprzedaży, liczbę zamówień, liczbę unikalnych klientów oraz maksymalną wartość pojedynczego zamówienia. Na wykresach przedstawiono również najlepiej sprzedające się produkty, zmienność sprzedaży w czasie oraz różnice w średnim czasie dostawy w zależności od wybranego sposobu wysyłki. Zastosowane wizualizacje ułatwiły zrozumienie danych i wyciąganie praktycznych wniosków.

Projekt pokazał, jak duży potencjał niesie ze sobą połączenie chmurowej hurtowni danych z nowoczesnymi narzędziami wizualizacji. Taki model pracy pozwala na elastyczne, interaktywne podejście do analizy danych, wspierające podejmowanie decyzji opartych na rzetelnych informacjach.