



Szkolenie autoryzowane

DP 203T00 Inżynieria danych na platformie Microsoft Azure[Strona szkolenia](#) | [Terminy szkolenia](#) | [Rejestracja na szkolenie](#) | [Promocje](#)

Opis szkolenia

Podczas tego szkolenia uczestnicy zapoznają się z wzorcami i praktykami inżynierii danych w trybie wsadowym i czasie rzeczywistym, w zakresie pracy z rozwiązaniami analitycznymi przy użyciu technologii platformy danych Azure.

Trener przedstawi, jak interaktywnie eksplorować zawartości przechowywane w plikach oraz zaprezentuje różne techniki, których można użyć do ładowania danych przy użyciu funkcji Apache Spark w usłudze Azure Synapse Analytics lub Azure Databricks oraz pozyskiwać dane przy użyciu Azure Data Factory lub Azure Synapse. Omówione też zostaną różne sposoby przekształcania danych przy użyciu tych samych technologii, które są używane do pozyskiwania danych. Kończąc kurs, uczestnik będzie wiedział jak ważne jest wdrożenie zabezpieczeń, aby zapewnić ochronę danych w spoczynku lub podczas przesyłania plików. Trener pokaże też w jaki sposób pliki w systemie analitycznym, można wykorzystać do tworzenia pulpitów nawigacyjnych lub budowania modeli predykcyjnych w usłudze Azure Synapse Analytics.

Wymagania:

Oprócz odpowiedniego doświadczenia zawodowego osoby, które chcą wziąć udział w tym szkoleniu, powinny mieć wiedzę techniczną w zakresie objętym tematyką szkolenia:

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

Podstawy platformy Azure.

Mile widziane: **AZ 900T00 Podstawy platformy Microsoft Azure** oraz **DP 900T00 Podstawy przetwarzania danych na platformie Microsoft Azure**

Umiejętności po szkoleniu:

- Poznanie możliwości obliczeniowych i pamięci masowej dla operacji związanych z inżynierią danych w Azure
- Projektowanie i implementacja warstwy obsługowej
- Zrozumienie zagadnień związanych z inżynierią danych
- Wprowadzenie interaktywnych zapytań przy użyciu bezserwerowej puli SQL
- Eksploracja, transformacja i ładowanie danych do Hurtowni Danych przy użyciu Apache Spark
- Przeprowadzanie eksploracji i transformacji danych w Azure Databricks
- Wprowadzanie i ładowanie danych do Hurtowni Danych
- Przekształcanie danych za pomocą Azure Data Factory lub Azure Synapse Pipelines
- Integracja danych z Notebooków z Azure Data Factory lub Azure Synapse Pipelines
- Optymalizacja wydajności zapytań dzięki dedykowanym pulom SQL w Azure Synapse
- Analizowanie i optymalizacja przechowywania w Hurtowni Danych
- Obsługa hybrydowego transakcyjnego przetwarzania analitycznego (HTAP) za pomocą Azure Synapse Link
- Zapewnienie bezpieczeństwa end-to-end za pomocą Azure Synapse Analytics
- Przeprowadzanie przetwarzania strumieniowego w czasie rzeczywistym za pomocą funkcji Stream Analytics

Harmonogram szkolenia

Moduł 1: Poznanie możliwości obliczeniowych i magazynowania dla procedur roboczych związanych z inżynierią danych

Moduł ten zapewnia przegląd opcji technologii obliczeniowej i przechowywania danych Azure, które są dostępne dla inżynierów danych budujących procedury analityczne. Moduł ten uczy sposobów strukturyzacji Data Lake oraz optymalizacji plików dla procedur eksploracji, strumieniowania i wsadowych. Studenci dowiedzą się jak zorganizować jezioro danych na poziomach przetwarzania danych, jak przekształcać pliki poprzez przetwarzanie wsadowe i strumieniowe. Następnie nauczą się jak tworzyć indeksy na swoich zbiorach danych, takich jak pliki CSV, JSON i Parquet, i używać ich do potencjalnego przyspieszenia

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

zapytań i procedur roboczych.

Teoria:

- Wprowadzenie do usługi Azure Synapse Analytics
- Opis Azure Databricks
- Wprowadzenie do pamięci masowej Azure Data Lake
- Opis architektury Delta Lake
- Praca ze strumieniami danych przy użyciu Azure Stream Analytics

Praktyka:

- Poznanie możliwości obliczeniowych i magazynowania dla procedur roboczych związanych z inżynierią danych
- Połączenie przetwarzania strumieniowego i wsadowego w jednym potoku.
- Zorganizowanie jeziora danych na poziomach transformacji plików
- Indeksowanie pamięci masowej jeziora danych w celu przyspieszenia zapytań i procedur roboczych.

Po ukończeniu tego modułu studenci będą potrafili:

- Opisać Azure Synapse Analytics
- Opisać Azure Databricks
- Opisać pamięć masową Azure Data Lake
- Opisać architekturę Delta Lake
- Opisać Azure Stream Analytics

Moduł 2: Projektowanie i implementacja warstwy serwerów

Moduł ten uczy jak projektować i implementować magazyny danych w nowoczesnej hurtowni danych w celu optymalizacji pracy analitycznej. Student dowie się jak zaprojektować wielowymiarowy schemat do przechowywania danych na temat faktów i wymiarów. Następnie dowie się jak zapełnić wolno zmieniające się wymiary poprzez przyrostowe ładowanie danych z Azure Data Factory.

Teoria:

- Projektowanie wielowymiarowego schematu w celu optymalizacji obciążeń analitycznych
- Bezkodowa transformacja w skali z Azure Data Factory
- Wypełnianie wolno zmieniających się wymiarów w potokach Azure Synapse Analytics

Praktyka:

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Projektowanie i implementacja warstwy obsługowych
- Projektowanie schematu gwiazd dla analitycznych procedur roboczych
- Wypełnianie wolno zmieniających się wymiarów za pomocą Azure Data Factory i mapowanie przepływów danych

Po ukończeniu tego modułu studenci będą potrafili:

- Projektować schemat gwiazdy dla obciążeń analitycznych
- Wypełniać powoli zmieniające się wymiary za pomocą Azure Data Factory i mapować przepływy danych

Moduł 3: Kwestie dotyczące inżynierii danych dla plików źródłowych

Moduł ten omawia zagadnienia inżynierii danych, które są powszechne podczas ładowania danych do nowoczesnej hurtowni danych analitycznych z plików przechowywanych w Azure Data Lake, a także zrozumienie względów bezpieczeństwa związanych z przechowywaniem plików przechowywanych w jeziorze danych.

Teoria:

- Projektowanie nowoczesnej hurtowni danych z wykorzystaniem Azure Synapse Analytics
- Zabezpieczanie hurtowni danych w Azure Synapse Analytics

Praktyka:

- Kwestie dotyczące inżynierii danych
- Zarządzanie plikami w jeziorze danych Azure
- Zabezpieczanie plików przechowywanych w jeziorze danych Azure

Po ukończeniu tego modułu studenci będą potrafili:

- Projektować nowoczesną hurtownię danych z wykorzystaniem Azure Synapse Analytics
- Zabezpieczać hurtownię danych w Azure Synapse Analytics

Moduł 4: Uruchamianie interaktywnych zapytań przy użyciu bezserwerowych puli SQL Azure Synapse Analytics

W tym module studenci dowiedzą się, jak pracować z plikami przechowywanymi w Data Lake i zewnętrznych źródłach plików, poprzez instrukcje T-SQL wykonywane przez bezserwerową pulę SQL w Azure Synapse Analytics. Studenci będą generować zapytania do plików Parquet przechowywanych w Data Lake, jak również do plików CSV przechowywanych

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

w zewnętrznym magazynie danych. Następnie, stworzą grupy bezpieczeństwa Azure Active Directory i wymuszą dostęp do plików w Data Lake poprzez kontrolę dostępu opartą na rolach (Role-Based Access Control - RBAC) i listy kontroli dostępu (Access Control Lists - ACL).

Teoria:

- Poznawanie możliwości bezserwerowych puli SQL w Azure Synapse
- Zapytywanie o dane w jeziorze przy użyciu bezserwerowych puli SQL Azure Synapse
- Tworzenie obiektów metadanych w bezserwerowych pulach SQL Azure Synapse
- Zabezpieczanie danych i zarządzanie użytkownikami w bezserwerowych pulach SQL Azure Synapse

Praktyka:

- Uruchamianie interaktywnych zapytań przy użyciu bezserwerowych puli SQL
- Zapytania o dane Parquet za pomocą bezserwerowych puli SQL
- Tworzenie zewnętrznych tabel dla plików Parquet i CSV
- Tworzenie widoków za pomocą bezserwerowych puli SQL
- Bezpieczny dostęp do danych w jeziorze danych przy użyciu bezserwerowych puli SQL
- Konfiguracja zabezpieczeń jeziora danych przy użyciu kontroli dostępu opartej na rolach (RBAC) i list kontroli dostępu

Po ukończeniu tego modułu studenci będą potrafili:

- Zrozumieć możliwości bezserwerowych puli SQL w Azure Synapse
- Zapytywać o Data Lake przy użyciu bezserwerowych puli SQL Azure Synapse
- Tworzyć obiekty metadanych w bezserwerowych pulach SQL Azure Synapse
- Zabezpieczać dane i zarządzać użytkownikami w bezserwerowych pulach SQL Azure Synapse

Moduł 5: Eksploracja, transformacja i ładowanie danych do Hurtowni Danych przy użyciu Apache Spark

Moduł ten uczy, jak eksplorować dane przechowywane w Data Lake, przekształcać dane i ładować dane do relacyjnego magazynu danych. Student poznaje pliki Parquet i JSON oraz techniki zapytań i transformacji plików JSON o strukturze hierarchicznej. Następnie student wykorzysta Apache Spark do załadowania danych do hurtowni danych i połączenia danych Parquet w Data Lake z danymi w dedykowanej puli SQL.

Teoria:

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Zrozumienie inżynierii Big Data z Apache Spark w Azure Synapse Analytics
- Pobieranie danych za pomocą notebooków Apache Spark w Azure Synapse Analytics
- Przekształcanie danych za pomocą DataFrames w pulach Apache Spark w Azure Synapse Analytics
- Integracja puli SQL i Apache Spark w Azure Synapse Analytics

Praktyka:

- Eksploracja, transformacja i ładowanie danych do Hurtowni Danych przy użyciu Apache Spark
- Przeprowadzanie eksploracji danych w programie Synapse Studio
- Pobieranie danych za pomocą notebooków Spark w Azure Synapse Analytics
- Przekształcanie danych za pomocą DataFrames w pulach Spark w Azure Synapse Analytics
- Integracja puli SQL i Spark w Azure Synapse Analytics

Po ukończeniu tego modułu studenci będą potrafili:

- Opisać inżynierię Big Data z Apache Spark w Azure Synapse Analytics
- Pobierać dane za pomocą notebooków Apache Spark w Azure Synapse Analytics
- Przekształcać dane za pomocą DataFrames w pulach Apache Spark w Azure Synapse Analytics
- Integracja puli SQL i Apache Spark w Azure Synapse Analytics

Moduł 6: Eksploracja i transformacja danych w Azure Databricks

Moduł ten uczy, jak używać różnych metod Apache Spark DataFrame do eksploracji i transformacji danych w Azure Databricks. Studenci nauczą się jak wykonywać standardowe metody DataFrame do eksploracji i transformacji danych. Dowiedzą się również, jak wykonywać bardziej zaawansowane zadania, takie jak usuwanie duplikatów danych, manipulowanie wartościami daty/czasu, zmianami nazw kolumn i agregowaniem danych.

Teoria:

- Opisywane Azure Databricks
- Odczyt i zapis danych w Azure Databricks
- Praca z DataFrames w Azure Databricks
- Praca z zaawansowanymi metodami DataFrames w Azure Databricks

Praktyka:

- Eksploracja i transformacja danych w Azure Databricks
- Używanie DataFrames w Azure Databricks do eksploracji i filtrowania danych
- Buforowanie DataFrame w celu przyspieszenia kolejnych zapytań
- Usuwanie zduplikowanych danych

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Manipulowanie wartościami daty/czasu
- Usuwanie i zmiana nazw kolumn DataFrame
- Agregowanie danych przechowywanych w DataFrame

Po ukończeniu tego modułu studenci będą potrafili:

- Opisać Azure Databricks
- Odczytywać i zapisywać dane w Azure Databricks
- Pracować z DataFrames w Azure Databricks
- Pracować z zaawansowanymi metodami DataFrames w Azure Databricks

Moduł 7: Wprowadzanie i ładowanie danych do hurtowni danych

Moduł ten uczy studentów jak wprowadzać dane do hurtowni danych za pomocą skryptów T-SQL oraz strumieni integracyjnych Synapse Analytics. Student dowie się jak ładować dane do dedykowanych puli SQL Synapse za pomocą PolyBase i COPY z wykorzystaniem T-SQL. Dowie się również jak wykorzystać zarządzanie obciążeniem wraz z aktywnością Copy w potoku Azure Synapse do ładowania danych w skali petabajtowej.

Teoria

- Wykorzystywanie najlepszych praktyk ładowania danych w Azure Synapse Analytics
- Wprowadzanie danych na skalę petabajtów za pomocą Azure Data Factory

Praktyka:

- Pobieranie i ładowanie danych do Hurtowni Danych
- Wykonywanie pobierania danych w skali petabajtów za pomocą potoków Azure Synapse Pipelines
- Import danych z PolyBase i COPY przy użyciu T-SQL
- Wykorzystywanie najlepszych praktyk ładowania danych w Azure Synapse Analytics

Po ukończeniu tego modułu studenci będą potrafili:

- Wykorzystywać najlepsze praktyki ładowania danych w Azure Synapse Analytics
- Pobierać dane na skalę petabajtów za pomocą Azure Data Factory

Moduł 8: Przekształcanie danych za pomocą Azure Data Factory lub Azure Synapse Pipelines

Moduł ten uczy studentów jak budować strumienie integracji danych w celu pobierania

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

danych z wielu źródeł, przekształcania danych przy użyciu mapowania przepływów danych i wykonywania przepływu danych do jednego lub więcej źródeł danych.

Teoria:

- Integracja danych za pomocą Azure Data Factory lub Azure Synapse Pipelines
- Bezkodowa transformacja w skali z Azure Data Factory lub Azure Synapse Pipelines

Praktyka:

- Przekształcanie danych za pomocą Azure Data Factory lub Azure Synapse Pipelines
- Wykonywanie bezkodowej transformacji w skali dzięki Azure Synapse Pipelines.
- Tworzenie potoku danych do importu źle sformatowanych plików CSV
- Tworzenie mapowania przepływów danych

Po ukończeniu tego modułu studenci będą potrafili:

- Wykonywać integrację danych za pomocą Azure Data Factory
- Wykonywać bezkodową transformację w skali dzięki Azure Data Factory

Moduł 9: Orkiestracja przepływu i transformacji danych w Azure Synapse Pipelines

W tym module dowiesz się, jak tworzyć połączone usługi oraz orkiestrację przepływu i transformacji danych przy użyciu notebooków w Azure Synapse Pipelines.

Teoria:

- Orkiestracja przepływu i transformacji danych w Azure Data Factory

Praktyka:

- Orkiestracja przepływu i transformacji danych w Azure Synapse Pipelines
- Integracja danych z notebooków z Azure Data Factory lub Azure Synapse Pipelines

Po ukończeniu tego modułu studenci będą potrafili:

- Orkiestrować przepływ i transformację danych w Azure Synapse Pipelines

Moduł 10: Optymalizacja wydajności zapytań dzięki dedykowanym pulom SQL w Azure Synapse

W tym module studenci poznają strategie optymalizacji przechowywania i przetwarzania

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

danych podczas korzystania z dedykowanych pul SQL w Azure Synapse Analytics. Student będzie wiedział, jak korzystać z funkcji deweloperskich, takich jak okienkowanie i funkcje HyperLogLog, stosować najlepsze praktyki ładowania danych oraz optymalizować i poprawiać wydajność zapytań.

Teoria:

- Optymalizacja wydajności zapytań do hurtowni danych w Azure Synapse Analytics
- Zrozumienie funkcji dewelopera hurtowni danych w Azure Synapse Analytics

Praktyka:

- Optymalizacja wydajności zapytań dzięki dedykowanym pulom SQL w Azure Synapse
- Zrozumienie funkcji deweloperskich oprogramowania Azure Synapse Analytics
- Optymalizacja wydajności zapytań do hurtowni danych w Azure Synapse Analytics
- Zwiększenie wydajności zapytań

Po ukończeniu tego modułu studenci będą potrafili:

- Optymalizować wydajności zapytań do hurtowni danych w Azure Synapse Analytics
- Rozumieć funkcje dewelopera hurtowni danych w Azure Synapse Analytics

Moduł 11: Analiza i optymalizacja przechowywania w hurtowni danych

W tym module studenci dowiedzą się jak analizować i optymalizować przechowywanie danych w dedykowanych pulach SQL Azure Synapse. Studenci poznają techniki pozwalające zrozumieć wykorzystanie przestrzeni tabel oraz szczegóły dotyczące przechowywania danych w kolumnach. Następnie student dowie się jak porównać wymagania dotyczące przechowywania danych pomiędzy identycznymi tabelami, które używają różnych typów danych. Wreszcie, student będzie obserwował wpływ widoków zmaterializowanych na wykonywanie złożonych zapytań oraz dowie się, jak uniknąć rozległego rejestrowania poprzez optymalizację operacji usuwania danych.

Teoria:

- Analizowanie i optymalizacja magazynowania danych w Azure Synapse Analytics

Praktyka:

- Analiza i optymalizacja przechowywania hurtowni danych

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Sprawdzanie zniekształconych danych i wykorzystania przestrzeni
- Zrozumienie szczegółów przechowywania danych w magazynie kolumnowym
- Badanie wpływu widoków zmaterializowanych
- Badanie zasady dotyczącej operacji z minimalną liczbą rejestracji

Po ukończeniu tego modułu studenci będą potrafili:

- Analizować i optymalizować magazynowanie danych w Azure Synapse Analytics

Moduł 12: Obsługa hybrydowego transakcyjnego przetwarzania analitycznego (Hybrid Transactional Analytical Processing - HTAP) za pomocą Azure Synapse Link

W tym module, studenci dowiedzą się jak Azure Synapse Link umożliwia bezproblemowe połączenie konta Azure Cosmos DB z przestrzenią roboczą Synapse. Student zrozumie jak włączyć i skonfigurować Synapse Link, a następnie jak wykonywać zapytania do magazynu analitycznego Azure Cosmos DB używając Apache Spark i SQL serverless.

Teoria:

- Projektowanie hybrydowego przetwarzania transakcyjnego i analitycznego przy użyciu Azure Synapse Analytics
- Konfiguracja Azure Synapse Link z Azure Cosmos DB
- Zapytania do Azure Cosmos DB za pomocą pul Apache Spark
- Zapytania do Azure Cosmos DB za pomocą bezserwerowych pul SQL

Praktyka:

- Obsługa hybrydowego transakcyjnego przetwarzania analitycznego (HTAP) za pomocą Azure Synapse Link
- Konfiguracja Azure Synapse Link z Azure Cosmos DB
- Zapytania do Azure Cosmos DB za pomocą Apache Spark dla Synapse Analytics
- Zapytania do Azure Cosmos DB za pomocą bezserwerowej puli SQL dla Azure Synapse Analytics

Po ukończeniu tego modułu studenci będą potrafili:

- Projektować hybrydowe przetwarzanie transakcyjne i analityczne przy użyciu Azure Synapse Analytics
- Konfigurować Azure Synapse Link z Azure Cosmos DB
- Przesyłać zapytania do Azure Cosmos DB za pomocą Apache Spark dla Azure Synapse Analytics
- Przesyłać zapytania do Azure Cosmos DB za pomocą bezserwerowej puli SQL dla Azure Synapse Analytics

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

Moduł 13: Bezpieczeństwo end-to-end z Azure Synapse Analytics

W tym module studenci dowiedzą się, jak zabezpieczyć przestrzeń roboczą Synapse Analytics i jej infrastrukturę wspierającą. Student będzie obserwował administratora SQL Active Directory, zarządzał regułami zapory IP, zarządzał sekretami z Azure Key Vault i uzyskiwał dostęp do tych sekretów poprzez usługę powiązaną z Key Vault i działania potokowe. Student będzie rozumiał jak zaimplementować bezpieczeństwo na poziomie kolumn, wierszy oraz dynamiczne maskowanie danych przy użyciu dedykowanych pul SQL.

Teoria:

- Zabezpieczanie hurtowni danych w Azure Synapse Analytics
- Konfiguracja i zarządzanie sekretami w Azure Key Vault
- Wdrożenie kontroli zgodności dla danych wrażliwych

Praktyka:

- Bezpieczeństwo end-to-end z Azure Synapse Analytics
- Bezpieczna infrastruktura wspierająca Azure Synapse Analytics
- Zabezpieczanie przestrzeni roboczej Azure Synapse Analytics i usług zarządzanych
- Zabezpieczanie dane przestrzeni roboczej Azure Synapse Analytics

Po ukończeniu tego modułu studenci będą potrafili:

- Zabezpieczać hurtownię danych w Azure Synapse Analytics
- Konfigurować i zarządzać sekretami w Azure Key Vault
- Wdrażać kontrolę zgodności dla danych wrażliwych

Moduł 14: Przetwarzanie strumieniowe w czasie rzeczywistym z wykorzystaniem Stream Analytics

W tym module studenci dowiedzą się, jak przetwarzać dane strumieniowe za pomocą Azure Stream Analytics. Będą wprowadzać dane telemetryczne pojazdów do Event Hubs, a następnie przetwarzać te dane w czasie rzeczywistym, używając różnych funkcji okienkowania w Azure Stream Analytics. Następnie, wyślą dane do Azure Synapse Analytics. Wreszcie, dowiedzą się, jak skalować zadanie Stream Analytics, aby zwiększyć przepustowość.

Teoria:

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Zapewnienie niezawodnego przesyłania wiadomości dla aplikacji Big Data za pomocą Azure Event Hubs
- Praca z danymi strumieniowymi przy użyciu Azure Stream Analytics
- Analiza strumieni danych za pomocą Azure Stream Analytics

Praktyka:

- Przetwarzanie strumieniowe w czasie rzeczywistym za pomocą Stream Analytics
- Używanie analizy strumienia do przetwarzania danych w czasie rzeczywistym z koncentratorów zdarzeń.
- Używanie funkcji okienkowania Stream Analytics do tworzenia agregatów i przesyłania danych do Synapse Analytics.
- Skalowanie zadania Azure Stream Analytics w celu zwiększenia przepustowości poprzez partycjonowanie
- Podział strumienia wejściowego w celu optymalizacji paralelizacji

Po ukończeniu tego modułu studenci będą potrafili:

- Zapewnić niezawodne przesyłanie wiadomości dla aplikacji Big Data za pomocą Azure Event Hubs
- Pracować z danymi strumieniowymi przy użyciu Azure Stream Analytics
- Analizować strumień danych za pomocą Azure Stream Analytics

Moduł 15: Tworzenie rozwiązań do przetwarzania strumieniowego przy użyciu Event Hubs i Azure Databricks

W tym module studenci dowiedzą się, jak pobierać i przetwarzać dane strumieniowe w skali za pomocą Event Hubs i Spark Structured Streaming w Azure Databricks. Studenci poznają kluczowe cechy i zastosowania Structured Streaming. Student zaimplementuje przesuwane okna do agregacji fragmentów danych oraz zastosuje znak wodny do usuwania starych danych. W końcu, student połączy się z Event Hubs w celu odczytu i zapisu strumieni.

Teoria:

- Przetwarzanie danych strumieniowych za pomocą strukturyzowanego strumieniowania Azure Databricks

Praktyka:

- Tworzenie rozwiązania do przetwarzania strumieniowego przy użyciu Event Hubs i Azure Databricks
- Poznanie kluczowych cechy i zastosowanie strumieniowania strukturalnego

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

- Strumieniowe przesyłanie danych z pliku i zapisywanie ich do rozproszonego systemu plików
- Używanie przesuwanego okna, aby agregować fragmenty danych, a nie wszystkie dane.
- Stosowanie znaku wodnego, aby usunąć nieaktualne dane
- Podłączenie do strumieni odczytu i zapisu Event Hubs

Po ukończeniu tego modułu studenci będą potrafili:

- Przetwarzanie danych strumieniowych za pomocą strukturyzowanego strumieniowania Azure Databricks.

Moduł 16: Tworzenie raportów przy użyciu integracji Power BI z Azure Synapse Analytics

W tym module student dowie się, jak zintegrować Power BI z obszarem roboczym Synapse w celu tworzenia raportów w Power BI. Utworzy nowe źródło danych i raport Power BI w Synapse Studio. Następnie dowie się, jak poprawić wydajność zapytań za pomocą widoków zmaterializowanych i buforowania zbioru wyników. Wreszcie, student eksploruje jezioro danych za pomocą bezserwerowych puli SQL i stworzy wizualizacje na podstawie tych danych w Power BI.

Teoria:

- Tworzenie raportów w Power BI przy użyciu integracji z Azure Synapse Analytics

Praktyka:

- Tworzenie raportów przy użyciu integracji Power BI z Azure Synapse Analytics
- Zintegrowanie przestrzeni roboczej Synapse Azure z Power BI
- Optymalizacja integracji z Power BI
- Zwiększenie wydajności zapytań dzięki widokom zmaterializowanym i buforowaniu zbioru wyników
- Wizualizacja danych za pomocą SQL serverless i tworzenie raportów Power BI

Po ukończeniu tego modułu studenci będą potrafili:

- Tworzenie raportów w Power BI przy użyciu integracji z Azure Synapse Analytics

Moduł 17: Wykonywanie zintegrowanych procesów uczenia maszynowego w Azure Synapse Analytics

Ten moduł omawia zintegrowane, kompletne doświadczenie Azure Machine Learning i Azure

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.

Cognitive Services w Azure Synapse Analytics. Dowiesz się, jak połączyć przestrzeń roboczą Azure Synapse Analytics z przestrzenią roboczą Azure Machine Learning za pomocą usługi połączonej, a następnie uruchomisz zautomatyzowany eksperyment ML, który wykorzystuje dane z tabeli Spark. Dowiesz się również, jak używać wyszkolonych modeli z Azure Machine Learning lub Azure Cognitive Services do wzbogacania danych w tabeli puli SQL, a następnie obsługiwać wyniki predykcji za pomocą Power BI.

Teoria:

- Wykorzystywanie zintegrowanego procesu uczenia maszynowego w Azure Synapse Analytics

Praktyka:

- Wykonywanie zintegrowanych procesów uczenia maszynowego w Azure Synapse Analytics
- Tworzenie usługi powiązanej z uczeniem maszynowym Azure
- Uruchamianie eksperymentu Auto ML przy użyciu danych z tabeli Spark
- Wzbogacanie danych za pomocą wyszkolonych modeli
- Podawanie wyników predykcji za pomocą Power BI

Po ukończeniu tego modułu studenci będą potrafili:

- Wykorzystywanie zintegrowanego procesu uczenia maszynowego w Azure Synapse Analytics

Tagi:

Adres korespondencyjny:

DAGMA Szkolenia IT | ul. Bażantów 6a/3 | Katowice (40-668)
tel. 32 793 11 80 | szkolenia@dagma.pl
szkolenia.dagma.eu

DAGMA Sp. z o.o. z siedzibą w Katowicach (40-478), ul. Pszczyńska 15
Sąd Rejonowy Katowice-Wschód w Katowicach Wydział VIII Gospodarczy
KRS pod numerem 0000130206, kapitał zakładowy 75 000 zł
Numer NIP 634-012-60-68, numer REGON: 008173852
DAGMA Sp. z o.o. posiada status dużego przedsiębiorcy
w rozumieniu art. 4c ustawy o przeciwdziałaniu nadmiernym
opóźnieniom w transakcjach handlowych.