

コースコード：MS-DP-203T00

税込価格：286,000円 (税抜価格：260,000円)

日数：4日間

トレーニング内容

Azureデータプラットフォームテクノロジーを使用したバッチおよびリアルタイムの分析ソリューションの操作に関連するデータエンジニアリングのパターンと実践について学習します。受講者は、分析ソリューションの構築に使用されるコアコンピューティングおよびストレージテクノロジーを理解することから始めます。次に、分析サービングレイヤーを設計する方法を探り、ソースファイルを操作するためのデータエンジニアリングの考慮事項に焦点を当てます。受講者は、データレイク内のファイルに保存されているデータをインタラクティブに探索する方法を学びます。彼らは、Azure Synapse AnalyticsまたはAzure DatabricksにあるApache Spark機能を使用してデータを読み込むために使用できるさまざまな取り込み手法、またはAzure Data FactoryまたはAzure Synapseパイプラインを使用して取り込む方法を学習して頂きます。受講者はまた、データの取り込みに使用されるのと同じテクノロジーを使用してデータを変換するさまざまな方法を学びます。受講者は、データロードまたはシステムに対して発行されるクエリのパフォーマンスを最適化できるように、分析システムのパフォーマンスをモニタリングおよび分析する方法を学習するコースに時間を費やします。彼らは、データが保管中または転送中に保護されることを保証するためにセキュリティを実装することの重要性を理解します。次に、分析システムのデータを使用してダッシュボードを作成したり、Azure Synapse Analyticsで予測モデルを構築したりする方法を学習して頂きます。

ここに注目!!

ワンポイントアドバイス

DP-203: Data Engineering on Microsoft Azure の受験予定の方に最適です。

受講対象者

このコースの受講対象者は次の通りです。

このコースの対象者は、Microsoft Azureに存在するデータプラットフォームテクノロジーを使用してデータエンジニアリングと分析ソリューションを構築する方法について学びたいデータプロフェッショナル、データアーキテクト、およびビジネスインテリジェンスプロフェッショナルです。このコースの二次対象者は、Microsoft Azure上に構築された分析ソリューションを使用するデータアナリストおよびデータサイエンティストです。

前提条件

このコースを受講する前に受講者が習得しておく必要がある知識およびスキルは次のとおりです。

成功する受講者は、クラウド コンピューティングとコアデータに関する知識とソリューションに関する専門的な経験を持った状態でこのコースを開始します。事前に次のコース受講を推奨します。

- AZ-900 - Azure Fundamentals
- DP-900 - Microsoft Azure Data Fundamentals

目的

このコースを修了すると次のことができるようになります。

- Azureのデータエンジニアリングワークロードのコンピューティングとストレージのオプションを調べる
- サービングレイヤーの設計と実装
- データエンジニアリングの考慮事項を理解する
- サーバーレスSQLプールを使用してインタラクティブクエリを実行する
- Apache Sparkを使用して、データを探索、変換、およびデータウェアハウスにロードします
- Azure Databricksでデータの探索と変換を実行する
- データを取り込み、データウェアハウスにロードします
- Azure DataFactoryまたはAzure Synapseパイプラインを使用してデータを変換する
- ノートブックのデータをAzure Data FactoryまたはAzure Synapseパイプラインと統合する
- Azure Synapseの専用SQLプールを使用してクエリパフォーマンスを最適化する
- データウェアハウスストレージの分析と最適化
- Azure Synapse Linkを使用したHybrid Transactional Analytical Processing (HTAP) のサポート
- Azure Synapse Analyticsを使用してエンドツーエンドのセキュリティを実行します
- StreamAnalyticsを使用してリアルタイムのストリーム処理を実行します
- Event HubsとAzure Databricksを使用してストリーム処理ソリューションを作成する
- Power BIとAzure Synapse Analyticsの統合を使用してレポートを作成する
- Azure Synapse Analyticsで統合された機械学習プロセスを実行する

アウトライン

■学習内容

1. データエンジニアリングワークロードのコンピューティングおよびストレージオプションの調査
 - Azure Synapse Analyticsの概要
 - Azure Databricksについての説明
 - Azure Data Lakeストレージの概要
 - Delta Lakeのアーキテクチャについての説明
 - Azure Stream Analyticsを使用してデータストリームの操作
2. サービングレイヤーを設計および実装
 - 分析ワークロードを最適化するための多次元スキーマの設計
 - Azure Data Factoryを使用した大規模なコードフリー変換
 - Azure Synapse Analyticsパイプラインにゆっくりと変化するディメンションの入力
3. ソースファイルのデータエンジニアリングに関する考慮事項
 - Azure Synapse Analyticsを使用して最新のデータウェアハウスの設計
 - Azure Synapse Analyticsでデータウェアハウスの保証
4. Azure Synapse AnalyticsサーバーレスSQLプールを使用してインタラクティブクエリを実行
 - Azure SynapseサーバーレスSQLプール機能の調査
 - Azure SynapseサーバーレスSQLプールを使用して湖のデータのクエリ
 - Azure SynapseサーバーレスSQLプールにメタデータオブジェクトの作成
 - Azure SynapseサーバーレスSQLプールでデータを保護してユーザーの管理
5. Apache Sparkを使用して、データを探索、変換、およびデータウェアハウスにロード

- Azure Synapse AnalyticsのApache Sparkを使用したビッグデータエンジニアリングの理解
- Azure Synapse AnalyticsでApache Sparkノートブックを使用してデータ取り込み
- Azure Synapse AnalyticsのApache SparkプールのData Framesを使用してデータの変換
- SQLプールとApache SparkプールをAzure Synapse Analyticsに統合

6. Azure Databricksでのデータの探索と変換

- Azure Databricksについての説明
- Azure Databricksでのデータの読み取りと書き込み
- Azure DatabricksでData Framesの操作
- Azure DatabricksでData Framesの高度なメソッドの操作

7. データを取り込んでデータウェアハウスにロード

- Azure Synapse Analyticsでデータ読み込みのベストプラクティスの使用
- Azure Data Factoryによるペタバイト規模の取り込み

8. Azure Data FactoryまたはAzure Synapseパイプラインを使用してデータの変換

- Azure Data FactoryまたはAzure Synapseパイプラインとのデータ統合
- Azure Data FactoryまたはAzure Synapse Pipelinesを使用した大規模なコードフリー変換

9. Azure Synapse Pipelinesでデータの移動と変換を調整

- Azure Data Factoryでデータの移動と変換を調整

10. Azure Synapseの専用SQLプールを使用してクエリのパフォーマンスを最適化

- Azure Synapse Analyticsでデータウェアハウスクエリのパフォーマンスを最適化
- Azure Synapse Analyticsのデータウェアハウス開発者機能の理解

11. データウェアハウストレージの分析と最適化

- Azure Synapse Analyticsでデータウェアハウストレージを分析および最適化

12. Azure Synapse Linkを使用したHybrid Transactional Analytical Processing (HTAP) のサポート

- Azure Synapse Analyticsを使用してトランザクションと分析のハイブリッド処理の設計
- Azure Cosmos DBを使用してAzure Synapse Linkの構成
- Apache Sparkプールを使用してAzure Cosmos DBにクエリの実行
- サーバーレスSQLプールを使用してAzure Cosmos DBにクエリの実行

13. Azure Synapse Analyticsによるエンドツーエンドのセキュリティ

- Azure Synapse Analyticsでデータウェアハウスの保護
- Azure KeyVaultでシークレットを構成および管理
- 機密データのコンプライアンス管理の実装

14. Stream Analyticsを使用したリアルタイムストリーム処理

- Azure Event Hubsを使用して、ビッグデータアプリケーションの信頼性の高いメッセージングの有効化
- Azure Stream Analyticsを使用してデータストリームの操作
- Azure Stream Analyticsを使用してデータストリームの取り込み

15. イベントハブとAzure Databricksを使用してストリーム処理ソリューションを作成する

- 構造化ストリーミングの主な機能と使用法の模索
- ファイルからデータをストリーミングし、分散ファイルシステムに書き出し
- スライディングウィンドウを使用して、かなりの量のデータの集約
- 透かしを適用した古いデータの削除
- Event Hubsの読み取りおよび書き込みストリームに接続

16. Power BIとAzure Synapse Analyticsの統合を使用してレポートの作成

- Azure SynapseワークスペースとPower BI統合
- Power BIとの統合の最適化
- マテリアライズドビューと結果セットキャッシングにより、クエリのパフォーマンスの向上化
- SQLサーバーレスでデータを視覚化し、Power BIレポートの作成

17. Azure Synapse Analyticsで統合された機械学習プロセスの実行 - Azure Synapse Analyticsの統合された機械学習プロセスの使用

■実習/演習環境

- Windows 10
- Microsoft Azure

■ご注意・ご連絡事項

●テキストおよび演習ガイドについて

本コースでは、マイクロソフト社から提供される電子テキストを使用します。

演習はオンラインラボで行います。オンラインラボ内のWindows OSの言語は英語です。

電子テキストのプラットフォームはMicrosoft

Learnを使用します。受講に際して事前にテキスト参照用に Microsoft
アカウントが必要となります。

●演習について

本コースをオンラインLive受講する場合、演習を実行するには以下の条件を満たす必要があります

。困難な場合は事前に主催会社にご相談ください。

- Microsoft Azureポータルサイトの利用
- その他Azureが提供するサービスへのHTTP/HTTPSアクセス
- Azure上の仮想マシンへのリモートデスクトップ接続(RDPプロトコルの利用)

●コースの内容は予告なく変更される可能性があります。