

Projet YDAYS



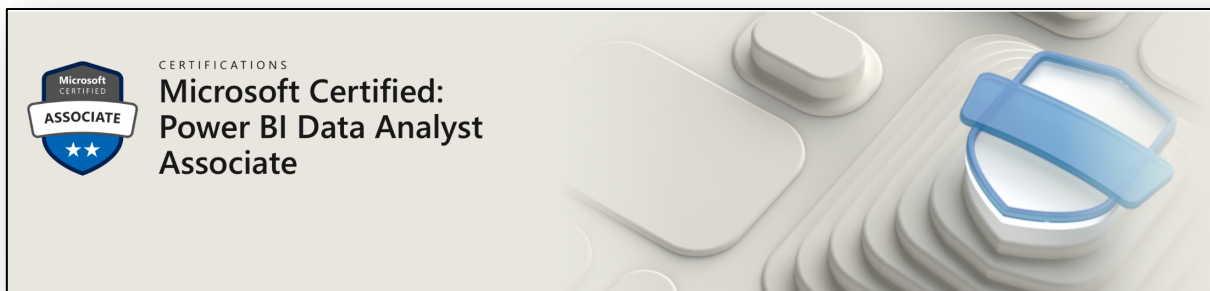
Mastère 1 Data Engineer & Data Scientist

Présenté par : Gérard KOUADIO

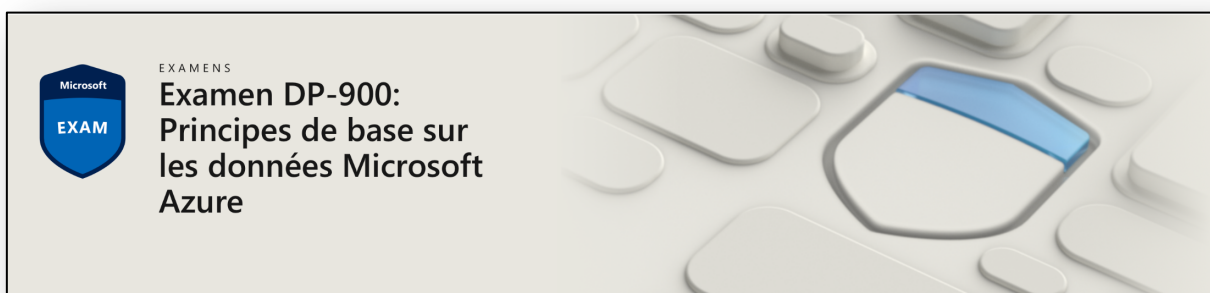
Ydays 1 : 08/11/2023: Choix des sujets

Lors de la première journée de Ydays, nous avons été invités à choisir des sujets sur lesquels nous souhaitons travailler. Me concernant, j'avais repéré des compétences et certifications que je souhaitais avoir avant la fin de ma formation. Mais pour faute de temps, je n'arrivais pas à avancer dans mes projets personnels. J'ai donc saisi cette opportunité en choisissant **le module Certifications**. Selon mon état d'avancement, je suivrai les 2 formations ci-dessous :

Certification 1: Microsoft Certified: Power BI Data Analyst Associate



Certification 2: Examen DP-900: Principes de base sur les données Microsoft Azure



Ydays 2 : 29/11/2023: Démarrage - Microsoft Certified: Power BI Data Analyst Associate

Compétences mesurées

- Préparer les données
- Modéliser les données
- Visualiser et analyser les données
- Déployer et tenir à jour les ressources

L'analyse de données est le processus qui consiste à identifier, nettoyer, transformer et modéliser les données pour révéler des informations pertinentes et utiles. Les données sont ensuite élaborées en histoire sous forme de rapports d'analyse qui aideront les décideurs à prendre des décisions critiques.

Tâches d'un analyste de données

En tant qu'analyste de données, Je dois prendre le temps de pleinement comprendre le problème que l'entreprise tente de résoudre. Déterminer si tous les points de données sont nécessaires, car une trop grande quantité de données peut compliquer la détection des points clés. Raconter une histoire courte et concise peut vous aider à trouver des insights rapidement.

Préparer

Avant la création d'un rapport, les données doivent être préparées. La préparation des données est le processus de profilage, de nettoyage et de transformation de vos données pour qu'elles soient prêtes à être modélisées et visualisées. Préparer les données consiste à prendre des données brutes et à les transformer en informations fiables et compréhensibles. Cela implique, entre autres tâches, de vérifier l'intégrité des données, de corriger les données inappropriées ou inexacts, d'identifier les données manquantes, de convertir les données d'une structure à une autre ou d'un type à un autre ou même tout simplement d'améliorer la lisibilité des données.

Modéliser (ils ont beaucoup insisté sur l'importance de cette partie)

La modélisation des données est le processus qui consiste à déterminer la façon dont vos tables sont liées les unes aux autres. Pour exécuter ce processus, vous définissez et créez des relations entre les tables. À partir de là, vous pouvez améliorer le modèle en configurant des métriques et en ajoutant des calculs personnalisés pour enrichir vos données.

étoileG: Du point de vue de Power BI, si votre rapport s'exécute lentement ou que vos actualisations prennent beaucoup de temps, vous devrez probablement revoir les tâches de préparation et de modélisation des données afin d'optimiser votre rapport.

Visualiser

Analyser

Gérer

Power BI se compose de nombreux éléments, parmi lesquels des rapports, des tableaux de bord, des espaces de travail, des jeux de données, etc. En tant qu'analyste de données, vous êtes responsable de la gestion de ces ressources Power BI : vous supervisez le partage et la distribution d'éléments tels que les rapports et les tableaux de bord, et vous garantissez la sécurité des ressources Power BI.

Rapports

Dans Power BI, un **rapport** est une collection de visualisations sur une ou plusieurs pages.

Tableaux de bord

Quand vous êtes prêt à partager un rapport ou un ensemble de visualisations, vous pouvez créer un **tableau de bord** Power BI. Tout comme le tableau de bord d'une voiture, un tableau de bord est un groupe sélectionné d'objets visuels qui fournissent des informations rapides et importantes sur les données ou l'histoire que vous essayez de présenter.

Limités à une seule page, les tableaux de bord permettent aux utilisateurs de suivre un objet visuel dans le rapport sous-jacent. Les utilisateurs interagissent avec les tableaux de bord au moyen du service Power BI ou sur un appareil mobile.

Vignettes

Dans Power BI, une **vignette** est une visualisation unique figurant dans un tableau de bord. Il s'agit de la zone rectangulaire qui contient un visuel individuel. Dans l'image suivante, vous pouvez voir une vignette, également entourée d'autres vignettes.

Power BI est également scalable ("évolutif", "extensible", "adaptatif" ou "modifiable"), ce qui vous permet de créer des jeux de données à partir de différentes sources de données, même en incorporant du code personnalisé. Le jeu de données peut ensuite être utilisé pour concevoir des rapports et des tableaux de bord interactifs qui mettent l'accent sur des données cruciales pour prendre des décisions métier éclairées.

Application Power BI

Une **application** est un moyen de regrouper des rapports et des tableaux de bord associés et de les distribuer aux audiences appropriées.

NB: Il ne peut y avoir qu'une seule application par espace de travail

Ydays 3 : 20/12/2023: Démarrage - Examen DP-900: Principes de base sur les données Microsoft Azure

Ce certificat contient plusieurs modules :

Module 1 Explore Core Data Concepts in Microsoft Azure.

▼ Introduction to the course

- ✓ Introduction to the program
Video • 4 min
- ✓ Introduction to the course
Video • 1 min
- ✓ Job roles in the world of data
Video • 5 min
- ✓ DP-900: Microsoft Azure AI Fundamentals certification
Video • 2 min
- ✓ Microsoft Certified: Azure Data Fundamentals – Skills Measured
Reading • 10 min
- ✓ Course syllabus
Reading • 10 min
- ✓ How to be successful in this course
Reading • 5 min
- ✓ Microsoft expert interview
Video • 4 min
- 📄 Meet and Greet
Discussion Prompt • 10 min
- ✓ Additional readings
Reading • 10 min

Resume

▼ Explore Core Data Concepts

✓ Complete

- ✓ Lesson Introduction
Video • 1 min
- ✓ Identify the need for data solutions
Video • 10 min
- ✓ Identify types of data and data storage
Video • 9 min
- ✓ Describe the difference between batch and streaming data
Video • 4 min
- ✓ Knowledge check
Practice Quiz • 9 questions
- ✓ Lesson summary
Video • 29 sec

▼ Explore roles and responsibilities in the world of data

✓ Complete

- ✓ Lesson introduction
Video • 45 sec
- ✓ Explore job roles in the world of data
Video • 2 min
- ✓ Review tasks and tools for database administration
Video • 3 min
- ✓ Review tasks and tools for data engineering
Video • 2 min
- ✓ Review tasks and tools for data visualization and reporting
Video • 3 min
- ✓ Knowledge check
Practice Quiz • 8 questions
- ✓ Test prep
Graded Quiz • 7 questions • Grade: 96.42%
- ✓ Lesson summary
Video • 21 sec

▼ Introduction to Microsoft Azure (Optional)

- ✓ Introduction to Azure
Video • 1 min
- ✓ What is cloud computing?
Ungraded Plugin • 3 min
- ✓ The benefits of cloud computing
Video • 1 min
- ✓ What is Azure?
Video • 4 min
- ✓ How does Azure work?
Ungraded Plugin • 3 min
- ✓ Tour of Azure Services
Ungraded Plugin • 3 min
- ✓ List of Azure services
Reading • 10 min
- ✓ Exercise: Get started with Azure accounts
Reading • 10 min
- ✓ Exercise quiz
Practice Quiz • 1 question
- ⏮ Case study introduction
Video • 1 min
- ✓ Knowledge check
Practice Quiz • 10 questions
- ✓ Additional readings

✓ **Cloud Computing Advantages (Optional)**

- ✓ The benefits of cloud computing
Video • 5 min
- ✓ Cloud service models
Video • 1 min
- ✓ Describe the different categories of cloud services
Reading • 10 min
- ✓ Types of cloud service models
Video • 57 sec
- ✓ What are public, private, and hybrid clouds?
Video • 1 min
- How has the cloud affected you?
Discussion Prompt • 10 min
- Knowledge check
Practice Quiz • 6 questions
- ✓ Additional readings
Reading • 10 min

✓ **Concepts of Relational Data**

✓ Complete

- ✓ Lesson Introduction
Video • 1 min
- ✓ The characteristics of relational data
Video • 5 min
- ✓ Explore relational data structures
Video • 5 min
- ✓ Choose the right platform for a relational workload
Video • 8 min
- ✓ Knowledge check
Practice Quiz • 7 questions
- ✓ Lesson summary
Video • 31 sec

✓ Concepts of non-relational data

✓ Complete

- ✓ Lesson Introduction
Video • 1 min
- ✓ Explore the characteristics of non-relational Data
Video • 8 min
- ✓ Explore non-relational data structures
Video • 5 min
- ✓ Choose the right platform for a non-relational workload
Video • 11 min
- ✓ Knowledge check
Practice Quiz • 8 questions
- ✓ Lesson summary
Video • 21 sec
- ✓ Additional readings
Reading • 10 min

✓ Explore concepts of data analytics

✓ Complete

- ✓ Lesson introduction
Video • 1 min
- ✓ Data ingestion and processing
Video • 8 min
- ✓ Data visualization
Video • 5 min
- ✓ Data analytics
Video • 3 min
- ✓ Knowledge check
Practice Quiz • 7 questions
- ✓ Test prep
Graded Quiz • 10 questions • Grade: 94.16%
- ✓ Weekly summary
Video • 16 sec
- ✓ Additional readings
Reading • 10 min

✓ Course Wrap-Up

✓ Complete

- ✓ Course review
Discussion Prompt • 10 min
- ✓ What to expect
Reading • 10 min

<https://coursera.org/share/452de82eac2481ba37a90bca914ea2a7>

