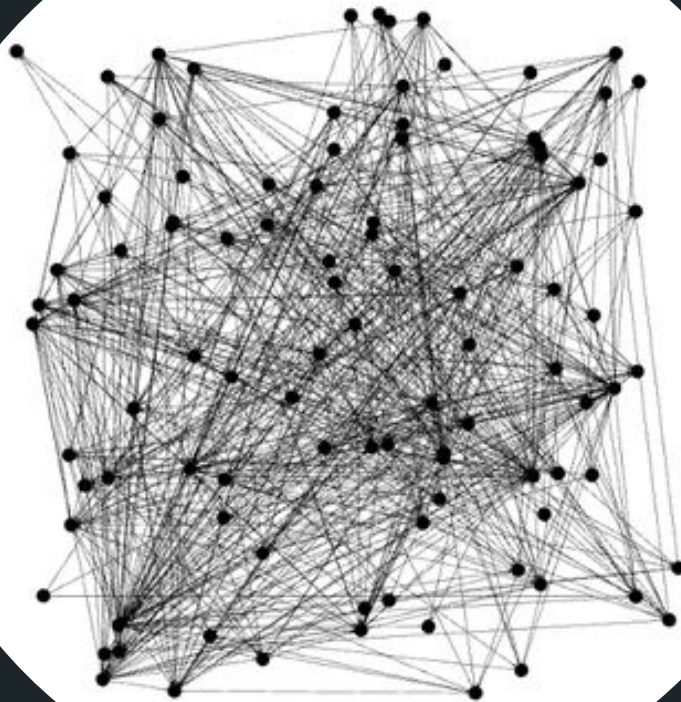


DATA MANAGEMENT & GOVERNANCE COURSE

M2 EFREI 2023 - Final Project



Prepared By:

Doha HAJJOU
Rahma ALBEKBASHY

Consignes du projet :

FINAL PROJECT

Efrei bank is a banking firm with more than 50 years of existence. Originally created by 3 students in final year of finance that still run the company. As of today, the bank have more than 2000 employees. Each entity have a deputy chief executive officer who also is part of the executive board. The company's key branches are in silos :

- Human resources are responsible for managing all activities related to the employees.
- Marketing Services that are leading direction on the development of the mark Efrei Bank, also they should manage all the commercial part from the onboarding's of new clients to the finalization of contract.
- The bank possess also a finance direction that are leading & managing all company's finance and accountability related subjects.
- Finally, there is also an IT direction that oversee all digital and information systems related fields (digital, software, data, infrastructure, websites etc.)

Recently the company's three founding partners decided to change their strategy and the way Efrei bank is using his data. As a bank they propose different services as insurances, loans, investment solutions, retail services etc.

Last year Efrei bank lost more than a thousand of important costumers. Most of them are under 30 years old. Also, they were under their yearly expected growth in term of selling insurances to the seniors, also expending the loans on the real estate field etc.

FINAL PROJECT

Following this observation, the executive board decide to invest in a global transformation to change habits of the company regarding their actual and historical data.

In term of IT organization Efrei bank have :

- An ERP for human resources management from a provider that never share the generated data and how it is organized in their cloud.
- A website that shows all the bank's products also allows online subscriptions to Efrei bank services.
- Finance direction is using Excel files for their analyses and reporting.
- There are also different IT tools for securing the information system, mechanism to manage the existing data as local devices for monitoring.

Efrei bank has recently named a local chief data officer for each entity and an IT Data leader that centralize and is responsible for all data activities in term of IT part.

The **objectives** of Efrei bank this year are to reduce churn in his younger costumers, sell more real estate loans and insurances then open new offices in where relevant for business growth.

All these initiatives should be driven by the company's Data asset.

As the new Chief data officer and member of the executive board, you have the responsibility to transform Efrei Bank to become data driven bank.

Questions et réponses :

Question 1 : • Please formalize the organizational chart of the Efrei bank with existing data roles. • Based on to the tools used by Efrei Bank what are the most relevant kind of data platform that will support Data usages and valuation. Justify your choice.

Organigramme de l'Efrei Bank avec les Rôles de Données Existants :

1. Associés Fondateurs (Conseil d'Administration) :

- Autorité décisionnelle finale.
- Définit la direction stratégique de la banque.
- Comprend les trois étudiants originaires qui ont fondé la banque.

2. Directeurs Généraux Adjoint :

- Chef de chaque branche clé :
 - Ressources Humaines
 - Services Marketing
 - Finance
 - IT

3. Département des Ressources Humaines :

- Responsable de la gestion de toutes les activités liées aux employés.
- Dirigé par le Directeur des RH.

4. Département des Services Marketing :

- Dirige le développement de la marque Efrei Bank.
- Gère toutes les activités commerciales, de l'onboarding des nouveaux clients à la finalisation des contrats.
- Dirigé par le Directeur Marketing.

5. Département Financier :

- Gère toutes les finances de l'entreprise et la responsabilité.
- Utilise des fichiers Excel pour les analyses et les rapports.
- Dirigé par le Directeur Financier.

6. Département IT :

- Supervise tous les domaines liés au numérique et aux systèmes d'information.
- Gère le numérique, les logiciels, les données, l'infrastructure, les sites web, etc.
- Utilise divers outils IT pour sécuriser le système d'information et gérer les données existantes.
- Dirigé par le Directeur IT.

7. Chefs de Données Locaux (Un pour Chaque Entité) :

- Responsable des activités liées aux données au sein de leurs entités respectives.
- Assure la conformité et la gouvernance des données.

8. Leader des Données IT :

- Centralise et supervise toutes les activités liées aux données au sein du département IT.
- Responsable de la gouvernance, de la gestion et de la stratégie des données.

Questions et réponses :

Question 1 : • Please formalize the organizational chart of the Efrei bank with existing data roles. • Based on to the tools used by Efrei Bank what are the most relevant kind of data platform that will support Data usages and valuation. Justify your choice.

Plateformes de Données les Plus Pertinentes pour l'Efrei Bank :

En tenant compte des outils actuellement utilisés par l'Efrei Bank et du besoin de gestion et d'utilisation complètes des données, les types de plateformes de données les plus pertinents seraient :

1. **Plateforme de Données Intégrée :** L'Efrei Bank a besoin d'une plateforme capable d'intégrer diverses sources de données provenant de différents départements tels que les RH, le marketing, la finance et l'IT. Cette plateforme devrait permettre un partage et une collaboration des données sans faille dans toute l'organisation.
2. **Entrepôt de Données Basé sur le Cloud :** Étant donné que l'Efrei Bank utilise un système ERP pour la gestion des RH et a des préoccupations concernant le partage des données avec son fournisseur actuel, un entrepôt de données basé sur le cloud fournirait une solution évolutive pour stocker et organiser les données de manière sécurisée. Il faciliterait également l'analyse et le reporting des données.
3. **Plateforme de Business Intelligence (BI) et d'Analytique :** Pour répondre au besoin d'analyse et de reporting des données, une plateforme de BI et d'analytique serait essentielle. Cette plateforme devrait prendre en charge des capacités d'analyse avancées, y compris l'analyse prédictive, pour aider l'Efrei Bank à comprendre le comportement des clients, à identifier les tendances et à prendre des décisions basées sur les données.
4. **Plateforme de Gouvernance des Données et de Conformité :** Compte tenu de l'environnement réglementaire dans le secteur bancaire, l'Efrei Bank bénéficierait d'une plateforme de gouvernance des données et de conformité. Cette plateforme garantirait la qualité, l'intégrité et la sécurité des données tout en aidant la banque à se conformer aux réglementations en matière de données telles que le RGPD.
5. **Outils d'Intégration de Données et ETL :** Étant donné que l'Efrei Bank utilise divers outils IT pour sécuriser le système d'information et gérer les données, des outils d'intégration de données et de transformation (ETL) seraient cruciaux pour intégrer des données provenant de différentes sources et formats dans la plateforme de données.

Questions et réponses :

Question 2 : • Identify 5 key data roles that will help you as a CDO to achieve the ambition of becoming data driven. • Describe their activities and responsibilities in Efrei bank context and identify tools that they will use in their everyday life (Ex : Data scientist: Developing models for business use in python with google collab)

1. Data Architect (Architecte des Données) :

- Activités et Responsabilités :
 - Concevoir et gérer l'architecture des données de l'entreprise, y compris les bases de données, les entrepôts de données et les pipelines de données.
 - Collaborer avec les équipes métier et techniques pour comprendre les besoins en matière de données et concevoir des solutions adaptées.
 - Assurer la cohérence, la sécurité et la conformité des données.
- Outils :
 - Outils de modélisation de données comme ERwin.
 - Outils de gestion des bases de données comme SQL Server Management Studio ou Oracle SQL Developer.
 - Outils d'intégration de données : Talend
 - Outils de surveillance et de gestion des performances : Splunk

2. Data Scientist (Scientifique des Données) :

- Activités et Responsabilités :
 - Analyser les données pour extraire des insights et des tendances exploitables.
 - Développer des modèles prédictifs et des algorithmes pour prendre des décisions basées sur les données.
 - Collaborer avec les équipes métier pour identifier les opportunités d'utilisation des données.
- Outils :
 - Langages de programmation comme Python et R.
 - Bibliothèques d'apprentissage automatique comme TensorFlow ou scikit-learn.
 - Plateformes de développement comme Google Colab ou Jupyter Notebook.

3. Data Engineer (Ingénieur des Données) :

- Activités et Responsabilités :
 - Concevoir, développer et maintenir les pipelines de données pour l'acquisition, le traitement et le stockage des données.
 - Optimiser les performances et la fiabilité des infrastructures de données.
 - Collaborer avec les équipes métier pour intégrer de nouvelles sources de données.
- Outils :
 - Outils ETL (Extract, Transform, Load) comme Talend.
 - Plateformes de traitement de données comme Snowflake.

Questions et réponses :

Question 2 : • Identify 5 key data roles that will help you as a CDO to achieve the ambition of becoming data driven. • Describe their activities and responsibilities in Efrei bank context and identify tools that they will use in their everyday life (Ex : Data scientist: Developing models for business use in python with google collab)

1. Data Quality Manager (Responsable de la qualité des Données) :

- Activités et Responsabilités :
 - Identifier les sources de données et évaluer leur qualité en termes de précision, d'exhaustivité, de cohérence et de fiabilité.
 - Développer des métriques et des KPIs pour mesurer la qualité des données et suivre les progrès dans l'amélioration de la qualité.
 - Élaborer des normes et des processus pour garantir la qualité des données tout au long de leur cycle de vie, depuis la collecte jusqu'à l'analyse.
 - Sensibiliser les employés à l'importance de la qualité des données et fournir une formation sur les bonnes pratiques de gestion des données.
- Outils :
 - Outils de gestion de la qualité des données comme Talend Data Quality.
 - Outils de Surveillance de la Qualité des Données comme Talend Data Quality.

2. Data Analyst :

- Activités et Responsabilités :
 - Créer et maintenir des rapports et des tableaux de bord pour suivre les performances commerciales.
 - Analyser les données pour fournir des insights aux décideurs.
 - Collaborer avec les équipes métier pour identifier les besoins en reporting.
- Outils :
 - Outils de BI et de visualisation de données comme Tableau, Power BI.
 - Langages de requête comme SQL pour l'extraction de données à partir des entrepôts de données.

Questions et réponses :

Question 3 : • Define the strategic reasons that motivate Efrei Bank to become Data driven. • Also list the difficulties that could represent huge challenges for becoming a data driven bank, please bring propositions on what data usage or use cases will you deploy for the business objectives.

Raisons Stratégiques pour Devenir Data Driven:

1. **Amélioration de la Prise de Décision** : En utilisant des données pertinentes et précises, l'Efrei Bank peut prendre des décisions plus éclairées et basées sur des preuves, ce qui peut conduire à des résultats commerciaux plus efficaces et rentables.
2. **Personnalisation des Services** : En comprenant mieux les besoins et les comportements des clients grâce à l'analyse des données, la banque peut offrir des produits et des services plus personnalisés, ce qui peut renforcer la fidélité des clients et augmenter la satisfaction.
3. **Optimisation des Opérations** : L'analyse des données peut aider à identifier les inefficacités opérationnelles et à améliorer les processus internes, ce qui peut conduire à des économies de coûts et à une utilisation plus efficace des ressources.
4. **Détection des Tendances du Marché** : En surveillant les tendances du marché et en anticipant les changements, l'Efrei Bank peut rester compétitive et réactive dans un environnement financier en évolution constante.
5. **Gestion des Risques Améliorée** : En utilisant des modèles prédictifs basés sur les données, la banque peut mieux évaluer et gérer les risques, ce qui peut contribuer à réduire les pertes et à renforcer la stabilité financière.

Difficultés Potentielles et Propositions de Solutions :

1. **Problèmes de Qualité des Données** : Les données de mauvaise qualité peuvent compromettre l'efficacité des analyses et des décisions. Proposition : Mettre en place des processus de gouvernance des données pour assurer la qualité, la fiabilité et la cohérence des données.
2. **Défis en Matière de Gouvernance et de Conformité** : Les réglementations strictes en matière de confidentialité des données peuvent rendre complexe l'utilisation et le partage des données. Proposition : Investir dans des technologies de gouvernance des données et établir des politiques de conformité claires pour garantir la sécurité et la confidentialité des données.
3. **Culture Organisationnelle** : La résistance au changement et le manque de sensibilisation à l'importance des données peuvent entraver l'adoption d'une approche axée sur les données. Proposition : Mettre en place des programmes de sensibilisation et de formation pour promouvoir une culture axée sur les données au sein de l'organisation.
4. **Infrastructure et Technologies Obsolètes** : Les systèmes informatiques vieillissants peuvent limiter la capacité de l'Efrei Bank à exploiter pleinement ses données. Proposition : Investir dans des technologies de pointe et mettre à niveau l'infrastructure informatique pour prendre en charge l'analyse des données à grande échelle.
5. **Siloisation des Données** : Les données dispersées dans différents systèmes et départements peuvent entraver la collaboration et l'intégration des données. Proposition : Mettre en œuvre une stratégie de gestion des données centralisée et intégrée pour favoriser le partage et l'utilisation efficace des données à travers l'organisation.

Questions et réponses :

. Question 4 : We shared the data management wheel with ten thematic during the course: • Choose 3 themes that you think are relevant to start with, in the case of Efrei bank • Identify the data domains according to the type of manipulated data. • Identify and list a minimum of three activities linked to the chosen themes that you will set up as a CDO. You can list more than three if relevant.

Thèmes Pertinents pour la Banque Efrei :

1. Data Security (Sécurité des données) :

- Domaines de données :
 - Données client.
 - Données financières.
 - Données de sécurité.
- Activités :
 - Développement et mise en œuvre de politiques et de procédures de sécurité des données pour protéger les actifs d'information contre les menaces internes et externes.
 - Gestion des accès aux données en définissant des stratégies d'authentification et d'autorisation appropriées pour limiter l'accès aux données sensibles.
 - Surveillance et audit réguliers des systèmes et des processus de sécurité des données pour détecter et répondre aux violations potentielles de la sécurité.

2. Data Warehousing & BI :

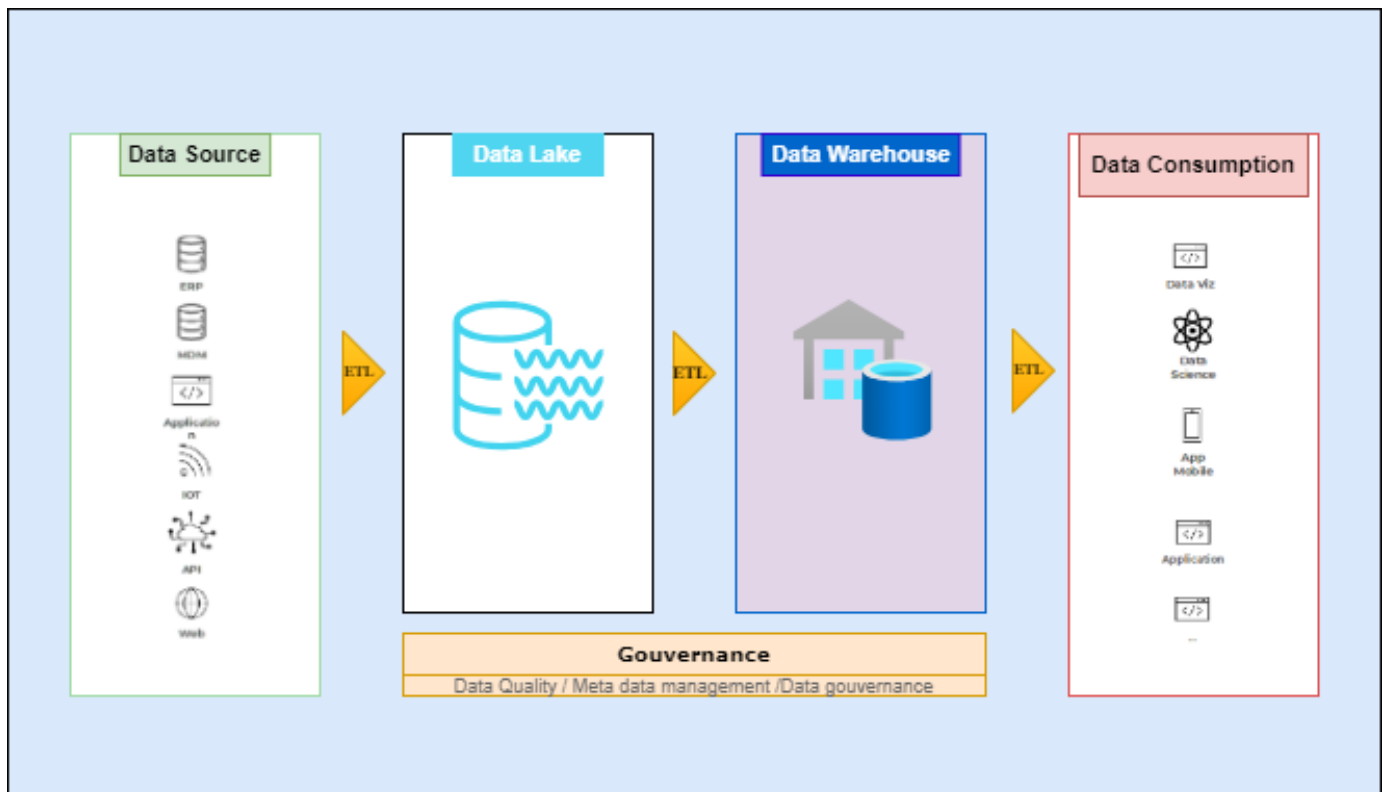
- Domaines de Données :
 - Données de Comportement Client
 - Données de Performance des Produits
 - Données de Marché
- Activités :
 - Stockage des données et Analyse des tendances et des modèles dans les comportements des clients pour améliorer la personnalisation des services.
 - Élaboration de tableaux de bord et de rapports pour suivre la performance des produits et identifier les opportunités d'amélioration.
 - Utilisation de l'analyse prédictive pour anticiper les changements de marché et prendre des décisions stratégiques.

3. Qualité des Données :

- Domaines de Données :
 - Données Clientes
 - Données Financières
 - Données Opérationnelles
- Activités :
 - Évaluation et nettoyage des données pour garantir la précision et la cohérence des informations.
 - Mise en place de processus de gestion de la qualité des données pour surveiller et améliorer en permanence la qualité des données.
 - Formation et sensibilisation des employés à l'importance de maintenir des données de haute qualité et à jour.

Questions et réponses :

Question 5 : Propose an architecture view from data acquisition to Visualization for a chosen use case in question 3 of the Efrei bank data. Explain your choices using (GCP, AWS or Azure platforms) in details



Outils de la Data Platform :

ETL : Talend

Data Lake: Azure Cloud

Data Warehouse: Snowflake

Data Gouvernance: Informatica

Data Consumption: Power BI, microstrategy, Anaconda et Jupyter

Langage Informatique: python ,spark

Questions et réponses :

Question 6 : Imagine & propose relevant KPIs that helps measure data management maturity into Efrei Bank. Find one KPI for each data management field. (Example : number of existing business data models in Data modeling & Design)

- 1. Intégration des Systèmes (Data Architecture) :**
 - Mesure le nombre de systèmes intégrés pour une gestion cohérente des données, facilitant ainsi la prise de décision.
- 2. Taux d'Exactitude des Données (Data Quality) :**
 - Évalue la précision des données pour garantir la fiabilité des informations utilisées.
- 3. Couverture des Métadonnées (Metadata Management) :**
 - Mesure le pourcentage de données avec des métadonnées pour faciliter leur découverte et leur analyse.
- 4. Temps de Génération de Rapports (Data Warehousing & BI) :**
 - Évalue le temps nécessaire pour produire un rapport, favorisant une prise de décision rapide.
- 5. Conformité des Données Maîtresses (Reference & Master Data) :**
 - Mesure le pourcentage de données maîtresses conformes pour assurer la cohérence des informations.
- 6. Accessibilité des Documents (Document & Content Management) :**
 - Évalue le pourcentage de documents accessibles pour favoriser la productivité et la collaboration.
- 7. Intégration des Systèmes (Data Integration & Interoperability) :**
 - Mesure le nombre de systèmes connectés pour faciliter l'échange d'informations.
- 8. Conformité aux Normes de Sécurité (Data Security) :**
 - Évalue le pourcentage de données protégées pour assurer la sécurité des informations.
- 9. Utilisation de l'Espace de Stockage (Data Storage & Operations) :**
 - Mesure le pourcentage d'espace de stockage utilisé pour optimiser les coûts.
- 10. Normalisation des Modèles de Données (Data Modelling & Design) :**
 - Mesure le nombre de modèles de données conformes pour garantir la cohérence des informations.

