

Plan de cours

Cours :	INF 735 – Entrepôt et forage de données
Trimestre :	Hiver 2020
Enseignant :	Robert J. Laurin

1. Mise en contexte

Dans la dernière décennie, l'utilisation stratégique de l'information a changé le monde des affaires. Ce changement s'est produit grâce à l'utilisation d'entrepôts de données comme outil d'aide à la décision et l'intelligence derrière certaines pratiques commerciales. Il est devenu essentiel pour le gestionnaire de connaître les enjeux que représentent les entrepôts de données et leur forage pour survivre dans le monde des affaires comme dans celui des technologies de l'information.

2. Place du cours dans le programme

Ce cours s'inscrit dans le programme de Diplôme de 2^e cycle de technologies de l'information. Il a comme préalable le cours INF732 « Bases de données ».

3. Descripteur du cours

Objectifs : Comprendre les concepts sous-jacents à la réalisation d'un entrepôt de données. Comprendre les concepts sous-jacents au forage de données et l'utilisation avec un entrepôt de données. Réaliser un entrepôt de données (Data Warehouse) et lui appliquer des techniques de présentation multidimensionnelle de l'information.

Contenu : Définitions et liens avec le contexte d'affaire d'une entreprise. Revue des enjeux et définition des besoins. Caractéristiques et architecture d'un entrepôt. Méta données. Modèles multidimensionnels. Schémas. Démarche de construction d'un entrepôt. Caractéristiques des outils OLAP et de forage. Limitations, défis et outils de présentation tel OLAP et tableaux de bords et forage.

4. Objectifs spécifiques

Au terme du cours, l'étudiante ou l'étudiant sera capable :

1. d'appliquer les principaux concepts sous-jacents à l'environnement de données de l'entreprise et leur organisation en entrepôt;
2. d'identifier les différentes approches informationnelles et de choisir l'architecture d'entrepôts de données appropriée;
3. de décrire les étapes cruciales dans la construction et la modélisation d'un entrepôt de données;
4. d'utiliser les outils d'exploitation et de forage de données.

5. Planification hebdomadaire

CONTENU DU COURS	Semaines	CHAPITRES DU LIVRE
Bloc 1 - Introduction: besoin, concepts et définitions <ul style="list-style-type: none">• Positionnement du cours• Évolution des T.I.• Besoin en information stratégique• Systèmes opérationnels .vs. informationnels• Entrepôts de données• Définition, étapes et enjeux	1-3	1, 2, 3
Bloc 2 - Modélisation (Entrepôt) <ul style="list-style-type: none">• Dimensions et schémas• Modèle étoile et Constellation• Flocons• Agrégations• ODS	4-5	10, 11
Bloc 3 – Outils de présentation, OLAP et Forage <ul style="list-style-type: none">• Outils de présentation de l'information• OLAP, Excel, Rapports• Tableaux de bord, Cartes de pointage• Techniques de forage et outils• Outils de développement	6-7	14, 15, 16, 17
Semaine 8 → Intra	8	
Bloc 3 – Outils de présentation, OLAP et Forage - Suite	9 - 10	

CONTENU DU COURS	Semaines	CHAPITRES DU LIVRE
Bloc 4 – ETC : Acquisition de données <ul style="list-style-type: none"> • Extraction de données • Méthode de ravitaillement • Transformation • Chargement et historisation • Qualité de la donnée et défis 	11	12, 13
Bloc 5 - Architecture et Méta Données <ul style="list-style-type: none"> • Architecture de l'entrepôt • Approches d'implantation • Méta données • Méta information et l'entrepôt 	12	7, 9
Bloc 6 – Gestion d'un projet d'entrepôt de données <ul style="list-style-type: none"> • Définition des besoins • Acteurs, rôles et responsabilités • Justifications • Gestion d'un projet d'entrepôt 	13	4, 5, 6
Bloc 7 - Techniques de réalisation et opération <ul style="list-style-type: none"> • Data Webhouse • Modèle physique • Infrastructure • Activités de déploiement • Optimisation • Gestion • Évolution 	14	8, 18, 19, 20
Semaine 15 – Examen final	15	

CONTENU DU LABORATOIRE	Semaines
Chargée d'exercices	
Laboratoires : <ol style="list-style-type: none"> Voir le modèle dans SQL PowerPivot MS Excel (requis: MS SQL 2008R2, MS Excel 2010 et Power Pivot) - en lien avec le TP2 ETC avec SSIS Designer (SQL Server Business Intelligence Development Studio) Cubes avec SSAS (Microsoft SQL) 	à déterminer

6. Approche pédagogique préconisée

Le cours permet à l'étudiante ou l'étudiant de prendre en charge de ses apprentissages. Le professeur est là comme guide et soutien dans cette démarche. Les notions théoriques indispensables à l'atteinte des objectifs seront exposées en classe en exploitant l'expérience vécue par chaque personne dans son milieu de travail. L'utilisation de logiciels de manipulation de données apportera une dimension pratique indispensable à l'atteinte des objectifs du cours. Avant chaque prestation, les étudiantes et les étudiants doivent lire les chapitres du livre et les notes de cours mentionnés dans la planification hebdomadaire.

La chargée d'exercice pourra, sur rendez-vous, vous assister dans vos travaux pratiques.

7. Évaluation de l'apprentissage

Les examens comportent des questions :

- à choix multiples,
- à développement court (moins d'une demi-page),
- à développement long.

L'enseignant évaluera seul les examens et les travaux.

Aucun travail en retard ne sera accepté, basé sur la date et l'heure de l'envoi électronique.

L'enseignant peut retrancher jusqu'à 5% de chaque évaluation pour la qualité du français.

Description sommaire	Pondération
TP1 - Travail pratique individuel ou en équipe	10 %
TP2 - Travail individuel ou en équipe	20 %
Examen mi-session -évaluation individuelle	30 %
Examen final -évaluation individuelle	40 %

8. Plagiat

Conformément à l'article 9 du Règlement des études de l'Université de Sherbrooke, le plagiat, soit le fait dans une activité évaluée de faire passer pour sien les idées et le travail d'autrui, est un délit académique qui peut être sanctionné par les autorités disciplinaires compétentes. Peuvent être imposées à titre de sanctions, l'une ou plusieurs des mesures suivantes :

- a) la réprimande simple ou sévère consignée au dossier étudiant pour la période fixée par l'autorité disciplinaire ou, à défaut, définitivement. En cas de réprimande fixée pour une période déterminée, la décision rendue demeure au dossier de la personne aux seuls fins d'attester de l'existence du délit en cas de récidive;
- b) l'obligation de reprendre une production ou une activité pédagogique, dont la note pourra être établie en tenant compte du délit survenu antérieurement;
- c) la diminution de la note ou l'attribution de la note E ou 0;
- d) le renvoi du dossier à la personne responsable de l'évaluation d'une production ou d'une activité pédagogique pour qu'elle attribue une nouvelle note en tenant compte du délit.

Par plagiat, on entend notamment :

- copier intégralement une phrase ou un passage d'un livre, d'un article de journal ou de revue, d'une page Web ou de tout autre document en omettant d'en mentionner la source ou de le mettre entre guillemets;
- reproduire des présentations, des dessins, des photographies, des graphiques, des données sans en préciser la provenance et, dans certains cas, sans en avoir obtenu la permission de reproduire;
- utiliser, en tout ou en partie, du matériel sonore, graphique ou visuel, des pages Internet, du code de programme informatique ou des éléments de logiciel, des données ou résultats d'expérimentation ou toute autre information en provenance d'autrui en le faisant passer pour sien ou sans en citer les sources;
- résumer ou paraphraser l'idée d'un auteur sans en indiquer la source;
- traduire en partie ou en totalité un texte en omettant d'en mentionner la source ou de le mettre entre guillemets;
- utiliser le travail d'un autre et le présenter comme sien (et ce, même si cette personne a donné son accord);
- acheter un travail sur le Web ou ailleurs et le faire passer pour sien;
- utiliser sans autorisation le même travail pour deux activités différentes (autoplégat).

9. Adresse électronique pour la remise des travaux

Robert.J.Laurin@usherbrooke.ca

10. Bibliographie

Livre facultatif

Data Warehousing Fundamentals : A Comprehensive Guide for IT Professionals

Auteur: Paulraj Ponniah

Hardcover: 571 pages

Éditeur: Wiley-Interscience; Second Edition (Août 2001)

ISBN: 0470462078

Autres références utiles

1. **The Data Warehouse Lifecycle Toolkit**, Expert Methods for designing, Developing, and Deploying Data Warehouses
Auteur: Ralph Kimbal, Laura Reeves, Margy Ross et Warren Thornthwaite
Paperback: 800 pages
Éditeur: John Wiley & Sons, édition Livre et Cd-rom (Août 1998)
ISBN: 0471255475
2. **The Data Warehouse Toolkit**, The complete guide to Dimensional Modeling
Auteur: Ralph Kimbal et Margy Ross



Paperback: 436 pages
Éditeur: John Wiley & Sons, Seconde édition (2002)
ISBN: 0471200247

3. **Building the Data Warehouse**

Auteur: William H. Inmon
Paperback: 541 pages
Éditeur: John Wiley & Sons, Quatrième édition (2005)
ISBN: 0764599445

4. **e-Data: Turning Data into Information with Data Warehousing**

Auteur: Jill Dyché
Paperback: 384 pages ; Dimensions (in inches): 0.72 x 9.22 x 7.33
Éditeur: Addison-Wesley Pub Co; 1 édition (Février 2000)
ISBN: 0201657805

5. **Building and Managing the Meta Data Repository: a full lifecycle guide**

Auteur: David Marco
Hardcover: 392 pages
Publisher: John Wiley & Sons, inc.
ISBN: 0471355232

11. Logiciels

SGBD relationnel MS SQL 2008R2 (minimum)
TABLEUR MS Excel 2010 (Minimum)

