## Sommaire

1.	Objectif BD	3
2.	Modèle logique	3
3.	Modèle relationnel	6
4.	Requêtes pour l'analyse de données	9
5.	Conclusion	12
	Annexe A : Méta-données pour l'explication des attributs	13
	Annexe B : Réalisation en Access	14

### 1. Objectif BD

<u>Domaine:</u> les données fournies par les différents services du Département de l'éducation du Royaume-Uni.

<u>Préconditions</u>: l'éducation scolaire en Royaume Uni se trouve dans une période de transformations globales qui incluent la création de Multi-Trusts pour la gestion des écoles. Comme résultat principal (pour l'objectif de BD), le financement des trusts est limité et les trusts (et les organisations publiques) doivent être plus précis avec les données financières et utiliser l'analyse de données, la modélisation de résultats, les rapports ainsi que la visualisation. Un autre problème, c'est la qualité des données initiales : les différents services du Département de l'éducation publient des données dans leurs sous-domaines d'activité, avec leurs méta-données et calculs. Par exemple, le salaire moyen et le budget pour les enseignants sont dans deux rapports différents, et le nombre d'élèves peut être différent dans les rapports différents.

Les données consistent en KPIs financiers, nombre d'élèves, information de management et de professeurs pour les écoles des niveaux KS2, KS4, KS5 (primaire, secondaire et 16-18 ans) comme des rapports annuels. Il est nécessaire de souligner que les données sont en forme de panel. Elles existent pour plusieurs années, donc on peut identifier les données par école et année conjointement.

<u>L'objectif</u> de la base de données est de rassembler des données pour simplifier le support décisionnel et l'analyse de données qui inclut des calculs de KPIs nécessaires.

<u>Le consommateur final</u> de la base de données : le management du trust ou des autorités locales qui doivent contrôler des résultats des écoles et prendre des décisions financières. En prenant en compte le consommateur final, on a <u>une condition supplémentaire</u>: on doit garder les données en accord avec les pratiques du Département de l'éducation. Par exemple, les données financières doivent être séparées des données de recensement, même s'il semble plus logique de tout ajouter dans un seul tableau.

Aussi, il est nécessaire de prévoir la possibilité de voir au minimum:

- si le directeur est presque retraité ou pas,
- l'expiration de certificats pour des enseignants,
- calculer des financements par élève,
- calculer un ratio d'enseignants par élève.

## 2. Modèle logique

Il est nécessaire de commencer par le modèle entités-associations. La liste des entités prévues est suivante:

Tableau 1. Les entités de la base de données

Entité	L'information explicative				
École	La description générale quantitative des écoles : le code national d'école,				
	le nom d'école, le code postal et l'autorité locale (pour la position				
	géographique d'école), et si l'école et associée avec le trust <sup>1</sup> ou pas.				
Directeur	Les données personnelles des directeurs d'écoles. On sait qu'un directeur				
	peut diriger juste une école. Mais il peut être également enseignant à				
	temps partiel. Il est important que la personne retraitée ne puisse pas être				
	directeur.				
Enseignant	Les données personnelles des enseignants doivent inclure le nom, le				
	prénom, la date de certification (pour vérifier s'il peut continuer son				
	travail). En plus, il est nécessaire de prévoir une différentiation p				
	l'expérience. On sait qu'après 10, 20, 25 ans la rémunération augmente				
	pour un pourcentage défini. Cette entité représente les places de trav				
	principal.				
Census	L'information du nombre d'élèves				
SWF	School work force (l'effectif scolaire): l'information du nombre des				
	enseignants, assistants etc. Les données doivent être présentes à plein				
	temps équivalent.				
CFR	Les données financières annuelles.				

Les entités Census, SWF et CFR sont descriptives par rapport à l'entité École.

Pour formuler les associations, je préfère utiliser les formules « obtenir des résultats financiers », « avoir l'info d'effectif », « avoir l'info d'élèves » qui sont un peu différentes d'un nom ou d'un verbe mais décrivent plus précisément des relations. Pourtant pour respecter les formalités on utilisera les mots « Réussir », « Posséder » et « Contenir ». On peut prévoir des associations suivantes :

Tableau 2. Les associations de la base de données

Entité 1	Entité 1 Action Entité 2		Explication	Type de lien
Enseignant	Enseigner	École	L'enseignant peut travailler	n:m
			dans 1+école, l'école peut	
			avoir 1+enseignants	
École Posséder SWF		L'école peut avoir l'effectif	1:n	
		pour 1+ année, l'information		
			de l'effectif caractérise 1 école	
			par an	

<sup>&</sup>lt;sup>1</sup>Conformément à la réforme éducative qui a été faite il y a plusieurs années, l'école peut rester séparée, sous gestion publique ou privée, peut être en fiduciaire unique ou en multi fiducie. Cette fiducie gère une ou plusieurs écoles, y compris la réception du budget des écoles. Cela a entraîné des exigences plus élevées en matière de gestion des données et de rapports dans le système éducatif.

École	Diriger	Directeurs	L'école peut avoir 1 directeur,	1:1
			le directeur peut travailler	
			comme directeur dans 1 école	
Directeur	Is-a	Enseignants	Le directeur peut être aussi un	Is-a, 1:1
			enseignant	
École	Contenir	Census	L'école peut avoir des élèves	1:n
			pour 1+ année. L'information	
			des élèves caractérise	
			seulement 1 école par an.	
École	Réussir	CFR	L'école peut obtenir des	1:n
			résultats pour 1+ années,	
			chaque résultat caractérise 1	
			école par an	
Enseignants	Recevoir	SWF	L'enseignant reçoit un salaire	1:n
			(en termes de salaire moyen,	
		identifiant). Le salaire peut être		
		changé chaque année. Donc on		
			peut obtenir 1+ salaires pour	
			enseignant	
Census	Demander	CFR	Les élèves demandent le	1:1
			financement et les dépenses.	
			Les dépenses peuvent être	
			changées chaque année et les	
			données de Census aussi, mais	
			elles sont toujours définies par	
			année et école	

On peut noter que dans cette base de données le lien 1:n apparaît grâce à des données du panel. La base d'information précédente, on peut formuler la liste des entités-associations et des propriétés prévues, et aussi définir les clefs (les dernières sont soulignées). Une clef doit permettre l'identification unique par chaque unité qui appartient à une entité. Dans ce cas on a aussi des clefs composites qui consistent en deux propriétés.

- 1. École {Nom d'école, AL, Trust, CodePostal, Date d'ouverture, Date de fermeture}
- 2. Directeurs {PrénomD, NomD, Date de naissanceD, <u>CodeD</u>}
- 3. Census {Nombre de garçons (FTE), Nombre de filles (FTE), % d'élèves ayant droit aux repas gratuits/ FSM, <u>URN, Année</u>}
  - 4. SWF {NTEA, NTEAAS, Salaire, URN, Année}
- 5. CFR {Elèves, FSM, GRANTFUNDING, SELFGENERATEDINCOME, TEACHINGSTAFF, SUPPLYTEACHERS, EDUCATIONSUPPORTSTAFF, PREMISES, BACKOFFICE, CATERING, OTHERSTAFF, ENERGY, LEARNINGRESOURCES, ICT, BOUGHTINPROFESSIONALSERVICES, OTHER, <u>URN</u>, <u>Année</u>}
- 6. Enseignants {NomE, PrénomE, Date de début d'emploi, Date de naissanceE, <u>CodeE</u>, Majeur, Date de certification}

L'explication des attributs est présentée dans l'Annexe A.

Les associations peuvent avoir les clefs des entités comme attributs. Pour les attributs d'associations la clef s'écrira sous la forme « URN+Année » (pour éviter la confusion avec un extra attribut).

- 7. Enseigner {CodeE, Nom d'école, Rôle}
- 8. Posséder {Nom d'école, URN+Année}
- 9. Diriger {CodeD, Nom d'école}
- 10. Is-a {Code D, CodeE}
- 11. Contenir {Nom d'école, URN+Année}
- 12. Réussir {Nom d'école, URN+Année}
- 13. Recevoir {CodeD,URN+Année}
- 14. Demander {URN+Année, URN+Année}

La représentation graphique du modèle logique est dans la Figure 1.

La Figure 1 et le Tableau 2 représentent les types de liens prévus. Il est à noter que les types de liens sont donnés en système 1:1, 1:n, n:m. Si la cardinalité maximale est connue (par exemple, le nombre d'ans ou on possède des données), on considère toujours qu'elle est indéterminée et vaut n.<sup>2</sup>

Les associations entre les données du panel (CFR, Census, SWF) et Ecole ou Enseignants ont le type n:1, parce que pour chaque école on possède des données pour les plusieurs années.

#### 3. Modèle relationnel

On a construit le modèle Entité-Association, ou le modèle logique. Pour passage au modèle relationnel on doit transformer toutes les entités aux tableaux avec les mêmes noms et mêmes attributs. Un lien qui a été défini pour le modelé logique peut disparaître dans le modelé relationnel. Généralement, toutes les bijections ou toutes les injections d'une entité vers une autre peut disparaître en ajoutant les clefs "étrangères" dans un des tableaux relationnels correspondants.

Les liens de type n:m demeurent toujours des tableaux relationnels importants dans les bases de données.

On va commencer le passage du modèle logique par les associations. L'association Enseigner {CodeE, Nom d'école, Rôle} a le type de relation n:m, donc on doit la garder comme le tableau Enseignement avec les attributs CodeE, Nom d'école et Rôle.

<sup>&</sup>lt;sup>2</sup> https://laurent-audibert.developpez.com/Cours-BD/?page=conception-des-bases-de-donnees-modele-a

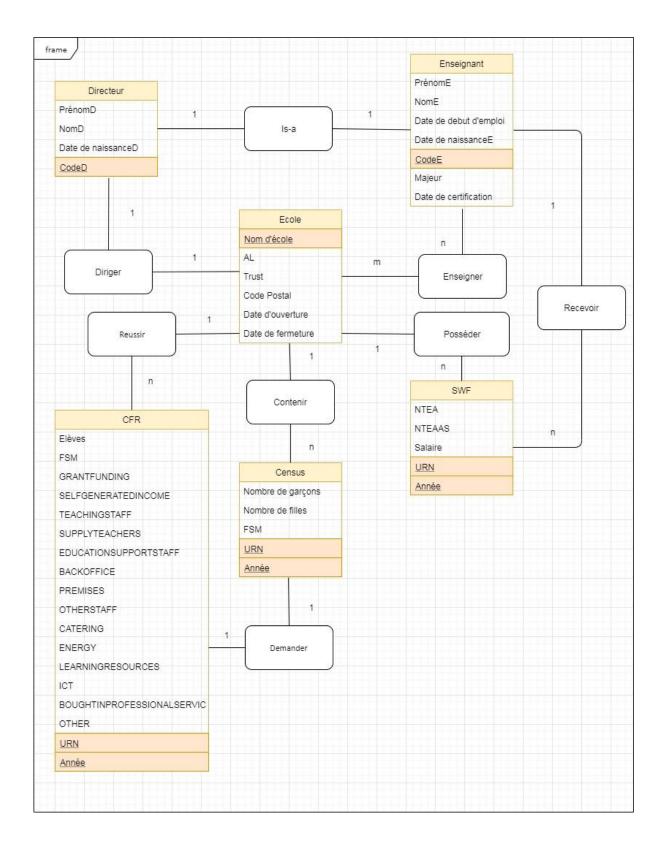


Figure 1. Représentation graphique du modèle logique

Toutes les autres associations doivent être simplifiées par le remplacement de clefs étrangères. Le tableau 3 présente les transformations nécessaires; la ligne EA correspond aux entités et associations du schéma logique, la ligne SR - avec le schéma relationnel, le résultat de

transformation. S'il y a plusieurs itérations, transformations entité-tableau, l'itération suivante inclut les changements précédents. Les clefs sont soulignées, les clefs étrangères sont en gros.

Tableau 3. Les transformations EA-Schéma physique

ГА		I	CAME (NITE A NITE A A C. C. L. LIDNI A	
EA	École {Nom	Recevoir	SWF{NTEA, NTEAAS, Salaire, <u>URN, Année</u> }	
	d'école, AL, Trust,	{URN,		
	Code Postal, Date	URN+Année		
	d'ouverture, Date de	}		
	fermeture}			
SR	•	· · · · · · · · · · · · · · · · · · ·	N, Année, Nom d'école}	
EA	Directeurs	Diriger	École {URN, Nom d'école, AL, Trust, Code	
	{PrénomD, NomD,	{CodeD,	Postal, Date d'ouverture, Date de fermeture}	
	Date de naissanceD,	Nom		
	CodeD}	d'école}		
SR	Directeurs {PrénomD		e naissanceD, <u>CodeD</u> , <b>Nom d'école</b> }	
EA	Directeurs	Is-a {Code	Enseignants {NomE, PrénomE, Date de début	
	{PrénomD, NomD,	D, CodeE}	d'emploi, Date de naissanceE, <u>CodeE</u> , Majeur,	
	Date de naissanceD,		Date de certification}	
	CodeD}			
SR	Directeurs {PrénomD	, NomD, Date d	le naissanceD, <u>CodeD</u> , <b>Nom d'école</b> , <b>CodeE</b> }	
EA	École {Nom	Contenir	Census {Nombre de garcons (FTE), Nombre de	
	d'école, AL, Trust,	{Nom	filles (FTE), % d'étudiants éligible aux repas	
	Code Postal, Date	d'école,	gratuits/ FSM, URN, Année}	
	d'ouverture, Date de	URN+Année		
	fermeture}	}		
SR	Census {Nombre de	garçons (FTE),	Nombre de filles (FTE), % d'étudiants éligible	
	aux repas gratuits/ FS	M, URN, Anné	e, Nom d'école}	
EA	École {Nom	Réussir	CFR{Elèves, FSM, GRANTFUNDING,	
	d'école, AL, Trust,	{Nom	SELFGENERATEDINCOME,	
	Code Postal, Date	d'école,	TEACHINGSTAFF, SUPPLYTEACHERS,	
	d'ouverture, Date de	URN+Année	EDUCATIONSUPPORTSTAFF, PREMISES,	
	fermeture}	}	BACKOFFICE, CATERING, OTHERSTAFF,	
	Í	-	ENERGY, LEARNINGRESOURCES, ICT,	
			BOUGHTINPROFESSIONALSERVICES,	
			OTHER, URN, Année}	
SR	CFR{Elèves, FSI	M, GRANT	FUNDING, SELFGENERATEDINCOME,	
	TEACHINGSTAFF,	,	ACHERS, EDUCATIONSUPPORTSTAFF,	
	· ·		CATERING, OTHERSTAFF, ENERGY,	
	LEARNINGRESOURCES, ICT			
		URN, Année, Nom d'école}		
EA	Enseignants	Recevoir	SWF{NTEA, NTEAAS, Salaire, URN, Année}	
	{NomE, PrénomE,		S (11121, 11121 113, Suidire, Ott., Millee)	
	Date de début			
l	Daic ac acoul	TATEMINE !		

	d'emploi, Date de		
	naissanceE, CodeE,		
	Majeur, Date de		
	certification}		
SR	SWF{NTEA, NTEAAS, Salaire, URN, Année, Nom d'école, CodeE}		

Finalement, on a réussi à minimiser le nombre de tableaux. Le schéma résultant avec des cardinalités prévus est présenté dans la Figure 2. Le lieu Enseigner est renommé comme Enseignement.

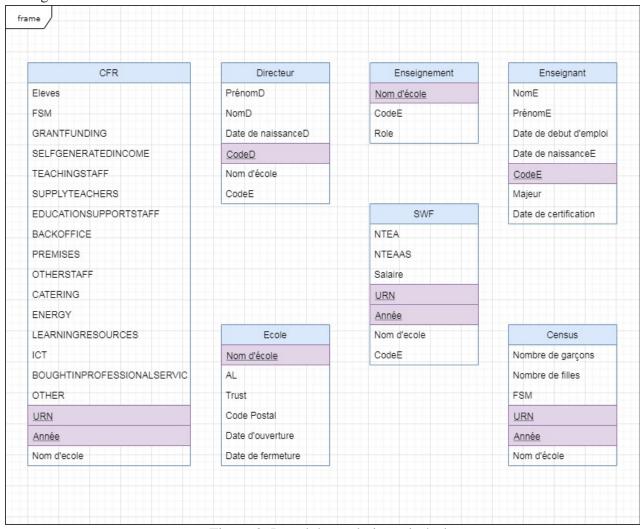


Figure 2. Le schéma relationnel résultant

L'implémentation en Access au niveau des données est présentée en Annexe B.

## 4. Requêtes pour l'analyse de données

Au niveau de la prise de décision, la base de données développée est utile pour les raisons suivantes : rassembler les données de différentes sources dans une seule base de données permet de voir les données de financement, d'effectif et d'élèves dans un seul endroit. En conséquence les ratios tels que les dépenses par élève ou bien par professeurs peuvent être calculés. Les tableaux

Directeurs et Enseignants aident à contrôler quand les enseignants doivent passer leur certification et vérifier que tout le personnel effectue les conditions de travail. L'expérience en tant que professeur montre les dépenses supplémentaires pour l'encadrement. Qui plus est l'âge des directeurs est important pour superviser les départs en retraite. La table Ecole est utile pour l'analyse spatiale.

A mon avis, la base de données est très utile pour MAT (Multi Academy Trust) car représente l'information de toutes les écoles nécessaires. L'analyse de données est faite en utilisant des requêtes.

Les exemples comment la base de données peut assister avec les décisions sont présentes dans le Tableau 4 :

Tableau 4. Les Requêtes utilisent: langue naturelle et SQL

Langue Naturelle	SQL			
Calculer le nombre des écoles	SELECT count(*) as NbEcole			
qui sont inclus aux trusts	FROM Ecole			
	WHERE Trust="Academy"			
Calculer le nombre des	SELECT count(b.CodeE) as NbEnsEC, c.Nomdecole			
enseignants par école	FROM Enseignement b, Ecole c			
	WHERE c.Nomdecole=b.Nomdecole			
	GROUP BY c.Nomdecole			
Calcule les extra de salaire pour	SELECT a.nomE & ' ' &a.prenomE as Enseign,			
toutes les enseignantes de la	datediff("yyyy",a.DDE,date()) as Experience,			
manière suivante 10 ans	iif(Experience<20 and Experience>=10, "10%",			
d'expérience - 10% d'extra	iif(Experience<=25 and Experience>=20, "20%",			
rémunération, 20 ans -	iif(Experience>=25, "35%", "0%"))) as Extra			
20%, >=25 - 35%	FROM Enseignant a			
Choisir tous les enseignants	SELECT a.nomE & ' ' &a.prenomE as Enseign, (2020-a.DC)			
avec les certificats qui ont déjà	-			
expiré ou va expirer en 2020. Le	FROM Enseignant a			
certificat est valide pendant 3	WHERE 2020-a.DC>3			
ans				
Prendre tout l'information qui	SELECT a.*, b.Majeur, b.DDE, b.DC			
existe pour les directeurs qui	FROM Directeur a			
aussi travaillent comme	INNER JOIN Enseignant b			
enseignants	ON a.CodeE=b.CodeE			
Calcule le moyen nombre élève				
par d'enseignants pour chaque				
école, base en Census et SWF	FROM Census c, SWF b			
	GROUP BY c.Nomdecole			
Mettre les données de tableaux	, , , , , , , , , , , , , , , , , , , ,			
Census, SWF et CFR dans 1	c.NTEAAS, c.Salaire INTO ALLDATA			
tableau	FROM (Census a INNER JOIN CFR b			
	ON a.URN=b.URN) INNER JOIN SWF c ON b.URN=c.URN			

Présenter le gendre d'école	SELECT c.URN, iif(count(c.NORG)>count(c.NORB),				
	"girls","boys") as gender				
	FROM Census c GROUP BY c.URN				
Présenter la liste avec les	SELECT a.Nomdecole, b.NomE, b.Majeur From				
majeurs et enseignants par les	Enseignement a, Enseignant b WHERE a.CodeE=b.CodeE				
écoles	Order BY a.Nomdecole				
Sélectionner toutes les écoles	SELECT URN, GRANTFUNDING+				
ou on a résultat annuel négatif	SELFGENERATEDINCOME as Tot_Income,				
pour l'année 2018	TEACHINGSTAFF+SUPPLYTEACHERS+				
(revenus <dépenses)< td=""><td>EDUCATIONSUPPORTSTAFF+PREMISES+BACKOFFIC</td></dépenses)<>	EDUCATIONSUPPORTSTAFF+PREMISES+BACKOFFIC				
	E+CATERING+OTHERSTAFF+ENERGY+				
	LEARNINGRESOURCES+ICT+BOUGHTIN+ OTHER as				
	Tot_Expend FROM CFR Where GRANTFUNDING+				
	SELFGENERATEDINCOME <teachingstaff+suppl< td=""></teachingstaff+suppl<>				
	YTEACHERS+EDUCATIONSUPPORTSTAFF+PREMISES				
	+BACKOFFICE+				
	CATERING+OTHERSTAFF+ENERGY+LEARNINGRESO				
	URCES+ICT+ BOUGHTIN+OTHER and Annee=2018				
	SELECT URN, GRANTFUNDING+				
	SELFGENERATEDINCOME as Tot_Income,				
	TEACHINGSTAFF+SUPPLYTEACHERS+				
	EDUCATIONSUPPORTSTAFF+				
	PREMISES+BACKOFFICE+CATERING+OTHERSTAFF+				
	ENERGY+ LEARNINGRESOURCES+ICT+				
	BOUGHTIN+OTHER as Tot_Expend FROM CFR Where				
	GRANTFUNDING+SELFGENERATEDINCOME <teachi< td=""></teachi<>				
	NGSTAFF+SUPPLYTEACHERS+EDUCATIONSUPPORT				
	STAFF+PREMISES+BACKOFFICE+CATERING+				
	OTHERSTAFF+ENERGY+LEARNINGRESOURCES+ICT				
	+BOUGHTIN+OTHER and Annee=2018				
Supprimer le tableau	DROP TABLE ALLDATA				
ALLDATA					
Somme des élevés au système	SELECT sum(c.Eleves) as SumEleves				
éducatif par année	, c.Annee				
	FROM CFR c				
	GROUP BY c.Annee				
Trouver, s'il y a les directeurs	SELECT a.nomD & ' ' &a.prenomD as Directeurs				
qui vont retraiter bientôt	, a.DatedenaissanceD				
	, datediff("yyyy",a.DatedenaissanceD,date()) as Age				
	, "Attention: Retrait bientôt!" as Status				
	FROM Directeur a				
	WHERE datediff("yyyy",a.DatedenaissanceD,date())>60				

On va décrire de l'algèbre relationnel relative à deux requêtes.

1:

SELECT a.Nomdecole, b.NomE, b.Majeur

FROM Enseignement a, Enseignant b

WHERE a.CodeE=b.CodeE

ORDER BY a.Nomdecole

Dans cette requête SQL on doit avoir la liste des enseignants et leurs majeurs pour chaque ecole. En algèbre relationnel on a obtenu:

 $\Pi_{NomE, NomedEcole, Majeur}$  ( $\sigma_{a.CodeE=b.CodeE}$  (Enseignement  $\alpha$  Enseignant))

2:

SELECT b.\*, a.NORG, a.NORB, a.NUMFSM, c.NTEA, c.NTEAAS, c.Salaire

INTO ALLDATA

FROM (Census a

INNER JOIN CFR b

ON a.URN=b.URN)

INNER JOIN SWF c

ON b.URN=c.URN

 $\Pi_{b.*, a.NORG, a.NORB, a.NUMFSM, c.NTEA, c.NTEAAS, c.Salaire}$  ((Census  $\bowtie a._{URN=b.URN}$  CFR)

⋈ b.URN=c.URN SWF)

Où ⋈ est l'opérateur de jointure conditionnelle, qui représente INNER JOIN<sup>3</sup>, ou en termes d'opérateurs basiques:

 $\Pi_{b.*, a.NORG, a.NORB, a.NUMFSM, c.NTEA, c.NTEAAS, c.Salaire} \sigma_{c.URN=b.URN} (\sigma_{a.URN=b.URN} (Census \alpha CFR) \alpha SWF)$ 

#### 5. Conclusion

En effectuant ce projet a été développée une base de données consistant en des KPIs pour les écoles du Royaume Uni. En prenant en compte les utilisateurs finaux nous avons continué à utiliser un système basé sur différentes entités pour les données financières, et les données quantitatives des élèves et enseignants. Cette entité consiste en des données de panel – les statistiques descriptives des écoles. Un échantillon de données a été utilisé dans ce projet car l'utilisation d'un jeu de données complet est impossible (en termes de limitation de volume final de projet) en raison de la taille des données (23K enregistrements par an).

Dans ce projet, le modèle Entité-Association a été formulé, sa transformation au modèle relationnel a été effectuée ainsi que l'implémentation dans Access de ce modèle. A l'aide d'Excel et Access ont été faites plusieurs requêtes SQL pour l'analyse de données et l'aide décisionnelle. Deux requêtes SQL de différents types sont présentes en termes d'algèbre relationnel.

<sup>&</sup>lt;sup>3</sup> https://www.geeksforgeeks.org/extended-operators-in-relational-algebra/

# Annexe A : Méta-données pour l'explication des attributs

Tableau A.1 CFR

Nombre d'élèves prévu
Nombre d'élèves ayant droit aux repas gratuits
Les parties de revenus
Les parties de dépenses

Tableau A.2 SWF

SWF	
URN	Identifiant unique
NTEA	Nombre total d'enseignants et d'enseignants dans l'équipe de management
NTEAAS	Nombre total d'assistants enseignants
NNONTEA	Nombre total de personnel non enseignant, à l'exclusion des assistants
Nalaire	Salaire brut de tous les enseignants (incluant les enseignants dans le l'équipe de management)

Tableau A.3 Census

Census	
URN	Identifiant unique
NORG	Nombre de filles en étude
NORB	Nombre de garçons en étude
PNUMFSM	Pourcentage d'élèves éligibles aux repas scolaires gratuits

Annexe B: Réalisation en Access

4	URN -	Year	•	NORG -	NORB -	NUMFSM -	Nomdecole -
	100000	2011				46	Sir John Cass's
	100000	2012		121	111	41	Sir John Cass's
	100000	2013		123	110	47	Sir John Cass's
	100000	2014		118	115	39	Sir John Cass's
	100000	2015		117	123		Sir John Cass's

Figure B.1 Réalisation de Census

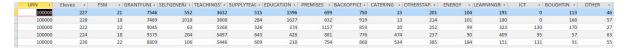


Figure B.2 Réalisation de CFR

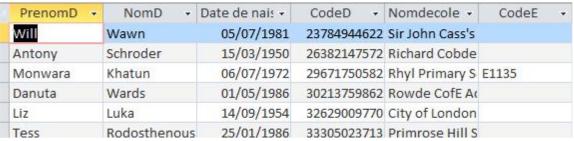


Figure B.3 Réalisation de Directeur



Figure B.4 Réalisation d'Ecole

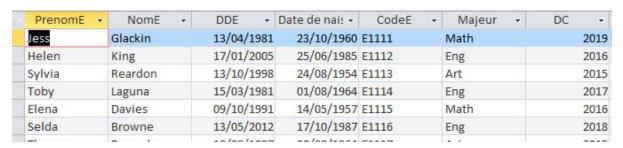


Figure B.5 Réalisation d'Enseignant



Figure B.5 Réalisation d'Enseignement

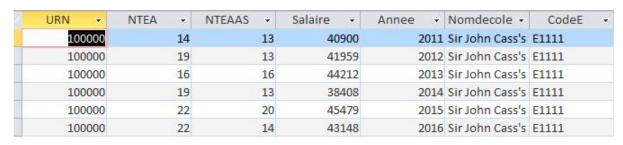


Figure B.5 Réalisation de SWF