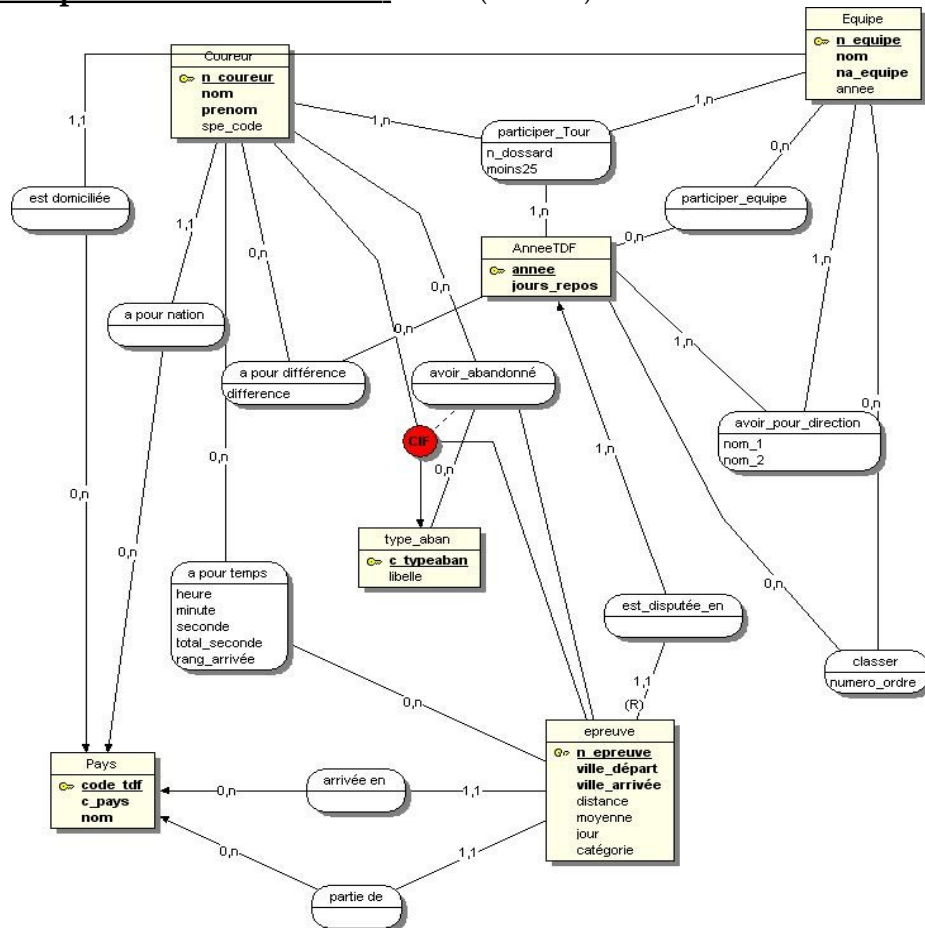


Étude de Cas TDF "Tour de France "

1 Introduction

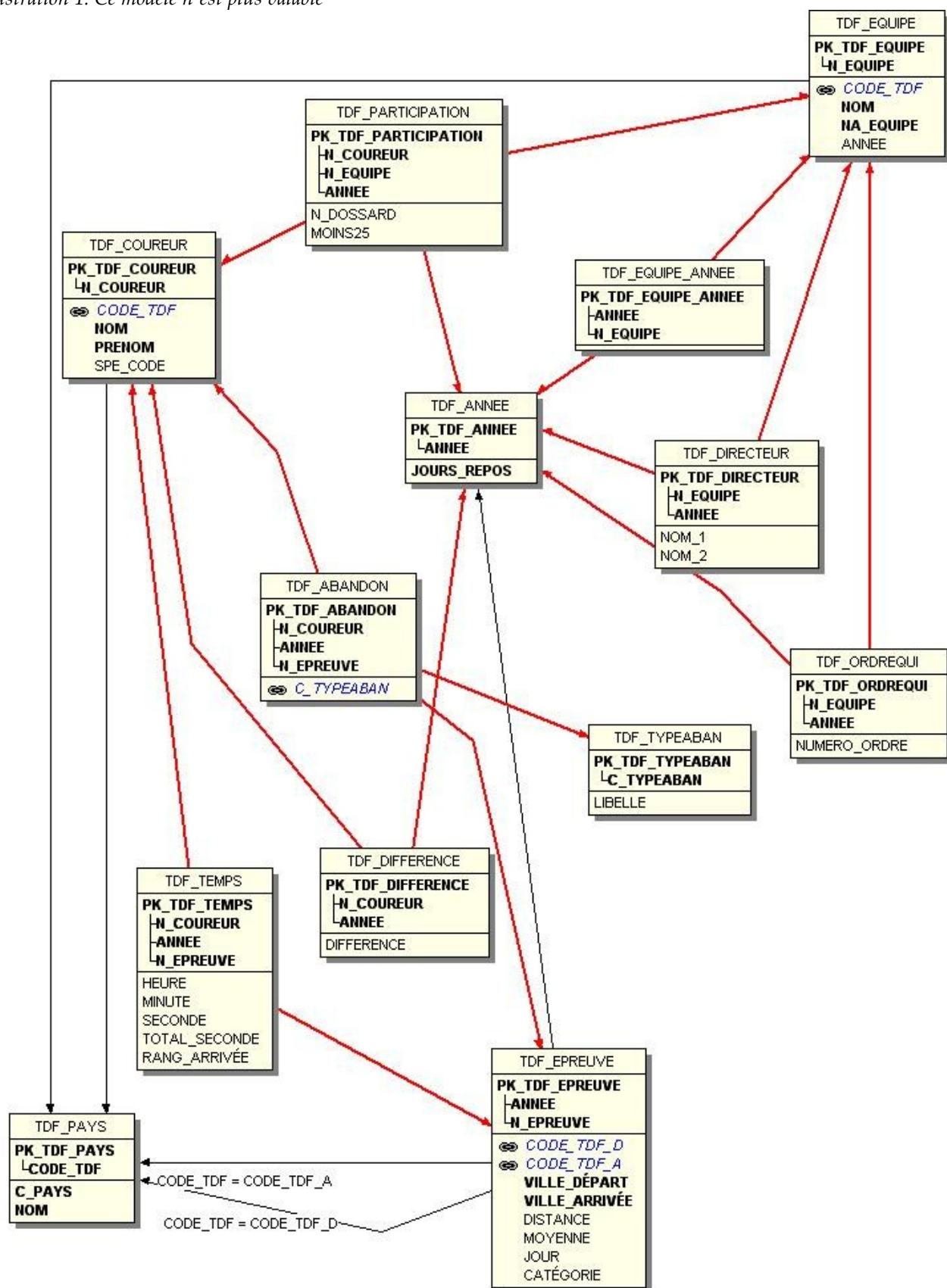
1. Un Tour de France est composé d'un certain nombre d'épreuves (" prologue ", " étape en ligne " et " course contre la montre " individuelle ou par équipe).
2. Les épreuves sont numérotées à partir de 1 (ou 0 quand un prologue est programmé).
3. Une épreuve est une distance à parcourir entre une ville départ et une ville arrivée. Depuis 1992, chaque épreuve se déroule à une date différente.
4. Les coureurs qui participent au Tour De France le font sous un nom d'équipe. Une équipe regroupe 9 coureurs maximum. Elle est dirigée par deux directeurs sportifs (max) et est classée au niveau international (voir chap 5).
5. Chaque coureur se voit attribuer un n° de dossard avant le départ du Tour.
6. Un coureur qui est inscrit au Tour et qui ne prend pas le départ d'une épreuve ou qui ne termine pas une épreuve est considéré comme ayant " abandonné ". Les causes d'abandons sont multiples.
7. Le coureur qui termine une épreuve dans les temps se voit attribuer un temps d'arrivée et un rang d'arrivée.
8. Les coureurs ont tous une nationalité ainsi que les sponsors appartenant à des équipes.
9. **Le nom d'une équipe est en fait le nom de son sponsor**
10. Une équipe peut changer de sponsors
11. Une équipe peut disparaître et éventuellement donner naissance à une nouvelle équipe (voir chap 6)
12. En général, les meilleures équipes du monde participe au tour.
13. La base de données Tour De France gérée avec le SGBD Oracle comprend 30 tours de 1988 à 2017. Elle contient environ 125 000 lignes (15MB).

2 Le modèle conceptuel de données V1.1 : MCD (obsolète)



3 Le modèle logique de données relationnel V1.1 : MLD (obsolète)

Illustration 1: Ce modèle n'est plus valable



4 Le modèle conceptuel de données V2.2 : MCD

Plusieurs défauts ont nécessité la mise à jour de la base de données

- quand un nom d'équipe change, ce n'est pas forcément une nouvelle équipe mais parfois seulement un nouveau sponsor
- les noms des directeurs étaient dupliqués dans tdf_directeur (avec de nombreuses erreurs)
- la participation des équipes nécessitaient la mise à jour de 3 tables (tdf_directeur, tdf_equipe_annee, tdf_participation)
- le champ moins25 a été remplacé par jeune ; le contenu a été mis à jour
- un champ année de naissance a été ajouté ; la mise à jour de la base reste à faire
- Ajout de la CIF sur " participer au tour " déjà présente dans la version 1.0 puis oubliée
- Ajout d'un attribut commentaire dans abandon (v2.2)
- Ajout d'un attribut commentaire dans type_aban (v2.2)
- Ajout d'un attribut valide dans participation (v2.2)
- Suppression de SPE_CODE car un coureur n'a pas la même spécialité toute sa carrière.

Une modification de la structure d'une base de données nécessite un travail assez complexe (voire fastidieux) de migration des données.

5 Le modèle conceptuel de données V2.3 : MCD

Le classement des équipes avait été oublié pendant quelques années car les règles de classement ont souvent changé et qu'il était difficile de retrouver les anciens classements. L'association « classer » donnant la relation « tdf_ordrequi » a été complétée pour tenir compte de ses différents classements. On distingue :

- La Coupe du monde de cyclisme sur route, qui a été organisée de 1989 jusqu'en 2004, rassemblait un ensemble d'épreuves de cyclisme sur route sous l'égide de l'UCI. Contrairement aux Championnats du monde, la coupe du monde regroupait plusieurs manches étalées dans le temps et dans l'espace. La Coupe du monde disparaît avec la création du ProTour et des circuits continentaux en 2005.
- En 1994, l'UCI met en place les classements UCI par équipe qui prennent en compte toutes les courses disputées sur les cinq continents. Jusqu'en 1998, les équipes cyclistes sont classées par l'UCI dans une division unique. En 1999 le classement UCI par équipes est divisé entre GSI, GSII et GSIII
- L'UCI met en place en 2005 un nouveau système d'élite qui contraint toutes les équipes à participer à toutes les épreuves du calendrier « UCI ProTour ». Aucun cycliste n'étant capable de courir toutes ces épreuves, un roulement se fera au sein de ces équipes. Chaque épreuve donne des points et un classement permettra de connaître la meilleure équipe de l'année.
- En partir de 2011, ce classement sera nommé UCI World Tour. Un maximum de dix-huit équipes professionnelles possèdent une licence l'UCI World Tour et sont appelées UCI ProTeams. Des équipes n'ayant pas le statut de ProTeam peuvent participer à ces courses mais sans y marquer de points.

On a ajouté au modèle un attribut score (nombre de points de l'équipe) et type_classement pour distinguer les classements détaillés ci-dessus.

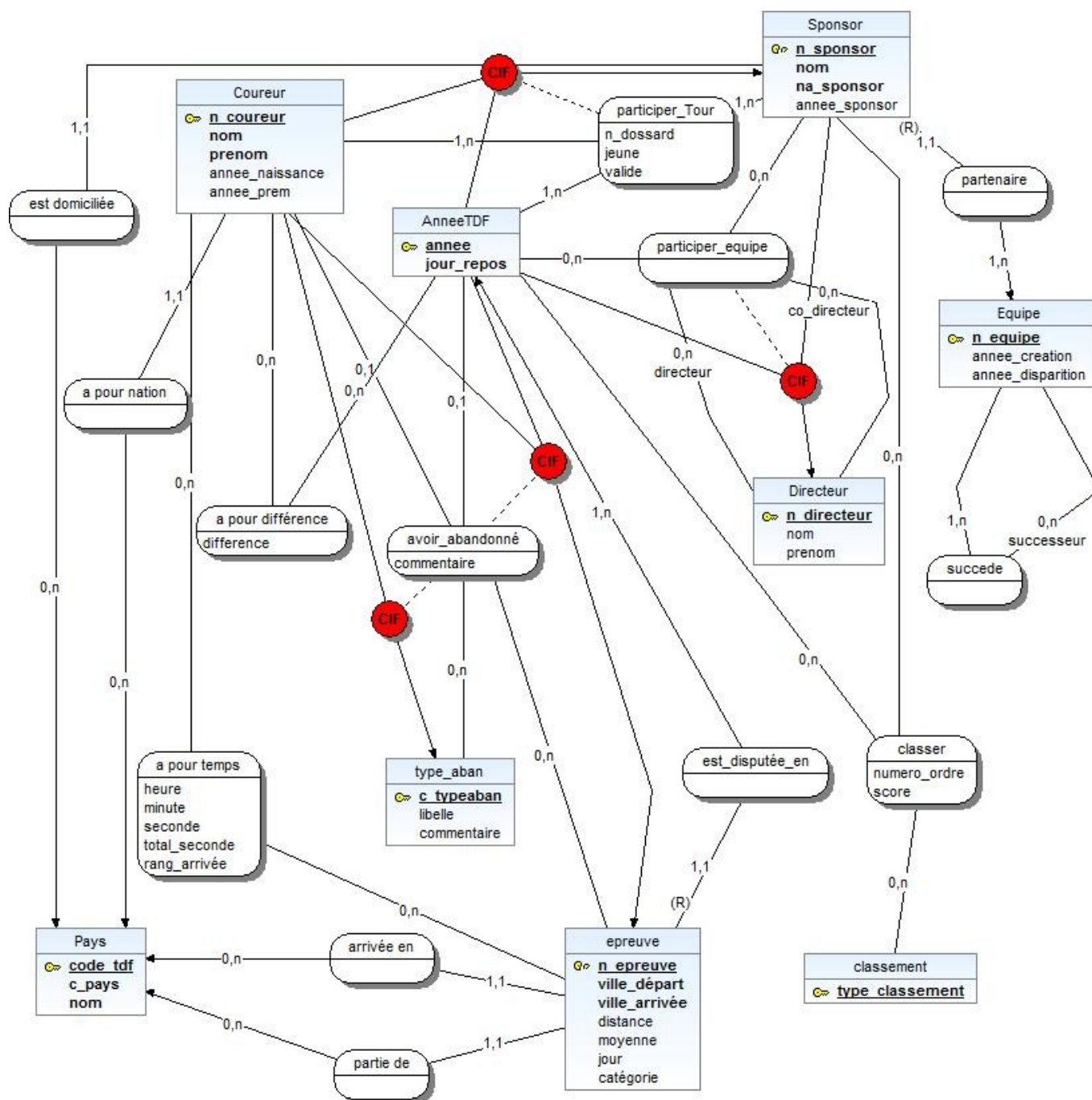
Toutes les équipes ayant un classement WORLD CUP sont enregistrées, **même celles qui n'ont jamais disputé le tour.**

Les équipes classées à l'UCI étant trop nombreuses, on n'a gardé que les 20 premières de chaque année ainsi que les classements de celles qui ont disputé au moins une fois le tour (même avec un sponsor différent).

Pour régler les modifications de classement des étapes dues aux sanctions post-tour, on a ajouté dans tdf_temps une colonne rang_initial qui contient le classement d'un coureur avant correction.

6 Le modèle conceptuel de données V2.4 (V2.5) : MCD

Il reste toujours une ambiguïté entre les équipes changeant de sponsor (ou de nationalité comme " Astana ") et les nouvelles équipes créées sur les cendres d'une ancienne. Il a donc été ajouté un lien successeur permettant de faire le lien entre ces équipes. Il arrive aussi qu'une équipe en engendre deux (ou plus ?) autres ou que 2 équipes fusionnent. Il a donc été décidé de créer une association réflexive avec l'entité équipe (" succède à ").

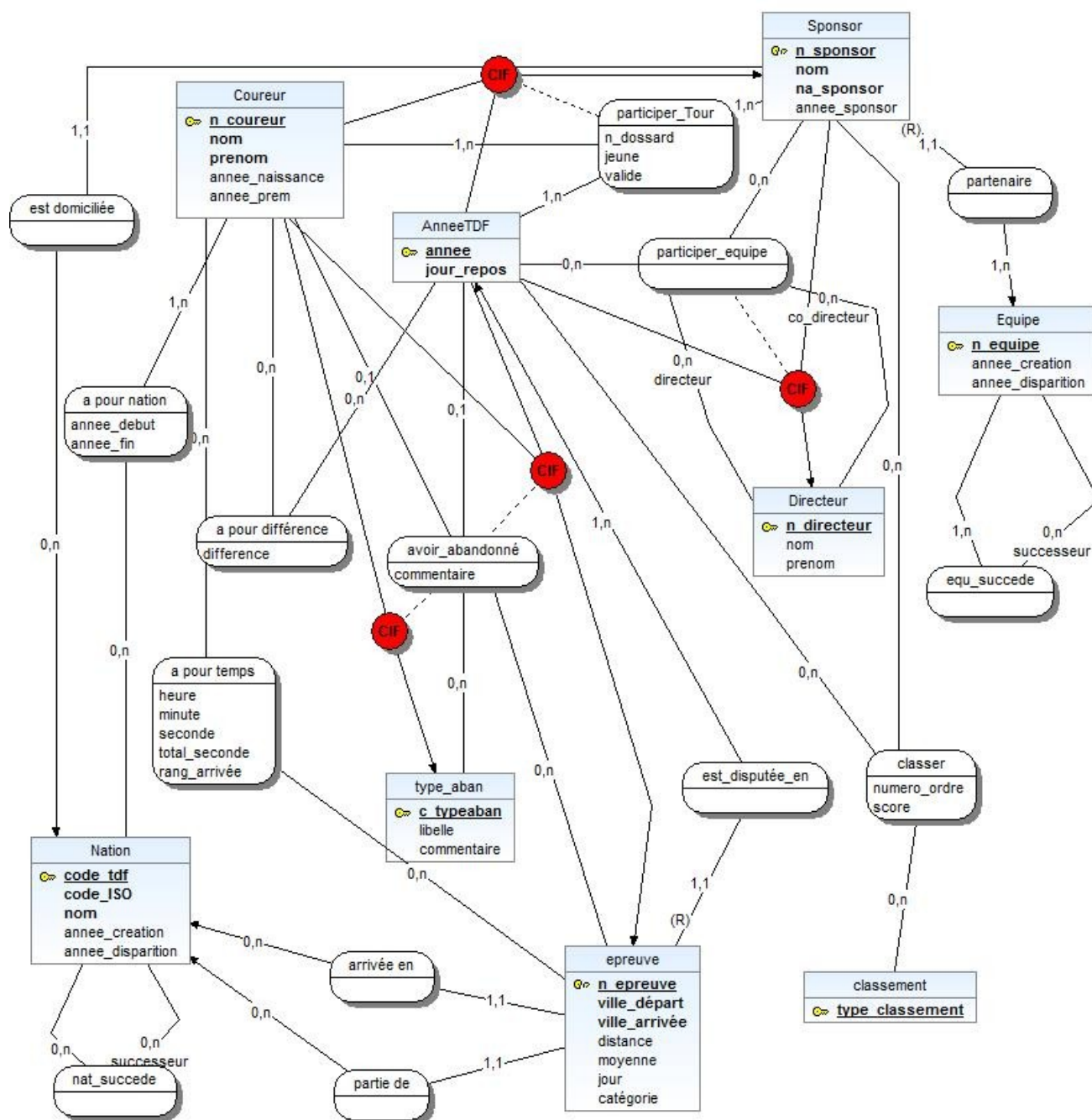


8 Le modèle conceptuel de données V2.6 : MCD

En février 2017, un groupe de projet 1A2A (Yves Le Mitouard, Léonard Commelin, Noé-Simon Etoungou N Gono, Colin Thorel) s'est constitué avec pour but de réaliser l'appli web du tour de France. En travaillant sur le sujet, Yves Le Mitouard a mis en évidence que la base ne permettait pas d'enregistrer les changements de nations des coureurs et des pays (ex URSS → Russie).

Il a donc fallu modifier le modèle :

- L'association "A pour nation" se voit ajouter une année de début et de fin permettant à un coureur de changer de nationalité (et même plusieurs fois).
- La propriété "code_tdf" doit être retirée de la table coureur.
- L'entité Pays devient l'entité Nation. On ajoute une année de création et de disparition de cette nation. Au passage, C_pays devient Code_Iso (Code ISO 3166).
- Une association réflexive "nat_succede" permettra de lier les nations en fonction de leur histoire.



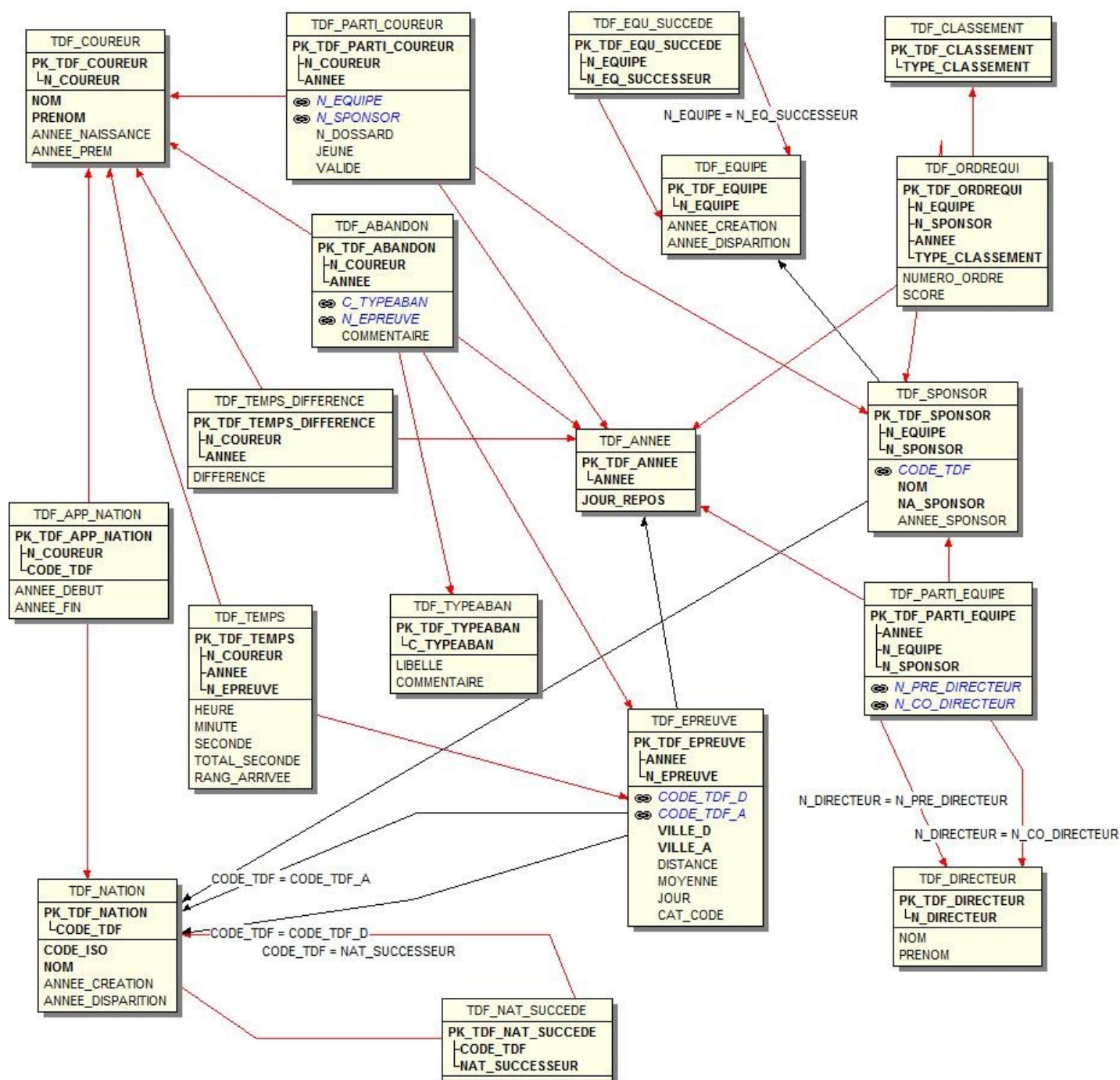
Les nations ayant changé de nom (entre autre) sont :

- L'URSS existe de 1922 à 1991. Ces pays la composait :
 - 1 L'Arménie de 1920 à 1991
 - 2 L'Azerbaïdjan de 1920 à 1991
 - 3 La Biélorussie de 1922 à 1991
 - 4 L'Estonie de 1940 à 1991
 - 5 La Géorgie de 1921 à 1991
 - 6 Le Kazakhstan de 1936 à
 - 7 Le Kirghizistan de 1924 à
 - 8 La Lettonie de 1940 à 1991
 - 9 La Lituanie de 1940 à 1991
 - 10 La Moldavie de 1940 à 1991
 - 11 La Russie de 1918 à 1991
 - 12 Le Tadjikistan de 1924 à 1991
 - 13 Le Turkménistan de 1922 à 1991
 - 14 L'Ukraine de 1919 à 1991
 - 15 L'Ouzbékistan de 1940 à 1991
- Une partie de la RFA fut la RDA de 1949 à 1990
- La Tchécoslovaquie a existé de 1918 à 1992. Depuis 1993, on distingue la Slovaquie de la république tchèque ou Tchéquie.
- La Yougoslavie de 1929 à 2003
 - 4 pays quittent La Yougoslavie : la Slovénie (25/06/1991), la Croatie (25/06/1991), la Bosnie-Herzégovine (01/03/1992) et la Macédoine (8/09/1991)
 - elle s'appelle à partir de 2003 la Communauté d'États Serbie-et-Monténégro
 - En 2006, La Serbie et le Monténégro se séparent

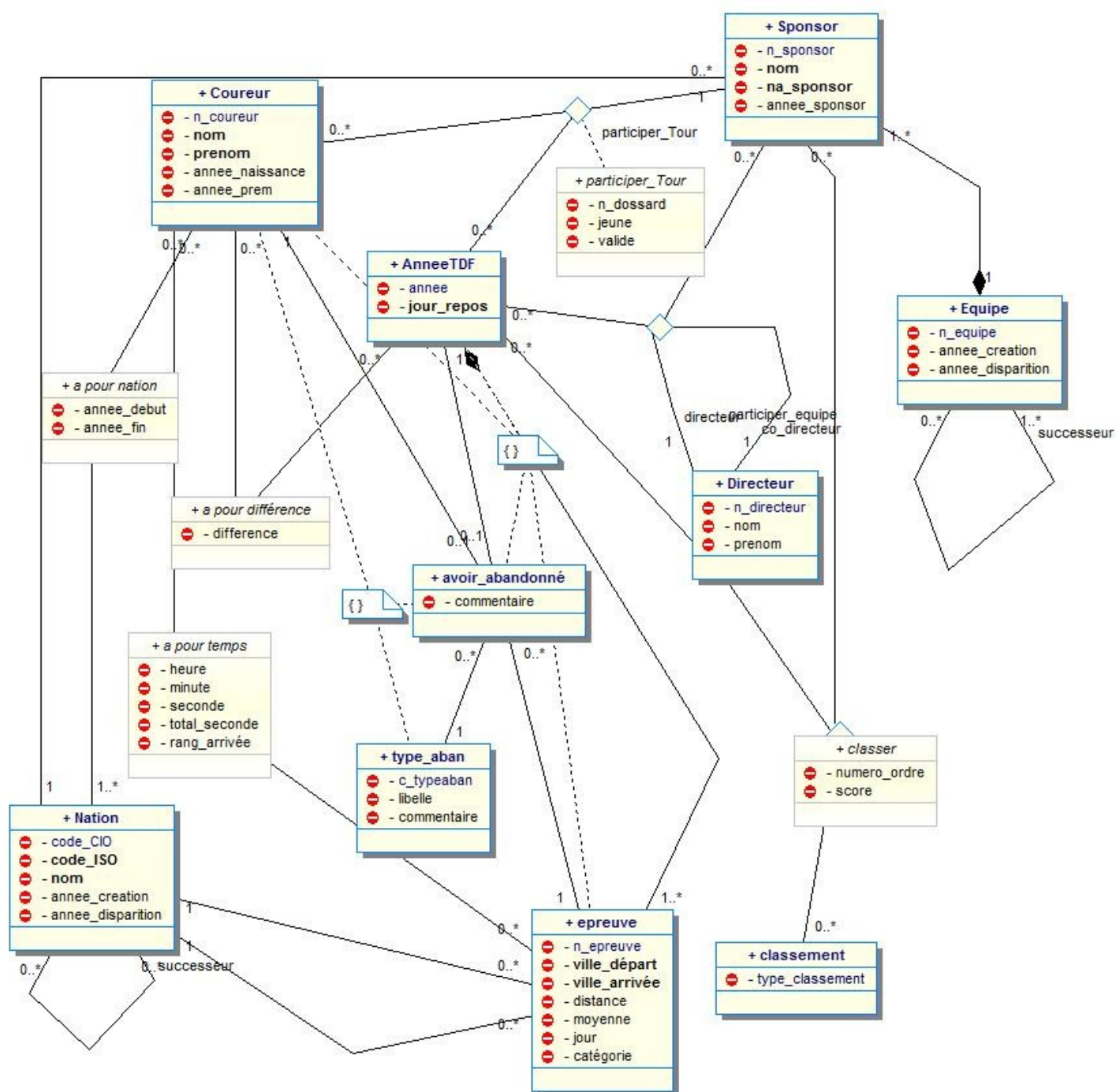
Cette version nécessiterait aussi un certain nombre d'améliorations (V2.6 ?)

- Les homonymes ne sont pas gérés
- Certains coureurs ont été connus sous différents noms. Exemple : Kanstantsin Siutsou ou Sivtsov (en biélorusse : Канстанцін Сіўцов),
- Certains anciens coureurs sont maintenant directeurs. Ils sont inscrits deux fois dans la base
- ...

9 Le modèle logique de données relationnel V2.6 : MLD



10 Le diagramme de classe UML : NON A JOUR



11 Le dictionnaire des données : liste des attributs issus du MLD/R

Noms des relations de la base « TDF »	commentaires
TDF_ABANDON	Cette relation référence les abandons des coureurs dans les différents « Tour de France ».
TDF_ANNEE	Cette relation référence les différences années de « Tour de France » ainsi que les jours de repos associés.
TDF_APP_NATION	Un coureur appartient à une nation à un moment donné
TDF_COMMENTAIRE	Cette relation référence les commentaires de saisie des données inscrites dans la base de données « Tour de France ».
TDF_COUREUR	Cette relation référence les coureurs.
TDF_DIRECTEUR	Cette relation référence les caractéristiques des directeurs d'équipe.
TDF_EPREUVE	Cette relation référence toutes les épreuves de tous les « Tour de France ».
TDF_EQU_SUCCEDE	Cette relation fait le lien entre une équipe et celle qui éventuellement lui succède. On parle bien de nouvelle équipe et pas de nouveau sponsor.
TDF_EQUIPE	Cette relation référence les équipes.
TDF_NAT_SUCCEDE	Cette relation fait le lien entre une nation et celle qui éventuellement lui succède. Ex : L'Allemagne succède à la RDA
TDF_NATION	Cette relation référence les pays d'origine des coureurs et des équipes. Il faut noter qu'il n'y a plus d'équipes nationales composées de coureurs de la même nationalité que celle de l'équipe.
TDF_ORDREQUI	Cette relation référence le classement des équipes au niveau international (classement non obtenu ni en 2004, ni en 2005).
TDF_PARTI_COUREUR	Cette relation référence aux coureurs qui ont participé au Tour de France avec une équipe.
TDF_PARTI_EQUIPE	Cette relation référence les équipes qui ont participé à différents "Tour de France" avec leurs directeurs sportifs.
TDF_SPONSOR	Cette relation référence les noms d'une équipe en fonction du sponsor et de l'année
TDF_TEMPS	Cette relation référence tous les temps d'arrivée de tous les coureurs dans tous les « Tour de France ».
TDF_TEMPS_DIFFERENCE	Cette relation référence les différences de temps constatées par rapport au temps du classement général. Ces différences sont dues au fait que les bonifications et les pénalités ne sont pas prises en compte dans la base de données.
TDF_TYPEABAN	Cette relation référence les différents types d'abandons possibles.

Nom attribut	signification	valeur
ANNEE ...	Année du Tour De France	19XX ... 2XXX
ANNEE_CREATION (EQUIPE)	Année de création d'une équipe	19XX-2XXX
ANNEE_CREATION (NATION)	Année de création d'une nation	
ANNEE_DEBUT	Année où commence une nationalité pour un coureur	
ANNEE_DISPARITION (EQUIPE)	Année de disparition d'une équipe	
ANNEE_DISPARITION (NATION)	Année de disparition d'une nation	
ANNEE_FIN	Année où se termine une nationalité pour un coureur	
ANNEE_NAISSANCE	Année de naissance du coureur	
ANNEE_PREM	Année première participation d'un coureur	19XX-20XX (redondant)
ANNEE_SPONSOR	Année de début d'un nom (sponsors) d'une équipe	
C_TYPEABAN	Code de l'abandon	
CAT_CODE	Catégorie de l'épreuve	PRO,ETA, CME et CMI
CODE_CIO (APP_NATION)	Nationalité d'un coureur	
CODE_CIO (EQUIPE)	Nationalité de l'équipe sponsorisée	
CODE_CIO_A	Code pays arrivée épreuve	
CODE_CIO_D	Code pays départ épreuve	
CODE_ISO	Code pays français sur 2 lettres	Codet ISO 3166-1-alpha-2
CODE-TDF (COUREUR)	Nationalité du coureur	
COMMENTAIRE (ABANDON)	Raisons de l'abandon	
COMMENTAIRE (COMMENTAIRE)	Informations sur un tour de France	
COMMENTAIRE (TYPEABAN)	Décodage du type d'abandon	
DISTANCE	Distance parcourue épreuve	
HEURE	Temps d'arrivée : heure	
JEUNE	Coureur ayant 25 ans au maximum	'o' si 25 ans max; null si plus de 25 ans
JOUR	Date de l'épreuve	
JOUR_REPOS	Nombre de jours de repos	
LIBELLE	Libellé de l'abandon	
MINUTE	Temps d'arrivée : minute	
MOYENNE	Moyenne de l'épreuve	null si moyenne non renseignée
N_COUREUR	N° du coureur	
N_DIRECTEUR	N° d'un directeur sportif (premier ou co-directeur)	
N_DOSSARD	N° de dossard lors d'une participation à un tour	
N_EPREUVE	N° de l'épreuve 0 à 21	
N_EQUIPE	N° de l'équipe	
N_SPONSOR	N° du sponsor pour une équipe donnée	
N_SUCCESSEUR	N° de l'équipe succédant à une ancienne	
NA_EQUIPE	Nom abrégé de l'équipe	
NAT_SUCCESSEUR	N° de la nation succédant à une ancienne	
NOM (DIRECTEUR)	Nom d'un directeur sportif	Null si non connu
NOM (COUREUR)	Nom du coureur	
NOM (EQUIPE)	Nom de l'équipe	
NOM (PAYS)	nom du pays	
NUMERO_D'ORDRE	Classement international UCI	
PRENOM	Prénom du coureur	
PRENOM (DIRECTEUR)	Prénom d'un directeur sportif	Null si non connu
RANG-ARRIVEE	Rang d'arrivée dans une épreuve	
SECONDE	Temps d'arrivée : seconde	
SPE_CODE	Spécialité du coureur	GPR,SPR,CPT (ce champ a disparu)
TOTAL_SECONDE	Temps d'arrivée total en secondes	
TYPE_CLASSEMENT	catégorie de classement d'une équipe (chap 5)	
VALIDE	La participation du coureur est elle valide (affaires ...)	{'O','R'}
VILLE_A	Ville d'arrivée de l'épreuve	
VILLE_D	Ville départ de l'épreuve	

12 Décisions disciplinaires (dopage)

12.1 Radiations

Levi Leipheimer (Omega Pharma-Quickstep), Christian Vande Velde (Garmin-Sharp), David Zabriskie (Garmin-Sharp), Tom Danielson (Garmin-Sharp), George Hincapie (BMC) et Michael Barry (Sky), tous ayant porté les couleurs de l'équipe US Postal et étant mis en cause par l'USADA dans un rapport accablant visant essentiellement Lance Armstrong, vont être suspendus pour une durée de six mois, avec rétroactivité pour le mois de septembre. Leurs palmarès est également remis en cause.

01/08/1998	vie	Lance Armstrong
01/06/1999	30/07/2006	Levi Leipheimer
07/07/2007	29/07/2007	
31/05/2004	31/07/2006	George Hincapie
04/06/2004	31/04/2006	Christian Vande Velde
31/05/2003	31/07/2006	David Zabriskie
01/03/2005	23/09/2006	Tom Danielson
13/05/2003	31/07/2006	Michael Barry
17/05/09	01/05/12	Franco Pellizotti

En février 2012, Jan Ullrich a également vu ses résultats obtenus du 1er mai 2005 au 26 février 2007 retirés par le Tribunal arbitral du sport.

En octobre 2012, l'UCI annonce qu'une procédure disciplinaire à l'encontre du coureur espagnol Carlos Barredo est ouverte suite à des anomalies dans son passeport biologique²³. Il est finalement suspendu deux ans et est disqualifié des courses auxquelles il a participé entre le 26 octobre 2007 et le 24 septembre 2011 dont sa 89^e place sur cette édition du Tour de France qui reste vacante

Certains résultats de ces coureurs sont annulés. Pas de reclassement des autres coureurs.

Dans la base TDF, ces coureurs ont un statut 'R' (radié) dans la table participation pour les années incriminées.

12.2 Disqualifications

1998 :

- L'affaire Festina commence peu avant cette édition du Tour et entraîne l'exclusion de la totalité des coureurs de l'équipe Festina pour dopage.
- Le 31 juillet 1998, l'équipe TVM se retire de la course suite à une nouvelle perquisition à l'hôtel de l'équipe après la mise en examen du directeur sportif et du médecin de l'équipe.
- Le 29 juillet 1998, les coureurs mettent pied à terre dès le kilomètre 32, afin de protester contre « les méthodes indignes de la police ». Après négociation, seulement une partie du peloton repart, alors que 5 équipes (Once, Banesto, Vitalicio, Kelme et Riso Scotti) préfèrent abandonner le Tour.

2001 :

- Le 14 juillet 2001, on apprend que le coureur espagnol Txema del Olmo a été contrôlé positif à l'EPO le 7 juillet, jour du prologue du Tour. Il est immédiatement exclu de la course

2005 :

- Evgueni Petrov non partant pour la 10^e étape pour cause de contrôle sanguin non conforme

2006 :

- Floyd Landis est exclus à partir de la première étape. Son rang d'arrivée aux étapes est placé à -1

2007 :

- Alexandre Vinokourov, leader de l'équipe Astana, a été contrôlé positif aux transfusions homologues samedi 21 juillet à l'issue du contre-la-montre individuel du Tour de France qu'il a remporté à Albi.

Son équipe quitte le tour à l'issue de l'étape de repos du 24/07

- À la suite du contrôle positif à la testostérone de Cristian Moreni sur la 11e étape, son équipe Cofidis se retire d'elle-même du Tour de France, le 24 juillet.

2008 :

- L'Espagnol Manuel Beltran, coureur de l'équipe italienne Liquigas, est contrôlé positif à l'issue de la première étape¹⁴. Il est mis hors-course par son équipe après la 7e étape.
- L'Espagnol Moisés Dueñas est également contrôlé positif à l'EPO à l'issue de l'étape du 8 juillet. Le coureur de l'équipe Barloworld, est mis hors course le 16 juillet par son équipe dès l'annonce de ce résultat.
- Le Kazakh Dmitriy Fofonov a été contrôlé positif à l'heptaminol à l'issue de la 18e étape et exclu quelques heures après l'arrivée de la dernière étape.
- Des contrôles de l'AFLD révèlent une positivité au CERA des échantillons de Stefan Schumacher, vainqueur de deux contre la montre, Leonardo Piepoli, vainqueur d'étape, et Bernhard Kohl, vainqueur du maillot à pois et troisième au classement général.
- Enfin l'Italien Riccardo Riccò est contrôlé positif au CERA, une EPO de troisième génération

2009

- L'espagnol Mikel Astarloza a été contrôlé positif à l'EPO suite à un contrôle hors compétition juste avant le tour, il risque deux ans de suspension, et sa victoire sur la 16e étape du Tour de France 2009 a été invalidée au profit de Sandy Casar
- Denis Menchov est exclus à partir de l'étape 1. Son rang d'arrivée aux étapes est placé à -1
- Franco Pellizotti : le 3 mai 2010, l'Union cycliste internationale demande l'ouverture d'une procédure disciplinaire à son encontre pour violation apparente du règlement antidopage. Il perd le bénéfice de ses résultats obtenus depuis le 17 mai 2009, date de début de la constatation de données anormales.

2010

- Alberto Contador est disqualifié à partir de l'étape 0. Son rang d'arrivée aux étapes est placé à -1. Les coureurs sont reclassés afin que tous les points UCI qu'offrent la course soient répartis mais la 21^e place reste vacante
- Denis Menchov est disqualifié à partir de l'étape 0. Son rang d'arrivée aux étapes est placé à -1
- Conformément à la révision du 1er septembre 2013 de l'article 2.6.037 du règlement de l'UCI, seules les trois premières places sont alors réattribuées, les autres places ne sont pas modifiées, Ainsi, comme 2 des 3 premiers sont disqualifiés, quatrième place reste vacante.

2011

- Alberto Contador est disqualifié à partir de l'étape 0 (suspendu de 2010 à 2011). Son rang d'arrivée aux étapes est placé à -1
- Le 11 juillet 2011, le coureur Alexandr Kolobnev (équipe Katusha) est exclu du Tour de France à la suite d'un contrôle positif à un diurétique, l'hydrochlorothiazide (HCT), le 6 juillet lors de la 5e étape

2012

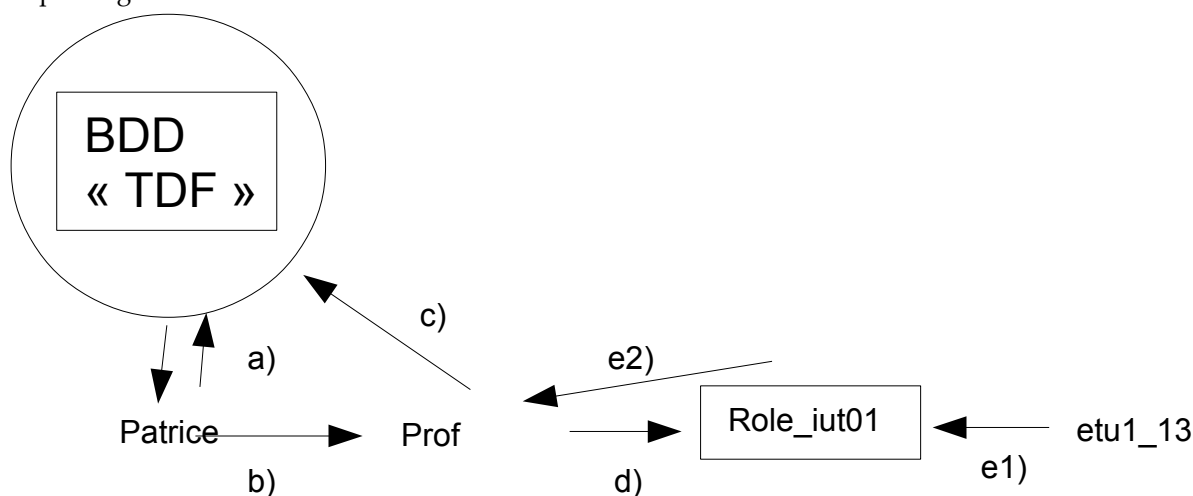
- Denis Menchov est disqualifié à partir de l'étape 0. Son rang d'arrivée aux étapes est placé à -1
- Le 10 juillet 2012, Rémy Di Grégorio est arrêté durant le Tour de France à l'hôtel de l'équipe Cofidis à Bourg-en-Bresse dans le cadre d'une affaire liée au dopage. Il aurait été interpellé suite à une conversation avec un fournisseur présumé
- En octobre 2012, Lance Armstrong, 23e, et Carlos Barredo, 41e, sont aussi déclassés, leurs places respectives au classement sont laissées vacantes.

13 Accès aux données de la base

13.1 Les privilèges "système" et les utilisateurs (administrateur et utilisateur d'application)

Privilèges Système	Utilisateurs	Remarque
DBA,CONNECT, RESOURCE ...	SYS	administrateur d'Oracle
DBA,CONNECT, RESOURCE ...	SYSTEM	administrateur des bases Oracle
DBA,CONNECT, RESOURCE	Patrice	propriétaire de la base TDF
CONNECT, RESOURCE ...	Eric	propriétaire des bases examen
CONNECT	Prof	propriétaire des vues TDF
CONNECT, RESOURCE	copie_tdf	propriétaire d'une copie des tables TDF
CONNECT, RESOURCE	ETU1_01 à ETU1_80	consultant des vues TDF
CONNECT, RESOURCE	ETU2_01 à ETU2_80	consultant des vues TDF

13.2 La base de données « Tour de France », utilisateurs d'application (propriétaire des données ou non), tables et privilèges



- (a) L'utilisateur " Patrice " est l'administrateur et le propriétaire des données de la base de données " TDF ". Il a donc la possibilité de mettre à jour la base de données.
- (b) L'utilisateur " Patrice " a accordé des privilèges de lecture à l'utilisateur "prof ".
- (c) "prof " peut consulter les données mis à disposition par l'utilisateur " Patrice ". Il a créé des vues sur les tables de " Patrice "
- (d) "prof " a accordé des privilèges de lecture sur ses vues à un groupe de privilèges (ou rôle) role_etu01.
- (e) Les utilisateurs attachés au groupe " ETU1_XX " peuvent, au travers du rôle, avoir accès en lecture à certaines données de la base " Tour de France ".
- (f)

Toutes les tables (TDF_XXX) appartiennent au propriétaire " Patrice ". "Prof " possède des vues (VT_XXX) sur ces tables. Ex : La vue VT_PAYS permet d'accéder à des données de la table TDF_PAYS.

Prof a accordé des privilèges de lecture. Les noms des objets cités dans les requêtes et permettant de visualiser les données seront donc préfixés par « prof » :

ex : select * from prof.vt_typeaban.

Prof possède en outre des vues vt_classement_XXXX et des vues vt_temps_XXXX permettant de simplifier les requêtes.

Remarque : il est possible de simplifier l'écriture en créant des synonymes.