

TP 1

Le Tour de France 2011

Ce premier TP porte exclusivement sur de l'extraction de données. Cette partie du langage est, de loin, la plus utilisée lorsque l'on travaille avec un SGBDR.

► Prérequis

Avoir compris le cours de la *séquence 2*, avoir fait tous les exercices et s'être autocorrigé.

► Capacités attendues en fin de TP

Être capable d'extraire des informations d'une base de données en SQL et être capable de modifier et d'adapter des scripts existants.

► Contenu

1. Contexte d'étude.....	6
2. Requêtes simples.....	8
3. Requêtes intermédiaires	10
4. Requêtes complexes.....	12

TP 1

Le Tour de France
2011

Page 5

1. Contexte d'étude



Le Tour de France est une compétition cycliste par étapes créée en 1903 qui se déroule chaque année en France, au mois de juillet. Cette épreuve de plus de 3 000 km est considérée comme la plus prestigieuse épreuve cycliste du monde.

Pour ce TP, vous allez imaginer que vous assistez un journaliste sportif qui prépare une émission sur le Tour de France. Votre prédécesseur a préparé des fiches avec les réponses à des questions précises sur le Tour, mais il a omis de recopier les requêtes SQL qui lui ont permis d'obtenir ces résultats.

Vous êtes donc chargé de trouver les requêtes SQL qui ont permis d'obtenir certains résultats. Pour cela, vous disposez d'une base de données qu'a conçue votre prédécesseur qui contient toutes les données recueillies pendant le Tour de France 2011 qui s'est déroulé du 2 au 24 juillet 2011. Vous disposez également des notes qu'il a laissées qui contiennent :

- Un schéma conceptuel des données (respectant le formalisme du MCD de la méthode d'analyse Merise) ;
- Un diagramme de classes (respectant le formalisme issu du langage UML) ;
- Un schéma relationnel ;
- Quelques requêtes.

Schéma conceptuel des données :

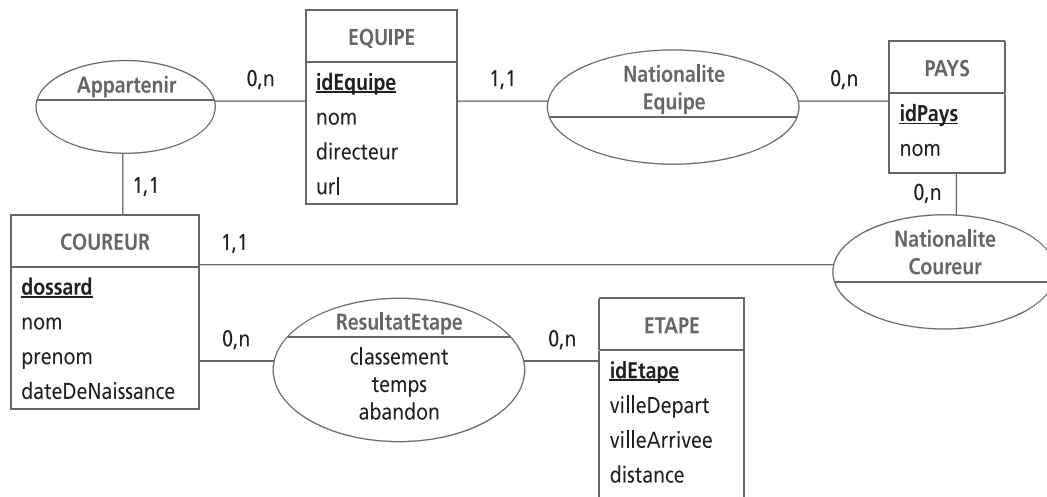
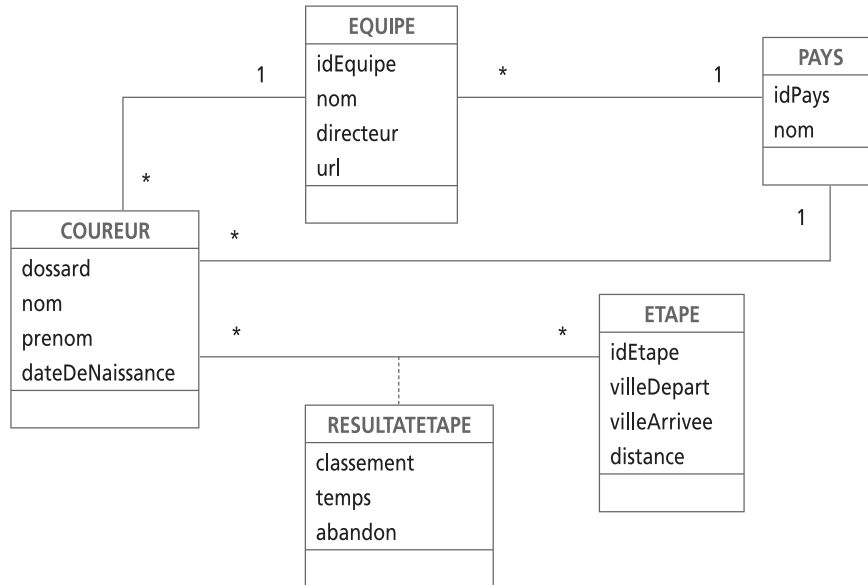


Diagramme de classes :



Règles de gestion présentes dans les deux schémas ci-dessus :

- Un coureur possède une nationalité et fait partie d'une équipe (qui n'est pas forcément de la même nationalité que le coureur).
- Une équipe possède une nationalité (celle de son principal sponsor).
- Un coureur participe à des étapes et en retire un classement et un temps (sauf en cas d'abandon).

TP 1

Le Tour de France
2011

Schéma relationnel :

```

PAYS (idPays, nom)
  idPays : Clé primaire

EQUIPE (idEquipe, nom, directeur, url, nationalite)
  idEquipe : Clé primaire
  nationalite : Clé étrangère en référence à idPays de pays

COUREUR (dossard, nom, prenom, dateDeNaissance, nationalite, idEquipe)
  dossard : Clé primaire
  nationalite : Clé étrangère en référence à idPays de pays
  idEquipe : Clé étrangère en référence à idEquipe de equipe

ETAPE (idEtape, villeDepart, villeArrivee, distance)
  idEtape : Clé primaire

RESULTATETAPE (idEtape, idCoureur, classement, temps, abandon)
  idEtape, idCoureur : Clé primaire
  idEtape : Clé étrangère en référence à idEtape de etape
  idCoureur : Clé étrangère en référence à dossard de coureur
  
```

Page 7

Importation des données :

- Commencez par créer une base de données "tourdefrance" en vous inspirant de ce que vous avez fait à la *séquence 1* (page 12).
- Importez ensuite le contenu de la base de données (structure et données) en utilisant le script "TP1_BDD.sql" en vous inspirant de ce que vous avez fait dans l'auto-correction de la *séquence 1* (page 137).

2. Requêtes simples

2.1. Quel est le parcours du Tour de France 2011 ?

N° Étape	Ville de départ	Ville d'arrivée	Distance (km)
1	Passage du Gois	Mont des Alouettes	191,5
2	Les Essarts	Les Essarts	23,0
3	Olonne-sur-Mer	Redon	198,0
4	Lorient	Mûr-de-Bretagne	172,5
... (Il y a 21 lignes au total)			

2.2. Quelles sont les villes qui, dans une même étape, sont "ville de départ" et "ville d'arrivée" ?

Villes à la fois départ et arrivée
Les Essarts
Grenoble

2.3. Quelle est la distance totale du Tour de France 2011 ?

Distance totale
3 430,0

2.4. Quelle est la distance de l'étape la plus longue ?

Distance de l'étape la plus longue
226,5

2.5. Quelle est la distance moyenne d'une étape ?

Distance moyenne d'une étape
163,3

2.6. Qui est parti avec le dossard n° 1 ?

Prénom	Nom
Alberto	CONTADOR

TP 1

Le Tour de France
2011

Page 8

2.7. Combien d'équipes participent au Tour 2011 ?

Nombre d'équipes participantes
22

2.8. Quels sont les directeurs d'équipes françaises ?

Directeurs d'équipes françaises
Jean-René BERNAUDEAU
Stéphane HEULOT
Franck PINEAU
Didier ROUS
Vincent LAVENU

2.9. Quelles équipes ont un site internet en ".fr" ?

Nom	URL
AG2R LA MONDIALE	www.cyclisme.ag2rlamondiale.fr

2.10. Quelle est la date de naissance du partant le plus âgé ?

Date de naissance du plus âgé
17/09/1971

2.11. Combien y a-t-il de nationalités représentées parmi les coureurs ?

Nombre de nationalités représentées
30

2.12. Qui sont les vainqueurs d'étapes (triés par noms) ?

Prénom	Nom
Mark	CAVENDISH
Alberto	CONTADOR
Rui Alberto	COSTA
... (Il y a 15 lignes au total)	

2.13. À quelle étape s'est produit le premier abandon ?

Étape du 1 ^{er} abandon
4

2.14. Combien d'abandons y a-t-il eu au total lors du Tour de France 2011 ?

Nombre total d'abandons
31

TP 1

Le Tour de France
2011

Page 9

3. Requêtes intermédiaires

3.1. Quelles sont les villes qui sont à la fois "villes de départ" et "villes d'arrivée" ?

Villes à la fois départ et arrivée
Les Essarts
Gap
Pinerolo
Grenoble

3.2. Quelle est la date de naissance du partant français le plus âgé ?

Date de naissance du Français le plus âgé
11/02/1974

3.3. Combien y a-t-il de coureurs par pays représenté (affichage des pays ayant le plus grand nombre de coureurs en premier et, en cas d'égalité, c'est l'ordre alphabétique sur le nom de pays qui prime) ?

Pays	Nombre de coureurs
France	45
Espagne	26
Belgique	15
Italie	15
Allemagne	12
... (30 lignes au total)	

TP 1

Le Tour de France
2011

Page 10

3.4. Quels sont les coureurs qui ont remporté plusieurs étapes ?

Prénom	Nom
Edvald Boasson	HAGEN
Thor	HUSHOVD
Mark	CAVENDISH

3.5. Quels sont les pays représentés par un seul coureur ?

Pays ayant un seul coureur
Autriche
Biélorussie
Canada
Costa Rica
... (9 lignes au total)

3.6. Quelles sont les sommes acquises par les victoires d'étapes sachant que chaque victoire rapporte au coureur 8 000 € (affichage des plus victorieux en premier, et, en cas d'égalité, c'est l'ordre alphabétique sur le nom du coureur qui prime) ?

Prénom	Nom	Sommes acquises en victoires d'étapes
Mark	CAVENDISH	40 000
Edvald Boasson	HAGEN	16 000
Thor	HUSHOVD	16 000
Alberto	CONTADOR	8 000
... (15 lignes au total)		

3.7. Quelle est l'étape la plus longue du Tour de France 2011 ?

N° Étape	Ville de départ	Ville d'arrivée	Distance
6	Dinan	Lisieux	226,5

3.8. Quel est le nombre d'abandons par étape (en donnant tous les numéros d'étapes) ?

N° Étape	Nombre d'abandons
1	
2	
3	
4	1
... (21 lignes au total)	

TP 1

Le Tour de France
2011

Page 11

3.9. Quelle est la date de naissance du plus jeune Français terminant le Tour 2011 ?

Date de naissance du plus jeune Français arrivant
11/09/1989

3.10. Quel est le directeur d'équipe du coureur qui a remporté l'arrivée sur les Champs-Élysées (étape n° 21) ?

Directeur du vainqueur de la dernière étape
Allan PEIPER

3.11. Quel est le classement par équipe de l'étape n° 2 (sachant que c'était un contre-la-montre par équipe, chaque coureur d'une même équipe a exactement le même temps. L'affichage se fera par meilleurs temps, et par noms d'équipe) ?

Équipe	Temps
TEAM GARMIN - CERVELO	00:24:48
BMC RACING TEAM	00:24:52
SKY PROCYCLING	00:24:52
TEAM LEOPARD-TREK	00:24:52
... (22 lignes au total)	

4. Requêtes complexes

4.1. Quel est le classement à l'arrivée de l'étape la plus longue ?

Classement	Prénom	Nom
1	Edvald Boasson	HAGEN
2	Matthew Harley	GOSS
3	Thor	HUSHOVD
4	Romain	FEILLU
... (195 lignes au total)		



Rappelez-vous que la requête 3.7 vous permettait de rechercher l'étape la plus longue.

4.2. Quelle distance les coureurs avaient-ils déjà parcourue avant d'attaquer l'étape la plus longue ?

Kilomètres parcourus avant la plus longue étape
749,5

TP 1

Le Tour de France
2011

Page 12



Comme pour la requête précédente, rappelez-vous que la requête 3.7 vous permettait de rechercher l'étape la plus longue.

4.3. Quelle est l'étape où il y a eu le plus d'abandons ?

Étape où il y a eu le plus d'abandons	Nombre d'abandons
9	8



Rappelez-vous que la requête 3.8 vous permettait de rechercher le nombre d'abandons par étape.

4.4. Quel est le pays ayant le plus de coureurs représentés au départ du Tour ?

Pays	Nombre de coureurs
France	45



Rappelez-vous que la requête 3.3 vous permettait de rechercher le nombre de coureurs par pays.

4.5. Quel est l'écart entre le premier et le dernier du Tour de France 2011 ?

Écart entre le premier et le dernier
03:57:36

4.6. Quel est le coureur qui a le plus grand nombre de victoires d'étapes ?

Prénom	Nom	Nombre d'étapes remportées
Mark	CAVENDISH	5



Vous disposez de la requête ci-après qui permet de trouver l'équipe qui a le plus grand nombre de victoires d'étapes :

Équipe	Nombre d'étapes remportées
HTC - HIGHROA	6

```
SELECT equipe.nom AS "Équipe",
       COUNT(*) AS "Nombre d'étapes remportées"
FROM coureur JOIN resultatEtape ON dossard = idcoureur
       JOIN equipe ON coureur.idEquipe = equipe.idEquipe
WHERE classement = 1
GROUP BY equipe.nom
HAVING COUNT(*) = ( SELECT MAX(countEtapesRemportees)
                    FROM ( SELECT idEquipe,
                                COUNT(*) AS countEtapesRemportees
                          FROM coureur JOIN resultatEtape
                                ON dossard = idcoureur
                          WHERE classement = 1
                          GROUP BY idEquipe
                        ) AS req
                  )
```

TP 1

Le Tour de France
2011

Page 13

4.7. Quel est le coureur qui a remporté le maillot blanc¹ du Tour de France 2011 avec le temps réalisé ?

Prénom	Nom	Temps total
Pierre	ROLLAND	86:23:05



Vous disposez, pour vous aider, de la requête, ci-après qui permet de trouver le coureur qui a remporté le maillot jaune¹² du Tour de France 2011 avec le temps réalisé :

Prénom	Nom	Temps total
Cadel	EVANS	86:12:22

1. Le maillot blanc, également appelé maillot du "meilleur jeune", est décerné au coureur ayant fait le meilleur temps sur l'intégralité du Tour et dont la naissance a eu lieu après le 1^{er} janvier 1986.
2. Le maillot jaune est décerné au coureur ayant fait le meilleur temps sur l'intégralité du Tour.

```

SELECT prenom, nom, SUM(temps) AS "Temps total"
FROM resultatEtape JOIN coureur ON idCoureur = dossard
GROUP BY prenom, nom
HAVING SUM(temps) = ( SELECT MIN(sumTP)
                      FROM ( SELECT idCoureur, SUM(temps) AS sumTP
                            FROM resultatEtape JOIN coureur
                                ON idCoureur = dossard
                            WHERE idCoureur
                                NOT IN ( SELECT idCoureur
                                        FROM resultatEtape
                                        WHERE abandon = true
                                    )
                            GROUP BY idCoureur
                        ) AS req
                )

```

4.8. Quel est le classement de l'étape n° 2 (contre la montre par équipe) avec les écarts de chaque équipe par rapport à la première ?

Équipe	Temps	Écarts
TEAM GARMIN - CERVELO	00:24:48	00:00:00
BMC RACING TEAM	00:24:52	00:00:04
SKY PROCYCLING	00:24:52	00:00:04
TEAM LEOPARD-TREK	00:24:52	00:00:04
HTC - HIGHROAD	00:24:53	00:00:05
... (22 lignes au total)		

TP 1

Le Tour de France
2011

Page 14



Rappelez-vous que la requête 3.11 vous permettait de rechercher le classement par équipe de l'étape n° 2 et que la requête 4.5 vous permettait de trouver un écart.

4.9. Quel est le classement général à l'arrivée du Tour de France 2011 avec les écarts de chaque coureur par rapport au premier ?

Prénom	Nom	Temps	Écarts
Cadel	EVANS	86:12:22	00:00:00
Andy	SCHLECK	86:13:56	00:01:34
Frank	SCHLECK	86:14:52	00:02:30
Thomas	VOECKLER	86:15:42	00:03:20
... (167 lignes au total)			