



MINI-PROJET BI

TOUR DE FRANCE

1. Compréhension du contexte métier

Le projet porte sur l'analyse historique du Tour de France, la plus célèbre course cycliste au monde.

Les données concernent les éditions, étapes, vainqueurs et finisseurs du Tour de France sur plusieurs décennies.

Objectifs BI par décideur

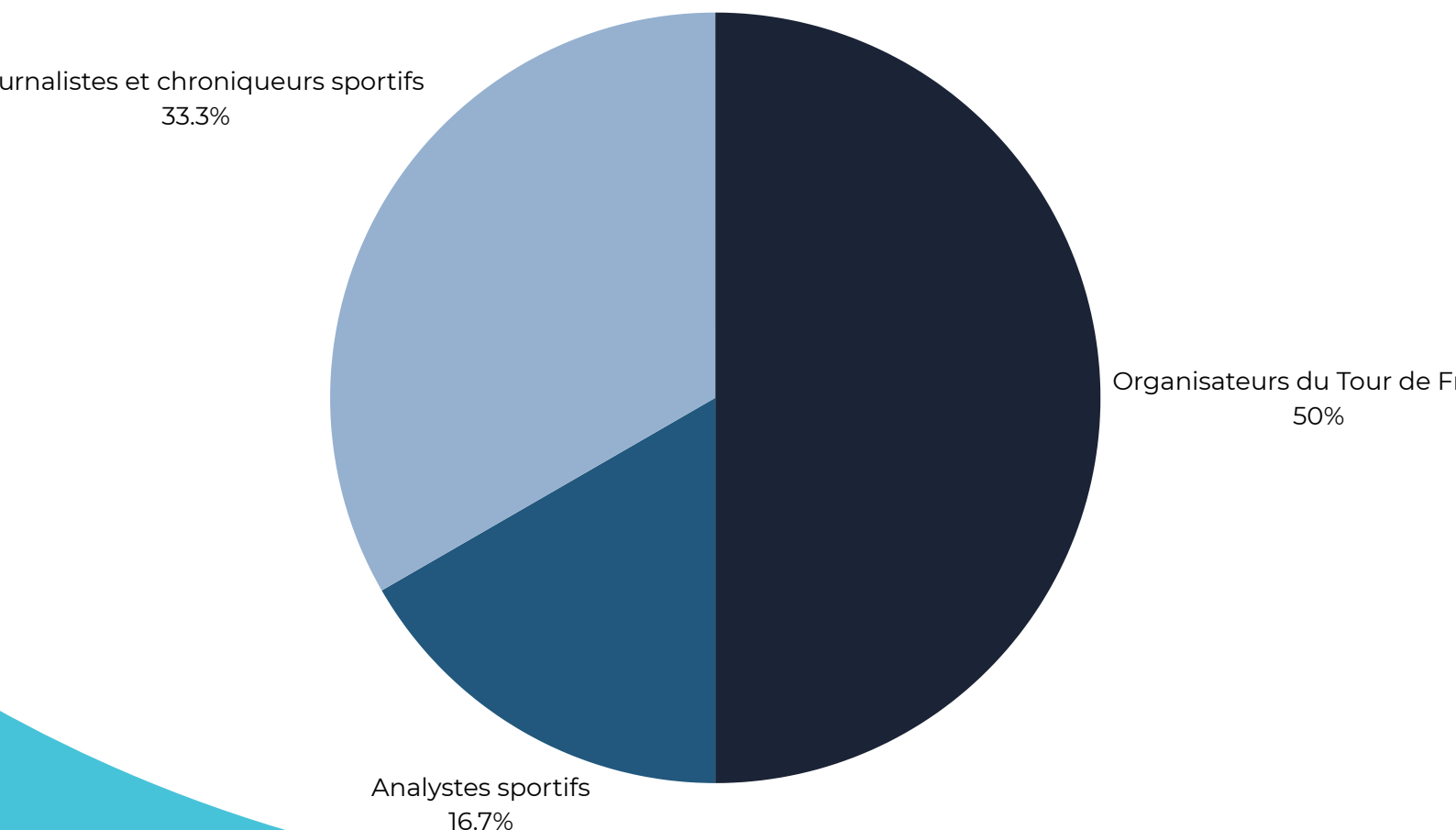
- Journaliste sportif :
 - Besoin de chiffres clés (étapes, éditions, pays).
 - Objectif : écrire des articles clairs avec des faits rapides à exploiter.
 - Exemple : Win Rate by Team.
- Manager du Tour :
 - Besoin d'analyser les performances globales et les risques (fatigue, logistique).
 - Objectif : optimiser l'organisation du Tour.
 - Exemple : Win Rate by Country, Rival Threat Level.
- Directeur sportif :
 - Besoin de données tactiques (vitesse, type d'étapes, poids).
 - Objectif : adapter l'entraînement des coureurs.
 - Exemple : Stages by Type.

2. Définition des objectifs métier

L'objectif métier de ce projet est de comprendre et analyser les performances des coureurs et des équipes au fil de l'histoire du Tour de France.

Cela permet d'identifier les facteurs clés de succès (âge, poids, équipe, pays, etc.) et de mieux visualiser les grandes évolutions du Tour (distance, nombre de participants, taux d'abandon...)

IDENTIFICATION DES DÉCIDEURS CIBLES



5. List des mesures

ID	Nom de la mesure	Description	Règle DAX
	Average Stage Distance	Donne la distance moyenne parcourue p	DIVIDE([Distance km], [Étapes])
	AvgSpeed_Overall	Moyenne globale des vitesses des vainq	AVERAGE('tdf_winners'[Avg Speed])
	Fatigue Pulse	Affiche un indicateur visuel du niveau de	DAX IF([Team Fatigue] = "High", "🔄 High
	Finisher Rate (%)	Taux de coureurs ayant terminé le Tour.	DIVIDE(tdf_tours[Arrivants], tdf_tours[P
	Has_Simple_Stage	Vérifie si l'étape est locale (non en "ville	DAX NOT(CALCULATE(CONTAINSSTRIN
	Multiple Wins	Nombre de victoires pour chaque coure	CALCULATE(COUNTROWS(tdf_winners),
	Nb_Editions	Nombre total d'éditions du Tour.	DISTINCTCOUNT(tdf_tours[Année])
	NbVictoiresParPays	Nombre de victoires par pays.	COUNTROWS(tdf_winners)
	Nom_Gagnant_Selectionne	Retourne le nom du gagnant d'une éditio	DAX CALCULATE(FIRSTNONBLANK(tdf_
	Nombre d'Étapes	Nombre d'étapes dans une édition.	DISTINCTCOUNT(tdf_tours[stage_id])
	Nombre Riders Par Team	Nombre de coureurs différents par équi	DAX COUNTROWS(SUMMARIZE(tdf_wir
	RépartitionÉtapes	Nombre d'étapes par type selon l'équipe	(long code, voir ton message)
	Rival Threat Level	Niveau de menace d'une équipe selon se	Rival Threat Level = VAR CurrentTeam = SELECTEDVALUE(tc
	Stage Category (Avg)	Catégorie moyenne d'étape selon la dist	Stage Category (Avg) = VAR avgDistance = AVERAGE(tdf_stages
	Stage Difficulty	Calcule une note de difficulté pour une e	Stage Difficulty = VAR SelectedStage = SELECTEDVALUE(t
	Stage Radar	Code numérique basé sur le type d'étap	Stage Radar = VAR SelectedStage = SELECTEDVALUE(t
	Stage Winner	Retourne le nom du vainqueur d'une éta	LOOKUPVALUE(tdf_stages[Winner], tdf_
	Stages_Won_Par_Coureur	Nombre d'étapes gagnées par un coureu	SUM(tdf_winners[Stages Won])
	Taux Victoire par etape	Ratio entre le nombre d'étapes et d'édit	DIVIDE([Nombre d'Étapes], COUNTROW
	Taux_Participations_Victoires	Ratio participations / victoires.	Taux_Participations_Victoires = DIVIDE(
	Taux_Performance_Vitesse	Rapport entre vitesse d'un pays et moye	Taux_Performance_Vitesse = DIVIDE(
	Taux_Victoires_Par_Equipe	Taux de victoires par équipe.	Taux_Victoires_Par_Equipe = DIVIDE(
	Taux_Victoires_Par_Pays	Taux de victoires par pays.	Taux_Victoires_Par_Pays = DIVIDE(
	Team Fatigue	Niveau de fatigue d'équipe selon les jou	Team Fatigue = VAR RacingDays = COUNTROWS(tdf_to
	Team Performance Score	Score de performance de l'équipe comp	Team Performance Score = VAR TeamAvgSpeed = [AvgSpeed
	Total_Participants	Total des participants tous Tours confor	SUM(tdf_tours[Partants])
	Victoires Garanties	Étapes gagnées par un coureur présent e	Victoires Garanties = COUNTROWS(

5. Tables disponibles

tdf_tours

Année

Stage_id

Winner_id

Développer 

tdf_stages

Course

Distance

Stage

Stage_id

Type

Winner

Year

Réduire 

countries

country

latitude

longitude

name

Réduire 

DimAnnéeUnique

Année

Développer 

tdf_winners

Σ Avg Speed

Σ Born

Country

Died

Height

Margin

Rider

Σ Stages Led

Σ Stages Won

Réduire 

tdf_finishers

Σ Arrondi à l'entier inférieur

Σ Arrondi à l'entier supérieur

Σ Finishiers_id

Σ Heures

Σ Minutes

Réduire 

6. Qualité des données

Types corrigés :

Distances converties en float.

Temps transformés (extraction pour calculs).

Dates transformées en format date.

Années normalisées en format numérique.

Gestion des valeurs manquantes :

Détection d'informations manquantes (ex: temps manquant, équipe manquante).

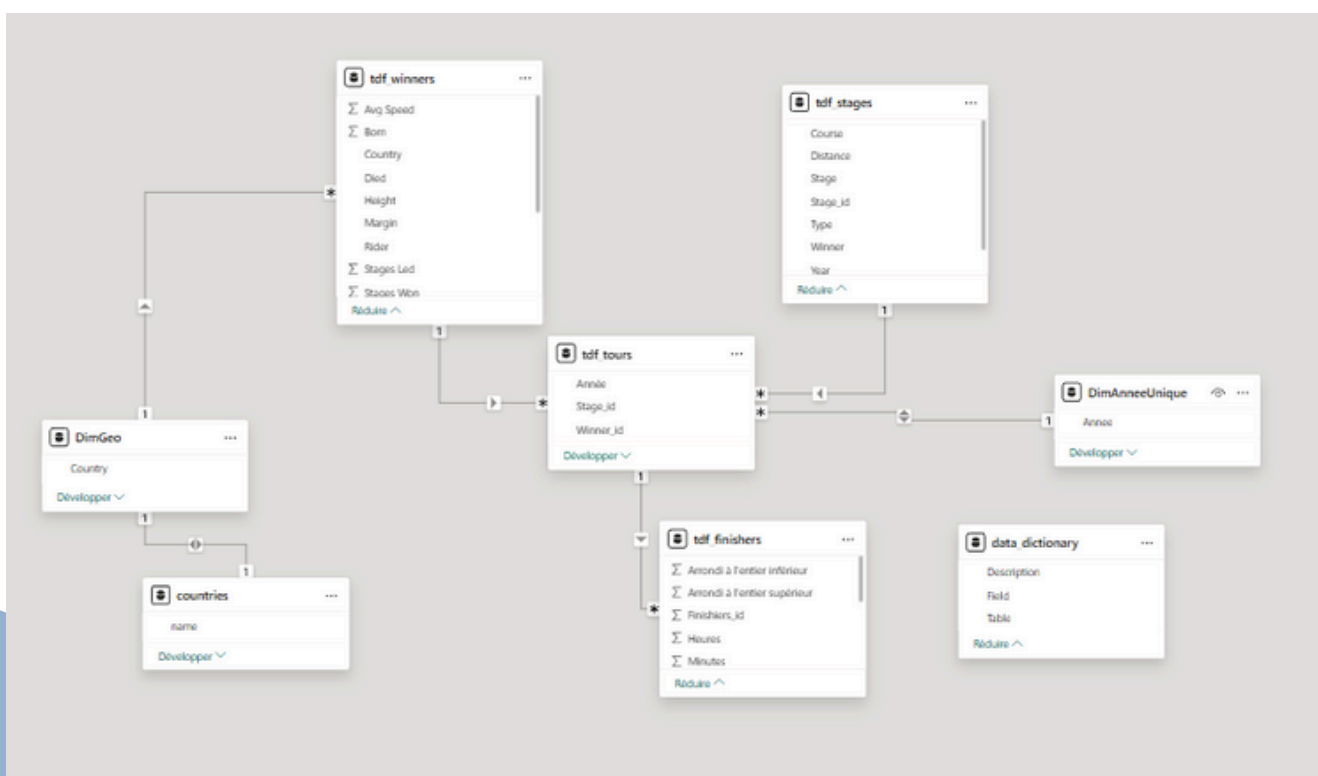
Suppression ou remplacement de valeurs incomplètes selon pertinence.

Normalisation :

Standardisation des distances (km).

Harmonisation des noms de pays, types d'étapes.

7. Modèle de données envisagé



8. Analyses supplémentaires déjà réalisées

Création des tables géographiques : Countries et DimGeo

Nouvelle table 1 : Countries

Contenu :

- Country : code pays à 3 lettres (ex : FRA, BEL).
- Latitude / Longitude : coordonnées géographiques du pays.
- Name : nom complet du pays (ex : France, Belgium).

Utilité :

- Permet d'afficher les données sur une carte Power BI.
- Sert de source géographique enrichie.

Nouvelle table 2 : DimGeo

Contenu :

- Country : liste unique des pays présents dans les données (extrait depuis les colonnes Rider, Team, etc.).

Utilité :

- Sert de table de dimension pour établir des relations dans le modèle.
- Rend l'analyse par pays plus flexible, lisible et performante.



- AvgSpeed_Overall = AVERAGE(tdf_winners[Avg Speed])

-->**Moyenne générale de la vitesse moyenne de tous les vainqueurs.**

- NbVictoiresParPays = COUNTROWS(tdf_winners)

-->**Nombre total de victoires (lignes) par pays (si filtré).**

- Nombre Riders Par Team = COUNTROWS(SUMMARIZE (tdf_winners, tdf_winners[Team], tdf_winners[Rider]))

-->**Nombre de coureurs différents par équipe.**

- Taux_Participations_Victoires = DIVIDE(COUNTROWS (tdf_finishers),CALCULATE(DISTINCTCOUNT (tdf_finishers[Year]),ALL(DimGeo)),0)

-->**Taux de participations par victoires**

- Vitesse_Moyenne = AVERAGE(tdf_winners[Avg Speed])
- NbVictoiresParPays = COUNTROWS(tdf_winners)
- Total_Participants = SUM(tdf_tours[Partants])
- Nb_Editions = DISTINCTCOUNT(tdf_tours[Année])
- NbVictoiresParPays = COUNTROWS(tdf_winners)
- Total_Participants = SUM(tdf_tours[Partants])
- Nb_Editions = DISTINCTCOUNT(tdf_tours[Année])

• Mesures DAX

ID	Nom de la mesure	Description	Règle DAX
	Average Stage Distance	Donne la distance moyenne parcourue p	DIVIDE([Distance km], [Étapes])
	AvgSpeed_Overall	Moyenne globale des vitesses des vainq	AVERAGE('tdf_winners'[Avg Speed])
	Fatigue Pulse	Affiche un indicateur visuel du niveau de	DAX IF([Team Fatigue] = "High", "🔄 High
	Finisher Rate (%)	Taux de coureurs ayant terminé le Tour.	DIVIDE(tdf_tours[Arrivants], tdf_tours[P
	Has_Simple_Stage	Vérifie si l'étape est locale (non en "ville	DAX NOT(CALCULATE(CONTAINSSTRIN
	Multiple Wins	Nombre de victoires pour chaque coure	CALCULATE(COUNTROWS(tdf_winners),
	Nb_Editions	Nombre total d'éditions du Tour.	DISTINCTCOUNT(tdf_tours[Année])
	NbVictoiresParPays	Nombre de victoires par pays.	COUNTROWS(tdf_winners)
	Nom_Gagnant_Selectionne	Retourne le nom du gagnant d'une éditio	DAX CALCULATE(FIRSTNONBLANK(tdf_
	Nombre d'Étapes	Nombre d'étapes dans une édition.	DISTINCTCOUNT(tdf_tours[stage_id])
	Nombre Riders Par Team	Nombre de coureurs différents par équi	DAX COUNTROWS(SUMMARIZE(tdf_wir
	RépartitionÉtapes	Nombre d'étapes par type selon l'équipe	(long code, voir ton message)
	Rival Threat Level	Niveau de menace d'une équipe selon se	Rival Threat Level = VAR CurrentTeam = SELECTEDVALUE(tc
	Stage Category (Avg)	Catégorie moyenne d'étape selon la dist	Stage Category (Avg) = VAR avgDistance = AVERAGE(tdf_stages
	Stage Difficulty	Calcule une note de difficulté pour une e	Stage Difficulty = VAR SelectedStage = SELECTEDVALUE(t
	Stage Radar	Code numérique basé sur le type d'étap	Stage Radar = VAR SelectedStage = SELECTEDVALUE(t
	Stage Winner	Retourne le nom du vainqueur d'une éta	LOOKUPVALUE(tdf_stages[Winner], tdf_
	Stages_Won_Par_Coureur	Nombre d'étapes gagnées par un coureu	SUM(tdf_winners[Stages Won])
	Taux Victoire par etape	Ratio entre le nombre d'étapes et d'édit	DIVIDE([Nombre d'Étapes], COUNTROW
	Taux_Participations_Victoires	Ratio participations / victoires.	Taux_Participations_Victoires = DIVIDE(
	Taux_Performance_Vitesse	Rapport entre vitesse d'un pays et moye	Taux_Performance_Vitesse = DIVIDE(
	Taux_Victoires_Par_Equipe	Taux de victoires par équipe.	Taux_Victoires_Par_Equipe = DIVIDE(
	Taux_Victoires_Par_Pays	Taux de victoires par pays.	Taux_Victoires_Par_Pays = DIVIDE(
	Team Fatigue	Niveau de fatigue d'équipe selon les jou	Team Fatigue = VAR RacingDays = COUNTROWS(tdf_to
	Team Performance Score	Score de performance de l'équipe comp	Team Performance Score = VAR TeamAvgSpeed = [AvgSpeed
	Total_Participants	Total des participants tous Tours confor	SUM(tdf_tours[Partants])
	Victoires Garanties	Étapes gagnées par un coureur présent e	Victoires Garanties = COUNTROWS(