

Exercices d'application en SQL

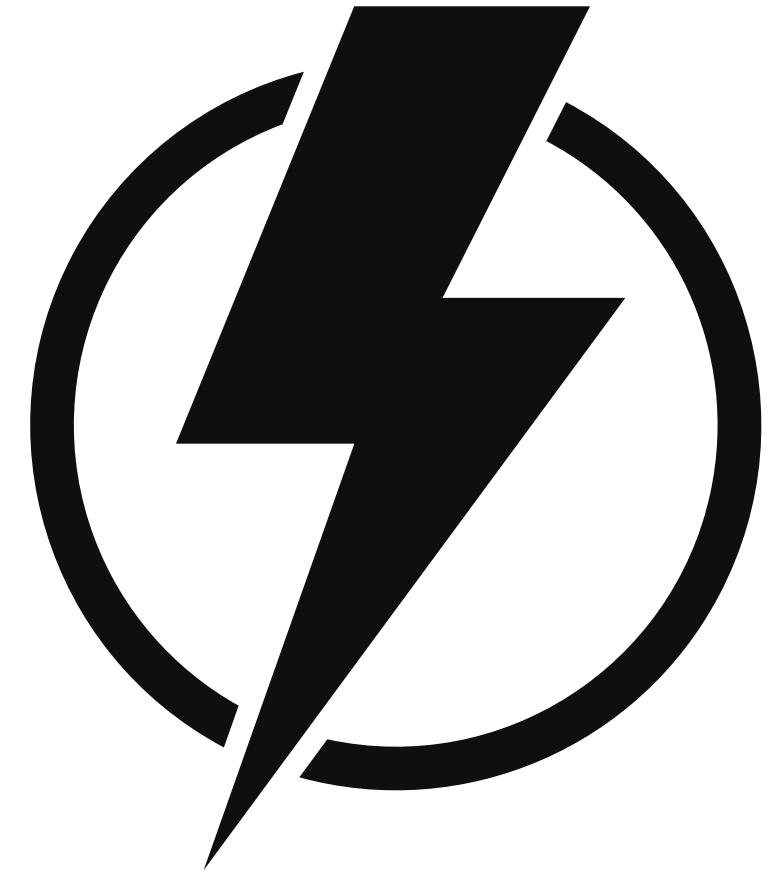
Jour 3 - Fonctions Aggrégatives, Groupements



Base de données TotaleEnergies_db

La base de données 'TotaleEnergies_db' représente une simulation imaginaire et détaillée de données.

Elle contient 4000 lignes, chacune correspondant à une filiale ou une unité opérationnelle distincte de l'entreprise.



Structure de la Base de Données

ID: Un identifiant unique pour chaque ligne (filiale).

Nom de la filiale: Le nom de la filiale ou de l'unité opérationnelle.

Pays: Le pays où la filiale est située.

Région: La région ou la ville spécifique où se trouve la filiale.

Secteur d'activité: Catégorise l'activité principale de la filiale, comme la production, le raffinage, etc.

Type de production: Le type d'énergie ou de ressource produite, comme le pétrole, le gaz, l'éolien, etc.

Volume de production: Quantité de production mesurée en unités appropriées.

Coût de production: Coût total de production de l'énergie ou des ressources.

Prix de vente: Prix auquel l'énergie ou les ressources sont vendues.

Revenu: Revenu total généré par la filiale.

Dépenses: Dépenses totales de la filiale.

Profit: Profit net réalisé par la filiale.

Nombre d'employés: Nombre total d'employés travaillant dans la filiale.

Taux d'accidents: Taux d'accidents enregistrés dans la filiale.

Investissements en R&D: Montant investi dans la recherche et le développement.

Emissions de CO2: Quantité totale d'émissions de CO2 produites.

Efficacité énergétique: Mesure de l'efficacité énergétique de la filiale.

Date de création de la filiale: Date à laquelle la filiale a été établie.

Date de dernier rapport financier: Date du dernier rapport financier publié par la filiale.

Note ESG: Score environnemental, social et de gouvernance attribué à la filiale.

Statut légal: Statut juridique de la filiale (publique, privée, partenariat).

Part de marché: Pourcentage estimé de part de marché détenu par la filiale.

01. Exercices guidés sur les fonctions agrégatives (COUNT, SUM, AVG, MAX, MIN)

01. Exercices guidés sur les fonctions agrégatives

Exercice 1 : Compter le nombre de filiales par pays

Problématique : Combien de filiales TotalEnergies y a-t-il dans chaque pays ?

- Ce qu'il faut faire : Utilisez la fonction COUNT pour compter le nombre de filiales dans chaque pays. Vous regrouperez les données par le nom du pays.

Exercice 2 : Calculer le revenu total

Problématique : Quel est le revenu total généré par toutes les filiales de TotalEnergies

- Ce qu'il faut faire : Utilisez la fonction SUM pour additionner toutes les valeurs dans la colonne "Revenu". Cela vous donnera le revenu total.

01. Exercices guidés sur les fonctions agrégatives

Exercice 3 : Calculer le total des investissements en R&D

Problématique : Quel est le montant total investi en recherche et développement (R&D) par TotalEnergies ?

- Ce qu'il faut faire : Utilisez la fonction SUM sur la colonne "Investissements en R&D" pour obtenir la somme totale des investissements en R&D.

Exercice 4 : Compter le nombre de filiales dans un secteur d'activité spécifique

Problématique : Combien de filiales de TotalEnergies sont engagées dans le secteur de l'"Exploration" ?

- Ce qu'il faut faire : Utilisez la fonction COUNT pour compter le nombre de filiales dont le "Secteur d'activité" est "Exploration". Filtrez les données en conséquence avant de compter.

01. Exercices guidés sur les fonctions agrégatives

Exercice 5 : Déterminer le nombre total d'employés

Problématique : Combien d'employés travaillent au total chez TotalEnergies ?

- Ce qu'il faut faire : Additionnez toutes les valeurs de la colonne "Nombre d'employés" en utilisant la fonction SUM pour obtenir le nombre total d'employés dans toutes les filiales.

Exercice 6 : Calculer la production moyenne

Problématique : Quelle est la production moyenne par filiale de TotalEnergies ?

- Ce qu'il faut faire : Utilisez la fonction AVG sur la colonne "Volume de production" pour calculer la moyenne de production.

01. Exercices guidés sur les fonctions agrégatives

Exercice 7 : Trouver le coût de production maximum

Problématique : Quel est le coût de production le plus élevé parmi toutes les filiales ?

Ce qu'il faut faire : Utilisez la fonction MAX sur la colonne "Coût de production" pour trouver la valeur la plus élevée.

Exercice 8 : Identifier le prix de vente minimum

Problématique : Quel est le prix de vente le plus bas enregistré par une filiale ?

Ce qu'il faut faire : Appliquez la fonction MIN à la colonne "Prix de vente" pour déterminer le prix de vente minimum.

01. Exercices guidés sur les fonctions agrégatives

Exercice 9 : Calculer le revenu moyen

Problématique : Quel est le revenu moyen généré par filiale ?

Ce qu'il faut faire : Utilisez la fonction AVG sur la colonne "Revenu" pour calculer le revenu moyen.

Exercice 10 : Trouver le nombre maximum d'employés dans une filiale

Problématique : Quelle filiale de TotalEnergies a le plus grand nombre d'employés ?

Ce qu'il faut faire : Utilisez la fonction MAX sur la colonne "Nombre d'employés" pour identifier la filiale avec le plus grand nombre d'employés.

01. Exercices guidés sur les fonctions agrégatives

Exercice 11 : Compter le nombre de pays uniques

Problématique : Combien de pays uniques sont représentés dans la base de données de TotalEnergies ?

Ce qu'il faut faire : Utilisez la fonction COUNT DISTINCT sur la colonne "Pays" pour déterminer le nombre de pays uniques où TotalEnergies a des filiales.

Exercice 12 : Compter le nombre de secteurs d'activité distincts

Problématique : Combien de secteurs d'activité différents sont présents dans la base de données ?

Ce qu'il faut faire : Appliquez la fonction COUNT DISTINCT à la colonne "Secteur d'activité" pour compter le nombre de secteurs d'activité distincts dans lesquels les filiales de TotalEnergies opèrent.

02. Exercices guidés sur GROUP BY et HAVING

02. Exercices guidés sur GROUP BY et HAVING

Exercice 13 : Revenu moyen par pays

Problématique : Quel est le revenu moyen généré par les filiales de TotalEnergies dans chaque pays ?

Ce qu'il faut faire : Utilisez GROUP BY sur la colonne "Pays" et appliquez la fonction AVG sur la colonne "Revenu".

Exercice 14 : Nombre total d'employés par secteur d'activité

Problématique : Combien d'employés travaillent dans chaque secteur d'activité au sein de TotalEnergies ?

Ce qu'il faut faire : Groupez les données par "Secteur d'activité" et utilisez SUM sur la colonne "Nombre d'employés".

02. Exercices guidés sur GROUP BY et HAVING

Exercice 15 : Nombre de filiales par type de production

Problématique : Combien de filiales opèrent dans chaque type de production (pétrole, gaz, éolien, etc.) ?

Ce qu'il faut faire : Appliquez GROUP BY à la colonne "Type de production" et utilisez COUNT pour compter les filiales dans chaque catégorie.

Exercice 16 : Investissement total en R&D par pays et type de production

Problématique : Quel est le montant total investi en R&D dans chaque pays pour chaque type de production ?

Ce qu'il faut faire : Utilisez GROUP BY sur les colonnes "Pays" et "Type de production", puis appliquez la fonction SUM sur "Investissements en R&D".

02. Exercices guidés sur GROUP BY et HAVING

Exercice 17 : Production moyenne par pays, région, secteur d'activité et type de production

Problématique : Quelle est la production moyenne pour chaque combinaison de pays, région, secteur d'activité et type de production ?

Ce qu'il faut faire : Utilisez GROUP BY sur les colonnes "Pays", "Région", "Secteur d'activité", et "Type de production". Appliquez ensuite la fonction AVG sur la colonne "Volume de production". Cela vous donnera une vue détaillée de la production moyenne en fonction de ces quatre dimensions.

Exercice 18 : Emissions de CO2 et efficacité énergétique par pays, région, type de production et statut légal

Problématique : Quelles sont les émissions moyennes de CO2 et l'efficacité énergétique moyenne pour chaque combinaison de pays, région, type de production et statut légal ?

Ce qu'il faut faire : Groupez les données par "Pays", "Région", "Type de production", et "Statut légal". Calculez ensuite la moyenne des "Emissions de CO2" et de l'"Efficacité énergétique" pour ces groupes. Cela fournira des insights sur la performance environnementale des filiales selon ces quatre catégories.

02. Exercices guidés sur GROUP BY et HAVING

Exercice 19 : Filiales avec un revenu moyen élevé par pays

Problématique : Dans quels pays les filiales de TotalEnergies ont-elles un revenu moyen supérieur à 500 000 € ?

Ce qu'il faut faire : Utilisez GROUP BY sur la colonne "Pays", calculez le revenu moyen avec AVG sur "Revenu", et utilisez HAVING pour filtrer les pays où ce revenu moyen est supérieur à 500 000 €.

Exercice 20 : Pays avec un nombre élevé de filiales dans le secteur de la 'Production'

Problématique : Quels pays ont plus de 10 filiales dans le secteur de la 'Production' ?

Ce qu'il faut faire : Groupez les données par "Pays" et "Secteur d'activité", comptez les filiales avec COUNT, et utilisez HAVING pour ne sélectionner que les pays où le nombre de filiales dans le secteur de la 'Production' est supérieur à 10.

02. Exercices guidés sur GROUP BY et HAVING

Exercice 21 : Filiales avec une performance élevée dans plusieurs domaines

Problématique : Identifiez les filiales de TotalEnergies qui ont un revenu moyen supérieur à 600 000 € et un taux d'accidents inférieur à 5 accidents par an dans chaque pays.

Ce qu'il faut faire :

Utilisez GROUP BY sur la colonne "Pays".

Calculez le revenu moyen avec AVG sur "Revenu" et le taux moyen d'accidents avec AVG sur "Taux d'accidents".

Appliquez HAVING pour sélectionner les pays où le revenu moyen est supérieur à 600 000 € ET où le taux moyen d'accidents est inférieur à 5.

02. Exercices guidés sur GROUP BY et HAVING

Exercice 22 : Secteurs d'Activité avec Haute Efficacité Énergétique et Faibles Émissions de CO2

Problématique : Quels secteurs d'activité de TotalEnergies dans différents pays ont une efficacité énergétique moyenne supérieure à 75 et des émissions de CO2 moyennes inférieures à 5 000 tonnes ?

Ce qu'il faut faire :

Groupez les données par "Pays" et "Secteur d'activité".

Utilisez AVG pour calculer l'efficacité énergétique moyenne et les émissions moyennes de CO2.

Appliquez HAVING pour filtrer les groupes ayant une efficacité énergétique moyenne supérieure à 75 ET des émissions moyennes de CO2 inférieures à 5 000 tonnes.

03. Exercices guidés sur CASE - GROUP BY - HAVING

03. Exercices guidés sur CASE – GROUP BY – HAVING

Exercice 23 : Catégorisation des revenus par pays

Problématique : Catégoriser les revenus moyens des filiales par pays en 'Élevé', 'Moyen' et 'Faible'.

Ce qu'il faut faire :

Utilisez GROUP BY sur la colonne "Pays".

Utilisez AVG pour calculer le revenu moyen par pays.

Employez CASE pour catégoriser chaque moyenne de revenu : 'Élevé' pour les revenus supérieurs à 700 000 €, 'Moyen' pour les revenus entre 300 000 € et 700 000 €, et 'Faible' pour les revenus inférieurs à 300 000 €.

03. Exercices guidés sur CASE – GROUP BY – HAVING

Exercice 24 : Évaluation de l'efficacité énergétique par secteur

Problématique : Déterminer si les secteurs d'activité ont une efficacité énergétique 'Bonne', 'Moyenne' ou 'Mauvaise'.

Ce qu'il faut faire :

Groupez les données par "Secteur d'activité".

Calculez l'efficacité énergétique moyenne avec AVG.

Utilisez CASE pour évaluer l'efficacité : 'Bonne' pour une efficacité supérieure à 80, 'Moyenne' pour une efficacité entre 50 et 80, et 'Mauvaise' pour une efficacité inférieure à 50.

03. Exercices guidés sur CASE – GROUP BY – HAVING

Exercice 25 : Segmentation des investissements en R&D par Région avec un critère de Performance

Problématique : Quelles régions investissent en moyenne plus de 150 000 € en R&D, et comment ces investissements sont-ils classés ?

Ce qu'il faut faire :

Utilisez GROUP BY sur "Région".

Calculez la moyenne des investissements en R&D avec AVG.

Appliquez CASE pour classer les investissements : 'Élevé' pour plus de 200 000 €, 'Moyen' pour 150 000 € à 200 000 €, 'Faible' pour moins de 150 000 €.

Utilisez HAVING pour sélectionner uniquement les régions dont l'investissement moyen en R&D est supérieur à 150 000 €.

03. Exercices guidés sur CASE – GROUP BY – HAVING

Exercice 26 : Efficacité énergétique et émissions de CO2 par pays et région

Problématique : Identifier les combinaisons de pays et de régions où l'efficacité énergétique moyenne est supérieure à 65 et les émissions de CO2 sont inférieures à 3000 tonnes.

Ce qu'il faut faire :

Groupez les données par "Pays" et "Région".

Calculez l'efficacité énergétique moyenne et les émissions moyennes de CO2 avec AVG.

Utilisez CASE pour évaluer l'efficacité et les émissions : 'Bonne' efficacité pour une valeur supérieure à 80, 'Moyenne' pour 65 à 80, 'Faible' pour moins de 65.

Employez HAVING pour sélectionner les groupes avec une efficacité énergétique moyenne supérieure à 65 et des émissions de CO2 inférieures à 3000 tonnes.

04. Exercices avancés

04. Exercices avancés

Exercice 27 : Analyse complète du secteur d'activité par pays et région

Problématique : Pour chaque combinaison de pays et de région, déterminez le secteur d'activité le plus rentable, en considérant uniquement les secteurs où le nombre moyen d'employés est supérieur à 3000 et les investissements en R&D sont supérieurs à 150 000 € en moyenne.

Exercice 28 : Segmentation des types de production par performance environnementale et économique

Problématique : Identifiez les types de production qui ont à la fois une efficacité énergétique moyenne supérieure à 70, des émissions de CO2 inférieures à 4000 tonnes en moyenne, et un revenu moyen supérieur à 500 000 € dans chaque pays.

04. Exercices avancés

Exercice 29 : Filtrage et analyse des filiales selon leur performance et secteur

Problématique : Pour les filiales créées après l'année 2000, déterminez dans chaque pays et secteur d'activité ceux qui ont un revenu moyen supérieur à 600 000 € et un taux d'accidents inférieur à 3, en considérant seulement les secteurs avec des investissements en R&D moyens supérieurs à 200 000 €.

Exercice 30 : Segmentation des filiales par efficacité énergétique et revenu dans certaines régions

Problématique : Identifiez les filiales dans les régions européennes qui ont une efficacité énergétique moyenne supérieure à 80 et un revenu moyen supérieur à 500 000 €, en excluant les filiales avec un taux d'accidents supérieur à 5.