

# L'analyse des open data de la SNCF

Soutenu par Mounir Ahmed Djibaba et Walid  
Benchabekh

**Encadré par :**  
**Mme c.Petitjean**



# Plan



## Introduction générale

- Problématique
- Objectifs



## Contexte du projet

- la SNCF
- SNCF Open Data



## Analyse et conception

- jeux de données utilisés



## Réalisation

- Technologies utilisées
- Démonstration



## Conclusion et perspectives

# Introduction générale

L'analyse des données est devenu crucial dans notre monde actuel, où nous sommes confrontés à une quantité croissante de données provenant de différentes sources, telles que les réseaux sociaux, les capteurs, les transactions en ligne, etc.



# Objectifs

- Améliorer la qualité de service
- Optimiser l'utilisation des ressources
- Prédire la demande



# SNCF

- La SNCF (Société nationale des chemins de fer français) a été créée en 1938 pour assurer l'exploitation du réseau ferroviaire national en France.
- Elle est devenue une entreprise publique en 1982
- La SNCF gère le réseau ferroviaire français
- Elle emploie plus de 150 000 personnes et dessert plus de 4 000 gares en France.



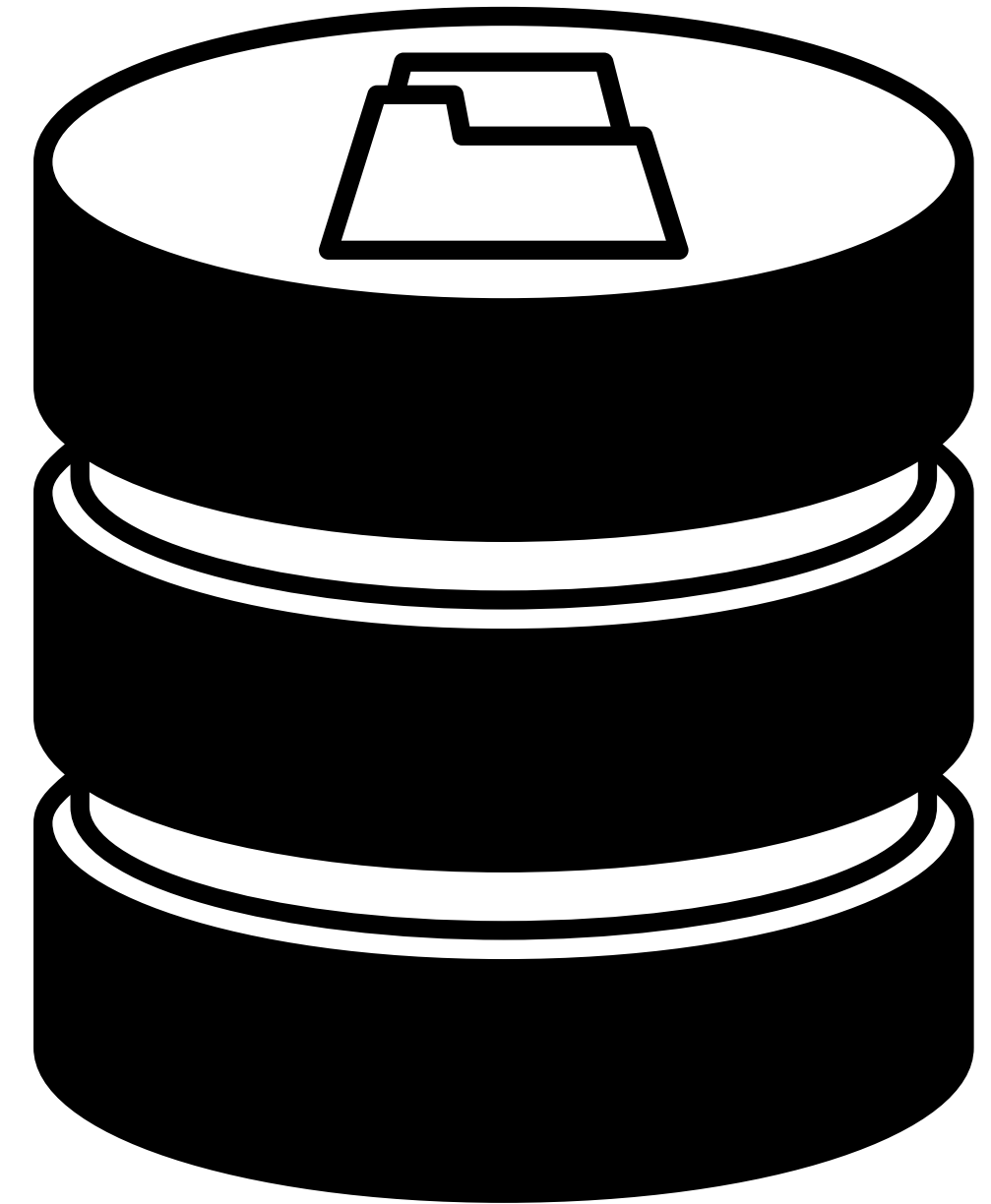
# SNCF Open Data

- La SNCF a adopté une politique d'open data en 2011, dans le but de rendre ses données plus accessibles au public et de stimuler l'innovation dans le secteur des transports.
- Ces données incluent des informations sur les horaires de train, les retards, les annulations...



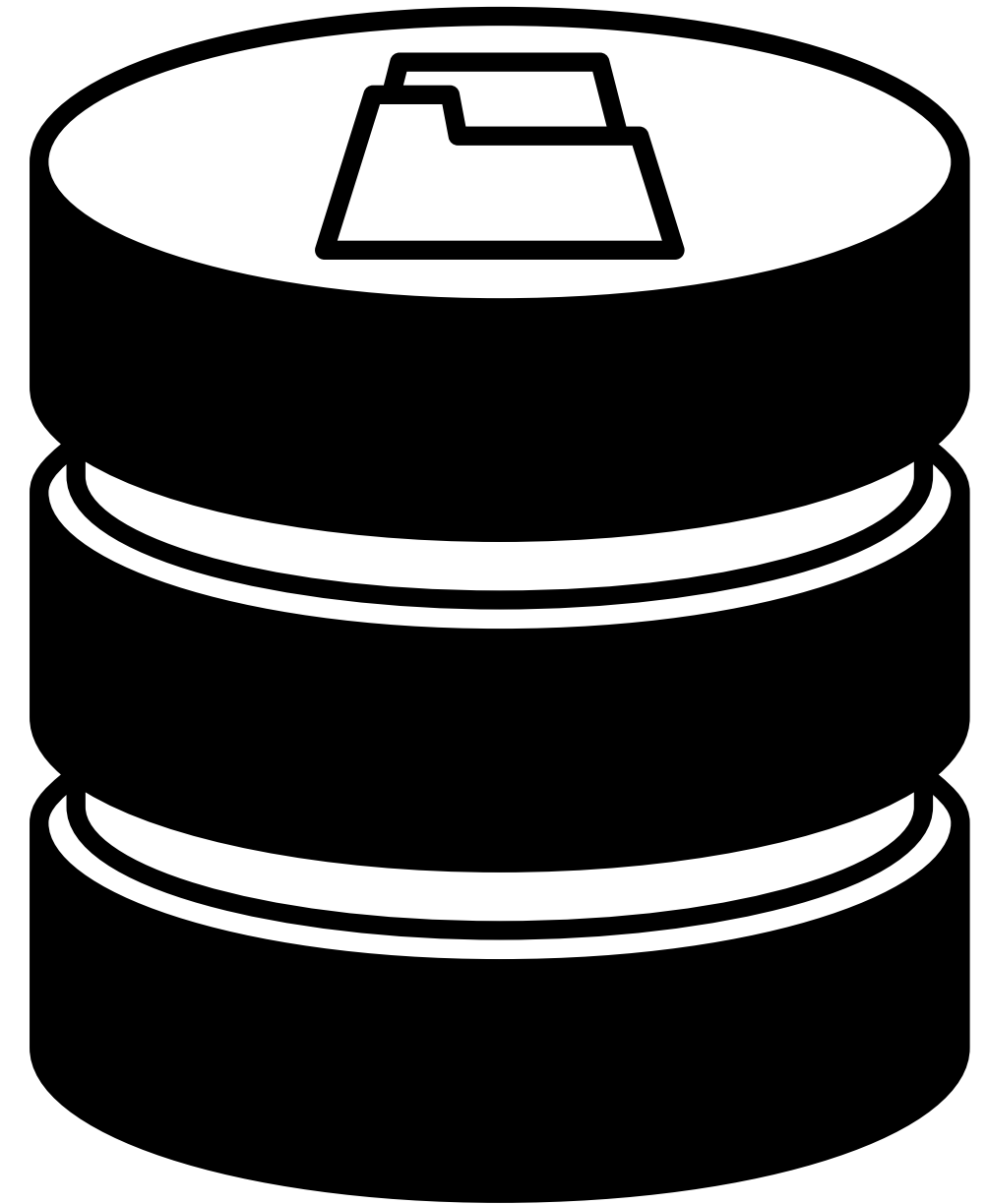
# Jeux de données utilisés

- Régularité mensuelle TER
- Comptage des voyageurs  
montants dans les trains  
Transilien
- Répartition des clients par âge



# Régularité mensuelle TER

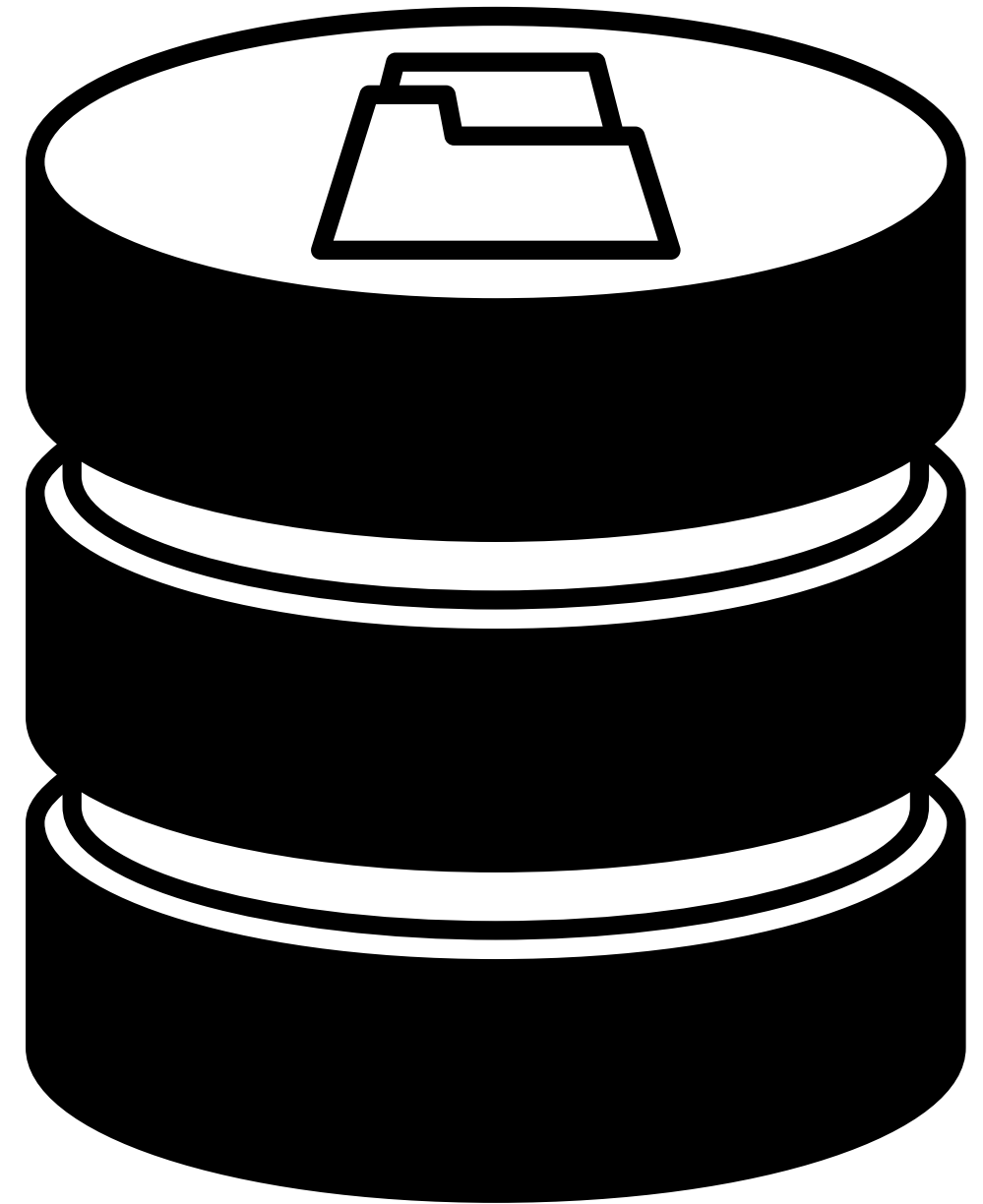
- Train programmé par années
- moyenne de taux de régularité par année





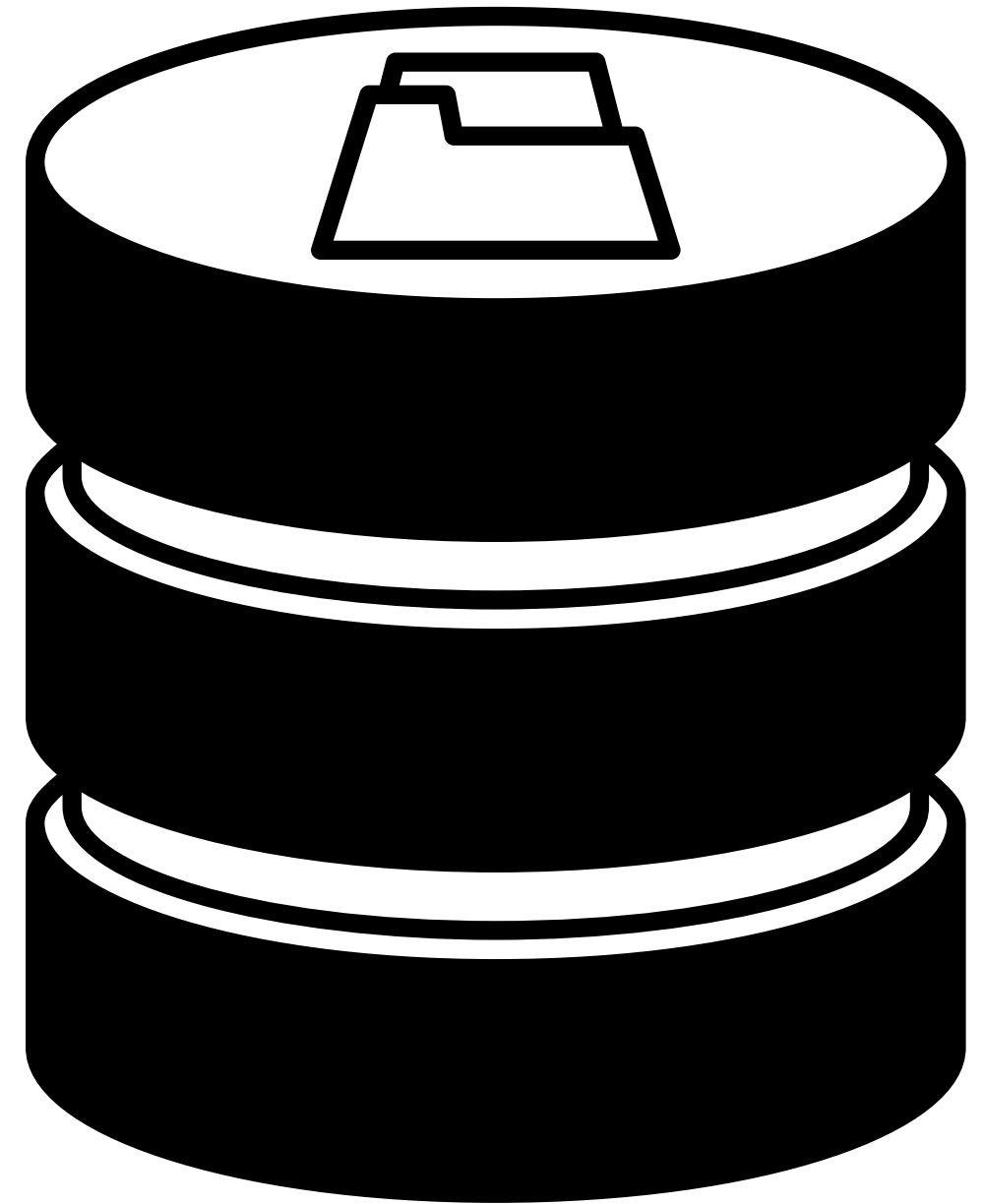
# Comptage des voyageurs montants dans le train transilien

- moyenne de voyageurs montants par tranche horaire
- Moyenne de voyageurs par mois par lignes



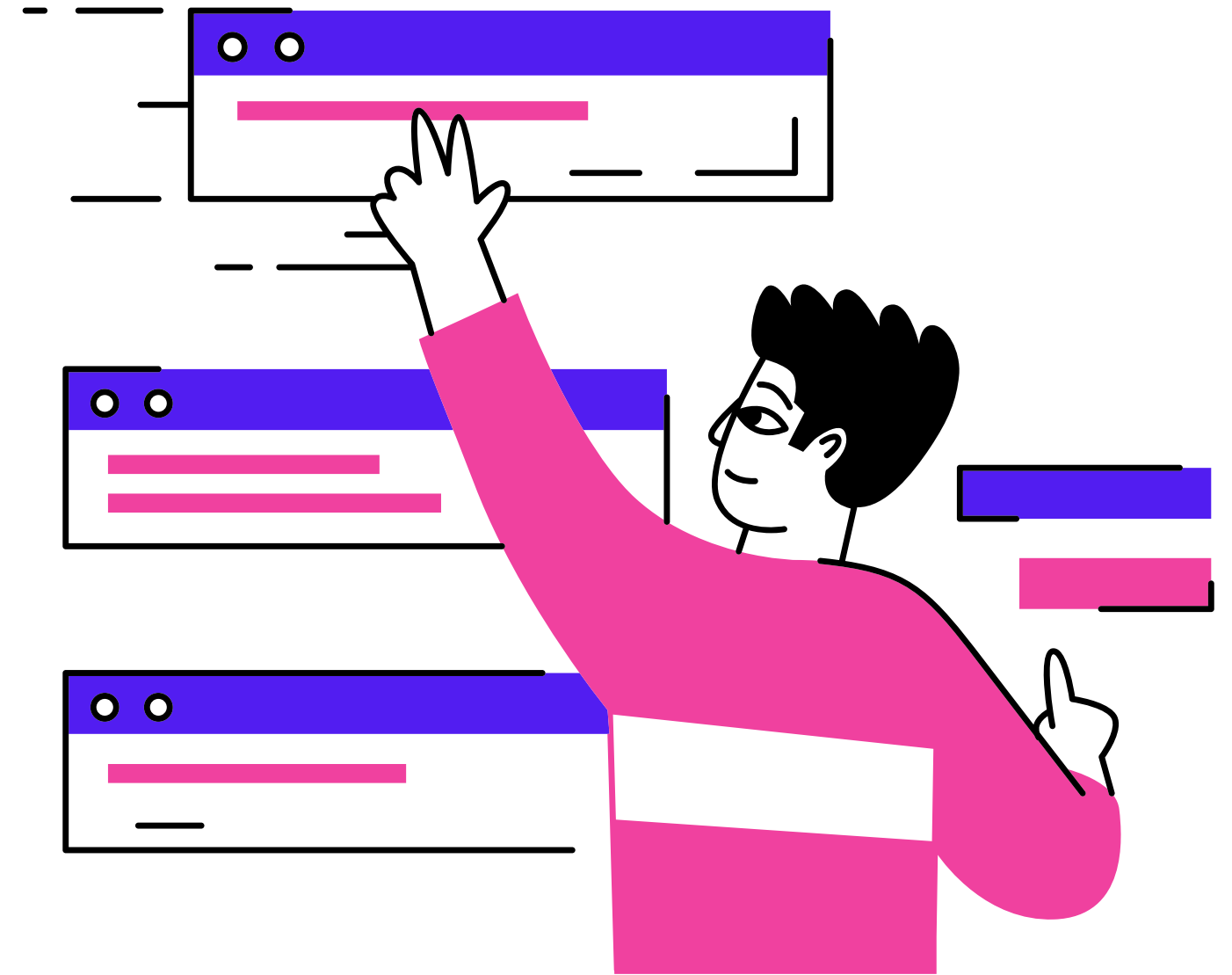
# Enquêtes en gares

- Répartition générale des clients par âge
- Répartition des clients par gare



# Dashboard

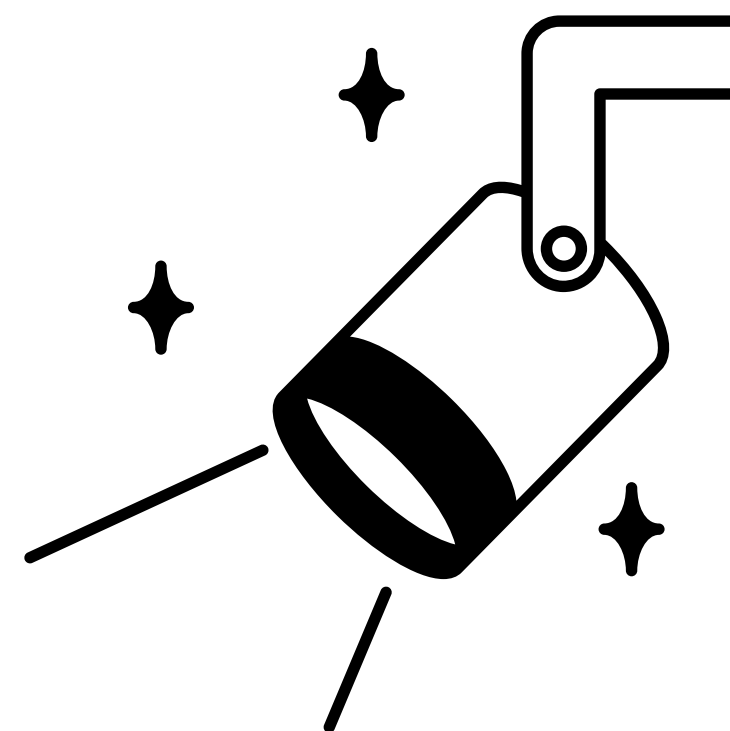
- Un dashboard est une interface graphique qui permet de visualiser des données de manière synthétique et interactive.
- Il est utilisé pour représenter des informations clés dans un format visuel facile à comprendre et à utiliser.



# Technologies utilisées



# Démonstration



# Régularité mensuelle TER

Taux de régularité minimal

86,75

Taux de régularité moyen

91,56

Taux de régularité maximal

98,03

Région

Languedoc Roussillon

## Taux de régularité par département

Année de Date

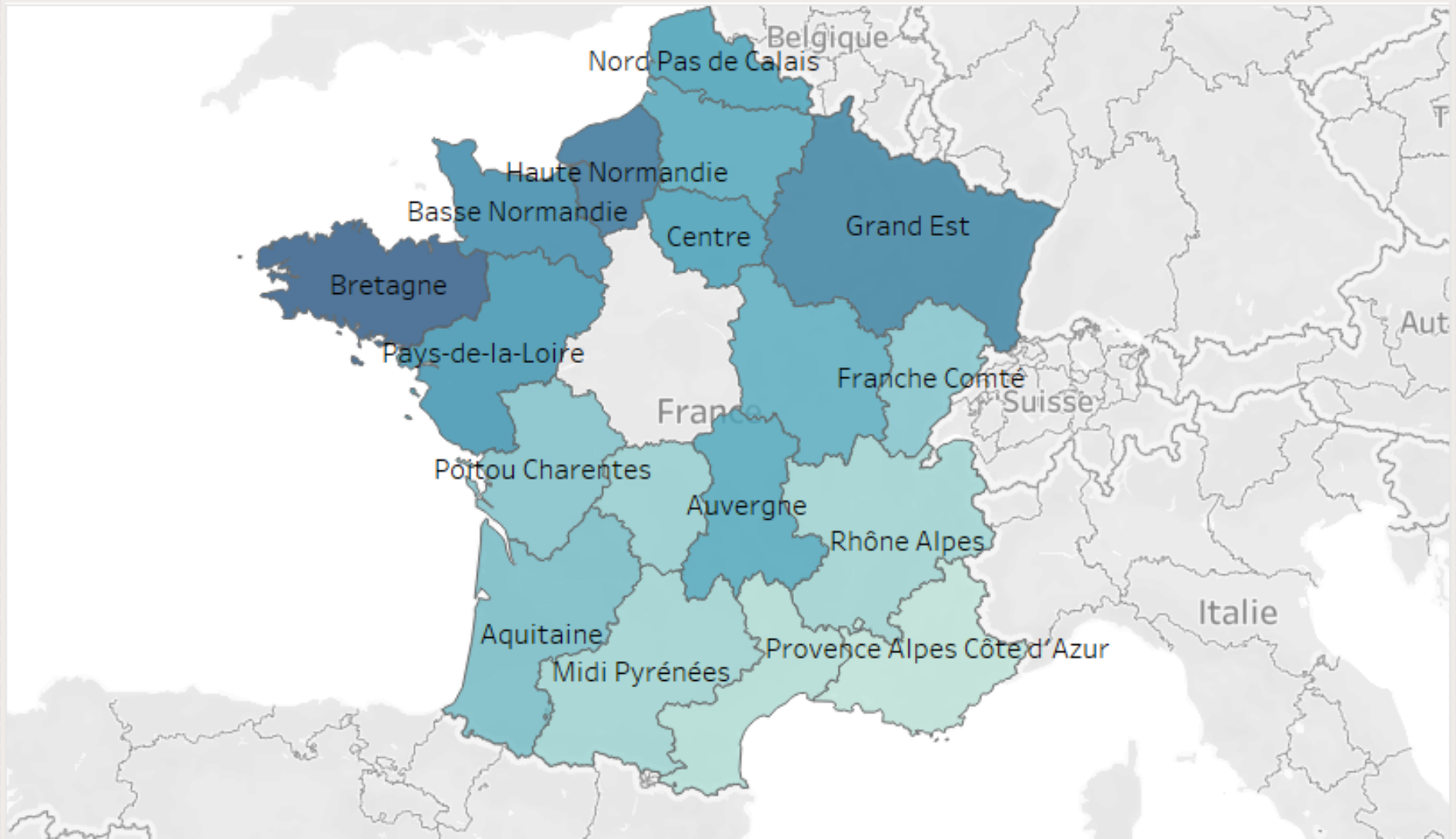
2017

Mois de Date

février

### Pourquoi ?

Ce mois de février est en progrès par rapport au mois précédent, mais l'objectif n'est toujours pas atteint. Les causes externes ont encore une fois très impactés nos résultats(Intempéries et Collision à un passage à niveau).



Taux de régularité

86,75

98,03

# Conclusion

- En conclusion, notre projet d'analyse de données de la open data de la SNCF a permis de mettre en évidence des tendances dans les données et surtout l'utilisation des bons outils de visualisation pour bien comprendre les données



# perspectives

**l'utilisation des méthodes avancées de l'analyse de données afin d'exploiter au mieux les différents jeux de données.**

**Exemple : l'utilisation du traitement automatique des langages pour analyser les commentaires sur le jeu de données de la régularité des TER**





*Merci pour votre attention*