

**DUPEYRON Robin – FAUVEAU Ewen**

## **SAE 2.03 – Mettre en place une solution informatique pour l'entreprise**

### **Compte Rendu du Projet : Fréquentation du réseau de bus Yélo de La Rochelle**



# **Sommaire**

*Tous les liens sont cliquables pour accéder au sujet*

**[Introduction](#)**

**[Description du sujet](#)**

**[Cahier des charges](#)**

**[Peuplement et Traitements](#)**

**[Répartition des tâches](#)**

**[Conclusion et bilan](#)**

## Introduction

Dans le cadre de notre formation en BUT R&T, nous avons entrepris un projet de peuplement et traitement d'une base de données SQL, qui s'inscrivait dans la compétence visée par la SAE 2.03 intitulée "Mettre en place une solution informatique pour l'entreprise". Ce projet avait pour objectif l'étude de la fréquentation du réseau de bus urbain de La Rochelle (Yélo), détaillant des informations telles que les lignes de bus, les arrêts, les chauffeurs, les validations, le parc de véhicules.... Nous avons conçu un fichier SQL, capable de peupler toutes les tables nécessaires par rapport aux informations que nous possédons, et ensuite traiter ces données afin de pouvoir les analyser sous forme de tableaux SQL.

## Description du sujet :

L'objectif principal de notre projet est la création, le peuplement et le traitement d'une base de données sur les données de fréquentation du réseau de bus Yélo de La Rochelle.

## Cahier de charge en français :

Le réseau de transport en commun Yélo cherche à optimiser l'offre de service en fonction de la demande réelle des usagers. L'analyse de la fréquentation des lignes de bus est essentielle pour ajuster la capacité des véhicules, modifier les horaires ou créer de nouveaux itinéraires en fonction des habitudes de déplacement. Ce projet vise à concevoir une base de données permettant de stocker et d'exploiter les données de montée et descente des passagers à chaque arrêt, en s'appuyant sur des informations issues des systèmes de billettique (cartes de transport, tickets électroniques) et des capteurs embarqués dans les bus.

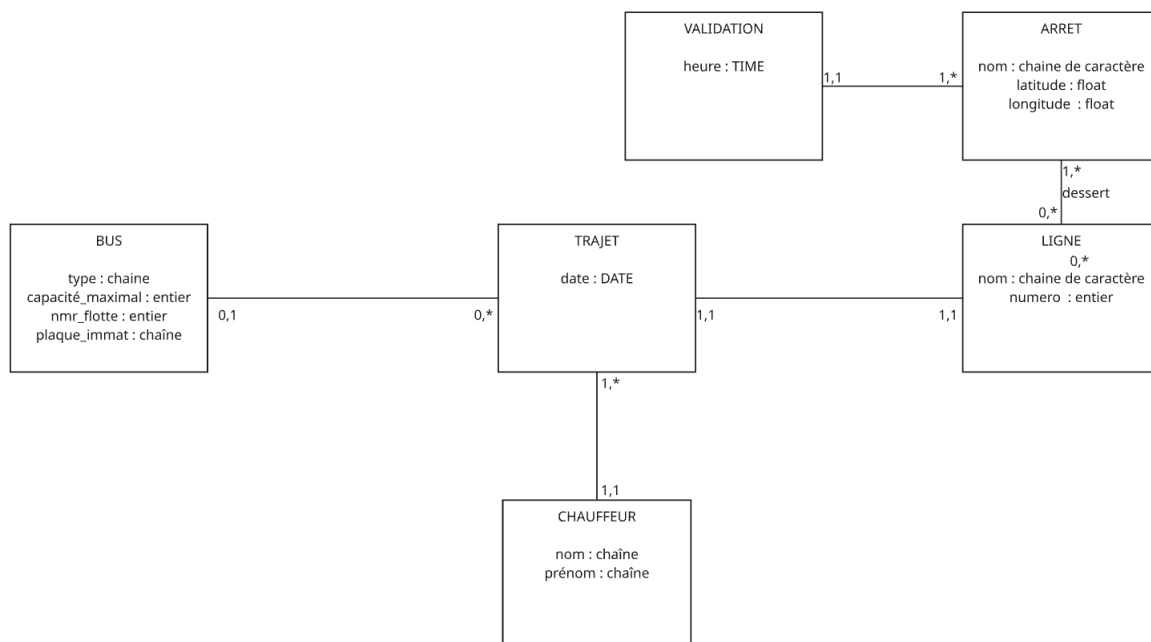
La base de données doit modéliser plusieurs entités interconnectées :

- Ligne de bus : chaque ligne possède un identifiant unique, un numéro, un nom et une liste ordonnée d'arrêts desservis.
- Arrêt de bus : chaque arrêt est identifié par un nom et des coordonnées GPS et peut être desservi par plusieurs lignes.
- Bus : chaque bus est identifié par un numéro de flotte et est associé à une ligne spécifique à un moment donné.
- Trajet : un trajet correspond au parcours d'un bus sur une ligne donnée à une date et un horaire précis.
- Passage : il enregistre le passage d'un bus à un arrêt donné à une heure donnée.
- Validation : lorsqu'un passager monte à bord, une validation est enregistrée avec l'arrêt d'embarquement, le bus concerné et l'horaire précis.

Héritage et relations spécifiques

- Les trajets sont liés à une ligne, un bus et une journée spécifique pour conserver l'historique des passages.
- Les validations sont liées aux trajets et aux arrêts, avec une possibilité de regrouper les données par heure, jour ou période pour des analyses statistiques.

### **Diagramme UML :**



## **Schéma relationnel :**

Bus : (id\_bus, type, numero\_flotte, capacite)

Lignes : (id\_ligne, numero, nom)

Arrêts : (id\_arret, nom, latitude, longitude)

Chauffeurs : (id\_chauffeur, nom, prenom)

Trajets : (id\_trajet, id\_ligne, id\_bus, date\_trajet, id\_chauffeur)

Passages : (id\_passage, id\_trajet, id\_arret, heure)

Validations : (id\_validation, id\_trajet, id\_arret, heure\_validation)

## **Traitements :**

Arrêts les plus utilisés

Taux de remplissage des bus

Top 5 des trajets avec le taux de remplissage le plus élevé

Arrêts les plus fréquentés pour une ligne spécifique

Bus les plus utilisés en nombre de trajets et de validations

Évolution quotidienne des validations sur le réseau

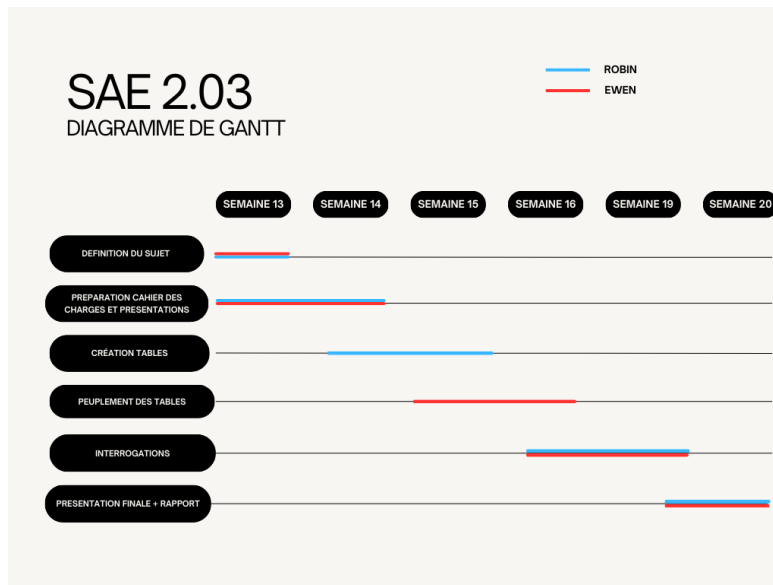
Lignes qui desservent le plus d'arrêts

Chauffeurs affectés à plusieurs lignes différentes

Temps moyen entre deux validations sur le même trajet (indicateur de densité passagers)

## Gestion de projet :

Ce projet a été réalisé selon la répartition des tâches suivantes :



## Éléments nécessaires :

- Base de données des véhicules bus de La Rochelle :  
<https://tc-infos.fr/reseau/83/parc-bus-cars>
- Données fréquentation lignes de bus de La Rochelle
- Notre cahier de charge
- Base de données PostgreSQL

## Peuplement et Traitements

Nous avons peuplé les tables en utilisant les données ouvertes disponibles sur internet du réseau Yélo. Nous nous sommes concentrés ici sur les 4 lignes principales urbaines du réseau (ILLICO 1 à 4), mais le projet est totalement fonctionnel avec d'autres lignes si l'on le voulait (lignes secondaires / interurbaines / scolaires)

Les véhicules ont été récupérés du site TC-Infos, une base de données ouverte et collaborative qui permet de répertorier les véhicules des réseaux de transports français. Nous avons ici récupéré un échantillon de véhicules (4-5 véhicules par type) pour éviter un surpeuplement pour nos tests.

Certaines données ne sont pas publiques sur internet (ex : chauffeurs en raison de la vie privée), il a donc fallu les inventer.

Le peuplement a pris moins de temps que prévu à la base, ce qui nous a permis de nous concentrer sur les traitements.

Le traitement des données s'est articulé autour de plusieurs axes d'analyse, chacun visant à mettre en lumière un aspect particulier de l'utilisation du réseau de bus de La Rochelle. Ces traitements ont été réalisés à l'aide de requêtes SQL, mobilisant des fonctions d'agrégation, des jointures entre tables.

Le premier axe a porté sur l'identification des arrêts les plus utilisés. L'objectif était de repérer les points névralgiques du réseau, ceux où la densité de validations est la plus forte. Pour cela, nous avons comptabilisé le nombre total de validations par arrêt, en croisant les données des trajets

et des validations. Le résultat a été synthétisé sous forme d'un tableau classé, mettant en avant les arrêts les plus fréquentés, ce qui constitue une base précieuse pour la planification des flux passagers.

Un deuxième axe majeur a été l'étude du taux de remplissage des bus, mesurant l'efficacité d'utilisation des véhicules. Nous avons comparé le nombre de validations par trajet au nombre de places disponibles dans chaque bus, permettant de calculer un ratio de remplissage moyen. Cette analyse, affinée pour chaque ligne et chaque tranche horaire, met en évidence les plages horaires les plus critiques en termes de saturation.

Dans la continuité, nous avons extrait le Top 5 des trajets avec le taux de remplissage le plus élevé, afin d'identifier les cas de surcharge potentielle. Ce traitement a nécessité une agrégation fine des validations par trajet, puis un classement descendant selon le taux de remplissage. Le résultat, présenté dans un tableau, aide à anticiper les besoins en renfort de capacité ou en réajustement d'horaires.

Un autre traitement a concerné les arrêts les plus fréquentés pour une ligne spécifique, permettant d'analyser la dynamique locale d'une ligne donnée. En filtrant les validations selon la ligne choisie, puis en les regroupant par arrêt, nous avons pu visualiser la fréquentation spécifique à cette ligne, ce qui est utile pour une optimisation ciblée.

L'analyse s'est également penchée sur les bus les plus utilisés, à la fois en nombre de trajets effectués et de validations enregistrées. En croisant les données des trajets avec celles des validations, puis en les regroupant par immatriculation, nous avons pu dresser un classement des bus les plus sollicités, tant en fréquence qu'en volume.

Un axe temporel a été exploré à travers l'évolution quotidienne des validations sur le réseau. Cette analyse visait à observer les variations de fréquentation au fil des jours, en regroupant les données par date. La

courbe obtenue permet de repérer des tendances générales (jours de forte affluence, weekends creux), et d'évaluer la régularité de la demande.

Un traitement complémentaire a ciblé les lignes qui desservent le plus d'arrêts, avec pour but d'identifier les parcours les plus longs ou les plus ramifiés. Cette requête s'est appuyée sur un comptage des arrêts différents rattachés à chaque ligne, et a débouché sur un tableau de classement des lignes selon leur couverture territoriale.

D'un point de vue organisationnel, une requête a permis de détecter les chauffeurs affectés à plusieurs lignes différentes, en croisant les affectations de trajets. Cette information est essentielle pour l'analyse de la polyvalence des conducteurs et l'optimisation des plannings.

Enfin, une analyse plus fine a été menée sur le temps moyen entre deux validations sur un même trajet, considéré comme un indicateur indirect de densité passagers. En calculant la différence temporelle moyenne entre les validations consécutives, nous avons pu estimer la cadence d'embarquement, révélatrice de l'intensité de fréquentation à bord.

## Répartition des Tâches

Le travail sur ce projet s'est déroulé sur une période de six semaines, durant lesquelles chaque membre de l'équipe a pris en charge des tâches spécifiques tout en collaborant étroitement pour garantir une bonne cohésion du projet. La première semaine a été consacrée à la description du sujet et à la création du cahier de charge. Ewen s'est occupé de la récupération des données, de l'identification des différentes tables à peupler et de la vérification de la qualité des données (correspondent bien au réseau de la Rochelle). De son côté, Robin a préparé le cahier des charges initial, en se basant sur les objectifs du projet, et a également

effectué des recherches pour identifier les types de traitements les plus pertinents.

La deuxième semaine a été marquée par la création des premières tables. Ewen a pris en charge l'implémentation des clés étrangères et de la cardinalité. Robin a de son côté codé la structure des tables.

La suite du projet a été dédiée au peuplement des tables, à la création des traitements et à l'interface graphique. Ewen a travaillé sur l'interface graphique, tandis que Robin s'est chargé des traitements.

Enfin, nous avons pu finaliser l'application, préparer la présentation finale et rédiger le compte-rendu. Ewen a préparé la présentation de soutenance, tandis que Robin a réalisé le compte-rendu. Nous nous sommes cependant concertés afin de prendre connaissance des deux supports.

## Conclusion et Bilan

Ce projet a constitué une expérience extrêmement enrichissante, permettant de mettre en pratique les compétences acquises en bases de données tout en apprenant à travailler en équipe de manière efficace. Il a offert l'opportunité de manipuler des données structurées en SQL et de concevoir des algorithmes de traitement pertinents.

Les points forts de ce projet sont nombreux : Le sujet porte sur des données réelles, ce qui permet la réutilisation de ce projet dans le cadre de vrais traitements de données. De plus, le code SQL contient toutes les tables peuplées et avec les traitements, ce qui permet d'être utilisé en "standalone". Le code est fourni avec beaucoup de commentaires, pour que quelqu'un qui ne le connaisse pas puisse le comprendre. Le travail en

équipe a été efficace, avec une répartition équilibrée des tâches, ce qui a permis de mener à bien le projet dans les délais impartis.

Cela dit, plusieurs pistes d'amélioration existent. L'une des principales améliorations possibles réside dans une interface graphique, permettant une analyse simple pour les utilisateurs débutants. Enfin, une version plus robuste du peuplement permettrait une vraie analyse du réseau de bus de La Rochelle, afin d'optimiser l'offre du réseau.

En conclusion, ce projet a non seulement permis d'approfondir nos connaissances techniques en matière de traitement de bases de données, mais aussi d'acquérir des compétences précieuses en matière de travail collaboratif. Ces compétences seront sans aucun doute utiles dans nos futurs projets et nous ont préparés à affronter des défis encore plus complexes.