



Projet Predit Mobilitic : Analyse de la mobilité par les données billettiques

Livrable 2 : Analyses préliminaires des données billettiques du réseau STAR de la ville de Rennes



Table des matières

Introduction	2
1 Présentation des données	2
1.1 Informations générales sur le réseau et les données billettiques	2
1.2 Titres de transport	3
1.3 Nombre d'usagers du réseau	4
1.4 Répartition des validations dans le temps	5
2 Interaction des variables avec le type de titre	9
A Plans du réseau	15
A.1 Plan des lignes urbaines	15
A.2 Plan des lignes métropolitaines	16
B Fréquentation par titre	16
B.1 Fréquentation des lignes par titre	16

Introduction

Dans le cadre de nos travaux sur les données billettiques, nous allons utiliser les données provenant du réseau de transport en commun de l'agglomération rennaise (Réseau STAR). En effet depuis 2006, la ville de Rennes dispose d'une carte sur laquelle l'utilisateur peut charger un titre de transport lui permettant de se déplacer dans l'agglomération, la carte Korrigo. À chaque validation du titre, lors de la montée dans un bus ou de l'entrée dans le métro, une trace de la transaction est alors enregistrée. Chaque trace contient un identifiant anonymisé unique, des informations sur le lieu de validation, la ligne, le titre et l'identifiant anonymisé de l'utilisateur, ainsi que les dates et heures de la transaction.

Dans ce rapport nous présentons une première étude statistique effectuée sur ces données, tout d'abord en nous intéressant à chaque information indépendamment (lignes, stations, titres, nombre de validations, etc...). Puis nous croiserons les informations issues de ces variables dans le but d'identifier dans quelle mesure celles-ci interagissent entre elles.

1 Présentation des données

Dans cette section nous allons nous intéresser plus particulièrement aux différentes informations extraites de notre jeu de données, notamment les lignes et les stations utilisées, les jours et heures de validation et le type d'utilisateur. Pour l'ensemble de cette section, sauf indication contraire, nous nous plaçons sur une période de 7 mois d'avril à octobre 2014. Le jeu de données sur cette période contient un total de 29569150 soit environ 4224164 validations par mois et 138174 validations par jour.

1.1 Informations générales sur le réseau et les données billettiques

Le réseau STAR ou Service de Transport en Commun de l'Agglomération Rennaise, est un Service de transport basé sur Rennes et son agglomération. Celui-ci est composé d'une ligne de métro (ligne a), d'un grand nombre de lignes de bus (une Ligne Chronostar, 7 lignes majeures, 2 lignes urbaines, 6 lignes inter-quartiers, 19 Lignes Express, 26 Lignes Métropolitaines, 3 lignes intercommunales, 1 ligne de nuit et 70 Services Scolaires) et de 5 parcs relais. Ainsi, 43 communes, soit 400.000 habitants bénéficient d'une desserte de transports en commun. Le plan des lignes urbaines et celui des lignes inter-urbaines se trouvent en annexe (A).

Sur la période étudiée, nous avons pu observer des validations sur 119 lignes différentes (certaines lignes n'ont donc pas été observées). La ligne la plus utilisée est l'unique ligne de métro, ou ligne a. On peut voir sur la figure 1 que le nombre de validations effectuées sur cette ligne est bien supérieur au nombre de validations effectuées sur les autres lignes.

Nous nous intéressons ici à la fréquentation des différentes stations du réseau STAR en différenciant les stations par leur nom. Cela permet de regrouper les stations bus et métro se trouvant au même endroit sous la même étiquette. Sur la figure 2 on peut noter que la station République concentre la majorité des validations. Le réseau de transports de la ville formant une étoile dont le centre est la station République, il est logique de retrouver cette station comme étant celle avec le plus grand nombre de validation. On peut également noter qu'un petit groupe de station reçoit un grand nombre de validations et que celles-ci correspondent presque toutes à des stations de métro. Il nous paraît alors intéressant d'analyser les stations les plus fréquentées une fois les validations dans le métro retirées.

On peut constater sur la figure 3 que, bien que la fréquentation ait notablement baissé à certaines stations (comme Charles de Gaulle), le pic de fréquentation reste situé à la station

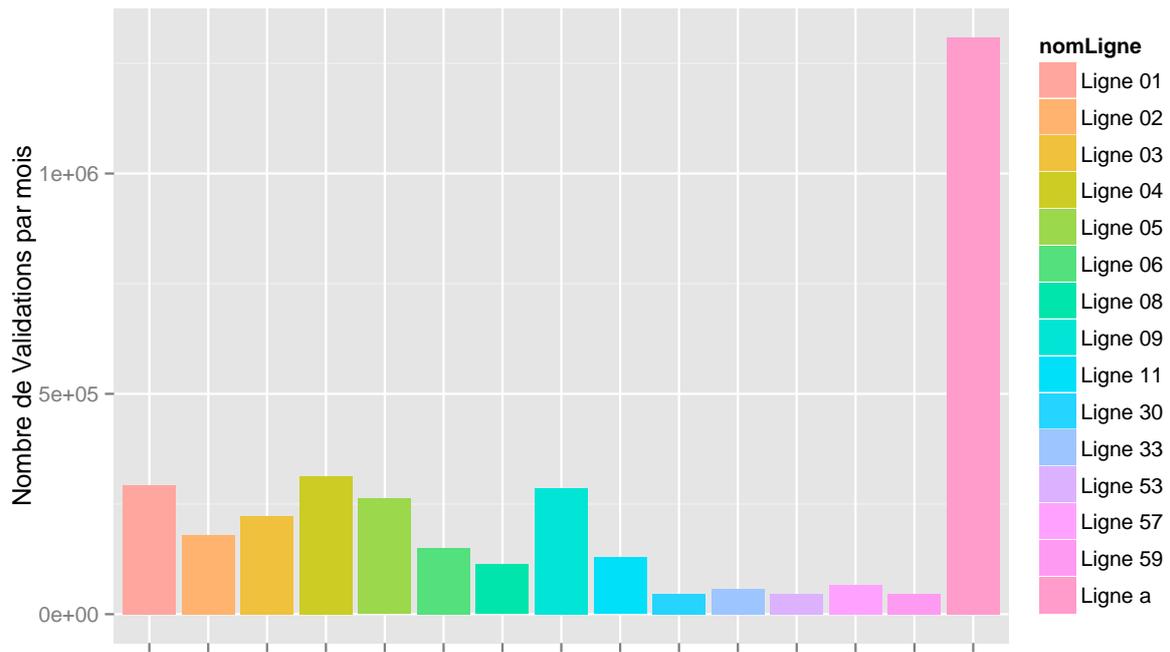


FIGURE 1 – Nombre moyen de validations par mois pour les 15 lignes les plus fréquentées (Avril-Octobre 2014)

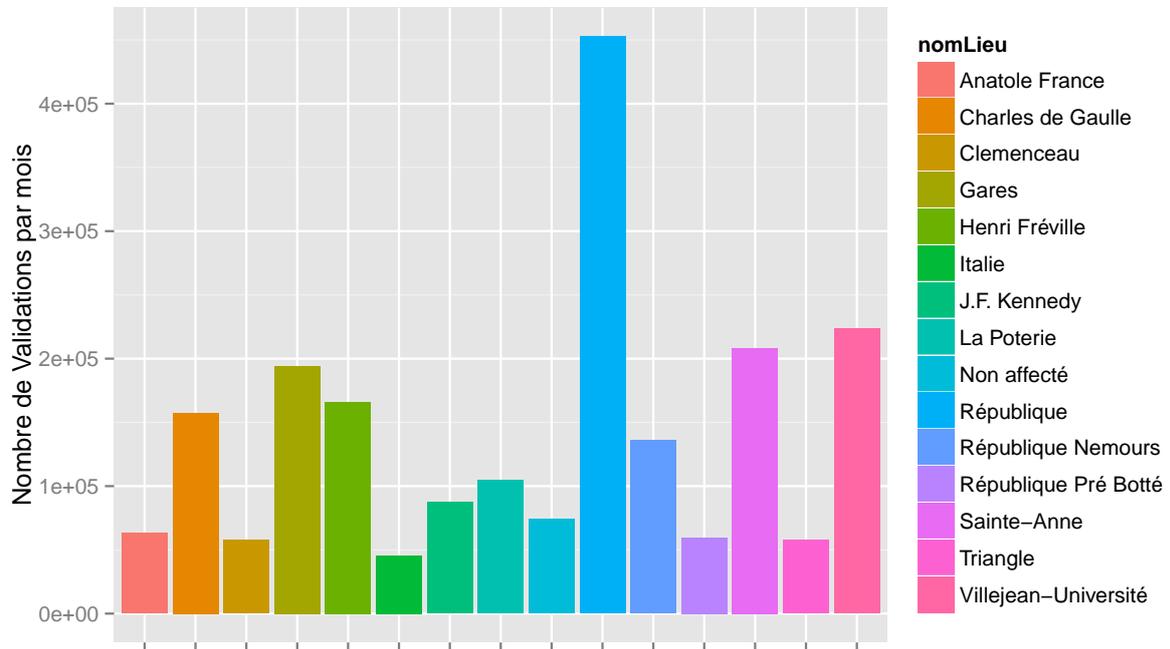


FIGURE 2 – Nombre moyen de validations bus et métro par mois pour les 15 stations les plus fréquentées (Avril-Octobre 2014)

république. On observe également un nouveau pic à la station République Nemours.

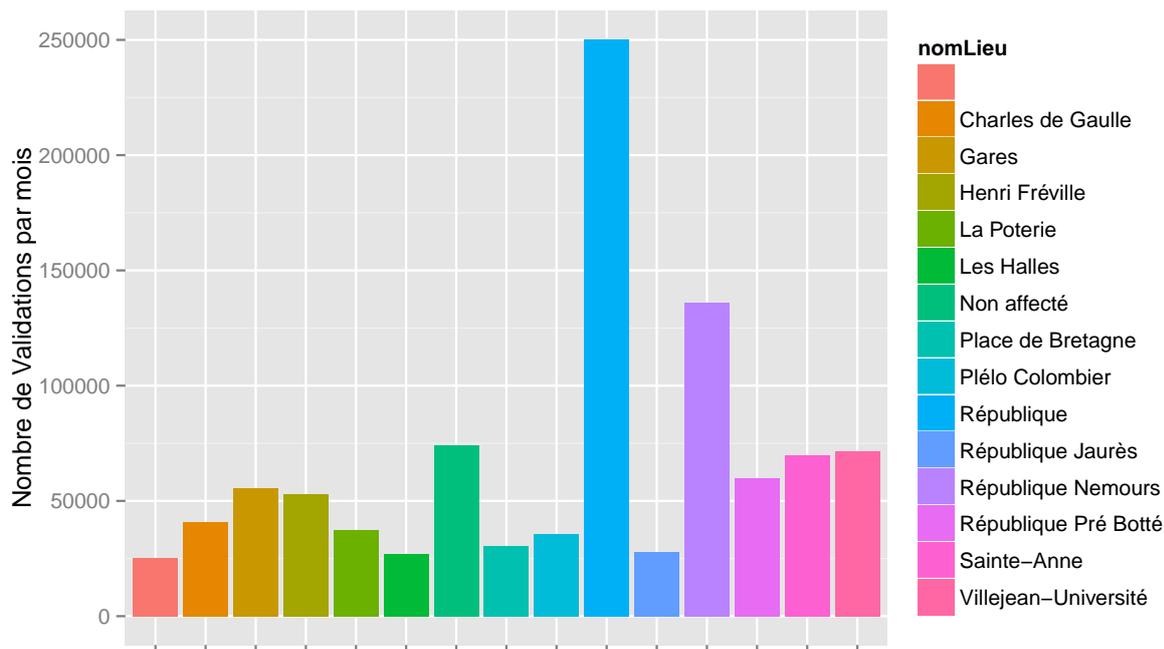


FIGURE 3 – Nombre moyen de validations bus par mois pour les 15 stations les plus fréquentées (Avril-Octobre 2014)

Un autre point qui mérite d’être examiné concerne la répartition géographique de ces stations. Sur le plan de la ville fourni figure 4, on retrouve bien le fait que la station République se trouve au centre de la ville et donc du réseau, d’où son grand nombre de validation. On peut voir également que la station Villejean Université, bien qu’excentrée attire elle aussi un grand nombre de voyageurs. Rennes étant une ville étudiante, il n’est pas étonnant que cette station enregistre un grand nombre de passage. Enfin on peut remarquer que les stations au sud du centre ville attirent elles aussi de nombreux voyageurs. Ces stations correspondant pour la plupart à des ensembles résidentiels (Triangle, Italie, ...), leur forte fréquentation est sans doute liée aux déplacements de leur population.

Un deuxième aspect à prendre en compte est la localisation de la ligne de métro. Comme nous avons pu le voir précédemment, cette ligne attire un nombre d’usagers bien supérieur à celui des autres lignes. Il est alors normale que les stations bus présentes sur cette ligne reçoivent un nombre plus important d’usagers et ce, même en retirant les validations propres aux stations de métro. En effet beaucoup d’usagers utilisent le métro, mais une fois sorti de celui-ci les usagers ont souvent besoin de prendre un bus, provoquant ainsi une validation à la station sans pour autant être une validation due au métro.

1.2 Titres de transport

Le réseau STAR dispose d’un très grand nombre de titres de transports différents, ceux-ci allant du simple ticket aux abonnements couplés STAR+TGV. Sur la période étudiée 108 titres différents ont été observés. Sur la figure 5, les 10 titres les plus utilisés ont été représentés. On peut voir que la majorité des validations est effectuée avec les tickets, les abonnements standards (abonnés -20 ans, 20-26 ans et 27-64 ans) et les abonnements subventionnés (gratuité), les autres titres bien que nombreux restent donc minoritaires.

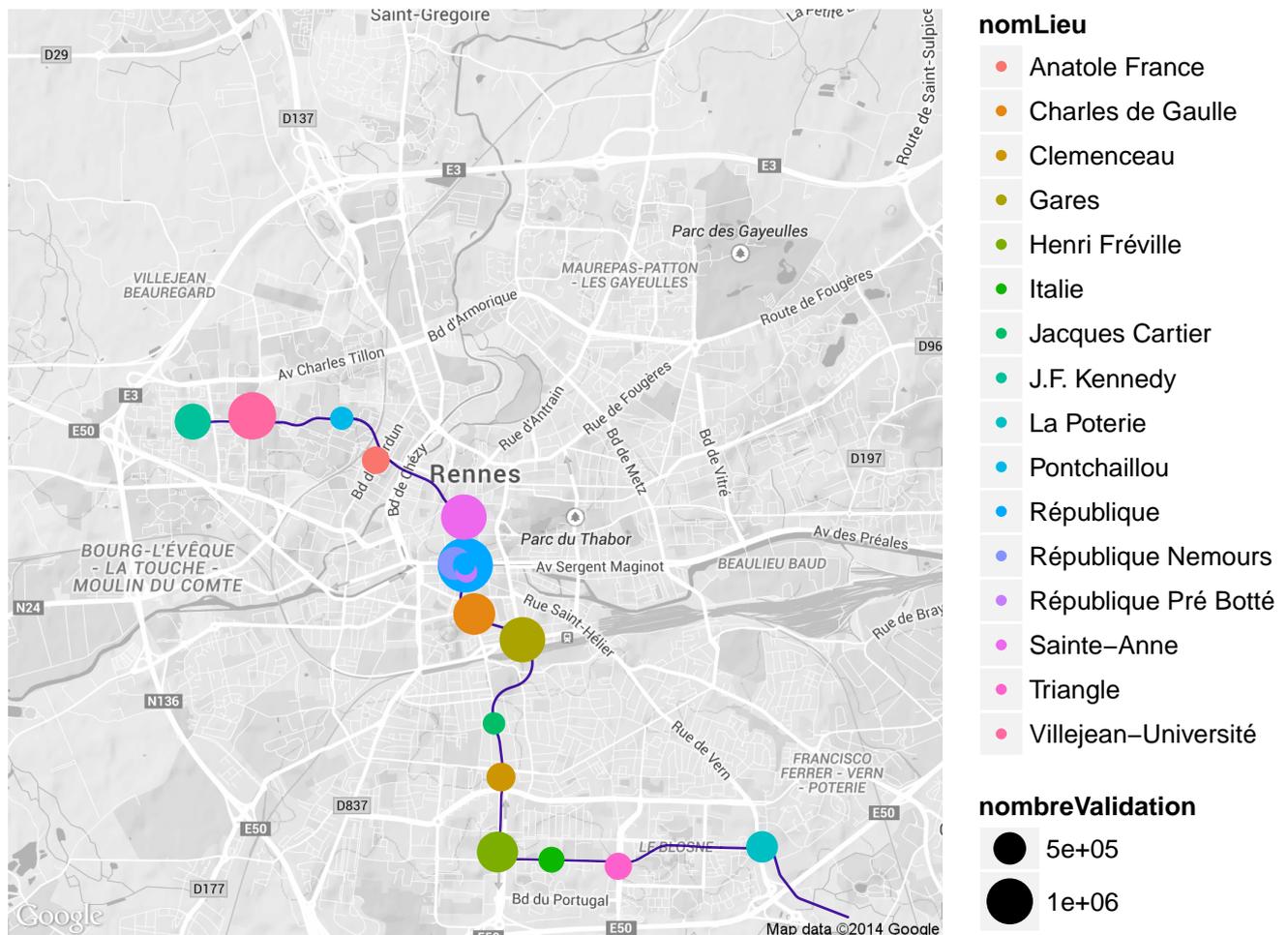


FIGURE 4 – Carte du réseau représentant les 15 stations (bus et métros) les plus actives (Avril-Octobre 2014)

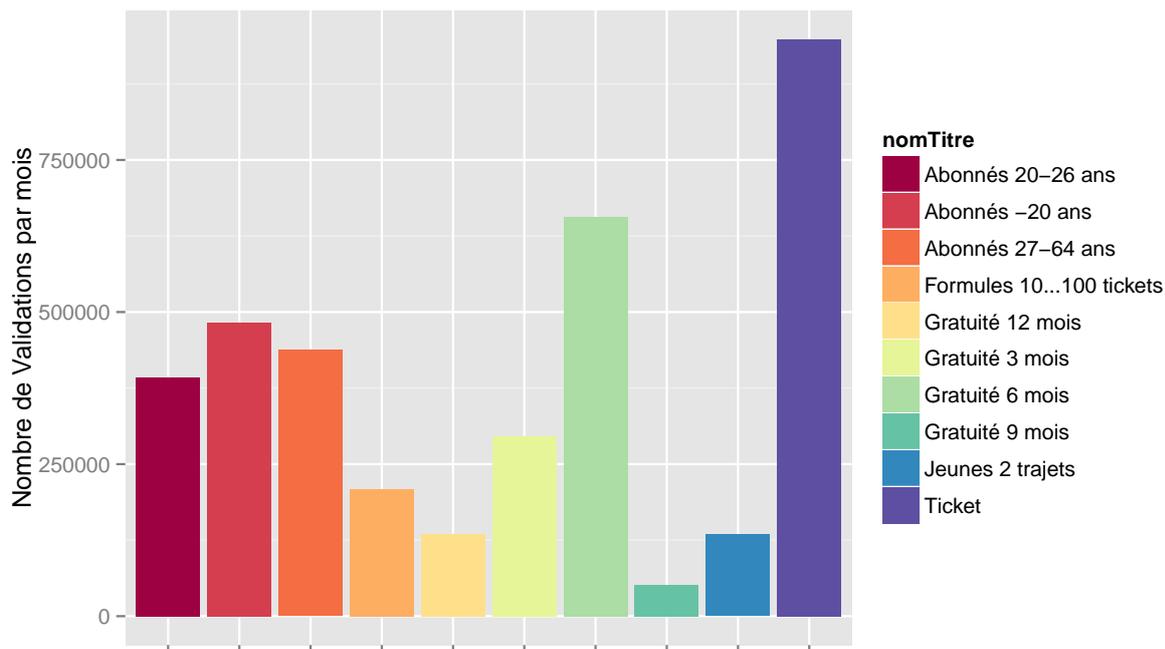


FIGURE 5 – Nombre moyen de validations par mois pour les 10 Titres les plus utilisés (Avril-Octobre 2014)

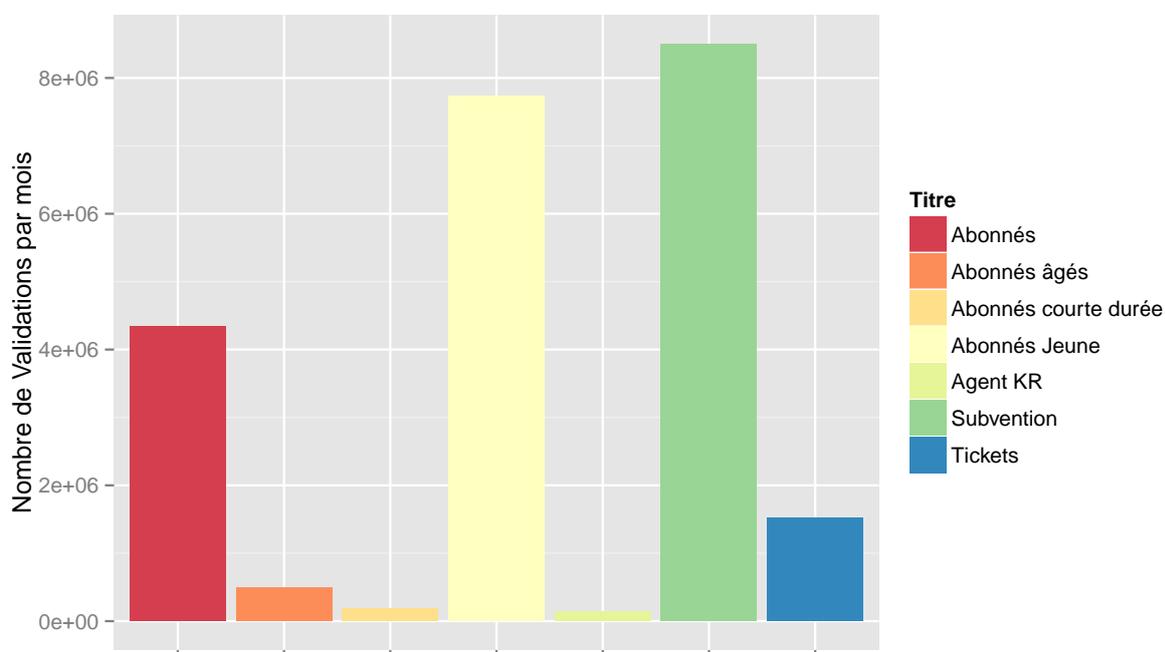


FIGURE 6 – Nombre moyen de validations par mois par type de titre (Avril-Octobre 2014)

Pour des raisons de lisibilité et à cause du grand nombre de titre, nous avons réduit ce nombre en classant les usagers en 8 catégories en fonction de leur titre : les abonnés, les abonnés jeunes,

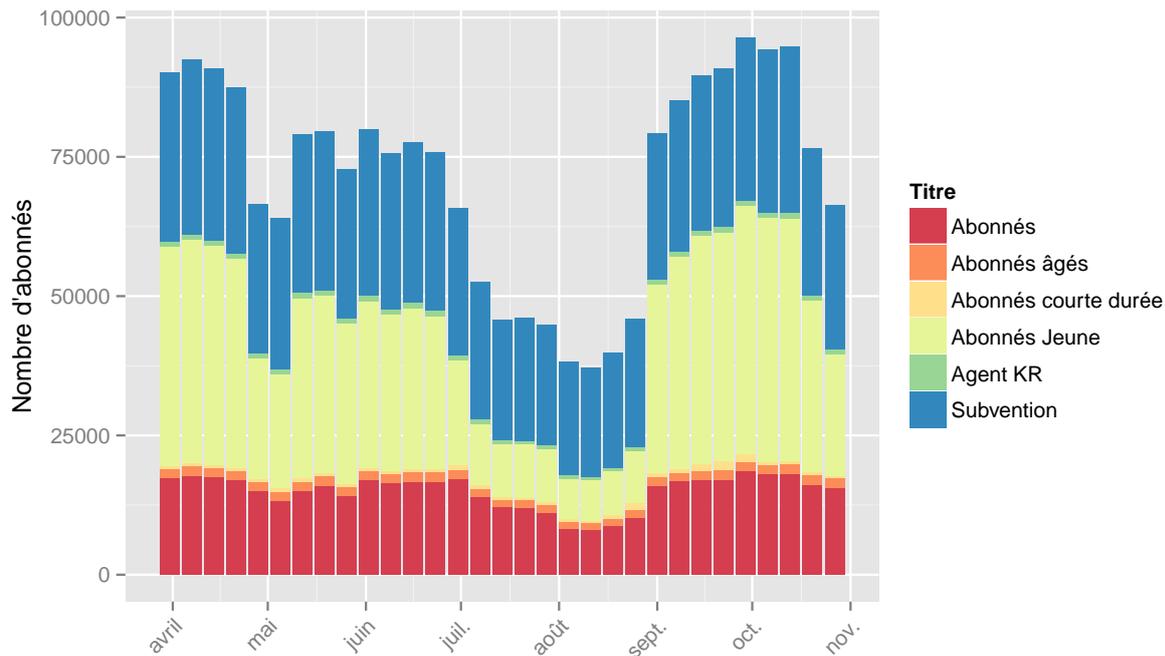


FIGURE 7 – Évolution du nombre d’usagers par semaine et par titre

les abonnés âgés, les abonnés courte durée (pass 1 à 7 jours), les agents KR, les subventionnés (personnes bénéficiant de la gratuité pour raison sociale ou scolaire) et enfin les usagers utilisant des tickets. On représente alors sur la figure 6 le nombre de validation enregistrées sur les mois d’Avril à Octobre pour l’ensemble des types de Titre. Il apparaît clairement que les validations sont principalement dues aux usagers utilisant des tickets, aux subventionnés, aux abonnés jeunes et aux abonnés. Rennes étant une ville connue pour son grand nombre d’étudiant, il n’est pas étonnant de retrouver un aussi grand nombre de validations d’abonnement de type jeune.

1.3 Nombre d’usagers du réseau

Sur toute la période de l’étude, nous suivons un total de 202717 voyageurs. Ceux-ci ne sont pas nécessairement actifs sur l’ensemble de notre étude. Pour mieux étudier leur activité nous avons analysé le nombre de début d’abonnement et de fin d’abonnement tout au long de la période analysée. Il est important de noter que nous appelons nouvel abonnement la date de première utilisation d’un abonnement et fin d’abonnement la date de fin d’utilisation d’un abonnement sur la période étudiée. Nous introduisons alors un biais en supposant que les abonnements ne sont pas utilisés en dehors de la période de l’étude, engendrant ainsi un grand nombre de nouvel abonnement en début de période et un grand nombre de fin d’abonnements en fin de période.

On retrouve sur la figure 7, la période de creux de juillet-août que nous avons déjà observée précédemment dans le nombre de validation. Cependant on peut voir ici que cette réduction du nombre de validation est principalement dûe aux jeunes et à certains abonnés alors que le nombre de validation par ticket reste à peu près stable.

Sur la figure 8, on observe le nombre de nouveaux abonnements cumulés par type de titre. Comme énoncé précédemment, on appelle "nouvel abonnement" la date de première valida-

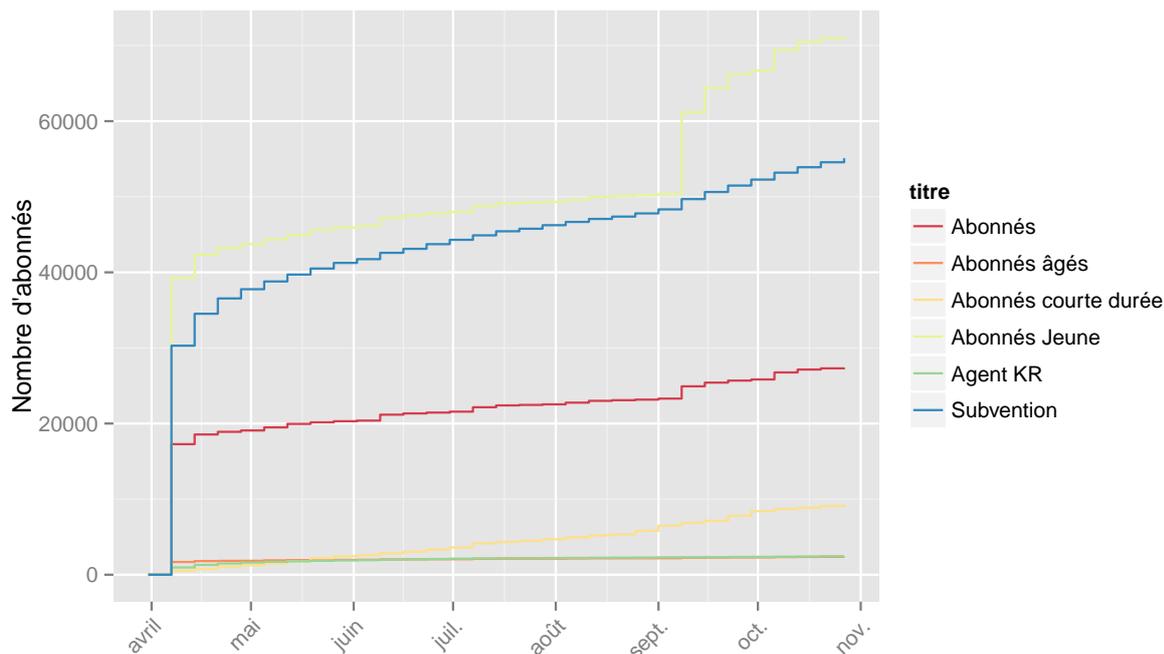


FIGURE 8 – Évolution du nombre de début d'abonnement

tion sur le réseau à partir du mois d'avril de l'année 2014. On peut ainsi observer un saut du nombre de nouveaux usagers en septembre pour les abonnements jeunes, date correspondant à la rentrée scolaire, et un saut plus faible pour les abonnements adulte toujours à la même date.

On appelle fin d'abonnement la date de dernière validation avant novembre de l'année 2014. On trace également (figure 9) le cumule des fins d'abonnement et on remarque une accentuation de celles-ci en juillet, date probable de départ en vacances et de départ définitif pour les étudiants présent à Rennes uniquement pour leurs études, ainsi qu'en octobre.

1.4 Répartition des validations dans le temps

Afin d'avoir une idée plus précise sur la répartition des validations dans le temps, nous avons tracé sur la figure 10 le nombre de validation pour chaque jour sur l'ensemble de la période étudiée (avril à octobre). Il apparaît qu'une grande majorité des validations est effectuée en semaine, les journées du samedi et du dimanche enregistrant presque toujours un nombre plus faible de validations. On note également une chute dans le nombre de validation durant les mois de juillet, août, qui correspondent aux vacances scolaires. Ainsi on peut voir une recrudescence des validations le premier septembre, ce qui correspond à la date de reprise du travail pour la plupart des usagers.

On retrouve sur la figure 11, qui représente le nombre de validation pour chaque jour sur l'ensemble de la période étudiée, que le nombre de validation est plus faible en fin de semaine (samedi et dimanche), qu'en semaine.

Nous traçons sur la figure 12 le nombre de validation par heure pour la première semaine de septembre (nous restreignons l'intervalle précédent pour plus de visibilité).

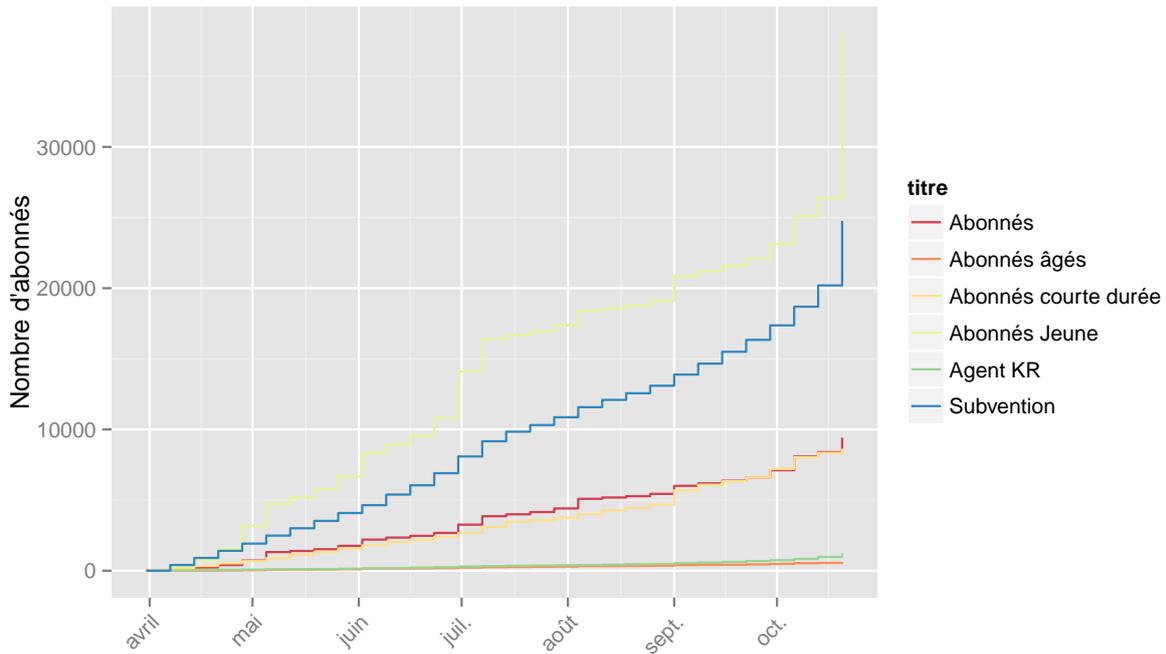


FIGURE 9 – Evolution du nombre de fin d’abonnement

On remarque rapidement la présence de trois pics en semaine, deux principaux le matin et le soir et un plus faible vers midi. Le pic du matin semble moins étalé que celui du soir. On peut également noter que ces pics sont principalement dûs aux abonnés, la consommation de ticket semblant régulière sur l’ensemble d’une journée en semaine.

En fin de semaine (samedi et dimanche), apparaît un motif plus diffus.

On voit également apparaître les vendredi et samedi soirs un plus grand nombre d’utilisation du réseau que les autres soirs de la semaine. Cette activité est majoritairement associée à des abonnements jeunes et des tickets. Ces soirs étant souvent associés aux soirs de sortie, il n’est pas étonnant de constater une plus forte activité.

2 Interaction des variables avec le type de titre

Toujours sur la même période on regroupe les données par type (titre, ligne, ...) avant d’effectuer la même analyse que précédemment. Sur la figure 19, on peut observer le nombre de validation par type de titre sur les 20 lignes les plus utilisées. Le but ici est de savoir dans quelle mesure certaines lignes sont utilisées par un groupe d’abonnés. On retrouve le fait que la ligne de métro (Ligne a) est la plus fréquentée. Pour mieux voir la proportion de chaque type d’usager on trace sur la figure 13 le taux d’utilisation de la ligne pour chaque usager. On remarque certaines différences selon les lignes mais ne pouvant pas les situer précisément nous ne pouvons rien conclure. Nous allons donc nous intéresser aux stations plutôt qu’aux lignes car celles-ci nous offrent des repaires géographiques fixes.

On trace sur la figure 14 le nombre de validations enregistrées pour chaque station et chaque titre. Toujours pour plus de visibilité nous traçons les taux de fréquentation par type d’usager sur la figure 15. Il apparaît alors que certaines stations attirent plus un certain type d’usagers que d’autres. On remarque notamment que la station Villejean-Université attire une proportion

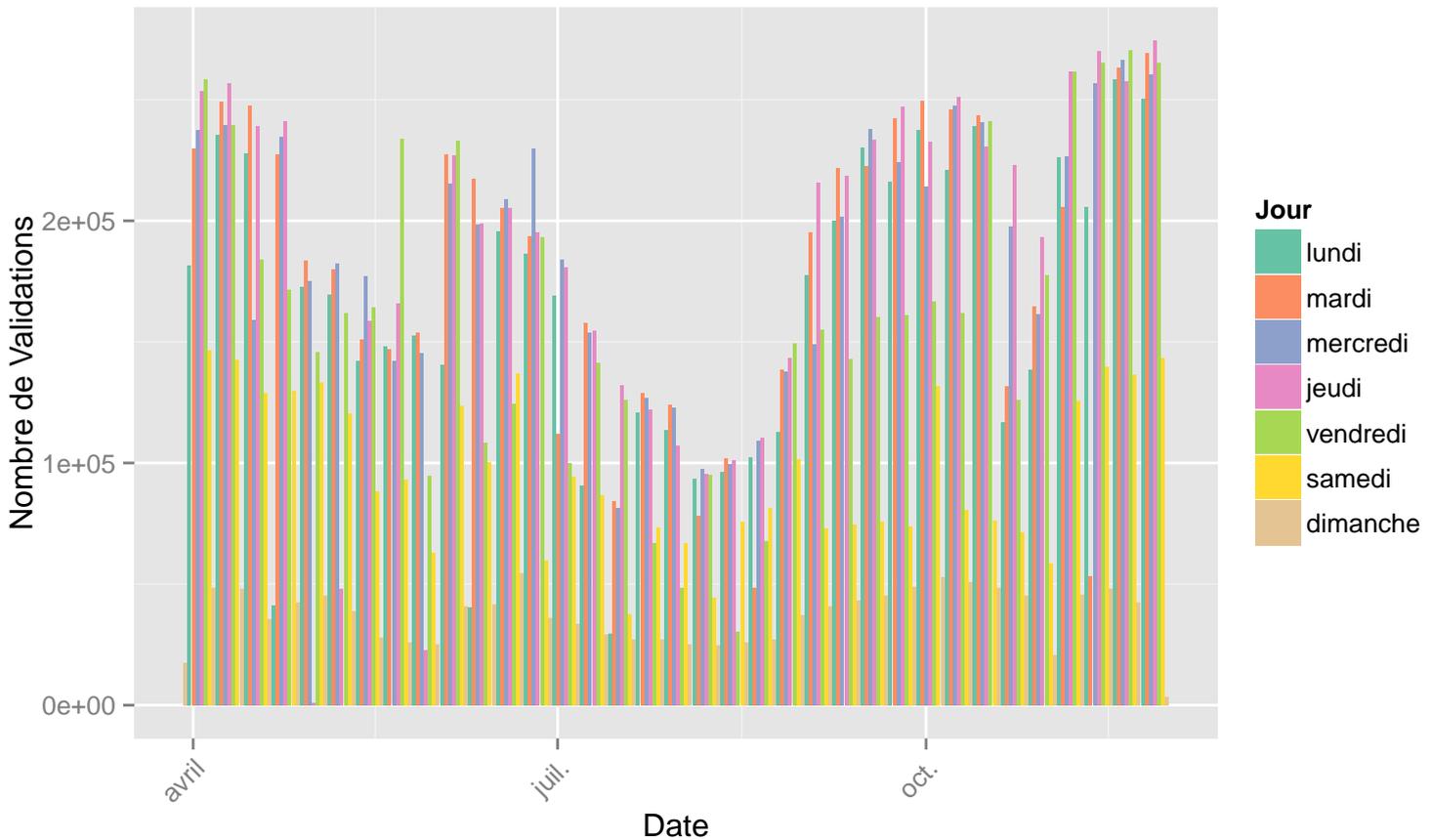


FIGURE 10 – Nombre de validations journalières d’avril à octobre

plus importante d’étudiant que les autres stations. Ou encore que les stations Triangle et Italie attirent un plus grand nombre d’usagers subventionnés. Le recouplement de ces données avec des données socio-économiques sur les zones entourant les stations permettraient de mieux interpréter nos résultats.

Nous nous intéressons ensuite au taux d’utilisation de chaque titre en fonction du jour de la semaine. On voit alors sans surprise (figure 16) que le résultat est très régulier les jours de semaine avec une forte proportion d’abonnés. En revanche en fin de semaine nous observons une chute de la proportion d’abonnés validants, au profit des usagers utilisant des tickets et des usagers subventionnés.

Toujours sur les 15 stations les plus fréquentées nous traçons le taux de fréquentation en fonction des jours de la semaine afin de voir si le samedi et le dimanche on une proportion moindre pour toute les stations (figure 17). Nous traçons également pour les mêmes stations le nombre de validation enregistré pour chaque jour de la semaine sur la période étudiée (figure 18).

On remarque que certaines stations comme Triangle, Italie enregistrent une différence moins grande entre la semaine et le week-end que des stations comme République ou Villejean-Université. Ces stations étant plus fréquentées par des usagers subventionnés qu’abonnés et ces usagers ayant montré une plus grande régularité d’utilisation quel que soit le jour de la semaine, il n’est alors pas étonnant de voir la fréquentation de ces stations baisser plus faiblement en fin de semaine. Cela est lié à l’usage qu’il est fait des stations, usage que l’on cherchera à

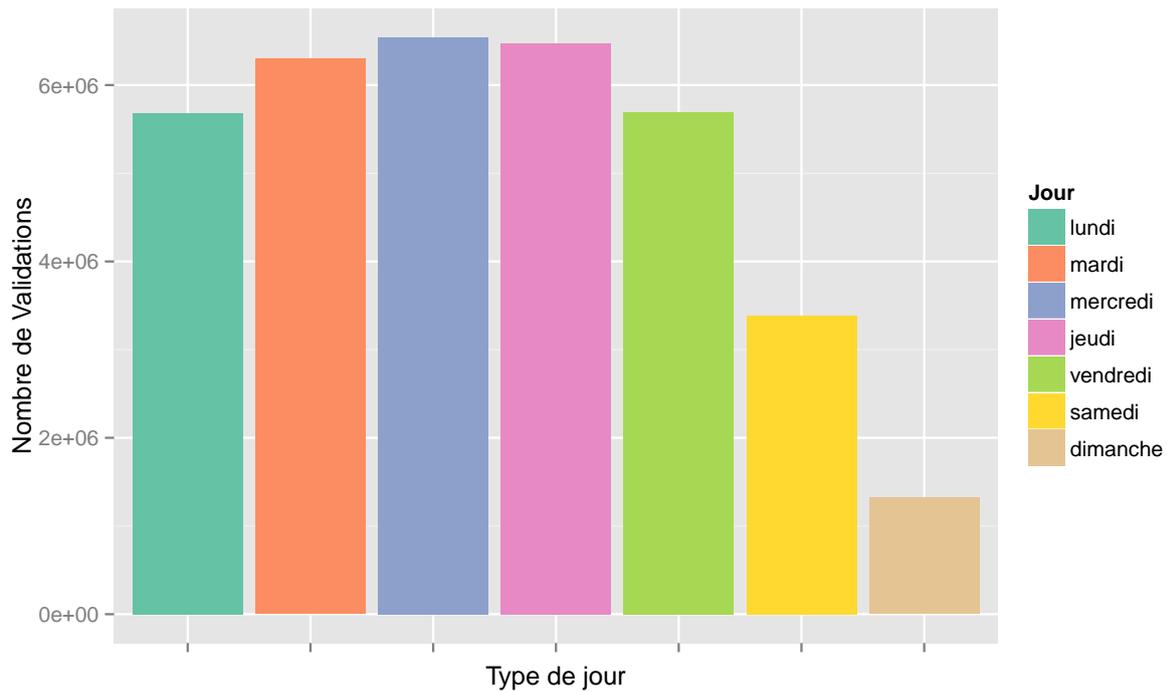


FIGURE 11 – Nombre de validations par jour de la semaine d’avril à octobre

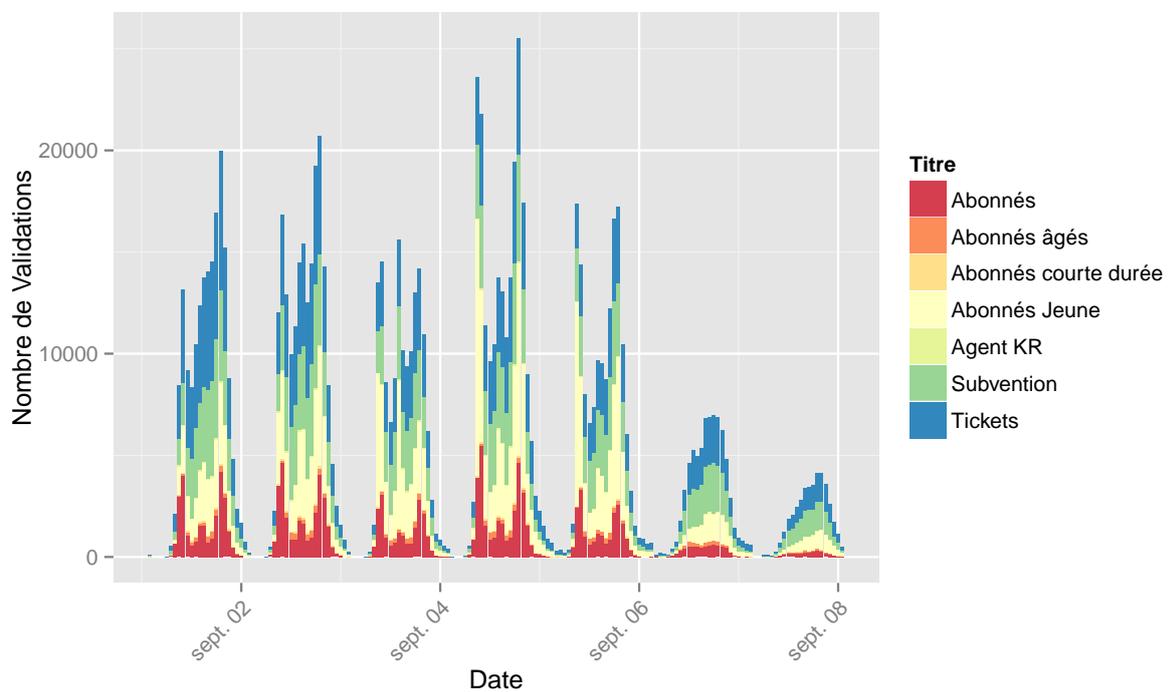


FIGURE 12 – Nombre de validations horaires par titre en septembre 2014

identifier avec des méthodes de clustering ou regroupement automatique).

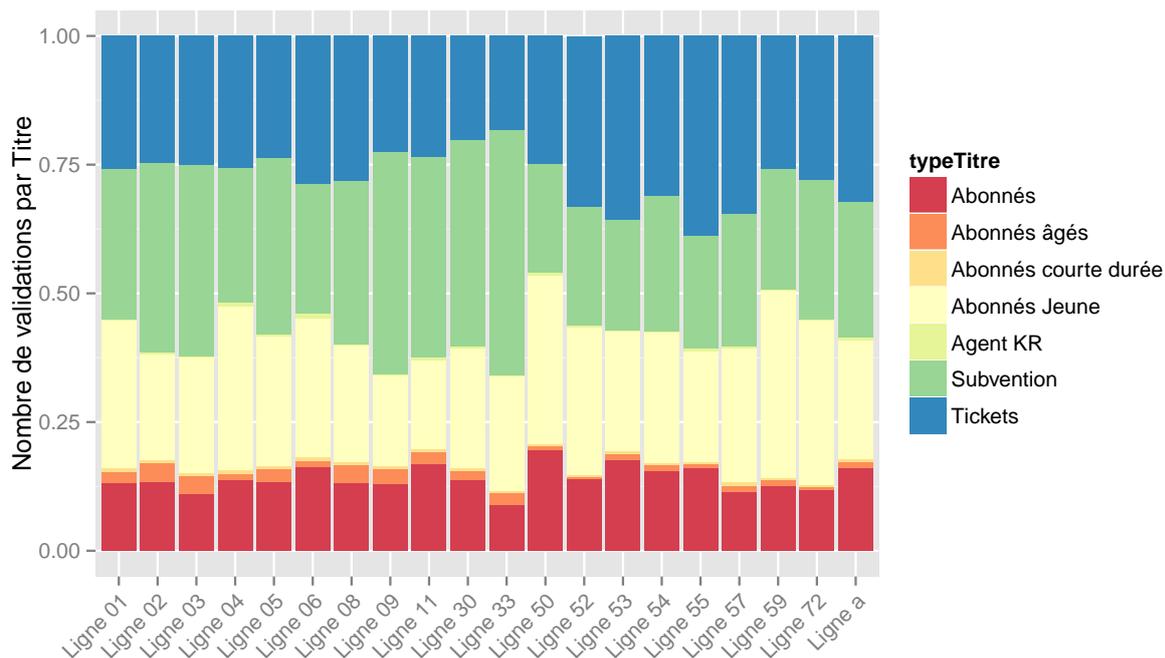


FIGURE 13 – Proportion d’utilisation des types de titre pour les 20 lignes les plus fréquentées

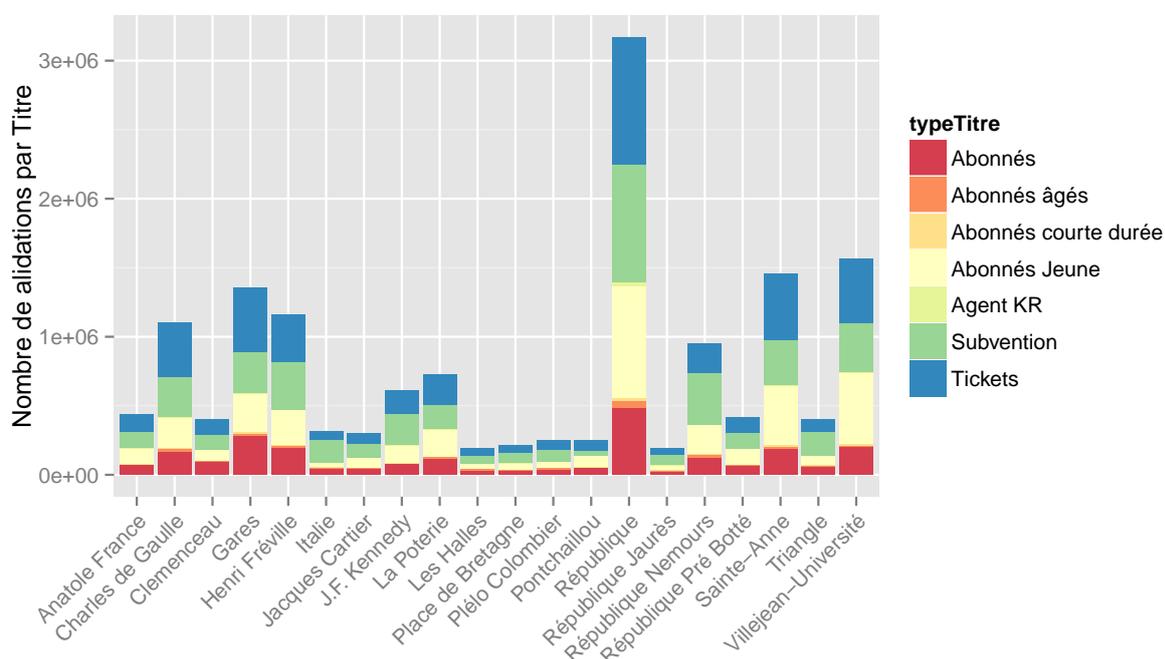


FIGURE 14 – Nombre de validations par type de titre sur les 20 stations les plus fréquentées

Conclusion

Ce rapport a présenté les résultats d’une première étude statistique réalisée sur les données billettiques du réseau STAR (ville de Rennes). Au travers de ces premiers résultats, nous avons acquis une meilleure connaissance du réseau de transport de la ville de Rennes sur lequel des

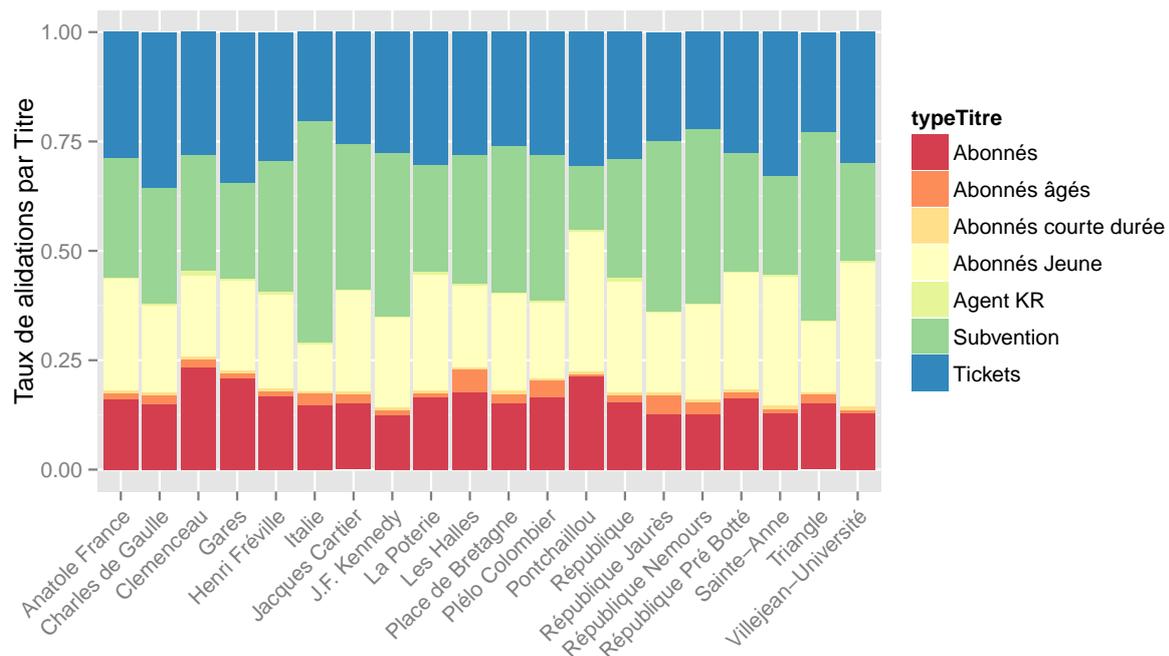


FIGURE 15 – Proportion d’utilisation des types de titre sur les 20 stations les plus fréquentées

analyses plus élaborées vont être mises en oeuvre.

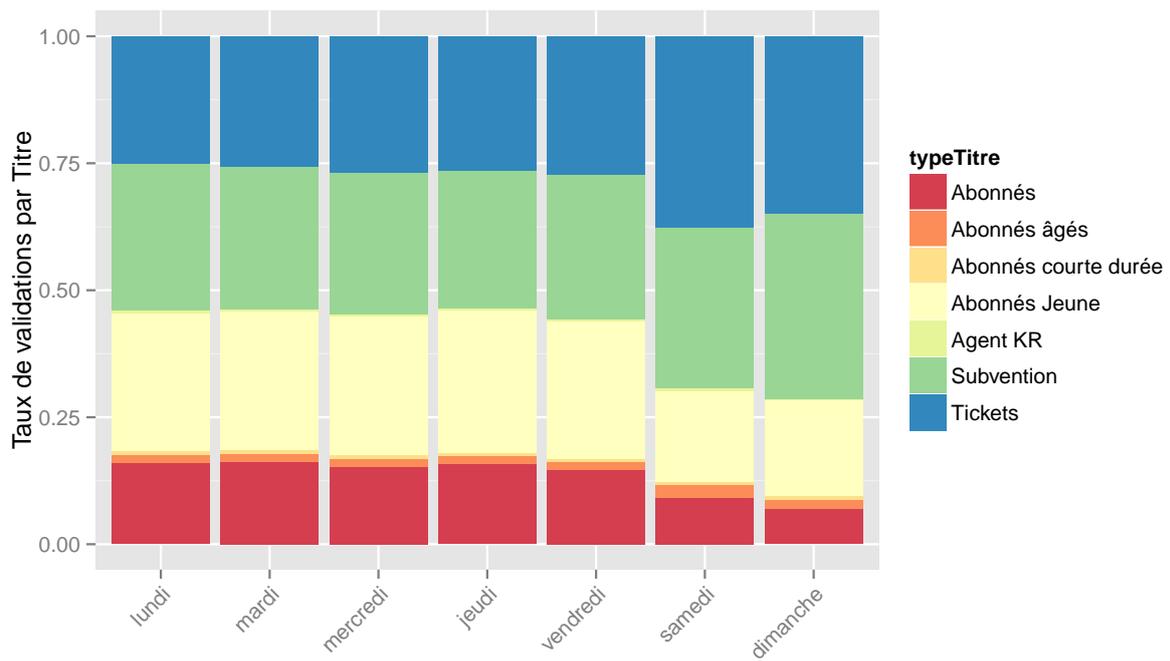


FIGURE 16 – Utilisation des transports en fonction du titre et du jour de la semaine

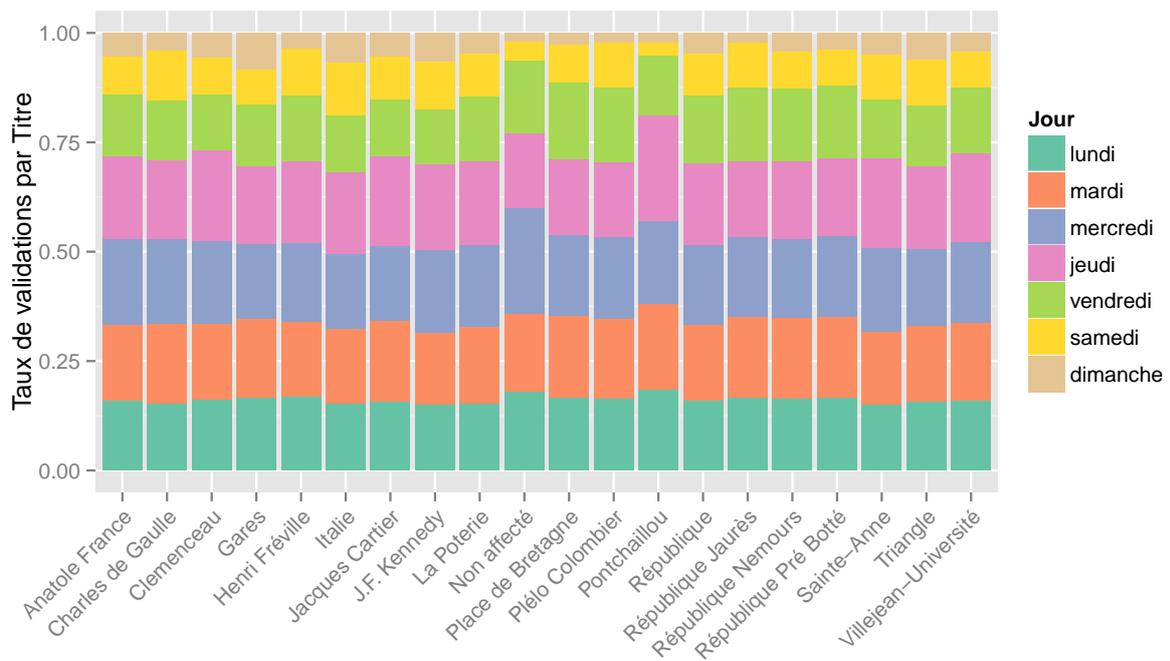


FIGURE 17 – Ratio d'utilisation des transports en fonction des jours de la semaine dans les 20 stations les plus fréquentées

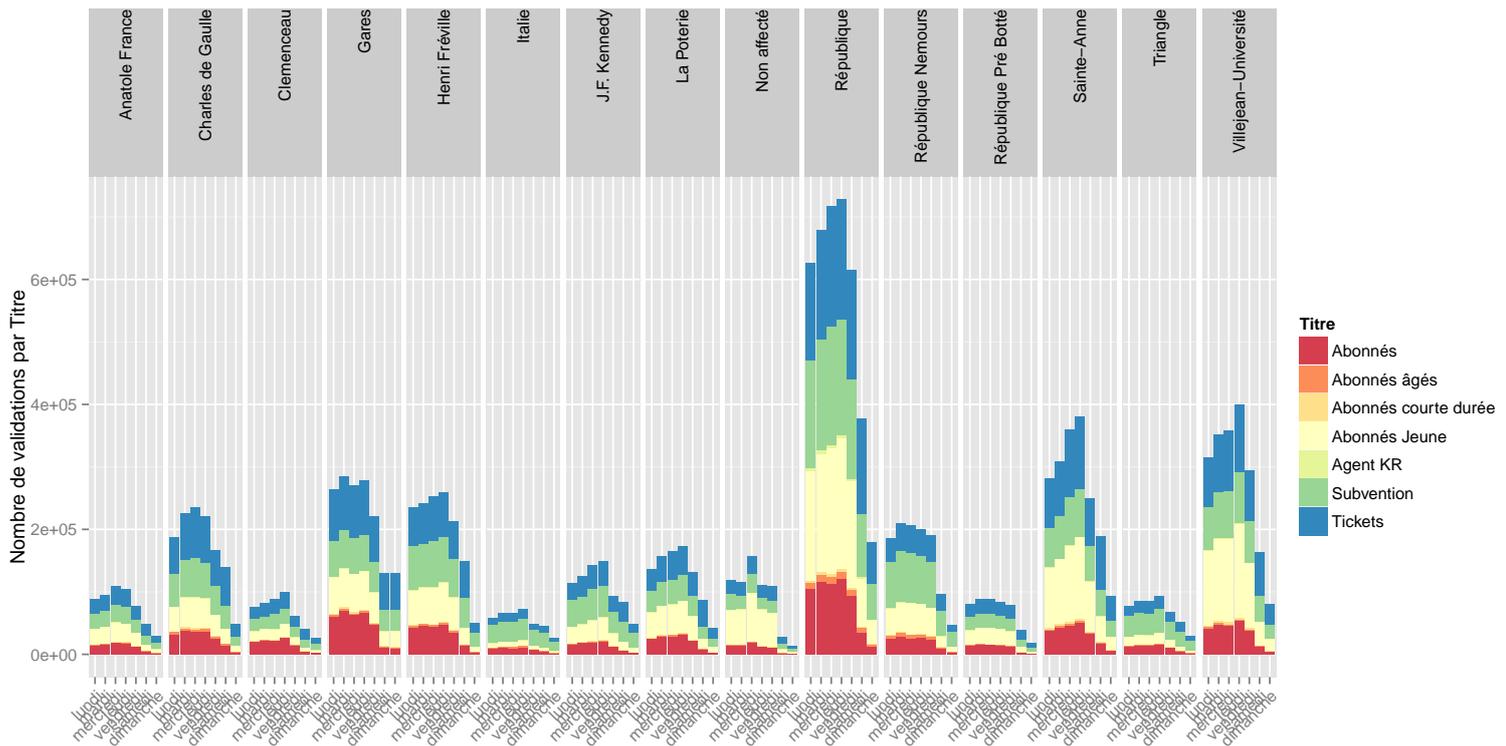
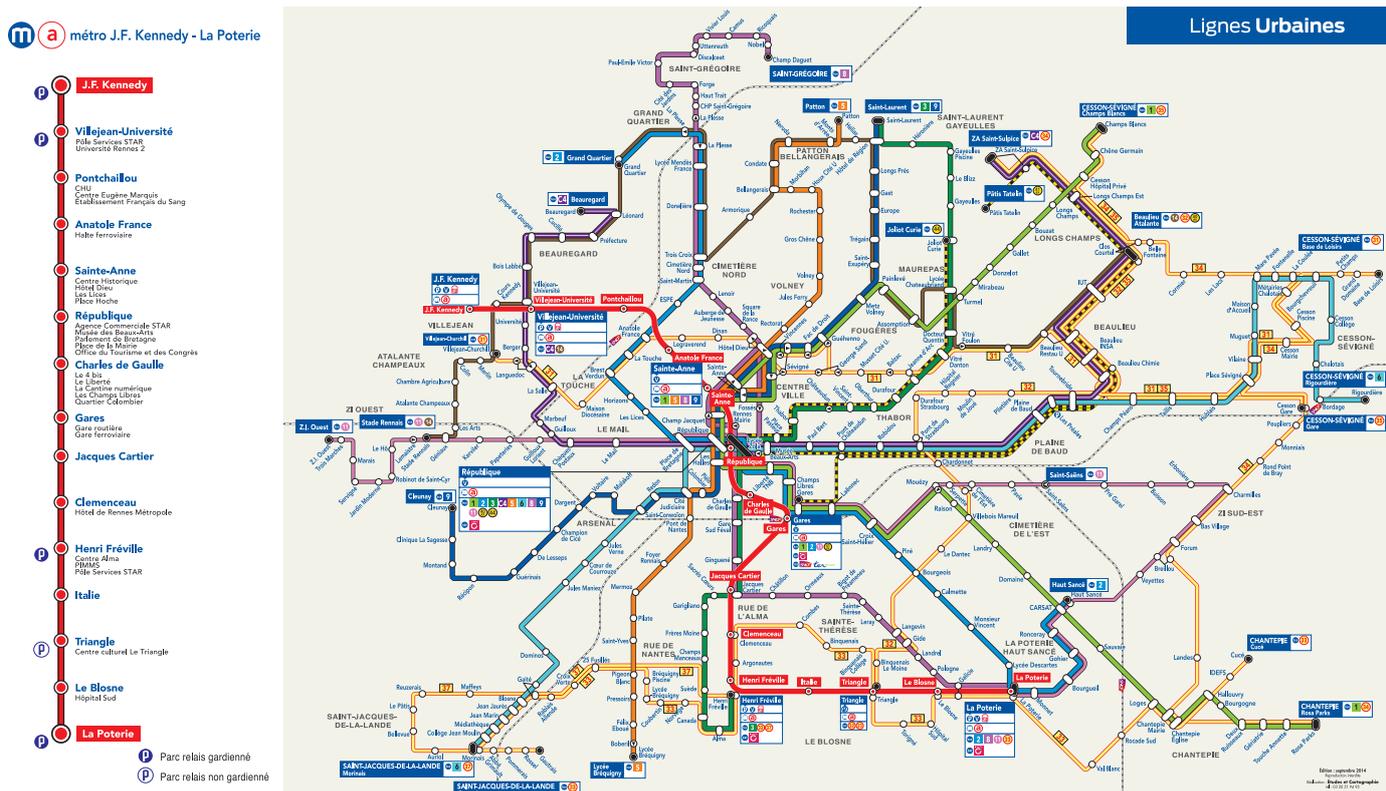


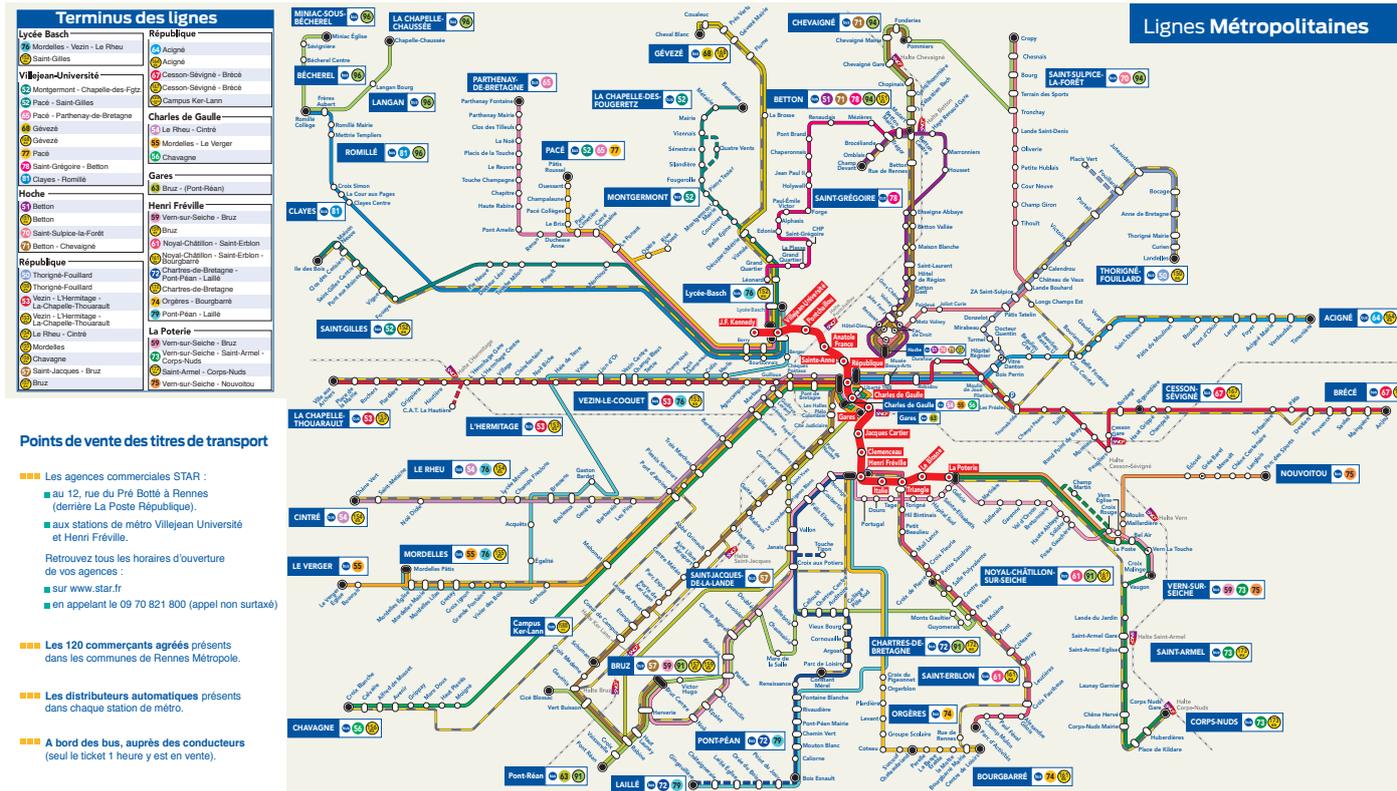
FIGURE 18 – Utilisation des transports en fonction des jours de la semaine et par titre sur les 15 stations les plus fréquentées

A Plans du réseau

A.1 Plan des lignes urbaines



A.2 Plan des lignes métropolitaines



B Fréquentation par titre

B.1 Fréquentation des lignes par titre

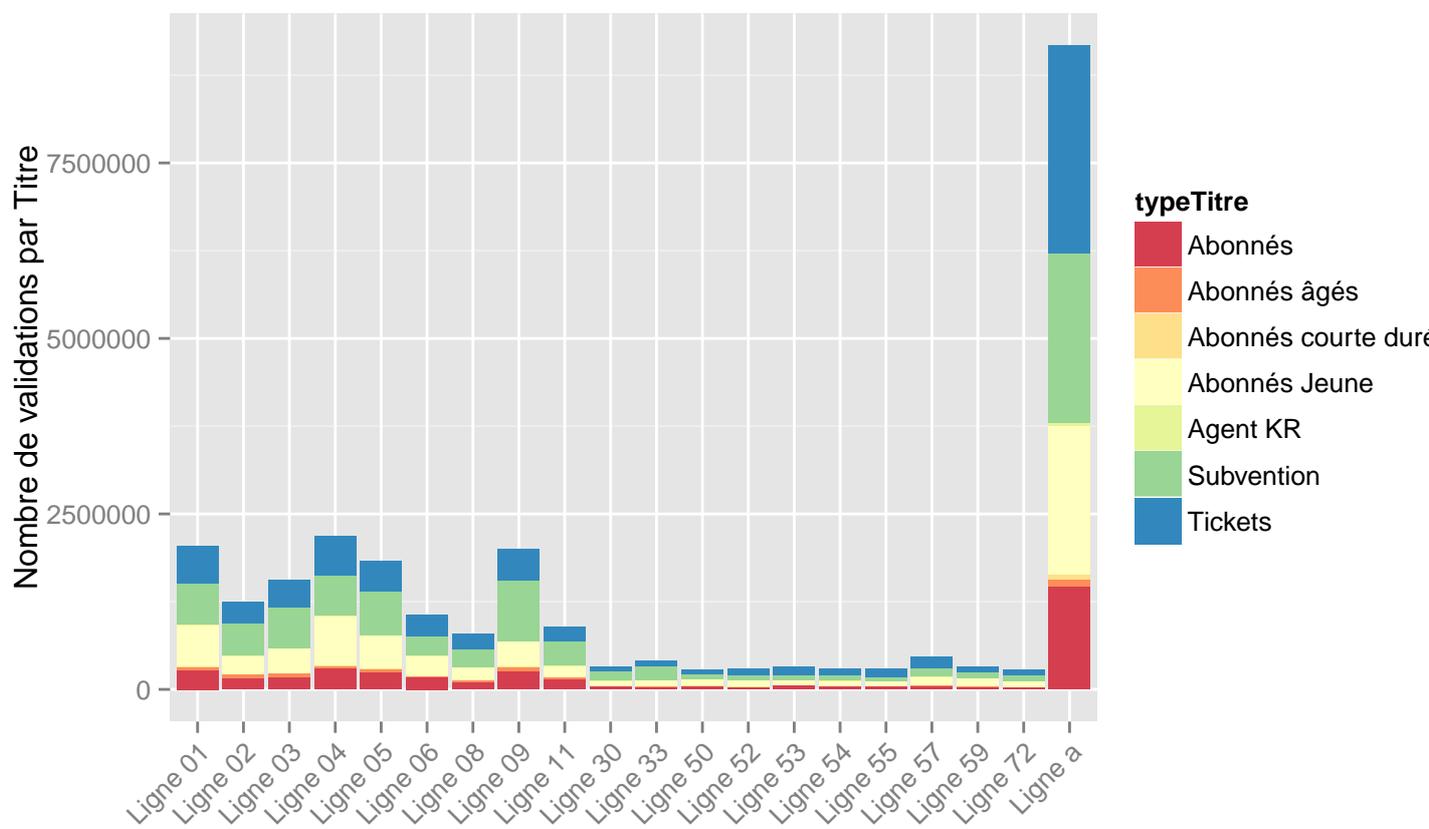


FIGURE 19 – Nombre de validations par type de titre pour les 20 lignes les plus fréquentées