

# Analyse de l'utilisation des vélos libre-service à Washington

Julien Schieler, Icham Lecorvaisier, Mathis Reslinger, Lila Mortier

---

## Analyse de l'utilisation des vélos libre-service à Washington

Julien Schieler, Icham Lecorvaisier, Mathis Reslinger, Lila Mortier

### Introduction

Ce rapport présente une analyse approfondie de l'utilisation des vélos libre-service à Washington, en exploitant quatre jeux de données pour identifier les tendances d'usage, les comportements des usagers et les facteurs influençant l'utilisation du service, afin de fournir des insights pertinents pour les décideurs publics et les acteurs de la mobilité durable.

### Données

Nous utilisons 4 jeux de données pour analyser l'utilisation des vélos libre-service à Washington :

**Information sur les locations de vélos** (source : [https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station\\_list.csv](https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station_list.csv))

- **Nom du dataset** : daily\_rent\_detail.csv
- **Origine des données** : Ces données proviennent de l'entreprise Capital Bike Share, l'entreprise de location de vélos libre service de Washington.
- **Pourquoi ces données ?** Elles permettent d'observer l'utilisation de ces vélos, en analysant les dates et heures d'utilisation, les lieux d'emprunt et de dépôt notamment.
- **Nombre d'observations** : 16 086 673
- **Nombre de colonnes** : 13 colonnes incluant :
  - Identifiant du trajet ("ride\_id")
  - Type du vélo ("rideable type") : vélo classique, électrique, autre (cargo)
  - Date et heure de l'emprunt ("started\_at")
  - Date et heure du retour ("ended\_at")
  - Nom de la station de départ ("start\_station\_name")
  - Identifiant de la station de départ ("start\_station\_id")
  - Nom de la station d'arrivée ("end\_station\_name")
  - Identifiant de la station d'arrivée ("end\_station\_id")
  - Latitude de départ ("start\_lat")
  - Longitude de départ ("start\_lng")
  - Latitude d'arrivée ("end\_lat")

- Longitude d'arrivée ("end\_lng")
- Type d'abonnement du client ("member\_casual") : utilisateur en abonnement casual ou membre.
- **Créateur et éditeur** : Entreprise Capital Bike Share
- **Format** : CSV
- **Sous-groupes** :
  - Type de vélo : vélo classique, électrique et cargo
  - Type d'utilisateur : occasionnel et membre
  - Stations : possibilité d'analyser par station d'origine ou de destination
  - Périodes temporelles : heures, jours, saisons, etc...

**Liste des stations de Washington** (source : [https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station\\_list.csv](https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station_list.csv))

- **Nom du dataset** : station\_list.csv
- **Origine des données** : Ces données proviennent de l'entreprise Capital Bike Share, l'entreprise de location de vélos libre-service de Washington.
- **Pourquoi ces données ?** Elles permettent d'avoir la liste des stations à Washington
- **Nombre d'observations** : 912 stations différentes
- **Nombre de colonnes** : 2 colonnes incluant :
  - Nom de la station
  - Identifiant de la station
- **Format** : CSV
- **Sous-groupes** :
  - Catégorisation géographique (ex: quartiers)

**Emprunts et dépôts des vélos par station** (source : [https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station\\_list.csv](https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=station_list.csv))

- **Nom du dataset** : usage\_frequency.csv
- **Origine des données** : Ces données proviennent de l'entreprise Capital Bike Share, l'entreprise de location de vélos libre-service de Washington.
- **Pourquoi ces données ?** Elles permettent d'avoir le nombre d'emprunt et de dépôt pour chaque station à Washington
- **Nombre d'observations** : 873 318
- **Nombre de colonnes** : 4 colonnes incluant :
  - Date
  - Nom de la station
  - Nombre d'emprunt sur cette station
  - Nombre de dépôt sur cette station
- **Format** : CSV
- **Sous-groupes** :
  - Date
  - Stations : possibilité d'analyser par stations
  - Activité : Emprunt et dépôt
  - Utilisation : stations très utilisées et peu utilisées

**Météo** (source : <https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408?select=weather.csv>)

- **Nom du dataset** : weather.csv
- **Origine des données** : Ces données proviennent de l'entreprise Visual Crossing qui est un service de météo de Virginia aux Etats Unis.
- **Pourquoi ces données ?** Elles permettent d'obtenir la météo de chaque jour afin de pouvoir croiser ces données avec l'utilisation des vélos à Washington
- **Nombre d'observations** : 1584, une pour chaque jour entre mai 2020 et août 2024
- **Nombre de colonnes** : 32 colonnes dont 7 que nous allons utiliser :
  - Date ("*datetime*")
  - Moyenne de température sur la journée ("*temp*")
  - Moyenne de température ressentie sur la journée ("*feelslike*")
  - Précipitations ("*precip*") : quantité de liquide tombé / prévu dans la période
  - Epaisseur de neige au sol ("*snowdepth*")
  - Couverture nuageuse ("*cloudcover*") : entre 0 et 100%
  - Vitesse du vent ("*windspeed*") : vitesse moyenne du vent soutenu, mesurée comme la vitesse moyenne du vent survenant au cours de la à deux minutes précédentes
- **Créateur et éditeur** : Visual Crossing (fournisseur de données météorologiques)
- **Format** : CSV
- **Sous-groupes** :
  - Conditions météorologiques : pluie, neige, vent, etc...
  - Températures : froid, tempéré, chaud
  - Saisons : printemps, été, automne, hiver
  - Jours : semaine ou week-end

## Plan d'analyse

Avant de commencer l'analyse, nous nous posons plusieurs questions que l'on peut catégoriser :

### 1. Utilisation générale des vélos

- Q1 - Combien de trajets sont effectués en moyenne par jour/mois ?
- Q2 - Quelle est la tendance générale de l'utilisation des vélos au cours de l'année ?
- Q3 - Quelles sont les durées moyenne des trajets ? Quels éléments influent sur cette durée (jour de la semaine, type d'utilisateur) ?
- Q4 - Quelle est l'évolution de l'utilisation des différents types de vélos ?
- Q5 - En se focalisant sur les utilisateurs, y'a-t-il des effets "heures de pointes" dans la journée ?

### 2. Utilisation des stations

- Q6 - Quelles stations sont les plus utilisées pour les départs/arrivées ?
- Q7 - Quelles stations sont les plus actives à différentes périodes de la journée ?
- Q8 - Quelles sont les stations les plus utilisées le week-end, en semaine ?
- Q9 - Existe-t-il des stations avec un fort déséquilibre entre départs et arrivées ?
- Q10 - Quels sont les trajets les plus fréquents (station de départ vers station d'arrivée) ?

### 3. Lien avec la météo

- Q11 - Comment la météo (pluie, température) influence-t-elle l'usage des vélos ?
- Q12 - Comment la météo (pluie, température) influence-t-elle les durées de trajets ?
- Q13 - En cas de températures extrêmes, est-ce que le volume de trajets diminue fortement ?
- Q14 - Les utilisateurs membres utilisent-ils plus les vélos par mauvais temps ?

#### 4. Question bonus

- Q15 - Est ce qu'il est possible de voir une différence d'utilisation des vélos lors de la prise du Capitole le 6 janvier 2021 ?

### Contraintes et limites :

- **Données agrégées** : L'absence d'identifiant unique pour chaque vélo limite les possibilités de reconstitution précise des itinéraires ou des distances parcourues.
- **Données contextuelles limitées** : Les événements exceptionnels (grèves, manifestations, travaux, etc.) ou les changements d'infrastructure (nouvelles pistes cyclables, réaménagements urbains) ne sont pas pris en compte dans les données disponibles.
- **Intention de l'utilisateur** : bien qu'il soit riche en information, les jeux de données ne reflète pas nécessairement les raisons des choix des usagers (par ex: s'il n'y a plus de vélos électriques disponibles, l'utilisateur va être contraint de prendre un vélo classique contre son gré)

### Importation du jeu de données et des librairies

## I. Utilisation générale des vélos

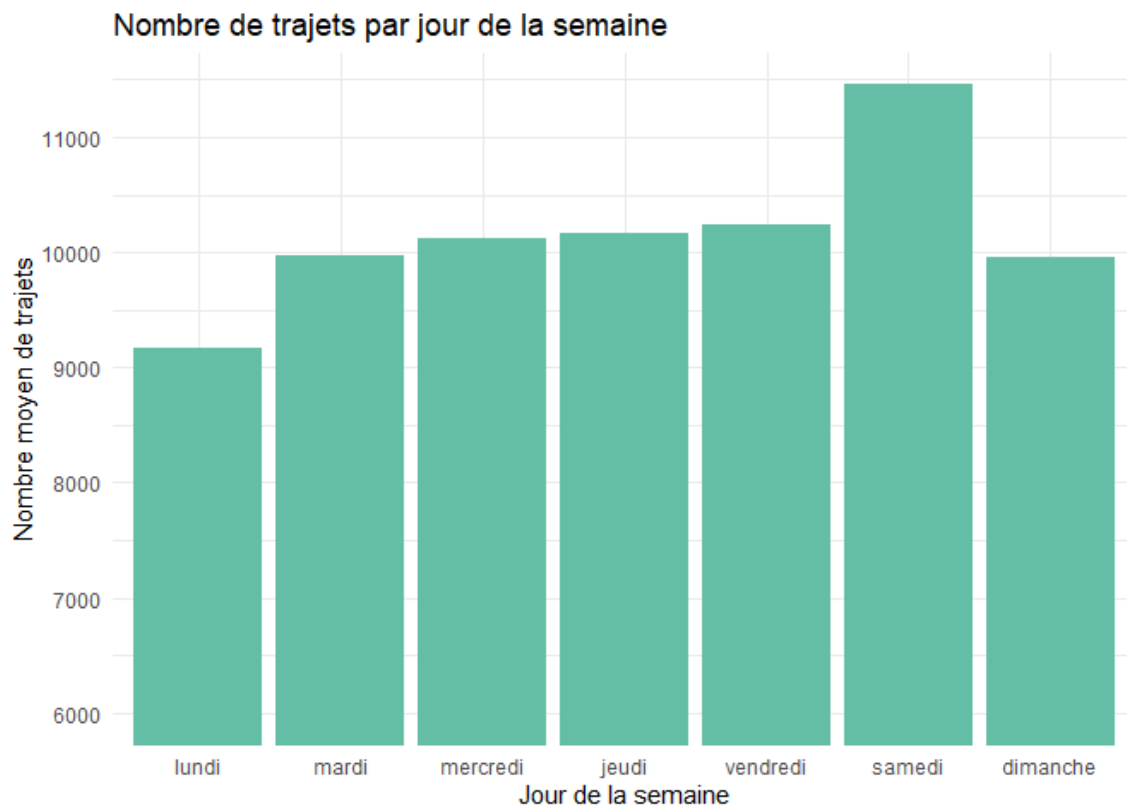
### Question 1 : Combien de trajets sont effectués en moyenne par jour/mois ?

### Question 1a : Y a-t-il une variation du nombre de trajets selon les jours de la semaine ?

#### Hypothèse :

Nous imaginons que les utilisateurs empruntent davantage les vélos en semaine (pour se rendre au travail notamment) que le week-end.

#### Graphique :



### Interprétation :

Contrairement à notre hypothèse initiale, les week-ends enregistrent en réalité un nombre très élevé de trajets, en particulier le samedi qui dépasse tous les autres jours de la semaine. Cela suggère que l'usage du vélo ne se limite pas à une fonction utilitaire (travail), mais qu'il est aussi très utilisé pour les loisirs. Le dimanche reste élevé, presque équivalent aux jours de travail.

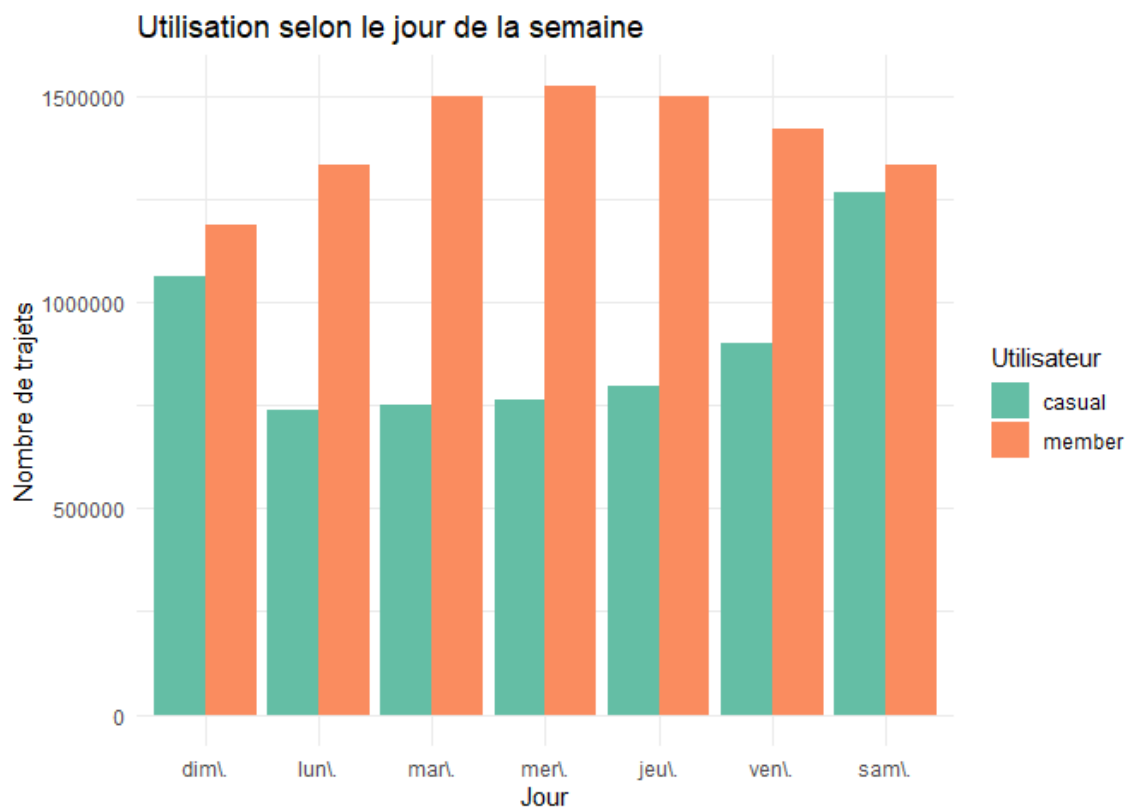
En semaine, le nombre de trajets reste relativement stable autour de 10 000 trajets par jour. Cependant, on observe une baisse notable le lundi, où la moyenne descend à environ 9 000 trajets. Cela peut s'expliquer par le fait que de nombreux commerces, établissements culturels ou services (comme les banques) sont souvent fermés le lundi, ce qui réduit potentiellement le besoin de déplacement.

Cette observation sera affinée par une analyse horaire (Les pics du matin et du soir existent-ils en semaine ?) dans la suite de notre étude. On peut également effectuer une analyse des catégories d'utilisateurs (membres ou occasionnels).

### Hypothèse :

On peut s'attendre à ce que les utilisateurs 'casual' utilisent plus les vélos le weekend.

### Graphique



## Interprétation

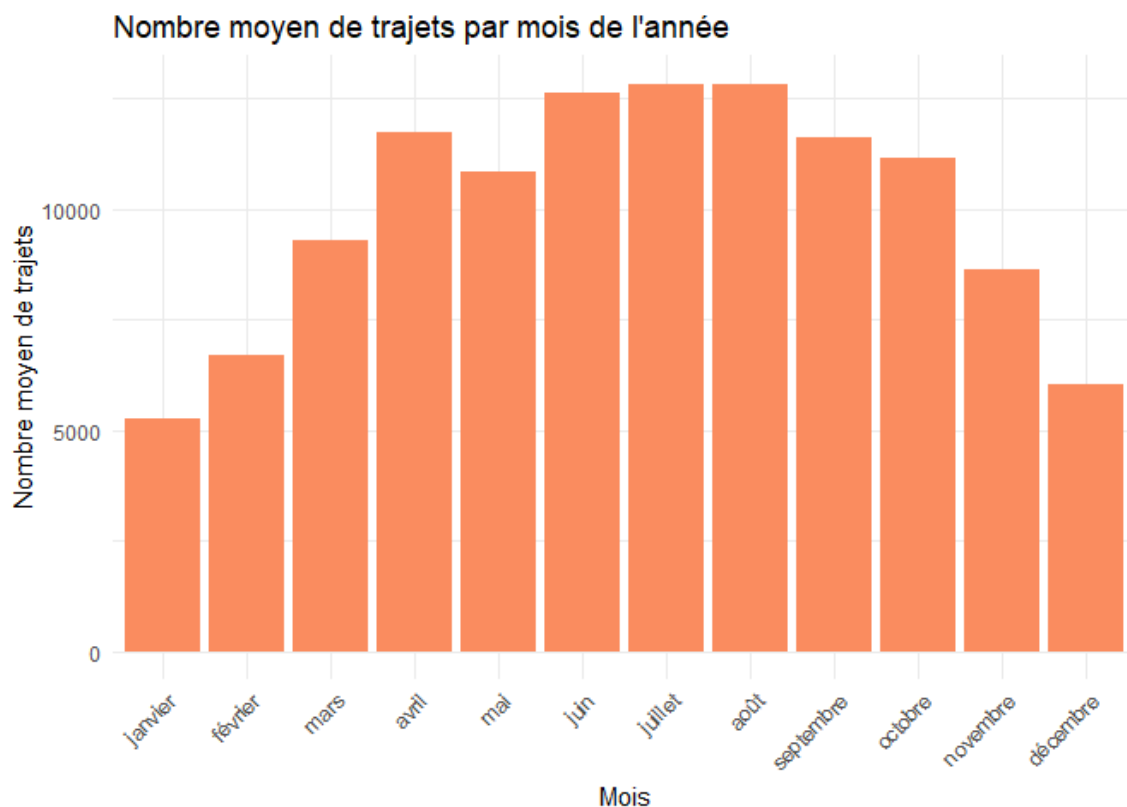
On peut voir sur le graphique que les utilisateurs casuels font plus de trajets le weekend tandis que les utilisateurs membres font plus de trajets la semaine. Ces résultats sont sûrement dûs au fait que les utilisateurs casual utilisent les vélos pour le loisir pour faire des balades le weekend, tandis que les membres les utilisent la semaine pour aller travailler, aller à l'école... Cela confirme donc nos hypothèses initiales

## Question 1b : Y a-t-il une variation du nombre de trajets selon les mois de l'année ?

### Hypothèse :

Nous supposons que la fréquentation des vélos est plus importante en été, en raison du beau temps, et qu'elle diminue pendant les mois froids et pluvieux.

### Graphique :



### Interprétation :

Le nombre moyen de trajets évolue clairement au rythme des saisons.

L'activité augmente fortement entre avril et septembre, avec un pic en juillet-août où l'on atteint près de 13 000 trajets en moyenne. Cette hausse s'explique sans doute par les vacances estivales et des conditions météorologiques plus favorables.

À l'inverse, les trajets diminuent nettement en hiver, notamment entre décembre et février, où le nombre moyen redescend autour de 5 000 trajets en décembre. Cette baisse reflète un usage moindre du vélo durant les mois les plus froids de l'année.

## Question 2 : Quelle est la tendance générale de l'utilisation des vélos au cours de l'année ?

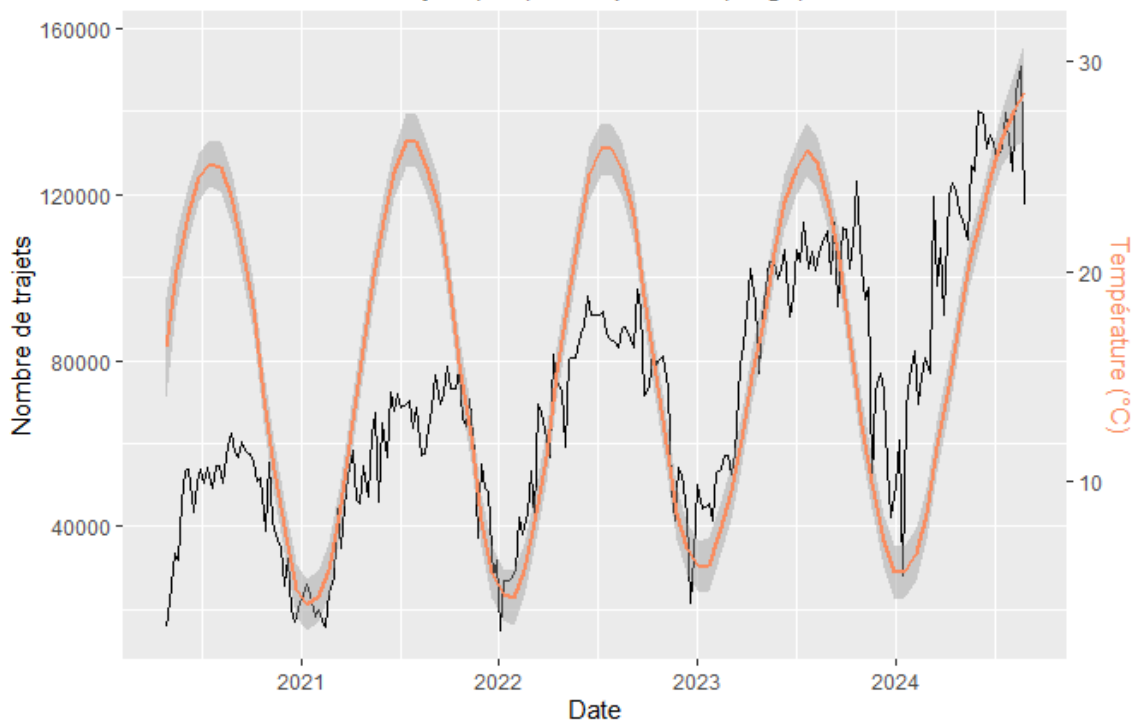
### Hypothèse :

En premier lieu, une observation intéressante à faire est d'observer la tendance générale de l'utilisation des vélos au cours des années étudiées.

### Graphique :

## Nombre de trajets par semaine et température

Trajets (noir) et température (rouge)



### Interprétation :

Nous constatons une augmentation au fil des années de l'utilisation des vélos et surtout une cyclicité. Nous observons une relation très forte entre la température et l'utilisation des vélos. Les utilisateurs privilégient l'utilisation des vélos lors de températures chaudes. Ainsi, nous pouvons voir que les utilisations majoritaires se font pendant le printemps, l'été et l'automne.

Certaines semaines ressortent plus sur le graphique comme la semaine du 3 janvier 2022. Avec une température moyenne de 0°, l'utilisation du vélo est la plus faible du dataset avec 15 083 utilisations. Ceci peut s'expliquer par une chute de neige tombée le 3 janvier où 17,5 cm sont tombés en une journée. On observe des phénomènes similaires lors de la semaine du 19 décembre 2022, en raison de conditions météorologiques défavorables comprenant des grêlons, de la neige et de la pluie. Pendant la semaine du 15 janvier 2024, on observe une chute de température importante, passant de 7°C la semaine précédente à -2,5°C. Cette semaine-là, on enregistre également des grêlons, de la neige et de la pluie, ce qui contribue à une utilisation réduite des vélos.

### Question 3: Quelles sont les durées moyennes de trajet ? Quels éléments influent sur cette durée ? (jour de la semaine, type d'utilisateur...)

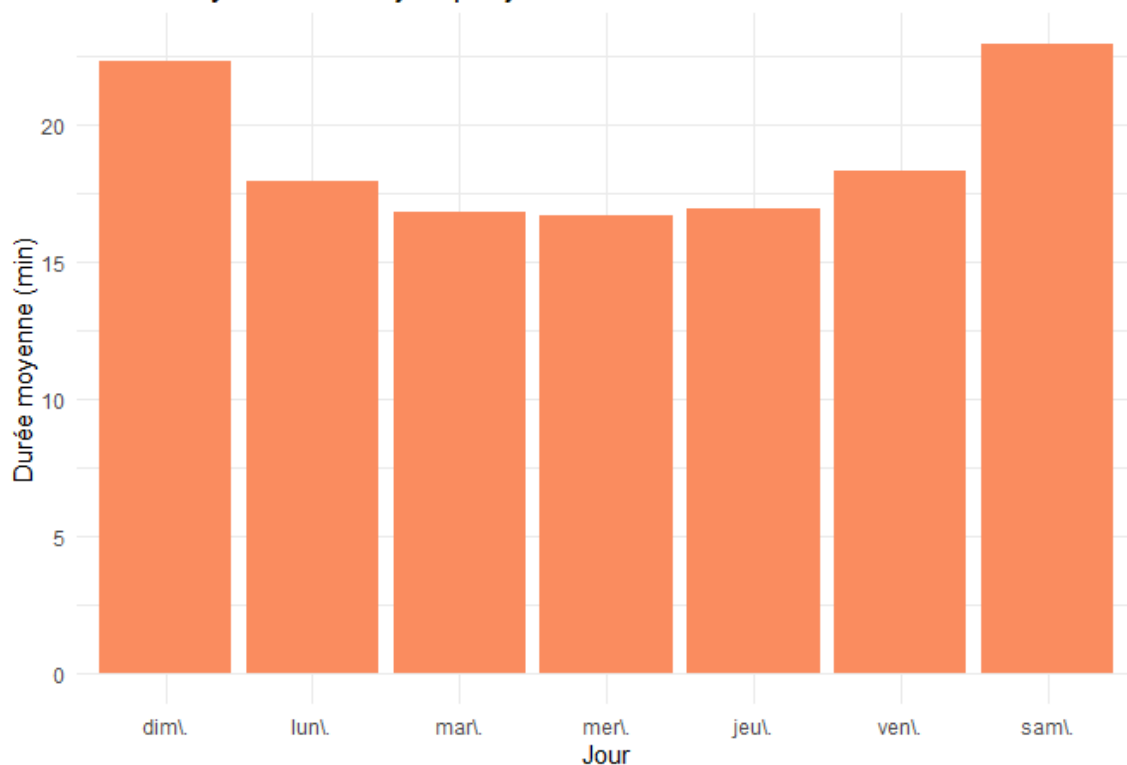
#### Hypothèse :

En étudiant les durées moyennes de trajet, on s'attend à ce que le jour de la semaine, et le type d'utilisateur influent. Nous pensons que les durées des trajets sont plus longues le weekend et pour les utilisateurs 'casual'.

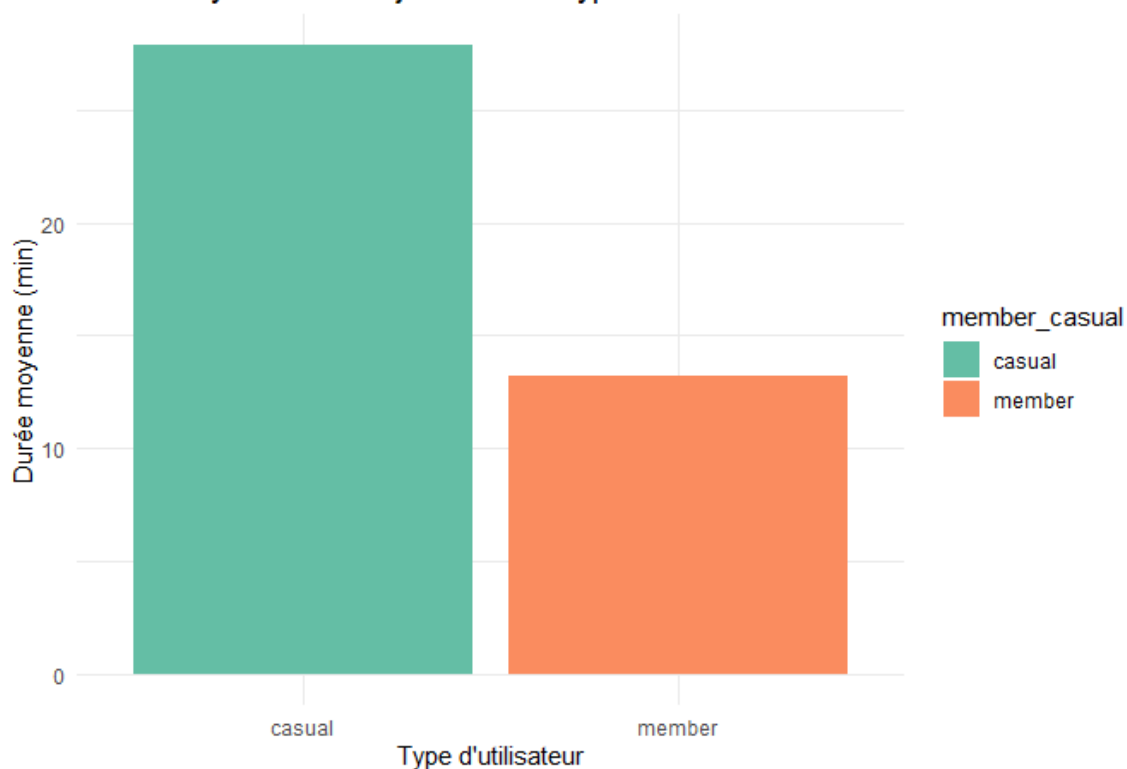
#### Graphique :



### Durée moyenne des trajets par jour de la semaine



### Durée moyenne des trajets selon le type d'utilisateur



### Interprétation :

On peut voir dans un premier temps, que la moyenne de trajet la semaine est d'environ 17 minutes tandis que celle durant le weekend est nettement plus élevée avec 25 minutes. Cela confirme donc nos attentes. Cela s'explique sûrement par l'utilisation différente des vélos : les gens les utilisent le weekend afin de se promener pour le loisir, tandis que durant la semaine, les utilisateurs s'en servent pour des trajet domicile-travail majoritairement.

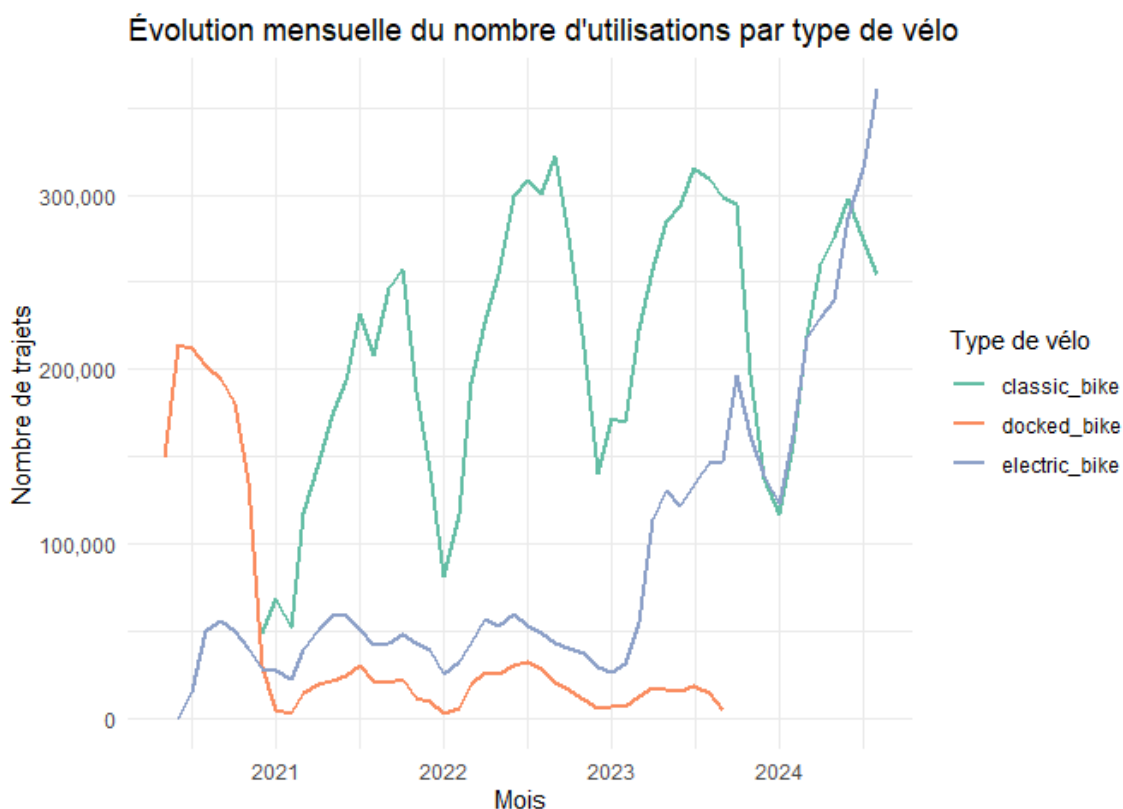
Concernant le type d'utilisateur, nos hypothèses étaient justes : les utilisateurs 'casual' se déplacent plus longtemps avec une moyenne de quasiment 30 min tandis que les utilisateurs 'members' utilisent les vélos pendant environ 12 minutes. Cela s'explique par le fait que les utilisateurs 'casual' font des balades, les utilisent pour visiter par exemple.

#### Question 4: Quelle est l'évolution de l'utilisation des différents types de vélos ?

##### Hypothèse :

On peut supposer une hausse de l'utilisation de vélos électriques. Les utilisateurs souhaitent probablement être le plus confortable possible.

##### Graphique :



##### Interprétation

L'évolution mensuelle du nombre d'utilisations par type de vélos permet de montrer plusieurs tendances intéressantes. On peut tout d'abord voir que les vélos « docked » c'est-à-dire les vélos qui se prennent et rendent à des stations ont été les plus populaires avant 2021 mais ont connu une grosse diminution au profit de vélos « classiques ». Ces vélos peuvent être posés et rendus n'importe où et ne nécessitent pas d'être rattachés à une station. Ensuite entre 2021 et 2024, le vélo classique a été largement le plus utilisé avant d'être dépassé par les vélos « électriques » qui ont connu une hausse significative à partir de 2023. Ces données suggèrent une modernisation de la flotte de Capital Bike Share, et une transition des utilisateurs vers des vélos plus confortables qui demandent moins d'efforts (électriques) ainsi que des vélos qui leur permettent plus de flexibilité en terme de localisation d'emprunt et de rendu (diminution des « docked »)

## **Question 5 : En se focalisant sur les utilisateurs, y a-t-il des effets “heures de pointe” dans la journée ?**

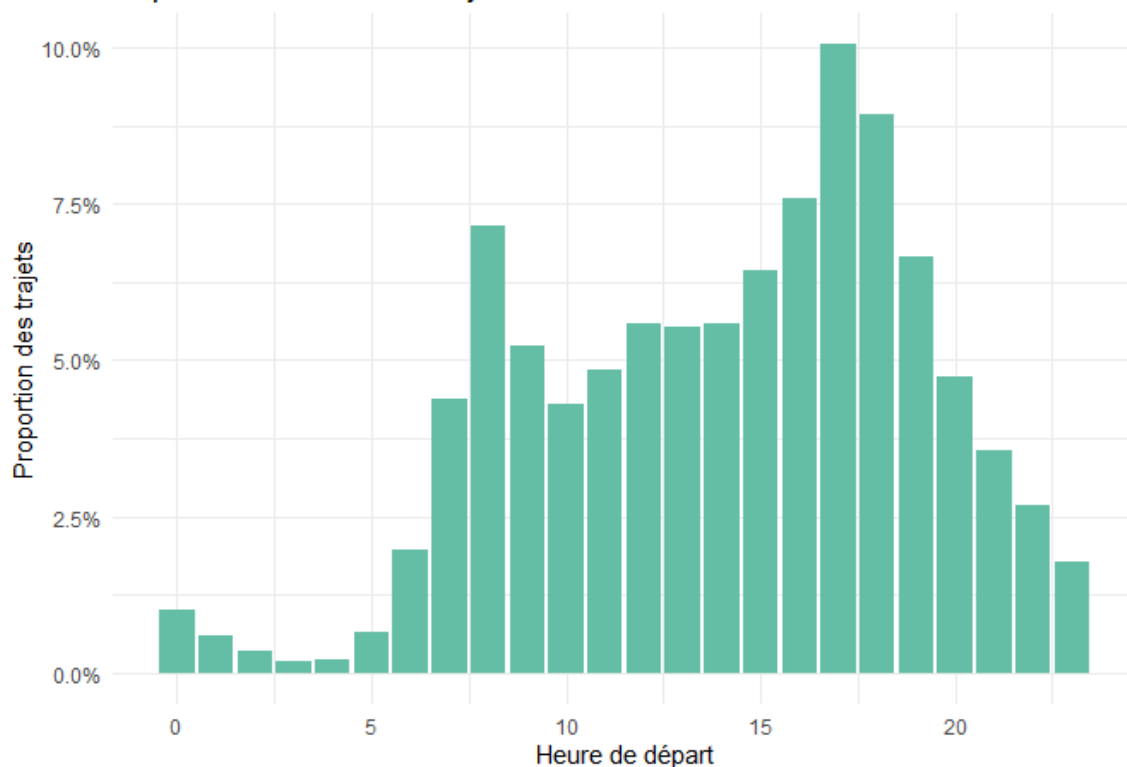
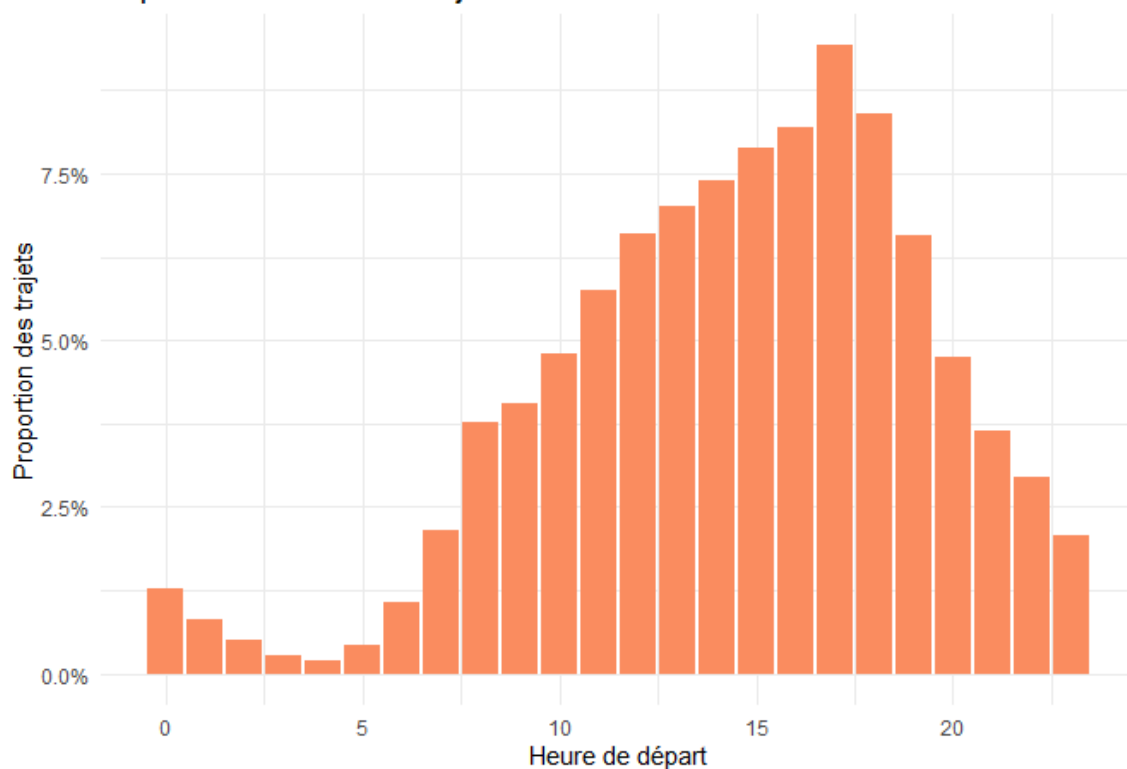
### **Hypothèse :**

Les utilisateurs abonnés adoptent très probablement des comportements liés à leur routine quotidienne (trajets domicile-travail par exemple). Ils seraient donc plus susceptibles d'utiliser les vélos pendant les heures de pointe, notamment en début de matinée (trajet vers le travail) et en fin de journée (retour).

À l'inverse, les utilisateurs occasionnels ont une utilisation plus flexible et dispersée dans la journée, et leur recours aux vélos ne devrait pas particulièrement se concentrer sur les heures de pointe.

On s'attend donc à observer des pics d'utilisation pendant les heures de pointe uniquement chez les abonnés, tandis que les utilisateurs occasionnels auront une répartition horaire plus homogène au cours de la journée.

### **Graphique :**

**Répartition horaire des trajets — Utilisateurs abonnés****Répartition horaire des trajets — Utilisateurs occasionnels****Interprétation :**

Le graphique montre la répartition horaire des trajets effectués selon le type d'utilisateur (abonné ou occasionnel), en pourcentage du total des trajets de chaque groupe.

On constate que :

Les abonnés présentent des pics très marqués à 8 h (7,1% des trajets) et 17-18 h (10% et 8,9%), ce qui correspond clairement aux heures de pointe traditionnelles. On note aussi une activité significative

dès 6 h du matin (2%), ce qui renforce l'idée d'un usage dans le cadre de déplacements domicile-travail.

À l'inverse, les occasionnels ont une distribution plus lissée dans la journée, avec une montée progressive de l'activité à partir de 7 h, avec un pic à 17 h (autour de 9.5%) et un déclin en soirée. Il n'y a pas de véritables heures de pointe, ce qui traduit un usage plus souple et orienté loisir.

Ces résultats confirment que les abonnés sont des usagers réguliers qui utilisent le service dans un cadre professionnel. En revanche, les cyclistes occasionnels exploitent davantage le service en dehors des pics, ce qui reflète une utilisation non contrainte par des horaires fixes.

## II. Utilisation des stations

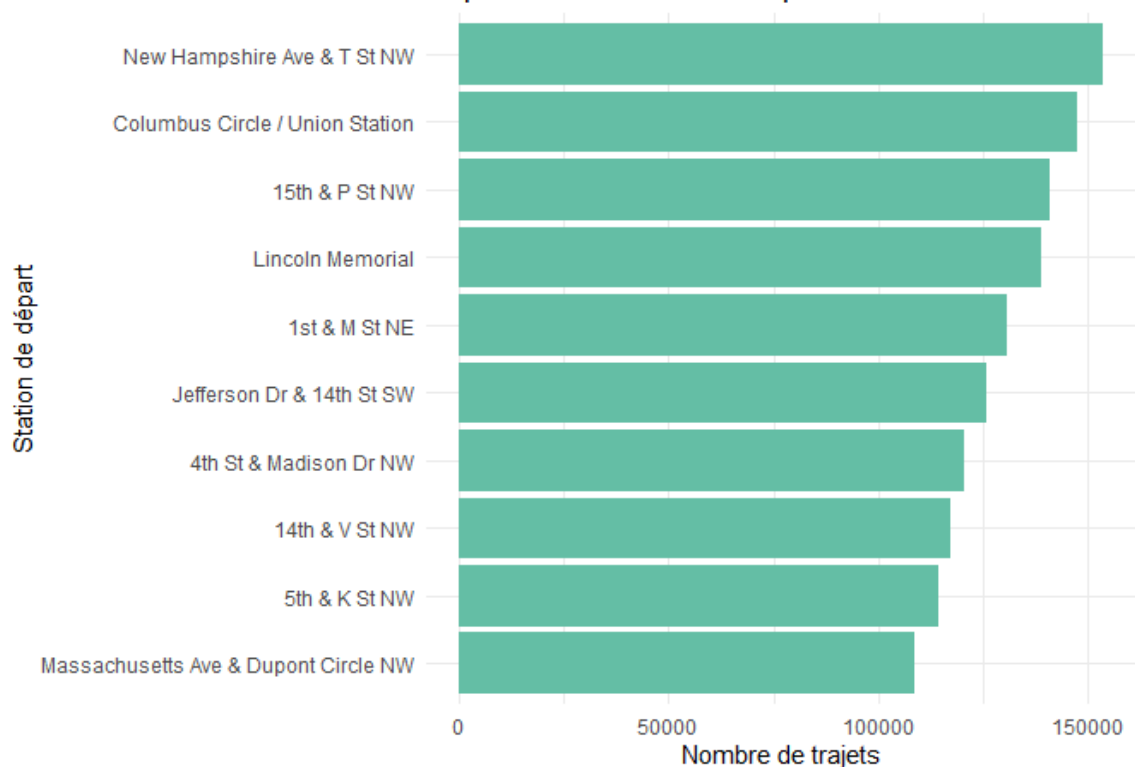
### Question 6 : Quelles stations sont les plus utilisées pour les départs/arrivées ?

#### Hypothèse :

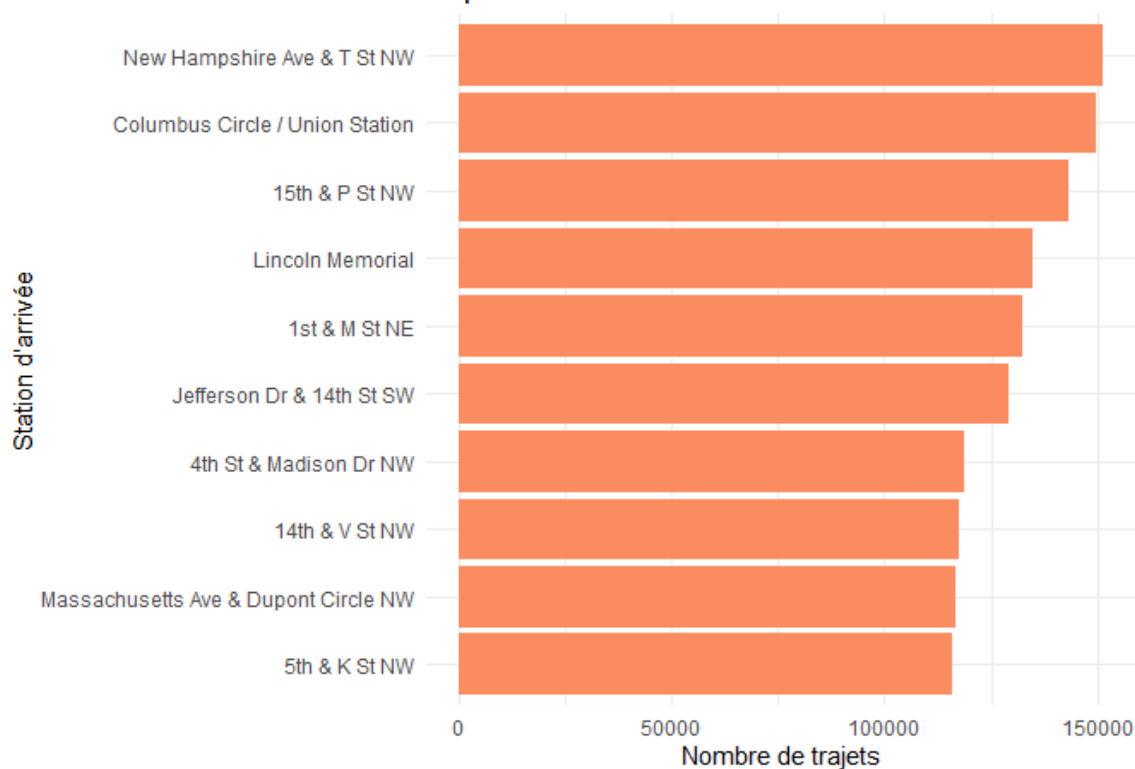
Notre hypothèse initiale est que les stations les plus utilisées sont probablement les mêmes pour les départs et les arrivées. Les stations situées près de secteurs dynamiques (gares, centres commerciaux, quartiers d'affaires) sont sûrement les plus utilisées.

#### Graphique :

### Top 10 des stations de départ

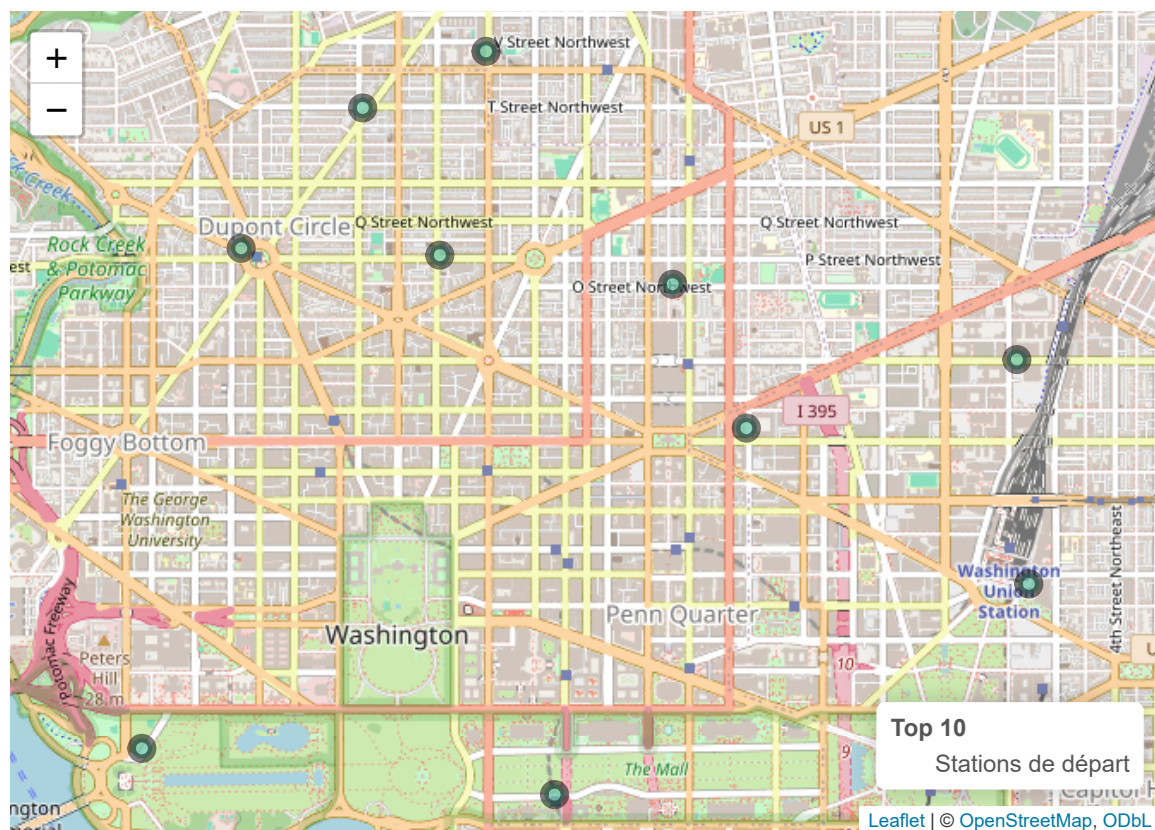


### Top 10 des stations d'arrivée



### Interprétation :

Premièrement, nous pouvons observer que les stations sont quasiment les mêmes pour les départs et les arrivées. Cela signifie qu'il n'y a pas de déséquilibres au niveau des stations les plus utilisées, il y a de nombreuses personnes qui empruntent des vélos à ces endroits mais il y a beaucoup de vélos qui sont redéposés ici aussi. On peut par contre se poser la question de pourquoi ces stations sont les plus utilisées. Pour répondre à cette question nous allons tracer ces stations sur une carte pour régarder les quartiers/les points d'intérêts près de ces stations.

**Graphique :****Interprétation :**

Les stations Lincoln Memorial, Jefferson Dr & 14th St SW et 4th St & Madison Dr NW sont localisés près du National Mall, qui est une zone très touristique de Washington.

Ensuite, la station Columbus Circle/Union Station est devant l'Union Station qui est un grand hub de transport (métro, gare routière, gare...)

Les autres stations sont dans des quartiers résidentiels ou des zones de bureaux très denses.

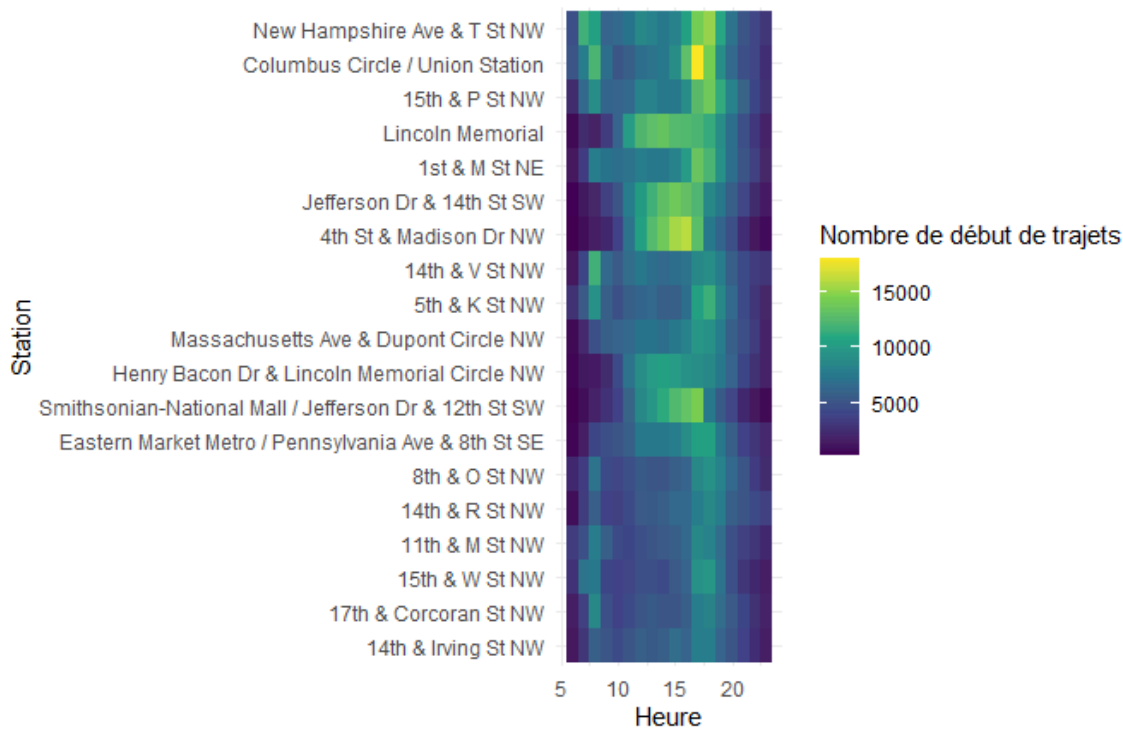
Cette identification des stations centrales soulève une question d'usage : ces volumes élevés correspondent-ils à des profils d'utilisation spécifiques ? L'analyse des durées de trajets par station permettra de comprendre si les stations les plus fréquentées sont utilisées pour des trajets courts (transit) ou longs (récréation).

**Question 7 : Quelles stations sont les plus actives à différentes périodes de la journée ?****Hypothèse :**

On suppose que l'activité des stations varie selon les heures de la journée. Cette question est essentielle pour comprendre les cycles d'utilisation quotidiens et optimiser la redistribution des vélos aux moments stratégiques.

**Graphique :**

### Stations de départ les plus actives dans la journée



### Interprétation :

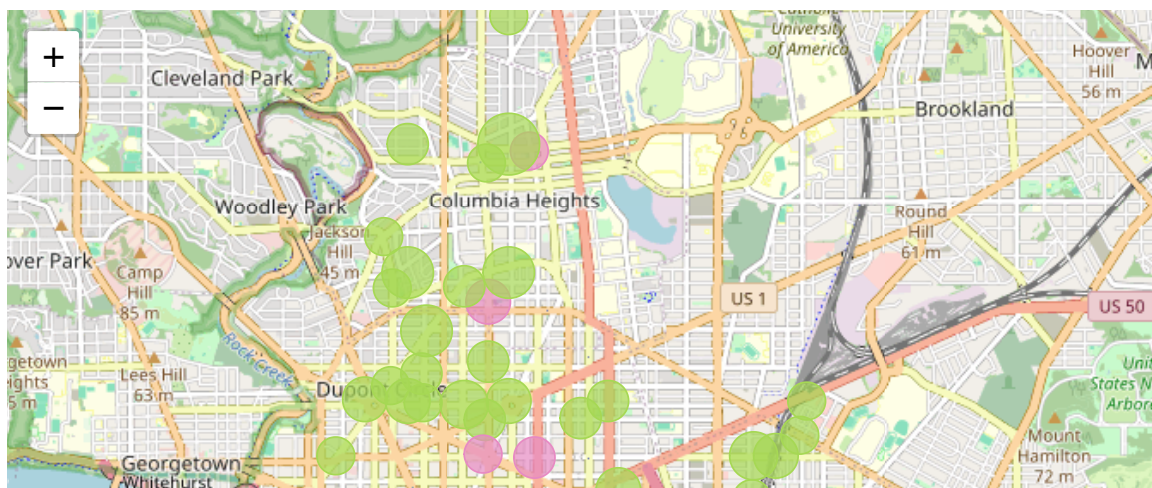
L'analyse des heatmaps représentant l'activité des stations de départ et d'arrivée en fonction de l'heure de la journée montre des modèles d'utilisation distincts.

Il y a deux profils d'utilisation différents :

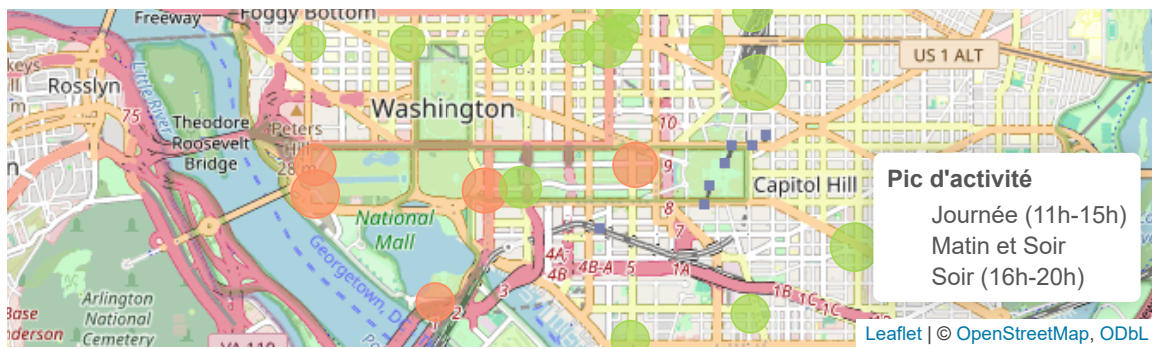
“New Hampshire Ave & T St NW” et “Columbus Circle/Union Station” présentent des pics d'activité marqués pendant les heures de pointe (7-9h pour les départs, 17-19h pour les arrivées). Cela témoigne d'une utilisation principalement liée aux déplacements domicile-travail, où les résidents utilisent ces stations comme points de départ le matin vers le centre-ville et points d'arrivée lors du retour en soirée.

À l'opposé, “Lincoln Memorial” et “Jefferson Dr & 14th St SW” montrent une activité plus étalée sur la journée avec des pics en milieu de journée (11h-15h). Ces stations sont utilisées pour le loisir et le tourisme.

### Graphique :







La carte des départs met en évidence les différences entre les quartiers :

- Stations vertes (pic matin 6h-10h) : localisées dans les zones résidentielles périphériques, elles confirment leur rôle de “gares de départ” pour les trajets pendulaires
- Stations roses (matin et soir) : situées dans des zones mixtes résidentiel/commercial, elles témoignent d’une double fonction dans les flux urbains
- Stations oranges (pic journée) : concentrées autour des sites touristiques et institutionnels

La carte des arrivées révèle les zones d’attraction urbaine :

- Stations vertes (pic soir 16h-20h) : marquent les destinations de retour vers les quartiers résidentiels
- Stations oranges (pic journée) : confirment les pôles touristiques et de loisir comme destinations privilégiées en milieu de journée

Cette double perspective géographique confirme l’hypothèse d’un système organisé autour de flux pendulaires structurés, avec des nuances selon la fonction urbaine des quartiers. La complémentarité spatiale entre les patterns de départ et d’arrivée illustre parfaitement les mouvements directionnels qui caractérisent la mobilité urbaine de Washington D.C.

Après avoir identifié ces variations horaires d’activité et leur distribution géographique, il devient pertinent d’analyser comment ces patterns se différencient entre la semaine et le week-end. Cette analyse temporelle complémentaire nous permettra de distinguer plus finement les stations principalement utilisées pour les trajets domicile-travail de celles fréquentées pour des activités de loisirs, et de valider les hypothèses émises sur les profils fonctionnels des différentes zones de la capitale.

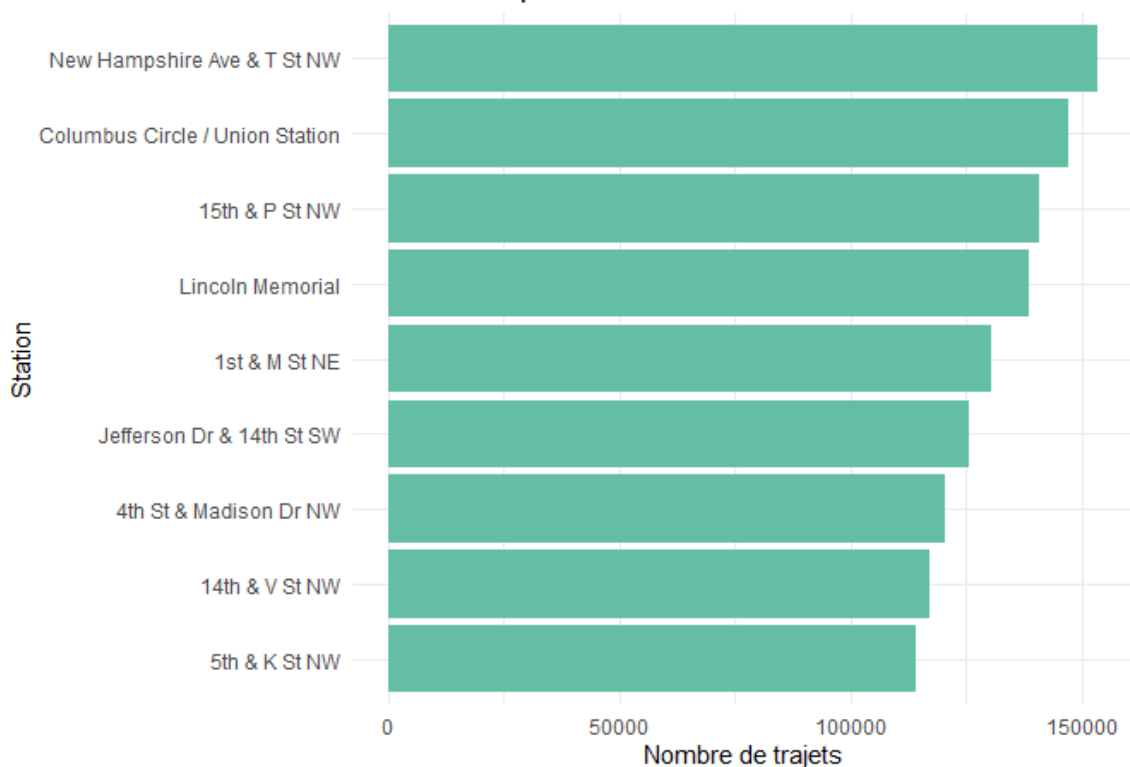
## Question 8 : Quelles sont les stations les plus utilisées le week-end, en semaine ?

### Hypothèse :

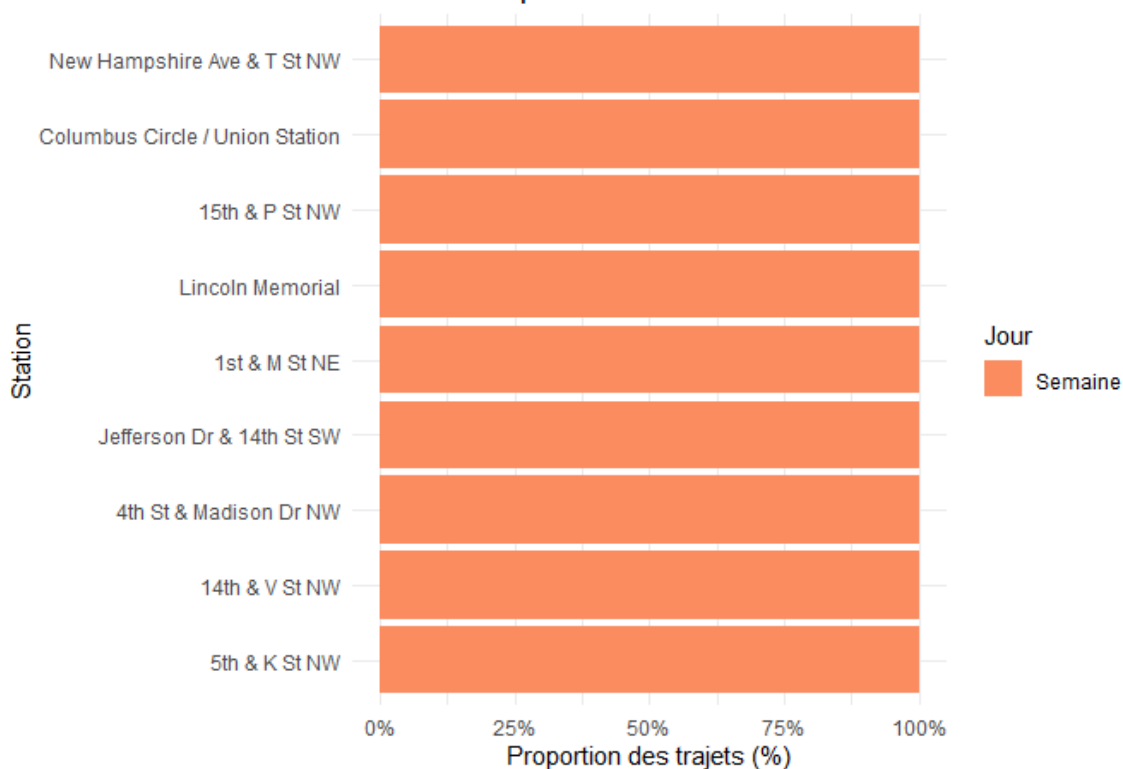
Nous avons voulu examiner comment l’utilisation des stations varie entre la semaine et le week-end. Cette question est pertinente car elle permet de distinguer les stations principalement utilisées pour les trajets domicile-travail (en semaine) de celles fréquentées pour des activités de loisirs (week-end). Nous nous attendons à ce que les stations “touristiques” soit plus utilisées le weekend.

### Graphique :

## Stations les plus utilisées : week-end vs semaine



## Stations les plus utilisées en % : week-end vs semaine



## Interprétation

L'analyse comparative semaine/week-end confirme les 3 types de stations trouvées dans l'analyse horaire.

- Stations domicile-travail : Columbus Circle/Union Station (80% semaine) et New Hampshire Ave & T St NW (75% semaine) sont utilisées pour aller au travail avec une chute drastique d'activité le week-end.

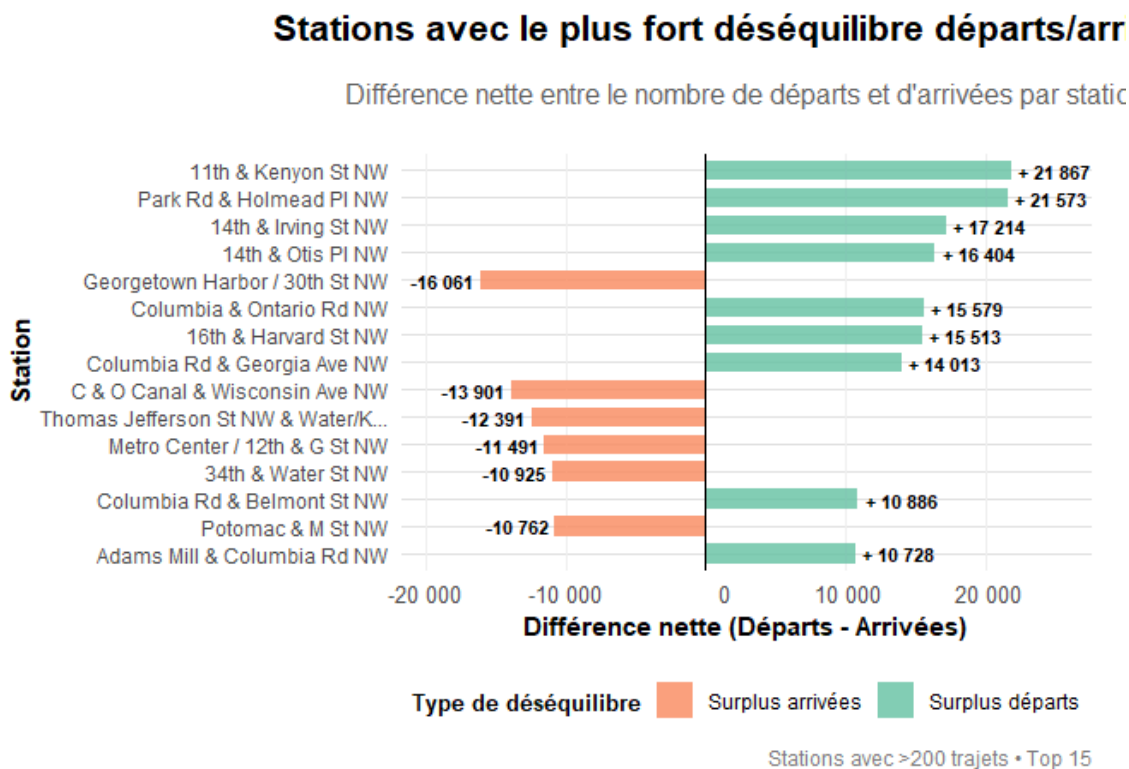
- Stations touristiques : Lincoln Memorial et Jefferson Dr & 14th St SW se distinguent par une utilisation week-end exceptionnellement élevée (35-40%), ce qui confirme qu'elles sont utilisées pour les tourisms et les loisirs.
- Stations mixtes : Les autres stations ont un usage multiple (20-25% week-end) ce qui montre qu'elles sont positionnées dans des zones résidentielles et d'attraction.

## Question 9 : Existe-t-il des stations avec un fort déséquilibre entre départs et arrivées ?

### Hypothèse :

Nous avons cherché à identifier les stations déséquilibrées entre le nombre de trajets qui y débutent et ceux qui s'y terminent. Nous supposons que les stations éloignées du centre ont beaucoup de vélos qui partent et qui ne reviennent pas.

### Graphique :



### Interprétation

On peut voir 2 types de stations différentes ;

- Stations résidentielles : Il y a plus de départs que d'arrivées. Par exemple : 11th & Kenyon St NW (+21 667) et Park Rd & Holmead Pl NW (+21 573). On peut voir que beaucoup de vélos partent le matin en direction du centre ville, et peu sont ramenés, ce qui peut créer un déficit important de vélos dans les banlieues. Ces stations sont celles qui connaissent les pics matin/soirs.
- Stations touristiques : À l'opposé, Georgetown Harbor / 30th St NW (-16 061) et C & O Canal & Wisconsin Ave NW (-13 901) est un exemple de station où il y a énormément d'arrivées mais peu de départs. C'est le cas de toutes les stations considérées comme "touristiques".

## Question 10 : Quels sont les trajets les plus fréquents (station de départ vers station d'arrivée)?

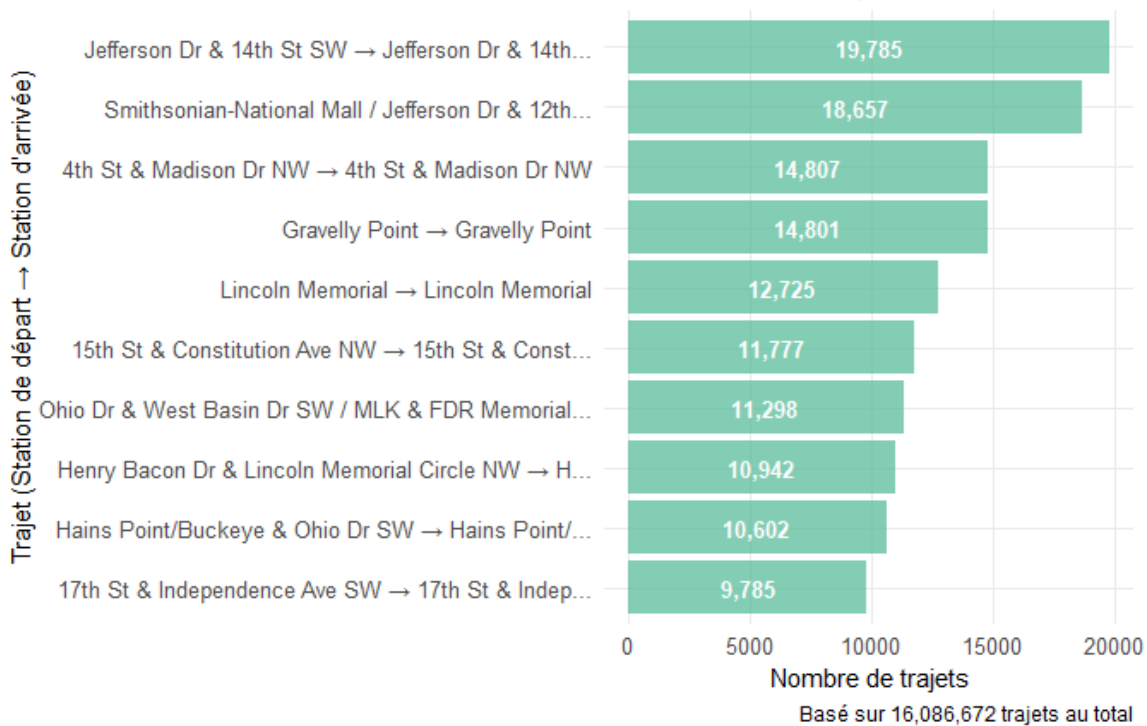
### Hypothèse :

Nous supposons qu'on pourra observer des flux pendulaires (domicile - travail).

### Graphique :

#### Top 10 des trajets les plus fréquents

Paires de stations départ → arrivée



### Interprétation

Nous pouvons voir que les 10 trajets les plus populaires sont tous des trajets en boucle (même station de départ et d'arrivée). Cela montre que les vélos sont beaucoup utilisés pour le loisir et le tourisme. Les stations concernées sont les stations près des sites historiques/touristiques de Washington DC. Jefferson Dr & 14th St SW domine largement avec 19 785 trajets, représentant 0,12% du trafic total, suivi de très près par Smithsonian-National Mall avec 18 657 trajets.

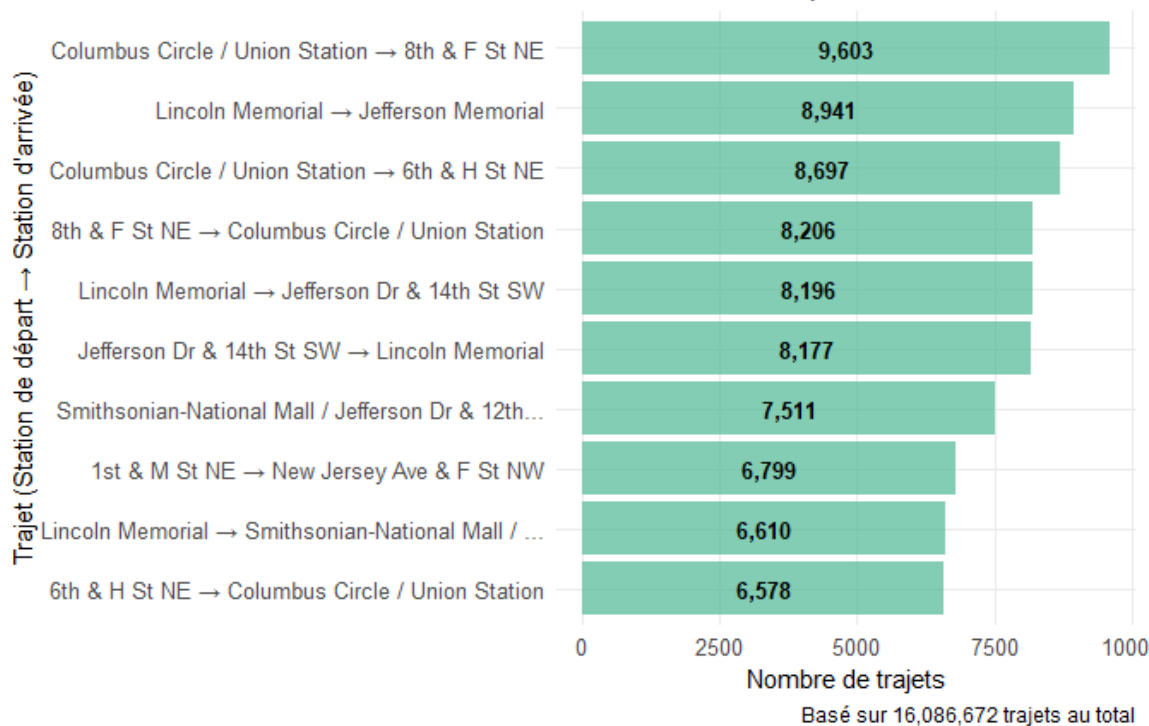
Le tarif des vélos encourage des trajets jusqu'à 45 minutes, ce qui pousse les utilisateurs à les emprunter pour visiter et pas seulement pour des trajets de déplacement.

On peut néanmoins s'interroger sur les trajets avec des départs et arrivées différentes.

### Graphique :

## Top 10 des trajets les plus fréquents

Paires de stations départ → arrivée



## Interprétation

La visualisation des trajets point à point nous permet de voir des schémas différents.

Columbus Circle / Union Station apparaît 4 fois dans les 10 trajets les plus fréquents. Cela peut s'expliquer car Union Station est la principale gare ferroviaire de Washington DC.

On observe qu'il existe un triangle touristique entre le Lincoln Memorial, Jefferson Memorial et Jefferson Dr & 14th St SW. Les visiteurs utilisent ces stations pour faire des trajets de 15-20 minutes entre elles pour visiter ces monuments historiques.

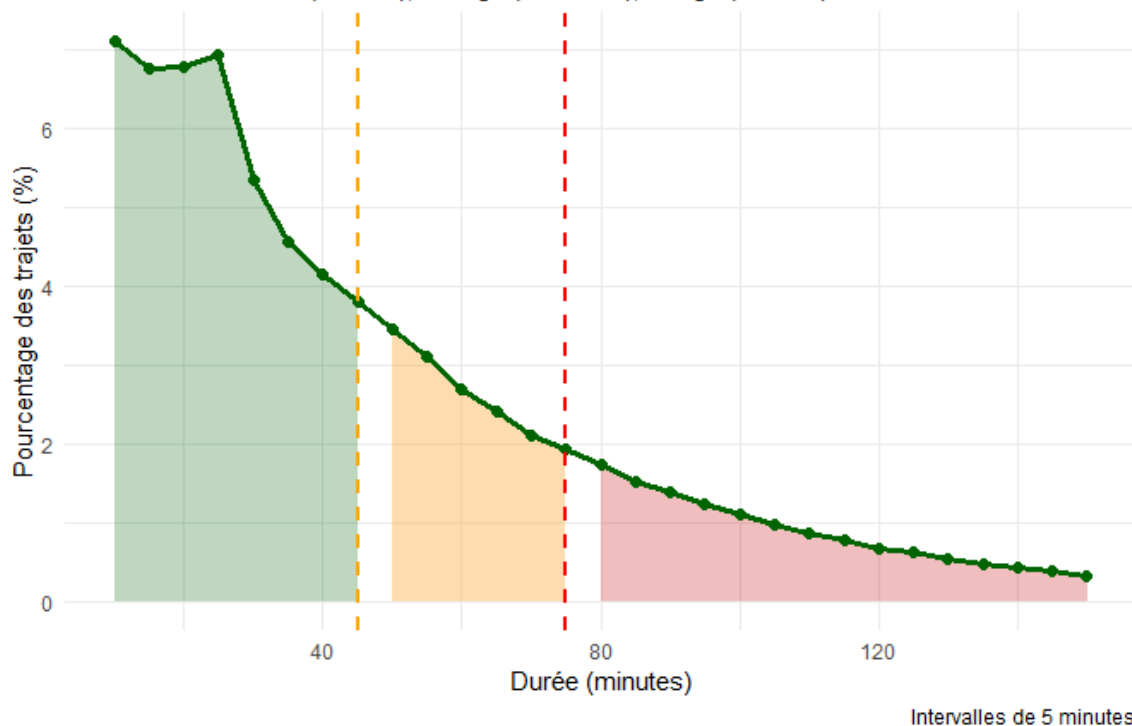
Maintenant que nous avons identifié la nature particulière des trajets en boucle, analysons plus précisément leurs caractéristiques temporelles pour comprendre les comportements d'usage qui les motivent

Nous allons essayer d'analyser les temps de trajets relatif au point de départ et d'arrivée

## Graphique :

### Distribution des temps de trajet (en pourcentages)

Zones colorées : Vert ( $\leq 45$ min), Orange (45-75min), Rouge ( $>75$ min)



### Interprétation

La distribution des temps de trajet pour les trajets en boucle (même station de départ et d'arrivée) montre un usage influencé par les tarifs de la plateforme de vélo Capital Bikeshare (Source : <https://capitalbikeshare.com/pricing>). On observe une concentration majeure des trajets dans la zone verte ( $\leq 45$  minutes), correspondant à la période gratuite pour les abonnés journaliers et annuels, ce qui suggère que de nombreux utilisateurs planifient délibérément leurs sorties récréatives, sportives ou touristiques pour rester dans cette fenêtre sans frais supplémentaires.

La diminution progressive dans la zone orange (45-75 minutes) s'explique par l'introduction des frais de 0,05\$/minute pour les vélos classiques, incitant les utilisateurs à raccourcir leurs balades ou à accepter un coût modéré pour des activités de loisir plus longues. Cette baisse devient encore plus marquée dans la zone rouge ( $>75$  minutes), où les coûts cumulés deviennent dissuasifs pour la plupart des utilisateurs occasionnels.

## III. Lien avec la météo

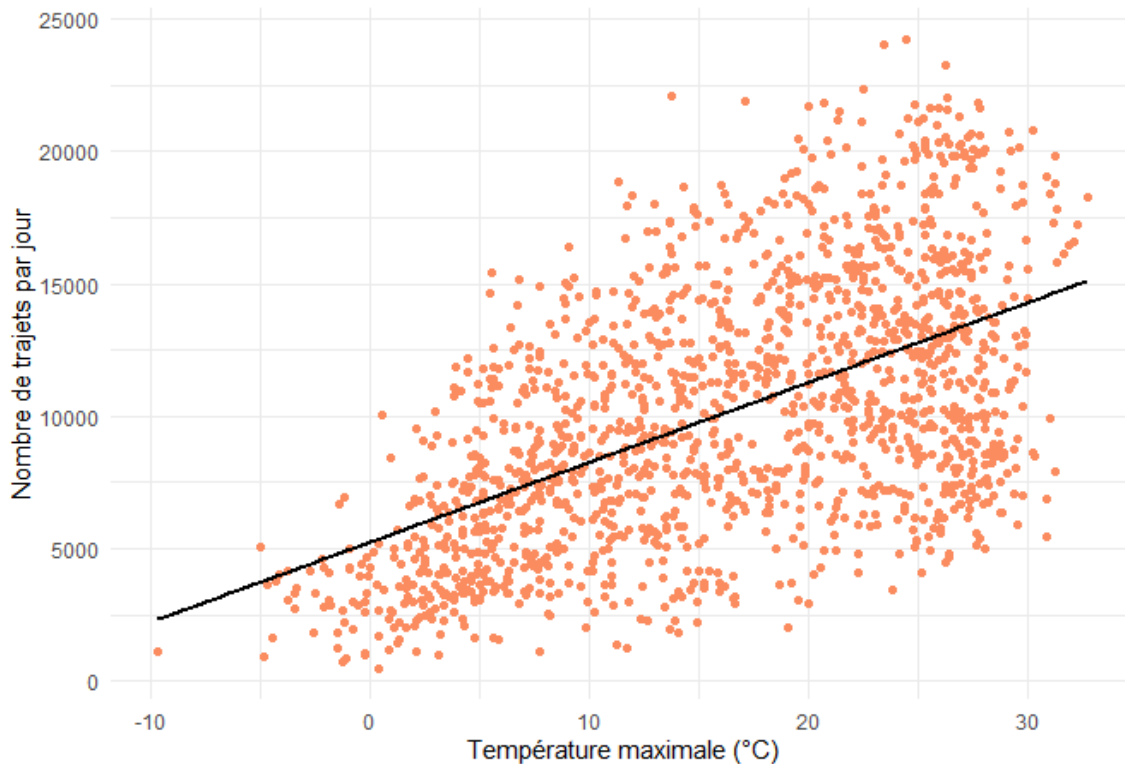
### Question 11: Comment la météo (pluie, température) influence-t-elle l'usage des vélos ?

#### Hypothèse :

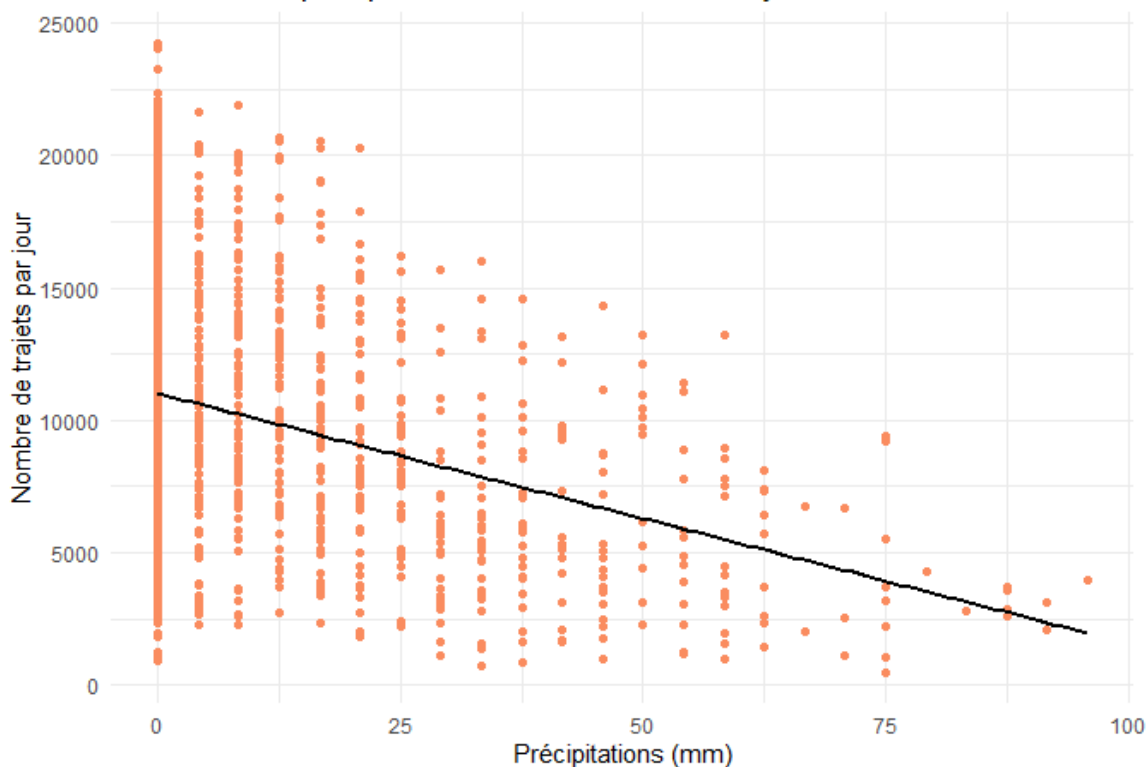
Comme on l'a vu précédemment en question 2, les vélos sont moins utilisés en hiver, sûrement parce qu'il fait plus froid. On souhaite maintenant voir de manière plus précise la relation entre la météo (température, pluie) sur l'influence de l'usage des vélos partagés.

#### Graphique :

### Influence de la température sur le nombre de trajets



### Influence des précipitations sur le nombre de trajets



### Interprétation :

La première visualisation est un nuage de point représentant le nombre de trajets par jour en fonction de la température maximale (en degré celsius) enregistrée dans la journée. La droite de régression en rouge permet d'illustrer la tendance générale. On observe une nette relation positive entre l'utilisation des vélos et la température. Cela signifie que lorsque la température augmente, l'utilisation des vélos augmente. Cependant, les points restent assez dispersés montrant que d'autres facteurs influencent sûrement également l'utilisation des vélos partagés.

La deuxième visualisation est elle aussi un nuage de points représentant le nombre de trajets par jour en fonction de la précipitation (en mm). On a aussi ajouté une courbe de régression en bleu montrant la tendance générale. A la différence de la première visualisation, la relation entre l'utilisation des vélos et les précipitations est négative. Le résultat est plutôt cohérent, personne n'aime faire du vélo sous la pluie. Comme pour la première visualisation, les points sont dispersés autour de la droite de régression montrant potentiellement l'importance d'autres facteurs dans l'utilisation des vélos.

## **Question 12 : Comment la météo (pluie, température) influence-t-elle les durées de trajets ?**

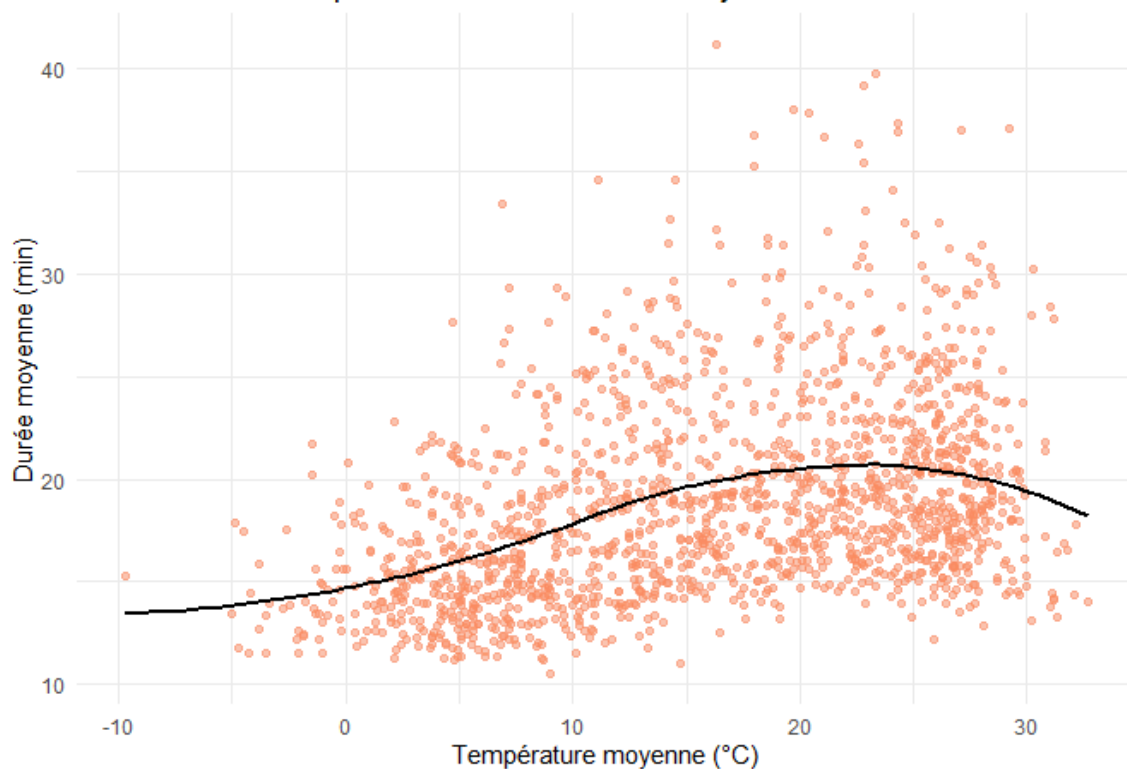
### **Hypothèse :**

Nous nous intéressons maintenant aux mêmes facteurs météorologiques mais cette fois sur leur impact sur les durées moyennes de trajets. Nous souhaitons voir si le mauvais temps défavorise l'utilisation des vélos. Nous supposons donc que la durée moyenne de trajet devrait être plus courte les jours de mauvais temps.

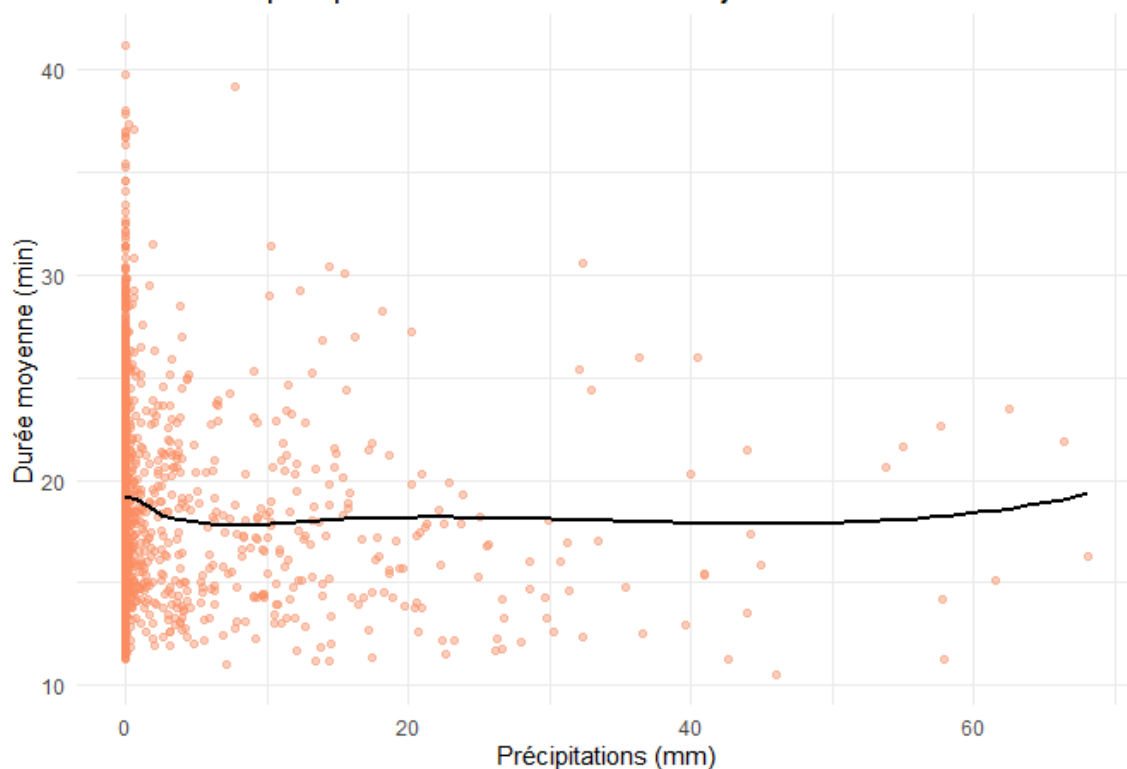
### **Graphique :**



### Influence de la température sur la durée des trajets



### Influence des précipitations sur la durée des trajets



### Interprétation :

La première visualisation met en relation la durée moyenne du trajet avec la température moyenne de la journée. On observe un pic de durée vers 22-23 degrés Celsius avec une durée moyenne comprise entre 15 et 30 minutes environ. On observe une diminution importante quand la température est inférieure à 15 degrés ou supérieure à 30 degrés. On observe donc que la chaleur n'influe pas beaucoup sur le nombre de trajets mais influe quand même sur la durée des trajets.

La deuxième visualisation met en relation les précipitations et la durée moyenne des trajets afin d'observer si les précipitations ont un impact sur les durées de trajets. La grande majorité des observations se trouvent à 0 de précipitations, où la durée moyenne des trajets est très variable. Ceci peut s'expliquer parce que les utilisateurs ne souhaitent juste pas faire du vélos quand il pleut. Il n'y a pas du tout d'influence au niveau de la quantité de pluie.

### **Question 13 : En cas de températures extrêmes, est-ce que le volume de trajets diminue fortement ?**

#### **Hypothèse :**

On cherche maintenant à s'intéresser à des cas extrêmes de température pour essayer d'observer un potentiel impact. Ainsi, nous avons cherché 4 évènements inclus dans les dates de nos données où les températures ont été particulièrement basses et hautes pendant une longue durée. En premier temps nous verrons l'impact de 2 évènements à température froide extrême, et ensuite 2 évènements à température chaude extrême.

Pour les évènements froid, nous en avons choisi un en février 2021, et l'autre en décembre 2022: En février 2021, Dû à un anticyclone arctique qui a soudainement déplacé de l'air glacial vers le sud. Une vague de froid a débuté au Canada et s'est propagé rapidement sur le territoire américain. À Washington DC, on observe une durée de 6 jours à Washington DC où les températures étaient nettement plus basse que les journées précédentes. Source :

<https://journals.ametsoc.org/view/journals/bams/103/12/BAMS-D-21-0266.1.xml>

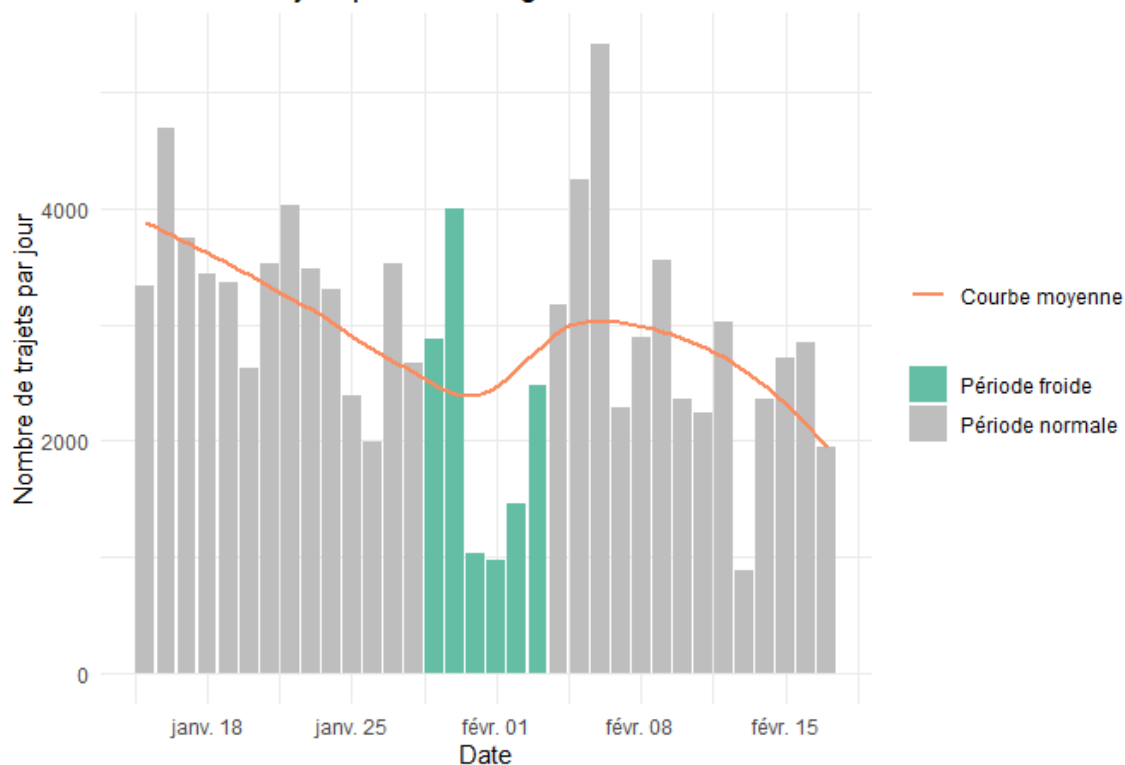
En décembre 2022, autour de Noël, un vent froid a balayé Washington DC entraînant des températures extrêmes en dessous de 0. Source :

[https://www.weather.gov/mrx/December\\_2022\\_Record\\_Cold](https://www.weather.gov/mrx/December_2022_Record_Cold)

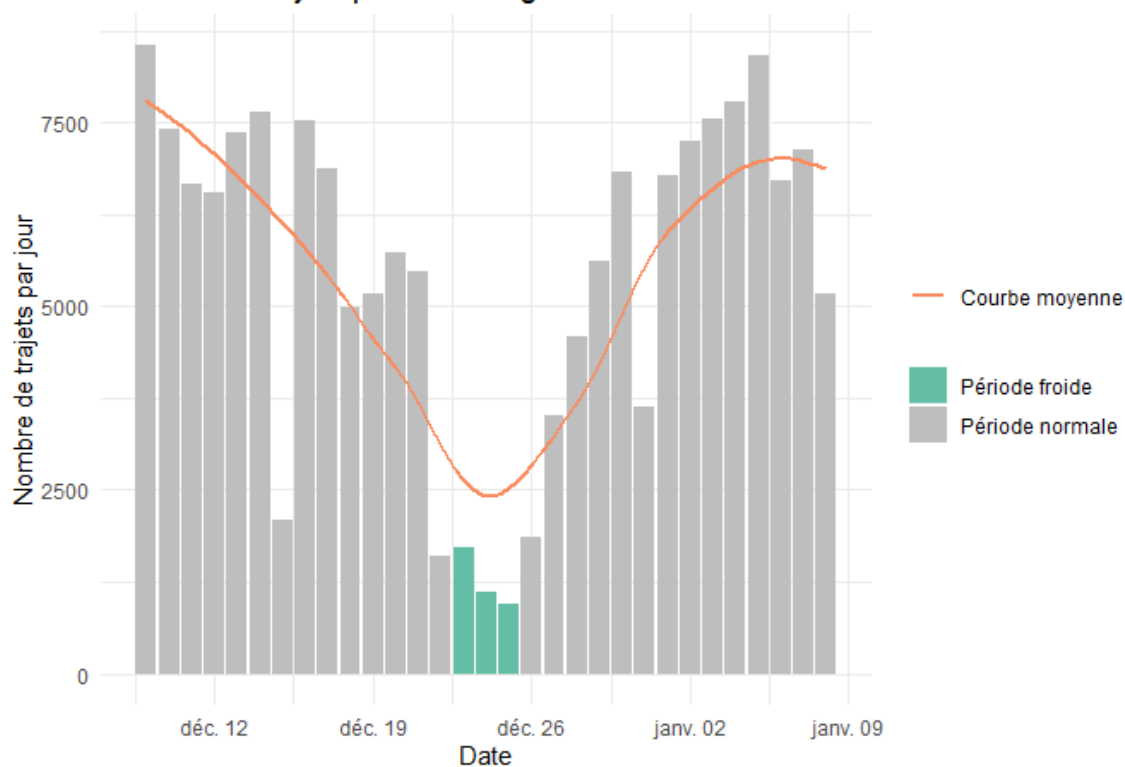
Comme on l'a observé dans la question précédente, la température a un effet sur l'utilisation des usagers, ainsi nous pensons que l'utilisation sera réduite lors de ces pics froid et chaud.

#### **Graphiques :**

### Baisse des trajets pendant les grands froids février 2021



### Baisse des trajets pendant les grands froids décembre 2022



### Interprétation :

Les deux visualisations sont faites sur 14 jours précédant et 14 jours suivant les périodes de températures extrêmes. Grâce à un barplot, on peut observer le nombre de trajet cumulé chaque jour ainsi que sa courbe de régression en rouge. La période de l'évènement est mis en valeur en colorant en bleu les barres la concernant. La première visualisation sur la période février 2021, on peut voir que les deux premiers jours n'ont pas tellement d'impact sur l'utilisation des vélos, mais à partir du 3ème. On voit une réelle chute pendant 3 jours, qui ensuite remonte le dernier jour. Pour le second graphique sur la période en décembre 2022, on voit une grosse chute d'utilisation. Cependant on

remarque aussi que les jours juste avant et après l'évènement suivent la même tendance. Ceci peut s'expliquer parce que les températures du 22 décembre (veille de l'évènement choisi) ont déjà atteint 5°C.

### **Hypothèse :**

Nous cherchons maintenant à voir l'impact que ces températures extrêmes chaudes auraient pu avoir sur l'utilisation des vélos.

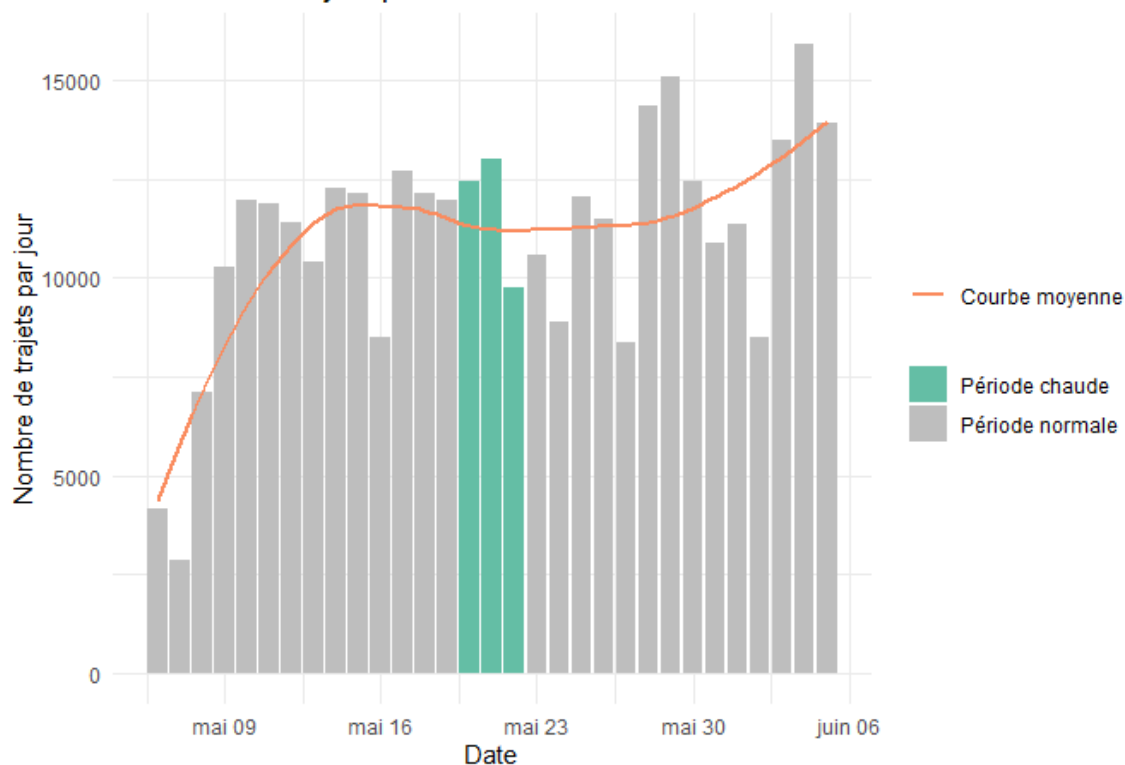
Pour les évènements chaud, nous en avons choisi un en mai 2022, et l'autre en juillet 2024: En mai 2022, Washington D.C. a connu une vague de chaleur marquante, enregistrant des températures exceptionnellement élevées pour cette période de l'année. Selon les prévisions de la National Weather Service, les températures ont commencé à grimper rapidement, atteignant des niveaux rarement observés si tôt dans la saison. Source : <https://www.washingtonpost.com/weather/2022/05/19/dc-may-heat-wave-record/>

En juillet 2024, Washington D.C. a subi une vague de chaleur particulièrement intense qui a marqué l'un des mois de juillet les plus chauds jamais enregistrés. Lors de cette vague de chaleur, la température a atteint 39 °C le 15 juillet, établissant un nouveau record pour cette date. Source : <https://www.washingtonpost.com/weather/2024/07/16/dc-heat-100-record-high-temperatures/>

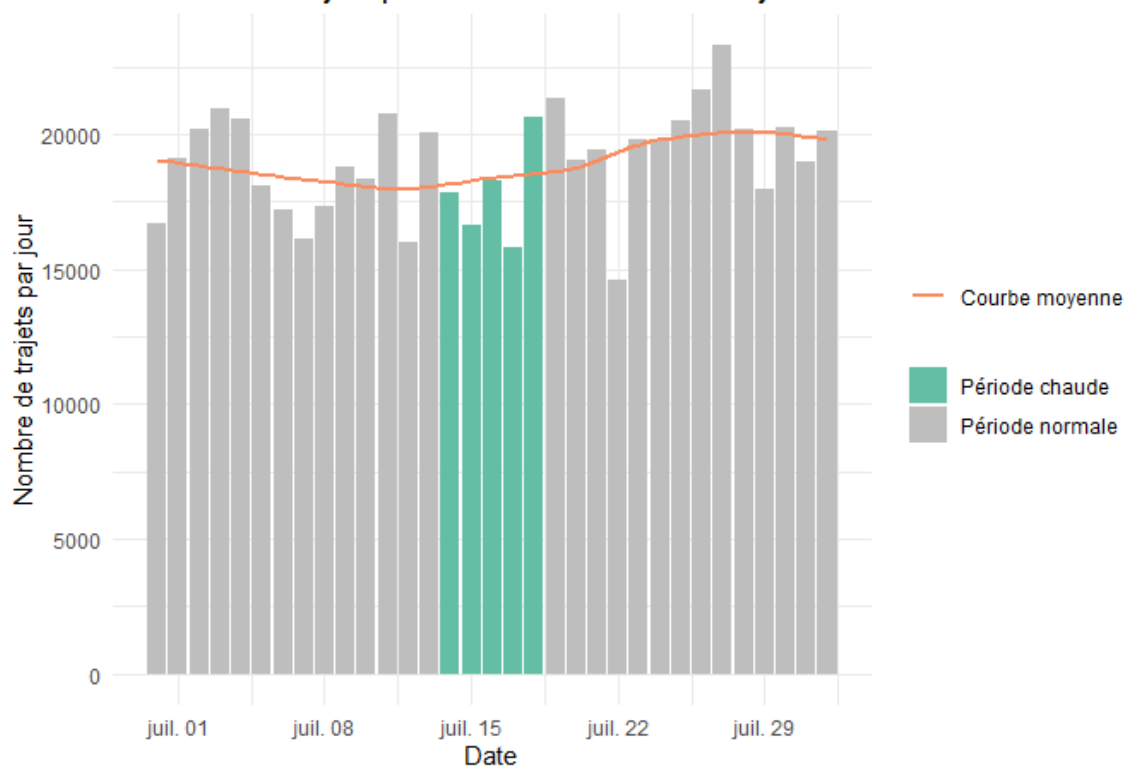
On peut supposer que le même effet de baisse d'utilisation des vélos aura lieu durant les pics de chaleur.

### **Graphiques :**

### Stabilité des trajets pendant les fortes chaleurs mai 2022



### Stabilité des trajets pendant les fortes chaleurs juillet 2024



#### Interprétation :

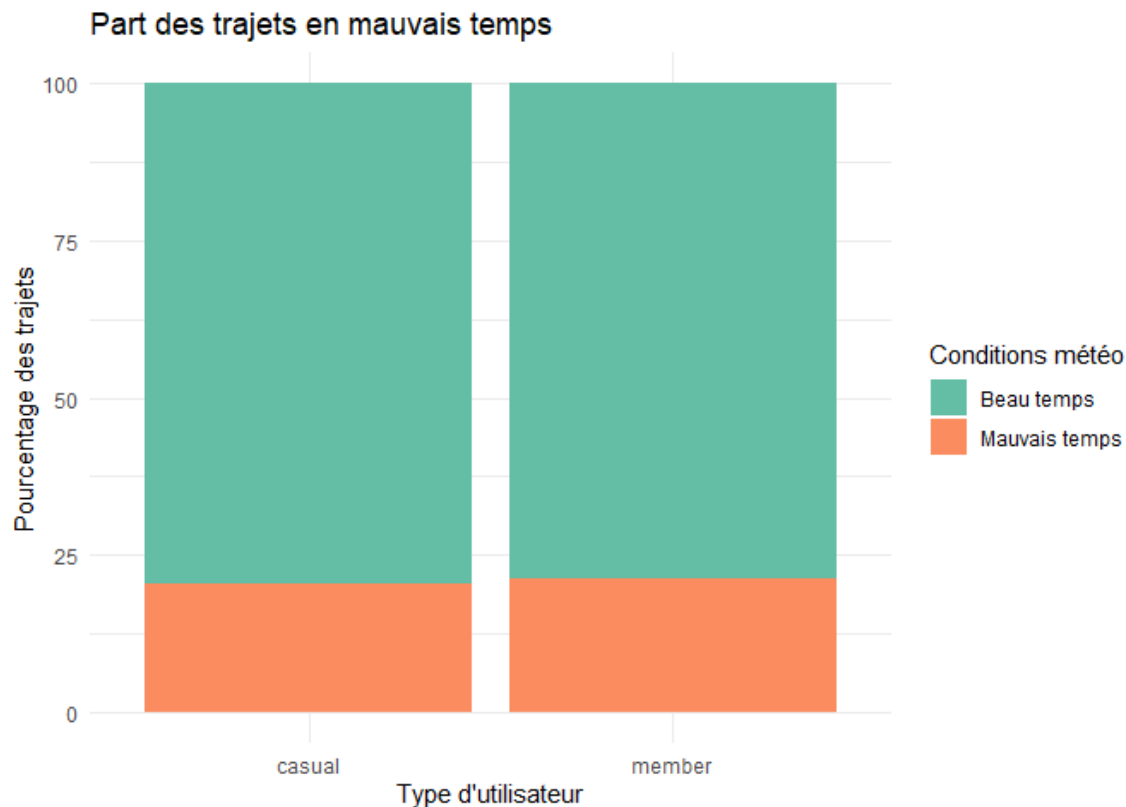
Pour les deux visualisations (mai 2022 / juillet 2024), à l'inverse des périodes froides, on voit une réaction totalement inverse. Étonnamment, les utilisateurs ne diminuent pas du tout et restent les mêmes, malgré des grosses chaleurs.

**Question 14 : Les utilisateurs membres utilisent-ils plus les vélos par mauvais temps ?**

### Hypothèse :

On se demande si les utilisateurs membres utilisent plus les vélos par mauvais temps. Nous aurions tendance à dire que oui, que les utilisateurs casual les utilisent pendant des périodes de beau temps pour faire des balades tandis que les utilisateurs membre les utilisent pour se déplacer quotidiennement, ce qui peut inclure davantage de trajets sous le mauvais temps.

### Graphique :



### Interprétation :

Sur ce graphique, nous avons représenté les jours de mauvais temps par les jours où il y a soit plus de 1mm de précipitation soit plus de 50km/h de vent. On peut voir que contrairement à nos hypothèses que le mauvais temps affecte quasiment autant les utilisateurs réguliers que les utilisateurs occasionnels.

## IV. Question bonus

### Question 15 : Est-ce qu'il est possible de voir une différence d'utilisation des vélos lors de la prise du Capitol le 6 janvier 2021 ?

#### Hypothèse :

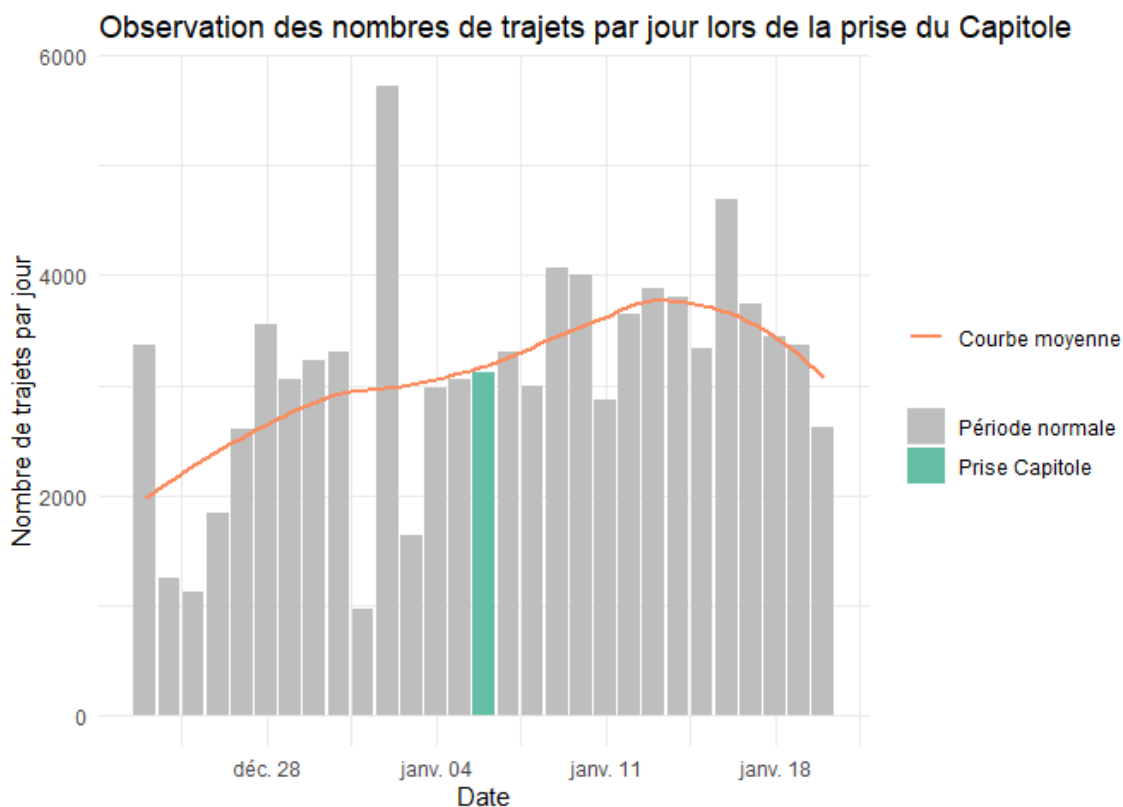
Dans la matinée du mercredi 6 janvier, les manifestants encerclent le Washington Monument et se rassemblent pour assister à des discours prononcés par Rudy Giuliani, conseiller de Trump, et par Trump lui-même. Il a réitéré des allégations infondées de fraude électorale et a exhorté ses partisans à marcher vers le Capitole, déclarant : "battons-nous comme des diables". En réponse, des milliers de manifestants ont convergé vers le Capitole (13h10), franchissant les barrières de sécurité et pénétrant dans le bâtiment (14h). Les forces de l'ordre, en sous-effectif, ont été rapidement débordées et le

capitole a été mis en confinement. A 14h48, le maire de Washington déclare un couvre-feu dans toute la ville. Ainsi on essaye alors de savoir s'il est possible de voir une différence d'utilisation des vélos lors de la prise du Capitole. On peut imaginer qu'à partir de l'information du couvre feu, les gens partent rapidement de leur travail pour rentrer.

Source : <https://fr.euronews.com/2022/01/06/l-assaut-du-capitole-retour-sur-le-deroule-de-cette-journee-noire-pour-la-democratie-ameri> <https://edition.cnn.com/2022/07/10/politics/jan-6-us-capitol-riot-timeline> <https://edition.cnn.com/interactive/2021/01/politics/us-capitol-siege/>

On souhaite d'abord voir sur l'ensemble de la journée s'il y a eu plus d'utilisation que les jours précédents et suivants le 6 janvier

### Graphique 1 :



### Interprétation :

La visualisation est faite sur 14 jours précédant et 14 jours suivant le 6 janvier. Grâce à un barplot, on peut observer le nombre de trajet cumulé chaque jour ainsi que sa courbe de régression en rouge. On observe aucun changement apparent dans l'utilisation quotidienne des vélos le 6 janvier.

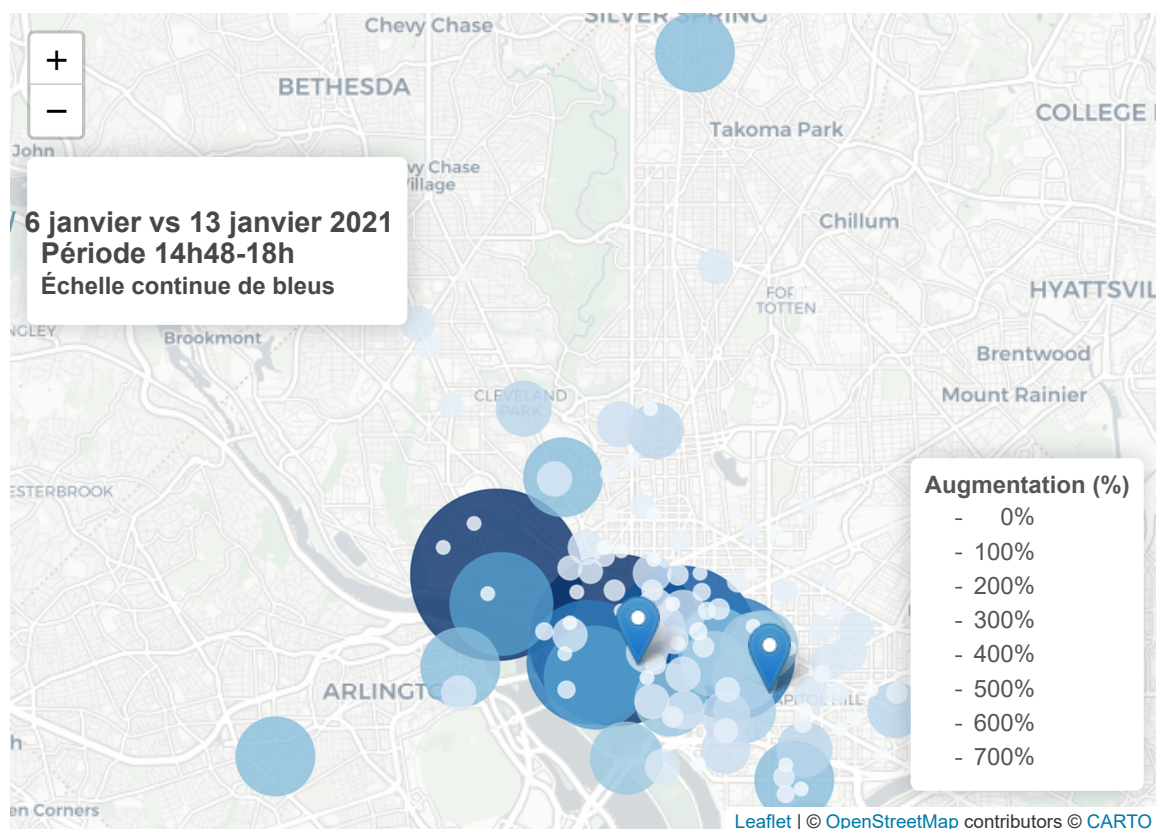
### Hypothèse :

Cependant on peut supposer que l'assaut du Capitole a entraîné une augmentation de l'utilisation des vélos autour du Capitole et du parc The Ellipse, où Donald Trump a prononcé son discours. (lieu du speech de Donald Trump). On peut donc s'intéresser à l'activité par station pour repérer des hausses à partir de ce moment là. On qualifie les augmentations par une fréquentation supérieure à la normale.

On commence la visualisation à partir du moment où les manifestants partent du Washington Monument et marchent sur la Pennsylvania Avenue (13h10) jusqu'à rejoindre le Capitole. Les

premières intrusions dans le bâtiment ont lieu vers 14h. Un couvre feu est annoncé à 14h48 et mis en place à 18h, fin de la visualisation.

## Graphique 2 :



## Interprétation :

Cette visualisation interactive montre uniquement les augmentations d'utilisation des stations de vélos en libre-service à Washington, DC, entre deux mercredis consécutifs : le 6 janvier 2021 (jour de l'assaut du Capitole) et le 13 janvier 2021, durant la même tranche horaire critique : 13h10 à 18h.

Seules les stations ayant enregistré une activité significative (au moins 3 trajets sur l'une ou l'autre des journées) sont prises en compte. Les stations ayant connu une forte baisse sont exclues pour se concentrer sur les hausses anormales. Pour chaque station, la variation de trafic est calculée en pourcentage par rapport au volume référent du 13 janvier. Échelle continue de bleus : plus l'augmentation du trafic est forte, plus la couleur de la station est foncée, et plus le cercle est large. Un clic sur un cercle affiche les trajets enregistrés pour chacun des deux jours, le pourcentage de variation, et le statut d'anomalie. Les emplacements du Capitole et du Washington Monument sont signalés pour repérer les zones d'intérêts principales.

On peut observer que les stations au nord de la Pennsylvania Avenue (avenue utilisée par les manifestants) ont des utilisations qui ont augmentées comparé à la normale. Les 3 stations avec les plus grosses augmentations autour de l'avenue sont "17th & G St NW", "10th & G St NW", et "4th & D St NW / Judiciary Square". En voyant la visualisation on remarque qu'il existe des augmentations inhabituelles autour de l'avenue potentiellement lié à la manifestation mais cependant il est difficile d'établir un lien direct.

## Conclusion



## **Nos observations :**

Lors de ce projet, nous avons pu analyser en détail l'utilisation des vélos en libre service de Washington. Nous avons pu en retirer de nombreuses informations.

Concernant l'utilisation générale des vélos, on a pu voir qu'ils étaient plus utilisés l'été, quand la température est clémente. Les utilisateurs membres les utilisent majoritairement en semaine pour aller travailler (on retrouve un effet heure de pointe particulièrement marqué) tandis que les utilisateurs occasionnels les utilisent plus le weekend pour du loisir ou du tourisme. On a pu également constater une augmentation du nombre de vélos électriques, ce qui traduit une volonté des utilisateurs de faire peu d'efforts et d'exécuter leurs trajets le plus rapidement possible.

Ensuite, en analysant l'utilisation des stations, nous avons pu voir qu'il y avait deux groupes distincts de stations : celles pour les déplacements domicile-travail et celles pour le tourisme. Les stations proche des quartiers résidentielles connaissent peu d'arrivées par rapport au nombre de départs, et sont donc souvent en déficit de vélos. Le phénomène s'inverse dans le centre-ville. Finalement, beaucoup de vélos sont utilisés dans des trajets "en boucle" pour visiter les monuments de Washington.

Concernant la météo, on voit que la température et les précipitations influent considérablement sur l'usage des vélos. Naturellement, peu d'utilisateurs se servent des vélos lorsqu'il pleut. Ce qui est néanmoins notable, c'est que ces derniers préfèrent faire du vélo pendant la canicule plutôt qu'à grand froid : la chaleur n'a donc pas l'air d'avoir une grande influence.

## **Répartition des tâches :**

Concernant l'organisation des tâches, nous nous sommes répartis les questions dès le début du projet. Nous avons fait environ 4 questions chacun. Mathis s'est occupé majoritairement des questions sur les stations, Icham et Lila sur la météo, Icham, Lila et Julien sur l'utilisation générale et Icham sur le Capitole.

Mathis s'est occupé de la Shiny App, Julien et Lila de Tableau, et Icham a fait le powerpoint de la soutenance.