



LONDON FIRE BRIGADE

ANALYSE DU TEMPS DE REPONSE



SASS Laszlo
JUN - JUILLET 2025

Table des matières

INTRODUCTION : LA LONDON FIRE BRIGADE	2
EXPLORATION, PRE-PROCESSING ET DATAVIZ	3
1. Données utilisées	3
Description et fiabilité	3
Choix des variables et limites	4
2. Pre-processing	4
3. Exploration des données	5
Catégories et sous-catégories d'incident	5
Catégories de propriété	7
Qualificatifs d'adresse	8
Incidents selon le mois	9
4. Data Visualisation	10
Temps de réponse du premier véhicule	10
Temps de réponse par catégorie d'incident	11
Temps de réponse par catégorie de propriété	13
Période de confinement	15
5. Corrélations	17
Matrice de corrélation	17
AttendanceTime et TravelTime	18
AttendanceTime et Delay	19
VISUALISATION AVANCEE	22
1. Préparation des données	22
2. Modélisation du dataset	23
Power Query	23
Schéma en étoile	23
Mesures	24
3. Rapport et visualisations	25
Rappel des objectifs	25
Carte de formes	26
Etude des objectifs	26
Focus sur les propriétés	30
Etude du retard et des appels	33
Etude du jour et de l'heure	34
CONCLUSION	35

INTRODUCTION : LA LONDON FIRE BRIGADE

"On 14 June 2017 a fire broke out at Grenfell Tower in North Kensington killing 72 people. It was one of the worst UK fires in living memory, and a tragedy for the survivors, their families, and the wider Grenfell community. The Grenfell Tower Inquiry was set up to examine the circumstances leading up to and surrounding the fire. It has since been a catalyst for deep institutional transformation within the Brigade."

Extrait du [site internet de la LFB](#)

Divisée en 33 boroughs (arrondissements), Londres est une ville couvrant 1587 km² et comptant 8 millions d'habitants, soit 1/7 de la population anglaise totale. Superficie et population équivalent à deux fois la petite couronne parisienne. La capitale est dotée d'une brigade de pompiers parmi les plus grandes au monde, la *London Fire Brigade*.

Créée en 1865, la LFB a pour mission de « [servir et protéger Londres](#) ». Elle emploie plus de 5 000 personnes, opérant depuis 109 sites, dont **102 casernes de pompiers**, la station fluviale et des locaux de soutien tels que le quartier général au centre de Londres. **Les véhicules et les équipements de la LFB doivent être fiables et facilement disponibles, l'indisponibilité des véhicules et des équipements ayant un impact significatif sur sa capacité à remplir ses obligations statutaires.**

En 2017, l'incendie de la Grenfell Tower ayant causé 72 victimes est venu remettre en cause l'organisation de la LFB et accélérer sa modernisation notamment par la mise en place de technologies télémétriques d'information.

Il s'ensuit qu'elle se fixe les objectifs suivants :

« L'approche opérationnelle de la LFB en matière de lutte contre les incendies et d'opérations de sauvetage est déterminée par le département **Operational Policy and Assurance (OP&A)**. **Les objectifs d'intervention actuels sont de 6 minutes en moyenne pour le premier appareil et de 8 minutes en moyenne pour le deuxième appareil. Un appareil doit arriver sur les lieux d'un incident partout dans Londres dans les 10 minutes pour 90 % des incidents et dans les 12 minutes pour 95 % des incidents. Un nombre prédéterminé d'appareils est placé stratégiquement dans des zones de Londres où ils peuvent être nécessaires pour atténuer le risque associé à cette zone spécifique.** L'une des principales exigences techniques des appareils de pompage est qu'ils doivent être **capables de pomper pendant 4h en continu sans être ravitaillés en carburant.** »

Bien que le nombre de victimes ait diminué au fil du temps, le nombre d'incidents est quant à lui en augmentation ces dernières années. L'atteinte des objectifs énoncés ci-dessus est donc primordiale puisque la catastrophe de Grenfell reste dans les esprits de par son ampleur et sa gestion sans doute perfectible.

Dans ce projet, j'étudie la période 2018-2020 qui suit directement l'incident. L'objectif est la réalisation d'une analyse des temps de réponse, c'est-à-dire des durées entre la réception d'un appel dans une caserne et l'arrivée d'un appareil de pompiers sur les lieux d'un incident, afin de comparer la performance de la LFB aux objectifs attendus et d'identifier les facteurs influençant le temps de réponse.

EXPLORATION, PRE-PROCESSING ET DATAVIZ

1. Données utilisées

Description et fiabilité

On dispose d'un jeu de données concernant les **incidents**, notamment leur catégorie, date et heure, l'arrondissement et le type de propriété, des informations géographiques ainsi que la durée d'arrivée du premier et du deuxième appareil. Ainsi qu'un jeu de données sur la **mobilitation** qui se concentre davantage sur le temps de réaction, le temps de trajet et le temps de réponse général des appareils.

En moyenne, on dénombre environ 100 000 incidents par an entre 2018 et 2020, ce qui en fait un jeu de données riche. Régi par le UK Code of Practice for Statistics et contrôlé par le UK Statistics Authority et le Statistics and Registration Service Act 2007, on peut considérer qu'il est [fiable](#). Les données sont disponibles en libre accès depuis le [London Datastore](#).

Choix des variables et limites

Les variables utilisées concernent les incidents, leur catégorie et sous-catégorie, les types de propriété, les qualificatifs d'adresse et les temps de réponse des différents appareils. J'ai également considéré les types de retard, le nombre d'appels ou encore l'arrondissement comme axes d'études pertinents.

Les variables sont pour la plupart qualitatives, une matrice de corrélation ne permettra donc pas d'identifier toutes les corrélations. D'autres axes d'analyse tels que la latitude et la longitude ou encore la caserne de départ pourraient s'avérer intéressants, cependant j'ai fait le choix de limiter le nombre d'axes plutôt que d'effectuer une étude très sommaire de chacun d'eux.

2. Pre-processing

Le jeu de données ne contient aucun doublon. Il y a en revanche quelques valeurs manquantes à traiter. Pour certaines variables, le nombre de valeurs manquantes est élevé mais il faut faire attention. Par exemple, la variable « Type de Service Spécial » n'a de sens que lorsque l'incident est de type « Service Spécial », ce qui ne concerne évidemment pas tous les incidents. Les remplacements nécessaires sont effectués sur les valeurs qui manquent réellement.

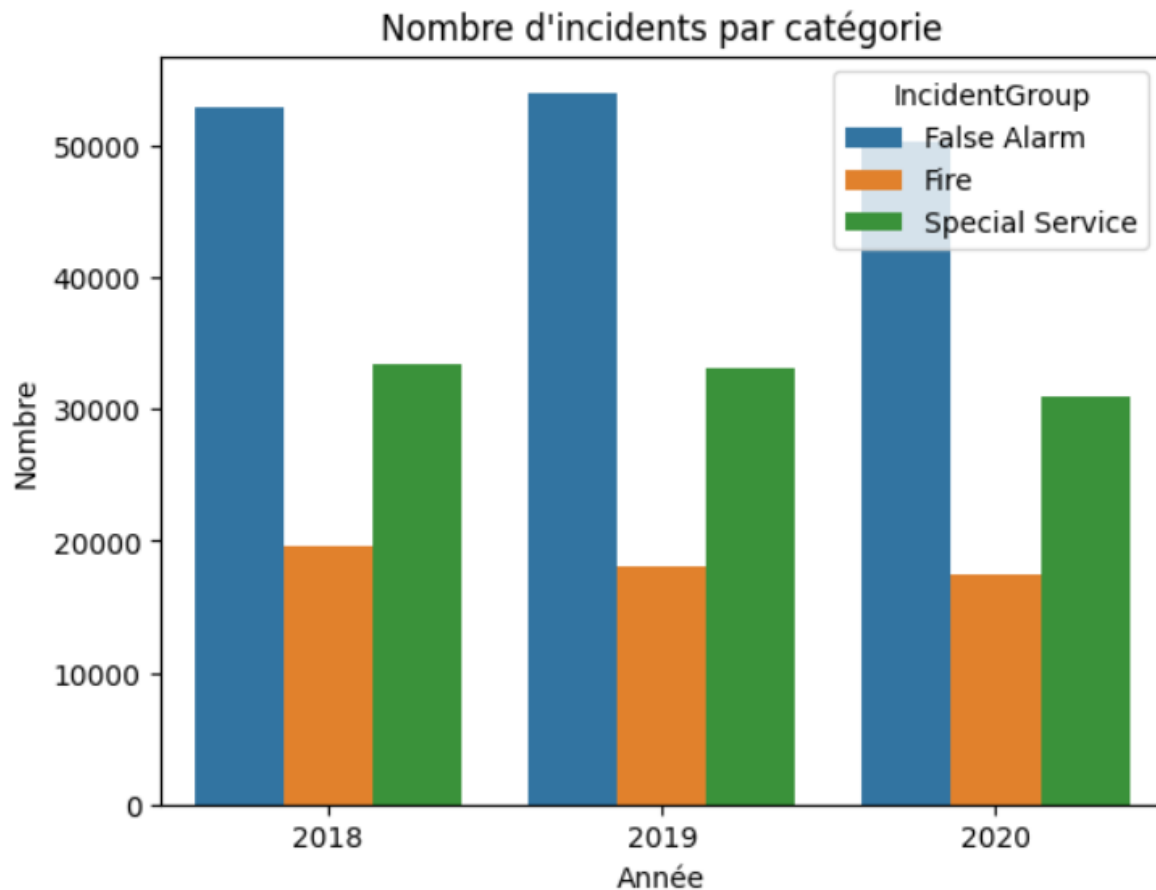
J'ai enrichi le jeu de données de nouvelles variables lors d'une étude préliminaire, notamment une variable « Mois » afin d'effectuer des explorations mensuelles. Pour celles-ci, j'ai également appliqué une standardisation des valeurs après avoir compté le nombre d'incidents par mois. Ceci a servi à titre personnel mais n'a pas été utilisé dans ce rapport.

D'autres transformations telles que la normalisation ou la réduction de dimension n'ont pas été requises pour les données de ce projet.

3. Exploration des données

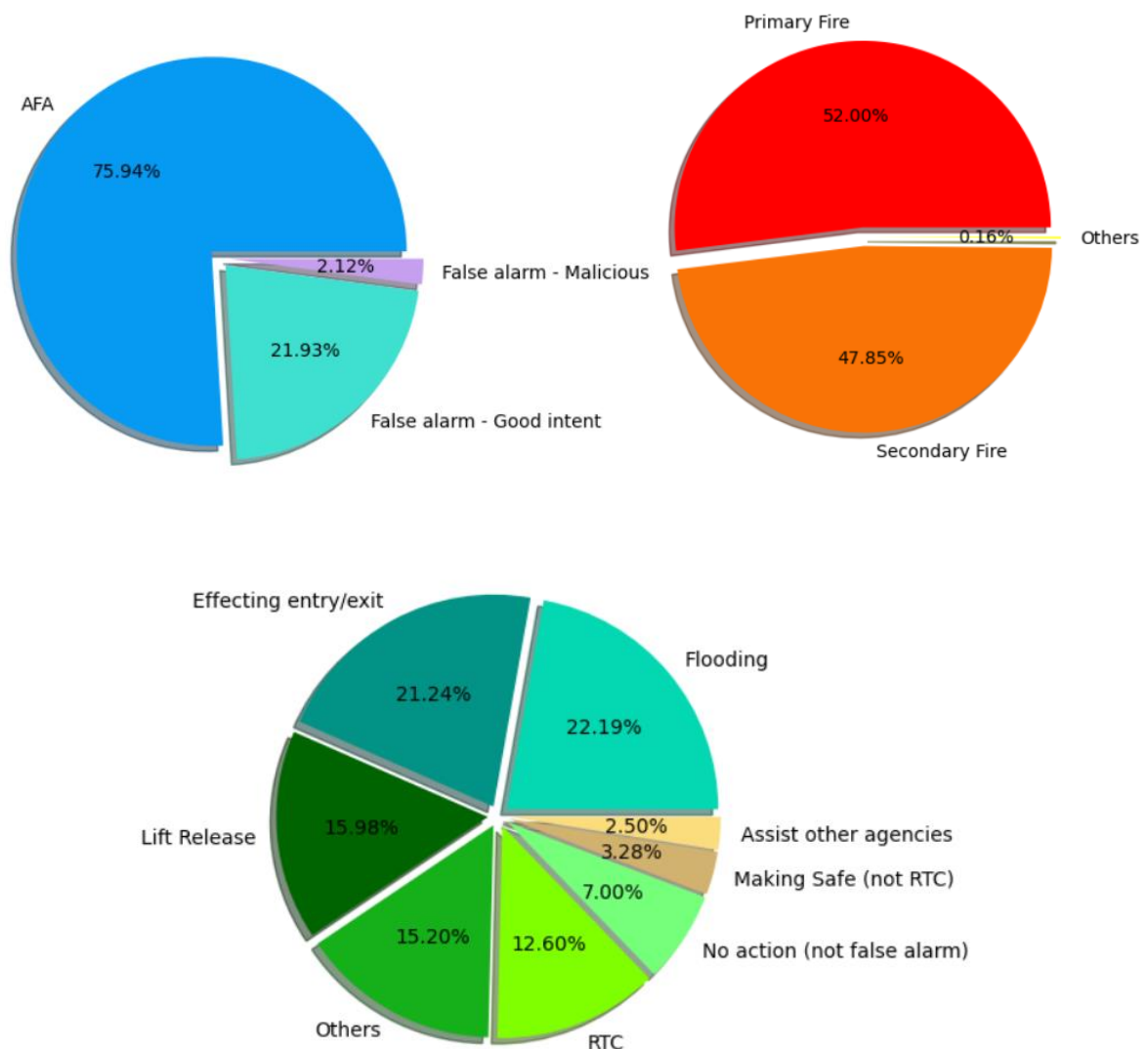
Catégories et sous-catégories d'incident

On commence par visualiser les types d'incidents qui sont regroupés en trois grandes catégories comme le montre le graphique ci-dessous.



- Les **fausses alarmes** représentent la majorité des incidents (51%), ce qui peut être surprenant à première vue. Les **services spéciaux** représentent une part également importante (31%), on verra dans la section suivante à quoi ils correspondent. Enfin les **feux** représentent la part la moins grande (18%).
- On note que le nombre d'incidents par catégorie est assez stable d'une année à l'autre. Il y a même une légère baisse de façon globale, excepté pour les fausses alarmes entre 2018 et 2019.

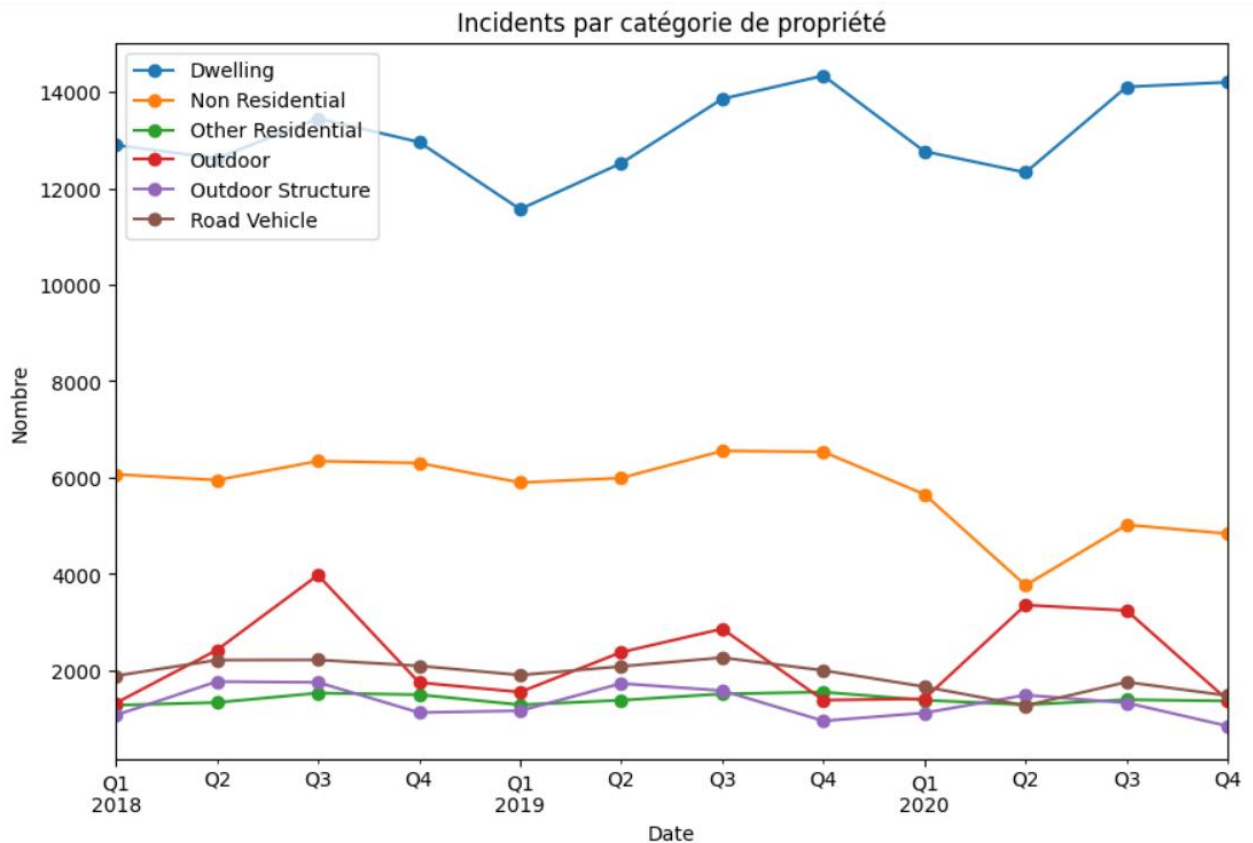
Les pie charts ci-dessous décrivent quant à eux les sous-catégories des incidents.



- La forte proportion de fausses alarmes provient des **Automatic Fire Alarm (AFA)**, ce qui suggère que de très nombreuses alarmes se déclenchent alors qu'elles ne le devraient pas. Cela entrave l'efficacité de la LFB.
- Les feux sont répartis de façon assez homogène entre **feux primaires** (sérieux, impliquant des décès ou ayant nécessité au moins cinq appareils, dont les feux d'habitation, de véhicules et de bâtiments en extérieur), et **feux secondaires** (moindres, ne concernant pas de propriété de personnes).
- Enfin les services spéciaux regroupent principalement les **inondations**, les **incidents affectant une entrée ou une sortie**, et les **pannes d'ascenseurs**. On ne représente que les principales sous-catégories ici.

Catégories de propriété

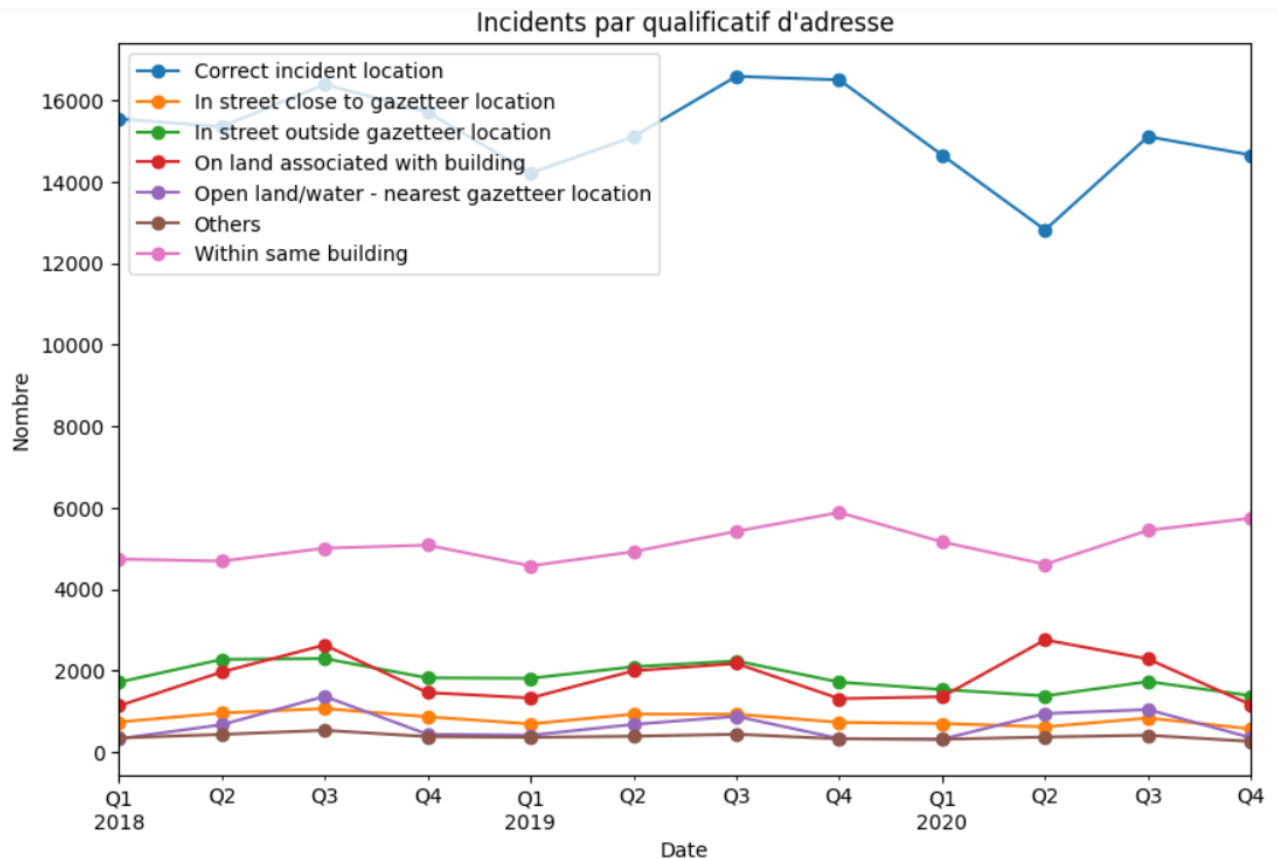
On a vu que les incidents peuvent concerner différentes catégories de propriété, mais lesquelles ? Le graphique suivant répond à cette question.



- Les **habitations (Dwelling)** représentent la majorité des propriétés, ce qui est intéressant car cela concerne les feux primaires sujets à des incidents graves. Il est donc d'autant plus important que le temps de réponse soit faible pour ce type de propriété.
- Les **propriétés non résidentielles** sont également assez représentées, elles aussi potentiellement sujettes à des feux de grande ampleur.
- D'autres catégories comme les **structures en extérieur** ou les **véhicules routiers** sont représentées mais ne constituent pas une majorité de cas.

Qualificatifs d'adresse

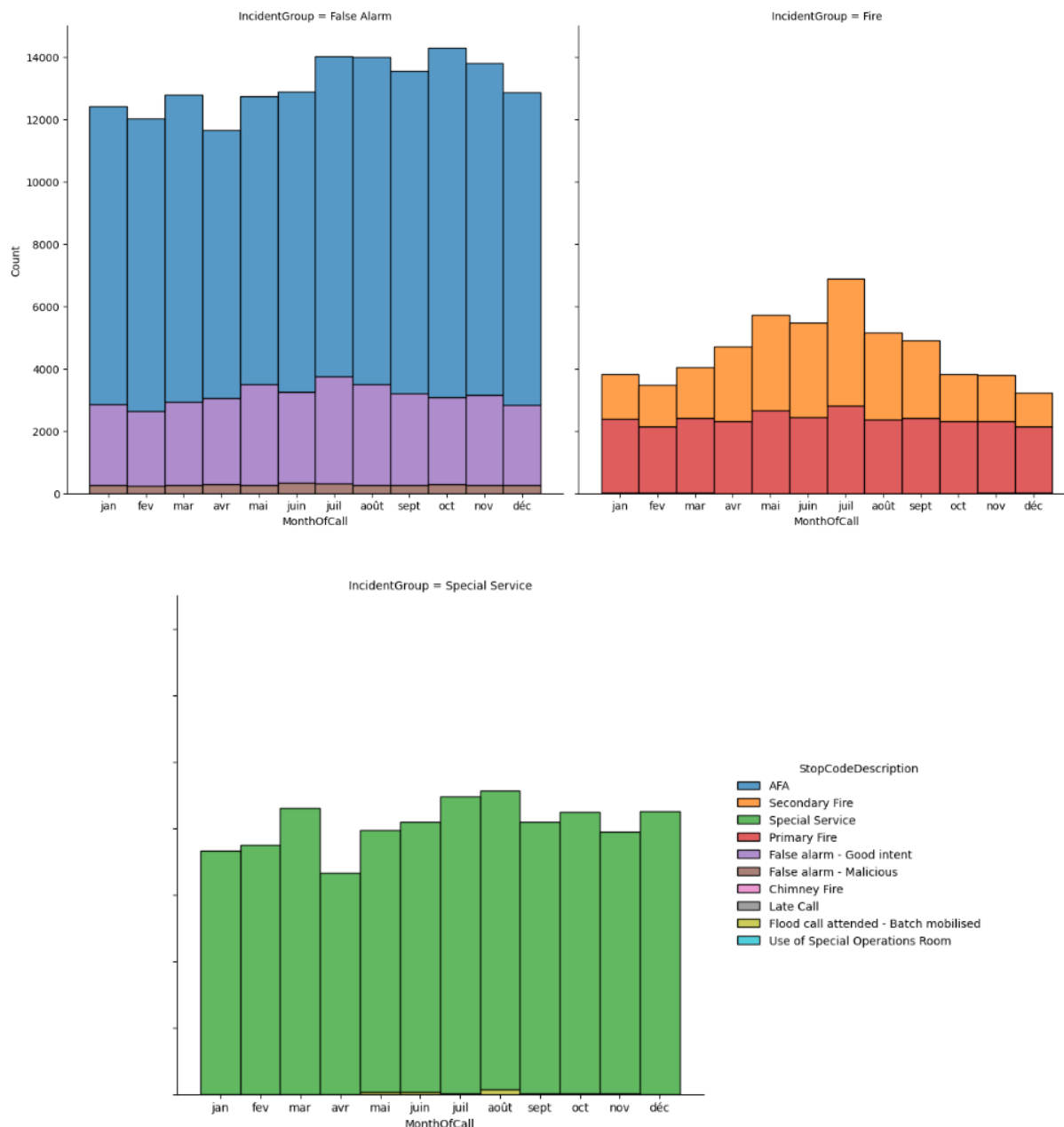
La variable **qualificatif d'adresse** indique si l'incident se situe à l'adresse renseignée ou non (le cas échéant, est-ce qu'il se situe dans le même bâtiment, la même rue, ...). On visualise la répartition des valeurs comme pour les propriétés.



- Le graphique montre que la plupart des incidents sont **bien renseignés**. L'efficacité de la LFB sera donc bonne et on s'attend à ce qu'il n'y ait pas trop de retard dû à une mauvaise localisation.
- En deuxième position, les incidents se situent **dans le même bâtiment** que le lieu indiqué, ce qui reste acceptable. En revanche, certains cas suggèrent qu'il se situe dans la même rue, voire plus loin parfois, ce qui augmentera le temps de réponse.

Incidents selon le mois

La question de l'incidence du mois et de la saison sur les incidents et le temps de réponse peut être étudiée en fonction de la catégorie de l'incident (IncidentGroup).



On constate un lien fort entre le mois de l'année et les catégories d'incidents observés. Plus précisément :

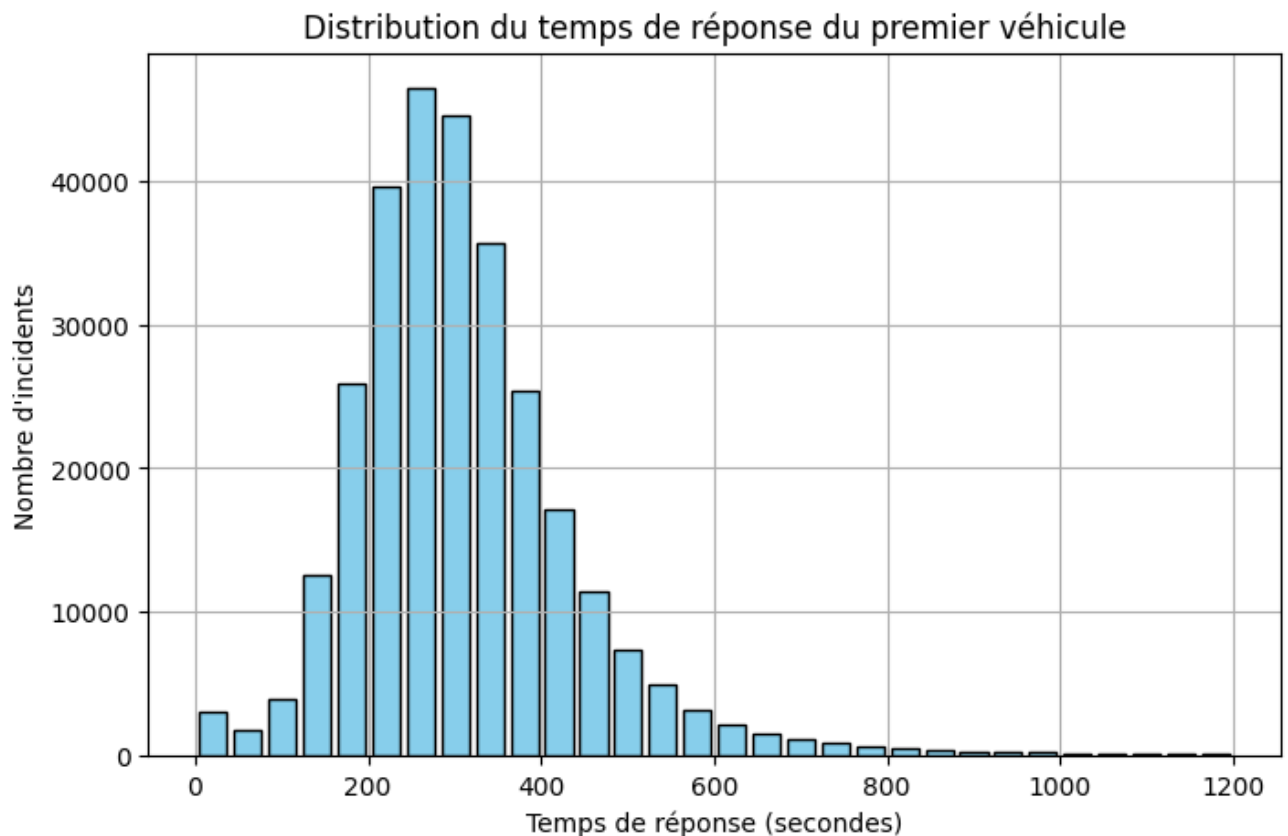
- Les **fausses alarmes** sont en nette augmentation au deuxième semestre ainsi qu'en fin d'année. Plus particulièrement, les **fausses alarmes avec mauvaises intentions** sont régulières sur l'année tandis que les **alarmes automatiques** sont en augmentation en fin d'année au contraire **des fausses alarmes avec bonnes intentions** dont le pic est atteint en été.

- Les **feux primaires** sont répartis de façon presque constante sur l'année. En revanche, on observe une forte augmentation des **feux secondaires** en été avec un net pic en juillet.
- Enfin les **services spéciaux** baissent de façon importante au printemps avec un pic en été poussé par les inondations.

4. Data Visualisation

Temps de réponse du premier véhicule

Le graphique ci-dessous présente la distribution du temps de réponse du premier véhicule de la LFB.



On observe une concentration des incidents avec un temps de réponse compris entre **200 et 400 secondes**, ce qui correspond globalement aux objectifs fixés par la brigade (un premier véhicule attendu en moins de 6 minutes en moyenne). Cependant, on note également la présence de valeurs extrêmes :

- 0,3% d'incidents ont un temps de réponse de 5 secondes ou moins, ce qui est aberrant sachant qu'il faut se préparer et se rendre sur les lieux de l'incident.

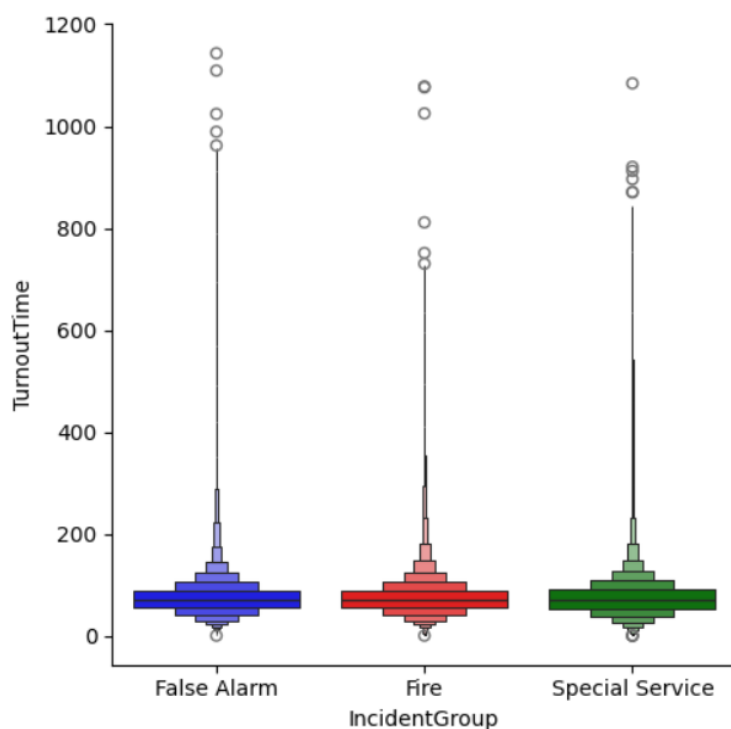
- 3% d'incidents ont des temps supérieurs à 10 minutes (600 secondes) et 0,4% dépassent même les 15 minutes (900 secondes). Ces valeurs témoignent de situations exceptionnelles (localisation imprécise, trafic, incidents dans des zones difficiles d'accès).

Ces résultats illustrent que, bien que la majorité des interventions se situent dans les délais attendus, des efforts restent à mener pour réduire les temps les plus longs, en particulier dans le cadre des objectifs post-Grenfell.

Temps de réponse par catégorie d'incident

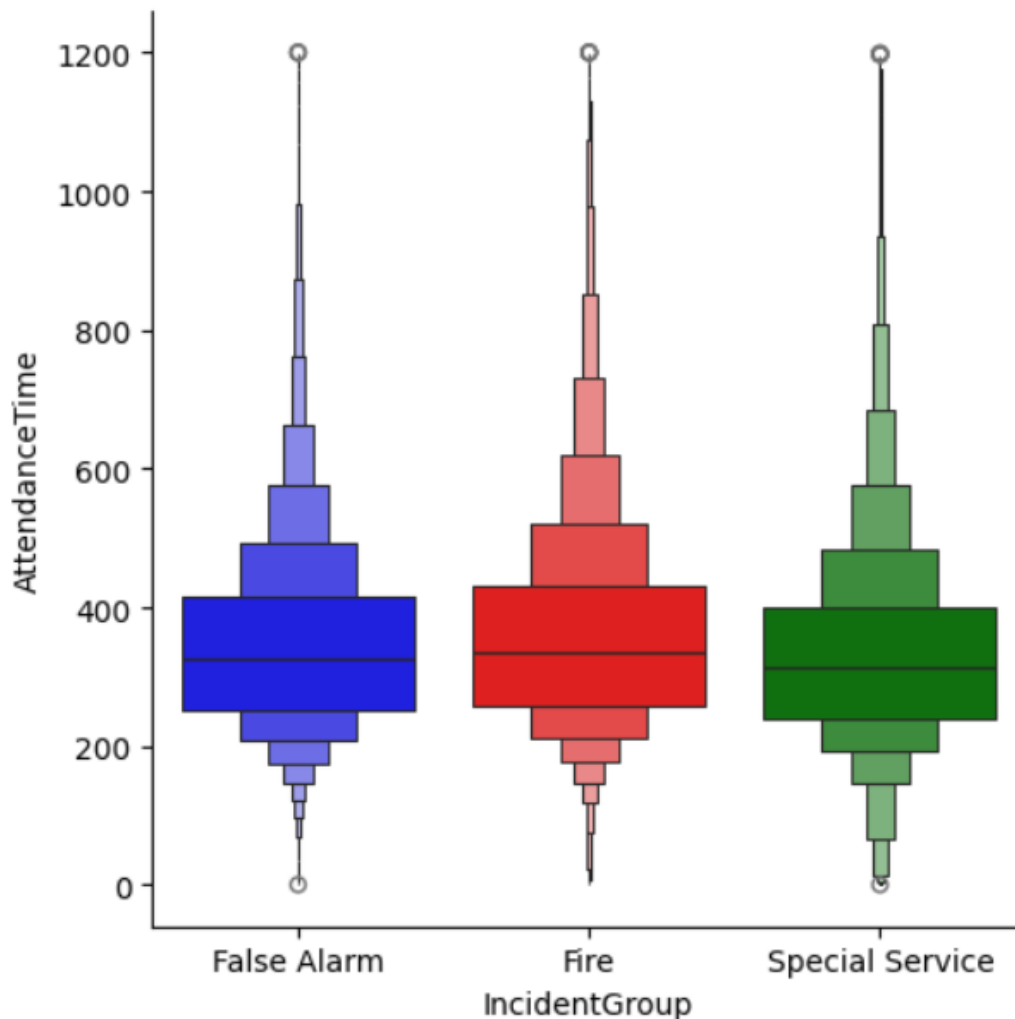
Les deux graphiques ci-dessous illustrent le temps de préparation et de réponse des véhicules de pompiers selon la catégorie des incidents.

- **Temps de préparation (TurnoutTime)**



Le temps de préparation est relativement homogène entre les trois types d'incident, avec une majorité des cas concentrés sous les 150 secondes (soit 2 à 2,5 minutes). Quelques valeurs extrêmes dépassent 600 secondes, voire 1000 secondes dans de rares situations. Cela suggère que, quelle que soit la nature de l'incident, les équipes quittent la caserne dans des délais similaires. Les valeurs très élevées pourraient correspondre à des erreurs d'enregistrement ou à des situations très spécifiques (problèmes techniques, attente d'instructions complémentaires, etc.).

- Temps de réponse (AttendanceTime)



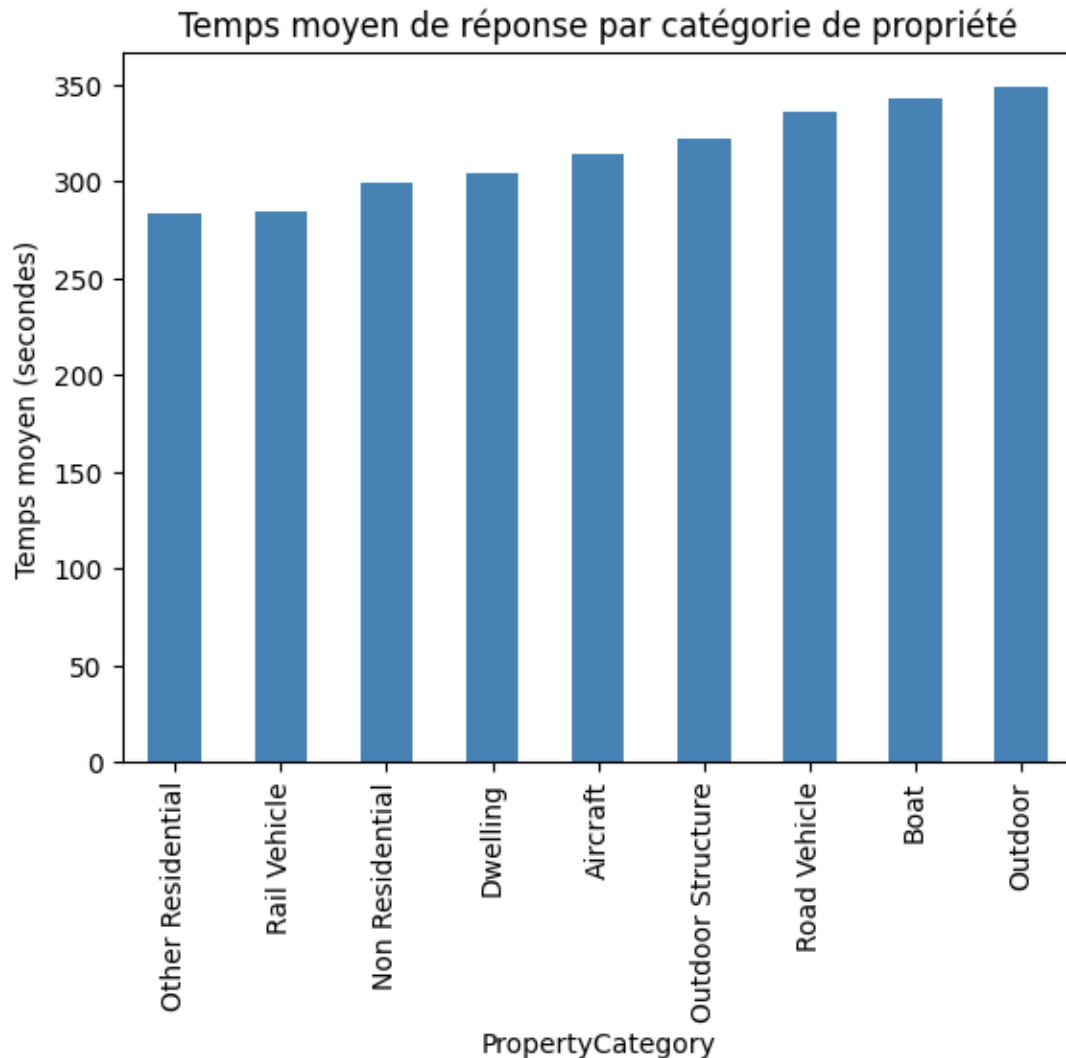
Le temps de réponse est globalement similaire quel que soit le type d'incident traité. Les médianes se situent autour des **300 secondes** pour les trois types d'incident et la dispersion des valeurs est comparable. Cela montre que la nature de l'incident a peu d'influence sur le délai moyen d'arrivée sur place.

On observe également la présence de valeurs extrêmes dépassant **1000 secondes**, ce qui reflète la diversité des situations rencontrées.

Ces résultats confirment que les fausses alarmes, les plus nombreuses dans le jeu de données, mobilisent les équipes de la même façon que les incidents réels, ce qui peut nuire à l'efficacité globale.

Temps de réponse par catégorie de propriété

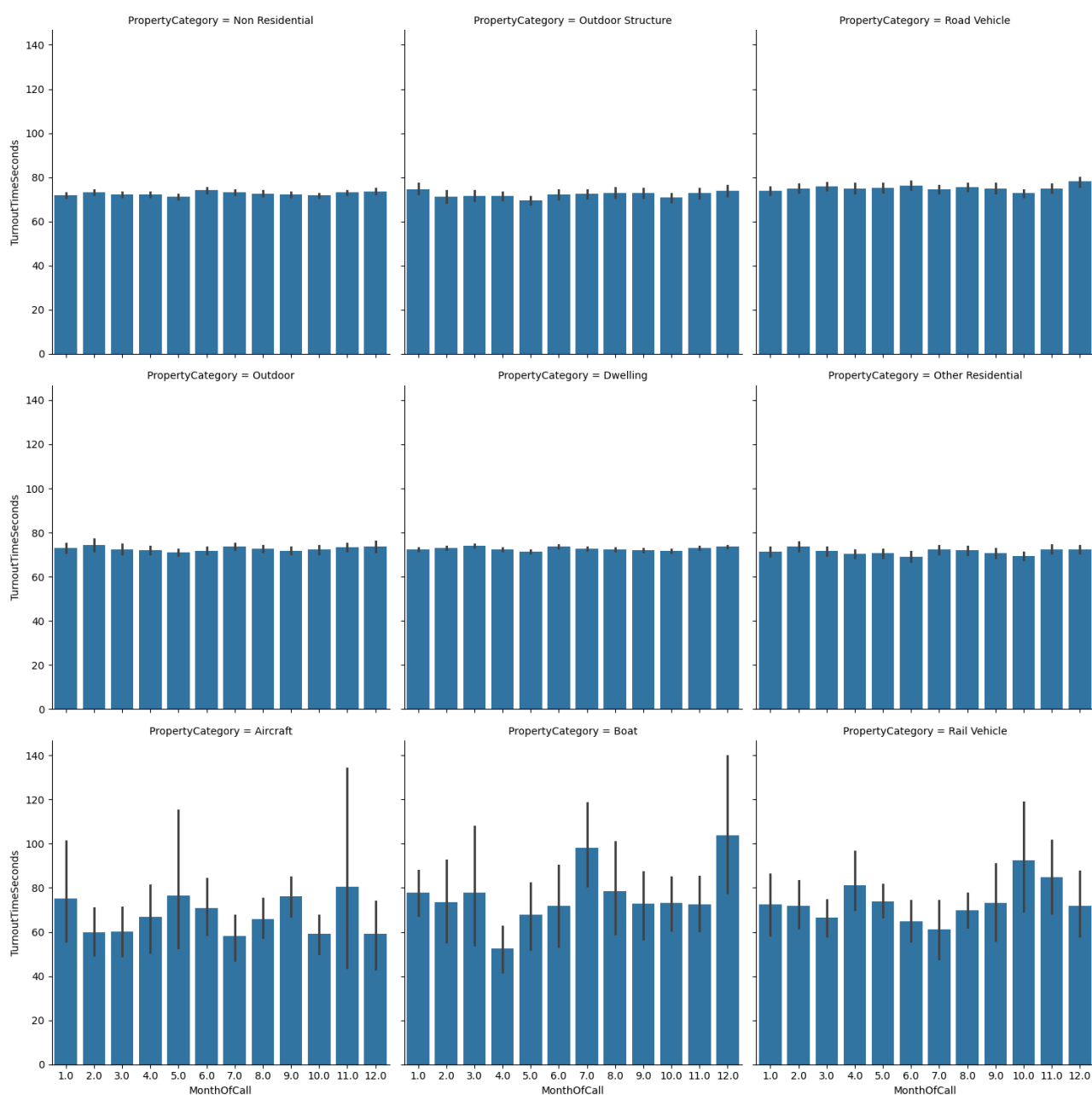
Le graphique suivant présente le temps moyen de réponse du premier véhicule selon la catégorie de propriété concernée.



On observe que :

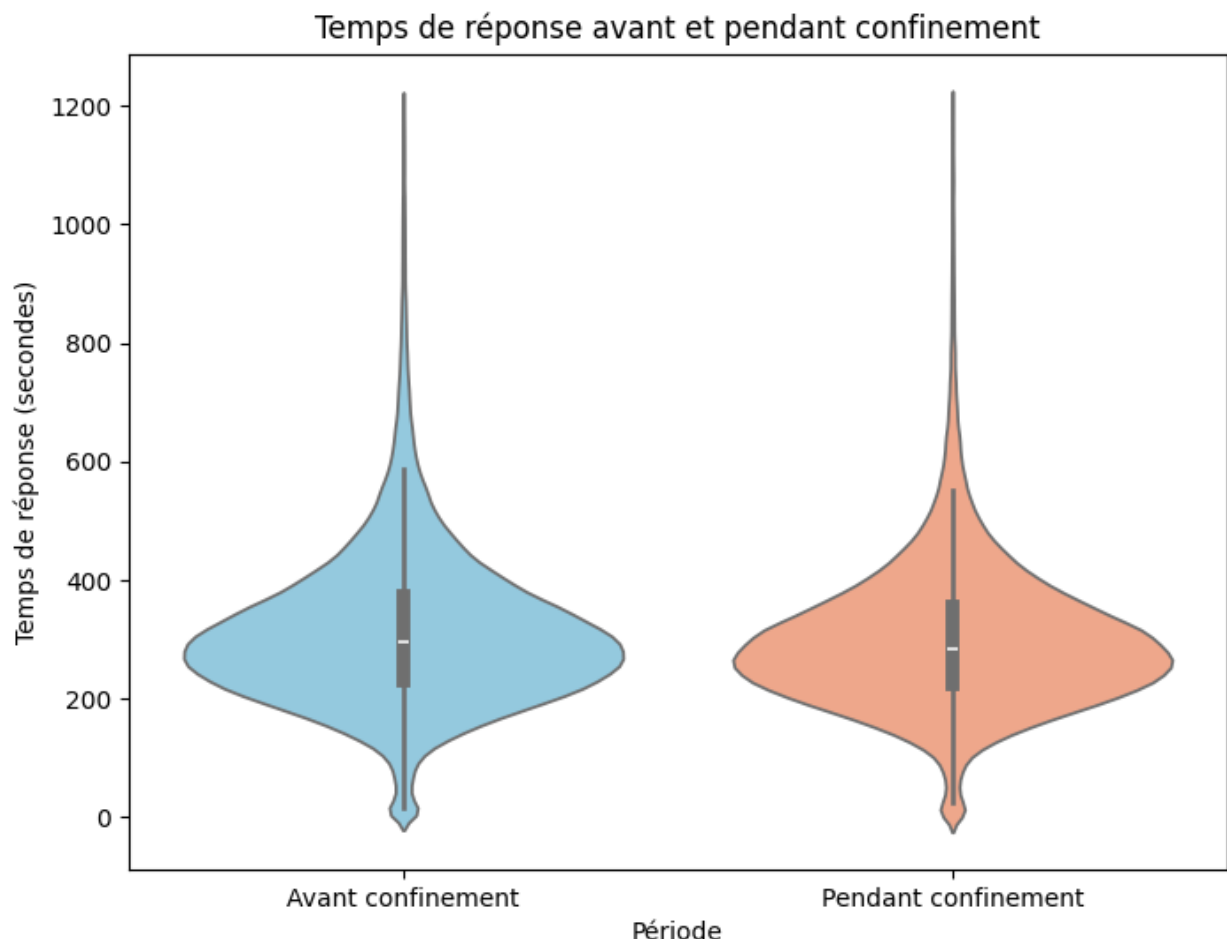
- Les habitations (**Dwelling**, **Other Residential**), les **propriétés non résidentielles**, les **véhicules ferroviaires** et l'**aéronef** affichent des délais moyens proches, entre 280 et 300 secondes.
- Les temps moyens sont légèrement plus élevés pour les interventions en zones **extérieures**, concernant les **bateaux** ou les **véhicules routiers**, dépassant les 330-340 secondes.

Ces résultats traduisent la capacité de la LFB à intervenir de façon similaire sur les bâtiments résidentiels et non résidentiels, dans lesquels les enjeux humains sont potentiellement les plus critiques, alors que les temps légèrement plus longs observés pour les interventions sur des lieux plus spécifiques (bateaux, espaces ouverts) peuvent s'expliquer par la nature des incidents ou leur localisation. On peut également noter que le temps d'intervention est fortement affecté par le mois en ce qui concerne les incidents en dehors des zones classiques accessibles par la route : avion, bateau et véhicules sur rail (graphique ci-dessous).



Période de confinement

Le graphique ci-dessous compare la distribution des temps de réponse du premier véhicule avant et pendant la période de confinement, laquelle a débuté à Londres le **23 mars 2020**.



On observe des distributions globalement similaires entre les deux périodes :

- Les temps médians sont proches, autour de **300 secondes**, et la concentration des temps de réponse se situe majoritairement entre **200 et 400 secondes**, avec une proportion comparable de valeurs dépassant les **600 secondes** dans les deux cas.
- Cela suggère que la période de confinement n'a pas eu d'impact majeur sur les délais d'intervention du LFB, malgré un contexte opérationnel particulier. Toutefois, une analyse statistique complémentaire est nécessaire afin de confirmer l'existence d'une différence significative entre les temps de réponse avant et pendant le confinement.

→ Calcul de la moyenne et de l'écart-type du temps de réponse

	mean	std	count
Period			
Avant confinement	312.987911	133.674860	219046
Pendant confinement	299.294711	123.294407	72980

On constate une légère diminution du temps moyen de réponse pendant la période de confinement (environ 14 secondes). L'écart-type est également plus faible, ce qui traduit une dispersion un peu moins marquée du temps de réponse. Cette tendance pourrait s'expliquer par un effet indirect du confinement (réduction du trafic, moindre complexité des interventions). Il convient de confirmer cela par une analyse statistique complémentaire telle qu'un test ANOVA.

→ Test ANOVA d'hypothèse

H0 : Le temps de réponse moyen est le même avant et pendant le confinement.
H1 : Le temps de réponse moyen est différent selon la période.

	df	sum_sq	mean_sq	F	PR(>F)
Period	1.0	1.026426e+07	1.026426e+07	596.677687	1.199674e-131
Residual	292024.0	5.023499e+09	1.720235e+04	NaN	NaN

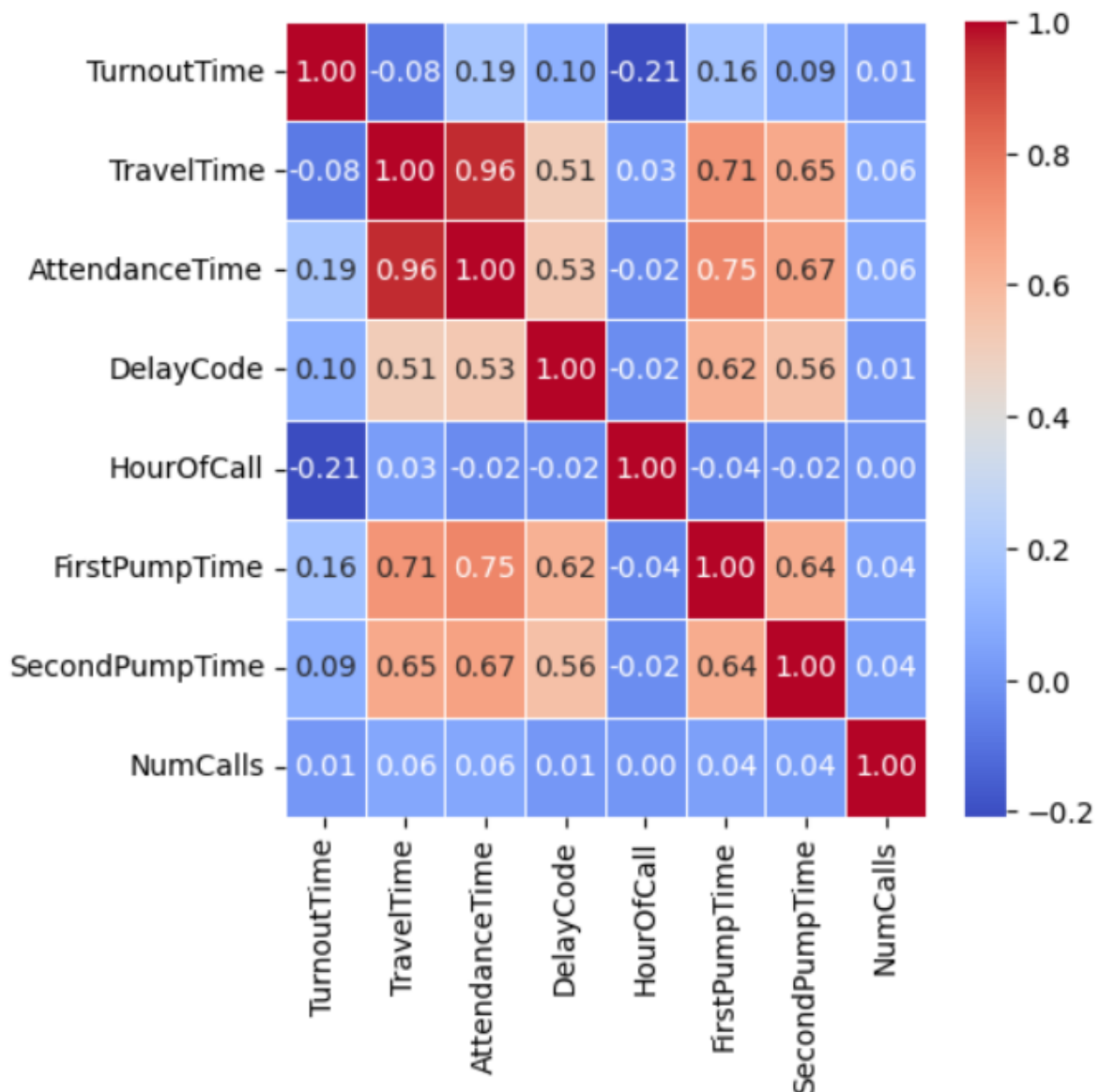
La p-value obtenue (**environ $1,20 \times 10^{-131}$**) est largement inférieure au seuil de 5%. Cela nous conduit à rejeter l'hypothèse nulle d'égalité des moyennes.

Conclusion : Le confinement a eu un effet significatif sur le temps de réponse de la LFB. Les statistiques descriptives, comme le test ANOVA, confirment une diminution du temps moyen de réponse ainsi qu'une dispersion légèrement moins marquée du temps d'intervention pendant le confinement, par rapport à la période précédente.

5. Corrélations

Matrice de corrélation

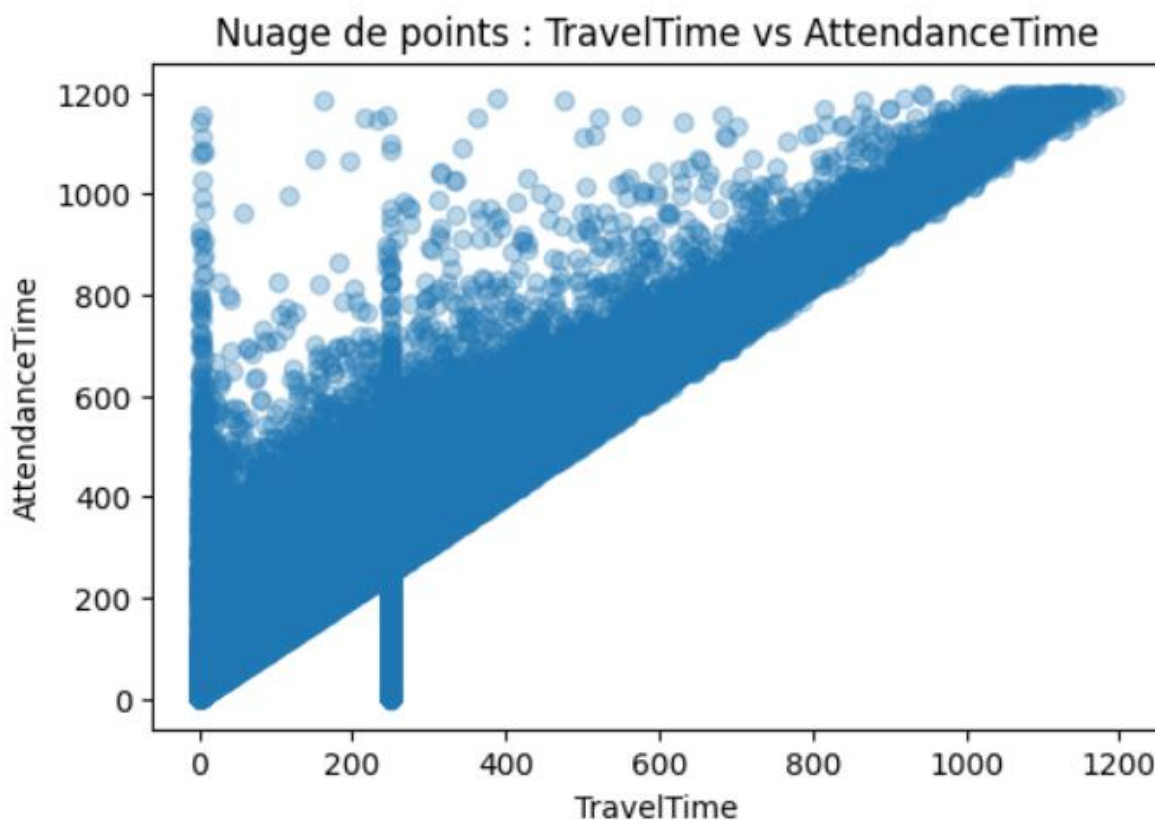
Pour compléter l'étude, il convient de s'intéresser aux corrélations entre le temps de réponse (AttendanceTime) et d'autres variables pouvant avoir un impact significatif dessus. Pour ce faire, on sélectionne les variables quantitatives pertinentes et on obtient la matrice de corrélation suivante.



- Les variables FirstPumpTime et SecondPumpTime montrent une corrélation élevée avec AttendanceTime, et c'est normal car AttendanceTime contient les temps de tous les appareils, donc en particulier du premier et du deuxième. On ne va donc pas effectuer d'étude complémentaire sur cette corrélation.
- Par définition, $\text{AttendanceTime} = \text{TravelTime} + \text{TurnoutTime}$. On s'attend donc à ce que ces variables soient corrélées et c'est bien le cas de **TravelTime** mais pas de TurnoutTime. On l'explique par le fait que TravelTime a beaucoup plus de poids que TurnoutTime (le trajet est plus long que la préparation en général). De plus, le temps de trajet est soumis aux aléas de la circulation, de la météo et de la localisation. Un allongement du temps de trajet impacte immédiatement le temps de réponse tandis que la durée de préparation reste la même.
- Une autre variable a une certaine corrélation : **DelayCode**. On va donc étudier la corrélation entre le type de retard et le temps de réponse.
- Enfin, les variables heures de l'incident et nombre d'appels ne montrent pas de corrélation significative.

AttendanceTime et TravelTime

Dans un premier temps, on a souhaité vérifier la linéarité entre les variables *AttendanceTime* et *TravelTime* afin de déterminer si un test de corrélation de Pearson était approprié.



Le nuage de points met en évidence une relation globalement linéaire et positive : **plus le temps de trajet augmente, plus le temps d'intervention augmente également**. Étant donné que la linéarité des deux variables est bien respectée, on réalise un test de Pearson afin de quantifier cette relation. Les résultats du test sont les suivants :

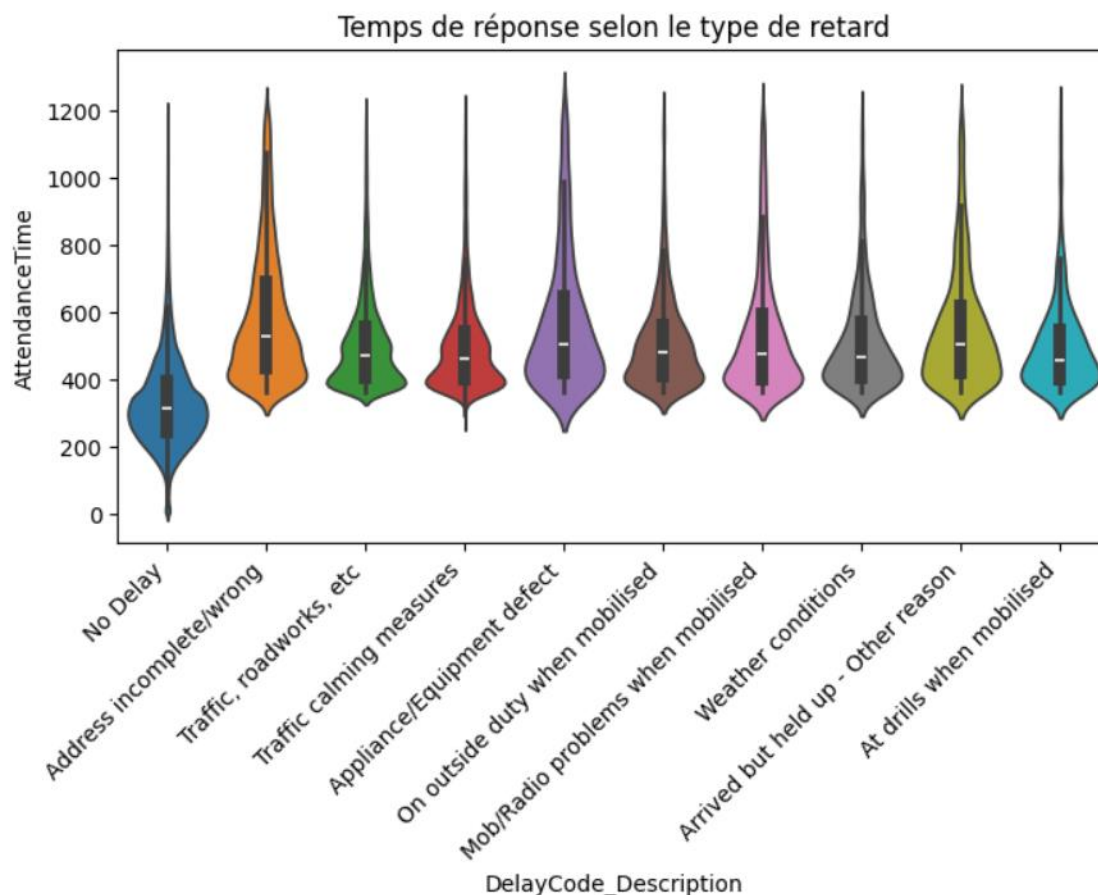
- **Coefficient de corrélation de Pearson : 0.965**
- **p-value : 0.000**

Ces résultats indiquent une **très forte corrélation linéaire positive** entre le temps de trajet et le temps de réponse. La p-value confirme que cette corrélation est hautement significative sur le plan statistique.

Conclusion : plus le temps de trajet augmente, plus le temps d'intervention augmente également, de manière presque proportionnelle.

AttendanceTime et Delay

Ci-dessous la répartition du temps de réponse en fonction des différentes modalités de retard.



- **Les incidents sans retard** (No Delay) présentent des temps d'intervention plus courts, avec une concentration des valeurs autour de **300 secondes**.
- Les différents retards entraînent quant à eux des temps de réponse plus élevés. Certains types tels que **Address incomplete** ou **Appliance/Equipment defect** sont associés à des distributions dont la médiane et la densité sont situées à des niveaux plus élevés, dépassant fréquemment 500 secondes.
- La variabilité des temps d'intervention est également plus grande en présence de retard, avec des distributions qui s'étirent davantage vers des temps longs.

Conclusion : Ce graphique met en évidence l'impact des différents types de retard sur le temps de réponse. Il confirme visuellement que la présence d'un retard, quel qu'il soit, tend à allonger le temps de réponse de la brigade. Ces observations seront confirmées par le calcul des statistiques descriptives et le test ANOVA.

→ Calcul de la moyenne et de l'écart-type du temps de réponse

	mean	std
DelayCode_Description		
Address incomplete/wrong	583.85	195.35
Appliance/Equipment defect	567.09	199.49
Arrived but held up - Other reason	550.85	173.75
At drills when mobilised	496.06	131.09
Mob/Radio problems when mobilised	533.53	181.87
No Delay	329.93	137.28
On outside duty when mobilised	505.72	132.70
Traffic calming measures	493.33	130.28
Traffic, roadworks, etc	502.19	135.65
Weather conditions	510.57	146.64

L'analyse des moyennes montre des différences marquées entre les types de retard.

- **Sans retard**, le temps moyen d'intervention est nettement plus bas (environ 302 secondes), ce qui confirme que l'absence de retard facilite une intervention rapide.
- **En présence de retards**, le temps moyen d'intervention est systématiquement plus élevé. Les retards liés à une adresse incomplète ou fausse (584 secondes) ou un défaut d'équipement (567 secondes) entraînent des temps de réponse plus longs que les retards dus aux conditions météo ou à des problèmes de radio générant également des moyennes (510 à 535 secondes).




Les écarts-types confirment que les situations avec retard traduisent une plus grande variabilité du temps de réponse selon les circonstances (par exemple 199 secondes pour les défauts d'équipement). À l'inverse, l'absence de retard s'accompagne d'un écart-type plus faible, ce qui traduit des interventions plus homogènes en durée.

Conclusion : Ces résultats illustrent clairement l'impact des différents types de retard sur la durée d'intervention des pompiers. Certains types de retard sont particulièrement pénalisants et pourraient faire l'objet d'actions spécifiques pour réduire leur incidence.

→ Test ANOVA d'hypothèse

On veut vérifier si le temps moyen de réponse varie de manière significative selon les différentes modalités de retard.

H0: Le temps moyen d'intervention est identique pour toutes les modalités de retard
H1: Il existe au moins une modalité de retard pour laquelle le temps moyen d'intervention est différent.

	df	sum_sq	mean_sq	F	PR(>F)	
C(DelayCode_Description)	9.0	1.425466e+09	1.583851e+08	8238.758089	0.0	
Residual	458787.0	8.819902e+09	1.922439e+04	NaN	NaN	 

La p-value obtenue est de 0.0, donc largement inférieure au seuil de 5%. Cela nous conduit à rejeter l'hypothèse nulle (H0), selon laquelle il n'existerait pas de différence significative du temps d'intervention entre les catégories de retard.

Conclusion : le type de retard a un impact statistiquement très significatif sur la durée des interventions des pompiers.

1. Préparation des données

L'objectif de cette partie est la production de visualisations permettant de décrire les facteurs qui influent sur le temps de réponse de la LFB dans une optique d'amélioration de ses interventions. J'ai également comparé les valeurs de la période étudiée aux objectifs par appareil. Commençons donc par les données.

Les fichiers de données brutes contiennent des colonnes qui ne m'intéressent pas. Au lieu de les injecter dans Power Query, j'ai effectué les prétraitements directement sur Python.

- Remplacement des valeurs manquantes et gestion des types
- Sélection des colonnes utiles
- Fusion des deux dataframes
- Sauvegarde du dataframe résultant au format CSV

L'essentiel de ces opérations, notamment le travail sur les valeurs manquantes et les types, a déjà été effectué lors de l'étape d'exploration des données. Néanmoins j'en ai effectué davantage ici et les étapes de fusion et de sauvegarde sont nouvelles.

L'étape de fusion (inner join) a conduit à la suppression d'environ 0,4% des données. En effet, certains incidents figurent uniquement dans le fichier Incident tandis que d'autres ne figurent que dans le fichier Mobilisation. La suppression de tels incidents est logique étant donné que les informations manquantes **ne sont pas remplaçables** par une médiane ou un mode (date par exemple).

Grâce à ces étapes, le fichier importé dans Power Query est propre et moins lourd.

2. Modélisation du dataset

Power Query

Une fois les données importées, il reste à finir la préparation des requêtes, c'est-à-dire

- Renommer les colonnes
- Vérifier les types (corrects ici)
- Créer trois tables de dimension : une pour la description de l'incident, une pour sa localisation, une pour la description du retard

Le dataset se présente finalement ainsi :

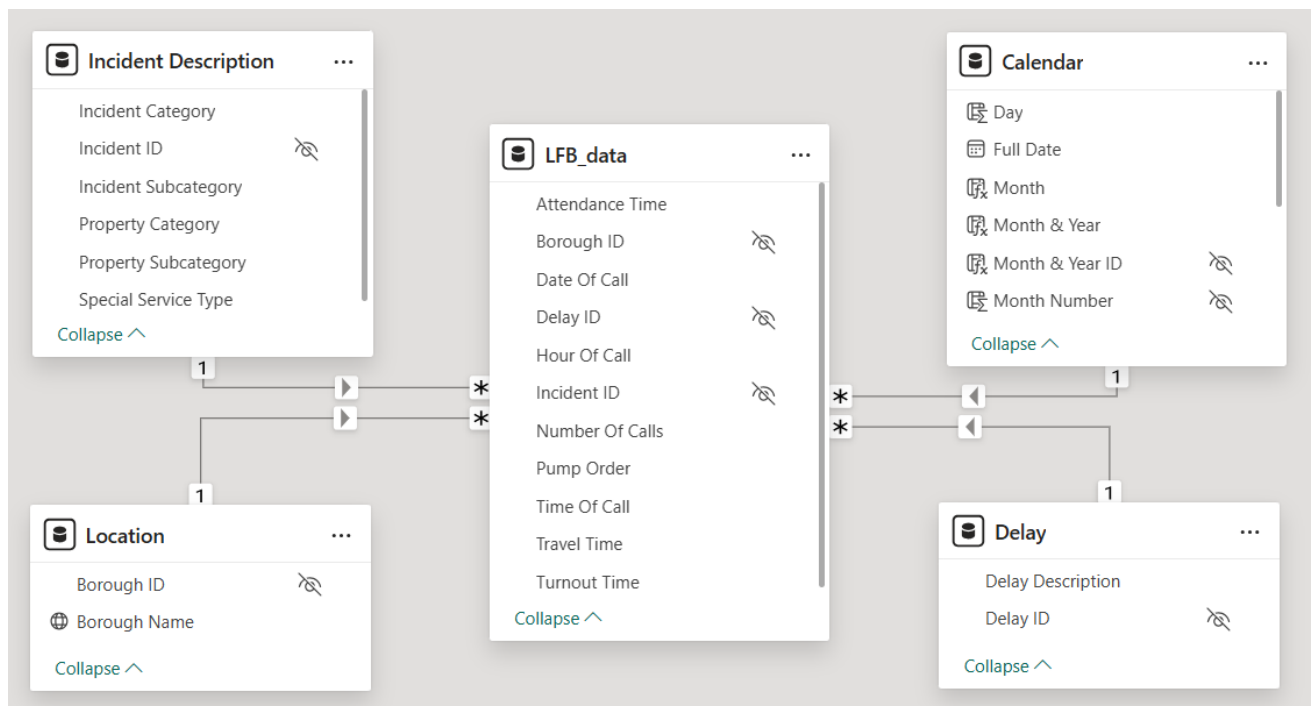
- **Table de faits** contenant les colonnes d'ID, les colonnes contenant les informations de date, et les colonnes de variables numériques
- **Table de dimension incident** contenant l'ID de l'incident et les colonnes de catégorie et sous-catégorie de l'incident, catégorie et sous-catégorie de propriété, types de services spéciaux
- **Table de dimension localisation** contenant l'ID de l'arrondissement et son nom
- **Table de dimension retard** contenant l'ID du retard et sa description

La table de dimension incident reste volumineuse car on considère près de 300 000 incidents. Les tables de dimension localisation et retard ne contiennent que quelques lignes chacune étant donné le peu d'arrondissements et de types de retard différents. La volumétrie du dataset est donc optimale du point de vue de la performance du modèle.

Schéma en étoile

Grâce aux colonnes d'ID, Power BI détecte les relations entre les tables. Il n'y a donc plus qu'à créer une table de dates et la relier à la table de faits via la colonne **Date Of Call**. Il en résulte le schéma en étoile suivant.

Les colonnes d'ID ont été masquées pour ne pas polluer l'interface de rapport.



Mesures

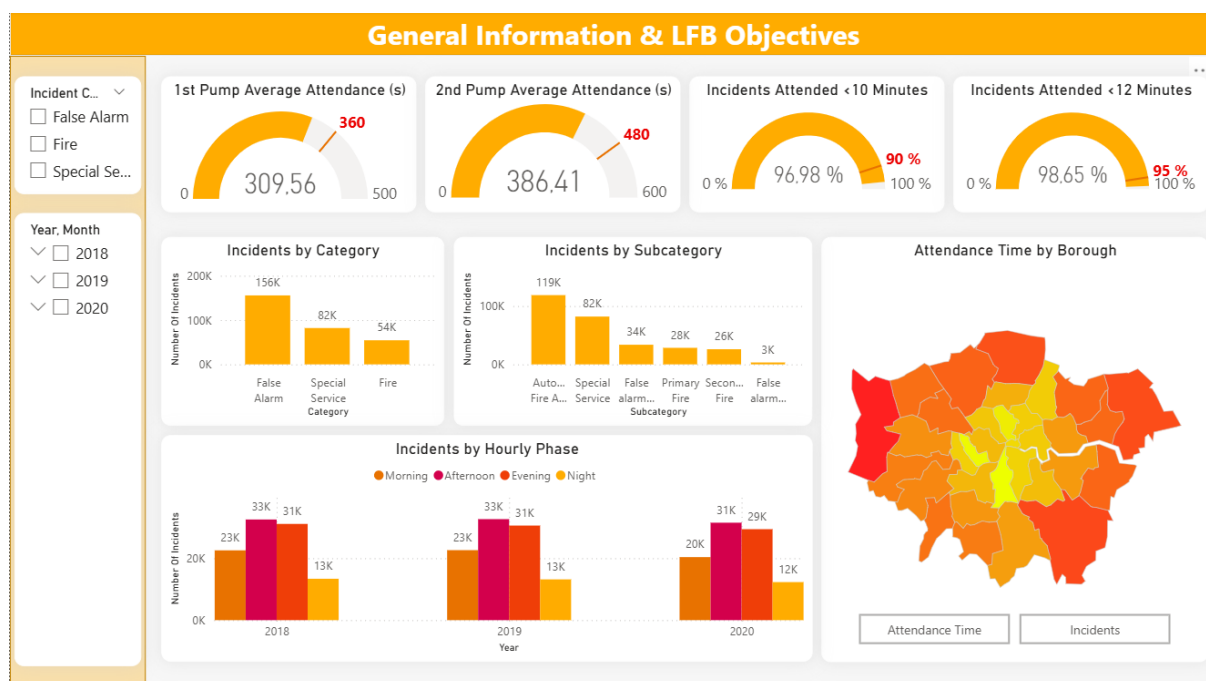
Pour finir, j'ai créé des mesures utiles au rapport :

- Temps moyen de réponse / de trajet / de réaction
- Temps moyen de réponse du premier / deuxième appareil
- Retard moyen lié au trajet
- Nombre total d'incidents
- Nombre moyen d'incidents par jour
- Nombre moyen d'incidents par phase horaire (matin, après-midi, soir, nuit)
- Proportion d'incidents dont la durée d'arrivée d'un appareil est inférieure à 10 minutes
- Proportion d'incidents dont la durée d'arrivée d'un appareil est inférieure à 12 minutes

La mesure principale est le temps moyen de réponse. Le nombre total d'incidents est également utilisé. Les proportions d'incidents seront utiles à l'étude des objectifs.

3. Rapport et visualisations

Le rapport présente l'analyse du temps de réponse de la LFB. Dans une première page, je m'intéresse aux catégories d'incident et à l'atteinte des objectifs.



On retrouve la répartition des incidents par catégorie et sous-catégorie qui illustre le nombre élevé de fausses alarmes et la répartition des feux en deux sous-catégories (primaires et secondaires). Le graphique des incidents par phase horaire montre quant à lui que c'est l'après-midi (entre 12h et 18h) et le soir (entre 18h et minuit) que le plus grand nombre d'incidents se produit.

Rappel des objectifs

Pour représenter l'atteinte des objectifs de la LFB, j'ai utilisé des jauges. On rappelle ces objectifs qui sont la durée moyenne d'arrivée du premier appareil fixée à 6 minutes (ou 360 secondes), la durée moyenne d'arrivée du second appareil à 8 minutes (480 secondes), un appareil sur les lieux en moins de 10 minutes pour 90% des incidents et un appareil sur les lieux en moins de 12 minutes pour 95% des incidents.

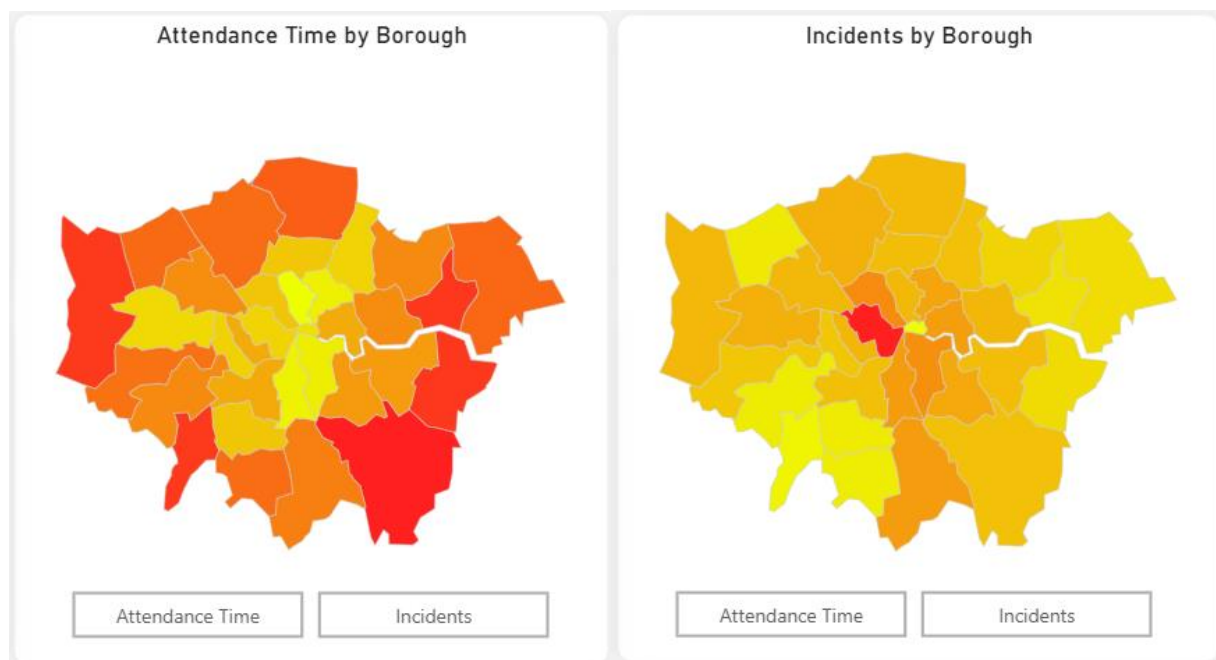
Sur l'ensemble du jeu de données, toutes années et tous arrondissements confondus, **ces objectifs sont respectés**. Cependant, on va voir que selon la catégorie d'incident et l'arrondissement, des différences surviennent.

Carte de formes

J'ai utilisé une carte de formes, ou **shape map**, pour étudier le temps de réponse et le nombre d'incidents par arrondissement (borough). Cette carte n'est pas disponible par défaut dans Power BI, j'ai dû la créer. Pour comprendre son fonctionnement, j'ai consulté cette [vidéo](#) et ce [site](#). La démarche consiste à créer sa propre carte à partir d'un site tel que [celui-ci](#), dans lequel j'ai choisi une **division par district**. La carte montre une division de tout le Royaume-Uni en 380 districts, il a donc fallu sélectionner ceux de Londres et sa périphérie. Une fois la carte téléchargée, il faut la formater de façon à ce qu'elle soit interprétable par Power BI. Le site [mapshaper](#) permet ce travail. Le fichier au format TopoJSON résultant est prêt à être utilisé.

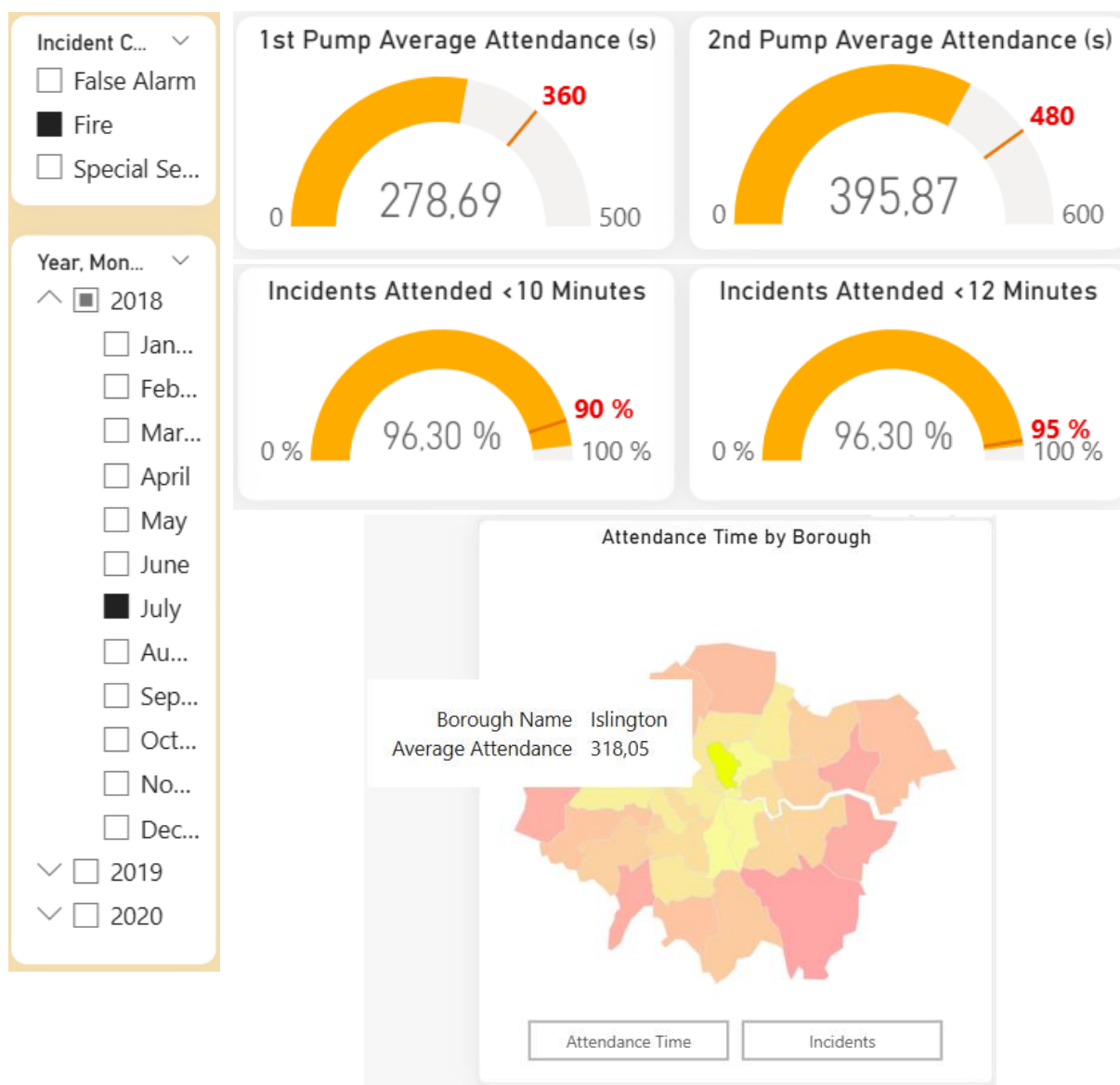
La dernière étape consiste à s'assurer qu'une variable du dataset correspond à l'une des clés de la shape map. Sans ce lien, Power BI est incapable de comprendre quelle section de la carte il faut colorer. Après des ajustements de majuscules avec Python, la variable **Borough Name** concorde avec la clé, ce qui a permis d'obtenir le résultat.

Etude des objectifs



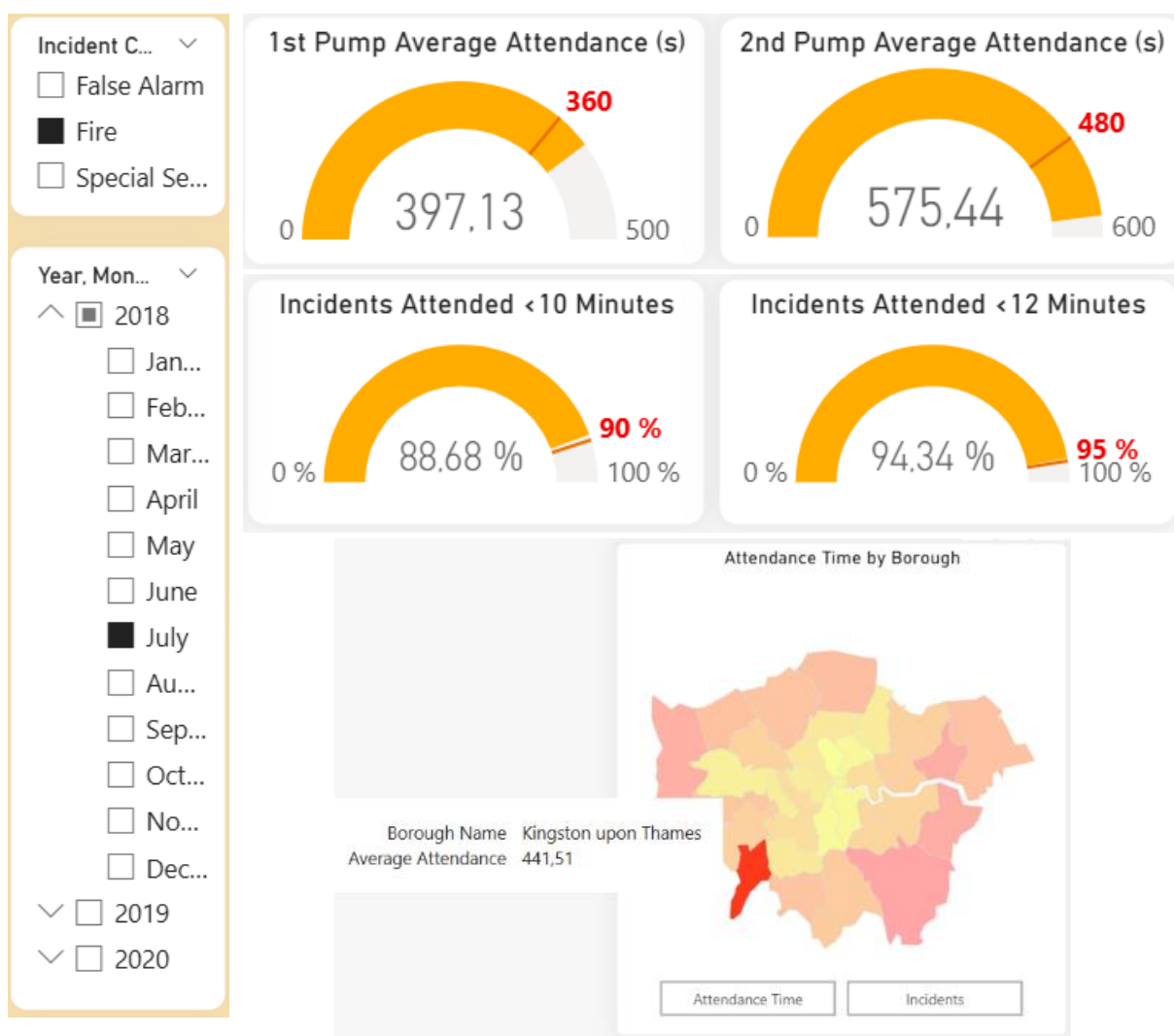
Une première étude visuelle révèle que le temps de réponse est plus faible pour les arrondissements situés au centre de Londres (coloration jaune, première carte) tandis que les arrondissements situés en périphérie ont des temps plus élevés (coloration rouge). Les incidents sont quant à eux plus nombreux dans le centre (coloration rouge, deuxième carte) qu'en périphérie.

Ce premier constat m'a conduit à m'intéresser à certains arrondissements particuliers. Par exemple, en sélectionnant la catégorie Fire pour l'arrondissement d'Islington ayant des temps de réponse bas d'après la carte, l'atteinte des objectifs est la suivante.



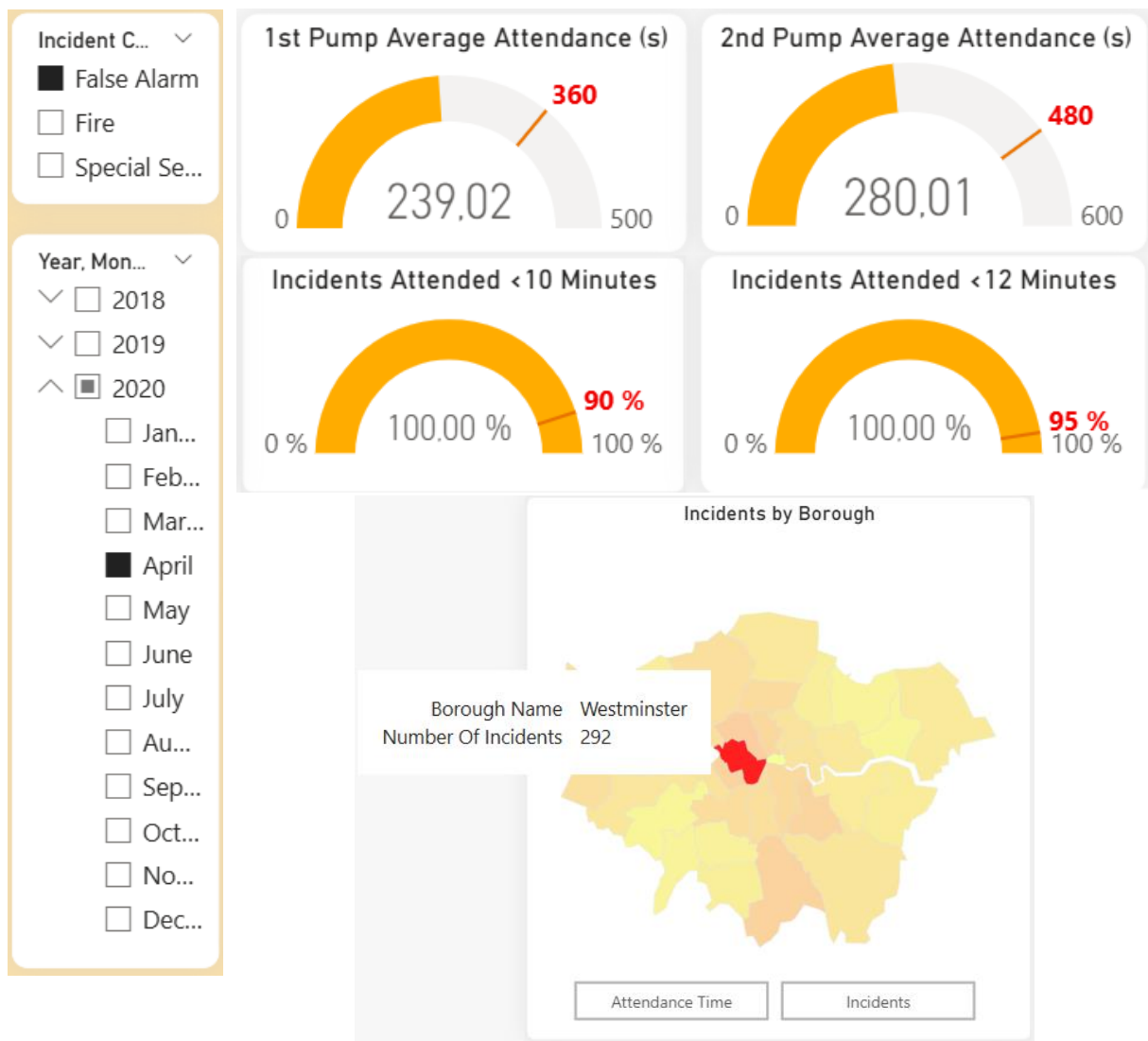
Comme attendu, les temps de réponse de la première pompe et de la deuxième pompe sont inférieurs à la limite. Les objectifs des 90% et 95% d'incidents en moins de dix et douze minutes respectivement sont également remplis. On peut supposer que l'égalité des deux valeurs est due au fait qu'hormis une majorité d'incidents traités rapidement, quelques-uns ont des temps très élevés résultant de situations particulières, difficulté d'accès ou encore erreur dans les enregistrements.

Je vais maintenant comparer ces résultats à un autre arrondissement situé cette fois en périphérie de Londres, Kingston upon Thames, pour lequel la carte suggère que le temps de réponse est très long.



Le temps de réponse du premier appareil dans cet arrondissement approche les 400 secondes et dépasse 575 secondes pour le deuxième appareil. Aucun objectif n'est atteint, ce qui suggère un défaut dans la gestion des incidents de cet arrondissement, ou un manque de moyens, une localisation difficile d'accès, etc.

Pour finir, j'étudie l'arrondissement de Westminster situé dans le centre de Londres qui enregistre le plus grand nombre d'incidents avec plus de 22 000 cas entre 2018 et 2020. Pour la catégorie des False Alarm, qui représente plus de la moitié des incidents, les objectifs pour cet arrondissement sont présentés ci-dessous.

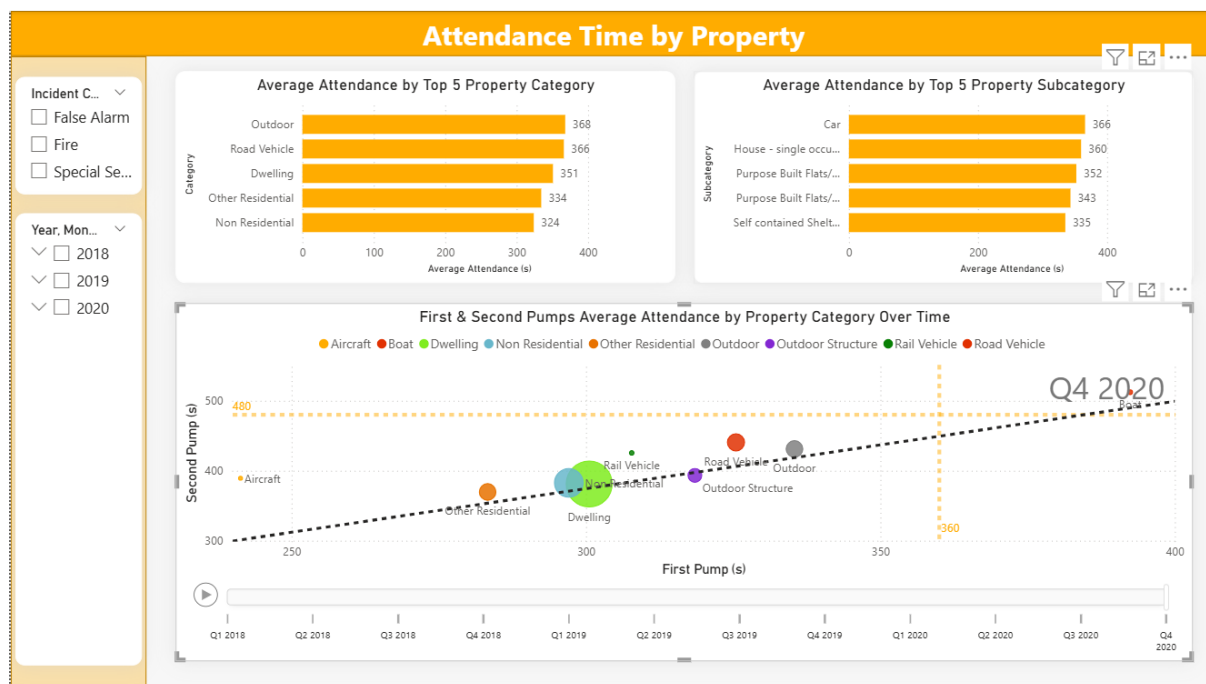


L'atteinte des objectifs dans cet arrondissement est excellente. Les temps de réponse sont très bas et 100% des incidents ont vu l'arrivée d'un appareil en moins de dix minutes, et ce malgré le nombre élevé d'incidents dans cet arrondissement.

Conclusion : cette analyse démontre l'influence de la zone géographique sur le temps de réponse et le fait que le nombre d'incidents n'a pas d'impact décisif. En effet, les arrondissements situés en périphérie de Londres ont généralement des temps plus longs que ceux situés au centre, malgré un nombre d'incidents plus faible.

L'explication derrière ce résultat est la suivante. Londres est divisée en deux grandes parties : **Inner London** constituée des arrondissements centraux, et **Outer London** pour les arrondissements périphériques. Les casernes sont majoritairement situées dans Inner London, ce qui explique l'efficacité des interventions dans Westminster. En revanche, dans Outer London, le nombre réduit de casernes augmente le chemin à parcourir jusqu'aux incidents et ainsi le temps de réponse.

Sur une deuxième page dont on peut voir un aperçu ci-dessous, j'ai étudié l'influence du type de propriété sur le temps de réponse.

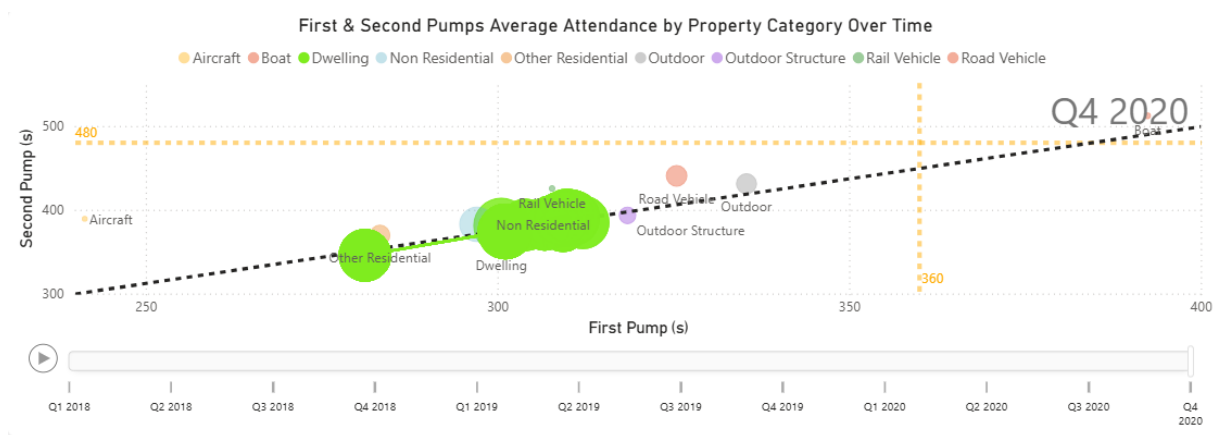


Les bar charts représentent les temps de réponse moyens pour les cinq catégories et sous-catégories de propriété les plus importantes. Les **incidents en extérieur** et impliquant des **véhicules routiers** sont ceux pour lesquels le temps est beaucoup plus long à cause de la difficulté d'accès dans la circulation. Les **résidences** par exemple ont des temps plus bas, y compris les **habitations** qui sont impliquées dans la plupart des incendies. La sous-catégorie **voiture** est celle avec le temps le plus long, suivie de près par les **maisons à un occupant** qui confirment l'enjeu autour des habitations.

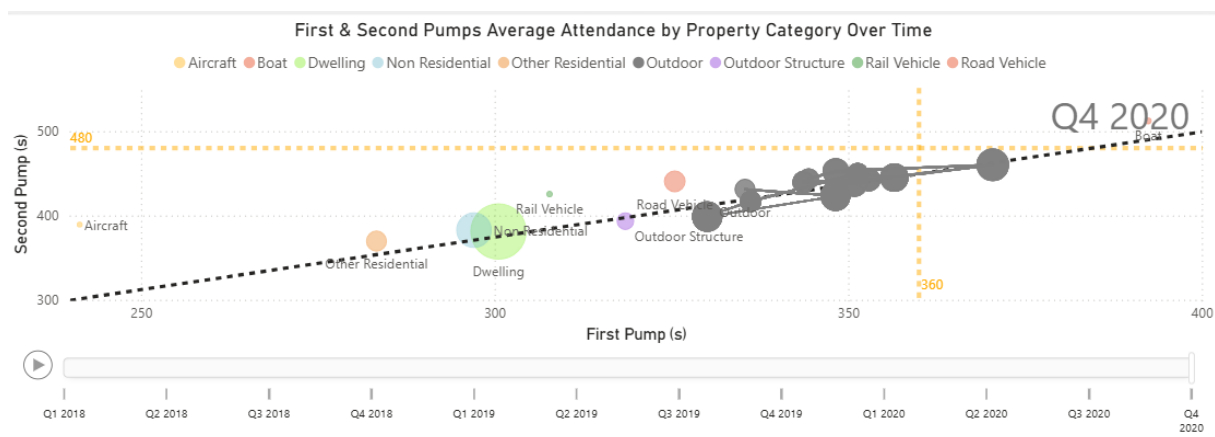
Focus sur les propriétés

L'animation en deuxième moitié de page montre l'évolution du temps de réponse de la première et de la deuxième pompe selon la catégorie de propriété. Plus la bulle est grande, plus il y a d'incidents de ce type. La bulle verte est donc celle des habitations. J'ai également représenté les valeurs des objectifs afin de savoir sur quelles propriétés mettre l'accent.

De façon générale, les bulles se situent dans le rectangle inférieur gauche de la figure, c'est-à-dire là où les objectifs sont remplis. Intéressons-nous maintenant à certaines catégories précises.

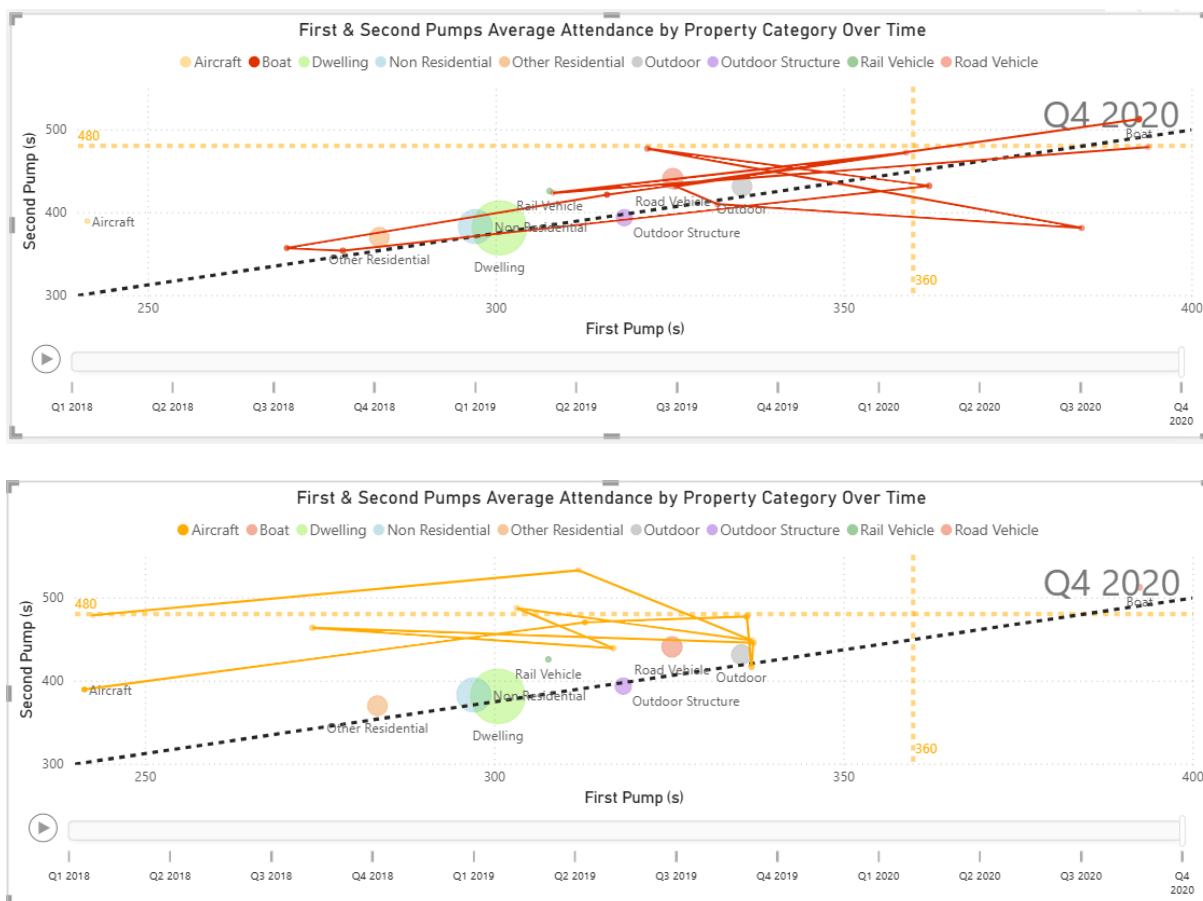


Pour les habitations (**Dwelling**) qui concernent environ la moitié des incidents globaux, le temps de réponse de la première pompe varie autour de 300 secondes et pour la deuxième pompe autour de 400 secondes. Les objectifs sont donc respectés pour la catégorie de propriété la plus importante du point de vue des incidents.



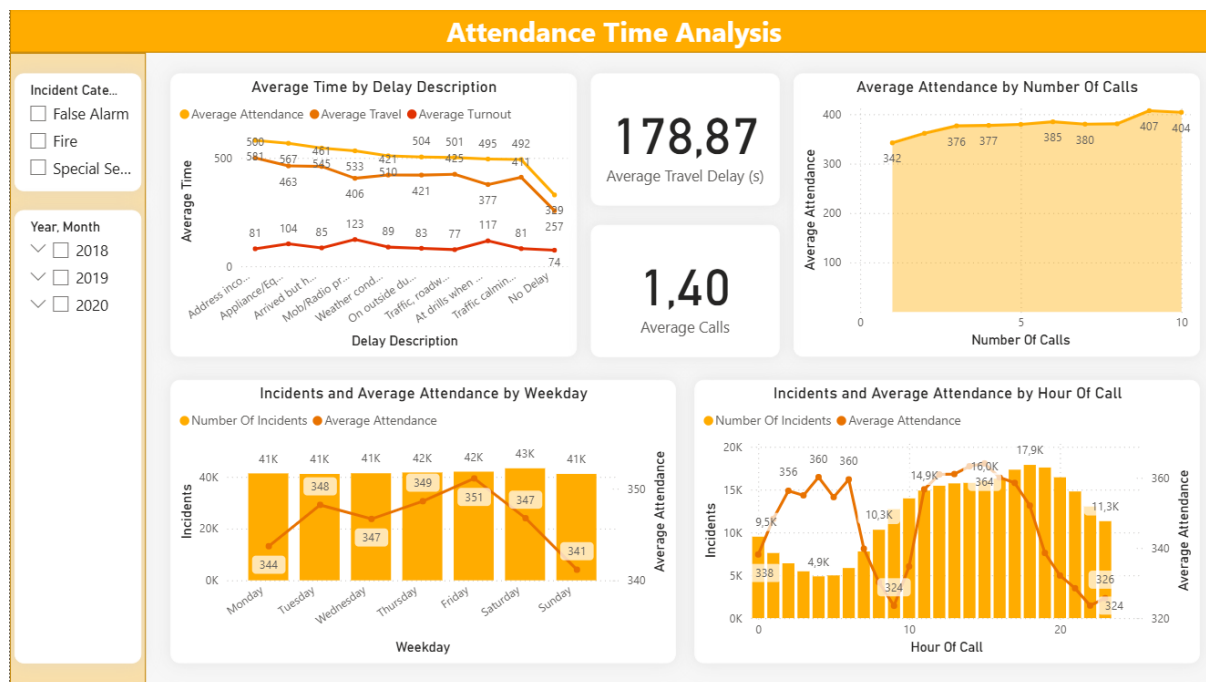
Pour les incidents en extérieur (**Outdoor**), qui représentent environ 10%, on remarque des temps de réponse plus élevés que pour les habitations. Les valeurs sont souvent 10 à 20 secondes en dessous du seuil seulement, et dépassent même ce dernier lors d'un trimestre. Ce sont des situations à surveiller avec plus d'attention.

Enfin les incidents de bateaux (**Boat**) et d'avions (**Aircraft**) montrent quant à eux des temps qui dépassent régulièrement les seuils limite. Malgré que peu d'incidents soient de ces types, il en va parfois de la vie des personnes et il est donc essentiel d'intervenir rapidement dans ces situations également.



Conclusion : les temps de réponse sont sujets à des variations importantes selon le type de propriété concerné. De nouveau, **l'influence du nombre d'incidents est à exclure** puisque les habitations ont des temps bas malgré qu'ils soient impliqués dans une majorité d'incidents. Néanmoins, et malgré que les résultats suivent globalement les objectifs, plusieurs situations restent à améliorer puisque les incidents d'extérieur qui comprennent par exemple des feux de buissons ne sont pas traités dans les délais. La difficulté d'accès ou une localisation imprécise sont des arguments qui peuvent expliquer ce résultat. Concernant les bateaux, la brigade ne possède que deux stations fluviales pour traiter tous les incidents, ce qui complique sa tâche. Les avions sont quant à eux situés en périphérie pour la plupart, et donc sujets à des temps plus longs.

Enfin sur une dernière page, j'ai analysé les autres facteurs qui influent sur le temps de réponse. Un aperçu est disponible ci-dessous.



Etude du retard et des appels

Le premier graphique étudie l'impact des situations de retard, quelles qu'elles soient. La première information est que sur l'ensemble des incidents, le temps de trajet dans les situations de retard est près de **179 secondes** plus long (3 minutes) que lorsqu'il n'y en a pas. La deuxième information est que le temps de préparation est également plus long dans les situations de retard puisqu'il varie entre 77 et 123 secondes contre 74 secondes en moyenne lorsqu'il n'y a pas de retard. Enfin on note que les types de retard les plus pénalisants sont les **adresses incomplètes ou erronées** ainsi que les **défauts d'équipement**. Pour ces types, le temps moyen de réponse est de 581 et 567 secondes respectivement alors qu'il varie entre 492 et 545 secondes pour les autres types de retard.

Conclusion : les situations de retard impactent fortement le temps de réponse et sont donc à éviter au maximum, dans la mesure du possible. Le retard provient d'un **temps de trajet** plus long en majorité, mais la **préparation** peut aussi être plus rapide afin de gagner des secondes précieuses. Le plus important est de se focaliser sur les grands retards puisque ce sont des situations améliorables. D'une part, un nouveau système de localisation entrera en vigueur en cette année 2025 afin de permettre aux appareils de se rendre au bon endroit grâce à la géolocalisation, d'autre part les équipements doivent être renforcés afin de connaître moins de pannes ou de problèmes affectant la rapidité d'arrivée sur place.

Le nombre d'**appels** peut-il avoir une influence sur les temps de réponse ? On observe que plus le nombre d'appels augmente, plus le temps de réponse augmente. En effet, lorsqu'il n'y a qu'un appel, le temps moyen est de 342 secondes alors qu'à dix appels, on dépasse les 400 secondes. Certains incidents enregistrent plus de dix appels, mais cela ne concerne que 0,5% des cas et donc des situations rares pouvant conduire à des temps très longs comme très courts. Ils ne sont pas représentés pour cette raison. Enfin, les feux reçoivent en moyenne 2,43 appels tandis que les fausses alarmes et les services spéciaux n'ont en moyenne que 1,15 appel !

Conclusion : lorsque le nombre d'appels est bas, l'intervention est souvent rapide. Une hausse du nombre d'appels peut signifier qu'un appareil n'est pas arrivé sur place ou que l'incident est visible par un grand nombre de personnes comme les feux. Ce résultat montre que le nombre d'appels traduit une certaine efficacité d'intervention mais reste un facteur à considérer avec précaution.

Etude du jour et de l'heure

Pour finir, j'ai comparé le temps de réponse par **jour de la semaine** et **heure de la journée**. Par jour de la semaine, le nombre d'incidents est équivalent et on note que les interventions sont plus rapides le **dimanche** et le **lundi**. Les autres jours, il faut compter cinq à dix secondes supplémentaires en moyenne.

Concernant l'heure de la journée, les incidents sont concentrés entre 11 heures et 21 heures. La plage de **2 heures à 6 heures** est celle ayant le moins d'incidents, pourtant, les temps sont plus longs à cette période avec 360 secondes. Entre **11 heures et 17 heures**, les temps sont également proches des 6 minutes. Enfin, autour de **9 heures**, les incidents sont traités plus rapidement, en 324 secondes en moyenne.

Conclusion : le jour de la semaine ne montre pas d'impact très significatif sur le temps de réponse. Les incidents étant équitablement répartis dans la semaine, on compare aisément les jours entre eux. La baisse des temps le dernier jour de la semaine et le premier jour de la semaine suivante montre une circulation plus fluide due à des comportements de population qui sortent moins en fin de weekend.

Concernant l'heure, l'augmentation des temps pendant la nuit démontre des conditions de visibilité moins bonnes obligeant les pompiers à rouler moins vite. Les pics matinaux et en fin de journée témoignent quant à eux des pics de circulation pour se rendre au travail. L'heure a donc un impact beaucoup plus important sur le temps de réponse que le jour, ce qui suggère que la préparation des pompiers doit être encore plus optimale à certains moments de la journée pour intervenir efficacement.

CONCLUSION

L'étude des temps de réponse de la LFB a permis de montrer que la catégorie des incidents ainsi que leur nombre avaient une influence modérée contrairement à ce que l'on aurait pu penser. La **localisation** de l'incident, en particulier si l'arrondissement se situe au centre ou dans la périphérie de Londres, a un impact bien plus grand en raison de la répartition inhomogène des casernes dans Londres.

Les objectifs fixés par la LFB sont plutôt respectés, mais on a identifié des situations, à nouveau géographiques ou selon le type de propriété concerné par l'incident, pour lesquelles les temps de réponse dépassaient ceux-ci, parfois largement. Les difficultés d'accès aux lieux des incidents, liées à la circulation et au moment de la journée, traduisent également la variabilité du temps de réponse. Enfin, on a mesuré l'impact des **retards** sur le temps de trajet ainsi que l'importance d'avoir des équipements opérationnels et des **adresses complètes** et correctes.

Des mesures palliatives sont en cours d'introduction afin de réduire les temps de réponse, notamment l'introduction d'un nouveau système de localisation pour les adresses, et des efforts sont à faire sur les périodes nocturnes durant lesquelles la mauvaise visibilité impacte la durée d'arrivée des appareils. D'autres mesures telles que l'amélioration de la fiabilité des équipements sont aussi à prévoir.

Dans ce projet, les données source constituaient une première limite puisque plusieurs variables étaient complexes ou contenaient uniquement des codes dont nous n'avons pas pu trouver la signification. Nous avons essayé de contacter la LFB afin d'obtenir des éclaircissements mais sommes hélas restés sans réponse. Par ailleurs, les données géographiques étaient manquantes pour les habitations (Dwellings) pour des raisons de confidentialité, ce qui concerne une bonne moitié des incidents. Il n'a donc pas été possible d'étudier la distance entre caserne et lieu de l'incident, ni d'utiliser la latitude et la longitude comme variables. Afin de garantir la fiabilité des analyses, nous avons été contraints de supprimer un grand nombre de variables pour ces raisons.

Les axes étudiés dans ce projet restent pertinents et ont permis d'éclairer plusieurs facteurs ayant un fort impact sur le temps de réponse ainsi que quelques mesures qui seront prises ou pourraient être prises afin d'améliorer l'efficacité des interventions. Un complément serait l'application de modèles de prévision sur le temps de réponse avec comme variables explicatives la date et l'heure de l'incident, la localisation, la catégorie et / ou le type de propriété concerné. D'autres variables non présentes ici comme les conditions météorologiques seraient intéressantes à inclure, mais il ne faut pas oublier que l'on parle de temps moyen et de phénomènes liés à de nombreux facteurs, dont certains ne sont pas prévisibles.