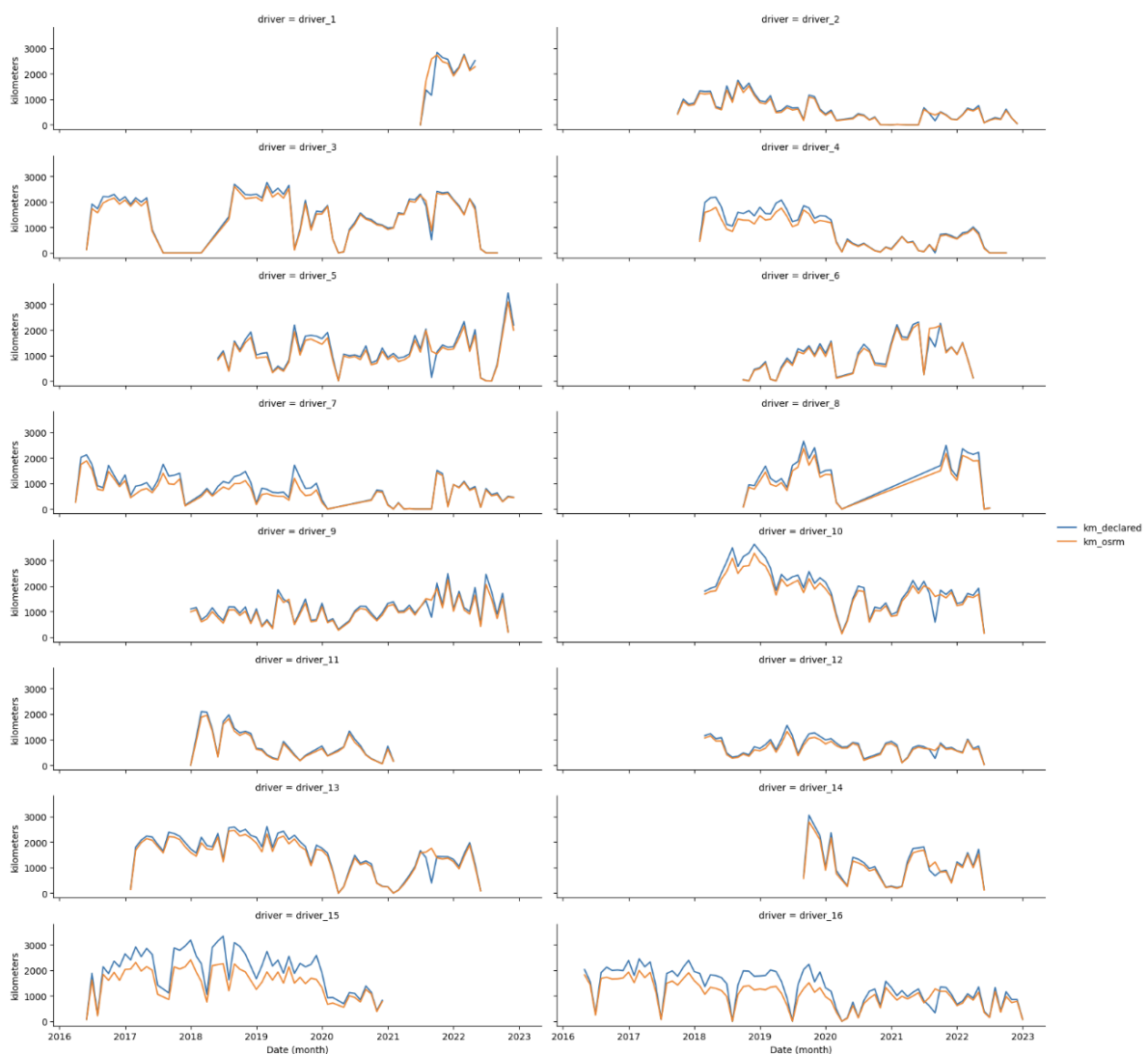


Résumé de la qualité des données kilométriques Uber

Le projet Digipower Academy, mené par l'association PersonalData.IO et l'entreprise hestia.ai, a mobilisé environ 100 chauffeurs Uber de la région genevoise pour récupérer leurs données personnelles de la plateforme, et mieux évaluer la qualité des données fournies par Uber. Nous avons pu analyser un échantillon d'environ 30 personnes.



Comparaison des données kilométriques de 16 chauffeurs Uber genevois, sommées par trajet, à partir des données récupérées par la "Méthode 2". En bleu les données déclarées

comme base de la compensation, en orange les données recalculées par un algorithme de routage (voir plus bas). Lien vers l'original ci-dessous.

Méthodologie

Nous confrontons les données obtenues par les chauffeurs via plusieurs méthodes:

- méthode validée par l'OCIRT (méthode "Etat")
- méthode d'accès aux données dans l'application mobile Uber ("méthode 1")
- méthode d'accès aux données, et portabilité RGPD ("méthode 2")
- relevés fiscaux mensuels sur le portail web Uber (auxquels nous n'avons pas accès, mais qui sont visibles pour les chauffeurs - ceux-ci sont ce qui s'approche le plus d'une fiche de paye)

A noter que dans tous les cas, la source des données est en fait Uber.

Ce travail a été mené dans l'extrême urgence face aux délais irréalistes et unilatéraux imposés par Uber aux chauffeurs pour décider. Ces données ne donnent donc qu'une vue parcellaire propre à chaque chauffeur, mais qui peut être confrontée entre chauffeurs pour identifier certaines tendances. Plusieurs tendances indiquent des erreurs dans les données récupérées par les chauffeurs, dont certaines ont des conséquences à plusieurs niveaux. Par exemple, certaines de ces données pourraient indiquer non seulement une erreur de calcul pour un chauffeur particulier, mais pourrait aussi indiquer des subtilités qui n'auraient pas été considérées par l'OCIRT dans la définition du taux correctif de 25.96 centimes/kilomètre. Ces effets agissent plusieurs fois, parfois systématiquement à l'encontre du chauffeur (erreur de validation par l'Etat de la formule, erreur dans les données propre à un chauffeur, erreur dans le raisonnement du chauffeur qui doit balancer différentes méthodes pour défendre ses intérêts, impact sur la collectivité dans le calcul des charges sociales,...).

Chacune des méthodes d'accès aux données mène à un calcul propre des kilomètres parcourus en course avec client, sur une unité de temps précise (trajet, mois, etc). Nous nous sommes concentrés là-dessus parce que l'OCIRT s'est basé sur ces kilomètres en course ("P3") pour raffiner et définir une formule de calcul d'arriérés (calcul qui inexplicablement se base seulement sur les kilomètres depuis octobre 2019, alors que la rétroactivité est normalement de 5 ans).

Les différents calculs possibles sont les suivantes:

- Les kilomètres "etat" sont fourni mensuellement via la méthode du même nom

- Les kilomètres “declared” sont fourni trajet par trajet par les méthodes 1 et 2, et peuvent donc être agrégées mensuellement
- Les kilomètres “osrm” sont recalculés à partir des points de prise et de dépose du client, en conjonction avec l’outil de routage de course OSRM (Open Source Routing Machine, <http://project-osrm.org/>). Nous avons configuré cet outil de telle sorte qu’il nous rende **les distances des trajets les plus rapides**, basé sur une connaissance des routes et règles de circulation de France et Suisse, indépendamment du trafic **et des frontières entre pays et cantons**.

Résultats

Ces calculs des kilomètres nous permettent donc de mener à ces comparaisons en paires “etat”-“declared”, “declared”-“osrm”, et enfin ”etat”-“osrm” .



Comparaison mois par mois des différences “état”-“declared”, “declared”-“osrm”, “état”-“osrm”. Lien vers l’original ci-dessous.

Ce qui ressort de ces graphes est une concordance générale entre les différentes techniques de calcul des kilomètres.

Cependant, 4 problèmes majeurs ont été identifiés à ce stade:

1. On voit que dans le cas d’au moins un chauffeur les données “etat” sont complètement divergentes. Chaque acteur a reçu donc une version différente des informations de la part d’Uber.
2. Cette concordance entre les courbes diverge nettement en août, septembre et octobre 2021. Après vérification avec les relevés fiscaux pour certains chauffeurs, il

semblerait que ceux-ci soient plus élevés que les kilomètres “etat” pour beaucoup de chauffeurs.

3. Une concordance entre les kilomètres “osrm” et les kilomètres “declared”, ainsi que la concordance entre les kilomètres “declared” et les kilomètres “état” devrait nous étonner: avoir les deux concordances en même temps nous paraît impossible. En effet les kilomètres intervenant dans l'accord validé par l'État concernent les kilomètres *parcourus*. Or nos kilomètres “osrm” sont des kilomètres *prédits*, et cette concordance “etat”-“osrm” nous paraît donc irréaliste. Les prédictions ne confirment pas le chemin effectivement parcouru par le chauffeur, ce qui suggère qu'Uber n'a pas transmis toutes les données nécessaires pour faire un calcul juste.
4. Uber a confirmé au syndicat SIT ne se baser que sur les kilomètres parcourus sur le canton de Genève en excluant ainsi la correction d'arriérés pour les trajets parcourus hors frontières cantonales. Nous soutenons le syndicat SIT qui affirme que si tel est le cas, Uber fait une lecture erronée de l'accord. Nous invitons, tout comme le syndicat SIT, à exiger de Uber un report de l'échéance de décision pour les chauffeurs, et à considérer favorablement la multiplicité des situations impossibles dans lesquels les chauffeurs se retrouvent (agonie de décision pendant le weekend, recommandé en transit ce mardi 31.01.2023, situation extrêmement confuse et manque de soutien juridique avec une expertise en analyse des données). **Cependant**, l'explication fournie par Uber du décomptage exclusif des kilomètres genevois nous paraît elle-même erronée: si tel était le cas, nous ne devrions pas trouver une telle concordance avec notre méthode de calcul, qui est aveugle sur les frontières entre pays et cantons, ainsi qu'aux travaux et autres détours dûs à des encombrements.

Additionnellement, nous nous sommes engagés dans la même démarche sur Paris avec plusieurs chauffeurs. Nos calculs préliminaires semblent indiquer que certaines de ces incohérences sont de fait présentes dans plusieurs pays en Europe.

Nous avons encore d'autres éléments en notre possession à ce stade, dont nous affinons l'analyse.

Conclusion

Nous invitons les autorités genevoises à:

- clarifier auprès d'Uber les données qui ont été fournies aux acteurs genevois à différents stades : maintenant aux chauffeurs, mais aussi plus tôt à l'OCIRT ou encore aux parties prenantes de la conciliation.
- retarder le délai de décision fixé au 31.01.2023.

Nous invitons tout acteur pouvant agir comme expert indépendant à valider nos méthodes.

Nous invitons tout avocat ou syndicat cherchant à représenter les chauffeurs à nous contacter, et à ne pas faire confiance de manière aveugle aux données fournies par Uber.

Contact:

- Paul-Olivier Dehaye, membre de l'association PersonalData.IO et directeur de hestia.ai paulolivier@hestia.ai, +41 76 407 57 96
- Jessica Pidoux, directrice de l'association PersonalData.IO, jessica@personaldata.io, +41 79 103 83 18

Originaux:

<https://tinyurl.com/tm9pheyf>

<https://tinyurl.com/2p8un4a6>