



100% Data Science – 16 novembre 2018

Telematics : Du Big Data au Business Model



Olivier
CLAEYS
Groupama



Rémi
BELLINA
Milliman



Antoine
LY
Milliman

**100% ACTUAIRES &
100% DATA SCIENCE**

INSTITUT DES
ACTUAIRES



16 Novembre 2018
Hôtel Marriott Rive Gauche
Paris 14ème



Groupama



Milliman

Atelier C

Telematics : Du Big Data au Business Model

Olivier CLAEYS, **Groupama**
Rémi BELLINA, Antoine LY, **Milliman**

100% Data Science, 16 novembre 2018

De quoi parlons-nous? Qu'est ce que le « **telematics** »?



Introduction

Les concepts

▪ Data

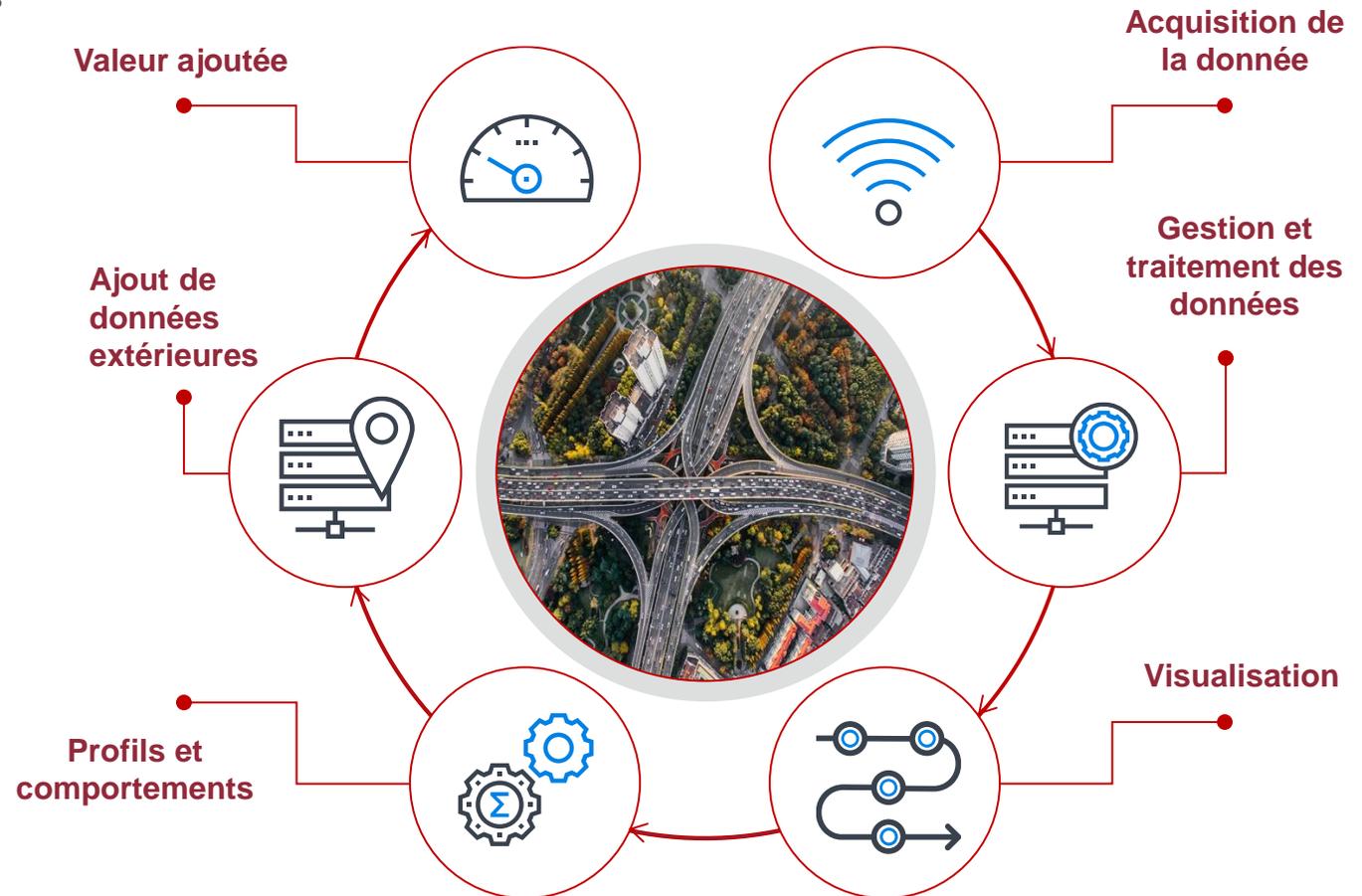
- Coordonnées GPS horodatées
- Accéléromètre
- Gyroscope
- Evènements
- Etc.

▪ Technologies

- BlackBox
- Dongle
- Smartphone, beacon
- Etc.

▪ Questions

- Où?
- Quand?
- Comment?
- Pourquoi et contexte?



Introduction

Les ambitions

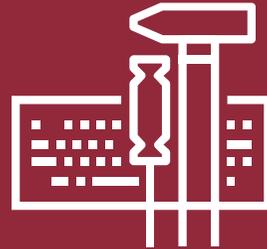
- Différents **mots-clés** reviennent :
 - Black Box Insurance
 - Usage Based Insurance (UBI)
 - Pay As You Drive (PAYD)
 - Pay How You Drive (PHYD)
 - Pay When You Go
 - Driving Behavior Data (DBD)
 - Tracking/geofencing
 - Comportement de conduite
 - Score de conduite
 - Assistance
 - Partage de véhicule
 - Gestion de flotte
 - Etc.
- Dans un environnement **disruptif**
 - **Assureurs** : augmenter la rentabilité, réduire les coûts, gagner des parts de marché, anticiper les besoins des clients, mieux comprendre les risques sous-jacents, etc.
 - **Assurés** : en attente de relations et d'offres personnalisées, de réactivité, de transparence, d'un tarif juste, de services de qualité, etc.
 - Avec de nouveaux **acteurs** : services VTC, partage de véhicules, équipementiers, etc.

- Il y a potentiellement 3 **niveaux** :

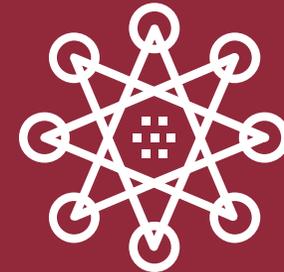
Conducteur



Véhicule



Infrastructures



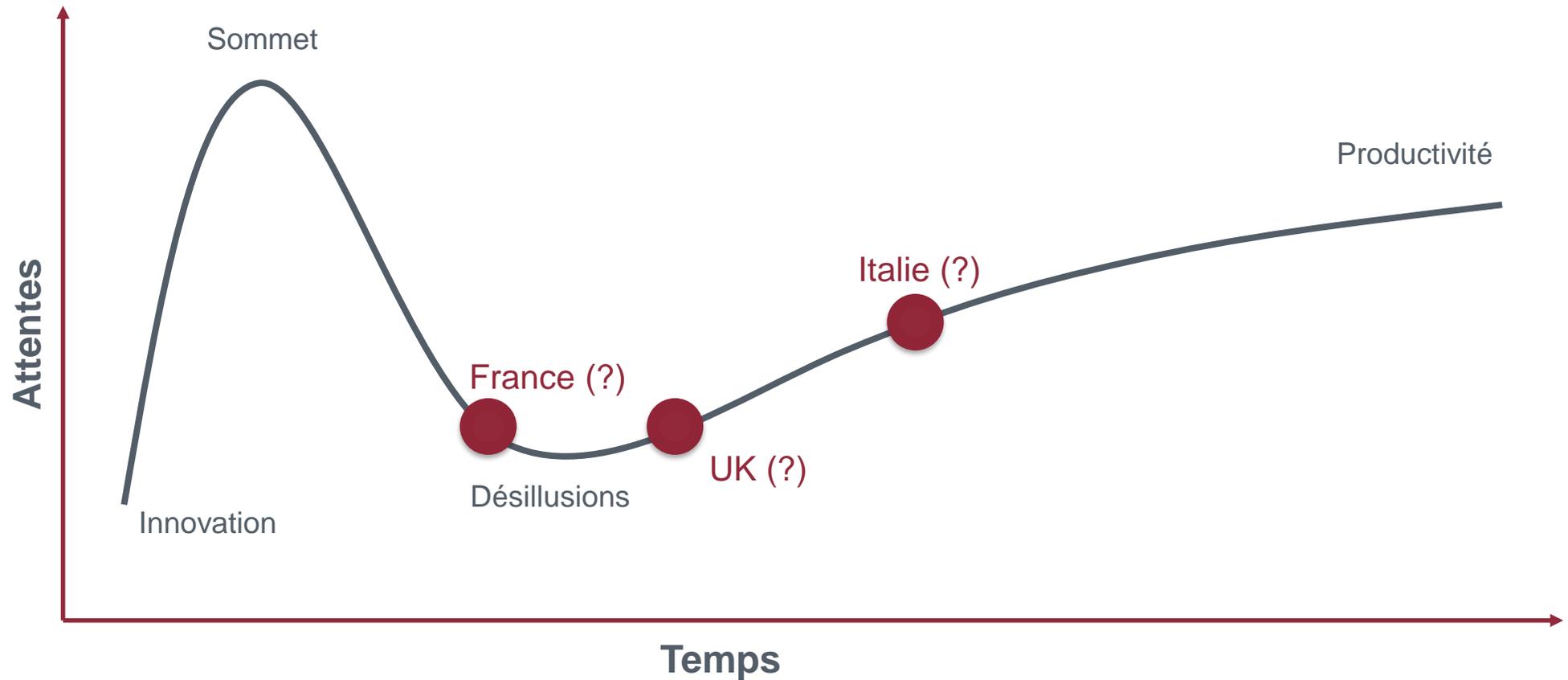
Le **marché** Telematics : L'**Italie** est-elle un bon élève? Où en sommes-nous sur le marché **français**?



Le marché Telematics

Une vision générale de l'Europe

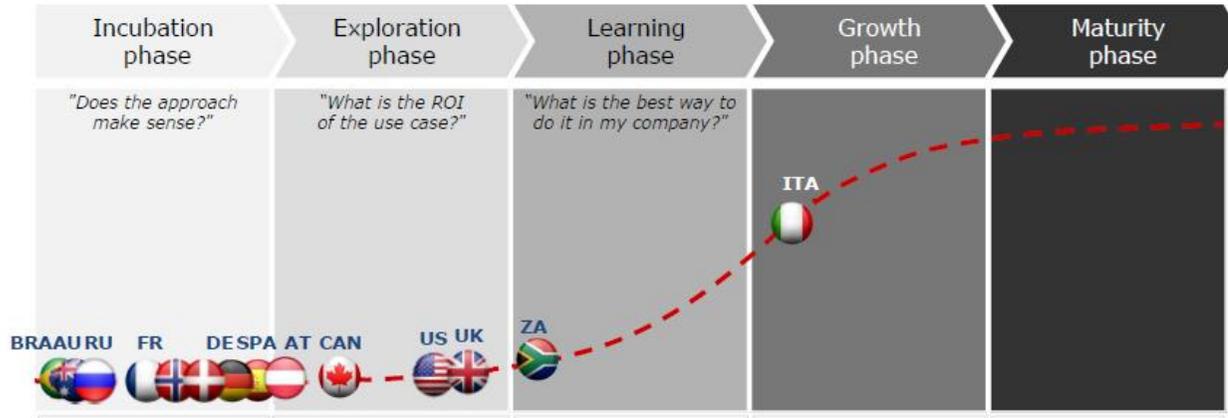
- Gartner Hype Cycle : Où placer la **France**, l'**Italie**, le **UK**?



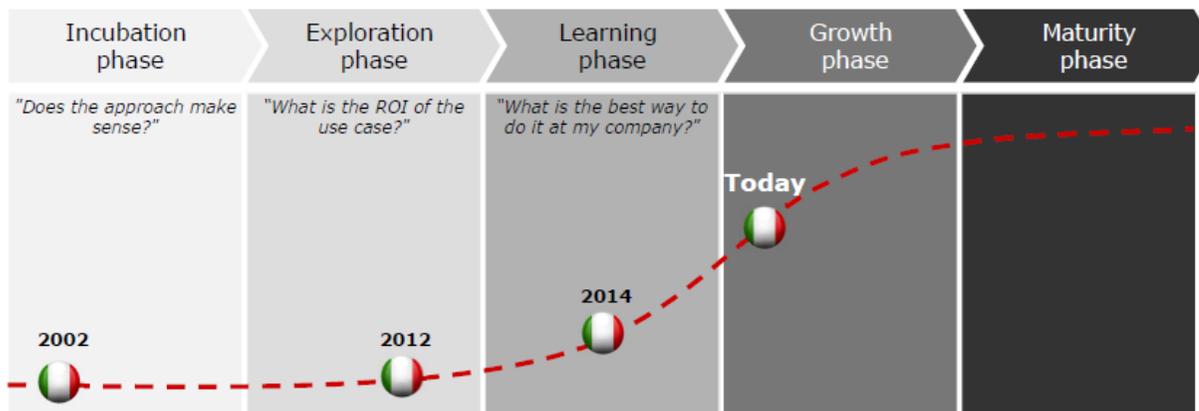
Le marché Telematics

Les spécificités du marché italien (1/2)

- **L'Italie** : Marché aujourd'hui **leader** sur la télématic automobile



- Retour sur la chronologie de ce développement italien :



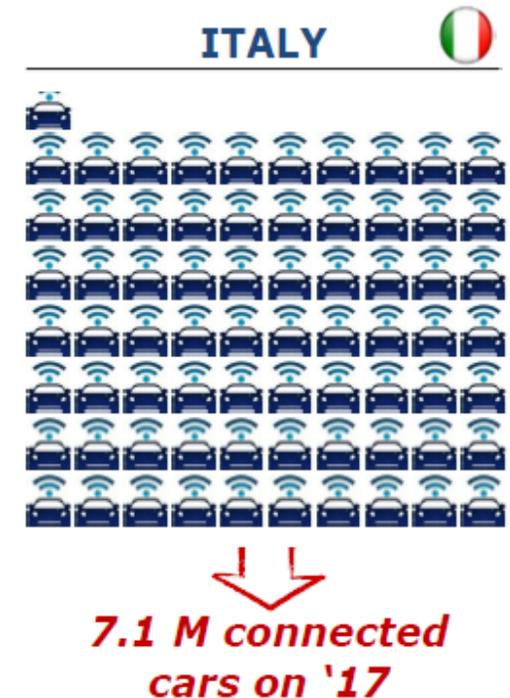
En nombre de voitures connectées, l'Italie est passée de 1,3 M véhicules en 2012 à 7,1 M aujourd'hui.

Source : IoT Insurance Observatory Estimates

Le marché Telematics

Les spécificités du marché italien (2/2)

- Les grandes **caractéristiques** du marché italien
 - **Premier marché** avec 7,1 M véhicules connectés
 - les deux marchés suivants étant : Les Etats-Unis avec 4,4 M et le Royaume-Uni avec 1,1 M
 - Taux de pénétration 17% du parc
 - avec des taux variant généralement autour de 40 à 60% de la production nouvelle
 - Type de « **device** » utilisé :
 - Installation par un professionnel : 60%
 - Auto-installation : 40%
 - Récente introduction des solutions à base de smartphone
 - **Usages** :
 - **Services** → Oui dans 100% des cas (a minima « Find my car »)
 - **Sélection du risque** → Oui de plus en plus généralisé au renouvellement
 - **Liquidation des sinistres** → Oui diffusion croissante de l'apport de la télématique (FNOL ; Coûts d'ouverture ; Fraude ; etc...)
 - **Prévention** → En cours de développement



Le marché Telematics

Le marché français (1/3)

- Où en sommes-nous sur le **marché français**?
 - Un business model difficile à trouver ou non assumé
 - Des technologies parfois jugées contraignantes (prix, installation, autonomie) ou même invasives
 - Des ROI encore mal identifiés
 - Une frontière floue entre tarification et surveillance

- Des **signaux forts** cependant :
 - Des solutions, même exploratoires, mises en œuvre
 - Des groupes de travail plus généraux sur la mobilité
 - Des coûts de partage et de stockage en baisse
 - RGPD : accent mis sur la transparence et le « droit » des utilisateurs à faire valoir leurs données

- Une **prise de conscience** générale :
 - L'écart entre, la situation d'hier et les voitures autonomes de demain, est grand
 - De nouveaux acteurs font leur apparition
 - Une volonté de « dépolssiérer » les offres

Le marché Telematics

Le marché français (2/3)

- Des **mutations** sont à prévoir, centrées autour de la **data**.
 - **UX/UI**
 - **Services**
 - **Sinistres**
 - **Tarifification et surveillance**
 - **Evolution de statut et d'image**
- Le **telematics** est l'un des moyens pour enclencher ces **changements**.
- **Etat du marché** :
 - Toutes les compagnies d'assurance automobile françaises du top 10 travaillent sur leur solution télématique !
 - De multiples acteurs s'emparent de ces sujets : Fournisseurs de services télématiques, Constructeurs automobiles, startups, etc.
 - Une évolution significative des coûts ces dernières années/mois

Le marché Telematics

Le marché français (3/3)

- Focus sur les **technologies**, ce que l'on entend :

BlackBox

« coût d'installation »
« pas de fraude »

Dongle OBD

« difficulté
d'installation »
« les constructeurs ne
suivent pas »
« indépendance »

Smartphone

« batterie »
« fraude? »
« universalité »

- Une fois les données acquises, il faut les **traiter**.

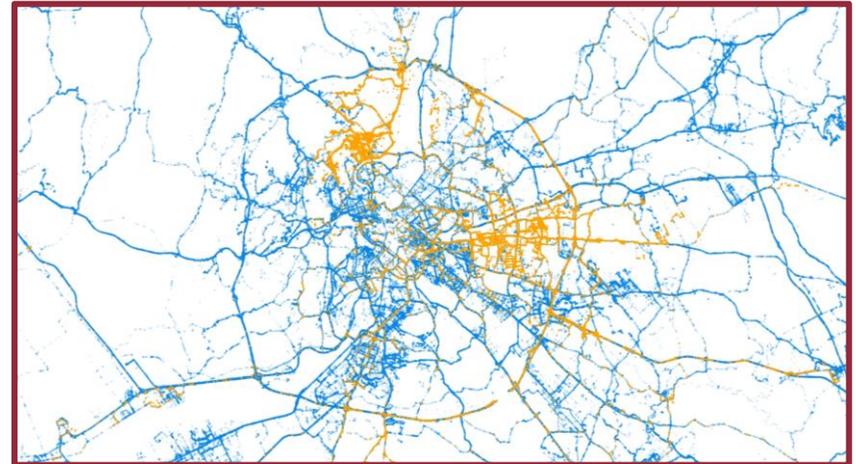
Retour d'**expérience** sur la **data** :
Quelles sont les **difficultés** rencontrées?
Quel est l'apport de la **Data Science**?



Retour d'expérience sur la data

La volumétrie des données

- **Volumétrie** : Cela soulève de nombreuses questions et enjeux, qu'il faut savoir **résoudre**.



GPS VS Accéléromètre : fréquence d'acquisition différente

Conducteurs X Trajets X Points

Optimiser les coûts :

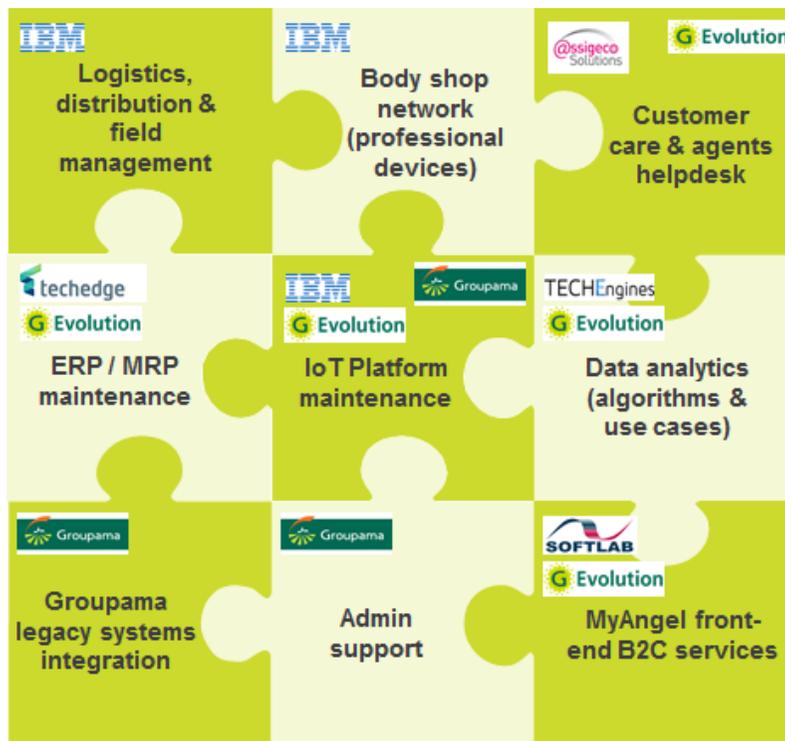
- infra locale ou sur un environnement cloud?

- Choix stratégiques sur la maintenance, la robustesse et l'intégration des services

Retour d'expérience sur la data

Un écosystème général

- La réussite d'un programme télématique passe par un **écosystème** complet et maîtrisé.
- Groupama Assicurazioni en Italie a développé une infrastructure 100 % dédiée à ces problématiques :
 - ✓ Choix, fourniture et **logistique du boîtier télématique** : OBD, Box batterie, Box châssis, etc...
 - ✓ **Plateforme Big Data** en partenariat avec IBM → Device agnostic ; Scalabilité ; RGPD compliant ; etc...
 - ✓ Moteur intégré de **Big Analytics** → codes open source et gestion en temps réel
 - ✓ Application **smartphone** pour communiquer avec le client



G-Evolution est la société de services télématiques de Groupama en Italie.

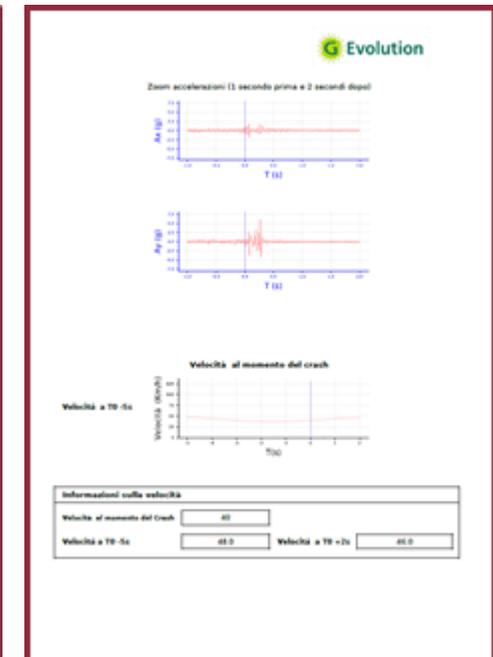
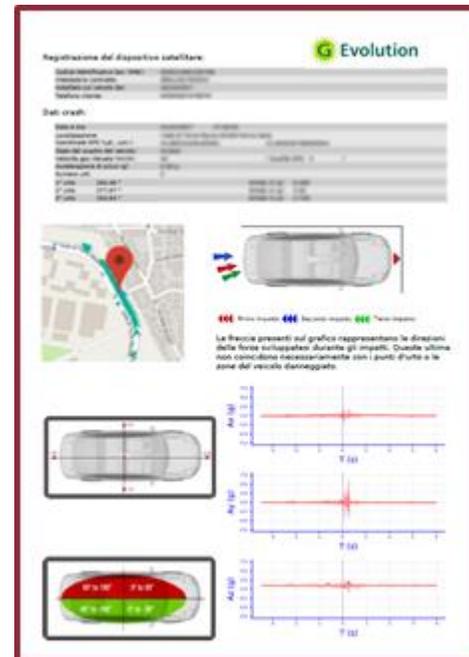
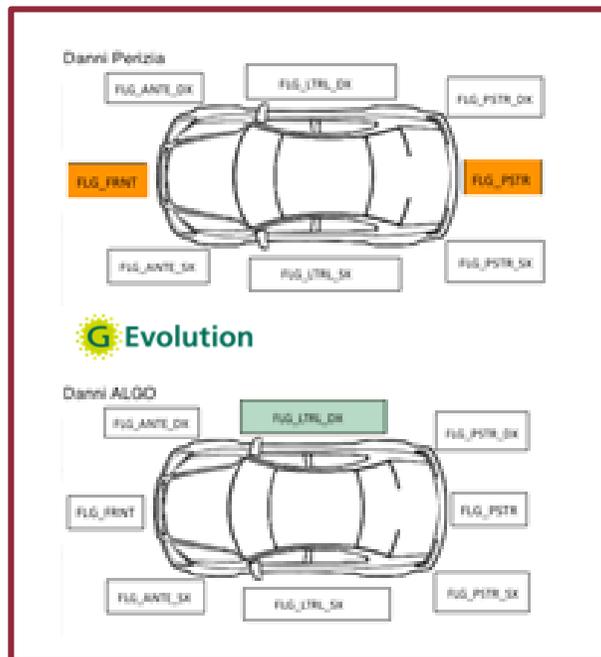
Elle a vocation à travailler en B2B avec Groupama Assicurazioni, les autres filiales du Groupe mais également avec toutes autres entités hors de Groupama.

1 – Vérification du point d'impact entre déclaration et données de la Box

2 – Sophistication du « Crash Report » et de la reconstruction du choc

→ Utilisation des données « Crash »

→ Utilisation des données « Crash » & « Trip »

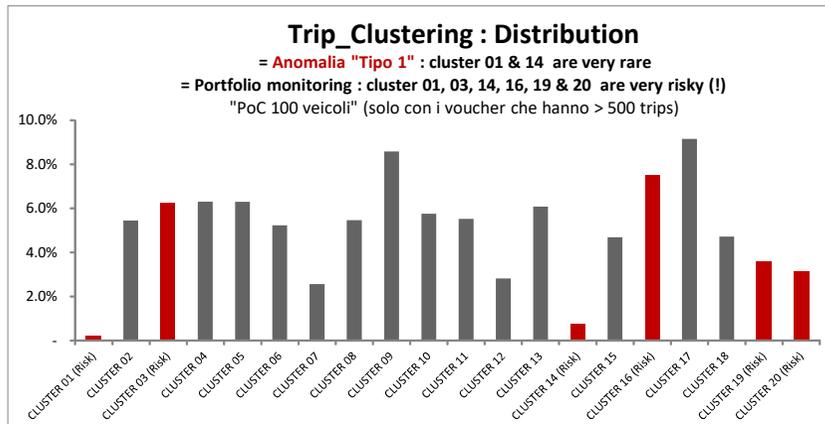


Retour d'expérience sur la data

Cas d'usage : La connaissance du risque (1/2)

1 – Classification automatique des parcours

Clustering du portefeuille complet :
(hors rouge : trajets dangereux)

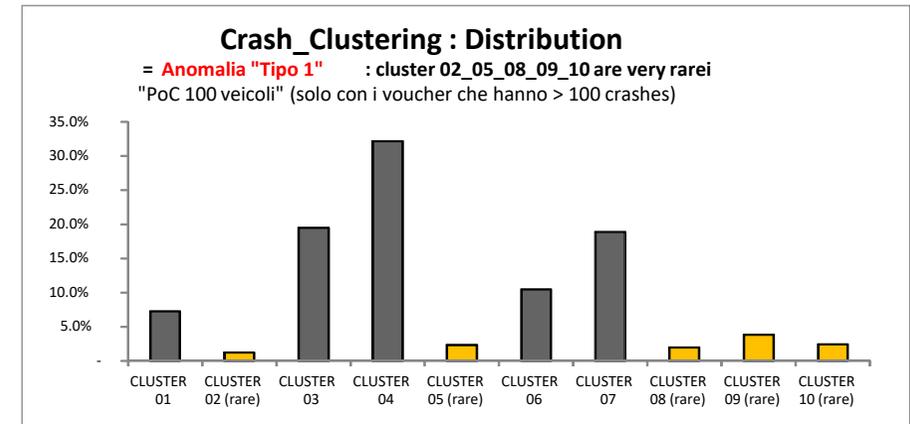


Cette classification doit notamment permettre :

- Le suivi de la qualité de conduite du portefeuille
- L'identification des conducteurs risqués
- L'identification des trajets anormaux (vol du véhicule ; prêt du véhicule ; etc...)

2 – Classification automatique des « Crash »

Clustering du portefeuille complet :
(hors orange : crash rares)



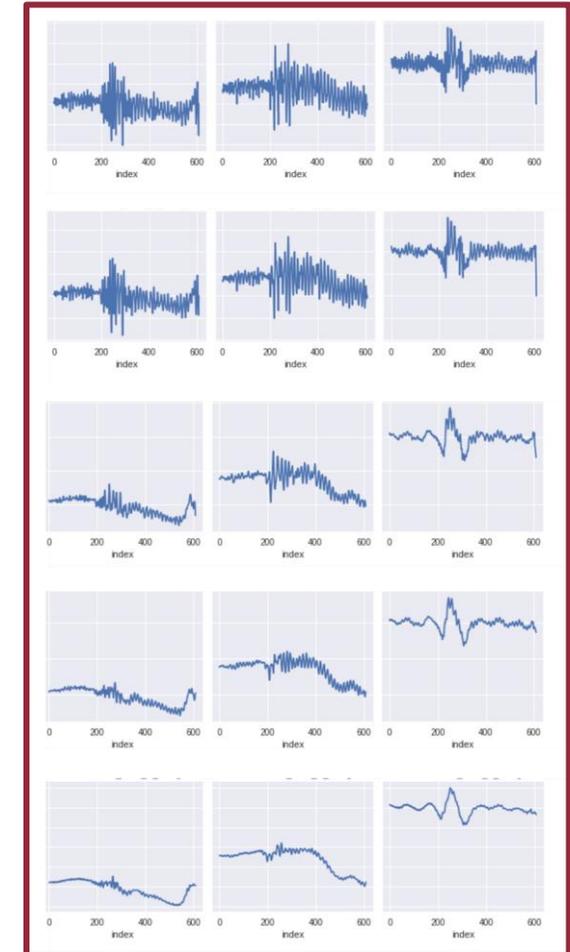
Cette classification doit notamment permettre :

- L'identification des crash anormaux (fraude ; box mal installée ; etc...)
- La création d'une variable « Crash » qui sera injectée en input de modèles prédictifs : probabilité de lésion ; cout moyen d'ouverture ; etc...

Retour d'expérience sur la data

Cas d'usage : La connaissance du risque (2/2)

Visualisations des résultats



Des enjeux techniques sur la fusion avec les données de crash :

- Gestion des ID
- Gestion des dates
- Gestion des multi-événements

Retour d'expérience sur la data

Cas d'usage : Les services et la prévention

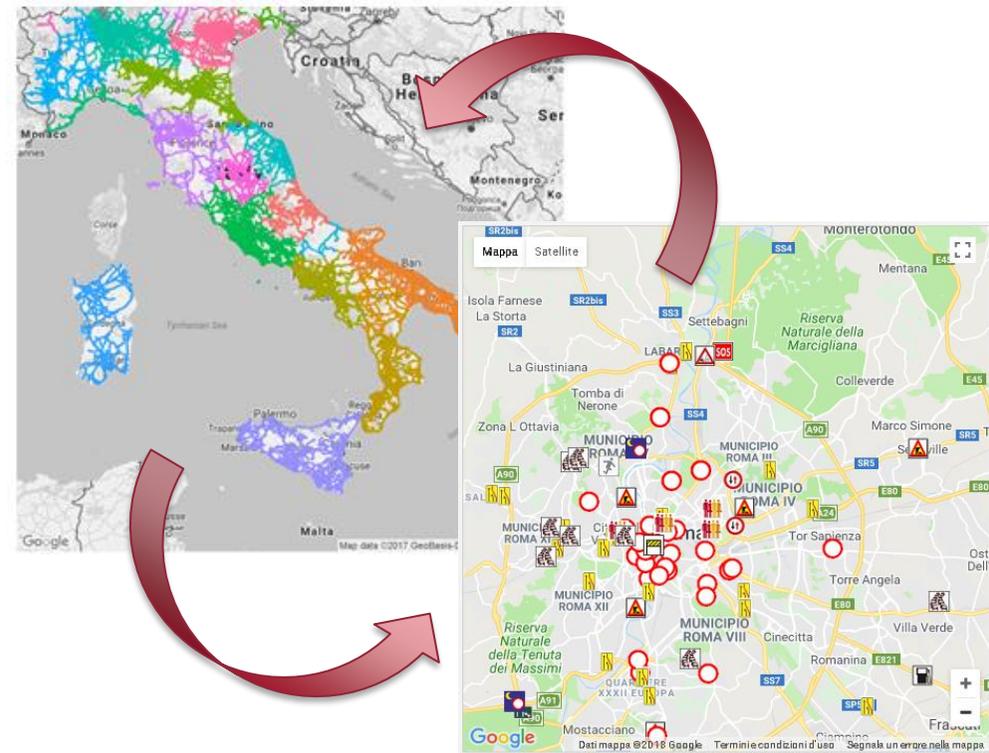
1 – Find my Car, Score et statistiques de conduite

→ **via une application Smartphone dédiée**



2 – Partenariat majeur avec le ministère italien des transports « CIS »

→ **Faire le lien « geofencing » entre les parcours et les événements sur route**

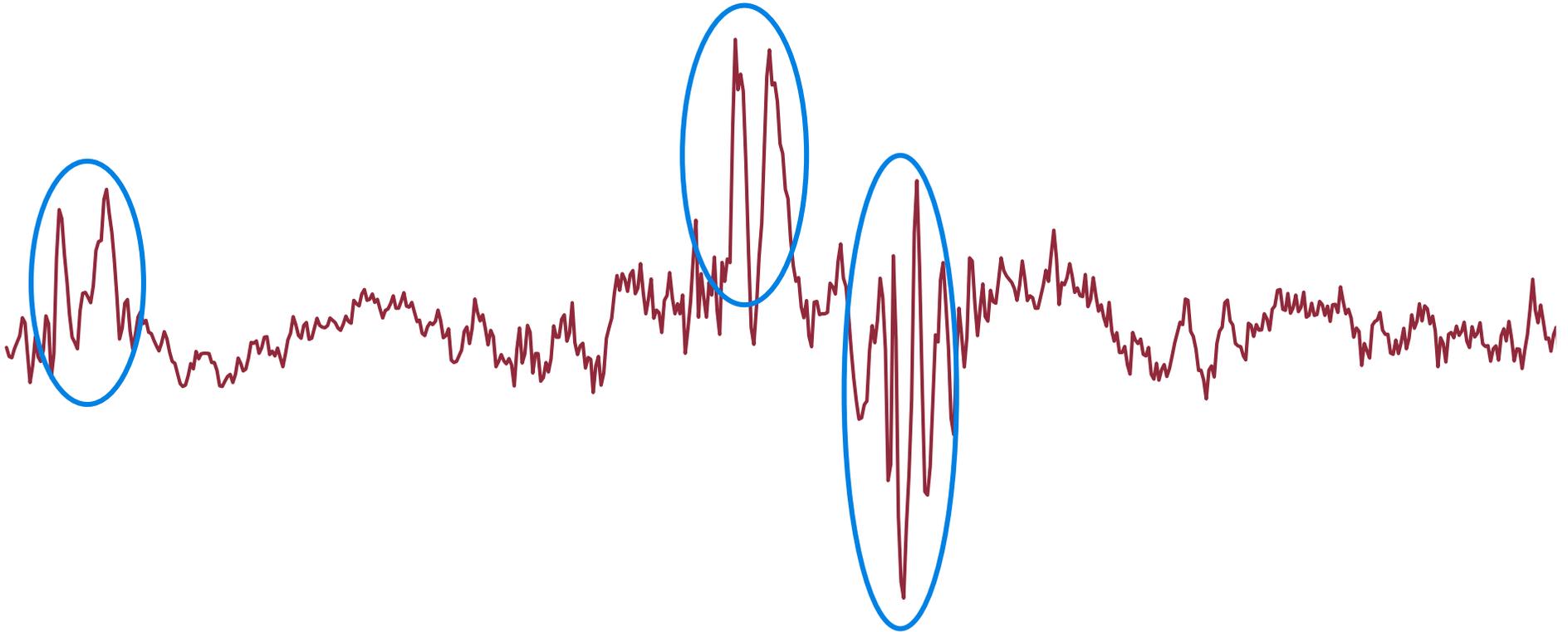


Retour d'expérience sur la data

Des enjeux Analytics et algorithmiques (1/3)

- Focus sur les crashes – trouver des « pics » et leur durée : trouver son propre algorithme

ACCELERATION 3D



Retour d'expérience sur la data

Des enjeux Analytics et algorithmiques (2/3)

- **Spark** permet d'utiliser certaines fonctions pré-existantes pour effectuer des opérations sur chaque ligne d'un dataframe (moyenne mobile, lag, etc.), cependant dans certains cas ces méthodes ne suffisent pas et il est nécessaire de coder une fonction spécifique (User Defined Function).
- Ici, un **exemple simple d'UDF** permettant de calculer la vitesse moyenne (en km/h) d'un véhicule à partir de sa distance parcourue (en m) et de la durée du trajet (en s).

1

```
#UDF
def avg_speed(distance,time):
    if time!=0:
        spd_tmp=distance/float(time)
        return spd_tmp*3600/1000
    else:
        return 0
```

Définition de l'UDF de manière analogue à n'importe quelle fonction sous **Python** :

```
def nom_fonction (param*) :
    ...
    return x
```

2

```
avg_speed_udf = udf(avg_speed, DoubleType())
```

Création de l'UDF via la fonction `udf()` de *PySpark*, en prenant soin de spécifier le **type retourné** (SQL Types)

3

```
dataframe=dataframe.withColumn('AVG_SPEED',avg_speed_udf(dataframe.DISTANCE,dataframe.DURATION))
```

Ajout de la nouvelle colonne/variable ('AVG_SPEED') ainsi calculée au dataframe existant.

Retour d'expérience sur la data

Des enjeux Analytics et algorithmiques (3/3)

- On obtient alors le dataframe suivant :

DISTANCE	DURATION	AVG_SPEED
3000	124	87.1
10342	3331	11.2
10	0.8	45
...		

Diagram illustrating the application of a User Defined Function (UDF) to calculate the average speed (AVG_SPEED) for each row of the dataframe. A yellow box labeled 'avg_speed_udf' is connected by lines to the 'AVG_SPEED' column of the table. An arrow points from the 'avg_speed_udf' box to the value '87.1' in the first row of the 'AVG_SPEED' column.

Répété pour chaque ligne du jeu de données (*process optimisé !*)

- Avantages** liés à l'utilisation des *User Defined Functions* :

- ✓ Permet de subvenir à des problématiques spécifiques dans l'éco-système Spark
- ✓ Réaliser des calculs ligne à ligne sur un jeu de données important de **manière optimisée** grâce à une infrastructure de calculs parallélisés

- Quelques limitations** selon l'utilisation souhaitée :

- ✗ Impossibilité de réaliser des UDF sur des dataframes agrégés (GroupBy) nativement sous PySpark
- ✗ Nombre important de types SQL définis sous PySpark notamment pour les sorties des UDF, ce qui rend parfois difficile la détermination du bon type

Retour d'expérience sur la data

Concrétiser ses expériences (1/2)

- Analytics : de la preuve de concept (PoC) à la **production** : deux approches complémentaires où l'une doit anticiper les besoins de l'autre.



Une idée, un problème à résoudre



Des résultats et de futures attentes

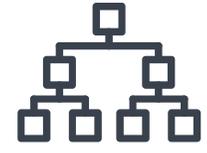
Preuve de concept



Efficacité



Durabilité



Évolutivité

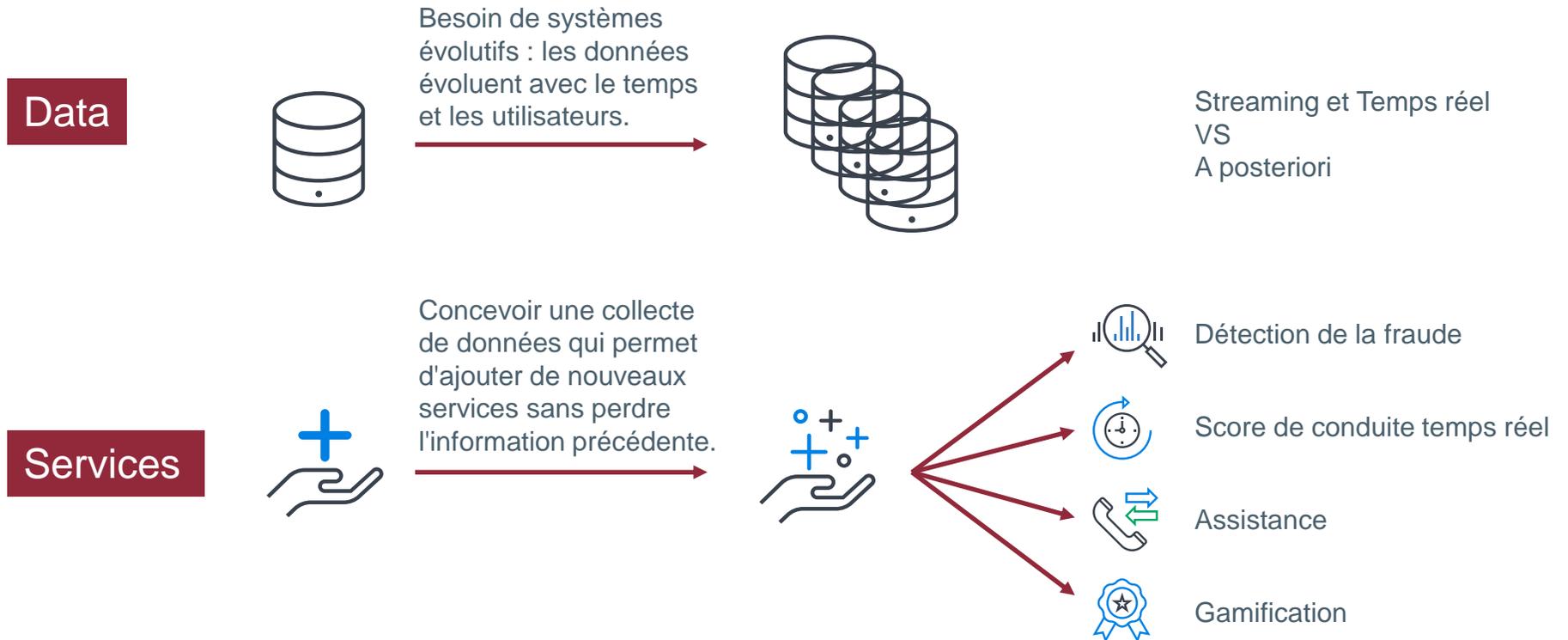
Production

- Bonne solution : celle qui anticipe l'utilisation future potentielle.
- Pour être rentable, un projet doit être conçu dans une **perspective générale**.

Retour d'expérience sur la data

Concrétiser ses expériences (2/2)

- En Telematics cela se rationalise :



Retour d'**expérience** sur le **business model** : Quels sont les modèles **émergents**? Comment les assureurs envisagent de créer de la **valeur**?

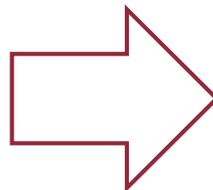


Retour d'expérience sur le business model

Quelles sont les questions posées?

Hier

- « Quel est le but? »
- « Cela ne fonctionne pas »
- « Ce n'est pas profitable »



Aujourd'hui

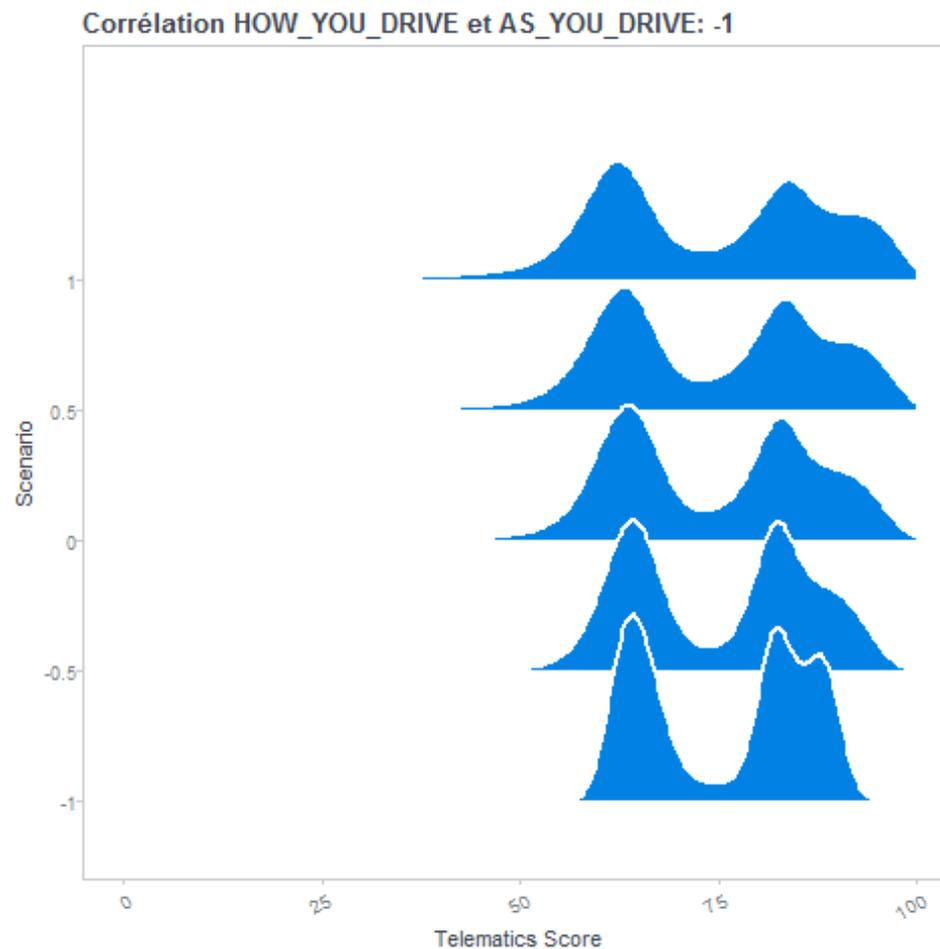
- « Il y a de nouveaux acteurs ? »
- « Combien d'autres le font? »
- « Nous ne voulons pas être les derniers à proposer une offre »
- « Avez-vous des preuves d'une corrélation entre le score et le taux de sinistres ? »
- « Les services pourraient aussi être intéressants »
- « Comment analyser les événements ? »
- « Quel pourrait être mon business model ? »

Retour d'expérience sur le business model

Tout commence par un score

→ **Retour d'expérience** sur une mission pour un assureur.

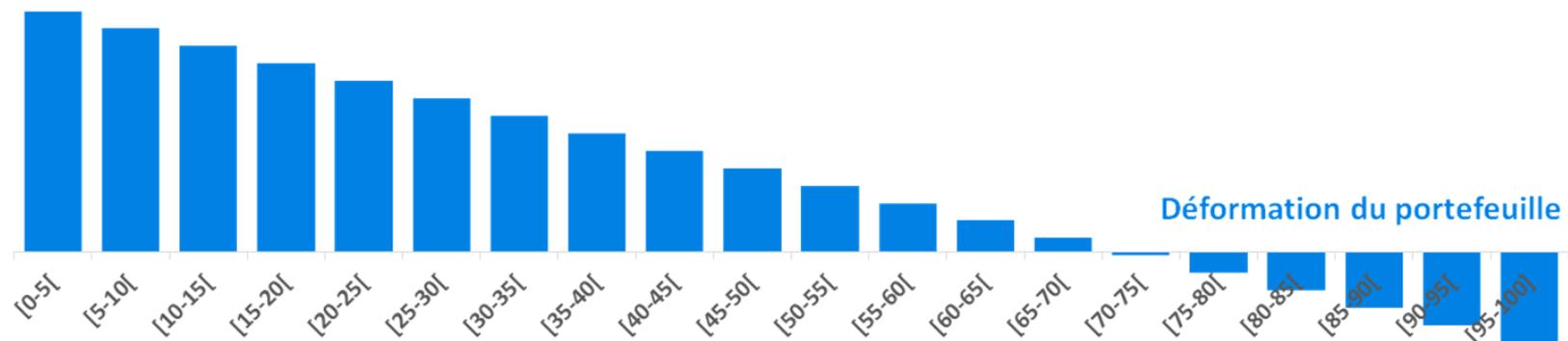
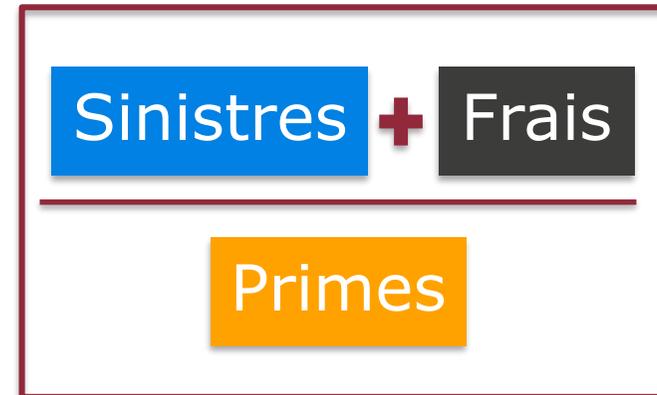
- Qu'est qu'un **score de conduite**?



Retour d'expérience sur le business model

Quel modèle tarifaire pour ce score?

- Quid du **Ratio Combiné** :
 - Comment gérer un environnement nouveau avec des **variables évolutives**?
 - Quel « **Discount** » envisager?
 - Un discours autour de la **causalité** – corrélation et des « incentives »

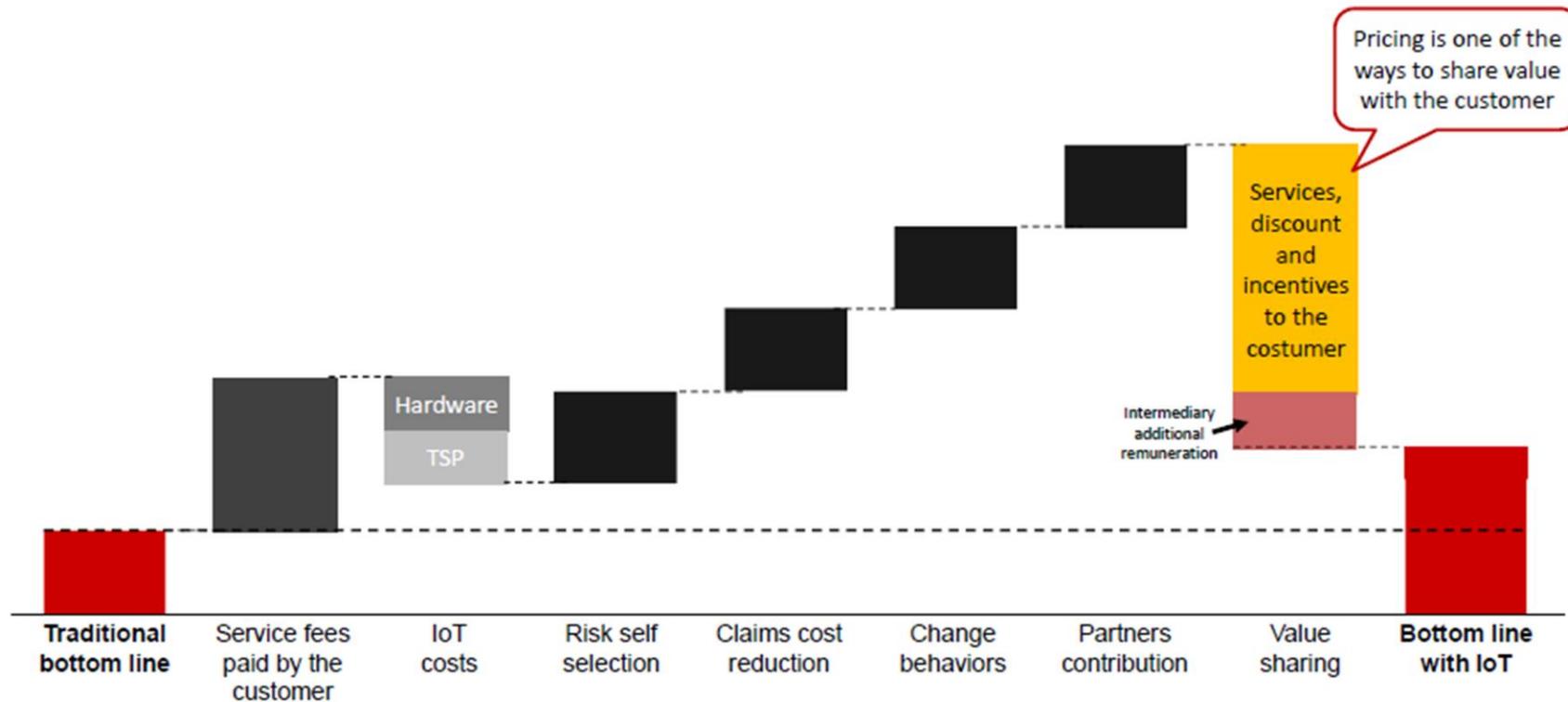


Retour d'expérience sur le business model

Zoom sur la marge

- Formation de la marge : de traditionnel à connecté

*** illustratif ***



Retour d'expérience sur le business model

Périmètre élargi : les éléments clés

Tarif adapté



Produits innovants



Souscription facilitée



Gestion de sinistres



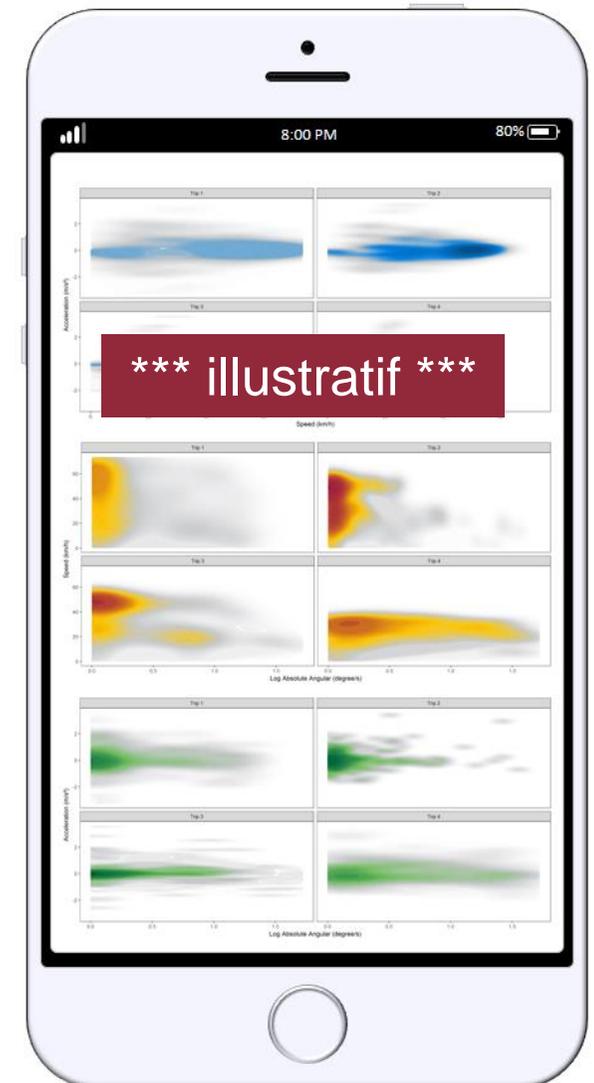
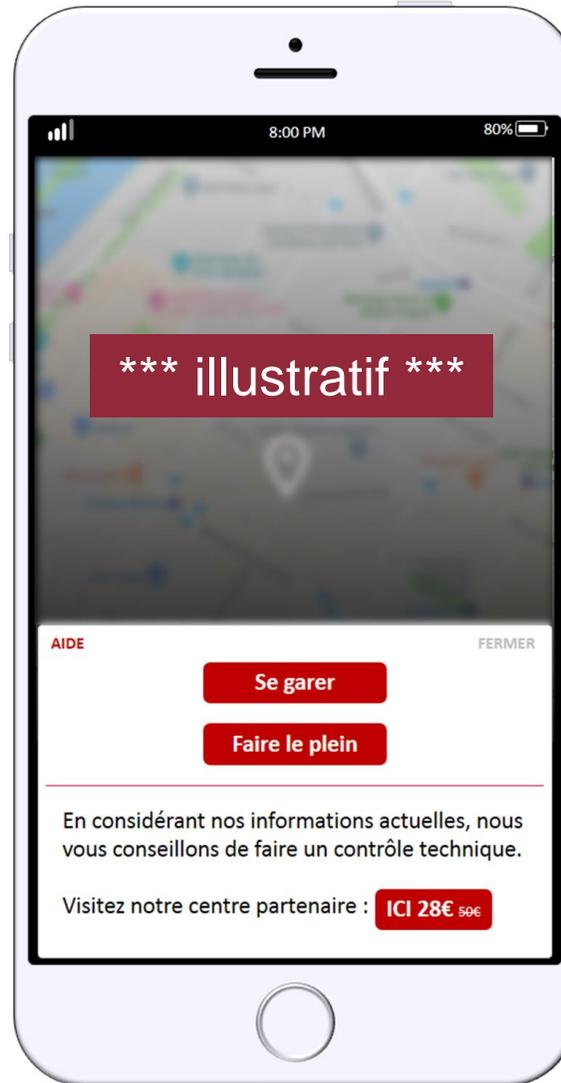
Services



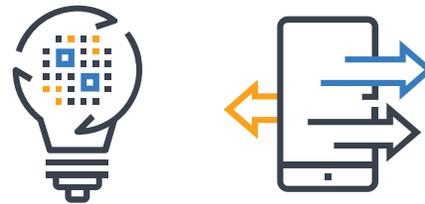
Retour d'expérience sur le business model

Zoom sur les services

- Assistance en temps réel
- Aide à la conduite
- Conseils préventifs
- Maintenance prédictive
- Contrôle parental
- Etc.



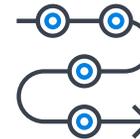
Quelle place pour le **telematics** et quel enjeu pour la **Data Science**? Quelles sont les futures **évolutions** en lien avec la **mobilité de demain**?



Conclusion

Trouver sa place dans la mobilité de demain

- Les **usages** et la **mobilité** évoluent :
 - Multimodale
 - Transition écologique
 - VTC, Covoiturage, partage de véhicule,
 - Etc.
- L'**Intelligence Artificielle** s'invite également :
 - Véhicule autonome
 - Taxi drone
 - Telematics « real-time »
 - Etc.
- Ainsi que le **digital** :
 - Génération connectée
 - En attente d'offres « à la demande »
- Des **initiatives** sur les évolutions de demain ont déjà été lancées



Conclusion

Une mutation en cours

- Les assureurs disposent désormais d'**outils toujours plus puissants** pour extraire l'information, l'analyser et créer de la valeur, que les données soient structurées ou non, avec une « **Toolbox** » toujours plus grande et adaptée :



- L'enjeu pour les assureurs est d'**anticiper l'avenir** et de bien se positionner dans ce nouveau paysage. Le **telematics** ouvre de nouvelles perspectives dans ce passage au **2.0**.



Groupama



Milliman

Contacts

Olivier CLAEYS, Responsable Digital Factory, **Groupama** –
olivier.claeys@ggvie.fr

Rémi BELLINA, Senior Consultant et Data Scientist, **Milliman** –
remi.bellina@milliman.com

Antoine LY, Senior Consultant et Data Scientist, **Milliman** –
antoine.ly@milliman.com