

## *BCG Gamma Datathon 2020*

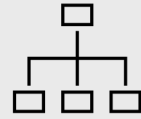
Bruno Mühlmann Holanda  
Caio de Próspero Iglesias  
Gabriel Pellegrino da Silva  
Marcel Makoto Kondo  
João Gabriel Lopes de Oliveira



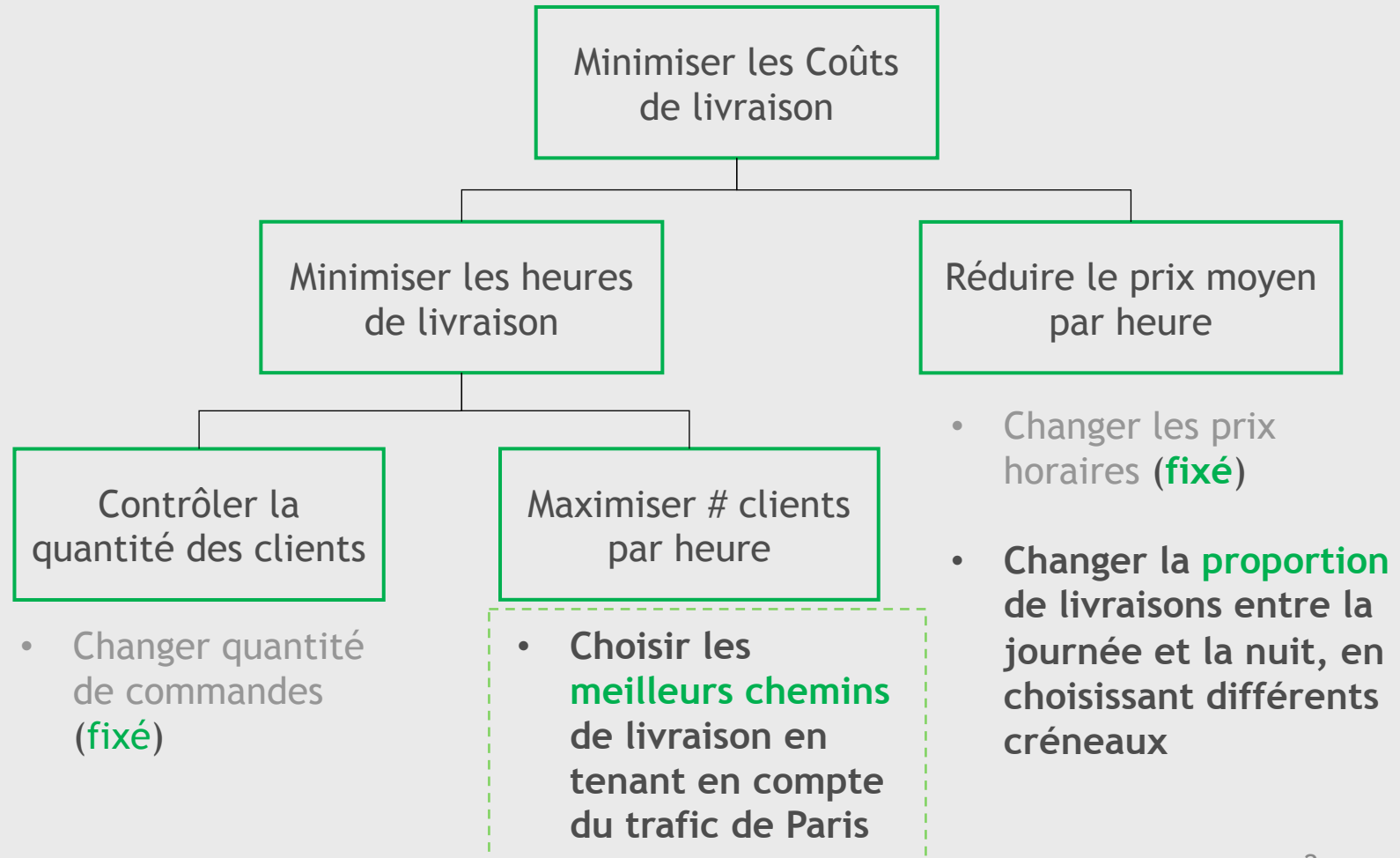
# LivraisonCo optimise son processus par minimisation des coûts

## Récap'

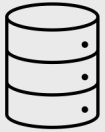
- LivraisonCo souhaite **optimiser son processus de livraison** à Paris
- Nos missions
  - **Prédire le trafic de Paris** (sur 3 rues pour le PoC)
  - **Construire une feuille de route** pour la suite du projet
- En considérant le nombre de commandes fixes
  - Optimisation des livraisons devient **minimisation des coûts**



## Structure du Problème



# Pour estimer le trafic, on a analysé les données qui pourraient l'influencer



## Collecte et préparation des données

Collecte de données supplémentaires

- **Open Data Paris** depuis 2016 - pour obtenir des comportements hors période de Covid
- **World Weather Online** - données avec la température et niveau de pluie
- Données avec les **vacances**<sup>1</sup> scolaires et jours fériés
- Données sur les périodes de **confinement**

Préparation de Données

- Élimination de l'**heure d'été**
- Élimination des heures qui n'avaient pas le débit ou le taux d'occupation enregistré



Collecte et Préparation des données



EDA



## Exploratory Data Analysis (EDA)

- Analyse d'**autocorrélation** du débit et Taux d'occupation
- Analyse des périodes **exceptionnelles** : vacances et confinement
- Inspection de l'importance de **données climatiques**



- **Forte corrélation** d'environ **0.4** avec le même jour de la semaine
- Réduction du **débit** et **taux d'occupation** en périodes exceptionnelles
- **Corrélation** entre les données **climatiques** et les différents **mois**



Construction et réglage de modèles



Sélection du modèle

# Construction du Modèle

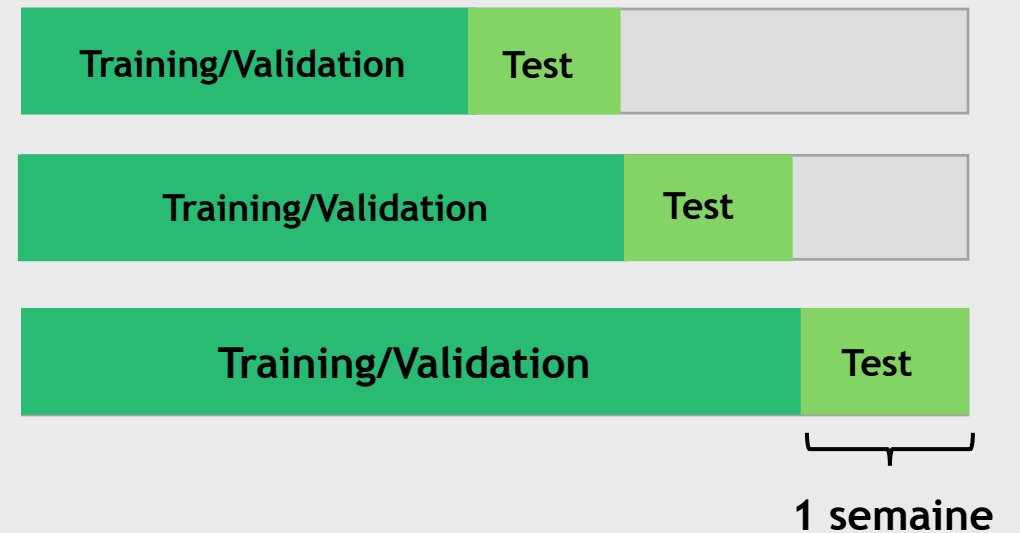
$$\hat{y}_{final} = \beta_0 + \beta_1 \times \hat{y}_{XGBoost} + \beta_2 \times \hat{y}_{KNN}$$

- **Stacking Regressor**
  - Utiliser les forces de chaque modèle
- **XGBoost** : Gradient Boosted Tree algorithm
  - **Rapidité** : calcul parallèle
  - **Évite l'overfit** : boosting régularisé
  - **Corrige bien ses erreurs** : entraînement itératif
  - **Gère bien les variables catégoriques** : adapté pour les données creuses
- **KNN**
  - Introduire une **mémoire** dans le modèle
  - Pondération sur les **valeurs historiques** les plus proche

# Réglage du modèle

- **Randomized Search CV**
  - Trouver un bon ensemble d'hyperparamètres
  - **Plus rapide** : explore des échantillons des paramètres

# Sélection du modèle



Collecte et Préparation des données



EDA



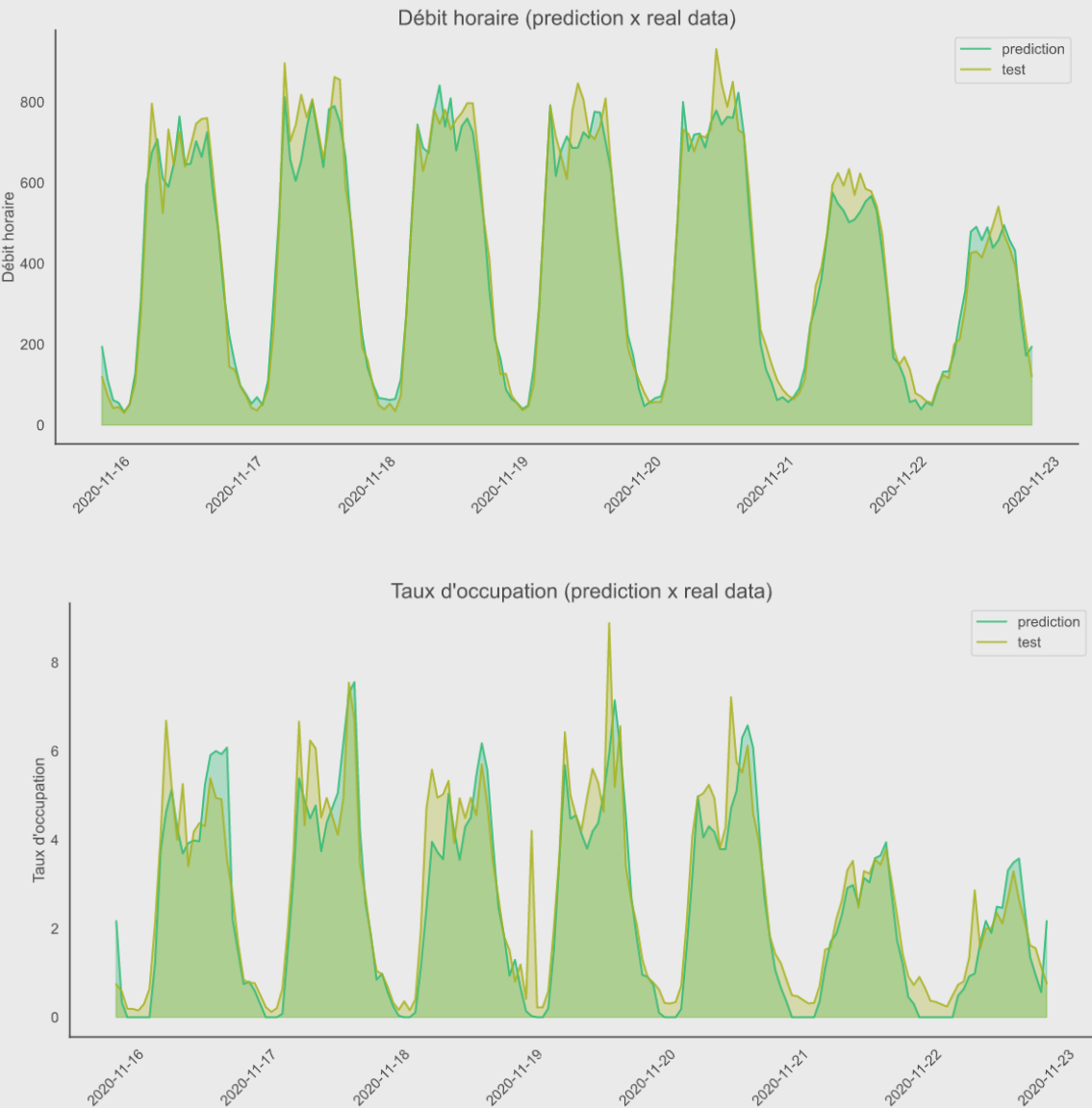
Construction et réglage de modèles



Sélection du modèle

# La robustesse du modèle est prouvée par le RMSE et le RMSE relatif faible

## Exemple pour la Rue de la Convention



## RMSE<sup>1</sup> pour chaque rue

	Débit Horaire	Taux d'occupation
Les Champs-Élysées	209	5.0
La rue de la Convention	75	1.7
La rue des Saints Pères	78	1.7



## RMSE<sup>2</sup> relatif pour chaque rue

	Débit Horaire	Taux d'occupation
Les Champs-Élysées	0.20	0.52
La rue de la Convention	0.13	0.24
La rue des Saints Pères	0.17	0.24

1- RMSE calculé en utilisant les 4 dernières semaines comme « Test set »  
2- Obtenu en divisant RMSE pour la moyenne de chaque rue

# Pour intégrer les estimations à une stratégie d'optimisation, des données complémentaires sont nécessaires



## Données nécessaires

- 1 Capacité des **véhicules** et taille des **colis**
- 2 Historique de **livraisons** hebdomadaires et nombre de **livreurs** disponibles par créneau<sup>1</sup>
- 3 Emplacement des **centres de distribution** (CD) et répartition des **produits** entre les CDs
- 4 Détails de la méthode de **paiement** des livreurs, principalement en ce qui concerne les voyages vers le CD



## Hypothèses de l'approche

- 1 Toutes les véhicules ont la **même** capacité et ils opèrent au **maximum** de leur capacité
- 2 **Nombre fixe** de livraisons par semaine, fourni par un modèle *XGBoost*, et de livreurs par créneau
- 3 Chaque livraison peut provenir **uniquement d'un CD**
- 4 Nous ne payons pas pour **les trajets maison-CD**, mais s'il a besoin de **prendre d'autres produits** sur le CD, cela sera payé

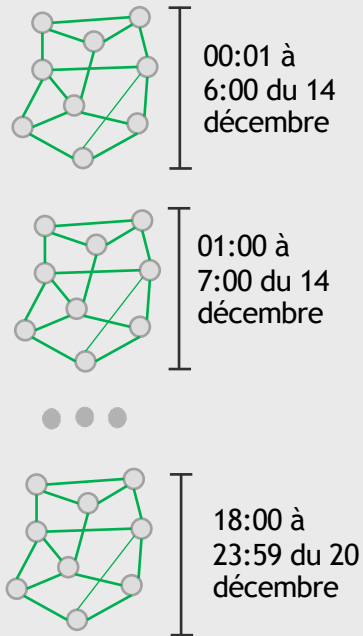
Optimiser la livraison d'un CD à la fois, en maximisant le nombre de livreurs aux meilleurs moments, sans qu'il soit nécessaire de revenir sur les CD

# La stratégie pour optimiser les livraisons consiste à choisir les meilleurs créneaux de livraison

1

## Construction d'un réseau pour chaque créneau

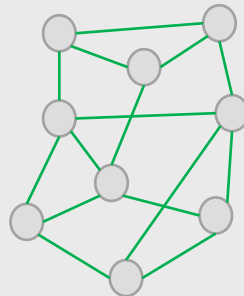
- **Créneau optimal** : temps moyen de livrer la pleine capacité du véhicule
  - Pas besoin de retourner au CD
- **Séparer en différent réseaux** : analyser toutes les possibilités pour le créneau optimal



3

## Sélection du meilleur réseau

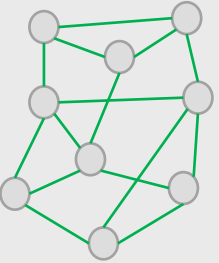
- **Heuristique** : sélectionner meilleur créneau et allouer le plus grand nombre possible de livreurs
- **Critère** : réseau avec des pondérations temps/prix plus faibles en moyenne



2

## Construction des poids des arrêts

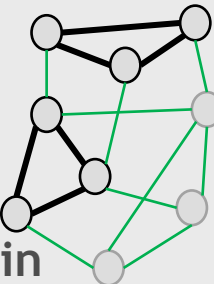
- **Nœuds** : Points des Livraisons
- **Arrêts** : Poids en fonction de temps de déplacement et prix<sup>1</sup>
- Calcul de temps de déplacement
  - **Distance** calculée en utilisant l'algorithme<sup>2</sup> A\*
  - **Vitesse**<sup>3</sup> calculée en utilisant le débit et le taux d'occupation estimés



4

## Séparation en *Clusters* dans le réseau

- Un cluster pour **chaque livreur** :
  - Des clusters qui prennent en compte les temps du créneau
  - Dès que les livreurs disponibles ont terminé, analyser le prochain créneau



# Prochaines étapes peuvent mitiger les risques associés à l'approche recommandée



L'**optimisation** des livraisons est obtenue avec une stratégie de choisir les meilleurs créneaux, en termes de **temps de déplacement** et **prix**

## Risques possibles <sup>1</sup>

- Dépendance des **modèles de prévision** pour l'organisation des créneaux (trafic et demande)
- Ne pas avoir une **priorité** entre les différents clients et demandes
- Les derniers points de livraison peuvent être **éloignées** entre eux

## Timeline





*Merci !*

*Des Questions ?*