*Nom de la team : RSA*

*Sprint n°1 - Projet N°6 Criminalité à Chicago*

*Scrum Master : Ryme Youb*

*Teammates : Ryme Youb, Sacha Cymermann, Adrien Golebiewski*

*Début : 20/09/2021*

*Fin : 29/11/2021*

**Sprint 1 - Migration des données dans Hadoop et Calcul Map Reduce avec l’outil Yarn**

1. **Transfert des données de criminalité dans l'éco-système Hadoop :**

Accès au serveur Hadoop du CNAM

Récupération des clé d’accès

Migration des données dans HDFS - Outil de gestion de fichiers et de clustering

1. **Prise en main de Hadoop et de l’outil Yarn**

Familiarisation des outils :

* montée en compétences des outils - cours “Base de données”
* recherche internet

Tests de requêtes depuis Yarn - Prise en main de JAVA (langage pour la réalisation des requêtes)

1. **Définition des requêtes Map Reduce**

Choix de 3 requêtes MAP REDUCE,qui montent en complexité, par exemple :

* Moyenne “simple” du nombre de crimes par “Location Description” (simple)
* Moyenne mobile du nombre de crimes (Fenêtre de données de 2001 à 2021 soit une intervalle de [2001-N ; 2021 + N]) (compliqué)

Le résultat sera une liste de couples (année, moyenne du nombre de crimess sur les ±N années autour)

* Moyenne mobile du nombre de crimes par “Location Description”. La clé de sortie du reducer sera le nom du “Location Description”, et il affichera l’année et le nombre de crimes. Quant à la sortie du mapper, il y aura à fois l’année et le “Location Description” (compliqué +)

1. **Analyse des résultats et commentaire**

Présentation des résultats des requêtes obtenus

Critiques sur les méthodes employées

**Backlog à l’état du Sprint 1**

