1. **Exploratory Data Analysis**

**Conclusions de l’analyse :**

* COMMENT GERER LES VARIABLES ‘TIME’ ?
* Variables de jours toutes identiques (quali & quanti)
* Target et distance corrélées (peut créer vitesse)
* Valeurs fausses dans la target

# Checklist de base

## Analyse de Forme :

### variable target :

Time from Pickup to Arrival (temps en secondes), numérique

### lignes et colonnes :

(21201, 33)

### types de variables :

* 19 numériques, 10 catégorielles

### Analyse des valeurs manquantes :

* Quelques manquantes sur Temperature et presque tout manquant sur precipitations

## Analyse de Fond :

### Visualisation de la target :

* Légèrement skewed à gauche
* Temps moyen de livraison = 1556.920947 s (=25 min)
* Valeurs fausses (durée de livraison = 1s)

### Signification des variables :

Variables quanti :

* Uniforme sur les days of month
* Distance : entre 0 et 20 km
* Temperature : majorité entre 15 et 30 (degrés celcius ?)
* Age à subi une transformation mathématique
* Notes des livreurs entre 12 et 16
* Majorité des livreurs notés entre 0 et 250 fois

Variables qualitatives :

* Tous les jours de la semaine semblent equidistribués mais moins de livraisons le weekend
* Les livreurs ne roulent qu’à vélo (variable useless)
* Majorité des commandes pour le buisness

### Relation Variables / Target :

Variables quanti :

* Pas de relation de corrélation (sauf peut-être Distance ?)

Variables quali :

* Rien de vraiment apparent, si ce n’est que le dimanche les livraisons sont légèrement plus longues

### Outliers :

* 775 outliers (3.65%)

## Analyse plus détaillée

### Relation Variables / Variables :

#### Quanti / Quanti :

* Majorité des variables : pas de relation (nuage de points uniforme)
* Faible corrélation (-0.2) entre température et précipitation
* Target et distance corrélés (0.6)
* Variables de jour corrélées entre elles
* Nombre de commandes et de notes corrélés
* Age et nombre de commandes corrélés

#### Quanti / Quali :

* Trajets les + longs pour le buisness
* Grande majorité sur plateforme 3, mais grand nombres de commandes à 45 km environ sur plateforme 1
* Particuliers notent ou mal ou bien (pas d’entre deux)
* Plateform 2 semble mal notée
* Particuliers commandent bcp quand il fait froid