

UNIVERSITÉ PARIS DAUPHINE

DATA ANALYTICS

---

# Implementation de KMeans avec Spark

---

*Etudiants*

Elie ABI HANNA DAHER  
Bilal EL CHAMI  
Badr ERRAJI

*Professeur*

M. Benjamin  
NEGREVERGNE

29 avril 2018



## Table des matières

<b>1</b>	<b>Project goal</b>	<b>2</b>
<b>2</b>	<b>Implementation</b>	<b>2</b>
<b>3</b>	<b>Execution</b>	<b>2</b>
<b>4</b>	<b>Data generator</b>	<b>2</b>
<b>5</b>	<b>Performance</b>	<b>2</b>

## **1 But du projet**

Le but du projet est d'implémenter l'algorithme KMean avec Spark (python) en évaluant la performance de notre implémentation basé sur des données que nous générons.

## **2 Implémentation**

## **3 Générateur des données**

## **4 Performance**