

Exercice 01 :

Une entreprise de télécommunications possède une chaîne de points de vente au niveau de différentes villes (un point de vente par ville), et elle prévoit l'ouverture d'un nouveau point. L'entreprise possède des données concernant ses bénéfices et les populations des différentes villes (fichier data1). Ayant le choix entre plusieurs villes, l'entreprise souhaite exploiter ces données afin de prendre une décision. Pour aider cette entreprise, proposer l'implémentation de la régression linéaire avec une variable pour prédire et comparer les éventuels bénéfices afférents aux différentes villes.

Le fichier data1.txt contient l'ensemble des données pour notre problème de régression linéaire. La première colonne représente la population d'une ville et la deuxième représente le bénéfice dans cette ville. Une valeur négative pour le bénéfice indique une perte.

Etapes à suivre :

- 1) Télécharger et visualiser les données.
- 2) Initialiser les différents paramètres m , X , θ , α , iteration et α .
- 3) Implémenter une fonction (cost) pour calculer le coût
- 4) Vérifier que le coût pour $\theta = [0 ; 0]$ est 32.07
- 5) Vérifier que le coût pour $\theta = [0 ; 0]$ est 54.24
- 6) Implémenter la fonction de la descente de gradient
- 7) Ajouter le code nécessaire à la fonction cost afin de sauvegarder le coût obtenu pour chaque itération.
- 8) Dessiner le coût en fonction des itérations et vérifier qu'il diminue à chaque pas.
- 9) Prédire les gains pour une ville dont la population est de 35 000.
- 10) Tester différentes valeurs pour α et commenter les résultats.
- 11) Réajuster votre code pour qu'il soit fonctionnel pour la régression multi-variables et en utilisant le fichier (data2) répondre aux questions 4), 5), 8) et 9.