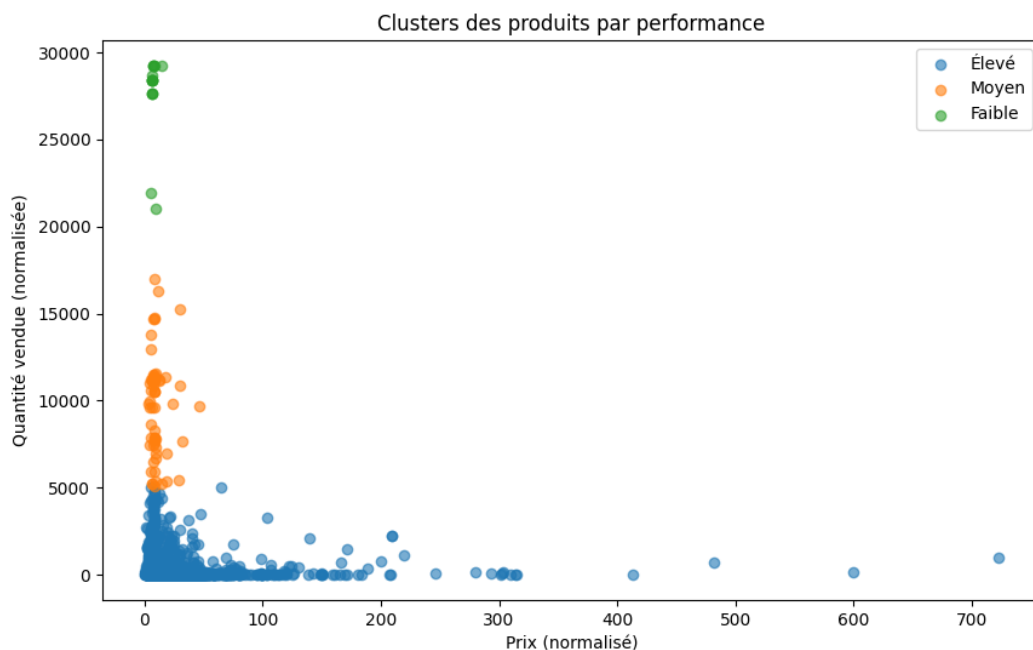


## RESULTATS VISUELLE DE L'EXERCICE AVEC PYTHON

1)J'ai chargé le fichier avec la fonction `read_csv` apres avoir installé le package pandas. Et tout les installations possibles

Sorties Attendues Diagramme de Dispersion

- Points regroupés en trois clusters, étiquetés comme produits : Élevés, Moyens, ou Faibles. Avec k-means k=3 .



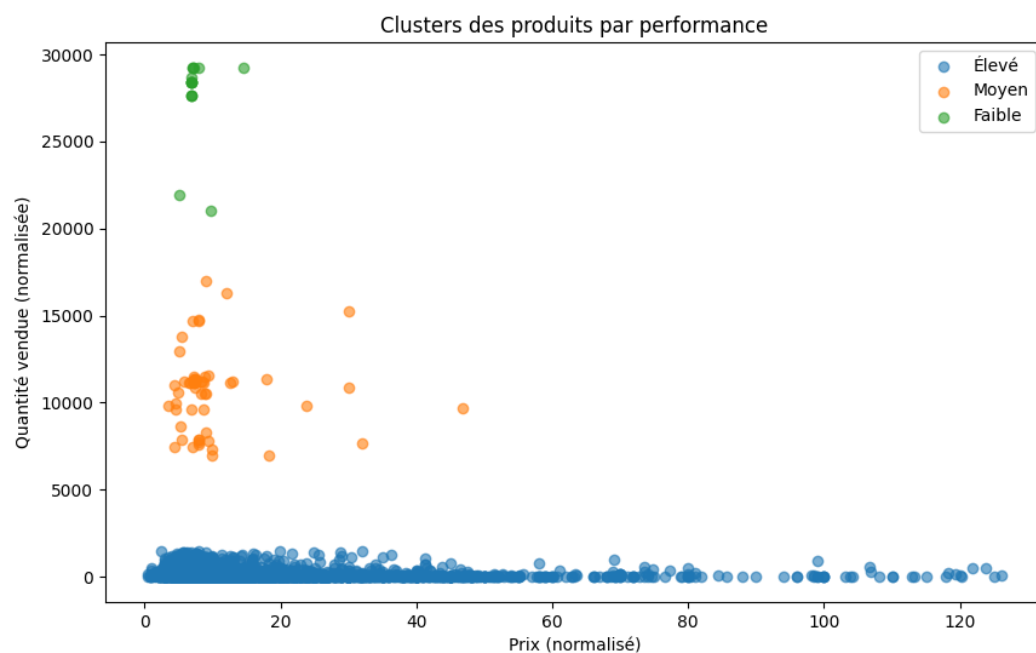
**Figure n°1 :** Le diagramme de dispersion qui montre les clusters dans un espace 2D, metant en évidence la relation entre **Price** et **Quantity\_Sold(quantité vendue)** .

2) Ce Tableau indiquant les moyennes de Price, Quantity\_Sold, et Inventory\_Level pour chaque cluster.

```
clustering_data['Cluster_Label'] = clustering_data['Cluster'].map(labels)
      Prix Moyen  Niveau de Stock Moyen  Quantité Vendue Moyenne
Cluster_Label
Faible          7.571250          11111111.0          27661.687500
Moyen          10.262714          11111111.0          9865.142857
Élevé          20.352167          11111111.0           221.481275
PS D:\EPSI\fouille_donnees>
```

## La modification du nombre de clusters (k) pour observer l'effet sur le regroupement.

Quand k= 5



**Figure n°1 :** Le diagramme de dispersion qui montre les clusters dans un espace 2D, mettant en évidence la relation entre **Price et Quantity\_Sold**(quantité vendue) . Quand le clusters k=5

**Le Tableau indiquant les moyennes de Price, Quantity\_Sold, et Inventory\_Level pour chaque cluster (k=5)**

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
clustering_data['Cluster_Label'] = clustering_data['Cluster'].map(labels)
      Prix Moyen  Niveau de Stock Moyen  Quantité Vendue Moyenne
Cluster_Label
Faible          7.571250          111111111.0          27661.687500
Moyen          10.053051          111111111.0          10655.542373
Élevé          17.895301          111111111.0           127.205337
PS D:\EPSI\fouille_donnees>
```

Ln 30, Col 29 Spaces:4

## QUAND LE CLUSTER K=1

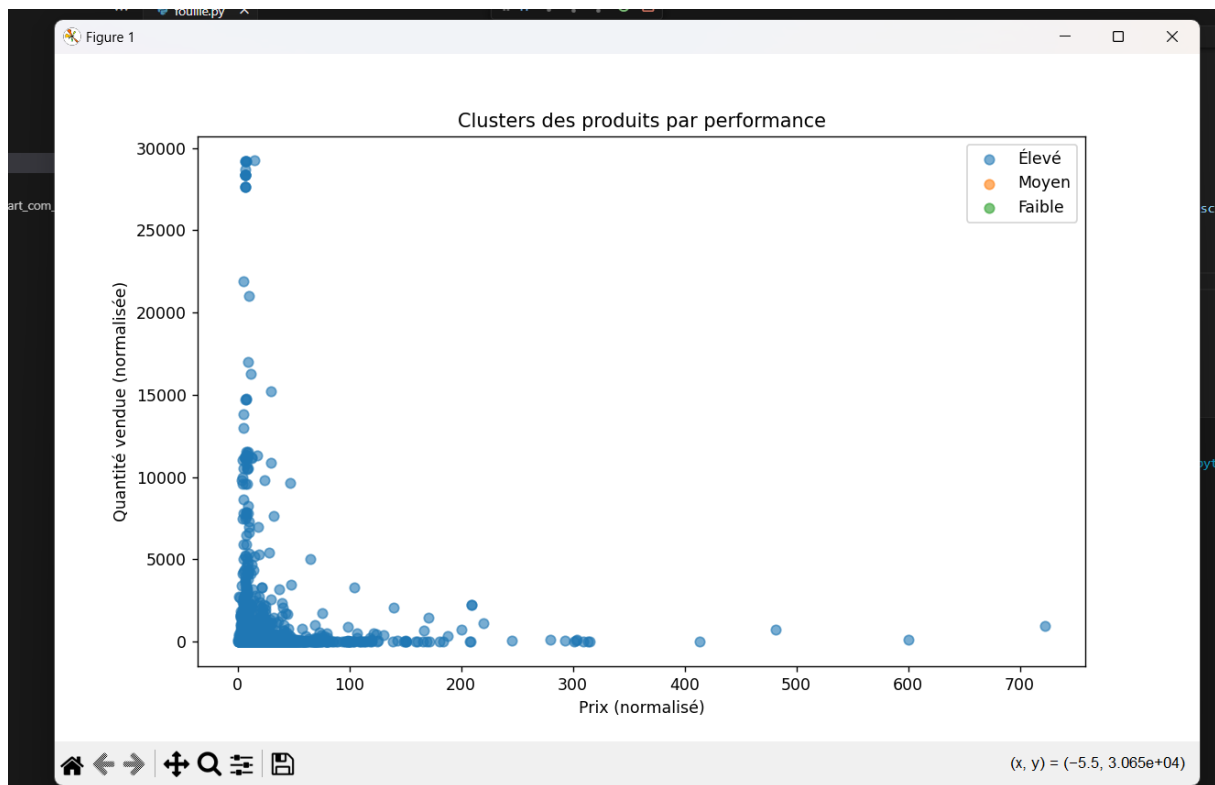


figure : Le diagramme de dispersion

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
clustering_data['Cluster_Label'] = clustering_data['Cluster'].map(labels)
Prix Moyen Niveau de Stock Moyen Quantité Vendue Moyenne
Cluster_Label
Élevé 20.077594 111111111.0 557.356949
```

Le Tableau qui montre le résultat de cluster k=1

Ajoute des fonctionnalités supplémentaires (comme ex.Category) pour explorer le clustering multidimensionnel.

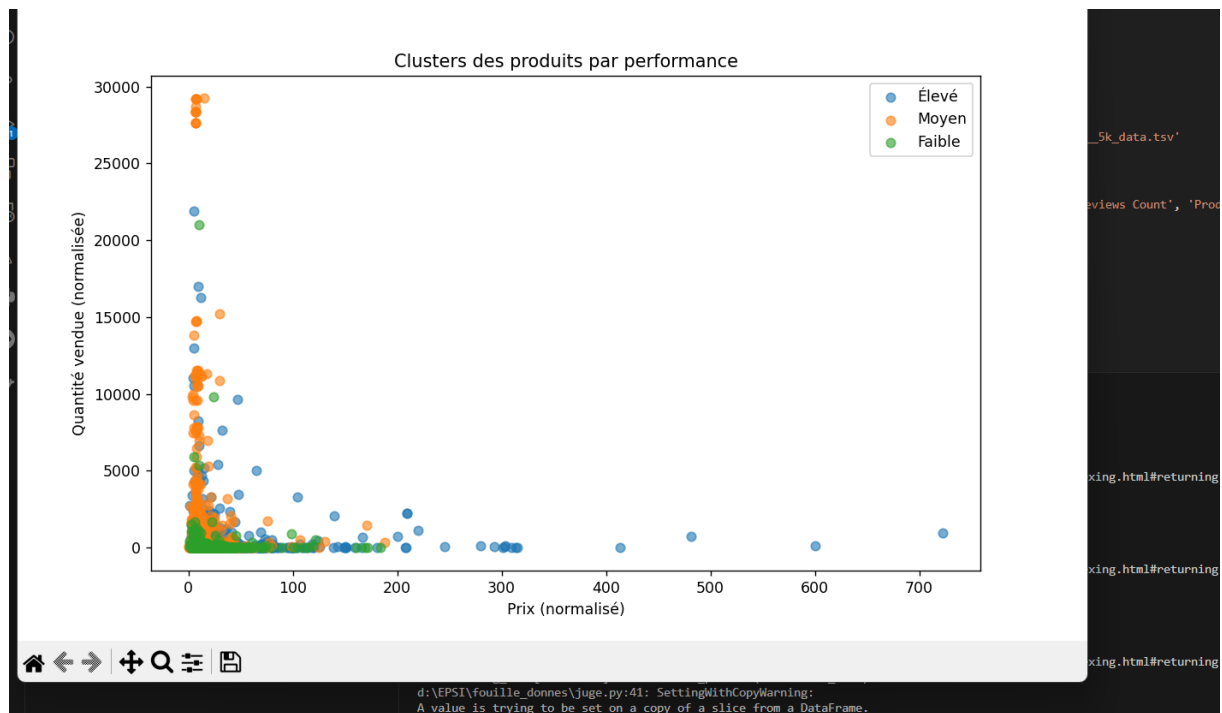


Figure : Ce diagramme montre la dispersion des clusters après avoir ajouté la fonctionnalité catégorie.

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
clustering_data['Cluster_Label'] = clustering_data['Cluster'].map(labels)
```

	Prix Moyen	Niveau de Stock Moyen	Quantité Vendue Moyenne	Catégorie Moyenne (encodée)
Cluster_Label				
Faible	17.178139	11111111.0	138.672222	451.381944
Moyen	13.272126	11111111.0	1360.221665	157.085256
Élevé	25.692388	11111111.0	244.247014	714.395349

PS D:\EPSI\fouille\_donnees>

Tableau montrant le résultat après l'ajout de Catégorie

Visualiser les clusters en utilisant différentes paires de fonctionnalités.

(le code est écrit en boucle pour présenté le diagramme en une seule page)

