

**Étude de la productivité du Riz au Bénin de 1998 à 2019 :
approche par Apprentissage non supervisé**

Adjimon VITOFFODJI

Table des matières

LISTE DES TABLEAUX ET GRAPHIQUES	3
<u>INTRODUCTION GENERAL</u>	<u>4</u>
<u>CHAPITRE 1 : CADRE THEORIQUE ET METHODOLOGIE</u>	<u>6</u>
SECTION 1 : CADRE DE L’ETUDE.....	6
SECTION 2 : METHODOLOGIE DE RECHERCHE	7
<u>CHAPITRE 2 : PRÉSENTATION ET ANALYSE DES RÉSULTATS</u>	<u>9</u>
<u>CONCLUSION</u>	<u>14</u>
<u>ANNEXE :.....</u>	<u>15</u>

Liste des tableaux et Graphiques

Tableau : Variables et Modalités	7
Graphique 1 : Nuage de point 1	9
Graphique 2 : Score et Plots	10
Graphique 3 : Distances	11
Graphique 4 : Matrice de Distance	11
Graphique 5 : Clustering Hiérarchique	12

INTRODUCTION GENERAL

Justification de l'étude

L'agriculture est retenue comme l'un des leviers principaux de la transformation structurelle de l'économie et un secteur de concentration pour dynamiser l'économie nationale. La contribution de ce secteur au PIB (part du PIB agricole dans le PIB global) a fluctué autour d'une moyenne de 27,8% sur la période 2015-2019 (Rapport de performance secteur agricole et gestion 2019).

Dans le but d'améliorer les performances de l'agriculture béninoise pour la rendre capable d'assurer de façon durable la souveraineté alimentaire et nutritionnelle et de contribuer au développement économique et social des hommes et des femmes du Bénin, à l'atteinte des Objectifs du Millénaire pour le Développement (OMD) et la réduction de la pauvreté, l'un des projets phares du Plan Stratégique de Relance du Secteur Agricole(PSDSA) concernent le renforcement des filières conventionnelles à savoir : le maïs, le manioc et le riz .

Le riz est un élément fondamental de l'alimentation de la moitié de la population mondiale avec près de 50millions de consommateurs en Afrique dont la plupart sont en milieu rural (Gnacadjia et al, 2017). C'est la deuxième céréale cultivée au monde et la troisième produite consommé et exportées après le blé et le maïs (Vigné. ,2011). En Afrique il vient après le sorgho et le maïs (Vigné, 2011). Au Bénin, le riz est de plus en plus consommé (SNDR, 2010). Il représente aujourd'hui une denrée de base dans le régime alimentaire de bon nombre de béninois.

Cependant malgré l'importance de la filière riz et les potentiels de production énormes (MAEP 2010), la production nationale n'arrive toujours pas à satisfaire tous ses besoins sans cesse croissants de la population. L'offre est nettement en dessous de la demande (MAEP, 2014). On assiste dès lors à un problème majeur qui est celui de la faible productivité. Ainsi pour combler le déficit engendré par l'insuffisance de la production nationale, le Bénin a recours à des importations massives. Pour renverser cette tendance il est alors primordial pour le Benin d'augmenter la production nationale à travers des stratégies pour la redynamisation de la filière riz.

En vue d'apporter une part contributive aux stratégies à mettre en œuvre pour la redynamisation de la filière riz au Bénin, nous avons jugé utile, dans le cadre de notre projet d'étude en Science de donnée de réfléchir sur le thème : « Étude de la productivité du riz au Benin de 1998 à 2019, approche par apprentissage supervisé ».

La présente étude s'articule autour de deux (02) chapitres. Le premier chapitre met en relief le cadre théorique et méthodologique de l'étude, le second présente les résultats, interprétation.

CHAPITRE 1 : CADRE THEORIQUE ET METHODOLOGIE

Le présent chapitre expose le cadre théorique et méthodologique de l'étude. Il s'articule autour de deux sections à savoir : les Objectifs et hypothèses de recherche et la méthodologie.

Section 1 : Cadre de l'étude

A-Problématique

Ciblée par le gouvernement comme étant une des filières à forte valeur ajoutée après le coton, l'anacarde, l'ananas, le manioc et le maïs, la filière riz joue un rôle important dans l'économie et le régime alimentaire de nombreux ménages. Considéré autrefois comme un aliment de luxe, le riz est devenu une denrée de consommation courante et un produit stratégique pour une bonne partie de la population mondiale à cause de la facilité et des faibles coûts de sa préparation, des prix bas et de la régularité de son approvisionnement (souvent satisfait par les importations) (Fall, 2005 ; Africaine, 2011).

Malgré les différentes actions que mène le gouvernement, l'offre domestique en riz demeure faible avec une présence accrue des importations.

Ainsi quel est l'effet des actions mené par les gouvernants de 1998 à 2019 sur la productivité du riz au Bénin ? quel est l'effet de l'accès au crédit agricole sur la productivité du riz ?

B-Objectifs de l'étude

Cette étude vise à :

- Regrouper les années semblables au cours desquelles nous avons une forte productivité du riz
- Identifier les facteurs déterminants en fonction des années qui expliquent au mieux la production du riz
- Analyser les raisons pour lesquelles certains groupe d'année ont été regrouper ensemble et d'autre non
- Examiner la relation existante entre l'accès au crédit agricole et la productivité du riz.

Section 2 : Méthodologie de recherche

Dans cette partie, nous présenterons dans un premier temps, la description de notre jeu donné, la nature et les sources de données puis, dans un second temps, les différentes méthodes de l'apprentissage non supervisé utilisées dans le cadre de notre étude.

Paragraphe 1 : Données utilisées et sources

A. Type et sources des données

Les données utilisées dans cette étude proviennent de l'Analyse Globale de la vulnérabilité de la Sécurité Alimentaire (AGVSA) réalisée en 2017. Ces données sont basées sur une enquête auprès de 15 000 ménages et une enquête auprès d'informateurs clefs dans 745 villages /quartiers.

B. Description des variables utilisées

La variable cible d'intérêt de notre étude est la variable quantitative année. Elle va nous permettre de voir comment est constitué les différents clusters au fil des années.

Les variables explicatives d'intérêt de notre étude sont : la variable quantitative superficie récoltée qui détermine la production du riz, la variable quantitative crédit agricole qui représente l'accès au crédit agricole reçu par les producteurs qui permet à ces derniers d'avoir accès aux engrais, pesticides, l'eau d'irrigation et les machines agricoles. L'accès aux crédits agricole permet d'améliorer la productivité du riz (Diamouténé et al, 2019)

Les autres variables explicatives de notre modèle sont : les variables quantitative rendement, engrais synthétiques et superficie total doté d'irrigation.

Tableau : Variables et Modalités

Variabes	Description	Modalité
Année (Cible)	Période de regroupement des clusters	Variable continue
Superficie récolté (intérêt)	Production du riz au cours d'une période	Variable continue
Crédit agricole (intérêt)	Subvention de l'état pour la production du riz	Variable continue

Paragraphe 2 : Méthode d'analyse

La méthodologie adoptée en vue d'atteindre notre objectif fait l'objet de ce paragraphe.

Spécification des clusters

Pour déterminer les différents groupes qui partitionnent au mieux les données de nos variables, nous allons explorer plus en détail les données que contient notre base de données à travers Data Table, nous allons afficher ces données sous la forme d'un nuage de points à l'aide d'un Scatter Plot. A l'aide de l'onglet Find informative Projections dans le Scatter plot nous allons retenir les deux (02) variables qui séparent au mieux les données en fonction de notre variable cible qui est Année. Nous allons ensuite centrer réduit (normaliser) nos différentes variables afin de rendre comparable nos différentes variables et de supprimer l'influence des variables susceptible d'influencer notre modèle. A la suite nous allons représenter la matrice de distance pour déterminer la distance Euclidienne qui sépare nos différents d'individu. La distance euclidienne entre U et V est donnée par la formule suivante :

$$P_i = (x_1, x_2, \dots, x_m)$$

$$P_j = (y_1, y_2, \dots, y_m)$$

$$d(P_i, P_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

La Matrice des distances pour n individus est donnée par la formule :

$$D = (d(P_i, P_j))_{1 \leq i < n, 1 \leq j < n}$$

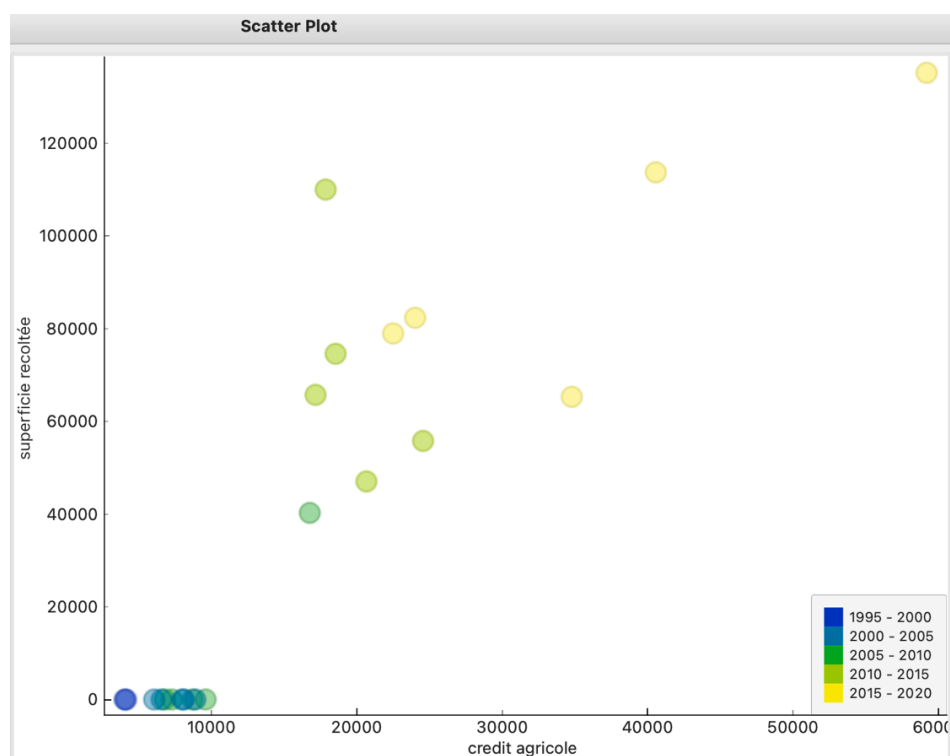
A la suite nous obtiendrons le dendrogramme produit par l'algorithme de clustering à l'aide de Hierarchical Clustering. Pour atteindre notre objectif nous allons calculer la distance moyenne entre les individus des deux Clusters. Cette dernière nous permettra ainsi de fusionner les groupes dont les individus sont les plus proches en moyenne. Les individus seront regroupés en classes à l'aide d'une ligne de coupe qui nous permettra d'avoir nos différents clusters.

CHAPITRE 2 : PRÉSENTATION ET ANALYSE DES RÉSULTATS

Le présent chapitre expose les résultats obtenus après le traitement des données et leurs interprétations.

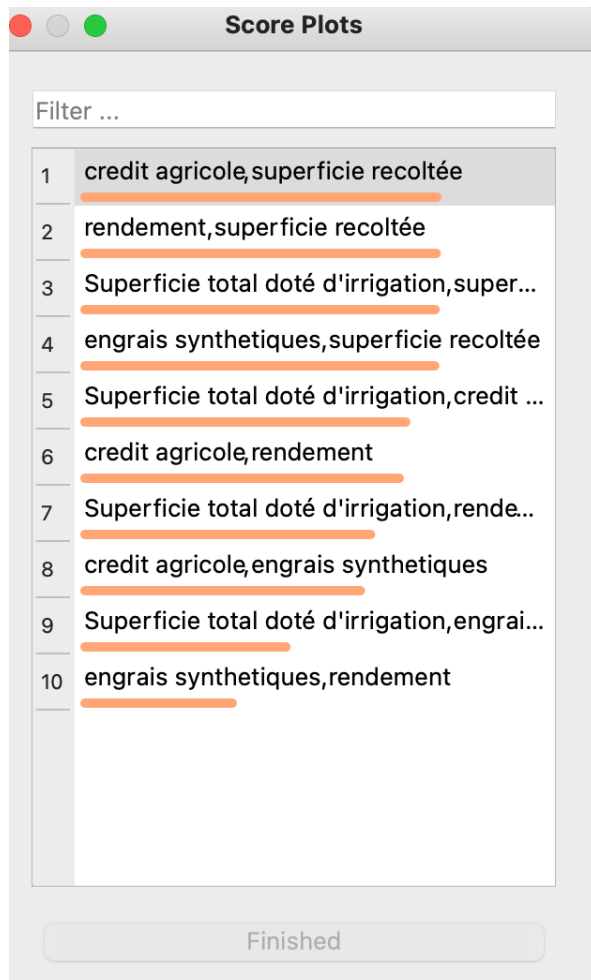
Nos différentes variables nous permettent d'avoir le nuage de points suivant :

Graphique 1 : Nuage de point 1



Le graphique 1 montre la représentation de la variable année à travers le nuage de point entre les variable crédit agricole sur l'axe des abscisses et la superficie récoltée sur l'axe des ordonnées. De l'analyse ce de graphique, on remarque que chaque année est représentée par un point. Sa position X correspond à la valeur de la variable Crédit agricole et sa position Y à la valeur Superficie récoltée. La couleur représente la variable quantitative année. On remarque certains points sont plus pochant que d'autre et qu'un groupe de points semblent se détacher des autres points.

Graphique 2 : Score et Plots



Le Graphique 2 montre les deux variables qui séparent au mieux les données en fonctions de l'année. De l'analyse de ce graphique on remarque que les deux variables qui séparent au mieux les données en fonctions de l'année sont le crédit agricole et la superficie récoltée.

Ainsi dans la suite de notre étude, nous utiliserons ces deux variables dans nos différentes analyses.

Graphique 3 : Distances

●
●
●
Distances

Distances between

☒ Rows
☐ Columns

Distance Metric

Euclidean

☒ Normalized

☐

Apply

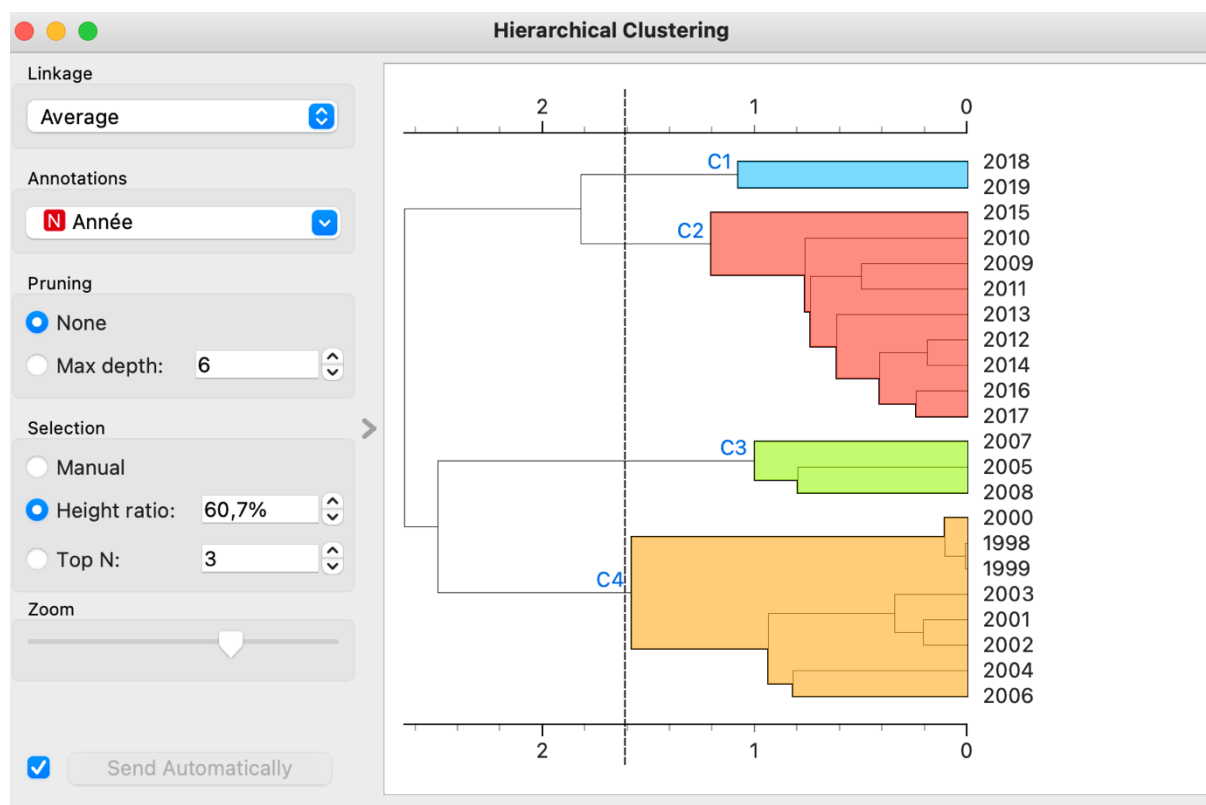
Le graphique 3 montre que nos variables seront standardisées et la mesure de la distance qui est utilisée est la distance euclidienne car nos données sont numériques.

Graphique 4 : Matrice de Distance

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1		0,006	0,107	1,282	1,378	1,498	1,868	2,672	1,888	3,413	3,044	2,948	2,557	3,220	2,967	3,293	2,974	3,353	3,221	3,227	4,554	3,870
2	0,006		0,102	1,282	1,377	1,497	1,867	2,671	1,888	3,412	3,044	2,946	2,555	3,218	2,966	3,292	2,973	3,350	3,220	3,225	4,550	3,868
3	0,107	0,102		1,275	1,365	1,483	1,859	2,664	1,883	3,409	3,036	2,925	2,522	3,185	2,944	3,270	2,948	3,302	3,188	3,196	4,486	3,818
4	1,282	1,282	1,275		0,205	0,441	0,989	2,114	1,207	3,012	2,383	2,070	1,944	2,355	2,256	2,668	2,315	2,746	2,525	2,501	4,033	3,241
5	1,378	1,377	1,365	0,205		0,240	0,854	2,028	1,029	2,921	2,244	1,881	1,760	2,171	2,078	2,519	2,143	2,576	2,353	2,328	3,889	3,083
6	1,498	1,497	1,483	0,441	0,240		0,729	1,945	0,805	2,822	2,087	1,668	1,556	1,976	1,877	2,355	1,950	2,404	2,168	2,140	3,760	2,930
7	1,868	1,867	1,859	0,989	0,854	0,729		1,235	0,819	2,109	1,395	1,548	1,517	1,874	1,775	2,274	1,863	2,109	2,081	1,977	3,695	2,874
8	2,672	2,671	2,664	2,114	2,028	1,945	1,235		1,840	0,931	0,794	2,209	2,202	2,440	2,361	2,756	2,429	2,194	2,592	2,408	3,918	3,251
9	1,888	1,888	1,883	1,207	1,029	0,805	0,819	1,840		2,563	1,707	1,251	1,139	1,665	1,435	2,021	1,536	2,108	1,819	1,773	3,624	2,716
10	3,413	3,412	3,409	3,012	2,921	2,822	2,109	0,931	2,563		1,067	2,795	2,788	2,992	2,880	3,213	2,937	2,541	3,082	2,876	4,252	3,673
11	3,044	3,044	3,036	2,383	2,244	2,087	1,395	0,794	1,707	1,067		1,791	1,905	2,038	1,966	2,425	2,062	1,873	2,218	2,016	3,675	2,937
12	2,948	2,946	2,925	2,070	1,881	1,668	1,548	2,209	1,251	2,795	1,791		0,707	0,497	0,520	1,162	0,690	1,390	0,818	0,751	2,735	1,735
13	2,557	2,555	2,522	1,944	1,760	1,556	1,517	2,202	1,139	2,788	1,905	0,707		0,826	0,555	1,104	0,570	1,204	0,774	0,794	2,538	1,611
14	3,220	3,218	3,185	2,355	2,171	1,976	1,874	2,440	1,665	2,992	2,038	0,497	0,826		0,603	1,030	0,680	1,133	0,559	0,533	2,274	1,287
15	2,967	2,966	2,944	2,256	2,078	1,877	1,775	2,361	1,435	2,880	1,966	0,520	0,555	0,603		0,709	0,186	1,254	0,456	0,433	2,500	1,468
16	3,293	3,292	3,270	2,668	2,519	2,355	2,274	2,756	2,021	3,213	2,425	1,162	1,104	1,030	0,709		0,575	1,420	0,554	0,608	2,240	1,217
17	2,974	2,973	2,948	2,315	2,143	1,950	1,863	2,429	1,536	2,937	2,062	0,690	0,570	0,680	0,186	0,575		1,205	0,359	0,386	2,378	1,350
18	3,353	3,350	3,302	2,746	2,576	2,404	2,109	2,194	2,108	2,541	1,873	1,390	1,204	1,133	1,254	1,420	1,205		1,059	0,960	1,820	1,206
19	3,221	3,220	3,188	2,525	2,353	2,168	2,081	2,592	1,819	3,082	2,218	0,818	0,774	0,559	0,456	0,554	0,359	1,059		0,237	2,059	1,015
20	3,227	3,225	3,196	2,501	2,328	2,140	1,977	2,408	1,773	2,876	2,016	0,751	0,794	0,533	0,433	0,608	0,386	0,960	0,237		2,144	1,126
21	4,554	4,550	4,486	4,033	3,889	3,760	3,695	3,918	3,624	4,252	3,675	2,735	2,538	2,274	2,500	2,240	2,378	1,820	2,059	2,144		1,075
22	3,870	3,868	3,818	3,241	3,083	2,930	2,874	3,251	2,716	3,673	2,937	1,735	1,611	1,287	1,468	1,217	1,350	1,206	1,015	1,126	1,075	

Le graphique 4 montre la distance entre les individus au lien avec leur année. Plus les cases sont vertes, plus les distances entre les variables sont grandes. Ainsi les individus 2 et 21 sont très éloignés environs 4,55 alors que les individus 1 et 2 sont plus proches environs 0,006.

Graphique 5 : Clustering Hiérarchique



Le graphique 5 montre les différentes classes les plus semblables qui sont formé. De ce graphique on note que quatre (04) clusters ont été constitué. Le clusters 1 regroupe les années 2018 et 2019. Il faut noter qu'au cours de ses année, le gouvernement dans sa nouvelle politique a mis à la disposition des producteurs une importante subvention dans le domaine de la production du riz local. Les crédits agricoles sont passée de 59 ;201 en 2018 à 40 ;571 en 2019. Ainsi les producteurs ont pu s'acquérir des engrais nécessaires à coût réduit pour pouvoir augmenter leur production, ce qui se traduit par une importante superficie récoltée pendant cette campagne. La production était aux alentours de 135.185ha en 2018 et 113.719ha en 2019, ce qui expliquent le regroupement de ses années en un cluster.

De même, le cluster 2 regroupent les années 2009, 2010, 2011, 2012, 2013, 2014, 2015 2016, et 2017. Cela s'explique par le fait que dans la continuité des actions menées par le gouvernement précédent, le nouveau gouvernement du nouveau président élu en 2014 dans sa politique a dans un premier temps continuer dans la même logique que le son prédécesseur, en accordant moins de

crédit agricole aux producteurs. Ce qui explique une moyenne production qui était de moins de 100.000ha observé durant ces années et une augmentation accrue des importations du riz au Bénin pour paliers à la production locale.

Le cluster 3 regroupe les années 2005, 2007 et 2008. Le nouveau gouvernement de l'ancien président de la république du Bénin le Dr Boni YAYI élu en 2006, a fourni d'importante effort dans la production locale. Ainsi dès le début de son mandat en 2006, l'ancien gouvernement a commencé à mettre en pratique son programme de subvention à la production locale. Ce qui explique une superficie de récolte considérable obtenue qui était de 28.904ha en 2005 à 33.294ha en 2008.

Le Cluster 4 regroupe les années 1998 à 2006. Cela s'explique par le fait que le programme d'action de l'ancien gouvernement de l'ancien président le Général Mathieu KÉRÉKOU était beaucoup plus basé sur la promotion de la production du Coton au détriment des autres produits. Ainsi durant ses années des subventions considérable ont été octroyé aux producteurs du Coton, ce qui a permis au pays d'être le premier producteur du l'or blanc (le coton) dans le monde. Dans le même temps le pays était très dépendant des importations des autres produits tels que riz. La production locale du riz était moins considérable. Elle était de mois de 25.000 de surface récoltée en durant toute ces périodes. Les difficultés d'accès aux crédits agricoles par les producteurs du riz s'expliquent par la faible du riz local durant ses années, c'est pour cela que ses années ont été regroupé dans le même clusters.

Conclusion

Tout au long de ce document, l'analyse a été centrée, sur l'approche hiérarchique de la productivité du riz au Bénin. La présente étude a eu pour objectif d'analyser la productivité du riz au Bénin. Elle s'est prioritairement intéressée à partir du dendrogramme produit par l'algorithme de clustering à constituer les groupes d'années les plus semblables en lien avec l'accès aux crédits agricoles et la superficie récoltée dans la productivité du riz au Bénin. Pour cela nous avons utilisé la classification par apprentissage non supervisé à partir de l'algorithme de clustering à l'aide du logiciel Orange pour constituer nos différentes classes. Des différents résultats, on retient que quatre (04) clusters ont été observés. Les années aux cours desquelles d'importantes subventions agricoles ont été octroyées sont regroupées ensemble et celles dont les subventions sont moins importantes sont mises ensemble.

Annexe :

Résumé de la chaîne mise en place

