



2021

Projet de fin d'année :
Développement touristique,
Croissance économique et
Réduction de la pauvreté : Une
analyse en termes de
Cointégration et Causalité

Ben Amor Aziz

2ème année

07/06/2021

Table des matières

I. Introduction	3
II. Croissance économique, développement humain et Activité touristique en Espagne.	4
A. La croissance économique en Espagne	5
B. Développement économique en Espagne : Pauvreté ...	5
C. Le tourisme en Espagne	6
III. Méthodologie d'analyse	6
A. Tests de racine unitaire	6
B. Processus VAR(), méthodologie de BJ, et phénomène de causalité (Sims, Granger-Engle, Toda-Yamamoto...).....	9
C. Approche de Cointégration et modèle à correction d'erreur par l'approche de Johansen	10
1. Approche de coïntégration de Johansen :	10
2. Modèle Vectorielle à Correction d'Erreur VECM :	12
IV. Application.....	12
A. Présentation des données.....	12
B. Analyse exploratoire des données	13
C. Étude des processus PIB-Tourisme-Pauvreté.....	18
1. La variable PIB (croissance économique).....	18
2. La variable représentant le tourisme.....	20
3. La variable représentant la pauvreté :.....	21
D. Processus VAR () & test de causalité de Sims, Granger- Engle.....	21
1. Processus VAR.....	21
2. Étude, Estimation et BJ-test.....	22

3. Sélection du retard optimale	27
4. Phénomène de causalité Sims, Granger-Engle	28
E. Phénomène de cointégration par l'approche de Johansen & modèle à correction d'erreur (ECM)	30
V. Conclusion	34
VI. Références et Bibliographie.....	36
A. Applications utilisées.....	36
B. Base de données	36
C. Bibliographie	36

I. Introduction

Dans ce projet, le principal objectif est de permettre une meilleure compréhension de ces liens de causalité. Plus spécifiquement, cette étude vise à :

Déterminer la relation de long terme entre les trois variables.

Déterminer la direction de la causalité entre les variables prises deux à deux.

Déduire les implications en termes de réduction de la pauvreté de la relation éventuelle entre croissance économique et développement du tourisme.

Donc, on doit répondre aux questions suivantes :

- Si le développement du tourisme engendre de la croissance à long terme, quel serait, dans ce cas, le niveau de croissance économique résultant d'une variation de la demande touristique ?
- S'il existe, une relation inverse entre croissance économique et pauvreté, quel serait, dans ce cas, le niveau de pauvreté qui résulterait d'une variation de l'activité économique ?

- Si une réduction de la pauvreté est observée grâce à l'activité touristique, quel serait dans ce cas, le niveau de cette réduction ?
- S'il existe une causalité unidirectionnelle / bidirectionnelle entre croissance économique et développement touristique, le développement touristique engendrerait-il dans ce cas de la croissance économique ou la croissance économique engendrerait-elle du développement touristique ?

Pour répondre à ces quatre questions, nous utiliserons les développements récents de l'économétrie des séries temporelles. La méthodologie adoptée est une approche en trois étapes :

Tests de racine unitaire (ADF), tests de cointégration de Johansen, tests de causalité de Granger dans le cadre d'un modèle vectoriel à correction d'erreurs.

Ce Projet est articulé comme suit :

La première section, nous présenterons les principales caractéristiques et les diverses phases d'évolution de la croissance économique, du développement humain et de l'activité touristique en Espagne, la deuxième sera consacrée à la méthodologie d'analyse, une troisième présentera les résultats obtenus et en finira par une conclusion.

II. Croissance économique, développement humain et Activité touristique en Espagne.

A. La croissance économique en Espagne

L'Espagne possède aujourd'hui une économie diversifiée, grâce à la croissance rapide de l'industrie depuis les années 1950 et à l'essor du tourisme. Entre 1995 et 2001, les emplois industriels ont augmenté de 38%. À partir de 1964, une série de plans de développement a contribué à l'expansion économique de ce pays.

L'Espagne est devenue au moins de 20 ans une grande puissance industrielle et agricole plus qu'au tourisme qui représente 5% du PIB.

En 2017, l'Espagne compte deux fois plus sur les super-riches qu'avant la crise de 2008 en lien avec la crise mondiale.

Près de 50% du PIB est détenu par 0.4% de la population.

B. Développement économique en Espagne : Pauvreté

Selon l'UNICEF, l'Espagne est l'un des pays développés où la pauvreté infantile est la plus élevée.

En 2017, plus de 1.400.000 enfants vivent dans un état de grande pauvreté.

En 2018, l'INS indique que 21.6% de la population vit dans la pauvreté.

Les régions les plus affectées sont :

- Estrémadure 38.9%
- Andalousie 31%
- Îles Canaries 30.5%

En 2019, la commission de l'ONU sur la pauvreté et les droits de l'homme indique dans son rapport que 26.1% de la population vit en pauvreté (29.5% pour les enfants).

C. Le tourisme en Espagne

La situation géographique de l'Espagne, ses paysages diversifiés, son historique, sa culture et ses excellentes infrastructures ont fait de l'industrie touristique internationale du pays l'un des plus importants au monde.

Au cours des 5 dernières décennies, le tourisme est devenu le 2^{ème} marché mondial en termes de dépenses, représentant environ 40 milliards d'euros (5% du PIB) en 2006.

En 2017, l'Espagne était le 2^{ème} pays le plus visité au monde, avec 82 millions de touristes.

III.Méthodologie d'analyse

A. Tests de racine unitaire

Une série chronologique est stationnaire si elle est la réalisation d'un processus stationnaire.

Plus généralement, la série ne comporte aucun facteur évoluant avec le temps.

Pour vérifier la stationnarité des séries, il faut appliquer des tests de stationnarité ou des tests de racine unitaire. Ces tests permettent d'identifier la présence de racine unitaire dans une série temporelle et donc de vérifier si elle est stationnaire.

Plus précisément, on cherche à vérifier l'hypothèse nulle d'existence de racine unitaire et que le processus autorégressif est non stationnaire contre l'hypothèse alternative.

Parmi les tests qui conviennent à cette fin, nous retenons le test **Augmented Dickey-Fuller (ADF)** (Smith et Fuertes, 2010).

Le test ADF prévoit trois modèles pour la régression selon les éléments déterministes rajoutés :

- Cas sans terme déterministe et sans constante
 1. $\Delta x_t = (\rho - 1)x_{t-1} + \sum_{j=2}^k \varphi_j \Delta x_{t-j+1} + \varepsilon_t$
- Cas avec constante
 2. $\Delta x_t = (\rho - 1)x_{t-1} + \sum_{j=2}^k \varphi_j \Delta x_{t-j+1} + \gamma + \varepsilon_t$
- Cas avec constante et tendance
 3. $\Delta x_t = (\rho - 1)x_{t-1} + \sum_{j=2}^k \varphi_j \Delta x_{t-j+1} + \gamma + \beta t + \varepsilon_t$

Les paramètres des relations sont estimés par MCO, les résultats obtenus permettent de calculer la statistique du test afin de déterminer s'il faut rejeter l'hypothèse nulle de la non-stationnarité.

Dans les trois cas, il revient de vérifier l'hypothèse $H_0 : \rho = 0$
Les tests sur les variables peuvent être menés en niveau et en différence.

Cependant, l'originalité du test réside dans le choix du nombre de retards du modèle VAR (p). Si le nombre de retards est faible, le test risque d'être biaisé et s'il est important, le test sera moins puissant.

Plusieurs méthodes peuvent être utilisées pour optimiser le nombre de retards (critère d'information d'**Akaike** [1973], de Gideon **Schwartz** [1978] et de **Hannan-Quinn** [1979]).

Nous retenons le critère d'information d'Akaike :

$$I_{AIC} = -2 \log \left(\frac{l}{T} \right) + 2 \left(\frac{k}{T} \right)$$

Avec :

- l Est le logarithme de la fonction de vraisemblance
- k Est le nombre de paramètres de la fonction
- T Est le nombre d'observations

Akaike propose de minimiser la valeur du critère en considérant des retards séquentiels après avoir estimé plusieurs relations du modèle.

Dans le cadre de ce travail, on doit déterminer l'ordre d'intégration $I(.)$ de la série. En effet, une série X_t est :

- Intégrée d'ordre 0, $I(0)$, si X_t est stationnaire.
- $I(1)$, si la première différence ΔX_t est stationnaire.

Un test sur l'intégration cherche à trouver le nombre de différentiations nécessaires à une variable pour qu'elle soit stationnaire. (Smith et Fuertes, 2010).

B. Processus VAR(), méthodologie de BJ, et phénomène de causalité (Sims, Granger-Engle, Toda-Yamamoto...)

La relation de causalité peut néanmoins nous indiquer des informations sur l'antériorité des événements entre deux variables.

Selon **Hurlin** (2007) deux principes doivent être vérifiés dans toute relation causale. Il s'agit du principe d'antériorité, « la cause précède l'effet » et en second lieu, la série causale contient de l'information sur l'effet, qui n'est contenue dans aucune autre série, au sens de la distribution conditionnelle.

Le test classique est celui de Granger (1969), le théorème repose sur la valeur des prévisions des variables à partir de leur passé.

Selon Bourbonnais (2002), supposer que la variable Y cause X veut dire que la prédictibilité de X est améliorée lorsque l'information relative à Y est incluse dans l'analyse.

La définition de **Sims** (1972) de la causalité affirme qu'une série cause une autre si les innovations de la première contribuent à la variance d'erreur de prévision de la seconde.

Selon le test de Toda et Yamamoto, l'objectif n'est pas d'étudier la coïntégration, mais plutôt de vérifier les hypothèses théoriques. Leur procédure consiste à étudier la causalité à partir d'un VAR d'ordre $(p + d_{\max})$ {où p est le nombre de retards optimal et d_{\max} est l'ordre d'intégration maximale} en appliquant les tests de restriction uniquement

sur les k premiers coefficients sans tenir compte de la cointégration.

La démarche de ce dernier implique l'utilisation d'une statistique de **WALD** de l'hypothèse de non-causalité du test classique de Granger (1969).

Le Test Jarque-Bera ayant comme hypothèse nulle la normalité de la distribution des résidus, utilise les valeurs de Skewness et Kurtosis pour évaluer cette hypothèse.

On rappelle qu'une distribution normale aie une Skewness et Kurtosis nulles.

Donc plus qu'ils s'approchent de 0 plus la distribution semble à une distribution normale.

Le signe de Skewness donne une information sur le sens d'asymétrie (gauche ou droite).

Celle du Kurtosis nous informe sur le poids du queue de la distribution (légère ou lourde).

C. Approche de Cointégration et modèle à correction d'erreur par l'approche de Johansen

1. Approche de cointégration de Johansen :

L'étude de la cointégration permet de tester l'existence d'une relation (une combinaison linéaire) de long terme entre deux variables non-stationnaires, en incluant des variables retards et des variables exogènes. Il existe plusieurs tests de la cointégration, le plus général étant celui de Johansen. Quel que soit le test retenu, il n'a de signification que sur des séries non stationnaires longues. Par conséquent, l'analyse de la cointégration permet d'identifier clairement la relation

véritable entre deux variables, en recherchant l'existence d'un vecteur de cointégration et en éliminant son effet le cas échéant.

Deux séries X et Y sont dites cointégrées si les deux conditions suivantes sont vérifiées : elles sont affectées d'une tendance stochastique de même ordre d'intégration et une combinaison linéaire de ces séries permet de se ramener à une série d'ordre d'intégration inférieur.

Le test de Johansen (1988) propose des estimateurs du maximum de vraisemblance pour tester la cointégration des séries, il effectue un test de rang de cointégration.

Ce dernier s'opère dans un univers multivarié des séries temporelles, la méthode propose une analyse des valeurs propres de la matrice du modèle pour tester la cointégration.

La forme générale d'un processus multivarié peut être exprimé de la sorte :

$$X_t = \pi_1 X_{t-1} + \dots + \pi_k X_{t-k} + \mu + \varphi D_t + \varepsilon_t$$

Où :

- $t = 1, \dots, T$
- ε_t est un bruit blanc
- D_t et μ sont des éléments déterministes ou exogènes.

Après interprétation de la matrice

$$\pi(p * p) = \pi_1 + \dots + \pi_k + I$$

Le rang 'r' de cette matrice indique le niveau de cointégration (le nombre de valeurs de cointégration).

À cet effet, il devient nécessaire de connaître les valeurs propres de la matrice pour les évaluer, elles doivent être significativement différentes de 0.

Deux statistiques nous permettent de calculer deux types de valeurs propres : la trace et la valeur propre maximale :

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^p \log(1 - \lambda_i)$$

$$\lambda_{max}(r, r + 1) = -T \log(1 - \lambda_{r+1})$$

Les hypothèses du test sont :

H_0 : il existe au plus r vecteurs de coïntégration

H_1 : il existe au moins r vecteurs de coïntégration

2. Modèle Vectorielle à Correction d'Erreur VECM :

Le VECM a pour but d'analyser la vitesse d'ajustement des variables en cas de déséquilibre à long terme.

On l'utilisera pour comparer les pouvoirs d'ajustement des trois variables.

IV. Application

A. Présentation des données

Les bases de données de la croissance économique ont été extraites du site (DATBANK), les variables ont été : PIB en milliards de dollars à partir de 1960, après correction et nettoyage de la base il y a eu construction d'une nouvelle

variable représentant l'évolution du PIB en utilisant la formule :

$$GPIB = \frac{PIB_t - PIB_{t-1}}{PIB_{t-1}} ; t=1961, \dots, 2019.$$

L'extraction des données pour la pauvreté ont été de la part du site (DATA.OCDE). les variables ont été :

- L'indice de GINI (estimé par World Bank)
- Ratio de population vivant sous le seuil de 1.9 \$ / jour
- Ratio de population vivant sous le seuil de 3.2 \$ / jour
- Ratio de population vivant sous le seuil de 5.5 \$ / jour

Sur une période de 1980 jusqu'à 2019.

Après j'ai construit les variables présentant l'évolution des variables ci-dessus.

'MacroTrends' a été la source des données représentant le développement touristique sur une période allant de 1995 jusqu'à 2019 représentant les recettes en billions de dollars et le nombre des touristes.

J'ai donc construit l'évolution de ces variables et le pourcentage des recettes touristique du PIB.

Le modèle contenant 3 variables représentant la croissance économique, le développement touristique et la pauvreté a été mis en place après analyse exploratoire (B) sur une période allant de 1995 jusqu'à 2019 et donc contient en total 24 observations.

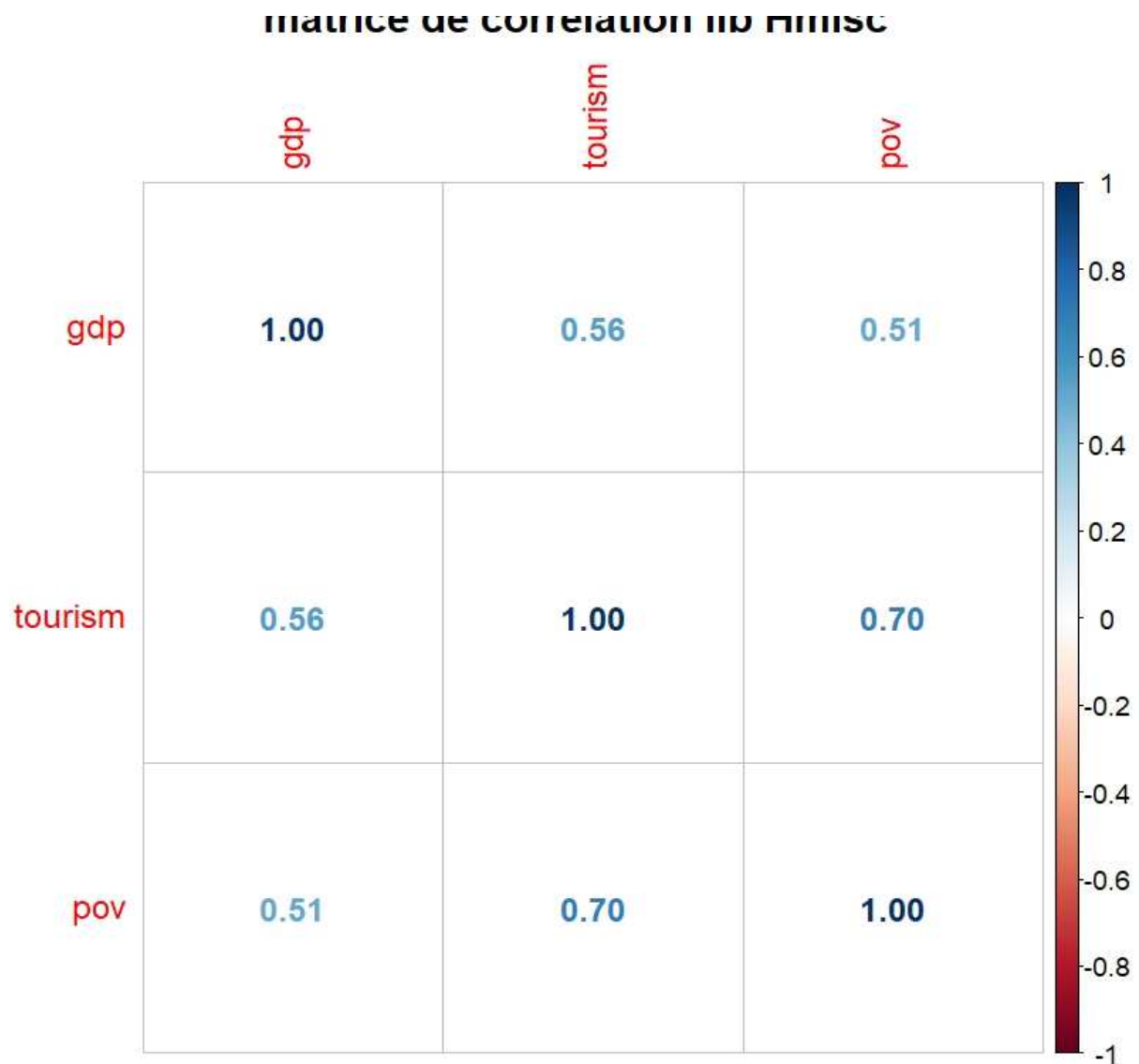
B. Analyse exploratoire des données

Pour la sélection des variables à utiliser pour représenter la croissance économique, le tourisme et la pauvreté.

Il fallait observer les matrices de corrélation de différentes composantes.

À l'aide d'une fonction R que j'ai créé, qui prendra en paramètre 3 variables de la base de données et retourne leur matrice de corrélation, j'ai obtenu les matrices suivantes contenant la corrélation entre les variables et le symbole (x) se trouve où la p-value est inférieur au seuil de signification que j'ai fixé à 5%. L'hypothèse nulle de ce teste (H_0) indique l'absence de corrélation entre les variables.

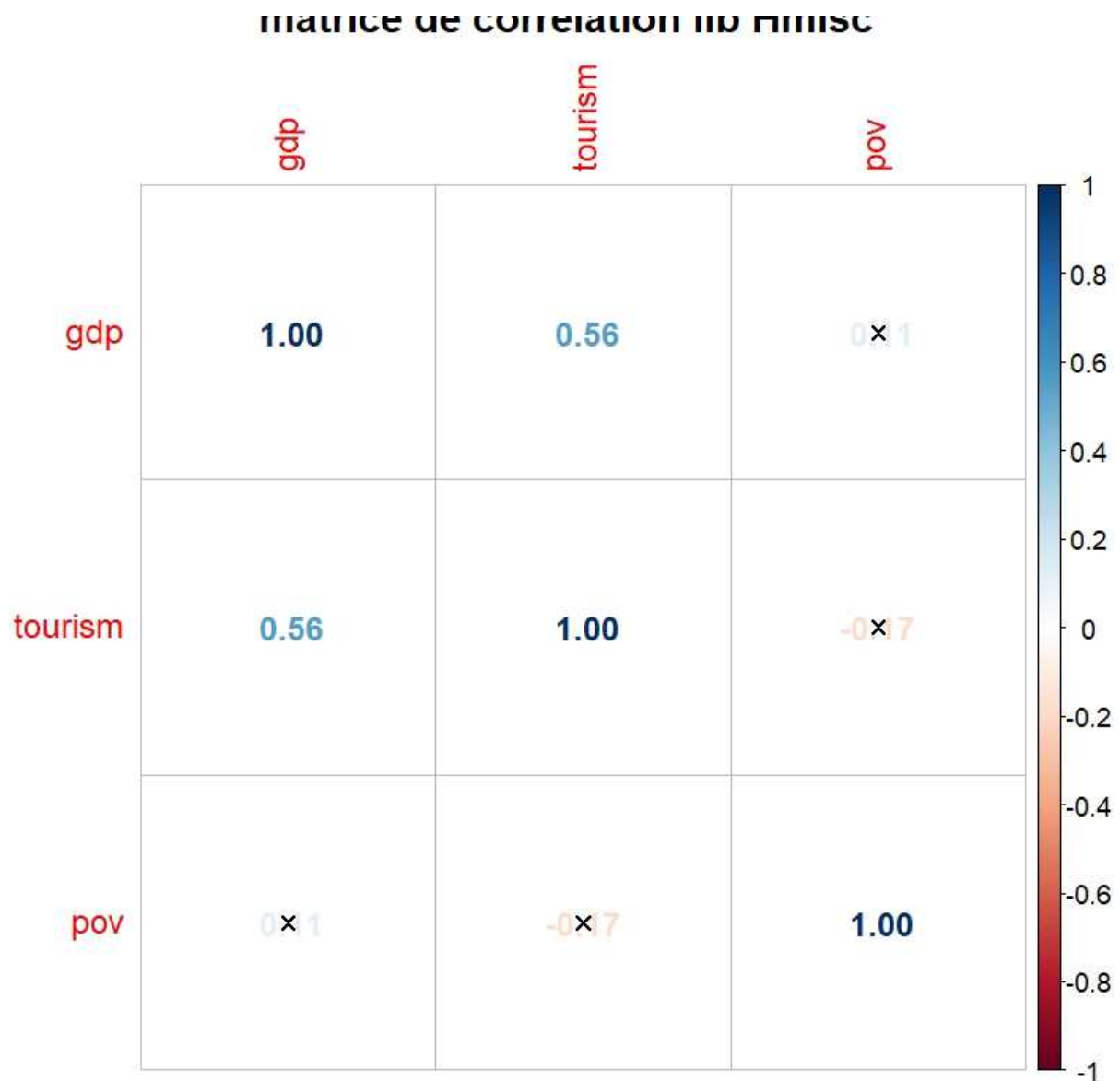
Ci-dessous les matrices obtenues :



APERÇU-1 : VARIABLES (GDB EN \$, RECETTE DE TOURISME EN \$, L'INDICE DE GINI)

Cette visualisation indique la présence d'une corrélation entre les variables statistiquement significative et importante en valeur :

70% entre pauvreté et tourisme ,56% entre tourisme et PIB et 51% entre PIB et pauvreté.



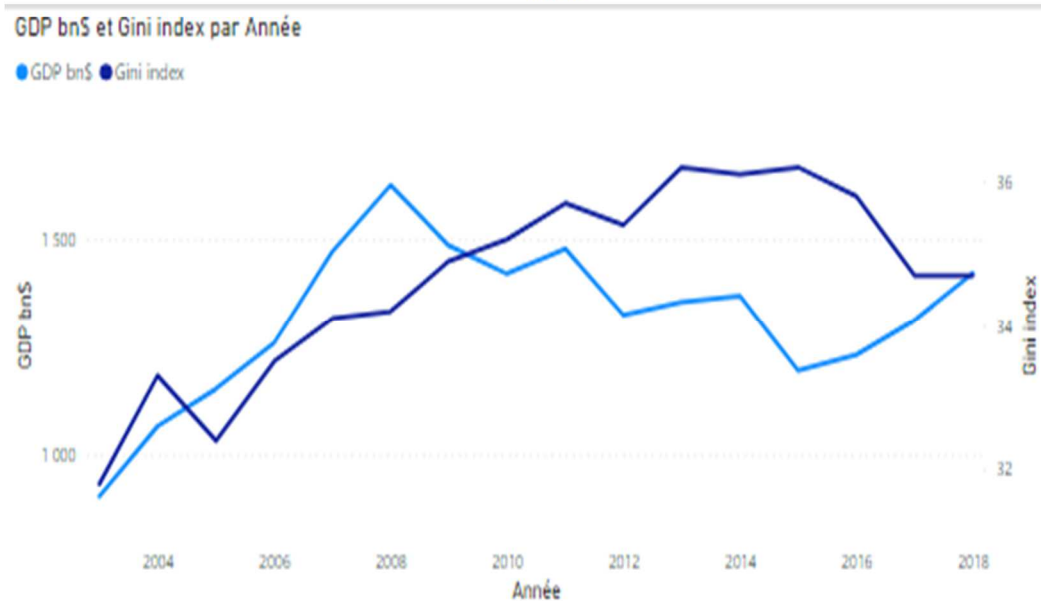
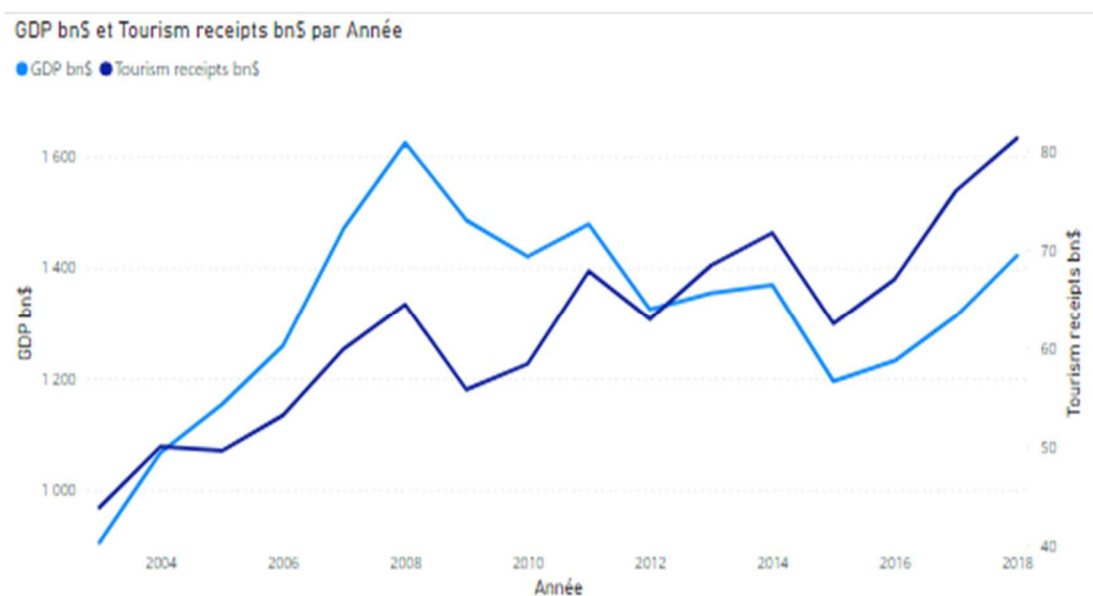
APERCU-2 : VARIABLES (PIB EN \$, RECETTE DE TOURISME \$, % POPULATION VIVANT À 5.5\$/JOUR)

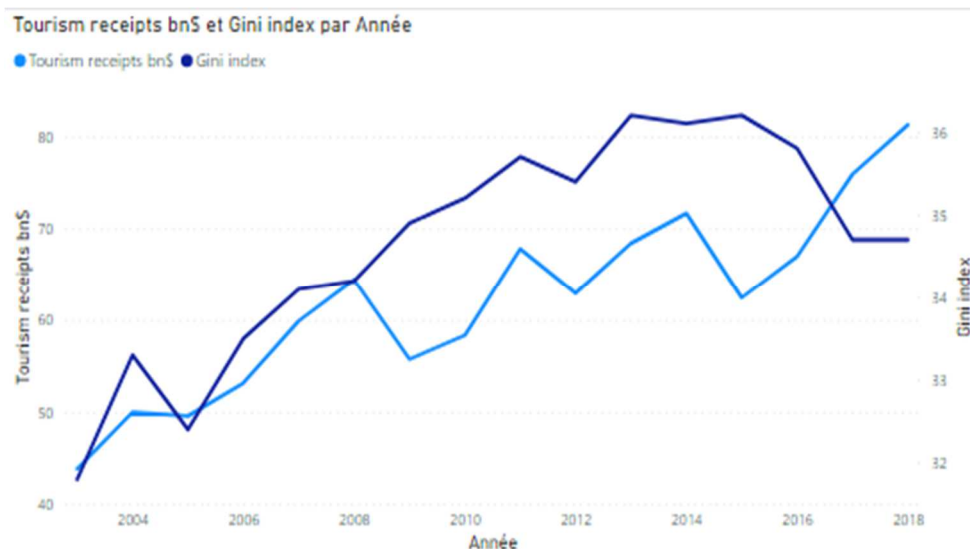
Cette matrice indique l'absence de corrélation entre la variable pauvreté et le reste des variables, donc on ne choisira pas ce modèle à cause de la présence d'un modèle ayant une corrélation significative entre les variables (Aperçu-1).

Après avoir tester des modèles contenant différentes variables.

On a sélectionné :

- ⇒ PIB pour représenter la croissance économique ;
 - ⇒ Les recettes du tourisme pour représenter leur domaine ;
 - ⇒ L'indice de Gini pour représenter la pauvreté ;
- Leurs représentations graphiques :





C. Étude des processus PIB-Tourisme-Pauvreté

On a effectué le test de racines unitaires augmenter de Dickey-Fuller pour tester la stationnarité des variables et leurs degrés d'intégration.

1. La variable PIB (croissance économique)

Après détermination de l'ordre de décalage optimale (lag=3), On peut tester la stationnarité de cette variable.

Ci-dessous le résultat obtenu :

Null Hypothesis: _GDP_BN\$ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on AIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.949963	0.1692
Test critical values: 1% level	-4.498307	
5% level	-3.658446	
10% level	-3.268973	

$P = 17\% > 5\%$ et la statistique $t = -2.9$ est supérieure aux valeurs critiques.

⇒ PIB est non stationnaire.

Donc on ressaye le test avec intégration de la variable qui aie un décalage optimal égale à 8.

Null Hypothesis: D(_GDP_BN\$) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 7 (Automatic - based on AIC, maxlag=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.598924	0.0649
Test critical values: 1% level	-4.728363	
5% level	-3.759743	
10% level	-3.324976	

Après intégration on trouve $P = 6\% > 5\%$ et la statistique $t = -3.59$ est supérieure à la valeur critique de niveau 10% qui est égale à -3.32.

⇒ La variable reste non stationnaire.

Donc on ressaye avec la deuxième différence qui aye un lag optimal égale à 8.

Null Hypothesis: _D2GDP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on AIC, maxlag=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.677254	0.0010
Test critical values: 1% level	-4.498307	
5% level	-3.658446	
10% level	-3.268973	

Finalement, on trouve un $p\text{-value} = 0.1\% < 5\%$ et une statistique $t = -5.67$ inférieur au valeurs critiques.

⇒ Le rejet de l'hypothèse nulle affirmant la non-stationnarité de la variable et donc on accepte que PIB soit stationnaire après intégration d'ordre 2.

2. La variable représentant le tourisme

Le décalage optimal est : 3

Ci-dessous le résultat obtenu à partir de l'ADF :

Null Hypothesis: _TOURISM_RECEIPTS_BN\$ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on AIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.273259	0.4280
Test critical values:		
1% level	-4.498307	
5% level	-3.658446	
10% level	-3.268973	

P = 42% >>5% et la statistique t = -2.27 est supérieurs aux valeurs critiques.

⇒ La variable niveau est non-stationnaire.

On essaye avec la différence de 1ere ordre avec lag=2 :

Null Hypothesis: _DTOURISM has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on AIC, maxlag=2)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.512689	0.0012
Test critical values:		
1% level	-4.467895	
5% level	-3.644963	
10% level	-3.261452	

On a P= 0.1%<<5% et la valeur de la statistique t = -5.51 est inférieur aux valeurs critiques (-4.46, -3.62, -3.26).

⇒ Cette variable est stationnaire après intégration de premier ordre.

3. La variable représentant la pauvreté :

Le décalage optimal est égal à 3

Ci-dessous les résultats présentés par l'ADF

Null Hypothesis: _GINI_INDEX has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on AIC, maxlag=3)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.730922	0.0051
Test critical values: 1% level	-4.416345	
5% level	-3.622033	
10% level	-3.248592	

*Mackinnon (1996) one-sided p-values.

$P = 0.5\% \ll 5\%$ et la statistique $t = -4.74$ est inférieure aux valeurs critiques $(-4.41, -3.62, -3.24)$.

Donc cette variable (niveau) est stationnaire ($I(0)$)

D. Processus VAR () & test de causalité de Sims, Granger-Engle

1. Processus VAR

Notre modèle présente trois variables (PIB, Pauvreté et Tourisme).

On distingue alors que l'équation représentant ce modèle soit de la forme suivante :

$$M_t = \begin{pmatrix} T_t \\ PIB_t \\ P_t \end{pmatrix} = \alpha + \sum_{i=1}^p \beta_i M_{t-i} + \varepsilon_t$$

Avec :

- M_t Vecteur de dimension 3
- E_t PIB de l'année numéro t en I(2) stationnaire
- T_t le Tourisme de l'année numéro t en I(1) stationnaire
- P_t Pauvreté de l'année numéro t stationnaire
- β_i Est le vecteur des paramètres du modèle dimension (3 x 3)
- ε_t Le vecteur d'erreur
- α Est le vecteur de dimension 3
- $p = 1$, le retard optimal (AIC, SC et HQ) (3)

2. Étude, Estimation et BJ-test

Estimation par MCO :

	_DTOURISM	_D2GDP	__GINI_INDEX
_DTOURISM(-1)	-0.843650 (0.37238) [-2.26559]	-26.88670 (7.62506) [-3.52610]	0.638471 (0.91699) [0.69627]
_D2GDP(-1)	0.034223 (0.01493) [2.29290]	0.576261 (0.30563) [1.88547]	-0.017782 (0.03676) [-0.48380]
__GINI_INDEX(-1)	0.029624 (0.07323) [0.40454]	0.026182 (1.49951) [0.01746]	0.536444 (0.18033) [2.97475]
C	3.492416 (2.17126) [1.60848]	64.68919 (44.4604) [1.45499]	12.51804 (5.34683) [2.34121]

R-squared	0.252276	0.489239	0.392789
Adj. R-squared	0.120325	0.399104	0.285635
Sum sq. resids	391.2808	164063.5	2372.784
S.E. equation	4.797553	98.23845	11.81421
F-statistic	1.911887	5.427880	3.665625
Log likelihood	-60.50919	-123.9143	-79.43432
Akaike AIC	6.143732	12.18232	7.946126
Schwarz SC	6.342689	12.38127	8.145083
Mean dependent	2.504762	7.720299	28.02381
S.D. dependent	5.115153	126.7307	13.97798
<hr/>			
Determinant resid covariance (dof adj.)	4657613.		
Determinant resid covariance	2470883.		
Log likelihood	-243.9540		
Akaike information criterion	24.37657		
Schwarz criterion	24.97344		
Number of coefficients	12		

On trouve entre crochets [] la valeur du test de Student t qui peuvent aussi être interprété par la formule : $t = \frac{coeff}{std\ error}$ avec un erreur standard écrit entre parenthèses ()).

Ayant 21 observations incluse et 3 paramètres on aura :

La statistique t suit la loi de Student à $(21-3-1 = 17)$ degrés de libertés. (Charpentier)

En lisant sur le tableau de Student on trouve à 5% une valeur de **2.11**.

On compare alors les valeurs du modèle à cette valeur pour accepter ou rejeté l'hypothèse nulle $H_0 : \beta_i = 0$

Pour le Tourisme : on accepte l'hypothèse nulle pour la pauvreté ($t = 0.4 < 2.11$)

Pour le PIB : on ne rejette H_0 que pour le Tourisme ($t=3.52$).

Pour la pauvreté on ne rejette H_0 que pour la variable de pauvreté et la constante.

Soient :

$$\beta_T = (-0.84 \quad 0.03 \quad 0)$$

$$\beta_{PIB} = (-26.88 \quad 0 \quad 0)$$

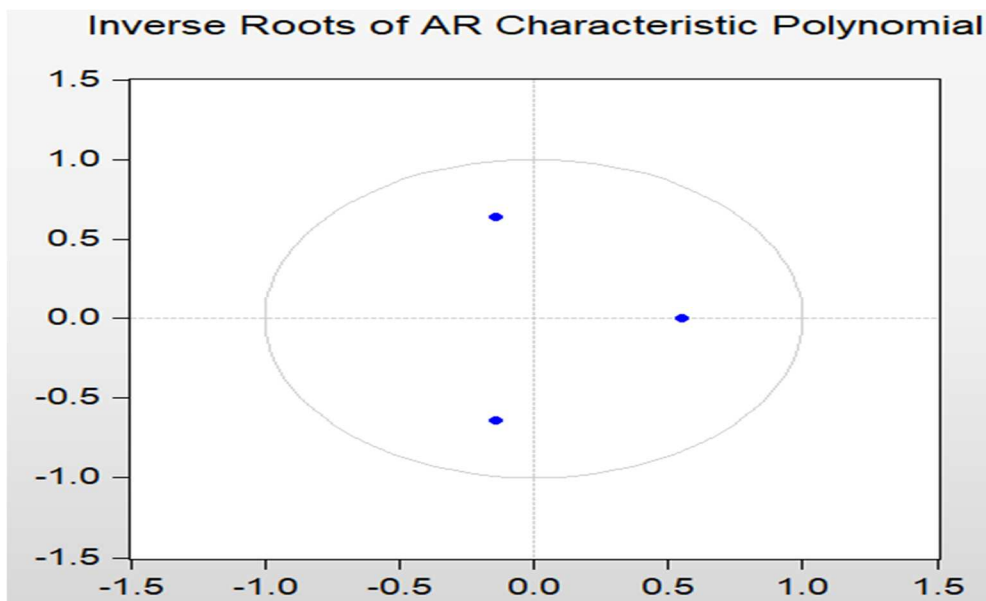
$$\beta_P = (0 \quad 0 \quad 0.53)$$

$$\alpha = \begin{pmatrix} 0 \\ 0 \\ 12.5 \end{pmatrix}$$

$$M_t = \begin{pmatrix} T_t \\ PIB_t \\ P_t \end{pmatrix}$$

On aura l'équation VAR du modèle suivante :

$$M_t = \alpha + \begin{pmatrix} \beta_T \\ \beta_{PIB} \\ \beta_P \end{pmatrix} M_{t-1}$$



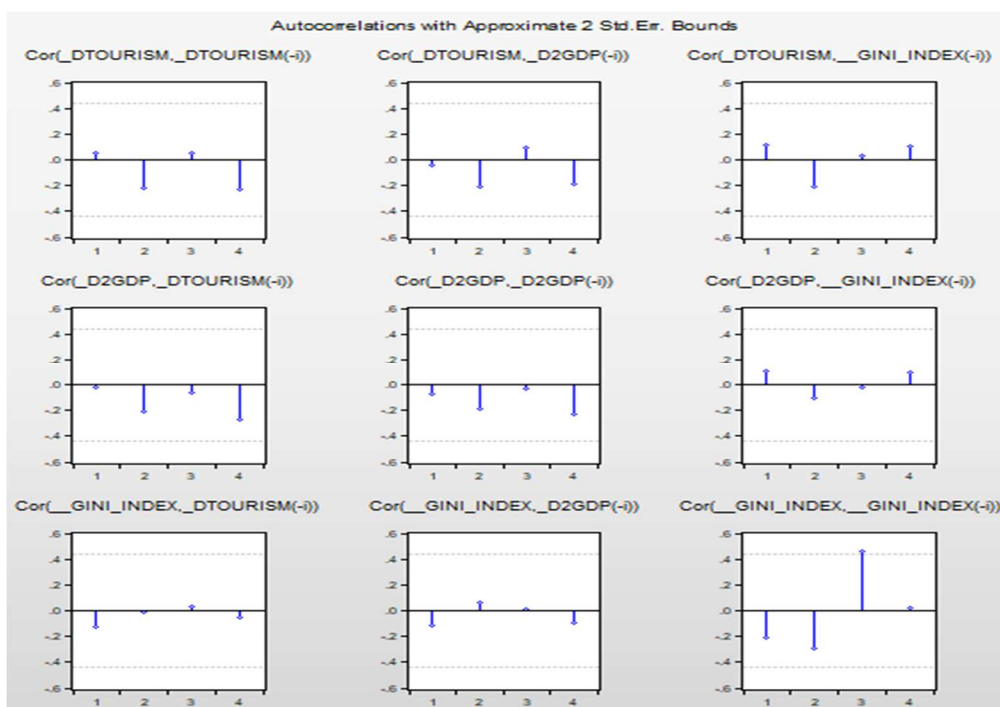
- ⇒ Les racines inverses du polynôme caractéristique du processus autorégressif (AR) du modèle sont toutes à l'intérieur du cercle unité donc **le processus estimé est stationnaire en 'covariance'**.
- ⇒ **On conclue que la VAR estimé ci-dessus est STABLE.**

Null hypothesis: No serial correlation at lag h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	8.118807	9	0.5222	0.914311	(9, 29.4)	0.5264
2	6.018008	9	0.7381	0.655651	(9, 29.4)	0.7409
3	7.687488	9	0.5659	0.859847	(9, 29.4)	0.5699
4	4.123789	9	0.9031	0.436189	(9, 29.4)	0.9042
5	3.595738	9	0.9360	0.377232	(9, 29.4)	0.9367

Ci-dessus, on observe les résultats du test LM d'autocorrélation ayant pour hypothèse nulle : l'absence d'autocorrélation.

Pour un retard de 1 ; on observe une p-value égale à 52.6% donc supérieure au seuil de significativité de 5%.

⇒ On accepte l'hypothèse nulle et alors **absence d'autocorrélation.**



⇒ Pour tous les couples des variables, on obtiendra des valeurs dedans l'intervalle de confiance (présenté sur les graphiques) [-4., 4.] dont le retard est égal à 1.

⇒ **Absence du MA. Donc on accepte ce modèle.**

Après on applique le test ADF sur les variables respectivement tourisme, PIB et pauvreté pour tester la stationnarité des modèles VAR estimer on obtient les résultats suivants :

Null Hypothesis: ZVARTOURISM has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on AIC, maxlag=1)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.990975	0.0268
Test critical values: 1% level	-4.498307	
5% level	-3.658446	
10% level	-3.268973	

T_t est bien stationnaire

Null Hypothesis: ZVARD2GDP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on AIC, maxlag=1)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.941281	0.0006
Test critical values: 1% level	-4.498307	
5% level	-3.658446	
10% level	-3.268973	

PIB_t est Stationnaire

Null Hypothesis: ZVARPOV has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on AIC, maxlag=1)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.825293	0.0045
Test critical values: 1% level	-4.440739	
5% level	-3.632896	
10% level	-3.254671	

P_t est Stationnaire

⇒ On affirme que le modèle M_t est Stationnaire

Component	Skewness	Chi-sq	df	Prob.*
1	-0.607081	1.289915	1	0.2561
2	-0.383701	0.515292	1	0.4729
3	-0.364707	0.465539	1	0.4950
Joint		2.270747	3	0.5181

Component	Kurtosis	Chi-sq	df	Prob.
1	2.835905	0.023561	1	0.8780
2	1.964139	0.938882	1	0.3326
3	3.706204	0.436384	1	0.5089
Joint		1.398828	3	0.7058

Component	Jarque-Bera	df	Prob.
1	1.313476	2	0.5185
2	1.454175	2	0.4833
3	0.901923	2	0.6370
Joint	3.669575	6	0.7213

Le test BJ retourne une P-value >> 5% pour toutes les variables du modèle donc on accepte l'hypothèse nulle qui énonce la normalité de la distribution des résidus.

Une 'Skewness' négative et au voisinage de 0 indique la quasi-symétrie des résidus plus qu'une faible asymétrie à gauche.

La 'Kurtosis' positive indique une distribution à queue lourde.

3. Sélection du retard optimale

Premièrement, on doit choisir le décalage (lag) optimale. Pour cela on utilise le critère de sélection de l'ordre de décalage VAR présenté sur Eviews.

VAR Lag Order Selection Criteria
 Endogenous variables: __GINI_INDEX _D2GDP _D2TOURISM
 Exogenous variables: C
 Date: 06/06/21 Time: 23:53
 Sample: 1 24
 Included observations: 18

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-227.0014	NA	25201598	25.55571	25.70410	25.57617
1	-208.4448	28.86568*	8911578.*	24.49387*	25.08745*	24.57572*
2	-202.7031	7.017656	14244472	24.85590	25.89467	24.99913
3	-191.7078	9.773635	15333067	24.63420	26.11815	24.83882
4	-184.9976	3.727886	40438954	24.88862	26.81776	25.15462

1 est le décalage optimal ayant :

Le critère d'information d'Akaike minimal : 24,49.

Le critère d'information de Schwartz minimal : 25,08.

Et Le critère d'information de Hannan-Quinn minimal : 24,57.

4. Phénomène de causalité Sims, Granger-Engle

Le test de causalité au sens de Granger indique ces deux hypothèses :

- H_0 : X ne cause pas Y au sens de Granger
- H_a : X cause Y au sens de Granger

En effectuant ce test avec l'ordre de décalage optimale déjà sélectionné, On obtient les résultats ci-dessous :

VAR Granger Causality/Block Exogeneity Wald Tests
Date: 06/06/21 Time: 22:12
Sample: 1 24
Included observations: 21

Dependent variable: _DTOURISM

Excluded	Chi-sq	df	Prob.
_D2GDP	5.257410	1	0.0219
__GINI_INDEX	0.163652	1	0.6858
All	5.283731	2	0.0712

Dependent variable: _D2GDP

Excluded	Chi-sq	df	Prob.
_DTOURISM	12.43335	1	0.0004
__GINI_INDEX	0.000305	1	0.9861
All	13.18021	2	0.0014

Dependent variable: __GINI_INDEX

Excluded	Chi-sq	df	Prob.
_DTOURISM	0.484785	1	0.4863
_D2GDP	0.234067	1	0.6285
All	0.511223	2	0.7744

Cette illustration étudie l'indépendance d'une variable des deux autres.

On ne trouvera une p-value <5% qu'entre (tourisme et PIB) et (PIB et Tourisme).

⇒ Il n'y a qu'une seule relation de causalité bidirectionnelle entre PIB et Tourisme.

⇒ Tourisme Granger cause PIB & PIB Granger cause Tourisme.

E. Phénomène de cointégration par l'approche de Johansen & modèle à correction d'erreur (ECM)

Les hypothèses ont été déjà annoncé ci-dessus (III-C)

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.788809	68.82776	42.91525	0.0000
At most 1 *	0.742963	37.72790	25.87211	0.0011
At most 2	0.410133	10.55718	12.51798	0.1040

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level

La trace indique l'existence de deux équations de cointégration

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.788809	31.09986	25.82321	0.0091
At most 1 *	0.742963	27.17072	19.38704	0.0030
At most 2	0.410133	10.55718	12.51798	0.1040

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level

La valeur propre maximale indique l'existence de deux équations de cointégration

⇒ Il existe deux équations de cointégration pour ce modèle déterminé ci-dessous avec :

1 Cointegrating Equation(s):		Log likelihood	-205.7918
Normalized cointegrating coefficients (standard error in parentheses)			
_TOURISM_R...	_GDP_BN\$	_GINI_INDEX	@TREND(2)
1.000000	-0.143275	4.850227	-5.634286
	(0.02562)	(0.89350)	(1.48482)
Adjustment coefficients (standard error in parentheses)			
D(_TOURISM...	0.107259		
	(0.12435)		
D(_GDP_BN\$)	3.873683		
	(2.45725)		
D(_GINI_IN...	-0.500432		
	(0.23236)		

$$T_t = 0.14PIB_t - 4.85P_t + 5.63t$$

On observe qu'à long terme, la croissance économique (PIB) et la tendance possèdent un effet positif sur le développement touristique.

L'indice de Gini est croissant par rapport au taux de pauvreté et il possède un effet négatif sur le développement touristique.

⇒ Le développement Touristique subit un effet positif par les trois autres variables.

La statistique t du modèle est calculé par la formule

$$t = \frac{coef}{std\ error}$$

La variable n'est significative que si t est supérieure à **2.086** (Tableau de Student) en valeur absolue

⇒ On a $t_{pib} = (-0.14/0.02) = -7$ donc significative.

⇒ Et $t_{pov} = (4.85/0.89) = 5.44$ donc significative.

⇒ Et $t_{trend} = (-5.63/1.48) = -3.8$ donc significative

Error Correction:	D(_TOURIS...	D(_GDP_BN\$)	D(_GINI_IN...
CointEq1	0.107259 (0.12435) [0.86258]	3.873683 (2.45725) [1.57643]	-0.500432 (0.23236) [-2.15370]

Les valeurs présentées ci-dessus par le VECM indiquent qu'en cas de déséquilibre à court terme, le PIB s'ajuste plus rapidement que les recettes touristiques.

En effet, la vitesse de convergence du PIB est d'environ 387% et celle des recettes touristiques est de 10,7%

2 Cointegrating Equation(s):		Log likelihood	-192.2064
Normalized cointegrating coefficients (standard error in parentheses)			
_TOURISM_R...	_GDP_BN\$	_GINI_INDEX	@TREND(2)
1.000000	0.000000	-0.261100 (0.11074)	-1.933939 (0.24411)
0.000000	1.000000	-35.67485 (5.05650)	25.82682 (11.1468)
Adjustment coefficients (standard error in parentheses)			
D(_TOURISM...	-2.278173 (0.60581)	0.005492 (0.01196)	
D(_GDP_BN\$)	-37.41413 (14.2022)	-0.193956 (0.28041)	
D(_GINI_IN...	-1.145625 (1.86456)	0.077342 (0.03681)	

- Tourisme-Pauvreté :

On a $(0.26/0.11 = 2.36)$ et $(1.93/0.24 = 8.04)$ don la significativité des coefficients.

$$T_t = 0.26P_t + 1.93t$$

On observe une influence positive de la part de la variable réduction de pauvreté et de la tendance sur le développement touristique.

Les statistiques t sont égaux respectivement à -2.36 et -8.04 et sont donc significatives.

- PIB-Pauvreté :

On a $(35/5 = 7)$ et $(25.82/11.14 = 2.31)$ donc la significativité des coefficients.

$$PIB_t = 35.67P_t - 25.82$$

On observe un effet positif de la pauvreté et négatif de la part de la tendance sur la croissance économique

Les statistiques t sont égaux respectivement à -7.06 et 2.31 et sont donc significatives

V. Conclusion

Ce projet s'est attaché à analyser les relations de causalité entre développement touristique, croissance économique et réduction de la pauvreté en Espagne.

Quatre questions ont été posées à cet effet afin d'identifier les relations entre les trois variables.

Pour répondre à ces questions, on a analysé la stationnarité des variables à l'aide de test augmenté de Dickey-Fuller, puis on a étudié le modèle VAR afin d'utiliser les tests de causalité de Granger, et on a utilisé le test de coïntégration de Johansen pour pointer l'existence de coïntégration entre ces variables et déterminer les équations.

Les résultats ont montré que :

⇒ Stationnarité :

- La série des variables pauvreté est stationnaire en niveau
- La série des variables développement touristique est stationnaire en différence première
- La série des variables croissance économique est stationnaire en seconde différence

⇒ Causalité :

- La relation entre croissance économique et développement touristique est bidirectionnelle, autrement dit, la présence de l'une de ces deux variables

dans l'analyse de l'autre engendre une meilleure prédictibilité de cette dernière.

- La réduction de la pauvreté n'admet aucune relation de causalité avec les deux autres variables, ce qui engendre une difficulté pour le gouvernement à cause de l'importance du domaine touristique dans l'économie de ce pays. De plus, la croissance économique est très importante à cette fin et même la variable dominante pour atteindre une réduction de pauvreté.

Peut-on dire alors que le taux de pauvreté dans ce pays riche a atteint une limite de sorte que sa réduction a été stagné ?

Est-ce qu'il y a d'autres domaines que le tourisme (agricole) qui peuvent influencer la réduction de pauvreté ?

⇒ Coïntégration

- Les trois variables sont cointégrées, elles évoluent ensemble et affichent par conséquent une relation de long terme entre eux.

Cette étude peut être clé pour le gouvernement afin de réduire le taux de pauvreté et de mieux gérer la croissance économique après l'interprétation des relations de causalité entre les différents domaines.

De plus, en cas d'Espagne, on peut étudier le modèle Croissance économique – développement agricole – réduction de pauvreté pour comparer entre les deux domaines et donc

dans lequel des deux on doit mieux investir pour trouver les meilleurs résultats.

Plus loin que ça, on peut exécuter une étude sur différents pays ou bien des continents pour acquérir un KPI (Indicateur de Performance Clé) qui clarifie la vision pour les décideurs du pays vers un future plus stable et profitant.

VI. Références et Bibliographie.

A. Applications utilisées

Eviews, RStudio, PowerBI, Excel.

B. Base de données

Macrotrends, data-bank, data.ocde ;

C. Bibliographie

Tourisme, croissance et intégration dans l'économie mondiale : les apports du concept de développement durable de (**Manhchien Vu**) ;

Kuersteiner,2008

Économétrie (**Régis Bourbonnais**) ;

Modèles linéaires Multivariés : VAR et cointégration, Introduction aux modèles ARCH et GARCH, Introduction à la notion de mémoire longue (**Arthur Charpentier**) (Volume 2) ;

MODELE A CORRECTION D'ERREUR (MCE) ET APPLICATIONS de (*Hélène Hamisultane*) ;

<https://en.wikipedia.org/wiki/Spain>

<http://www.eviews.com/help/>

Article de Toda et Yamamoto (1995) en anglais :

We proposed a simple way to test economic hypotheses expressed as restrictions on the parameters of V AR models without pretest for a unit root(s) and a cointegrating rank(s). Hypothesis tests such as (5) in levels V AR's, in general, involve not only nonstandard distributions but also nuisance parameters if the processes are integrated or cointegrated, and critical values for the tests cannot conveniently be tabulated. So the usual way to proceed is formulating equivalent ECM's in which most hypothesis testing can be conducted using the standard asymptotic theory. But this requires pretests of a unit root and cointegrating rank, which one may wish to avoid if the cointegrating relation itself is not one's interest since those tests are known to have low power. Hence, our simple method of adding extra lags intentionally in the estimation should be very useful in practice. (Ibid., p. 246)
